# An Approach to Generating Mutually Independent Initial Perturbations for Ensemble Forecasts: Orthogonal Conditional Nonlinear Optimal Perturbations

WANSUO DUAN

*State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, and Ningbo Collaborative Innovation Center of Nonlinear Hazard System of Ocean and Atmosphere, Ningbo University, Ningbo, China*

ZHENHUA HUO

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

## ABSTRACT

Conditional nonlinear optimal perturbation (CNOP) is the initial perturbation that satisfies a certain physical constraint and causes the largest nonlinear evolution at prediction time. To yield mutually independent initial perturbations in ensemble forecasts, orthogonal CNOPs are developed. Orthogonal CNOPs are then applied to a Lorenz-96 model to generate initial perturbations for ensemble forecasting, as compared with orthogonal singular vectors (SVs). When the initial analysis errors are fast growing, the ensemble forecasts generated by orthogonal CNOPs of the control forecasts perform much more skillfully. Nevertheless, for slow-growing initial analysis errors, the ensemble forecasts generated by orthogonal SVs achieve higher skill when the ensemble initial perturbations are large, whereas the ensemble forecasts generated by orthogonal CNOPs achieve almost the same forecast skill as those generated by orthogonal SVs when the ensemble initial perturbations are sufficiently small. The initial analysis errors that possess much faster growth behavior are easily influenced by nonlinearity, and extreme events (extreme here refers to strong), because of strong nonlinear instability, may be much more likely to cause fast growth of initial analysis errors. Therefore, the ensemble forecasts generated by orthogonal CNOPs may have higher skill than those generated by orthogonal SVs for extreme events; in particular, the ensemble forecasts generated by orthogonal CNOPs, compared with those generated by orthogonal SVs, require a much smaller number of ensemble members to achieve high skill. Therefore, orthogonal CNOPs may provide another useful technique to generate initial perturbations for ensemble forecasting.

## 1. Introduction

A forecast is an estimate of the future state of the atmosphere or ocean. Forecasts are generally conducted by estimating the current states using observations and then investigating how these states evolve using numerical models. Because of the effect of instability and related nonlinearity, very small errors in initial states can be nonlinearly amplified and lead to large errors in the forecast results (Lorenz 1963). Because we cannot observe every detail of the atmospheric and oceanic initial states, we cannot construct a perfect forecast system, and the initial uncertainties result in large forecast errors. Therefore, there is a limit to how far ahead we can make predictions.

To estimate the forecast uncertainty, Epstein (1969) suggested explicitly integrating the Liouville equation to obtain a probability distribution of the atmospheric state, which could then describe the uncertainties of the forecast results. However, for complex weather and climate models, this approach is computationally unfeasible. Later, Leith (1974) introduced the Monte Carlo forecasting (MCF) method, which generates a

*Corresponding author address*: Prof. Wansuo Duan, 40# Hua Yan Li, Qi Jia Huo Zi, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China.
E-mail: duanws@lasg.iap.ac.cn

group of forecast members to estimate the probability distribution function of forecast states by superimposing random initial perturbations on the initial analysis. This is the basic idea of ensemble forecasting, which produces an estimate of the forecast uncertainties and indicates the probability of the occurrence of weather and climate events. The ensemble mean of the forecasting members is often regarded as the result of a deterministic forecast. The ensemble mean may filter the unpredictable parts and leave the common parts of the forecasting members, ultimately decreasing the uncertainties of single forecast results (Leith 1974; Leutbecher and Palmer 2008).

With its benefit of producing probabilistic distribution information of forecast results, ensemble forecasting has become a major technique in numerical weather and climate forecasting. Various approaches have been introduced to generate initial perturbations for ensemble forecasting and have been applied in operational weather and climate predictions. Among these approaches, singular vectors (SVs) (Lorenz 1965; Molteni et al. 1996; Mureau et al. 1993) have been adopted in operational forecasts by the European Centre for Medium-Range Weather Forecasts (ECMWF) and have achieved great success in reducing forecast uncertainties. SVs are a group of orthogonal initial perturbations that possess the largest growth rate in different but mutually orthogonal subspaces of initial perturbations in linearized models, which are consequently expected to dominate the forecast errors at prediction time and to describe the probabilistic distribution of the forecast results. However, Gilmour and Smith (1997) noted that there are limits in the construction of ensemble perturbations based on their linear approximations. Anderson (1997) indicated that SVs are only sensitive to the evolution of the perturbations in a linear regime, and they fail to provide information about the likelihood of extreme perturbations in nonlinear fields. Therefore, SVs are unable to consider the effect of nonlinearity, and limitations exist in estimating forecast uncertainties.

Considering the limitations of the linear theory of SVs, Mu et al. (2003) proposed the approach of conditional nonlinear optimal perturbation (CNOP) to study the growth behavior of prediction errors caused by initial errors. CNOP is a nonlinear generalization of the leading SV (LSV) that corresponds to the largest growth rate of the initial perturbations in a linearized model and represents the initial perturbations that satisfy a certain physical constraint and possess the largest nonlinear evolution at prediction time (Duan et al. 2004; Mu and Zhang 2006; Duan and Mu 2009; Zhou and Mu 2011; Dijkstra and Viebahn 2015). CNOP

can be approximated by LSV when the initial perturbations are sufficiently small and/or the forecast time period is short; however, with increasing initial perturbation magnitudes and forecast time periods, considerable differences occur between CNOP and LSV gradually. In such a case, CNOP cannot be approximated by LSV because of the effect of nonlinearity. For ensemble forecasting, to consider the effect of nonlinearity on the ensemble initial perturbations, Mu and Jiang (2008a,b) applied the CNOP method to yield ensemble initial perturbations by replacing LSV with CNOP (see also Jiang and Mu 2009) and attempted to improve the related ensemble prediction skill. However, such an approach still involves linear approximation because nonleading SVs are regarded as ensemble initial perturbations. This limitation encouraged us to further investigate the application of CNOP in ensemble forecasting.

Ensemble forecasts, if the ensemble initial perturbations are orthogonal, may provide a better estimation of forecast uncertainties (Annan 2004; Feng et al. 2014). SVs, although they are orthogonal, do not consider the effect of nonlinearity and may not be sufficient for sampling effective initial errors and the related forecast uncertainties in nonlinear models. As such, SVs are limited in their ability to yield ensemble initial perturbations. As mentioned above, CNOP is a natural generalization of LSV in nonlinear regimes. Therefore, it would be useful to calculate orthogonal CNOPs to generate ensemble initial perturbations and to estimate forecast uncertainties. However, how to compute orthogonal CNOPs is an unresolved question. In this paper, we first explore an approach to calculating orthogonal CNOPs, then apply the results to producing ensemble initial perturbations and, finally, study the validity of orthogonal CNOPs in improving ensemble forecast skill. In section 2, we introduce a method to compute orthogonal CNOPs. Then we describe the Lorenz-96 model (Lorenz 1996) adopted in this study in section 3. The experimental strategy is described in section 4. The skill of ensemble forecasts generated by orthogonal CNOPs and SVs is evaluated in section 5. In section 6, we discuss the implications of the results from section 5. Finally, we present a summary and discussion in section 7.

## 2. Orthogonal CNOPs

If we denote the state vector as $\mathbf{U}$ (which can represent, e.g., the atmospheric temperature, surface current, and sea surface temperature), then the evolution of the state vector can be described by the following nonlinear partial differential equation:

$$\begin{cases} \dfrac{\partial \mathbf{U}}{\partial t} = F[\mathbf{U}(\mathbf{x},t)], \\ \mathbf{U}|_{t=0} = \mathbf{U}_0, \end{cases} \quad (1)$$

where $\mathbf{U}(\mathbf{x}, t) = [U_1(\mathbf{x}, t), U_2(\mathbf{x}, t), \ldots, U_n(\mathbf{x}, t)]$ is the state vector, $F$ is a nonlinear differential operator, $\mathbf{U}_0$ is the initial state, $(\mathbf{x}, t) \in \Omega \times [0, T]$, $\Omega$ is a domain in $\mathbb{R}^n$, $0 < T < +\infty$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, and $t$ is the time. Assuming that the dynamic system Eq. (1) and the related initial states are known exactly, then the solution $\mathbf{U}(\mathbf{x}, t)$ at time $\tau$ can be given by

$$\mathbf{U}(\mathbf{x}, \tau) = M_\tau(\mathbf{U}_0), \quad (2)$$

where $M_\tau$ is the propagator of Eq. (1). If $\mathbf{u}_0$ represents the initial errors, the resultant prediction errors $\mathbf{u}_\tau$ at prediction time $\tau$ can be estimated by

$$\mathbf{u}_\tau = M_\tau(\mathbf{U}_0 + \mathbf{u}_0) - M_\tau(\mathbf{U}_0). \quad (3)$$

Based on Eqs. (2) and (3), Mu et al. (2003) defined the CNOP, which represents the initial error that satisfies a certain physical constraint and causes the largest prediction error at the prediction time. Specifically, an initial perturbation $\mathbf{u}_0^*$ is called CNOP if and only if

$$J(\mathbf{u}_0^*) = \max_{\|u_0\|_i \leq \delta} J(\mathbf{u}_0),$$

where

$$J(\mathbf{u}_0) = \|M_\tau(\mathbf{U}_0 + \mathbf{u}_0) - M_\tau(\mathbf{U}_0)\|_f, \quad (4)$$

where $\|\ \|_i$ is a norm to measure the magnitude of the initial errors, $\|\mathbf{u}_0\|_i \leq \delta$ is the constraint condition of the initial error magnitudes ($\delta$ is a constant), and $\|\ \|_f$ is the norm to measure the magnitude of the forecast errors at the prediction time $\tau$, which is also the optimization period associated with the maximization of Eq. (4). From Eq. (4), it is seen that the CNOP is superimposed on the solution $\mathbf{U}(\mathbf{x}, t)$, and the deviation from $\mathbf{U}(\mathbf{x}, t)$ caused by the CNOP is explored. Therefore, the solution $\mathbf{U}(\mathbf{x}, t)$ can be understood as a reference state when investigating the effect of CNOP.

