JGR Atmospheres

RESEARCH ARTICLE
10.1029/2024JD043261

# Improving Tropical Cyclone Track Forecast Skill Through Assimilating Target Observation Achieved by AI-Based Conditional Nonlinear Optimal Perturbation

**Yonghui Li[1,2,3], Wansuo Duan[1,2], Wei Han[4,5], Hao Li[3,6], and Xiaohao Qin[7]**

[1]State Key Laboratory of Earth System Numerical Modeling and Application, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]Shanghai Academy of Artificial Intelligence for Science (SAIS), Shanghai, China, [4]State Key Laboratory of Severe Weather Meteorological Science and Technology (LASW), Beijing, China, [5]CMA Earth System Modeling and Prediction Centre (CEMC), China Meteorological Administration, Beijing, China, [6]Artificial Intelligence Innovation and Incubation Institute, Fudan University, Shanghai, China, [7]State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

**Abstract** The artificial intelligence (AI)-based weather forecasting model named FuXi and its data assimilation (DA) system FuXi-En4DVar has been developed for high-efficiently forecasting high-impact weather events such as tropical cyclones (TCs). Besides conventional observations, target observations are essential to further improve initial field accuracy and then increasing high-impact weather event forecasting skills. The identification of the sensitive area, where the additional observations should be deployed, is the key to implementing target observations. In this paper, a sensitive area identification system is established for the FuXi model on the basis of FuXi-En4DVar, based on the fully nonlinear method of conditional nonlinear optimal perturbation (CNOP). The CNOP represents the optimally growing initial perturbation and can be calculated by using the adjoint of numerical models in numerical forecast models, but in the AI-based FuXi model, it is solved by directly using the automatic differential algorithm embedded in the FuXi model. Such an approach of calculating CNOP significantly increases the computational efficiency. Applying this system to the forecasts of 11 TCs demonstrates that the additional target observations can significantly improve TC track forecast skills, as compared with the other additional observations. Moreover, a small number of additional target observations can be expected to achieve the forecast skill comparable to, or even surpassing to, that obtained by tens of times more observations. This validation shows the potential of applying dynamical CNOP to AI-based model for highly effectively identifying the sensitive area for target observations associated with TC forecasting.

**Plain Language Summary** This study developed a conditional nonlinear optimal perturbation (CNOP)-based sensitive area identification system for the AI model, FuXi, to enhance the quality of the initial fields and lead to more reliable forecasts. By conducting observing system simulation experiments on tropical cyclones (TCs), we assimilate observations in the CNOP sensitive areas using the FuXi-En4DVar system and forecast with FuXi. The results show that assimilating observations within CNOP sensitive areas leads to better TC track predictions than observations elsewhere. This end-to-end process—ranging from sensitive area identification to observation assimilation and forecasting—is fully AI-driven and offers a promising approach for advancing future numerical weather prediction systems.

**Key Points:**

- AI optimization system for conditional nonlinear optimal perturbation (CNOP) is developed for identifying the sensitive areas of tropical cyclones (TC).
- Assimilating target observation obviously improves TC track forecast skills.
- The end-to-end forecast is realized in FuXi by En4DVar with target observations.

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

W. Duan and W. Han,
duanws@lasg.iap.ac.cn;
hanwei@cma.gov.cn

**Citation:**

Li, Y., Duan, W., Han, W., Li, H., & Qin, X. (2025). Improving tropical cyclone track forecast skill through assimilating target observation achieved by AI-based conditional nonlinear optimal perturbation. *Journal of Geophysical Research: Atmospheres*, *130*, e2024JD043261. https://doi.org/10.1029/2024JD043261

Received 7 JAN 2025
Accepted 28 APR 2025

**Author Contributions:**

**Conceptualization:** Yonghui Li, Wansuo Duan, Wei Han, Hao Li
**Data curation:** Yonghui Li, Wansuo Duan
**Formal analysis:** Yonghui Li, Wansuo Duan
**Funding acquisition:** Yonghui Li, Wansuo Duan
**Investigation:** Yonghui Li, Wansuo Duan, Wei Han
**Methodology:** Yonghui Li, Wansuo Duan, Wei Han, Hao Li
**Project administration:** Yonghui Li, Wansuo Duan, Wei Han
**Resources:** Yonghui Li, Wansuo Duan, Wei Han
**Software:** Yonghui Li, Wansuo Duan, Hao Li
**Validation:** Yonghui Li, Wansuo Duan, Xiaohao Qin
**Visualization:** Yonghui Li, Wansuo Duan



## 1. Introduction

Accurate weather forecasting plays a crucial role in many aspects of society. Currently, the vast majority of operational forecasting centers rely on traditional numerical models, known as numerical weather prediction (NWP), for weather forecasting. These numerical models are built upon the physical principles governing atmospheric dynamic. Running an NWP model depends on high-performance computing systems, but limited computational resources have constrained its development (Brotzge et al., 2023). Alternatively, artificial intelligence (AI)-based weather forecasting models offer promising alternatives with enhanced computational efficiency, several orders of magnitude faster than traditional NWP models. Currently, AI-based weather forecasting models have made significant progress, with examples including FourCastNet (2022) (Pathak et al., 2022), GraphCast (2023) (Lam et al., 2023), NowcastNet (2023) (Y. Zhang et al., 2023), Pangu-Weather (2023) (Bi

**Writing – original draft:** Yonghui Li, Wansuo Duan
**Writing – review & editing:** Yonghui Li, Wansuo Duan, Wei Han, Hao Li, Xiaohao Qin

et al., 2023), Fengwu (2023) (K. Chen et al., 2023), FuXi (2023) (L. Chen et al., 2023), and AIFS (2024) (Lang et al., 2024). These models have demonstrated remarkable advantages and great potential in improving forecasting accuracy, which match or even surpass traditional NWP models such as the Integrated Forecasting System (IFS) from the European Centre for Medium-Range Weather Forecasts (ECMWF). For instance, FuXi has demonstrated forecast performance comparable to the ensemble mean of ECMWF in 15-day forecasts.

The aforementioned AI-based models focus on short- to medium-term weather forecasting, particularly in predicting synoptic-scale systems, where the primary source of forecast errors originates from initial condition errors (Ek, 2018), which is the predictability problem of the first kind (Lorenz, 1975). In the past, AI-based weather forecasting models often directly used ECMWF Reanalysis v5 (ERA5) data (Hersbach et al., 2020) to reduce initial condition errors. Alternatively, forecast fields from operational forecasting centers are used as initial conditions to drive the AI-based models, which would lead to greater initial errors. To enable AI models to gradually reduce dependence on traditional NWP analysis fields and improve the accuracy of the initial conditions, FuXi-En4DVar (Y. Li, Han, Li, et al., 2024), FengWu-4DVar (Xiao et al., 2023), and FuXi-DA (Xu et al., 2024) have made preliminary progress, enabling the assimilation of observational data, which enhances the accuracy of the initial conditions.

As data assimilation (DA) techniques continue to advance and the observational capabilities of Earth observation systems rapidly improve, new opportunities arise to enhance the quality of initial conditions in both numerical forecasting and AI-based weather forecasting models. However, the data obtained from existing systems remain sparse relative to model resolution and are unevenly distributed, especially in oceanic regions. Consequently, the efficient use of limited observations has become a critical focus. This is where the concept of targeted observations proposed by meteorologists (Bergot, 1999; Emanuel et al., 1997; Rabier et al., 1996; Snyder, 1996), also known as adaptive observations, comes into play. Targeted observation involves identifying sensitive areas at decision time ($\mathbf{t}_d$) through objective methods to improve the quality of weather forecasts for specific times (verification times $\mathbf{t}_v$) and areas (verification areas). Enhanced observations, using tools such as satellites, airborne radars, and dropsondes, are conducted in these sensitive areas and then assimilated at analysis time ($\mathbf{t}_a$) to reduce forecast errors at verification time ($\mathbf{t}_v$) (Figure 1). It has been employed in various field observation experiments, such as The Fronts and Atlantic Storm-Track Experiment (FASTEX) (Joly et al., 1997), the North Pacific experiment (NORPEX) (Langland et al., 1999), the 2000 Winter Storm Reconnaissance (WSR00) (Szunyogh et al., 2002), and the Dropsonde Observations for Typhoon Surveillance near the Taiwan Region (DOTSTAR) (C.-C. Wu et al., 2005, 2007), recently, several field experiments focusing on targeted observations of tropical cyclones (TCs) have been conducted (Feng et al., 2022; Qin et al., 2023).

In the past, identifying sensitive areas for targeted observations was typically done using traditional NWP (Phase 1–3 in Figure 1) with the calculation of these sensitive areas dependent on models. Thus, the sensitive areas identified by NWP are not suitable for AI models. To further improve the accuracy of the initial conditions in AI-based models, an effective sensitive area identification system for AI-based models is crucial. In this study, we construct such a system based on the FuXi model and use FuXi-En4DVar (Y. Li, Han, Li, et al., 2024) for the assimilation of observations, thereby improving forecasting skills of high-impact weather (Phase 4 in Figure 1). This approach realizes the entire chain of targeted observations using AI-based models—from the initial design of the observation array to the final forecast—entirely driven by AI. Currently, the methods for identifying these sensitive areas can be broadly categorized into two types. The first type focuses on the dynamics of error growth, identifying initial errors with rapid growth or calculating the gradient of an objective function with respect to the initial state to identify sensitive areas. Examples include the leading singular vector (LSV) method (Palmer et al., 1998), conditional nonlinear optimal perturbation (CNOP) method (Duan et al., 2018; Duan & Huo, 2016; Mu et al., 2003, 2009), breeding vector (BV) method (Toth & Kalnay, 1997), and adjoint sensitivity (ADS) method (Ancell & Mass, 2006; Baker & Daley, 2000; Kim et al., 2004). The second type involves ensemble-based methods, such as ensemble sensitivity analysis (ESA) (Torn & Hakim, 2008), ensemble Kalman filter (ENKF) (Hamill & Snyder, 2002), and ensemble transform Kalman filter (ETKF) (Bishop et al., 2001). Each method has its own strengths and weaknesses. For instance, the ESA, ENKF, ETKF, LSV, and ADV are linear approaches that cannot describe well the nonlinear growth of error, while BV, though nonlinear, lacks a solid theoretical foundation. CNOP, on the other hand, is a nonlinear method with significant theoretical advantages (Duan et al., 2004; Duan & Mu, 2009; Mu et al., 2003, 2007), though its computation is costly in traditional NWP. In AI-based weather forecasting models, the issue of slow CNOP computation can be mitigated by avoiding the execution of tangent linear and adjoint models. In this work, a CNOP-based targeted observation system will be

