



Ensemble-based sensitivity analysis of track forecasts of typhoon In-fa (2021) without and with model errors in the ECMWF, NCEP, and CMA ensemble prediction systems

Liangying Liu ^a, Jie Feng ^{a,b,c,*}, Li Ma ^a, Yanru Yang ^a, Xiaotian Wu ^a, Chao Wang ^b

^a Department of Atmospheric and Oceanic Sciences and Institute of Atmospheric Sciences, Fudan University, Shanghai, China

^b Key Laboratory of Meteorological Disaster (KLMD), Ministry of Education & Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters (CIC-FEMD), Nanjing University of Information Science and Technology, Nanjing, China

^c Shanghai Key Laboratory of Ocean-land-atmosphere Boundary Dynamics and Climate Change, Fudan University, Shanghai, China

ARTICLE INFO

Keywords:

Ensemble sensitivity analysis
Tropical cyclone track forecast error
Model comparison
Model performance

ABSTRACT

Ensemble-based sensitivity analysis (ESA) has been widely applied to identifying and investigating the sources of forecast uncertainty in tropical cyclone (TC) track. The standard ESA used in most preceding studies involves calculating the time-lagged covariance of ensemble perturbations by removing the ensemble mean. This method primarily focuses on the influence of initial errors. However, such studies ignore two critical dimensions of ESA. One is how ensemble sensitivity is influenced by the varying forecast performance across different ensemble prediction systems (EPSs). Secondly, the impact of model errors on forecast uncertainties remains unaddressed. An in-depth examination of these two aspects provides a more comprehensive understanding of the factors contributing to TC track uncertainties.

This study employed the standard ESA to analyze and compare the sources of uncertainty in the track forecast of Typhoon In-fa (2021) in three representative operational EPSs. Our findings reveal that the EPS's specific performance in ensemble spread markedly influence ensemble sensitivity. We identified that variations in the shape and location of key synoptic systems, such as the western Pacific subtropical high and monsoon trough, across ensemble members were notably distinct. These variations played a significant role in shaping the uncertainty for In-fa's track forecast within each system. Furthermore, we introduced a modified ESA to better account for the influence of model errors on TC track uncertainties. The modified ESA, when applied to ensembles with substantial systematic deviations, predominantly reflects the impact of model errors on track forecast inaccuracies, offering a notably different perspective compared to the standard ESA.

Significance statement: Ensemble-based sensitivity analysis (ESA) serves as an effective method for identifying the origins of forecast uncertainty in tropical cyclone (TC) tracks. Most preceding studies based on the standard ESA have predominantly concentrated on examining key physical processes within individual ensemble prediction system (EPS) and the impact of initial condition uncertainties. They often overlook the dependence of the ensemble sensitivity on the varying forecast performance of EPSs, as well as the impact of model errors on sensitivity assessment. This study revealed that the ensemble forecasts across different EPSs exhibit different performance in terms of ensemble spread in the shape and location of primary weather systems. This variability significantly contributes to the uncertainty in TC track forecasts through diverse processes. Furthermore, this research introduced a modified ESA, which effectively identifies the impact of model systematic deviations on TC track forecast errors. This approach along with the standard ESA provide a more comprehensive and nuanced analysis of the forecast error sensitivity associated with both initial condition and model related uncertainties.

1. Introduction

Tropical cyclones (TCs) present as a significant natural disaster, often

resulting in substantial economic damage and loss of life. To mitigate these impacts, it is essential to develop accurate predictions of TC tracks, especially regarding the timing and location of landfall. Over recent

* Corresponding author at: Department of Atmospheric and Oceanic Sciences and Institute of Atmospheric Sciences, Fudan University, Shanghai, China.
E-mail address: fengjief@fudan.edu.cn (J. Feng).

decades, the accuracy of TC track forecasting has improved substantially attributed to advancements in observing systems, the refinement of numerical weather prediction (NWP) models, and the expansion of computing capabilities. Specifically, the 2-day forecast track errors for typhoons (TCs occurring in the western North Pacific) have reduced from over 500 km in the 1980s to roughly 100 km in recent years (Goerss et al., 2004; Yu et al., 2022; Chen et al., 2023). However, despite these advancements, some TC cases present complex motions such as looping, meandering, and stalling that could lead to track forecast errors extending several hundred kilometers (e.g., Heming et al., 2019; Magnusson et al., 2019; Leonardo and Colle, 2021; Tang et al., 2021). Therefore, understanding the origins of these substantial track errors remains a pivotal area of investigation in the field of TC track prediction.

It has long been recognized that TC movement is predominantly controlled by large-scale environmental wind flow (George and Gray, 1976; Franklin et al., 1996; Sobel and Camargo, 2005; Galarneau and Davis, 2013; Zhang et al., 2020). The effect of the TC environmental flow is commonly approximated by the “steering wind,” a concept that refers to the deep-layer mean wind field, calculated as an average over specified horizontal areas and vertical levels. Based on a statistical analysis of observational data, Chan and Gray (1982) found that the steering flow between 500 and 700 hPa in the mid-troposphere shows the strongest correlation with typhoon tracks. Subsequent studies (e.g., Holland, 1984; Wu et al., 2011) also focused on identifying the optimal steering wind but used the average within a slightly different vertical depth, i.e., from 850 to 300 hPa. Velden and Leslie (1991) demonstrated the dependence of the selection of pressure levels on a TC’s intensity and structure.

Ensemble forecasts, which involve initialized forecasts from slightly perturbed states, have been widely applied to studying the predictability of TC tracks. The approach of ensemble-based sensitivity analysis (ESA) was notably advanced by Torn and Hakim (2008), who demonstrated its effectiveness in assessing the impacts of small variations in initial conditions on TC predictions. Generally, ESA for TC track predictability is calculated using the time-lagged ensemble covariance between the forecasted position of a TC center at a specified lead time and the prognostic variable fields at preceding times (e.g., Chang et al., 2013; Ito and Wu, 2013; Torn et al., 2018; Ashcroft et al., 2021; Hazelton et al., 2023; Nakano et al., 2023).

Specific applications of ESA provide deeper insights into TC track predictability. For instance, Ito and Wu (2013) applied ESA to the forecast track uncertainties of typhoons Shanshan (2006) and Dolphin (2008). Their findings revealed that the most sensitive regions affecting the TC track were predominantly located in the mid-troposphere, characterized by a horizontal dipole pattern of vorticity near the TC center. Similarly, Chang et al. (2013) employed a sensitivity analysis based on the principal component decomposition of the 50-member ensemble of the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble prediction system (EPS). This study highlighted the influential role of an upper-tropospheric trough and ridge in determining the track uncertainty of a winter storm. Torn et al. (2015) investigated the sources of the erroneous eastward track prediction of Hurricane Sandy (2012) using ESA. They concluded that variations in an upper-tropospheric ridge north of Sandy, associated with moisture perturbation, may induce the track forecast uncertainty. Extending this line of research, Torn et al. (2018) examined the sensitivity of the track forecasts of hurricanes Debby (2012) and Joaquin (2015) in deformation steering zones. Their analysis, utilizing ECMWF ensemble forecast data, revealed a close relationship between the position and structure of the steering flow near the TC (within 500 km) and the forecast track uncertainty.

Sensitivity analysis methods can be broadly divided into two categories: the dynamic and ensemble-based techniques. The dynamic approaches include the adjoint sensitivity analysis such as Singular Vectors (Diaconescu and Laprise, 2012) and Conditional Nonlinear Optimal Perturbations (Mu et al., 2003; Duan and Huo, 2016). The other type

termed as ESA stands out due to its proficiency in rapidly and effectively identifying the origins of uncertainty in TC track forecasts (Zheng et al., 2013; Torn et al., 2015; Feng et al., 2024). However, it’s notable that the majority of relevant research, including works by Zhang and Yu (2017); Huang et al. (2020); Li et al. (2021); Peng et al. (2022b), predominantly utilizes the ensemble members from a single EPS. This approach inherently limits the understanding of TC track uncertainty to the specific model configurations and ensemble initialization schemes of the chosen EPS. Therefore, the time-lagged ensemble covariance in ESA, which encapsulates potential physical and dynamic regimes, can vary substantially across different EPSs (Alaka et al., 2019). However, comparative studies that evaluate ESA for TC track uncertainty across multiple EPSs are markedly scarce. Such comparative analyses could illuminate the nuanced differences in the sources of TC track uncertainty and elucidate how these differences are interlinked with the setup and performance of various EPSs.

The calculation of the time-lagged ensemble covariance in standard ESA simply relies on the differences among parallel ensemble forecasts, specifically the perturbations from the ensemble mean. This approach does not account for potential discrepancies between the model’s ensemble forecasts and actual atmospheric conditions. As a result, the conclusions drawn from ESA-based studies about the origins of track uncertainty may not be directly transferable to real-world forecasting scenarios. In practical forecasts, the uncertainties in the prognostic fields and TC positions are calculated relative to actual observations, as opposed to the mean state of the ensemble members. Therefore, it is important to examine how potential model inaccuracies influence the uncertainty in forecasted TC tracks and how these influences diverge from those observed in standard ESA. This investigation could provide a more comprehensive understanding of the error sources in practical TC track forecasts. Notably, the impact of model errors on forecast uncertainty has been largely overlooked in prior research.

Typhoon In-fa, a formidable TC, emerged and intensified in the western North Pacific Ocean in July 2021. In-fa maintained a high intensity, exceeding 30 m s^{-1} , for several days subsequent to its escalation to a tropical storm around 1200 UTC on July 18. After making a nearly right-angled turn from a westward to a northward trajectory, In-fa reached its peak intensity at approximately 55 m s^{-1} . Then, In-fa veered northwestward and made landfall over China’s eastern coast (see Fig. 1), causing heavy rainfall and strong gales in various coastal cities. Several major operational forecasting centers encountered challenges in accurately predicting Typhoon In-fa’s trajectory, with errors spanning several hundred kilometers (Wang et al., 2022). Specifically, the European Centre for Medium-Range Weather Forecasts (ECMWF) presented a 72-h deterministic forecast error of around 270 km. In comparison, both the regional operational system of China Meteorological Administration (CMA) and the global operational system of the U.S. National Centers for Environmental Prediction (NCEP) demonstrated even larger errors, each exceeding 330 km for the same forecast duration. These inaccuracies are above the 2021 average global and regional TC position error of about 300 km across major operational forecast centers (Chen et al., 2023).

This specific case prompted this study to explore the principal factors and processes leading to significant inaccuracies in the forecast track. Different from previous research, this study concentrates on examining and contrasting the sources of TC track uncertainty across three diverse EPSs, i.e., the ECMWF (Palmer, 2019), NCEP (Zhou et al., 2022), and CMA (Peng et al., 2022a). These three operational forecasting centers, each employing distinct ensemble initialization schemes and NWP models (see details in Table 1), provide a robust framework for this comparative analysis. Additionally, this study introduces and applies a modified ESA that incorporates model errors, enhancing the assessment of how these errors influence the uncertainty in forecasting TC tracks. The analysis specifically focuses on two critical phases of TC In-fa’s trajectory: its westward progression and the subsequent sharp northward turn followed by northwestward movement.

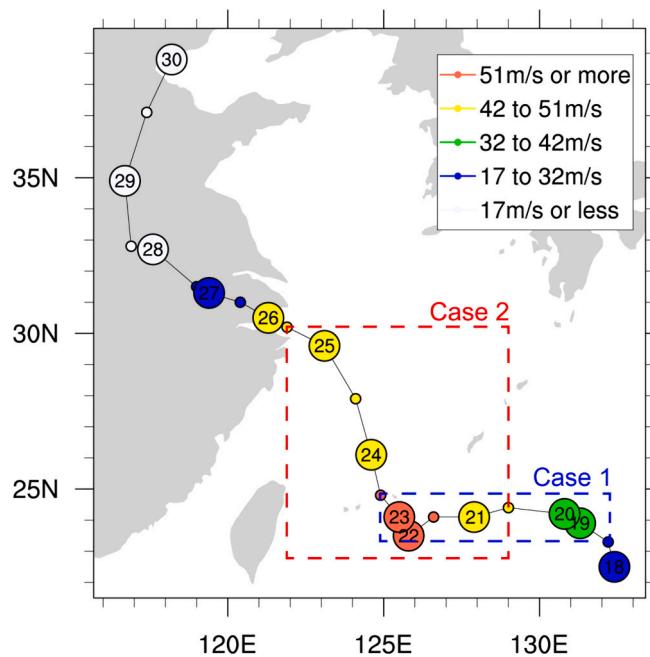


Fig. 1. Observed track and intensity of Typhoon In-fa from the CMA best-track datasets, with dots and numbers indicating the 12-h interval and date, respectively. The dotted boxes indicate the periods of the two cases in this study.

