

# JGR Atmospheres

## RESEARCH ARTICLE

10.1029/2022JD038143

### Key Points:

- Reflectivity structure predicted by the 4-km Weather Research and Forecasting model are evaluated using 3D radar observations
- The model overpredicts the intensity of reflectivity at lower levels and fails to reproduce the diurnal variation of object height
- The failure to correctly reproduce the structure of smaller reflectivity objects is primarily responsible for the deviations

### Correspondence to:

M. Xue,  
[mxue@ou.edu](mailto:mxue@ou.edu)

### Citation:

Zhu, K., Xue, M., Yang, N., & Zhang, C. (2023). How well does 4-km WRF model predict three-dimensional reflectivity structure over China as compared to radar observations? *Journal of Geophysical Research: Atmospheres*, 128, e2022JD038143. <https://doi.org/10.1029/2022JD038143>

Received 8 NOV 2022  
Accepted 24 MAR 2023

## How Well Does 4-km WRF Model Predict Three-Dimensional Reflectivity Structure Over China as Compared to Radar Observations?

Kefeng Zhu<sup>1</sup> , Ming Xue<sup>2,3</sup> , Nan Yang<sup>4</sup> , and Chenyue Zhang<sup>5</sup>

<sup>1</sup>Nanjing Joint Institute for Atmospheric Sciences and CMA Key Laboratory of Transportation Meteorology, Nanjing, China, <sup>2</sup>Key Laboratory of Mesoscale Severe Weather, Ministry of Education and School of Atmospheric Sciences, Nanjing University, Nanjing, China, <sup>3</sup>Center for Analysis and Prediction of Storms and School of Meteorology, University of Oklahoma, Norman, OK, USA, <sup>4</sup>CMA Basin Heavy Rainfall Key Laboratory, Hubei Key Laboratory for Heavy Rain Monitoring and Warning Research, Institute of Heavy Rain, China Meteorological Administration, Wuhan, China, <sup>5</sup>Aviation Meteorological Center, Air Traffic Management Bureau, Civil Aviation Administration of China, Beijing, China

**Abstract** The performance of 4-km Weather Research and Forecasting model in predicting reflectivity structure in China over three summer months is investigated using three-dimensional (3D) reflectivity observations. Three verification domains, namely Southern China (SC), Central and Eastern China, and North China, that correspond to the three major rainfall centers of mainland China are selected. Results show that the forecasts reproduce the distribution and diurnal variation of precipitation well, but significant differences exist in the vertical distributions of the predicted and observed reflectivities. In observations, the highest frequency (of reflectivity  $\geq 35$  dBZ) occurs between 3 and 6 km, whereas it is at the surface in the forecast. The forecasts tend to over-predict reflectivity intensity at the lower levels, especially in SC. Further evaluation using object-based verification methods show that the forecasts greatly underestimate the afternoon peak frequency of precipitation clouds with reflectivity  $> 30$  dBZ. The forecasts fail to reproduce the diurnal variation of 35 dBZ mean and maximum height of the objects, producing less variation than observations. Analyses show that the failure in properly reproducing small-scale reflectivity objects (with diameter  $< 100$  km) is primarily responsible for the underestimation of the mean and maximum object heights. Evaluation using additional 3D information show that the forecasts tend to produce a greater proportion of faster-moving small-scale objects. This study reveals that the simulation of the 3D structure of precipitation clouds in terms of reflectivity remains a great challenge, especially for smaller convective cells.

**Plain Language Summary** The simulation of precipitation clouds is associated carries the largest uncertainty in numerical weather prediction and climate models. With continued advancement in observations and numerical models, forecasts of horizontal distribution, intensity, diurnal variation and propagation of precipitation have been greatly improved. However, faithful representation of three-dimensional (3D) cloud structure remains a great challenge, with the lack of systematic evaluations being one of the key causes. In this study, 3D gridded radar reflectivity observations are used to evaluate the precipitation clouds structure (with reflectivity  $\geq 30$  dBZ) predicted at a convection-permitting-resolution (CPR) over three summer months of 2016 in China. The evaluation reveals differences between the simulated and observed reflectivity structures, including the reflectivity intensity profile, the 35 dBZ mean and maximum height, the fraction of fast-moving objects. CPR models still have difficulties in simulating 3D reflectivity structures especially for relatively small-scale objects (with diameter  $< 100$  km).

### 1. Introduction

Precipitation is an important component of hazardous weather events. With improvements in model resolution (Saito et al., 2006; Zhu et al., 2018) and the assimilation of satellite (Lin et al., 2017; Zhu et al., 2019) and radar data (Benjamin et al., 2016; Ding et al., 2022; Hu & Xue, 2007), as well as better parameterization of microphysics (B. R. Brown et al., 2017; Morrison et al., 2005, 2015), forecasting of precipitation has been greatly improved in recent decades. Evaluations show that most convection-permitting-resolution (CPR) models can well predict the horizontal distribution of precipitation, as well as its diurnal variation and propagation (Li et al., 2020; Yu et al., 2020). However, CPR models still suffer from great uncertainty in cloud simulation. In-depth evaluations

of simulated three-dimensional (3D) cloud structure is an effective way to expose the potential issues of microphysical processes, and help improve precipitation forecasts.

Many studies use satellite observations to reveal potential issues in cloud simulations (Griffin et al., 2021; Otkin & Greenwald, 2008). For instance, Jankov et al. (2011) evaluated different microphysics schemes using satellite observations and found that, whilst there was good agreement with respect to cumulative precipitation, there were notable differences in the simulation of cloud properties. Specifically, the Lin microphysics scheme tended to overestimate the high cloud, while the other schemes underestimated the mid-level cloud. A similar result was found by Yang et al. (2022) in a low-echo-centroid (LEC) extreme rainfall case. Moreover, all microphysics schemes significantly overestimated the cloud top height, suggesting overestimation of cloud ice in the upper level. Jones et al. (2018) evaluated a short-range forecast system using synthetic satellite brightness temperature and cloud products. They found that the double-moment NSSL variable density scheme overestimated the areal coverage of convectively generated cirrus clouds. Adjustments to the fall speed and collection efficiencies were made, after which the simulation of upper-level cloud properties improved. In short, evaluation using satellite observations provides insights into a variety of issues related to forecast uncertainty.

Unlike satellite observations, radar observations can detect in-cloud information and shed light on potential issues regarding the microphysical processes related to precipitating particles. Aligo et al. (2018) made a series of improvements to the Ferrier-Aligo microphysical scheme in the North American Mesoscale Forecast System based on an evaluation of reflectivity structure, including the ice nucleation activation process, the transport of rimed ice to the extremely cold region below  $-40^{\circ}\text{C}$ , and the fall speed of ice particles. The simulation of deep convection and stratiform cloud structure were slightly improved. Gallus and Pfeifer (2008) compared and evaluated five microphysical schemes of the Weather Research and Forecasting (WRF) model and found that they yielded large differences in simulated reflectivity in the melting layer. Moreover, all five schemes tended to overestimate the intensity, due to the overestimation of the snow mixing ratio and the assumption of a very broad snow size distributions near the zero line. Snow particles can directly affect the raindrop distribution in the lower layer, resulting in stronger simulated reflectivity, especially in the stratiform region. B. R. Brown et al. (2017) reported a similar issue. By adjusting the tunable parameter in the snow melting processes, the simulated drop sizes of the Thompson-Eidhammer microphysics scheme (Thompson & Eidhammer, 2014) in deep convective regions were slightly degraded.

Other studies have used more advanced dual-polarization Doppler radars to evaluate the microphysical processes in clouds. Polarimetric observations are far more meaningful in their representation of particle properties than reflectivity observations. For example, differential reflectivity ( $Z_{\text{DR}}$ ) is sensitive to the shape and orientation, while specific differential phase ( $K_{\text{DP}}$ ) is sensitive to the number concentration (Bringi & Chandrasekar, 2001). Whether a numerical model can correctly simulate the observed polarization signature indirectly indicates the model's ability in simulating the dynamics (e.g., the  $Z_{\text{DR}}$  column corresponds to the strong updraft region) and particle properties (e.g., large  $Z_{\text{DR}}$  indicates large raindrop sizes) in clouds (Kumjian et al., 2014). By comparing the retrieved spatial pattern of ice particles with the forecast, potential issues relating to microphysical processes that produce the particles can be identified. For example, too many wet aggregates indicate overly active melting process, too many graupel/rimed aggregates indicate excessive riming and deposition. Barnes and Houze (2016) found that the ice microphysical processes, such as aggregation and riming, were generally too strong in all the microphysics schemes they investigated. Excessive riming was responsible for the overestimation of snow and graupel in the upper level as mentioned above, which ultimately led to extremely large raindrops at low levels (B. R. Brown et al., 2016; Chen et al., 2021). Large raindrops can also be caused by tunable parameters within the raindrop breakup parameterization (Morrison et al., 2012). By decreasing the cutoff radius of rain breakup and increasing the breakup efficiency, the simulated raindrop size distribution can be greatly improved (Zhou et al., 2022).

The above avenues of research have thus far mostly drawn upon case studies. There are common issues, such as overly frequent riming and excessively large raindrop sizes at the surface, resulting in exaggerated reflectivity strength. However, the biases in simulated cloud properties, such as the radar echo top and cloud top, are highly sensitive to the selected cases and microphysics schemes. Until now, there have been few statistical evaluations of precipitation clouds using 3D radar observations. Wang et al. (2021) evaluated the 3D radar reflectivity structure simulated by the Energy Exascale Earth System Model, which runs at a horizontal grid resolution of approximately  $1^{\circ}$ , and found that it failed to capture mesoscale convective systems observed over the central United

States. The model also significantly underestimated the radar echo top. Unlike global forecasts, CPR models can explicitly resolve most convection systems and provide greatly improved precipitation forecasts. In a recent study, our group evaluated the performance of a CPR model (the 4 km WRF model) running in real-time over China (Zhu et al., 2018). Results showed that the model simulated the distribution and diurnal variation of precipitation well. In the present study, 3D radar observations were used as additional observations to evaluate the forecasted precipitation clouds (with reflectivity  $\geq 30$  dBZ). The purpose was to assess and discuss the general performance of the widely used double-moment Morrison microphysics scheme in the prediction of 3D reflectivity structure over China.

The rest of the paper is organized as follows. In Section 1, the data and the verification metric for evaluating the precipitation clouds are introduced. Section 2 compares and discusses the simulated and observed 3D reflectivity structures, followed by a summary and conclusions in Section 3.

## 2. Forecasts and Observations

### 2.1. Forecast Data

The forecast data were from real-time 4 km WRF forecasts run at Nanjing University, China, every summer since 2013 (Zhu et al., 2018). These forecasts cover the entire mainland of China and have 51 vertical levels. The initial and lateral boundary conditions were supplied by Global Forecast System real-time analyses and forecasts, respectively (grid spacing of  $0.5^\circ$  before 27 June 2016 and  $0.25^\circ$  thereafter). The key physics parameterizations were the Morrison double-moment microphysics scheme (Morrison et al., 2005); the Asymmetrical Convective Model, version 2, planetary boundary layer scheme (Pleim, 2007); the Pleim-Xiu land surface and surface layer schemes (Pleim & Xiu, 1995); and the Community Atmosphere Model radiation scheme (Collins et al., 2004). More details on the configurations can be found in Zhu et al. (2018).

The present study focuses on evaluating the reflectivity structure during summer (June–August) 2016. In a recent study, we have used this data set to study the thermal and dynamical factors that lead to the biases of precipitation forecast in South China (Zhu et al., 2021), but have not discussed the microphysical factors. This study is a supplement to the previous research, which aims to understand the possible issues with microphysical processes. The 12–36 hr forecasts initiated every day at 12:00 UTC were examined. Figure 1 shows the three selected verification domains for evaluation, which are Southern China (SC), Central and Eastern China (CEC), and North China (NC). As the summer monsoon proceeds to the north, the rain band moves from South China to North China. These selected domains are the three major rainfall centers during the summer period. Details will be introduced in Section 3.1.

