SCIENCE CHINA
Earth Sciences

•RESEARCH PAPER•

October 2022 Vol.65 No.10: 1903–1920
https://doi.org/10.1007/s11430-022-9991-4

Uncertainties and error growth in forecasting the record-breaking rainfall in Zhengzhou, Henan on 19–20 July 2021

Yunji ZHANG[1], Huizhen YU[2], Murong ZHANG[3], Yawen YANG[4,5] & Zhiyong MENG[5*]

[1] Department of Meteorology and Atmospheric Science, Pennsylvania State University, University Park, Pennsylvania 16802, USA;
[2] Qingdao Meteorological Bureau, Qingdao 266003, China;
[3] State Key Laboratory of Marine Environmental Science, College of Ocean and Earth Sciences, Xiamen University, Xiamen 361102, China;
[4] Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China;
[5] Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

Received February 21, 2022; revised August 12, 2022; accepted August 23, 2022; published online September 19, 2022

Abstract This study explores the controlling factors of the uncertainties and error growth at different spatial and temporal scales in forecasting the high-impact extremely heavy rainfall event that occurred in Zhengzhou, Henan Province China on 19−20 July 2021 with a record-breaking hourly rainfall exceeding 200 mm and a 24-h rainfall exceeding 600 mm. Results show that the strengths of the mid-level low-pressure system, the upper-level divergence, and the low-level jet determine both the amount of the extreme 24-h accumulated and hourly rainfall at 0800 UTC. The forecast uncertainties of the accumulated rainfall are insensitive to the magnitude and the spatial structure of the tiny, unobservable errors in the initial conditions of the ensemble forecasts generated with Global Ensemble Forecast System (GEFS) or sub-grid-scale perturbations, suggesting that the predictability of this event is intrinsically limited. The dominance of upscale rather than upamplitude error growth is demonstrated under the regime of $k^{-5/3}$ power spectra by revealing the inability of large-scale errors to grow until the amplitude of small-scale errors has increased to an adequate amplitude, and an apparent transfer of the fastest growing scale from smaller to larger scales with a slower growth rate at larger scales. Moist convective activities play a critical role in enhancing the overall error growth rate with a larger error growth rate at smaller scales. In addition, initial perturbations with different structures have different error growth features at larger scales in different variables in a regime transitioning from the $k^{-5/3}$ to k^{-3} power law. Error growth with conditional nonlinear optimal perturbation (CNOP) tends to be more upamplitude relative to the GEFS or sub-grid-scale perturbations possibly owing to the inherited error growth feature of CNOP, the inability of convective parameterization scheme to rebuild the $k^{-5/3}$ power spectra at the mesoscales, and different error growth characteristics in the $k^{-5/3}$ and k^{-3} regimes.

Keywords Extremely heavy rainfall, Forecast error, Predictability, Ensemble forecast, Henan

Citation: Zhang Y, Yu H, Zhang M, Yang Y, Meng Z. 2022. Uncertainties and error growth in forecasting the record-breaking rainfall in Zhengzhou, Henan on 19–20 July 2021. Science China Earth Sciences, 65(10): 1903–1920, https://doi.org/10.1007/s11430-022-9991-4

1. Introduction

Wide-spread torrential rainfall hit Henan province on 17−22 July 2021. The most intense episode of rainfall occurred on 19−20 July 2021 in Zhengzhou, the capital city of Henan Province, and the surrounding area. The 24-h accumulated rainfall from 0000 UTC 20 July to 0000 UTC 21 July (LST=UTC+8) exceeded 600 mm and several stations recorded their respective historically highest daily accumulated rainfall (Ran et al., 2021; Shi et al., 2021). The highest hourly rainfall of 201.9 mm in metropolitan Zhengzhou city, which occurred from 0800 UTC to 0900 UTC 20 July, marks

* Corresponding author (email: zymeng@pku.edu.cn)

a new record for hourly rain rate in Chinese mainland (Shi et al., 2021; Sun et al., 2021; Zhuang and Xing, 2022). Associated hazards on that day, especially inland flash flooding, leading to 380 casualties.

The geographical locations and topographical features of Henan Province make it prone to many different types of heavy-rain-producing weather systems (e.g., Liang et al., 2020). The devastating rainfall event on 19–20 July 2021 has been found to be contributed by various environmental forcing such as liftings associated with upper-level troughs and mid-level vortices, moisture transportations associated with the subtropical high to the east, Typhoon In-Fa to the southeast, and Typhoon Cempaka to the south, as well as low-level convergence associated with low-level jet and local topography (Ran et al., 2021; Sun et al., 2021). Many dynamical and thermodynamical parameters during this event deviated significantly from the climatology of major torrential rainfall events in this area, especially low-level vorticity and column-integrated precipitable water (Zhang et al., 2021).

In spite of the extremeness of this event, operational weather forecast offices of Henan Province and Zhengzhou predicted the occurrence of this extreme rainfall event a few days before and issued several warnings for a vast region in Henan Province in the following days prior to this event. However, the most intense rainfall centers in the operational forecasts are several hundred kilometers away from the actual epicenter. Similarly, forecasts from several global and regional numerical weather prediction (NWP) models show large variations in terms of the location and the intensity of the highest accumulated rainfall (Shi et al., 2021). This work aims to understand these discrepancies between forecasted and observed rainfall in this record-breaking rainfall event by examining the forecast uncertainties and the associated error growth mechanism.

The extent of accuracy in numerical weather forecasting is often referred to as "atmospheric predictability" and was first proposed by Lorenz (1963). Lorenz (1996) categorized this problem into practical predictability (Lorenz, 1982), or the forecast capability given currently available knowledge and techniques, and intrinsic predictability (Lorenz, 1969), or the longest possible forecast extent given nearly perfect knowledge and techniques. One important aspect of practical predictability originates from uncertainties in representing key environmental forcing. Key environment forcing for a heavy rainfall event is generally identified using ensemble-based sensitivity analysis (ESA), which measures linear relationships between a scalar forecast metric and atmospheric state variables through ensemble statistics (Hakim and Torn, 2008), or the conditional nonlinear optimal perturbation (CNOP; Mu and Duan, 2003) method, which is an adjoint-based method that takes nonlinear processes into account. By diagnosing key environmental factors in extreme rainfall events, understanding of the rainfall forecast uncertainties can be improved (e.g., Hawblitzel et al., 2007; Lynch and Schumacher, 2014; Yu and Meng, 2016; Zhang and Meng, 2018). Based on ESA, Zhang and Meng (2018) revealed the importance of well-forecast low-level jet locations in determining the performance of ensemble rainfall forecast during a persistent heavy rainfall event in Guangdong, China, in early spring of 2014. Based on both ESA and CNOP, Yu and Meng (2016) consistently demonstrated the essential role of the mid-level trough in the westerly flow and the associated low-level low in the high-impact rainfall event in Beijing, China, on 21 July 2012. Yu and Meng (2022) found that the CNOP with moist physics identified the sensitive areas at both the lower levels and upper levels for four typical heavy rainfall events in north China. The upper-level sensitive area, which corresponds to the upper-level weather systems, is associated with high baroclinicity, while the lower-level sensitive area, which corresponds to the lower-level weather systems, is associated with moist physics. Although there have been studies revealing weather systems that may have affected this record-breaking extremely heavy rainfall in Henan (e.g., Ran et al., 2021; Sun et al., 2021), the key environmental factors and their associated forecast uncertainties remain unknown.

Unlike practical predictability, which is primarily controlled by uncertainties in NWP models and initial conditions, intrinsic predictability is primarily limited by error growth mechanisms that are inherently embedded in the dynamical and thermodynamical processes of the weather (e.g., Melhauser and Zhang, 2012; Sun and Zhang, 2016). Zhang et al. (2007) presented the conceptual model of how tiny, unobservable errors will limit the predictability at the mesoscales: those small-amplitude small-scale errors will grow upscales and rapidly spread with the help of moist convective processes, saturate at smaller scales and transfer to progressively larger scales through geostrophic adjustments, and eventually limit the predictability of mesoscale and synoptic scales. This conceptual model has been proved by many following studies (e.g., Judt, 2018; Selz, 2019; Selz and Craig, 2015; Sun and Zhang, 2016, 2020; Sun et al., 2017; Zhang et al., 2016; Zhang et al., 2019).

Several studies argue that large-scale errors are just as important as, if not more than, small-scale errors (Durran and Gingrich, 2014; Durran and Weyn, 2016; Nielsen and Schumacher, 2016; Zhang, 2021), and errors grow upamplitude at all model-resolved scales simultaneously rather than transfer upscales (Weyn and Durran, 2017; Judt, 2018, 2020). It should be noted that these different disagreements are essentially equivalent: small-scale errors are more important if the errors are governed by upscale growth, because the upscale growth of small-scale errors will dominate the existing large-scale errors (Zhang et al., 2007); while large-scale errors are more important if the errors are governed by upamplitude growth, because large-scale errors can grow to greater amplitudes owing to the greater base energy at these scales (Durran and Weyn, 2016). Understanding the relative importance of errors at different spatial scales will facilitate a better understanding of the error growth mechanisms. Therefore, many of the previous studies have used high-resolution, convection-permitting ensemble forecasts that incorporate initial condition uncertainties of different amplitudes and/or spatial scales to examine the error growth mechanisms (e.g., Melhauser and Zhang, 2012; Nielsen and Schumacher, 2016; Zhang et al., 2016; Weyn and Durran, 2019). However, previous studies either examined the sensitivity of the forecast error growth to different amplitudes and horizontal scales of homogeneous initial uncertainties or did not examine this sensitivity when flow-dependent initial uncertainties were imposed, while how sensitive the forecast error growth is to different amplitudes and horizontal scales of flow-dependent unobservable initial uncertainties in a high-resolution convection-permitting ensemble forecast on a real-world high-impact rainfall event remains unknown.

In addition to the scale and amplitude, the structure of initial perturbations may influence the forecast uncertainty and error growth features as well. Initial perturbations with different structures are mainly generated through breeding vectors, singular vectors, and random sampling from a climatologically based background error covariance such as CV3 from the WRFDA package, and the CNOP method. Mu et al. (2007) found that CNOP-type error tends to have a seasonal dependent evolution and produces the most considerable negative effect on the forecast results. Adding CNOP to the initial condition yields a spring predictability barrier phenomenon, while adding perturbations with the same magnitude but a different structure from the CNOP does not. How sensitive the forecast error growth is to different structures of initial uncertainties in a real-world extremely heavy rain event is also one interesting question to answer.

Therefore, to explore the uncertainties and error growth in forecasting this high-impact torrential rainfall event at different spatial and temporal scales, we present a suite of analyses using forecasts from numerical models ranging from global models to regional, convection-permitting models in this study. This includes ESA using the Observing System Research and Predictability Experiment (THORPEX) Interactive Grand Global Ensemble (TIGGE; Bougeault et al., 2010), the CNOP method using coarse-resolution simulations from the Pennsylvania State University–National Center for Atmospheric Research (PSU–NCAR) fifth-generation Mesoscale Model (MM5; Grell et al., 1995), and high-resolution convection-permitting ensemble simulations from the Weather Research and Forecasting (WRF) model with initial perturbations of different amplitudes and spatial scales.

2. Data and methodology

2.1 Observed 24-h accumulated rainfall

Hourly rain gauge data provided by the China Meteorological Administration with an average site spacing of ~5–10 km was interpolated to a 0.1°×0.1° grid using a Cressman interpolation method (Cressman, 1959). Figure 1a shows the 24-h accumulated rainfall from 1200 UTC 19 July to 1200 UTC 20 July 2021. The accumulated rainfall intensely concentrated over northern Henan Province, with a maximum of 505.54 mm, a considerable area that exceeds 400 mm, and an area-average of 74.49 mm over the inner box of Figure 1a.