Table 1
Basic information on the ECMWF, NCEP, and CMA ensemble prediction systems.

Ensemble prediction system	ECMWF	NCEP	CMA
Initial perturbation scheme	Singular vector	Ensemble Kalman filter	Singular vector
Model perturbation	Yes	Yes	Yes
Maximum lead time (d)	15	16	15
Daily initialization time (UTC)	00, 12	00, 06, 12, 18	00, 12
Ensemble size (members)	51	30	30

The remainder of this paper is organized as follows. Section 2 describes the dataset and the calculation of the ESA used in this study. Section 3 provides an overview of Typhoon In-fa (2021). In Section 4, the primary results including the ensemble track forecasts and the standard and modified ESA for the three operational EPSs are presented. Finally, the conclusions and a discussion are provided in Section 5.

2. Data and methodology

2.1. Data

The ensemble and control forecast data in this study were from the operational global forecast products of the ECMWF, NCEP, and CMA, which are accessible through the THORPEX Interactive Grand Global Ensemble archive available at <https://apps.ecmwf.int/datasets/data/tigge>. The basic information of each of the three operational EPSs considered in this study is summarized in Table 1. In terms of initial perturbation generation, the ECMWF and CMA employ a dynamical perturbation scheme, i.e., singular vectors (Palmer, 2019; Peng et al., 2022a), whereas NCEP adopts posterior members that are updated by the ensemble Kalman filter recentered around the control analysis (Zhou et al., 2022). Operationally, the EPSs of the ECMWF, NCEP, and CMA consist of 51, 30, and 30 members, respectively. For a fair comparison, this study considered the same number of ensemble members ($N = 30$)

for each EPS and interpolated these to the same horizontal resolution of 0.5° . The 30 ensemble forecast members for the ECMWF were randomly selected from the total 51 members. The forecast fields of the prognostic variables from each operational system were all verified against a reference which was derived by taking the mean state of the control analyses from the three operational centers. This is to mitigate potential biases that might arise from utilizing different verification references. The best track data for verification of the TC track forecasts were downloaded from the CMA Tropical Cyclone Data Center (https://tcdata.typhoon.org.cn/zjljsjj_zlhq.html).

In this study, two phases in the lifecycle of Typhoon In-fa were analyzed (see Fig. 1 and Table 2) to elucidate the case dependence of ESA. The first phase, commencing at 1200 UTC on July 18, 2021, is characterized by In-fa's predominantly westward trajectory (refer to Fig. 1). The second phase began with a notable northward turn of In-fa at 1200 UTC on July 20, 2021 and culminated in its progression toward the East China Sea (see Fig. 1). Each phase was analyzed with the 120-h control and ensemble forecasts at 12-h intervals from each operational center.

2.2. Algorithms of ensemble-based sensitivity analysis

The standard ESA has been well defined in several prior studies (e.g., Ancell and Hakim, 2007; Torn and Hakim, 2008; Torn et al., 2018), which is fundamentally related to the time-lagged cross-variable ensemble covariance. Generally, the ESA for TC track predictability is to evaluate the spatiotemporally coherent relationship between the TC position at a specific forecast lead time and the prognostic variables at earlier times. This approach facilitates the efficient identification of both the strength and spatial pattern of physical factors that possibly contribute to the uncertainty in forecasting the position of a TC. Specifically, given an ensemble of N members, the sensitivity of the forecast metric J_t at time t to state variable $x_{i,t-\delta t}$ at location i and at an earlier lead time $t - \delta t$ can be expressed as follows:

$$\frac{\partial J_t}{\partial x_{i,t-\delta t}} = \frac{\text{cov}(J_t, x_{i,t-\delta t})}{\text{var}(x_{i,t-\delta t})} \quad (1)$$

where $\text{cov}(\bullet)$ denotes the covariance between two variables, and $\text{var}(\bullet)$ is the variance. Torn and Hakim (2008) stated that the above equation represents a linear regression where the independent variable is an analysis grid point and the dependent variable is the forecast metric. They also pointed out that J_t and $x_{i,t-\delta t}$ represent perturbations with the ensemble mean removed. Without incorporating a true reference field, the impact of model errors on forecast uncertainties cannot be explicitly addressed. This indicates that the standard ESA primarily accounts for the forecast error sensitivity associated with initial condition uncertainties and does not account for the sensitivity rooted in model errors.

The standard ESA used in our study closely aligns with Eq. (1), with a notable difference: the variables J_t and $x_{i,t-\delta t}$ are normalized in the computation of the standard ESA, following the definition in Zheng et al. (2013), Chang et al. (2013) and Ancell and Coleman (2022). The algorithm defining the standard ESA in this study is written as follows:

Table 2
Initial time, forecast lead time, and forecast interval for each case.

	Stage	Initial time	Forecast lead time (h)	Forecast interval (h)
Case 1	Westward movement	1200 UTC July 18, 2021	120	12
Case 2	Sharp turn and northward movement	1200 UTC July 20, 2021	120	12

$$\begin{aligned} r(J_t, x_{i,t-\delta t}) &= \frac{\text{cov}(J_t, x_{i,t-\delta t})}{\sqrt{\text{var}(x_{i,t-\delta t})} \cdot \sqrt{\text{var}(J_t)}} \\ &= \frac{\sum_{k=1}^N (J_{t,k} - \bar{J}_t)(x_{i,t-\delta t,k} - \bar{x}_{i,t-\delta t})}{\sqrt{\sum_{k=1}^N (J_{t,k} - \bar{J}_t)^2} \sqrt{\sum_{k=1}^N (x_{i,t-\delta t,k} - \bar{x}_{i,t-\delta t})^2}} \end{aligned} \quad (2)$$

where \bar{J}_t and $\bar{x}_{i,t-\delta t}$ are the ensemble mean states of J_t and $x_{i,t-\delta t}$, respectively. The primary objective of normalizing variables in ESA is to mitigate the impact of varying ensemble variance among the three ensemble systems. The normalization process is essential for enabling a direct comparison of the standard ESA across the three EPSs.

As shown in Eqs. 1 and 2, the standard ESA is computed with the ensemble anomalies (or perturbations) from the ensemble mean. It means the standard ESA is unrelated to the true reference and does not account for ensemble forecast uncertainties rooted from model-related errors. In this study, a modified ESA that considers the discrepancy of ensemble forecasts from the true reference is introduced as follows:

$$r'(J_t, x_{i,t-\delta t}) = \frac{\sum_{k=1}^N (J_{t,k} - J_t^a)(x_{i,t-\delta t,k} - x_{i,t-\delta t}^a)}{\sqrt{\sum_{k=1}^N (J_{t,k} - J_t^a)^2} \sqrt{\sum_{k=1}^N (x_{i,t-\delta t,k} - x_{i,t-\delta t}^a)^2}} \quad (3)$$

where J_t^a and $x_{i,t-\delta t}^a$ denote the true reference (generally the analysis state for a proxy) of variables J_t and $x_{i,t-\delta t}$ instead of the ensemble mean. Specifically, J_t^a represents the best track position at time t , and $x_{i,t-\delta t}^a$ represents the multi-model analysis fields at time $t - \delta t$. To elucidate the modified ESA, Eq. 3 can be expanded as the following:

$$\begin{aligned} r'(J_t, x_{i,t-\delta t}) &= \\ &\frac{\sum_{k=1}^N [(J_{t,k} - \bar{J}_t) + (\bar{J}_t - J_t^a)] [(x_{i,t-\delta t,k} - \bar{x}_{i,t-\delta t}) + (\bar{x}_{i,t-\delta t} - x_{i,t-\delta t}^a)]}{\sqrt{\sum_{k=1}^N [(J_{t,k} - \bar{J}_t) + (\bar{J}_t - J_t^a)]^2} \sqrt{\sum_{k=1}^N [(x_{i,t-\delta t,k} - \bar{x}_{i,t-\delta t}) + (\bar{x}_{i,t-\delta t} - x_{i,t-\delta t}^a)]^2}} \end{aligned} \quad (4)$$

In Eq. 4, two terms $(\bar{J}_t - J_t^a)$ and $(\bar{x}_{i,t-\delta t} - x_{i,t-\delta t}^a)$ represent the ensemble mean error relative to the true reference field. When the ensemble mean error is relatively minor compared to the ensemble perturbations, the modified ESA would yield similar results to the standard ESA. However, if the ensemble mean error is significant, for example, the ensemble forecasts present remarkable systematic deviations from the true reference, the modified ESA would predominantly show the impact of forecast uncertainties from model-related errors.

It should be noted that the modified ESA is intended as a complement to the standard ESA. Employing both the standard and modified ESA can provide a more comprehensive and nuanced analysis of the forecast error sensitivity associated with initial condition and model related uncertainties. In addition, because the modified ESA takes the true reference field into account, it cannot be conducted in real time but is intended for retrospective analysis. Thus, the modified ESA serves as a diagnostic tool rather than a predictive tool. In contrast, the standard ESA can be computed in real-time, making it a potential tool for immediate evaluations, such as during targeted observation experiments and field campaigns (Ancell and Hakim, 2007; Zack et al., 2010; Limpert and Houston, 2018; Ancell and Coleman, 2022).

In our application of ESA to TC track predictability below, we follow the definitions of J_t and $x_{i,t-\delta t}$ as in various previous studies (Torn et al., 2018; Ashcroft et al., 2021; Hazelton et al., 2023). Specifically, J_t is defined as the position of the TC center projected onto the specified major axis at lead time t . The major axis represents the direction in which the ensemble TC positions have the largest variability (i.e., standard deviation), e.g., as defined by Hamill et al. (2011). This approach allows the position variability to extend beyond the limitations of the Cartesian coordinate framework, making it particularly

advantageous for cases where the greatest variability spans both zonal and meridional directions. The orientation of the major axis is independently determined at each lead time. The variable $x_{i,t-\delta t}$ represents a scalar state variable at location i and lead time $t - \delta t$. For a vector variable such as wind, $x_{i,t-\delta t}$ is its projection onto the major axis.

3. Case overview

Typhoon In-fa, which struck in 2021, was notably destructive and exhibited a complex trajectory. It maintained a robust intensity over its unusually long lifespan of nearly ten days in the western North Pacific before its eventual landfall on China's eastern coast (Fig. 1). In-fa developed from a weak tropical depression over the Pacific Ocean, east of the Philippines, in mid-July. By 1200 UTC on July 17, 2021, the tropical depression had strengthened to a tropical storm, and the intensification continued as it navigated westward through the path between the Western Pacific Subtropical High (WPSH) and the monsoon trough (MT). The WPSH is depicted by a pink contour in Fig. 2e-h, while the MT, a critical near-equatorial zone where the southwest and southeast monsoons converge, is represented by a pink line in Fig. 2i-l. This convergence zone, as delineated by Ritchie and Holland (1999), Briegel (2002), and Wu et al. (2012), is a significant factor in the typhoon's development and trajectory.