As described in the introduction, Anderson (1997) argued that orthogonal SVs are optimal initial perturbations in linearized models, and they fail to provide information about the likelihood of extreme perturbations in ensemble forecasts. CNOP is the initial perturbation that has the largest nonlinear evolution at the final time of the interval $[0, \tau]$, so it may capture information about extreme perturbations, which therefore encouraged us to explore orthogonal CNOPs associated with ensemble forecasts. The CNOP can be calculated along the descending direction of the gradient of the objective function with respect to the initial perturbations by optimization solvers, such as sequential quadratic programming (SQP; Powell 1983), limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS; Liu and Nocedal 1989), and spectral projected gradient 2 (SPG2; Birgin et al. 2000). CNOP $\mathbf{u}_0^*$ is the global maximum of $J(\mathbf{u}_0)$ in the constraint condition $\|\mathbf{u}_0\|_i \leq \delta$, and there may also exist a local maximum $\mathbf{u}_0^l$. In this case, we call CNOP a global CNOP and $\mathbf{u}_0^l$ a local CNOP. Previous studies showed that the spatial structures of the global CNOP have significant similarities to local CNOPs (Mu et al. 2003; Duan et al. 2004; Jiang et al. 2008). The similarities between these CNOPs are unfavorable for the diversity of ensemble members and may result in a failure to track the error between the control forecast and the true state when they evolve. Moreover, orthogonal initial perturbations lead to a relatively large spread for ensemble members and are favorable for the diversity of ensemble members. In particular, there is a greater possibility that the spread covers the actual atmospheric state to be predicted.

To guarantee the diversity of the ensemble initial perturbations and to take nonlinearity into consideration, orthogonal CNOPs are explored. To obtain orthogonal CNOPs, we first calculate the global (or the first) CNOP using an optimization solver (see last paragraph). Then the second CNOP can be obtained in the subspace orthogonal to the first CNOP; the third CNOP can be calculated in the subspace orthogonal to the first and second CNOPs; the fourth CNOP can be achieved in the subspace orthogonal to the first, second, and third CNOPs; and so on. Mathematically, the related optimization problem for calculating orthogonal CNOPs is as follows. The $j$th CNOP is the initial perturbation $\mathbf{u}_{0j}^*$, satisfying the following optimization equation:

$$J(\mathbf{u}_{0j}^*) = \max_{\mathbf{u}_{0j} \in \Omega_j} \|M_\tau(\mathbf{U}_0 + \mathbf{u}_{0j}) - M_\tau(\mathbf{U}_0)\|_f, \quad (5)$$

where

$$\Omega_j = \begin{cases} \{\mathbf{u}_{0j} \in \mathbb{R}^n | \|\mathbf{u}_{0j}\|_i \leq \delta\}, & j = 1 \\ \{\mathbf{u}_{0j} \in \mathbb{R}^n | \|\mathbf{u}_{0j}\|_i \leq \delta, \mathbf{u}_{0j} \perp \Omega_k, k = 1, \ldots, j-1\}, & j > 1 \end{cases}. \quad (6)$$

Here, $\Omega_j$ is one subspace of the whole space, the symbol $\{\ \}$ here represents the set of vectors, $\perp$ is the orthogonality of vector spaces, $\mathbf{u}_{0j}$ is the initial perturbation in the subspace $\Omega_j$, and $\|\mathbf{u}_{0j}\|_i \leq \delta$ is the constraint condition of the initial perturbations ($\delta$ is the constraint

radius). According to Eqs. (5) and (6), the first CNOP (i.e., $\mathbf{u}_{01}^*$) possesses the largest nonlinear evolution in the first subspace (i.e., the whole space) and the $j$th CNOP (i.e., $\mathbf{u}_{0j}^*$) possesses the largest nonlinear evolution in the subspace orthogonal to the $j-1$ CNOPs (i.e., $\mathbf{u}_{01}^*, \mathbf{u}_{02}^*, \ldots, \mathbf{u}_{0j-1}^*$). Orthogonal CNOPs, compared with orthogonal SVs, further consider the effect of nonlinearity.

In this paper, we use the SPG2 solver to compute orthogonal CNOPs. We adopt the L2 norm for measuring the initial errors and the related prediction errors: that is,

$$\|\mathbf{x}\|_{L2} = \sqrt{\sum_{i=1}^{n} x_i^2}, \qquad (7)$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ is the vector to be measured.

## 3. The Lorenz-96 model

In this paper, we use the Lorenz-96 model (Lorenz 1996) to compute orthogonal CNOPs and to evaluate the related ensemble forecast skill. This model has been used to study various questions associated with predictability. For example, it has been applied in studies of data assimilation (Fertig et al. 2007; Whitaker and Hamill 2002; Hunt et al. 2004; Ott et al. 2004) and adaptive observation (Khare and Anderson 2006). In particular, the model has been used to explore the theory of ensemble forecasting (Roulston and Smith 2003; Revelli et al. 2010; Koyama and Watanabe 2010; Descamps and Talagrand 2007; Basnarkov and Kocarev 2012; Li et al. 2013). Therefore, we chose the Lorenz-96 model to investigate the role of orthogonal CNOPs in generating ensemble initial perturbations.

The model is governed by the following differential equation:

$$\frac{dX_j}{dt} = (X_{j+1} - X_{j-2})X_{j-1} - X_j + F, \qquad (8)$$

where $j = 1, \ldots, m$ with cyclic boundary conditions. The variables in Eq. (8) are nondimensional. Despite the model being artificial, it describes the main basic characteristics of atmospheric motion and is commonly used to simulate atmospheric dynamics over a single latitudinal circle, such as the dynamical behavior of vorticity, temperature, and gravitational potential (Lorenz 1996; Lorenz and Emanuel 1998; Basnarkov and Kocarev 2012). The three terms on the right-hand side of Eq. (8) stand for the nonlinear advection, damping, and forcing of the atmosphere, respectively. The nonlinear

behavior of this system changes with the number (i.e., $m$) of variables and the magnitude of the external forcing $F$. Lorenz and Emanuel (1998) demonstrated that the model is chaotic when $m = 40$ and $F = 8$, and in such a case, one time unit is 5 days in reality. For the parameter values used here, once the system has reached its attractor, the expectation and standard deviation of $X_j$ ($j = 1, \ldots, m$) are approximately 2.3 and 3.6, and the cross correlation between $X_j$ ($j = 1, \ldots, m$) is negligible. In addition, Lorenz and Emanuel (1998) demonstrated that if the model is discretized with a fourth-order Runge–Kutta scheme and the related time step is 0.05 time units (i.e., 6 h), the error-doubling time is approximately 0.4 time units (i.e., 2 days) which is consistent with realistic weather forecast models. These results suggest that the Lorenz-96 model, with the above configuration, is acceptable for studying ensemble forecasts associated with orthogonal CNOPs.

## 4. Experimental strategy

After a spinup run of 4000 time steps (i.e., 1000 model days), we continue to integrate the Lorenz-96 model for 730 000 time steps (i.e., 500 model years) and obtain time series of $X_j$ ($j = 1, \ldots, m$), where $X_j$ can be regarded as being the discrete component of the variable $\mathbf{X}$ along one latitudinal circle. From the time series of the variable $\mathbf{X}$, we take the state values of $\mathbf{X}$ every 1460 time steps (i.e., 1 model year) as initial values and integrate the model over 40 time steps (i.e., 10 model days), ultimately obtaining 500 "truth runs," which are regarded as the future states to be forecasted.

To forecast the truth run, an initial analysis field is determined. Here, we use the four-dimensional variational data assimilation (4DVAR) method to generate the initial analysis fields. The corresponding observations are produced by adding random-noise (observational) errors with a standard normal distribution, $N(0, I)$, to the truth run every 6 h (i.e., one time step), where the standard deviation of the observational errors is 28% of the standard deviation of $X_j$ ($j = 1, \ldots, m$). Therefore, the magnitude of the observational errors is acceptable. We also used other magnitudes of observational errors and found that the results were qualitatively less sensitive to the magnitudes of observational errors. Therefore, in this paper, we use the above observational error magnitude to describe the results. When forecasting the truth run, we use 4DVAR to generate the initial analysis by assimilating the observations at the initial and the next time step. With this initial analysis field, we integrate the Lorenz-96 model and obtain a forecast for the truth run. The difference between the initial analysis field and the initial value of

the truth run to be predicted is the initial analysis error. The related prediction errors are estimated by evaluating the differences between the forecasts and the truth runs. For convenience, the corresponding initial analysis error is referred to as ''4DVAR-type analysis error'' (the related amplitude is denoted by $\delta_a$), and the resultant forecast is called the control forecast.

For each of the chosen 500 truth runs (i.e., cases), we regard the control forecast as a reference state to compute the orthogonal CNOPs. CNOPs depend on the amplitudes of the initial perturbation (indicated by the constraint radius $\delta$ in section 2) and optimization time period $T$. Therefore, we take different combinations of $\delta$ and $T$ to calculate the orthogonal CNOPs of each reference state and conduct ensemble forecast experiments. We adopted 16 combinations of $\delta$ and $T$ (see Table 1), and the first 15 orthogonal CNOPs of the reference state are obtained for each combination. The 15 orthogonal CNOPs are then superimposed on the initial analysis field of the corresponding control forecast (i.e., the reference state) to yield 15 perturbed initial analysis fields. In addition, we also superimpose the initial perturbations on the initial analysis with patterns opposite to the orthogonal CNOPs to generate another 15 perturbed initial analysis fields. As a result, 30 perturbed initial analysis fields are obtained for the control forecast. By integrating the Lorenz-96 model with each perturbed initial analysis field, we obtain a forecast of the corresponding truth run. For 30 perturbed initial analysis fields, 30 forecasts can be obtained for the truth run. Combined with the control forecast, 31 forecasts are obtained. We regard these forecasts as ensemble members to evaluate the skill of the ensemble forecasts associated with orthogonal CNOPs. For the 500 truth runs, a total of 15 500 forecast members are obtained. Based on these forecast members, we evaluate the skill of the ensemble forecast of the orthogonal CNOPs.

The skill of an ensemble forecast is often evaluated using the root-mean-square error (RMSE; Murphy and Epstein 1989), anomaly correlation coefficient (ACC; Murphy and Epstein 1989), Brier score (BS; Brier 1950), and relative operating characteristic curve area (ROCA; Mason 1982). The RMSE and ACC are generally used to evaluate the forecast skill of the ensemble mean. The former measures the difference between the ensemble mean and the observation (here, the ''truth state''), and the smaller the RMSE is, the more accurate the ensemble mean. The latter estimates the anomaly correlation between the ensemble mean and the truth state, and the larger the value of the ACC is, the higher the forecast skill. BS and ROCA are often applied to estimate the probabilistic forecast skill of an ensemble forecast. The BS is the mean-square error of the

TABLE 1. Shown are 16 combinations of perturbation magnitude $\delta$ and optimization time $T$ ($\delta_a$: magnitude of the initial analysis error; $S$–$i$: combination of $\delta$ and $T$).

| | $\delta$ | | | |
|---|---|---|---|---|
| $T$ (days) | $30\%\delta_a$ | $50\%\delta_a$ | $80\%\delta_a$ | $100\%\delta_a$ |
| 2 | $S$–1 | $S$–2 | $S$–3 | $S$–4 |
| 3 | $S$–5 | $S$–6 | $S$–7 | $S$–8 |
| 4 | $S$–9 | $S$–10 | $S$–11 | $S$–12 |
| 5 | $S$–13 | $S$–14 | $S$–15 | $S$–16 |

probability forecasts for a binary event and comprehensively evaluates the forecast reliability, the forecast resolution, and the observational uncertainty of the ensemble forecast. Smaller BS values indicate better probability forecasts. ROCA measures the resolution of the forecast system and evaluates the probabilistic forecast skill of the occurrence of a binary event. Larger values of ROCA indicate higher skill for ensemble forecasts. Generally, forecasts can be regarded as skillful when the value of ROCA is larger than 0.5. A more detailed description of the four scores is provided in appendixes A through D.

