Front. Earth Sci. 2019, 13(4): 758–777
https://doi.org/10.1007/s11707-019-0802-8

RESEARCH ARTICLE

BMA probability quantitative precipitation forecasting of land-falling typhoons in south-east China

Linna ZHAO[1], Xuemei BAI (✉)[2], Dan QI (✉)[3], Cheng XING[4]

1 Chinese Academy of Meteorological Sciences, Beijing 100081, China
2 Heilongjiang Meteorological Observatory, Harbin 150001, China
3 National Meteorological Centre, Beijing 100081, China
4 Heilongjiang Mulan County Meteorological Observatory, Mulan 151900, China

Abstract The probability of quantitative precipitation forecast (PQPF) of three Bayesian Model Averaging (BMA) models based on three raw super ensemble prediction schemes (i. e., A, B, and C) are established, which through calibration of their parameters using 1–3 day precipitation ensemble prediction systems (EPSs) from the China Meteorological Administration (CMA), the European Centre for Medium-Range Weather Forecasts (ECMWF) and the National Centers for Environmental Prediction (NCEP) and observation during land-falling of three typhoons in south-east China in 2013. The comparison of PQPF shows that the performance is better in the BMA than that in raw ensemble forecasts. On average, the mean absolute error (MAE) of 1 day lead time forecast is reduced by 12.4%, and its continuous ranked probability score (CRPS) of 1–3 day lead time forecast is reduced by 26.2%, respectively. Although the amount of precipitation prediction by the BMA tends to be underestimated, but in view of the perspective of probability prediction, the probability of covering the observed precipitation by the effective forecast ranges of the BMA are increased, which is of great significance for the early warning of torrential rain and secondary disasters induced by it.

Keywords Bayesian model averaging, probabilistic quantitative precipitation forecasting, ensemble prediction, typhoon precipitation

Received May 30, 2019; accepted October 9, 2019

E-mails: b_xmei@163.com (Xuemei BAI); qidan@cma.gov.cn (Dan QI)

1 Introduction

China is one of the most tropical cyclone prone countries in the world, with an average of 7 land-falling typhoons (including tropical storm, severe tropical storm, typhoon, severe typhoon and super typhoon, hereafter) in the mainland of China every year since 1949 (Hurricane Research Division, 2015). The coastal provinces from south to north are probably attacked and influenced by typhoons, and the impacts could even extend to the inland area in China. Typhoon is a kind of severe weather with intensely destructive power, which is generally accompanied by strong wind, heavy rainfall, storm surge and some other secondary disasters. There were one or two Tropical Cyclones (TCs) in the list of the worst 10 natural disasters in China every year from 2007 to 2012 (Chen et al., 2013). These often result in enormous casualty and property loss in coastal areas. In an average year, such TCs cause 44.2 billion yuan of normalized direct economic losses (Wang et al., 2016). In 2016, statistics show that these TCs have affected 15.616 million people in 14 provinces, autonomous regions or municipalities of the mainland of China. The TCs have resulted in 198 deaths or missing, 2.547 million people evacuated, 1.577 million hectares of crops affected with 183200 ha of croplands producing no yield at all, 181000 houses collapsed, and 169000 houses damaged to different degrees, with direct economic losses worth 72.24 billion yuan. Meanwhile, 0.149 million people have received emergency daily assistance due to the attack of typhoon (Gu et al., 2017). In the past 20 years, researchers and forecasters have made encouraging progress in the forecasting of TC track and intensity, the ability of typhoon operational numerical forecast model in China has been greatly improved (Chen et al., 2013; Gu et al., 2017), but the quantitative refined forecast of the wind and rainfall induced by typhoons still

needs to be urgently developed. Considering that the objective, refined and quantitative forecast plays a vital role in the effective prevention and defense of typhoons, the accurately quantitative forecast for typhoon precipitation is of great significance for the prevention and mitigation of typhoon disasters in China.