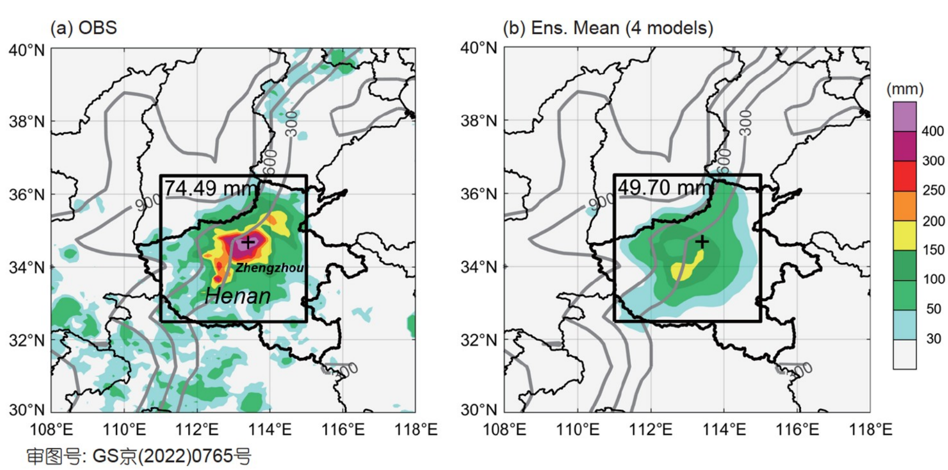

Figure 1 Rainfall distribution in observation and the mean of ensemble forecasts. (a) Observed 24-h accumulated rainfall from 1200 UTC 19 July to 1200 UTC 20 July 2021 (shading; units: mm) and terrain height (grey contour; units: m). (b) Same as (a) but for ensemble mean rainfall forecast initialized at 0000 UTC 19 July based on 4 best-performed models (BoM, NCMRWF, UKMO, and KMA; see text for details). The black box denotes the focused region, and the area-averaged 24-h accumulated rainfall is given on the top left of the box. The location of Zhengzhou City is marked as black cross and Henan Province is outlined in solid black.

2.2 The TIGGE ensemble and the ensemble sensitivity analysis

Forecasts on the 24-h accumulated rainfall from 1200 UTC 19 July to 1200 UTC 20 July over the focused region given in Figure 1a were quantitatively evaluated using the TIGGE ensembles, with a forecast initialization of 0000 UTC 19 July. Twelve global models from TIGGE evaluated in the present study are listed in Table 1. The TIGGE-derived fields were interpolated into a 0.1°×0.1° grid to facilitate the comparison with observations.

Four global models with better forecast performances (see Section 3 for details) were then selected to identify the key factors for extremely heavy rainfall using ESA. We calculated the area-averaged 24-h accumulated rainfall over the focused regions (32.5°–36.5°N, 111°–115°E, the inner box in Figure 1a) from 1200 UTC 19 July to 1200 UTC 20 July as the forecast metric (P). The Pearson correlation coefficient (R) was used to measure the correlation between the forecast metrics and the variables of interest (X) at different forecast times and pressure levels, and was calculated as follows (Hakim and Torn, 2008):

$$R = \frac{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)\left(P_i - \bar{P}\right)}{\sqrt{\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \sum_{i=1}^{n}\left(P_i - \bar{P}\right)^2}}, \tag{1}$$

where the overbar represents the ensemble mean and n is the ensemble size (73 herein combining the 4 selected models).

2.3 Configurations and experiment design of the regional convection-permitting ensemble

High-resolution convection-permitting ensemble simulations using the Advanced Research WRF (ARW/WRF; Skamarock et al., 2021) dynamical core were used to examine the intrinsic predictability and error growth in forecasting this event. Three one-way nested domains using the ARW/WRF model, version 4.2, are configured with horizontal grid spacings of 27, 9, and 3 km, and 210×130, 340×280, and 301×301 horizontal grids, respectively. There are 51 hybrid terrain-pressure levels, and the upper-most level is located at 50 hPa. Physical parameterization schemes are selected after trial-and-error tests, including the aerosol-aware Thompson and Eidhammer (2014) microphysics scheme, modified Tiedke cumulus scheme (Zhang and Wang 2017; only applied in the 27-km domain), revised MM5 scheme for surface layer processes (Jiménez et al., 2012), thermal diffusion scheme for land surface processes, Yonsei University PBL scheme (Hong et al., 2006), and RRTMG schemes for longwave and shortwave radiation (Iacono et al., 2008).

In order to examine the influence of the amplitude and scale of the initial uncertainties on the intrinsic predictability of this event, a total of four ensemble forecasts, each containing 40 ensemble members that run from 0600 UTC 19 July to 1200 UTC 20 July, are designed. Two of them incorporate initial uncertainties from relatively large scales. We first derive the perturbations from the 20-member 0.5°×0.5° Global Ensemble Forecast System (GEFS) analyses valid at 0600 UTC 19 July and the 20-member GEFS 6-h forecasts from 0000 UTC 19 July (also valid at 0600 UTC 19 July) by subtracting their respective ensemble mean from each of the respective 20 members. Temperature, water vapor mixing ratio, and the two components of the horizontal wind are processed. Then, the 40 perturbations are scaled by a factor of 0.1 and added to the GFS analysis valid at 0600 UTC 19 July to generate 40 initial conditions (ICs) with uncertainties that are an order of magnitude smaller than current global model analysis uncertainties, which is necessary because intrinsic predictability examines error growth mechanisms resulted from tiny, unobservable initial uncertainties. These 40 ICs are used to initialize the “LARGE” ensemble forecast. The initial perturbations of the LARGE ensemble are further scaled by a factor of 0.1 (therefore a 0.01 factor from their original values) to form the ICs that initialize the “LARGE0.1” ensemble forecast.

The other two ensembles contain initial uncertainties that are concentrated at smaller scales. To facilitate this purpose, we first run a short-term deterministic forecast from the GFS analysis valid at 0600 UTC 19 July 2021 using a configuration of model domains that cover exactly the same region but using horizontal grid spacings that are 1/3 of their original values (i.e., 9, 3, and 1 km for the three domains). Then, values of temperature, water vapor mixing ratio, and the two horizontal components of the wind at each grid point of the original model domains are replaced by randomly, non-repetitively selected values from the adjacent 3×3 grid points

Table 1 Descriptions of TIGGE models used in the study[a)]

	Model	Original resolution (km)	Ensemble size
1	*BoM*	*30–45*	*17*
2	CMA	50	30
3	DWD	40	40
4	ECCC	39	21
5	ECMWF	16/32 (after day 10)	51
6	IMD	12	21
7	JMA	139	51
8	*KMA*	*33*	*26*
9	Météo France	7.5–37	35
10	NCEP	25	31
11	*NCMRWF*	*13*	*12*
12	*UKMO*	*21*	*18*

a) More details about the TIGGE models can be found at https://confluence.ecmwf.int/display/TIGGE/Models. The models in italics are used by ESA.

in the higher-resolution 9-3-1-km simulation, similar to the generation of initial perturbations of Zhang et al. (2016). Since each grid point in the original model domain corresponds to 8 surrounding grid points in the 9-3-1-km domain (excluding the grid point that are collocated), each 9-3-1-km simulation output can be used to generate 8 different perturbations. Five model outputs from 0655 to 0700 UTC 19 July 2021 (when small-scale structures are sufficiently developed while no significant precipitation occurs) from the 9-3-1-km simulation, each 72 s apart, are used to generate 40 ICs (8 for each output) that contain uncertainties that represent flow-dependent features that the original model resolutions are not able to resolve, and these ICs are used to initialize the "SMALL" ensemble forecast. Similar to the LARGE0.1 ensemble, the initial perturbations of the SMALL ensemble are also multiplied by 0.1 to initialize the "SMALL0.1" ensemble forecast. Although the perturbations in SMALL and SMALL0.1 are drawn from simulation outputs close to 0700 UTC, they are nonetheless added to the GFS analysis valid at 0600 UTC 19 July, consistent with LARGE and LARGE0.1.

2.4 Description of CNOP and its experiment design

CNOP is the initial perturbation that maximizes the cost function under certain initial constraint conditions (Mu and Duan, 2003). The cost function is defined as $J(\delta x_0) = M(x_0 + \delta x_0) - M(x_0)$, and the initial perturbation δx_0^* is called CNOP, if and only if $J(\delta x_0^*) = \max_{\delta x_0 \le \beta} J(\delta x_0)$. M is the nonlinear operator. x_0 is the state vector x at the initial time and the $M(x_0)$ represents the value of x at forecast time t. β is used to constrain the values of the initial perturbations.

The norm used to constrain the cost function and the initial perturbations is the total moist energy (TME) norm (Ehrendorfer et al., 1999), which is calculated as follows:

$$\mathrm{TME} = \frac{1}{2}\left(u^2 + v^2 + \frac{C_\mathrm{p}}{T_\mathrm{r}} T^2 + \frac{L^2}{C_\mathrm{p} T_\mathrm{r}} q^2 + R_\mathrm{a} T_\mathrm{r} \left(\frac{P_s}{P_\mathrm{r}} \right)^2 \right), \tag{2}$$

where C_p(1005.7 J kg^{-1} K^{-1}) is the specific heat at constant pressure, T_r(270 K) is the reference temperature, R_a(287.04 J kg^{-1} K^{-1}) is the gas constant of dry air, L(2.5104×10^6 J kg^{-1}) is the latent heat of condensation per unit mass, P_r(1000 hPa) is the reference pressure, and u, v, T, q and P_s are the two horizontal wind components, temperature, water vapor mixing ratio, and surface pressure, respectively. The sensitive area, in which area the weather systems can be regarded as the key weather systems to the heavy rainfall (Yu and Meng, 2016, 2022), is defined as the location of the top 1% vertically integrated TME of the entire model simulation domain.

In this study, the CNOP is calculated based on the MM5 model (Grell et al., 1995) and its tangent linear and adjoint models (Zou et al., 1997) using the spectral projected gradient 2 (SPG2; Birgin et al., 2001) optimization algorithm. The model domain has 90×65 horizontal grids with a horizontal resolution of 60 km and 21 terrain-following levels in the vertical from the surface to 50 hPa. The initial and boundary conditions are provided by the National Centers for Environmental Prediction (NCEP) final analysis (FNL) of 1°×1° at a 6-h interval. The large-scale precipitation scheme, the Anthes-Kuo cumulus parameterization scheme, and the bulk planetary boundary layer scheme are used. In order to reveal the sensitive areas of the extreme hourly rainfall at 0800 UTC 20 July, the starting and ending times are 0600 UTC 19 July and 0800 UTC 20 July 2021, respectively. The verification area covers the location of the heavy rainfall (the inner green box in Figure 2a).

In order to examine the evolution of initial perturbations with different structures and their impact on the rainfall forecast, the starting time of CNOP is the same as that in the convection-permitting ensemble forecast experiments. In detail, the perturbations of CNOP and three random members from the LARGE ensemble are added to the GFS analysis at 0600 UTC 19 July 2021 as initial conditions to calculate the perturbation development using the WRF model with the same physical parameterization schemes as those used in the LARGE ensemble, except for using a domain coverage and horizontal grid spacing the same as those for the MM5 model used to calculate the CNOP. The perturbations of the three random members from the outer-most domain of the high-resolution LARGE ensemble (see Section 2.3) are interpolated to the CNOP model grid, and the magnitude of the CNOP perturbations is scaled down to be the same as the LARGE perturbations in terms of the mean TME in the area of interests (the black inner box in Figure 2a).