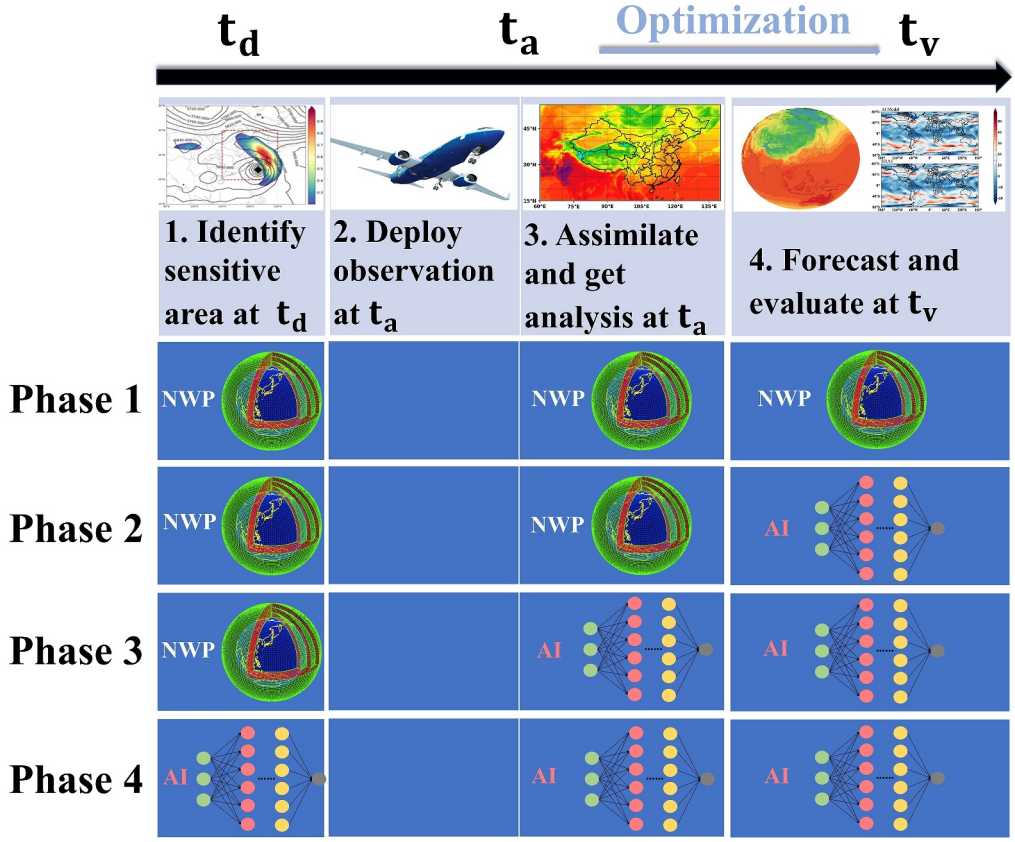


**Figure 1.** Four development phases of targeted observation. "NWP" refers to the use of the traditional numerical weather prediction model at a stage of targeted observations, while "AI" represents the application of the AI-based meteorological forecasting model.

developed for the FuXi model, and its effectiveness is validated through multiple TC cases. A chain from the AI-based FuXi model and its DA system to greatly improve TC forecast skills is connected through the CNOP sensitive area identification system.

The structure of the article is organized as follows: Section 2 introduces the development of the CNOP-based sensitive area identification system for the FuXi model. Section 3 describes the experimental design, while the results of the experiments are presented in Section 4. Finally, Section 5 provides the conclusions and discussions.

## 2. Method

### 2.1. The Model and Assimilating System

The FuXi Model (L. Chen et al., 2023), developed by the Artificial Intelligence Innovation and Incubation Institute at Fudan University, features a cascaded architecture designed to provide a 15-day global forecast with a temporal resolution of 6 hr and a spatial resolution of 0.25°. It incorporates 70 weather variables, utilizing the input data combining both current and previous time steps with dimensions of $2 \times 70 \times 721 \times 1440$. This model is developed utilizing a 39-year ERA5 reanalysis data set. Forecast performance is evaluated using latitude-weighted root mean square error (RMSE) and anomaly correlation coefficient (ACC). Notably, its 15-day forecasts are comparable to those of the ECMWF ensemble mean, making it the first AI-based model to achieve such accuracy.

The FuXi-En4DVar (Y. Li, Han, Li, et al., 2024) system, built upon the FuXi model, employs a 6-hr assimilation window, considering observations only at the current time and 6 hr later. Unlike traditional systems, this assimilation process does not rely on tangent linear and adjoint models to compute the gradient of the objective function. Instead, it uses automatic differentiation to calculate the gradient. The background error covariance matrix is generated through ensemble calculations, taking advantage of the ability of the AI-based model to

quickly produce a substantial number of samples. The analysis increments produced by this system satisfy basic physical constraints and exhibit flow-dependent characteristics. When assimilating observations with this system, the analysis field shows reduced errors compared to the background field.

### 2.2. The CNOP Method

Assuming the state vector is denoted as $\mathbf{x}$, representing variables in the model, the evolution of the state vector can be described by the following system of nonlinear partial differential equations:

$$\begin{cases} \dfrac{\partial \mathbf{x}}{\partial t} = F(\mathbf{x}) \\ \mathbf{x}|_{t=0} = \mathbf{x}_0 \end{cases} \tag{1}$$

where $F$ is a nonlinear differential operator, and $\mathbf{x}_0$ represents the initial state. Therefore, at time $\tau$, the solution to the system of Equation 1 can be expressed as follows:

$$\mathbf{x}_\tau = M_\tau(\mathbf{x}_0) \tag{2}$$

where $M_\tau$ is a nonlinear propagation operator (In this paper, $M$ is the FuXi model). Thus, at the forecast time, the forecast error $\delta\mathbf{x}_\tau$ caused by the initial error $\delta\mathbf{x}_0$ can be calculated by the following equation:

$$\delta\mathbf{x}_\tau = M_\tau(\mathbf{x}_0 + \delta\mathbf{x}_0) - M_\tau(\mathbf{x}_0) \tag{3}$$

According to the above equation defining CNOP, it represents all initial perturbations that satisfy given physical constraints and cause the maximum forecast error within the optimization period. In other words, the initial perturbation $\delta\mathbf{x}_0^*$ is referred to as CNOP if and only if

$$J(\delta\mathbf{x}_0^*) = \max_{\|\delta\mathbf{x}_0\|_{C_2} \le \epsilon} \|M_\tau(\mathbf{x}_0 + \delta\mathbf{x}_0) - M_\tau(\mathbf{x}_0)\|_{C_1} \tag{4}$$

where $\|\|_{C_2}$ is the norm measuring the magnitude of the initial perturbation, $\|\delta\mathbf{x}_0\|_{C_2} \le \epsilon$ represents the constraint condition on the initial perturbation (where $\epsilon$ is a constant), and $\|\|_{C_1}$ is the norm measuring the forecast error magnitude at the forecast time (i.e., the objective function), with $\tau$ being the optimization duration.

In this study, the initial perturbation $\delta\mathbf{x}_0$ is composed of $\mathbf{u}_0, \mathbf{v}_0, \mathbf{T}_0, \mathbf{p}_{s0}$, which are the perturbed zonal and meridional wind components, temperature, and mean sea level pressure, respectively, at the initial time. $\|\|_{C_2}$ and $\|\|_{C_1}$ represent the norm of total dry energy and kinetic energy, respectively. The reason for choosing kinetic energy as the objective function is that the TC track is mainly determined by the momentum field. The continuous form is expressed as follows:

$$\|\delta\mathbf{x}_0\|_{C_2}^2 = \frac{1}{D_2}\int_\sigma\int_D \left[\mathbf{u}_0'^2 + \mathbf{v}_0'^2 + \frac{c_p}{T_r}\mathbf{T}_0'^2 + R_a T_r\left(\frac{\mathbf{p}_{s0}'}{p_r}\right)^2\right] dD_2 d\sigma, \tag{5}$$

and

$$J = \frac{1}{D_1}\int_\sigma\int_{D_1} \left[\mathbf{u}_\tau'^2 + \mathbf{v}_\tau'^2\right] dD_1 d\sigma, \tag{6}$$

where $c_p$ and $R_a$ are the specific heat at constant pressure and the gas constant of air, respectively (with numerical values of 1,005.7 J kg$^{-1}$ K$^{-1}$ and 287.04 J kg$^{-1}$ K$^{-1}$). The reference parameters are the following. $T_r = 270$ K and $p_r = 1000$ hPa. $J$ is the cost function defined as the kinetic energy over the verification area $D_1$. Here, $\mathbf{u}_\tau'$ and $\mathbf{v}_\tau'$ are components of the state vector $\delta\mathbf{x}_\tau$, which is the nonlinear development of $\delta\mathbf{x}_0$ at time $\tau$. The integration extends over the full domain $D_2$ and the vertical direction $\sigma$. The initial constraint value $\epsilon$ is chosen as 0.3 J kg$^{-1}$, and the optimization time $\tau$ is 1 day.

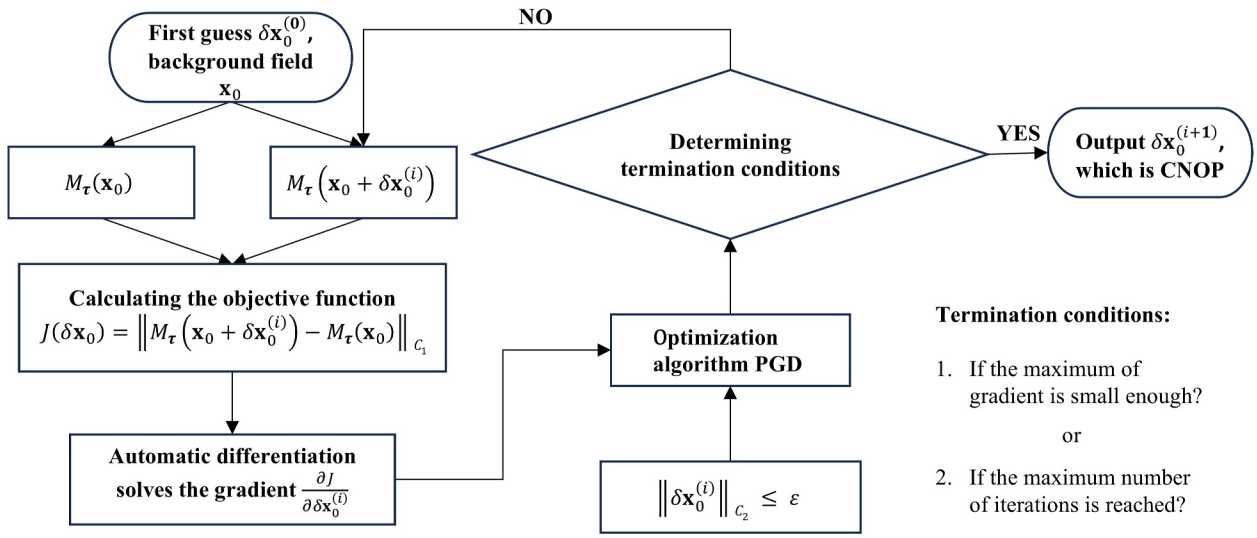


**Figure 2.** Calculation process of conditional nonlinear optimal perturbation based on the FuXi model.

### 2.3. CNOP Calculation Process

According to Equation 4, the calculation framework of CNOP is established based on the FuXi model. The flow chart of CNOP calculation is as shown in Figure 2. It is worth nothing that the calculation of CNOP based on the FuXi model does not rely on tangential linearity and the adjoint model but through the automatic differential solution gradient (Paszke et al., 2017), which reduces the engineering amount of scripting and improves the computational efficiency. And the optimization algorithm is projected gradient descent (PGD). The termination condition in Figure 2 is if the number of iterations reaches 50 or the gradient maximum is less than $10^{-6}$.

