Using the Orthogonal Conditional Nonlinear Optimal Perturbations Approach to Address the Uncertainties of Tropical Cyclone Track Forecasts Generated by the WRF Model

HAN ZHANG,[a,b] WANSUO DUAN,[a,b] AND YICHI ZHANG[a,b]

[a] *LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*
[b] *University of Chinese Academy of Sciences, Beijing, China*

ABSTRACT: The orthogonal conditional nonlinear optimal perturbations (O-CNOPs) approach for measuring initial uncertainties is applied to the Weather Research and Forecasting (WRF) Model to provide skillful forecasts of tropical cyclone (TC) tracks. The hindcasts for 10 TCs selected from 2005 to 2020 show that the ensembles generated by the O-CNOPs have a greater probability of capturing the true TC tracks, and the corresponding ensemble forecasts significantly outperform the forecasts made by the singular vectors, bred vectors, and random perturbations in terms of both deterministic and probabilistic skills. In particular, for two unusual TCs, Megi (2010) and Tembin (2012), the ensembles generated by the O-CNOPs successfully reproduce the sharp northward-turning track in the former and the counterclockwise loop track in the latter, while the ensembles generated by the other methods fail to do so. Moreover, additional attempts are performed on the real-time forecasts of TCs In-Fa (2021) and Hinnamnor (2022), and it is shown that O-CNOPs are very useful for improving the accuracy of real-time TC track forecasts. Therefore, O-CNOPs, together with the WRF Model, could provide a new platform for the ensemble forecasting of TC tracks with much higher skill.

KEYWORDS: Tropical cyclones; Uncertainty; Ensembles; Probability forecasts/models/distribution; Nonlinear models

1. Introduction

Tropical cyclones (TCs), a type of high-impact weather event occurring in the western North Pacific basin (WNP), cause great losses in terms of human lives and property every year from high winds, heavy rainfall, and storm surges along coastlines. Improving the accuracy of TC forecasts helps to minimize property damage and casualties from TCs. One of the most important aspects of predicting a TC is its future track (Wang 2014); furthermore, accurate track forecasts would also help to improve other aspects of TC forecasts, such as wind speed and precipitation (Majumdar and Finocchio 2010).

In recent decades, TC track forecast errors, especially for short-range forecasts, have substantially decreased due to progress in numerical weather prediction models and the greater number of observations obtained by satellites and aircrafts (Goerss et al. 2004; Rappaport et al. 2009). However, providing more skillful medium-range (approximately 5 days) track forecasts to the public for the early warning of TCs is also essential. In addition, although the forecast errors of TC tracks have been reduced on average, there are still cases that have very large forecast errors, particularly for some unusual cases (Puri et al. 2001; Yamaguchi and Majumdar 2010; Titley et al. 2020). There is potential for further reducing the annual mean of TC track errors by reducing the track errors of these particular cases. Large track errors are often related to initial uncertainties in the TC steering flow associated with nearby synoptic-scale systems, such as the subtropical high (SH) and midlatitude troughs (Yamaguchi and Majumdar 2010; Torn

et al. 2018; Magnusson et al. 2019). In particular, Miyachi and Enomoto (2021) noted that the prediction of the steering flow is sensitive to initial conditions and that the uncertainties significantly affect the forecasting skills of the recurvature tracks of TCs. There are inherent difficulties in accurately forecasting TC tracks by a single deterministic forecast given the inevitable initial errors, which motivates the use of ensemble forecasts to describe and reduce the forecast uncertainties associated with initial uncertainties. Most operational centers have built and developed their ensemble prediction systems to improve forecast skill, provide uncertainty information and deliver probability forecasts for TC tracks (Puri et al. 2001; Yamaguchi et al. 2009; Dube et al. 2020).

Currently, several perturbation methods have been developed to represent the initial uncertainties. Singular vectors (SVs) and bred vectors (BVs) (Lorenz 1965; Epstein 1969; Leith 1974) are two popular methods. They both aim to properly sample the fastest-growing components of the initial errors, which are believed to be potentially responsible for a large part of forecast uncertainties (Molteni et al. 1996; Magnusson et al. 2008; Diaconescu and Laprise 2012). The SVs method has been successfully applied at the European Centre for Medium-Range Weather Forecasts (ECMWF) to generate initial perturbations (Molteni and Palmer 1993; Molteni et al. 1996). Benefiting from the application of target SVs on TCs, the track forecast skill of the ECMWF has been improved (Puri et al. 2001; Yamaguchi and Majumdar 2010; Titley et al. 2020). Under the assumption that perturbations grow linearly, SVs are able to capture a set of orthogonal initial perturbations that have the largest growth rates within a finite time interval (Lorenz 1965). However, SVs sometimes cannot capture the most unstable directions in nonlinear models, resulting in an

underestimation of the forecast uncertainty (Anderson 1997; Hamill et al. 2000). For example, Buizza (1994) noted that some leading SVs represent spurious modes that have a large growth rate in the linear model but decay quickly in the nonlinear model. Puri et al. (2001) also observed that similar modes appeared when computing tropical SVs and target SVs for TCs. Ehrendorfer et al. (1999) noted that the tangent linear assumption is only qualitatively accurate for small-scale SVs structures in regions dominated by moist convection. The contradiction in which the evolution of initial errors is highly nonlinear while SVs are based on linear hypotheses has limited the role of SVs in improving ensemble forecasting skill.

To overcome the linear limitation of SVs, Mu et al. (2003) proposed the conditional nonlinear optimal perturbation (CNOP) method. The CNOP method is a natural extension of the leading SV (LSV) method into the nonlinear regime and represents the initial perturbation having the largest nonlinear evolution within a finite time interval (Duan et al. 2004; Duan and Mu 2009). The CNOP method has been successfully applied to predictability studies of high-impact weather and climate events, such as El Niño–Southern Oscillation (ENSO; Duan et al. 2018), the Indian Ocean dipole (IOD; Feng and Duan 2014), and TCs (Mu et al. 2009; Qin et al. 2013). To consider the influence of nonlinearity on ensemble forecasts, Jiang et al. (2009) replaced LSVs with CNOPs to yield initial perturbations in a two-dimensional quasigeostrophic model and consequently made the forecast quality better than the SVs method. This indicates the benefits from the initial perturbations that describe the nonlinearly fast-growing component of the initial errors of the ensemble forecast. However, such an approach is still insufficient to estimate forecast uncertainties in nonlinear models because the other SVs do not consider the effect of nonlinearity.

To fully consider the influence of nonlinearity, Duan and Huo (2016) further proposed the orthogonal conditional nonlinear optimal perturbations (O-CNOPs) method. The O-CNOPs generalize SVs to fully nonlinear fields and represent a group of mutually independent initial perturbations that have the maximum nonlinear evolution in their respective subphase spaces within a finite time interval (Duan and Huo 2016; Huo et al. 2019; Wang and Duan 2019). These characteristics illustrate the potential of the O-CNOPs method to generate initial perturbations for ensemble forecasts. Duan and Huo (2016) first applied the O-CNOP method to a simple Lorenz-96 model (Lorenz 1996) to perturb the initial conditions, leading to a much higher forecast skill than the SVs method. Huo et al. (2019) then adopted O-CNOPs to conduct ensemble forecast experiments for TC tracks using the fifth-generation Pennsylvania State University, National Center for Atmospheric Research Mesoscale Model (MM5) and showed advantages over the forecasts made by the random perturbation (RP), BV, and SV methods. However, this previous study did not examine the performance of O-CNOPs for the forecasting of unusual TC tracks, which still poses challenges for operational forecasts; furthermore, O-CNOPs have not yet been used in real-time TC forecasting. In addition, the MM5 has been thought of as falling behind the ranks and performs much worse than the Weather Research and Forecasting (WRF) model in simulations of TC tracks

(Pattanayak and Mohanty 2008). Hence, further applying the O-CNOPs method to the WRF Model is of more practical significance for improving the forecasts of TC tracks. The CNOP have been applied in the WRF version 3.6 (WRFV3.6) model to identify the sensitive area for targeted observations associated with TC forecasts in a real-time field campaign, and when assimilating the corresponding targeted data, unanimously positive effects on the TC track forecasting skill have been achieved (Qin et al. 2022). This useful sensitivity revealed by the CNOP is also the requirement of ensemble forecasting with respect to initial perturbations and therefore encourages us to use the WRFV3.6 model to examine the usefulness of O-CNOPs in ensemble forecasts for the first time. Therefore, we naturally ask the following: can O-CNOPs still provide improvements in the forecasting skill of TC tracks? To answer this question, we conducted ensemble forecasting experiments to provide both deterministic and probabilistic forecasts for TC tracks in the present study and evaluate the performance of the O-CNOPs method in forecasting TC tracks, particularly concerning unusual TCs. Moreover, we implemented the O-CNOPs method in the real-time forecasts of TC In-Fa (2021) and TC Hinnamnor (2022), and in the present study, we also report these forecasts and indicate the potential of O-CNOPs for enhancing the accuracy of real-time TC track forecasts in the future.

The remainder of the paper is organized as follows. Section 2 introduces the TC cases, model configuration, and ensemble forecast methods, and section 3 presents the O-CNOPs configuration for the ensemble forecasting of TC tracks. Section 4 applies O-CNOPs in the hindcasts of strong TCs during 2018–20, hindcasts of the two unusual TC tracks and real-time forecasts of TCs IN-FA (2021) and Hinnamnor (2022). In section 5, the performance of the O-CNOPs method is statistically evaluated by comparison with the SV, BV, and RP methods, and in section 6, the performance of the O-CNOPs for determining TC track forecasts in the MM5 and WRF Model is discussed. Finally, a summary is provided in section 7.

2. TC cases, model, and methods

a. Case overview

Six TC cases that had great impact on China from 2018 to 2020 and reached the strong typhoon (STY) or super typhoon (Super TY) levels are selected for testing O-CNOPs, and two others, STY Matsa (2005) and Super TY Sepat (2007), which were the only two strong TC cases adopted by Huo et al. (2019) for examining the application of O-CNOPs in the MM5, are also selected to compare the ensemble forecasting skills in the MM5 and WRF Model. Among these TCs, five [i.e., Matsa (2005), Sepat (2007), Maria (2018), Mangkhut (2018), and Lekima (2019)] tended to move northwestward or westward toward land in the coastal areas of southeastern China, while the other three [i.e., Bavi (2020), Maysak (2020), and Haishen (2020)] moved straight northward and influenced Northeast China.

Super TY Megi (2010) and STY Tembin (2012) are two additional TC cases. They exhibited unusual movements during their passages (see Fig. 1), and their operational forecasts
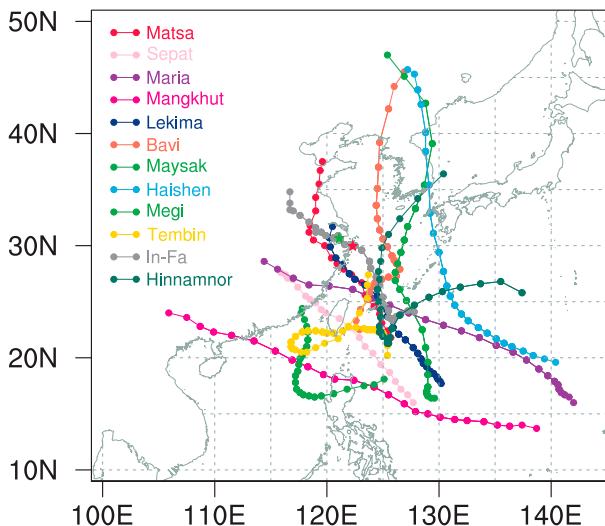
FIG. 1. The best tracks obtained from the China Meteorological Administration (CMA) for the 12 selected TCs in this study. The position at the initialized time of the first forecast period of TC forecasts is used as the beginning of the track, and its subsequent track is marked by dots every 6 h. The red and green stars represent the first and second landfall points of TC In-Fa (2021).

were challenging (Qian et al. 2013; Tallapragada et al. 2015; Moon and Ha 2019). Megi (2010) moved westward into the South China Sea from 13 to 19 July 2010, abruptly turned almost 90° to the north at 0000 UTC 20 October, and subsequently made landfall on the coast of Fujian Province, China, at 0455 UTC 23 October (Qian et al. 2013). For Tembin (2012), its erratic track appeared to be associated with binary interactions with TC Bolaven (2012) when they were close to each other. Specifically, Tembin (2012) moved westward from 22 August 2012 and entered the northern part of the South China Sea on 24 August; then, it executed a slow counterclockwise loop in the following 2 days and moved northeastward after 27 August (Liu et al. 2015).

In addition to the above 10 TCs, we made additional attempts to forecast STY In-Fa (2021) and Super TY Hinnamnor (2022) in real time. In-Fa (2021) had the longest-lasting effect on the Chinese mainland on record. It first made landfall in Zhoushan (on the east coast of China's Zhejiang Province) at 0430 UTC 25 July 2021 and then made landfall in Zhejiang at 0150 UTC 26 July; afterward, In-Fa (2021) continued to move northwestward slowly, lingered for a long time on land and brought continuous precipitation to China. Hinnamnor (2022) had an unusual V-shaped track during the period from 28 August to 6 September 2022. At the early stage, Hinnamnor (2022), with extreme wind, steadily moved southwestward to approach the east coast of China. Suddenly, Hinnamnor (2022) turned to the northeast on 2 September and did not directly make landfall in China but brought heavy rain and wind to China's coastal areas. Subsequently, Hinnamnor (2022) continued to move northeastward and made landfall in South Korea on 6 September 2022.

In summary, a total of 12 TCs are investigated in the present study to evaluate the performance of the O-CNOPs in

TABLE 1. List of TC names, initialized times (UTC), and end times of 29 forecasts for 12 TCs carried out in this study.

TC names	Initialized time of forecasts	End time of forecasts
Matsa	1200 UTC 3 Aug 2005	1200 UTC 8 Aug 2005
Sepat	1200 UTC 15 Aug 2007	1200 UTC 20 Aug 2007
Maria	0000 UTC 6 May 2018	0000 UTC 11 May 2018
	1200 UTC 6 May 2018	1200 UTC 11 May 2018
	0000 UTC 7 May 2018	0000 UTC 12 May 2018
Mangkhut	1200 UTC 11 Sep 2018	1200 UTC 16 Sep 2018
	0000 UTC 12 Sep 2018	0000 UTC 17 Sep 2018
	1200 UTC 12 Sep 2018	1200 UTC 17 Sep 2018
Lekima	1800 UTC 4 Aug 2019	1800 UTC 9 Aug 2019
	0600 UTC 5 Aug 2019	0600 UTC 10 Aug 2019
	1800 UTC 5 Aug 2019	1800 UTC 10 Aug 2019
Bavi	1800 UTC 21 Aug 2020	1800 UTC 26 Aug 2020
	0600 UTC 22 Aug 2020	0600 UTC 27 Aug 2020
	1800 UTC 22 Aug 2020	1800 UTC 27 Aug 2020
Maysak	1800 UTC 28 Aug 2020	1800 UTC 2 Sep 2020
	0600 UTC 29 Aug 2020	0600 UTC 3 Sep 2020
	1800 UTC 29 Aug 2020	1800 UTC 3 Sep 2020
Haishen	1200 UTC 2 Sep 2020	1200 UTC 7 Sep 2020
	0000 UTC 3 Sep 2020	0000 UTC 8 Sep 2020
	1200 UTC 3 Sep 2020	1200 UTC 8 Sep 2020
Megi	1200 UTC 17 Oct 2010	1200 UTC 23 Oct 2010
Tembin	0000 UTC 21 Aug 2012	0000 UTC 29 Aug 2012
	0000 UTC 22 Aug 2012	0000 UTC 29 Aug 2012
	0000 UTC 23 Aug 2012	0000 UTC 29 Aug 2012
In-Fa	0000 UTC 21 May 2021	0000 UTC 26 May 2021
	0000 UTC 22 May 2021	0000 UTC 27 May 2021
	0000 UTC 23 May 2021	0000 UTC 28 May 2021
	0000 UTC 24 May 2021	0000 UTC 29 May 2021
Hinnamnor	0000 UTC 30 Aug 2022	0000 UTC 6 Sep 2022

the ensemble forecasts of TC tracks. The forecast periods are chosen to be 5 days for the TC cases to cover their landfall processes except that the forecast periods of Megi (2010) and Hinnamnor (2022) are 6 and 7 days, respectively, because this much better captures not only their sharp turn but also their landfall after turning, and even longer forecast periods are chosen for Tembin (2012) to cover its complete loop track. Then, a total of 29 forecast periods are determined. The specific initial and end times of the forecast periods are shown in Table 1, and the best tracks obtained from the China Meteorological Administration (CMA) of these TCs are plotted in Fig. 1.

b. Model configuration

As mentioned in the introduction, we use the WRFV3.6 model in the present study. The model domain is configured with 151 × 96 model grid points in the west–east and south–north directions and, similar to Huo et al. (2019), has a horizontal grid spacing of 60 km, which makes it easy to compare the results in the present study with those in Huo et al. (2019). The simulation area covers the region of the WNP and South China Sea. A coarser vertical resolution of 15 vertical levels with the top level at 50 hPa is experimentally adopted for less use of computing resources. The nonlinear physics parameterizations

selected in the WRF Model include the Lin microphysics scheme (Lin et al. 1983), Kain–Fritch cumulus parameterization scheme (Kain 2004), Dudhia shortwave radiation scheme (Dudhia 1989), and Rapid Radiative Transfer Model longwave radiation scheme (Mlawer et al. 1997). The WRFV3.6, together with its adjoint model, is used to calculate the O-CNOPs (see section 2c), where only three simplified physics packages, surface friction, cumulus parameterization, and large-scale condensation, are available for the adjoint model (Xiao et al. 2008; Zhang et al. 2013). This configuration of the WRF and its adjoint model have been widely used in studies of data assimilation or ensemble forecasting.

