ADVANCES IN ATMOSPHERIC SCIENCES, VOL. 40, MAY 2023, 937–951

• **Original Paper** •

Evaluation of the Added Value of Probabilistic Nowcasting Ensemble Forecasts on Regional Ensemble Forecasts※

Lu YANG*, Cong-Lan CHENG, Yu XIA, Min CHEN, Ming-Xuan CHEN, Han-Bin ZHANG, and Xiang-Yu HUANG

Institute of Urban Meteorology, *China Meteorological Administration*, *Beijing* 100089, *China*

(Received 28 April 2022; revised 23 September 2022; accepted 26 October 2022)

ABSTRACT

Ensemble forecasting systems have become an important tool for estimating the uncertainties in initial conditions and model formulations and they are receiving increased attention from various applications. The Regional Ensemble Prediction System (REPS), which has operated at the Beijing Meteorological Service (BMS) since 2017, allows for probabilistic forecasts. However, it still suffers from systematic deficiencies during the first couple of forecast hours. This paper presents an integrated probabilistic nowcasting ensemble prediction system (NEPS) that is constructed by applying a mixed dynamic-integrated method. It essentially combines the uncertainty information (i.e., ensemble variance) provided by the REPS with the nowcasting method provided by the rapid-refresh deterministic nowcasting prediction system (NPS) that has operated at the Beijing Meteorological Service (BMS) since 2019. The NEPS provides hourly updated analyses and probabilistic forecasts in the nowcasting and short range (0–6 h) with a spatial grid spacing of 500 m. It covers the three meteorological parameters: temperature, wind, and precipitation. The outcome of an evaluation experiment over the deterministic and probabilistic forecasts indicates that the NEPS outperforms the REPS and NPS in terms of surface weather variables. Analysis of two cases demonstrates the superior reliability of the NEPS and suggests that the NEPS gives more details about the spatial intensity and distribution of the meteorological parameters.

Key words: integration, ensemble nowcasting, probabilistic prediction, evaluation and verification

Citation: Yang, L., C.-L. Cheng, Y. Xia, M. Chen, M.-X. Chen, H.-B. Zhang, and X.-Y. Huang, 2023: Evaluation of the added value of probabilistic nowcasting ensemble forecasts on regional ensemble forecasts. *Adv. Atmos. Sci.*, **40**(5), 937–951, https://doi.org/10.1007/s00376-022-2056-8.

Article Highlights:

- The NEPS system provides hourly updated analyses and probabilistic forecasts in the nowcasting and short range (0–6 h) with a grid spacing of 500 m.
- The added value of the probabilistic forecast of the NEPS was attributed to the mixed dynamic-integrated method.
- The NEPS is needed to provide end users with more details about the spatial intensity and distribution of the meteorological parameters.

1. Introduction

Weather forecasting is a key factor that supports decision-making in various applications. Therefore, the public, private sector, and other stakeholders have high requirements for accurate weather forecasts. In operational weather forecasting, assessing forecast uncertainty is a key and challenging issue. At present, the primary means to achieve this goal is through numerical weather forecasting. Numerical model output is generated based on complex mathematics and the input of many observations obtained by various conventional and remote sensing technologies. However, influential factors such as the uncertainty of the model's initial field, dynamics, and physical processes may cause a large deviation between a single deterministic numerical forecast and actual weather, which also limits the predictability of a single model (Lorenz, 1965, 1969).

Over the past ten years, ensemble prediction systems (EPSs) have become the standard method of accounting for uncertainties in initial conditions and model formulations. The development of ensemble forecasts heralds the transition from deterministic forecasts to probabilistic forecasts, which can provide users with more comprehensive and com-

※ This paper is a contribution to the special issue on the 14th International Conference on Mesoscale Convective Systems and High-Impact Weather.

* Corresponding author: Lu YANG
Email: lyang@ium.cn

plete information. Therefore, ensemble forecasting technology has played an important role in numerical weather forecasting operations in various countries (Toth and Kalnay, 1993; Buizza and Palmer, 1995; Houtekamer et al., 1996; Molteni et al., 1996; Li and Chen, 2002; Buizza et al., 2005; Chen et al., 2005; Chen and Shen, 2006).

The European Centre for Medium-Range Weather Forecasts (ECMWF; Molteni et al., 1996) and the National Centers for Environmental Prediction (NCEP; Toth and Kalnay, 1993) successively put the ensemble forecast system into operation. Various ensemble perturbation schemes have since been developed to address the uncertainties in global ensemble prediction systems (GEPSs; Buizza and Palmer, 1995; Houtekamer et al., 1996; Buizza et al., 2005). A new generation of numerical forecasting systems of the Meteorological Administration Global Assimilation Forecast System (CMA-GFS, formerly GRAPES Global Assimilation Forecast System, GRAPES-GFS) in China was independently developed and officially put into operation in 2009 (Li and Chen, 2002; Chen et al., 2005; Chen and Shen, 2006; Zhang and Shen, 2008; Deng et al., 2010).

The development of regional ensemble prediction lags behind that of global ensemble prediction. However, the need for regional ensemble forecast systems (REPSs) to provide mesoscale weather forecasts is obvious (Peralta et al., 2012; Kühnlein et al., 2014; Kim and Kim, 2017; Kiktev et al., 2017). Regional ensemble prediction is of great significance for improving the forecast accuracy of severe weather, and research on related technologies has received more attention in recent years. The Beijing Meteorological Service has been routinely running a REPS since 2017. The REPS has a better representation of the convective structure and probabilistic precipitation forecasts than coarser models using convective parameterizations (Clark et al., 2009; Johnson et al., 2014). However, the system still faces problems, such as a lack of ensemble spread, obvious systematic bias, and imperfect prediction of mesoscale uncertainty (Zhang et al., 2014, 2017). Therefore, post-processing methods are still needed to further correct system deviations and adjust dispersion.

In observation-based or nowcasting forecasting, the emphasis has recently shifted from deterministic extrapolation to the inclusion of probabilistic methods (Dance et al., 2010). The NPS (Nowcasting Prediction System) mentioned in the abstract is a multivariable analysis and nowcasting system based on the core algorithm of the Integrated Nowcasting through Comprehensive Analysis (INCA) system developed by the Austrian Meteorological Bureau (Haiden et al., 2011). It has operated at BMS since 2017. The basic concept of the nowcasting method in NPS is to improve upon numerical weather prediction (NWP) model output by downscaling and bias correction, using the latest surface observational data and high-resolution topographic data (Chen et al., 2020; Cheng et al., 2019; Song et al., 2019a, b; Yang et al., 2019, 2021, 2022). Thus the advantages of the system are mainly reflected during the nowcasting period as it directly assimilates and incorporates multi-source observation data, which contributes to a forecast performance that is better than NWP models in the first 2–3 hours. However, this method does not take into account uncertain information from the REPS. At longer lead times, the NWP output it uses is affected by uncertainties due to improper distribution of initial conditions and physical parameterizations. In addition, extreme weather has become more common and persistent over the past decade. However, extreme weather is a low-probability event with great uncertainty in its development, so it is far from enough to rely only on extrapolation prediction and a single numerical model to make a deterministic prediction (Gao et al., 2019).

This study aims to introduce a new integrated probabilistic nowcasting ensemble system to quantify these uncertainties and to provide reliable site-specific, probabilistic short-range forecasts that cannot be provided by a single deterministic nowcasting forecast. More specifically, our attention is focused on the evaluation of the NEPS (Nowcasting Ensemble Prediction System) to see if the forecast of the NEPS could add value to the forecast of the REPS and NPS that is already available.

The remainder of this paper is organized as follows: The REPS and NPS configuration are introduced in section 2; in section 3, a brief description of the method and implementation are demonstrated; section 4 demonstrates verification data, methods, and results of a comparison of the NEPS to the REPS and NPS. Two cases are given in section 5. A conclusion and discussions are given in section 6.

2. System description

2.1. *The regional ensemble prediction system*

The regional ensemble prediction products we used in this paper are based on the output of a convective-scale ensemble forecast system developed by the BMS (Zhang et al., 2017, 2019, 2022). The REPS uses the Weather Research and Forecast model (WRF version-4.1.2). The key model physics configuration of all members are the same, i.e., Thompson microphysics (Thompson et al., 2008), Kain-Fritsch Cumulus parameterization (Kain and Fritsch, 1993), Mellor-Yamada-Janjic (MYJ) planetary boundary layer (PBL) scheme (Mellor and Yamada, 1982), and the RRTMG longwave and shortwave radiation schemes (Iacono et al., 2008). The domain and topography of the REPS are shown in Fig. 1a. The REPS is run with a grid spacing of 3 km, with 550 and 423 grid points in the x- and y-directions, respectively. The REPS is run twice daily at 0000 and 1200 UTC, up to a lead time of 48 h. It consists of one control forecast and 20 perturbed members running on a large domain that covers most areas of North China.

The system process framework is shown in Fig. 1b. First, initial conditions are obtained by the dynamical downscaling (Weidle et al., 2016) of the 12-h forecast field of GEFS. For example, the initial field at 0000 UTC on 2 July 2021 is obtained from, the dynamical downscaling of the 12-h forecast field of GEFS at 1200 UTC on 1 July 2021.