### 2.4. Definition of Sensitive Areas

In previous studies, the horizontal sensitive areas were defined by vertically integrating the total dry energy (Buizza et al., 2007; Zhou & Mu, 2012). This integration was performed using the following equation:

$$f(i,j)=\int_0^1 E_d(i,j,\sigma)d\sigma, \tag{7}$$

where $E_d(i,j,\sigma)$ is the total dry energy of the perturbation at the grid point $(i,j,\sigma)$. The horizontal grid points where the value $f(i,j)$ is larger than a certain value $c$ are defined as the sensitive areas. $E_d(i,j,\sigma)$ can be changed according to the specific problem, not necessarily the total dry energy. In this study, $E_d(i,j,\sigma)$ represents the total dry energy and the threshold of the sensitive areas $c$ is determined to be 10 J kg$^{-1}$.

## 3. Experiment Design

To validate the effectiveness of the CNOP-identified sensitive areas based on the FuXi model, 11 TC cases are selected for analysis, as detailed in Table 1 (The atmospheric circulation of the 11 TCs is presented in Supplementary Information (SI) Figures S1–S4 in Supporting Information S1). All these cases meet rigorous selection criteria: They exhibited severe typhoon strength during the forecast period and remained within China's 48-hr alert window. Besides, these TC cases are intentionally outside the temporal scope of the FuXi model's training data set, which is much realistic and helpful for examining the performance of the FuXi model in predicting the TCs except for those in the training period. In this study, the verification areas are designed with a proper size according to the location of each TC at prediction time by consulting the simulations from the FuXi model and the tropical TC data released by China Meteorological Administration. The TC tracking algorithm employed in this study is sourced from ECMWF (Der Grijn, 2002). Zhou and Mu (2011) studied the impact of different verification areas on CNOP

**Table 1**
*Tropical Cyclone Cases*

| Cases | TC | Verification areas | Analysis time |
|---|---|---|---|
| Case_I | Maysak | 26°N–31°N, 124°E–129°E | 0000 UTC 01 Sep 2020 |
| Case_II | In-fa | 21°N–26°N, 126°E–132°E | 0000 UTC 20 July 2021 |
| Case_III | | 22°N–27°N, 124°E–129°E | 0000 UTC 22 July 2021 |
| Case_IV | Chanthu | 20°N–25°N, 120°E–125°E | 0000 UTC 11 Sep 2021 |
| Case_V | | 23°N–30°N, 121°E–126°E | 0000 UTC 12 Sep 2021 |
| Case_VI | | 28°N–33°N, 121°E–126°E | 0000 UTC 13 Sep 2021 |
| Case_VII | | 28°N–33°N, 122°E–127°E | 0000 UTC 14 Sep 2021 |
| Case_VIII | Khanun | 18°N–23°N, 130°E–135°E | 0000 UTC 30 July 2023 |
| Case_IX | | 21°N–26°N, 128°E–133°E | 0000 UTC 31 July 2023 |
| Case_X | | 23°N–28°N, 125°E–133°E | 0000 UTC 01 Aug 2023 |
| Case_XI | | 24°N–29°N, 123°E–128°E | 0000 UTC 02 Aug 2023 |

and its identified sensitive areas. Their findings indicated that once the general positions of the verification areas were determined, minor variations in size or position had minimal influence on the results.

To assess the impact of assimilating observations from the sensitive area on TC analysis and prediction, six experiments are designed for comparison for each TC case. Since the sensitive areas for TCs are mostly over the ocean, where observations are sparse, an observing system simulation experiment is conducted to simulate observations and evaluate the forecasts. The simulated observations are based on ERA5 data, including temperature, specific humidity, u-wind, and v-wind at the 150, 200, 250, 300, 500, 700, 850, and 1,000 hPa levels, similar to the dropsonde observation. To reflect real-world observational uncertainties, Gaussian noise with a mean of 0 and a variance equal to 0.001 times the standard deviation of observation errors is added to the ERA5 data. These observations are recorded every 6 hr, aligning with the temporal resolution of the FuXi model.

The six experiments are as follows (see Table 2 for detail). Nature Run (hereafter called "NR"): ERA5 data are directly used as the truth. Control experiment (hereafter called "CTRL"): A forecast of the FuXi model is initiated using the ERA5 analysis from 2 days before the initial time as the initial condition. This forecast field serves as the initial condition at the initial time and there are no observations for assimilation. Large-scale observation experiment (hereafter called "ALL"): Observations are arranged at 1°intervals within the region (10°N–60°N, 105°E–145°E). The FuXi-En4DVar (Y. Li, Han, Li, et al., 2024) system is then used to assimilate the observations and forecast. CNOP experiment (hereafter called "EXP_CNOP"): Observations are placed at 1° intervals within the CNOP-sensitive area. If the sensitive area extended beyond 10°N–60°N and 105°E–145°E, observations outside this region are excluded. Random experiment (hereafter called "RAN"): Observations are randomly distributed within 10°N–60°N and 105°E–145°E with the number of observations matching that of the EXP_CNOP. No sensitive area random experiment (hereafter called "NoSA_RAN"): Similar to the RAN, but the random observations are placed within 10°N–60°N and 105°E–145°E, excluding the CNOP-sensitive areas with the same number of observations as in the EXP_CNOP.

## 4. Results

### 4.1. Sensitive Areas for Each Cases

Using the above framework to calculate CNOP, we can identify the sensitive areas for each case, as shown in Figure 3. For all cases, the CNOP-sensitive areas are primarily concentrated around the verification areas and near the TC itself. When this type of error is reduced, the accuracy of the initial fields for the TC and its surrounding areas improves, leading to better TC track forecasts one day later. This partly explains why reducing CNOP-type

**Table 2**
*Experiment Setup*

| Experiment names | Observations assimilated | Forecast length (days) |
|---|---|---|
| NR | ERA5 data as "truth" | 3 |
| CTRL | No observations | 3 |
| EXP_CNOP | Assimilating the observation of the intersection of the CNOP- sensitive areas and (10°N–60°N, 105°E–145°E), where the observation is placed at 1°interval | 3 |
| RAN | Assimilating random observations in (10°N–60°N, 105°E–145°E), where the number of observations is consistent with the number observed in the EXP_CNOP | 3 |
| NoSA_RAN | Assimilating random observations in specific areas, where specific areas are (10°N–60°N, 105°E–145°E) subtracting the CNOP-sensitive areas, and the number of observations is consistent with the number observed in the EXP_CNOP | 3 |
| ALL | Assimilating observations in (10°N–60°N, 105°E–145°E), where the observation is placed at 1°interval | 3 |

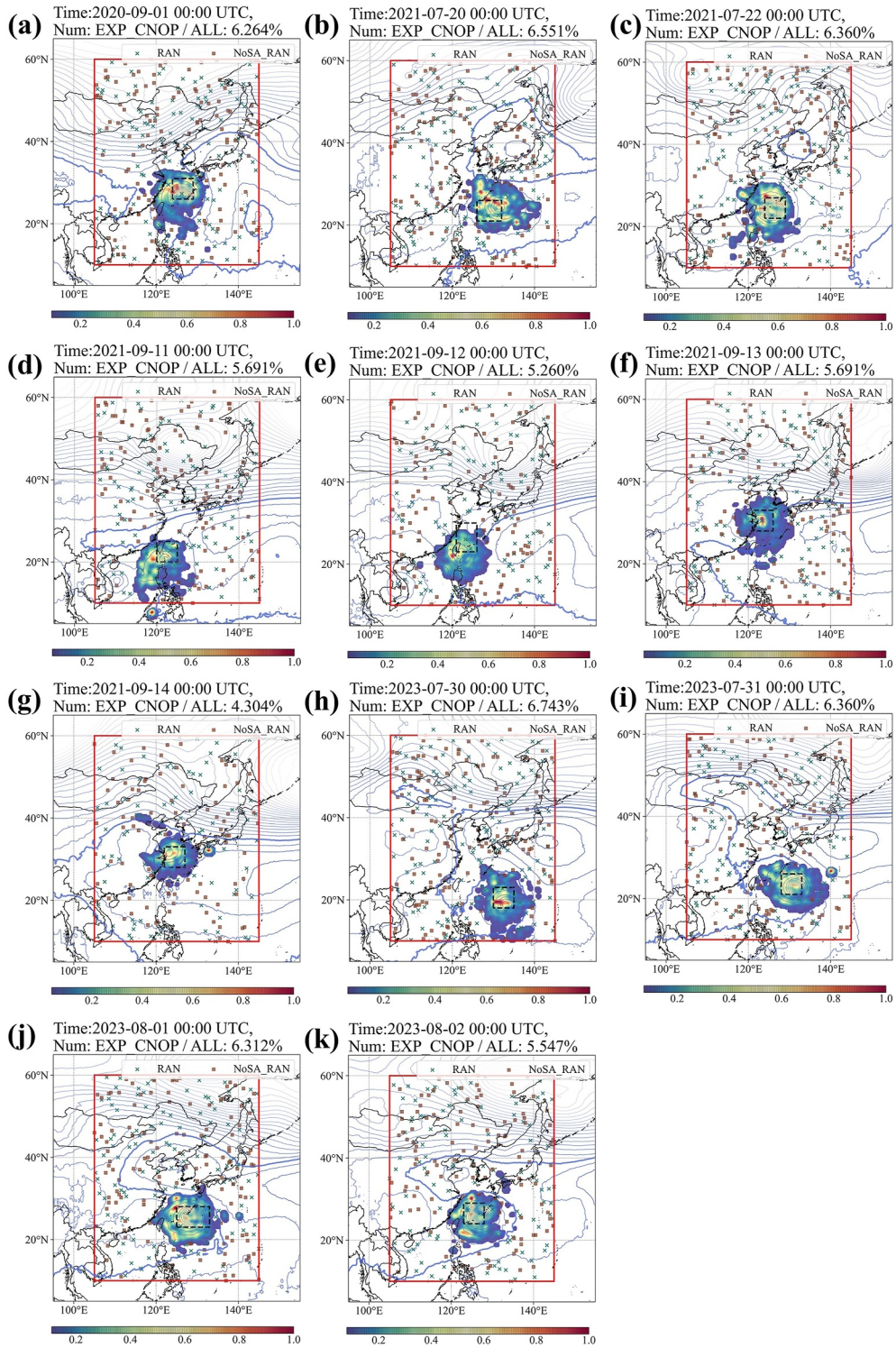


**Figure 3.** Conditional nonlinear optimal perturbation (CNOP) sensitive areas. In the figure, the shaded areas represent the CNOP sensitive areas, while the black dashed box indicates the verification areas. The red box represents the observation areas for the ALL, and green x-mark denotes the observation points for the RAN, and the brown square symbols denote the observation points for the NoSA_RAN. The observation points in EXP_CNOP correspond to the intersection between the shaded area and the red box. The light blue lines represent the contour of the geopotential height and the thick blue line represents the 5,880 gpm contour at the 500 hPa. The label "Num" in the upper left corner shows the percentage of observation points in the EXP_CNOP relative to the observation points in the ALL (For more details, see Table S1 in Supporting Information S1). Figures (a)–(k) correspond to cases I-XI in Table 1. The energy values have been normalized.