BS and ROCA are computed for the frequent event ev1, $X_j > 2.0$ ($j = 1, \ldots, 40$), and the less frequent event ev2, $X_j > \mu_j + \sigma_j$ ($j = 1, \ldots, 40$), where $\mu_j$ and $\sigma_j$ are the corresponding climatological mean and standard deviation of $X_j$. The former indicates that $X_j$ occurs with a frequency of 0.5, whereas the latter implies a frequency of approximately 0.175. Here, the climatological mean is obtained by taking the mean of 10-yr integrations of the model.

## 5. Results

In this section, we adopt the Lorenz-96 model and the experimental strategy shown in section 4 to evaluate the validity of orthogonal CNOPs for improving the ensemble forecast skill in a perfect model scenario, which is then compared with orthogonal SVs to investigate whether the former is better than the latter in improving forecast skill.

### a. The ensemble forecast experiments associated with orthogonal CNOPs

The skill level of 10-day ensemble forecasts is first assessed by RMSE and ACC, which measure the forecast skill of the ensemble mean. For each truth run and lead time (6 h, 12 h, 18 h, 1 day, 30 h, . . . , 10 days), we calculate the RMSE and ACC for each $S$–$i$ ($i = 1, 2, \ldots, 16$) and plot the mean of RMSE (and ACC) for all truth runs and lead times in Fig. 1. The ensemble forecast skill measured by the mean RMSE (and ACC) varies with
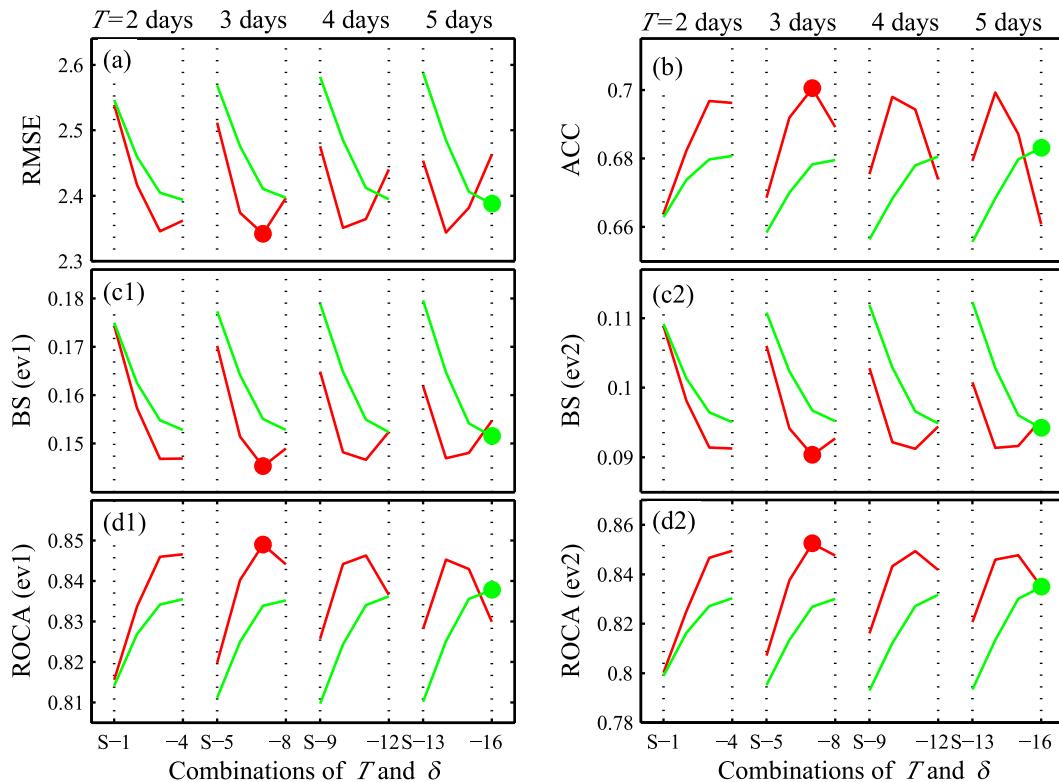
FIG. 1. The scores of the ensemble forecasts generated by orthogonal CNOPs (red) and SVs (green) measured by (a) RMSE, (b) ACC, BS for (c1) the frequent event ev1 and (c2) the less frequent event ev2, and ROCA for (d1) the frequent event ev1 and (d2) the less frequent event ev2, averaged over 500 truth runs and all lead times in 10 days. The horizontal axis denotes the combinations of the optimization the time period $T$ and the initial perturbation magnitude $\delta$, and the vertical axis indicates the scores measured by the corresponding measurements. The intervals with dashed lines divide the $S$–$i$ that correspond to the same optimization time period $T$. In each interval of $S$–$i$, the values of $\delta$ increase with increasing $i$ values. The dots indicate the combination of $T$ and $\delta$ that corresponds to the highest skill for the ensemble forecasts generated by orthogonal CNOPs (red) and SVs (green).

$\delta$ and $T$ ($\delta$ measures the amplitude of CNOPs, and $T$ is the optimization time period used to calculate the CNOPs; see section 4). In particular, for each $T$, the RMSE (ACC) presents the minimum (maximum) values of $\delta$. Of the minima (maxima) corresponding to different $T$, the value with $T = 3$ days and $\delta = 80\%\delta_a$ ($\delta_a$ is the amplitude of the initial analysis error; see section 4) is the smallest (largest), which indicates that the ensemble forecast is generally much more skillful when the orthogonal CNOPs of $T = 3$ days and $\delta = 80\%\delta_a$ are used to generate the ensemble initial perturbations.

We also use BS and ROCA to measure the probability prediction skill of the ensemble forecasts, where all forecast variables and truth runs are combined to compute the BS and ROCA at every lead time within 10 days (6 h, 12 h, 18 h, 1 day, 30 h, 36 h, . . . , 10 days; Fig. 1), and the number of realizations of the prediction processes is $N = 40 \times 500 = 20\,000$. The results illustrate that the ensemble forecasts with CNOPs of $T = 3$ days and $\delta = 80\%\delta_a$ provide a better estimation of the forecast

uncertainties and have higher forecast skill, which is accordant with the results obtained by RMSE and ACC, which measure the forecast skill of the ensemble mean.

Orthogonal CNOPs are developed based on orthogonal SVs and take the effect of nonlinearity into consideration. As such, orthogonal CNOPs and SVs should have different spatial structures. As an example, we plot in Fig. 2 the first and seventh CNOPs (and SVs) for one truth run with constraint bound $\delta$ being the 4DVAR-type analysis error $30\%\delta_a$ and $\delta_a$ and the optimization time period being 2 days. The CNOPs with large magnitudes are different from the corresponding SVs, which may indicate that the ensemble forecasts made by CNOPs and SVs have different skill scores. To confirm this result, we compute orthogonal SVs for the control forecast of each truth run and take the first 15 orthogonal SVs, similar to orthogonal CNOPs. We scale the SVs to possess the same amplitude as the CNOPs. If $\mathbf{u}_L$ is the $i$th SV for an optimization time period $T$, then the scaled $i$th SV $\hat{\mathbf{u}}_L$ can be defined as follows:
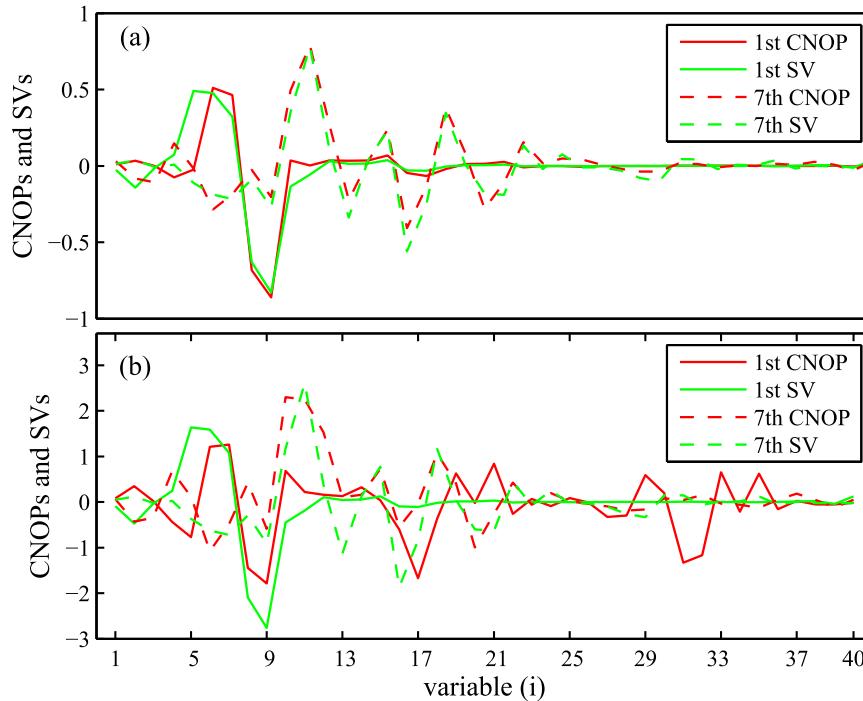
FIG. 2. The first and seventh CNOPs and SVs for one of the truth runs with a 4DVAR-type analysis error of $\delta$ of (a) $30\%\delta_a$ and (b) $\delta_a$ and an optimization period of 2 days. For $\delta$ equal to $30\%\delta_a$, the similarity coefficient (see appendix E) between the CNOPs and SVs in (a) exceeds 0.88; and for $\delta$ equal to $\delta_a$, the similarity coefficient between the CNOPs and SVs in (b) is less than 0.76. This indicates that the CNOPs are different from the SVs when the initial perturbations are large.

$$\hat{\mathbf{u}}_L = (-1)^n \frac{\|\mathbf{u}_\delta\|}{\|\mathbf{u}_L\|} \mathbf{u}_L, \quad n = 1, 2, \qquad (9)$$

where $\mathbf{u}_\delta$ is the $i$th CNOP, with the optimization period $T$ being the same as the corresponding $i$th SV. Thus,

$$\|\hat{\mathbf{u}}_L\| = \|\mathbf{u}_\delta\|; \qquad (10)$$

that is, the scaled $i$th SV $\hat{\mathbf{u}}_L$ possesses the same amplitude as that of the $i$th CNOP $\mathbf{u}_\delta$. By this approach, we can obtain orthogonal SVs with different combinations of $T$ and $\delta$. From Eq. (9), there are two scaled SVs for one SV; furthermore, they are of opposite signs. That is to say, for the first 15 SVs, we can obtain 15 pairs of SVs (i.e., a total of 30 SVs). These 30 scaled SVs, together with the control forecast, are used in ensemble forecasting to compare the forecast skill between SVs and CNOPs. To facilitate the description, we continue to refer to the scaled SVs hereafter as SVs.