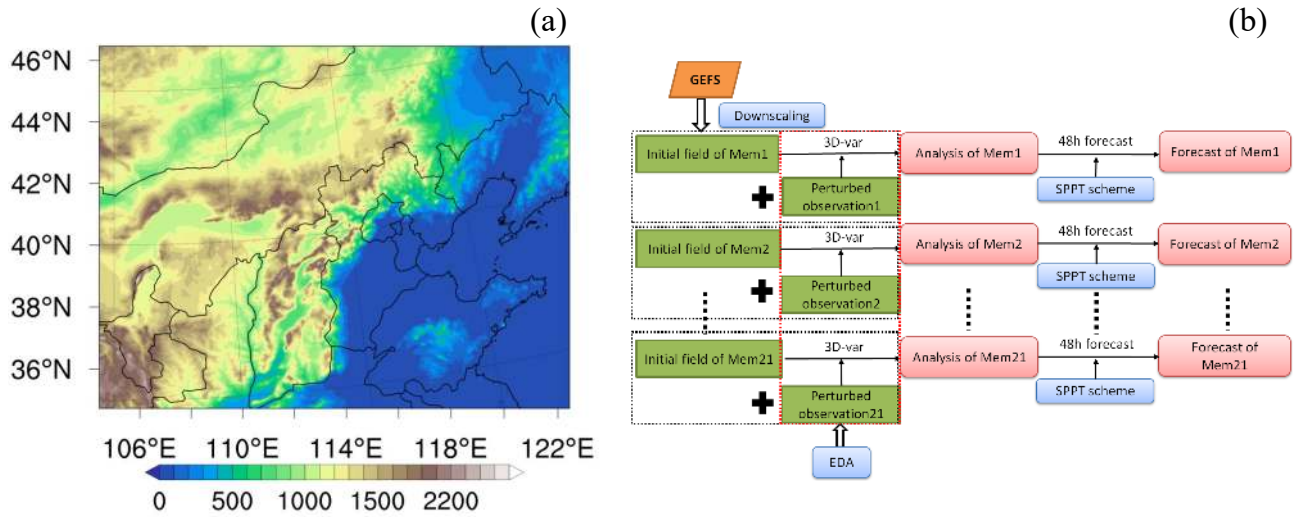

Fig. 1. (a) The domain and topography of the REPS (color shading shows terrain height; units: m) and (b) the process framework of the regional ensemble prediction system.

The dynamical downscaling approach obtains the regional ensemble initial condition perturbations by interpolating an ensemble of global initial conditions to the regional model domain and grid spacing (Kühnlein et al., 2014). It is an effective way to extract local-scale information by regional models with coarse global data used as boundary conditions (Cannon and Whitfield, 2002).

There are many kinds of observation data in North China, which can be used to carry out three-dimensional variation (3DVAR) based on the GEFS forecast product to improve the quality of initial conditions. Thus, the initial uncertainty is provided by the Ensemble Data Assimilation (EDA) method based on 3DVAR algorithms from the WRF data assimilation (WRFDA, version-4.1.2) system and perturbed sets of observations (Zhang et al., 2022). The ensemble members of the REPS assimilate various kinds of observations, including land-based synoptic reports (SYNOP), aviation routine weather reports (METAR), radiosondes, aircraft reports, and ship and buoy reports over the sea. In the EDA process, different ensemble members are updated using different sets of perturbed observations (Houtekamer and Mitchell, 1998). This process is carried out at the forecast time and started twice daily (cold-started) at 0000 and 1200 UTC, respectively. In the REPS, the background error covariance is estimated by the National Meteorological Center (NMC, USA) method (Parrish and Derber, 1992). Pairs of forecast differences valid for the same time but at different lengths for three months yield samples to estimate background error covariance. Model uncertainties due to physical parameterizations are simulated using a stochastic perturbation of physics tendencies (SPPT) scheme (Buizza et al., 1999; Bouttier et al., 2012), and the side boundary uncertainty is provided by downscaling the GEFS global background field.

2.2. *The deterministic nowcasting prediction system*

The analysis and forecasting methodology in the NPS has also been described elsewhere (Chen et al., 2018; Yang et al., 2021), so only a short summary is given here. It was developed primarily as a means of providing modified numerical forecast products, such as temperature, humidity, wind, precipitation, and precipitation type, by using high-resolution topographic data and real-time multi-source observations in the nowcasting range. Outside the nowcasting range, the nowcasts are blended into downscaled NWP forecasts. The NPS system provides frequently updated analyses and forecasts for a domain covering the Jing-Jin-Ji (Beijing-Tianjin-Hebei) area, the horizontal grid spacing is 500 m, and the forecast range is 24 h. The domain, topography (color shading), locations of surface observation stations (black dots), and radar stations (red circles) used in the blending are shown in Fig. 2. The topography used in the NPS system is constructed using a bilinear interpolation of the 30-m global terrain height data onto the 500-m NPS grid.

Figure 3 briefly introduces the input, the primary fields, their sources for the analysis and forecast fields, and the nowcasting and blending method in NPS. The main nowcasting methods of the system include the blending algorithm of multi-source data extrapolated prediction with NWP, the bias correction algorithm, and the downscaling algorithm under complex terrain (Haiden et al., 2011; Cheng et al., 2019; Song et al., 2019a, b; Yang et al., 2019, 2021, 2022; Chen et al., 2020). The analysis fields at 10-minute intervals on a 500 m × 500 m grid are generated from a combination of NWP model output, observation data, radar data, and high-resolution topographic data. In the current version, the NWP model output is taken from the deterministic CMA-BJ model (formerly known as RMAPS-ST) (He et al., 2019; Xie et al., 2019), which has a horizontal grid spacing of 3 km. The domain of the deterministic CMA-BJ model is the same as the REPS in Fig. 1. Surface observations are provided by 4323 real-time automatic weather stations (Fig. 2). Radar data are taken from ten S-band radars and one C-band radar in North China (Fig. 2).

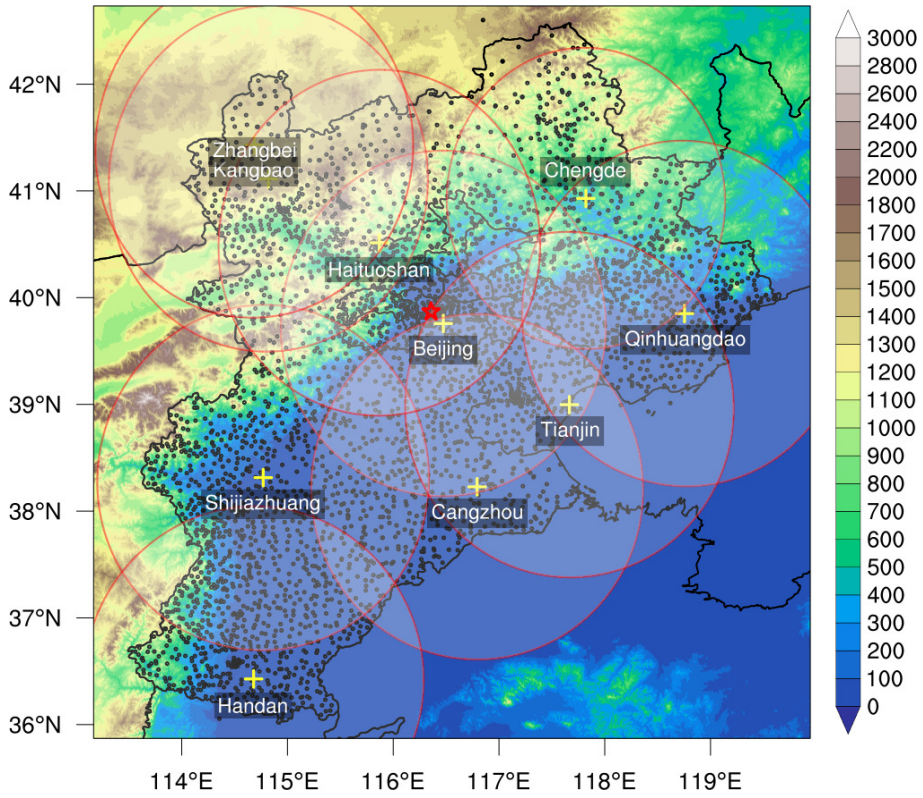

Fig. 2. The domain and topography of the NPS, locations of surface observation stations and radar stations participating in the blending (black dots and red circles indicate observation and radar stations, respectively; color shading shows terrain height; units: m).

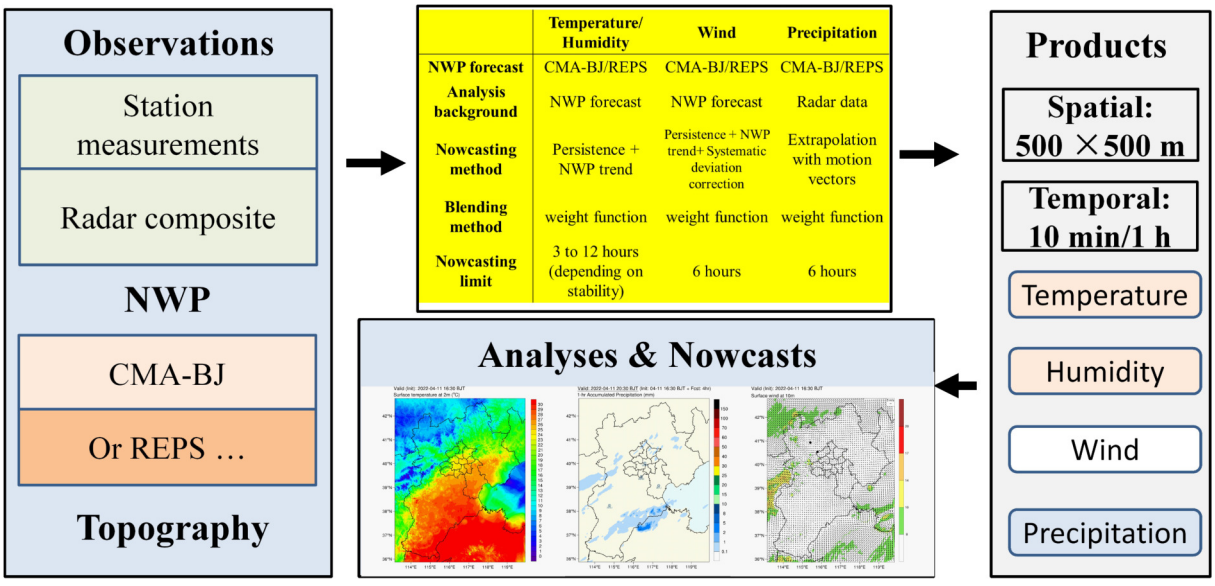

Fig. 3. A flowchart showing the input, the primary fields, their sources for the analysis and forecast fields, and the nowcasting and blending method in the deterministic nowcasting prediction system.

