*atmosphere*

MDPI

*Article*

# Optimal Initial Error and Targeted Observation Sensitive Area for Predicting the Northeast China Cold Vortex Revealed by a Deep Learning Model

**Chen Zhang [1,2], Junkai Qian [1,2] and Qiang Wang [1,2,*]**

[1] Key Laboratory of Marine Hazards Forecasting, Ministry of Natural Resources, Hohai University, Nanjing 210024, China; 231311010004@hhu.edu.cn (C.Z.); 240217010005@hhu.edu.cn (J.Q.)
[2] College of Oceanography, Hohai University, Nanjing 210024, China
* Correspondence: wangq@hhu.edu.cn

**Abstract**

The Northeast China Cold Vortex (NECV) is a key circulation system affecting weather patterns over North China, frequently triggering thunderstorms, hail, and other severe convective weather. Accurate prediction of NECVs is therefore of great importance. However, substantial forecast errors still remain, largely due to uncertainties in the initial conditions. To improve NECV forecast skills, we investigate the optimal initial errors and targeted observation sensitive areas using a sampling-based approximation of the conditional nonlinear optimal perturbation (CNOP) method together with the Pangu-Weather deep learning model. We first evaluate the model's performance over Northeast China and find that Pangu-Weather exhibits forecast skill generally comparable to the ECMWF Integrated Forecasting System (IFS) during the May–August 2022 period over Northeast China. Then the CNOP-based approach is used to capture the optimal initial errors with the greatest impact on NECV forecasts. The largest error amplitudes are primarily located upstream of the vortex and near upper-level jet-entrance regions, which are identified as the targeted observation sensitive areas. Perturbation kinetic-energy diagnostics further indicate that baroclinic conversion is the dominant mechanism for error growth. Observing system simulation experiments suggest that, under an idealized assumption of completely eliminating errors in a given region, targeted observations over the sensitive area can produce the largest forecast improvement, with an average error reduction of approximately 13% relative to other areas. This study contributes to a deeper understanding of NECV predictability and may help improve forecasting capability.

**Keywords:** northeast China cold vortex; deep learning; predictability; targeted observation

Check for updates

Academic Editor: Xiaorui Niu

Received: 24 March 2026
Revised: 26 May 2026
Accepted: 28 May 2026
Published: 30 May 2026

**Copyright:** © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

## 1. Introduction

The Northeast China Cold Vortex (NECV) is a cut-off low system that frequently occurs over East Asia [1,2]. During the boreal summer, prevailing northerly flow over the mid-to-high latitudes favors the formation and maintenance of NECVs, leading to a substantial increase in their frequency [3,4]. NECVs are characterized by cold lows in the mid-to-upper troposphere, marked by frequent generation, slow movement, and intense convective activity [5]. Associated severe weather—including heavy rainfall, thunderstorm winds, hail, and summer cold damage—seriously threatens regional socio-economic activities and agricultural production [5,6]. Enhancing NECV predictability is, therefore, essential for climate security and disaster mitigation in northern China.

NECV forecasting has traditionally relied on numerical models, which integrate atmospheric dynamical and thermodynamical equations based on specified initial and boundary conditions [7–9]. Major operational centers such as the European Centre for Medium-Range Weather Forecasts (ECMWF) [10] and the National Centers for Environmental Prediction (NCEP) have established high-resolution global models capable of capturing mid-latitude systems. However, forecast accuracy remains constrained by initial errors, which are often amplified by observational limitations, data assimilation uncertainties, and atmospheric nonlinearity [11–13]. Furthermore, this challenge is particularly acute for rapidly evolving, structurally complex mesoscale systems such as NECVs [14].

Rapid advances in deep learning have opened new avenues for weather and climate prediction [9,15–17]. By leveraging large-scale datasets, deep learning models can directly learn nonlinear atmospheric evolution from reanalysis data and observations, bypassing the need for explicit dynamical equations [18–20]. Notably, the Pangu-Weather model [21] has demonstrated forecast performance comparable to, or even surpassing, that of the ECMWF Integrated Forecasting System (IFS), while drastically reducing computational costs. However, similar to traditional numerical models, deep learning forecasts are also sensitive to initial-condition errors [20].

To mitigate the limitations imposed by initial-error growth, previous studies introduced the concept of targeted observations [22,23], in which additional observations were deployed in the most sensitive regions to maximize forecast accuracy at the target time. Common sensitive-area identification methods include ensemble Kalman filter [12,24], singular vectors [25,26], and conditional nonlinear optimal perturbation (CNOP) [27–29]. The former two methods rely on linear approximations and therefore have limitations in representing nonlinear error growth. In contrast, the latter method is fully nonlinear and can effectively capture the optimal initial error, i.e., the initial perturbation that produces the largest forecast-error growth under a prescribed constraint, thereby efficiently identifying targeted observation sensitive areas.

Traditional targeted observation studies mainly rely on numerical weather prediction models, where repeated integrations for sensitivity analysis are computationally expensive. In contrast, deep-learning weather models such as Pangu-Weather provide substantially faster forecast inference. Compared with traditional numerical weather prediction (NWP) integrations that may require minutes to hours per forecast cycle, Pangu-Weather completes inference within seconds, enabling hundreds of perturbation experiments at feasible cost. This computational advantage makes large-sample sensitivity experiments feasible. Therefore, integrating the CNOP framework with a deep-learning forecasting model provides an efficient framework for predictability diagnosis and targeted observation design. In this study, we identify the optimal initial errors and targeted observation sensitive areas using the Pangu-Weather model and a CNOP-based approach, and then employ observing system simulation experiments (OSSEs) to evaluate the effectiveness of reducing errors in sensitive regions for improving forecast accuracy.

The remainder of this paper is organized as follows. Section 2 introduces the datasets and methods. Section 3 evaluates model forecast performance, identifies the fastest-growing (optimal) initial errors, examines the physical mechanisms of the error growth, and finally assesses targeted observation sensitive areas through OSSEs. Section 4 summarizes the main findings and discusses the methodological implications and limitations of this study.

## 2. Materials and Methods

### *2.1. Data*

This study primarily utilizes data from the ECMWF Fifth Generation Global Atmospheric Reanalysis product (ERA5; [30]). Operated by ECMWF on behalf of the Copernicus

Climate Change Service (C3S), ERA5 integrates multi-source historical observations (including satellite, radar, radiosonde, aircraft, and surface stations) into a consistent global climate estimate using four-dimensional variational data assimilation (4D-Var) and an advanced model system, with time coverage extending from 1940 to the present.

ERA5 provides an extensive suite of atmospheric, land-surface, and oceanic variables at an hourly resolution. The vertical structure is represented by 137 model levels extending from the surface to approximately 80 km, with standard pressure-level products covering 37 isobaric heights. To meet the input requirements of the Pangu-Weather model, we extract key meteorological variables for the period of May–August 2022. The three-dimensional atmospheric field is constructed using five upper-air variables—geopotential height, specific humidity, temperature, and U/V wind components—across 13 mandatory pressure levels. Additionally, four surface variables—mean sea-level pressure, 10 m U/V wind components, and 2 m temperature—are selected. All ERA5 data are processed on a uniform $0.25^\circ \times 0.25^\circ$ grid to align with the input resolution of Pangu-Weather model.

To evaluate the deep learning model's performance against traditional numerical methods, the forecast products from the ECMWF-IFS are employed as a baseline. To ensure a rigorous and fair comparison, the ECMWF-IFS data are selected for the same variables, time periods, and spatial resolution ($0.25^\circ \times 0.25^\circ$) as the ERA5-driven Pangu-Weather forecasts.

### *2.2. Pangu-Weather Deep Learning Model*

Pangu-Weather is a global deep-learning weather forecasting model developed by the Huawei Cloud team [21]. Trained on ERA5 reanalysis data, the model features a specialized framework tailored to the three-dimensional structure of the atmosphere. This is achieved through a 3D block strategy, local 3D self-attention, and relative position encoding adapted to the Earth's spherical geometry, enabling end-to-end modeling of complex atmospheric spatiotemporal evolution. Based on a single initial state, Pangu-Weather outputs key meteorological variables—including temperature, wind fields, geopotential height, and specific humidity—across multiple vertical levels covering the troposphere. The model provides a global uniform grid resolution of $0.25^\circ \times 0.25^\circ$, with a forecast lead time extending to approximately 10 days. Its medium-range performance has been shown to be comparable to, or even surpass, traditional numerical models across various evaluation metrics.

During inference, Pangu-Weather bypasses the step-by-step integration of dynamical and physical processes used in traditional NWP models. Instead, it predicts the atmospheric 3D state at subsequent intervals via a single forward pass through the deep learning network. This study utilizes the publicly available 24 h time-step model, where each forward calculation yields a forecast for the following 24 h. This strategy minimizes the number of iterative steps required for multi-day forecasts, significantly reducing cumulative errors and computational overheads.

### *2.3. Study Area and Case Selection*

The model evaluation domain in this study is defined as 110°–145° E and 35°–60° N, corresponding to the NECV activity region [31]. Within this broader domain, Liu et al. [32] identified 120°–130° E and 40°–50° N as the key region for NECV, where variations in 500-hPa geopotential height effectively characterize the system's life cycle. Accordingly, the regional mean geopotential height within this key region is used to define the NECV Activity Index (NECVI). Three typical NECV cases that occurred between May and August 2022 are selected for analysis (Cases 1–3). The mature stages of these cases are 7–9 May, 26–28 May, and 8–10 July 2022, respectively.