The 3D simulated reflectivity was generated using the built-in algorithm in WRF (Stoelinga, 2005), in which the reflectivity factor  $Z$  is calculated as the sixth moment of the drop size distribution

$$Z = \int N(D)D^6 dD, \quad (1)$$

where

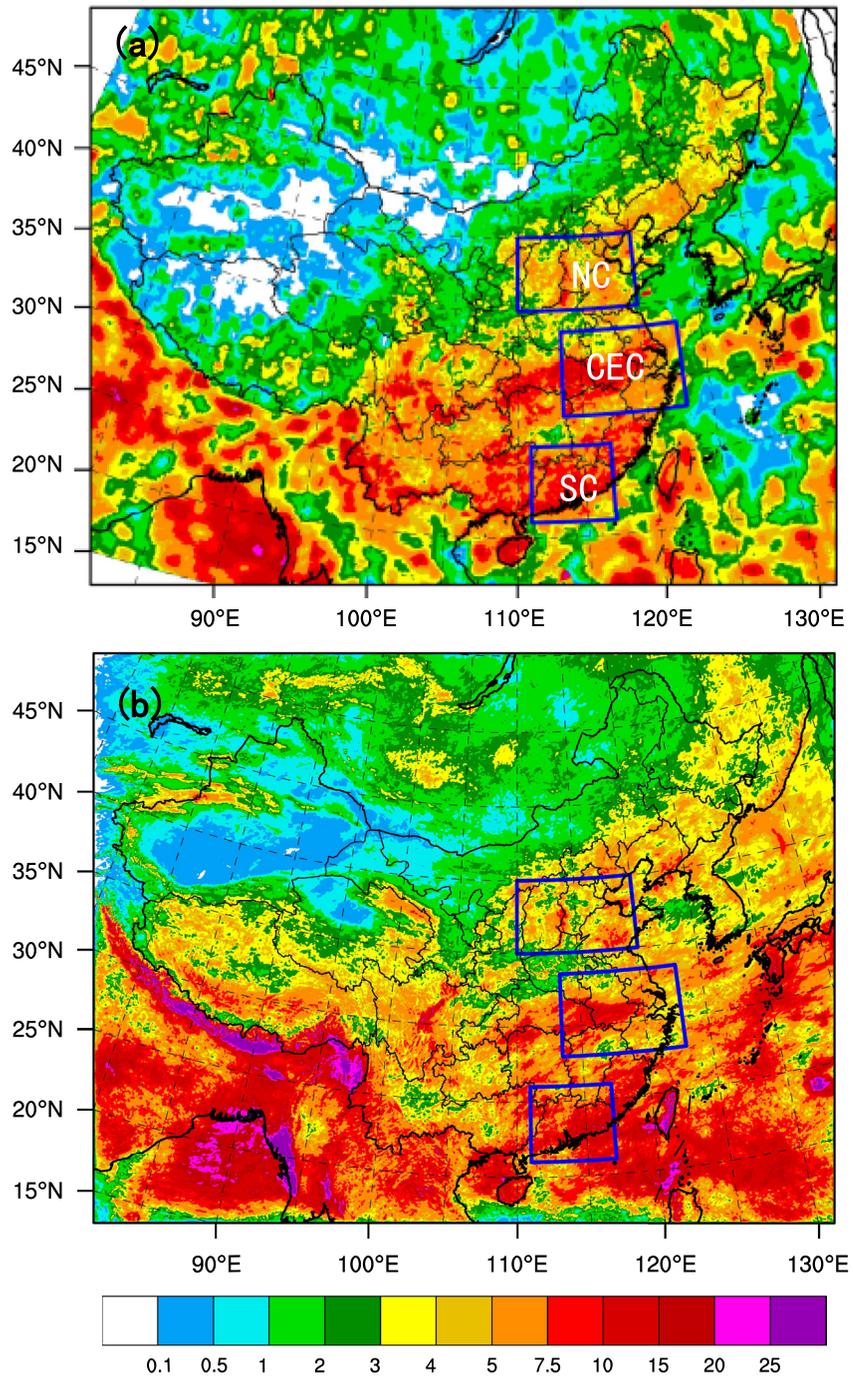
$$N(D) = N_0 D^\alpha e^{-\lambda D}. \quad (2)$$

Here,  $N_0$ ,  $\alpha$ , and  $\lambda$  are the intercept, shape, and slope parameters of the gamma distribution, respectively. Using Equations 1 and 2, the equivalent reflectivity factor of rain can be deduced as follows:

$$Z_{e,r} = N_{0r} \times \frac{\Gamma(7 + \alpha_r)}{(\lambda_r)^{7+\alpha_r}}. \quad (3)$$

By assuming Rayleigh scattering of an irregular ice particle to be the same as a solid ice sphere of equivalent mass (Marshall & Gunn, 1952), the equivalent reflectivity factor of snow ( $Z_{e,s}$ ) and graupel ( $Z_{e,g}$ ) can be calculated as follows:

$$Z_{e,s} = \frac{|K_i|^2}{|K_l|^2} \times \left(\frac{\rho_s}{\rho_l}\right)^2 \times N_{0s} \times \frac{\Gamma(7 + \alpha_s)}{(\lambda_s)^{7+\alpha_s}}, \quad (4)$$



**Figure 1.** (a) Observed and (b) 4 km Weather Research and Forecasting simulated summer-mean precipitation (unit:  $\text{mm}\cdot\text{day}^{-1}$ ) in 2016. The blue frames mark the locations of three selected evaluation regions, which are Northern China (NC), Central and Eastern China (CEC), and Southern China (SC).

$$Z_{e-g} = \frac{|K_i|^2}{|K_l|^2} \times \left(\frac{\rho_g}{\rho_l}\right)^2 \times N_{0g} \times \frac{\Gamma(7 + \alpha_g)}{(\lambda_g)^{7+\alpha_g}}, \quad (5)$$

where

$$\lambda_x = \left(\frac{\pi}{6} \rho_x \times \frac{\Gamma(4 + \alpha_x)}{\Gamma(1 + \alpha_x)} \times \frac{N_{Tx}}{q_x}\right)^{1/3}, \quad (6)$$

$$N_{0x} = N_{Tx} \frac{(\lambda_x)^{(1+\alpha_x)}}{\Gamma(1+\alpha_x)}. \quad (7)$$

Here,  $x$  represents the hydrometer species of rain, snow, and graupel,  $\Gamma$  is the value of the gamma function,  $|K_i|^2 = 0.176$  and  $|K_l|^2 = 0.93$  are the dielectric factors of ice and liquid water, respectively.  $\rho_i$  is the density of ice, and  $\rho_x$ ,  $N_{Tx}$ , and  $q_x$  are the density, forecasted total number concentration, and mixing ratio of each species, respectively. For the current double-moment Morrison scheme, the shape parameter  $\alpha_x = 0$ .

The equivalent reflectivity (in dBZ) can then be summarized as follows:

$$Z_e(\text{in dBZ}) = 10 \times \log_{10}(10^{18} \times (Z_{e,r} + Z_{e,s} + Z_{e,g})). \quad (8)$$

## 2.2. Observations

Radar observations within the three selected domains were collected. Raw data were first quality-controlled using software developed at Nanjing University (Huang et al., 2018). Then, the radar mosaic tools developed by the Advanced Regional Prediction System (Xue et al., 2000) were used to generate observations on a  $0.036^\circ \times 0.036^\circ$  ( $\sim 4$  km) grid. The remapping process from the radar volume in polar coordinates to a Cartesian grid was accomplished via a least squares fit to a local polynomial function, which was quadratic in the horizontal and linear in the vertical. The final 3D grid of reflectivity observations had 31 vertical levels (Zhang et al., 2005), 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 10, 11, 12, 13, 14, 15, 16, and 18 km. The composite reflectivity was calculated as the maximum value within a column. Here, to facilitate comparison, WRF forecasts are also interpolated to the same vertical grid before comparison.

Figure 2 shows both the radar-observed and the forecasted reflectivity for a Meiyu Front case. In general, high-quality observations are apparent in both the horizontal and vertical directions. With a dense radar network, observations cover the entire Meiyu Front system. The only issue is at the lower level. Due to the limitation of the lowest elevation angle and the occlusion effect of surface objects, observations show a small amount of missing data near the ground (Figures 2c and 2e). In comparison, for the forecasts, only data below the ground are set to missing values (Figures 2d and 2f). Therefore, in Section 3.2, considering the inconsistency in valid grid numbers between the observation and forecast, the intensity and normalized frequency are used for comparison. The intensity is defined as the mean of grid points with reflectivity larger than 35 dBZ. Reflectivity ( $R$ ) for a given time and location could be expressed as  $R(iday, ih, ilon, ilat, iz)$ . The variables  $iday$ ,  $ih$ ,  $ilon$ ,  $ilat$ , and  $iz$  represent day, hour of the day, longitude, latitude, and vertical level, respectively. The intensity ( $I$ ) for a given latitude and height is calculated as follows:

$$I(ilat, iz) = \frac{\text{sum}(\text{where}(R(:, :, :, ilat, iz) \geq 35, R(:, :, :, ilat, iz), 0))}{\text{sum}(\text{where}(R(:, :, :, ilat, iz) \geq 35, 1, 0))}. \quad (9)$$

The frequency is normalized by the maximum frequency of all points. The normalized frequency ( $F_{nor}$ ) for a given latitude and height is calculated as follows:

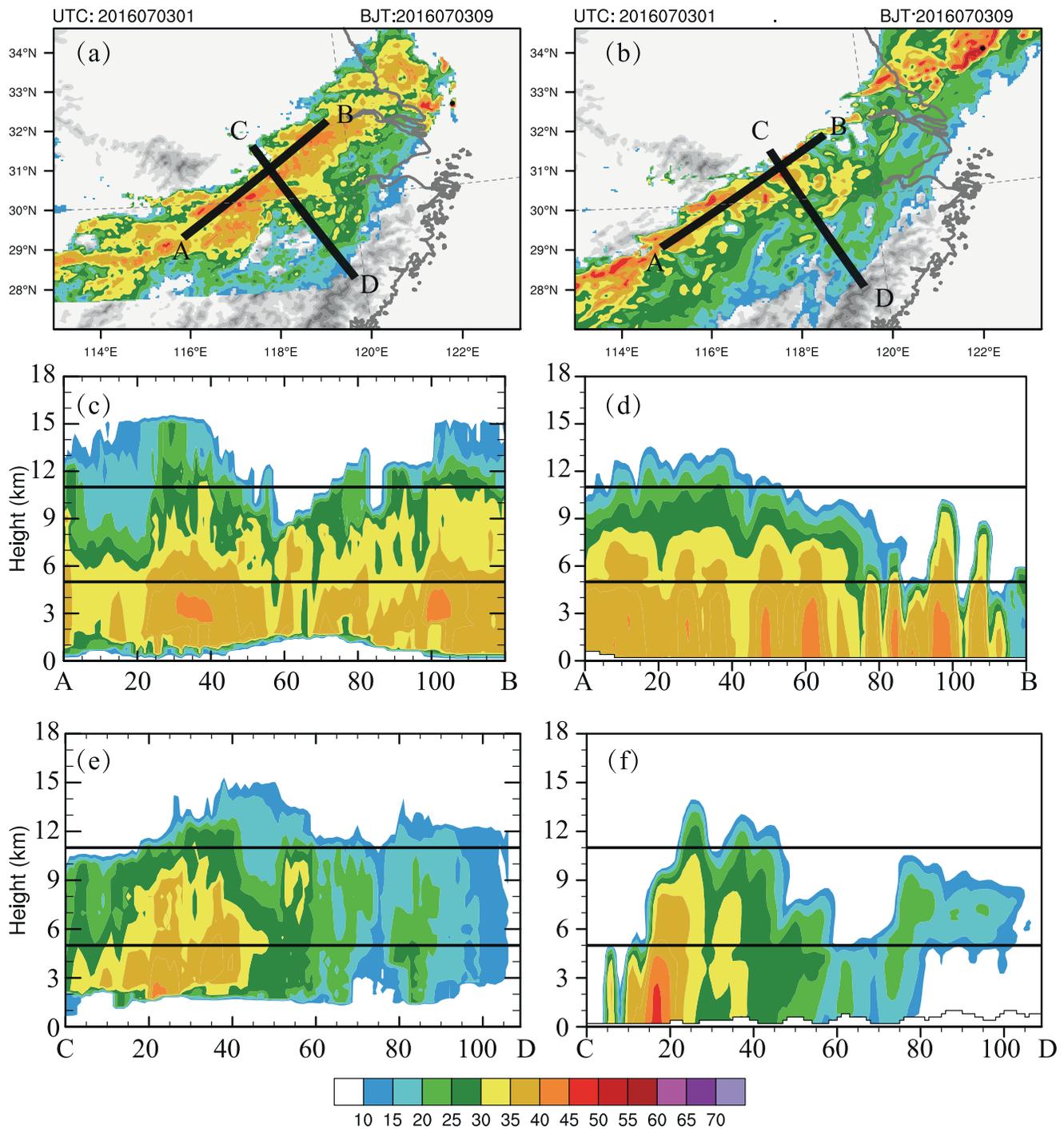
$$F_{nor}(ilat, iz) = \frac{\text{sum}(\text{where}(R(:, :, :, ilat, iz) \geq 35, 1, 0))}{\max(\text{sum}(\text{where}(R(:, :, :, ilat, iz) \geq 35, 1, 0)))}. \quad (10)$$

The function “where” performs array assignment based on a conditional expression (the first variable). If the reflectivity value is larger than 35 dBZ, a true value (the second variable) is assigned to the array. Otherwise, it is set to 0.

Other verification data used in this study were the gridded hourly precipitation data from a  $0.05^\circ \times 0.05^\circ$  station-satellite-radar multi-source merged precipitation analysis product ( $15^\circ$ – $60^\circ$ N,  $70^\circ$ – $140^\circ$ E) produced by the China Meteorological Administration (Pan et al., 2015; Shen et al., 2014). Many studies have used this data set to evaluate the CPR model performance in capturing the distribution and diurnal variation of precipitation over China (Cai et al., 2021; Li et al., 2021, 2020; Wu et al., 2018; Y. Zhao et al., 2020; Zhu et al., 2021).

