Evaluation of ECMWF Precipitation Predictions in China during 2015–18

CUI LIU,[a,b] JIANHUA SUN,[a,b,c] XINLIN YANG,[d] SHUANGLONG JIN,[e,f] AND SHENMING FU[g]

[a] *Key Laboratory of Cloud-Precipitation Physics and Severe Storms (LACS), Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*
[b] *University of the Chinese Academy of Sciences, Beijing, China*
[c] *Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Zhuhai, China*
[d] *Key Laboratory of Regional Climate-Environment for Temperate East Asia, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*
[e] *State Key Laboratory of Operation and Control of Renewable Energy and Storage Systems, China Electric Power Research Institute, Beijing, China*
[f] *Electric Power Meteorology State Grid Corporation of China Joint Laboratory, Beijing, China*
[g] *International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China*

(Manuscript received 13 August 2020, in final form 26 March 2021)

ABSTRACT: Precipitation forecasts from the ECMWF model from March to September during 2015–18 were evaluated using observed precipitation at 2411 stations from the China Meteorological Administration. To eliminate the influence of varying climatology in different regions in China, the stable equitable error in probability space method was used to obtain criteria for 3- and 6-h accumulated precipitation at each station and classified precipitation into light, medium, and heavy precipitation. The model was evaluated for these categories using categorical and continuous methods. The threat score and the equitable threat score showed that the model's forecasts of rainfall were generally more accurate at shorter lead times, and the best performance occurred in the middle and lower reaches of the Yangtze River basin. The miss ratio for heavy precipitation was higher in the northern region than in the southern region, while heavy precipitation false alarms were more frequent in southwestern China. Overall, the miss ratio and false alarm ratio for heavy precipitation were highest in northern China and western China, respectively. For light and medium precipitation, the model performed best in the middle and lower reaches of the Yangtze River basin. The model predicted too much light and medium precipitation, but too little heavy precipitation. Heavy precipitation was generally underestimated over all of China, especially in the western region of China, South China, and the Yungui Plateau. Heavy precipitation was systematically underestimated because of the resolution and the related parameterization of convection.

SIGNIFICANCE STATEMENT: Quantitative precipitation forecast is an important reference for operational weather forecasting. Verification of model-forecast precipitation in China is complicated because of its complex climatology. To reveal the spatiotemporal performance of the ECMWF model for 3- and 6-h precipitation forecasts in different regions of China, we defined thresholds for different rainfall categories from the cumulative precipitation at each station and evaluated the model after eliminating the influence of climatology. These verification results will help numerical model developers to improve their models and will also help forecasters have a better understanding of model predictions. Future research will focus on the accuracy of the model's predictions under different circulation situations.

KEYWORDS: Forecast verification/skill; Numerical weather prediction/forecasting; Precipitation

1. Introduction

Precipitation is one of the most important variables in weather forecasting. Currently, as a result of many factors, the accuracy of quantitative precipitation forecasts (QPF) is insufficient to satisfy societal requirements. Many countries have set up special projects to improve QPF, such as the Working Group on Numerical Experimentation in the twentieth century (Ebert et al. 2003), the Improvement of Microphysical Parameterization through Observational Verification Experiment in 2001 (Stoelinga et al. 2003), and the Southern China Monsoon Rainfall Experiment in China during 2013–15 (Luo et al. 2017). With the rapid increases in computational power of massively parallel computers and improvements in the assimilation of high-resolution atmospheric observations over the last decade, quantitative precipitation prediction using numerical models has become important in operational weather prediction (Golding 2000; Benjamin et al. 2019). Evaluation of the performance of these model predictions could help researchers improve the various processes in the models and help users understand their reliability (Casati et al. 2004; Kober et al. 2012; Kniffka et al. 2020). Many operational weather forecast centers take QPF verification scores as important indices to represent the performance of precipitation models. Since the outputs from numerical models have become widely used in operational QPF, researchers and operational forecasters have conducted in-depth evaluations of QPF accuracy (Mullen and Buizza 2001; Rodwell et al. 2010; Haiden et al. 2012). Owing to the complexity of its topography and its highly

Corresponding author: Jianhua Sun, sjh@mail.iap.ac.cn

DOI: 10.1175/WAF-D-20-0143.1

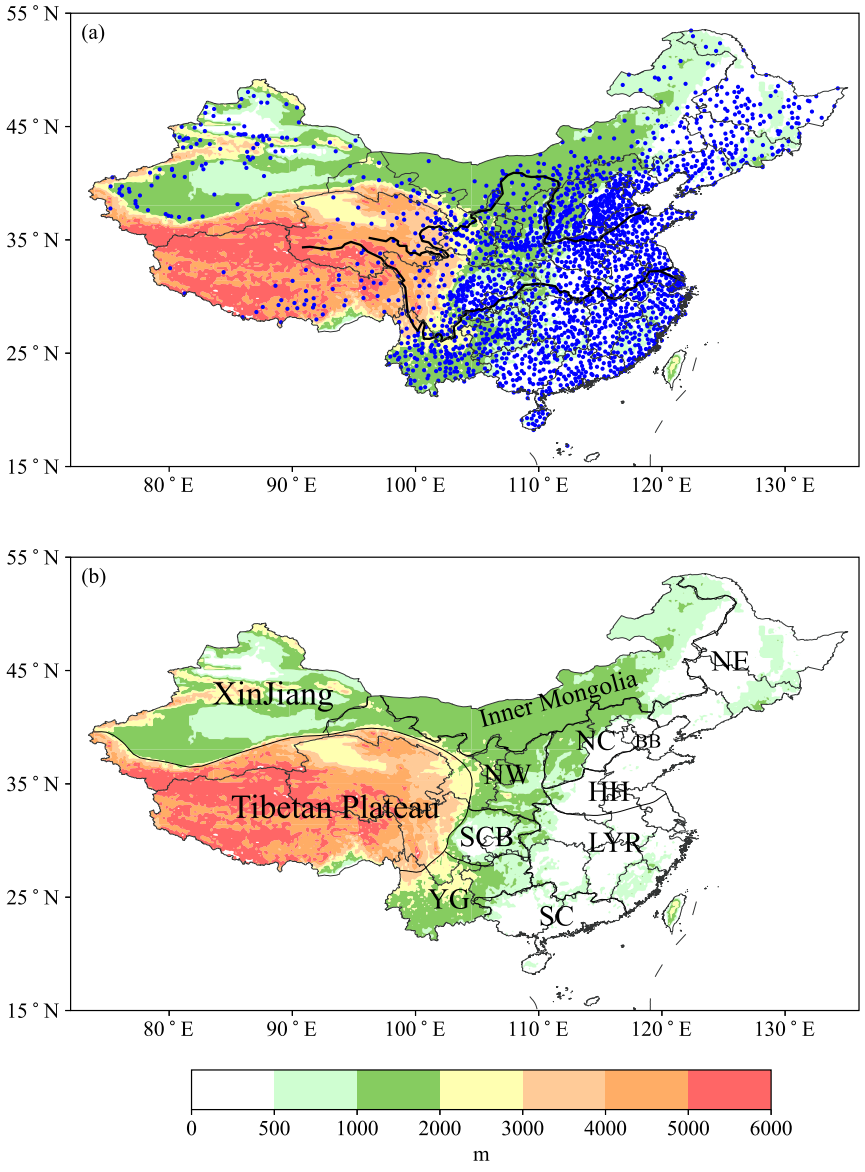

FIG. 1. (a) Geographical distribution of 2411 observation stations (blue dots) and elevation (shading; m) in China. (b) Regional divisions and elevation of China (shading; m): Sichuan basin (SCB), North China (NC), Northeast China (NE), Huang–Huai River basin (HH), middle and lower reaches of Yangtze River basin (MLYR), South China (SC), Yungui Plateau (YG), Northwest China (NW), and Bohai Bay (BB).

variable climate, China lacks in-depth and long-term verifications of precipitation models, and there are no objective quantitative correction methods and indexes pertaining to the country, which has led to an insufficient understanding of the results from numerical models and limited QPF correction performance (Bi et al. 2016). Verification of numerical model precipitation forecasts is therefore important for improving QPF in China.

Generally, for QPF verification scores to be useful, they need to be clearly and intuitively understood by weather forecasters. Some categorical scores are used to verify the accuracy of operational forecasts, such as the critical success index, also known as the threat score (TS) (Schaefer 1990), modified Heidke skill score (Doswell et al. 1990), odds ratio (Stephenson 2000), and the equitable threat score (ETS) (Baldwin and Kain 2006;

TABLE 1. Rainfall classification standards of the CMA.

Category	1 h (mm)	12 h (mm)	24 h (mm)
Light rainfall	0.1–2.5	0.1–5	0.1–10
Moderate rainfall	2.5–8	5–15	10–25
Heavy rainfall	8–15	15–30	25–50
Torrential rainfall	>16	>30	>50

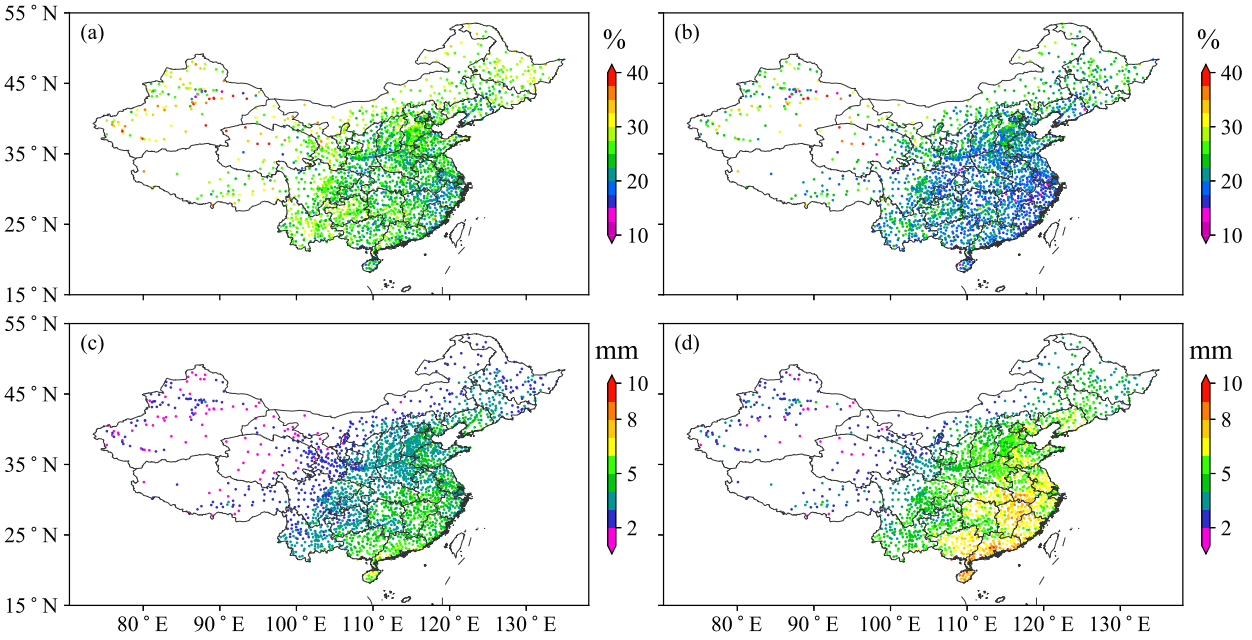

FIG. 2. The number of light precipitation events (accumulated precipitation no greater than 0.2 mm) as a percentage of the total number of precipitation events (%): (a) 3 and (b) 6 h. Threshold of medium and heavy precipitation of stations (mm): (c) 3 and (d) 6 h.

Brill and Mesinger 2009). Among these, the TS and the ETS are the most commonly used scores for precipitation evaluation because their physical meaning is clear. However, these scores only indicate whether the predicted precipitation is correctly forecast at a specific location, but do not reflect the degree of precipitation deviation between the observation and prediction. Consequently, two categorical methods associated with spatial diagnostics have been developed: filtering and displacement (Harris et al. 2001; Roberts and Lean 2008; Wang and Lu 2010; Liu et al. 2011). The filtering method is mainly used to evaluate the ability of models to predict multiscale weather systems, while the displacement method takes more account of the deviations in location, area, and orientation, between forecasts and observations. The common verification techniques are the grid-to-grid and site-to-site methods. The site-to-site method is simple and practical (Harris et al. 2001; Tiziana et al. 2002; Accadia et al. 2005) and can provide users with the basic features and average accuracy of the precipitation forecast. This method has been used widely for a long time by the National Meteorological Center of the China Meteorological Administration (CMA) (Wang and Yan 2007; Xiong 2011).