As the major meteorological operation centers in the world have successively established the ensemble prediction systems (EPSs) with continuously improved resolutions, the ensemble forecast based on global medium-term forecast system has been a new trend in the development of multi-model super ensemble forecast for TCs, especially since the beginning of the Observing System Research and Predictability Experiment (THORPEX) project (Nakazawa et al., 2010). Because all models are imperfect and no single model is expected to perform well under all circumstances, multimodel postprocessing methods have emerged as a way to achieve better forecast skills (Ajami et al., 2006). There is no doubt that, till now, the ensemble forecasting approach has great popularity to reduce the various uncertainties in numerical predictions (Buizza et al., 1999). The ensemble numerical precipitation forecast is not only be used routinely by experts and the public to prepare for weather conditions in coming days, but also be used in decision making of early warning for the precautionary measures against disasters such as floods and debris flows. However, it has relatively limited space to ameliorate the ensemble forecast by improving the physical processes in the numerical model or increasing the spatial resolutions. There is systematic biases and the under- or over-discrete forecasting from ensemble models in operation of weather forecasts, which could be attributed to shortcomings including the errors in the model initial conditions, method of data assimilation, the imperfections of model-self, insufficient model physical mechanism, limited ensemble members and the deficiency of ensemble methods (Hamill and Colucci, 1998; Li et al., 2015) and inadequate representation of the uncertainties of the atmosphere (Hamill et al., 2000 and 2003). Beside this, the outputs of ensemble numerical forecast do not generate a full predictive probability density function; rather, they give only probabilities for certain specific events. Therefore, the ensemble models used directly by the end users is far from satisfactory because of various biases in them and also be hardly make use of all the information available in an ensemble forecast. Besides, it cannot support decision making of other downstream applications such as early warning of floods and rainfall induced geological hazards (Krzysztofowicz et al., 1993; Krzysztofowicz, 1998; Zhao et al., 2011). Hence, in order to improve the reliability and forecasting skills of the ensemble forecast, it must be treated by statistical postprocessors before the application of outputs from ensemble forecast (Shrestha et al., 2015).

Precipitation forecasting is one of the most important meteorological elements for either weather forecast or early warnings for floods and debris flows. However, the statistical post-processing modeling of the precipitation forecast has evident difficulties and challenges compared to those of other elements, particularly for the precipitation with time scale shorter than one day, because: 1) the series of precipitation observation are discontinuous, those are a certain probability at the zero precipitation and while it shows a skewed distribution at nonzero precipitation. Therefore, the statistical post-processing model of precipitation forecast should be established by probability distributions of mixed discrete-continuous distributions. 2) The uncertainty of precipitation forecasts tends to be intensified with the increasing precipitation, which leads to heteroscedasticity in the statistical modeling. 3) Moreover, due to the relatively small number records of the extreme rainfall events, a collection of more training samples is required for the modeling (Scheuerer and Hamill, 2015).

In view of the great importance of the quantitative rainfall forecast, plenty of investigations have been carried out. The statistical models are usually employed for the statistical post-processing of numerical forecasts. So far, the traditional model output statistics (MOS) method has been widely used (Glahn and Lowry, 1972), which is suitable for the post-processing of deterministic models but not applicable for the ensemble forecasts (Cui et al., 2012). They also do not make use of all the information available in an ensemble forecast. In recent years, many post-processing methods have been developed for the ensemble models, such as the spatio-temporal neighborhood method (Theis et al., 2005), the bayesian model averaging (BMA) method (Raftery et al., 2005; Sloughter et al., 2007) based on the Bayesian statistical theory, the reliability ensemble mean method (Smith et al., 2009; Li et al., 2017), etc.

The BMA method provides the weighted average of each ensemble model forecast after the linear calibration. The member weights are defined as the Bayesian posterior probability of each model when it performances the best in the training period, BMA reflects the performance of each model prediction during the training period. The BMA model cannot only present the greatest forecasting possibility, but also realistically describe the uncertainty of the precipitation forecast. Since the BMA model was extended to the post-processing of the precipitation ensemble forecast, promising and delightful developments have been obtained (Sloughter et al., 2007). Yang et al. (2012) applied the BMA method to generate the daily probabilistic precipitation forecast based on the ensemble prediction system in the Huaihe Basin for a period of 3 months, further confirming that the precipitation revised by the BMA method can capture the information of extreme precipitation. Liang et al. (2013) employed the BMA model to conduct the probability integration and bias correction for the ensemble forecast of the T213 model by using the daily precipitation observation from automatic stations in the Huaihe Basin. By using the THORPEX