To hindcast the 10 TCs (see section 2a), the analysis data derived from the National Centers for Environment Prediction final operational global analysis (NCEP FNL) dataset are adopted to provide the initial and boundary conditions, where the horizontal resolution is $1° \times 1°$, and the time interval is 6 h. It is noted that the present study focuses on the performance of the O-CNOPs and mainly investigates the effect of the initial uncertainties on TC track forecasts; thus, a more accurate analysis boundary condition could highlight the role of the initial conditions in the TC track forecasts. Strictly speaking, these forecasts are not in retrospective circumstances and very unrealistic as compared with those that use forecast boundary conditions. However, in contrast with the subsequent real-time forecasts, we still call them "hindcasts." In fact, when we conduct the real-time forecasts of the two additional TC cases of STY In-Fa (2021) and Super TY Hinnamnor (2022), we have to adopt more realistic forecast data to provide initial and boundary conditions; specifically, we take the data from the NCEP Global Forecast System (GFS) forecast dataset with a horizontal resolution of $0.25° \times 0.25°$ and a time interval of 6 h. In this case, we investigate the role of having a much more accurate estimation of the initial uncertainties in improving the ensemble forecast skill under the existence of large boundary condition errors, which indicates that the O-CNOPs method is examined under more realistic circumstances. Finally, for the verification of the TC track forecasts, the best track dataset from the CMA (Lu et al. 2021) is used in this study.

c. The O-CNOPs method

The O-CNOPs method (Duan and Huo 2016) is adopted to perturb the initial conditions of the control forecast generated by the WRF Model for the ensemble forecast experiments of the TC track. The resultant forecast skills are compared with those of three traditional methods: the RPs and SVs methods, as mentioned in the introduction, and the orthogonal BVs method, which is similar to the orthogonal nonlinear local Lyapunov vectors proposed by Feng et al. (2016, 2018). This subsection gives a brief overview of the O-CNOPs method adopted in this study, while descriptions of the traditional methods are given in appendix A.

Consider a nonlinear model M acting on an N-dimensional state vector $\mathbf{X}_0 \in \mathbb{R}^n$, such that $\mathbf{X}_\tau = M(\mathbf{X}_0)$, where the subscript τ refers to the integration time, and its corresponding time interval $[0, \tau]$ is named the optimization time interval

(OTI) for O-CNOPs. Let $\delta\mathbf{X}_0$ and $\delta\mathbf{X}_\tau$ represent the initial and final perturbation states, respectively; then, $\mathbf{X}_\tau + \delta\mathbf{X}_\tau = M(\mathbf{X}_0 + \delta\mathbf{X}_0)$. If two norms, $\mathbf{C}_1$ and $\mathbf{C}_2$, are chosen to measure the amplitude of the perturbations at the initial and final times, the O-CNOPs can be formulated by Eq. (1), where the O-CNOPs identify a group of orthogonal initial perturbations, i.e., 1st-CNOP, 2nd-CNOP, 3rd-CNOP, …, jth-CNOP … , etc. The jth-CNOP belongs to the subspace Ω_j and is referred to as the initial perturbation $\mathbf{x}_{0j}^*$ that satisfies the constraint in Eq. (2) and has the largest nonlinear evolution at time τ (Duan and Huo 2016):

$$
J(\mathbf{x}_{0j}^*) = \max_{\mathbf{x}_{0j}\in\Omega_j} [PM_\tau(\mathbf{X}_0 + \mathbf{x}_{0j}) - PM_\tau(\mathbf{X}_0)]^T \mathbf{C}_2 [PM_\tau(\mathbf{X}_0 + \mathbf{x}_{0j})
$$
$$
- PM_\tau(\mathbf{X}_0)], \tag{1}
$$

where subspace Ω_j is represented by

$$
\Omega_j = \begin{cases} \{\mathbf{x}_{0j} \in \mathbb{R}^n | \mathbf{x}_{0j}^T \mathbf{C}_1 \mathbf{x}_{0j} \le \delta\}, & j = 1, \\ \{\mathbf{x}_{0j} \in \mathbb{R}^n | \mathbf{x}_{0j}^T \mathbf{C}_1 \mathbf{x}_{0j} \le \delta, \mathbf{x}_{0j} \perp \Omega_k, k = 1, ..., j-1\}, & j > 1, \end{cases}
$$
$$
\tag{2}
$$

where P is a local projection operator, which has a value of 1 when the final perturbations are in the verification region and otherwise, it has a value of 0, the superscript "T" is the transposition symbol, and δ is a positive number that constrains the amplitude of the initial perturbations. Note that the norms $\mathbf{C}_1$ and $\mathbf{C}_2$, local projection operator P, OTI, and initial perturbation amplitude δ need to be determined experimentally according to the problem of interest (see section 3 for TC track forecasts).

Based on the WRF and its adjoint model, we adopt the Spectral Projected Gradient 2 optimization algorithm (Birgin et al. 2000) to compute the O-CNOPs along the fastest-descending direction of the gradient of the cost function in Eq. (1), where the norms ($\mathbf{C}_1$ and $\mathbf{C}_2$ at the initial and optimization times both use the total moist energy (TME, also adopted to compute the RPs, BVs and SVs) and are expressed by Eq. (3):

$$
\|\delta\mathbf{X}\|_{C_1}^2 = \|\delta\mathbf{X}\|_{C_2}^2
$$
$$
= \frac{1}{D}\int_D\int_0^1 \left[\mathbf{u}'^2 + \mathbf{v}'^2 + \left(\frac{g}{\overline{N}\theta}\right)^2\boldsymbol{\theta}'^2 + R_a T_r\left(\frac{\mathbf{p}_s'}{p_r}\right)^2 \right.
$$
$$
\left. + \frac{L^2}{c_p T_r}\mathbf{q}'^2 \right]d\sigma dD, \tag{3}
$$

where $\delta\mathbf{X}$ is composed of $\mathbf{u}'$, $\mathbf{v}'$, $\boldsymbol{\theta}'$, $\mathbf{p}_s'$, and $\mathbf{q}'$, representing the perturbed zonal wind, meridional wind, potential temperature, surface pressure, and water vapor mixing ratio, respectively; the constants g, $\overline{N}$, R_a, L, and c_p are the gravitational acceleration, Brunt–Väisälä frequency, dry air gas constant, latent heat of condensation per unit mass, and specific heat at a constant pressure, respectively; $\overline{\theta} = 300\,\mathrm{K}$, $T_r = 270$ K and $p_r = 1000$ hPa are prescribed physical parameters; D is the horizontal domain; and σ represents the vertical coordinate. The vertical integration of the kinetic energy and potential energy term extends up to the top level, while that of the

water vapor term is limited up to approximately 500 hPa. By employing the local projection operator P, the verification region is determined to cover a $10° \times 10°$ box centered on the TC position at the optimization time, where the boxed region is sufficient to interpret the positional uncertainties of the TC circulation at the optimization time (Tseng and Lai 2020), while the initial perturbation region covers the whole domain, including the TC itself and the environmental flow.

3. Configuration of the O-CNOPs for TC track forecasts generated by the WRF Model

The ensemble forecasting skill, as mentioned in section 2c, is sensitive to the ensemble size N, initial perturbation amplitude δ, and OTI for computing the optimal initial perturbations (Wang and Duan 2019). However, these essential parameters are different for different models and methods (Puri et al. 2001; Yamaguchi and Majumdar 2010), and there is not a universal approach to determine them. In the present study, we design sensitivity experiments to evaluate the sensitivity of the ensemble forecasting skill to these parameters and determine their values for TC track forecasting. The perturbation amplitude δ is experimentally chosen from the range 0.12 to 1.8, which avoids unreasonable perturbation amplitudes that can trigger numerical instabilities and ensures that the initial perturbation amplitudes are physically reasonable; the OTI is restricted so that it does not exceed 36 h and includes 6, 12, 24, and 36 h. Therefore, a total of 28 combinations of δ and OTIs are shown in Table 2. For each combination, the O-CNOPs are calculated for each control forecast and then added to the initial field of the control forecast with a positive/negative perturbation pair. With these perturbed initial fields, the WRF Model is integrated, and a group of perturbed forecasts is obtained, which, together with the control forecasts, are composed of ensemble forecasting members. Next, we finalize the ensemble parameters of ensemble size N, initial perturbation amplitude δ, and OTI, which would help the ensemble forecasts of TC tracks achieve much higher ensemble forecasting skill.

We determine the ensemble parameters by evaluating the ensemble forecasting skills of TCs from both ensemble mean forecasting and TC strike probability forecasting. The ensemble mean is calculated by taking the mean of all ensemble members. The TC strike probability focuses on a spatial location and is defined as the probability that a TC will pass within a 120-km radius from the location during the next 120 h (WMO 2013). We compute the ensemble mean forecasting error of the TC track to evaluate the deterministic forecasting skill, which is defined as the great-circle distance between the best track and the ensemble mean forecasting at a lead time (Heming 2017); for the TC strike probability, we compute the Brier skill score (BSS) to measure the relative skill of the probabilistic forecast over that of the control forecast in terms of forecasting whether an event occurred. Larger BSS values indicate higher probabilistic forecasting skills. The details of these measurements are introduced in appendix B.

We select one TC case from each year during 2018–20 shown in Table 1, and obtain three TC cases of Manghut

TABLE 2. Twenty-eight combinations of initial perturbation amplitudes δ and OTIs.

OTI (h)	Initial perturbation amplitude (J kg^{-1})						
	0.12	0.3	0.6	0.9	1.2	1.5	1.8
6	E1	E2	E3	E4	E5	E6	E7
12	E8	E9	E10	E11	E12	E13	E14
24	E15	E16	E17	E18	E19	E20	E21
36	E22	E23	E24	E25	E26	E27	E28

(2018), Lekima (2019) and Bavi (2020), which present westward, northwestward, and straight northward moving tracks, respectively (see Fig. 1). Additionally, the two TC cases of Matsa (2005) and Sepat (2007) adopted in Huo et al. (2019) are selected. With these five TC cases, we optimize the ensemble parameters by performing the forecasts over the predetermined forecast periods (see Table 1). For each E_i in Table 2, we conduct ensemble forecasting experiments for the predetermined five TC cases by assigning varying OTIs in Table 2 and varying ensemble sizes N (see Fig. 2). We first evaluate the ensemble mean forecasting errors of TC tracks averaged over all lead times in 5 days for all five TC cases and find that similar results are obtained for each value of δ as N increases. For simplicity, we only plot the results of $\delta = 1.8$ in Fig. 2. It is evident that the ensemble mean forecasts of the TC tracks, despite presenting different forecasting errors and improvements against the control forecast for different OTIs and ensemble sizes N, all significantly improve the control forecast, with the largest improvement being for an OTI of 6 h, and in this case, ensemble sizes from 13 to 21 yield better results than other sizes [see Figs. 2a(1),(2)]. Afterward, we examine the ensemble mean forecasting errors for an OTI of 6 h with different perturbation amplitudes δ and find that the error for $\delta = 1.8$ is the smallest and decreases the control forecasting error to the greatest extent [see Figs. 2b(1),(2)]. The BSS of the strike probability for an OTI of 6 h and perturbation amplitude $\delta = 1.8$ is also superior to those of other values of the OTI and δ [see Figs. 2a(3),b(3)]. Therefore, we determine the ensemble parameters of the OTI and δ as 6 h and 1.8, respectively. For the ensemble sizes, it has been shown that the ensemble mean forecasting errors for an OTI of 6 h are much smaller for ensemble sizes N from 13 to 21; nevertheless, when we look at the BSS as a function of ensemble sizes, such ensemble sizes cannot support the largest BSS of the strike probability but merely approach the largest score [see Fig. 2a(3)]. In this situation, we have to balance the OTI, δ, and N to achieve much higher ensemble forecasting skill. As a result, we select an ensemble size of $N = 21$ because the corresponding skill is nearest to the largest score. Although this ensemble size is finalized for the O-CNOPs using the WRFV3.6, there are previous studies and even operational forecasts that adopt approximately the same ensemble sizes (Zhou et al. 2017; Li et al. 2019; Mamgain et al. 2019; Sarkar et al. 2021). This implies that the ensemble size of 21 could be feasible. Therefore, we determine the ensemble parameters of the O-CNOPs as OTI = 6 h, perturbation amplitude $\delta = 1.8$, and ensemble size $N = 21$.
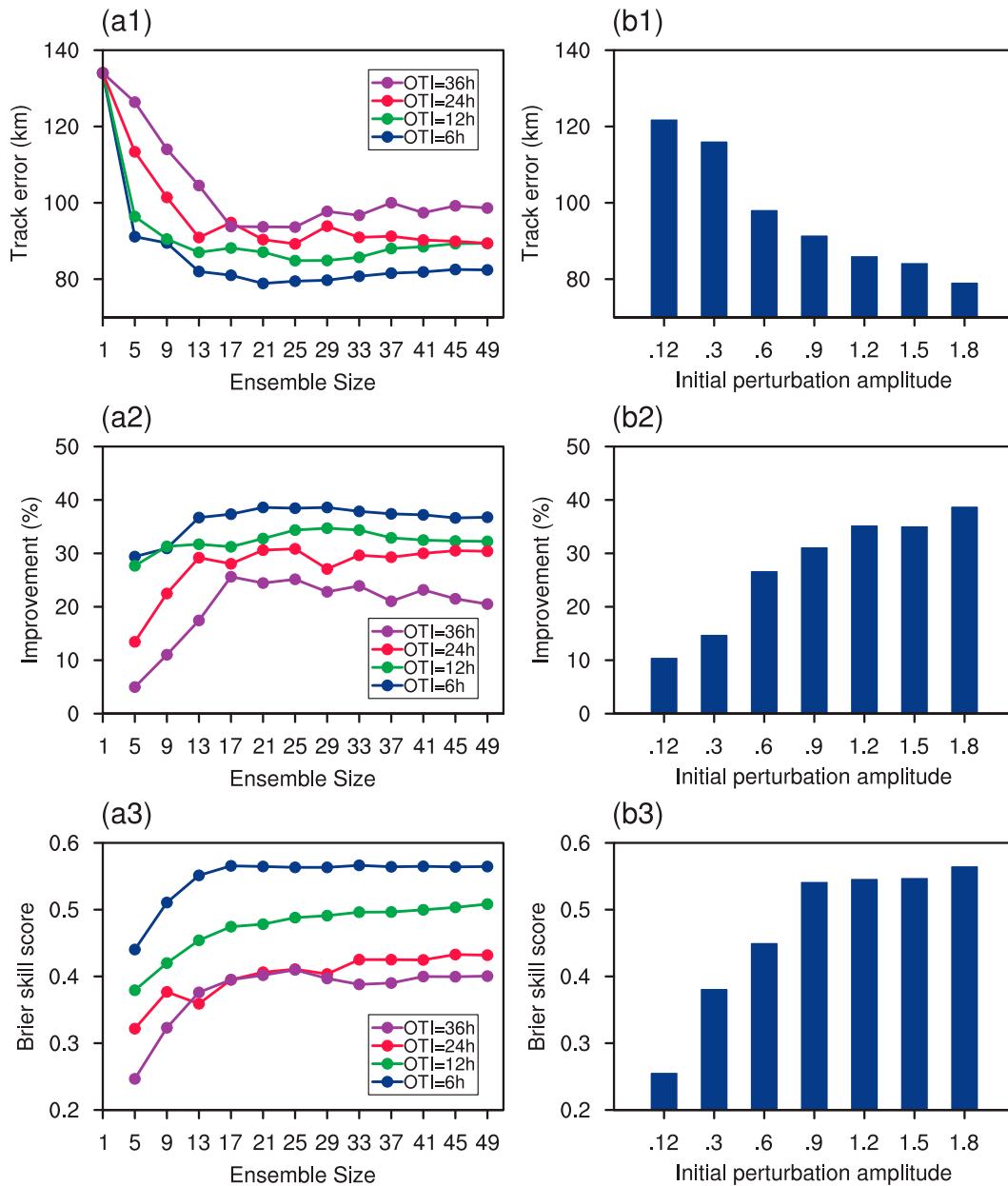
FIG. 2. Effect of the ensemble size and OTI on (a1) the ensemble mean TC track forecasting error, (a2) its improvement against the control forecast, and (a3) the BSS for TC strike probability. These skills are obtained by averaging all lead times over 5 days and the 5 selected TC cases when the initial perturbation amplitude is 1.8 and the OTIs are 6, 12, 24, and 36 h, respectively. (b1)–(b3) The corresponding skills as a function of the initial perturbation amplitude when the ensemble size is fixed to 21 and the OTI is 6 h. A paired t test is performed on the track errors of the five TC cases at all lead times, and significance at the 95% confidence level is achieved for the improvement of the ensemble mean against the control forecast.

4. Ensemble forecasting experiments for TC cases

In the last section, we finalized the configuration of the O-CNOPs for the ensemble forecast experiments with the WRF Model using the five TC cases formed in the WNP, including one in 2005, one in 2007 and three during 2018–20. In fact, there were a total of six TCs in the WNP during 2018–20, which all reached at least STY intensity and greatly influenced

China (see section 2a). Therefore, it should be investigated whether the above configuration of the O-CNOPs also greatly helps increase the track forecasting skill in the other three TC cases, i.e., Maria (2018), Maysak (2020), and Haishen (2020). In addition, it is worth examining whether the usefulness of the O-CNOPs with the above configurations can be reflected in improving the forecasts of two other TCs, Megi (2010) and

Tembin (2012), which have unusual tracks. After addressing these questions, we make additional attempts to implement O-CNOPs in the real-time track forecasts of TC In-Fa (2021) and TC Hinnamnor (2022), and the results are also reported in this section.

a. Hindcasts of the other three TCs

The three TC cases, as mentioned above, include the westward-moving TC Maria (2018) and the northward-moving TCs Maysak (2020) and Haishen (2020). Each of these TCs includes three forecast periods (see Table 1). The ensemble members for each TC case and their respective ensemble mean and best tracks are shown in Fig. 3. It is shown that the ensembles for any case tend to be located on the two sides of the best track, rather than the control forecast, which certainly causes the ensemble mean to be farther from the control forecast but close to the best track. From Fig. 3, it can be seen that almost all of the TC forecasts present ensemble mean forecasting in which the TC landfall points and times are much closer to those in the best tracks compared to the control forecasts, except for the forecast of the first forecast period of TC Haishen (2020); in particular, there are always several members that capture the landfall points accurately. Quantitatively, for the three forecasts of each TC case, the landfall positions of the ensemble mean reduce their counterpart errors in the control forecasts from 371 to 194 km for Maria (2018), from 130 to 90 km for Maysak (2020), and from 113 to 60 km for Haishen (2020), which all decrease the position errors of the control forecasts by at least 31%. We also investigate the performance of the O-CNOPs with the above configuration at different lead times. As shown in Fig. 3, the TC tracks in the ensemble mean often have forecasting errors that are obviously smaller than those in the control forecasts, especially when the lead times exceed 72 h.

b. Hindcasts of the two unusual TC tracks

Now, we investigate whether the above O-CNOPs also perform well in the hindcasts of unusual TC tracks. It is known that accurately forecasting the unusual movement of a TC is very difficult. Therefore, if the ensemble forecasts generated by the O-CNOPs can reproduce the unusual tracks, it would further confirm the usefulness of the O-CNOPs in improving the TC track forecasting skill. To test this, we conduct ensemble forecast experiments for the two unusual TCs, Super TY Megi (2010) with a sharp northward-turning track and STY Tembin (2012) with a counterclockwise loop track.