The NPS system has proven to be valuable compared to NWP. On the one hand, the NWP data are topographically downscaled onto a 500 m grid spacing using the 30 m global terrain height data. A more realistic simulation of topographic downscaling is crucial for wind and temperature, at least in mountainous regions. On the other hand, the systematic errors of NWP can be greatly reduced by the NPS algorithm with the assimilation and incorporation of additional multi-source observations, especially in the nowcasting and short range (Haiden et al., 2011; Yang et al., 2019; Chen et

al., 2020). The primary fields include surface meteorological parameters like temperature, humidity, wind, and precipitation. However, there is limited interdependency between the fields and it cannot be used as the initial field for other models. Besides, considering the requirements of operational efficiency and computational expense, the system focuses mainly on the forecast of surface weather variables.

In the case of temperature, the NWP forecast for the temperature at each pressure level from CMA-BJ is trilinearly interpolated to the three-dimensional NPS 500-m grid and used as the first guess. In the next step, the first guess is corrected according to the differences between the observed data and the NWP forecast at each station. The difference field obtained by interpolating the observed increments is added to the analysis field (0 h), and the nowcasting is based on the trend given by the NWP model to the latest NPS analysis. At a lead time of several hours, the trend extrapolation is smoothly blended into the downscaled NWP forecast according to the weight function. The actual nowcasting limit of temperature varies between 3 and 12 h, depending on the stability of the valley atmosphere (Haiden et al., 2011).

In the case of wind forecast, a basic bias correction algorithm is also an integral part (Yang et al., 2022) of the NPS. The 10-m wind field of NWP forecasts from CMA-BJ is first downscaled and interpolated to the two-dimensional NPS 500-m grid and taken as a first guess of NPS. Then the downscaled 10-m wind field with a 500-m grid spacing is modified by multiplying the bias correction coefficient to eliminate systematic errors. Next, it will be corrected based on the differences in the u and v components between the model and the observations (Yang et al., 2019). In the first 6-h lead times, the wind forecast is based on the trend given by the NWP model to the NPS analysis, and the weight of the NWP model increases with time. At a longer lead time, the wind forecast is purely the downscaled NWP 10-m wind forecast with the systematic errors eliminated.

In the case of precipitation, the analysis field is based on the weighted fusion of observational rain gauge data and radar quantitative precipitation estimation (Cheng et al., 2019; Song et al., 2019a, b). The rain gauge and maxCAPPI product are first bilinearly interpolated onto the NPS grid. To eliminate the systematic errors of the maxCAPPI products, a climatological scaling factor has been calculated for each month (Song et al., 2019a) by using the accumulated precipitation obtained from radar and station observations. The rescaled radar field and the observational rain gauge data are finally aggregated to accumulated precipitation amounts. In the first 2-h lead time, the precipitation forecast is based on a correlation-based extrapolation with motion vectors. Between 3–6 h, the precipitation forecast is obtained by blending the NWP output and correlation-based extrapolation results. At longer lead times, the precipitation forecast is purely the downscaled NWP precipitation forecast.

3. Methodology

The main goal of this research work is to create a new integrated probabilistic nowcasting ensemble prediction system by applying a blending dynamical-integrated approach. The main approach is to blend the ensemble variance information from the REPS with the nowcasting methods in NPS. The output forecast of the 21 members from the REPS is used as a background for creating the probabilistic forecast. The nowcasting and blending methods of NPS are used for creating the nowcasting forecast.

A general overview of the technical route of the NEPS is shown in Fig. 4. The NEPS meteorological (temperature/wind/precipitation) analysis and forecast start with the meteorological forecast of the 21 members of the REPS from the two cycles at 0000 and 1200 UTC as a first-guess field, which is different from NPS with the single initial field. Then, the meteorological fields from the 21 members of the REPS are topographically downscaled onto a grid with

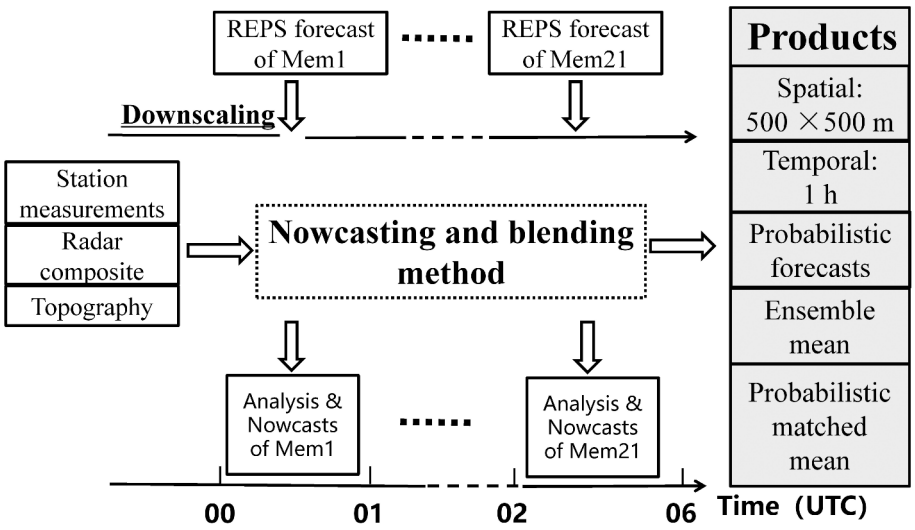

Fig. 4. A general overview of the technical route of the probabilistic nowcasting ensemble prediction system.

500-m spacing using the 30-m global terrain height data, consistent with the NPS. In the next step, applying nowcasting methods like multi-source data (Observations from radar and AWS) blending, bias correction algorithm, and extrapolation with vector motions are applied to the 21 downscaled first-guess fields. This leads to a variety of nowcasts that combines the advantages of the two systems: on the one hand, the observation basis on nowcasting is at very high resolution; on the other hand, the probabilistic short-range forecasting is based on the REPS.

Because of current computational limitations, the update frequency of the NEPS is set to 1 hour instead of 10 min. The domain and topography of the NEPS (Fig. 2) are the same as in the NPS. For every hour, the latest observation and radar data are used to modify the first guess field provided by the REPS. The NEPS provides the 1-h updated analyses and the probabilistic forecasting in the short range (0–6 h) at a spatial grid spacing of 500 m in the Jing-Jin-Ji area. The products include the precipitation, wind, and temperature output of 21 different members, as well as the probabilistic forecasts and probability-matched (PM) mean forecasts (Clark, 2017).

4. Verification data, methods, and results

To evaluate the improved performance of the NEPS and the extent to which it can further add application value to weather forecast services in comparison with the NPS and REPS, a more comprehensive evaluation of the performance of the NEPS was done based on an almost one-month period from 1 July 2021 to 25 July 2021. The sourced forecasts from the NEPS and REPS have different temporal resolutions and spatial grid spacings. The REPS runs start at 0000 and 1200 UTC to generate a forecast with 1-hourly output with a grid spacing of 3 km. The NEPS and NPS, which have a grid spacing of 500 m, are run once an hour to generate a forecast with 1-hourly output. For a proper comparison, the REPS forecast was bilinearly interpolated to the NEPS domain and 500-m grid spacing. Since hourly initializations (0000, 0100, …, 2300 UTC) of the NEPS are both from the two cycles at 0000 and 1200 UTC of the REPS, our primary interest is to evaluate the added value brought by the NPS. Therefore, hourly initializations of the NEPS and NPS were all considered and combined for each of the 6-h forecast lead times to compare the NEPS to the REPS and NPS.

Since the NEPS is mainly concerned with forecasting surface weather variables, we verified the 2-m temperature, 10-m wind, and precipitation from deterministic and probabilistic comparative tests, respectively. There are about 4000 automatic stations in the verification domain, including national and regional automatic stations, marked with black dots in Fig. 2. The automatic stations cover most of the topographic elevation range from 0 to 2194 m. The observation is matched to the nearest grid point, and observation uncertainties are not considered.

For the deterministic forecast of the 2-m temperature and 10-m wind, we compared the ensemble mean product of the NEPS with the NPS. The root-mean-square error (RMSE, Yang et al., 2019) was calculated for 2-m temperature and 10-m wind. Deterministic forecasts of precipitation in the NEPS were obtained by using the PM technique (Clark, 2017). To test the forecast skill of precipitation and how well the areal coverage of precipitation matched the observations, the neighborhood-based equitable threat score (ETS, Wang and Yan, 2007; Clark et al., 2010) and bias score (BIAS, Wilks, 2006) were calculated for precipitation using accumulation thresholds of 0.1 mm h^{-1}, 1 mm h^{-1}, 5 mm h^{-1} and 10 mm h^{-1}.

The formulation of a neighborhood-based ETS is described by Clark et al. (2010). In this study, we set the neighborhood radius (r) to 5 km so that if a given precipitation threshold (q) that is observed at a grid point is met, it is considered a hit if the event is forecast at any grid point within the neighborhood radius (r). If an event is observed or forecasted at a grid point, but no grid points within radius (r) forecast or observe the event, it is considered as a miss or false alarm. Correct negatives are calculated when an event is neither observed nor forecasted at a single point. Then, the neighborhood-based ETS can be computed according to the hits, misses, false alarms, and correct negatives. Using these elements, a neighborhood-based ETS is expressed as:

$$\mathrm{ETS} = \frac{\mathrm{hits} - \mathrm{chance}}{\mathrm{hits} + \mathrm{misses} + \mathrm{false\ alarms} - \mathrm{chance}}, \tag{1}$$

$$\mathrm{chance} = \frac{(\mathrm{hits} + \mathrm{misses})(\mathrm{hits} + \mathrm{false\ alarms})}{\mathrm{hits} + \mathrm{misses} + \mathrm{correctnegatives} + \mathrm{false\ alarms}}. \tag{2}$$

Similarly, BIAS can be calculated as follows:

$$\mathrm{BIAS} = \frac{\mathrm{hits} + \mathrm{misses}}{\mathrm{hits} + \mathrm{false\ alarms}}. \tag{3}$$

The spatial representation of the PM product is given by the ensemble mean and the rainfall amounts are given by the 90-th percentile value in the distribution of ensemble member quantitative precipitation forecasts (QPFs). We chose the 90% value in the distribution of ensemble member QPFs because it is the best frequency distribution of rainfall amount according to the neighborhood-based ETS and BIAS score.