### *2.4. Identification of the Optimal Initial Error Using the CNOP Approach*

To capture the optimal initial error (the fastest-growing initial error), we employ the CNOP approach proposed by Mu et al. [27]. For simplicity, the CNOP-type optimal initial errors are obtained by the re-forecast experiments, following Feng and Duan [33]. This sampling-based strategy provides a practical approximation of the CNOP. It effectively identifies perturbations with the largest nonlinear growth under the prescribed constraints. Previous predictability studies have shown that this approach can capture robust sensitive-error structures [33,34]. The experimental domain spans 80°–150° E and 10°–80° N, covering the main circulation systems of the mid-to-high latitudes and the evolution of NECVs.

The specific experimental procedure is as follows. First, after determining the initial forecast time, two datasets within a three-month period and separated by 3–5 days are randomly selected. Differences in temperature, zonal wind, and meridional wind at the surface and across 13 upper-air pressure levels are then calculated. These differences are scaled to 0.3 times their standard deviation, serving as a physical constraint for constructing a random initial error field. The imposed perturbation amplitude (0.3 σ) represents a moderate perturbation amplitude. Comparison with the calculated perturbation magnitudes indicates that they are broadly comparable to typical observational uncertainty ranges for atmospheric temperature and wind. The resulting perturbation is then superimposed onto the original model input, and Pangu-Weather is used to perform 5-day forecasts. The above procedure is repeated 500 times.

To quantify the impact of initial errors on forecast results, the dry energy of the forecast error is calculated over the region where NECVs frequently occur. The forecast error is defined as the difference between the perturbed forecast and the control forecast (i.e., the forecast without superimposed perturbations). This metric characterizes the temporal evolution of dry-energy error growth at different forecast lead times. The dry-energy norm (E) of the error is defined as:

$$E = \frac{1}{D_1}\int_{D_1}\int_0^1 \left[u'^2 + v'^2 + \frac{c_p}{T_r}T'^2 + R_a T_r\left(\frac{p'_s}{p_r}\right)^2\right] d\sigma dD_1, \quad (1)$$

In this equation, the dry-energy error $E$ consists of kinetic-energy and effective potential-energy components. The zonal and meridional wind errors ($u'$ and $v'$) represent the kinetic-energy contribution associated with wind-field perturbations, while the temperature and surface-pressure error terms ($T'$ and $p'_s$) describe the effective potential-energy contribution associated with thermal and mass-field anomalies. Here, $R_a$ is the gas constant for dry air, with a value of 287 J $kg^{-1}K^{-1}$. $T_r$ is the reference temperature, and $c_p$ is the specific heat at constant pressure.

We first calculate the squared error terms of the key variables over region $D_1$ and perform area averaging. Next, mass integration is carried out in the vertical σ coordinates. This procedure yields a scalar value E, which represents the vertically integrated, mass-weighted dry-energy error per unit mass within the target region $D_1$. Physically, E corresponds to the sum of the kinetic energy and effective potential energy of the error. Because NECV evolution involves coupled thermodynamic and dynamical processes, the dry-energy norm is suitable for measuring forecast-error growth associated with NECV development.

Based on the above re-forecast experiments, the 20 perturbations with the most pronounced dry-energy growth at the 5-day lead time are selected. These 20 "rapid-growth initial-error" fields are regarded as the optimal initial error. Their spatial structures are used to identify the targeted observation sensitive regions for improving the prediction of the NECV.

### *2.5. Perturbation Kinetic-Energy Budget Analysis*

To reveal the energy sources and conversion mechanisms during the growth of initial errors (perturbations), a perturbation kinetic-energy budget equation is employed for diagnostic analysis. All variables are defined as differences between the perturbed experiment and the reference state.

#### 2.5.1. Definition of Perturbation Kinetic Energy

The perturbation kinetic energy is defined as the kinetic energy associated with wind-field errors:

$$K = \frac{1}{2}(u'^2+v'^2), \tag{2}$$

where $u'$ and $v'$ denote the zonal and meridional wind error components, respectively.

#### 2.5.2. Perturbation Kinetic-Energy Budget Equation

The temporal tendency of perturbation kinetic energy can be decomposed into barotropic conversion, baroclinic conversion, and advection terms:

$$\frac{\partial K}{\partial t} = C_{BT} + C_{BC} + C_{ADV}, \tag{3}$$

where $C_{BT}$ is the barotropic conversion term, $C_{BC}$ is the baroclinic conversion term, and $C_{ADV}$ is the advection term.

- Barotropic Conversion Term (BT)

Barotropic conversion represents the momentum exchange between the mean flow and perturbation kinetic energy:

$$C_{BT} = -u'^2\frac{\partial U}{\partial x} - u'v'\frac{\partial U}{\partial y} - u'v'\frac{\partial V}{\partial x} - v'^2\frac{\partial V}{\partial y}, \tag{4}$$

where U and V are the zonal and meridional components of the reference-state wind field.

- Baroclinic Conversion Term (BC)

Baroclinic conversion describes the conversion of available potential energy into kinetic energy:

$$C_{BC} = g\frac{w'T'}{T}, \tag{5}$$

where $w'$ is the perturbation vertical velocity, $T'$ is the perturbation temperature, and T is the reference-state temperature.

It should be noted that Pangu-Weather does not explicitly output some variables, such as vertical velocity. Therefore, the $C_{BC}$ term in the budget analysis is estimated approximately using the mass continuity equation, which may introduce some uncertainty into the quantitative results. Nevertheless, this method can still capture the primary energy-conversion pathways during perturbation growth and provides useful insight into the mechanisms of forecast-error amplification.

- Advection Term (ADV)

The advection term represents the transport of perturbation kinetic energy by the reference-state flow, including both horizontal and vertical advection:

$$C_{ADV} = -U\frac{\partial K}{\partial x} - V\frac{\partial K}{\partial y} - W\frac{\partial K}{\partial z}, \tag{6}$$

Because the model output does not provide sufficient diagnostic variables to accurately quantify subgrid-scale and frictional processes, the dissipation term is not explicitly calculated in this study.

## 3. Results

### *3.1. Evaluation of Pangu-Weather's Performance in Predicting NECV*

To evaluate the forecasting capabilities of Pangu-Weather in the Northeast China region, we compare its predictions and those of ECMWF-IFS, including the key atmospheric circulation variables during the period of May–August 2022. Figure 1 shows the root-mean-square error (RMSE) and anomaly correlation coefficient (ACC) calculated from forecasts by the Pangu and IFS models, alongside the observational data. The results imply that for 1–7-day forecasts of geopotential height, specific humidity, temperature, and wind field, Pangu-Weather generally yields comparable or slightly lower RMSE than ECMWF-IFS. Although this improvement is evident across all parameters, there are inter-variable variations: the reductions in RMSE for temperature and humidity are more substantial, whereas the improvement for geopotential height is relatively modest. Consistent with the RMSE results, the ACC metrics reveal a modest advantage for the deep learning model; Pangu-Weather maintains higher ACC scores than the IFS for nearly all variables throughout the 7-day forecast period. From Figure 1k, we can also see that the predicted NECVIs of the Pangu-Weather model and IFS model are very close to the observed index.

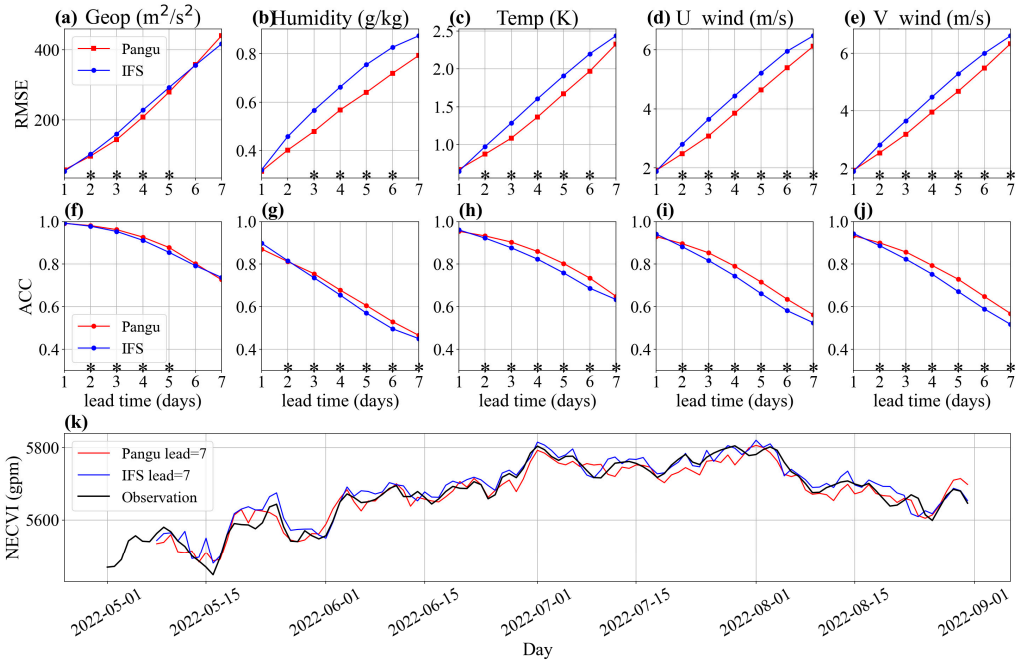


**Figure 1.** (**a**–**e**) Comparison of root-mean-square error (RMSE) for all upper-level variables (geopotential height, specific humidity, temperature, and U/V winds) in Pangu-Weather at different forecast lead times. (**f**–**j**) Corresponding anomaly correlation coefficient (ACC) comparison for the same variables, with blue representing the Integrated Forecasting System (IFS) and red representing Pangu-Weather. (**k**) Time series of the observed and forecasted northeast China Cold Vortex Index (NECVI) based on 7-day forecasts from the Pangu-Weather and IFS models. Asterisks indicate lead times at which the differences between the two models are statistically significant at the 95% confidence level.