For Megi (2010), we use the initial and boundary conditions derived from the NCEP FNL dataset as described in section 2 and integrate the WRF Model from 1200 UTC 17 October to 1200 UTC 23 October 2010. Then, the control forecast with a lead time of 6 days is obtained. This control forecast reproduces the motion of Megi (2010) well, except for an earlier turn than that in the best track. In this case, the use of the O-CNOPs reduces the track error averaged over 6 days from 130 km in the control forecast to 94 km in the ensemble mean forecasting, which corrects the turning position much more but does not correct the turning angle (see Fig. 4). However,

when we additionally use the corresponding fifth-generation ECMWF reanalysis (ERA5) dataset to provide the initial and boundary conditions and yield the control forecast, similar to the operational forecast from the ECMWF global model, we obtain a much larger forecast error of the TC track and fail to catch the sharp northward turning position and time of Megi (2010) (Qian et al. 2013). Then, in this situation, can the O-CNOPs correct this large track error? Fig. 4 also shows the ensemble members of Megi's track generated by the O-CNOPs with a lead time of 6 days, which covers the northward turn and final landing processes. Note that these ensemble members are generated on the control forecast associated with ERA5. It is inspiring that the ensembles tend to be located on the two sides of the best track, especially after the turning point for this unusual TC. As a result, Megi's track in the ensemble mean forecasting reproduces the sudden northward movement and associated subsequent landfall in the best track much more accurately. On average, Megi's track in the ensemble mean forecasting corrects the bias of the control forecast by approximately 41%.

For Tembin (2012), we attempt to conduct ensemble forecasts initialized at 0000 UTC 21–23 August 2012, all finalized at 0000 UTC 29 August 2012, which have 8-, 7- and 6-day lead times, respectively, and all cover the complete loop track of Tembin (2012). Practically, Tembin (2012) in the best track moved westward along the periphery of the SH in the early stage, and the much larger Super TY Bolaven (2012), which was located east of Tembin (2012), gradually approached Tembin (2012), affecting its environmental steering flow and forcing Tembin (2012) to execute a slow cyclonic loop (Xian and Chen 2019). To reveal the forecast errors in the speed and direction of TC movement, we decompose the track error into one part along the best track, which is denoted as the along-track error (ATE), and another part orthogonal to the best track, which is called the cross-track error (CTE) (see appendix B for the definition). Figure 5 plots ATE and CTE for the control forecasts initialized at 0000 UTC 21 August with the NCEP FNL and the associated ensemble mean for Tembin (2012), where a positive (negative) ATE indicates that the forecasted storm moves faster (slower) than indicated by the best track and a positive (negative) CTE indicates that the forecasted track is to the right (left) of the best track. Figure 5 shows that the control forecast initialized at 0000 UTC 21 August has a slow bias in the westward movement of Tembin (2012) in the early stage (as indicated by the negative ATE before 0000 UTC 25 August shown in Fig. 5), which may result from a weak bias in simulating the SH in the control forecast and a stronger bias in simulating the continental high (CH), thus causing a much smaller distance between Tembin (2012) and Bolaven (2012). This could result in a stronger interaction between the two TCs than the actual interaction, and it yields a premature onset near Taiwan of Tembin's loop track; thereafter, the track error grows rapidly as Tembin (2012) moves northeastward after 0000 UTC 27 August (as indicated by the positive CTE shown in Fig. 5). Nevertheless, regarding the ensemble forecast generated by the O-CNOPs, the real track of Tembin (2012) is much better replicated when averaged over the entire forecast period,
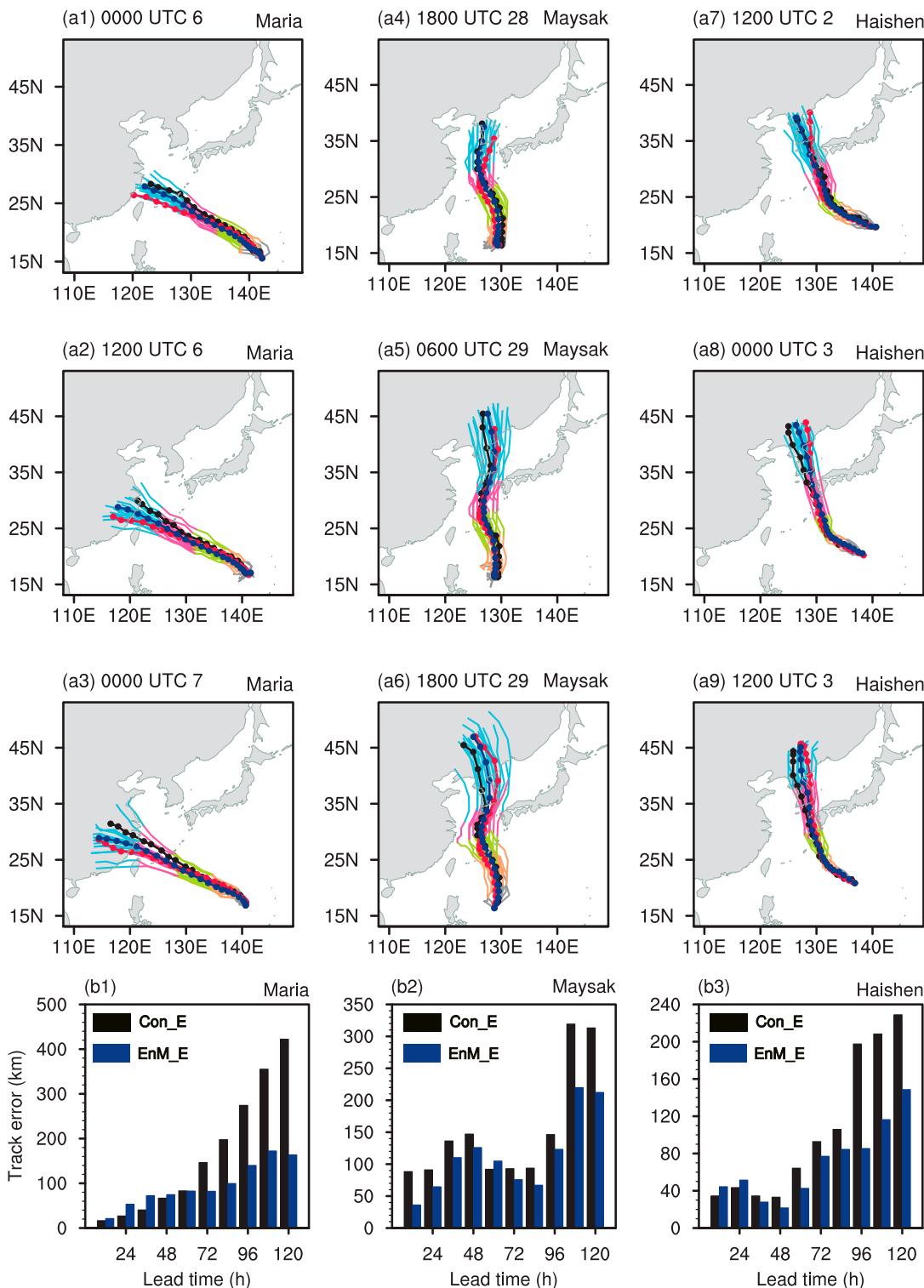
FIG. 3. The 5-day tracks for the control (black) and three ensemble mean (blue) forecasts of TCs (a1)–(a3) Maria (2018), (a4)–(a6) Maysak (2020), and (a7)–(a9) Haishen (2020) and their respective best tracks (red), where any two consecutive dots located on a track define a time interval of 6 h, and the ensemble members are time phases marked with different colors representing 0–24 h (gray), 24–48 h (orange), 48–72 h (green), 72–96 h (magenta), and 96–120 h (light blue). (b1)–(b3) The track errors at different forecast times for the control forecast (black bar) and the ensemble mean (blue bar) when averaged over all three forecasts for each of the three TC cases.
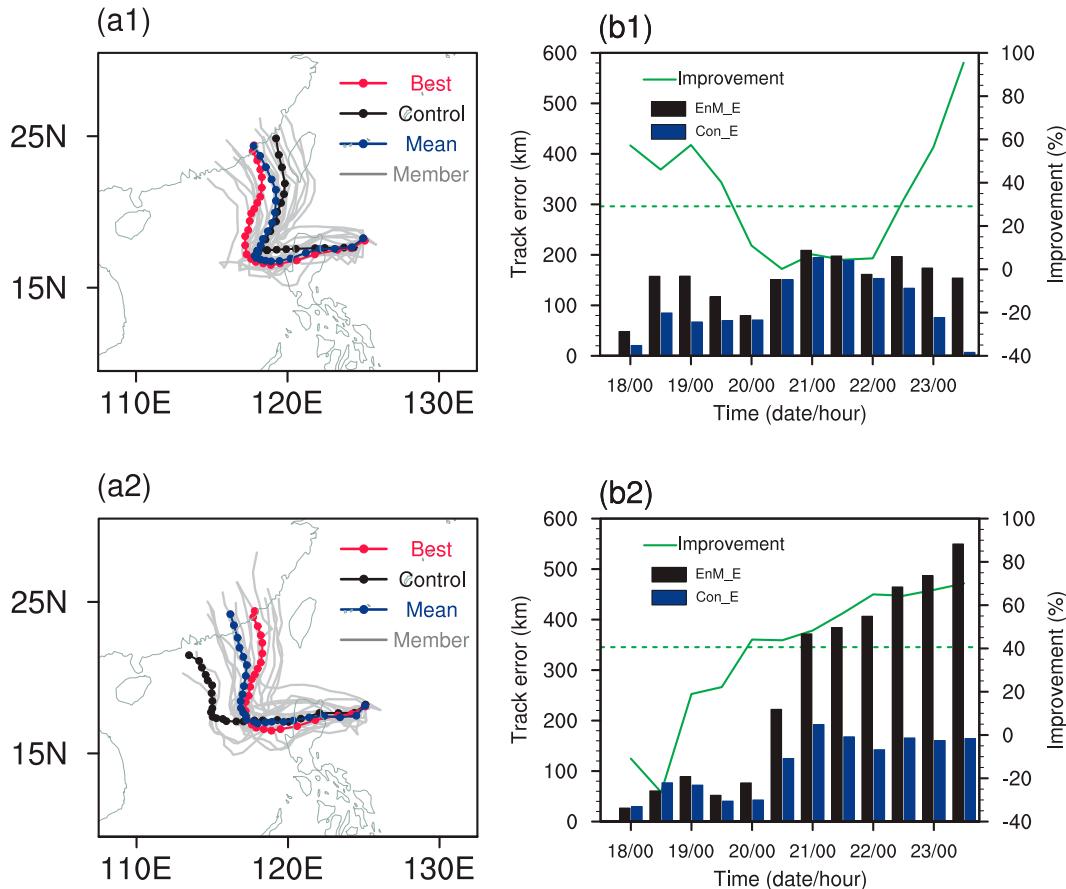
FIG. 4. (a1) The 6-day best track (red) and 6-day tracks of the control forecast (black), ensemble mean forecast (blue), and ensemble members (gray) for Megi (2010) generated by the O-CNOP method initialized at 1200 UTC 17 Oct 2010 using the NCEP FNL dataset to provide the initial and boundary conditions. (b1) Track errors at different forecast times for the control forecast (black bar) and ensemble mean forecast (blue bar) and the error reduction rate (i.e., skill improvement; solid green line) due to the ensemble mean, where the dashed line represents the mean of the improvements averaged over a forecast period of 6 days. (a2),(b2) As in (a1) and (b1), but for the ensemble forecasts using the ERA5 dataset to provide the initial and boundary conditions.

even though there is a slow bias at day 8 (see Fig. 5); in particular, there exist some ensemble members possessing a stronger SH and weaker CH and thus correct the slow-moving bias of Tembin (2012) in the early stage as well as the strength and speed of the westerly trough (see Fig. 6). These members realistically describe the interaction of Tembin (2012) with Bolaven (2012) and ultimately enhance the improvement of the forecasting skill of Tembin's track.

The ensemble forecast initialized at 0000 UTC 22 August for Tembin (2012) exhibits a dramatic range across the ensemble members (see Fig. 5). While some bad members, for example, the control forecast, present tracks that drift northeastward quickly and make landfall in the Korean Peninsula or Japan earlier than the best track, there exist other good members, which loop slowly and present TC movements that are similar to those of the best track and do not affect Korea or Japan before 29 August. Regardless, the ensemble mean improves the skill of the control forecast for Tembin's track. More precisely, the control forecast possesses an ATE that

grows from almost zero on 26 August to approximately 1000 km on 29 August and a CTE from approximately −500 to 373 km, indicating that Tembin (2012) moves northeastward much faster than the best track after 26 August, while the ensemble mean reduces the 7-day track forecasting errors by approximately 40% on average despite presenting a much faster and more eastward movement compared with the best track. When the lead time is shortened as initialized at 0000 UTC 23 August, the track error in the control forecast obviously decreases (see Fig. 5); correspondingly, more ensemble members reproduce the slow loop track of Tembin (2012) than the forecasts initialized at 0000 UTC 21 and 22 August and cause the ensemble mean to present a track much closer to the best track even though the moving speed is still slightly faster.

From the above, it is clear that the O-CNOPs are still useful for improving the forecasting skill of unusual TC tracks, especially because they capture the turning positions and times of unusual TCs well.
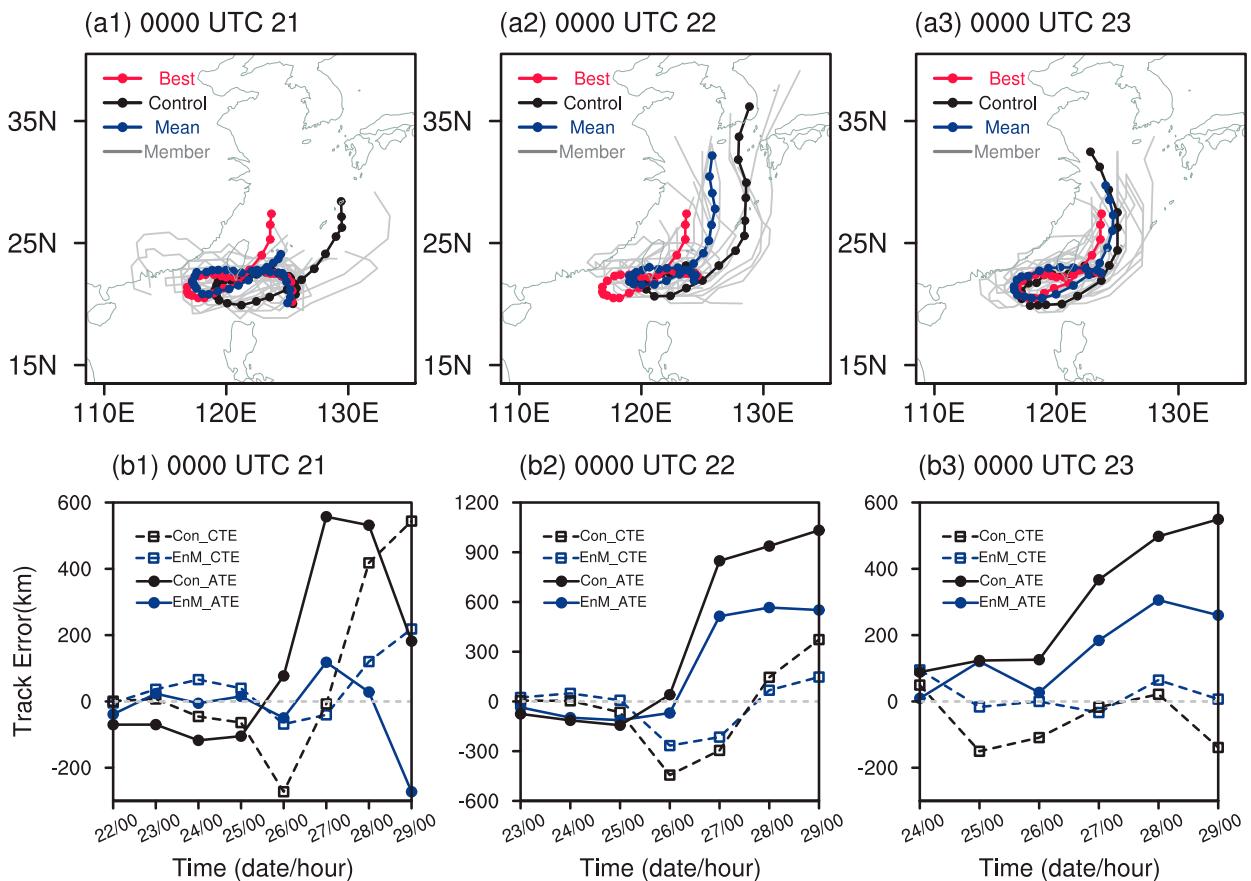
FIG. 5. Best track (red) and tracks of the control forecast (black), ensemble mean (blue), and ensemble members (gray) for Tembin (2012) generated by the O-CNOPs initialized at 0000 UTC 21–23 Aug 2012, which represent (a1) 8-, (a2) 7-, and (a3) 6-day lead times, respectively. The corresponding ATEs (dots) and CTEs (boxes) are for the control forecasts (black) and the ensemble mean (blue) during the lead times of (b1) 8, (b2) 7, and (b3) 6 days.

c. Real-time forecasts of STY In-Fa (2021) and Super TY Hinnamnor (2022)

To further evaluate the role of the O-CNOPs in improving TC track forecasting skill, we also conducted real-time forecasts for STY In-Fa (2021) and Super TY Hinnamnor (2022) using the GFS forecast data as the initial and boundary conditions of the WRF Model. Because the GFS data are delayed in terms of their availability from the National Oceanic and Atmospheric Administration (NOAA; see the data availability statement at the end of the article) and time is needed to calculate the O-CNOPs, the real-time forecasts are initialized approximately 8 hours later, but we focus on the future tracks after this 8-h period. Taking In-Fa (2021) as an example, the 5-day track forecast is initialized at 0000 UTC 21 July 2021 but focuses on the future track from 0800 UTC 21 July 2021 to 0000 UTC 26 July 2021.