## 2.3. Verification Metrics

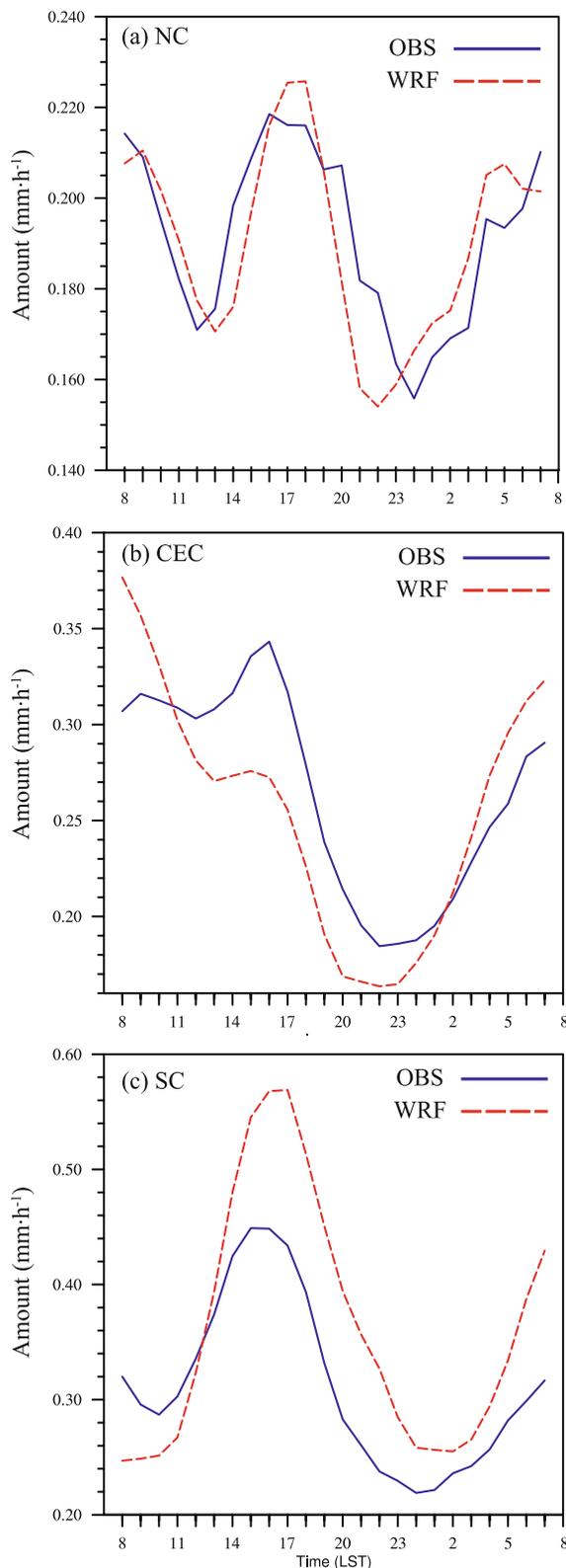
CPR models can successfully predict precipitation structures, but usually cannot exactly match observations in terms of size, intensity, structure, orientation, location, and timing beyond several hours of forecasting. Therefore, in this study, the object-based diagnostic evaluation (MODE), proposed by Davis et al. (2006), in the Model



**Figure 2.** A Meiyou Front case at 01:00 UTC 3 July 2016, showing the composite reflectivity of the (a) radar observations and (b) 4 km Weather Research and Forecasting forecast. The black rectangles AB and CD indicate the locations of 10 slices used for the calculation of vertical composite reflectivity in panels (c, d) and (e, f), respectively. The horizontal black lines in (c–f) indicate the mean height of forecasted temperature of 0°C and –40°C.

Evaluation Tools (B. Brown et al., 2021) was employed. MODE is a spatial verification method (Gilleland et al., 2009) that can avoid the double penalty issue in point-to-point verification metrics and provides more informative measures of finer grid forecasts (Gallus, 2010; Skinner et al., 2018).

The properties of objects such as area and intensity, as well as the number of objects, were calculated using two-dimensional composite reflectivity. The convolution threshold for identifying objects was set to 30 dBZ, the recommended threshold by Xu et al. (2008). Reflectivity below this threshold mainly corresponds to light rain.



**Figure 3.** Time series of observed (blue) and forecasted (red) hourly mean precipitation during summer 2016 for regions (a) North China (NC), (b) Central and Eastern China (CEC), and (c) Southern China (SC) in Figure 1.

Also, for the purpose of tracking a precipitation system through its lifetime, a low threshold as 30 dBZ is preferred so that the object's representation of a precipitation system can be as complete as possible. Too large a threshold may cause a split of the object, making system tracking difficult. The convolution radius is another key parameter for object identification, which is set to four grid points.

As soon as the object is identified, the object's boundary could be obtained. We refill the object with 3D gridded reflectivity. The object's properties including the mean and maximum height of the 35 dBZ were calculated. First, the maximum height of 35 dBZ of a column within the object was calculated. To avoid cases with a singular point, at least two continuous points in a column larger than 35 dBZ were required. Otherwise, the maximum height of 35 dBZ for that column is assigned as missing. Then, the mean and maximum height were calculated using all columns within the object.

Additionally, to evaluate the performance of WRF in predicting the lifetime and speed of the precipitation system, the newly developed MODE Time Domain (MTD) tools, which incorporate additional information in time, were employed. For this evaluation, the composite reflectivity at 10 min intervals was used. All settings for identifying objects were the same as the MODE configuration.

### 3. Results

#### 3.1. Evaluation of WRF Precipitation Forecasts

Precipitation in China is closely associated with the intraseasonal movement of the East Asian summer monsoon (Zhu et al., 2018). Figure 1 shows the observed and forecasted daily mean precipitation during summer 2016. For the observation, there are three main precipitation zones. The first zone is in SC. It is associated with the pre-summer rainy season (Jiang et al., 2017; Luo et al., 2017) in early June and typhoon precipitation over the summer period. The second zone is in CEC. It is mainly contributed by Meiyu front precipitation as the western Pacific subtropical high (WPSH) makes its first jump from south to north. The Meiyu season usually lasts from the middle of June to the middle of July (Bao et al., 2011; Y. Zhao et al., 2020). Affected by a super El Niño, the Meiyu season of 2016 produced much more precipitation than the climatological annual mean, resulting in one of the four most severe summers of flooding over the Yangtze River region since 1990 (J. Zhao et al., 2018). The third zone is in NC. It usually enters its rainy season in mid-July as the WPSH makes its second jump. In general, the rainfall distribution was successfully simulated and the rainfall centers are also well captured. The biggest issues, as we have pointed out in a previous study, are the overestimation of rainfall amount and the slight mislocation of the predicted rainfall center in South China (Zhu et al., 2021).

Figure 3 presents the diurnal variation in precipitation of the three selected domains (NC, CEC, and SC). Both the observation and forecast show a clear bimodal pattern, with one peak in the afternoon and the other in the early morning. In region SC, the rainfall amount is mostly overestimated. In region CEC, the precipitation amount during the afternoon is underestimated, whereas it is overestimated at nighttime. Comparing the above two regions, the predicted rainfall amount in region NC is clearly better (see also Figure 1). However, the timing of the afternoon peak is delayed by roughly 2 hr, whereas the morning peak is 3 hr early.

Overall, the WRF forecast successfully simulated the rainfall distribution and its diurnal variation, except for some biases in intensity and timing. Among the three verification domains, SC was subject to the largest intensity bias, while NC witnessed the lowest.

### 3.2. Evaluation of WRF 3D Reflectivity Forecasts

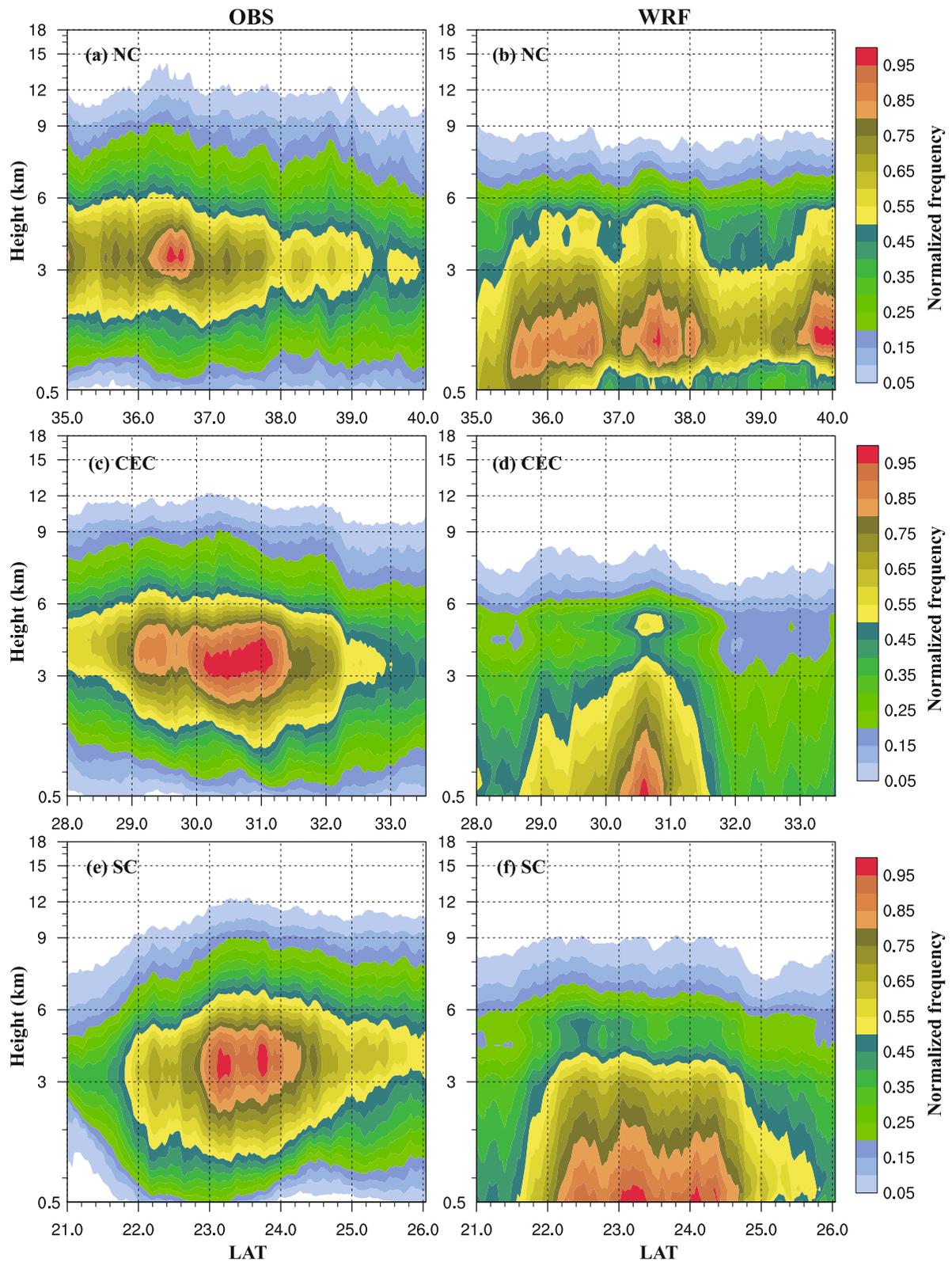
Figure 4 shows the normalized frequency of reflectivity larger than 35 dBZ during the afternoon period of summer 2016. For the observation, the highest frequency is found between 3 and 5 km, and height decreases somewhat as we move from south to north. We see higher reflectivity intensity at around 3 km for the NC and CEC regions, and above 6 km in SC (Figure 5). These layers of higher reflectivity correspond to the melting level. As ice particles fall through the melting level, the partially melt ice particles become coated with liquid water but still retain the lower density and larger sizes of the ice particles, leading to increased reflectivity because the dielectric coefficient of water is larger than that of ice (see Equations 3–5). Meanwhile, it is easier for snowflakes wrapped in water to stick and hook together; that is, the processes of collision and polymerization between particles are enhanced, resulting in larger particles. Both greatly increase the reflectivity. For heights below 3 km, when an ice particles completely melt, they turn into smaller spherical raindrops under the action of surface tension. The fall speed of raindrops is much larger than that of ice crystals and snowflakes, which greatly reduces the number concentration below the melting layer and in turn also reduces the intensity of reflectivity. In terms of the intensity, we also see another high reflectivity zone at 6–9 km in the observations (left column of Figure 5). The high mean intensity at those levels occurs because only relatively few points related with deep convection can reach those levels and be counted (see Figure 4). In comparison, at around 3 km the intensity may include many stratiform clouds. This explains why the intensity between 6 and 9 km is high even though the frequency of high reflectivity is low. For the three regions, the overall distribution pattern of observed normalized frequency and intensity (high frequency between 3 and 5 km and two high reflectivity bands with one around 3 km and the other between 6 and 9 km) are similar.

In the forecast, the highest frequency of reflectivity exceeding 35 dBZ is however found below the 3 km level, with the maximum to be at the ground level in the CEC and SC regions (right column of Figure 4). The lower frequency approaching the ground in NC is due to the predicted afternoon convection occurring mostly over mountains. In the other two regions, the height of the mountains is mostly below 500 m (not shown). In the forecast, the mean intensity of reflectivity exceeding 35 dBZ is also the largest near ground (right column of Figure 5).

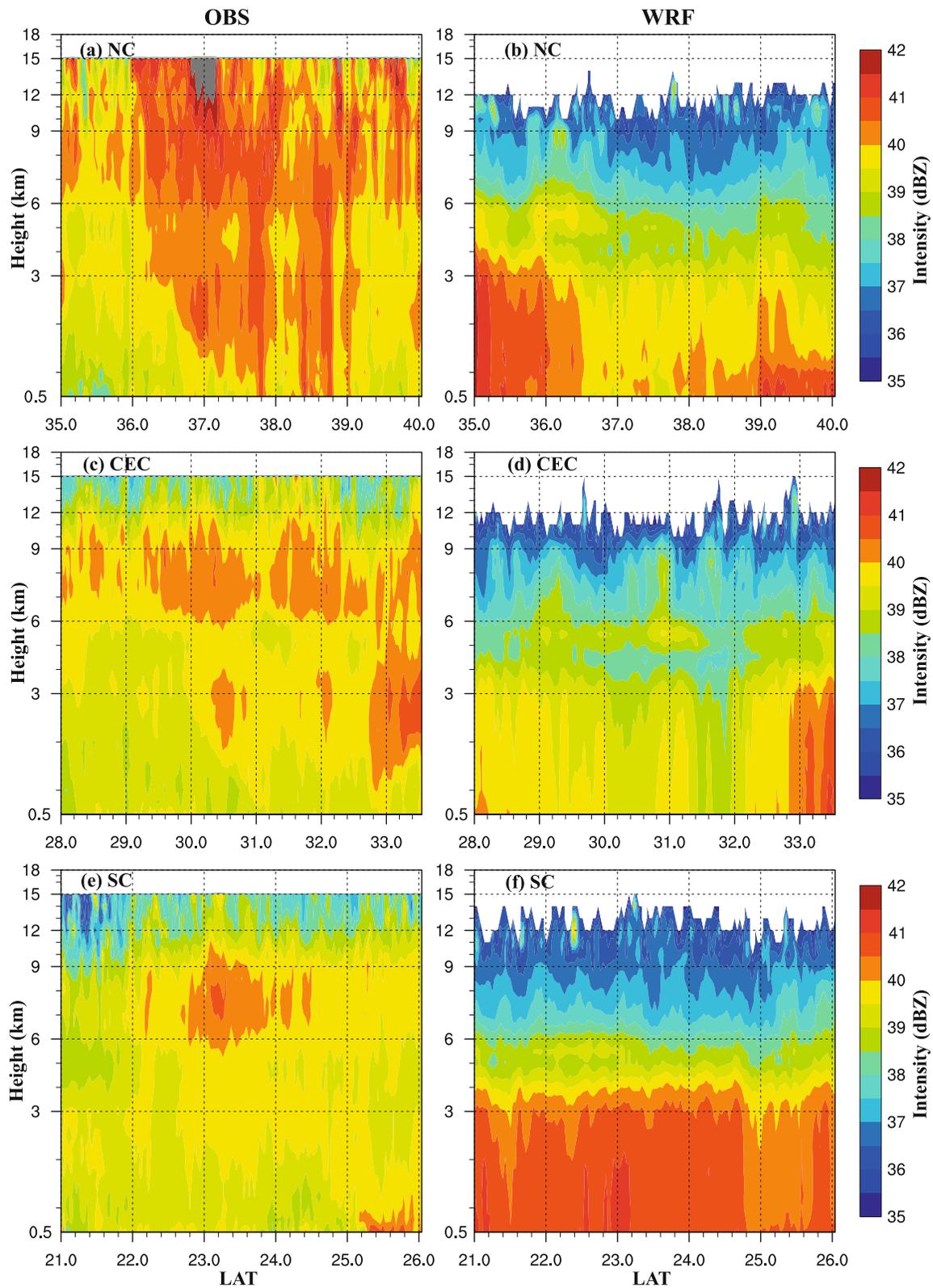
There are several possible reasons for the stark difference between the forecast and observations. First, the frequency of reflectivity exceeding 35 dBZ can be significantly affected by the near-ground data coverage. Due to the 0.5° elevation of the lowest radar beam, the Earth curvature effect as well as possible terrain blockage, there may not be radar data coverage below 2 km height (see e.g., Figures 2c and 2e); this will cause under-estimation of the reflectivity frequency in the observations, although this should not affect height above 2 km much. In the observations, the frequency starts to decrease from around 3 km height (Figure 4), placing a maximum above 3 km; this result should be trustable (see also Figure 2c). In comparison, the mean intensity of reflectivity near the surface is less affected by the data coverage, because the average is performed only at locations observations are available.