On July 22, as Typhoon In-fa neared longitude 125°E, it exhibited an abrupt deceleration, followed by a directional shift to the north and subsequently a predominant northwest trajectory for the remainder of its lifespan. In-fa made landfall in Zhejiang, an eastern province of China, at approximately 04:30 UTC on July 25, 2021. The path of Typhoon In-fa was influenced by a confluence of multiple meteorological systems, including the WPSH, the MT, and the concurrent Typhoon Cempaka, which collectively rendered its trajectory challenging to predict with precision (Huang et al., 2022; Rao et al., 2022; Xu et al., 2022). The ensemble forecasts displayed significant variances among the ECMWF, NCEP, and CMA EPSs (see Figs. 3 and 4). To examine the case dependence of ESA concerning the uncertainty in the typhoon's track, this study focused on two distinct forecast scenarios: the period of westward movement and the period encompassing the sharp directional shift and subsequent northwestward progression.

4. Results

4.1. Ensemble track forecasts

In Fig. 3, we present the five-day ensemble (gray lines) and control (pink lines) forecasts for Typhoon In-fa's track from the ECMWF, NCEP, and CMA initialized at 1200 UTC on July 18 (i.e., Case 1), overlaid by the observed typhoon track (red lines). The blue ellipses in Fig. 3 were derived from a bivariate normal distribution fit, as elaborated by Hamill et al. (2011), applied to the 96-h ensemble TC positions (blue hollow circles). The longest and shortest diameters of the ellipse, aligned along the major (the blue line) and minor axes respectively, represent the distances of two standard deviations, capturing the maximum and minimum variability in the ensemble TC positions.

The ensemble forecasts of the three operational EPSs all capture the overall westward movement of In-fa, yet they exhibit distinct characteristics. Both the ECMWF and NCEP demonstrate comparable performance, with their ensemble forecast tracks encompassing the observed path of In-fa, as indicated by the gray and red lines. Nevertheless, the 96-h ensemble members of the ECMWF and NCEP present a northwestern deviation of approximately 200 km from the observed track (cf. the blue ellipse and the red dot). This deviation likely stems from the members' accelerated easterly steering wind from 60 to 96 h (see Fig. 6a). Furthermore, the ECMWF ensemble's TC position spread at 96 h is notably greater in the east-west direction (as indicated by the blue line in Fig. 3a) compared to the north-south direction. In contrast, the spread of the NCEP ensemble TC positions maintains a relatively uniform

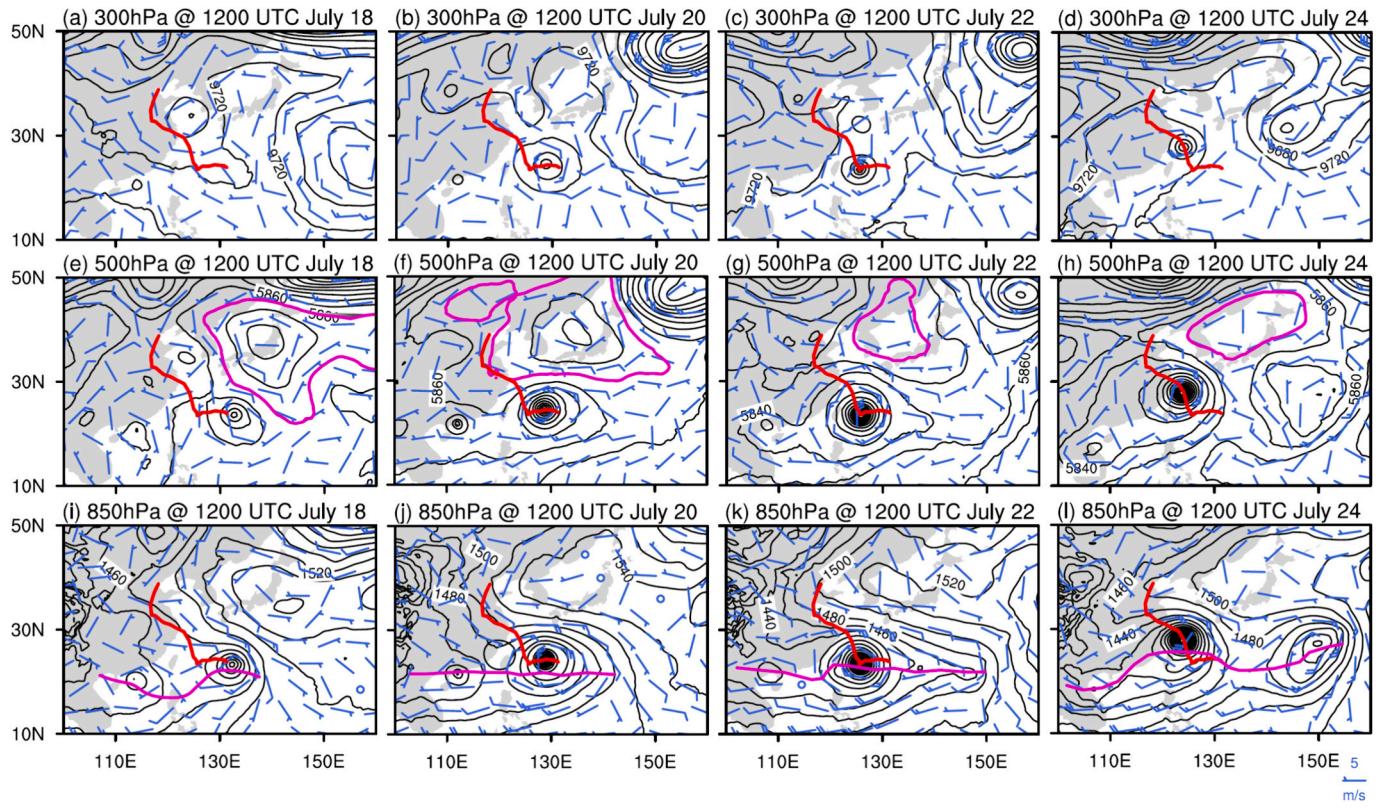


Fig. 2. Observed synoptic circulations associated with Typhoon In-fa at (a)–(d) 300 hPa, (e)–(h) 500 hPa, and (i)–(l) 850 hPa at 48-h intervals during July 18–24, 2021. The observation uses the mean state of the control analyses from the three operational systems. Red lines represent Typhoon In-fa's track in best track data. Black contours and blue wind bars represent the geopotential height (unit: gpm) and environmental wind (unit: m s⁻¹), respectively. Pink contours in (e)–(h) delineate the western Pacific subtropical high with geopotential height of 5860 gpm; pink lines in (i)–(l) denote the trough line of the monsoon trough. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

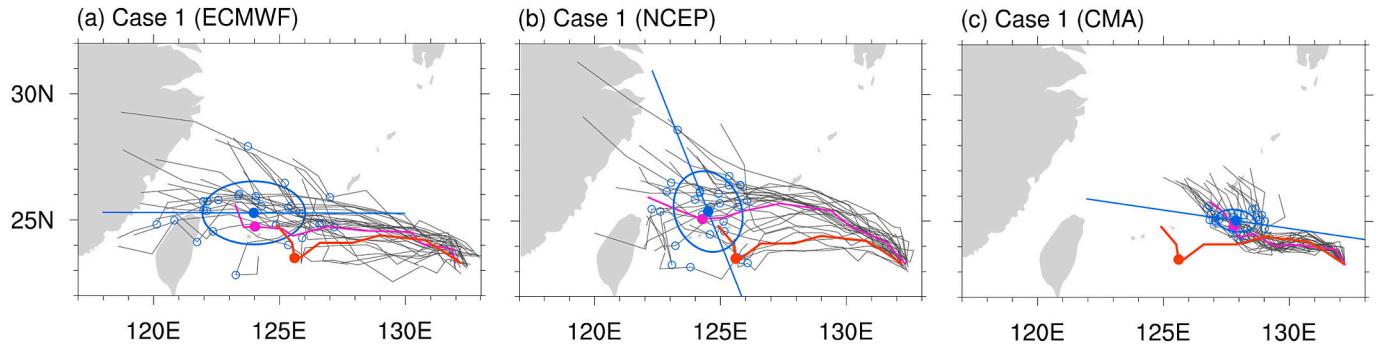


Fig. 3. Ensemble (gray lines) and control (pink lines) forecasts of the track of Typhoon In-fa by the (a) ECMWF, (b) NCEP, and (c) CMA EPSs initialized at 1200 UTC on July 18, 2021. Blue hollow circles represent the 96-h TC positions of each ensemble member and blue ellipses indicate bivariate normal fits to these positions. Red lines represent the best track positions. Colored filled circles denote the corresponding 96-h TC positions of the best track (red), control forecast (pink), and ensemble mean (blue). Blue lines show the major axis of the 96-h ensemble TC positions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

distribution in both the zonal (east-west) and meridional (north-south) directions.

In contrast to the ECMWF and NCEP ensembles, the CMA ensemble's TC tracks exhibit a notably smaller spread, as well as a significant eastward deviation of approximately 300 km relative to the observations. The reduced variability is likely due to an excessively small initial spread in the TC's position (Fig. 5a) and the steering flow at the Typhoon center (Fig. 6a). Moreover, the eastward deviation can be attributed to the relatively slower easterly steering flow within the CMA ensemble (Fig. 6a). These findings indicate a substantial systematic deviation in the CMA EPS when simulating the track of Typhoon In-fa.

Similar to Fig. 3, Fig. 4 shows the ensemble (gray lines) and control (pink lines) track forecasts for Case 2, initialized at 1200 UTC on July 20, 2021, along with the observed track (red lines). In this particular instance, Typhoon In-fa exhibited a deceleration in its westward trajectory near 125°E longitude, subsequently altering its course northwards before progressing toward the East China Sea. The northward shift was broadly replicated in all three EPSs, albeit with notable discrepancies in the timing and the exact location of this directional change. These discrepancies resulted in significant divergence within the ensembles for In-fa's northward path. This is exemplified by the 96-h ensemble TC positions (blue hollow circles) and the near east–west

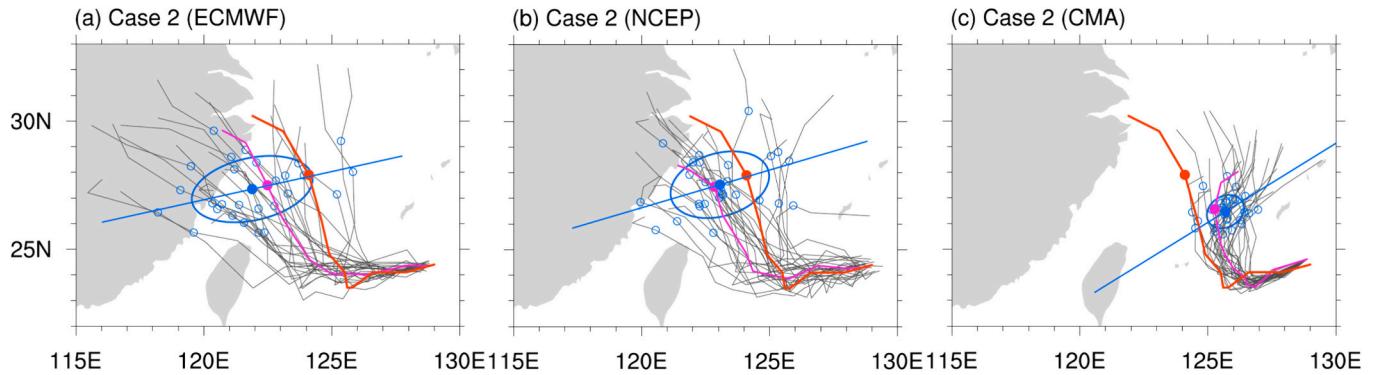


Fig. 4. Same as in Fig. 3 but for Typhoon In-fa in Case 2 initialized at 1200 UTC on July 20, 2021.