3. Key environmental factors related with the forecast uncertainties of the rainfall

3.1 Evaluation of rainfall forecast in TIGGE ensembles

Due to the relatively coarse horizontal resolutions of the global models (Table 1), the TIGGE ensemble forecast generally underestimates the rainfall amount (Figure 3a). Nevertheless, BoM, NCMRWF, UKMO, and KMA stand out from the 12 TIGGE models with higher P (Figure 3a) and are utilized for further ESA. Ensemble mean rainfall of the 4 models generally reproduces the rainfall distribution at the threshold of 50 mm (Figure 1b), with a much smaller maximum of 186.88 mm located more to the south compared with observation (Figure 1a). In terms of forecast skills, KMA has the highest equitable threat score (ETS; Wilks 1995) at thresholds of 100 mm and 150 mm, while NCMRWF still retains some skills at higher thresholds such

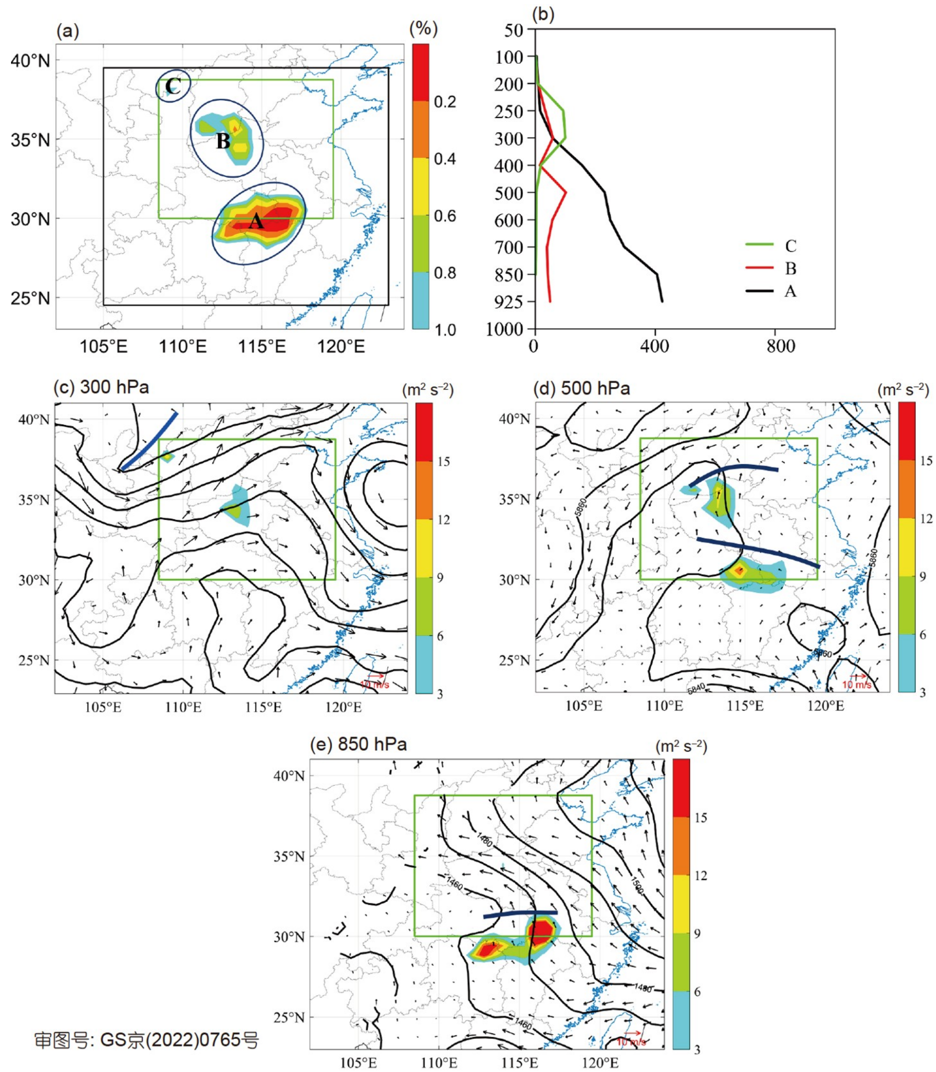

Figure 2 (a) Sensitive areas (shading) identified by the CNOP and (b) the vertical distribution of hTME (units: $m^2\ s^{-2}$) over sensitive area A (black line), B (red line) and C (green line) in panel (a). Wind (vector; units: $m\ s^{-1}$), Z (contour; units: gpm), and TME (shading; units: $m^2\ s^{-2}$) of CNOP at (c) 300 hPa, (d) 500 hPa, and (e) 850 hPa. The blue line in (c) denotes the trough extending from the cold vortex, and the blue line in (d) and (e) denotes the shear line from the low vortex. The inner green square box in (a), (c), (d), (e) denotes the verification area, and the inner black box in (a) denotes the area for the scale analysis.

as 250 mm and 300 mm (Figure 3b–3e).

Typical members with good and poor rainfall forecasts are selected based on ETSs of individual members (Appendix Figure S1, https://link.springer.com) and subjective comparison with the observed rainfall pattern. Members 01, 17, 49, and 71 are eventually chosen as good members, while members 35, 42, 72, and 40 are chosen as poor members (Figure S2). Comparisons are then performed between good and poor members to obtain more physical insights into the correlation patterns of the ESA.

3.2 Results from ensemble-based sensitivity analysis

Synoptic circulation systems in mid-troposphere are highly correlated with the extremely heavy rainfall. There are prominent negative correlations between 500-hPa Z and P in central China, especially in northern Henan Province from 0600 UTC 19 July to 1200 UTC 20 July, with the strongest negative correlation of ~−0.8 occurring at 1800 UTC 19 July (Figure 4a, 4d). Consistently, the composite of good members is characterized by a deeper mid-level low compared with that of poor members (Figure 4e, 4f). This result suggests that the mid-level low significantly contributes to a large rainfall accumulation. We thus choose 1800 UTC 19 July to examine the key environmental forcing for the 24-h accumulated rainfall. Figure 4a also shows positive correlation over the subtropical high as well as the ridge region near the southwest periphery of the subtropical high, but with a much smaller area with a confidence level of 95% and above. Together with the comparison on 500 hPa between typical members (Figure 4b, 4c), these results consistently suggest that a stronger subtropical high with a deeper ridge

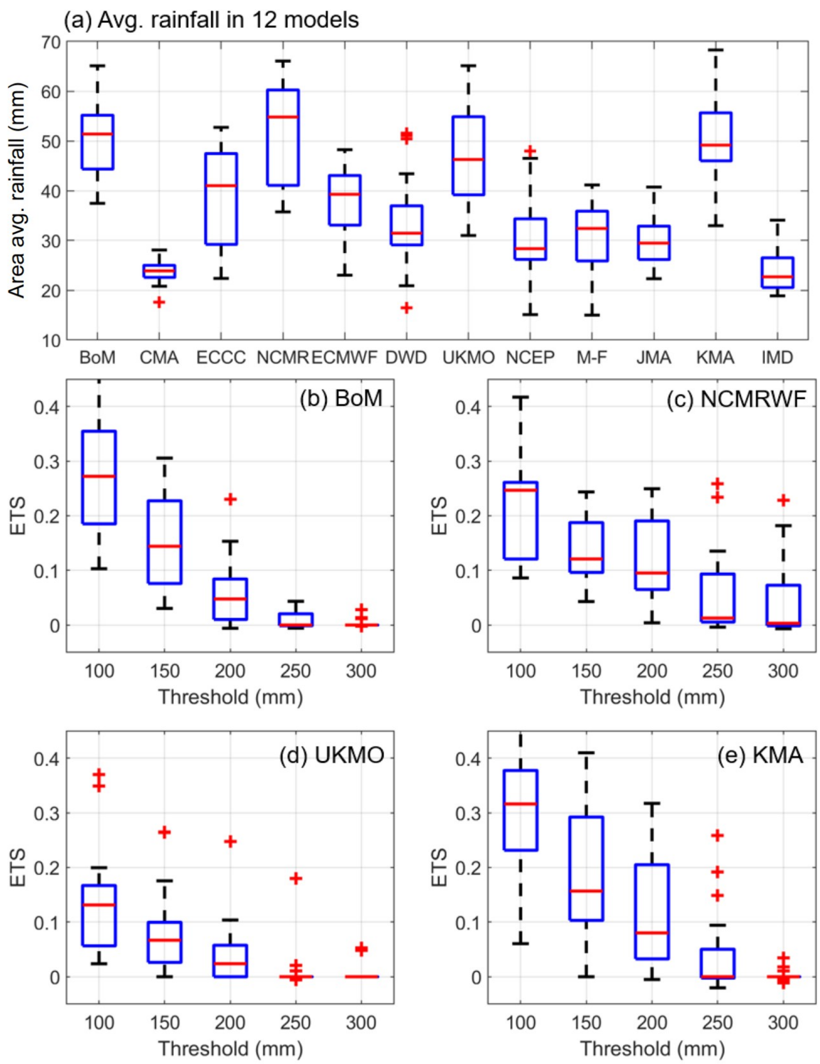

Figure 3 Quantitative precipitation evaluation of ensemble forecasts. (a) Boxplot of area-averaged 24-h accumulated rainfall of ensemble members from the 12 TIGGE models. NCMRWF is denoted by its first four letters (NCMR) owning to the limited space on the *x* axis. Boxplot of ETSs of ensemble members from (b) BoM, (c) NCMRWF, (d) UKMO, and (e) KMA models at different thresholds from 100 to 300 mm over the focused region. The whiskers extend to the most extreme data points that are not considered outliers. Points are identified as outliers if they are larger than $q_3+1.5(q_3-q_1)$ or smaller than $q_1-1.5(q_3-q_1)$, where q_1 and q_3 are the 25th and 75th percentiles.

on its southwestern flank is favorable for a larger rainfall accumulation, possibly through preventing the low pressure vortex from rapidly moving to the east.

The upper-level jet stream also plays an essential role in the extremely heavy rainfall process. In the correlation map between 200-hPa Z and P (Figure 5a), there are significant positive correlations of ~0.4 to the north of Henan Province and negative correlations of ~−0.3 downstream. This result indicates that a deeper ridge and trough on 200 hPa, namely, a wavier upper-level circulation is beneficial for the rainfall accumulation. The deeper ridge is associated with a stronger upper-level northwesterly jet stream to the north of the focused region (Figure 5b), which may provide more sufficient upper-level divergence that favors the heavy rainfall process.

The correlation map between 850-hPa Z and P (Figure 5c, 5e) is generally similar to that on 500 hPa (Figure 4a, 4d), reinforcing that lower geopotential height (herein a deeper low-level trough) over the focused region and a stronger subtropical high to the east are favorable for the rainfall accumulation. In the correlation map between 850-hPa horizontal wind speed (Figure 5d, 5f), there are positive correlations of ~0.6 to the south and east of the focused region, as well as over the northern periphery of Typhoon In-Fa. This result implies that the southerly and southeasterly low-level jets upstream of the focused region, which could be strengthened by the warm ridge extending southwestward from the subtropical high (Ran et al., 2021), are essential during the extremely heavy rainfall process by providing abundant moisture from the south.

The locations of the two tropical cyclones are remotely relevant to the heavy rainfall accumulation. The difference in the composite 850-hPa relative vertical vorticity between

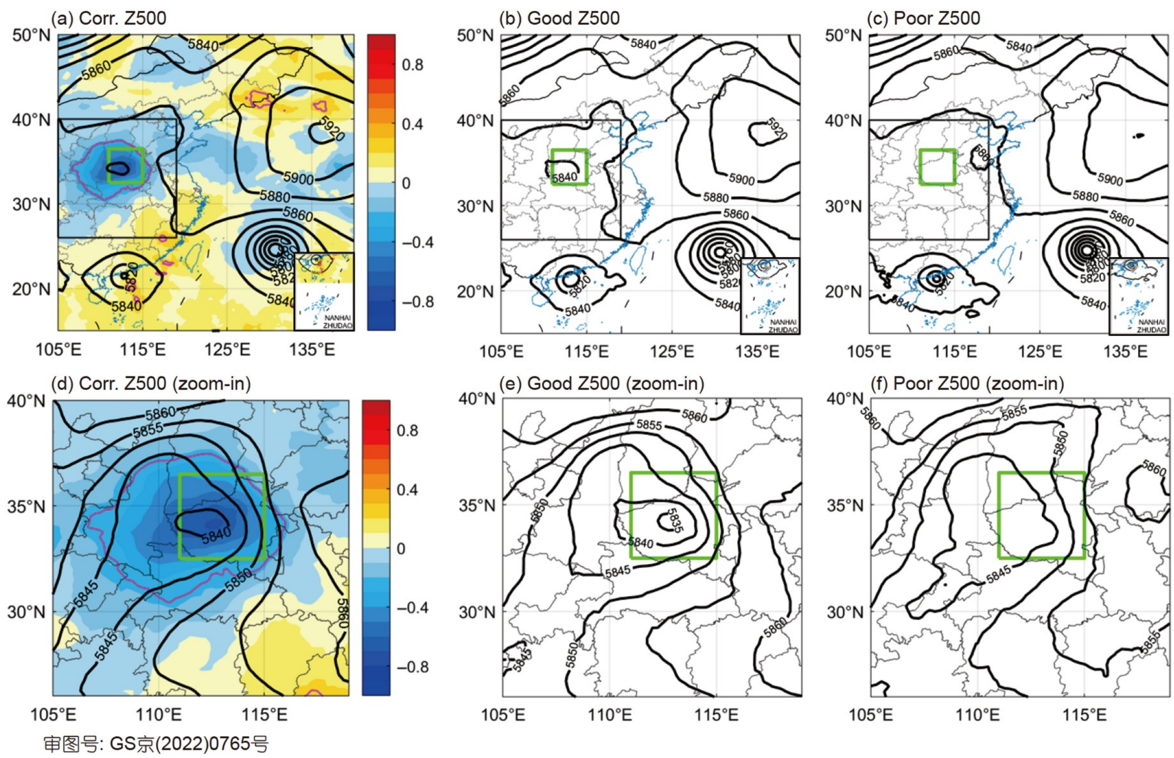

Figure 4 (a) Correlation coefficients (shading; magenta contours for 95% confidence) between the 500-hPa *Z* at 1800 UTC 19 July and the area-averaged 24-h accumulated rainfall, with ensemble mean 500-hPa *Z* at respective time contoured in bold black (in gpm); the green box denotes the focused region same as the black one in Figure 1; the grey box denotes the spatial range of (d)–(f). (b) Composite of good members on 500-hPa Z at 1800 UTC 19 July contoured in black (in gpm). (c) is the same as (b) but for composite of poor members. (d)–(f) are the same as (a)–(c) but zoomed in over the grey box given in (a)–(c).

good members at 1800 UTC 19 July are given in Figure 6a. Compared with poor members, the good members are characterized by stronger vorticity over the focused region, which is corroborated by the deeper low-level vortex suggested in Figure 5c. Moreover, there are dipoles of vorticity differences over two tropical cyclones (Figure 6a), and the dipoles become stronger (~6×10^{-5} s^{-1}) later on at 0000 UTC 20 July (Figure 6b), indicating the locations of tropical cyclones are associated with the rainfall accumulation over the focused region. Consistently, the composite of good members is featured with a Cempaka located more to the southeast and an In-Fa located slightly more to the south, which is especially true at 0000 UTC 20 July in Figure 6b. These variations of tropical cyclone locations may be closely related to the southwestward intrusion of the subtropical high. The southwestward intrusion of the subtropical high may increase the geopotential height gradient over the northern area of Typhoon In-Fa, which enhances the low-level easterlies as revealed in Figure 5d and thus facilitates more moisture transportation to the focused region.