Errors with microphysics schemes can also cause the very different reflectivity structure near surface. The Morrison scheme tends to overestimate raindrop size owing to imperfect microphysical representations, such as a low breakup efficiency of raindrops (Morrison et al., 2012) and a thicker melting layer (Barnes & Houze, 2016). The bigger the raindrop size, the higher the efficiency of raindrop collection and the weaker the evaporation, resulting in larger raindrops near the surface and hence higher reflectivity. This can be seen with the Meiyu front case shown in Figure 2. The observed reflectivity maximum is at around 3 km and decreases toward the ground but the predicted reflectivity core is at the ground. Among the three regions, the overestimation near the surface is the most obvious in SC. This is because water vapor content in SC is greatly overestimated in the forecasts (Zhu et al., 2021), which in turn affects the subsequent microphysical processes. More water vapor will produce more clouds and, in the meantime, suppress raindrop evaporation, resulting in larger intensity.

Figure 6 presents the normalized frequency distributions of the observation and forecast at night. In general, the observed frequency distribution in the vertical direction is quite similar to the afternoon period, with a high-frequency area between 3 and 5 km. For the NC and CEC domains, the high-frequency area moves northwards with the prevailing southeasterly wind at night (Xue et al., 2018; Figures 6a and 6c). For SC, the local circulation turns from a sea breeze during the day to a land breeze at night. Convection forms offshore where the land breeze converges with the prevailing onshore winds, forcing the high-frequency area to move southwards (Figure 6e). The forecast successfully



**Figure 4.** Normalized frequency of reflectivity larger than 35 dBZ of radar observations (left column) and the Weather Research and Forecasting (WRF) forecasts (right column) during the afternoon (13:00–19:00 LST) of summer 2016 in different regions. The top, middle and bottom rows present statistics for the North China (NC), Central and Eastern China (CEC), and Southern China (SC) regions, respectively.



**Figure 5.** As Figure 4 but for the intensity of reflectivity larger than 35 dBZ. The intensity is the average of reflectivity exceeding 35 dBZ in latitude and time.

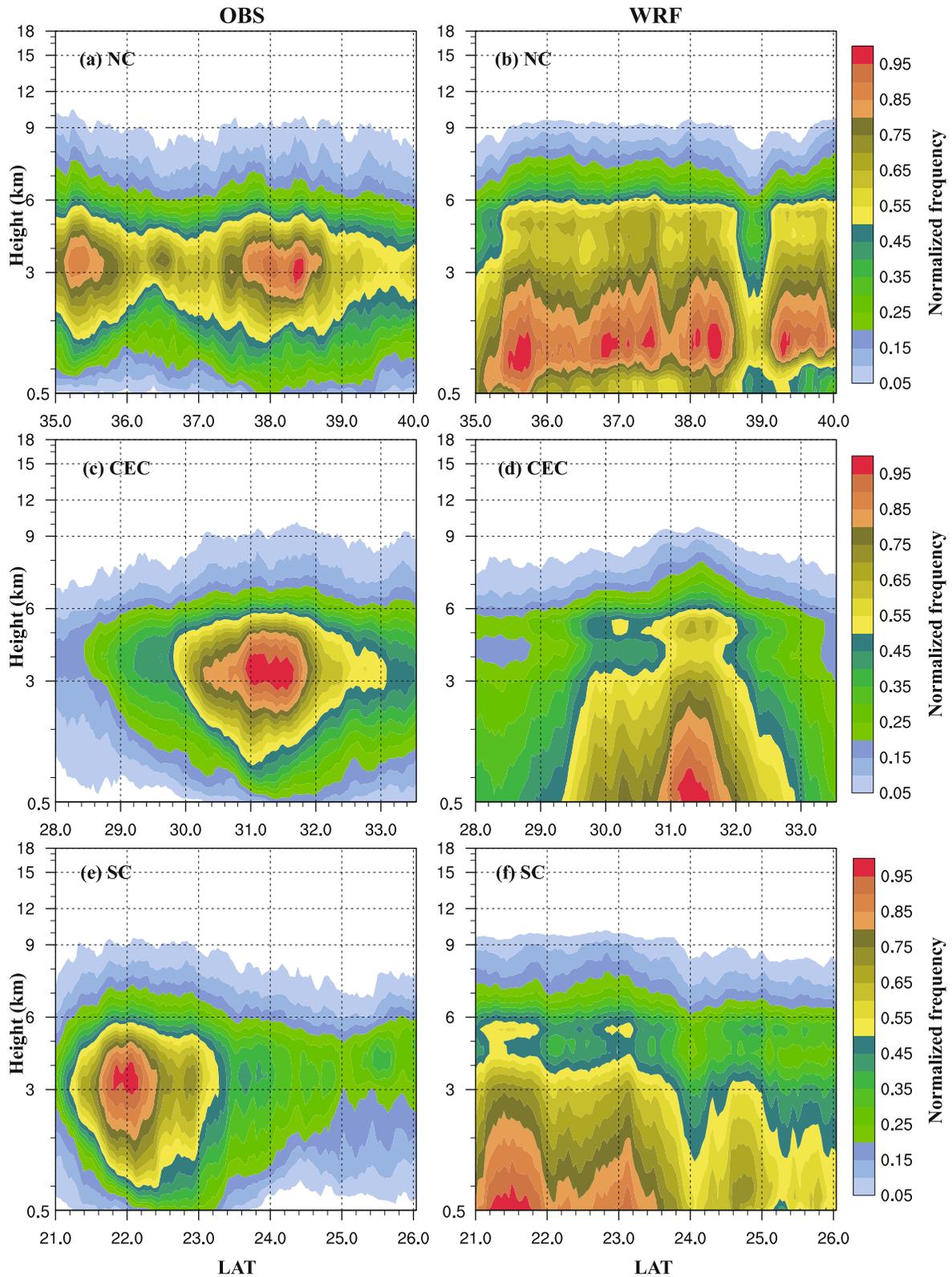


Figure 6. As Figure 4 but for the nighttime (01:00–07:00 LST).

predicts the movement of the high-frequency area, but the predicted high-frequency area is clearly larger than that in the observations in SC and NC. This suggests that the forecast overestimates either the precipitation area or the intensity. In terms of the vertical distribution of frequency, the same issue is found as in the afternoon. The predicted highest frequency area is at the surface, while the observed highest frequency is between 3 and 5 km.

Regarding the intensity, the observation shows a substantial difference between the afternoon and nighttime (see the differences in the left column between Figures 5 and 7). A notable change is that the intensity of the observed reflectivity is clearly reduced, especially in the NC region. This is because precipitation during the afternoon is mostly contributed by thermal convection, whereas nighttime rainfall is mostly from organized precipitation systems. The inclusion of stratiform clouds reduces the average reflectivity intensity at night. However, intensity is increased during the nighttime for the forecast compared to the afternoon, especially for the SC and CEC regions. The increased intensity corresponds well to the increased rainfall amount at night (see Figure 3), which may explain why the forecast tends to overestimate rainfall amount at night.

Overall, the forecast shows a considerable difference from the observation with respect to the 3D reflectivity structure, indicating that the vertical distributions of hydrometeor species were not well reproduced.

### 3.3. Evaluation of WRF-Predicted Objects

Figure 8 shows the properties of forecasted and observed objects identified with composite reflectivity. Unlike the precipitation curve, which shows a bimodal structure, the properties of objects show a single peak, either in the afternoon (around 16:00 LST) or in the morning (around 08:00 LST). This is because the main precipitation systems in the two periods are different. The afternoon peak is mostly contributed by thermal convection, corresponding to higher numbers of objects (middle column of Figure 8), whereas the morning peak is contributed by organized precipitation systems, corresponding to a larger total area (left column of Figure 8), but smaller numbers of objects (middle column of Figure 8). Consistent with the analysis in Section 2.2, the intensity during the afternoon period is stronger than that in the early morning (right column of Figure 8).

For the total area, the forecast matches the observation quite well, except for a slight overestimation in the morning (from 08:00 LST to 13:00 LST in NC, and from 08:00 LST to 11:00 LST in CEC) and a slight underestimation in the afternoon (from 12:00 LST to 20:00 LST in CEC, and from 10:00 LST to 19:00 LST in SC; see Figures 8a, 8d, and 8g). Here, the total area is the sum of all object areas at the current time. The underestimation of total area in the afternoon is due to the forecast's underestimation of the number of objects (Figures 8e and 8h). This indicates that the current 4 km configuration of the WRF model still has difficulty in predicting afternoon convection. It can be seen that the afternoon peak for the number of objects is completely missed in the CEC region (Figure 8e) and greatly underestimated in the SC region (Figure 8h). For the 50th percentile of intensity, the forecast overestimates the intensity in the NC and SC regions. The higher intensity in SC may help explain why, even with smaller total area, the precipitation amount during the afternoon period is still overestimated (Figure 3c). In CEC, with lower intensity and lower frequency, the precipitation amount during the afternoon period is greatly underestimated (Figure 3b). Among the three regions, the properties of NC are best predicted, which is also consistent with the performance in terms of the daily variation of precipitation.

Figure 9 shows the mean and maximum height of the 35 dBZ contour of isolated objects. There are three main obvious differences. The first is that the predicted (red) mean and maximum heights are clearly lower than those of the observation (blue). The second is that the forecast fails to reproduce the afternoon peak. Due to the relatively large thermal instability in the afternoon, the heights of the observed objects are generally higher than that at night. However, the predicted height shows almost no variation between day and night. The third difference is that the spread of the observed maximum height is much wider than that of the forecast. The observed heights of different precipitation types may vary greatly. For example, the height of LEC (Vitale & Ryan, 2013) convection is significantly lower than that of deep convection, although the rainfall amount generated by these two systems may almost be the same. In a study by Yang et al. (2022), the forecast demonstrated almost the same height for LEC and subsequent deep convection on the following day. The development of a microphysics scheme is usually based on a single type of precipitation system and typically represents the mean distribution of convection. As such, the empirical parameters may not be applicable to all types of precipitation, which explains the small variations in the predicted maximum height.