Interactive Grand Global Ensemble (TIGGE) prediction products, Liu and Xie (2014) evaluated the BMA model performance based on the probabilistic forecast of the daily precipitation in the Huaihe Basin. The results suggested that the BMA model for exchangeable members outperformed those for single-center EPSs for all 1–7-day ensemble forecasts, and mostly improved the quality of probabilistic quantitative precipitation forecasting (PQPF). To study the impact of training sample heterogeneity on the performance of BMA, Zhu et al. (2015) used stratified sampling BMA to the Advanced Research version of the Weather Research and Forecasting (ARW-WRF) Model in the northern China region, and found that stratified sampling BMA outperformed the raw ensemble and conventional BMA. The above results reveal that the BMA model has a good correction performance on the statistical post-processing of non-typhoon precipitation. For the ensemble forecast of typhoon rainstorms, Chen et al. (2016) improved the ensemble prediction skill of the European Centre for Medium-Range Weather Forecasts (ECMWF) by the optimal selection of ensemble members, indicating the some advantage compared with the deterministic forecast of ECMWF.

As discussed above, the BMA method is mainly applied in the post-process of ensemble forecasts for non-typhoon precipitation. So far, there are relatively small numbers of studies focusing on the statistical post-processing modeling in the typhoon probability quantitative precipitation ensemble forecasts due to the relevant difficulties. Therefore, in this paper we extend the study of BMA by instilling realistic spatial characteristics in calibrated PQPFs of typhoon precipitation forecast, the specific objectives are for the three typhoon processes from July to September of 2013, three ensemble schemes are constructed for the precipitation forecast by the multicenter ensemble prediction systems from the China Meteorological Administration (CMA), the European Centre for Medium-Range Weather Forecasts (ECMWF) and the National Centers for Environmental Prediction (NCEP) and observation during the typhoon landing period in China. Besides, the BMA model for the probabilistic forecast is established for the typhoon precipitation based on the three ensemble schemes, with the parameters calibrated during the training period. The associated results are then evaluated to investigate the advantages of the BMA model based on different ensemble prediction systems in improving the prediction skills for the typhoon PQPF. Moreover, the adaptability of each ensemble prediction system and the super ensemble forecast products for the study area are discussed.

The rest of the paper is organized as follows. Section 2 briefly introduces the theoretical framework of the BMA method. Data and study area are described in Section 3. In Section 4, the experiment results are analyzed. Finally, the summary and discussion are expressed in Section 5.

2 Methodology

2.1 The BMA and the parameter estimation

According to the BMA method, the single model is not considered as the best, while the entire statistical model ensemble should be first considered. It provides the forecasting probability density function (PDF) obtained from the weighted average of ensemble members. The used weights are consistent with the posterior probabilities of the model forecasts, which indicate the contribution of each ensemble member to the final prediction skill over the whole training period. Ultimately, the BMA model intends to generate a well-defined and highly concentrated PDF, which can efficiently predict the percentiles of weather elements. However, the statistical model is mainly applied to weather elements that approximately follow the normal distribution, such as temperature, sea level pressure, etc., (Raftery et al., 2005; Wilson et al., 2007) but it is not exactly suitable for the precipitation forecast, which is discontinuous and non-normally distributed. To solve this problem, Sloughter et al. (2007) applied the BMA model to the statistical post-processing of precipitation forecast by establishing a prediction PDF mixed by zero value and gamma distribution for the given ensemble members yielding calibrated and sharp predictive PDFs of future weather quantities. Fraley et al. (2010) extended the BMA model for situations with exchangeable ensemble members or missing members.

The BMA predictive PDF is a mixture of the component PDFs: Assuming that y is the "realistic" value of a specific meteorological element which needs to be predicted, and f_k is the ensemble forecast of member k, $h_k(y|f_k)$ represents the PDF of y when the initial forecast result of f_k is the best in the ensemble. The predicted f_k of each member corresponds to a specific $h_k(y|f_k)$. Then the PDF of the BMA model is calculated by the weighted average of the PDFs of multiple members,

$$P(y|f_1,..., f_k) = \sum_{k=1}^{K} w_k h_k(y|f_k), \quad (1)$$

where w_k is the posterior probability when f_k is the best forecast. The w_k represents the weight of each member, which can be calculated from the historical ensemble predictions and observations. It also reflects the relative contribution of member k to the prediction during the training period, and follows the equation $\sum_{k=1}^{K} w_k = 1$.