3.3 Results from CNOP

Key environmental forcings for the 24-h accumulated rainfall identified using ESA are generally consistent with those for the hourly extreme rainfall at 0800 UTC 20 July identified using CNOP. Three sensitive areas are identified by the CNOP (Figure 2a). The vertical distribution of the horizontally integrated TME of CNOP (hereafter referred to as hTME) over the sensitive area B, which is located in the middle of the verification area, peaks at the middle (~500 hPa) and upper level (~300 hPa, Figure 2b). This sensitive area corresponds to the low pressure vortex and its associated shear line at 500 hPa (Figure 2d) and the ridge at 300 hPa (Figure 2c). The hTME over the sensitive area A, which is located to the south of the verification area (Figure 2a), peaks at lower (~850 hPa) and middle levels (~500 hPa, Figure 2b). This sensitive area corresponds to the southeasterly flow to the south of the shear line extending from the low vortex at 850 hPa (Figure 2e) and 500 hPa (Figure 2d). The environmental systems in sensitive areas A and B identified by CNOP are consistent with those identified by ESA. The CNOP also identifies the sensitive area C in the northwest of the verification area (Figure 2a) at ~300 hPa (Figure 2b), which corresponds to the westerly trough (Figure 2c) neighboring the ridge.

4. Error growth features and their sensitivities to different scales, amplitudes, and structures

All four convection-permitting ensemble forecasts show

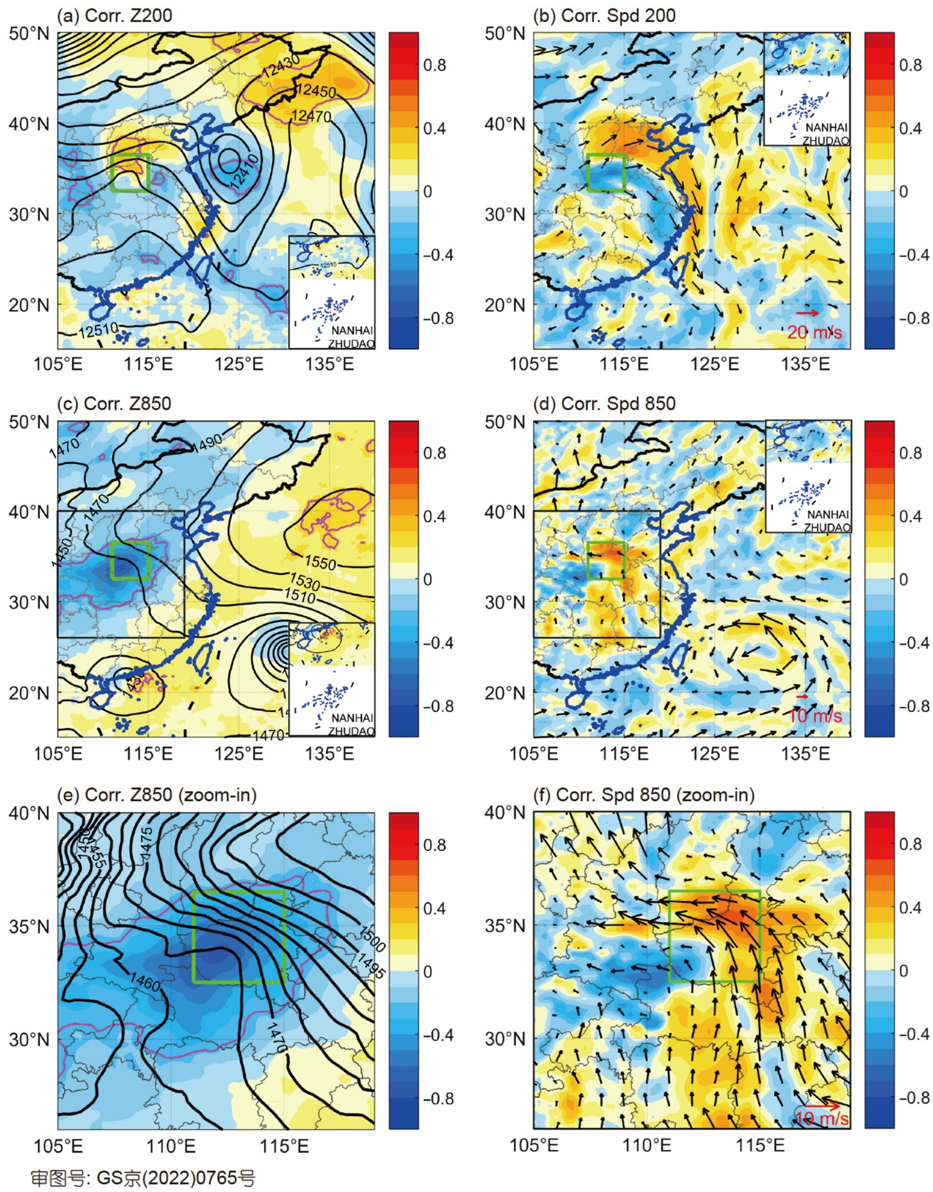

Figure 5 (a) Correlation coefficients (shading; magenta contours for 95% confidence) between the 200-hPa *Z* at 1800 UTC 19 July and the area-averaged 24-h accumulated rainfall, with ensemble mean 200-hPa *Z* at respective time contoured in bold black (in gpm). (b) Correlation coefficients (shading) between the 200-hPa horizontal wind speed at 1800 UTC 19 July and the area-averaged 24-h accumulated rainfall, with ensemble mean 200-hPa horizontal wind vector at respective time (reference vector given in bottom right). (c) and (d) are the same as (a) and (b) but for 850 hPa, where the grey box denotes the spatial range of (e) and (f). (e) and (f) are the same as (c) and (d) but zoomed in over the grey box given in (c) and (d).

very similar distribution, structure, and values of the accumulated rainfall as well as the uncertainties across their ensemble members, and higher accumulated rainfall amounts are generally collocated with greater uncertainties (Figure 7). On the one hand, the similarity of rainfall region across all ensemble forecasts suggests that the general location of where rainfall will occur is quite predictable. However, the large uncertainties of the 24-hour accumulated precipitation in these forecasts with minute initial perturbations, as well as the insensitivity of these forecast uncertainties to the spatial scale or amplitude of the initial perturbations, suggest that the predictability of the extreme rainfall during this event is intrinsically limited and highly unpredictable in a deterministic forecasting system.

4.1 Overall power spectrum and error growth features with respect to different scales and amplitudes

Although all ensemble forecasts show very similar uncertainties for their 24-hour accumulated precipitation forecasts as well as the power spectra of various state variables at the end of the ensemble forecast at 1200 UTC 20 July (figure

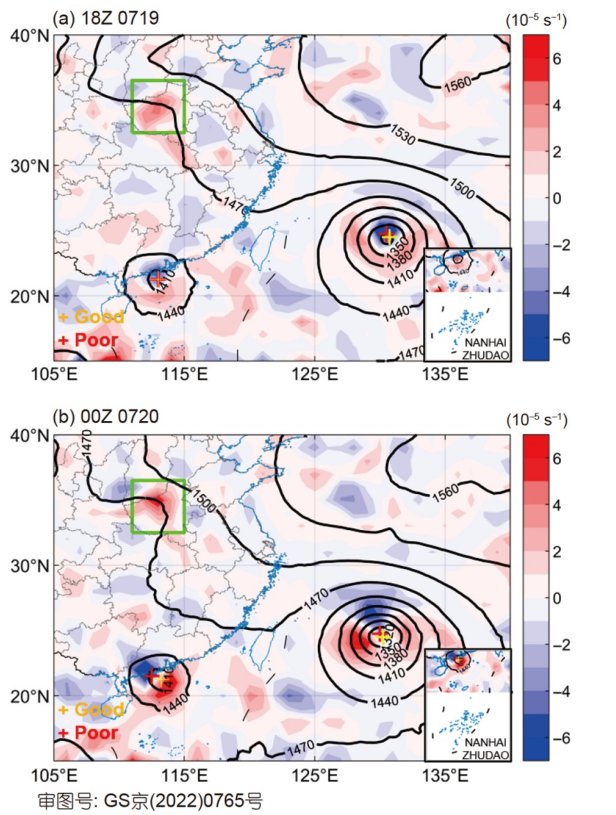

Figure 6 (a) Differences of 850-hPa vertical relative vorticity between composite of good members and poor members (Good − Poor) on 1800 UTC 19 July; ensemble mean 850-hPa Z is contoured in black (in gpm); locations of tropical cyclones (identified based on the minimum of 850-hPa Z) in the composite of good (poor) members were denoted by yellow (red) crosses. (b) is the same as (a) but for 0000 UTC 20 July. The green box indicates the focused region same as the black one in Figure 1.

not shown), the growth of the ensemble spread at the first several hours shows different characteristics associated with the spatial scale and amplitude of the initial perturbations, which is shown in Figure 8 for the energy spectra of the ensemble spread (simply "error energy" hereafter) of the U-wind component (the power spectra of temperature and water vapor mixing ratio are qualitatively similar and therefore omitted).

For LARGE and LARGE0.1, because of the 0.5°×0.5° horizontal grid spacing of the GEFS analyses that are used to generate their initial perturbations, most of the error energy is concentrated in relatively large scales, and the energy decreases rapidly for wavelengths below ~200 km (Figure 8a, 8b), consistent with the statement that the smallest resolvable features of a numerical model are roughly 4 to 6 times of its horizontal grid spacing (e.g., Skamarock, 2004). The missing error energy at shorter wavelengths is quickly filled as the simulation goes on. However, the error energy "plateau" at wavelengths longer than ~200 km does not increase for the first 3 to 4 hours (Figure 8a, 8b). The error energy at relatively large scales only starts to increase when the error energy at relatively small scales has grown to an amplitude that is comparable to the large-scale errors, and LARGE0.1 starts to increase slightly earlier than LARGE due to its smaller initial error energy (Figure 8b). If we look at how much error at each scale grows every hour by examining the ratios of the error energy spectra of two consecutive hours, it is clear that error growth is greater in smaller scales at earlier times, and shifts to larger scales at later times, for both LARGE and LARGE0.1 ensembles (Figure 8e, 8f). Furthermore, accompanying this shift from smaller to larger scales of error growth peaks, the amplitude of the peaks also gradually decreases as they move toward larger scales, suggesting a slower growth at larger scales (Figure 8e, 8f). The behavior of error growth of the LARGE and LARGE0.1 ensembles suggests that the most prominent growth of errors—even when only large—scale uncertainties are imposed at the initial conditions-occurs at smaller scales first, then gradually transitions to larger scales ("upscale growth"), and the speed of error growth at smaller scales is faster than later at larger scales, consistent with the three-stage error growth model of Zhang et al. (2007).