To further evaluate the capability of Pangu-Weather in predicting the NECV, three representative cases are analyzed: Case 1 (7–9 May), Case 2 (26–28 May), and Case 3 (8–10 July). Figures 2–4 compare the observed 500-hPa temperature fields (shading) and geopotential height fields (black contours) with Pangu-Weather forecasts at various lead times.

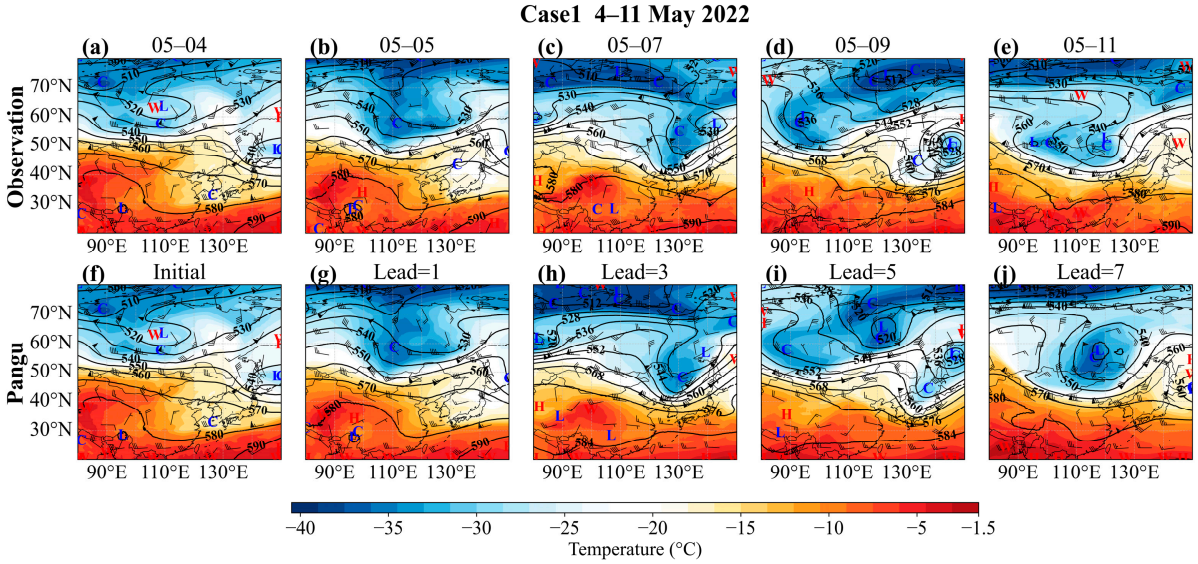


**Figure 2.** Comparison of 500 hPa geopotential height (black contours, unit: gpm) and temperature (shading, unit: °C) between ERA5 observations and Pangu-Weather forecasts at forecast lead times of 0, 1, 3, 5, and 7 days. Panels (**a**–**e**) show ERA5 observations on different dates, while panels (**f**–**j**) show the corresponding Pangu-Weather forecasts initialized on the starting date of the case. The contour labels indicate the corresponding 500 hPa geopotential–height values. Blue "L" and "C" denote the low-pressure center and cold center, respectively, while red "H" and "W" denote the high-pressure center and warm center, respectively. The panels illustrate the evolution of the NECV structure for Case 1.

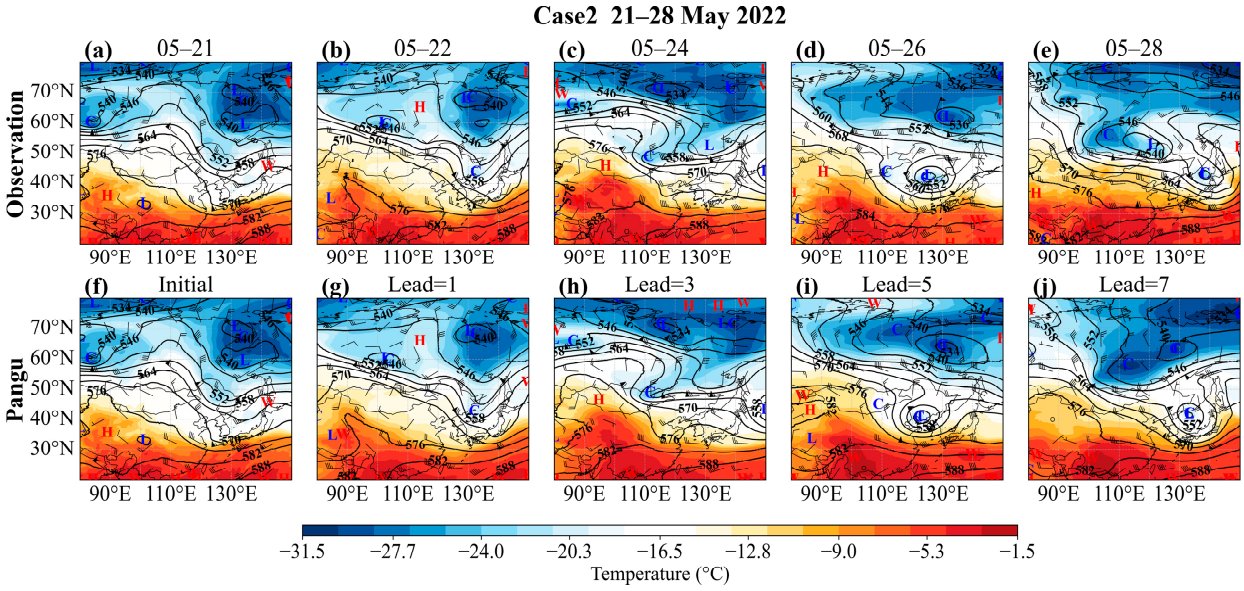


**Figure 3.** Comparison of 500 hPa geopotential height and temperature between ERA5 observations and Pangu-Weather forecasts for Case 2. Panels (**a**–**e**) show ERA5 observations on different dates, while panels (**f**–**j**) show the corresponding Pangu-Weather forecasts initialized on the starting date of the case. Other variables, symbols, and plotting conventions are the same as in Figure 2.

Pangu-Weather reproduces the main structural characteristics of the cold vortex, including its location, morphology, and intensity, and maintains reasonable forecast skill for both quasi-stationary and rapidly evolving NECV cases up to a 7-day lead time. These results support the use of Pangu-Weather in the subsequent sensitivity experiments.

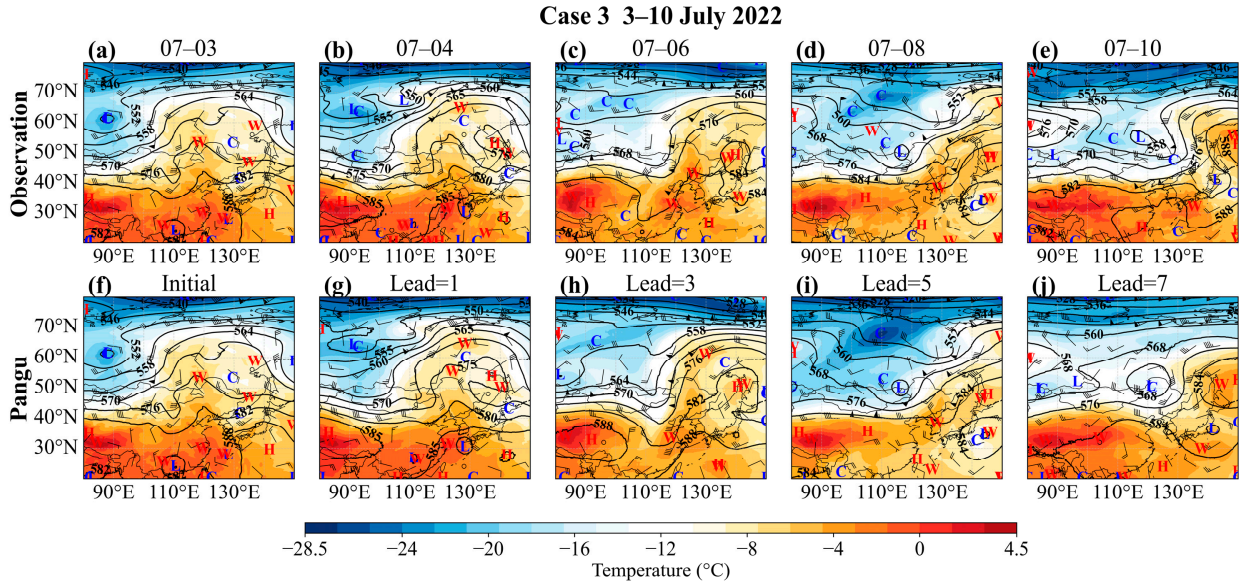


**Figure 4.** Comparison of 500 hPa geopotential height and temperature between ERA5 observations and Pangu-Weather forecasts for Case 3. Panels (**a**–**e**) show ERA5 observations on different dates, while panels (**f**–**j**) show the corresponding Pangu-Weather forecasts initialized on the starting date of the case. Other variables, symbols, and plotting conventions are the same as in Figure 2.

### *3.2. Optimal Initial Errors in the NECV Prediction Based on Pangu-Weather Model*

Figure 5 illustrates the energy evolutions of the initial errors for three representative NECV cases. By computing the dry-energy norm of forecast errors within the primary NECV activity region (indicated by the red box), the growth trajectories of the initial-error energy are obtained. Overall (Figure 5a–c), a subset of initial errors exhibits rapid amplification, reaching peak magnitudes around forecast Days 2–3, a clear manifestation of nonlinear error growth. In contrast, the remaining errors show relatively limited growth. We further isolate the 20 initial errors with the most pronounced energy growth at the 5-day lead time (Figure 5d–f). These optimal initial errors exhibit a consistent evolution process, characterized by an initial stage of rapid intensification followed by a gradual saturation phase as the forecast lead time increases.