To verify the probabilistic forecast, we compared the probabilistic forecast results of the NEPS and REPS. The Relative Operating Characteristic (ROC) measures the combined effect of the Probability Of Detection (POD) and the False Alarm Rate (FAR). The area under the ROC curve (AROC, Zhong et al., 2017) is often calculated to determine whether the forecast is skillful, and forecasting systems with a ROC area greater than 0.7 are considered useful (Stensrud and Yussouf, 2007).

The RMSE, the proper continuous ranked probability score (CRPS, Gneiting and Raftery, 2007), the percentage of outliers (OUTLIERS, Suklitsch et al., 2015), and the Talagrand histogram (Talagrand et al., 1997; Hamill, 2001) were used to assess the probabilistic 2-m temperature and 10-m wind products.

The CRPS score measures the probabilistic skill of the ensemble forecasting system, and it measures the overall ensemble forecast performance compared to the observations (Hersbach, 2000). A zero value of CRPS translates to the best forecast. The higher the CRPS score, the worse the performance of the ensemble forecasting system. The Talagrand histogram is a tool for testing the reliability of the ensemble forecast system. If the ensemble forecast is reliable, the predicted and the observed values of the ensemble forecast member at a given point should be regarded as a random sample subject to the same probability distribution. Furthermore, the Talagrand histogram indicates bias, with an L-shaped (U-shaped) rank histogram indicating a tendency for members to over-forecast (under-forecast) the variable being examined (Hamill, 2001). The percentage of outliers is the sum of the probabilities at the two ends of the Talagrand histogram. The percentage of outliers shows how many observed values lie outside the full forecasted range. The smaller the outliers, the better the reliability of the ensemble forecasting system.

4.1. *Validation results*

4.1.1. *Comparative verification of deterministic products*

Table 1 shows the 1–6 h aggregate neighborhood-based ETS and BIAS scores of the NEPS and NPS using different accumulation thresholds of 0.1 mm h^{-1}, 1 mm h^{-1}, 5 mm h^{-1}, and 10 mm h^{-1} and the improvement percentage of the ETS scores of the NEPS compared to NPS for the period of 1 to 25 July 2021. In addition, Fig. 5 presents the neighborhood-based ETS and BIAS scores for various hourly fixed precipitation thresholds of the NEPS and NPS.

It can be seen from Table 1 and Fig. 5 that the NEPS produces more skillful forecasts at the 0.1 mm h^{-1}, 1 mm h^{-1}, and 5 mm h^{-1} thresholds and that the improvements persist through the 6-h validation period. The improvement rates of ETS scores were respectively 123%, 55%, and 122%. The ETS score of the NEPS PM product at the 0.1 and 1 mm h^{-1} threshold were about 0.7 and 0.45, respectively, within the 6-h forecast lead time, which was better than that of the NPS (about 0.3). The ETS score of the NEPS at the 5 mm h^{-1} threshold was about 0.27 within the 6-h forecast lead time, slightly better than the NPS (about 0.12), and there was not much difference in the BIAS score. At the 10 mm h^{-1} threshold, the NPS performed better and had higher ETS values. The BIAS score of the NEPS was close to 1 at the 0.1 and 1 mm h^{-1} thresholds through the 6-h forecast period, while the BIAS score of the NEP increased with increasing precipitation threshold. The BIAS was about 2 at the 5 mm h^{-1} and 10 mm h^{-1} thresholds for forecast lead times of 2–6 h. The NEPS overpredicted moderate and heavy precipitation amounts, especially at the 10 mm h^{-1} threshold.

The NPS appeared to have lower ETS scores than the NEPS, indicating that most of the skill comes from ensemble variance information from the REPS. The PM product represents an improvement relative to the deterministic forecast in the NPS. The problem with the NPS is that light precipitation amounts are overpredicted, whereas the BIAS of the NEPS is closer to 1, and the areal coverage of precipitation at the 0.1 mm h^{-1} and 1 mm h^{-1} thresholds match better with the observations. However, very little skill is obtained for the NEPS at the 10 mm h^{-1} accumulation threshold, and the ETS scores of the NEPS are higher than the NPS in the first two hours while they are lower than the NPS in the following forecast hours.

Table 2 presents the 1–6 h aggregate RMSE of the 2-m temperature and 10-m wind field for the NEPS and NPS and the RMSE reduction rate of the NEPS compared with the NPS. In addition, Fig. 6 shows the RMSE and the reduction rate of the NEPS compared with the NPS for different lead times. Table 2 and Fig. 6 show that the deterministic forecasts of 2-m temperature and 10-m wind speed computed from the ensemble mean of the NEPS perform better than the NPS, especially for the 10-m wind speed. The NEPS' forecasts of 2-m temperature (Fig. 6a) are more skillful than those of the NPS, whose forecast errors are slightly smaller than the NEPS. The NPS showed an RMSE of around 1.7°C, which was reduced to about 1.62°C in the NEPS the first six hours. The RMSE of the NEPS for the 10-m wind field (Fig. 6b) was lower compared to NPS. The RMSE of the 10-m wind field was reduced from 1.2 m s^{-1} to 1.0 m s^{-1}.

The RMSE of the ensemble mean describes the correctness of the average estimate from the ensemble. The RMSE reduction rate for the 2-m temperature and 10-m wind speed was 4.25% and 19.02%, respectively. The NEPS and NPS sys-

Table 1. 1–6 h aggregate neighborhood-based ETS and BIAS scores at different precipitation thresholds for the NEPS and NPS and the improvement rates of ETS scores of the NEPS compared with the NPS for the period of 1 to 25 July 2021. Deterministic forecasts of precipitation in the NEPS were obtained from the PM technique.

	1–6 h aggregate ETS score			1–6 h aggregate BIAS score	
Precipitation thresholds	NEPS	NPS	Improvement rates of ETS scores	NEPS	NPS
0.1 mm h^{-1}	0.70	0.32	123%	0.96	1.79
1 mm h^{-1}	0.48	0.31	55%	1.2	1.60
5 mm h^{-1}	0.32	0.14	122%	1.53	1.42
10 mm h^{-1}	0.09	0.10	–13%	1.60	1.04

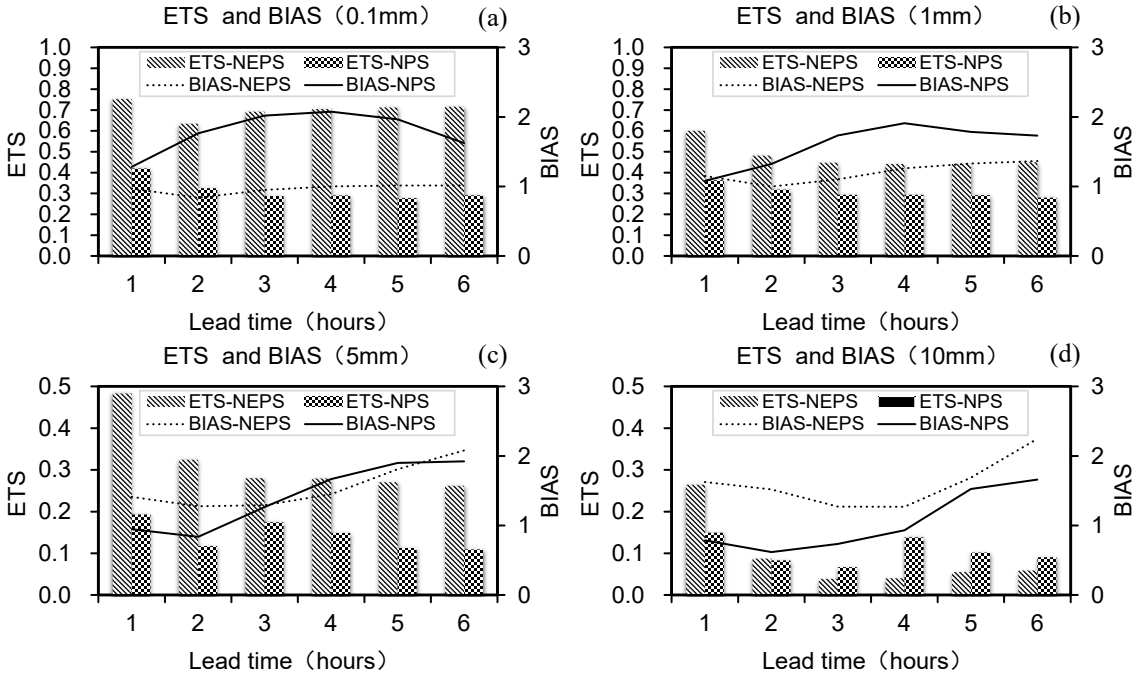

Fig. 5. Neighborhood-based ETS and BIAS scores of the NEPS and NPS, plotted as a function of lead time for the different accumulation thresholds of 0.1 mm h^{-1}, 1 mm h^{-1}, 5 mm h^{-1}, and 10 mm h^{-1}, aggregated for the period of 1 to 25 July 2021. The columns and the lines denote the corresponding ETS and BIAS of the NEPS and NPS, respectively (refer to the legend at the top of the figure). Deterministic forecasts of precipitation in the NEPS were obtained from the PM technique.

Table 2. 1–6 h aggregate RMSE of the 2-m temperature and 10-m wind field for the NEPS and NPS and the RMSE reduction rate of the NEPS compared with NPS for the period of 1 to 25 July 2021. Deterministic forecasts of 2-m temperature and 10-m wind field in the NEPS were obtained from the ensemble mean product.