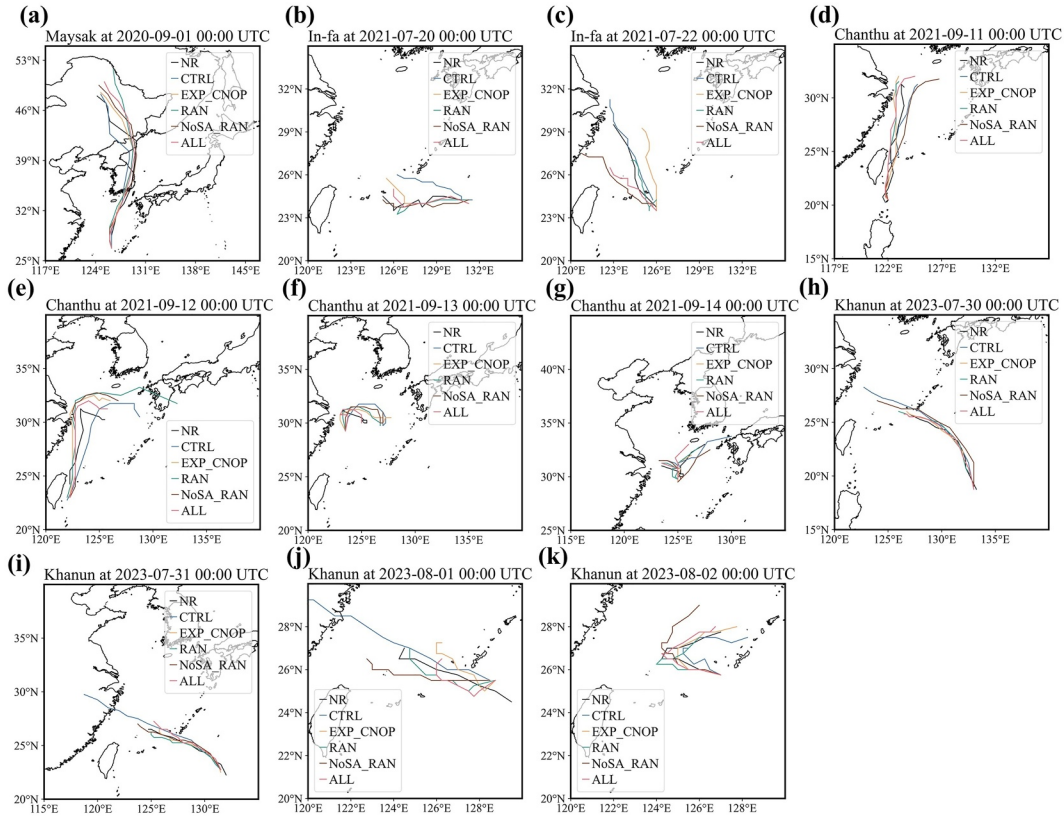


**Figure 4.** The forecasts of the tropical cyclone (TC) track. The black line represents the NR TC track, and the blue, yellow, green, brown, and red lines represent CTRL, EXP_CNOP, RAN, NoSA_RAN, and ALL TC track forecasts, respectively. Figures (a)–(k) correspond to cases I-XI in Table 1. The forecast initialization time is indicated at the top of each figure.

initial errors can reduce TC forecast errors after one day. It also underscores that, with a one-day optimization period, the accuracy of the TC's core physical state is a critical factor influencing the forecast accuracy for the following day, consistent with previous studies (G. Chen et al., 2018; Qin et al., 2020; Wang et al., 2025). And except for case III (For detailed information, please refer to Figure S2 in Supporting Information S1), the movement of TCs in the other cases is mainly guided by the subtropical high (thick blue solid line).

### 4.2. Track Forecasting

TC track forecasting is a critical metric for evaluating TC forecasting skills. It is important to highlight that, on average, the number of observations in EXP_CNOP, RAN, and NoSA_RAN accounts for only 5.9% of those in ALL across the 11 cases, as shown in Figure 3. Figures 4a–4k and 5a–5k compare the TC track forecasts from EXP_CNOP, RAN, NoSA_RAN, CTRL, and NR for the 11 cases. Overall, CTRL produces the poorest track forecasts compared to the other experiments that assimilate observations. Except for cases I and IV (Figures 4a, 4d, 5a, and 5d), the experiments with assimilated observations show similar forecasting skills to the CTRL, possibly due to dynamic inconsistencies in the adjusted analysis field caused by an imperfect assimilation system in this cases. For most cases, the ALL yielded the best track forecasts. Notably, in case VIII (Figures 4h and 5h), EXP_CNOP provided the most accurate track forecast, demonstrating that reducing perturbations in the sensitive areas not only have the greatest impact on improving the forecast after one day but also remained effective for later forecast. When observations are not located in the sensitive areas, as in NoSA_RAN, the forecast results are worse than those from EXP_CNOP, ALL, and RAN, especially for cases VII-VIII (Figures 4g–4h and 5g–5h), where the initial forecast even deteriorated compared to the CTRL, leading to generally poor subsequent forecasts. In RAN, the results are consistently poorer than those of EXP_CNOP for the first 24 hr. However, in the later forecast, the gap between RAN and EXP_CNOP narrows or even reverses (e.g., in cases II, VI, and X,

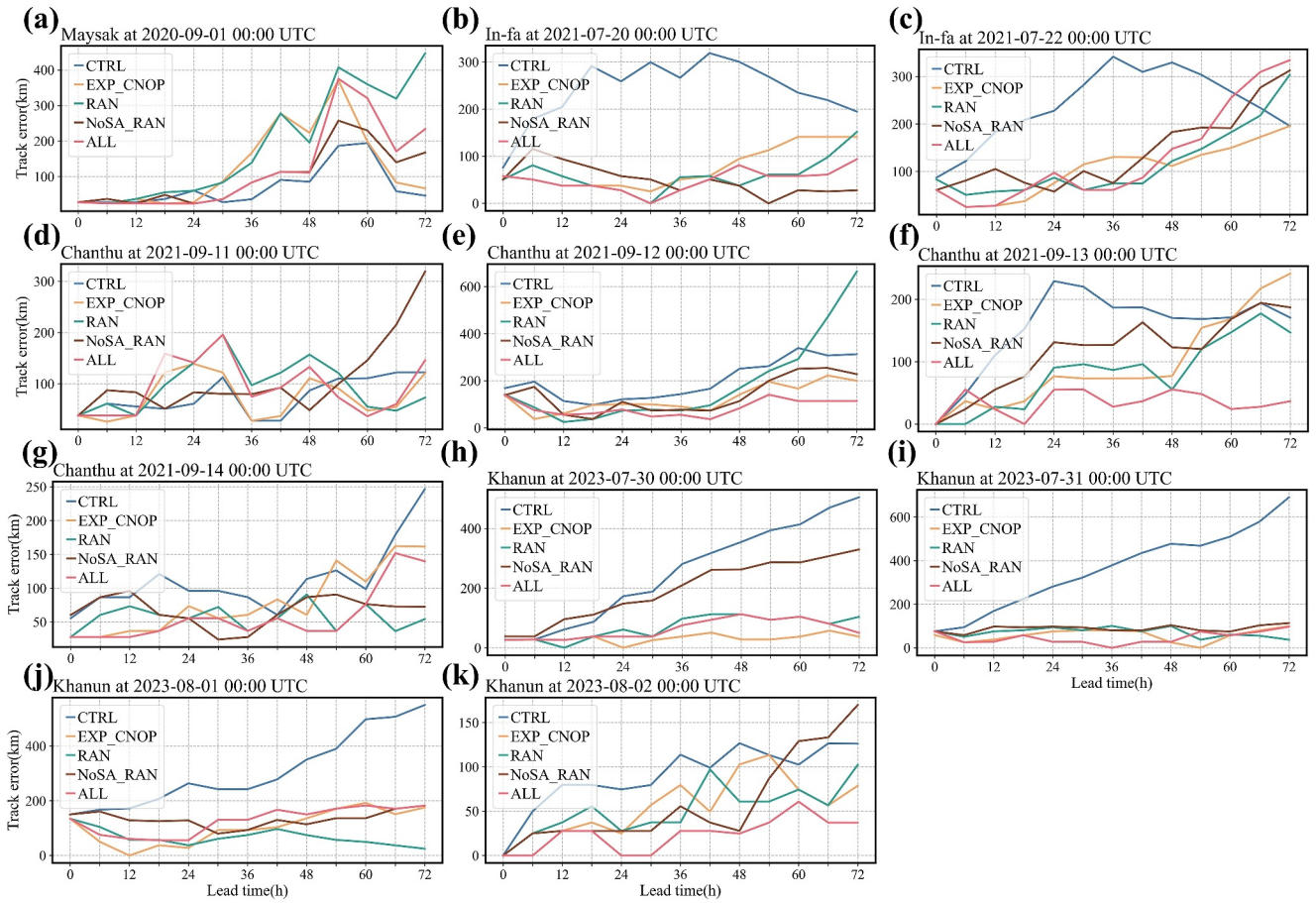


**Figure 5.** Tropical cyclone track forecast errors (unit: km) for CTRL (blue), EXP_CNOP (yellow), RAN (green), NoSA_RAN (brown), and ALL (red) verified against the NR track for the 11 cases. Figures (a)–(k) correspond to cases I-XI in Table 1. The forecast initialization time is indicated at the top of each figure.

Figures 4b, 4f, 4j, 5b, 5f, and 5j). Nonetheless, over the entire forecast period, RAN performs worse than EXP_CNOP.

Averaging the results across the 11 cases (Figures 6a and 6b), it becomes evident that CTRL yields the poorest forecasts, directly demonstrating the effectiveness of the FuXi-En4Var assimilation system. Following CTRL, NoSA_RAN produces the next poorest results, with RAN performing slightly better. Notably, EXP_CNOP exhibited highest forecasting skills compared to other experiments during the first 24 hr, with an average improvement of 55.92% for EXP_CNOP, outperforming ALL at 52.69%, RAN at 45.67%, and NoSA_RAN at 26.67%. This difference is particularly pronounced at the 12-hr forecast mark, where EXP_CNOP achieved a 72.04% improvement. It noted that the optimization time is 24 hr. Beyond 24 hr, ALL maintains the highest forecasting skill among all experiments, but the performance of EXP_CNOP is only marginally lower with an average difference in the improvement percentage of about 6.49%. In contrast, NoSA_RAN and RAN exhibit a more significant gap, with average differences of 15.18% and 10.95%, respectively. Given that the number of observations in EXP_CNOP accounted for only 5.91% of those in ALL, these results suggest that reducing errors in CNOP-sensitive areas can significantly decrease forecast errors, achieving forecasting skills comparable to that of large-scale observations.

It can be observed that at the 54-hr forecast (Figure 6a), EXP_CNOP underperforms compared to RAN. This reason may be that EXP_CNOP is focused solely on reducing forecast errors in the verification area 24 hr ahead. In contrast, RAN may include some observations in sensitive areas over a 54-hr optimization period, leading to slightly better results for RAN compared to EXP_CNOP.