For comparison, the mean evolutions of the top 100 and bottom 400 perturbation groups are also shown. The top 100 group displays a growth pattern broadly similar to that of the top 20 group, although with weaker amplitudes, whereas the bottom 400 group remains nearly unchanged throughout the forecast period. This contrast suggests that the rapidly growing perturbations may share similar error-growth characteristics, and that the selected top 20 subset provides a useful approximation of the fastest-growing perturbations.

Subsequently, a composite analysis of these 20 optimal initial errors is conducted to identify their spatial structures. The results reveal prominent localized distribution characteristics (Figure 5g–i), with initial-error maxima primarily concentrated over Northeast China, the Mongolian Plateau, and the region south of Lake Baikal. These results suggest that NECV forecast errors may be influenced by initial-field uncertainties in these upstream regions.

To evaluate the impact of the optimal initial errors on NECV predictability, sensitivity experiments are conducted within the Pangu-Weather model. For each case, the temporal evolution of the 20 initial perturbations exhibiting the most pronounced energy growth is composited. Figures 6–8 show the control forecasts, the corresponding composite perturbed forecasts, and their difference fields for the 500-hPa atmospheric state.

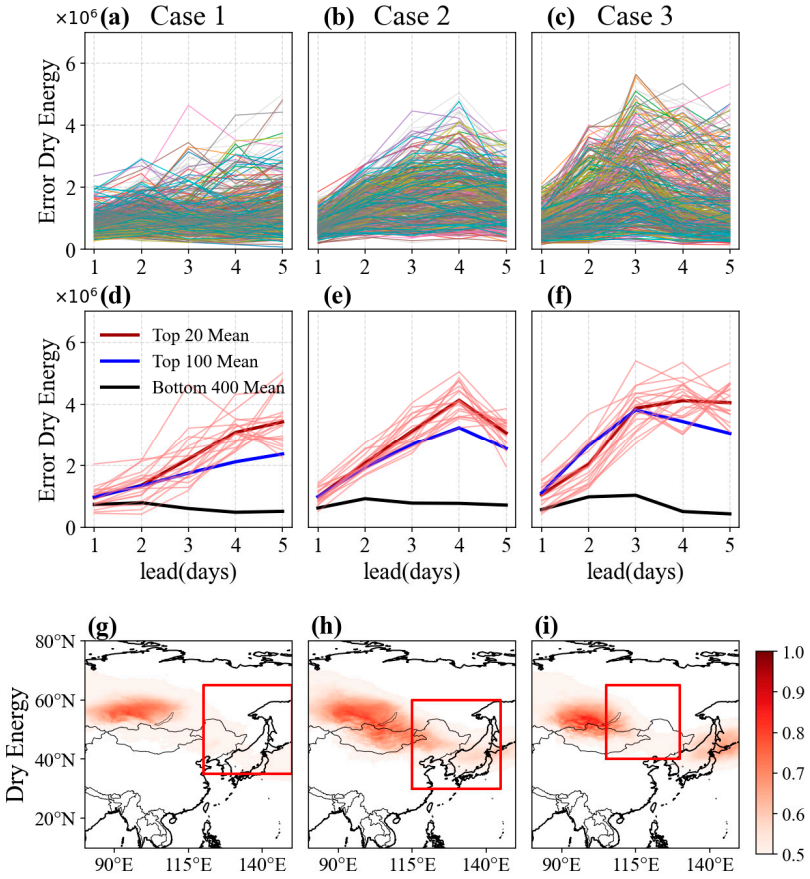


**Figure 5.** (**a**–**c**) Temporal evolution of vertically integrated dry-energy errors ($m^2/s^2$) relative to the control forecasts for 500 perturbations within the target region, shown for the three NECV cases at forecast lead times of 1–5 days. (**d**–**f**) The 20 perturbations with the largest dry-energy growth for each case are shown as thin red curves, and their mean is shown as the thick red curve. The thick blue and black curves denote the mean evolutions of the top 100 and bottom 400 perturbations, respectively. (**g**–**i**) Spatial composites (shading) of the dry-energy fields corresponding to the optimal initial error structures. The red box indicates the region of the NECV.

Overall, the optimal initial errors are associated with substantial changes in the forecast evolution and intensity of the cold vortex. For Cases 1 and 3, the composite perturbed forecasts (Figures 6 and 8) exhibit deeper geopotential height contours and enhanced temperature gradients, suggesting a more intense cold trough that deepens and propagates downstream. These features indicate that the identified optimal initial errors tend to amplify the cold-vortex system in these cases. In contrast, for Case 2 (Figure 7), the perturbed geopotential height field becomes shallower and the cold core weakens. This response suggests that the initial errors act to suppress the cold-vortex structure, consistent with the eastward propagation and decay stage of this case.

Despite these contrasting intensity responses, a coherent spatiotemporal pattern of error growth is evident across all three cases. Forecast errors initially concentrate near the frontal zone ahead of the developing cold trough, and subsequently amplify rapidly along the westerly jet before propagating downstream. This evolution follows a characteristic pattern of localized perturbation amplification followed by downstream development [35]. Throughout this process, the error fields are advected by the background wind while maintaining a high degree of spatial coherence. These results suggest that the identified optimal initial-error structures can promote nonlinear error growth in Pangu-Weather and subsequently affect NECV forecasts.

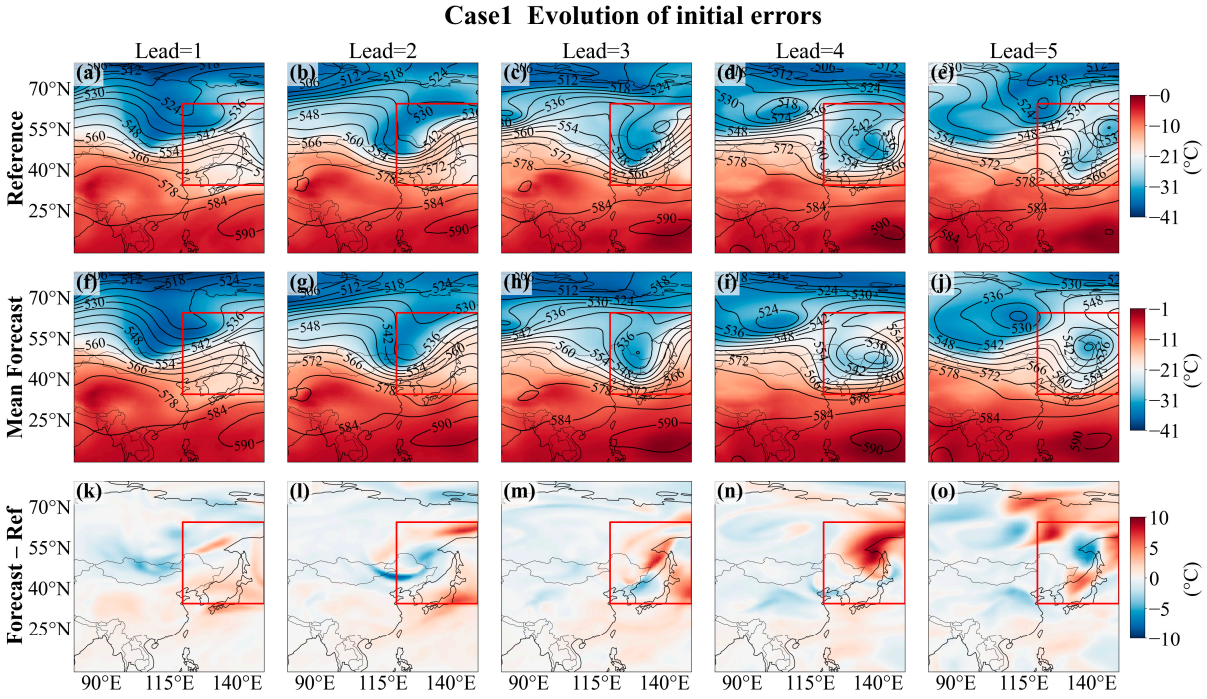


**Figure 6.** (**a**–**e**) Control forecast of the cold vortex in Case 1 at 500 hPa. (**f**–**j**) Composite-mean forecasts from the 20 optimal initial-error experiments. (**k**–**o**) Differences in temperature between the composite and control forecasts. Shading indicates temperature (°C), and black contours denote geopotential height (gpm). Contour labels indicate 500 hPa geopotential-height values, and the red box denotes the cold-vortex area.