It is already known that the precipitation is a non-continuous variable, which contains plenty of zero values and highly skewed-distributed nonzero values. Therefore, the logistic regression model and the gamma distribution model is applied in the precipitation distribution, which describes the precipitation probability and the nonzero precipitation amount, respectively.

The first part is the logistic regression model as follows,

$$\text{logit}(P(y=0|f_k)) = \log\frac{P(y=0|f_k)}{P(y>0|f_k)}$$

$$= a_{0k} + a_{1k}f_k^{1/3} + a_{2k}\delta_k, \quad (2)$$

where δ_k is the indicator of f_k, if f_k equals 0, $\delta_k = 1$, otherwise $\delta_k = 0$. If the precipitation is nonzero, then use gamma distribution to fit the nonzero precipitation amount. The parameters a_{0k}, a_{1k} and a_{2k} are member specific. They are determined separately for each ensemble member, using the forecasts from that ensemble member only and the associated verifying observations. They are estimated by logistic regression with precipitation/no precipitation as the dependent variable, and $f_k^{1/3}$ and δ_k as the two predictor variables. The probability $P(y>0|f_k)$ is the precipitation probability when the forecast member f_k is optimal (Sloughter et al., 2007). Hamill et al. (2004) and Sloughter et al. (2007) have emphasized that the logistic regression model is more effective when the predicted variable is taken as the cubic root of f_k. Therefore, the form $f_k^{1/3}$ is utilized in the Eq. (2).

The second part is the conditional PDF of the nonzero precipitation amount. Use gamma distribution to fit the nonzero precipitation amount, where y is the cube root of the

$$g_k(y|f_k) = \frac{1}{\beta_k^{\alpha_k}\Gamma(a_k)}y^{\alpha_k-1}\exp(-y/\beta_k), \quad (3)$$

precipitation amount (Sloughter et al., 2007), where α_k and β_k represent the shape parameter and scale parameter of the Gamma distribution. The mean and variance of the gamma distribution in Eq. (3) are as follows.

$$\mu_k = \alpha_k\beta_k = b_{0k} + b_{1k}f_k^{1/3}; \quad (4)$$

$$\sigma_k^2 = \alpha_k\beta_k^2 = c_{0k} + c_{1k}f_k, \quad (5)$$

where b_{0k}, b_{1k}, c_{0k} and c_{1k} are the parameters to be estimated. As the two variance parameters between the ensemble members represented by c_{0k} and c_{1k} do not differ a lot with each other, the two parameters are considered consistent, that is, $\sigma_k^2 = c_0 + c_1f_k$. Hence, in the case that the member f_k is the optimal, the conditional PDF of precipitation can be expressed as follows.

$$h_k(y|f_k) = P(y=0|f_k)I[y=0]$$

$$+P(y>0|f_k)g_k(y|f_k)\text{I}[y>0], \quad (6)$$

where the $I[\ ]$ denotes the indicator function. When y equals 0, $I[y=0]$ is 1 and $I[y>0]$ is 0. Conversely, when y is greater than 0, $I[y=0]$ is 0 and $I[y>0]$ is 1. Overall, the obtained PDF of the precipitation ensemble forecast is shown.

$$P(y|f_1,...,f_k) = \sum_{k=1}^{\text{K}} w_k[P(y=0|f_k)I[y=0]$$

$$+P(y>0|f_k)g_k(y|f_k)\text{I}[y>0]], \quad (7)$$

In the Eq. (7), the relation between $P(y=0|f_k)$ and $P(y>0|f_k)$ is given in the above Eq. (2).

The parameters of a_{0k}, a_{1k} and a_{2k} are associated with the member k in the ensemble system, which can be estimated from the logistic regression model composed in Eq. (2). The parameters of b_{0k} and b_{1k} can be measured by the generalized linear model in Eq. (4). The parameters of c_0 and c_1 as well as the weight parameter of w_k ($k = 1, 2,..., \text{K}$) can be obtained from the training period by the maximum likelihood method (Sloughter et al., 2007). Sloughter et al. (2007) revealed that the longer the training period is, the better the forecast of the BMA model will make. The PQPF obtains its nearly optimal outcome from the BMA model when the training period is set as 30 days. However, even if the training period keeps increasing after that, the forecast results cannot be significantly improved any more. Accordingly, the training period is determined as 30 days in the study.