	1–6 h aggregate RMSE		
Verification variables	NEPS	NPS	RMSE reduction rate
2-m temperature (°C)	1.62	1.70	4.25%
10-m wind field (m s^{-1})	0.99	1.24	19.02%

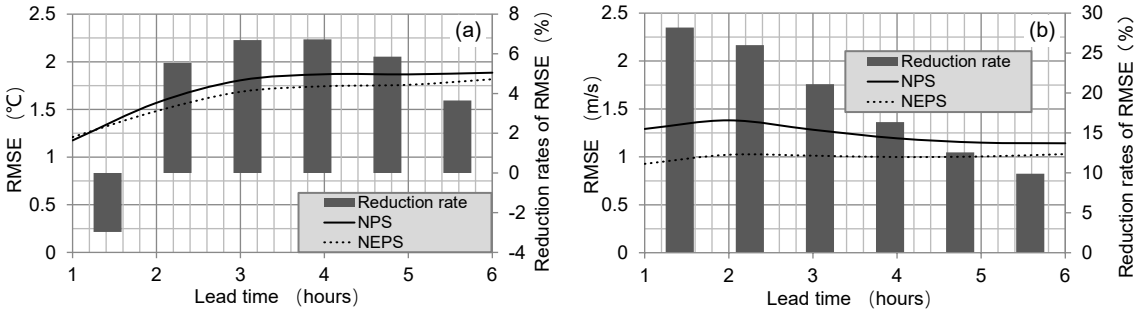

Fig. 6. RMSE of 2-m temperature (a) and 10-m wind field (b) for 1–25 July 2021 of 6-hour forecasts with the NEPS (dashed line) and NPS (solid line), the grey histogram represents the RMSE reduction rate of the NEPS compared with the NPS. Deterministic forecasts of the 2-m temperature and 10-m wind field in the NEPS were obtained from the ensemble mean product.

tems use the same underlying topography and observational data. The underperformance of the NPS data in estimating the 2-m temperature and 10-m wind field may be largely due to the improper distribution of initial conditions and physical parameterizations from a single numerical model. So the added value of deterministic products comes from the ensemble mean product of the REPS system, which describes the central estimate produced by the ensemble, and the most uncertain aspects of the individual member forecast are filtered out by computing the ensemble mean (Leith,

1974; Holton, 2004).

4.1.2. *Comparative verification of probabilistic products*

Figure 7 compares the aggregate AROC for various 1–6 h hourly fixed precipitation thresholds for the REPS and NEPS. The REPS generally has lower AROC scores than the NEPS for each precipitation threshold. Thus, most of the increase in the AROC is realized from the NPS. Using a ROC area of 0.7 as a threshold to determine forecast skill, the REPS could not produce useful forecasts when the precipitation threshold was equal to 5 mm h^{-1}, 10 mm h^{-1}, and 25 mm h^{-1}. However, the NEPS provided useful information at all thresholds except 25 mm h^{-1}. This finding indicates that the NEPS can improve the skill of probabilistic precipitation forecasts and yield an added value over the REPS forecasts in predicting the hourly rainfall.

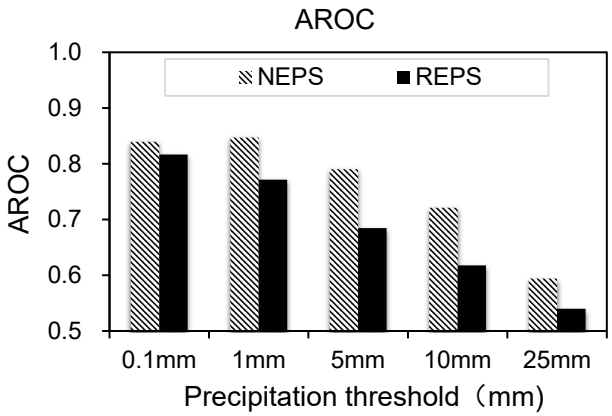

Fig. 7. The area under the ROC curve (AROC) using data aggregated from 1 to 25 July 2021 using accumulation thresholds of 0.1 mm h^{-1}, 1 mm h^{-1}, 5 mm h^{-1},10 mm h^{-1}, and 25 mm h^{-1}, with the REPS and NEPS.

For the 2-m temperature (Fig. 8a), the REPS and NEPS had a very similar U-shaped rank histogram, with both ensembles exhibiting a lack of variability. However, it appeared that the probability of ensemble forecast members falling outside the maximum and minimum values of the NEPS was slightly lower than that of the REPS, as indicated by a flatter rank histogram in the NEPS compared to the REPS. In the case of a 10-m wind, it was found that the REPS had a greater overestimated bias for wind speed than the NEPS by comparing the Talagrand histogram of the REPS and NEPS. The REPS results showed an L-shaped distribution (Fig. 8b), indicating that the ensemble forecasts were systemically large for wind speed. The distribution probability of the NEPS system was relatively flat (Fig. 8b), indicating that the systematic bias was recalibrated to some extent with the NEPS system, thus achieving improved probabilistic forecasts. The added value of probabilistic forecasts comes from the increased horizontal resolution of the underlying topography, the blending of multi-source observation data, and the integrated nowcasting methods. However, the probability of the observation point appearing in the last box is slightly higher than the others, showing a slightly inverse L-shaped distribution. It also provides information that may be used in the future to recalibrate ensemble forecasts through ensemble post-processing methods.

The lines in Fig. 9 show the CRPS values, and the columns represent the percentage of outliers for the 2-m temperature and 10-m wind. Evidently, the additional skill in 2-m temperature and 10-m wind of the NEPS compared with

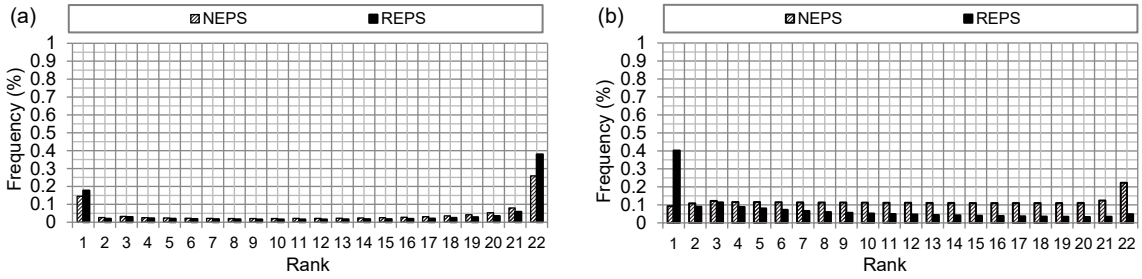

Fig. 8. Talagrand diagrams using data aggregated for 1 to 25 July 2021 for the (a) 2-m temperature and (b) 10-m wind for the REPS and the NEPS.

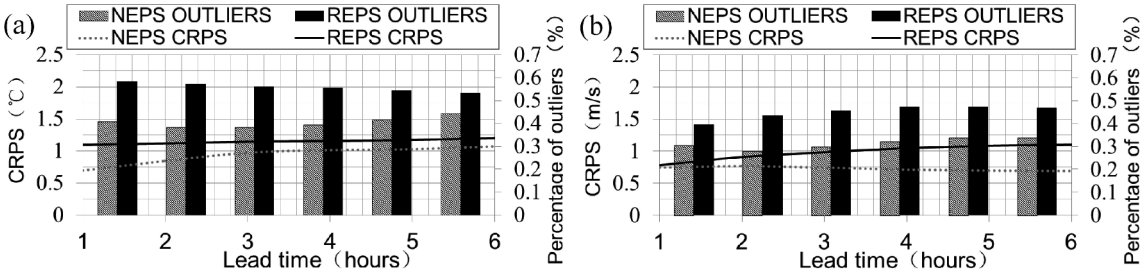

Fig. 9. Continuous Ranked Probability Score (CRPS; lines) for the (a) 2-m temperature and (b) 10-m wind with the REPS (solid line) and the NEPS (dashed line), and the percentage of outliers (columns) using data aggregated from 1 to 25 July 2021 for the (a) 2-m temperature and (b) 10-m wind in the REPS and NEPS.

the REPS corresponds to lower CRPS scores. In the case of temperature, the CRPS score decreased from 1.2°C to 1.0°C in the nowcasting ranges of up to 6 h. For 10-m wind, there was a considerable improvement by the NEPS, as the CRPS score was reduced from 1.1 m s^{-1} to 0.7 m s^{-1}. Still, the NEPS had smaller outliers in both parameters for the first six hours. The percentage of outliers was reduced from 0.47 (REPS) to about 0.3 (NEPS) for temperature; for the 10-m wind, it was reduced from 0.5 (REPS) to 0.38 (NEPS).

This result shows that the integrated probabilistic nowcasting system of the NEPS can improve the skill of ensemble forecasts, providing probabilistic nowcasting with high spatial and temporal resolution. Deterministic ensemble mean values over long periods of time could also be improved and add value to the NPS and the coarser REPS forecasts.

5. Two case studies

5.1. *Case 1: 11–13 July 2021*

A large-scale rainstorm affected most areas of Beijing from 11 to 13 July 2021. The maximum total accumulation precipitation in this event exceeded 200 mm. As shown in the observed 6-h accumulated precipitation fields (Figs. 10a, d), the precipitation range of this case is relatively large. The area with 6-h cumulative precipitation exceeding 25 mm is mainly concentrated in most areas of Beijing except urban areas. The 6-h mean accumulated precipitation with the REPS (Fig. 10b) corresponded well in southwest Beijing and the area with a probability greater than 80% (Fig. 10e) covered the southwest part of Beijing. However, only a few members showed signs of precipitation exceeding 25 mm in the north and northeast of Beijing, culminating in a < 10% chance for such an event (Fig. 10e); therefore, heavy precipitation was not indicated in the 6-h mean accumulated precipitation in the REPS, as seen in Fig. 10b.

Compared with the REPS, the NEPS contained the latest observation and live radar data; thus, the NEPS was able to provide more information about the development of the precipitation event. The 6-h mean accumulated precipitation output from the NEPS (Fig. 10c) corresponded well with the observations, especially in northeastern Beijing, showing the precipitation intensity exceeding 50 mm. The probability of precipitation greater than or equal to 25 mm for the same period in the NEPS (Fig. 10f) exceeded 90% in the corresponding area, which was consistent with the spatial structures of the observed precipitation.