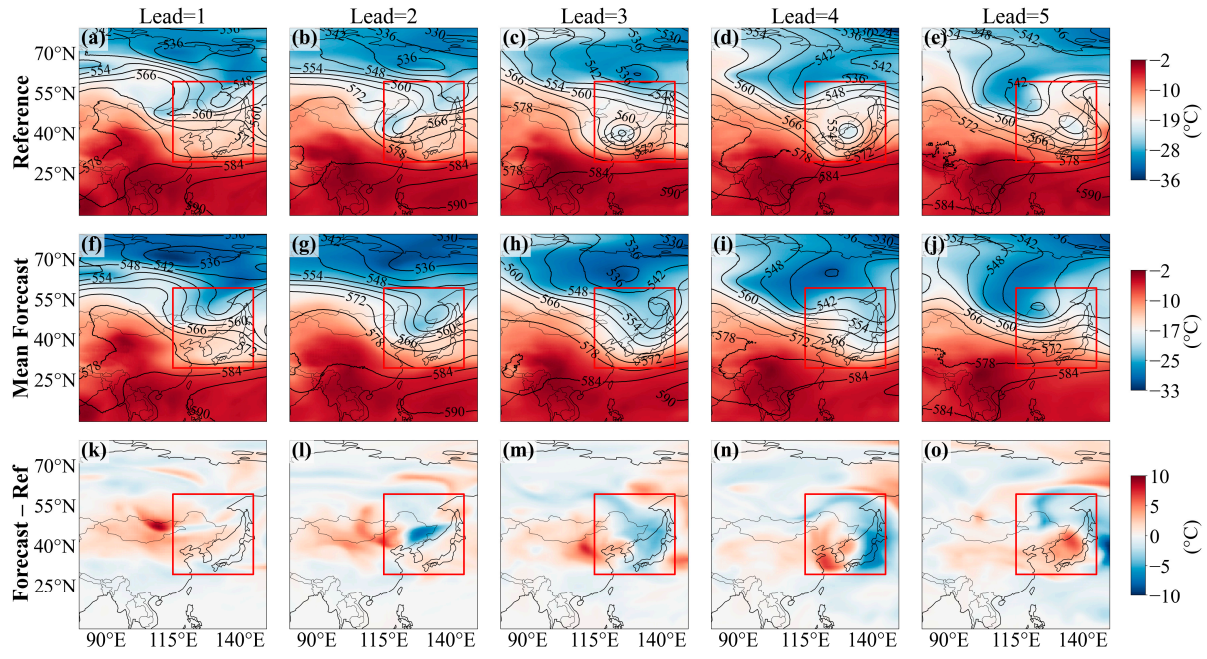


**Figure 7.** Same as Figure 6, but for Case 2. Panels (**a**–**e**) show the control forecast of the cold vortex at 500 hPa, panels (**f**–**j**) show the composite-mean forecasts from the 20 optimal initial-error experiments, and panels (**k**–**o**) show the temperature differences between the composite and control forecasts. Other variables, symbols, and plotting conventions are the same as in Figure 6.

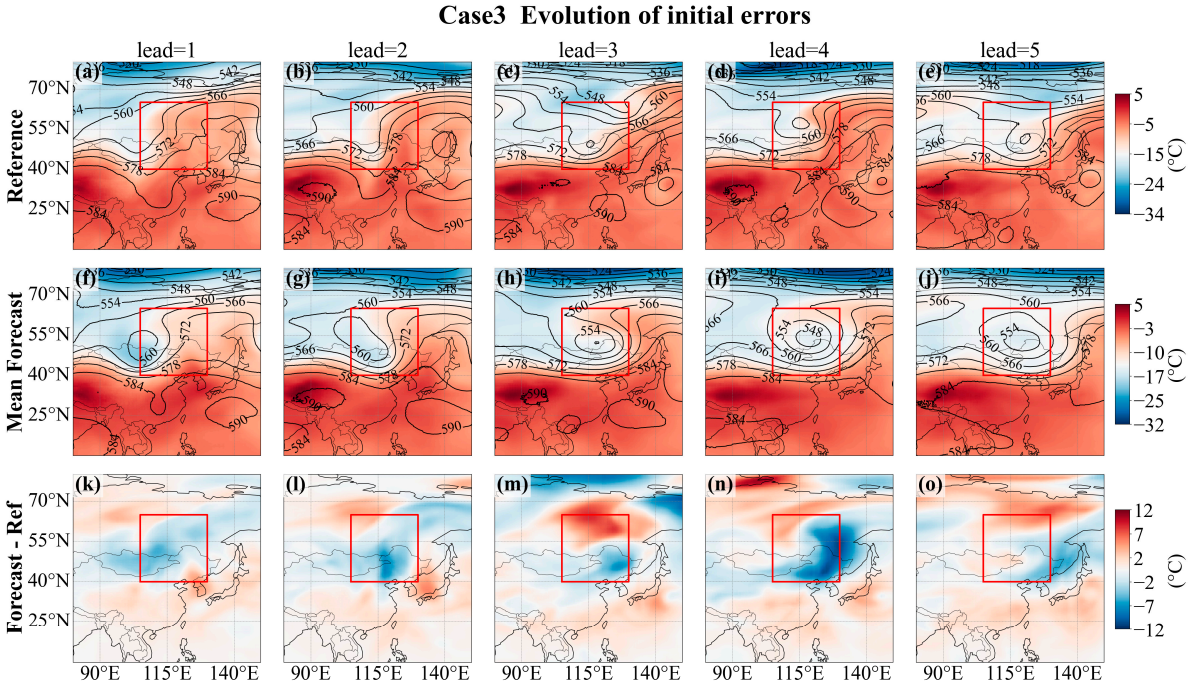


**Figure 8.** Same as Figure 6, but for Case 3. Panels (**a**–**e**) show the control forecast of the cold vortex at 500 hPa, panels (**f**–**j**) show the composite-mean forecasts from the 20 optimal initial-error experiments, and panels (**k**–**o**) show the temperature differences between the composite and control forecasts. Other variables, symbols, and plotting conventions are the same as in Figure 6.

### *3.3. The Development Mechanism of the Optimal Initial Error*

To further understand the energy-evolution mechanisms during the error growth, this subsection examines the perturbation kinetic-energy budget (see Section 2.5), focusing on the contributions of BC, BT, and ADV, as shown in Figures 9–11. The time series (panel m in each figure) show a consistent feature among the three cases: BC is the dominant source of perturbation kinetic-energy growth, whereas BT and ADV play comparatively secondary roles.

In Case 1, perturbation kinetic-energy growth is mainly driven by BC, which gradually intensifies and reaches its maximum during the later stage, indicating continuous conversion of mean-flow available potential energy into kinetic energy. BT remains weak, and ADV stays close to zero, suggesting limited contributions from momentum transport and advection. Spatially, strong BC is concentrated ahead of the trough and near the warm sector, consistent with regions of strong temperature gradients.

Case 2 exhibits a similar structure, with BC still dominating the growth process. BT shows only weak positive contributions in the early stage. In contrast to Case 1, ADV becomes increasingly negative, implying that advection suppresses system development through outward energy transport, mainly in the downstream region.

Case 3 shows the strongest baroclinic signature. BC increases continuously and more strongly than in the other two cases, indicating more efficient energy extraction from the background field. BT remains weak and becomes slightly negative in the later stage, while ADV also shows persistent negative contributions, suggesting clear outward energy dispersion.

Overall, the three cases suggest that baroclinic conversion is the primary mechanism for perturbation kinetic-energy growth, whereas barotropic conversion is generally weak and advection may suppress development in some cases. These results suggest that baroclinic instability likely plays a major role in NECV evolution, while initial errors modulate system intensity and structure by influencing the energy-conversion processes.

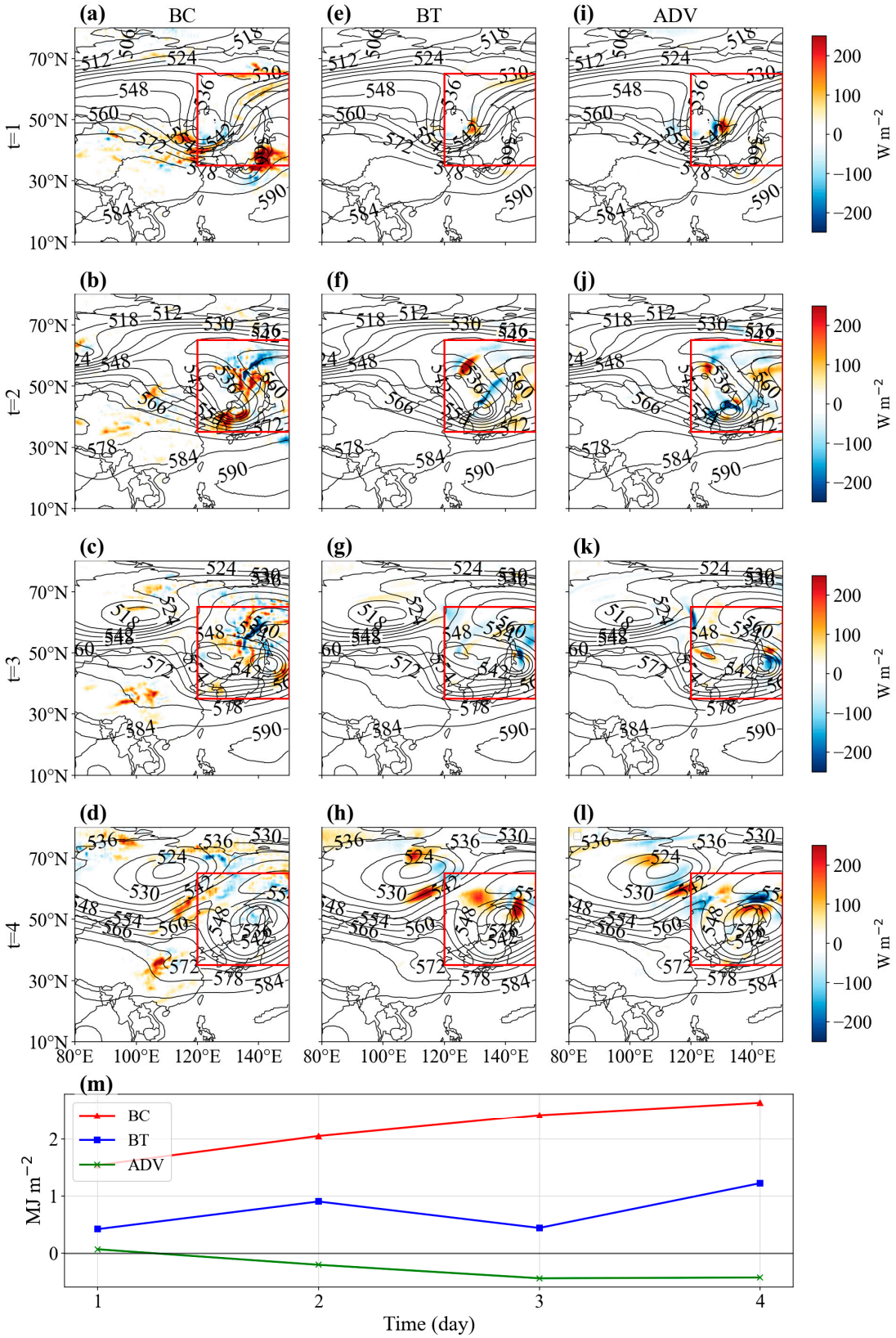


**Figure 9.** Evolution characteristics of the perturbation kinetic-energy budget terms for Case 1. (**a**–**l**) show the spatial distributions of BC, BT, and ADV from Day 1 to Day 4, respectively. Shading indicates the energy–conversion rate (W $m^{-2}$), black contours denote the 500 hPa geopotential height (gpm), the contour labels indicate the corresponding geopotential-height values, and the red box marks the cold-vortex area. (**m**) shows the cumulative vertically and regionally integrated contributions of BC, BT, and ADV within the red box.