Heavy rainfall in mountainous regions more frequently leads to secondary disasters such as landslides and debris flows. Some research has focused on the evaluation of precipitation prediction by the ECMWF model in eastern regions of China (e.g., Huang and Luo 2017; Cao et al. 2019), but there has been a lack of attention to mountainous regions and the Tibetan Plateau in western China. It is therefore necessary to investigate the performance of the ECMWF precipitation model in China as a whole. Previous studies attempting to verify precipitation forecasts in China have mostly used the same thresholds for different categories of rainfall, as is defined and used operationally by the CMA (Su et al. 2014; Fu et al. 2014; Pan et al. 2017a; Cao et al. 2019). However, climatological differences lead to great differences in the total amount and numbers of precipitation in different regions of China (Qian and Lin 2005; Li et al. 2015). When evaluating a precipitation model, the threshold at a given station should therefore be defined taking into account the climatological characteristics of precipitation at that station (Rodwell et al. 2010). In the present study, the stable equitable error in probability space (SEEPS) method is employed to obtain thresholds for different categories of precipitation at each station. The thresholds obtained using this method are objective and reflect the precipitation climatology of each station (Rodwell et al. 2010; Haiden et al. 2012).

The present study evaluates ECMWF model precipitation forecasts with lead times up to 168 h from March to September during 2015–18, using a time resolution of 3 h for lead times up to 72 h and a time resolution of 6 h for lead times greater than 72 h. This evaluation of a model with high spatial and temporal resolution over the whole of China is useful to both users and model developers. The remainder of the paper is organized as follows. Section 2 introduces the data and methods in detail. Section 3 focuses on the evaluation results for different categories of precipitation and lead time in China. Finally, the conclusions and a discussion are given in section 4.

TABLE 2. Two-category contingency table.

	Forecast	
Observation	Yes	No
Yes	N_A	N_C
No	N_B	N_D

TABLE 3. Evaluation of scores and their definitions.

Definition	Formula	Reference
Threat score	$\mathrm{TS} = \frac{N_A}{N_A + N_B + N_C}$	Schaefer (1990)
Equitable threat score	$\mathrm{ETS} = \frac{N_A - N_{\mathrm{ref}}}{N_A - N_{\mathrm{ref}} + N_B + N_C}$, where $N_{\mathrm{ref}} = \frac{(N_A + N_B)(N_A + N_C)}{(N_A + N_B + N_C + N_D)}$	Baldwin and Kain (2006); Brill and Mesinger (2009)
Special miss ratio	$\mathrm{SMR} = \frac{N_{SC}}{N_{SA} + N_{SC}}$	Rossa et al. (2008)
Special false alarm ratio	$\mathrm{SFAR} = \frac{N_{SB}}{N_{SA} + N_{SB}}$	Huang and Luo (2017)
Bias score	$\mathrm{BS} = \frac{N_A + N_B}{N_A + N_C}$	Haiden et al. (2012)
Mean error	$\mathrm{ME} = \frac{\sum_{i=1}^{N} (F_i - O_i)}{N}$	Hong (2003); Chien et al. (2006)
Normalized root-mean-square error	$\mathrm{NRMSE} = \frac{\sqrt{\frac{\sum_{i=1}^{N} (F_i - O_i)^2}{N}}}{R}$	

2. Data and methodology

a. Observational precipitation data

Observational hourly precipitation data provided by the CMA from Chinese national stations for the period 2015–18 were used to verify model precipitation predictions. Figure 1a shows the geographical distribution of the 2411 national CMA sites. Most of these stations are distributed across the central and eastern regions, while the station density is relatively low in the northwestern regions and the Tibetan Plateau. In this study, the area east of 105°E is referred to as the eastern region of China, and the area north of 35°N is

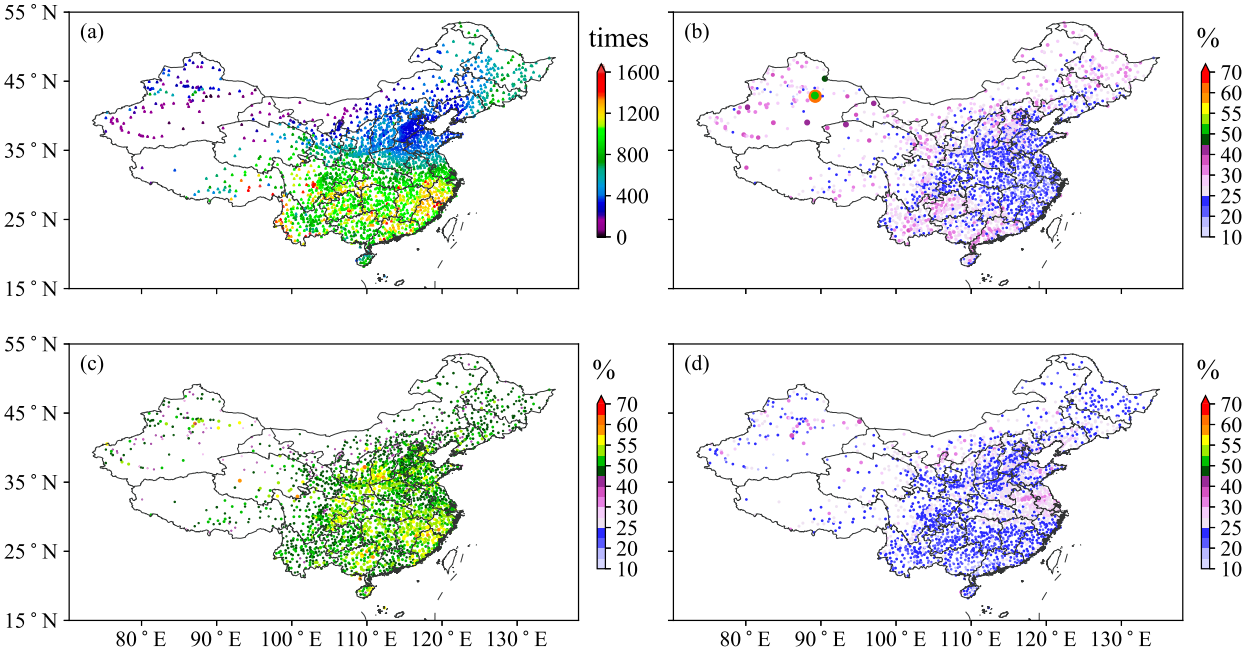

FIG. 3. (a) The number of total precipitation events, and proportion of (b) light precipitation events, (c) medium precipitation events, and (d) heavy precipitation events (%).

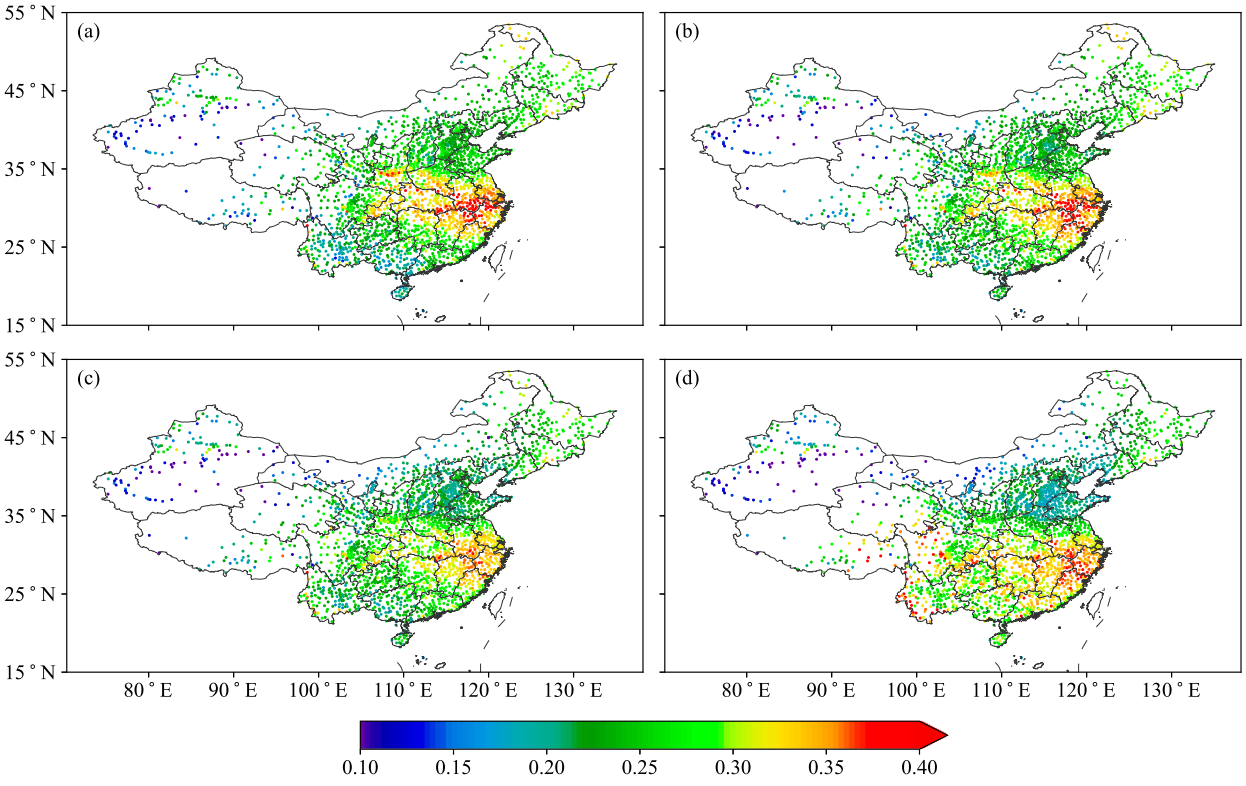

FIG. 4. Spatial distribution of the TS for "rain or no rain": (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

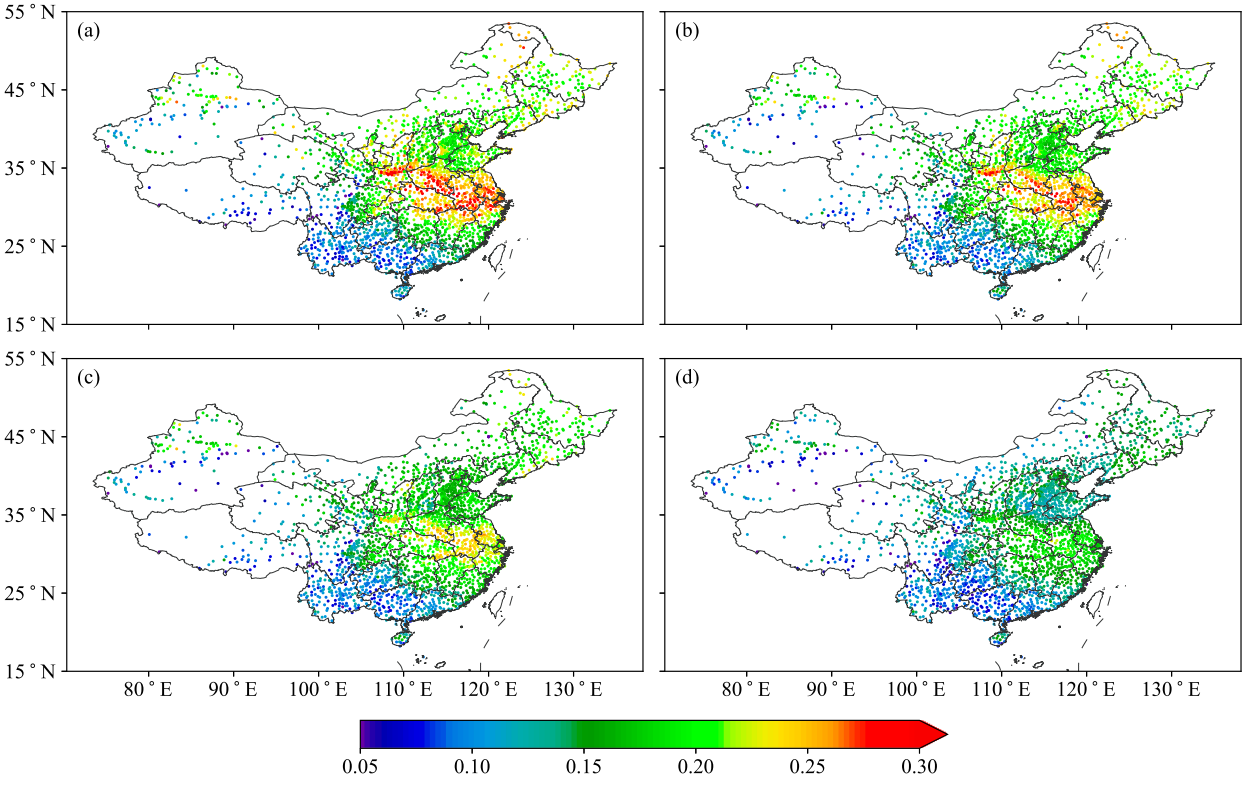

FIG. 5. Spatial distribution of ETS for "rain or no rain": (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