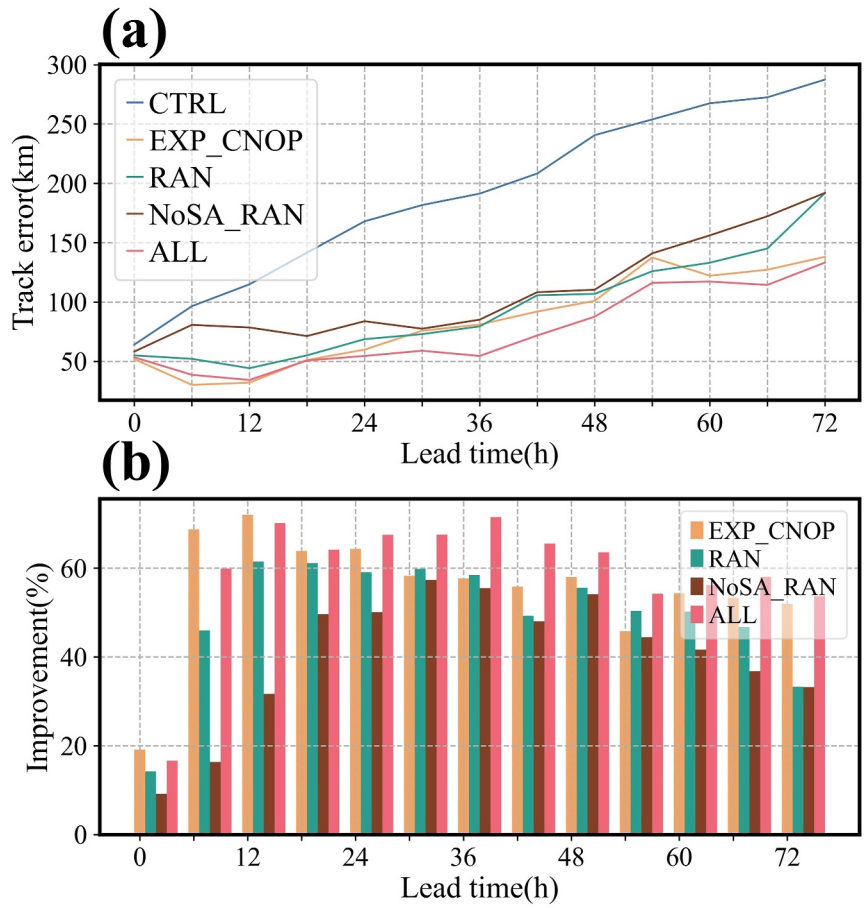


**Figure 6.** Figure (a) shows the average track error across the 11 cases with different colors corresponding to different experiments, consistent with Figure 5. Figure (b) displays the percent of improvement in the track error for EXP_CNOP (yellow), RAN (green), NoSA_RAN (brown), and ALL (red) compared to the CTRL. $Improvent = 100 * ((Track_{err})_{CTRL} - (Track_{err})_{xa})/(Track_{err})_{CTRL}$, where $(Track_{err})_{xa}$ is the track error of the experiment that assimilates observation.

### 4.3. Mechanism Analysis

The differences in the TC track forecasting results among EXP_CNOP, RAN, and NoSA_RAN can be attributed to the strategic placement of observations with an equal number of observation points in each case. The determination of both the observation locations and their quantity is guided by the CNOP-sensitive areas. From the previous section, EXP_CNOP almost consistently outperforms RAN and NoSA_RAN. To understand why observations placed in CNOP-sensitive areas are more effective than random placements, it is essential to analyze the rationale behind the CNOP-sensitive area's structure and its advantages.

To illustrate this, we will examine case V, where intensified observations in the sensitive area had a notable impact. This case will be analyzed to uncover why perturbations with a CNOP structure are prone to rapid development and how assimilating observations in these sensitive areas enhances track forecasts.

#### 4.3.1. The Structure of CNOP-Sensitive Areas

Figure 7 shows the distribution of total dry energy for CNOP. Horizontally, the high values of total dry energy are primarily concentrated in regions with large horizontal pressure gradients, specifically near the TC. This indicates that in the vicinity of the TC, perturbations are more likely to draw energy from the basic flow. Vertically, the high values are mainly concentrated in the mid-to-lower troposphere, which aligns with previous studies (Torn et al., 2018; H. Zhang et al., 2023).

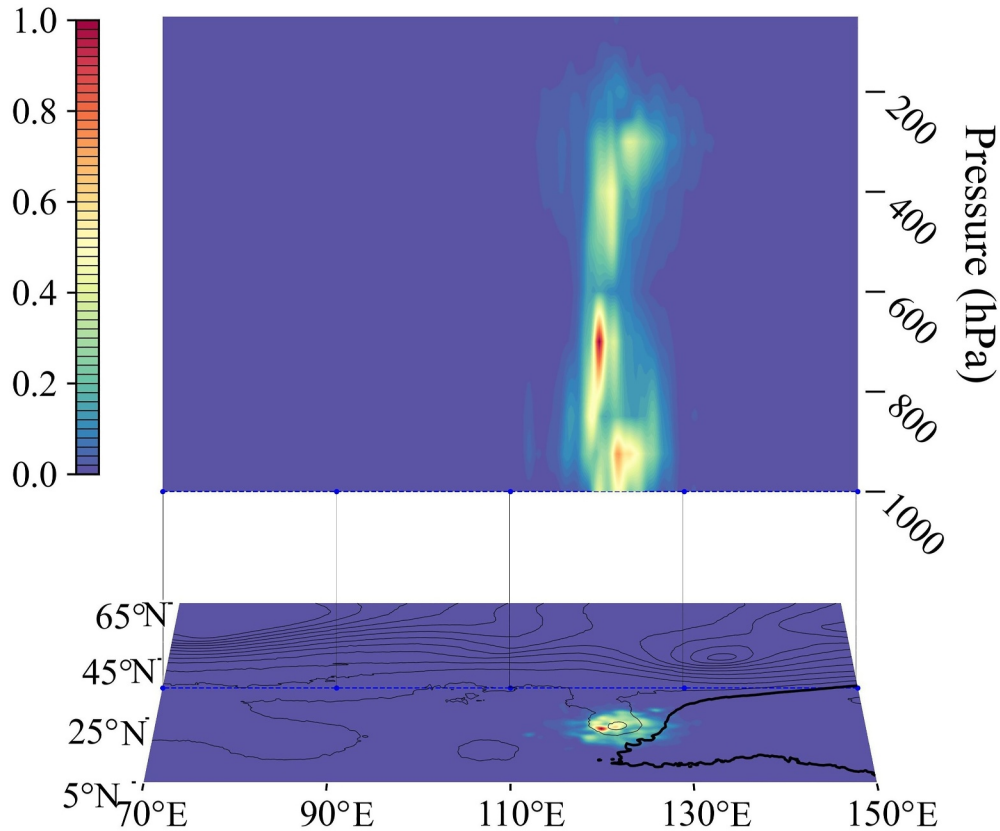


**Figure 7.** The dry total energy for conditional nonlinear optimal perturbation (CNOP) for case V. The upper figure represents the zonal integral of the dry total energy for CNOP, while the lower figure depicts the vertical integral of the dry total energy, as given by Equation 5. And the black line in the lower figure is the geopotential height of 500 hPa, and the thick black line is the contours of the geopotential height of 5,880 gpm. All energy values have been normalized.

Different perturbation variables exhibit distinct structures, leading to varying patterns of development. To further analyze this, the dry total energy is decomposed into kinetic and full potential energy, as illustrated in Figure 8. For kinetic energy (Figure 8a), the concentration is mainly in the mid-to-upper troposphere near the TC, aligning with the warm-core height (Figure 8b). The strong horizontal temperature gradients associated with the warm-core height enhance atmospheric baroclinicity. The background field releases available potential energy via baroclinic conversion, which is transformed into kinetic energy. Specifically, the perturbation kinetic energy continuously extracts energy from the baroclinic processes, enabling the rapid growth of the disturbance. From another perspective, the altitude of maximum kinetic energy is also closely associated with the upper-level divergence of the TC, with the location of the divergence center providing a key indication of the TC's position. The upper-level divergence center of TC is situated at the edge of the subtropical high, and its position can indirectly influence the TC's movement trajectory by adjusting the direction of the surrounding steering flow (Webster, 2020). Specifically, perturbation structures such as CNOP can affect the TC track forecast by altering the upper-level wind field. This underscores the importance of accurately capturing these upper-level wind errors to enhance the reliability of TC track forecasts.

Regarding full potential energy (Figure 8c), high values are mainly concentrated in the mid-to-lower troposphere near the TC center. In fact, within the boundary layer of a TC, there are pronounced inflow and Ekman pumping phenomena, which accelerate the vertical transport of moist air in the TC's core region (T.-H. Li & Wang, 2021). Moreover, there is a strong coupling between the vertical redistribution of absolute angular momentum near the TC center and the release of latent heat. Deep convection, particularly in the eyewall region, lifts high-angular-momentum and high-entropy air from the lower levels to the upper troposphere (Peng et al., 2019). During these processes, a significant amount of latent heat is released, and then the total energy of CNOP gains energy and

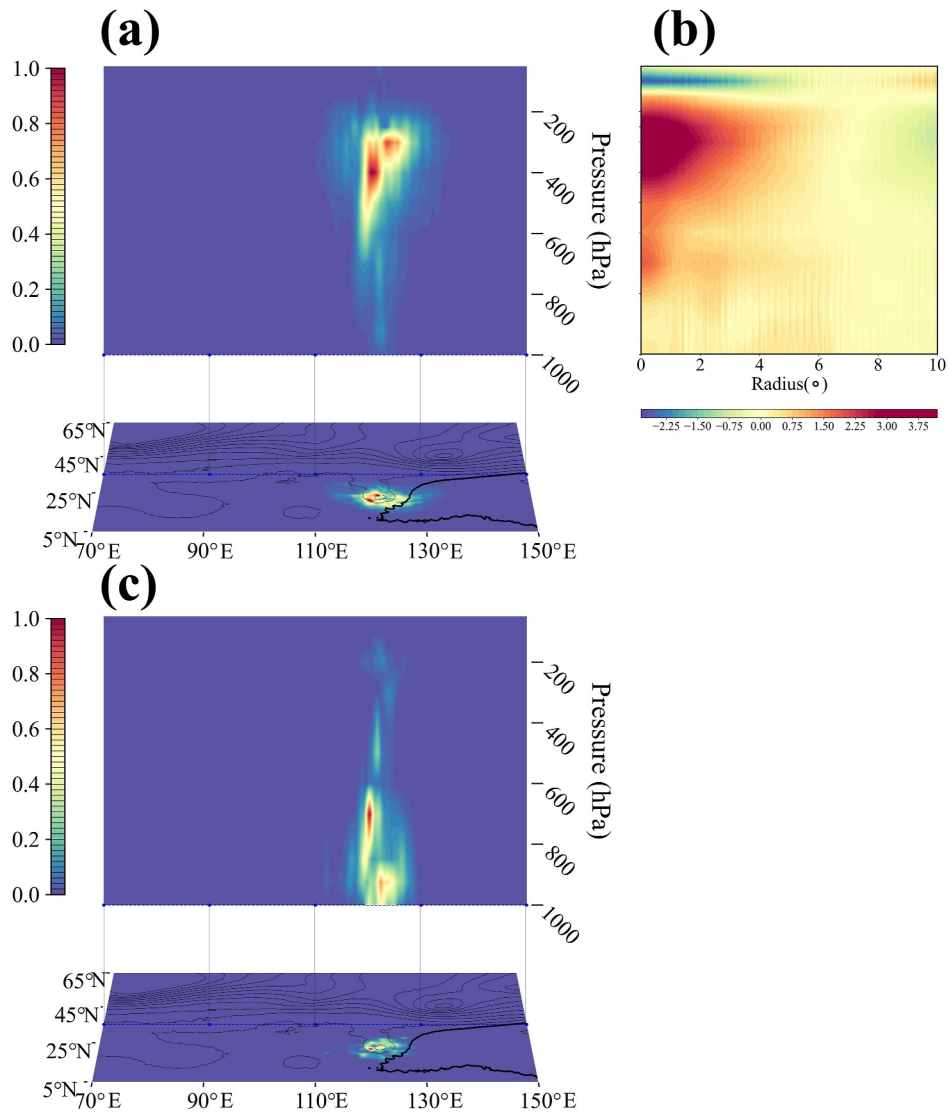


**Figure 8.** Figures (a) and (b) are similar to Figure 7 and also focus on case V, but with differences in the displayed quantities. Figure (a) shows the kinetic energy, while Figure (c) presents the full potential energy. Figure (b) illustrates the radius-height cross-sections of the azimuthally averaged temperature anomaly field (unit: °C), centered on the tropical cyclone.

grows rapidly. This energy growth process has a good correspondence with the position of TC changes, which is of great significance for understanding the variability of TC tracks.