For In-Fa (2021), the challenges of operational track forecasts are its northward movement before landfall and slow movement on land (Wang et al. 2022; Xiang et al. 2022). We initialize the forecasts of In-Fa (2021) at 0000 UTC 21–24 July 2021 and finalize them with a lead time of 5 days. To facilitate the discussion, we also call these forecasts control forecasts,

similar to those in the hindcast experiments in sections 4a and 4b. Figure 7 shows that the control forecasts initialized at 0000 UTC 21 and 22 July for the northwestward movement of the TC in the best track present a southward bias reaching 200 and 130 km before landfall, as indicated by the negative CTEs; however, the ensemble mean forecasts made by the O-CNOPs correct the biases to below 80 and 60 km and accurately capture the first landfall point of the TC. For the strike probability, the ensembles made by the O-CNOPs exhibit 50%–70% probabilities for the actual landfall point when the forecasts are initialized at 0000 UTC 21 and 22 July (see Fig. 7). As expected, when the initialized times are much closer to the landfall time, for example, when the forecasts are initialized at 0000 UTC 23 and 24 July, an increasing number of ensemble members capture the landfall positions, and the strike probabilities across the landfall positions also increase to 80%–100% (see Fig. 7). When In-Fa (2021) landed, the SH and CH retreated to the east and west, respectively, which trapped In-Fa (2021) into a saddle pattern and caused a very weak steering flow that slowed it down (Yang et al. 2022). From Fig. 7, we can also see that after TC landfall at 0430 UTC 25 July, the control forecasts initialized at
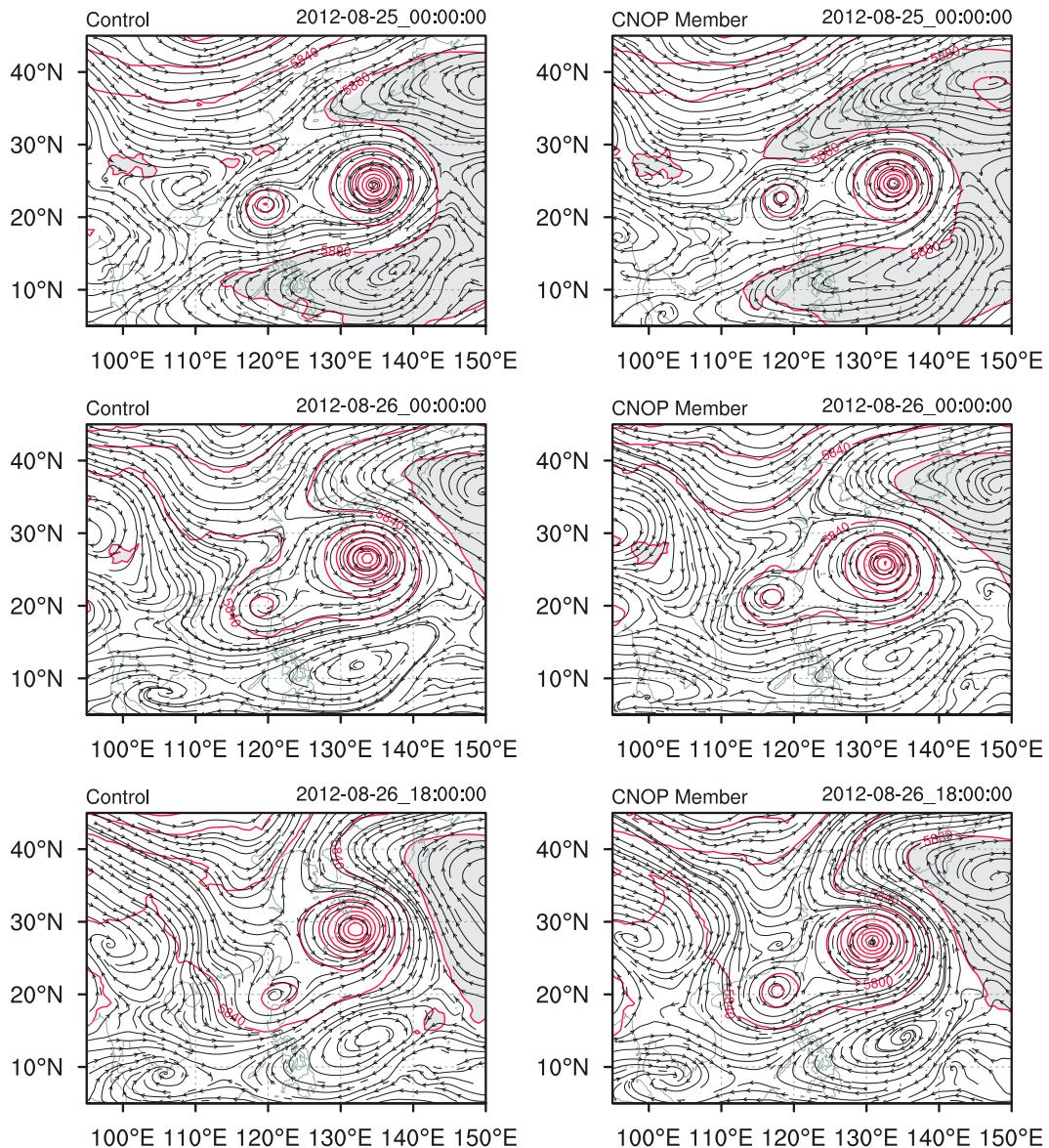
FIG. 6. Deep-layer mean winds averaged vertically between 850 and 250 hPa (black streamlines) and 500-hPa geopotential heights (red contours, interval of 40 gpm) at 0000 UTC 25 Aug, 0000 UTC 26 Aug, and 1800 UTC 26 Aug 2012 for Tembin (2012) for (left) the control forecast and (right) one of the ensemble members generated by the O-CNOPs that possesses a stronger SH and weaker CH. The areas above 5880 gpm are shaded. The forecasts are initialized at 0000 UTC 21 Aug 2012.

0000 UTC 22 and 23 July move faster than the best track, while the ensemble mean forecasts correct the erroneously fast-moving tracks to nearly overlay the best track. However, the control forecast initialized at 0000 UTC 24 July already reproduces the movement speed of In-Fa (2021) much better, and the track error only includes a small bias of northward movement. Although the bias is small, it is still reduced by the ensemble mean of the O-CNOPs; in particular, more ensemble members are also found to be much closer to the best track, giving the best track a more than 80% strike probability at lead times ranging from 6 h to 4 days.

The real-time forecast for Hinnamnor (2022) is initialized at 0000 UTC 30 August 2022 and finalized with a lead time of 7 days to see whether it could greatly influence China at the longer lead times. This TC also has an unusual track, and the ensemble forecasts generated by the O-CNOPs perform well regarding the track forecasts. Figure 8 shows the control forecast, the ensemble mean and the best track of Hinnamnor (2022). It is shown that the control forecast presents a much smaller track error in the first 3 days, and the ensembles generated by the O-CNOPs exhibit a much smaller divergence at this stage. However, after the TC turns northward, the track
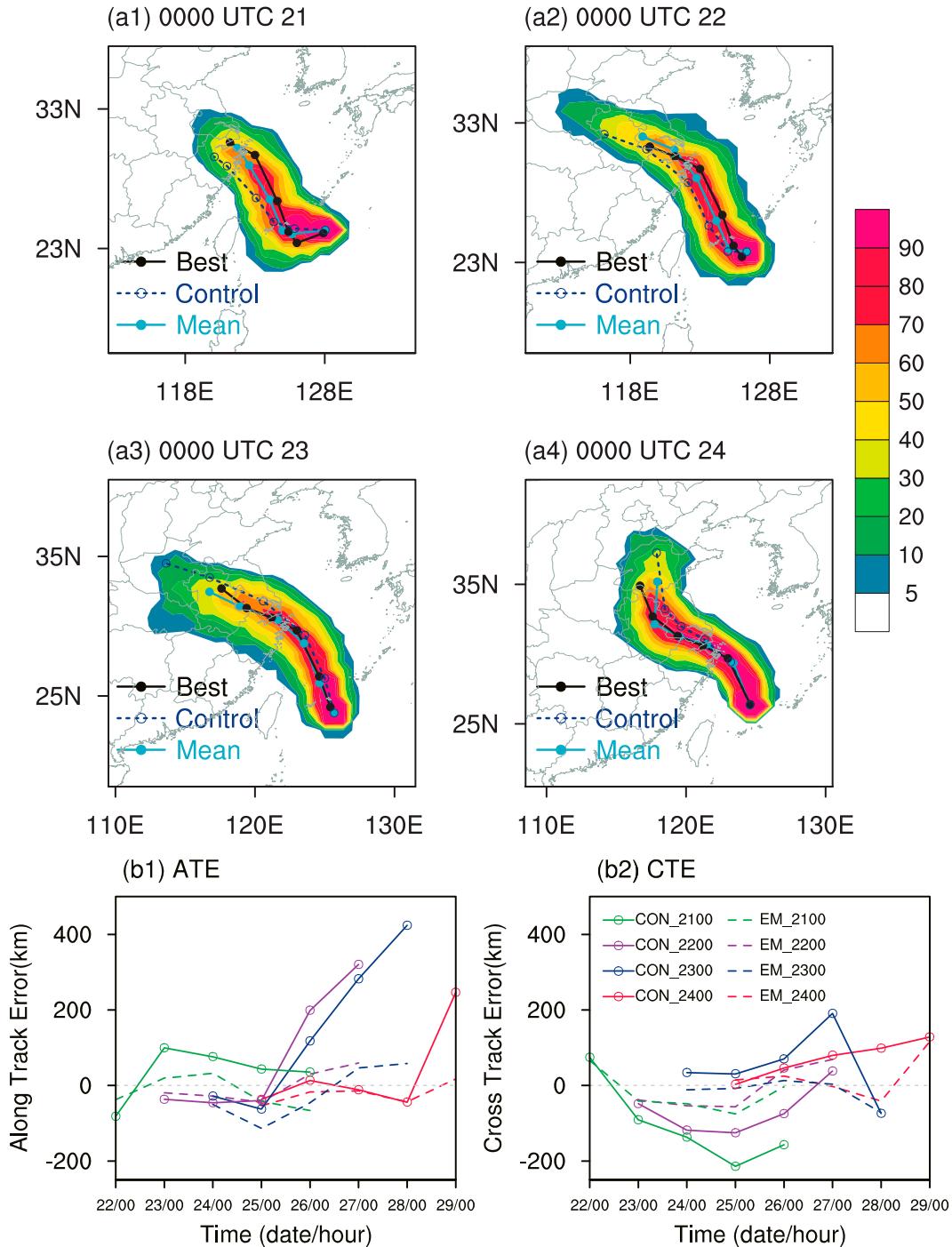
FIG. 7. The 5-day tracks in the best forecasts (black), control forecasts (blue), and ensemble mean forecasts (light blue) generated by the O-CNOPs for In-Fa (2021) and corresponding strike probabilities (shaded), where the initialized times are at (a1) 0000 UTC 21 Jul, (a2) 0000 UTC 22 Jul, (a3) 0000 UTC 23 Jul, and (a4) 0000 UTC 24 Jul 2021. The (b1) ATE and (b2) CTE for the control forecasts (solid lines) and the ensemble mean forecasts (dotted lines).

error in the control forecast grows rapidly, and a slow bias occurs, finally reaching an error of approximately 600 km (see the negative ATE in Fig. 8); in this situation, the ensemble mean forecast corrects the bias to less than 150 km (see Fig. 8) and captures the landfall in Korea, which does not occur in the control forecast. Therefore, the ensemble forecast better evaluates the influences of Hinnamnor (2022) on the eastern coast of China, which are overestimated in the control forecast at the later stage.
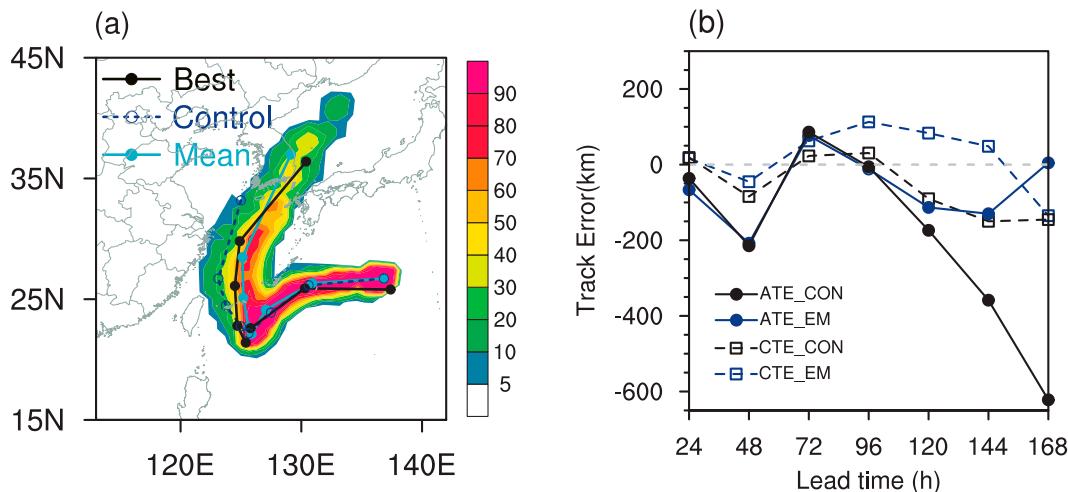
FIG. 8. (a) The 7-day track for the best forecast (black), control forecast (blue), and ensemble mean forecast (light blue) for Hinnamnor (2022) generated by the O-CNOP method initialized at 0000 UTC 31 Aug 2022 and the corresponding strike probability forecast (shaded). (b) ATEs (dots) and CTEs (boxes) for the control (black) and ensemble mean (blue).

It is now obvious that the O-CNOPs method shows the potential to greatly improve the track forecast skill in real-time forecasts of TCs.

5. A statistical metric for the O-CNOPs' performance and comparison with SVs, BVs, and RPs

As stated in section 4, a total of 18 ensemble forecast experiments are conducted for the 7 TCs to investigate the performance of O-CNOPs. In this section, we combine these ensemble forecasts, together with the 11 ensemble forecasts of the 5 TCs adopted in section 3, to statistically evaluate the performance compared with that of traditional RPs, BVs, and SVs, where the same configurations are used for the ensemble size, amplitude of initial perturbations, and OTI, particularly for the SVs, as those in the O-CNOPs. In fact, we test different configurations of the RPs, BVs, and SVs as in the O-CNOPs in numerical experiments (see section 3) and find that they are always shown to have much lower ensemble forecast skill than the O-CNOPs; for simplicity, the same configuration as that for the O-CNOPs is chosen to compare them.

A comparison of the ensemble forecast skills is performed from both the ensemble mean forecasts and TC strike probability forecasts. In this section, the ensemble mean forecast error and the BSS, as described in section 3 and appendix B, respectively, are adopted to assess the quality of the deterministic forecasts and strike probabilities. Additionally, the reliability diagram (RD), the relative operating characteristic (ROC) curve and the area under the ROC curve (ROCA) are also adopted to evaluate the skill of the strike probabilities. Specifically, the RD displays how well the forecasted probabilities correspond to their observed frequencies; the ROC curve acts as a function of the hit rate with respect to the false alarm rate, and its associated ROCA measures the

probabilistic forecast skill of a binary event, where a forecast can be regarded as being skillful when the ROCA is larger than 0.5, and the larger the ROCA is, the more skillful the ensemble forecasts. The details of these measurements are also introduced in appendix B.