On the other hand, the SMALL and SMALL0.1 ensembles first show apparent adjustment from 0 to 1 h resulting from the unbalances in the sub-grid-scale initial perturbations. Unlike the error energy spectra of LARGE and LARGE0.1 that drastically decrease for wavelengths smaller than ~200 km, the error energy spectra of SMALL and SMALL0.1 at 0 h and 1 h are almost flat across the entire range of the wavelengths (Figure 8c, 8d); however, since larger scales have greater base energy than smaller scales, therefore the overall "flat" error energy spectra of SMALL and SMALL0.1 actually indicate that errors are more concentrated at smaller scales than larger scales, opposite to the power spectra of LARGE and LARGE0.1 that errors are more concentrated at larger scales. Compared with the growth of error energy of LARGE and LARGE0.1 (Figure 8a, 8b), the errors seem to be growing at all scales simultaneously for SMALL and SMALL0.1 (Figure 8c, 8d), similar to the error growth after 4 h in LARGE0.1 (Figure 8b, cyan color). However, if we look at the error growth ratios, there is also a shift of error growth peaks from the smaller scale with larger amplitude at earlier times to the larger scale with smaller amplitude at later times (Figure 8g, 8h), similar to what we have already observed for the LARGE and LARGE0.1 ensembles (Figure 8e, 8f), especially for SMALL0.1 which has smaller initial errors (Figure 8h).

In general, although the error grows up at all scales in some of the circumstances, the wavelengths at which the fastest error growth occurs shift upscales for initial condition perturbations from both small and large scales with different initial amplitudes. This remains true in sensitivity experiments that completely remove initial perturbations at scales

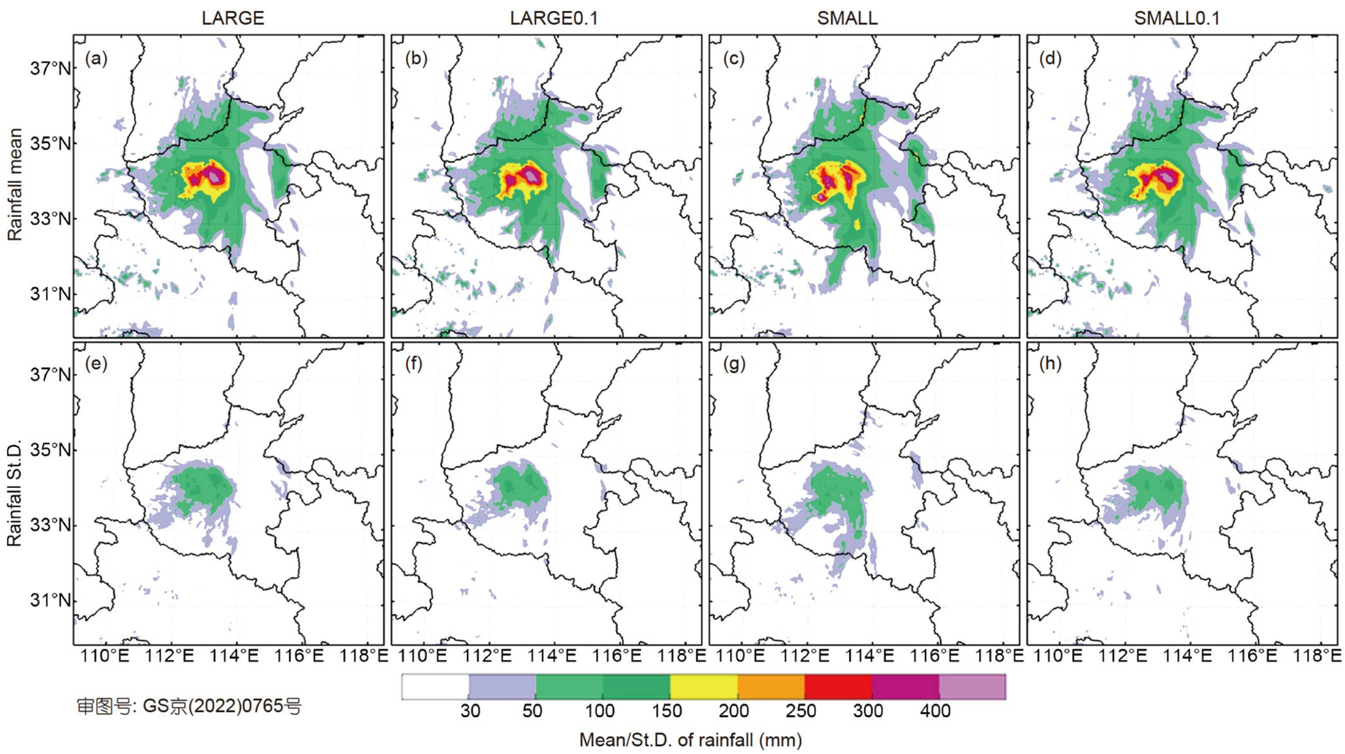

Figure 7 (upper panels) Ensemble probability-matched mean and (lower panels) ensemble standard deviation of 24-hour accumulated rainfall from 1200 UTC 19 July 2021 to 1200 UTC 20 July 2021 for the (first column) LARGE, (second column) LARGE0.1, (third column) SMALL, and (fourth column) SMALL0.1 ensemble forecasts.

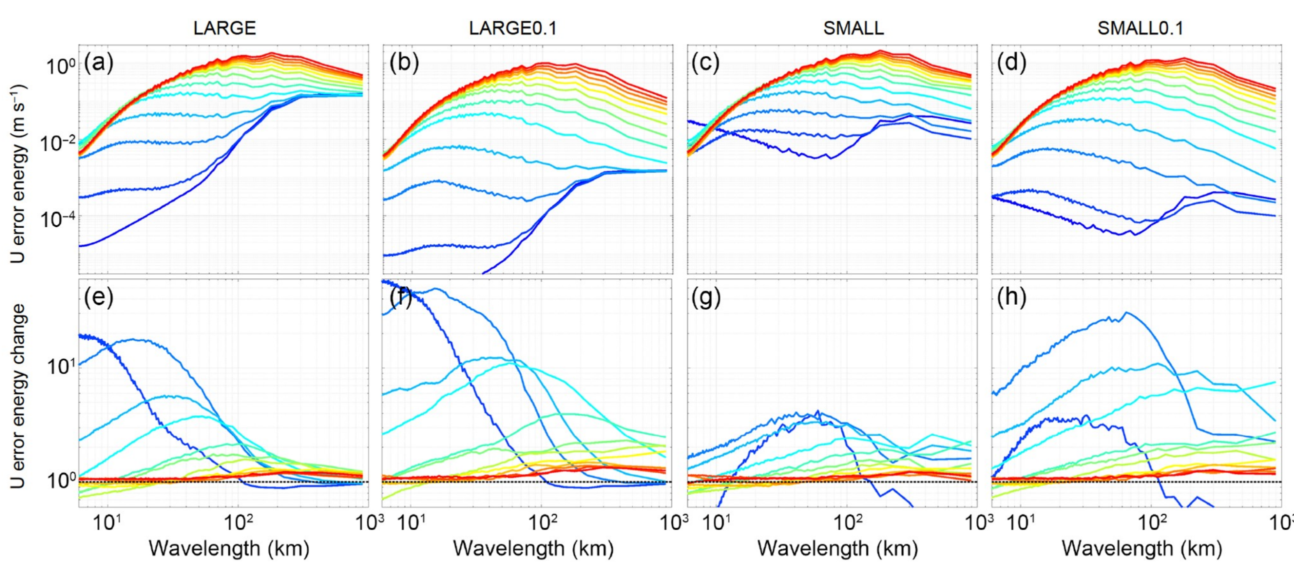

Figure 8 (upper panels) Hourly ensemble-mean power spectra and (lower panels) hourly growth of the ensemble-mean power spectra of the U-wind perturbations from 0600 UTC to 1800 UTC of 19 July 2021 (0 to 12 hours of the ensemble simulation) for the (first column) LARGE, (second column) LARGE0.1, (third column) SMALL, and (fourth column) SMALL0.1 ensemble forecasts. Blue colors denote earlier times (shorter simulation lengths) and red colors denote later times (longer simulation lengths).

smaller than 200 km of LARGE and SMALL ensembles, while another sensitivity experiment that only keeps SMALL's initial perturbations at scales smaller than 200 km shows that error at smaller scales can grow without larger-scale errors (see supplement Figure S3).

4.2 Impact of moist process on the power spectrum and error growth features with respect to different scales and amplitudes

Error growth rates in regions with and without precipitation are distinct due to the dominant role of moist convective processes in error growth at mesoscales (Zhang et al., 2007). Figure 9a shows the root-mean difference kinetic energy (RMDKE; e.g., Zhang et al., 2002) averaged over regions with ("moist" in Figure 9a) and without ("dry" in Figure 9a) precipitation, defined as the ensemble mean precipitation rate exceeding or lower than 10^{-6} mm h^{-1}, in the ensembles. The characteristics of temperature and moisture are generally the same as RMDKE. RMDKE is defined as

$$\mathrm{RMDKE} = \sqrt{\frac{\sum_i \frac{1}{2}\left(u_i'^2 + v_i'^2\right)}{n}}, \tag{3}$$

where u' and v' are the differences between an ensemble member and the ensemble mean for the two components of the horizontal wind, i is all the grid points within the moist or the dry region from all the 40 ensemble members, and n is the quantity of all i grid points. It is apparent from Figure 9a that the error growth rate for the first 6 to 8 hours is much faster in the moist region than in the dry region for all the ensembles, proving the critical role of moist convective processes in boosting the error growth at mesoscales. Furthermore, while the four ensembles contain initial perturbations from different scales with different amplitudes, the curves of their respective RMDKE, both in the moist and the dry region, are generally parallel with each other. This suggests that, at least for the four ensembles that we have examined for this event, the error growth mechanisms at the first 6 to 8 hours are likely independent of the scales and amplitudes (when they are already very small) of the initial perturbations.

We further decompose the horizontal winds into three different scales with partitions at 30 and 200 km, then categorize the decomposed wind components into moist and dry regions and examine how RMDKE grows with and without precipitation at different scales (Figure 9b). Similar to what we have observed in the temporal evolution of the error energy spectra (Figure 8), the distinctions of error growth rates in moist and dry regions are greater at smaller scales than at larger scales: the 0–6-h error growth rate in the moist region is 3.76 times of the dry region growth rate at the small scales (<30 km), while this ratio of moist-region versus dry-region error growth rate is 2.77 at the medium scale (30–200 km), and it becomes almost comparable in these two regions at the large scale (>200 km). The "stall" of error growth at the first several hours at the largest scales (Figure 8) is also apparent in Figure 9b that the large-scale RMDKE almost does not grow at the beginning of the forecasts, unlike the other two scales. The small-scale RMDKE at 6-hour lead-time is smaller than the medium-scale RMDKE (Figure 9b), because the fastest error growth scale moves beyond 30 km after about 2 h (Figure 8e–8h).

In short, ensemble forecasts show that the predictability of this rainfall event is intrinsically limited. Reducing initial error amplitudes will not lead to improved forecasts. No matter what the spatial scales and amplitudes the initial perturbations are, the error energy spectra have no difference after 6–8 h. This time scale is consistent with many other studies (e.g., Durran and Gingrich, 2014, Durran and Weyn, 2016). There is also an apparent upscale growth of errors, with errors at smaller scales growing faster, and errors grow faster in regions where precipitation occurs than that in the no-precipitation region due to the dominant role of moist convective processes in mesoscale error growth, both consistent with the three-stage error growth conceptual model of Zhang et al. (2007). Additionally, it is shown for the first time in peer-reviewed literatures that the error growth at larger scales depends on the smaller-scale errors and that larger-scale errors will not grow until smaller-scale errors have grown to an amplitude that is comparable to larger-scale errors, while error growth at smaller scales is independent of larger-scale errors. This suggests that the mechanism governing error growth of this event in our ensembles is primarily the upscale growth rather than the upamplitude growth.

4.3 Error growth features with respect to different structures

Larger error growth is observed in the CNOP than the LARGE (it should be noted that in this section, "LARGE" refers to the simulation with perturbations derived from the LARGE ensemble in Sections 4.1 and 4.2, rather than the LARGE ensemble itself) in the whole integration time from 0600 UTC 19 July to 1200 UTC 20 July (Figure 10a). The CNOP grows faster thans the LARGE in the first several hours, and the growth rates of the CNOP and the LARGE become similar after that (Figure 10b).