The above analysis partially explains why perturbations with CNOP structures develop rapidly, highlighting the reason behind identifying such sensitive areas based on the FuXi model.

#### 4.3.2. Assimilation Analysis

Next, we analyze how assimilating different observations impacts the TC track. Effective assimilation should make the analysis closer to the truth or independent verifying observations relative to the background. We compared the fitting of the background before assimilation and the analysis after assimilation to the observation because there are no independent observations for verification. Due to the limitations of the assimilation system, which cannot assimilate unconventional variables, we opted to use gridded data as observations. Therefore, the above mentioned fitting method employed is a simple interpolation approach.

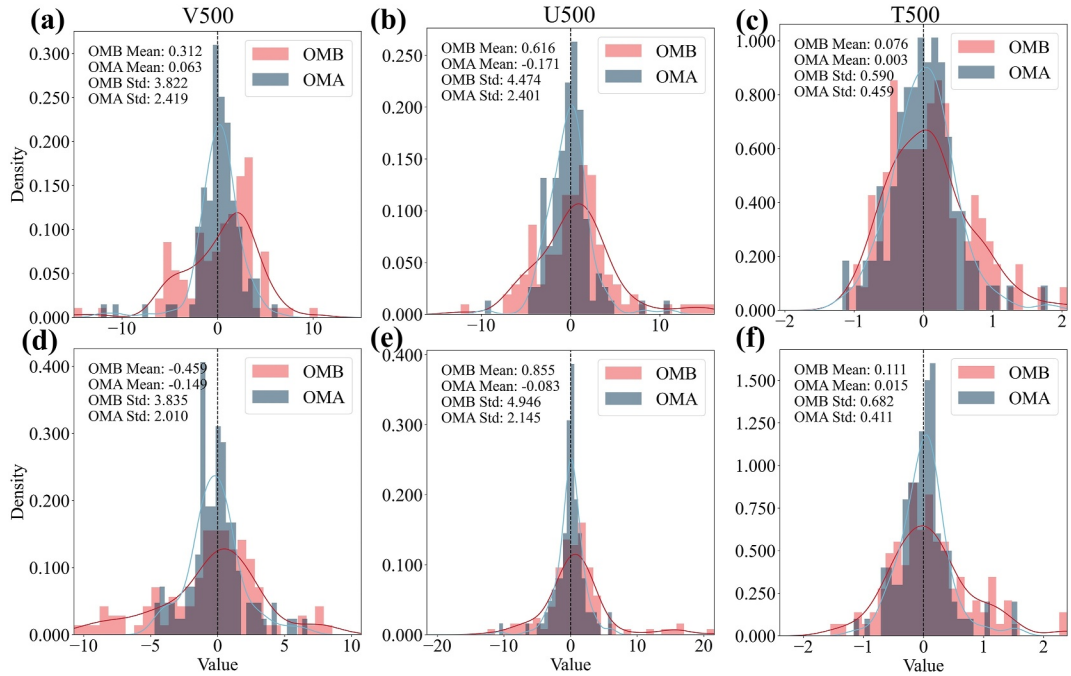


**Figure 9.** The probability density function of observations minus background (OMB) and observations minus analysis (OMA) are shown. These represent the OMB and OMA for the EXP_CNOP of case V. Specifically, Figures (a) and (d) show the radial wind at 500 hPa (V500, units: m/*s*), Figures (b) and (e) depict the zonal wind at 500 hPa (U500, unit: m/*s*), and Figures (c) and (f) illustrate the temperature at 500 hPa (T500, unit: °C). Figures (a)–(c) correspond to the initial time of the assimilation window, while Figures (d)–(f) represent the end of the assimilation window.

Figure 9 shows the probability density function (PDF) of observations minus background (OMB) and observations minus analysis (OMA) for v-wind (Figures 9a and 9d), u-wind (Figures 9b and 9e), and temperature variables at 500 hPa (Figures 9c and 9f) in case V for the EXP_CNOP. It is evident that both at the start (Figures 9a–9c) and the end (Figures 9d–9f) of the assimilation window, the mean and standard deviations of OMA have significantly decreased compared to OMB. The distribution of OMA is closer to a Gaussian distribution with a mean of zero and smaller uncertainty. Quantitatively, the absolute mean of OMA reduces compared to the absolute mean of OMB: for v-wind (V500), the average reduces to 73.67%, for u-wind (U500), the average reduces to 81.25%, and for temperature (T500), the average reduces to 91.26%. The standard deviation has slightly decreased with reductions of 42.15%, 51.48%, and 30.97% for V500, U500, and T500, respectively. Overall, the analysis field is more aligned with the observations after assimilation, demonstrating that the assimilation system enhances the accuracy of the initial conditions by incorporating observations.

Assimilated observations effectively convey observational information to the background field, resulting in analysis increments that improve the accuracy of the background field. Figure 10 illustrates the analysis increments for V500 and Z500 in the EXP_CNOP (Figures 10a and 10e), RAN (Figures 10b and 10f), NoSA_RAN (Figures 10c and 10g), and ALL (Figures 10d and 10h). For EXP_CNOP, the regions of large analysis increments are primarily concentrated near the TC. For RAN, large analysis increments are found not only near the TC but also in the surrounding trough and ridge systems. In the NoSA_RAN, the large analysis increments are not concentrated around the TC, while in the ALL, they are found both near the TC and in regions such as the subtropical high and westerly troughs. It is important to note that the Z500 analysis increments are influenced by changes in other variables, as observations do not include geopotential. Analysis increments of EXP_CNOP for V500 and Z500 near the TC are more similar to those of the ALL compared to RAN and NoSA_RAN. Furthermore, the Z500 analysis increments from EXP_CNOP (Figure 10e) and ALL (Figure 10h) result in a slower northward movement of the TC, more akin to NR, whereas RAN and NoSA_RAN do not affect the TC's northward movement. This leads to better forecast results for EXP_CNOP and ALL compared to RAN and NoSA_RAN.

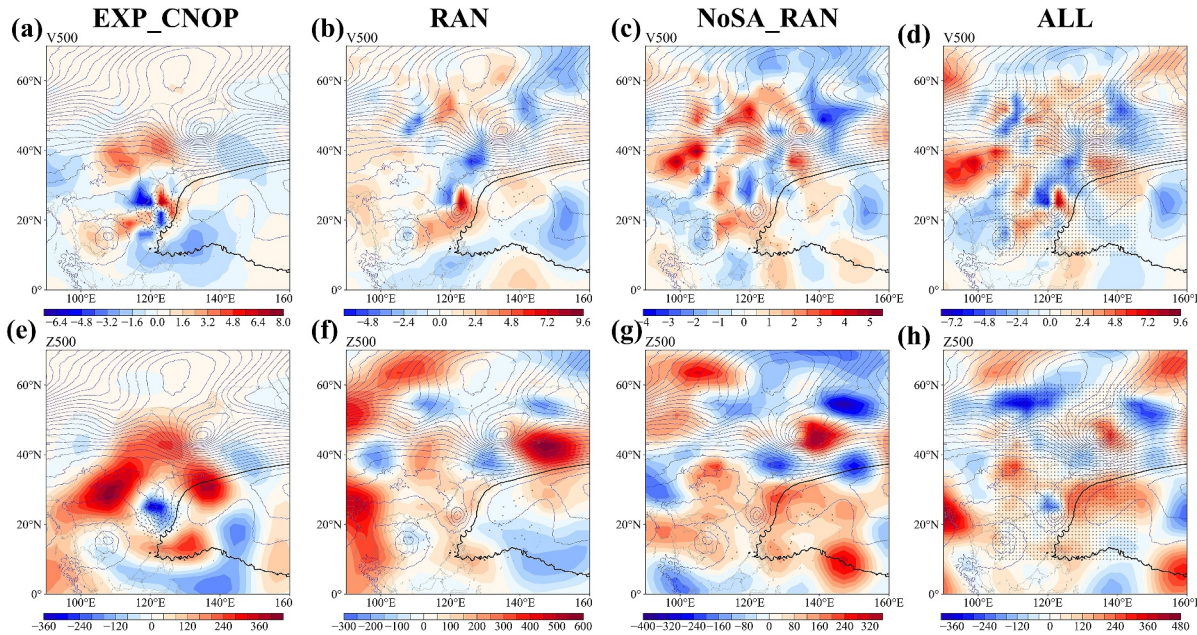


**Figure 10.** Analysis increments of case V. Figures (a) and (e) represent the analysis increments for EXP_CNOP, Figures (b) and (f) show the analysis increments for RAN, Figures (c) and (g) display the analysis increments for NoSA_RAN, and Figures (d) and (h) display the analysis increments for ALL. Figures (a)–(d) depict V500 (unit: m/$s$), while Figures (e)–(h) illustrate the geopotential at 500 hPa (Z500, unit: $m^2/s^2$). The black dots represent the observation points.

### 4.3.3. Steering Flow

It is well known that the motion of TC is mainly controlled by steering flow in the middle and lower atmosphere. The steering flow is generally defined as the average wind vector in the TC environment within a certain radius in the mid-troposphere. The steering flow is calculated by averaging the wind vectors with a TC as the center and a radius of 1,000 km (T.-C. Wu et al., 2014). To better understand how does assimilating observations in sensitive areas affect TC track forecasts, it is essential to study how the momentum field responds to the increment in the initial field.

Figure 11 displays the steering flow at various levels under different lead times of CTRL (blue arrows) and experiment assimilating observations (orange arrows) for case V along with their difference in steering flow (red arrows). At the initial time (the left side of each plot in Figure 11), EXP_CNOP (Figure 11a), RAN (Figure 11b), NoSA_RAN (Figure 11c), and ALL (Figure 11d) show a westerly steering flow in the mid-to-lower troposphere (500–700 hPa) compared to CTRL, indicating that all experiments correct CTRL toward NR (Figure 4e). However, EXP_CNOP and ALL show a southerly steering flow compared to the CTRL, which helps to slow down the northward movement of the TC and aligns more closely with NR. In contrast, RAN and NoSA_RAN exhibit a northward steering flow compared to CTRL, leading to a northward bias in the subsequent TC track forecasts (the right side of each plot in Figure 11), resulting in larger track forecasting errors.

## 5. Conclusion and Discussion

In recent years, AI-based weather forecasting models have rapidly developed and demonstrated significant effectiveness. There have been successful implementations of AI-based assimilation systems, showing that AI models have the potential to operate independently from traditional assimilation systems and perform cyclical forecasting on their own. However, due to the limitations in the spatial and temporal resolution of current observations, targeted observation remains essential for further improving the accuracy of initial fields, especially for high-impact weather events. In this study, we developed a CNOP-sensitive area identification system based on the FuXi model and applied it to 11 TC cases. This approach allows the FuXi model to enhance the accuracy of initial fields as much as possible, even with limited observations, thereby improving forecasting skills. It is important to note that in the process of calculating CNOP, the gradient is calculated through automatic differentiation rather than relying on tangent linear and adjoint models, which increases the computational efficiency.