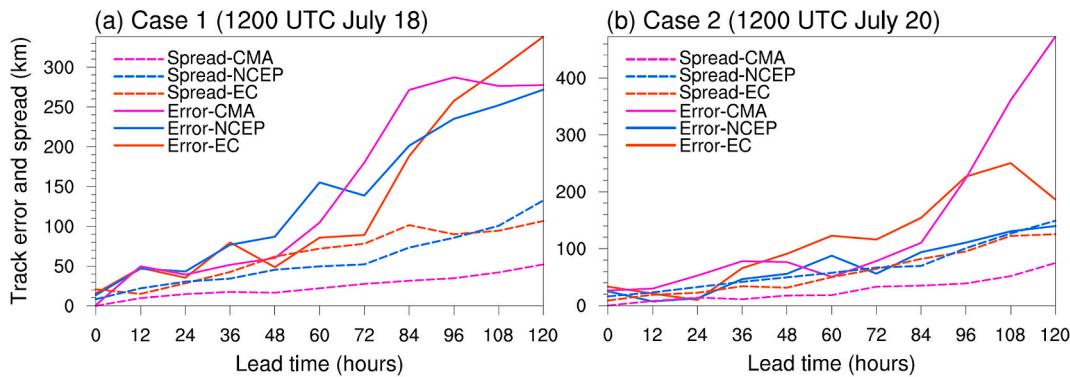


Fig. 5. Variation of the ensemble mean track error (colored solid lines) and spread (colored dashed lines) as a function of lead time in (a) Case 1 and (b) Case 2.

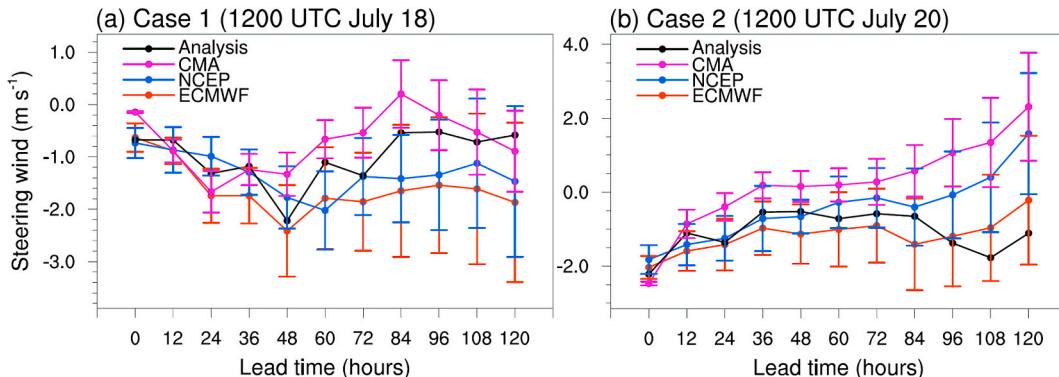


Fig. 6. Variation in the zonal component of the steering wind (negative and positive values for easterly and westerly winds, respectively) in the ensemble forecasts as a function of lead time: (a) Case 1 and (b) Case 2. Colored lines and black lines denote the ensemble mean forecasts and the reference observations, respectively; the colored vertical bars denote the ensemble spread.

orientation of the major axis (blue lines) of the ellipses. The ensemble forecasts from the ECMWF and NCEP (gray lines) demonstrated a considerable spread. This spread not only encompassed the observed track (red lines) but also highlighted the variability in terms of the possibility of landfall.

Similar to Fig. 3, the ensemble members of the CMA EPS exhibit a predominantly anomalous southeastward track relative to the observed track after their northward turn. This deviation is characterized by a notably smaller spread (~ 50 km) compared to the wider spreads observed in both the ECMWF and NCEP EPSs (~ 100 km), as illustrated in Fig. 4c. Therefore, none of the CMA ensemble members suggest a probable landfall by Typhoon In-fa. The performance of the CMA ensemble can be attributed to two primary factors: a significant underestimation in the initial spread of the TC position, as evidenced in

Fig. 5b, and a positive deviation in the weak easterly steering flow (cf. pink and black lines in Fig. 6b). This deviation may be linked to model deficiencies in simulating the environmental fields of In-fa. Our analysis predominantly focuses on evaluating the impact of these ensemble forecast performance by various EPSs on the ESA of TC track uncertainty.

The ensemble mean error and the ensemble spread of the TC track forecasts from the three operational EPSs were analyzed for the two specified cases in Fig. 5. In Case 1, the three operational EPSs have similar ensemble mean track errors in the first 36 h. Beyond this period, from 48 to 84 h, the ECMWF EPS demonstrated the smallest ensemble mean track error, averaging approximately 60 km less than those of the NCEP and CMA EPSs. Additionally, the ECMWF EPS exhibited an ensemble spread closely matching its ensemble mean error during the

first 72 h, suggesting superior accuracy in quantifying TC track forecast errors in this case. In contrast, the NCEP EPS followed with an ensemble spread about 30% lower than that of the ECMWF. The CMA EPS showed the lowest ensemble spread of TC track forecasts throughout all lead times, nearly one-third of the ECMWF's spread. The significant underestimation of ensemble spread in the CMA EPS's TC track forecasts may be attributed to the smallest initial spread in both the TC position (see Fig. 5a) and the environmental flow (see Fig. 6a).

In Case 2, the NCEP EPS overall performs the best in terms of the ensemble mean track. The ensemble mean track error of the CMA EPS is comparable to that of the NCEP EPS (cf. pink and blue solid lines in Fig. 5b) and lower than that of the ECMWF EPS (cf. pink and red solid lines in Fig. 5b) from 36 to 84 h. However, the ensemble mean track errors of the CMA EPS significantly exceed those of the NCEP and ECMWF EPSs beyond 84 h. This can be attributed to the pronounced eastward and southeastward deviations in the ensemble TC tracks of the CMA EPS (see Fig. 4c) with the lowest ensemble spread (see dashed lines in Fig. 5b), especially at later lead times. The NCEP EPS presents the closest ensemble mean error and the spread of TC tracks in Case 2.

To further address the differences in the forecasting performance of the three EPSs for TC tracks, we conduct a comparative analysis of the steering flow, a dynamical indicator for TC movement. In this study, the steering flow is calculated as the average wind in a $600 \text{ km} \times 600 \text{ km}$ region surrounding the TC center, spanning pressure layers from 850 to 300 hPa as in Chan and Gray (1982) and Akter and Tsuboki (2021). This analysis focuses on the zonal components of the TC steering flow (negative and positive values representing easterly and westerly winds, respectively) in the ensemble members (vertical bars) and their mean values (lines) for the ECMWF, NCEP, and CMA EPSs as a function of lead time for Case 1 and Case 2, as illustrated in Fig. 6a and b.

The comparative analysis reveals a significantly smaller initial spread of the steering wind in the CMA ensemble across the three EPSs for both cases. This discrepancy likely contributes to the notably smaller forecast spread of the steering flow in the CMA EPS, which is approximately half that of the ECMWF and NCEP EPSs throughout all lead

times. This result is consistent with the performance differences in the TC track spread among the three systems (cf. Figs. 5 and 6). Furthermore, the NCEP and the ECMWF EPSs demonstrate superior accuracy in capturing the easterly component of the steering wind in Case 1 and Case 2, respectively, compared to the reference analysis. The CMA EPS exhibits a significant positive deviation in the overall easterly steering flow relative to the observations in both cases. This deviation explains the overall eastward deviation of the TC tracks in the CMA EPS as illustrated in Figs. 3c and 4c.

4.2. Ensemble sensitivity analysis without model errors

Fig. 7 shows the standard ensemble sensitivity [i.e., r in Eq. (2)] of the TC center position at the 96-h forecast lead time to the 500-hPa environmental wind in the forecasts at 24-, 48-, and 72-h lead times for the three operational EPSs in Case 1. The TC center position and the environmental wind in the ensembles are all projected onto the major axis at 96 h (blue straight lines in Fig. 3), along which the ensemble TC positions have the largest spread. It should be noted that the major axes of the ECMWF and CMA ensembles in Case 1 are predominantly aligned in the west–east direction (see Fig. 3a and c). Therefore, the near north–south major axis of the NCEP ensemble in Case 1 (i.e., Fig. 3b) is substituted with the mean direction (see the blue straight line in Fig. 7d–f) of those of the ECMWF and CMA ensembles for projection purpose. This facilitates a relatively uniform and meaningful comparison of the ESA results along a similar axis direction among the three EPSs.

Fig. 7 shows that the ensemble sensitivity for the individual EPSs overall exhibits a similar pattern, which notably intensifies with increasing lead time from 24 to 72 h. This indicates an increased linear impact of uncertainties in the environmental flow on the uncertainties in the TC position prediction over shorter lead times. Notably, the high sensitivities within the three EPSs are mainly concentrated around the environment ($\sim 1000 \text{ km}$) of Typhoon In-fa. These includes areas south of the WPSH and north of the MT (marked by A and B, respectively), as

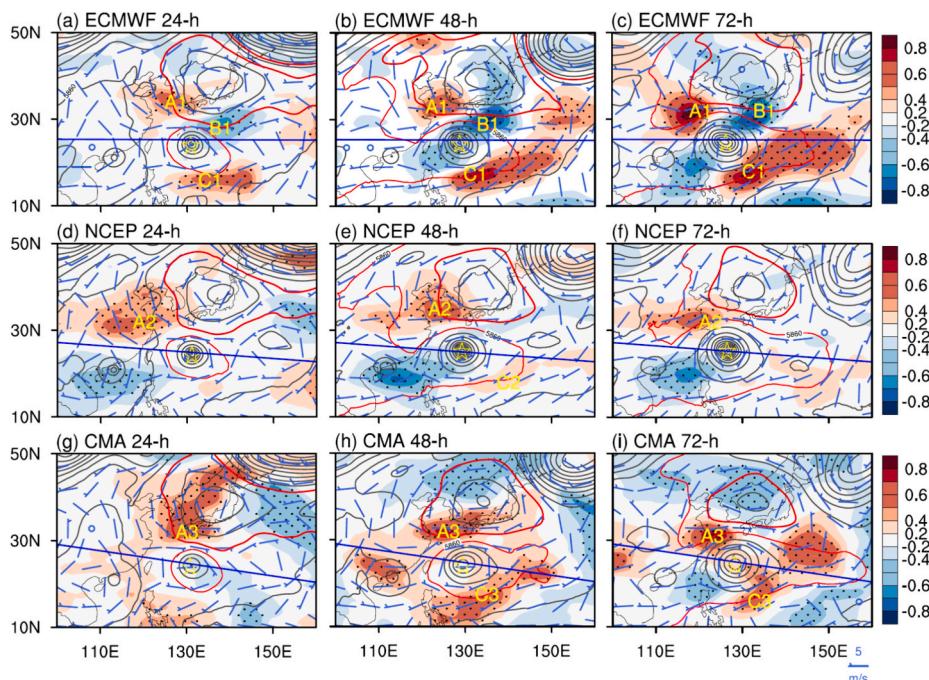


Fig. 7. Ensemble sensitivity (shading) of the 96-h TC position to the 500-hPa environmental wind at (a) 24 h, (b) 48 h, and (c) 72 h, projected on the 96-h major axis (blue lines), for the ECMWF EPS in Case 1. The panels of the middle and lower rows are similar to those of the upper row but for the NCEP and CMA EPSs, respectively. Stippled regions indicate where the sensitivity is statistically significant at the 95% confidence level. Blue bars and black contours denote the ensemble mean wind and the geopotential height at 500 hPa, respectively. Yellow stars represent the ensemble mean position of In-fa. Red contours highlight the geopotential height of 5880 and 5860 gpm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

well as the area south of the MT (marked by C). This implies that the WPSH and MT might be the primary weather systems influencing the forecast uncertainty of the typhoon's position. Nevertheless, the ensemble sensitivity reveals distinct patterns for each EPS, highlighting their differences in how the precise processes of environmental flow modulate the uncertainty in TC position prediction.