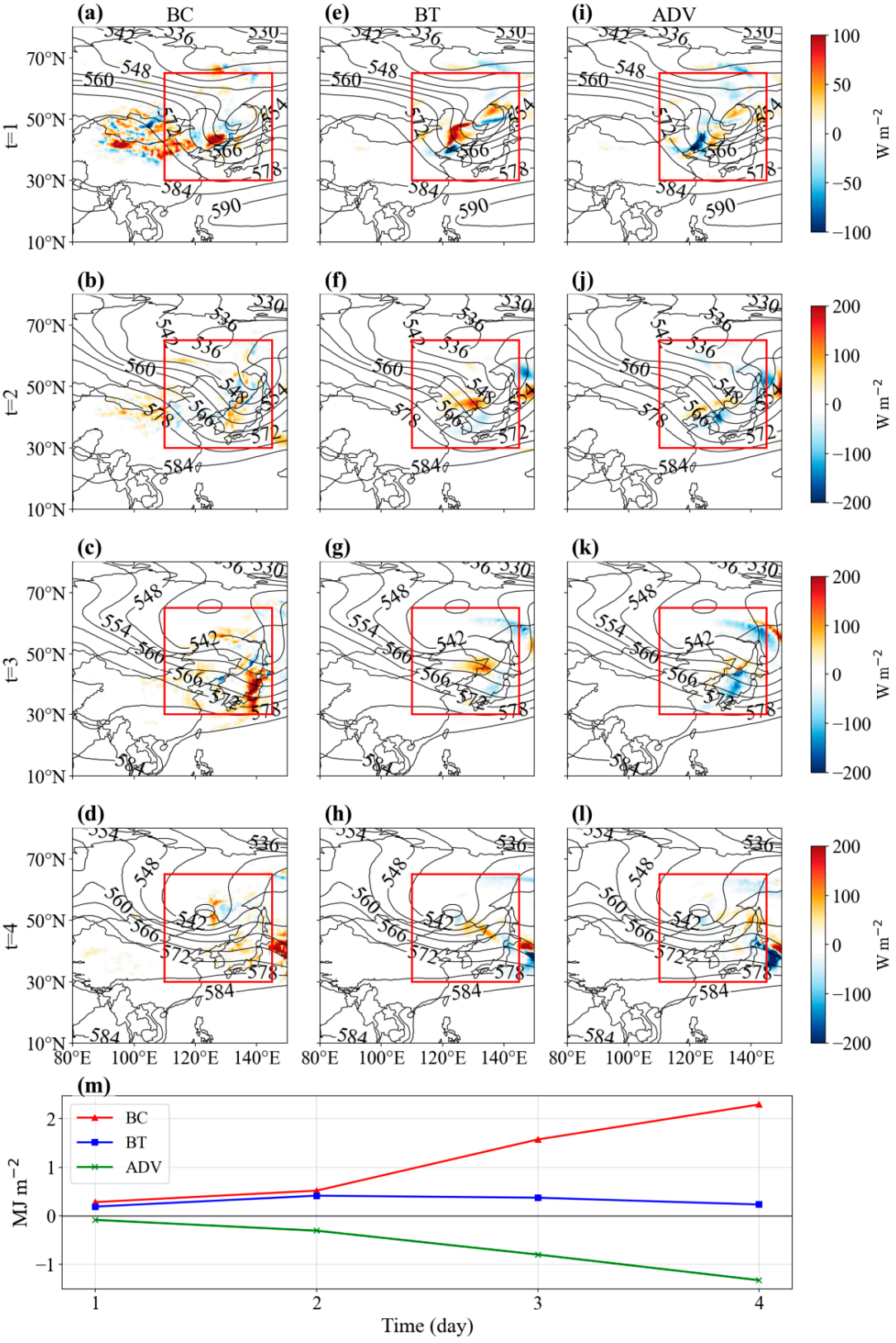


**Figure 10.** Same as Figure 9, but for Case 2. Panels (**a**–**l**) show the spatial distributions of BC, BT, and ADV from Day 1 to Day 4, respectively, and panel (**m**) shows the cumulative vertically and regionally integrated contributions of BC, BT, and ADV within the cold-vortex area. Other variables, symbols, and plotting conventions are the same as in Figure 9.

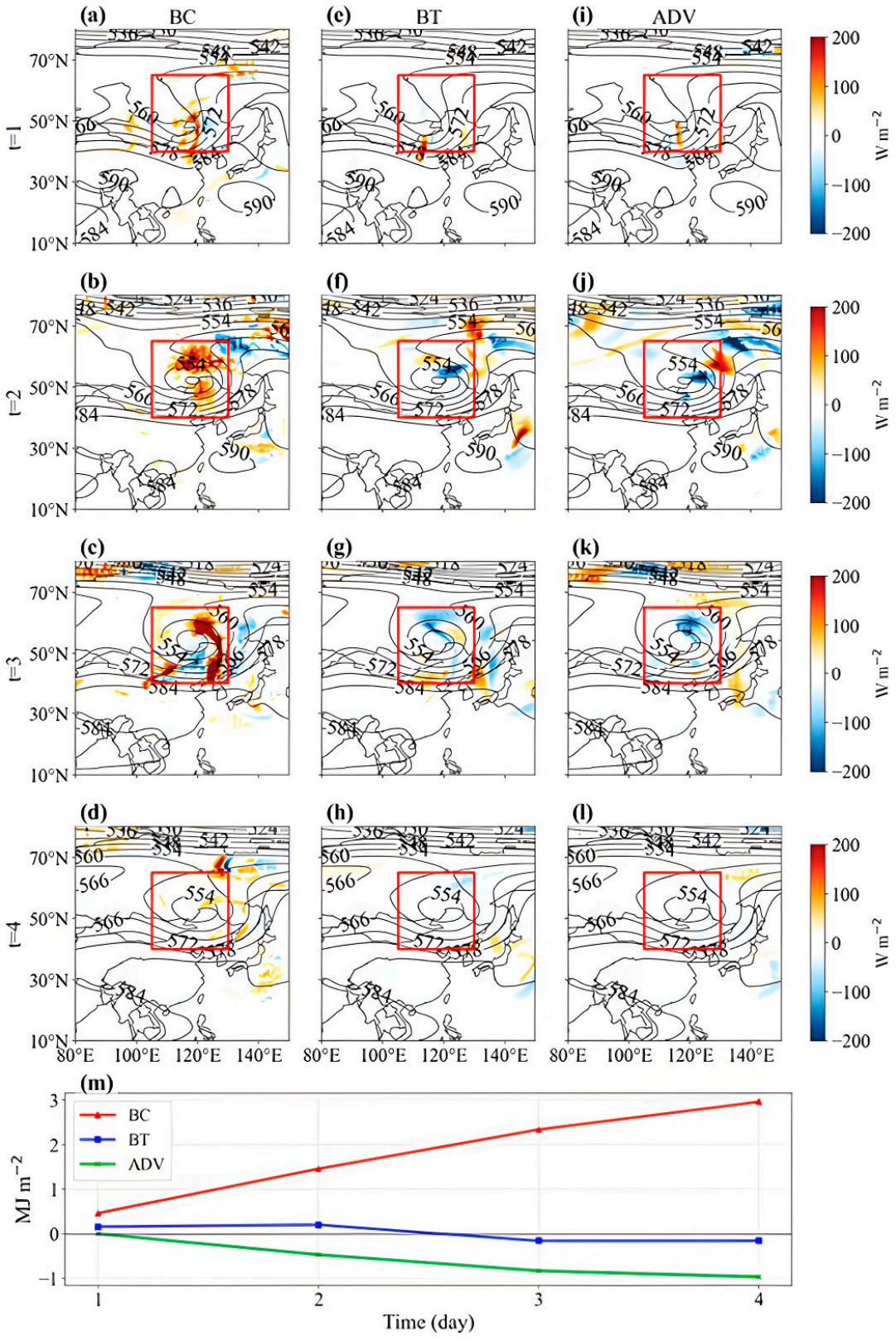


**Figure 11.** Same as Figure 9, but for Case 3. Panels (**a**–**l**) show the spatial distributions of BC, BT, and ADV from Day 1 to Day 4, respectively, and panel (**m**) shows the cumulative vertically and regionally integrated contributions of BC, BT, and ADV within the cold-vortex area. Other variables, symbols, and plotting conventions are the same as in Figure 9.