Similar to CNOPs, the skills of the ensemble forecasts generated by the SVs are also evaluated using RMSE, ACC, BS, and ROCA for 500 truth runs and all of the related lead times within 10 days (see Fig. 1). The results show that the ensemble forecasts have different skill scores

for different $S$–$i$; for each $T$, the skill scores present a minimum for RMSE and BS and a maximum for ACC and ROCA for values of $\delta$. Among these minima or maxima, the one showing with a skill score of $T = 5$ days and $\delta = 100\%\delta_a$ is the smallest for RMSE and BS and the largest for ACC and ROCA. However, when comparing the forecast skill associated with SVs and CNOPs, almost all ensemble forecasts generated by the CNOPs have much higher skill than those generated by the SVs. In particular, the ensemble forecast generated by CNOPs with $T = 3$ days and $\delta = 80\%\delta_a$ has the highest skill among all forecasts generated by both CNOPs and SVs. These results indicate that orthogonal CNOPs may be more applicable than orthogonal SVs in generating ensemble initial perturbations and improving forecast skill.

Therefore, the skill of ensemble forecasts generated by CNOPs is higher than that generated by SVs. However, the skill is evaluated by taking the mean of the forecast skill scores for the 500 truth runs and/or all of the related lead times within 10 days. To validate the ensemble forecast skill of the orthogonal CNOPs, we also calculate the skill scores of the ensemble forecasts corresponding to each lead time for 500 truth runs. Figure 3 illustrates the results, which show that, as the lead time gradually
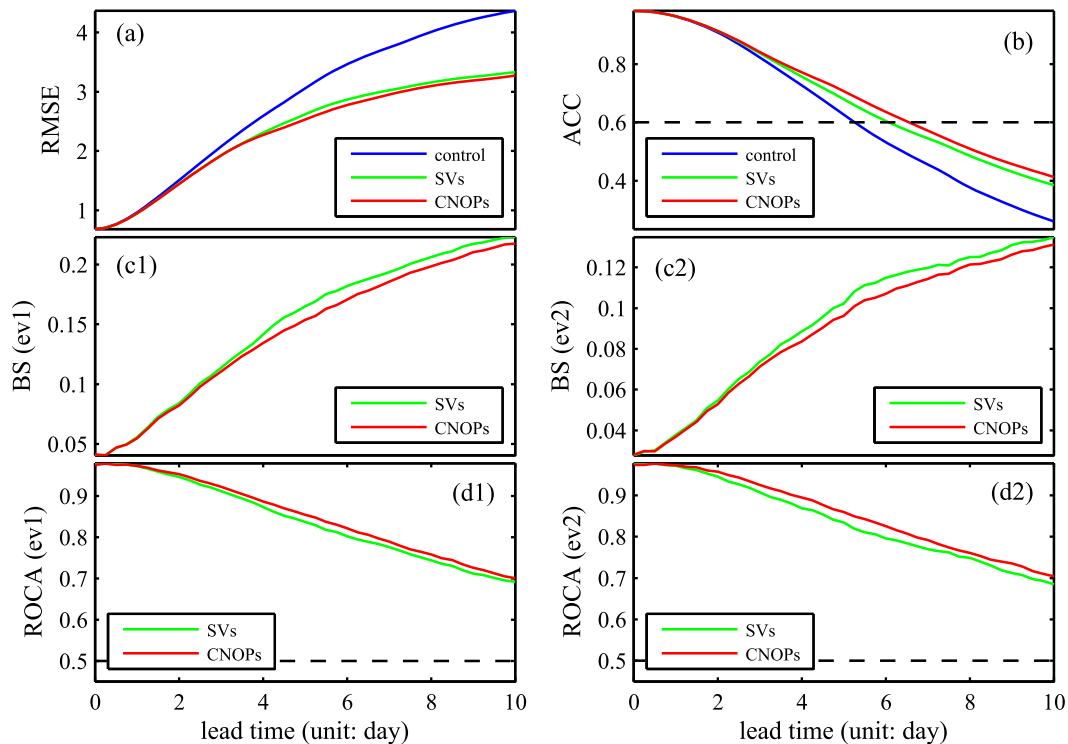
FIG. 3. The skill scores of the ensemble forecasts generated by orthogonal CNOPs ($T = 3$ days and $\delta = 80\%\delta_a$; red lines) and SVs ($T = 5$ days and $\delta = 100\%\delta_a$; green lines) for lead times of $1, 2, 3, \ldots, 10$ days, measured by (a) RMSE, (b) ACC, BS for (c1) the frequent event ev1 and (c2) the less frequent event ev2, and ROCA for (d1) the frequent event ev1 and (d2) the less frequent event ev2. The horizontal axis denotes the lead time, and the vertical axis indicates the skill score. The blue lines represent the skill scores of the control forecasts for 500 truth runs. The dashed lines are the reference lines, denoting the skill scores of 0.6 for ACC and 0.5 for ROCA.

increases, the skill of the ensemble forecasts generated by orthogonal CNOPs and SVs is substantially different, with the former being higher than the latter. Nonlinearities have greater effects when the lead times are long. CNOPs are directly derived from a nonlinear model and contain the effect of nonlinearities, which may cause the related ensemble forecasts to possess higher skill than those generated by SVs. We also explore the dependence of the ensemble forecast skill associated with CNOPs on truth runs to be predicted. However, the ensemble forecasts related to CNOPs for different truth runs are not always of higher skill than those related to SVs. Therefore, we naturally ask the following question: What conditions are responsible for the ensemble forecasts generated by CNOPs being more skillful than those generated by SVs? In the next section, we address this question.

### b. Conditions responsible for the superior performance of the ensemble forecasts generated by CNOPs

For each of the 500 truth runs, we evaluate the ensemble forecast skill by combining all variables to calculate the forecast skill scores measured by RMSE, ACC, BS, and ROCA for 31 forecast members generated by the CNOPs and SVs at each lead time. We then use the mean of the forecast skill scores obtained at all lead times to compare the ensemble forecast skill generated by CNOPs with that generated by SVs for each truth run. The ensemble forecasts generated by CNOPs are not always of higher skill than those generated by SVs for different truth runs. Specifically, there are 236 (76) truth runs whose ensemble forecasts generated by CNOPs (SVs) possess much higher skill than those generated by SVs (CNOPs), in terms of each of the four forecast skill scores. For convenience, we classify the 236 truth runs as category 1 and the 76 truth runs as category 2.

For the truth runs in category 1 and category 2, we use the empirical orthogonal function (EOF) to extract the leading EOF mode (EOF1) of the time-dependent prediction errors (measured by the L2 norm) generated by the corresponding control forecasts. The EOF1 for the truth runs in category 1 and category 2 explains more than 92% of the total variance, which indicates
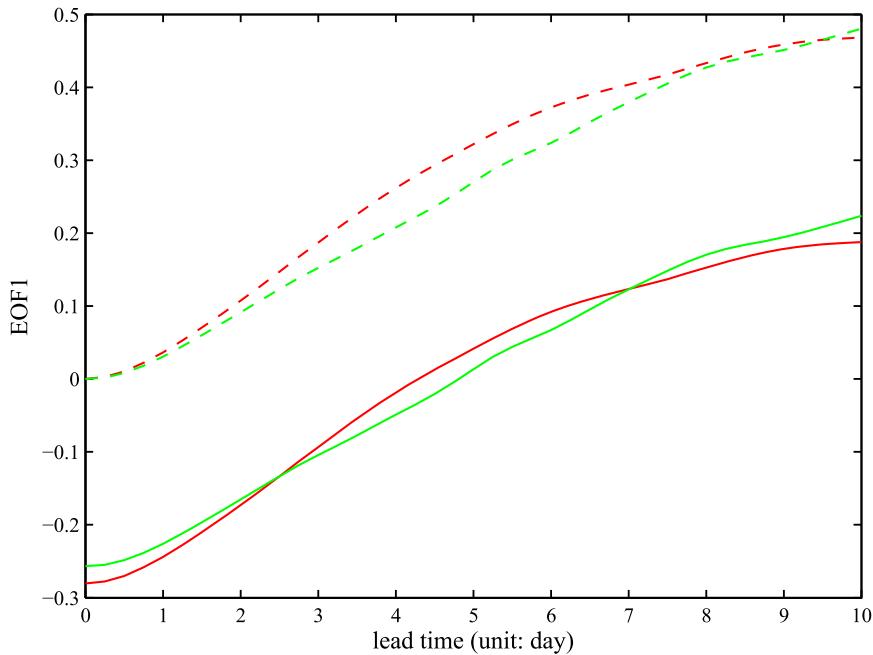
FIG. 4. The EOF1 of the time-dependent prediction errors generated by the control forecasts for the 236 truth runs in category 1 (red solid line) and the 76 truth runs in category 2 (green solid line). The EOF1s of the truth runs in category 1 and category 2 are each moved to the dashed lines to have the same value at the initial time.

that the respective EOF1s can describe the common characteristics of the evolutionary tendency of the prediction errors generated by the control forecasts. To compare the evolutionary tendencies of prediction errors generated by control forecasts of the truth runs in category 1 and category 2, we shift their EOF1s to have the same value at the initial time (Fig. 4). The results show that the prediction errors generated by the control forecasts for the truth runs in category 1 tend to grow faster than those in category 2; furthermore, for the truth runs in category 1, the ensemble forecasts generated by CNOPs have much higher skill than those generated by SVs, which may indicate that orthogonal CNOPs are more useful than orthogonal SVs in achieving high ensemble forecast skill for truth runs with control forecasts possessing faster-growing forecast errors. In particular, the initial analysis errors of the control forecasts in category 1 are more similar to the CNOP-type error of the corresponding truth runs than those in category 2 (see Table 2). Here, the so-called CNOP-type error is superimposed on the truth runs and acts as the fastest-growing initial error for the truth runs, which is a global CNOP and is computed with the optimization time periods being those of the orthogonal CNOPs of the control forecasts and the initial perturbation magnitude being the magnitude of the initial analysis errors. The similarities between the CNOP-type errors and initial

analysis errors in category 1 may explain why the initial analysis errors in category 1 grow much faster than those in category 2, where the similarities are measured by a similarity coefficient (see the appendix E). As a result, the faster-growing CNOPs of the control forecasts may be more likely than the SVs to describe the initial analysis errors in category 1; therefore, the ensemble forecasts generated by CNOPs in category 1 are of higher skill than those generated by SVs. That is to say, the orthogonal CNOPs are more appropriate than the orthogonal SVs for describing initial analysis errors that grow faster.