Similar to Fig. 10, Fig. 11 showed the observation and

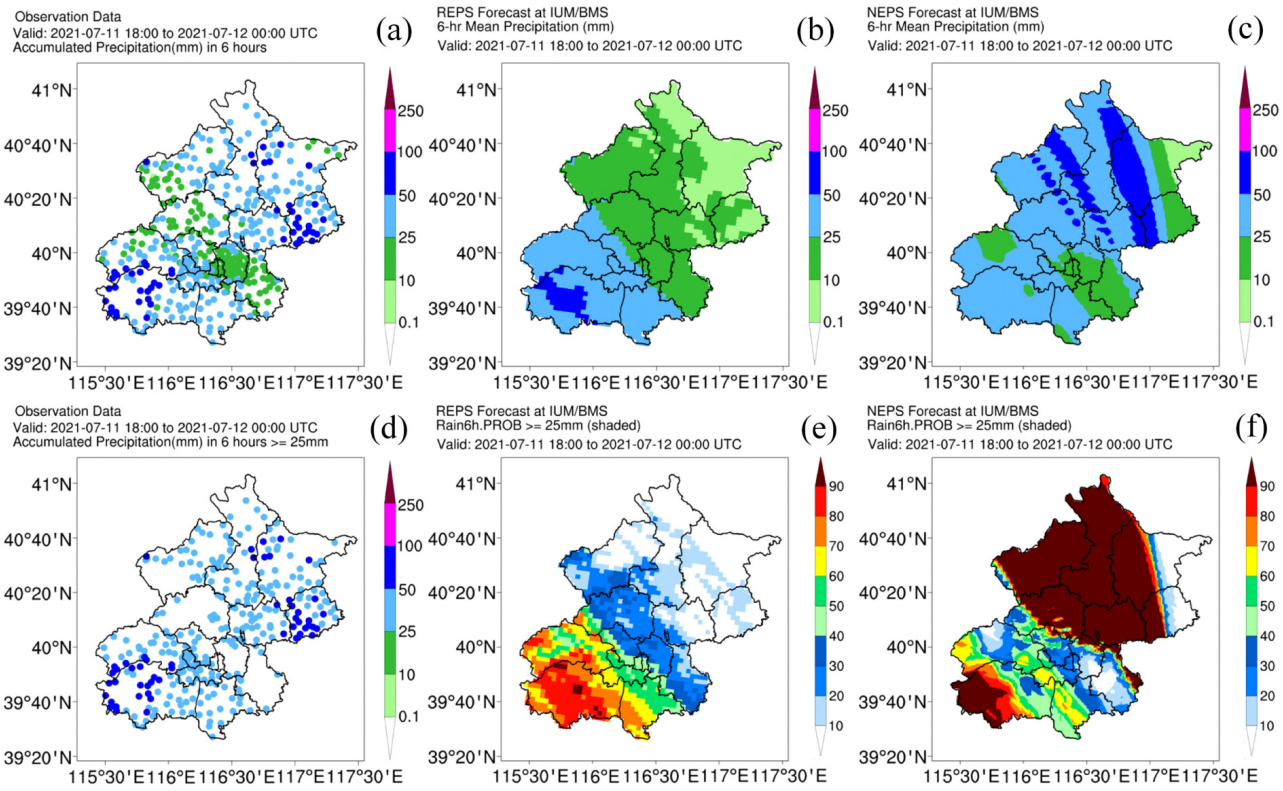

Fig. 10. (a) The observed 6-h accumulated precipitation fields, (b) the 6-h mean accumulated precipitation with the REPS and NEPS, (c) the observed 6-h accumulated precipitation fields exceeding 25 mm, and (d) the corresponding heavy precipitation probability for 6-h accumulated precipitation with the (e) REPS and (f) NEPS from 1800 UTC 11 July to 0000 UTC 12 July 2021. Color-shaded areas in panels (b) and (c) represent the 6-h mean accumulated precipitation with the NEPS and REPS, respectively. Color-shaded areas in panels (e) and (f) present the probability of 6-h accumulated precipitation over 25 mm with the REPS and NEPS, respectively.

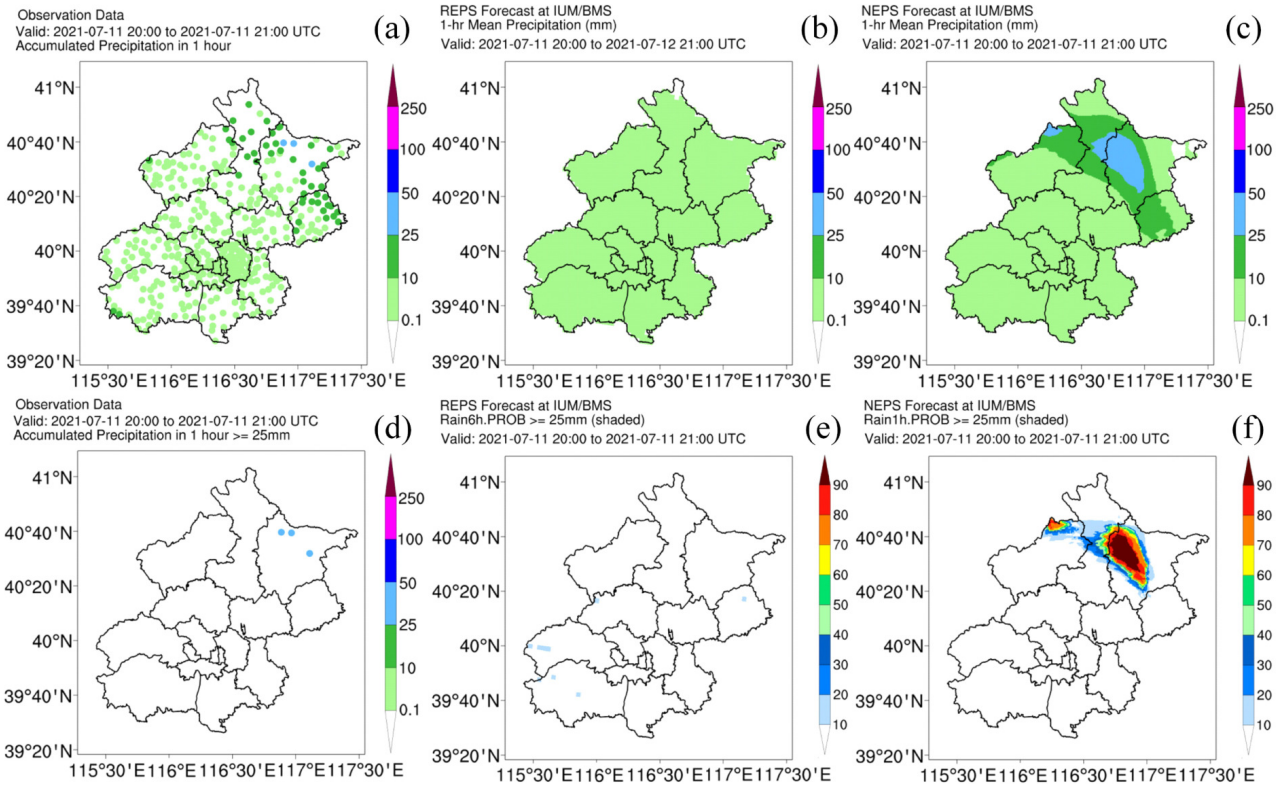

Fig. 11. (a) The observed 1-h accumulated precipitation, (b and c) the 1-h mean accumulated precipitation with the REPS and NEPS, respectively, and (d) the observed 1-h accumulated precipitation exceeding 25 mm and the corresponding heavy precipitation probability for 1-h accumulated precipitation with the (e) REPS and (f) NEPS from 2000 UTC to 2100 UTC 11 July 2021. Color-shaded areas in panels (b) and (c) represent the 6-h mean accumulated precipitation with the NEPS and REPS, respectively. Color-shaded areas in panels (e) and (f) present the probability of 6-h accumulated precipitation exceeding 25 mm with the REPS and NEPS, respectively.

forecast results with the REPS and NEPS from 2000 to 2100 UTC for 11 July 2021. It should be noted that REPS is updated twice daily at 0000 and 1200 UTC, while the NEPS is updated every hour from 0000 UTC to 2300 UTC. Specifically, we compared the forecast products updated by the REPS at 1200 UTC and those updated by the NEPS at 1800 UTC. From the observed 1-h accumulated precipitation, we can see that the precipitation in most parts of Beijing was below 10 mm, except for a few stations in the Miyun district, where the precipitation exceeded 25 mm. The 1-h accumulated precipitation amounts of the REPS were all below 10 mm, further noting that the forecast skill was very poor in the Miyun district of Beijing. Although the location of the heavy precipitation output from the NEPS was predicted to the southwest of the observation, it was still possible to track the evolution of the precipitation field with the aid of the NEPS products. At this event, these products showed a high probability of precipitation greater than or equal to 25 mm (Fig. 11f) in the Miyun district. The probability information indicated that the highest impact of the storm was expected to be in the northeast part of Beijing, consistent with where it was actually observed.

5.2. *Case2: 21–22 November 2021*

An estimate of the high wind probability is vitally important to many end users. Figure 12 illustrates the impact of the REPS and NEPS for the case of 21–22 November 2021. This case was characterized by a strong northerly flow leading to pronounced windy weather in Beijing.

The comparison of the 10-m wind speed output (Fig. 12b) from the REPS with the observational data (Fig. 12a) reveals significantly higher wind speeds in the case of the REPS, especially in the western areas of Beijing. The NEPS was able to simulate the complex local wind situation in this area, especially the 10-m wind speed forecasted 3 h in advance by the NEPS (Fig. 12c), much more realistically because of better orographic representation and observation-based nowcasting techniques. Though the NEPS did not catch the pattern in its entirety, it still delivered a better forecast for this case.