### *3.4. Identification of the Targeted Observation Sensitive Areas*

Figure 5g–i indicate that the patterns of the optimal initial errors exhibit pronounced spatial localization, which motivates the identification of targeted observation sensitive areas. More specifically, based on the dry-energy norm, the 1600 grid points with the largest error amplitudes are selected for each case (accounting for approximately 2% of the total domain grid points) to constitute the sensitive region of CNOP-type optimal initial errors (RC). As shown in Figure 12a–c, the RC spatial distribution exhibits remarkable consistency, with all three cases steadily located along the upstream pathway of the cold-vortex formation and development region. This spatial distribution is dynamically reasonable. Previous studies have shown that NECV activity is closely linked to upstream precursor disturbances and Rossby wave-energy propagation along the Eurasian mid-latitude waveguide. In addition, jet-entrance regions often favor perturbation growth through enhanced baroclinic conversion and upper-level flow interaction [36]. To verify the effectiveness of RC in improving forecast skill, four additional comparison subdomains (R1–R4), independent of the error-growth analysis, are manually selected (indicated by blue boxes). Based on these regions, a series of OSSEs is designed to quantitatively evaluate the forecasting improvement of reducing initial errors in different regions.

First, the initial field $X_{in}$ is used as model input to obtain the reference forecast field $F_1 = M_t(X_{in})$. Then, a random perturbation field is sampled from a standard normal distribution and constrained to 0.3 times the standard deviation, denoted as $x_{re}$. This perturbation is added to the initial field, and Pangu-Weather is used to generate the resulting forecast $F_2 = M_t(X_{in} + x_{re})$. Furthermore, by selectively removing errors in the sensitive regions (RC and R1–R4) and re-running Pangu-Weather, the resulting forecast is obtained as $F_3 = M_t(X_{in} + x_{ren})$. It should be noted that the forecast error in this study is computed using the geopotential height field at 500 hPa. To quantify the degree to which reducing initial perturbations in a specific region suppresses error growth, the regional improvement rate is computed as:

$$I_r = \frac{|F_2 - F_1| - |F_3 - F_1|}{|F_2 - F_1|} \times 100\%, \quad (7)$$

The value of $I_r$ provides a direct comparison of the relative effectiveness of enhanced observations across different regions under identical initial-error conditions. To avoid randomness associated with individual perturbations, 50 independent perturbation experiments are conducted for each NECV case. During the OSSEs, errors are removed in RC and in the four comparison regions R1–R4 for each perturbation field, resulting in a total of $5 \times 50 = 250$ re-forecast experiments. The final standard deviation maps are based on statistics from all experiments.

The statistical results shown in Figure 12d–f further confirm that, in all three cases, targeted observations within RC—by effectively reducing initial errors—produce the most substantial forecast improvements. Compared with R1–R4, correcting the initial errors within RC produces much larger forecast improvements. Under the idealized assumption that all initial errors in the sensitive region are removed, the mean improvement reaches about 13%. In contrast, the improvements associated with R3 and R4 are very small in all experiments.

The experiments with partial error removal show similar results. When 70%, 50%, or 30% of the initial errors within RC are removed, the forecast improvements remain consistently larger in the RC region than those obtained for R1–R4. As the proportion of removed errors decreases, the improvement gradually weakens. However, RC still maintains a clear advantage over the artificially selected regions.

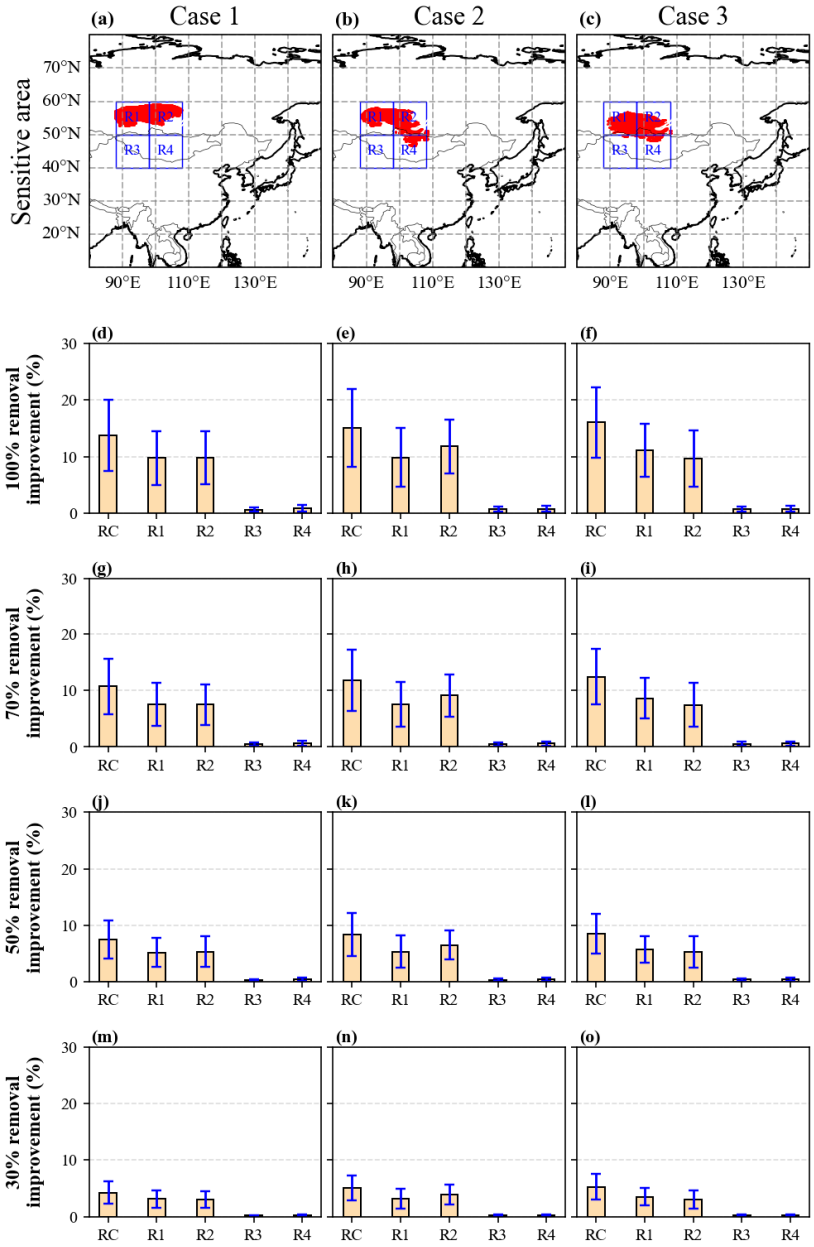


**Figure 12.** (**a**–**c**) Sensitive regions (red shading) identified from the fastest-growing initial-error structures. (**d**–**f**) Relative forecast improvement rates of NECV in the OSSEs after removing 100% of the initial errors within each corresponding region. (**g**–**o**) are the same as (**d**–**f**), but for experiments in which 70%, 50%, and 30% of the initial errors are removed, respectively. RC denotes the sensitive region, and R1–R4 represent the artificially defined comparison boxes. Bars indicate the mean forecast improvement, and error bars denote plus/minus one standard deviation ($\pm 1$ $\sigma$).

## 4. Discussion

The present study suggests the feasibility of combining a deep-learning forecasting model with a CNOP-based framework to diagnose NECV predictability and targeted observation sensitivity. For the selected NECV cases, Pangu-Weather exhibits forecast skill that is generally slightly superior to ECMWF-IFS, which supports its use in subsequent nonlinear sensitivity experiments. The identified optimal initial errors are mainly concentrated upstream of the cold vortex and near the entrance region of the upper-level jet. Subsequently, these regions are recognized as targeted observation sensitive areas for NECV prediction. Perturbation kinetic-energy budget analysis indicates that baroclinic energy conversion is the dominant mechanism for forecast-error growth. OSSEs further

suggest that constraining initial errors within these sensitive areas can improve cold-vortex forecast skill by up to approximately 13% under idealized conditions, indicating their potential practical value for targeted observing strategies over Northeast Asia.

This study helps address a gap in previous research by providing, to our knowledge, the first CNOP-based predictability analysis of NECVs. More broadly, the results suggest that predictability diagnostics can also be applied to artificial intelligence (AI) forecasting systems. Compared with conventional NWP-based studies, the use of Pangu-Weather offers much higher computational efficiency, making large-sample perturbation experiments and sensitivity analyses more practical. The overall consistency between the identified sensitive regions and previous dynamical studies further suggests that AI-based frameworks can capture physically meaningful sources of forecast uncertainty and may provide an efficient new tool for future predictability research.

Several limitations should also be noted. First, Pangu-Weather remains a black-box model, and its internal forecast processes cannot yet be fully interpreted in dynamical terms. Second, the CNOP-type perturbations are identified through sampling-based search rather than formal optimization, and therefore represent practical approximations to the theoretical optimum. Third, because the OSSE framework is idealized, the reported forecast improvement should be regarded as a theoretical upper bound. In practice, factors such as instrument errors and uneven observational coverage may reduce the actual achievable improvement. In addition, the evaluation dataset is limited in both temporal coverage and sample size, as only May–August 2022 data and three representative NECV cases are examined. Model performance may also vary across seasons, regions, variables, and synoptic conditions. Finally, only one deep learning model is investigated here. Future work should include larger samples, longer evaluation periods, partial-error reduction scenarios, and cross-model comparisons among multiple AI and NWP systems.