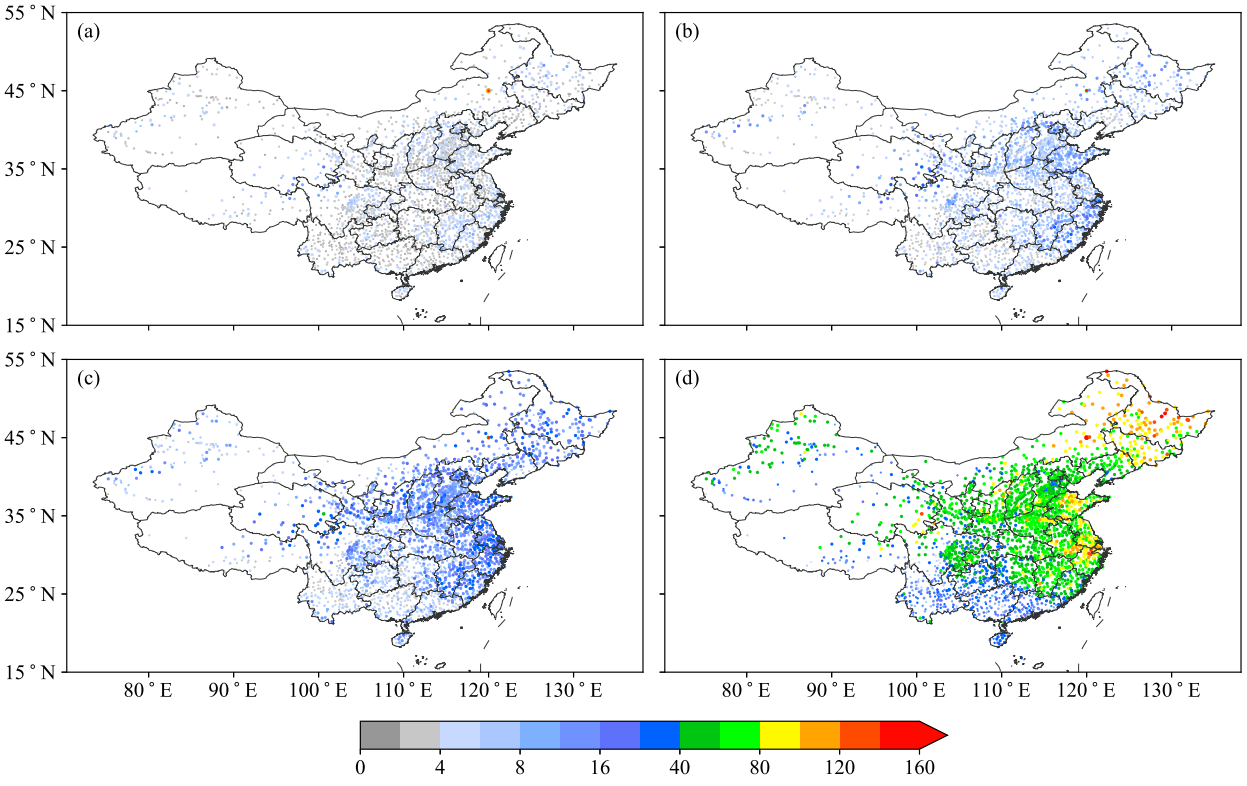

FIG. 6. Number of special miss events: (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

referred to as the northern region of China. Figure 1b shows the detailed regional division of China used in the present study. This observational hourly precipitation dataset has been used in various previous weather and climate studies (Yu et al. 2010; Li et al. 2015; Zheng et al. 2016). Although the hourly precipitation data provided by the CMA is quality controlled, further quality control was carried out to ensure its accuracy for this study. Our quality control focused on samples whose hourly precipitation was more than 30 mm h^{-1}. If there was no precipitation at the surrounding stations within 3 h before and after such samples, it was considered that they were incorrect. Based on this quality-control process, 138 precipitation records were excluded.

b. *Model precipitation data*

The forecast data used in this study are from the ECMWF operational-forecast model, at 0.125° × 0.125° resolution, initialized daily at 0000 UTC, and with lead times of 7 days (168 h). Precipitation forecasts from March to September 2015–18 (i.e., spring, summer, and early autumn) were evaluated. The time resolution is 3 h during days 1–3 and 6 h during days 4–7. The gridded precipitation forecast data were interpolated to the CMA stations (Fig. 1a) using linear interpolation (North et al. 2013).

c. *Verification methods*

To examine the predictive ability of the ECMWF operational model for different regions and precipitation categories in China, precipitation thresholds should be obtained for the different regions on the basis of local climatology. However, in CMA operations, all stations in China are assigned the same threshold (Table 1), which does not account for the geographical characteristics of climatological cumulative precipitation. It is necessary to define thresholds for different 3- and 6-h precipitation categories for each station separately to account for this geographical variation.

The SEEPS method proposed by the ECMWF can better account for the geographical variation in precipitation climatology (Rodwell et al. 2010), as it provides thresholds for different precipitation categories under different geographical conditions and can be used to verify the accuracy of precipitation forecasts under varying climates (Rodwell et al. 2010; Haiden et al. 2012; North et al. 2013). Rainfall in China mainly occurs in the warm season, controlled by the East Asian summer monsoon (Tao and Chen 1987; Xie and Arkin 1997; Wang and Lin 2002), but the rainfall begins in April over South China. Based on observed precipitation data from April to September of 2013–17, the SEEPS was employed to classify precipitation into light, medium, and heavy precipitation. All precipitation events in this station are sorted by rainfall amount, Light precipitation is defined as having a 3- or 6-h accumulation no greater than 0.2 mm ($P \leq 0.2$ mm), while the number of the medium precipitation events is defined to be twice that of heavy precipitation events. For instance, all precipitation event (>0.2 mm) is 120 in station A, which is sorted as $P(1), \ldots, P(80), \ldots, P(120)$ based on precipitation amount.

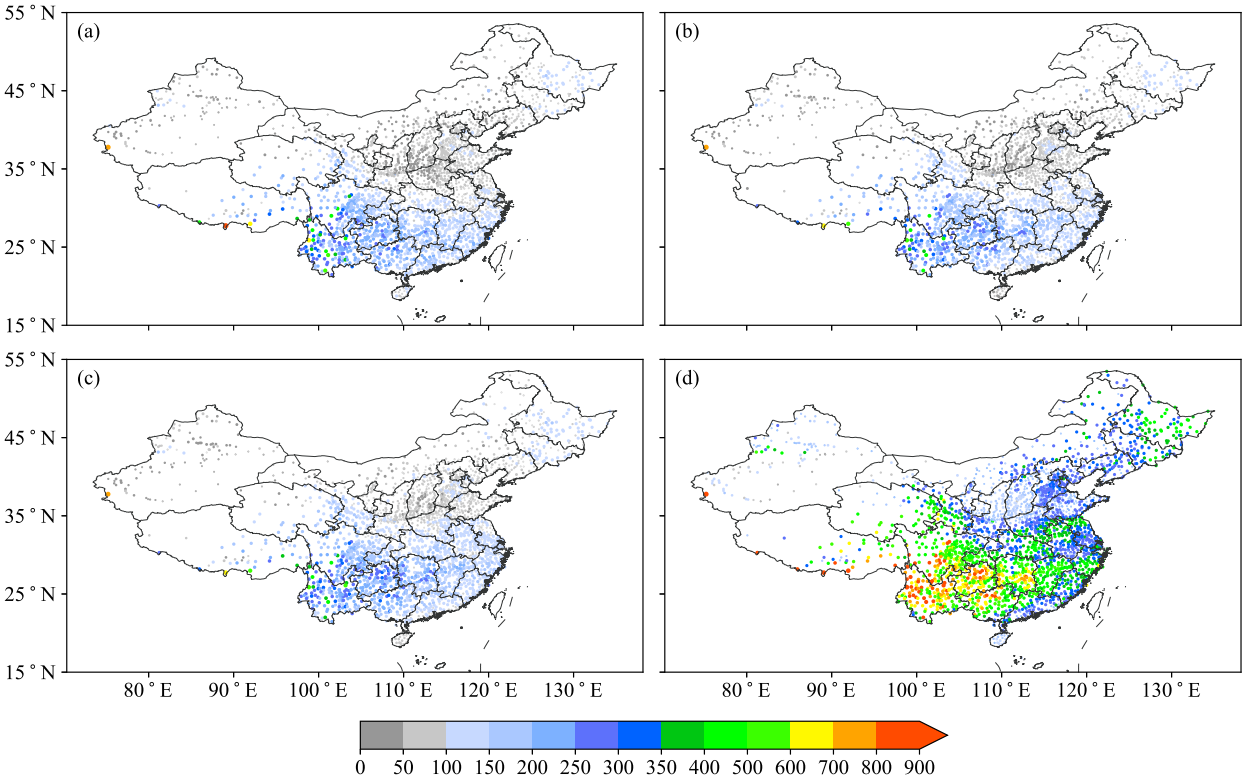

FIG. 7. Number of special false alarm events: (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168.

medium precipitation is 0.2 mm $< P <$ P (80), heavy precipitation events is $P \geq P$ (80). Thereby, the threshold between medium and heavy precipitation can be obtained from all the nonlight precipitation data. In this paper, we only used the SEEPS to classify the categories of precipitation, and used the commonly operational skill scores to evaluate precipitation prediction in China.

Figures 2a and 2b show the number of light precipitation events (accumulated precipitation no greater than 0.2 mm) as a percentage of the total numbers of precipitation events for 3- and 6-h accumulations, respectively. The proportion of 3-h light precipitation events in the eastern region of China is generally lower than that in the western region (Fig. 2a); the proportion in the middle and lower reaches of the Yangtze River basin (YRB) and South China was less than 25%, and that in the other eastern stations was about 25%–30%. The proportion from the west of the Tibetan Plateau to the middle of Xinjiang was highest, at about 32%–40%. The distribution of the proportion of 6-h accumulation was similar to that of 3-h accumulation, but the values of the 6-h proportion were less than the 3-h values (Fig. 2b). The geographic distribution patterns of the medium/heavy threshold were similar for 3- and 6-h accumulations; the maximum thresholds were about 7 mm for 3 h and 10 mm for 6 h (Figs. 2c,d), and the values in South China and the middle and lower reaches of the YRB were similar, with corresponding thresholds of 5–7 and 6–10 mm, respectively. The medium/heavy thresholds at most stations in the western region were less than 3 mm for 3-h and 4 mm for 6-h accumulation (Figs. 2c,d). The topography and latitude are the major factors which determine the thresholds. For example, thresholds in Yungui Plateau and Sichuan basin were similar, North China had similar thresholds with Northwest China. These results are consistent with the geographic distributions of precipitation climatology obtained in previous studies (Yu et al. 2007; He and Zhang 2010; Li et al. 2015), indicating that the criteria used reflect the climatology of the stations, and so these station-dependent thresholds will produce a more practically applicable verification than applying one threshold to all stations in China.

Based on these thresholds, precipitation forecasts from the ECMWF operational model were evaluated using various methods (Table 3). The categorical evaluation includes the TS, ETS, special miss ratio (SMR), special false alarm ratio (SFAR), and bias score (BS), and the continuous analysis includes the mean error (ME), and normalized root-mean-square error (NRMSE).

1) THREAT SCORE AND EQUITABLE THREAT SCORE

The TS is also known as the critical success index. As shown in Table 2, N_A is the number of hits, N_B is the number of false alarms, and N_C is the number of misses. A higher TS means a better performance. The ETS includes all the situations in Table 2; N_D is the number of correct negatives and N_{ref} is the number of possible correct forecast under random conditions. The ETS expresses the hit rate and the total numbers of rainfall, relative to random incidents.