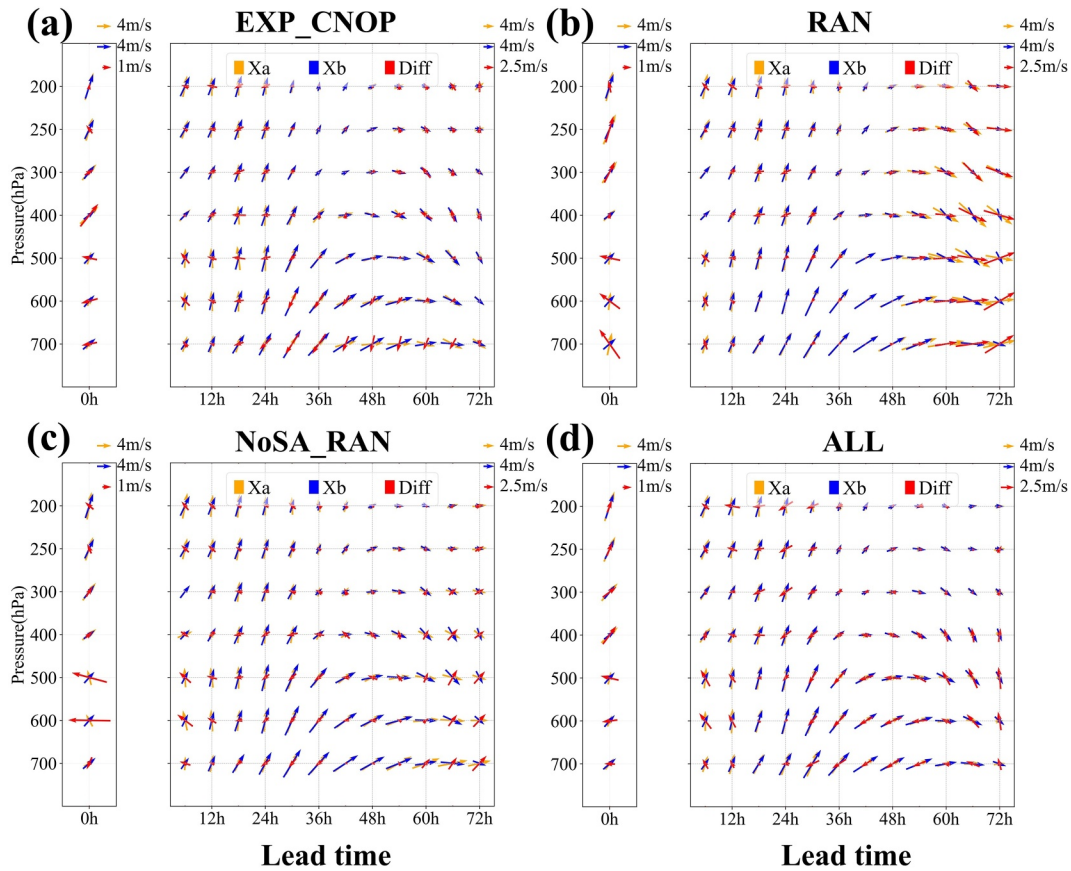


**Figure 11.** Steering flow of case V. Figures (a)–(d) correspond to EXP_CNOP, RAN, NoSA_RAN, and ALL, respectively. In these figures, orange arrows represent the results from the analysis field, blue arrows show the results from the background field, and red arrows indicate the difference between the analysis field and the background field.

The development of this system marks a significant step forward, connecting from the AI-based FuXi model and its DA system to greatly improve TC forecast skills through the AI-based sensitive area identification system (Phase 4 in Figure 1).

The main results indicate that enhancing and assimilating observations in sensitive areas (EXP_CNOP) significantly improves TC track forecasting skill compared to the CTRL. To further evaluate the effectiveness of observations in these sensitive areas, three sensitivity experiments were conducted (RAN, NoSA_RAN, and ALL). In these experiments, the number of observations for EXP_CNOP, RAN, and NoSA_RAN accounted for only 5.91% of those used in the ALL with observation points of three experiments located within the ALL observation range. The averaged results across cases showed that EXP_CNOP had the highest forecasting skill during the first 24 hr, with an average improvement of 55.92% compared to other experiments. This time corresponds to an optimization time of 1 day. In the subsequent forecast periods, the performance of EXP_CNOP remained comparable to that of ALL, while RAN showed moderate skill, and NoSA_RAN performed the worst among the four experiments. During the optimization period, EXP_CNOP is nearly optimal, and outside the optimization period, it still maintains a high level of forecasting skill. This demonstrated two key findings. First, strategically deploying a small number of observations in sensitive areas can achieve forecasting skill comparable to using a wide range of observations. This approach optimizes resource utilization while maintaining high forecasting accuracy. Second, perturbations structured according to the CNOP method, as calculated by the AI-based model, tend to develop rapidly, and significantly influence future forecasts. Effectively reducing errors associated with these CNOP-structured perturbations yields higher forecasting skill compared to mitigating errors in randomly selected or nonsensitive areas.

A more detailed analysis confirms the validity of the CNOP calculated by the FuXi model. This perturbation effectively captures unstable structures in the atmosphere and rapidly develops. When observations within the sensitive area are assimilated, the resulting analysis increment can reasonably adjust the steering flow, producing changes similar to those seen in the ALL experiment. The initial alteration of the steering flow in EXP_CNOP leads to subsequent forecasts that more closely align with NR.

This study demonstrates the effectiveness of AI-based CNOP-sensitive areas, showing that deploying observations in these regions is more beneficial than placing them elsewhere, thus maximizing the use of limited human, material, and financial resources. The positive outcomes from these experiments directly validate the feasibility of identifying sensitive areas and conducting assimilation forecasts entirely through AI, offering a new paradigm for the future of NWP, and show the potential of applying dynamical CNOP to AI-based models for highly effectively identifying sensitive areas of TC. The full-chain AI implementation also presents a promising approach to improving computational efficiency and reducing resource consumption.

While the results affirm the viability of an AI-based full-chain system, challenges remain, such as the current limitation to assimilating only conventional observations. The development of AI neural radiative transfer operators (Z. Li et al., 2024; Y. Li et al., 2025) holds potential to enable satellite DA within this system. Furthermore, it provides a means for targeted observations using satellites in subsequent applications (Y. Li, Han, & Duan, 2024). Future work will focus on refining the AI-based assimilation system and exploring whether the AI-driven targeted observation system can effectively improve TC intensity forecasts, which will involve parameter selection for CNOP-sensitive areas. Moreover, the selected case studies in this work primarily focus on the TCs that impact the China region. In fact, given the universal applicability of the CNOP method and the global coverage of the FuXi model, similar targeted observation experiments can be extended to other regions (e.g., the Bay of Bengal and Atlantic) for relevant TC forecasts. It is expected that future work involves forecasts of TCs in different ocean basins, thereby enhancing the robustness of our findings. Through these efforts, AI models may eventually operate independently of traditional numerical prediction systems, enabling the effective assimilation of multisource data and real-time targeted observations of high-impact weather events.

## Data Availability Statement

The ERA5 data (Hersbach et al., 2020) utilized in this paper are sourced from the official website of Copernicus Climate Data, and the ERA5 data are publicly accessible.

**Acknowledgments**
We thank the three anonymous reviewers for their insightful comments and suggestions. This study was jointly supported by the National Natural Science Foundation of China (Grant 41930971) and the International Partnership Program of the Chinese Academy of Sciences (Grant 060GJHZ2022061MI).

## References

Ancell, B. C., & Mass, C. F. (2006). Structure, growth rates, and tangent linear accuracy of adjoint sensitivities with respect to horizontal and vertical resolution. *Monthly Weather Review*, *134*(10), 2971–2988. https://doi.org/10.1175/mwr3227.1

Baker, N. L., & Daley, R. (2000). Observation and background adjoint sensitivity in the adaptive observation-targeting problem. *Quarterly Journal of the Royal Meteorological Society*, *126*(565), 1431–1454. https://doi.org/10.1002/qj.49712656511

Bergot, T. (1999). Adaptive observations during FASTEX: A systematic survey of upstream flights. *Quarterly Journal of the Royal Meteorological Society*, *125*(561), 3271–3298. https://doi.org/10.1256/smsqj.56106

Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X., & Tian, Q. (2023). Accurate medium-range global weather forecasting with 3D neural networks. *Nature*, *619*(7970), 533–538. https://doi.org/10.1038/s41586-023-06185-3

Bishop, C. H., Etherton, B. J., & Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Monthly Weather Review*, *129*(3), 420–436. https://doi.org/10.1175/1520-0493(2001)129<0420:aswtet>2.0.co;2

Brotzge, J. A., Berchoff, D., Carlis, D. L., Carr, F. H., Carr, R. H., Gerth, J. J., et al. (2023). Challenges and opportunities in numerical weather prediction. *Bulletin of the American Meteorological Society*, *104*(3), E698–E705. https://doi.org/10.1175/bams-d-22-0172.1

Buizza, R., Cardinali, C., Kelly, G., & Thepaut, J. (2007). The value of targeted observations—Part II: The value of observations taken in singular vectors based target areas. *ECMWF research department technical memorandum n. 512, ECMWF, Shinfield Park*.

Chen, G., Wu, C.-C., & Huang, Y.-H. (2018). The role of near-core convective and stratiform heating/cooling in tropical cyclone structure and intensity. *Journal of the Atmospheric Sciences*, *75*(1), 297–326. https://doi.org/10.1175/jas-d-17-0122.1

Chen, K., Han, T., Gong, J., Bai, L., Ling, F., Luo, J.-J., et al. (2023). Fengwu: Pushing the skillful global medium-range weather forecast beyond 10 days lead. *arXiv preprint arXiv:2304.02948*.

Chen, L., Zhong, X., Zhang, F., Cheng, Y., Xu, Y., Qi, Y., & Li, H. (2023). Fuxi: A cascade machine learning forecasting system for 15-day global weather forecast. *npj Climate and Atmospheric Science*, *6*(1), 190. https://doi.org/10.1038/s41612-023-00512-1

Der Grijn, V. (2002). Tropical cyclone forecasting at ECMWF: New products and validation. *ECMWF Tech. Memo.*, *386*, 1.