2.2 Verification methods

2.2.1 Mean Absolute Error

The Mean Absolute Error (*MAE*) is an indicator directly reflecting the forecast biases, defined as follows (Wilks, 2006),

$$MAE = \frac{1}{n}\sum_{k=1}^{n}|\hat{x}_k - x_k|, \quad (8)$$

where $\hat{x}_k$ and x_k are the values of the forecast and observation, respectively. Only if all the predictions are the same as the observations, *MAE* is equal to 0, which is the "perfect" forecast. Smaller *MAE* values indicate more skillful forecasts. In the current study, the *MAE* for the raw ensemble prediction system is calculated by absolute differences between the initial ensemble mean and the observation. While for the forecast results calibrated by the BMA model, the *MAE* is obtained from absolute differences between the ensemble median of BMA output and the observation.

2.2.2 Brier Score

The Brier Score (*BS*) is widely used to examine the accuracy of the ensemble forecast at different thresholds (in a specific weather event, e.g. precipitation). With respect to a specific dichotomous event, the *BS* is defined as follows,

$$BS = \frac{1}{N}\sum_{n=1}^{N}(P_n - O_n)^2, \quad (9)$$

where N is the total sample numbers accumulated in the space during the whole inspection period. P_n is the forecast probability of the verified event for sample n, ranging from 0 to 1. Correspondingly, O_n is the observation frequency of the event. When the verified event is consistent with the observation, O_n is taken as 1, otherwise it equals 0. The *BS* is negatively oriented, and it represents that the ensemble forecast is absolutely accurate when it equals 0, whereas larger *BS* values denote lower accuracy of the ensemble system.

2.2.3 Continuous Ranked Probability Score

The Continuous Ranked Probability Score (*CRPS*) can measure the overall probabilistic performance of the ensemble prediction system, which is described by the accumulation of squared errors between the cumulative distribution function of observation and prediction. It can be calculated via the formula as follows (Wilks, 2006).

$$CRPS = \int_{-\infty}^{\infty} (p_i(x) - o_i(x))^2 \mathrm{d}x, \tag{10}$$

where $p_i(x)$ and $o_i(x)$ denote cumulative distributions of the ensemble probabilistic forecast and the observation, respectively. The *CRPS* equaling 0 indicates the "perfect" forecast, while larger values of *CRPS* imply the relatively lower capability of the ensemble prediction system.

3 Data and study area

3.1 Typhoon cases

Previous studies (Ren et al., 2011) have reported that typhoons are fairly active over the western North Pacific from July to September every year, with the most intense activities in August. The 2013 western North Pacific tropical cyclone (TC) season was marked by above-average TC activity, the 2013 season featured greater cyclone activity than the previous seasons since 1995 (Ying et al., 2014a). Accordingly, the period from July to September in 2013 is taken as the study period in the paper. There were three landed typhoons from July 1 to September 30, 2013, illustrated as follows. 1) Typhoon "Utor" landed in Yangjiang, and made influence from August 10 to 18, 2013 (Cheng et al., 2014); 2) Typhoon "Trami" landed in Fuqing, and made influence from August 18 to 25, 2013 (Guo et al., 2014); 3) Typhoon "Usagi" landed in Shanwei, and made influence from September 18 to 25, 2013 (Li et al., 2013). The distribution of their tracks is displayed in Fig. 1, with the pentagrams representing the selected typical meteorological stations of "Yangjiang" and "Xinyi."

3.2 Data and research area

The observation data are the hourly precipitation of 818 national surface meteorological stations from July 1 to September 30, 2013, provided by the National Meteorological Information Center of China. In this study, the TCs best track data of TCs are applied as the TCs track observations, which are obtained from the Shanghai Typhoon Institute of the China Meteorological Administration (CMA-STI) data set in the China Typhoon Online, with a time interval of 6 h (Ying and Zhang W et al., 2014b). The research area and the definition of precipitation of landfall typhoons flowing as Zhao et al. (2017). The study area of (15–34°N, 105–130°E) is displayed as dashed line box in Fig. 1.