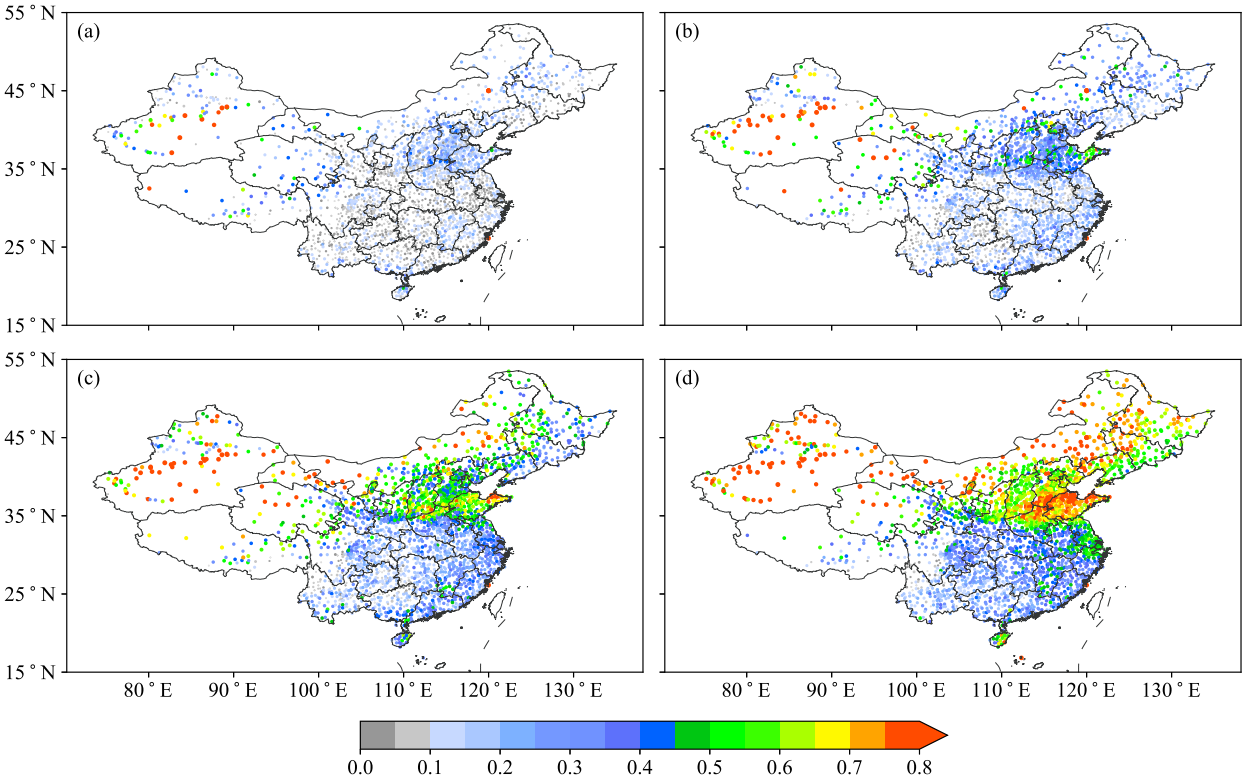

FIG. 8. Spatial distribution of SMR: (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

2) SPECIAL MISS RATIO AND SPECIAL FALSE ALARM RATIO

The miss and false alarm ratios are important for analyzing incorrect forecasts (Rossa et al. 2008; Huang and Luo 2017). Missing heavy precipitation can make disasters more serious; for example, the ECMWF model missed the heavy rainstorm in Fujian Province on 7–8 May 2016, which induced a landslide and 41 deaths (Dai et al. 2016). Generally, the heavier the precipitation, the more difficult it is to predict, and the more likely it is that extreme precipitation events are missed (Lu et al. 2013; Mao et al. 2016). To further verify the predictive capabilities of the ECMWF model for heavy precipitation, this study evaluated two particular situations that could lead to a more impact on the heavy precipitation prediction.

The first situation is where heavy precipitation is observed, but the forecast precipitation is 0 mm, so that the model does not predict the heavy precipitation at all: this situation is defined as a special miss, N_{SA} is the total number of heavy precipitation hits, N_{SC} is the total number of heavy precipitation misses (Table 3). The second situation is where heavy precipitation is predicted, but observed precipitation is 0 mm. This situation is defined as a special false alarm. As a heavy precipitation alarm is issued by CMA in this case, the special false alarm is likely to cause economic losses. Here, N_{SB} is the total number of heavy precipitation false alarms (Table 3). Both the SMR and the SFAR are used to evaluate incorrect forecasts of heavy precipitation.

3) BIAS SCORE

The BS quantifies the ratio of the number of forecast events to the number of observed events (Haiden et al. 2012). BS was calculated for precipitation thresholds of light, medium and heavy precipitation in this evaluation. BS varies from 0 to infinity, with a full score of 1.

4) MEAN ERROR AND NORMALIZED ROOT-MEAN-SQUARE ERROR

The ME is defined as the mean of the forecast precipitation minus observed precipitation (Hong 2003; Chien et al. 2006), which represents the bias between the forecast and the observations. In the expression for ME in Table 3, F_i indicates the model precipitation, O_i indicates the observed precipitation, and N is the number of the precipitation.

The RMSE is a statistical measure of the error between the observations and the forecast that has been used by many previous studies (Hong 2003; Chien et al. 2006; Su et al. 2014). To compare the relative forecast errors for different precipitation categories at each station, the station climatology must be eliminated, so the NRMSE is defined as the root-mean-square error at the station divided by its threshold for medium/heavy precipitation (R). Additionally, in order to investigate the spatial distribution of poor performing stations, the stations with the top 10% highest NRMSE in each precipitation category are selected to reveal the geographic characteristics of highest NRMSE.

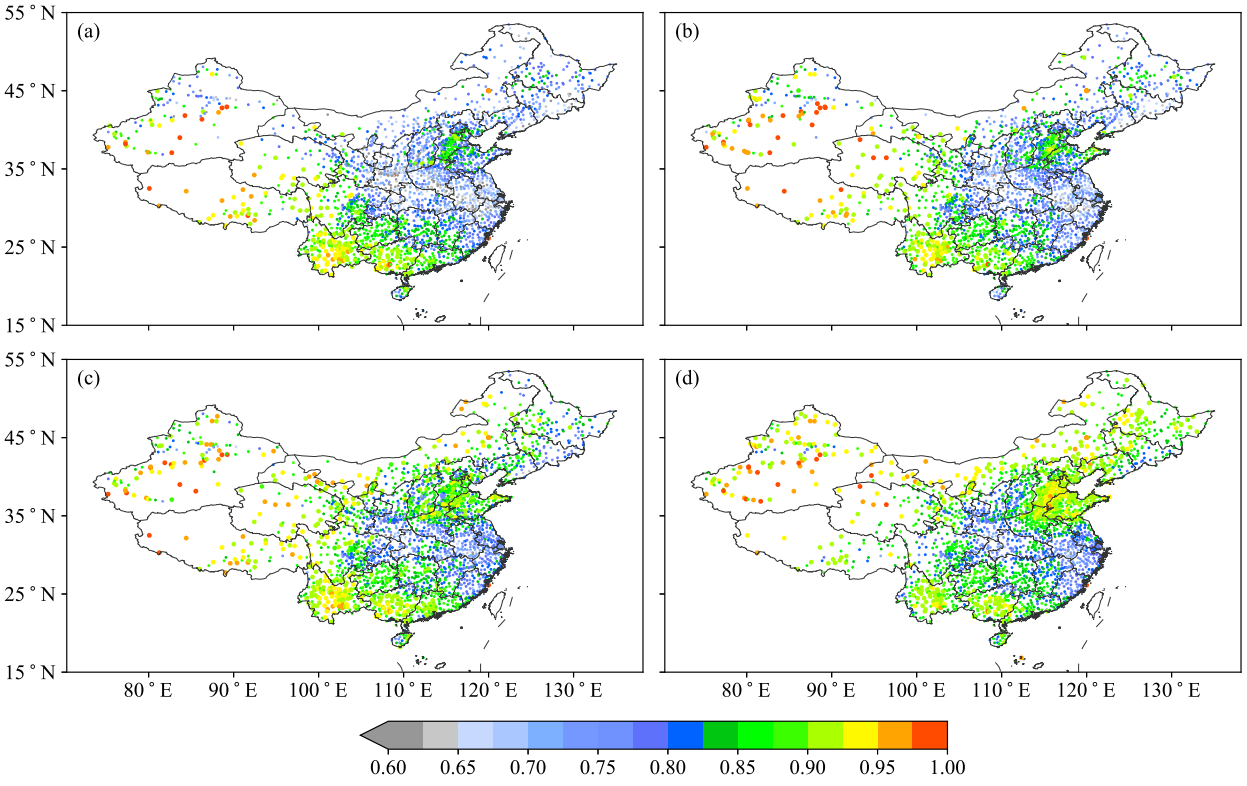

FIG. 9. Spatial distribution of SFAR: (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

In this study, the 3-h threshold was applied to forecasts for days 1–3 and the 6-h threshold to forecasts for days 4–7. To further understand the performance from day 1 to day 3, the scores computed for each contingency table of separate 3h time periods for days 1–3, and then averaged for day 1, day 2 and day 3, respectively. The scores computed for each contingency table of separate 6 h time periods for days 4–7, and then averaged as a whole. The following analysis refers to 0–24, 24–48, 48–72 and 72–168 h, respectively, as the first, second, third, and fourth periods.

3. Results

a. Spatial distribution of precipitation

Figure 3 illustrates the total number of precipitation events and the proportion of the three categories. This shows that the numbers of precipitation events in the southern region of China were higher than those in the northern region. In particular, the Sichuan basin, the Yungui Plateau, and the middle and lower reaches of the YRB had higher times, while the western part of Inner Mongolia Autonomous Region (hereafter referred to as Inner Mongolia) and Xinjiang Uygur Autonomous Region (hereafter referred to as Xinjiang) had the lowest times (Fig. 3a). The distribution was consistent with the precipitation climatology of 1982–2012 shown by Li et al. (2015). The proportion of light precipitation is lower than 40%, the middle and lower reaches of YRB had lowest proportion, about 10%–25% (Fig. 3b). The proportion of medium and heavy precipitation was different to that of light precipitation (Figs. 3b–d). The highest proportion of medium precipitation, about 55% appeared over the middle and lower reaches of YRB, northwest China and South China (Fig. 3c). South China had the lowest proportion of heavy precipitation (Fig. 3d). In summary, the number in the southern region of China was higher than that in the northern region, and the lowest was in southern Xinjiang.

b. Threat score and equitable threat score of "rain or no rain" events

The ability to forecast the occurrence or nonoccurrence of rain is a crucial factor in numerical modeling and operational forecasting. The geographic distribution of TS for "rain or no rain" events was similar over days 1–3 (Figs. 4a–c). The TS in the eastern region of China was relatively high, especially in the middle and lower reaches of the YRB, where it was higher than 0.3. The TS in the western region of China was generally lower than in the eastern region, which is consistent with the results of Zhang et al. (2014). The TS in southern Xinjiang was the lowest, with values around 0.1. The high-value region of the ETS for "rain or no rain events" was similar to that of the TS (Fig. 5); both were located in the middle and lower reaches of YRB, while low values appeared over the Yungui Plateau and South China. The ETS of "rain or no rain events" decreased with lead time. In conclusion, the ECMWF model showed different skill in the regions with similar thresholds between the middle and lower reaches of the YRB and South China. The model performed best in the middle and lower reaches of