Similar vertical distributions are observed in CNOP and the 3 members of LARGE in the first several hours (Figure 11c), while the horizontal distributions are greatly different from each other. The sensitive areas identified by CNOP are corresponding to the key synoptic weather systems, while those of LARGE are not (Figure 11a, 11b; Figure 2c–2e). Large vertically-integrated TME of CNOP is generally collocated with the hourly precipitation simulation at both the initial hours (Figure 11d) and times after that (Figure 11g). The large perturbation developments at different vertical levels correspond to the key synoptic weather systems that are associated with the rainfall at the whole integration time, which are the low-level water vapor convergence area, the mid-level low and the upper-level divergence area (Figure 12a, 12c, 12e, 12g, 12i, 12k). However, at the first several hours in LARGE (~10 h), the large vertically-integrated TME is not quite consistent with the simulated hourly precipitation (Figure 11e) and the large perturbation developments at different vertical levels are not associated with the key synoptic weather systems as good as in CNOP (Figure

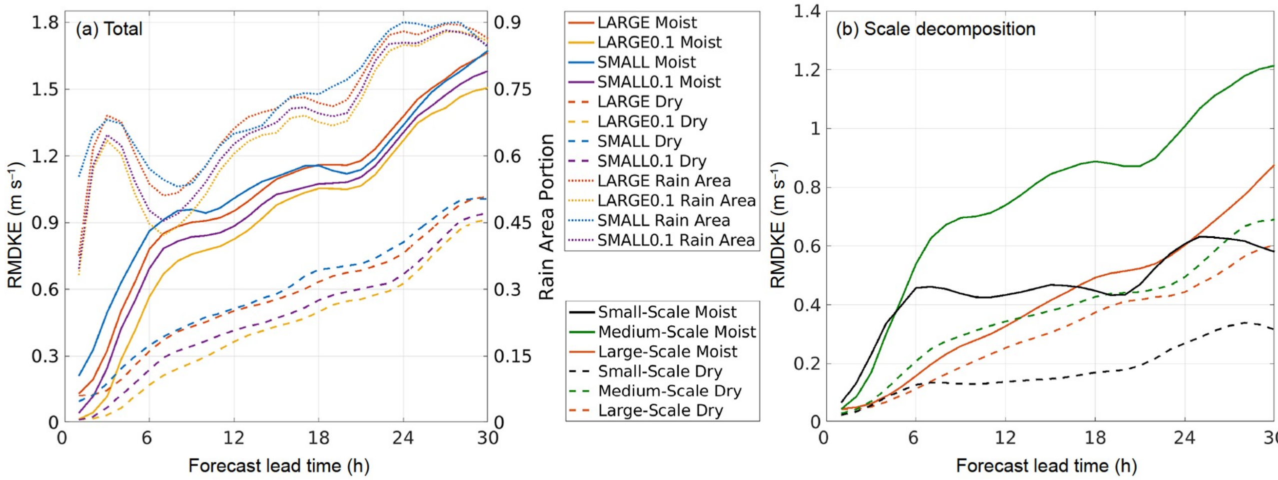

Figure 9 RMDKE of (a) the original model forecast fields, and (b) the model forecast fields decomposed into small scales (<30 km), medium scales (30–200 km), and large scales (>200 km) for regions with (solid lines) and without (dashed lines) precipitation. The portion of the model domain that is covered by precipitation is also plotted in (a) as the dotted lines. The lines in (b) are the averages of respective, scale-decomposed RMDKE values of the four ensembles (i.e., LARGE, LARGE0.1, SMALL, and SMALL0.1).

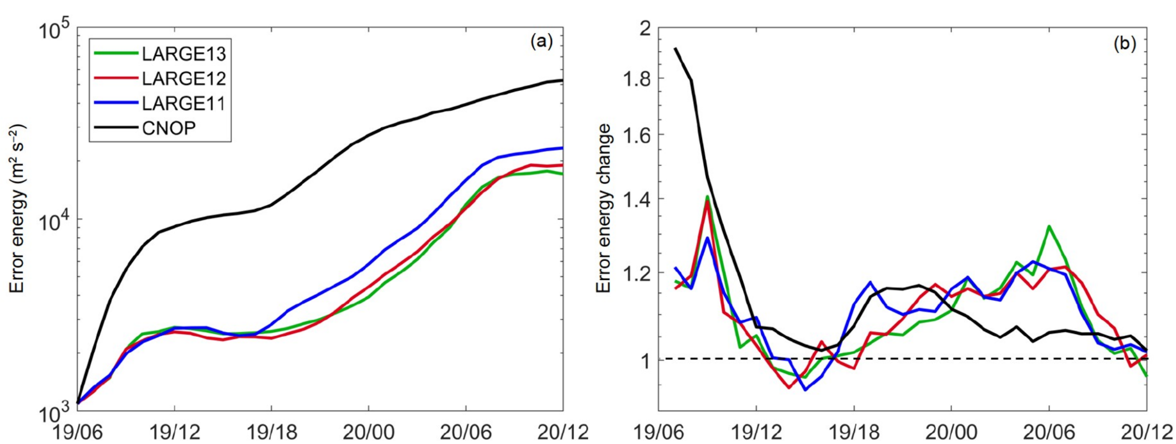

Figure 10 (a) TME (units: $m^2 s^{-2}$) and (b) hourly growth of the TME over the area interested (the inner black box in Figure 2a) for the CNOP (black line) and LARGE (red, blue and green lines).

12b, 12f, 12j). After the first several hours, the LARGE development patterns become similar to the CNOP (Figure 11h) and are better corresponding to the key synoptic weather systems mentioned above (Figure 12d, 12h, 12l).

Faster error growth at smaller scales than that at larger scales is also observed in CNOP and LARGE in the first several hours (Figure 13). However, different error growth features are found at larger scales in different variables for these two types of perturbations with different structures. While the characteristics of error growth of temperature in both CNOP and LARGE forecasts are similar to those observed in the previous subsection that larger scale errors stall when smaller scale errors grow for the first few hours ("upscale"; Figure 13e–13h), error growth of Q_v in both CNOP and LARGE forecasts are more uniform ("upamplitude"; Figure 13i, 13j), although smaller scale errors grow slightly faster than larger scale errors at the beginning (Figure 13k, 13l). On the other hand, error growth of the wind (the U-component and V-component are similar) shows different behavior in the two forecasts: it is more upamplitude with the CNOP initial perturbations (Figure 13a), while more upscale with the LARGE initial perturbations (Figure 13b). This result suggests that the error growth of wind is more sensitive to the structure of initial perturbations than those of temperature and Q_v.

The reasons for the more upamplitude features are two folds related to the structure of the initial perturbation and the wavelength regime used for forecasts. On the one hand, the more upamplitude features in wind and Q_v of the CNOP may be contributed partly by the inherent faster error growth as-

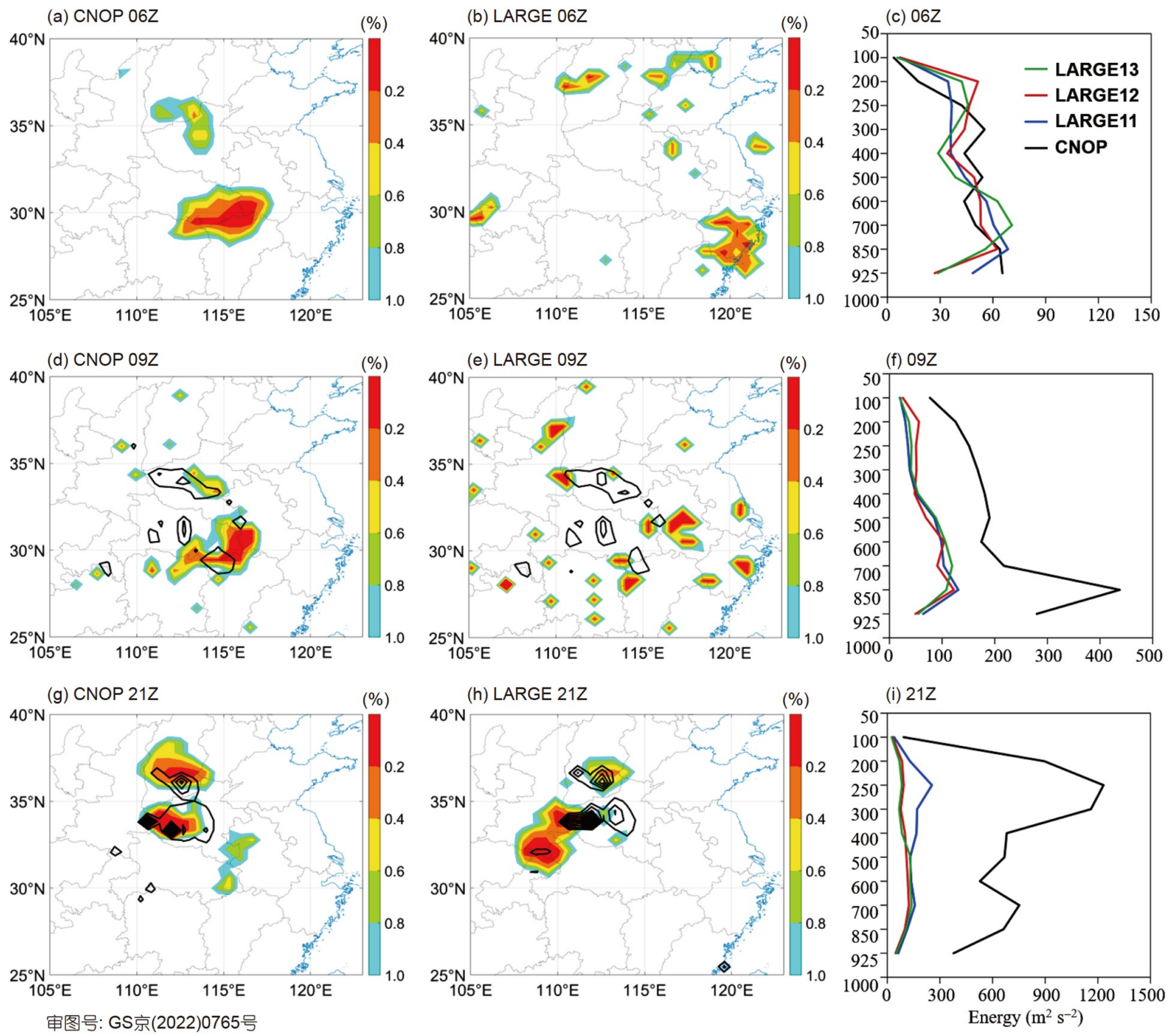

Figure 11 Development of the CNOP and LARGE at (upper panels) 0600 UTC, (middle panels) 0900 UTC, and (bottom panels) 2100 UTC 19 July 2021. The vertical integrated TME (shading, the top 1% energy) for (first column) CNOP and (second column) LARGE_11, and (third column) the vertical distribution of the hTME (units: $m^2 s^{-2}$) for CNOP and LARGE are shown. The simulated 1-h precipitation (contour, every 2 mm) at the corresponding time are also shown in (d), (e), (g), (h).

sociated with the large-scale flow patterns that are well collocated with rainfall. On the other hand, the distribution of atmospheric kinetic energy with respect to wavelengths has already transitioned from a $k^{-5/3}$ power law at the smallest scales of the CNOP and LARGE forecasts to a k^{-3} power law at the largest scales of these forecasts (e.g., Skamarock, 2004), and Rotunno and Snyder (2008) and Durran and Gingrich (2014) show that error grows more upscale in the $k^{-5/3}$ regime while more upamplitude in the k^{-3} regime. Additionally, Skamarock (2004) shows that forecasts with parameterized convection (like the CNOP and LARGE forecasts) are not able to build the $k^{-5/3}$ energy spectrum and hinders error growth at smaller scales compared with forecasts with explicit convection, which may enhance the upamplitude tendency for the forecasts. Therefore, mixed behavior of both upscale and upamplitude error are observed in the three variables, with Q_v showing the strongest upamplitude characteristics as the convective parameterization scheme directly impacts it while it is more of an indirect impact on the temperature and wind through modified convective activities.