To further validate the above argument, we conduct two additional experiments. Because the CNOP-type errors of truth runs grow much faster than the SV-type errors, we directly take the CNOP-type errors (and SV-type errors) of the truth runs as the initial analysis error, where the SV-type errors are superimposed on the truth

TABLE 2. Mean similarity coefficient between the 4DVAR-type analysis errors for the truth runs in category 1 and category 2 and the CNOP-type errors of truth runs with different optimization times.

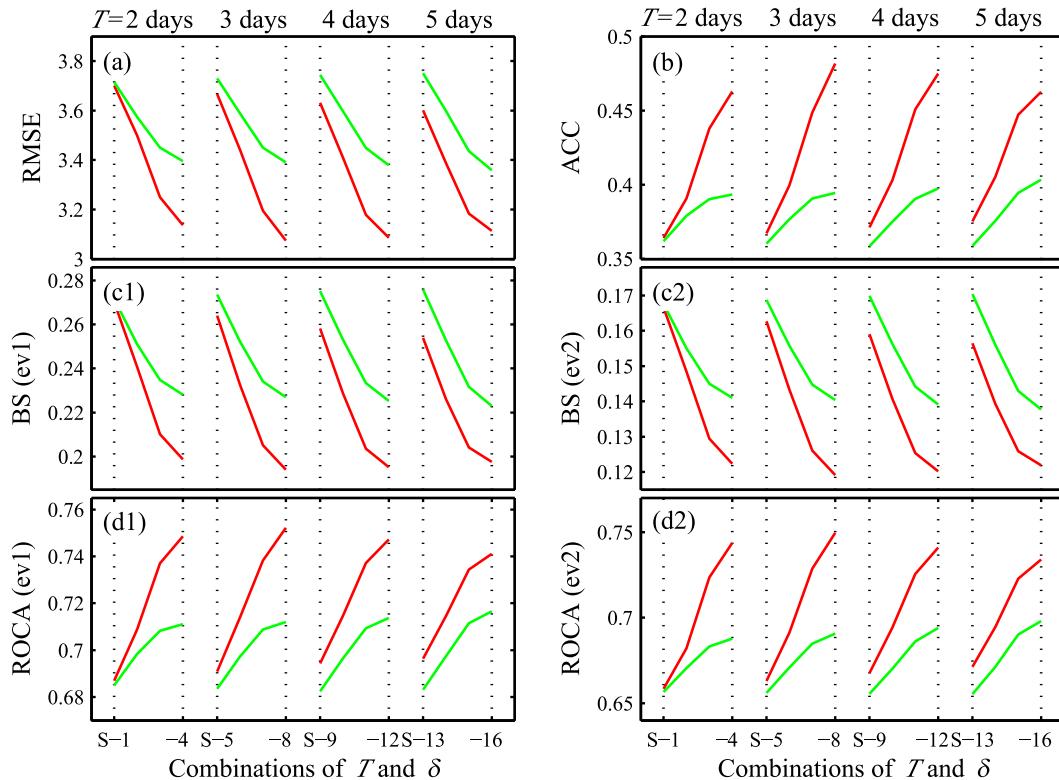| Optimization time (days) | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| Category 1 | 0.0291 | 0.0224 | 0.0007 | 0.0243 |
| Category 2 | −0.0215 | −0.0051 | 0.0004 | −0.0185 |

FIG. 5. As in Fig. 1, but for 500 truth runs with CNOP-type analysis errors, where the magnitudes of the CNOP-analysis errors are the same as those of the 4DVAR-type errors and the optimization time period is 2 days (see section 5b).

run and represent the lead one of SVs. Now we investigate whether the ensemble forecasts generated by orthogonal CNOPs (and SVs) have higher skill than those generated by orthogonal SVs (and CNOPs). The first experiment involves taking the CNOP-type errors of the truth run as the initial analysis errors, and the second experiment involves regarding the SV-type error of the truth run as the initial analysis error. To facilitate the following description, we use "CNOP-type analysis error," and "SV-type analysis error" to denote the initial analysis errors obtained by CNOP-type errors and SV-type errors, respectively. We have denoted the initial analysis error of the aforementioned control forecasts as 4DVAR-type analysis errors in section 4.

1) EXPERIMENT 1

In this experiment, we use the CNOP-type errors as the initial analysis error to yield control forecasts for the truth runs. Specifically, the CNOP-type errors of the truth runs are calculated with optimization time periods of 2 days and an initial perturbation amplitude constrained by the corresponding 4DVAR-type analysis error. Based on these control forecasts corresponding to CNOP-type analysis errors, we compute the orthogonal

CNOPs and SVs for all combinations of $T$ and $\delta$ in Table 1 and use the results to generate initial perturbations of ensemble forecasts for 500 truth runs. The results show that the ensemble forecasts generated by the orthogonal CNOPs have higher skill than those generated by orthogonal SVs (Fig. 5). In particular, for each optimization time period $T$ adopted in computing the orthogonal CNOPs and SVs, the ensemble forecast skill generated by CNOPs gradually increases with $S$–$i$ as $i$ increases; furthermore, the skill becomes increasingly higher than that generated by SVs, which indicates that, for each $T$, the extent to which the ensemble forecast skill generated by CNOPs is higher than that generated by SVs becomes increasingly significant as the magnitudes of CNOPs increase toward those of the 4DVAR analysis errors.

We also use the CNOP-type analysis errors with optimization time periods of 3, 4, and 5 days to yield control forecasts and to calculate their orthogonal CNOPs and SVs. We then compare the related ensemble forecast skill associated with CNOPs and SVs and find similar results (i.e., the ensemble forecasts generated by CNOPs show greatly improved skill compared with the ensemble forecasts generated by SVs). For simplicity,
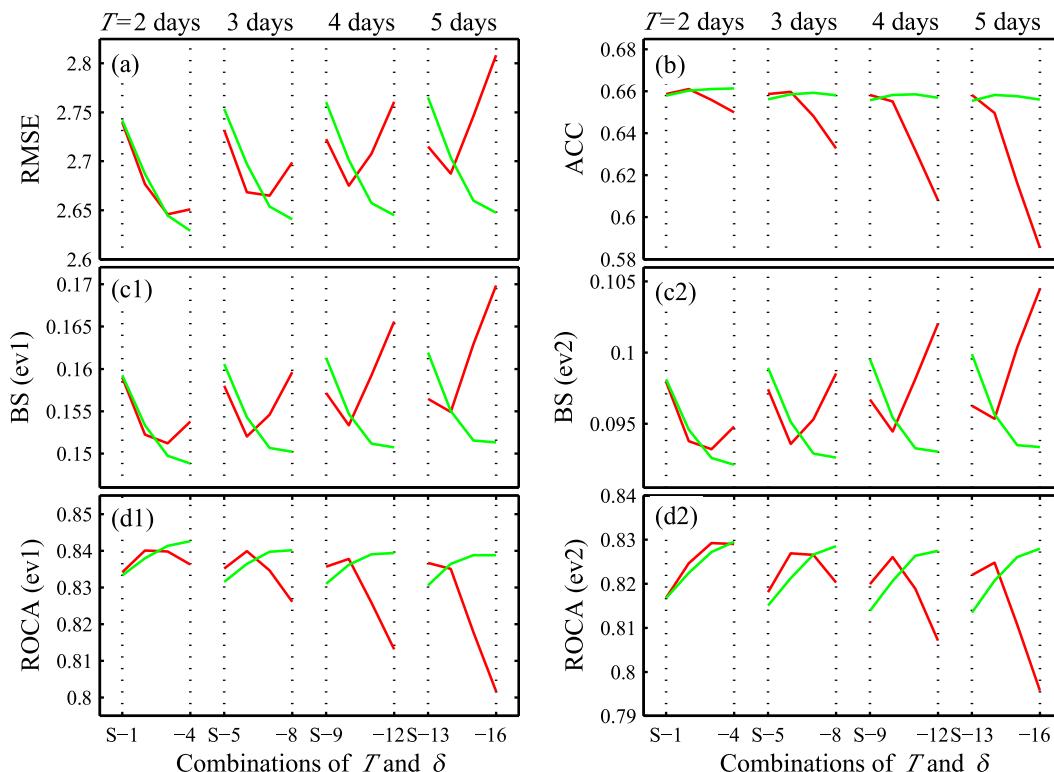
FIG. 6. As in Fig. 1, but for 500 truth runs with SV-type analysis errors, where the magnitudes of the SV-type analysis errors are the same as those of the 4DVAR-type errors and the optimization time period is 2 days (see section 5b).

the related figures and tables are omitted here. These results indicate that the orthogonal CNOPs of the control forecasts may be better able to capture the fast growth behavior of the CNOP-type analysis errors, indicating that the related ensemble forecasts would possess higher skill than those generated by SVs.

2) EXPERIMENT 2

In this experiment, we examine whether the ensemble forecasts generated by orthogonal SVs possess higher skill when the initial analysis errors are taken as the SV-type analysis errors.