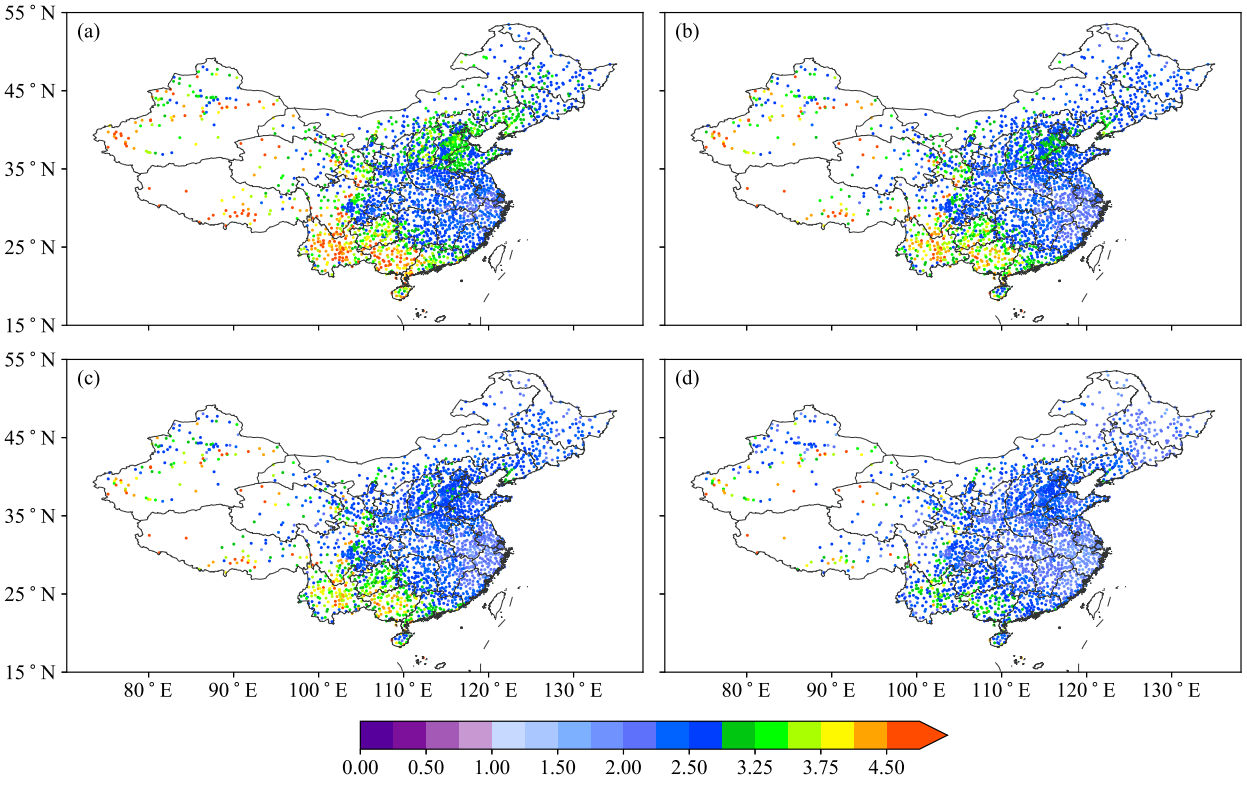

FIG. 10. Distribution of BS for precipitation (≥light precipitation threshold): (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

the YRB in the eastern region of China and worst in South China. The possible reasons for this result can be understood by further studying the characteristics of atmospheric circulation and the underlying surface effects in different regions of China.

c. *Special miss ratio and special false alarm ratio*

Figure 6 shows the spatial distribution of the number of special misses. In the first three periods, the number and range of special misses increased with lead time, especially in the lower reaches of the YRB, North China, Northeast China, and the Tibetan Plateau, and the number of special misses in the Yungui Plateau was the lowest in the southern region of China (Figs. 6a–c). In the fourth period, the number of special misses was more than 60 in the northeast China, followed by eastern Huang–Huai River basin and the lower reaches of the YRB, and the lowest numbers occurred in South China and the southern Yungui Plateau (Fig. 6d). Figure 7 shows the spatial distribution of the number of special false alarm. The number was higher in the southern region of China than in the northern region in the four evaluation periods. There was a considerable different number between the special miss and the special false alarm, that is, the number of special false alarm did not increase with the lead time, and the high-number area of special false alarm was in southwestern China in all evaluation periods.

The SMR in China generally decreased from the northern region to the southern region and increased with lead time (Fig. 8). The SMR in southern Xinjiang was highest in all evaluation periods, indicating that the prediction of heavy rainfall in this area is very difficult. However, in the past, little attention has been paid to heavy rainfall in Xinjiang, and the major characteristics of heavy rainfall and their formation mechanisms in Xinjiang should be studied in the future. The SMR in the middle and lower reaches of the Yellow River and Inner Mongolia increased considerably in the third period (Fig. 8c), and the SMR in some regions of northern China increased to more than 0.5 in the fourth period (Fig. 8d). Comparing the number of special misses and the SMR, it was found that the SMR was not high in some areas with high number of special miss in eastern China, because there were more heavy rainfall events (Fig. 3d).

Unlike the situation for special misses, the number of special false alarm was higher in the Yungui Plateau (Figs. 6 and 7). However, the spatial distributions of special false alarm and SFAR are different (Figs. 7 and 9). The South China, the Yungui Plateau, and the western part of China, had a higher SFAR in the first period and the second period (Figs. 9a,b). In the third period, there was a large area with SFAR values higher than 0.8 (Fig. 9c). In the fourth period, the range of the large-value area of SFAR further expanded, and the SFAR in Huang–Huai River basin increased considerably, the SFAR values for all stations were higher than 0.8, expect the lower reaches of the YRB (Fig. 9d).

d. *Bias score*

The BS is the ratio of the forecasted number to observed rainfall events. BS < 1 indicates underestimation tendency, and BS > 1 indicates overestimation tendency.

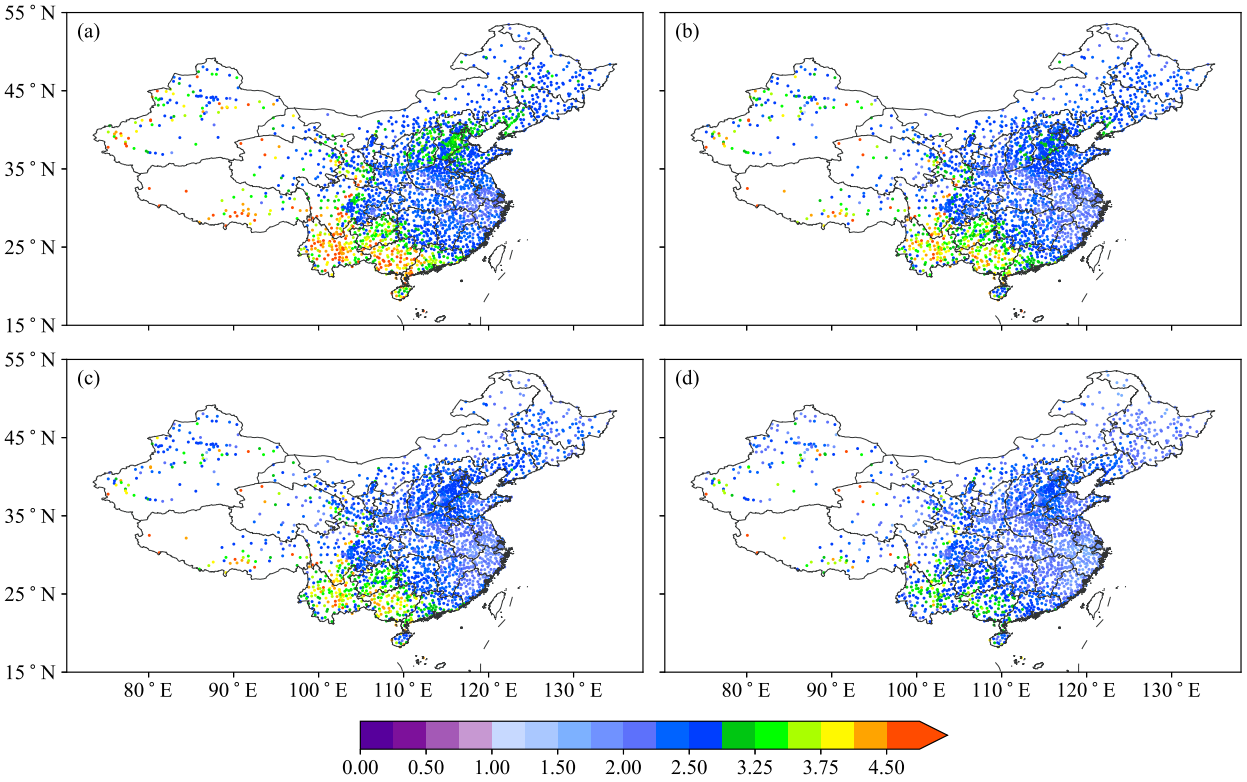

FIG. 11. Distribution of BS for precipitation (≥medium precipitation threshold): (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

ECMWF model overestimated precipitation (≥light precipitation threshold) in whole China in all estimation periods, and the BS decreased with lead time (Fig. 10). During the first three periods, the stations with BS values more than 3 were concentrated in the Yungui Plateau, South China, and the western region (Figs. 10a–c). The middle and lower reaches of the YRB has lowest BS values in all estimation periods. ECMWF model is better at predicting precipitation (≥light precipitation threshold) in the middle and lower reaches of the YRB. The BS values remained above 1 in the southern region for all periods (Fig. 10). Previous studies have shown that the ECMWF model overestimates light rainfall (Haiden et al. 2012; Pan et al. 2017b), which is consistent with the results of this study. For precipitation (≥medium precipitation threshold), the distribution of BS were same as precipitation (≥light precipitation threshold), but the BS values were relative lower than that of light precipitation (Fig. 11). Stations with BS above 3 are still concentrated in the Yungui Plateau and South China.

The BS values for heavy precipitation were less than 2 at most stations in all evaluation periods, especially in the South China (Fig. 12). Comparing with the light and medium precipitation, the overestimation tendency for heavy precipitation is weaker. The underestimation tendency in South China with BS less than 1, increased considerably with lead time. In the first period, stations with BS values higher than 2 were only distributed across Bohai Bay and the Yungui Plateau (Fig. 12a). During the second and three period, the BS in South China decreased considerably (Figs. 12b,c). In the fourth period, the BS values of stations in South China were considerably less than 1, which indicated that the heavy rainfall in South China was underestimated seriously. The underestimation of heavy precipitation became worse with lead time. Some previous studies have also shown that the ECMWF model underestimates heavy precipitation and this underestimation becomes worse with increasing lead time (Haiden et al. 2012; Pan et al. 2017b; Fu et al. 2014). However, BS only give the tendency of overestimation or underestimation for different categories of precipitation, ME could reveal detail information of overestimation or underestimation.

e. *Mean error*

The above analyses of TS, ETS, and BS reveal the basic performance of the model precipitation forecast, but the continuous error has not been revealed. Therefore, continuous estimation methods, the ME and the NRMSE, are evaluated here. The larger the absolute value of the ME and the NRMSE, the worse the performance of the model in predicting precipitation is.

The BS analysis indicated that light precipitation was greatly overestimated in whole China. The spatial distribution of the ME for light precipitation in most parts of China ranged from 0 to 3 mm, and the ME in southern China was remarkably larger than in northern China. In South China, the ME at most stations was greater than 1 mm, and the maximum of the ME was 2–3 mm. In northern China, apart from Bohai Bay, where it

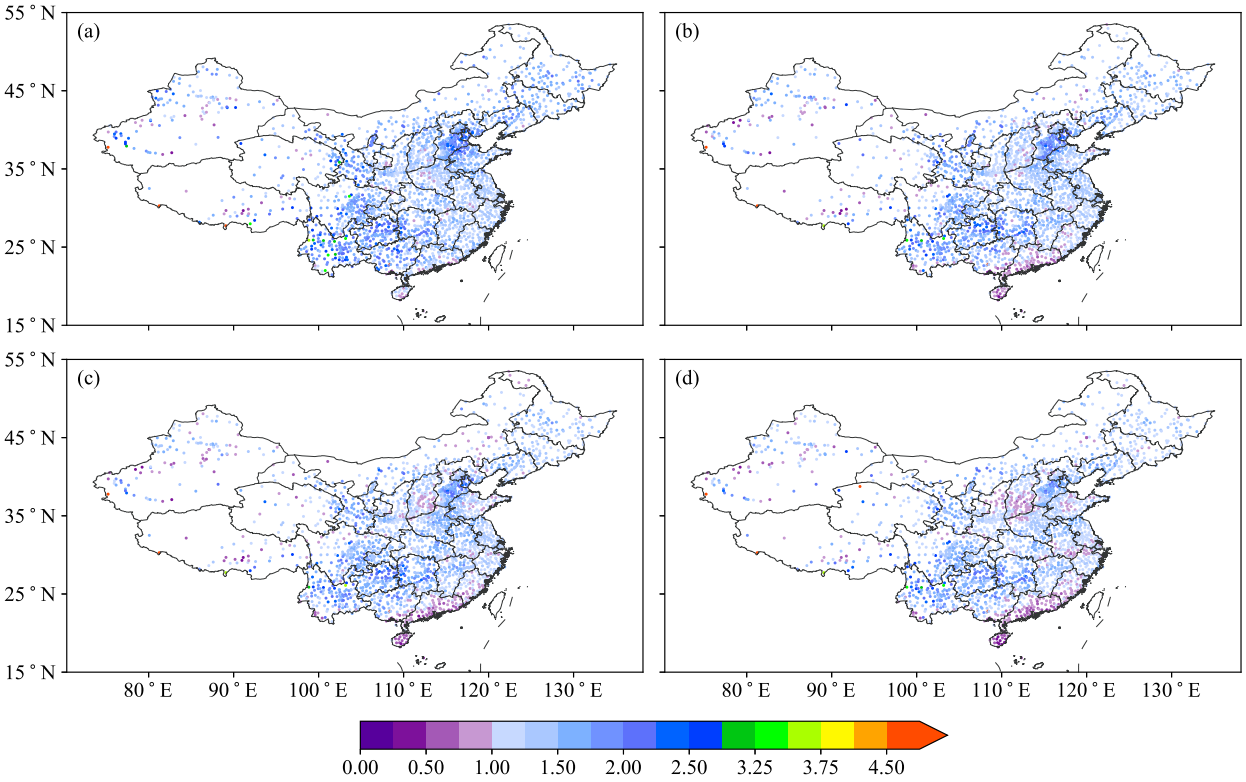

FIG. 12. Distribution of BS for precipitation (≥heavy precipitation threshold): (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

was 1–2 mm, the ME at most stations was less than 1 mm. The ME for light precipitation was positive over the whole of China, which means the overestimation was considerable for some light precipitation events. The geographical distribution of the ME for light precipitation found in this study was the same as that found by Zhang et al. (2017). For medium precipitation, the ME in the southern region was generally higher than in the northern region, and that in the eastern region was generally higher than that in the western region (Fig. 13b). The maximum positive error was between 1 and 2 mm in the middle and lower reaches of the YRB and the Yungui Plateau. The maximum negative error, between approximately −1 and −0.5 mm, appeared in the central part of Xinjiang.