Figures 10 and 11 further explain why the predicted height is lower than in the observation. The kernel density estimation (KDE), which employs a Gaussian weighting function (Zucchini, 2003), is displayed. The three contours represent the envelopment of 25%, 50%, and 75% of the objects, respectively. The innermost contour represents

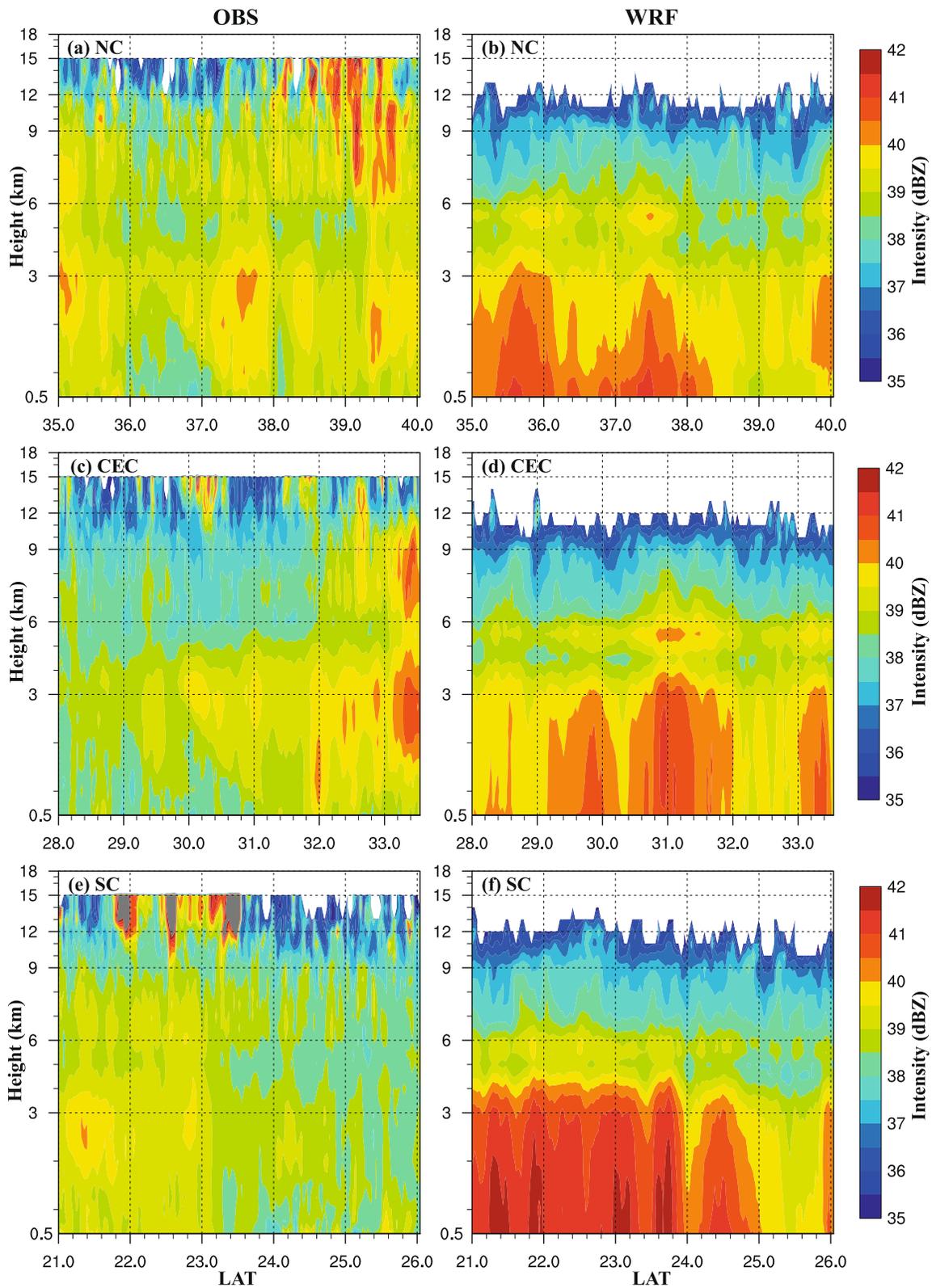
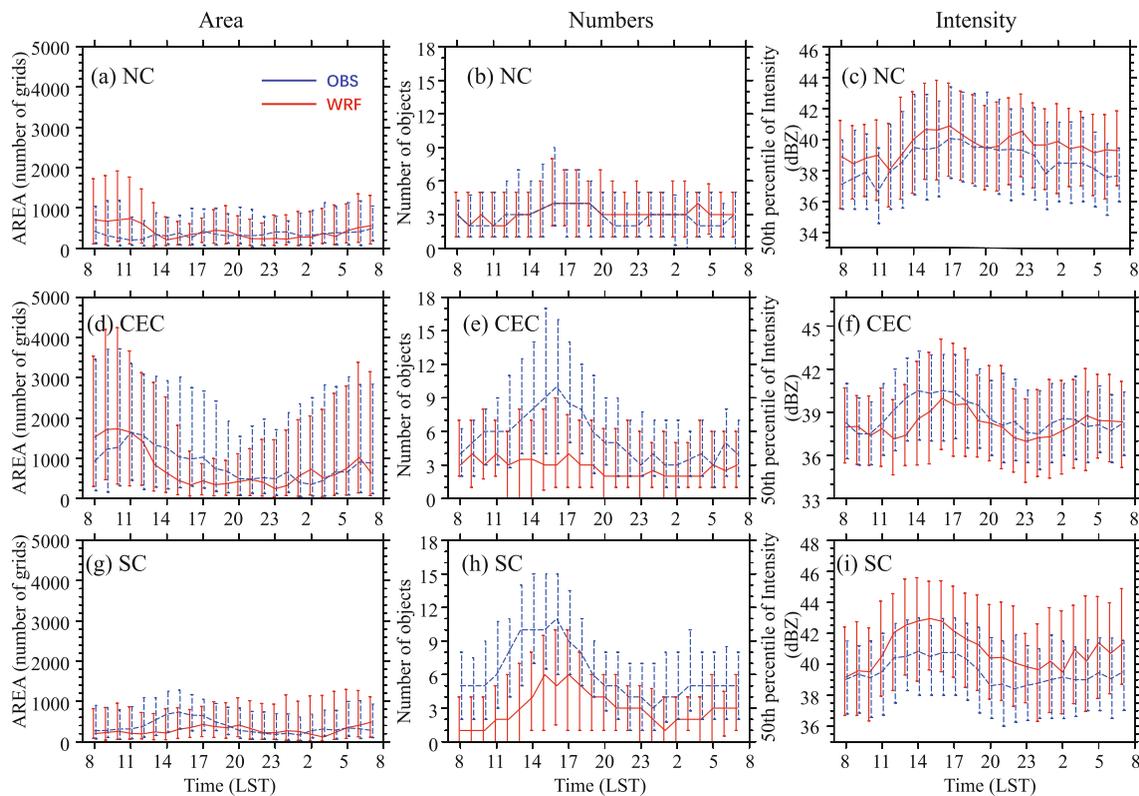


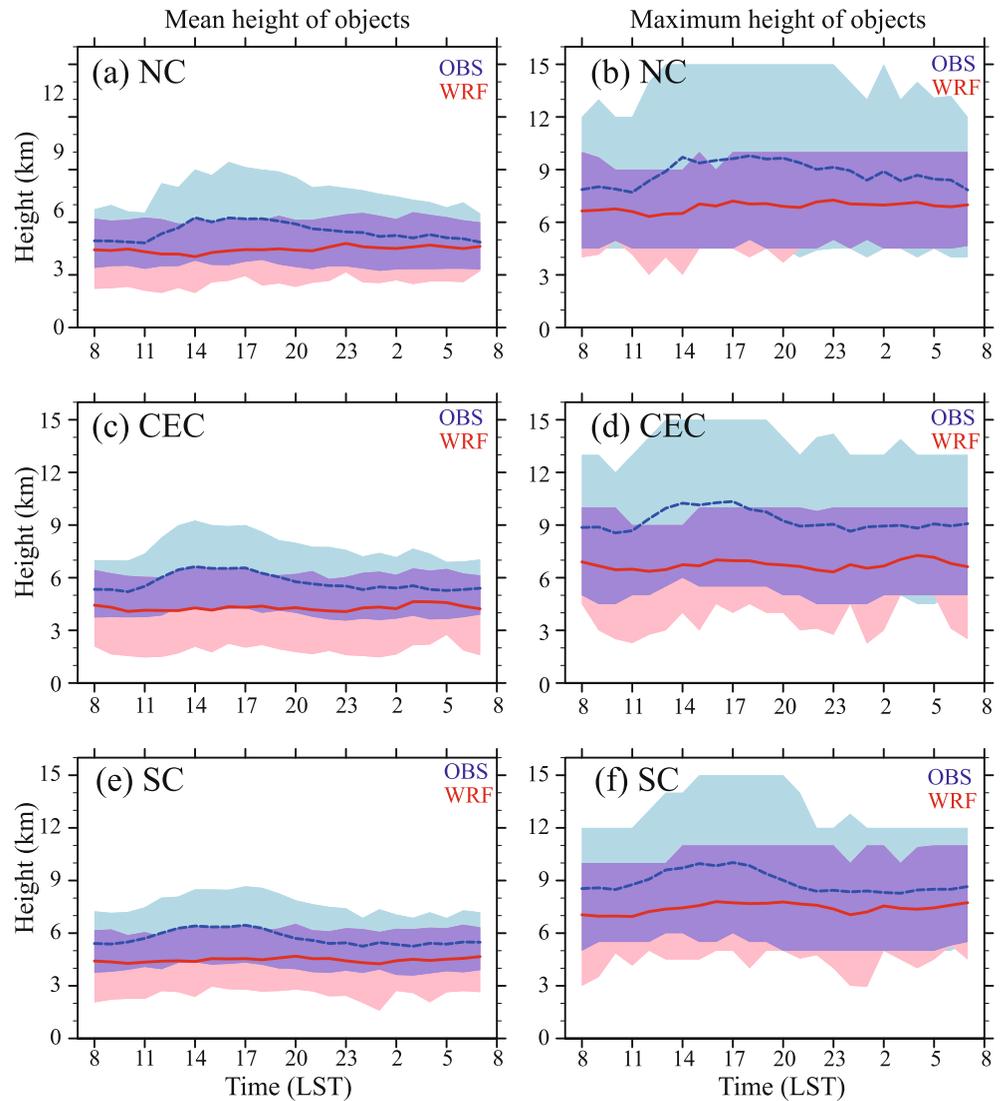
Figure 7. As Figure 5 but for the nighttime (01:00–07:00 LST).



**Figure 8.** The properties of objects (>30 dBZ), including the area (left column), number (middle column), and the 50th percentile of intensity (right column), in the North China (NC; top row), the Central and Eastern China (CEC; middle row), and the Southern China (SC) regions (bottom row). Here, area is measured in terms of grid squares, that is 1 grid square = 16 km<sup>2</sup>. The vertical bars represent the 25th and 75th percentile of the full range.

the highest density (Zhang et al., 2023). The KDE distribution shows that, in the afternoon, the forecasted objects appear vertically limited while the observed objects appear vertically extended (see the 75% contour in the left column of Figure 10). The area difference between the observation and forecast is reduced at nighttime, but the height of small objects (<500 grids area  $\approx$  100 km in diameter) is still underestimated (see the 75% contour in the left column of Figure 11). This result is consistent with that of the diurnal variation, which shows the height difference decreasing at night (see also Figure 10). For both the observation and the forecast, there are mostly small objects, as demonstrated by the maximum of the 75% fraction contours being mostly small objects. For the small objects, the forecast clearly underestimates their mean and maximum heights, and the underestimation is much more obvious in the afternoon (Figure 10) than at night (Figure 11). For the large objects (>500 grids area), the underestimation of mean height is not obvious, but the underestimation of maximum height is still very obvious. The maximum heights of large precipitation systems are mostly contributed by the embedded meso- $\beta$  and - $\gamma$  systems. The underestimation indicates that the forecast was not able to accurately predict the depth of these embedded deep convection systems. Overall, the large difference between the forecast and observation was mainly due to the forecast's inability to reproduce the height of relatively small-scale objects.

There are several possible reasons for the underestimation of the height of small-scale deep convection. One is that the vertical velocity is possibly underestimated due to the relatively coarse resolution. Bryan and Morrison (2011) examined the sensitivity of simulations to the horizontal resolution. Compared to the higher resolutions (1 and 0.25 km), the 4 km resolution greatly underestimated the maximum vertical velocity, suggesting the current 4 km grid spacing may be not sufficient. Another reason might be the considerable uncertainty involved in simulating super-cooled clouds. The reflectivity in the upper level is mostly contributed by snow and graupel. A low height indicates that snow and graupel in the upper level are underestimated, and a major source for these hydrometeors is the collection of super-cooled water through the riming process. Chen et al. (2021) investigated several microphysics schemes, including the Morrison, MY2, and Thompson schemes, for a Meiyu front case. They found that all of them underestimated the content of super-cooled water. Meanwhile, in a squall line case, the height of deep convection was improved after a new cloud analysis scheme produced a proper representation of

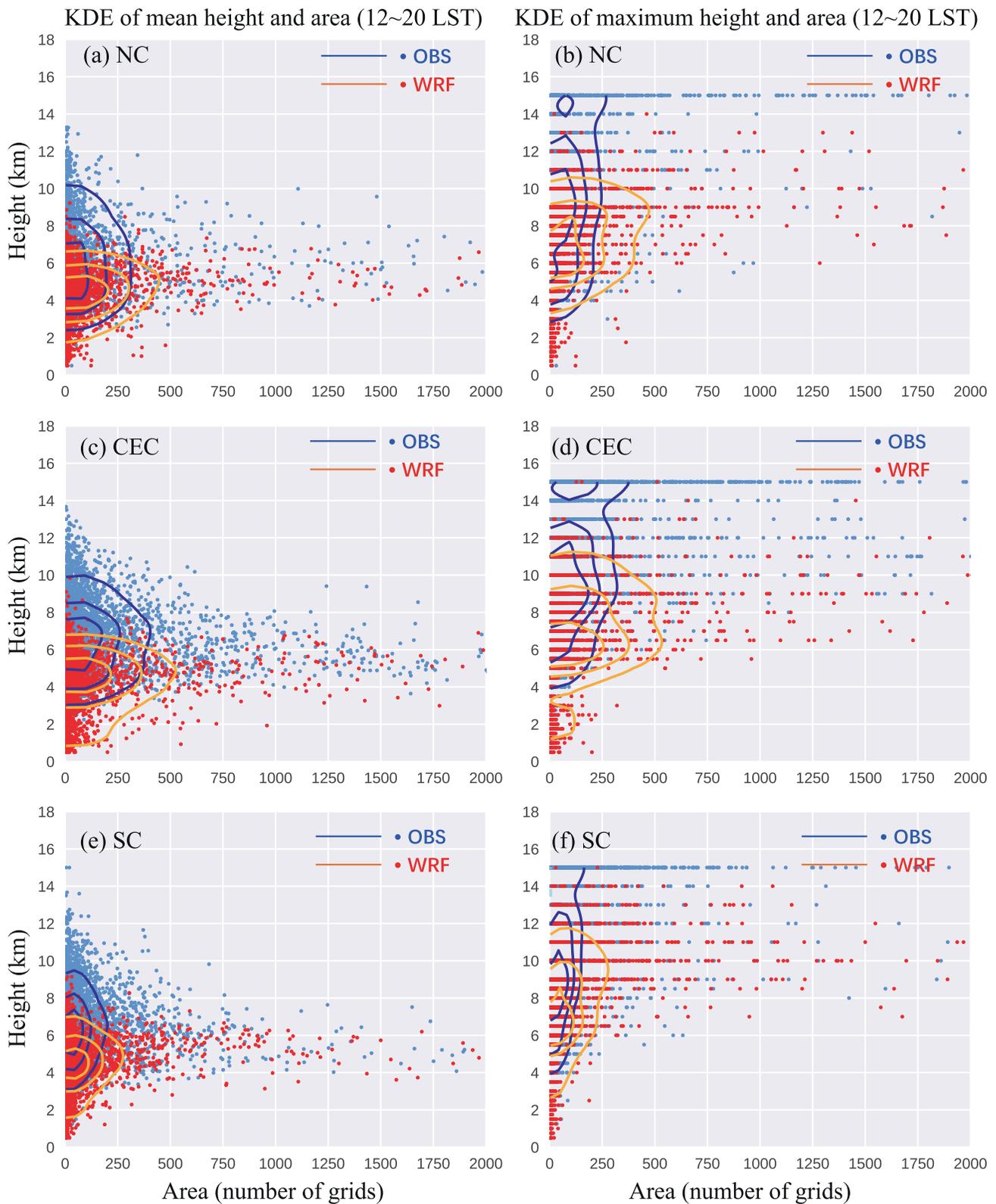


**Figure 9.** Time series of the mean (left column) and the maximum height (right column) of the observed (dashed blue lines) and forecasted objects (red lines) during summer 2016 in the North China (NC; top row), the Central and Eastern China (CEC; middle row) and the Southern China (SC) regions (bottom row). Shading denotes the range from the 10th to 90th percentile. Here, light blue and light pink represent observations and forecasts, respectively. The areas where they overlap are colored with purple.