The SV-type analysis errors, similar to the CNOP-type analysis errors, are computed for the optimization time periods of the orthogonal SVs superimposed on the control forecasts in section 5a. Then the obtained SV-type analysis errors are scaled to have the same amplitude (measured by the L2 norm) as the corresponding 4DVAR-type analysis error. We use these SV-type analysis errors to yield control forecasts for the corresponding truth runs. Based on these control forecasts, we compute the orthogonal CNOPs and SVs for all combinations of $T$ and $\delta$ in Table 1 and use the results to generate ensemble initial perturbations for ensemble

forecasts. The results show that, for each optimization time period associated with calculating the CNOPs and SVs, the ensemble forecasts generated by the SVs are not always of higher forecast skill than those generated by CNOPs for different magnitudes (denoted by $\delta$; see section 4) of ensemble initial perturbations, despite the initial analysis errors being taken as the SV-type errors of the truth runs. Specifically, for each optimization time period $T$, when $\delta$ is smaller, the skill of the ensemble forecasts generated by the CNOPs and SVs is nearly the same, with that generated by CNOPs a little higher than that generated by SVs. When $\delta$ is larger, the skill of the ensemble forecasts generated by SVs is significantly higher than that generated by CNOPs (Fig. 6). The SV-type analysis errors are relatively slowly growing and present more stable dynamical growing behavior. Therefore, when $\delta$ is sufficiently small, the nonlinear evolution of the initial perturbations can be approximated by the linear counterparts (Duan et al. 2009), and the nonlinear effect can be neglected. In this case, CNOPs are trivially different from the SVs (also see Fig. 2a); therefore, the difference between the ensemble forecast skill associated with CNOPs and SVs is small. Nevertheless, when $\delta$ is larger, the CNOPs include the
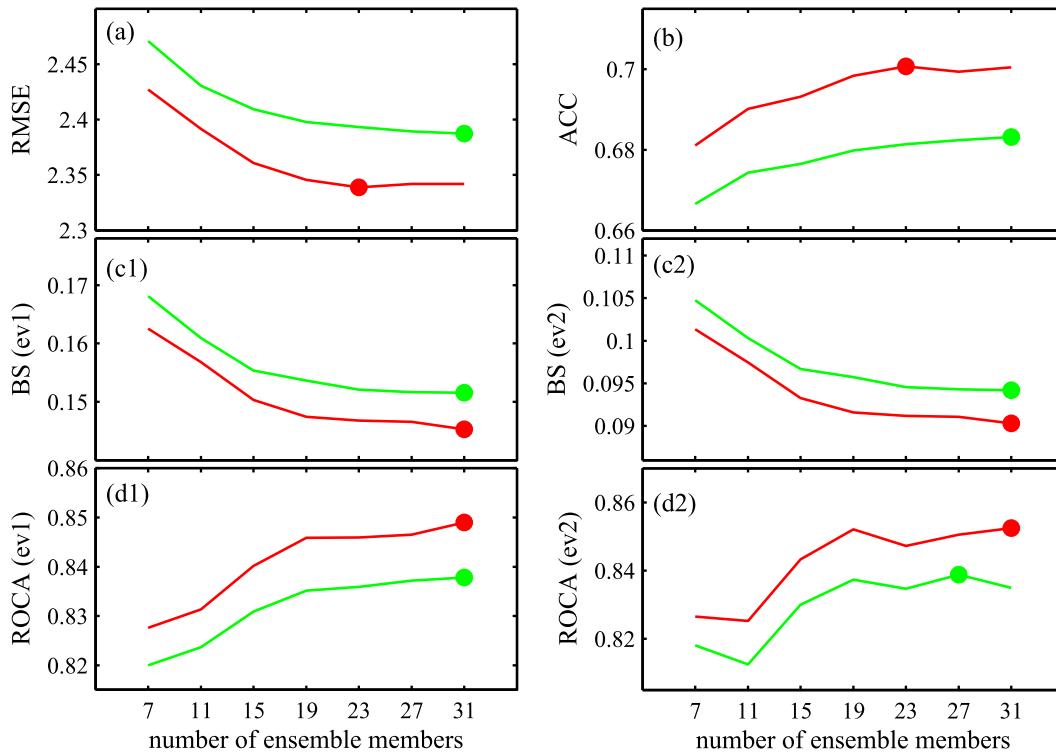
FIG. 7. The averaged skill scores over 500 truth runs and all lead times in terms of the score when the forecast skill associated with CNOPs (red lines) and SVs (green lines) is the highest for different combinations of $T$ and $\delta$. Here, the related control forecast is generated by the 4DVAR-type analysis errors. The horizontal axis represents the numbers of ensemble members, and the vertical axis denotes different evaluation scores, as in Fig. 1. The dots indicate the ensemble forecast members' number that corresponds to the highest skill for different numbers of ensemble members.

full nonlinear effects, whereas the SVs are completely linear; therefore, the CNOPs are significantly different from the SVs (see Fig. 2b). Consequently, the corresponding orthogonal CNOPs superimposed on the control forecasts may overestimate the nonlinearity of the growth of the SV-type analysis error, ultimately causing the corresponding ensemble forecast skill to be lower than that generated by SVs.

### c. Influence of the number for ensemble members associated with orthogonal CNOPs and SVs on forecast skill

In sections 5a and 5b, we compare the role of orthogonal CNOPs and SVs in improving ensemble forecast skill by taking the equivalent number of ensemble members. In this section, we explore the impact of the change of ensemble members' number on the forecast skill associated with orthogonal CNOPs and SVs. Specifically, we adopt the approach shown in sections 4 and 5a to yield the orthogonal CNOPs and SVs and use the results to obtain 7, 11, 15, 19, 23, 27, and 31 ensemble members for each control forecast of the truth runs. For

each of these ensemble members' numbers, we conduct numerical experiments similar to section 5a. The results demonstrate that the combination of $T$ and $\delta$ with the highest ensemble forecast skill varies with the ensemble members' number. Based on the combination of $T$ and $\delta$ showing the highest forecast skill for each number of ensemble members, we study the effect of the number of ensemble members on the related forecast skill.

The results show that, for the control forecast generated by the 4DVAR-type analysis errors, the CNOPs always perform better than SVs regardless of the number of ensemble members (Fig. 7), which indicates that even if we obtain a small number of ensemble members associated with CNOPs, there could be a high forecast skill compared with that associated with SVs. In addition, we notice that the skill of the ensemble forecasts is higher when the number of ensemble members becomes large. This does not mean that the larger the number of ensemble members is, the higher the ensemble forecast skill. In fact, there may exist an upper limit to the numbers of ensemble members for achieving a high forecast skill. For example, the highest skill of ensemble
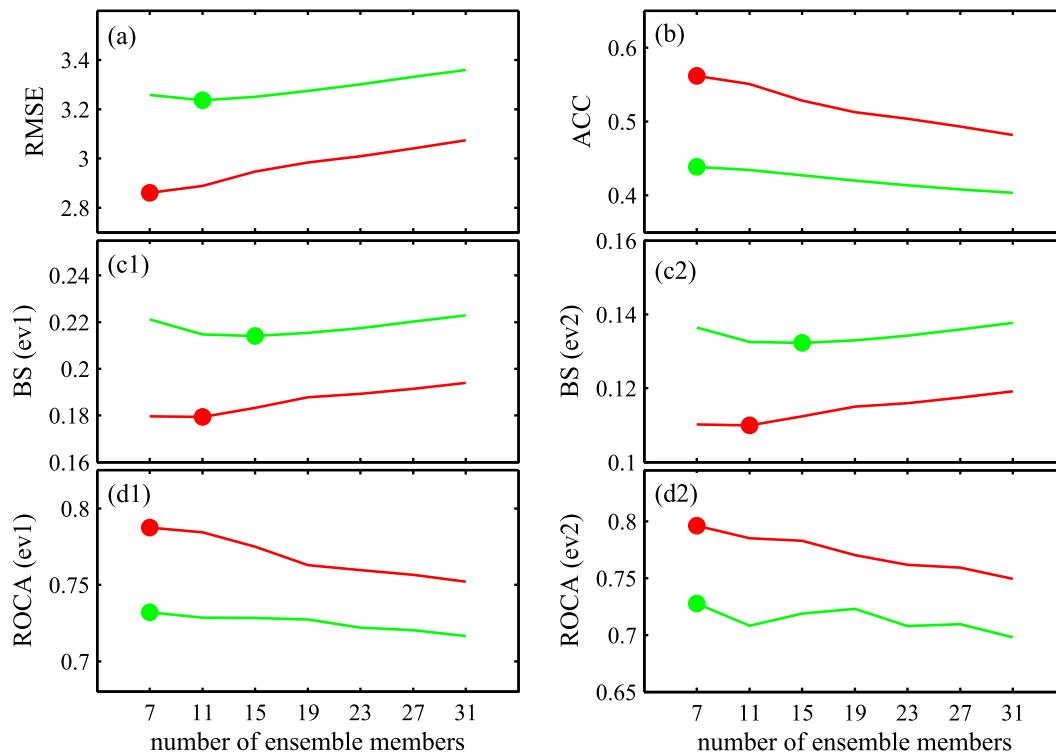
FIG. 8. As in Fig. 7, but with the related control forecast generated by CNOP-type analysis errors, where the magnitudes of the CNOP-type analysis errors are the same as those of the 4DVAR-type errors and the optimization time period is 2 days (also see section 5b).

forecasts associated with orthogonal CNOPs in terms of RMSE and ACC is achieved when the number of ensemble members is 23 rather than 27 or 31.

The CNOP- and SV-type analysis errors are similarly used to examine the effect of the ensemble members' number. The results associated with the CNOP-type analysis errors demonstrate that the ensemble forecasts generated by CNOPs always perform better than those generated by SVs regardless of number of ensemble members (Fig. 8). In particular, when the number of ensemble members is only seven, the ensemble forecasts generated by the CNOPs achieve the highest forecast skill in terms of RMSE, ACC, and ROCA measurements, which is much higher than the highest forecast skill that the SVs achieve. For the results associated with the SV-type analysis errors, the ensemble forecasts generated by orthogonal SVs always behave better than those generated by orthogonal CNOPs, regardless of the number of ensemble members, which further validates the conclusion in section 5b: that is, the ensemble forecasts generated by orthogonal SVs achieve higher forecast skill when the initial analysis errors are slow growing. For simplicity, the related figure is omitted here.

In sections 5a, 5b, and 5c, we demonstrate that orthogonal CNOPs are superior to orthogonal SVs in

yielding ensemble initial perturbations for the control forecast with fast-growing analysis errors. In this case, only a small number of ensemble members generated by orthogonal CNOPs are required to achieve higher forecast skill.