The negative ME for heavy precipitation was considerably greater than that for the other two categories (Fig. 13), which also indicates that the model considerably underestimates heavy precipitation in China (Fig. 12). Some studies have revealed that the ECMWF model is poor at forecasting heavy rainfall (Accadia et al. 2005; Haiden et al. 2012; North et al. 2013). The distribution of the ME for heavy precipitation shows that error increased from the western region to the eastern region (Fig. 13c). The ME in the western region was between about −5 and 0 mm, and that in the eastern region was between −5 and −20 mm because the amount and the numbers of precipitation were higher in the eastern region than in the western region. The maximum negative ME, from about −15 to −20 mm, occurred in South China, followed by from about −10 to −15 mm in North China and the middle and lower reaches of the YRB, and the Yungui Plateau was from about −4 to −12 mm. It is clear that underestimation of heavy precipitation by the model was more serious in North China and South China than in other regions of China.

The reason of underestimation of heavy precipitation is complicated, and convective parameterization in model may impact the precipitation prediction seriously. Model precipitation is an average value in the grid box. Localized extreme values in precipitation totals are systematically underestimated because of the resolution and also the related parameterization of convection. The design of convective parameterization faces the difficulty of separating convective and nonconvective processes (Yano et al. 2003; Piriou et al. 2007). Detail is lost within the grid box due to subgrid variability, particularly in convective situations when the individual showers might be heavy but the displayed average precipitation is low. This leads to an overestimate of coverage, and an underestimate of heterogeneity and maximum intensity.

The above analysis shows that the model generally overestimates light precipitation and medium precipitation and underestimates heavy precipitation in whole China, which is consistent with previous evaluations of the ECMWF model (Fu et al. 2014; Moore et al. 2015). For light precipitation and medium precipitation, the areas with considerable overestimation areas were in southern China, while the areas with considerably underestimated heavy precipitation were South China and parts of North China.

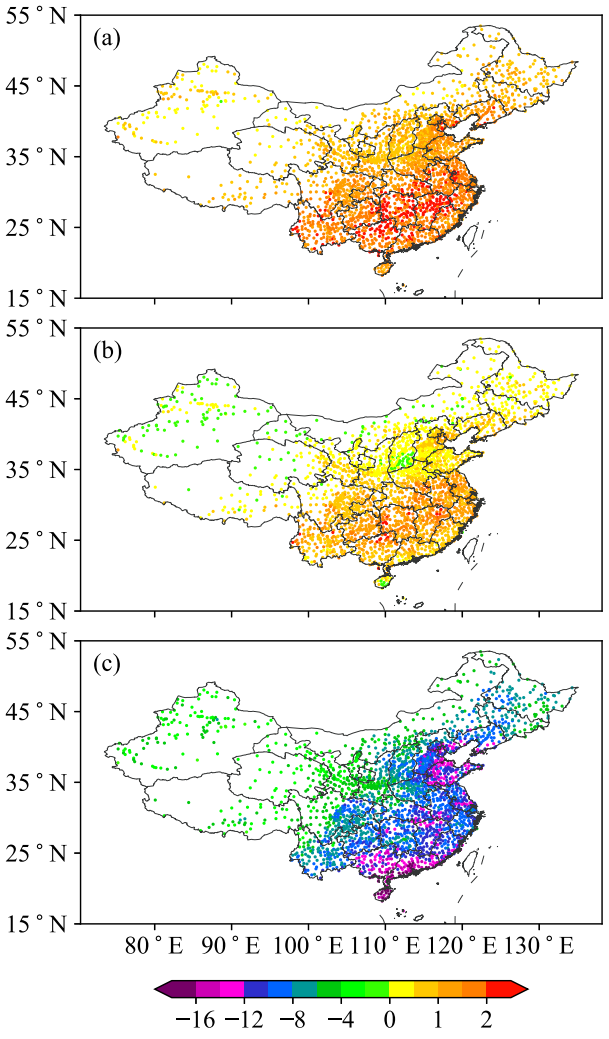

FIG. 13. Distribution of ME for different categories of precipitation (mm): (a) light precipitation, (b) medium precipitation, and (c) heavy precipitation.

f. Normalized root-mean-square error

The precipitation number and total precipitation amount varied greatly between different regions of China (Fig. 3). For example, the absolute value of the ME in South China was higher than that in North China (Fig. 13), which seems to suggest that the performance of the model in North China is better than that in South China. However, due to the high intensity of precipitation events in summer in North China, it is difficult to forecast precipitation (Tao 1980). The performance of the model in different regions cannot be compared without considering the climatology. The NRMSE is the RMSE divided by the threshold for distinguishing medium and heavy precipitation at each station (Fig. 2), which eliminates the influence of the climate at the station (Table 3).

Figure 14 shows that the NRMSE of light precipitation in the eastern region was generally greater than in the western region. The relatively higher NRMSE stations were scattered during the first three periods, and the stations with high values greater than 1.2 were concentrated in North China. In the fourth period, the high NRMSE areas were relatively more concentrated, but most of the stations had values less than 1.0 (Fig. 14d), which means the RMSE values at these stations were less than the threshold between medium and heavy precipitation. For medium precipitation, the distribution of the NRMSE was similar to that for light precipitation (Fig. 15). It was higher in the eastern region than in the western region, and the high-value areas were North China, the middle and lower reaches of the YRB, and the Sichuan basin. During the first period, the NRMSE values at eastern stations were mostly less than 1.0 (Fig. 15a). From the second to fourth period, the number of stations with an NRMSE greater than 1.0 increased considerably, mainly in the Huang–Huai River basin and North China, as well as in the Sichuan basin (Fig. 15b).

The NRMSE values for heavy precipitation were considerably higher than those for light and medium precipitation, but the values and ranges did not increase with lead time (Fig. 16). China is located in the East Asian monsoon region, with complex terrain and severe convective precipitation in summer (Tao and Chen 1987), so the accurate prediction of heavy precipitation is very difficult. In general, the NRMSE of heavy precipitation in the eastern region was higher than that in the western region, and the NRMSE values in the eastern region were greater than 2.5. North China, the Sichuan basin, and the Yungui Plateau had higher NRMSE values that changed little with lead time (Fig. 16). The complexity of the topography and the influence of systems in the Sichuan basin may be important reasons that the model had poor forecasting performance for rainstorms in the Sichuan basin (Zong et al. 2013).

To further understand the geographical distribution of the maximum error for the three precipitation categories, Fig. 17 shows the distribution for the three categories where the NRMSE reached the 90% quantile of all 2411 station values. The results show that the prediction error of the ECMWF model for North China, the Huang–Huai River basin, and the Yungui Plateau was high in each precipitation category, and higher than in other regions of China.

4. Conclusions and discussion

The high-spatiotemporal-resolution ECMWF model precipitation from March to September during 2015–18 was evaluated using hourly observation CMA precipitation data from 2411 national stations. Due to the great variations in precipitation climatology between different regions of China, the classification criteria of 3- and 6-h accumulated precipitation at each station were obtained using the SEEPS method. The threshold obtained by this method takes into account the climatological characteristics of different regions. Based on the threshold, the observed precipitation was divided into three types, light, medium, and heavy, and the accuracy of the precipitation forecasts of the ECMWF model in China was evaluated. The main findings can be summarized as follows.

In general, the middle and lower reaches of the Yangtze River basin had similar threshold with South China, the

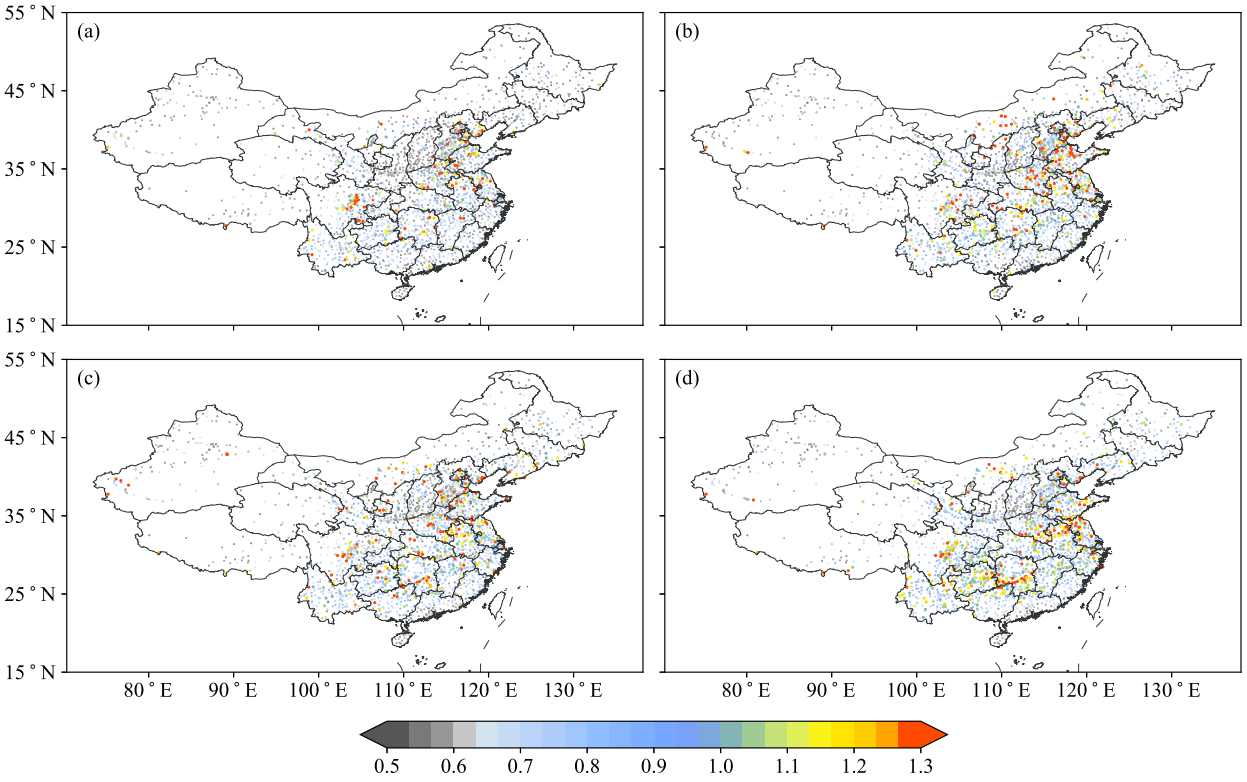

FIG. 14. Distribution of NRMSE for light precipitation: (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