In short, CNOP has larger error growth at the whole integration time and a much faster growth rate at the first several hours than the LARGE. In addition, error growth tends to be more upamplitude in these coarse resolution forecast, especially with the CNOP. The error growth at larger scales may be related to both the inherited feature of CNOP perturbation, the inability of the convective parameterization scheme to rebuild the $k^{-5/3}$ atmospheric power spectra at the mesoscales, and different error growth char-

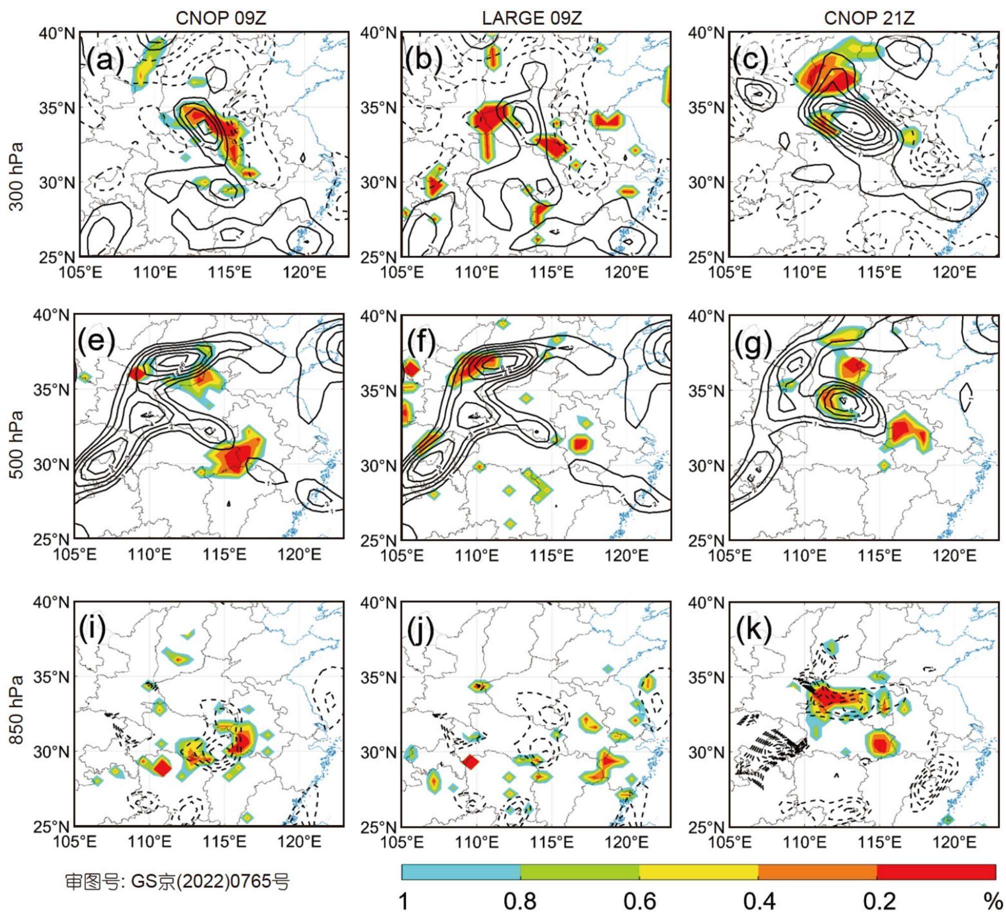

Figure 12 Development of the (first and third columns) CNOP and (second and fourth columns) LARGE_11 at (first and second columns) 0900 UTC and (third and fourth columns) 2100 UTC 19 July 2021. The TME (shading, the top 1% energy) at (upper panels) 300 hPa, (middle panels) 500 hPa, and (bottom panels) 850 hPa. Also shown are (upper panels) the horizontal wind divergence (contour, units: 10^{-5} s^{-1}) at 300 hPa, (middle panels) the vertical relative vorticity (contour, units: 10^{-5} s^{-1}) at 500 hPa, and (bottom panels) the water vapor flux divergence (contour, units: 10^{-4} kg (kg s)$^{-1}$) at 850 hPa for the (first and third columns) CNOP and (second and fourth columns) LARGE_11. The solid contour line denotes the divergence area at 300 hPa and positive vorticity at 500 hPa, and the dotted line denotes the convergence area at 850 hPa.

acteristics in the $k^{-5/3}$ and k^{-3} regimes.

5. Concluding remarks

This study explores the controlling factors of the uncertainties and error growth features with various initial scales, amplitudes, and structures in forecasting the high-impact extremely heavy rainfall event that occurred in Henan Province China, on 17–22 July 2021. The most intense events happened during 19–20 July 2021, when the metropolitan area of Zhengzhou, the capital city of Henan Province, and the surrounding area received record-breaking hourly rainfall of 201.9 mm and 24-hour accumulated rainfall of over 600 mm. In spite of warnings that were issued several days prior to this event, large uncertainties exist in operational forecasts in the location and intensity of the highest accumulated rainfall of this event. A suite of analyses, including ensemble sensitivity analysis (ESA) using an ensemble of global models, conditional nonlinear optimal perturbation (CNOP) method using a coarse-resolution regional model, and ensemble simulations using a high-resolution convection-permitting regional model, is designed in this study.

Using the four models that most accurately predicted the rainfall amount and location of this event from the TIGGE ensemble, ESA reveals several dominating synoptic features that determine the forecast uncertainties of this event. The most significant contributor is found to be the mid-to-lower low-pressure system directly over Henan Province. The upper-level deeper ridge and trough that are associated with a stronger jet stream are found to provide stronger upper-level divergence and hence stronger lifting and more favorable for heavy rainfall. In addition, the positions of the two tropical cyclones and the associated low-level jets are also important for rainfall. Likely being associated with the southwestward extending of a ridge from the subtropical high, when Typhoon Cempaka is more located to the southeast or Typhoon In-Fa is more located to the south, the low-level jets are enhanced, and so is the total amount of precipitation in He-

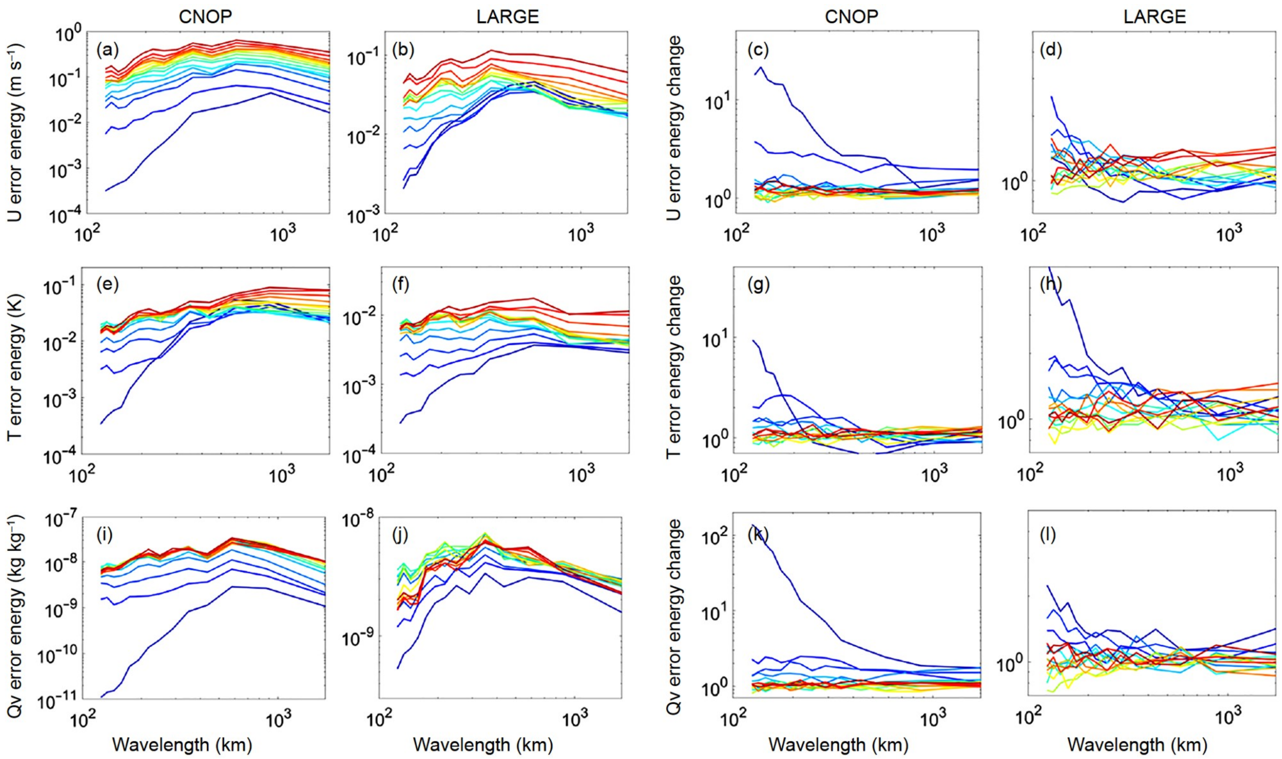

Figure 13 (first and second columns) Hourly power spectra and (third and fourth columns) hourly growth of the power spectra of (upper panels) the U-wind perturbations, (middle panels) the T perturbations, and (bottom panels) the Q_v perturbations from 0600 UTC to 2100 UTC 19 July 2021 (0 to 15 hours of the simulation) for the (first and third column) CNOP, (second and fourth column) LARGE_11. Blue colors denote earlier times (shorter simulation lengths) and red colors denote later times (longer simulation lengths).

nan Province. Similar to the 24-h accumulated rainfall, the hourly extreme rainfall at 0800 UTC 20 July is also sensitive to the upper-level ridge, mid-level low, and low-level trough extending from the low pressure vortex, as revealed by our CNOP analysis.

High-resolution convection-permitting ensemble forecasts with flow-dependent, unobservably small initial perturbations show that rainfall area is quite predictable, but the predictability of this event is intrinsically limited in terms of the maximum values of 24-hour accumulated precipitation. Reducing initial perturbations by order of magnitude will not lead to reduced forecast uncertainties, no matter the spatial scales of the initial perturbations are relatively large (from a global model) or small (from sub-grid-scale unresolved uncertainties).

The evolution of the energy spectra of the forecast errors is insensitive to the amplitudes or spatial scales, or structures of the initial perturbations after 6 to 8 hours. The intrinsically limited predictability of convectively driven extreme rainfall events is widely recognized, and this insensitivity of forecast errors to the amplitudes and spatial scales in initial perturbations is aligned with previous studies of other extreme rainfall events under different synoptic regimes, including Mei-yu rainfall (Bei and Zhang, 2007), warm-sector rainfall (Wu et al., 2020), frontal and pre-frontal rainfall (Weyn and Durran, 2019), and organized convective systems (Nielsen and Schumacher, 2016, Weyn and Durran, 2019). However, for the initial perturbation generated with GEFS or sub-grid-scale uncertainties, one outstanding discovery is the behavior of large-scale flow-dependent errors when they have much larger amplitudes than the small-scale errors: to the knowledge of the authors, this is the first study that shows the inability of large-scale errors to grow until the amplitude of small-scale errors have increased to an adequate amplitude, confirming that errors of smaller scale grow faster than those of larger scale.

In addition, the error growth rate with respect to different spatial scales and time—despite whether large-scale or small-scale initial uncertainties are imposed—also shows an apparent transfer of the fastest growing scale from smaller to larger scales with a slower growth rate at larger scales. This result suggests that although upamplitude growth and upscale growth coexist, the dominant mechanism controlling the error growth is their upscale transfer, at least for the ensemble forecasts of this high-impact event examined in this study. Faster error growth is also observed in regions where precipitation occurs, suggesting the importance of moist convective processes in controlling the error growth of this event. Whether this behavior of large-scale flow-dependent errors holds true for other events and how sensitive this behavior is to different strengths of synoptic forcing remain unknown and deserve further studies.

The sensitivity of the error growth to different structures of initial perturbations was also examined with the distribution of atmospheric kinetic energy transitioning from the $k^{-5/3}$ to k^{-3} regimes. Results show that CNOP has larger error growth at the whole integration time and a much faster growth rate at the first several hours than the GEFS or sub-grid-scale perturbations. Different error growth features at larger scales are observed in different variables for the perturbations with different structures. CNOP pattern initial perturbations, whose error growth well corresponds to the rainfall associated key synoptic weather systems at the whole integration time, show more upamplitude feature with an error growth at the initial hours at both smaller and larger scales for wind and water vapor mixing ratio. However, the error growth feature of temperature is not quite sensitive to the structure of initial perturbations. The error growth at larger scales may be owning to the inherited feature of CNOP perturbation, the inability of the convective parameterization scheme to rebuild the $k^{-5/3}$ power spectra at the mesoscales, and different error growth characteristics in the $k^{-5/3}$ and k^{-3} regimes.