Taking the ensemble sensitivity of the 96-h TC position to the 48-h environmental flow as an example (i.e., Fig. 7b, e, and h), distinct ESA patterns are observed for different EPSs. Specifically, the ensemble forecasts of the ECMWF and CMA EPSs display significantly positive sensitivity to the southeast of the MT (i.e., C1 in Fig. 7b and C3 in Fig. 7h), which is much stronger than that of the NCEP EPSs in similar regions (C2 in Fig. 7e). This positive sensitivity implies a dynamical process where the position of Typhoon In-fa shifts eastward (westward) in response to the strengthening (weakening) of the westerly flow south of the MT. In contrast, the dynamic processes influencing TC position north of the MT and south of the WPSH display variations across the three systems, despite an overall similarity in their sensitivity regions. The sensitive region on the southern side of the WPSH in the 48-h ensemble of the ECMWF EPS displays a west-positive (A1) and east-negative (B1) dipole mode (Fig. 7b). Differently, the NCEP ensemble shows a sole region of positive sensitivity near the southwest of the WPSH (A2), while the CMA ensemble exhibits a longer and narrower region of positive sensitivity stretching from the west to the east (A3). The mechanisms underlying these divergent sensitivities are elucidated in Fig. 8.

To elucidate the ensemble sensitivity in Fig. 7, Figs. 8(a)-(c) feature a 48-h ensemble spaghetti plot for 500-hPa geopotential height of 5880 and 5860 gpm (gray lines) for individual EPSs, overlaid by their respective 96-h ensemble TC positions (black dots). The contours of 5880 and 5860 gpm approximately outline the structures of the WPSH and the MT. Three representative members of each EPS are highlighted, with the colored lines indicating the 48-h geopotential heights and the corresponding colored hollow circles marking the 96-h TC positions. It is important to note that selecting an alternative set of three members would not fundamentally alter the derived conclusions. Considering the inherent limitations of spaghetti diagrams, especially in regions with weaker geopotential height gradients, the placement of contours may falsely suggest higher variability than actually present, Figs. 8(d)-(e) also include the spatial distribution of the 30-member standard deviation of the 500-hPa geopotential height field as a supplementary reference.

As shown in Fig. 7b and e, the TC position uncertainties within the ECMWF and NCEP ensembles exhibit significantly positive sensitivity to the environmental flow southwest of the WPSH (see A1 and A2 in Fig. 7). This sensitivity is closely related to the east-west positional displacement of the WPSH among the ensemble members (cf. pink, red, and blue solid contours at A1 and A2 in Fig. 8). Specifically, an eastward shift in the western boundary of the WPSH correlates with a weakening of the easterly environmental flow in the southwest of the WPSH near A1 and A2, while a westward shift correlates with a strengthening of this flow. These variations in flow intensity directly cause corresponding eastward or westward shifts in the TC position observed 48 h later (see the solid contours and hollow circles), demonstrating positive sensitivity.

Compared to the ECMWF and NCEP ensembles, the CMA ensemble (Fig. 8c) demonstrates a closer resemblance in depicting the features of the WPSH. This is evidenced by a notably smaller spread in the 500-hPa geopotential height near area A3, when contrasted with the spread near areas A1 and A2 (cf. Fig. 8f and Figs. 8d, 8e). As a result, unlike the ECMWF and NCEP ensembles, the CMA ensemble uniquely exhibits a pronounced and elongated east–west extension of high positive sensitivity along the southern boundary of the WPSH (see A3 in Fig. 7h). Furthermore, the center of maximum positive sensitivity within the CMA ensemble is located further eastward (cf. A3 and A1, A2 in Fig. 7). This suggests that, for the CMA ensemble, the TC positional uncertainty might partially arise from variations in the strength of the easterly environmental wind on the southern flank of the WPSH in addition to the effect of the positional uncertainty of the WPSH. Additionally, there is a region of negative sensitivity (identified as B1 in Fig. 7b) for the ECMWF ensemble to the northeast of the TC, situated between the WPSH and the MT. This feature is absent in both the NCEP and CMA ensembles. This negative sensitivity could be related to the positional deviations of the eastern boundary of the MT among the members of the ECMWF ensemble (Fig. 8a). In contrast to the ECMWF ensemble, the NCEP ensemble displays more significant variations in both the position and morphology of the MT, as exemplified by the split and integrated structure of the MT shown in blue and red, respectively (Fig. 8b). This distinction is further underscored by the larger ensemble spread of the NCEP EPS than that of the ECMWF, particularly in the vicinity of region B1 (cf. Fig. 8e and 8d). Such variability in the NCEP ensemble could possibly weaken the overall linear impact of the environmental flow associated with the MT on the TC position (see Fig. 7).

The analysis of ensemble predictions from three different EPSs reveals a notably positive sensitivity in the southeastern region of the MT

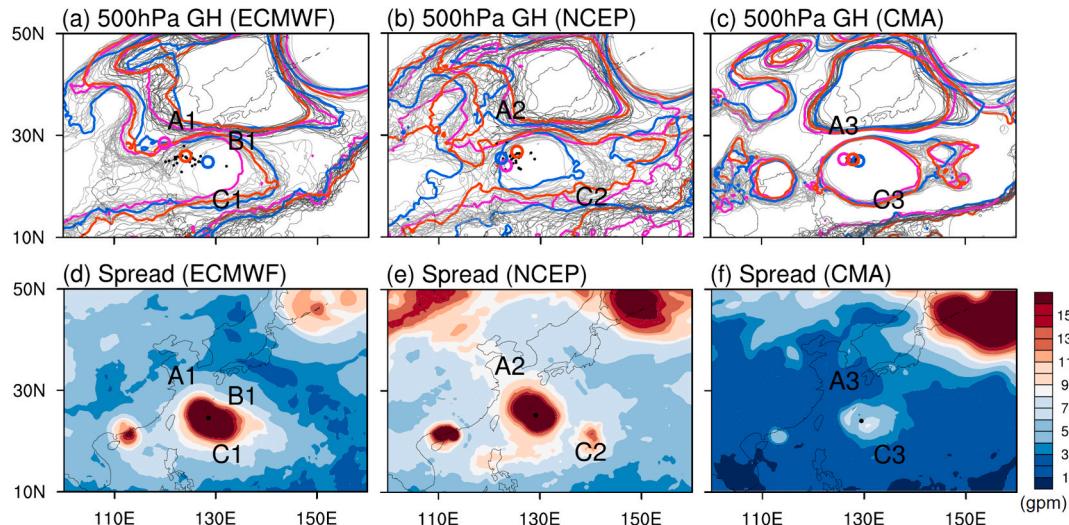


Fig. 8. (a)-(c): Ensemble spaghetti plot for 500-hPa geopotential height of 5880 and 5860 gpm (gray contours) at 48 h for the three EPSs. Black dots represent their respective 96-h ensemble TC positions. Colored contours and hollow circles highlight the geopotential height and TC positions of three members in each EPS. (d)-(f): The ensemble spread of the 500 hPa geopotential height field at 48 h for the three EPSs.

(C1 in Fig. 7b, C2 in Fig. 7e, and C3 in Fig. 7h). This sensitivity indicates that the prevalent westerly winds in this region (see Fig. 6) tend to slow the westward movement of the TC. Noticeably, the ensemble sensitivity in this area is much weaker for the NCEP EPS than for the ECMWF and CMA EPSs (cf. C1, C3, and C2 in Fig. 7). This discrepancy could be attributed to the more pronounced deviations in the position and the shape of the MT in the NCEP ensemble, relative to the other two ensembles (cf. C1, C2, and C3 in Fig. 8a-c, respectively), as evidenced by their respective ensemble spread (see Fig. 8d-f). Such deviations introduce a more complex and nonlinear interaction with the MT, affecting the TC's trajectory. The variance in the performance of these EPSs underscores the critical role of ensemble characteristics in determining how the environmental flow influences TC track uncertainty, as revealed by the ESA. It highlights how ensemble spread in the representation of large-scale synoptic systems can differentially influence the uncertainties in TC position.

Similar to Fig. 7, Fig. 9 shows the ensemble sensitivity of the 96-h TC position to the environmental flow from 24 to 72 h in Case 2. Consistent with Case 1, the ensemble correlation intensifies as the lagged time decreases. Notably, the ensemble sensitivity linked to the WPSH at earlier lead times (e.g., 24 h) is markedly less pronounced in Case 2 compared to Case 1. The ensemble correlation adjacent to the southern boundary of the WPSH in Case 2 ranges between -0.3 and 0.3 at 24 and 48 h across all three EPSs, which contrasts with the higher correlation in Case 1. This is possibly related to the significant evolution of the WPSH in terms of shape and size as exemplified by the 5880 gpm contour highlighted in red. Such changes may exert a stronger nonlinear impact of the WPSH on the TC's trajectory. The large uncertainties in the TC's northward turn and subsequent path (refer to Fig. 3) could be associated with the rapid diminution and eastward movement of the WPSH from 24 to 72 h. In Case 2, the areas of high sensitivities predominantly located near the MT to the east of typhoon In-fa (marked by B), similar to Case 1 (see Fig. 7). This is probably attributable to the relatively stable structure and the maintenance of the MT during the evolution of In-fa.

Fig. 10 is the same as Fig. 8 but for Case 2. Overall, the ensemble

diversity of the ECMWF and NCEP EPSs is larger than that of the CMA EPS, as evidenced by the detailed analysis of the spaghetti diagrams (Fig. 10a-c) and the ensemble spread distribution (Fig. 10d-f). These findings align with the observations in Fig. 8. The region of negative sensitivity at 48 h, identified in the south of the WPSH for the ECMWF ensemble (see A1 in Fig. 9b), could be attributed to the variances in the longitudinal positioning of the WPSH across the ensembles (see A1 in Fig. 10a). The structure of the 500-hPa geopotential height near the east of the MT shows greater diversity in the NCEP ensemble compared to the ECMWF and CMA ensembles (cf. B2 and B1, B3 in Fig. 10). This may account for the relatively lower positive ensemble correlation in the NCEP ensemble (cf. B2 in Fig. 9e, B1 in Fig. 9b, B3 in Fig. 9h).

4.3. Ensemble sensitivity analysis with model errors

Errors in ensemble forecasts, verified against the truth (generally using the analysis state as a proxy), are jointly contributed by initial condition errors and model-related errors. However, the standard ESA does not account for the effect of model errors because the perturbations of the physical variables are computed relative to the ensemble mean state (Ancell and Hakim, 2007; Torn and Hakim, 2008). In this study, we propose a modified ESA that considers the combined effects of initial and model-related errors on TC track forecast errors. This is achieved by calculating the deviations of ensemble forecasts from observations (or analysis), rather than from the ensemble mean (see Eq. 3). According to Eq. 3, if there are no errors in the ensemble mean forecasts of synoptic variables (i.e., $\bar{J}_t - J_t^a = 0$) and TC positions (i.e., $\bar{x}_{i,t-\delta t} - x_{i,t-\delta t}^a = 0$), the modified ESA would be equivalent to the standard ESA. It should be emphasized that the modified ESA is not designed to replace the standard ESA, but rather to complement it. Employing both the standard and modified ESA can provide a more comprehensive and nuanced analysis of forecast error sensitivity associated with initial condition and model-related uncertainties.