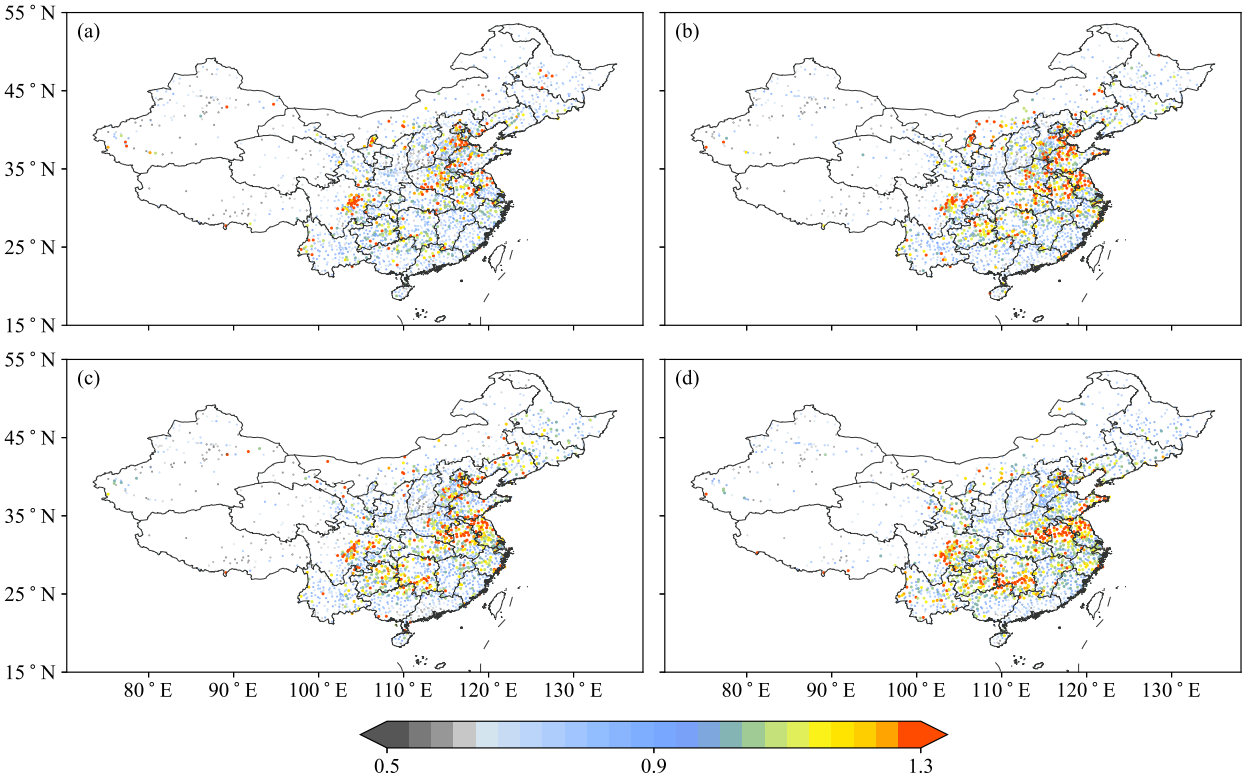

FIG. 15. Distribution of NRMSE for medium precipitation: (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

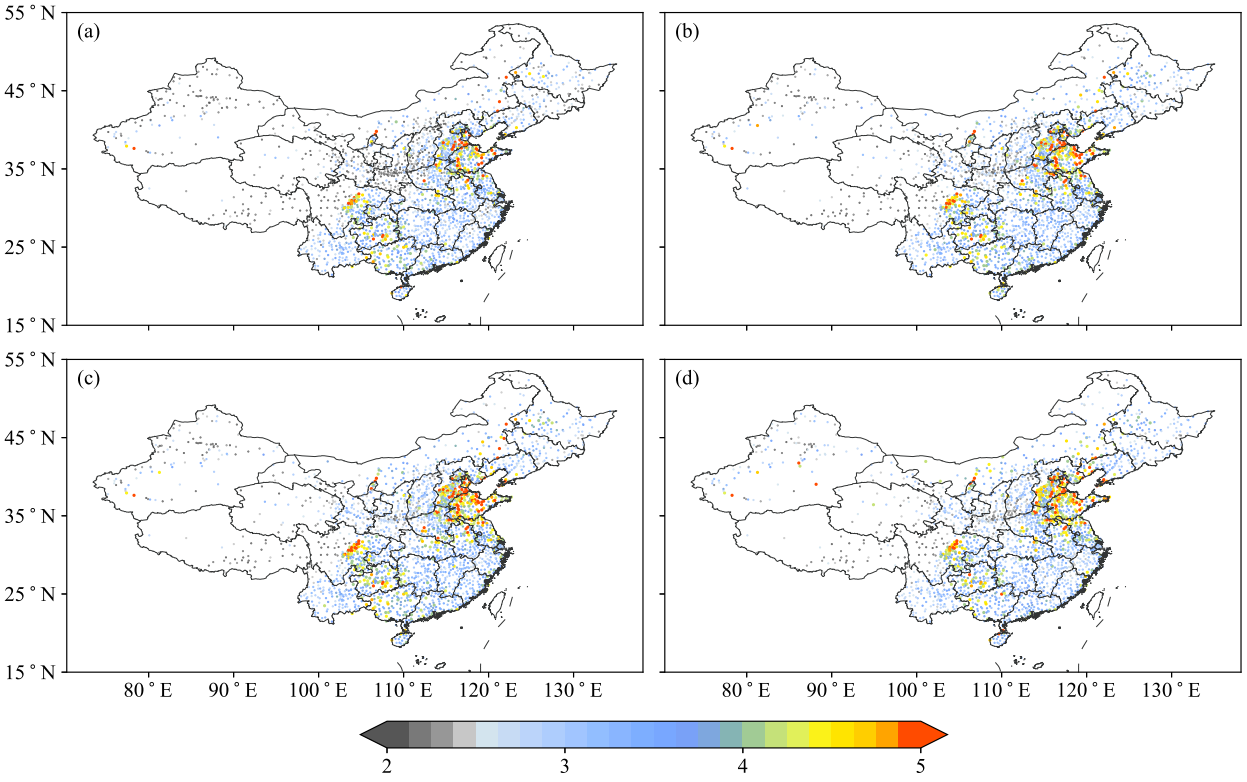

FIG. 16. Distribution of NRMSE for heavy precipitation: (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

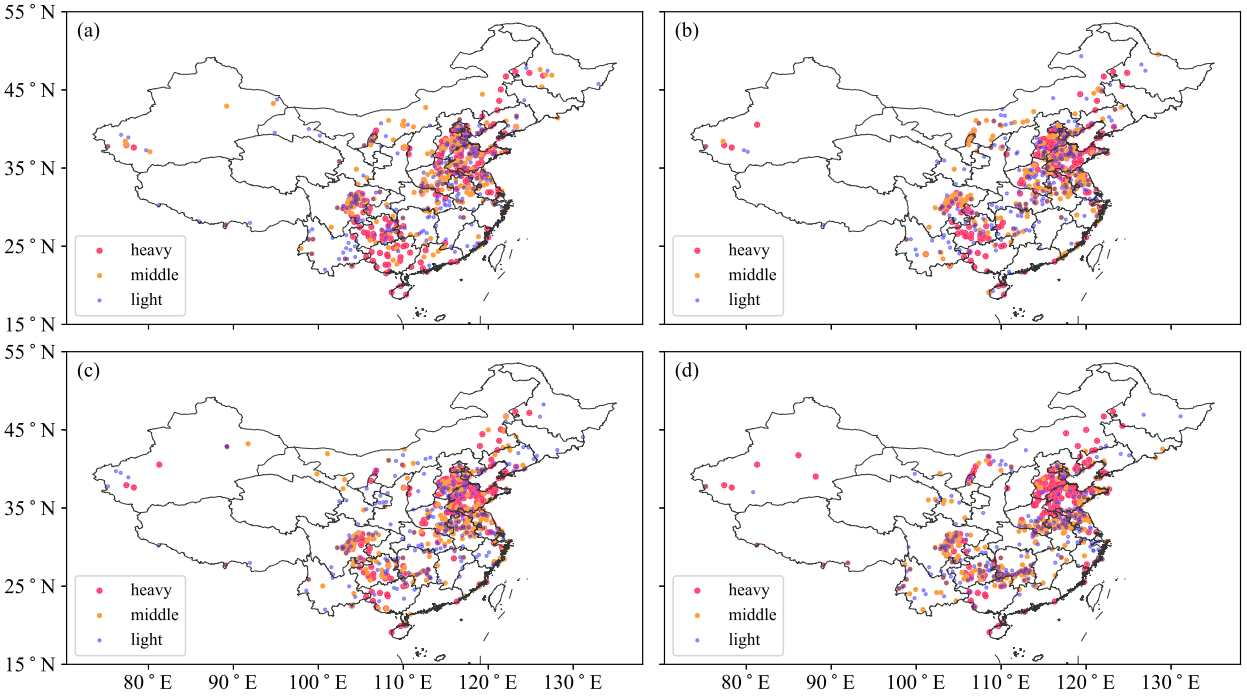

FIG. 17. The stations with the top 10% highest NRMSE in each precipitation category. Red spots represent heavy precipitation, orange spots represent medium precipitation, and blue spots represent light precipitation: (a) 0–24, (b) 24–48, (c) 48–72, and (d) 72–168 h.

thresholds in the Yungui Plateau were similar with Sichuan basin, while those of North China and Northwest China are close. The ECMWF model had better forecast skills in the middle and lower reaches of the YRB than in South China, model showed better predication skills in northwest China than in North China.

The results for the TS and the ETS of ''rain or no rain'' generally suggest that the best performance appeared in the middle and lower reaches of the Yangtze River basin, and the scores in South China and southwestern China were lower than those in other regions in eastern China. The shorter the lead time, the better the forecast accuracy. The number of special misses in eastern China was higher than in other regions, but the SMR in the northern region was higher than in the southern region, and it gradually decreased going from north to south. The number of special false alarms was highest in southwestern China, and the SFAR was higher in the western China, the Huang–Huai River basin, and Inner Mongolia. The results show that, for heavy precipitation, the miss ratio and false alarm ratio were mainly higher in northern China and western China.

The distributions of the BS, the ME and the NRMSE of the model precipitation show that the prediction error increased with lead time. Generally, the model overpredicted light and medium precipitation, but underpredicted heavy precipitation. Model precipitation is an average value in the grid box and the related parameterization of convection lead to a systematic underestimate of heavy precipitation intensity. The geographical distribution of the NRMSE shows that the highest forecast errors occurred in North China, the Huang–Huai River basin, and the Yungui Plateau.

On the basis of the present study, it can be explained that the ECMWF model misses more heavy precipitation in the northern region than that in the southern region of China, but the number of false alarms of heavy precipitation is highest in southwestern China. This may be due to the complex topography of the southwestern mountainous area, North China, and the Huang–Huai region (Chen et al. 2012; Sun and Zhang 2012; Bao and Zhang 2013). In addition, it will also be useful to study the errors in precipitation forecasts under different circulation situations, and this will form the basis of future work.

Acknowledgments. This research was supported by the Science and Technology Foundation of State Grid Corporation of China (Grant 5200-201955490A-0-0-00).

REFERENCES

Accadia, C., S. Mariani, M. Casaioli, M. A. Lavagnini, and A. Speranza, 2005: Verification of precipitation forecasts from two limited-area models over Italy and comparison with ECMWF forecasts using a resampling technique. *Wea. Forecasting*, **20**, 276–300, https://doi.org/10.1175/WAF854.1.

Baldwin, M. E., and J. S. Kain, 2006: Sensitivity of several performance measures to displacement error, bias, and event frequency. *Wea. Forecasting*, **21**, 636–648, https://doi.org/10.1175/WAF933.1.

Bao, X., and F. Zhang, 2013: Impacts of the mountain-plains solenoid and cold pool dynamics on the diurnal variation of precipitation over Northern China. *Atmos. Chem. Phys.*, **13**, 6965–6982, https://doi.org/10.5194/acp-13-6965-2013.

Benjamin, S. G., J. M. Brown, G. Brunet, P. Lynch, K. Saito, and T. W. Schlatter, 2019: 100 years of progress in forecasting and NWP applications. *A Century of Progress in Atmospheric and Related Sciences: Celebrating the American Meteorological Society Centennial*, *Meteor. Monogr.*, No. 59, Amer. Meteor. Soc., https://doi.org/10.1175/AMSMONOGRAPHS-D-18-0020.1.

Bi, B. G., K. Dai, Y. Wang, J. L. Fu, Y. Cao, and C. H. Liu, 2016: Advances in techniques of quantitative precipitation forecast (in Chinese). *Yingyong Qixiang Xuebao*, **27**, 534–549.

Brill, K. F., and F. Mesinger, 2009: Applying a general analytic method for assessing bias sensitivity to bias-adjusted threat and equitable threat scores. *Wea. Forecasting*, **24**, 1748–1754, https://doi.org/10.1175/2009WAF2222272.1.

Cao, Y., L. N. Zhao, Y. F. Gong, D. B. Xu, and Y. J. Gao, 2019: Evaluation and error analysis of precipitation forecast capability of the ECMWF high-resolution mode (in Chinese). *Torrential Rain Disasters*, **38**, 249–258.