6. Conclusion

In this study, an integrated probabilistic nowcasting system (NEPS) was introduced by combining the high-resolution deterministic nowcasting approach of the NPS system with the uncertainty information provided by the REPS ensemble prediction system in North China. Here, the NEPS system provides the 1-hour updated analyses and the probabilistic forecasts in the nowcasting and short range (0–6 h) with a grid

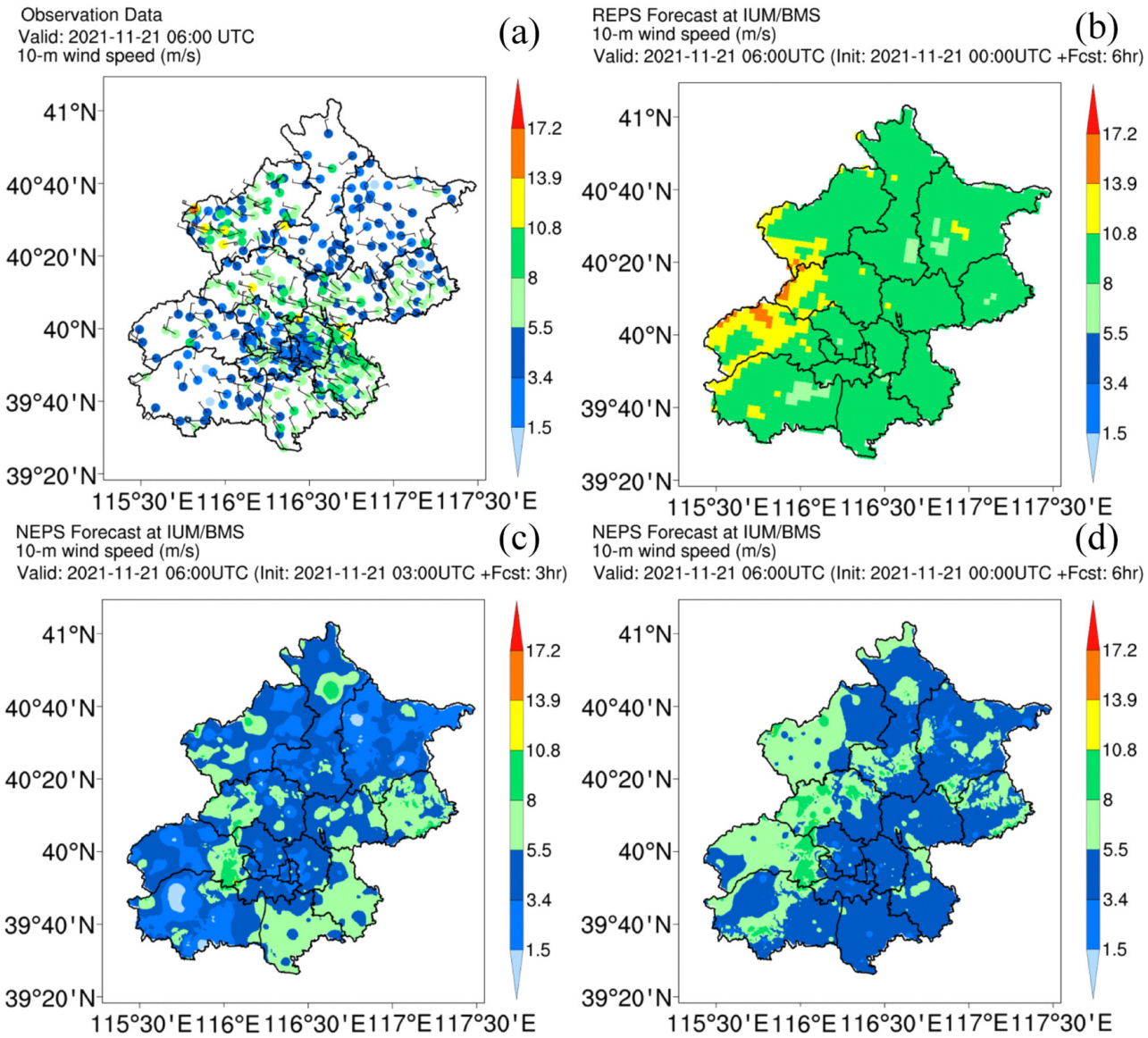

Fig. 12. (a) Observations, (b) the 10-m wind speed output from REPS 6 h in advance for 0600 UTC 21 November 2021, and the 10-m wind speed output for the NEPS (c) 3 h in advance, and (d) 6 h in advance for 0600 UTC 21 November 2021.

spacing of 500 m. It forecasts three parameters: temperature, wind, and precipitation. The comparison was made between the NEPS and NPS/REPS for the period of 1–25 July 2021. The evaluation results presented in this study showed that the NEPS performed quite favorably on surface weather variables compared to the REPS and NPS. The important results are summarized as follows:

The RMSE of mean values of wind and temperature of the NEPS were lower than those of the NPS model; the ETS and BIAS scores of precipitation demonstrated notable improvement with the NEPS. The added value of deterministic verification in the short range (0–6 h) may be attributed to the ensemble mean products, which describes the correctness of the ensemble mean estimate. Probabilistic comparison results showed that the mixed dynamic-integrated method improved the skill of the raw probabilistic forecast from the REPS. The CRPS and the percentage of outliers of aggregating wind speed and temperature were lower than those of the REPS system, and the AROC score of precipitation was greatly improved.

Finally, it was found that a subjective comparison of forecast quality in the NEPS and REPS for two selected cases also indicates that the NEPS forecasts were better than those of the REPS, as was generally evident by the differences in forecast quality implied by the ETS, CRPS, and other evaluation indicators. The probabilistic information can help in the estimation of extremes.

Generally, the role of the NEPS is to give more details about spatial intensity and distribution of the meteorological parameters. The results from this study are encouraging for the probabilistic forecast products of the NEPS system, as well as for the mean values of the deterministic forecast. In

an area with complex terrain, like Beijing, such an integrated probabilistic nowcasting system is needed to provide end users with high spatial and temporal resolution weather forecasts and the corresponding forecast uncertainties. The improved mean value and probability information can help in planning and decision-making regarding civil protection.

In the future, the accuracy of the NEPS could be further improved with a combination of post-processing methods like the neighborhood method or Bayesian model averaging, reducing the impact of systematic model errors. Besides improving the forecast skill of the NEPS, it is also of urgent concern to effectively translate and adapt the probability information to the end user requirements.

Acknowledgements. This study was supported by National Key Research and Development Program of China (Grant No. 2018YFC1506804), the Beijing Natural Science Foundation (Grant No. 8222051), the Key Innovation Team of China Meteorological Administration (CMA2022ZD04).

REFERENCES

Bouttier, F., B. Vié, O. Nuissier, and L. Raynaud, 2012: Impact of stochastic physics in a convection-permitting ensemble. *Mon. Wea. Rev.*, **140**(11), 3706–3721, https://doi.org/10.1175/MWR-D-12-00031.1.

Buizza, R., and T. N. Palmer, 1995: The singular-vector structure of the atmospheric global circulation. *J. Atmos. Sci.*, **52**, 1434–1456, https://doi.org/10.1175/1520-0469(1995)052<1434:TSVSOT>2.0.CO;2.

Buizza, R., M. Milleer, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908, https://doi.org/10.1002/qj.49712556006.

Buizza, R., P. L. Houtekamer, G. Pellerin, Z. Toth, Y. J. Zhu, and M. Z. Wei, 2005: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon. Wea. Rev.*, **133**, 1076–1097, https://doi.org/10.1175/MWR2905.1.

Cannon, A. J., and P. H. Whitfield, 2002: Downscaling recent streamflow conditions in British Columbia, Canada using ensemble neural network models. *J. Hydrol.*, **259**, 136–151, https://doi.org/10.1016/S0022-1694(01)00581-9.

Chen, D. H., and X. S. Shen, 2006: Recent progress on GRAPES research and application. *Journal of Applied Meteorological Science*, **17**(6), 773–777, https://doi.org/10.3969/j.issn.1001-7313.2006.06.014. (in Chinese with English abstract)

Chen, J., J. S. Xue, and H. Yan, 2005: A new initial perturbation method of ensemble mesoscale heavy rain prediction. *Chinese Journal of Atmospheric Sciences*, **29**(5), 717–726. (in Chinese with English abstract)

Chen, K. K., L. Y. Song, L. Yang, M. X. Chen, M. Chen, L. Han, and W. H. Cao, 2020: Research and application of a three-dimensional interpolation method for high-resolution temperature in complex terrain based on gaussian fuzzy. *Plateau Meteorology*, **39**(2), 367–377, https://doi.org/10.7522/j.issn.1000-0534.2019.00108. (in Chinese with English abstract)

Chen, M. X., and Coauthors, 2018: Enhanced weather research and forecasting in support of the Beijing 2022 Winter Olympic and Paralympic Games. *WMO Bulletin*, **67**, 58–61.

Cheng, C. L., M. Chen, M. X. Chen, F. Gao, L. Y. Song, R. Qin, L. Yang, and Y. Wang, 2019: Comparative experiments on two high spatiotemporal resolution blending algorithms for quantitative precipitation nowcasting. *Acta Meteorologica Sinica*, **77**(4), 701–714, https://doi.org/10.11676/qxxb2019.017. (in Chinese with English abstract)

Clark, A. J., 2017: Generation of ensemble mean precipitation forecasts from convection-allowing ensembles. *Wea. Forecasting*, **32**(4), 1569–1583, https://doi.org/10.1175/WAF-D-16-0199.1.

Clark, A. J., W. A. Gallus Jr., M. Xue, and F. Y. Kong, 2009: A comparison of precipitation forecast skill between small convection-allowing and large convection-parameterizing ensembles. *Wea. Forecasting*, **24**(4), 1121–1140, https://doi.org/10.1175/2009WAF2222222.1.

Clark, A. J., W. A. Gallus Jr., and M. L. Weisman, 2010: Neighborhood-based verification of precipitation forecasts from convection-allowing NCAR WRF Model simulations and the operational NAM. *Wea. Forecasting*, **25**, 1495–1509, https://doi.org/10.1175/2010WAF2222404.1.

Dance, S., E. Ebert, and D. Scurrah, 2010: Thunderstorm strike probability nowcasting. *J. Atmos. Oceanic Technol.*, **27**, 79–93, https://doi.org/10.1175/2009JTECHA1279.1.

Deng, G., and Coauthors, 2010: Development of mesoscale ensemble prediction system at national meteorological center. *Journal of Applied Meteorological Science*, **21**(5), 513–523, https://doi.org/10.3969/j.issn.1001-7313.2010.05.001. (in Chinese with English abstract)

Gao, L., J. Chen, J. W. Zheng, and Q. L. Chen, 2019: Progress in researches on ensemble forecasting of extreme weather based on numerical models. *Advances in Earth Science*, **34**(7), 706–716, https://doi.org/10.11867/j.issn.1001-8166.2019.07.0706. (in Chinese with English abstract)

Gneiting, T., and A. E. Raftery, 2007: Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, **102**, 359–378, https://doi.org/10.1198/016214506000001437.