a. Comparison of ensemble forecast skills

Figures 9–11 show the track forecasts made by the O-CNOPs, RPs, BVs, and SVs for the 12 selected TC cases. By comparison, we show that the ensembles for the 12 TC cases made by the RPs, BVs, and SVs are always much closer to the control forecasts and do not capture the best track. However, the ensemble members generated by the O-CNOPs, as shown in section 4, are often closer to the best tracks than the control forecasts. As a result, the ensemble mean forecasts made by the RPs, BVs, and SVs are almost identical to the control forecasts, while those generated by the O-CNOPs often remain away from the control forecasts and approach the best tracks because they have much faster growth than the RPs, BVs, and SVs in the WRF Model. In particular, the O-CNOPs capture the sharp northward turn for Megi (2010) much more accurately and reproduce the counterclockwise loop track for Tembin (2012). Qualitatively, the ensemble mean forecast generated by the O-CNOPs, averaged for the 12 TC cases and a total of 29 forecast periods, reduces the track forecast error at day 5 from 316 to 163 km (i.e., the skill is improved by 48%; see Fig. 12). This error is even smaller than the error in the control forecast at a lead time of 84 h. However, the ensemble mean forecasts generated by the SVs only reduce the track errors in the control forecasts by approximately 12% at day 5 (see Fig. 12). It is conceivable that the inclusion of nonlinearities in the O-CNOPs method contributes to its superiority over SVs in the forecasts of nonlinear TC motions. Although the BVs method is also a nonlinear approach, it yields an ensemble mean for the TC tracks that is initially close to the best track but after 2 days is very similar to that of the RPs method
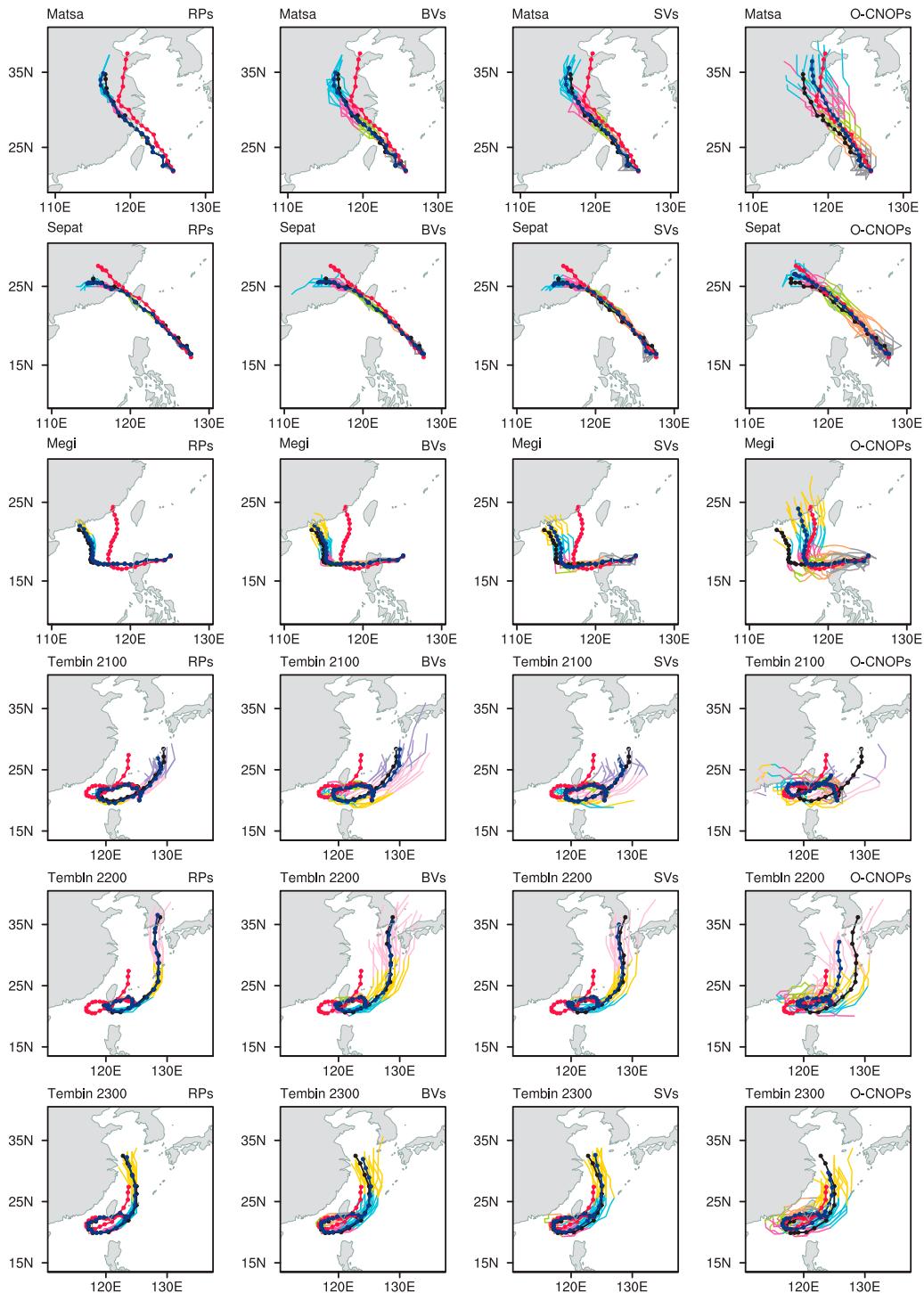
FIG. 9. Forecasts of all four TC cases before 2018 adopted in the present study. The figure plots the best tracks (red), tracks of control forecasts (black), and ensemble mean forecasts (blue) generated by the RPs, BVs, SVs, and O-CNOPs, where any two consecutive dots located on the tracks define a time interval of 6 h, and the time periods of each of the ensemble members are marked with different colors, representing 0–24 h (gray), 24–48 h (orange), 48–72 h (green), 72–96 h (magenta), 96–120 h (light blue), 120–144 h (yellow), 144–168 h (pink), and 168–192 h (purple). The four TCs are Matsa (2005), Sepat (2007), Megi (2010), and Tembin (2012), and each of the former three TCs includes one forecast period, while the fourth TC includes three forecast periods.
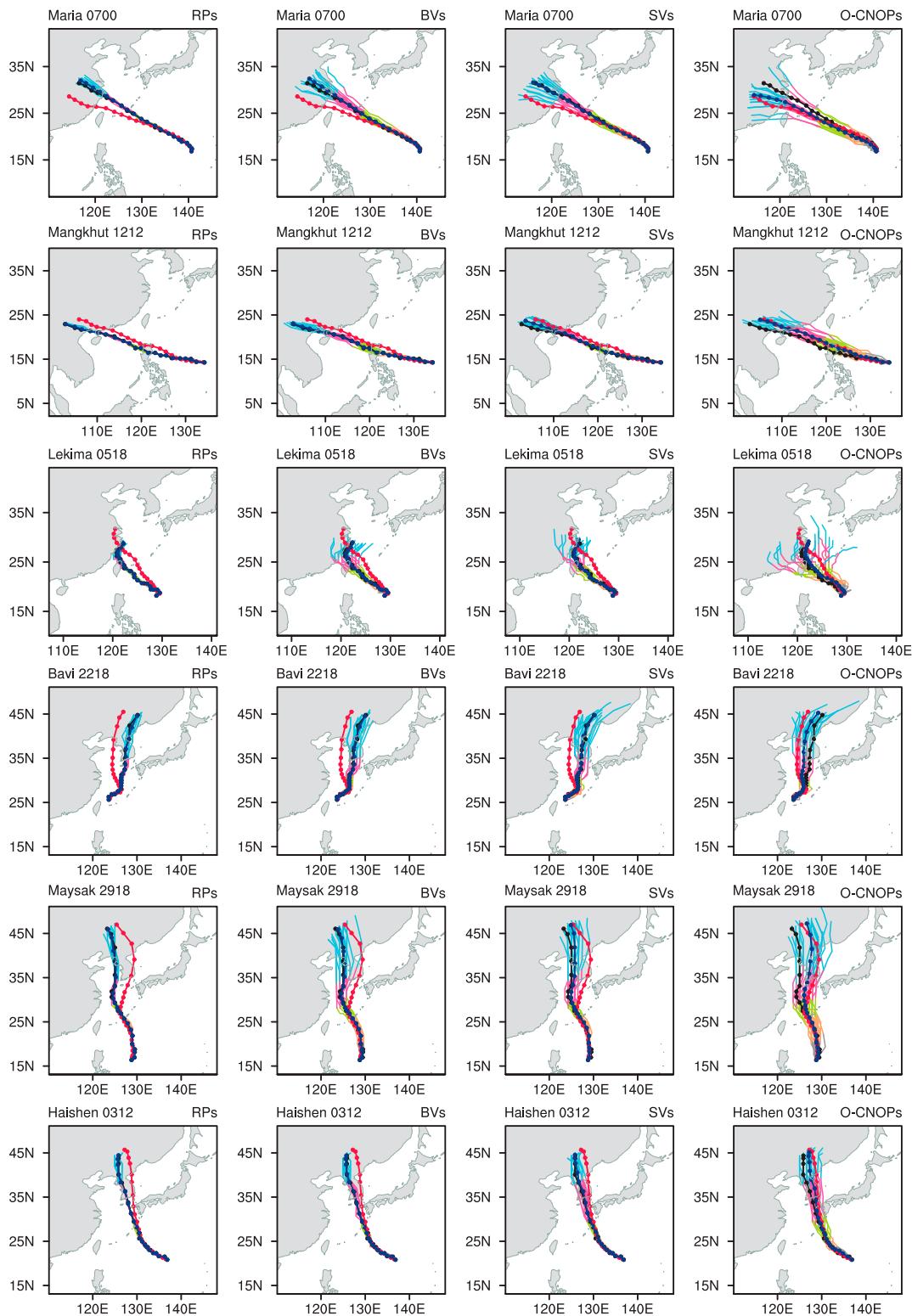
FIG. 10. As in Fig. 9, but for the six TC cases during 2018–20, which are Maria (2018), Mangkhut (2018), Lekima (2019), Bavi (2020), Maysak (2020), and Haishen (2020). Each of these TCs includes three forecast periods, but the figure only shows the tracks during the latest forecast period.
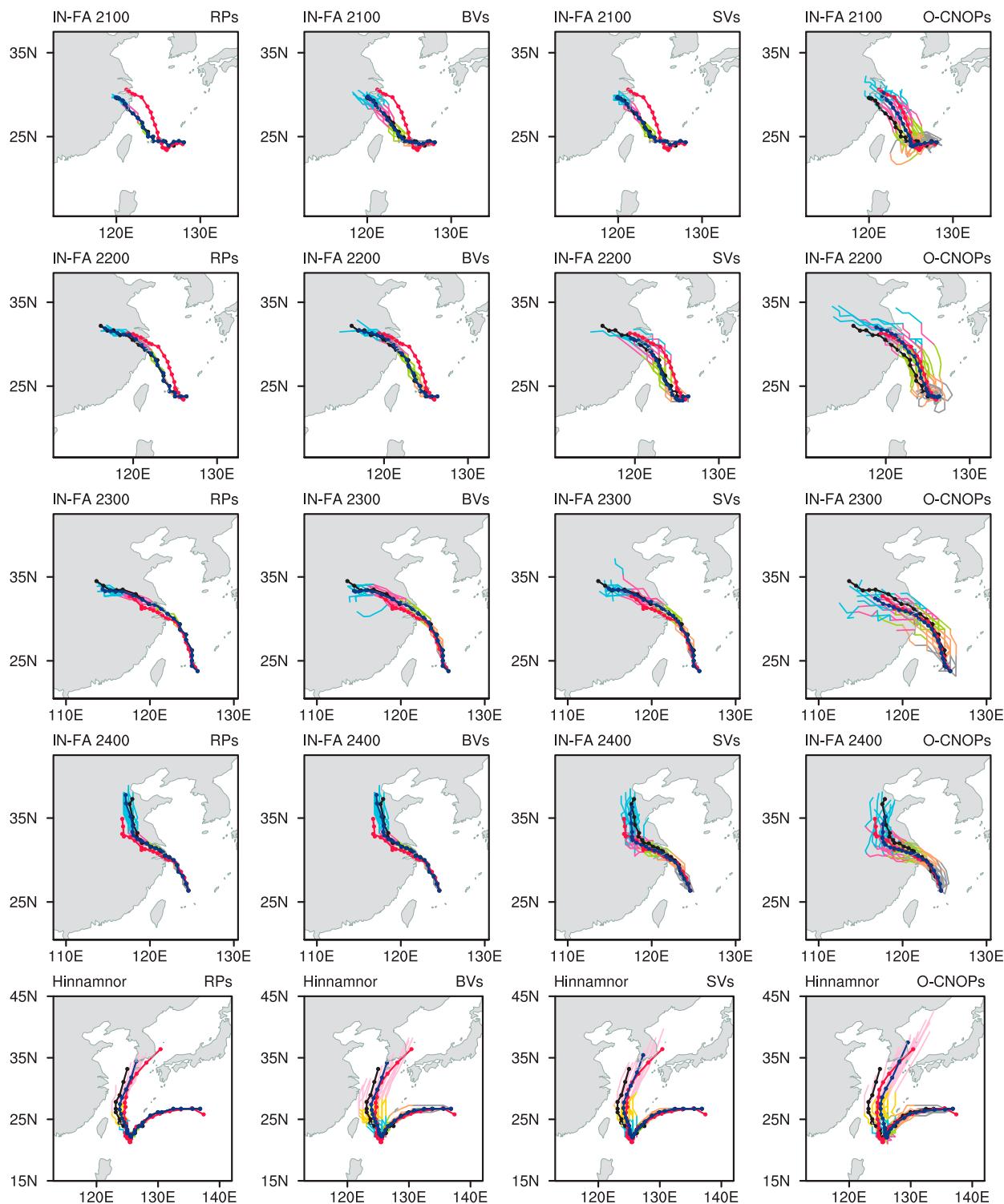
FIG. 11. As in Figs. 10 and 11, but for the real-time forecasts for In-Fa (2021) and Hinnamnor (2022). The figure shows the tracks for four forecast periods for In-Fa (2021) and one forecast period for Hinnamnor (2022).

and remains away from the best track. In fact, the BVs, as a group of orthogonal growing initial perturbations, are generated by sampling the growth behaviors of the initial perturbations during a time period before the initial time of the forecasts. Therefore, the

BVs cannot ensure that they capture the fast-growing modes of the initial errors during the forecast period as required by ensemble forecasts and thus cannot guarantee a high forecasting skill at a much longer lead time.
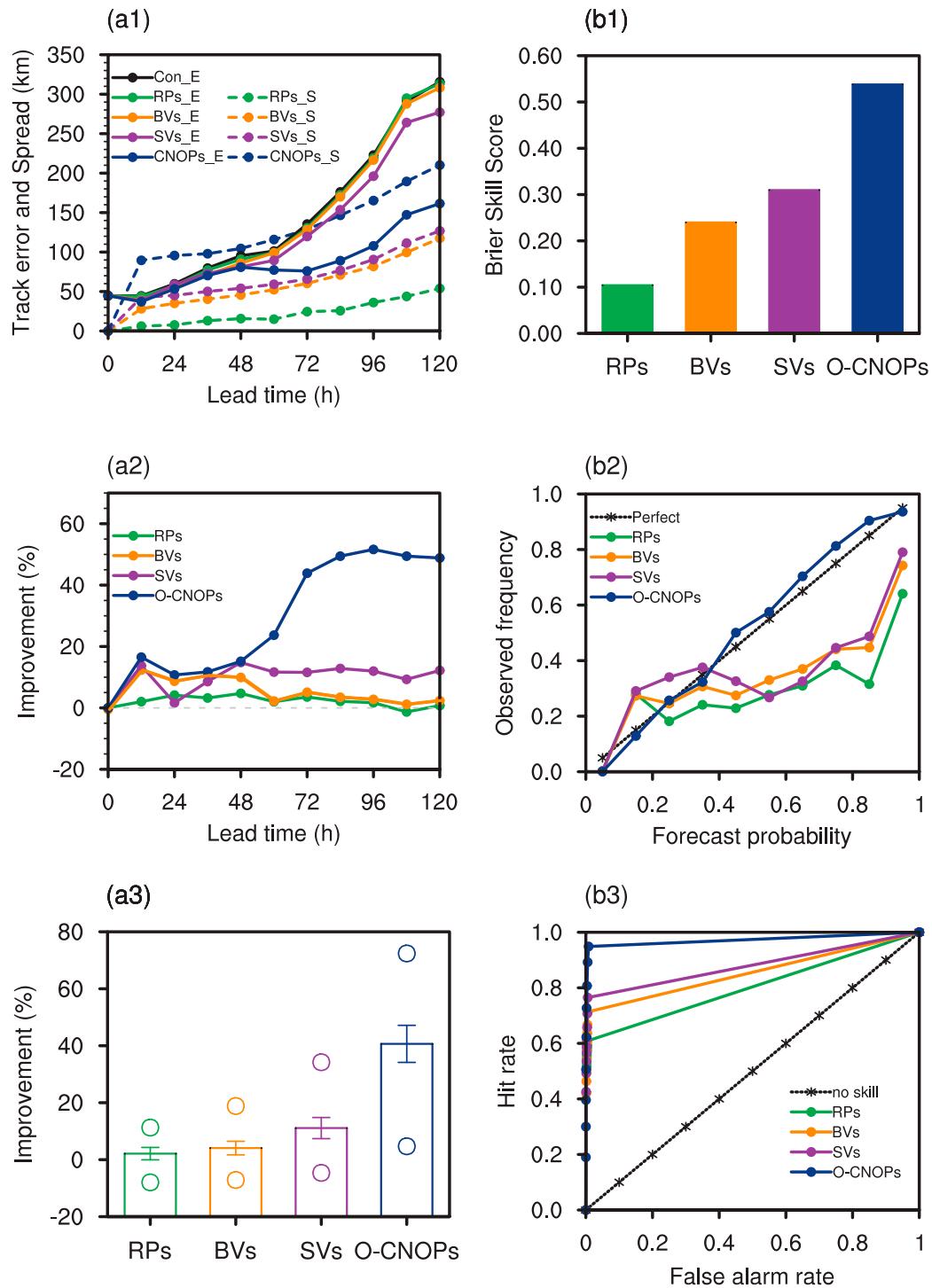
FIG. 12. (a1) Track errors (solid lines) at different forecast times for the control forecast and the ensemble mean forecasts averaged for 12 TCs and the corresponding ensemble spreads (dotted lines) generated by the RPs (green), BVs (orange), SVs (purple), and O-CNOPs (blue); (a2) error reduction rate (i.e., skill improvement) due to the ensemble mean; (a3) box-and-whisker plot for the skill improvement averaged for 12 TCs and all lead times with a 95% confidence interval, where the circles denote the maximum and minimum improvements for the 12 TCs; (b1) BSSs, (b2) RDs, and (b3) ROC curves are plotted for the TC strike probabilities generated by the four methods. The improvement of the ensemble mean generated by the O-CNOPs against those generated by the SVs, BVs, and RPs is statistically significant at the 95% confidence level.

The strike probabilities of the TC ensemble forecasts are also compared among the O-CNOPs, RPs, SVs, and BVs by evaluating the BSS, reliability diagram (RD), and ROC curve (see appendix B for their calculations). The results, averaged for the 12 TC cases at all lead times, are illustrated in Fig. 12. Specifically, the BSS achieved by the O-CNOPs is significantly larger than those obtained by the RPs, SVs, and BVs, which implies that the O-CNOPs achieve a strike probability far superior to those of the RPs, SVs, and BVs. For the reliability illustrated by the RD, the O-CNOPs provide the most reliable ensembles, which is indicated by the fact that their reliability curve is closest to the diagonal line, while the other methods exhibit reliability lines that are below the line of the O-CNOPs. From Fig. 12, we can also see that the O-CNOPs produce the highest hit rate and the lowest false alarm rate, which is also indicated by the largest ROCA value of 0.972 for the O-CNOPs, followed by 0.880 for the SVs, 0.855 for the BVs and 0.803 for the RPs.

In addition, we know that a perfect ensemble forecasting system is expected to provide an ensemble spread that is almost equal to the corresponding ensemble mean forecasting error (Bowler 2006; Leutbecher and Palmer 2008; Hopson 2014), which allows the ensemble spread to correctly indicate the forecast uncertainty. A comparison is also made to this spread-skill relationship for the above four methods that are used to generate the ensemble in the present study (see Fig. 12). The ensemble spreads generated by the RPs, BVs and SVs are significantly smaller than the corresponding ensemble mean forecast errors. In contrast, the O-CNOPs provide additional benefits for increasing the ensemble spread, which helps the ensemble mean forecast achieve a much higher skill, especially reducing the distance between the ensemble spread and the ensemble mean forecast error and making it smaller than those of the SVs, BVs, and RPs when the lead time exceeds 36 h. This illustrates that the O-CNOPs provide the best spread-skill relationship of the ensembles among the above four methods. However, the relationship is not perfect. The O-CNOPs aim to represent initial uncertainties that grow rapidly, and for a much larger amplitude of initial perturbations, the O-CNOPs may overestimate the growth of initial errors when they are applied to ensemble forecasts, which arises that the ensemble spreads are larger than the track errors of the ensemble mean. In fact, we find that further reduction of the initial perturbation amplitude by 50% can lead to an improvement in the consistency between the ensemble spreads and the track errors of the ensemble mean (the figure is omitted); however, the track forecast skills of the ensemble mean are substantially reduced at longer lead times.

From the above comparison, it is clear that the O-CNOPs presents the best performance in improving TC track forecasts among the four ensemble forecasting methods, followed by SVs and then BVs and RPs. Specifically, the O-CNOPs method is not only able to provide more accurate ensemble mean forecasts of TC tracks but can also produce more credible strike probability estimates, indicating that the O-CNOPs represent the initial uncertainties of TC track forecasts much better and provide both more reliable deterministic forecast results and more credible risk assessments for TCs.

b. Why do the O-CNOPs provide much higher ensemble forecast skill for the TC tracks?

It has been shown that the RPs, compared with the O-CNOPs, SVs, and BVs, generate ensemble members that are much closer to the control forecast and much farther from the best track. In fact, ensemble forecasts, to achieve much higher forecast skill, generally require initial perturbations that possess unstable spatial structures for fast growth (Toth and Kalnay 1997; Zhou and Chen 2006; Duan et al. 2016); however, RPs are generated by perturbing the initial analysis with random noise and lack unstable spatial structures, which certainly cause RPs to grow slowly or even fail to develop (Cheung 2001a,b). This may be the reason why the RPs have the lowest ensemble forecast skill for the TC tracks among the four methods. The O-CNOPs, SVs, and BVs all aim to estimate the subspace of fast-growing perturbations, but we show that there always exist more ensemble members generated by the O-CNOPs that better reproduce the best tracks for different TC cases. It is therefore inferred that the O-CNOPs are more appropriate for estimating the uncertainties of the key processes that dominate TC motion and are inclined to achieve much higher forecasting skill for TC tracks. In this section, we take Megi (2010) as an example to explain why the O-CNOPs achieve a higher ensemble forecast skill than the SVs and BVs.