super-cooled water above the 0°C line (Ding et al., 2022). The release of latent heat through the riming process greatly enhanced the updraft velocity; hence more water vapor was transported to the upper level, resulting in greater ice content and an improved deep convective cloud top height. The lack of cloud information in the initial field and the imperfect ice microphysics may be reasons for the underestimation of object height. Besides, a legitimate question to ask is if the cloud height prediction problem we are seeing with the Morrison scheme is unique to this scheme, and if other schemes share the same problem. Limited case studies we have conducted as well as in Yang et al. (2022) with different microphysical schemes do not show significant differences in the predicted cloud height, which suggests that the problem is not specific to particular microphysics schemes. Further studies with more cases using additional microphysics schemes are needed to robustly answer this question.

### 3.4. Evaluation of Object Properties Using MTD

As shown in Figure 12, MTD tools were used to further evaluate the lifetime and speed of the simulated objects using composite reflectivity at 10 min intervals. In terms of lifetime, the forecast shows a similar distribution as



**Figure 10.** The kernel density estimation (KDE) of 35 dBZ mean (left column) and maximum (right column) height and area of composite reflectivity larger than 30 dBZ between 12:00 LST and 20:00 LST of summer 2016 in the North China (NC; top row), the Central and Eastern China (CEC; middle row) and the SC regions (bottom row).

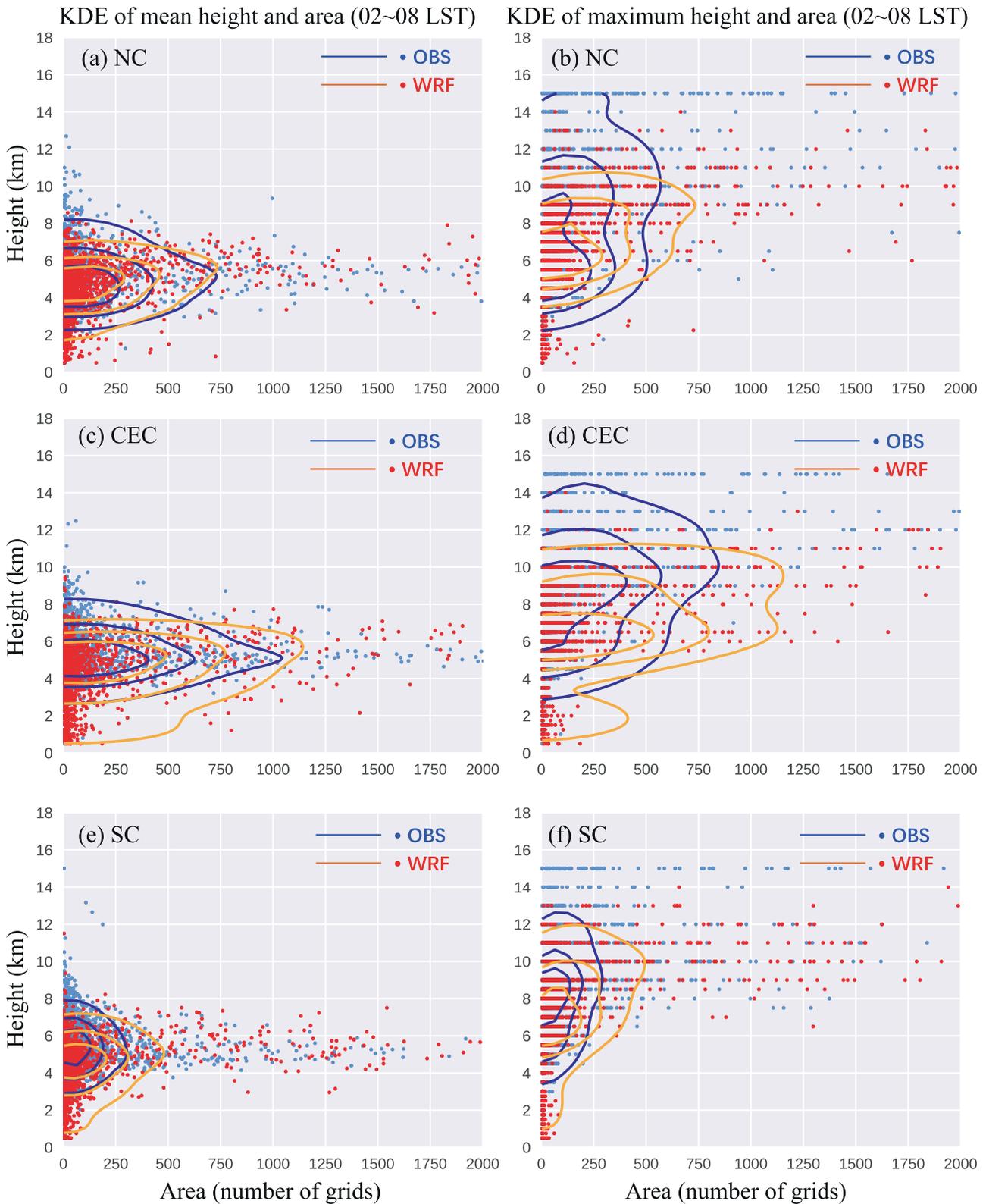
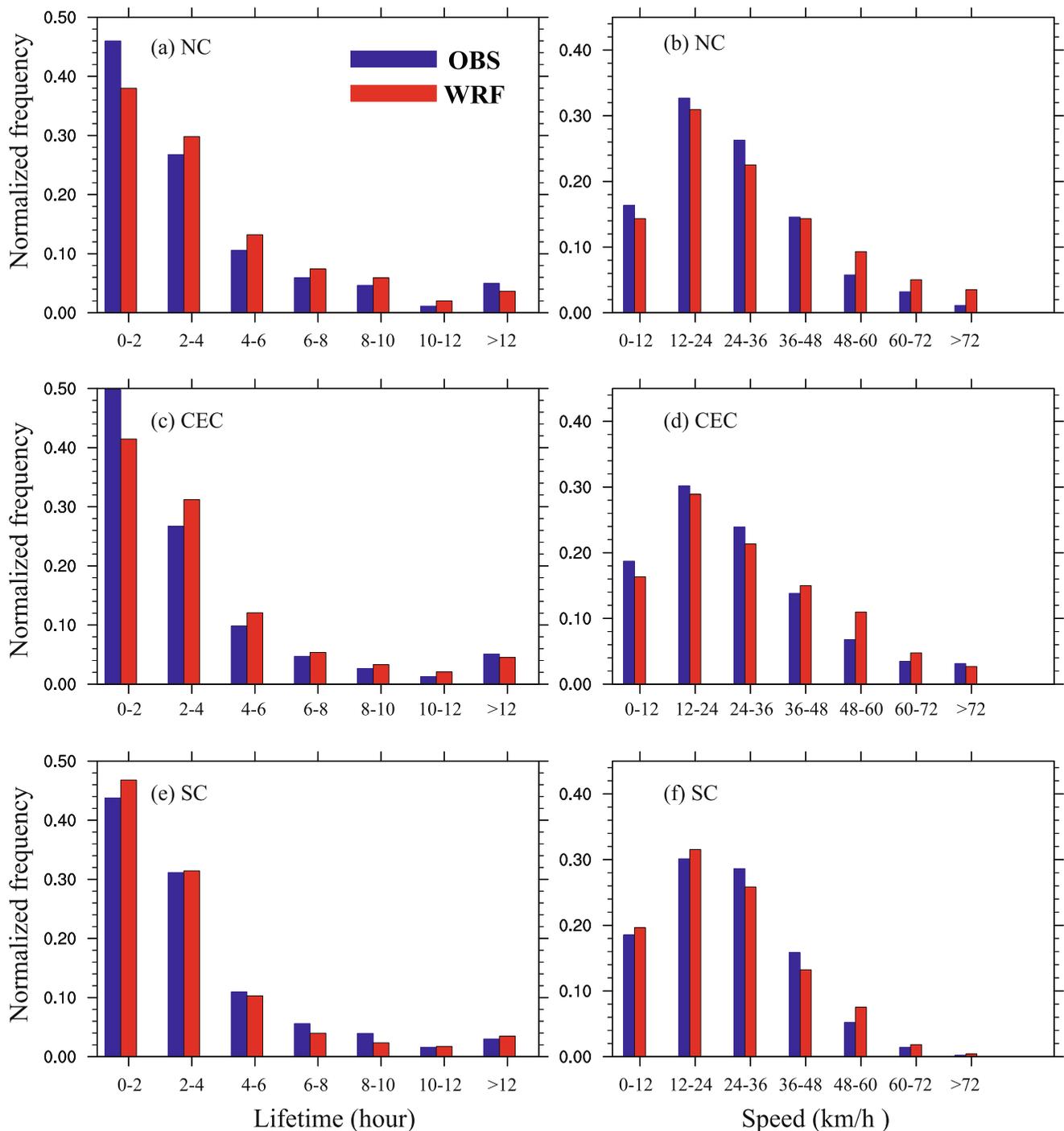
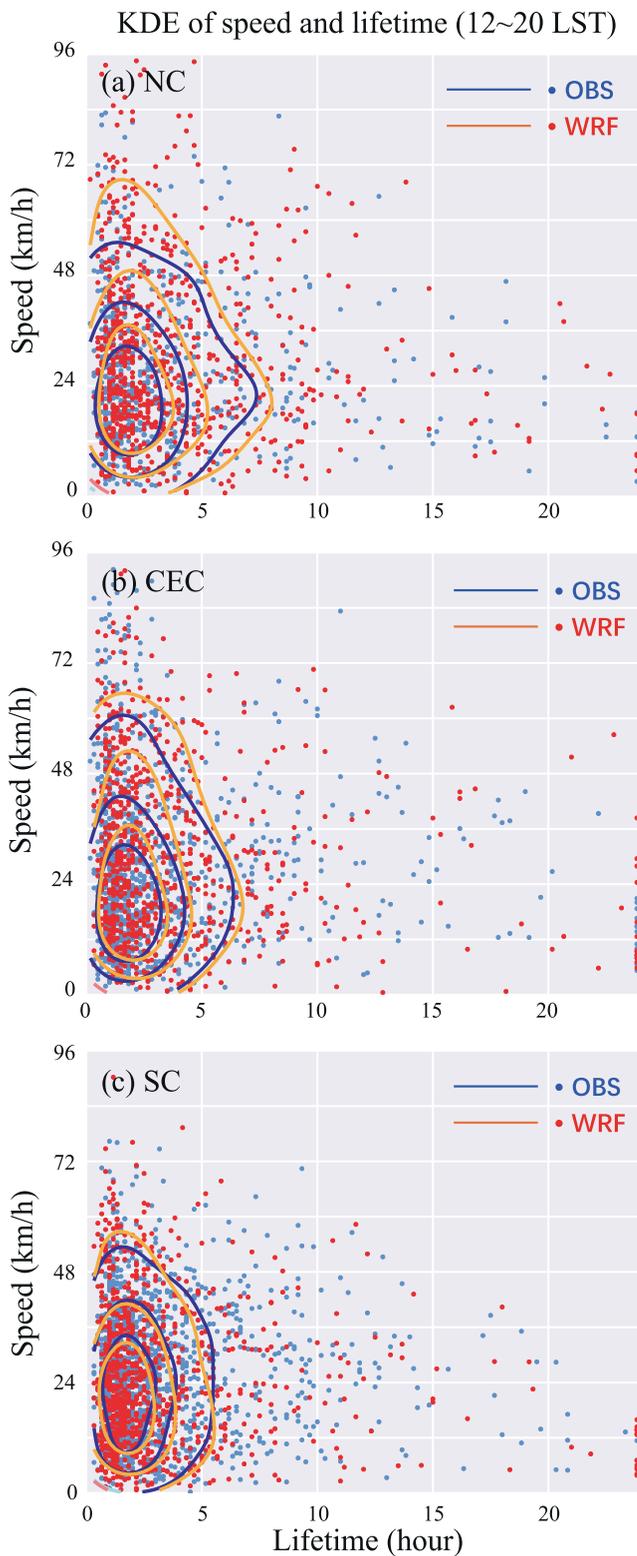


Figure 11. As Figure 10 but for 02:00–08:00 LST.