Casati, B., G. Ross, and D. Stephenson, 2004: A new intensity-scale approach for the verification of spatial precipitation forecasts. *Meteor. Appl.*, **11**, 141–154, https://doi.org/10.1017/S1350482704001239.

Chen, M. X., Y. C. Wang, F. Gao, and X. Xiao, 2012: Diurnal variation in convective storm activity over contiguous North China during the warm-season based on radar mosaic climatology. *J. Geophys. Res.*, **117**, D20115, https://doi.org/10.1029/2012JD018158.

Chien, F. C., Y. C. Liu, and B. J. D. Jou, 2006: MM5 ensemble mean forecasts in the Taiwan area for the 2003 mei-yu season. *Wea. Forecasting*, **21**, 1006–1023, https://doi.org/10.1175/WAF960.1.

Dai, K., Y. Cao, Q. F. Qian, S. Gao, S. G. Zhao, Y. Chen, and C. H. Qian, 2016: Situation and tendency of operational technologies in short- and medium-range weather forecast (in Chinese). *Meteor. Monogr.*, **42**, 1445–1455.

Doswell, C. A., R. Davies-Jones, and D. L. Keller, 1990: On summary measures of skill in rare event forecasting based on contingency tables. *Wea. Forecasting*, **5**, 576–585, https://doi.org/10.1175/1520-0434(1990)005<0576:OSMOSI>2.0.CO;2.

Ebert, E. E., U. Damrath, W. Wergen, and M. E. Baldwin, 2003: The WGNE assessment of short-term quantitative precipitation forecasts. *Bull. Amer. Meteor. Soc.*, **84**, 481–492, https://doi.org/10.1175/BAMS-84-4-Ebert.

Fu, J. L., Z. P. Zong, K. Dai, F. H. Zhang, and D. B. Gao, 2014: Application of a verification method on bias analysis of quantitative precipitation forecasts (in Chinese). *Meteor. Monogr.*, **40**, 796–805.

Golding, B. W., 2000: Quantitative precipitation forecasting in the UK. *J. Hydrol.*, **239**, 286–305, https://doi.org/10.1016/S0022-1694(00)00354-1.

Haiden, T. M., M. J. Rodwell, and D. S. Richardson, 2012: Intercomparison of global model precipitation forecast skill in 2010/11 using the SEEPS score. *Mon. Wea. Rev.*, **140**, 2720–2733, https://doi.org/10.1175/MWR-D-11-00301.1.

Harris, D., E. Foufoulageorgiou, K. K. Droegemeier, and J. J. Levit, 2001: Multiscale statistical properties of a high-resolution precipitation forecast. *J. Hydrometeor.*, **2**, 406–418, https://doi.org/10.1175/1525-7541(2001)002<0406:MSPOAH>2.0.CO;2.

He, H., and F. Zhang, 2010: Diurnal variations of warm-season precipitation over Northern China. *Mon. Wea. Rev.*, **138**, 1017–1025, https://doi.org/10.1175/2010MWR3356.1.

Hong, J. S., 2003: Evaluation of the high-resolution model forecasts over the Taiwan area during GIMEX. *Wea. Forecasting*, **18**, 836–846, https://doi.org/10.1175/1520-0434(2003)018<0836:EOTHMF>2.0.CO;2.

Huang, L., and Y. Luo, 2017: Evaluation of quantitative precipitation forecasts by TIGGE ensembles for south China during the presummer rainy season. *J. Geophys. Res. Atmos.*, **122**, 8494–8516, https://doi.org/10.1002/2017JD026512.

Kniffka, A., and Coauthors, 2020: An evaluation of operational and research weather forecasts for southern West Africa using observations from the DACCIWA field campaign in June–July 2016. *Quart. J. Roy. Meteor. Soc.*, **146**, 1121–1148, https://doi.org/10.1002/qj.3729.

Kober, K., G. C. Craig, C. Keil, and A. Dornbrack, 2012: Blending a probabilistic nowcasting method with a high-resolution numerical weather prediction ensemble for convective precipitation forecasts. *Quart. J. Roy. Meteor. Soc.*, **138**, 755–768, https://doi.org/10.1002/qj.939.

Li, D. S., J. H. Sun, S. M. Fu, J. Wei, D. Wang, and F. Y. Tian, 2015: Spatiotemporal characteristics of hourly precipitation over central eastern China during the warm season of 1982–2012. *Int. J. Climatol.*, **36**, 3148–3160, https://doi.org/10.1002/joc.4543.

Liu, Y., J. Brown, J. Demargne, and D. Seo, 2011: A wavelet-based approach to assessing timing errors in hydrologic predictions. *J. Hydrol.*, **397**, 210–224, https://doi.org/10.1016/j.jhydrol.2010.11.040.

Lu, X. N., S. Z. Niu, C. F. Yuan, and X. C. Yuan, 2013: Verification and error analysis of quantitative precipitation estimation and forecast products in SWAN (in Chinese). *Torrential Rain Disasters*, **32**, 142–150.

Luo, Y., and Coauthors, 2017: The Southern China Monsoon Rainfall Experiment (SCMREX). *Bull. Amer. Meteor. Soc.*, **98**, 999–1013, https://doi.org/10.1175/BAMS-D-15-00235.1.

Mao, M., J. H. Dai, B. P. Li, and X. Zhang, 2016: Object-based verification and evaluation for different types of severe convection forecasting products (in Chinese). *Meteor. Monogr.*, **42**, 389–397.

Moore, B. J., K. M. Mahoney, E. M. Sukovich, R. Cifelli, and T. M. Hamill, 2015: Climatology and environmental characteristics of extreme precipitation events in the southeastern United States. *Mon. Wea. Rev.*, **143**, 718–741, https://doi.org/10.1175/MWR-D-14-00065.1.

Mullen, S. L., and R. Buizza, 2001: Quantitative precipitation forecasts over the United States by the ECMWF ensemble prediction system. *Mon. Wea. Rev.*, **129**, 638–663, https://doi.org/10.1175/1520-0493(2001)129<0638:QPFOTU>2.0.CO;2.

North, R., M. Trueman, M. Mittermaier, and M. J. Rodwell, 2013: An assessment of the SEEPS and SEDI metrics for the verification of 6 h forecast precipitation accumulations. *Meteor. Appl.*, **20**, 164–175, https://doi.org/10.1002/met.1405.

Pan, L. J., C. F. Xue, H. F. Zhang, J. P. Wang, and J. Yao, 2017a: Comparison of three verification methods for high-resolution grid precipitation forecast (in Chinese). *Climatic Environ. Res.*, **22**, 45–58.

——, ——, ——, X. T. Chen, L. W. Qu, and Y. Yuan, 2017b: Evaluation of precipitation probability forecasts of ECMWF ensemble prediction system in central China (in Chinese). *Plateau Meteor.*, **36**, 138–147.

Piriou, J., J. Redelsperger, J. Geleyn, J. Lafore, and F. Guichard, 2007: An approach for convective parameterization with memory: Separating microphysics and transport in grid-scale equations. *J. Atmos. Sci.*, **64**, 4127–4139, https://doi.org/10.1175/2007JAS2144.1.

Qian, W., and X. Lin, 2005: Regional trends in recent precipitation indices in China. *Meteor. Atmos. Phys.*, **90**, 193–207, https://doi.org/10.1007/s00703-004-0101-z.

Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, https://doi.org/10.1175/2007MWR2123.1.

Rodwell, M. J., D. S. Richardson, T. D. Hewson, and T. Haiden, 2010: A new equitable score suitable for verifying precipitation in numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **136**, 1344–1363, https://doi.org/10.1002/qj.656.

Rossa, A., P. Nurmi, and E. Ebert, 2008: Overview of methods for the verification of quantitative precipitation forecasts. *Precipitation: Advances in Measurement, Estimation and Prediction*, S. Michaelides, Ed., Springer, 419–452, https://doi.org/10.1007/978-3-540-77655-0_16.

Schaefer, J. T., 1990: The critical success index as an indicator of warning skill. *Wea. Forecasting*, **5**, 570–575, https://doi.org/10.1175/1520-0434(1990)005<0570:TCSIAA>2.0.CO;2.

Stephenson, D. B., 2000: Use of the "odds ratio" for diagnosing forecast skill. *Wea. Forecasting*, **15**, 221–232, https://doi.org/10.1175/1520-0434(2000)015<0221:UOTORF>2.0.CO;2.

Stoelinga, M. T., and Coauthors, 2003: Improvement of microphysical parameterization through observational verification experiment. *Bull. Amer. Meteor. Soc.*, **84**, 1807–1826, https://doi.org/10.1175/BAMS-84-12-1807.

Su, X., H. Yuan, Y. Zhu, Y. Luo, and Y. Wang, 2014: Evaluation of TIGGE ensemble predictions of Northern Hemisphere summer precipitation during 2008–2012. *J. Geophys. Res. Atmos.*, **119**, 7292–7310, https://doi.org/10.1002/2014JD021733.

Sun, J. H., and F. Q. Zhang, 2012: Impacts of mountain–plains solenoid on diurnal variations of rainfalls along the mei-yu front over the East China plains. *Mon. Wea. Rev.*, **140**, 379–397, https://doi.org/10.1175/MWR-D-11-00041.1.

Tao, S. Y., 1980: *Heavy Rainfalls in China* (in Chinese). Science Press, 225 pp.

——, and L. X. Chen, 1987: A review of recent research on the East Asian summer monsoon in China. *Monsoon Meteorology*, C. P. Chang and T. N. Krishnamurti, Ed., Oxford University Press, 60–92.

Tiziana, C., A. Ghelli, and F. Lalaurette, 2002: Verification of precipitation forecasts over the Alpine region using a high-density observing network. *Wea. Forecasting*, **17**, 238–249, https://doi.org/10.1175/1520-0434(2002)017<0238:VOPFOT>2.0.CO;2.

Wang, B., and H. Lin, 2002: Rainy season of the Asian–Pacific summer monsoon. *J. Climate*, **15**, 386–398, https://doi.org/10.1175/1520-0442(2002)015<0386:RSOTAP>2.0.CO;2.

Wang, N., and C. Lu, 2010: Two-dimensional continuous wavelet analysis and its application to meteorological data. *J. Atmos. Oceanic Technol.*, **27**, 652–666, https://doi.org/10.1175/2009JTECHA1338.1.

Wang, Y., and Z. H. Yan, 2007: Effect of different verification schemes on precipitation verification and assessment conclusion (in Chinese). *Meteor. Monogr.*, **33**, 53–61.

Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558, https://doi.org/10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2.

Xiong, Q. F., 2011: Verification of GRAPES_meso precipitation forecasts based on fine-mesh and station datasets (in Chinese). *Meteor. Monogr.*, **37**, 185–193.

Yano, J., P. Bechtold, and J. Redelsperger, 2003: "Renormalization" approach for subgrid-scale representations. *J. Atmos. Sci.*, **60**, 2029–2038, https://doi.org/10.1175/1520-0469(2003)060<2029:RAFSR>2.0.CO;2.

Yu, R., T. Zhou, A. Xiong, Y. Zhu, and J. Li, 2007: Diurnal variations of summer precipitation over contiguous China. *Geophys. Res. Lett.*, **34**, L01704, https://doi.org/10.1029/2006GL028129.

——, J. Li, W. Yuan, and H. Chen, 2010: Changes in characteristics of late-summer precipitation over eastern China in the past 40 years revealed by hourly precipitation data. *J. Climate*, **23**, 3390–3396, https://doi.org/10.1175/2010JCLI3454.1.

Zhang, Y. F., L. J. Pan, and X. Yang, 2014: Comparative analysis of precipitation forecasting capabilities of ECWMF and Japan high-resolution models (in Chinese). *Meteor. Monogr.*, **40**, 424–432.

——, ——, S. Lu, and X. X. Ju, 2017: Performance analysis on deterministic precipitation forecasting in surrounding areas of Qinling Mountains by ECMWF ensemble prediction system (in Chinese). *Climatic Environ. Res.*, **22**, 551–562.

Zheng, Y. G., M. Xue, B. Li, J. Chen, and Z. Y. Tao, 2016: Spatial characteristics of extreme rainfall over China with hourly through 24-hour accumulation periods based on national-level hourly rain gauge data. *Adv. Atmos. Sci.*, **33**, 1218–1232, https://doi.org/10.1007/s00376-016-6128-5.

Zong, Z. P., T. J. Chen, and Y. Guan, 2013: Analysis and forecast verification of two southwest vortex torrential rain events in Sichuan Basin in early autumn of 2012 (in Chinese). *Meteor. Monogr.*, **39**, 567–576.