Fig. 11 shows the modified ESA of the TC position forecasts at a 96-h lead time in relation to forecast errors in the 72-h environmental flow

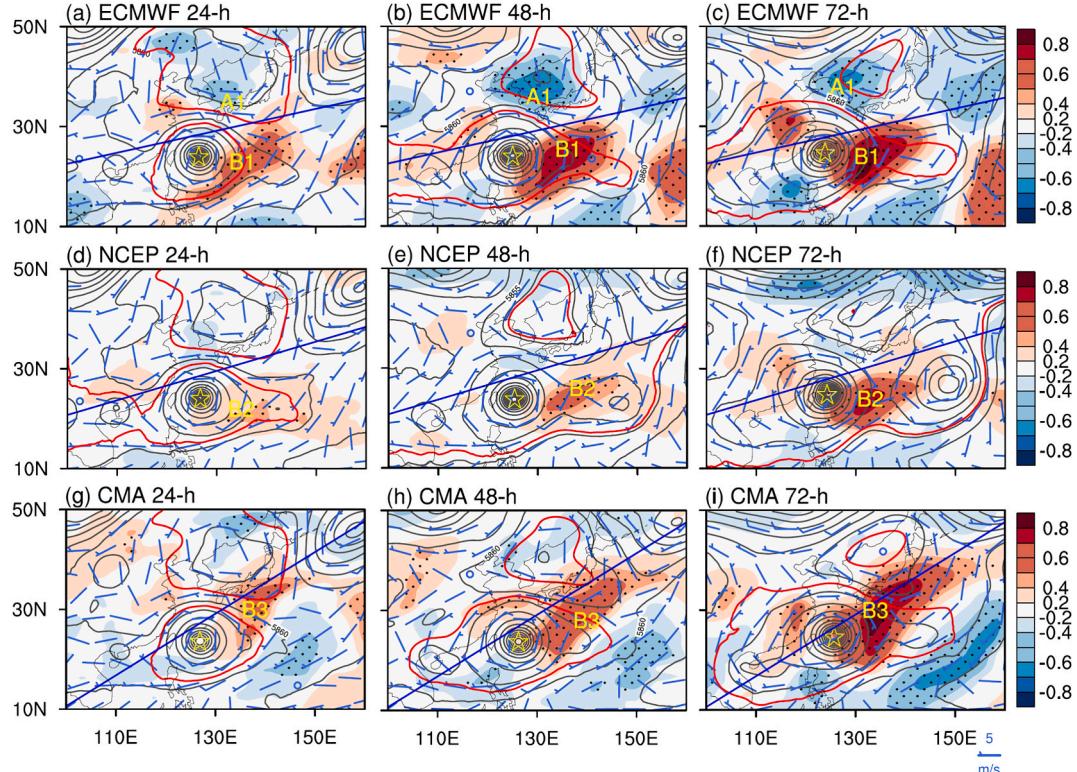


Fig. 9. Same as in Fig. 7 but for Case 2.

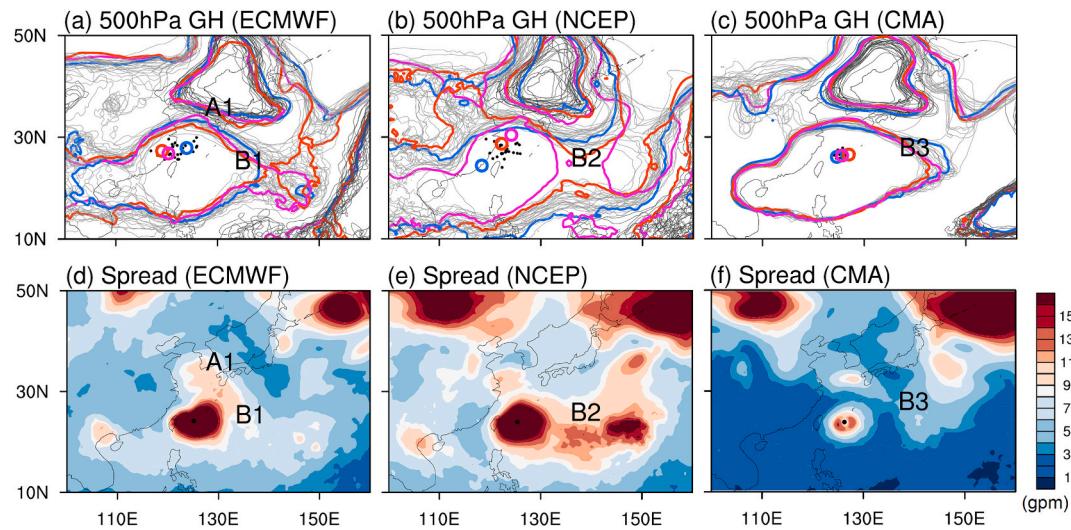


Fig. 10. Same as in Fig. 8 but for Case 2.

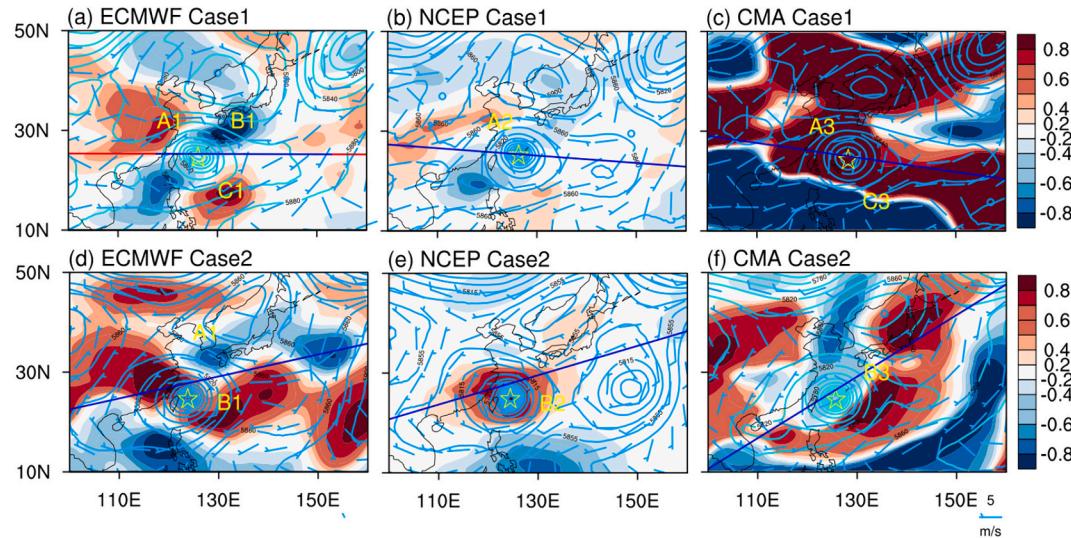


Fig. 11. Modified ESA (shading) of the TC position at 96 h to the 72-h environmental flow projected on the 96-h major axis (blue lines) for the three EPSs in both cases. Blue barbs and contours denote the ensemble mean wind and 500-hPa geopotential height, respectively. Yellow stars represent the ensemble-mean positions of In-fa. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

averaged between 850 and 300 hPa for the three EPSs in both cases. In the modified ESA, the environmental flow and the TC's center positions are projected onto the same major axis, consistent with the approach in Figs. 7 and 9. The reference for the environmental flow uses the mean state of the control analyses from the three operational centers, while the TC center positions are compared against the best track data. The modified ESA for other lead times produced qualitatively similar results and thus are not shown. It is noticeable that the modified ESA presents different results from the standard ESA for each EPS, primarily due to the use of different reference states in the ESA calculation. In both cases, the modified and standard ESA fields of the ECMWF ensemble exhibit a high degree of similarity, with the highest spatial correlation of nearly 0.7 (cf. Figs. 11a and 7c, Figs. 11d and 9c). This is mainly explained by the fact that the ensemble spread of the environmental wind for the ECMWF EPS is considerably larger than the absolute error of the ensemble mean in both cases, as indicated by Fig. 13a and d. In such a situation, according to Eq. 4, the terms associated with the ensemble covariance of variable perturbations [i.e., $(J_{t,k} - \bar{J}_t)$ and $(x_{i,t-\delta t,k} - \bar{x}_{i,t-\delta t})$] play a more important role in determining the

modified ESA, resulting in a performance more similar to the standard ESA. These results also suggest a relatively smaller systematic deviation of the environmental flow and TC position in the ECMWF ensemble forecasts from the observation reference.

The similarity between the modified and standard ESA is followed by the NCEP ensemble with a spatial correlation of approximately 0.4 in Case 1 and 0.6 in Case 2. In contrast, the spatial correlation for the CMA ensemble is below 0.2 in Case 1 and approximately 0.4 in Case 2. Moreover, the modified ESA applied to the CMA ensemble demonstrates a notably higher sensitivity and much sharper gradients between positive and negative sensitivities compared to the results of the standard ESA (cf. Figs. 11c and 7i, Figs. 11f and 9i). This is attributed to the much larger model systematic deviation in the CMA ensemble than in other ensembles in these two cases, as detailed in Figs. 12 and 13.

To explain the differences in the results of the modified ESA across the three EPSs in Fig. 11, Fig. 12 shows the ensemble forecasts (gray contours) of 500-hPa geopotential height at the 72-h lead time, overlain by the ensemble mean (blue) and observation analysis (red) at the same lead time for the individual EPSs. Notably, the 72-h ensemble mean

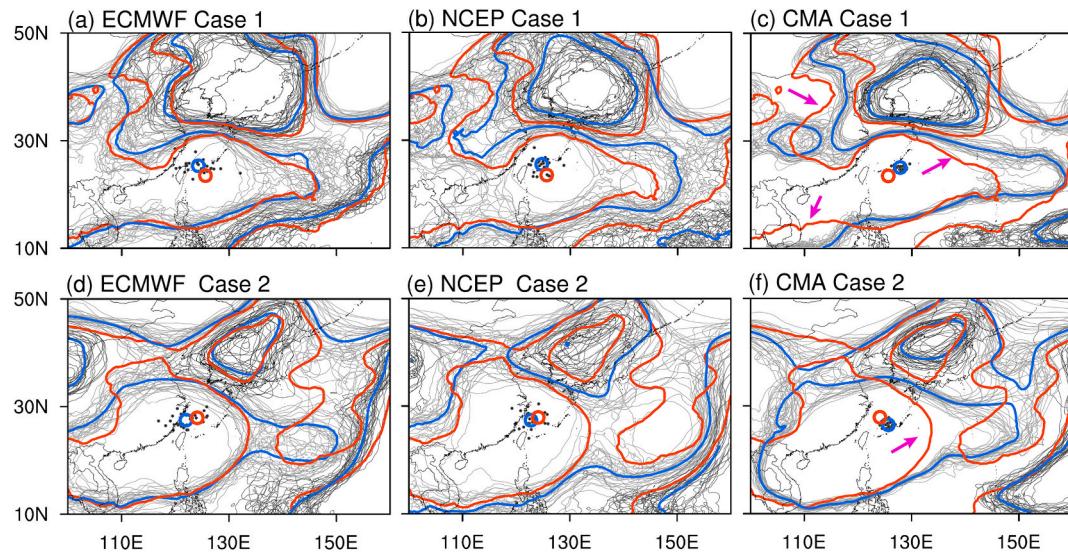


Fig. 12. Ensemble spaghetti plot for 500-hPa geopotential height of 5880 gpm and 5860 gpm (gray contours) at 72 h for the three EPSs overlain by the ensemble mean (blue contours) and observation analysis (red contours). Red and blue hollow circles represent the 96-h ensemble mean and best track TC positions, respectively; black dots are the 96-h ensemble TC positions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