Figure 13 shows the vertical distribution profiles of the perturbed moist energies (MEs) averaged for all ensemble members and over the entire domain (see section 2) for different lead times using the O-CNOPs, SVs, and BVs. The MEs of the BVs capture the dual peaks at both the surface and upper layers, as illustrated in Thanh et al. (2016), while the MEs for the SVs have a maximum in the lower boundary, and those for the O-CNOPs are mainly concentrated in the middle-to-lower troposphere. Torn et al. (2018) demonstrated that TC track forecasts are very sensitive to the uncertainty in the mid-to-lower troposphere [see also Wang and Ni (2011)]. It seems that the sensitivity provided by the MEs of the O-CNOPs ensemble captures this uncertainty. Furthermore, it could be this plausible sensitivity of the O-CNOPs that promotes the evolution of the MEs more significantly than those of the SVs and BVs at all layers (see Fig. 13). Therefore, the O-CNOPs are more likely to yield ensemble members that substantially depart from the control forecasts and provide a much larger ensemble spread that can most likely encompass the best track (see Fig. 12). This may be one of the reasons why the O-CNOPs ensemble yields a much higher ensemble forecast skill for the TC tracks.

Figure 14 provides the horizontal distributions of the vertically averaged MEs for the first three BVs, SVs and O-CNOPs. It is clear that their distributions are quite different. The BVs are often widely dispersed within the TC vicinity and the large-scale flow in which the TCs are embedded; however, the SVs and CNOPs generally identify the dynamically unstable structures located in the outer region of the TCs, which could reflect the effect on TC flows of the surrounding synoptic-scale systems. Nevertheless, due to the inclusion of nonlinear processes, the CNOPs provide unstable structures that are different from those of the SVs. Specifically, the maximum ME of the SVs is close to the CH and locates west of the TC center, while that of
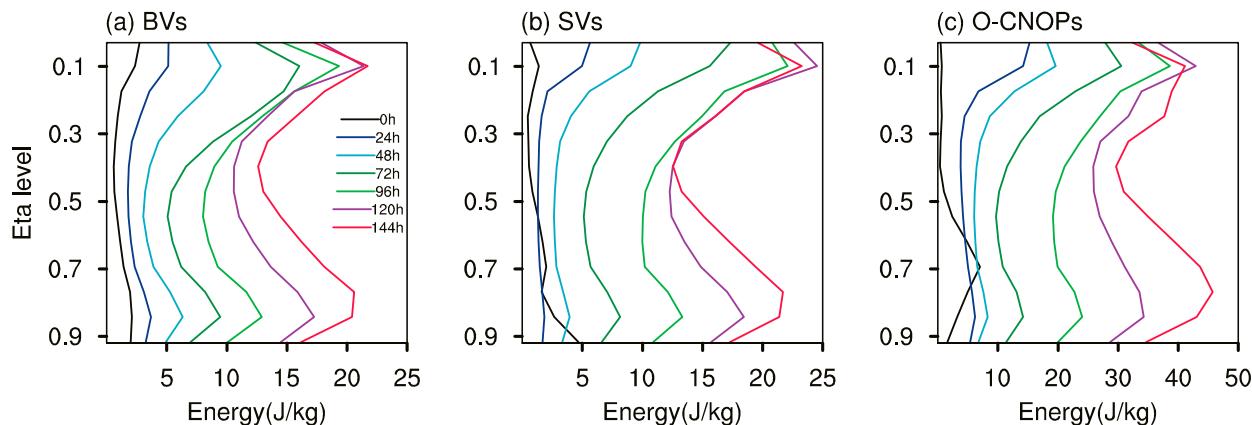
FIG. 13. Vertical distributions of the moist energies (MEs; J kg^{-1}) averaged for all ensemble members of TC Megi (2010) and over the entire domain at different lead times (colored curves). (a) BVs, (b) SVs and (c) O-CNOPs.

the CNOPs looks like an annulus around the TC and is relevant to not only the CH but also the SH, whose combined effects can determine when and where Megi (2010) began to deflect (Shi et al. 2014). This indicates that the control forecast, which fails to identify the turning position and time of Megi (2010), is very sensitive to the uncertainties that occurred in the annulus region identified by the CNOPs, because the CNOPs represent the optimally growing initial perturbations in the nonlinear model. Furthermore, these uncertainties, if they are described by the ensemble generated by the O-CNOPs, have a greater probability of capturing the combined effect of the SH and the CH on the deflection of Megi (2010) and result in a much better forecasted TC track. In Fig. 15, the environmental steering flows are illustrated in terms of the ensemble mean and ensemble spread obtained by the O-CNOPs, SVs, and BVs, where the steering flows are obtained by calculating the wind vertically averaged from 850 to 250 hPa (Luan et al. 2015). The uncertainties represented by the ensemble spread associated with the O-CNOPs locate the TC vicinity at the initial time when Megi (2010) in the control forecast is situated below the high pressure ridge (as indicated by the 5880-gpm isopleth in Fig. 14), which is composed of the CH over China and the SH over the WNP. In fact, when the positive/negative pairs of O-CNOPs presenting the dominant MEs in the annulus around Megi (2010) are superimposed on the initial analysis of the control forecast, they could cause uncertainties to occur in the TC vicinity and strengthen or weaken the cyclonic circulation of Megi (2010), thus perturbing the TC size. The change in the size of Megi (2010) would further modulate the intensity of the high pressure ridge and then the track of Megi (2010) (Qian et al. 2013; Wang et al. 2017). Specifically, if an enlarged Megi (2010) occurs in one ensemble member, it can erode both the SH and the CH and accelerate the breakdown of the high pressure ridge, which may in turn weaken the westward steering flow to decrease the westward moving speed of Megi (2010) and strengthen the meridional steering flows to make Megi (2010) turn northward earlier (Sun et al. 2015). The ensemble members generated by the O-CNOPs either include an enlarged TC or a diminished TC with respect to the TC in the control forecast because they are

formed by superimposing the positive/negative pairs of CNOPs on the control forecast. As a result, there is an ensemble spread of the steering flows associated with the SH and CH. Figure 15 shows that a much larger spread occurs in the northward steering flow relative to the SH after 48 h, which tends to modulate the uncertainty of Megi's turning northward. It could be that the O-CNOPs ensemble describes this uncertainty well and thereby provides more chances to capture the northward turning position and time of Megi (2010). As the lead time increases, the ensemble presents a much larger spread around the TC and starts to perturb the synoptic systems of the easterly wave and the midlatitude trough, consequently pushing Megi (2010) northeastward after deflection. Therefore, the O-CNOPs ensemble depicts the uncertainties of the key processes and systems that affect TC motions during the forecast periods; thus, a higher ensemble forecast skill for Megi (2010) is achieved. For the SVs and BVs, it is shown in Fig. 15 that although the ensembles also focus on the uncertainties of the key systems that influence TC motions, their spreads are much smaller than those of the O-CNOPs, especially at longer lead times, resulting in the ensemble members generally missing the TC best track. In fact, the SVs are generated by a linearized WRF and do not represent the optimally growing initial perturbations in the nonlinear WRF at longer lead times, causing the resultant ensemble to have a small spread and fail to depict the strong uncertainties of the TC track at longer lead times; additionally, the BVs are not optimized for growth during forecast periods and therefore cannot depict the strong TC track uncertainties at longer lead times. These could be the reasons why neither the SVs nor BVs generate a large spread or much higher ensemble forecast skill for the track of Megi (2010).

Finally, we interpret what dynamic process is responsible for the growth of the O-CNOPs, SVs, and BVs and reveal with what manner the O-CNOPs yield a large ensemble spread. From the horizontal ME patterns shown in Fig. 14, it is easily found that the O-CNOPs, SVs, and BVs exhibit quite different scale characteristics. A spectrum analysis using the two-dimensional discrete Fourier transform method (Rao et al. 2010) is further performed to quantify the evolution of
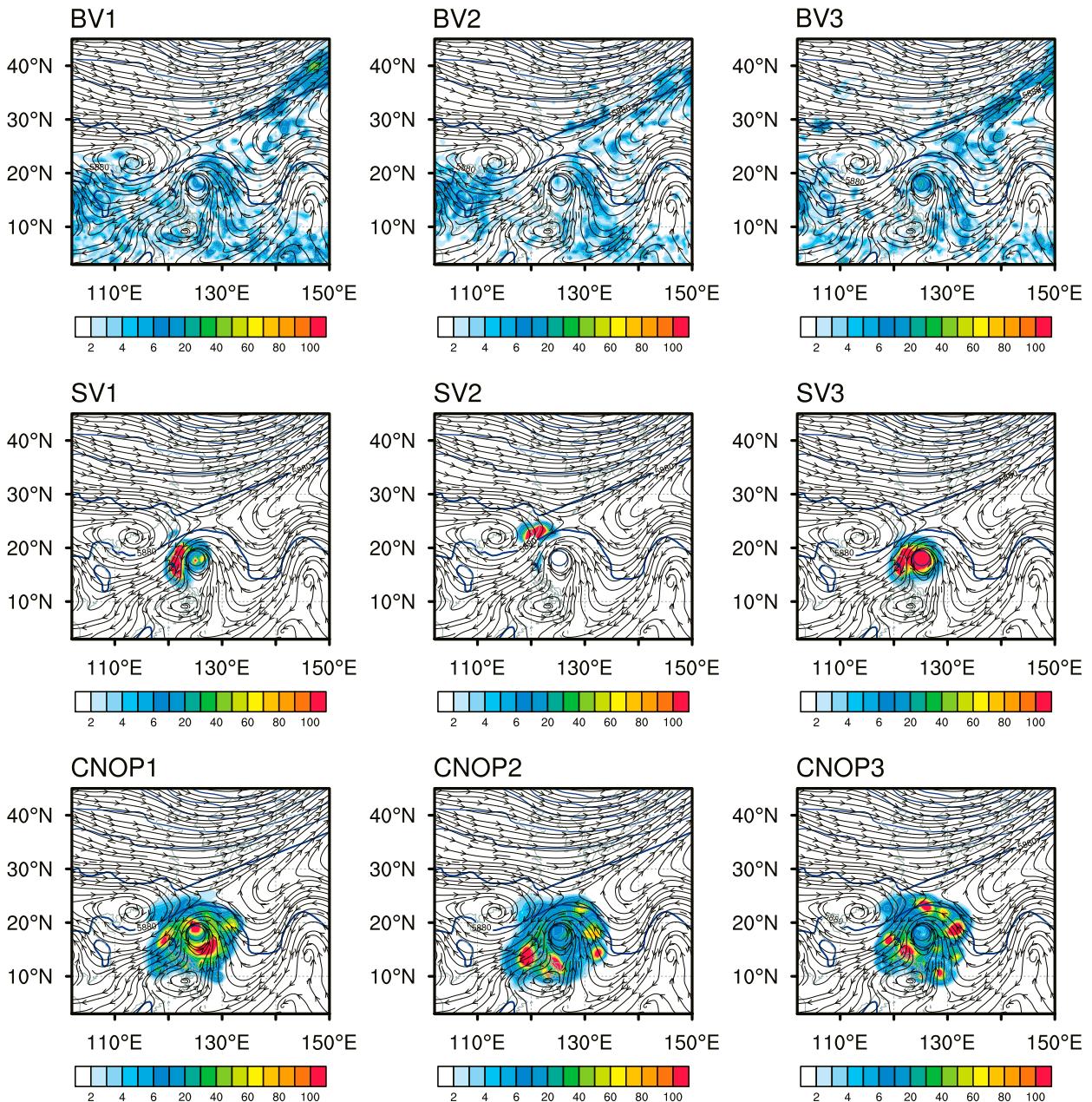
FIG. 14. Horizontal distributions of vertically averaged moist energies (MEs; J kg^{-1}; shaded) for the first three BVs, SVs, and O-CNOPs. The black streamlines denote the environmental steering flows (m s^{-1}), which are the deep-layer mean winds averaged vertically between 850 and 250 hPa. The blue contours denote the geopotential heights at 500 hPa with an interval of 80 gpm, where the high pressure ridge composed of the CH and the SH are indicated by the 5880-gpm bold isopleth.

the BVs, SVs and O-CNOPs at different scales. From Fig. 16a, it can be seen that for the BVs, SVs, and O-CNOPs, the peaks of the power spectra are all located at the mesoscale wavelength of 1200 km at the initial time. However, when comparing their powers, it can be found that the power of the BVs is much larger at wavelengths over 1500 km, whereas that of the O-CNOPs exhibits larger powers at mesoscales from 500 to 1500 km, while the SVs present larger powers at much smaller scales. It is obvious that both the SVs and

O-CNOPs trivially include large-scale flow information. However, the upscale growth of the O-CNOPs within the first 24 h compensates for this effect. Figure 16 shows that the powers induced by the O-CNOPs grow much more rapidly at large scales within a short lead time of 24 h and then are significantly larger than those yielded by the BVs and SVs at scales larger than 400 km. The smaller-scale SVs can also grow upscale but induce large-scale powers to a much smaller degree than the O-CNOPs (see Fig. 16). For the BVs, the mesoscale modes
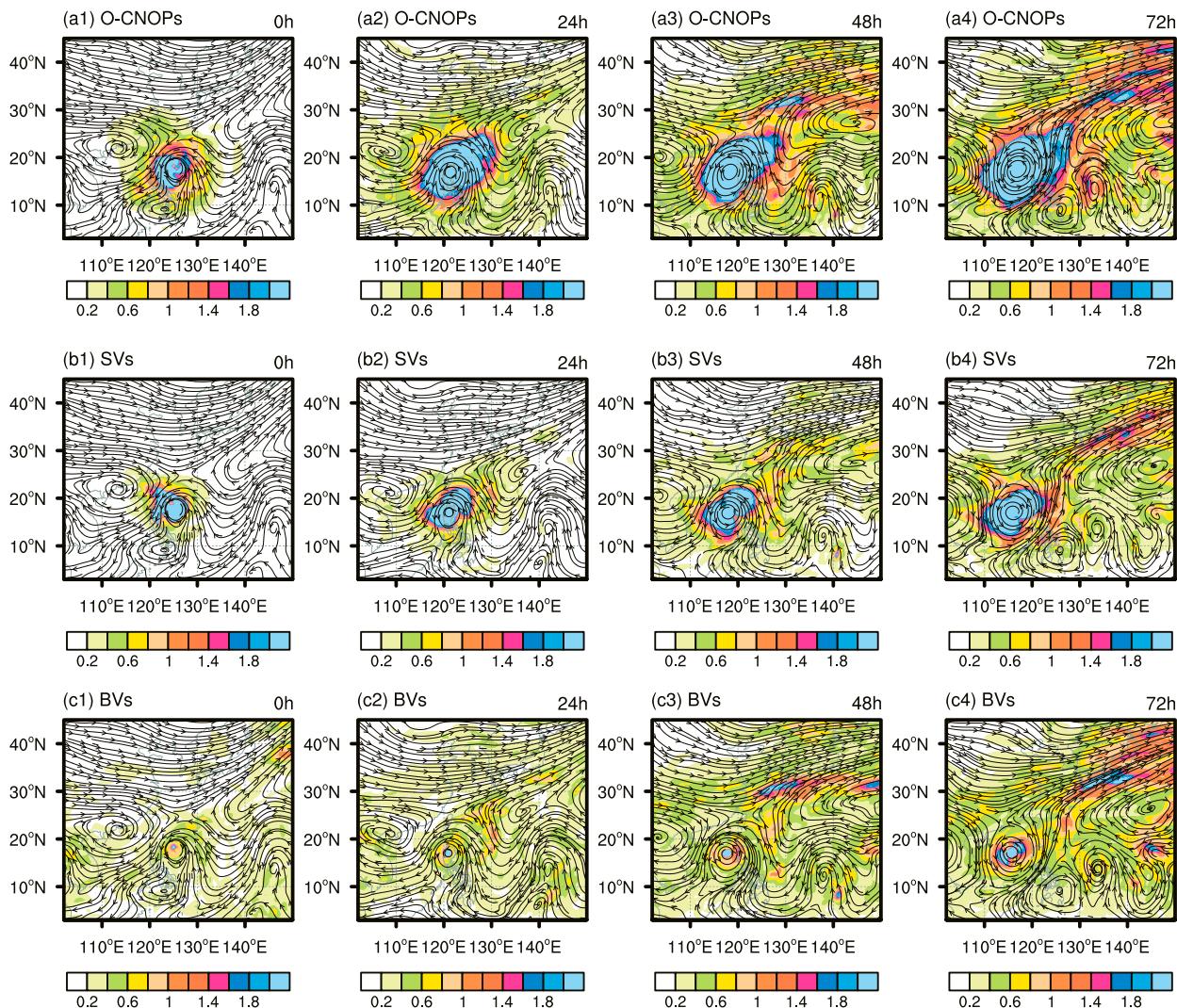
FIG. 15. Ensemble mean of the deep-layer (250–850 hPa) mean winds (black streamlines) and its spread (shaded; m s^{-1}) generated by (a) the O-CNOPs, (b) SVs, and (c) BVs at lead times of 1) 0, 2) 24, 3) 48, and 4) 72 h.

from 200 to 2000 km grow within a 48-h lead time, and after 48 h, the powers at much larger scales start to grow relatively obviously; however, they are far smaller than those of the SVs and O-CNOPs. In this situation, the BVs ensemble may not have the ability to perturb the large-scale flow around Megi (2010) (Thanh et al. 2016). That is, the BVs ensemble cannot depict the uncertainties of the interaction of the TC with the large-scale environment well, which controls the TC motion (Chan and Li 2005). While for the O-CNOPs and SVs, their upscale growth indicates the effect on the nonlinear interaction between the TC and its surrounding environmental flow, and they adjust the steering flows that influence Megi's motion. Nevertheless, since the O-CNOPs possess the largest growth in the power for both the TC circulation and large-scale flows, they yield a much larger ensemble spread of the TC track, which may better capture the uncertainties of TC motion controlled by the nonlinear interaction between the

TC and its surrounding environmental flow and provide much higher ensemble forecast skill.