**Figure 12.** Normalized frequency distribution of the lifetime (left column) and speed of objects (right column) during summer 2016 in the North China (NC; top row), the Central and Eastern China (CEC; middle row) and the Southern China (SC) regions (bottom row). The frequency distribution is normalized by the total number of objects.

the observation, but with some biases that vary by region. For NC and CEC, the forecast tends to underestimate the proportion of lifetimes within 0–2 hr and >12 hr but overestimates the proportion within 2–10 hr. The result in SC is just the opposite. For speed, the forecast tends to produce a greater proportion of fast-moving objects. It can be seen that, in all three regions, the proportions of speed  $>48 \text{ km}\cdot\text{h}^{-1}$  are generally larger than observed. Figure 13 shows the KDE distribution of the speed and duration of the objects. The fast-moving objects are mostly contributed by short-duration precipitation systems. The predicted speeds of short-duration precipitation



**Figure 13.** The Kernel density estimation (KDE) of the average moving speed and lifetime of isolated objects during summer 2016 in (a) Northern China (NC), (b) Central and Eastern China (CEC), and (c) SC regions.

systems are slightly faster than observed (see differences in the 75% contour), especially in the NC region.

Figure 14 presents the properties of long-lifetime (>12 hr) precipitation systems. The purpose is to see whether the 4-km WRF model can simulate the diurnal variation of properties well for long-lived precipitation systems. In general, the longer the lifetime of the precipitation system, the larger the area of objects. The WRF forecast still shows notable differences to the observation, and different regions present different behaviors. In NC and CEC, the observation shows the areas of precipitation systems gradually increase overnight, and the forecast successfully reproduces the diurnal variation of precipitation area. However, the predicted intensity in these two regions differs greatly to the observed. The observation shows the largest intensity as occurring in the afternoon, whereas it is overnight in the forecast. The predicted intensity seems to be closely correlated with the precipitation area, but the opposite is true for the observation. This high intensity in the observation, as mentioned above, may be contributed by the embedded micro- $\beta$  or  $\gamma$  convection systems in the larger precipitation system. Their properties are still extremely hard for the 4 km WRF model to predict.

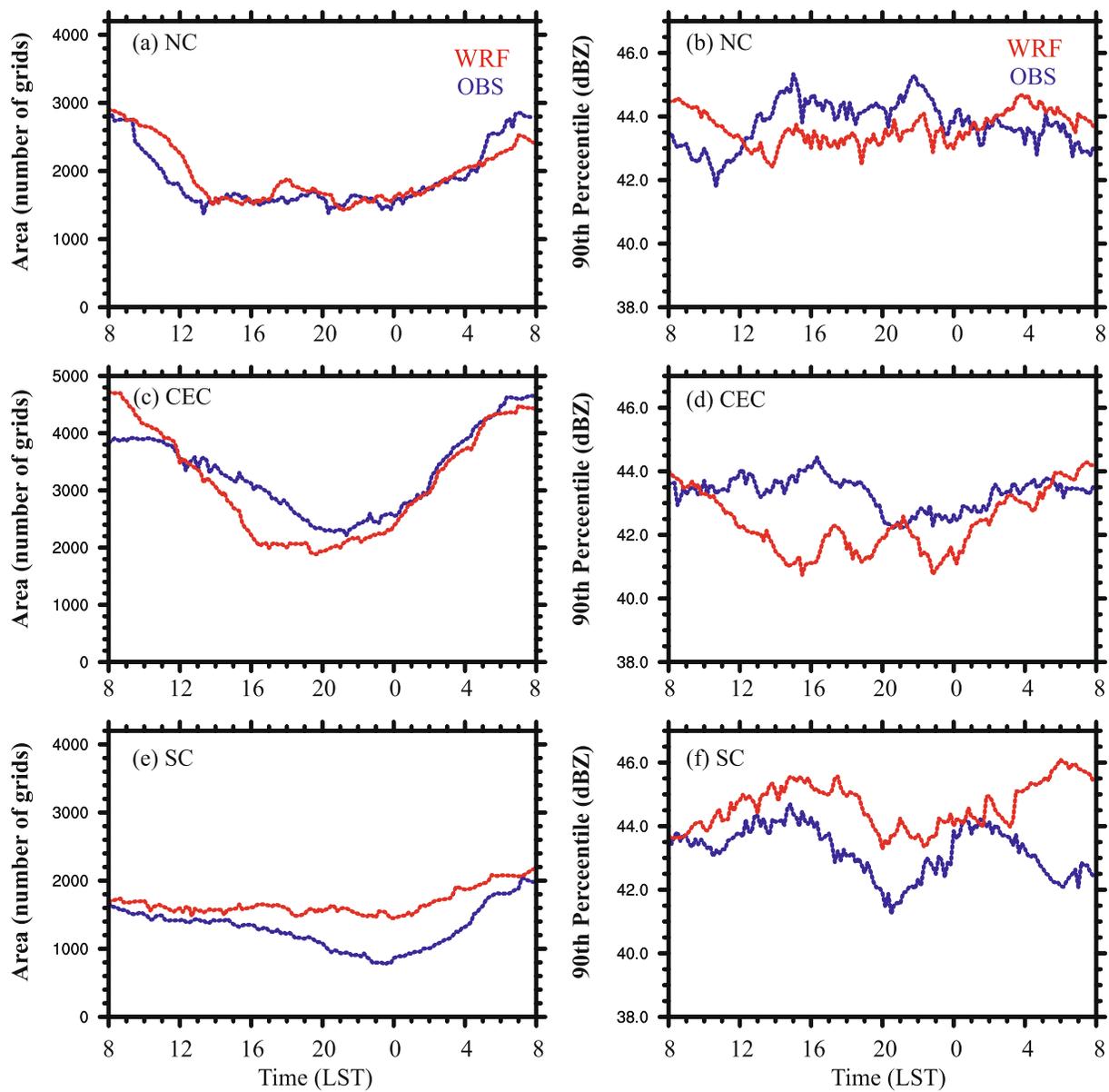
Compared to NC and CEC, the differences in SC are more obvious. In terms of precipitation area, the forecast and the observation match well in NC and CEC. But in SC, there is a rapid decrease in area between 16:00 LST and 00:00 LST in the observation, followed by an increase. The forecast, however, shows almost no variation before 00:00 LST. Regarding the intensity, the observation shows two peaks. One at around 15:00 LST and the other at around 02:00 LST. The forecast captures the afternoon peak well, but predicts a nighttime peak at around 06:00 LST. Again, the predicted peak intensity at night is slightly larger than that in the afternoon, which is contrary to the observation. The larger difference in SC is probably because, in addition to possible microphysical factors, the water vapor forecast biases in SC is more pronounced than that of in NC and CEC (Zhu et al., 2021). For the forecast, excessive water vapor in SC greatly increases the predicted precipitation area and intensity. More sensitivity experiments are needed to reveal the impact of water vapor forecast bias on the simulated cloud structure.

#### 4. Summary and Future Work

This study used 3D reflectivity observations to investigate how well a CPR model predicted reflectivity structure over China. The forecasts were from the 4 km resolution WRF model run in real time at Nanjing University for the summer of 2016. Three rainfall centers (SC, CEC, and NC) in China were selected for evaluation. While the precipitation distribution and diurnal variation were predicted well by the model, there were major differences between the predicted and observed 3D reflectivity.

The highest observed frequency (with reflectivity  $\geq 35$  dBZ) was between 3 and 6 km, whereas that in the forecast is found at the surface. Compared to the observation, the forecast greatly overestimated the intensity at low levels, especially in SC where the water vapor content is overestimated. The closer to the surface, the higher the predicted reflectivity. This may indicate that the model tends to produce too many large raindrops at lower levels.

Object-based verification methods were used to investigate the properties of simulated reflectivity objects (with reflectivity  $\geq 30$  dBZ), including the frequency, area, and intensity, as well as the mean and maximum height of the 35 dBZ reflectivity. Although WRF successfully simulated the diurnal



**Figure 14.** Time series of the mean area (left column) and the 90th percentile (right column) of intensity of long-lived precipitation systems (lifetime >12 hr) during summer 2016 in the Northern China (NC; top row), the Central and Eastern China (CEC; middle row) and the Southern China (SC) regions (bottom row).

variation of precipitation, the simulated properties still showed large differences. The forecast tended to underestimate the frequencies of afternoon precipitation but overestimate the intensity, especially at night. The low frequency of afternoon convection was the reason for the underestimated precipitation amount during the afternoon in CEC. The relatively high intensity and large area were the reasons for the overestimated rainfall in SC. Compared to the observation, the forecast failed to reproduce the diurnal variation in the mean and maximum height of the 35 dBZ reflectivity for the identified objects. The observation showed an afternoon peak at around 16:00–17:00 LST, which was correlated with deep convection in the afternoon. The forecast failed to reproduce the afternoon peak and showed a near-uniform height throughout the day and night. Further analysis showed that the underestimation of small-scale objects was the main reason for the underestimated mean and maximum object heights. Moreover, the predicted object height showed less variation than observed. In other words, the current microphysics scheme may not reflect all precipitation types.

Additionally, the forecast was evaluated using MTD, which incorporates additional information in the temporal dimension. Results showed that, although the forecast reproduced the overall distribution of object duration and

speed, there were still some differences. The predicted average speed of small objects (0–3 hr short duration) was faster than observed, especially in NC, and the forecast failed to reproduce the diurnal variation in the intensity of large precipitation systems (>12 hr long duration).

It is important to highlight that the current study only investigated the performance of the Morrison microphysics scheme, and that the simulation of precipitation cloud properties is highly sensitive to the choice of microphysics scheme. Accordingly, the plan for future work is to use a small simulation domain for sensitivity experiments to reveal the key microphysical processes that lead to the uncertainty in simulating 3D reflectivity. In addition, the current study focuses on evaluating forecasts at 4 km grid spacing. Given that 3–4 km grid spacings remain the prevailing grid spacings used by high-resolution convection-allowing regional operational prediction systems in a number of countries, such evaluations are of practical importance. Still, as computational power increases, the use of higher resolutions will happen in not too distant future. Similar evaluations for forecasts at higher resolutions should be performed also, and should be done when a sufficient number of such forecasts can be produced in the future. Also, the current study did not examine polarimetric radar observations. This is because the upgrading of single-Doppler radar to dual-Doppler radar is not yet complete in China. Compared to single-Doppler radar, polarimetric radar is more useful in uncovering the potential forecasting issues related to the representation of microphysical processes. Regions such as SC, where upgrades to dual-Doppler radar are complete, will be used for future evaluations.

### Data Availability Statement

The hourly precipitation observation could be obtained from the website: <http://data.cma.cn/>. The attribute data of simulated and observed reflectivity objects are available in the study of Zhu (2023).

### References

- Aligo, E. A., Ferrier, B., & Carley, J. R. (2018). Modified NAM microphysics for forecasts of deep convective storms. *Monthly Weather Review*, *146*(12), 4115–4153. <https://doi.org/10.1175/mwr-d-17-0277.1>
- Bao, X., Zhang, F., & Sun, J. (2011). Diurnal variations of warm-season precipitation east of the Tibetan Plateau over China. *Monthly Weather Review*, *139*(9), 2790–2810. <https://doi.org/10.1175/mwr-d-11-00006.1>
- Barnes, H. C., & Houze, R. A. (2016). Comparison of observed and simulated spatial patterns of ice microphysical processes in tropical oceanic mesoscale convective systems. *Journal of Geophysical Research: Atmospheres*, *121*(14), 8269–8296. <https://doi.org/10.1002/2016jd025074>
- Benjamin, S. G., Weygandt, S. S., Brown, J. M., Hu, M., Alexander, C. R., Smirnova, T. G., et al. (2016). A North American hourly assimilation and model forecast cycle: The rapid refresh. *Monthly Weather Review*, *144*(4), 1669–1694. <https://doi.org/10.1175/mwr-d-15-0242.1>
- Bringi, V. N., & Chandrasekar, V. (2001). Polarimetric Doppler weather radar (p. 636).
- Brown, B., Jensen, T., Gotway, J. H., Bullock, R., Gilleland, E., Fowler, T., et al. (2021). The Model Evaluation Tools (MET): More than a decade of community-supported forecast verification. *Bulletin of the American Meteorological Society*, *102*(4), E782–E807. <https://doi.org/10.1175/bams-d-19-0093.1>
- Brown, B. R., Bell, M. M., & Frambach, A. J. (2016). Validation of simulated hurricane drop size distributions using polarimetric radar. *Geophysical Research Letters*, *43*(2), 910–917. <https://doi.org/10.1002/2015gl067278>
- Brown, B. R., Bell, M. M., & Thompson, G. (2017). Improvements to the snow melting process in a partially double moment microphysics parameterization. *Journal of Advances in Modeling Earth Systems*, *9*(2), 1150–1166. <https://doi.org/10.1002/2016ms000892>
- Bryan, G. H., & Morrison, H. (2011). Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Monthly Weather Review*, *140*(1), 202–225. <https://doi.org/10.1175/mwr-d-11-00046.1>
- Cai, S. X., Huang, A. N., Zhu, K. F., Yang, B., Yang, X. Y., Wu, Y., & Mu, X. Y. (2021). Diurnal cycle of summer precipitation over the Eastern Tibetan Plateau and surrounding regions simulated in a convection-permitting model. *Climate Dynamics*, *57*(1–2), 611–632. <https://doi.org/10.1007/s00382-021-05729-5>
- Chen, G., Zhao, K., Huang, H., Yang, Z., Lu, Y., & Yang, J. (2021). Evaluating simulated raindrop size distributions and ice microphysical processes with polarimetric radar observations in a Meiyu front event over Eastern China. *Journal of Geophysical Research: Atmospheres*, *126*(22), e2020JD034511. <https://doi.org/10.1029/2020jd034511>
- Collins, W. D., Rasch, P. J., Boville, B. A., Hack, J. J., Mccaa, J. R., Williamson, D. L., et al. (2004). *Description of the NCAR Community Atmosphere Model (CAM 3.0)*. NCAR Technical Note NCAR/TN-464+STR. <https://doi.org/10.5065/D63N21CH>
- Davis, C., Brown, B., & Bullock, R. (2006). Object-based verification of precipitation forecasts. Part I: Methodology and application to mesoscale rain areas. *Monthly Weather Review*, *134*(7), 1772–1784. <https://doi.org/10.1175/mwr3145.1>
- Ding, Z., Zhao, K., Zhu, K., Feng, Y., Huang, H., & Yang, Z. (2022). Assimilation of polarimetric radar observation with GSI cloud analysis for the prediction of a squall line. *Geophysical Research Letters*, *49*(16), e2022GL098253. <https://doi.org/10.1029/2022gl098253>
- Gallus, W. A., Jr. (2010). Application of object-based verification techniques to ensemble precipitation forecasts. *Weather and Forecasting*, *25*(1), 144–158. <https://doi.org/10.1175/2009waf2222274.1>
- Gallus, W. A., Jr., & Pfeifer, M. (2008). Intercomparison of simulations using five WRF microphysical schemes with dual-Polarization data for a German squall line. *Advances in Geosciences*, *16*, 109–116. <https://doi.org/10.5194/adgeo-16-109-2008>
- Gilleland, E., Ahijevych, D., Brown, B. G., Casati, B., & Ebert, E. E. (2009). Intercomparison of spatial forecast verification methods. *Weather and Forecasting*, *24*(5), 1416–1430. <https://doi.org/10.1175/2009waf2222269.1>
- Griffin, S. M., Otkin, J. A., Nebuda, S. E., Jensen, T. L., Skinner, P. S., Gilleland, E., et al. (2021). Evaluating the impact of planetary boundary layer, land surface model, and microphysics parameterization schemes on cold cloud objects in simulated GOES-16 brightness temperatures. *Journal of Geophysical Research: Atmospheres*, *126*(15), e2021JD034709. <https://doi.org/10.1029/2021jd034709>

### Acknowledgments

This work was primarily supported by the National Natural Science Foundation of China (Grants 41975124, 41730965) and the National Key Research and Development Program of China under Grant 2018YFC1507303.