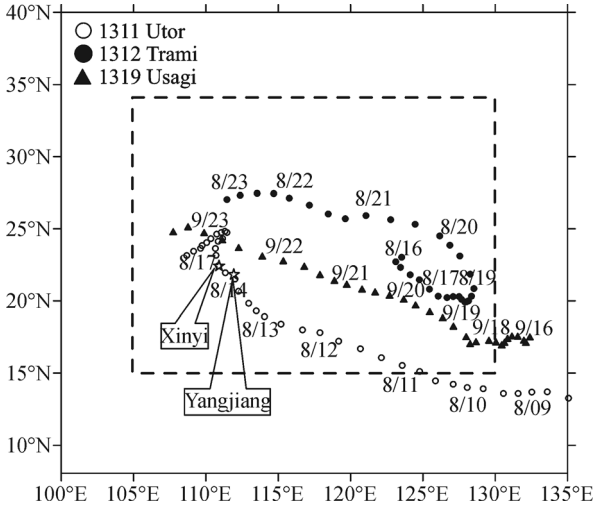

Fig. 1 Tracks of typhoon (Pentagrams represent the meteorological stations) and study domain (dash line box represents study domain).

The raw ensemble forecasts are the outputs of precipitation forecasts in the same period with lead time of 24-h, 48-h and 72-h, which are obtained from the ensemble prediction systems of CMA, NCEP and ECMWF. The initial forecast time is 00:00 UTC every day.

Table 1 Comparison of single-center ensemble prediction systems used in this study

Center	Country	Model spectral resolution	Ensemble Members (perturbed)	Spatial resolution	Forecast length (in days)
CMA	China	T213L31	14	0.5625° × 0.5625°	15
ECMWF	Europe	TL639L91/TL319L91	50	0.28125° × 0.28125°	10
NCEP	America	T126	20	1° × 1°	10

Detailed information regarding of the three ensemble prediction systems is provided in Table 1. The outputs of the ensemble model precipitation forecast are interpolated into stations in the study area by the bilinear interpolation method, which is operationally used (ECMWF, 2001; Bankman, 2009).

3.3 Schemes of multicenter ensemble forecast design

The ensemble forecast outputs from the CMA, NCEP and ECMWF are employed in the ensemble forecast of the study, which contains 87 forecast members in total. Three construction schemes are designed for the multicenter (model) ensemble forecasts. 1) The ensemble forecast consisting of the 87 members from CMA, NCEP and ECMWF, is hereafter noted as Scheme A. 2) After the comparison analysis by using the mean absolute errors, 4 members selected from each prediction center compose a multicenter ensemble forecast including 12 members, which is hereafter recorded as Scheme B. 3) The super ensemble forecast system framed by ensemble means of each prediction center is noted as Scheme C.

More details of the construction can be found in Table 2. Afterwards, BMA models are established based on the obtained multi-model ensemble forecasts for the investigating the precipitation forecast capability.

4 Results and analysis

4.1 24 h lead time skill of the PQPF by BMA

4.1.1 Verification of MAE for precipitation forecasts

To examine the difference between general precipitation errors of multicenter ensemble schemes and that of BMA, the MAE comparison between ensemble predictions based on the three multicenter ensemble schemes and the predictions after BMA calibration for the forecast lead-time of 24-h are displayed in Table 3, Table 4, and Table 5, during the influence periods of "Utor," "Trami" and "Usagi," respectively. The results show that before the BMA calibration of precipitation forecast for Typhoon "Utor," Scheme A and C are characterized by the smallest and largest MAEs, respectively. After the deviation correction by the BMA model, pronounced improvements can be revealed for the three ensemble schemes and their MAE values become fairly close. During the period of 15–18 August 2013, the BMA model shows limited capability of optimization for Scheme A and B, with their MAEs higher than that of the BMA-calibrated Scheme C. For the 8 days from August 11 to 18, the BMA model obviously improves the ensemble forecasts of Scheme C for the whole period except August 15, whereas for both schemes A and B it shows no improvement in 3 days. Overall, the BMA model has different calibration effects on ensemble forecasts based on the three schemes, with the best performance on Scheme C.

During the 7 days calibration of precipitation forecast influenced by Typhoon "Trami" (Table 4), the BMA model shows no evident improvement for 2 days in Scheme A and B, while 1 day in Scheme C. According to the calculated MAE scores, the forecast results from Scheme C calibrated by the BMA model are superior to those from the calibrated or non-calibrated Scheme A and B, except for August 23 and 24. Additionally, during the 7 days influenced by Typhoon "Usagi" (Table 5), the forecast capability is not enhanced by the BMA calibration in 4, 3 and 2 days for Scheme A, B and C, respectively. After revising by the BMA model, the MAE of Scheme C is the lowest.