**Author Contributions:** Conceptualization, C.Z., J.Q. and Q.W.; methodology, C.Z.; writing—original draft preparation, C.Z.; writing—review and editing, C.Z., J.Q. and Q.W.; supervision, Q.W.; funding acquisition, Q.W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Open Fund Project of Key Laboratory of Marine Environmental Information Technology, Ministry of Natural Resources of the People's Republic of China, and the Fundamental Research Funds for the Central Universities (B250202005).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The datasets analyzed in this study are publicly available from ERA5 (http://doi.org/10.24381/cds.f17050d7 (accessed on 6 December 2024)). The Pangu-Weather model code is accessible at https://github.com/198808xc/Pangu-Weather (accessed on 5 December 2024), and the code implementing the nonlinear optimization framework can be obtained from the corresponding author upon request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

| | |
|---|---|
| NECV | Northeast China Cold Vortex |
| NCEP | National Centers for Environmental Prediction |
| NWP | numerical weather prediction |
| CNOP | conditional nonlinear optimal perturbation |
| ECMWF | European Centre for Medium-Range Weather Forecasts |
| IFS | integrated forecasting system |

| | |
|---|---|
| OSSEs | observing system simulation experiments |
| C3S | Copernicus Climate Change Service |
| 4D-Var | four-dimensional variational data assimilation |
| RMSE | root-mean-square error |
| ACC | anomaly correlation coefficient |
| BC | baroclinic conversion |
| BT | barotropic conversion |
| ADV | advection |
| AI | artificial intelligence |

## References


1. Price, J.D.; Vaughan, G. The potential for stratosphere–troposphere exchange in cut-off low systems. *Q. J. R. Meteorol. Soc.* **1993**, *119*, 343–365. [CrossRef]
2. Nieto, R.; Gimeno, L.; De la Torre, L.; Ribera, P.; Gallego, D.; García-Herrera, R.; Agustín García, J.; Nuñez, M.; Redaño, A. Climatological features of cutoff low systems in the Northern Hemisphere. *J. Clim.* **2005**, *18*, 3085–3103. [CrossRef]
3. Zhang, C.; Zhang, Q.; Wang, Y.; Liang, X. Climatology of warm season cold vortices in East Asia: 1979–2005. *Meteorol. Atmos. Phys.* **2008**, *100*, 291–301. [CrossRef]
4. Chen, H.; Xie, Z.; He, X.; Zhao, X.; Gao, Z.; Wu, B.; Zhang, J.; Zou, X. Northeast China cold vortex amplifies extreme precipitation events in the middle and lower Yangtze River basin. *Atmosphere* **2024**, *15*, 819. [CrossRef]
5. Xie, Z.; Bueh, C. Different types of cold vortex circulations over Northeast China and their weather impacts. *Mon. Wea. Rev.* **2015**, *143*, 845–863. [CrossRef]
6. Meng, F.; Xie, K.; Liu, P.; Chen, H.; Wang, Y.; Shi, H. Extreme precipitation trends in Northeast China based on a non-stationary generalized extreme value model. *Geosci. Lett.* **2024**, *11*, 13. [CrossRef]
7. Kalnay, E. *Atmospheric Modeling, Data Assimilation and Predictability*; Cambridge University Press: Cambridge, UK, 2003.
8. Leutbecher, M.; Palmer, T.N. Ensemble forecasting. *J. Comput. Phys.* **2008**, *227*, 3515–3539. [CrossRef]
9. Bauer, P.; Thorpe, A.; Brunet, G. The quiet revolution of numerical weather prediction. *Nature* **2015**, *525*, 47–55. [CrossRef]
10. Palmer, T. The ECMWF ensemble prediction system: Looking back (more than) 25 years and projecting forward 25 years. *Q. J. R. Meteorol. Soc.* **2018**, *145*, 12–24. [CrossRef]
11. Buizza, R.; Palmer, T.N. The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.* **1995**, *52*, 1434–1456. [CrossRef]
12. Houtekamer, P.L.; Zhang, F. Review of the ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.* **2016**, *144*, 4489–4532. [CrossRef]
13. Leutbecher, M.; Lock, S.-J.; Ollinaho, P.; Lang, S.T.K.; Balsamo, G.; Bechtold, P.; Bonavita, M.; Christensen, H.M.; Diamantakis, M.; Dutra, E.; et al. Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Q. J. R. Meteorol. Soc.* **2017**, *143*, 2315–2339. [CrossRef]
14. Weyn, J.A.; Durran, D.R. The Dependence of the Predictability of Mesoscale Convective Systems on the Horizontal Scale and Amplitude of Initial Errors in Idealized Simulations. *J. Atmos. Sci.* **2017**, *74*, 153–170. [CrossRef]
15. Reichstein, M.; Camps-Valls, G.; Stevens, B.; Jung, M.; Denzler, J.; Carvalhais, N. Deep learning and process understanding for data-driven Earth system science. *Nature* **2019**, *566*, 195–204. [CrossRef]
16. Scher, S.; Messori, G. Weather and climate forecasting with neural networks: Using general circulation models (GCMs) with different complexity as a study ground. *Geosci. Model Dev.* **2019**, *12*, 2797–2809. [CrossRef]
17. Rasp, S.; Dueben, P.D.; Scher, S.; Weyn, J.A.; Mouatadid, S.; Thuerey, N. WeatherBench: A benchmark dataset for data-driven weather forecasting. *J. Adv. Model. Earth Syst.* **2020**, *12*, e2020MS002203. [CrossRef]
18. Schultz, M.G.; Betancourt, C.; Gong, B.; Kleinert, F.; Langguth, B.; Leufen, L.H.; Mozaffari, A.; Stadtler, S. Can deep learning beat numerical weather prediction? *Phil. Trans. R. Soc. A* **2021**, *379*, 20200097. [CrossRef]
19. Pathak, J.; Subramanian, S.; Harrington, P.; Raja, S.; Chattopadhyay, A.; Mardani, M.; Kurth, T.; Hall, D.; Li, Z.; Azizzadenesheli, K.; et al. FourCastNet: A global data-driven high-resolution weather model using adaptive Fourier neural operators. *arXiv* **2022**, arXiv:2202.11214. [CrossRef]
20. Lam, R.; Sanchez-Gonzalez, A.; Willson, M.; Wirnsberger, P.; Fortunato, M.; Alet, F.; Ravuri, S.; Ewalds, T.; Eaton-Rosen, Z.; Hu, W.; et al. Learning skillful medium-range global weather forecasting. *Science* **2023**, *382*, 1416–1421. [CrossRef]
21. Bi, K.; Xie, L.; Zhang, H.; Chen, X.; Gu, X.; Tian, Q. Accurate medium-range global weather forecasting with 3D neural networks. *Nature* **2023**, *619*, 533–538. [CrossRef]
22. Langland, R.H. Issues in targeted observing. *Q. J. R. Meteorol. Soc.* **2005**, *131*, 3409–3425. [CrossRef]

23. Palmer, T.N.; Gelaro, R.; Barkmeijer, J.; Buizza, R. Singular vectors, metrics, and adaptive observations. *J. Atmos. Sci.* **1998**, *55*, 633–653. [CrossRef]
24. Torn, R.D.; Hakim, G.J. Ensemble-based sensitivity analysis. *Mon. Wea. Rev.* **2008**, *136*, 663–677. [CrossRef]
25. Buizza, R.; Montani, A. Targeting observations using singular vectors. *J. Atmos. Sci.* **1999**, *56*, 2965–2985. [CrossRef]
26. Ancell, B.C.; Hakim, G.J. Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.* **2007**, *135*, 4117–4134. [CrossRef]
27. Mu, M.; Duan, W.S.; Wang, B. Conditional nonlinear optimal perturbation and its applications. *Nonlin. Process. Geophys.* **2003**, *10*, 493–501. [CrossRef]
28. Duan, W.; Mu, M.; Wang, B. Conditional nonlinear optimal perturbations as the optimal precursors for El Niño–Southern Oscillation events. *J. Geophys. Res.* **2004**, *109*, D23105. [CrossRef]
29. Wang, Q.; Mu, M.; Pierini, S. The fastest growing initial error in prediction of the Kuroshio Extension state transition processes and its growth. *Clim. Dyn.* **2020**, *54*, 1953–1971. [CrossRef]
30. Hersbach, H.; Bell, B.; Berrisford, P.; Hirahara, S.; Horányi, A.; Muñoz-Sabater, J.; Nicolas, J.; Peubey, C.; Radu, R.; Schepers, D.; et al. The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1999–2049. [CrossRef]
31. Sun, L.; Zhen, X.Y.; Wang, Q. The climatological characteristics of northeast cold vortex in China. *J. Appl. Meteor. Sci.* **1994**, *5*, 297–303. (In Chinese). Available online: https://qikan.camscma.cn/article/id/YYQX403005 (accessed on 20 November 2025).
32. Liu, H.B.; Wen, M.; He, J.H.; Zhang, R.H. Characteristics of the northeast cold vortex at intraseasonal timescale and its impact. *Chin. J. Atmos. Sci.* **2012**, *36*, 959–973. (In Chinese) [CrossRef]
33. Feng, R.; Duan, W.S. Indian Ocean Dipole-related predictability barriers induced by initial errors in the tropical Indian Ocean in a CGCM. *Adv. Atmos. Sci.* **2019**, *36*, 658–668. [CrossRef]
34. Feng, R.; Duan, W.S. Sensitive area in the tropical Indian Ocean for advancing beyond the summer predictability barrier of Indian Ocean Dipole. *Dyn. Atmos. Ocean.* **2025**, *110*, 101552. [CrossRef]
35. Hoskins, B.J.; Ambrizzi, T. Rossby Wave Propagation on a Realistic Longitudinally Varying Flow. *J. Atmos. Sci.* **1993**, *50*, 1661–1671. [CrossRef]
36. Song, J. Northeast China cold vortex: Three-dimensional structure, life cycle, and dynamical understanding. *J. Meteor. Res.* 2026, *in press*. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.