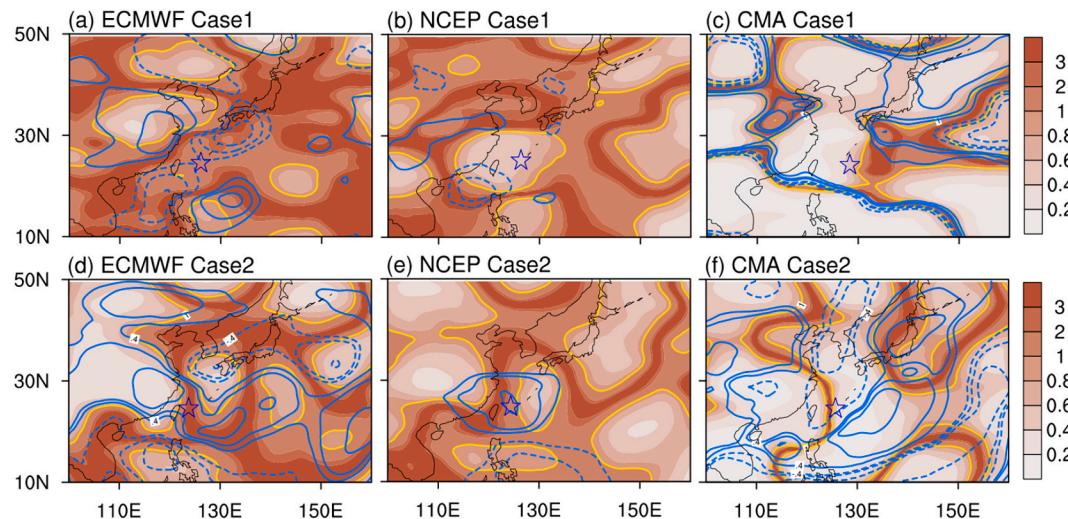


Fig. 13. Spread–error ratio of the environmental flow projected on the major axis for the 72-h ensemble forecasts of the three EPSs in both cases. The spread–error ratio of 1 is highlighted by yellow contour. Blue contours represent modified ESA of the TC position at 96 h to the 72-h environmental flow projected on the 96-h major axis. Black stars represent the 72-h ensemble mean TC positions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

forecasts of the 500-hPa geopotential height of the ECMWF EPS exhibit greater resemblance to the reference observations compared to the forecasts from the NCEP and CMA systems in both cases, with a lower root mean square error (9.73 vs. 14.73 and 12.56 gpm in case 1; 11.96 vs. 18.1 and 13.41 gpm in case 2). This is probably associated with the appropriate spread in the ECMWF ensembles, in terms of the shape and the position of synoptic systems such as the WPSH and MT. Such a spread facilitates more effective sampling or covering of the true state within the ensemble distribution. The ensemble members of the NCEP display a spread comparable in magnitude to that of the members of the ECMWF. However, the NCEP ensemble members display notably lower similarity to the observations in terms of the shape of the environmental flow compared to the ECMWF ensemble members. This discrepancy might explain the relatively reduced correlation between the modified and standard ESAs for the NCEP ensemble (see details in Fig. 11). In contrast, for the CMA ensemble, the observations fall outside of the

ensemble forecast distribution in both cases (see areas highlighted by pink arrows). This phenomenon is attributed to the underestimation of the ensemble spread (see Figs. 6 and 13). As a result, the TC position forecasts of the CMA ensemble also exhibit the smallest spread across the three EPSs (see Fig. 12c and f).

As indicated by Eq. 4, the performance of the modified ESA is closely related to the relative magnitudes of ensemble perturbations versus ensemble mean error. Therefore, Fig. 13 shows the spatial distribution of the ratio between the ensemble spread and the ensemble mean error magnitude for the 72-h environmental wind. To align with preceding results, the environmental winds are represented by the component of the 300–850 hPa averaged wind projected on the major axis of the ensemble TC positions for the individual ensembles in both cases. In both cases, most regions (over 70%) for the ECMWF and NCEP EPSs have a ratio exceeding 1.0, indicating relatively smaller systematic ensemble deviations in comparison to ensemble perturbations. In

contrast, the spread–error ratio of the CMA ensemble is overall considerably lower than that of the ECMWF and NCEP EPSs, with values below 0.4 in almost 80% of the regions. This suggests that the term associated with the ensemble mean errors would predominantly determine the performance of the modified ESA for the CMA ensemble. This was evidenced by the result that the spatial distribution of the spread–error ratio of the CMA ensemble shows a notable correlation with that of the modified ESA (cf. shadings and blue contours in Fig. 12c and f). In other words, the smaller spread–error ratios correspond to higher modified ensemble sensitivity, suggesting that the high sensitivity in the modified ESA for the CMA ensemble is primarily attributable to the large model systematic deviations in the In-fa case.

5. Conclusions and discussion

Understanding the origins of uncertainty in TC track forecasts, especially in cases with significant track forecast errors, is imperative. The ESA has been widely applied as a practical and effective tool for investigating the determinants that contribute to uncertainties in TC track forecasting. Nonetheless, the majority of existing research has primarily depended on a single EPS. Within the framework of the standard ESA, perturbations are calculated relative to the ensemble mean to estimate the ensemble sensitivity (or covariances). This methodology, however, fails to account for the impact of both the configurations and performance of different EPSs, as well as model errors on the ESA. This constrains a more comprehensive understanding of the factors influencing uncertainty in TC track forecasts.

This research employed the ESA to explore the dynamic factors and processes that contribute to uncertainty in TC track forecasts. To address the dependence of the ensemble sensitivity on the performance of ensemble forecasts, ensemble forecast data from three operational forecasting centers, i.e., the ECMWF, NCEP, and CMA, were utilized. In addition to the standard ESA approach commonly used in prior studies, this study introduced a novel modified ESA algorithm. This algorithm calculates the deviations of ensemble variables in relation to the true reference (typically the analysis state), rather than the ensemble mean as is customary in the standard ESA. This modification allows for a more comprehensive assessment of uncertainty sources in TC positions, taking into account not only initial errors but also model inaccuracies. The study involved a comparative analysis of the outcomes derived from both the modified and standard ESA methods. The case of Typhoon In-fa was selected due to its notably large forecast track errors, and two specific stages during its lifetime were investigated.

The main conclusions derived from the study are summarized in the following.

- (1) The accuracy of the ensemble mean forecast for TC tracks is overall comparable across the three operational systems. However, the CMA ensemble presents notably lower spread of the TC track and environmental flow than those of ECMWF and NCEP ensembles through all lead times in both cases.
- (2) The standard ESA based on the three operational ensemble systems consistently indicate that the WPSH to the north and the MT to the south of Typhoon In-fa are the primary weather systems that contribute to the uncertainty in the typhoon's position. However, due to varying degrees of ensemble spread (or diversity) among these systems, distinct differences in the results of the standard ESA are observed. In the early lead times of the ECMWF ensembles, the WPSH and MT exhibit similar patterns with notable positional deviations, which partially account the later uncertainties in TC track. In contrast, the CMA ensemble displays a significantly lower spread in both the position and structure of the WPSH and MT, leading to a scenario where the typhoon's positional uncertainty is primarily ascribed to variations in the strength of the environmental flow. The NCEP ensemble demonstrates more pronounced nonlinear differences

in the features of environmental fields compared to the ECMWF and CMA ensembles. This results in a comparatively weaker overall sensitivity of the NCEP ensemble to the variation of these features. Therefore, while all three operational ensemble systems identify the WPSH and MT as key factors influencing the uncertainty in the typhoon's trajectory, their varying ensemble spreads and sensitivities lead to different assessments of these influences.

- (3) The ensemble results from the ECMWF demonstrate the highest consistency between the modified and standard ESA, surpassing those of the NCEP and the CMA ensembles. The performance of the ECMWF EPS can be primarily attributed to its significantly lower ensemble mean error of the TC track and environmental flow compared to their ensemble spread. In contrast, the modified ESA of the CMA ensemble exhibits the most pronounced differences from the standard ESA. Detailed analysis of the spread–error ratio indicates that the modified ensemble sensitivity of the CMA mainly reflects the relationship between the model systematic deviations of the environmental flow and TC positions.

Our study systematically examined the influence of ensemble forecast performance on the ESA results by comparing outcomes derived from various EPSs. This study underscores the importance of EPS selection when using ESA to discern the sources of forecast uncertainty. It is imperative for researchers and forecasters to carefully assess the performance of different EPSs, considering their appropriateness for specific research objectives or operational applications. Additionally, our findings emphasize the persistent need for an enhanced representation of the environmental flow in TC models. Advancements in parameterization schemes and physical processes related to the environmental flow could reduce model biases and increase the representativeness of ensemble members, consequently refining the precision of TC track predictions.

In this study, we introduced a novel approach by modifying ESA to incorporate model errors alongside initial errors. This methodology assists in evaluating the impact of model-related errors on forecast error sensitivities and aids in the practical applications of ESA-derived findings. However, it is crucial to note that the modified ESA cannot be implemented in real-time, as it relies on the availability of verifying analysis for the computation of forecast errors. Furthermore, to extract meaningful insights, the results derived from the modified ESA must be compared and analyzed against those obtained from the standard ESA. The modified ESA serves as a tool for in-depth post-analysis and diagnostic studies. Further research could explore quantitative expressions and assessments of the impact of model errors, as well as the associated dynamic processes. For instance, it is possible to consider model errors in ESA by including the model error covariance matrix in the multivariate ESA (Hacker and Lei, 2015; Ren et al., 2019). Another approach is to adopt optimization algorithms to identify the model error structure that has the most significant impact on forecast errors (e.g., Duan et al., 2022; Zhang et al., 2023). Such advancements could lead to significant improvements in the accuracy and reliability of forecast uncertainty assessments. Additionally, exploring the influence of ensemble data assimilation techniques on ESA, such as various perturbation methods, could further enhance our understanding.

CRediT authorship contribution statement

Liangying Liu: Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jie Feng:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Conceptualization. **Li Ma:** Writing – review & editing, Formal analysis, Conceptualization. **Yanru Yang:** Writing – review & editing, Formal analysis, Conceptualization. **Xiaotian Wu:** Writing – review & editing. **Chao Wang:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The ensemble and control forecast data used in this study can be downloaded from the THORPEX Interactive Grand Global Ensemble (TIGGE) website available at <http://apps.ecmwf.int/datasets/data/tigge>. The best track data are issued by the CMA Tropical Cyclone Data Center (https://tcdatalab.typhoon.org.cn/zjlsjj_zlhq.html).

Acknowledgments

This study was jointly supported by the National Natural Science Foundation of China (Grant No. 42105054, No. 42288101, and No. 42375058) and the open project (KLME202003) from the Key Laboratory of Meteorological Disaster (KLME).