## 6. Discussion

We have shown that for fast-growing initial analysis errors, orthogonal CNOPs have higher forecast skill than orthogonal SVs. Several studies showed that fast-growing initial analysis errors are easily influenced by nonlinearity (Mu et al. 2003; Duan and Mu 2006). Mu et al. (2007) and Duan et al. (2009) showed that, whether the initial analysis error grows significantly is dependent on both its spatial structure and the related reference-state events. That is to say, one initial analysis error may be significantly growing for some events but slowly growing for other events. Because the high skill of the ensemble forecasts generated by CNOPs is determined by the dynamically growing behavior of the initial analysis errors, the skill is also dependent on the related reference-state events. Extreme events are much more likely to induce fast growth of initial errors (Mu et al. 2007; Duan et al. 2009), which, combined with the

conclusion that CNOPs possess higher forecast skill than orthogonal SVs for fast-growing initial analysis errors, indicates that the ensemble forecasts generated by CNOPs have higher skill than those generated by SVs in forecasting evolutions of extreme events.

In addition, Duan et al. (2013) demonstrated that the spatial structure of precursory disturbance may indicate the dynamic growth behavior of future events (also see Qin et al. 2013). Therefore, if we can observe the precursor of an event in advance in realistic forecasts, we may estimate whether the event will be strong or weak and then determine which method (CNOPs or SVs) should be chosen to yield the initial perturbations for the ensemble forecasts. For a precursor whose resultant event is weak and whose evolution can be approximately described by a linear dynamic system, orthogonal SVs should be used to generate the ensemble initial perturbations; otherwise, orthogonal CNOPs should be selected to yield the ensemble initial perturbations, yielding a higher skill.

## 7. Summary

In this paper, we extend orthogonal SVs to the nonlinear regime and propose the concept of orthogonal CNOPs. Orthogonal CNOPs describe a group of orthogonal initial perturbations, each of which represents the initial error that causes the largest prediction error at the prediction time in the related initial perturbation subspace. We use orthogonal CNOPs to yield ensemble members and investigate their role in improving forecast skill by comparing the related ensemble forecast skill with that using orthogonal SVs to generate the initial perturbations, in which the Lorenz-96 model is used as a platform for the ensemble forecast experiments.

Three types of initial analysis error are used. The first is obtained through the 4DVAR approach and is referred to as "4DVAR-type analysis errors." The related results show that the ensemble forecast skill associated with orthogonal CNOPs is statistically higher than that associated with orthogonal SVs. Further investigation shows that orthogonal CNOPs are more applicable than orthogonal SVs for yielding mutually orthogonal initial perturbations for ensemble forecasts when the initial analysis errors are fast growing. To further address this issue, the CNOP errors and the leading SV errors of the truth runs are continuously used as the second and third initial analysis errors, referred to as "CNOP-type analysis errors" and "SV-type analysis errors," respectively. The ensemble forecasts generated by orthogonal CNOPs greatly improve the forecast skill when using the CNOP-type analysis errors. However, for the SV-type analysis errors, the related ensemble forecasts generated

by orthogonal SVs are not always of higher forecast skill than those generated by CNOPs. Specifically, for small magnitudes of orthogonal initial perturbations, the evolution is weakly influenced by nonlinearity, causing the ensemble forecasts generated by orthogonal CNOPs to possess almost the same skill as those generated by orthogonal SVs. Nevertheless, for each of the given optimization time periods, as the magnitude of the orthogonal initial perturbations increases, orthogonal CNOPs overestimate the nonlinearity of the growth of SV-type analysis errors and have substantially worse forecast skill than the orthogonal SVs. These comparisons further validate that orthogonal CNOPs are more appropriate than orthogonal SVs in describing the uncertainties in the evolution of fast-growing initial analysis errors. Generally, fast-growing initial analysis errors are related to extreme events (here referred to as strong events); therefore, ensemble forecasts generated by orthogonal CNOPs have higher skill than the orthogonal SVs in forecasting the evolution of such extreme events. We demonstrate that much smaller numbers of ensemble members generated by orthogonal CNOPs are required to achieve the highest skill. For different atmospheric and oceanic phenomena, the numbers of ensemble members to achieve the highest skill are different and depend on the attractor dimension of the related systems (Carrassi et al. 2009). In addition, if one can observe the precursor of an event in advance in realistic forecasts, whether orthogonal CNOPs or SVs should be used to generate the initial perturbations for the ensemble forecast can be determined, and a higher forecast skill for the evolution of the precursor can be expected.

Bowler (2006) compared random initial perturbations (RPs) with SVs using the Lorenz-96 model and demonstrated that, because of the simplicities of the model, the superiority of SVs in the ensemble forecast cannot be revealed. The present study also compared RPs with CNOPs using the Lorenz-96 model. Unfortunately, the result fails to reveal the superiority of CNOPs in ensemble forecasts. We follow Bowler (2006) in proposing that the simplicity of the Lorenz-96 model may be responsible for the result. Toth and Kalnay (1993, 1997) showed that RPs cannot describe the spatial structure of initial analysis errors and fails to develop baroclinic unstable modes in the baroclinically unstable regions in realistic numerical weather forecast models, which indicates the usefulness of the initial perturbations of particular structures in ensemble forecasts. Therefore, to demonstrate the superiority of CNOPs in ensemble forecasting compared to RPs, realistic weather models, such as the Fifth-generation Pennsylvania State University–National Center for Atmospheric Research

Mesoscale Model (MM5; Dudhia 1993) or Weather Research and Forecasting (WRF; Skamarock et al. 2005) Model, should be used. Such studies are under investigation, and satisfactory results have been obtained, which will be reported in a future paper.

Bred vectors (BVs; Toth and Kalnay 1993, 1997) represent another approach that yields ensemble initial perturbations and have been used at the National Centers for Environmental Prediction. Nevertheless, BVs and SVs possess different dynamic characteristics in terms of optimal perturbation growth. So comparisons between ensemble forecasts generated by orthogonal CNOPs and BVs must be performed in future work. In addition, model errors also influence the forecast skill, so the application of ensemble forecast in decreasing model errors should be investigated. Duan and Zhou (2013) proposed the approach of nonlinear forcing singular vectors (NFSVs; also see Duan and Zhao 2015), which describes the model tendency error that causes the largest prediction error at the prediction time. If one considers the orthogonal NFSVs and disturbs the model in mutually independent subspaces of the model tendency perturbations, the forecast errors induced by the model errors are significantly decreased by ensemble forecasts. If one combines orthogonal CNOPs and NFSVs and applies them in ensemble forecasts, the impact of both the initial errors and the model errors would be reduced, and the forecast skill may be significantly increased. To perform this combination, there are several theoretical and technical problems to be addressed. For example, how to obtain orthogonal NFSVs? What is the theoretical basis associated with orthogonal NFSVs? In addition, computations of orthogonal CNOPs and NFSVs are expensive, although the results shown in this paper demonstrate only a small number of ensemble members generated by orthogonal CNOPs are required to achieve the highest forecast skill. Reducing computation costs is important, and effective algorithms should be developed to calculate the orthogonal CNOPs and NFSVs. It is expected that ensemble forecasts can be improved by the application of orthogonal CNOPs and NFSVs.

## APPENDIX A

### Root-Mean-Square Error

The RMSE is commonly adopted to measure the difference between the forecast (here, the mean of ensemble members; i.e., the ensemble mean) and the observation. The smaller the RMSE is, the more accurate the ensemble mean. The RMSE is calculated by the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{m} \sum_{i=1}^{m} (y_i - o_i)^2}, \qquad \text{(A1)}$$

where $m$ is the number of spatial grids, and $y_i$ and $o_i$ are the forecast value and observation at grid $i$, respectively.

## APPENDIX B

### Anomaly Correlation Coefficient

The ACC is one of the most widely used measures in the verification of spatial fields (Murphy and Epstein 1989). ACC is the correlation between forecasts (here, the ensemble mean of the forecast results) and observed anomalies [as shown in Eqs. (B1) and (B2)]. The larger the ACC is, the higher the forecast skill:

$$\text{ACC} = \frac{\sum_{i=1}^{m} (y_i' - \overline{y}')(o_i' - \overline{o}')}{\sqrt{\sum_{i=1}^{m} (y_i' - \overline{y}')^2} \sqrt{\sum_{i=1}^{m} (o_i' - \overline{o}')^2}}, \qquad \text{(B1)}$$

with

$$y_i' = y_i - c_i \quad \text{and} \quad o_i' = o_i - c_i, \qquad \text{(B2)}$$

where for grid $i$, $y_i$ and $o_i$ are the forecast value and observation, respectively; $c_i$ is the climatological state; $y_i'$ and $o_i'$ are the forecast anomaly and observed anomaly, respectively; and $\overline{y}'$ and $\overline{o}'$ are the spatial mean of the forecast and observed anomaly field, respectively.

## APPENDIX C

### Brier Score

The BS is the mean square error of the probability forecasts (Brier 1950) with the representation [Eq. (C1)]. BS comprehensively evaluates the forecast reliability, the forecast resolution, and the observational uncertainty of the ensemble forecasting system for

probabilistic prediction of the occurrence of a binary event. BS is negatively oriented (i.e., it gives smaller values for better probability forecasts):

$$\text{BS} = \frac{1}{N} \sum_{i=1}^{N} (f_i - o_i)^2, \tag{C1}$$

where $N$ is the number of realizations of the prediction process, and $f_i$ and $o_i$ are the forecast and observational probability for the $i$th prediction process, respectively. The observational probability $o_i$ is equal to 1 or 0 depending on whether the binary event has been observed to occur.

# APPENDIX D

## Relative Operating Characteristic Area

The ROCA [i.e., the area under the relative operating characteristic (ROC) curve (Mason 1982)], is used to represent the forecast skill according to a contingency table. By considering whether an event happens at every grid and checking the forecasts with the truth (or the observations), the result is one of the following outcomes: a hit (i.e., an event occurred and a warning was provided); a false alarm (i.e., an event did not occur but a warning was given); a miss (i.e., an event occurred but a warning was not given); or a correct rejection (i.e., an event did not occur and a warning was not given). A two-by-two contingency table can then be generated, as illustrated in Table D1.

In Table D1, $X$ is the number of hits, $Y$ is the number of misses, $Z$ is the number of false alarms, and $W$ is the number of correct rejections. The hit rate $H$ and the false alarm rate $F$ can be represented as

$$H = \frac{X}{X+Y} \quad \text{and} \quad F = \frac{Z}{Z+W}. \tag{D1}$$

Different values of $H$ and $F$ can be obtained according to different probability thresholds, which means that an event is deemed to occur when the probability forecast is not smaller than the given threshold. The ROC curve can then be drawn with paired values of $(F, H)$ related to different thresholds. ROCA can then be calculated as follows:

$$\begin{aligned}
\text{ROCA} &= \int_0^1 H(x) \, dx \\
&= \sum_{i=1}^{M} \frac{1}{2}(H_{i+1} + H_i)(F_{i+1} - F_i),
\end{aligned} \tag{D2}$$

where $M$ is the number of categories relative to the probability threshold. ROCA is positively oriented (i.e.,

TABLE D1. Two-by-two contingency table for verification of a binary event.

| Observation | Forecast | |
| --- | --- | --- |
| | Occur | Not occur |
| Occur | $X$ | $Y$ |
| Not occur | $Z$ | $W$ |

it gives larger values for better forecasts). Generally, forecasts are skillful when the value of ROCA is greater than 0.5.