6. Discussion

Huo et al. (2019) adopted O-CNOPs to conduct ensemble forecast experiments for TCs using the MM5, and among the TC cases selected there, Matsa (2005) and Sepat (2007) are included in the scenario of strong TCs considered in the present study. Therefore, for comparison, these two TCs are further forecasted using the WRF Model (see section 2b). By comparing the results shown in Huo et al. (2019), it is found that the control forecasts in the WRF Model greatly improve over those in the MM5, and even the ensemble mean forecasts further significantly increase the forecasting skill of the TC tracks. Specifically, taking Matsa (2005) as an example, it is shown that from the MM5 to the WRF Model, the track error
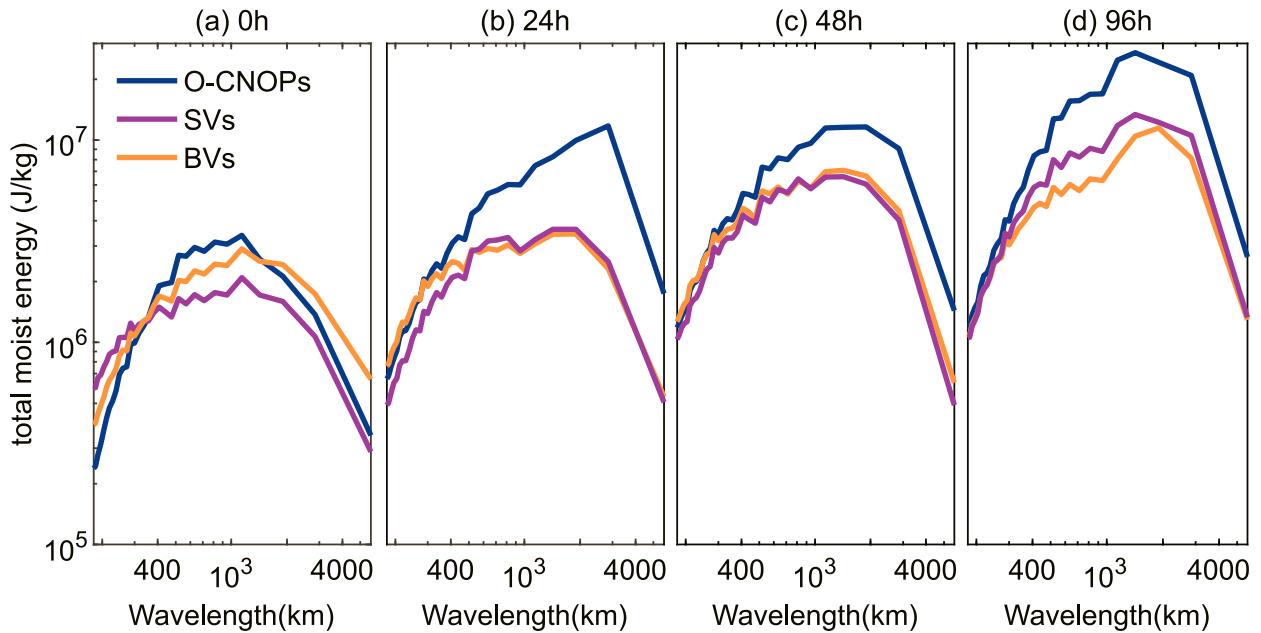
FIG. 16. Ensemble mean of the power spectra of the moist energies generated by the O-CNOPs, SVs, and BVs as a function of wavelength at different lead times.

in the control forecast for a lead time of 5 days is reduced from 311 to 163 km, while the ensemble mean generated by the O-CNOPs further reduces the track error from 163 to 96 km, with the skill improvement increasing to 41% in the WRF Model (which is significantly larger than 14% in the MM5). A possible reason for this is related to the model errors. First, a significant component of the track errors in the MM5 is generated by inherent model uncertainties, causing the reduction in track errors generated by the O-CNOPs to be small relative to the track errors generated by both the initial errors and significant model errors. On the other hand, the model errors are greatly reduced in the WRF Model, and thus, the reduction in track errors by the O-CNOPs becomes more significant relative to the track errors generated by both the initial errors and small model errors.

However, the ensemble mean track errors in the WRF Model generated by the SVs, BVs, and RPs are up to 147, 162, and 165 km, respectively, and the skill is slightly improved or becomes worse than the control forecast. The results indicate that the track errors produced by the initial uncertainties have not been effectively removed with the three traditional methods. It is clear that there exist small model errors in the WRF, and the control forecast skill is greatly improved. The O-CNOPs still contribute more than the SVs, BVs, and RPs to the improvement of the TC track forecast skill. Therefore, it seems that a much more accurate model is favorable for manifesting the positive effect of the O-CNOPs in estimating initial uncertainties and then improving the TC track forecast skill.

7. Summary

In this paper, we apply a new ensemble forecasting method, i.e., the O-CNOPs, to the WRF Model to provide more

skillful forecasts of tropical cyclone (TC) tracks. First, we conduct sensitivity experiments to determine the configurations of the O-CNOPs for achieving a much higher ensemble forecast skill, which, specifically, consists of the ensemble size, the initial perturbation amplitude and the optimization time interval for calculating the O-CNOPs. Then, ensemble forecasts are conducted for 12 TCs, 10 from 2010 to 2022 and 2 of those selected in Huo et al. (2019), with lead times of at least 5 days. Most of these TC cases directly made landfall in China, while the others landed in Korea but greatly influenced China. In particular, these TCs include two unusual tracks and another two that were realistically forecasted by the authors in 2021 and 2022. The results show that the ensemble mean forecasts for the TC tracks generated by the O-CNOPs can significantly improve the skill of the control forecasts in terms of the track errors, and even the landfall position errors are reduced by more than 31% for most TC cases. In particular, the ensembles generated by the O-CNOPs successfully reproduce the sharp northward-turning track of Megi (2010) and the counterclockwise loop track of Tembin (2012) and demonstrate the usefulness of the O-CNOPs in forecasting unusual TC tracks. In addition, the real-time track forecasts of TC In-Fa (2021) are also significantly enhanced by taking the ensembles of the O-CNOPs in terms of the moving track before, during, and after landfall, and the real-time forecast of TC Hinnamnor (2022) is also improved at long lead times. It is clear that the O-CNOPs have the potential to improve the TC track forecast skill in real-time TC forecasts.

Statistically, we evaluate the superiority of the O-CNOPs compared with the traditional SVs, BVs, and RPs according to the ensemble forecasts of the 12 TC cases. The results show that the ensembles generated by the O-CNOPs generally

exhibit a much larger spread and often include the best track within them, while those made by the SVs, BVs, and RPs always miss the best tracks, ultimately causing the O-CNOPs to possess the highest skill in TC track forecasts. A comparison of the skill–spread relationships among the four ensemble forecasting methods suggests that the ensemble spreads generated by the RPs, BVs, and SVs are significantly smaller than the corresponding ensemble mean forecasting errors; however, the O-CNOPs provide additional benefits for increasing the ensemble spread and making it much closer to the ensemble mean forecasting error. For the TC strike probability forecasts, the O-CNOPs also display the highest skill as measured by the BSS and ROCA and provide the greatest reliability in terms of the RD. Therefore, the O-CNOPs can provide a much more credible risk assessment of TCs than the SVs, BVs, and RPs. Furthermore, from a comparison of the results in the MM5 and the WRF, it seems that the O-CNOPs can produce much better TC track forecasting skill using a much more advanced model.

Taking Megi (2010) as an example, we explore why the O-CNOPs achieve a higher ensemble forecast skill than the SVs and the BVs. Specifically, the O-CNOPs are shown to be located mainly in the middle-to-lower troposphere, and the sensitivity provided by them captures the uncertainty in the mid-to-lower troposphere, to which the TC track forecasts are most sensitive. Simultaneously, the O-CNOPs identify the uncertainties that occurred over the annulus region around the TC center that determine when and where Megi (2010) began to deflect. In addition, the O-CNOPs characterize a rapid upscale growth that disturbs the steering flow responsible for Megi's movement much more and better describes the uncertainties of TC motion controlled by the nonlinear interaction between the TC and its surrounding environmental flow. Due to these advantages of the O-CNOPs, they provide much higher ensemble forecast skill for TC tracks than SVs and BVs. However, it is noted that localized maxima exist in the O-CNOPs patterns, and we presently still have no clear idea to interpret this. It is expected that more analysis and thought will contribute to addressing this question in the following study.

Although the ensemble forecasts generated by the O-CNOPs achieve much higher skill in deterministic and probabilistic forecasts, the spread–skill relationship is not yet perfect for measuring the reliability of the ensembles. In the present study, we only consider the effect of initial uncertainties and adopt the O-CNOPs to obtain the ensemble forecasts of TC tracks; however, the effects of model errors are inevitable. The present study regards the O-CNOPs as a new method and, for the purposes of showing "proof of concept" of this method, follows Huo et al. (2019) to adopt a relatively coarse horizontal resolution of 60 km to explore its performance in forecasting TC tracks. For the 15 vertical levels adopted, it is also a very low resolution; however, we have adopted a 60-km horizontal resolution but increased the vertical levels from 15 to 31 and 60 to make comparisons and validated the applicability of the 15 vertical levels (the details are omitted). Certainly, these coarse resolutions would cause model errors and deteriorate TC track forecasts. Duan et al. (2022) demonstrated that

dynamically coordinated growth of initial and model perturbations helps to obtain ensembles of the strongest reliability when initial and model errors exist in forecasts. Therefore, it is expected that a combined mode of initial and model perturbations, such as the C-NFSVs in Duan et al. (2022), will be adopted to perform ensemble forecasts for TC tracks, not only to acquire much more reliable ensembles but also to acquire much higher forecast skill. In addition, the calculation of the O-CNOPs presently requires a large amount of time, which is another reason why much higher resolution in the WRF Model is not applied in the present study. In fact, even for the WRF Model with the above coarse resolution, the calculation of the above O-CNOPs with 64 computing kernels for one forecast takes nearly 2 h when we compute the CNOPs one by one, which is twice the amount of time needed to calculate the SVs. It is understandable that the O-CNOPs are much more expensive to compute due to their consideration of nonlinear processes. Of course, this does not mean that the O-CNOPs cannot be efficiently produced. In fact, we are currently developing a new efficient algorithm, which is similar to the Lanczos algorithm (Simon 1984) for calculating the SVs. We have recently used an idealized model to test it, and the preliminary results are encouraging. Its advantage is that it allows all the CNOPs to be computed at once. That is, we can use the computation time for calculating one CNOP to compute all O-CNOPs. This new efficient algorithm for the O-CNOPs will be reported in a follow-up paper. It is expected that this efficient algorithm will be used at much higher resolutions and extend the capability of O-CNOPs. This also inspires us to further investigate the role of the O-CNOPs in improving the forecast skill of TC intensity and even precipitation using WRF Model with a finer resolution. The performances of O-CNOPs and other methods can also be more clearly clarified at a finer resolution. Besides, TC systems are often related to multiscale interactions, which certainly influence TC track and intensity forecasts (Chan and Li 2005). Therefore, effectively blending multiscale information to the O-CNOPs is also an important issue for ensemble forecasts of TCs, which will be explored in future work.

The O-CNOPs method, as a new method, is theoretically reasonable, but its functions should be gradually revealed and improved step by step. For TC track forecasts of interest, there are different types of TCs (e.g., strong TC cases, weak TC cases, or TC cases with unusual tracks). For each type of TCs, the configuration of the O-CNOPs for achieving high forecast skill could be different. Thus, what is the law of the configurations from strong TCs to weak TCs or other types of TCs? This question should be addressed by comparing the forecasts of different types of TC cases. In the present study, we focus on the forecasts of strong TCs that severely influenced mainland China. The results show that the ensemble forecasts generated by the O-CNOPs outperform the forecasts made by the SVs, BVs, and RPs in terms of both deterministic and probabilistic skills. This result encourages us to further adopt more weak TC cases to examine the O-CNOPs and compare the performance with that of the strong TC cases, thus revealing a law on how to determine the configurations of the O-CNOPs for achieving high forecast skill. It is

believed that the O-CNOPs method will be a useful ensemble forecasting method by gradually improving its functions.

Data availability statement. The datasets generated and/or analyzed during the study are stored on the computers at the State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics (LASG; https://www.lasg.ac.cn) and will be available to researchers upon request. The NCEP FNL dataset for this study can be obtained from the National Center for Atmospheric Research via https://rda.ucar.edu/datasets/ds083.3/. The NCEP GFS forecast data can be downloaded from the National Oceanic and Atmospheric Administration via http://nomads.ncep.noaa.gov. The ECMWF ERA5 data can be obtained from the Climate Data Store via https://cds.climate.copernicus.eu/. The CMA TC best track dataset is available at http://tcdata.typhoon.org.cn.

APPENDIX A

SVs, RPs, and BVs

Singular vectors (SVs) represent a group of initial perturbations that have the largest amplification rates in their respective subphases in the tangent linear model (TLM) and satisfy Eq. (A1):

$$J(\mathbf{x}_0^*) = \max_{x_0 \in \Omega_j} \frac{(PL\mathbf{x}_0)^T \mathbf{C}_2 (PL\mathbf{x}_0)}{\mathbf{x}_0^T \mathbf{C}_1 \mathbf{x}_0}, \qquad (A1)$$

where $\mathbf{x}_0$ is the initial perturbation, $\Omega_j = \{\mathbf{x}_0 \in \mathbb{R}^n | \mathbf{x}_0^T \mathbf{C}_1 \mathbf{x}_0 \leq \delta\}$ is the constraint condition of the initial perturbations, L is the forward propagator of the TLM, and $\mathbf{C}_1$ and $\mathbf{C}_2$ refer to the total moist energy norm. With the TLM and its adjoint model, the SVs can be computed using the Lanczos algorithm (Simon 1984).

Random perturbations (RPs) in the present study can be expressed by Eq. (A2) (see Yu and Zhang 2005):

$$RP(z) = \omega \mathbf{r} E(z), \qquad (A2)$$

where $RP(z)$ is the random perturbations in the vertical layer z, ω is a coefficient that controls the amplitude of the perturbations, $\mathbf{r} = (\mathbf{r}_i)_N$ denotes the ith group of random numbers on horizontal grids (the grid number is N), and $E(z)$ is the root-mean-squared error in layer z, which is calculated by the 6-h forecast field with respect to the NCEP FNL.

The bred vectors (BVs) are a generalization of the local Lyapunov vectors (LLVs) in a nonlinear field (Toth and Kalnay 1997), which solve the fast-growing initial perturbations during a time interval prior to the initial time (Cheung 2001b). The orthogonal BVs used in this study are not reordered according to

their growth rates, which are more similar to nonlinear local Lyapunov vectors (Feng et al. 2016, 2018). In the present study, the BVs are computed 72 h ahead of the initial time, and the breeding cycle period is taken as 6 h. Specifically, the RPs are initially added to the state field 72 h ahead of the initial time of the control forecast, and with this perturbed field as the initial value, the WRF Model is integrated forward for 6 h; thus, a group of perturbed forecasts is obtained. Then, the difference between the control forecast and the perturbed forecast at this time is taken as a new set of perturbations. These new perturbations are further rescaled to the amplitude of the initially superimposed RPs and then orthogonalized. Such a process is often referred to as a breeding cycle. After 12 6-h breeding cycles, a group of orthogonal initial perturbations can be obtained, which are the orthogonal BVs for ensemble forecasting. In fact, we have tried more breeding cycles and found that the resultant BVs possess similar structures when the breeding period exceeds 72 h. To save computing costs, we chose a 72-h breeding period.

APPENDIX B

Forecast Errors of the TC Tracks, Ensemble Spread, and Strike Probability

The forecast errors of the TC tracks are measured by the great-circle distance between a TC center in the best track and that in the forecasts (Neumann and Pelissier 1981), which can be expressed by Eq. (B1):

$$\text{Track error} = 111.11 \cos^{-1}[\sin\varphi_0 \sin\varphi_s + \cos\varphi_0 \cos\varphi_s \cos(\lambda_0 - \lambda_s)], \qquad (B1)$$

where λ_0 and φ_0 are the longitude and latitude of the TC center in the best track, respectively, and λ_s and φ_s are the longitude and latitude in the control forecast or the ensemble mean forecast, respectively. The forecast error of each TC track can be decomposed into two components: one is along the best track (ATE), and the other crosses the best track (CTE). A positive CTE value indicates that the forecasted TC track lies to the right of the best track, and a negative value indicates that it lies to the left; a positive ATE indicates that the forecasted TC moves faster than indicated by the best track, and a negative ATE indicates slower movement (WMO 2013; Heming 2017).

The ensemble spread of the TC track forecasts is defined by the standard deviation of the ensemble members with respect to the ensemble mean (Hopson 2014), which can be calculated by Eq. (B2):

$$\text{spread} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} |F_i - \overline{F}|^2}, \qquad (B2)$$

where N represents the ensemble size, F_i is the track of the ith ensemble member, $\overline{F}$ is the ensemble mean, and $|F_i - \overline{F}|$ is the great-circle distance between F_i and $\overline{F}$. A perfect ensemble forecast is proven to have a relationship in which

the ensemble spread is equal to the ensemble mean forecasting error (Bowler 2006; Hopson 2014).

The TC strike probability at a spatial location is defined as the probability that a TC will pass within a 120-km radius from the location during the next 120 h (WMO 2013). The TC strike probability skill can be evaluated by the Brier score (BS), Brier skill score (BSS), relative operating characteristic (ROC), and reliability diagram (RD).

(i) The BS is the mean squared error of the probability forecasts defined by Eq. (B3):

$$BS = \frac{1}{N}\sum_{i=1}^{N}(f_i - o_i)^2,   \text{(B3)}$$

where N is the number of realizations of the prediction processed, and f_i and o_i are the probabilities of the forecast and observation for the prediction processes, respectively. A smaller BS indicates a better probability forecast skill (Brier 1950).

(ii) The BSS assesses the improvement of the BS for a given forecast with respect to a reference forecast (i.e., the control forecast in the present study). A BSS of 0 indicates no skill, and a BSS of 1 represents a perfect forecast with the highest skill (Titley et al. 2020).

(iii) The ROC curve measures the ability of a forecast to discriminate between events and nonevents. This curve is a function of the hit rate and false alarm rate. Thus, higher skill is indicated by an ROC curve that is closer to the top-left corner of the diagram (implying a low false alarm rate and high hit rate) or a larger area under the curve (ROCA) (Mason and Graham 2002).

(iv) The RD curve displays how well the forecast probabilities correspond to the observed frequencies. Therefore, for a perfect ensemble forecast, the RD curve should be on the diagonal. A line above the diagonal indicates under-forecasting, meaning that the forecast probabilities are lower than the observed frequencies; in contrast, a line below the diagonal indicates over-forecasting, i.e., that the event is forecasted more often than it is actually observed (Dube et al. 2020).

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, https://doi.org/10.1175/1520-0493(1997)125<2969:TIODCO>2.0.CO;2.