Haiden, T., A. Kann, C. Wittmann, G. Pistotnik, B. Bica, and C. Gruber, 2011: The Integrated Nowcasting through Comprehensive Analysis (INCA) system and its validation over the Eastern Alpine region. *Wea. Forecasting*, **26**, 166–183, https://doi.org/10.1175/2010WAF2222451.1.

Hamill, T. M., 2001: Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Wea. Rev.*, **129**, 550–560, https://doi.org/10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2.

He, J., M. Chen, J. Q. Zhong, and X. Y. Hong, 2019: A study of three-dimensional radar reflectivity mosaic assimilation in the regional forecasting model for North China. *Acta Meteorologica Sinica*, **77**, 210–232, https://doi.org/10.11676/qxxb2019.005. (in Chinese with English abstract)

Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. *Wea. Forecasting*, **15**, 559–570, https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

Holton, J. R., 2004: An Introduction to Dynamic Meteorology. 4th ed. Academic Press, 535 pp.

Houtekamer, P. L., and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126**, 796–811, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2.

Houtekamer, P. L., L. Lefaivre, J. Derome, H. Ritchie, and H. L.

Mitchell, 1996: A system simulation approach to ensemble prediction. *Mon. Wea. Rev.*, **124**, 1225–1242, https://doi.org/10.1175/1520-0493(1996)124<1225:ASSATE>2.0.CO;2.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.: Atmos.*, **113**, D13103, https://doi.org/10.1029/2008JD009944.

Johnson, A., and Coauthors, 2014: Multiscale characteristics and evolution of perturbations for warm season convection-allowing precipitation forecasts: Dependence on background flow and method of perturbation. *Mon. Wea. Rev.*, **142**(3), 1053–1073, https://doi.org/10.1175/MWR-D-13-00204.1.

Kain, J. S., and J. M. Fritsch, 1993: Convective parameterization for mesoscale models: The Kain-Fritsch scheme. *The Representation of Cumulus Convection in Numerical Models*, K. A. Emanuel and D. J. Raymond, Eds., American Meteorological Society, 165–170, https://doi.org/10.1007/978-1-935704-13-3_16.

Kiktev, D., and Coauthors, 2017: FROST-2014: The sochi winter olympics international project. *Bull. Amer. Meteor. Soc.*, **98**(9), 1908–1929, https://doi.org/10.1175/BAMS-D-15-00307.1.

Kim, S. H., and H. M. Kim, 2017: Effect of considering sub-grid scale uncertainties on the forecasts of a high-resolution limited area ensemble prediction system. *Pure Appl. Geophys.*, **174**(5), 2021–2037, https://doi.org/10.1007/s00024-017-1513-2.

Kühnlein, C., C. Keil, G. C. Craig, and C. Gebhardt, 2014: The impact of downscaled initial condition perturbations on convective-scale ensemble forecasts of precipitation. *Quart. J. Roy. Meteor. Soc.*, **140**(682), 1552–1562, https://doi.org/10.1002/qj.2238.

Leith, C. E., 1974: Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, **102**, 409–418, https://doi.org/10.1175/1520-0493(1974)102<0409:TSOMCF>2.0.CO;2.

Li, Z. C., and D. H. Chen, 2002: The development and application of the operational ensemble prediction system at national meteorological center. *Journal of Applied Meteorological Science*, **13**(1), 1–15. (in Chinese with English abstract)

Lorenz, E. N., 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**(3), 321–333, https://doi.org/10.3402/tellusa.v17i3.9076.

Lorenz, E. N., 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, https://doi.org/10.1111/j.2153-3490.1969.tb00444.x.

Mellor, G. L., and T. Yamada, 1982: Development of a turbulence closure model for geophysical fluid problems. *Rev. Geophys.*, **20**, 851–875, https://doi.org/10.1029/RG020i004p00851.

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

Parrish, D. F., and J. C. Derber, 1992: The National Meteorological Center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763, https://doi.org/10.1175/1520-0493(1992)120<1747:TNMCSS>2.0.CO;2.

Peralta, C., Z. Ben Bouallègue, S. E. Theis, C. Gebhardt, and M. Buchhold, 2012: Accounting for initial condition uncertainties in COSMO-DE-EPS. *J. Geophys. Res.: Atmos.*, **117**(D7), D07108, https://doi.org/10.1029/2011JD016581.

Song, L. Y., M. X. Chen, C. L. Cheng, F. Gao, and M. Chen, 2019b: Characteristics of summer QPE error and a climatological correction method over Beijing-Tianjin-Hebei region. *Acta Meteorologica Sinica*, **77**(3), 497–515, https://doi.org/10.11676/qxxb2019.022. (in Chinese with English abstract)

Song, L. Y., M. X. Chen, F. Gao, C. L. Cheng, M. Chen, L. Yang, and Y. Wang, 2019a: Elevation influence on rainfall and a parameterization algorithm in the Beijing area. *J. Meteor. Res.*, **33**(6), 1143–1156, https://doi.org/10.1007/s13351-019-9072-3.

Stensrud, D. J., and N. Yussouf, 2007: Reliable probabilistic quantitative precipitation forecasts from a short-range ensemble forecasting system. *Wea. Forecasting*, **22**, 3–17, https://doi.org/10.1175/WAF968.1.

Suklitsch, M., A. Kann, and B. Bica, 2015: Towards an integrated probabilistic nowcasting system (En-INCA). *Advances in Science and Research*, **12**, 51–55, https://doi.org/10.5194/asr-12-51-2015.

Talagrand, O., R. Vautard, and B. Strauss, 1997: Evaluation of probabilistic prediction systems. *Proc. Workshop on Predictability*, Reading, United Kingdom, ECMWF, 1–25.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Toth, Z., and E. Kalnay, 1993: Ensemble forecasting at NMC: The generation of perturbations. *Bull. Amer. Meteor. Soc.*, **74**, 2317–2330, https://doi.org/10.1175/1520-0477(1993)074<2317:EFANTG>2.0.CO;2.

Wang, Y., and Z. H. Yan, 2007: Effect of different verification schemes on precipitation verification and assessment conclusion. *Meteorological Monthly*, **33**(12), 53–61, https://doi.org/10.3969/j.issn.1000-0526.2007.12.008. (in Chinese with English abstract)

Weidle, F., Y. Wang, and G. Smet, 2016: On the impact of the choice of global ensemble in forcing a regional ensemble system. *Wea. Forecasting*, **31**(2), 515–530, https://doi.org/10.1175/WAF-D-15-0102.1.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed., Academic Press, 467 pp.

Xie, Y. H., S. Y. Fan, M. Chen, J. C. Shi, J. Q. Zhong, and X. Y. Zhang, 2019: An assessment of satellite radiance data assimilation in RMAPS. *Remote Sensing*, **11**, 54, https://doi.org/10.3390/rs11010054.

Yang, L., M. Chen, M. X. Chen, F. Gao, R. Qin, L. Y. Song, and C. L. Cheng, 2019: Fusion of 3D high temporal and spatial resolution wind field and its application in nowcasting of severe convective weather. *Acta Meteorologica Sinica*, **77**(2), 243–255. (in Chinese with English abstract)

Yang, L., M. X. Chen, X. L. Wang, L. Y. Song, M. L. Yang, R. Qin, C. L. Cheng, and S. T. Li, 2021: Classification of precipitation type in North China using model-based explicit fields of hydrometeors with modified thermodynamic conditions. *Wea. Forecasting*, **36**, 91–107, https://doi.org/10.1175/WAF-D-20-0005.1.

Yang, L., L. Y. Song, H. Jing, M. X. Chen, W. H. Cao, and J. K. Wu, 2022: Fusion prediction and correction technique for high-resolution wind field in Winter Olympic Games area under complex terrain. *Meteorological Monthly*, **48**(2), 162–176, https://doi.org/10.7519/j.issn.1000-0526.2021.

092902. (in Chinese with English abstract)

Zhang, H. B., J. Chen, X. F. Zhi, Y. L. Li, and Y. Sun, 2014: Study on the application of GRAPES regional ensemble prediction system. *Meteorological Monthly*, **40**, 1076–1087, https://doi.org/10.7519/j.issn.1000-0526.2014.09.005. (in Chinese with English abstract)

Zhang, H. B., Y. H. Li, S. Y. Fan, J. Q. Zhong, and B. Lu, 2017: Study on initial perturbation construction method for regional ensemble forecast based on dynamical downscaling. *Meteorological Monthly*, **43**(12), 1461–1472, https://doi.org/10.7519/j.issn.1000-0526.2017.12.002. (in Chinese with English abstract)

Zhang, H. B., S. Y. Fan, M. Chen, and X. Sun, 2019: Study on a synthetic model perturbation method based on SKEB and multi physics for regional ensemble forecast. *Meteorological Monthly*, **45**(1), 17–28, https://doi.org/10.7519/j.issn.1000-0526.2019.01.002. (in Chinese with English abstract)

Zhang, H. B., Y. X. Ji, M. Chen, X. Sun, and Y. Xia, 2022: Study on the EDA initial condition perturbation method for ensemble prediction system based on observation perturbation. *Meteorological Monthly*, **48**(4), 406–417, https://doi.org/10.7519/j.issn.1000-0526.2021.102301. (in Chinese with English abstract)

Zhang, R. H., and X. S. Shen, 2008: Development of new generation operational numerical prediction system GRAPES. *Chinese Science Bulletin*, **53**(20), 2393–2395. (in Chinese)

Zhong, Y. L., J. Chen, J. Wang, K. Lü, and X. L. Li, 2017: Evaluation of the forecast for landed Typhoons by GRAPES-REPS regional ensemble prediction system. *Journal of Tropical Meteorology*, **33**(6), 953–964, https://doi.org/10.16032/j.issn.1004-4965.2017.06.016. (in Chinese with English abstract)