# APPENDIX E

## Similarity Coefficient

The similarity coefficient is defined as follows:

$$S = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \, \|\mathbf{y}\|}, \tag{E1}$$

where $S$ is the similarity coefficient between the vectors $\mathbf{x}$ and $\mathbf{y}$, $\langle \cdot, \cdot \rangle$ is the Euclidean inner product, and $\|\cdot\|$ is the L2 norm.

# 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.

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.

Basnarkov, L., and L. Kocarev, 2012: Forecast improvement in Lorenz 96 system. *Nonlinear Processes Geophys.*, **19**, 569–575, doi:10.5194/npg-19-569-2012.

Birgin, E. G., J. M. Martinez, and M. Raydan, 2000: Nonmonotone spectral projected gradient methods on convex sets. *SIAM J. Optim.*, **10**, 1196–1211, doi:10.1137/S1052623497330963.

Bowler, N. E., 2006: Comparison of error breeding, singular vectors, random perturbations and ensemble Kalman filter perturbation strategies on a simple model. *Tellus*, **58A**, 538–548, doi:10.1111/j.1600-0870.2006.00197.x.

Brier, G. W., 1950: Verification of forecasts expressed in terms of probabilities. *Mon. Wea. Rev.*, **78**, 1–3, doi:10.1175/1520-0493(1950)078<0001:VOFEIT>2.0.CO;2.

Carrassi, A., S. Vannitsem, D. Zupanski, and M. Zupanski, 2009: The maximum likelihood ensemble filter performances in chaotic systems. *Tellus*, **61A**, 587–600, doi:10.1111/j.1600-0870.2009.00408.x.

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.

Dijkstra, H. A., and J. P. Viebahn, 2015: Sensitivity and resilience of the climate system: A conditional nonlinear optimization approach. *Commun. Nonlinear Sci. Numer. Simul.*, **22**, 13–22, doi:10.1016/j.cnsns.2014.09.015.

Duan, W. S., and M. Mu, 2006: Investigating decadal variability of El Niño–Southern Oscillation asymmetry by conditional nonlinear optimal perturbation. *J. Geophys. Res.*, **111**, C07015, doi:10.1029/2005JC003458.

——, and ——, 2009: Conditional nonlinear optimal perturbation: Applications to stability, sensitivity, and predictability. *Sci. China*, **52D**, 883–906, doi:10.1007/s11430-009-0090-3.

——, and F. F. Zhou, 2013: Non-linear forcing singular vector of a two-dimensional quasi-geostrophic model. *Tellus*, **65A**, 18 452, doi:10.3402/tellusa.v65i0.18452.

——, and P. Zhao, 2015: Revealing the most disturbing tendency error of Zebiak–Cane model associated with El Niño predictions by nonlinear forcing singular vector approach. *Climate Dyn.*, **44**, 2351–2367, doi:10.1007/s00382-014-2369-0.

——, M. Mu, and B. Wang, 2004: Conditional nonlinear optimal perturbation as the optimal precursors for El Niño–Southern Oscillation events. *J. Geophys. Res.*, **109**, D23105, doi:10.1029/2004JD004756.

——, X. C. Liu, K. Y. Zhu, and M. Mu, 2009: Exploring the initial errors that cause a significant ''spring predictability barrier'' for El Niño events. *J. Geophys. Res.*, **114**, C04022, doi:10.1029/2008JC004925.

——, Y. S. Yu, H. Xu, and P. Zhao, 2013: Behaviors of non-linearities modulating the El Niño events induced by optimal precursory disturbances. *Climate Dyn.*, **40**, 1399–1413, doi:10.1007/s00382-012-1557-z.

Dudhia, J., 1993: A nonhydrostatic version of the Penn State–NCAR Mesoscale Model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, **121**, 1493–1513, doi:10.1175/1520-0493(1993)121<1493:ANVOTP>2.0.CO;2.

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

Feng, J., R. Q. Ding, D. Q. Liu, and J. P. Li, 2014: The application of nonlinear local Lyapunov vectors to ensemble predictions in Lorenz systems. *J. Atmos. Sci.*, **71**, 3554–3567, doi:10.1175/JAS-D-13-0270.1.

Fertig, E. J., J. Harlim, and B. R. Hunt, 2007: A comparative study of 4D-VAR and a 4D Ensemble Kalman Filter: Perfect model simulations with Lorenz-96. *Tellus*, **59A**, 96–100, doi:10.1111/j.1600-0870.2006.00205.x.

Gilmour, I., and L. A. Smith, 1997: Enlightenment in shadows. *Applied Nonlinear Dynamics and Stochastic Systems near the Millennium*, J. B. Kadtke and A. Bulsara, Eds., American Institute of Physics, 335–340.

Hunt, B. R., and Coauthors, 2004: Four-dimensional ensemble Kalman filtering. *Tellus*, **56A**, 273–277, doi:10.1111/j.1600-0870.2004.00066.x.

Jiang, Z. N., and M. Mu, 2009: A comparisons study of the methods of conditional nonlinear optimal perturbations and singular vectors in ensemble prediction. *Adv. Atmos. Sci.*, **26**, 465–470, doi:10.1007/s00376-009-0465-6.

——, ——, and D. H. Wang, 2008: Conditional nonlinear optimal perturbation of a T21L3 quasi-geostrophic model. *Quart. J. Roy. Meteor. Soc.*, **134**, 1027–1038, doi:10.1002/qj.256.

Khare, S. P., and J. L. Anderson, 2006: An examination of ensemble filter based adaptive observation methodologies. *Tellus*, **58A**, 179–195, doi:10.1111/j.1600-0870.2006.00163.x.

Koyama, H., and M. Watanabe, 2010: Reducing forecast errors due to model imperfections using ensemble Kalman filtering. *Mon. Wea. Rev.*, **138**, 3316–3332, doi:10.1175/2010MWR3067.1.

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

Leutbecher, M., and T. N. Palmer, 2008: Ensemble forecasting. *J. Comput. Phys.*, **227**, 3515–3539, doi:10.1016/j.jcp.2007.02.014.

Li, S., X. Y. Rong, Y. Liu, Z. Y. Liu, and K. Freadrich, 2013: Dynamic analogue initialization for ensemble forecasting. *Adv. Atmos. Sci.*, **30**, 1406–1420, doi:10.1007/s00376-012-2244-z.

Liu, D. C., and J. Nocedal, 1989: On the limited memory BFGS method for large scale optimization. *Math. Program.*, **45B**, 503–528, doi:10.1007/BF01589116.

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

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

——, 1996: Predictability: A problem partly solved. *Proc. Workshop on Predictability*, Reading, United Kingdom, ECMWF, 18 pp.

——, 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.

Mason, I., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30**, 291–303.

Molteni, F., 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.

Mu, M., and Z. Y. Zhang, 2006: Conditional nonlinear optimal perturbation of a two-dimensional quasigeostrophic model. *J. Atmos. Sci.*, **63**, 1587–1604, doi:10.1175/JAS3703.1.

——, and Z. N. Jiang, 2008a: A method to find perturbations that trigger blocking onset: Conditional nonlinear optimal perturbation. *J. Atmos. Sci.*, **65**, 3935–3946, doi:10.1175/2008JAS2621.1.

——, and ——, 2008b: A new approach to the generation of initial perturbations for ensemble prediction: Conditional nonlinear optimal perturbation. *Chin. Sci. Bull.*, **53**, 2062–2068, doi:10.1007/s11434-008-0272-y.

——, W. S. Duan, and B. Wang, 2003: Conditional nonlinear optimal perturbation and its application. *Nonlinear Processes Geophys.*, **10**, 493–501, doi:10.5194/npg-10-493-2003.

——, H. Xu, and W. S. Duan, 2007: A kind of initial errors related to ''spring predictability barrier'' for El Niño events in Zebiak–Cane model. *Geophys. Res. Lett.*, **34**, L03709, doi:10.1029/2006GL027412.

Mureau, R., F. Molteni, and T. N. Palmer, 1993: Ensemble prediction using dynamically conditional perturbations. *Quart. J. Roy. Meteor. Soc.*, **119**, 299–323, doi:10.1002/qj.49711951005.

Murphy, A. H., and E. S. Epstein, 1989: Skill scores and correlation coefficients in model verification. *Mon. Wea. Rev.*, **117**, 572–581, doi:10.1175/1520-0493(1989)117<0572:SSACCI>2.0.CO;2.

Ott, E., and Coauthors, 2004: A local ensemble Kalman filter for atmospheric data assimilation. *Tellus*, **56A**, 415–428, doi:10.1111/j.1600-0870.2004.00076.x.

Powell, M. J. D., 1983: VMCWD: A Fortran subroutine for constrained optimization. *ACM SIGMAP Bull.*, No. 32, Association for Computing Machinery, New York, NY, 4–16, doi:10.1145/1111272.1111273.

Qin, X. H., W. S. Duan, and M. Mu, 2013: Conditions under which CNOP sensitivity is valid for tropical cyclone adaptive

observations. *Quart. J. Roy. Meteor. Soc.*, **139**, 1544–1554, doi:10.1002/qj.2109.

Revelli, J. A., M. A. Rodriguez, and H. S. Wio, 2010: The use of Rank Histograms and MVL diagrams to characterize ensemble evolution in weather forecasting. *Adv. Atmos. Sci.*, **27**, 1425–1437, doi:10.1007/s00376-009-9153-6.

Roulston, M. S., and L. A. Smith, 2003: Combining dynamical and statistical ensembles. *Tellus*, **55A**, 16–30, doi:10.1034/j.1600-0870.2003.201378.x.

Skamarock, W. C., J. B. Klemp, J. Duhia, D. O. Gill, D. M. Barker, W. Wang, and J. G. Powers, 2005: A description of the Advanced Research WRF version 2. NCAR Tech. Note NCAR/TN-468+STR, 88 pp. [Available online at http://www2.mmm.ucar.edu/wrf/users/docs/arw_v2_070111.pdf.]

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 and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319, doi:10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, doi:10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

Zhou, F. F., and M. Mu, 2011: The impact of verification area design on tropical cyclone targeted observations based on the CNOP method. *Adv. Atmos. Sci.*, **28**, 997–1010, doi:10.1007/s00376-011-0120-x.