Birgin, E. G., J. M. Martínez, and M. Raydan, 2000: Nonmonotone spectral projected gradient methods on convex sets. *SIAM J. Optim.*, **10**, 1196–1211, https://doi.org/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, https://doi.org/10.1111/j.1600-0870.2006.00197.x.

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

Buizza, R., 1994: Sensitivity of optimal unstable structures. *Quart. J. Roy. Meteor. Soc.*, **120**, 429–451, https://doi.org/10.1002/qj.49712051609.

Chan, J. C. L., and K. K. Li, 2005: Ensemble forecasting of tropical cyclone motion using a barotropic model. Part III: Combining perturbations of the environment and the vortex. *Meteor. Atmos. Phys.*, **90**, 109–126, https://doi.org/10.1007/s00703-004-0092-9.

Cheung, K. K. W., 2001a: A review of ensemble forecasting techniques with a focus on tropical cyclone forecasting. *Meteor. Appl.*, **8**, 315–332, https://doi.org/10.1017/S1350482701003073.

——, 2001b: Ensemble forecasting of tropical cyclone motion: Comparison between regional bred modes and random perturbations. *Meteor. Atmos. Phys.*, **78**, 23–34, https://doi.org/10.1007/s007030170003.

Diaconescu, E. P., and R. Laprise, 2012: Singular vectors in atmospheric sciences: A review. *Earth-Sci. Rev.*, **113**, 161–175, https://doi.org/10.1016/j.earscirev.2012.05.005.

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

——, and Z. Huo, 2016: An approach to generating mutually independent initial perturbations for ensemble forecasts: Orthogonal conditional nonlinear optimal perturbations. *J. Atmos. Sci.*, **73**, 997–1014, https://doi.org/10.1175/JAS-D-15-0138.1.

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

——, P. Zhao, J. Hu, and H. Xu, 2016: The role of nonlinear forcing singular vector tendency error in causing the "spring predictability barrier" for ENSO. *J. Meteor. Res.*, **30**, 853–866, https://doi.org/10.1007/s13351-016-6011-4.

——, X. Li, and B. Tian, 2018: Towards optimal observational array for dealing with challenges of El Niño-Southern Oscillation predictions due to diversities of El Niño. *Climate Dyn.*, **51**, 3351–3368, https://doi.org/10.1007/s00382-018-4082-x.

——, J. Ma, and S. Vannitsem, 2022: An ensemble forecasting method for dealing with the combined effects of the initial errors and the model errors and a potential deep learning implementation. *Mon. Wea. Rev.*, **150**, 2959–2976, https://doi.org/10.1175/MWR-D-22-0007.1.

Dube, A., R. Ashrit, S. Kumar, and A. Mamgain, 2020: Improvements in tropical cyclone forecasting through ensemble prediction system at NCMRWF in India. *Trop. Cyclone Res. Rev.*, **9**, 106–116, https://doi.org/10.1016/j.tcrr.2020.04.003.

Dudhia, J., 1989: Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107, https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2.

Ehrendorfer, M., R. M. Errico, and K. D. Raeder, 1999: Singular-vector perturbation growth in a primitive equation model with moist physics. *J. Atmos. Sci.*, **56**, 1627–1648, https://doi.org/10.1175/1520-0469(1999)056<1627:SVPGIA>2.0.CO;2.

Epstein, E. S., 1969: Stochastic dynamic prediction. *Tellus*, **21**, 739–759, https://doi.org/10.3402/tellusa.v21i6.10143.

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. *Adv. Atmos. Sci.*, **33**, 1036–1046, https://doi.org/10.1007/s00376-016-6003-4.

——, ——, ——, and Z. Toth, 2018: Comparison of nonlinear local Lyapunov vectors and bred vectors in estimating the spatial distribution of error growth. *J. Atmos. Sci.*, **75**, 1073–1087, https://doi.org/10.1175/JAS-D-17-0266.1.

Feng, R., and W.-S. Duan, 2014: The spatial patterns of initial errors related to the "winter predictability barrier" of the Indian Ocean dipole. *Atmos. Oceanic Sci. Lett.*, **7**, 406–410, https://doi.org/10.3878/j.issn.1674-2834.14.0018.

Goerss, J. S., C. R. Sampson, and J. M. Gross, 2004: A history of western North Pacific tropical cyclone track forecast skill. *Wea. Forecasting*, **19**, 633–638, https://doi.org/10.1175/1520-0434(2004)019<0633:AHOWNP>2.0.CO;2.

Hamill, T. M., C. Snyder, and R. E. Morss, 2000: A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles. *Mon. Wea. Rev.*, **128**, 1835–1851, https://doi.org/10.1175/1520-0493(2000)128<1835:ACOPFF>2.0.CO;2.

Heming, J. T., 2017: Tropical cyclone tracking and verification techniques for Met Office numerical weather prediction models. *Meteor. Appl.*, **24**, 1–8, https://doi.org/10.1002/met.1599.

Hopson, T. M., 2014: Assessing the ensemble spread–error relationship. *Mon. Wea. Rev.*, **142**, 1125–1142, https://doi.org/10.1175/MWR-D-12-00111.1.

Huo, Z., W. Duan, and F. Zhou, 2019: Ensemble forecasts of tropical cyclone track with orthogonal conditional nonlinear optimal perturbations. *Adv. Atmos. Sci.*, **36**, 231–247, https://doi.org/10.1007/s00376-018-8001-1.

Jiang, Z., M. Mu, and D. Wang, 2009: Ensemble prediction experiments using conditional nonlinear optimal perturbation. *Sci. China.*, **52D**, 511–518, https://doi.org/10.1007/s11430-009-0042-y.

Kain, J. S., 2004: The Kain–Fritsch convective parameterization: An update. *J. Appl. Meteor.*, **43**, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418, https://doi.org/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, https://doi.org/10.1016/j.jcp.2007.02.014.

Li, H., J. Luo, and M. Xu, 2019: Ensemble data assimilation and prediction of typhoon and associated hazards using TEDAPS: Evaluation for 2015–2018 seasons. *Front. Earth Sci.*, **13**, 733–743, https://doi.org/10.1007/s11707-019-0794-4.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22**, 1065–1092, https://doi.org/10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

Liu, J. C., Y. A. Liou, M. X. Wu, Y. J. Lee, C. H. Cheng, C. P. Kuei, and R. M. Hong, 2015: Analysis of interactions among two tropical depressions and Typhoons Tembin and Bolaven (2012) in Pacific Ocean by using satellite cloud images. *IEEE Trans. Geosci. Remote Sens.*, **53**, 1394–1402, https://doi.org/10.1109/TGRS.2014.2339220.

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

——, 1996: Predictability: A problem partly solved. *Proc. Workshop on Predictability*, Shinfield Park, Reading, ECMWF, 1–18, https://www.ecmwf.int/en/elibrary/75462-predictability-problem-partly-solved.

Lu, X. Q., H. Yu, M. Ying, B. Zhao, S. Zhang, L. Lin, L. Bai, and R. Wan, 2021: Western North Pacific tropical cyclone database created by the China Meteorological Administration. *Adv. Atmos. Sci.*, **38**, 690–699, https://doi.org/10.1007/s00376-020-0211-7.

Luan, X., X. P. Cui, S. T. Gao, and Y. J. Huang, 2015: Cause analysis of sudden track change of Typhoon Megi (2010). *Daqi Kexue Xuebao*, **38**, 658–669.

Magnusson, L., M. Leutbecher, and E. Källén, 2008: Comparison between singular vectors and breeding vectors as initial perturbations for the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, **136**, 4092–4104, https://doi.org/10.1175/2008MWR2498.1.

——, J. D. Doyle, W. A. Komaromi, R. D. Torn, C. K. Tang, J. C. L. Chan, M. Yamaguchi, and F. Zhang, 2019: Advances in understanding difficult cases of tropical cyclone track forecasts. *Trop. Cyclone Res. Rev.*, **8**, 109–122, https://doi.org/10.1016/j.tcrr.2019.10.001.

Majumdar, S. J., and P. M. Finocchio, 2010: On the ability of global ensemble prediction systems to predict tropical cyclone track probabilities. *Wea. Forecasting*, **25**, 659–680, https://doi.org/10.1175/2009WAF2222327.1.

Mamgain, A., A. Sarkar, and E. N. Rajagopal, 2019: Medium-range global ensemble prediction system at 12 km horizontal resolution and its preliminary validation. *Meteor. Appl.*, **27**, e1867, https://doi.org/10.1002/met.1867.

Mason, S. J., and N. E. Graham, 2002: Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quart. J. Roy. Meteor. Soc.*, **128**, 2145–2166, https://doi.org/10.1256/003590002320603584.

Miyachi, T., and T. Enomoto, 2021: Tropical cyclone track forecasts using NCEP-GFS with initial conditions from three analyses. *SOLA*, **17**, 140–144, https://doi.org/10.2151/sola.2021-025.

Mlawer, E. J., S. J. Taubman, P. D. Brown, M. J. Iacono, and S. A. Clough, 1997: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.*, **102**, 16663–16682, https://doi.org/10.1029/97JD00237.

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

——, R. Buizza, T. N. Palmer, and T. Petroliagis, 1996: The ECMWF ensemble prediction system: Methodology and validation. *Quart. J. Roy. Meteor. Soc.*, **122**, 73–119, https://doi.org/10.1002/qj.49712252905.

Moon, M., and K.-J. Ha, 2019: Effect of typhoon-generated cold wake on the subsequent Typhoon Tembin and its sensitivity to horizontal resolutions. *Atmosphere*, **10**, 644, https://doi.org/10.3390/atmos10110644.

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

——, F. Zhou, and H. Wang, 2009: A method for identifying the sensitive areas in targeted observations for tropical cyclone prediction: Conditional nonlinear optimal perturbation. *Mon. Wea. Rev.*, **137**, 1623–1639, https://doi.org/10.1175/2008MWR2640.1.

Neumann, C. J., and J. M. Pelissier, 1981: Models for the prediction of tropical cyclone motion over the North Atlantic: An operational evaluation. *Mon. Wea. Rev.*, **109**, 522–538, https://doi.org/10.1175/1520-0493(1981)109<0522:MFTPOT>2.0.CO;2.

Pattanayak, S., and U. C. Mohanty, 2008: A comparative study on performance of MM5 and WRF models in simulation of tropical cyclones over Indian seas. *Curr. Sci.*, **95**, 923–936.

Puri, K., J. Barkmeijer, and T. N. Palmer, 2001: Ensemble prediction of tropical cyclones using targeted diabatic singular vectors. *Quart. J. Roy. Meteor. Soc.*, **127**, 709–731, https://doi.org/10.1002/qj.49712757222.

Qian, C., F. Zhang, B. W. Green, J. Zhang, and X. Zhou, 2013: Probabilistic evaluation of the dynamics and prediction of Super Typhoon Megi (2010). *Wea. Forecasting*, **28**, 1562–1577, https://doi.org/10.1175/WAF-D-12-00121.1.

Qin, X., 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, https://doi.org/10.1002/qj.2109.

——, W. Duan, P.-W. Chan, B. Chen, and K.-N. Huang, 2022: Effects of dropsonde data in field campaigns on forecasts of tropical cyclones over the western North Pacific in 2020 and the role of CNOP sensitivity. *Adv. Atmos. Sci.*, **40**, 791–803, https://doi.org/10.1007/s00376-022-2136-9.

Rao, K. R., D. N. Kim, and J. J. Hwang, Eds., 2010: Two-dimensional discrete Fourier transform. *Fast Fourier Transform—Algorithms and Applications*, Springer, 127–184.

Rappaport, E. N., and Coauthors, 2009: Advances and challenges at the National Hurricane Center. *Wea. Forecasting*, **24**, 395–419, https://doi.org/10.1175/2008WAF2222128.1.

Sarkar, A., S. Kumar, A. Dube, S. K. Prasad, A. Mamgain, P. Chakraborty, R. Ashrit, and A. K. Mitra, 2021: Forecasting of tropical cyclone using global and regional ensemble prediction systems of NCMRWF: A review. *MAUSAM*, **72**, 77–86, https://doi.org/10.54302/mausam.v72i1.131.

Shi, W., J. Fei, X. Huang, X. Cheng, J. Ding, and Y. He, 2014: A numerical study on the combined effect of midlatitude and low-latitude systems on the abrupt track deflection of Typhoon Megi (2010). *Mon. Wea. Rev.*, **142**, 2483–2501, https://doi.org/10.1175/MWR-D-13-00283.1.

Simon, H. D., 1984: The Lanczos algorithm with partial reorthogonalization. *Math. Comput.*, **42**, 115–142, https://doi.org/10.2307/2007563.

Sun, Y., Z. Zhong, L. Yi, T. Li, M. Chen, H. Wan, Y. Wang, and K. Zhong, 2015: Dependence of the relationship between the tropical cyclone track and western Pacific subtropical high intensity on initial storm size: A numerical investigation. *J. Geophys. Res. Atmos.*, **120**, 11 451–11 467, https://doi.org/10.1002/2015JD023716.

Tallapragada, V., and Coauthors, 2015: Forecasting tropical cyclones in the western North Pacific basin using the NCEP operational HWRF: Real-time implementation in 2012. *Wea. Forecasting*, **30**, 1355–1373, https://doi.org/10.1175/WAF-D-14-00138.1.

Thanh, C., T. T. Tien, and K. Q. Chanh, 2016: Application of breeding ensemble to tropical cyclone track forecasts using the regional atmospheric modeling system (RAMS) model. *Appl. Math. Model.*, **40**, 8309–8325, https://doi.org/10.1016/j.apm.2016.04.010.

Titley, H. A., R. L. Bowyer, and H. L. Cloke, 2020: A global evaluation of multi-model ensemble tropical cyclone track probability forecasts. *Quart. J. Roy. Meteor. Soc.*, **146**, 531–545, https://doi.org/10.1002/qj.3712.

Torn, R. D., T. J. Elless, P. P. Papin, and C. A. Davis, 2018: Tropical cyclone track sensitivity in deformation steering flow. *Mon. Wea. Rev.*, **146**, 3183–3201, https://doi.org/10.1175/MWR-D-18-0153.1.

Toth, Z., and E. Kalnay, 1997: Ensemble forecasting at NCEP and the breeding method. *Mon. Wea. Rev.*, **125**, 3297–3319, https://doi.org/10.1175/1520-0493(1997)125<3297:EFANAT>2.0.CO;2.

Tseng, J. C.-H., and Y.-S. Lai, 2020: Perturbation structure and evolution in tropical cyclones Noul and Nepartak based on singular vectors. *Tellus*, **72A**, 1–26, https://doi.org/10.1080/16000870.2020.1814589.

Wang, C. X., 2014: Ensemble prediction of Typhoon Muifa's track and intensity. *J. Trop. Oceanogr.*, **33**, 48–60.

——, and Y. Q. Ni, 2011: Sensitivity experiments of impacting tropical cyclone track. *Acta Meteor. Sin.*, **69**, 757–769.

Wang, H., L. Dong, Y. Xu, and G. Nie, 2022: Analysis on main characteristics of Typhoon In-fa and difficulties in its track forecast. *J. Mar. Meteor.*, **42**, 83–91.

Wang, Y., and W. Duan, 2019: Influences of initial perturbation amplitudes and ensemble sizes on the ensemble forecasts made by CNOPs method. *Chin. J. Atmos. Sci.*, **43**, 915–929, https://doi.org/10.3878/j.issn1006-9895.1810.18219.

——, Y. Sun, Q. Liao, Z. Zhong, Y. Hu, and K. Liu, 2017: Impact of initial storm intensity and size on the simulation of tropical cyclone track and western Pacific subtropical high extent. *J. Meteor. Res.*, **31**, 946–954, https://doi.org/10.1007/s13351-017-7024-3.

WMO, 2013: Verification methods for tropical cyclone forecasts. WWRP/WGNE Joint Working Group on Forecast Verification Research, WWRP 2013-7, 89 pp., https://filecloud.wmo.int/share/s/Fotf-7H1RLyK1lkSC_onBA.

Xian, Z., and K. Chen, 2019: Numerical analysis on the effects of binary interaction between Typhoons Tembin and Bolaven in 2012. *Adv. Meteor.*, **2019**, 7529263, https://doi.org/10.1155/2019/7529263.

Xiang, C., Y. Xu, S. Gao, Q. Wang, and H. Wang, 2022: Analysis of the characteristics and forecast difficulties of TCs over the western North Pacific in 2021. *Meteor. Mon.*, **48**, 1195–1208, https://doi.org/10.7519/j.issn.1000-0526.2022.053001.

Xiao, Q., and Coauthors, 2008: Application of an adiabatic WRF adjoint to the investigation of the May 2004 McMurdo, Antarctica, severe wind event. *Mon. Wea. Rev.*, **136**, 3696–3713, https://doi.org/10.1175/2008MWR2235.1.

Yamaguchi, M., and S. J. Majumdar, 2010: Using TIGGE data to diagnose initial perturbations and their growth for tropical cyclone ensemble forecasts. *Mon. Wea. Rev.*, **138**, 3634–3655, https://doi.org/10.1175/2010MWR3176.1.

——, R. Sakai, M. Kyoda, T. Komori, and T. Kadowaki, 2009: Typhoon ensemble prediction system developed at the Japan Meteorological Agency. *Mon. Wea. Rev.*, **137**, 2592–2604, https://doi.org/10.1175/2009MWR2697.1.

Yang, S., B. Chen, F. Zhang, and Y. Hu, 2022: Characteristics and causes of extremely persistent heavy rainfall of tropical cyclone In-Fa (2021). *Atmosphere*, **13**, 398, https://doi.org/10.3390/atmos13030398.

Yu, Y.-F., and L.-F. Zhang, 2005: A study of initial perturbation saturation in ensemble prediction based on the "breeding of growing modes" method. *Chin. J. Atmos. Sci.*, **29**, 955–964, https://doi.org/10.3878/j.issn.1006-9895.2005.06.11.

Zhang, X., X.-Y. Huang, and N. Pan, 2013: Development of the upgraded tangent linear and adjoint of the Weather Research and Forecasting (WRF) Model. *J. Atmos. Oceanic Technol.*, **30**, 1180–1188, https://doi.org/10.1175/JTECH-D-12-00213.1.

Zhou, X., and J. C. L. Chen, 2006: Ensemble forecasting of tropical cyclone motion using a baroclinic model. *Adv. Atmos. Sci.*, **23**, 342–354, https://doi.org/10.1007/s00376-006-0342-5.

——, Y. Zhu, D. Hou, Y. Luo, J. Peng, and R. Wobus, 2017: Performance of the New NCEP global ensemble forecast system in a parallel experiment. *Wea. Forecasting*, **32**, 1989–2004, https://doi.org/10.1175/WAF-D-17-0023.1.