Duan, W., & Huo, Z. (2016). An approach to generating mutually independent initial perturbations for ensemble forecasts: Orthogonal conditional nonlinear optimal perturbations. *Journal of the Atmospheric Sciences*, *73*(3), 997–1014. https://doi.org/10.1175/jas-d-15-0138.1

Duan, W., Li, X., & Tian, B. (2018). Towards optimal observational array for dealing with challenges of El Niño-Southern oscillation predictions due to diversities of El Niño. *Climate Dynamics*, *51*(9–10), 3351–3368. https://doi.org/10.1007/s00382-018-4082-x

Duan, W., & Mu, M. (2009). Conditional nonlinear optimal perturbation: Applications to stability, sensitivity, and predictability. *Science in China, Series A: Earth Sciences*, *52*(7), 883–906. https://doi.org/10.1007/s11430-009-0090-3

Duan, W., Mu, M., & Wang, B. (2004). Conditional nonlinear optimal perturbations as the optimal precursors for El Nino–southern oscillation events. *Journal of Geophysical Research*, *109*(D23), D23105. https://doi.org/10.1029/2004jd004756

Ek, M. B. (2018). Land surface hydrological models. In Q. Duan, F. Pappenberger, J. Thielen, A. Wood, H. L. Cloke, & J. C. Schaake (Eds.), *Handbook of hydrometeorological ensemble forecasting* (pp. 1–42). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-40457-3_24-1

Emanuel, K., Kalnay, E., Bishop, C., Elsberry, R., Gelaro, R., Keyser, D., et al. (1997). Observations in aid of weather prediction for North America: Report of prospectus development team seven. *Bulletin of the American Meteorological Society*, *78*(12), 2859–2868. https://doi.org/10.1175/1520-0477-78.12.2859

Feng, J., Qin, X., Wu, C., Zhang, P., Yang, L., Shen, X., et al. (2022). Improving typhoon predictions by assimilating the retrieval of atmospheric temperature profiles from the FengYun-4A's geostationary interferometric infrared sounder (GIIRS). *Atmospheric Research*, *280*, 106391. https://doi.org/10.1016/j.atmosres.2022.106391

Hamill, T. M., & Snyder, C. (2002). Using improved background-error covariances from an ensemble Kalman filter for adaptive observations. *Monthly Weather Review*, *130*(6), 1552–1572. https://doi.org/10.1175/1520-0493(2002)130<1552:uibecf>2.0.co;2

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis [Dataset]. *Wiley Online Library*, *146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

Joly, A., Jorgensen, D., Shapiro, M. A., Thorpe, A., Bessemoulin, P., Browning, K. A., et al. (1997). The fronts and Atlantic storm-track experiment (FASTEX): Scientific objectives and experimental design. *Bulletin of the American Meteorological Society*, *78*(9), 1917–1940. https://doi.org/10.1175/1520-0477(1997)078<1917:tfaast>2.0.co;2

Kim, H. M., Morgan, M. C., & Morss, R. E. (2004). Evolution of analysis error and adjoint-based sensitivities: Implications for adaptive observations. *Journal of the Atmospheric Sciences*, *61*(7), 795–812. https://doi.org/10.1175/1520-0469(2004)061<0795:eoaeaa>2.0.co;2

Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., et al. (2023). Learning skillful medium-range global weather forecasting. *Science*, *382*(6677), 1416–1421. https://doi.org/10.1126/science.adi2336

Lang, S., Alexe, M., Chantry, M., Dramsch, J., Pinault, F., Raoult, B., et al. (2024). AIFS-ECMWF'S data-driven forecasting system. *arXiv preprint arXiv:2406.01465*.

Langland, R., Toth, Z., Gelaro, R., Szunyogh, I., Shapiro, M., Majumdar, S., et al. (1999). The north pacific experiment (NORPEX-98): Targeted observations for improved North American weather forecasts. *Bulletin of the American Meteorological Society*, *80*(7), 1363–1384. https://doi.org/10.1175/1520-0477(1999)080<1363:tnpent>2.0.co;2

Li, T.-H., & Wang, Y. (2021). The role of boundary layer dynamics in tropical cyclone intensification. Part II: Sensitivity to initial vortex structure. *Journal of the Meteorological Society of Japan. Ser. II*, *99*(2), 555–573. https://doi.org/10.2151/jmsj.2021-028

Li, Y., Han, W., & Duan, W. (2024). Dynamic channel selection based on vertical sensitivities for the assimilation of FY-4A geostationary interferometric infrared sounder targeted observations. *Quarterly Journal of the Royal Meteorological Society*, *150*(763), 3305–3321. https://doi.org/10.1002/qj.4760

Li, Y., Han, W., Duan, W., Li, Z., & Li, H. (2025). A machine learning-based observation operator for FY-4B GIIRS brightness temperatures considering the uncertainty of label data. *Journal of Geophysical Research: Machine Learning and Computation*, *2*(1), e2024JH000449. https://doi.org/10.1029/2024jh000449

Li, Y., Han, W., Li, H., Duan, W., Chen, L., Zhong, X., et al. (2024). Fuxi-en4dvar: An assimilation system based on machine learning weather forecasting model ensuring physical constraints. *Geophysical Research Letters*, *51*(22), e2024GL111136. https://doi.org/10.1029/2024gl111136

Li, Z., Han, W., Xu, X., Sun, X., & Li, H. (2024). All-sky microwave radiance observation operator based on deep learning with physical constraints. *Journal of Geophysical Research: Atmospheres*, *129*(23), e2024JD042436. https://doi.org/10.1029/2024jd042436

Lorenz, E. (1975). *Climate predictability: The physical basis of climate modeling* (Vol. 16). WMO, GARP Publications Series.

Mu, M., Duan, W., & Wang, B. (2003). Conditional nonlinear optimal perturbation and its applications. *Nonlinear Processes in Geophysics*, *10*(6), 493–501. https://doi.org/10.5194/npg-10-493-2003

Mu, M., Duan, W., & Wang, B. (2007). Season-dependent dynamics of nonlinear optimal error growth and El Niño-southern oscillation predictability in a theoretical model. *Journal of Geophysical Research*, *112*(D10), D10113. https://doi.org/10.1029/2005jd006981

Mu, M., Zhou, F., & Wang, H. (2009). A method for identifying the sensitive areas in targeted observations for tropical cyclone prediction: Conditional nonlinear optimal perturbation. *Monthly Weather Review*, *137*(5), 1623–1639. https://doi.org/10.1175/2008mwr2640.1

Palmer, T., Gelaro, R., Barkmeijer, J., & Buizza, R. (1998). Singular vectors, metrics, and adaptive observations. *Journal of the Atmospheric Sciences*, *55*(4), 633–653. https://doi.org/10.1175/1520-0469(1998)055<0633:svmaao>2.0.co;2

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., et al. (2017). Automatic differentiation in pytorch.

Pathak, J., Subramanian, S., Harrington, P., Raja, S., Chattopadhyay, A., Mardani, M., et al. (2022). Fourcastnet: A global data-driven high-resolution weather model using adaptive Fourier neural operators. *arXiv preprint arXiv:2202.11214*.

Peng, K., Rotunno, R., Bryan, G. H., & Fang, J. (2019). Evolution of an axisymmetric tropical cyclone before reaching slantwise moist neutrality. *Journal of the Atmospheric Sciences*, *76*(7), 1865–1884. https://doi.org/10.1175/jas-d-18-0264.1

Qin, X., Duan, W., Chan, P.-W., Chen, B., & Huang, K.-N. (2023). Effects of dropsonde data in field campaigns on forecasts of tropical cyclones over the western north pacific in 2020 and the role of CNOPsensitivity. *Advances in Atmospheric Sciences*, *40*(5), 791–803. https://doi.org/10.1007/s00376-022-2136-9

Qin, X., Duan, W., & Xu, H. (2020). Sensitivity to tendency perturbations of tropical cyclone short-range intensity forecasts generated by WRF. *Advances in Atmospheric Sciences*, *37*(3), 291–306. https://doi.org/10.1007/s00376-019-9187-6

Rabier, F., Klinker, E., Courtier, P., & Hollingsworth, A. (1996). Sensitivity of forecast errors to initial conditions. *Quarterly Journal of the Royal Meteorological Society*, *122*(529), 121–150. https://doi.org/10.1256/smsqj.52905

Snyder, C. (1996). Summary of an informal workshop on adaptive observations and FASTEX. *Bulletin of the American Meteorological Society*, *77*(5), 953–961. https://doi.org/10.1175/1520-0477-77.5.953

Szunyogh, I., Toth, Z., Zimin, A. V., Majumdar, S. J., & Persson, A. (2002). Propagation of the effect of targeted observations: The 2000 winter storm reconnaissance program. *Monthly Weather Review*, *130*(5), 1144–1165. https://doi.org/10.1175/1520-0493(2002)130<1144:poteot>2.0.co;2

Torn, R. D., Elless, T. J., Papin, P. P., & Davis, C. A. (2018). Tropical cyclone track sensitivity in deformation steering flow. *Monthly Weather Review*, *146*(10), 3183–3201. https://doi.org/10.1175/mwr-d-18-0153.1

Torn, R. D., & Hakim, G. J. (2008). Ensemble-based sensitivity analysis. *Monthly Weather Review*, *136*(2), 663–677. https://doi.org/10.1175/2007mwr2132.1

Toth, Z., & Kalnay, E. (1997). Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*, *125*(12), 3297–3319. https://doi.org/10.1175/1520-0493(1997)125<3297:efanat>2.0.co;2

Wang, H., Han, W., Li, J., Chen, H., & Yin, R. (2025). Impact of assimilation of FY-4A GIIRS three-dimensional horizontal wind observations on typhoon forecasts. *Advances in Atmospheric Sciences*, *42*(3), 467–485. https://doi.org/10.1007/s00376-024-4051-8

Webster, P. J. (2020). *Dynamics of the tropical atmosphere and oceans*. John Wiley & Sons.

Wu, C.-C., Chen, J.-H., Lin, P.-H., & Chou, K.-H. (2007). Targeted observations of tropical cyclone movement based on the adjoint-derived sensitivity steering vector. *Journal of the Atmospheric Sciences*, *64*(7), 2611–2626. https://doi.org/10.1175/jas3974.1

Wu, C.-C., Lin, P.-H., Aberson, S., Yeh, T.-C., Huang, W.-P., Chou, K.-H., et al. (2005). Dropwindsonde observations for typhoon surveillance near the Taiwan region (DOTSTAR) an overview. *Bulletin of the American Meteorological Society*, *86*(6), 787–790.

Wu, T.-C., Liu, H., Majumdar, S. J., Velden, C. S., & Anderson, J. L. (2014). Influence of assimilating satellite-derived atmospheric motion vector observations on numerical analyses and forecasts of tropical cyclone track and intensity. *Monthly Weather Review*, *142*(1), 49–71. https://doi.org/10.1175/mwr-d-13-00023.1

Xiao, Y., Bai, L., Xue, W., Chen, K., Han, T., & Ouyang, W. (2023). Fengwu-4dvar: Coupling the data-driven weather forecasting model with 4D variational assimilation. *arXiv preprint arXiv:2312.12455*.

Xu, X., Sun, X., Han, W., Zhong, X., Chen, L., & Li, H. (2024). Fuxi-da: A generalized deep learning data assimilation framework for assimilating satellite observations. *arXiv preprint arXiv:2404.08522*.

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 and Forecasting*, *38*(10), 1907–1933. https://doi.org/10.1175/waf-d-22-0175.1

Zhang, Y., Long, M., Chen, K., Xing, L., Jin, R., Jordan, M. I., & Wang, J. (2023). Skilful nowcasting of extreme precipitation with Nowcastnet. *Nature*, *619*(7970), 526–532. https://doi.org/10.1038/s41586-023-06184-4

Zhou, F., & Mu, M. (2011). The impact of verification area design on tropical cyclone targeted observations based on the CNOP method. *Advances in Atmospheric Sciences*, *28*(5), 997–1010. https://doi.org/10.1007/s00376-011-0120-x

Zhou, F., & Mu, M. (2012). The time and regime dependencies of sensitive areas for tropical cyclone prediction using the CNOP method. *Advances in Atmospheric Sciences*, *29*(4), 705–716. https://doi.org/10.1007/s00376-012-1174-0