To conclude, this study suggests that the forecast uncertainties of the record-breaking extreme rainfall event that occurred in Henan Province China on 19–20 July 2021 are associated with many different factors across different spatial scales. Practically, because of incomplete knowledge of the atmosphere, model deficiencies, and imperfect data assimilation techniques, initial conditions of different models disagree in terms of their representations of the upper-level ridge and trough, the mid-level low-pressure system directly over Henan Province, as well as the low-level jet associated with the warm ridge of the subtropical high and the two distant tropical cyclones to the southeast, which leads to diverse forecasts of the total accumulated rainfall. However, even we have a nearly perfect model with nearly perfect estimations of the atmospheric conditions, tiny, unobservable errors will grow upscale and, to a slightly lesser extent, upamplitude with the help of moist convective processes, and intrinsically prevent the accurate predictions of the location and strength of the accumulated rainfall in a deterministic sense. Although the universality of some of these conclusions needs to be further examined under different scenarios, they nonetheless highlight the importance of further developing advanced data assimilation techniques that can make better use of existing but underutilized observations, as well as the benefits of ensemble forecasts that consider uncertainties in initial conditions over deterministic forecasts, in improving practical predictability of extreme weather events and providing more useful numerical weather predictions as forecast guidance in the future.

Acknowledgements *The WRF ensemble forecasts were performed on the Stampede 2 supercomputer of the Texas Advanced Computing Center (TACC) through the Extreme Science and Engineering Discovery Environment (XSEDE) program supported by the NSF. The CNOP simulations were performed on the Tianhe Supercomputer at the National Supercomputing Center of Tianjin, China. This work was supported by the National Natural Science Foundation of China (Grant Nos. 42030604, 41875051), the National Science Foundation (Grant No. AGS-1712290), and China Postdoctoral Science Foundation (Grant No. 2021M702725). Dr. Murong Zhang is sponsored by the MEL Outstanding Postdoctoral Scholarship from Xiamen University. Dr. Huizhen Yu is also supported by the East China Regional Meteorological Science and Technology Collaborative Innovation Fund (Grant No. QYHZ201801) and the project from the Qingdao Meteorological Bureau (Grant No. 2021qdqxz01).*

References

Bei N, Zhang F. 2007. Impacts of initial condition errors on mesoscale predictability of heavy precipitation along the Mei-yu front of China. Q J R Meteorol Soc, 133: 83–99

Birgin E G, Martínez J M, Raydan M. 2001. Algorithm 813: SPG—Software for convex-constrained optimization. ACM Trans Math Softw, 27: 340–349

Bougeault P, Toth Z, Bishop C, Brown B, Burridge D, Chen D H, Ebert B, Fuentes M, Hamill T M, Mylne K, Nicolau J, Paccagnella T, Park Y Y, Parsons D, Raoult B, Schuster D, Dias P S, Swinbank R, Takeuchi Y, Tennant W, Wilson L, Worley S. 2010. The THORPEX interactive grand global ensemble. Bull Amer Meteorol Soc, 91: 1059–1072

Cressman G P. 1959. An operational objective analysis system. Mon Weather Rev, 87: 367–374

Durran D R, Gingrich M. 2014. Atmospheric predictability: Why butterflies are not of practical importance. J Atmos Sci, 71: 2476–2488

Durran D R, Weyn J A. 2016. Thunderstorms do not get butterflies. Bull Am Meteorol Soc, 97: 237–243

Ehrendorfer M, Errico R M, Raeder K D. 1999. Singular-vector perturbation growth in a primitive equation model with moist physics. J Atmos Sci, 56: 1627–1648

Grell G A, Dudhia J, Stauffer D R. 1995. A description of the fifth-generation Penn State/NCAR Mesoscale Model (MM5). NCAR Tech Note NCAR/TN-398+STR. 121

Hakim G J, Torn R D. 2008. Ensemble synoptic analysis. In: Bosart L F, Bluestein H B, eds. Synoptic-Dynamic Meteorology and Weather Analysis and Forecasting: A Tribute to Fred Sanders. Boston: American Meteorological Society

Hawblitzel D P, Zhang F, Meng Z, Davis C A. 2007. Probabilistic evaluation of the dynamics and predictability of the mesoscale convective vortex of 10–13 June 2003. Mon Weather Rev, 135: 1544–1563

Hong S Y, Noh Y, Dudhia J. 2006. A new vertical diffusion package with an explicit treatment of entrainment processes. Mon Weather Rev, 134: 2318–2341

Iacono M J, Delamere J S, Mlawer E J, Shephard M W, Clough S A, Collins W D. 2008. Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. J Geophys Res, 113: D13103

Jiménez P A, Dudhia J, González-Rouco J F, Navarro J, Montávez J P, García-Bustamante E. 2012. A revised scheme for the WRF surface layer formulation. Mon Weather Rev, 140: 898–918

Judt F. 2018. Insights into atmospheric predictability through global convection-permitting model simulations. J Atmos Sci, 75: 1477–1497

Judt F. 2020. Atmospheric predictability of the tropics, middle latitudes, and polar regions explored through global storm-resolving simulations. J Atmos Sci, 77: 257–276

Liang Y, Qiao C, Dong J. 2020. Spatial-temporal distribution and impact analysis of the first rainstorm in Henan Province of the recent 34 years (in Chinese). Meteorol Environ Sci, 43: 26–32

Lorenz E N. 1963. Deterministic nonperiodic flow. J Atmos Sci, 20: 130–141

Lorenz E N. 1969. The predictability of a flow which possesses many scales of motion. Tellus, 21: 289–307
Lorenz E N. 1982. Atmospheric predictability experiments with a large numerical model. Tellus, 34: 505–513
Lorenz E N. 1996. Predictability—A problem partly solved. In: Proceedings of Seminar on Predictability. Reading, United Kingdom, ECMWF. 1−18
Lynch S L, Schumacher R S. 2014. Ensemble-based analysis of the May 2010 extreme rainfall in Tennessee and Kentucky. Mon Weather Rev, 142: 222–239
Melhauser C, Zhang F. 2012. Practical and intrinsic predictability of severe and convective weather at the mesoscales. J Atmos Sci, 69: 3350–3371
Mu M, Duan W. 2003. A new approach to studying ENSO predictability: Conditional nonlinear optimal perturbation. Chin Sci Bull, 48: 1045–1047
Mu M, Xu H, Duan W. 2007. A kind of initial errors related to "spring predictability barrier" for El Niño events in Zebiak-Cane model. Geophys Res Lett, 34: L03709
Nielsen E R, Schumacher R S. 2016. Using convection-allowing ensembles to understand the predictability of an extreme rainfall event. Mon Weather Rev, 144: 3651–3676
Ran L, Li S, Zhou Y, Yang S, Ma S, Zhou K, Shen D, Jiao B, Li N. 2021. Observational analysis of the dynamic, thermal, and water vapor characteristics of the "7·20" extreme rainstorm event in Henan Province (in Chinese). Chin J Atmos Sci 45: 1366−1383
Rotunno R, Snyder C. 2008. A generalization of Lorenz's model for the predictability of flows with many scales of motion. J Atmos Sci, 65: 1063–1076
Selz T. 2019. Estimating the intrinsic limit of predictability using a stochastic convection scheme. J Atmos Sci, 76: 757–765
Selz T, Craig G C. 2015. Upscale error growth in a high-resolution simulation of a summertime weather event over Europe. Mon Weather Rev, 143: 813–827
Shi W, Li X, Zeng M, Zhang B, Wang H, Zhu K, Zhuge X. 2021.Multi-model comparison and high-resolution regional model forecast analysis for the "7·20" Zhengzhou severe heavy rain. Trans (in Chinese). Atmos Sci, 44: 688−702
Skamarock W C. 2004. Evaluating mesoscale NWP models using kinetic energy spectra. Mon Weather Rev, 132: 3019–3032
Skamarock W C, Klemp J B, Dudhia J, Gill D O, Liu Z, Berner J, Wang W, Powers J G, Duda M G, Barker D, Huang X Y. 2021. A Description of the Advanced Research WRF Model Version 4.3. NCAR Technical Note NCAR/TN-556+STR
Sun Y Q, Zhang F. 2016. Intrinsic versus practical limits of atmospheric predictability and the significance of the butterfly effect. J Atmos Sci, 73: 1419–1438
Sun Y Q, Zhang F. 2020. A new theoretical framework for understanding multiscale atmospheric predictability. J Atmos Sci, 77: 2297–2309
Sun Y Q, Rotunno R, Zhang F. 2017. Contribution of moist convection and internal gravity waves to building the atmospheric −5/3 kinetic energy spectra. J Atmos Sci, 74: 185–201
Sun Y, Xiao H, Yang H, Ding J, Fu D, Guo X, Feng L. 2021. Analysis of dynamic conditions and hydrometeor transport of Zhengzhou super heavy rainfall event on 20 July 2021 based on optical flow field of remote sensing data (in Chinese). Chin J Atmos Sci, 45: 1384−1399
Thompson G, Eidhammer T. 2014. A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. J Atmos Sci, 71: 3636–3658
Weyn J A, Durran D R. 2017. The dependence of the predictability of mesoscale convective systems on the horizontal scale and amplitude of initial errors in idealized simulations. J Atmos Sci, 74: 2191–2210
Weyn J A, Durran D R. 2019. The scale dependence of initial-condition sensitivities in simulations of convective systems over the southeastern United States. Q J R Meteorol Soc, 145: 57–74
Wilks D S. 1995. Statistical Methods in the Atmospheric Sciences: An Introduction. San Diego, CA: Academic Press. 467
Wu N, Zhuang X, Min J, Meng Z. 2020. Practical and intrinsic predictability of a warm-sector torrential rainfall event in the south China monsoon region. J Geophys Res-Atmos, 125: e31313
Yu H, Meng Z. 2016. Key synoptic-scale features influencing the high-impact heavy rainfall in Beijing, China, on 21 July 2012. Tellus A-Dynamic Meteorol Oceanography, 68: 31045
Yu H, Meng Z. 2022. The impact of moist physics on the sensitive area identification for heavy rainfall associated weather systems. Adv Atmos Sci, 39: 684–696
Zhang C, Wang Y. 2017. Projected future changes of tropical cyclone activity over the western North and South Pacific in a 20-km-mesh regional climate model. J Clim, 30: 5923–5941
Zhang F, Bei N, Rotunno R, Snyder C, Epifanio C C. 2007. Mesoscale predictability of moist baroclinic waves: Convection-permitting experiments and multistage error growth dynamics. J Atmos Sci, 64: 3579–3594
Zhang F, Snyder C, Rotunno R. 2002. Mesoscale predictability of the "surprise" snowstorm of 24–25 January 2000. Mon Weather Rev, 130: 1617–1632
Zhang F, Sun Y Q, Magnusson L, Buizza R, Lin S J, Chen J H, Emanuel K. 2019. What is the predictability limit of midlatitude weather? J Atmos Sci, 76: 1077–1091
Zhang M, Meng Z. 2018. Impact of synoptic-scale factors on rainfall forecast in different stages of a persistent heavy rainfall event in south China. J Geophys Res-Atmos, 123: 3574–3593
Zhang X. 2021. Impacts of different perturbation methods on multiscale interactions between multisource perturbations for convection-permitting ensemble forecasting during SCMREX. Q J R Meteorol Soc, 147: 3899–3921
Zhang X, Yang H, Wang X, Shen L, Wang D, Li H. 2021. Analysis on characteristic and abnormality of atmospheric circulations of the July 2021 extreme precipitation in Henan (in Chinese). Trans Atmos Sci, 44: 672−687
Zhang Y, Zhang F, Stensrud D J, Meng Z. 2016. Intrinsic predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma at storm scales. Mon Weather Rev, 144: 1273–1298
Zhuang Y, Xing A. 2022. History must not repeat itself-urban geological safety assessment is essential. Nat Hazards, 111: 2141–2145
Zou X L, Vandenberghe F, Pondeca M, Kuo Y H. 1997. Introduction to adjoint techniques and the MM5 adjoint modeling system. NCAR Tech. Note NCAR/TN2435+STR. 110

(Responsible editor: Guixing CHEN)