References

- Akter, N., Tsuboki, K., 2021. Recurvature and movement processes of tropical cyclones over the Bay of Bengal. *Q. J. R. Meteorol. Soc.* 147, 3681–3702. <https://doi.org/10.1002/qj.4148>.
- Alaka, G.J., Zhang, X., Gopalakrishnan, S.G., Zhang, Z., Marks, F.D., Atlas, R., 2019. Track uncertainty in high-resolution HWRF ensemble forecasts of hurricane Joaquin. *Weather Forecast.* 34, 1889–1908. <https://doi.org/10.1175/WAF-D-19-0028.1>.
- Ancell, B.C., Coleman, A.A., 2022. New perspectives on ensemble sensitivity analysis with applications to a climatology of severe convection. *Bull. Am. Meteorol. Soc.* 103, E511–E530. <https://doi.org/10.1175/BAMS-D-20-0321.1>.
- Ancell, B., Hakim, G.J., 2007. Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Weather Rev.* 135, 4117–4134. <https://doi.org/10.1175/2007MWR1904.1>.
- Ashcroft, J., Schwendike, J., Griffiths, S.D., Ross, A.N., Short, C.J., 2021. The impact of weak environmental steering flow on tropical cyclone track predictability. *Q. J. R. Meteorol. Soc.* 147, 4122–4142. <https://doi.org/10.1002/qj.4171>.
- Briegel, L.M., 2002. The Structure and Variability of the ITCZ of the Western North Pacific. *The Pennsylvania State University.*
- Chan, J.C.L., Gray, W.M., 1982. Tropical Cyclone Movement and Surrounding Flow Relationships. *Mon. Weather Rev.* 110, 1354–1374. [https://doi.org/10.1175/1520-0493\(1982\)110<1354:TCMASF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<1354:TCMASF>2.0.CO;2).
- Chang, E.K.M., Zheng, M., Raeder, K., 2013. Medium-range ensemble sensitivity analysis of two extreme pacific extratropical cyclones. *Mon. Weather Rev.* 141, 211–231. <https://doi.org/10.1175/MWR-D-11-00304.1>.
- Chen, G., Li, T., Yang, M., Zhang, X., 2023. Evaluation of western north pacific typhoon track forecasts in global and regional models during the 2021 typhoon season. *Atmosphere* 14. <https://doi.org/10.3390/atmos14030499>.
- Diaconescu, E.P., Laprise, R., 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., Huo, Z., 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-15-0188.1>.
- Duan, W., Ma, J., Vannitsem, S., 2022. An ensemble forecasting method for dealing with the combined effects of the initial and model errors and a potential deep learning implementation. *Mon. Weather Rev.* 150, 2959–2976. <https://doi.org/10.1175/MWR-D-22-0007.1>.
- Feng, J., Judt, F., Zhang, J., Wang, X., 2024. Influence of region-dependent error growth on the predictability of track and intensity of Typhoon Chan-hom (2020) in high-resolution HWRF ensembles. *Atmos. Res.* 308, 107536. <https://doi.org/10.1016/j.atmosres.2024.107536>.
- Franklin, J.L., Feuer, S.E., Kaplan, J., Aberson, S.D., 1996. Tropical cyclone motion and surrounding flow relationships: searching for beta gyres in omega dropwindsonde datasets. *Mon. Weather Rev.* 124, 64–84. [https://doi.org/10.1175/1520-0493\(1996\)124<0064:TCMASF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<0064:TCMASF>2.0.CO;2).
- Galarneau, T.J., Davis, C.A., 2013. Diagnosing forecast errors in tropical cyclone motion. *Mon. Weather Rev.* 141, 405–430. <https://doi.org/10.1175/MWR-D-12-00071.1>.
- George, J.E., Gray, W.M., 1976. Tropical cyclone motion and surrounding parameter relationships. *J. Appl. Meteorol. Climatol.* 15, 1252–1264. [https://doi.org/10.1175/15200450\(1976\)015<1252:TCMASP>2.0.CO;2](https://doi.org/10.1175/15200450(1976)015<1252:TCMASP>2.0.CO;2).
- Goerrs, J.S., Sampson, C.R., Gross, J.M., 2004. A history of western north pacific tropical cyclone track forecast skill. *Weather Forecast.* 19, 633–638. [https://doi.org/10.1175/1520-0434\(2004\)019<0633:AHOWNP>2.0.CO;2](https://doi.org/10.1175/1520-0434(2004)019<0633:AHOWNP>2.0.CO;2).
- Hacker, J.P., Lei, L., 2015. Multivariate ensemble sensitivity with localization. *Mon. Weather Rev.* 143, 2013–2027. <https://doi.org/10.1175/MWR-D-14-00309.1>.
- Hamill, T.M., Whitaker, J.S., Fiorino, M., Benjamin, S.G., 2011. Global ensemble predictions of 2009's tropical cyclones initialized with an ensemble kalman filter. *Mon. Weather Rev.* 139, 668–688. <https://doi.org/10.1175/2010MWR3456.1>.
- Hazelton, A., Alaka Jr., G.J., Fischer, M.S., Torn, R., Gopalakrishnan, S., 2023. Factors influencing the track of Hurricane Dorian (2019) in the West Atlantic: analysis of a HAFS ensemble. *Mon. Weather Rev.* 151, 175–192.
- Heming, J.T., et al., 2019. Review of recent progress in tropical cyclone track forecasting and expression of uncertainties. *Tropical Cyclone Res. Rev.* 8, 181–218. <https://doi.org/10.1016/j.tcr.2020.01.001>.
- Holland, G.J., 1984. Tropical cyclone motion. A comparison of theory and observation. *J. Atmos. Sci.* 41, 68–75. [https://doi.org/10.1175/1520-0469\(1984\)041<0068:TCMACO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1984)041<0068:TCMACO>2.0.CO;2).
- Huang, L., Wan, Q., Liu, C., Huang, H., 2020. Ensemble based diagnosis of the track errors of super Typhoon Mangkhut (2018). *J. Meteorol. Res.* 34, 353–367. <https://doi.org/10.1007/s13351-020-9086-x>.
- Huang, S.-H., Wen, Z.-P., Chen, X.-D., Guo, Y.-Y., Wang, Z.-W., 2022. The Henan extreme rainfall in July 2021: Modulation of the northward-shift monsoon trough on the synoptic-scale wave train. *Adv. Clim. Chang. Res.* 13, 819–825. <https://doi.org/10.1016/j.accre.2022.11.001>.
- Ito, K., Wu, C.-C., 2013. Typhoon-position-oriented sensitivity analysis. Part I: Theory and verification. *J. Atmos. Sci.* 70, 2525–2546. <https://doi.org/10.1175/JAS-D-12-0301.1>.
- Leonardo, N.M., Colle, B.A., 2021. An investigation of large cross-track errors in North Atlantic tropical cyclones in the GEFS and ECMWF ensembles. *Mon. Weather Rev.* 149, 395–417. <https://doi.org/10.1175/MWR-D-20-0035.1>.
- Li, J., Zhang, Z., Liu, L., Zhang, X., Qu, J., Wan, Q., 2021. The simulation of five tropical cyclones by sample optimization of ensemble forecasting based on the observed track and intensity. *Adv. Atmos. Sci.* 38, 1763–1777. <https://doi.org/10.1007/s00376-021-0353-2>.
- Limpert, G.L., Houston, A.L., 2018. Ensemble sensitivity analysis for targeted observations of supercell thunderstorms. *Mon. Weather Rev.* 146, 1705–1721. <https://doi.org/10.1175/MWR-D-17-0029.1>.
- Magnusson, L., Doyle, J.D., Komaromi, W.A., Torn, R.D., Tang, C.K., Johnny, C.L., Chan, M., Yamaguchi, Zhang, F., 2019. Advances in understanding difficult cases of tropical cyclone track forecasts. *Tropical Cyclone Res. Rev.* 8, 109–122. <https://doi.org/10.1016/j.tcr.2019.10.001>.
- Mu, M., Duan, W.S., Wang, B., 2003. Conditional nonlinear optimal perturbation and its applications. *Nonlinear Process. Geophys.* 10, 493–501. <https://doi.org/10.5194/npg-10-493-2003>.
- Nakano, M., Chen, Y.-W., Satoh, M., 2023. Analysis of the Factors that Led to Uncertainty of Track Forecast of Typhoon Krosa (2019) by 101-Member Ensemble Forecast Experiments Using NICAM. *Journal of the Meteorological Society of Japan. Ser. 101*, 191–207. <https://doi.org/10.2151/jmsj.2023-013>.
- Palmer, T., 2019. The ECMWF ensemble prediction system: looking back (more than 25 years and projecting forward 25 years. *Q. J. R. Meteorol. Soc.* 145, 12–24. <https://doi.org/10.1002/qj.3383>.
- Peng, F., Li, X., Chen, J., 2022a. Stochastically perturbed parameterizations for the process-level representation of model uncertainties in the CMA global ensemble prediction system. *J. Meteorol. Res.* 36, 733–749. <https://doi.org/10.1007/s13351-022-0118-8>.
- Peng, F., Li, X., Zhao, B., Chen, J., 2022b. Preliminary exploration on the origin of large-scale medium-range forecast errors over East ASIA in the CMA global numerical prediction system: a case study. *Meteorol. Monogr.* 48, 665–676. <https://doi.org/10.7519/j.issn.1000-0526.2021.090601>.
- Rao, J., Xie, J., Cao, Y., Zhu, S., Lu, Q., 2022. Record flood-producing rainstorms of July 2021 and August 1975 in Henan of China: comparative synoptic analysis using ERA5. *J. Meteorol. Res.* 36, 809–823. <https://doi.org/10.1007/s13351-022-2066-6>.
- Ren, S., Lei, L., Tan, Z.-M., Zhang, Y., 2019. Multivariate ensemble sensitivity analysis for super typhoon haiyan (2013). *Mon. Weather Rev.* 147, 3467–3480. <https://doi.org/10.1175/MWR-D-19-0074.1>.
- Ritchie, E.A., Holland, G.J., 1999. Large-scale patterns associated with tropical cyclogenesis in the western pacific. *Mon. Weather Rev.* 127, 2027–2043. [https://doi.org/10.1175/1520-0493\(1999\)127<2027:LSPAWT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2027:LSPAWT>2.0.CO;2).
- Sobel, A.H., Camargo, S.J., 2005. Influence of Western North Pacific Tropical cyclones on their large-scale environment. *J. Atmos. Sci.* 62, 3396–3407. <https://doi.org/10.1175/JAS3539.1>.
- Tang, C.K., Chan, J.C.L., Yamaguchi, M., 2021. Large tropical cyclone track forecast errors of global numerical weather prediction models in western North Pacific basin. *Tropical Cyclone Res. Rev.* 10, 151–169. <https://doi.org/10.1016/j.tcr.2021.07.001>.
- Torn, R.D., Hakim, G.J., 2008. Ensemble-based sensitivity analysis. *Mon. Weather Rev.* 136, 663–677. <https://doi.org/10.1175/2007MWR2132.1>.
- Torn, R.D., Whitaker, J.S., Piegion, P., Hamill, T.M., Hakim, G.J., 2015. Diagnosis of the source of GFS medium-range track errors in hurricane sandy (2012). *Mon. Weather Rev.* 143, 132–152. <https://doi.org/10.1175/MWR-D-14-00086.1>.
- Torn, R.D., Elless, T.J., Papin, P.P., Davis, C.A., 2018. Tropical cyclone track sensitivity in deformation steering flow. *Mon. Weather Rev.* 146, 3183–3201. <https://doi.org/10.1175/MWR-D-18-0153.1>.
- Velden, C.S., Leslie, L.M., 1991. The basic relationship between tropical cyclone intensity and the depth of the environmental steering layer in the Australian region. *Wea. Fore.* 6, 244–253.
- Wang, H., Lin, D., Xu, Y., Nie, G., 2022. Analysis on main characteristics of Typhoon In-fa and difficulties in its track forecast. *J. Marine Meteorol.* 42, 83–91. <https://doi.org/10.19513/j.cnki.issn2096-3599.2022.01.009>.

- Wu, L., Liang, J., Wu, C.-C., 2011. Monsoonal influence on typhoon morakot (2009). Part I: Observational analysis. *J. Atmos. Sci.* 68, 2208–2221. <https://doi.org/10.1175/2011JAS3730.1>.
- Wu, L., Wen, Z., Huang, R., Wu, R., 2012. Possible linkage between the monsoon trough variability and the tropical cyclone activity over the western north pacific. *Mon. Weather Rev.* 140, 140–150. <https://doi.org/10.1175/MWR-D-11-00078.1>.
- Xu, J., Li, R., Zhang, Q., Chen, Y., Liang, X., Gu, X., 2022. Extreme large-scale atmospheric circulation associated with the “21·7” Henan flood. *Sci. China Earth Sci.* 65, 1847–1860. <https://doi.org/10.1007/s11430-022-9975-0>.
- Yu, H., et al., 2022. Are we reaching the limit of tropical cyclone track predictability in the Western North Pacific? *Bull. Am. Meteorol. Soc.* 103, E410–E428. <https://doi.org/10.1175/BAMS-D-20-0308.1>.
- Zack, J., Natenberg, E., Young, S., Manobianco, J., Kamath, C., 2010. Application of ensemble sensitivity analysis to observation targeting for short-term wind speed forecasting. <https://doi.org/10.2172/972845>.
- Zhang, X., Yu, H., 2017. A probabilistic tropical cyclone track forecast scheme based on the selective consensus of ensemble prediction systems. *Weather Forecast.* 32, 2143–2157. <https://doi.org/10.1175/WAF-D-17-0071.1>.
- Zhang, G., Murakami, H., Knutson, T.R., Mizuta, R., Yoshida, Kohei, 2020. Tropical cyclone motion in a changing climate. *Sci. Adv.* 6, eaaz7610 <https://doi.org/10.1126/sciadv.aaz7610>.
- Zhang, H., Duan, W., Zhang, Y., 2023. Using the orthogonal conditional nonlinear optimal perturbations approach to address the uncertainties of tropical cyclone track forecasts generated by the WRF model. *Weather Forecast.* 38, 1907–1933. <https://doi.org/10.1175/WAF-D-22-0175.1>.
- Zheng, M., Chang, E.K.M., Colle, B.A., 2013. Ensemble sensitivity tools for assessing extratropical cyclone intensity and track predictability. *Weather Forecast.* 28, 1133–1156. <https://doi.org/10.1175/WAF-D-12-00132.1>.
- Zhou, X., et al., 2022. The development of the NCEP global ensemble forecast system version 12. *Weather Forecast.* 37, 1069–1084. <https://doi.org/10.1175/WAF-D-21-0112.1>.