- Hu, M., & Xue, M. (2007). Implementation and evaluation of cloud analysis with WSR-88D reflectivity data for GSI and WRF-ARW. *Geophysical Research Letters*, 34(7), L07808. <https://doi.org/10.1029/2006GL028847>
- Huang, H., Zhao, K., Zhang, G., Lin, Q., Wen, L., Chen, G., et al. (2018). Quantitative precipitation estimation with operational polarimetric radar measurements in Southern China: A differential phase-based variational approach. *Journal of Atmospheric and Oceanic Technology*, 35(6), 1253–1271. <https://doi.org/10.1175/jtech-d-17-0142.1>
- Jankov, I., Grasso, L. D., Sengupta, M., Neiman, P. J., Zupanski, D., Zupanski, M., et al. (2011). An evaluation of five ARW-WRF microphysics schemes using synthetic GOES Imagery for an atmospheric river event affecting the California Coast. *Journal of Hydrometeorology*, 12(4), 618–633. <https://doi.org/10.1175/2010jhm1282.1>
- Jiang, Z., Zhang, D.-L., Xia, R., & Qian, T. (2017). Diurnal variations of presummer rainfall over Southern China. *Journal of Climate*, 30(2), 755–773. <https://doi.org/10.1175/jcli-d-15-0666.1>
- Jones, T. A., Skinner, P., Knopfmeier, K., Mansell, E., Minnis, P., Palikonda, R., & Smith, W. (2018). Comparison of cloud microphysics schemes in a warn-on-forecast system using synthetic satellite objects. *Weather and Forecasting*, 33(6), 1681–1708. <https://doi.org/10.1175/waf-d-18-0112.1>
- Kumjian, M. R., Khain, A. P., Benmoshe, N., Ilotoviz, E., Ryzhkov, A. V., & Phillips, V. T. J. (2014). The anatomy and physics of Z(DR) columns: Investigating a polarimetric radar signature with a spectral bin microphysical model. *Journal of Applied Meteorology and Climatology*, 53(7), 1820–1843. <https://doi.org/10.1175/jamc-d-13-0354.1>
- Li, P. X., Furtado, K., Zhou, T. J., Chen, H. M., & Li, J. (2021). Convection-permitting modeling improves simulated precipitation over the central and eastern Tibetan Plateau. *Quarterly Journal of the Royal Meteorological Society*, 147(734), 341–362. <https://doi.org/10.1002/qj.3921>
- Li, P. X., Furtado, K., Zhou, T. J., Chen, H. M., Li, J., Guo, Z., & Xiao, C. (2020). The diurnal cycle of East Asian summer monsoon precipitation simulated by the Met Office Unified Model at convection-permitting scales. *Climate Dynamics*, 55(1–2), 131–151. <https://doi.org/10.1007/s00382-018-4368-z>
- Lin, H. D., Weygandt, S. S., Benjamin, S. G., & Hu, M. (2017). Satellite radiance data assimilation within the hourly updated rapid refresh. *Weather and Forecasting*, 32(4), 1273–1287. <https://doi.org/10.1175/waf-d-16-0215.1>
- Luo, Y. L., Zhang, R. H., Wan, Q. L., Wang, B., Wong, W. K., Hu, Z. Q., et al. (2017). The Southern China Monsoon Rainfall Experiment (Scmrex). *Bulletin of the American Meteorological Society*, 98(5), 999–1013. <https://doi.org/10.1175/bams-d-15-00235.1>
- Marshall, J. S., & Gunn, K. L. S. (1952). Measurement of snow parameters by radar. *Journal of the Atmospheric Sciences*, 9(5), 322–327. [https://doi.org/10.1175/1520-0469\(1952\)009<0322:mospbr>2.0.co;2](https://doi.org/10.1175/1520-0469(1952)009<0322:mospbr>2.0.co;2)
- Morrison, H., Curry, J. A., & Khvorostyanov, V. I. (2005). A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description. *Journal of the Atmospheric Sciences*, 62(6), 1665–1677. <https://doi.org/10.1175/jas3446.1>
- Morrison, H., Milbrandt, J. A., Bryan, G. H., Ikeda, K., Tessendorf, S. A., & Thompson, G. (2015). Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part II: Case study comparisons with observations and other schemes. *Journal of the Atmospheric Sciences*, 72(1), 312–339. <https://doi.org/10.1175/jas-d-14-0066.1>
- Morrison, H., Tessendorf, S. A., Ikeda, K., & Thompson, G. (2012). Sensitivity of a simulated midlatitude squall line to parameterization of raindrop breakup. *Monthly Weather Review*, 140(8), 2437–2460. <https://doi.org/10.1175/mwr-d-11-00283.1>
- Otkin, J. A., & Greenwald, T. J. (2008). Comparison of WRF model-simulated and MODIS-derived cloud data. *Monthly Weather Review*, 136(6), 1957–1970. <https://doi.org/10.1175/2007mwr2293.1>
- Pan, Y., Shen, Y., Yu, J. J., & Xiong, A. Y. (2015). An experiment of high-resolution gauge-radar-satellite combined precipitation retrieval based on the Bayesian merging method (in Chinese). *Acta Meteorologica Sinica*, 73, 177–186.
- Pleim, J. E. (2007). A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *Journal of Applied Meteorology and Climatology*, 46(9), 1383–1395. <https://doi.org/10.1175/jam2539.1>
- Pleim, J. E., & Xiu, A. (1995). Development and testing of a surface flux and planetary boundary layer model for application in mesoscale models. *Journal of Applied Meteorology*, 34(1), 16–32. <https://doi.org/10.1175/1520-0450-34.1.16>
- Saito, K., Fujita, T., Yamada, Y., Ishida, J.-i., Kumagai, Y., Aranami, K., et al. (2006). The operational JMA nonhydrostatic mesoscale model. *Monthly Weather Review*, 134(4), 1266–1298. <https://doi.org/10.1175/mwr3120.1>
- Shen, Y., Zhao, P., Pan, Y., & Yu, J. J. (2014). A high spatiotemporal gauge-satellite merged precipitation analysis over China. *Journal of Geophysical Research*, 119(6), 3063–3075. <https://doi.org/10.1002/2013jd020686>
- Skinner, P. S., Wheatley, D. M., Knopfmeier, K. H., Reinhart, A. E., Choate, J. J., Jones, T. A., et al. (2018). Object-based verification of a prototype warn-on-forecast system. *Weather and Forecasting*, 33(5), 1225–1250. <https://doi.org/10.1175/waf-d-18-0020.1>
- Stoelinga, M. (2005). Simulated equivalent reflectivity factor as currently formulated in RIP: Description and possible improvements. Retrieved from <https://www.researchgate.net/publication/242107593>
- Thompson, G., & Eidhammer, T. (2014). A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *Journal of the Atmospheric Sciences*, 71(10), 3636–3658. <https://doi.org/10.1175/jas-d-13-0305.1>
- Vitale, J., & Ryan, T. (2013). Operational recognition of high precipitation efficiency and low-echo-centroid convection. *Journal of Operational Meteorology*, 1(12), 128–143. <https://doi.org/10.15191/nwajom.2013.0112>
- Wang, J., Fan, J., Houze, R. A., Jr., Brodzik, S. R., Zhang, K., Zhang, G. J., & Ma, P. L. (2021). Using radar observations to evaluate 3-D radar echo structure simulated by the Energy Exascale Earth System Model (E3SM) version 1. *Geoscientific Model Development*, 14(2), 719–734. <https://doi.org/10.5194/gmd-14-719-2021>
- Wu, Y., Huang, A. N., Huang, D. Q., Chen, F., Yang, B., Zhou, Y., et al. (2018). Diurnal variations of summer precipitation over the regions east to Tibetan Plateau. *Climate Dynamics*, 51(11–12), 4287–4307. <https://doi.org/10.1007/s00382-017-4042-x>
- Xu, X., Howard, K., & Zhang, J. (2008). An automated radar technique for the identification of tropical precipitation. *Journal of Hydrometeorology*, 9(5), 885–902. <https://doi.org/10.1175/2007jhm954.1>
- Xue, M., Droegemeier, K. K., & Wong, V. (2000). The Advanced Regional Prediction System (ARPS)—A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteorology and Atmospheric Physics*, 75(3–4), 161–193. <https://doi.org/10.1007/s007030070003>
- Xue, M., Luo, X., Zhu, K. F., Sun, Z. Q., & Fei, J. F. (2018). The controlling role of boundary layer inertial oscillations in Meiyu frontal precipitation and its diurnal cycles over China. *Journal of Geophysical Research*, 123(10), 5090–5115. <https://doi.org/10.1029/2018jd028368>
- Yang, N., Zhu, K., & Xue, M. (2022). Investigation of the convection-allowing prediction error of an extreme precipitation event of China using CRTM-simulated brightness temperature. *Journal of Geophysical Research: Atmospheres*, 127(18), e2022JD036760. <https://doi.org/10.1029/2022jd036760>
- Yu, B., Zhu, K., Xue, M., & Zhou, B. (2020). Using new neighborhood-based intensity-scale verification metrics to evaluate WRF precipitation forecasts at 4 and 12 km grid spacings. *Atmospheric Research*, 246, 105117. <https://doi.org/10.1016/j.atmosres.2020.105117>

- Zhang, C., Xue, M., Zhu, K., & Yu, X. (2023). Climatology of significant tornadoes within China and comparison of tornado environments between United States and China. *Monthly Weather Review*, *151*, 465–484. <https://doi.org/10.1175/MWR-D-22-0070.1>
- Zhang, J., Howard, K., & Gourley, J. J. (2005). Constructing three-dimensional multiple-radar reflectivity mosaics: Examples of convective storms and stratiform rain echoes. *Journal of Atmospheric and Oceanic Technology*, *22*(1), 30–42. <https://doi.org/10.1175/jtech-1689.1>
- Zhao, J., Chen, L., & Wang, D. (2018). Characteristics and causes analysis of abnormal Meiyu in China in 2016 (in Chinese). *Chinese Journal of Atmospheric Sciences*, *42*, 1055–1066.
- Zhao, Y., Huang, A. N., Kan, M. Y., Dong, X. N., Yu, X. J., Wu, Y., et al. (2020). Characteristics of hourly extreme precipitation along the Yangtze River Basin, China during warm season. *Scientific Reports*, *10*(1), 5613. <https://doi.org/10.1038/s41598-020-62535-5>
- Zhou, A., Zhao, K., Lee, W.-C., Ding, Z., Lu, Y., & Huang, H. (2022). Evaluation and modification of microphysics schemes on the cold pool evolution for a simulated bow echo in Southeast China. *Journal of Geophysical Research: Atmospheres*, *127*(2), <https://doi.org/10.1029/2021jd035262>
- Zhu, K. (2023). How well does 4 km WRF model predict three-dimensional reflectivity structure over China as compared to radar observations? [Dataset]. Zenodo. <https://doi.org/10.5281/zenodo.7613988>
- Zhu, K., Xue, M., Pan, Y., Hu, M., Benjamin, S. G., Weygandt, S. S., & Lin, H. (2019). The impact of satellite radiance data assimilation within a frequently updated regional forecast system using a GSI-based ensemble Kalman Filter. *Advances in Atmospheric Sciences*, *36*(12), 1308–1326. <https://doi.org/10.1007/s00376-019-9011-3>
- Zhu, K., Xue, M., Zhou, B., Zhao, K., Sun, Z., Fu, P., et al. (2018). Evaluation of real-time convection-permitting precipitation forecasts in China during the 2013–2014 summer season. *Journal of Geophysical Research: Atmospheres*, *123*(2), 1037–1064. <https://doi.org/10.1002/2017jd027445>
- Zhu, K., Yu, B., Xue, M., Zhou, B., & Hu, X.-M. (2021). Summer season precipitation biases in 4 km WRF forecasts over Southern China: Diagnoses of the causes of biases. *Journal of Geophysical Research: Atmospheres*, *126*(23), e2021JD035530. <https://doi.org/10.1029/2021jd035530>
- Zucchini, W. (2003). *Applied smoothing techniques. Part 1: Kernel density estimation* (p. 20). Temple University. Retrieved from <http://staff.ustc.edu.cn/~zwp/teach/Math-Stat/kernel.pdf>