In general, the BMA model is characterized by higher prediction skill than the raw ensemble forecast, with the averaged MAE reduced by 12.4%. Besides, the BMA calibration is most effective for Scheme C, while weakest for Scheme A. Much lower MAE values (i.e., more significant improvement induced by the BMA method) feature the ensemble precipitation forecast before the typhoon landfall than those after landfall. The precipitation in the study area due to the typhoon landfall, as well as the heavy rainfall in many stations, results in the relatively large biases of the precipitation forecast after the landfall for the ensemble schemes of A, B and C. To a great extent, the calibration effect of BMA model depends on prediction results from the three ensemble schemes.

4.1.2 Skill of PQPF by BMA

To further illustrate the difference forecast skill in four precipitation thresholds between the PQPFs by multicenter ensemble schemes and by BMA model, the BS of the BMA PQPF based on the three-scheme ensemble forecasts during the influence period of Typhoon "Utor," "Trami" and "Usagi" are presented in Figs. 2–4, respectively, with the precipitation thresholds of 0.1 mm, 10.0 mm, 25.0 mm and 50.0 mm.

During the 8 days influenced by Typhoon "Utor," the three scheme BSs for the precipitation probability forecast

Table 2 Configuration schemes of the multi-model ensemble forecasts in this study

Scheme	Ensemble Approach	Members	Lead time /hours
A	CMA15 members + NCEP21 members + ECMWF51 members	87	24
B	The best four members selected from EPS of CMA, NCEP, ECMWF, respectively	12	24
C	The ensemble mean of CMA, NCEP, ECMWF	3	24, 48, and 72

Table 3 MAE scores of the 24 h probabilistic precipitation forecasts using three ensemble schemes and the BMA during the Typhoon "Utor" (unit: mm)

Date	Scheme A		Scheme B		Scheme C	
	Raw	BMA	Raw	BMA	Raw	BMA
11 Aug	2.65	1.95	2.81	1.95	2.87	1.95
12 Aug	3.47	3.12	2.5	1.75	2.6	1.76
13 Aug	2.27	1.78	2.11	1.07	2.19	1.05
14 Aug	1.87	1.02	3.54	3.12	3.79	2.96
15 Aug	6.9	7.99	6.98	7.98	7.32	7.5
16 Aug	7.02	7.55	6.9	7.29	7.57	6.97
17 Aug	7.68	8.76	8.43	8.56	9.67	8.13
18 Aug	8.71	8.12	8.44	8.12	9	7.83

Table 4 The same as Table 3, but for Typhoon "Trami"

Date	Scheme A		Scheme B		Scheme C	
	Raw	BMA	Raw	BMA	Raw	BMA
19 Aug	7.34	6.8	7.66	6.81	7.93	6.58
20 Aug	6.22	5.23	6.68	5.29	6.84	5.1
21 Aug	5.18	4.11	5.67	4.17	5.82	4.11
22 Aug	5.75	5.49	6.49	5.41	6.56	5.24
23 Aug	9.63	10.74	9.76	10.92	10.2	10.09
24 Aug	10.49	11.89	10.68	12.01	10.99	11.12
25 Aug	9.04	8.79	9.36	8.77	9.47	8.42

Table 5 The same as Table 3, but for Typhoon "Usagi"

Date	Scheme A		Scheme B		Scheme C	
	Raw	BMA	Raw	BMA	Raw	BMA
19 Sept	1.99	2.1	2.41	2.15	3.26	1.98
20 Sept	1.13	0.93	1.34	0.95	1.51	0.92
21 Sept	0.66	0.35	0.74	0.37	0.91	0.35
22 Sept	1.35	0.96	1.39	0.99	1.52	0.96
23 Sept	5.38	6.56	5.57	6.67	5.6	5.97
24 Sept	9.05	10.63	9.04	10.68	9.43	9.56
25 Sept	9.24	10.30	10.12	10.47	10.09	10.00

of the 0.1 mm threshold are reduced by more than 0.3 after calibration by the BMA model (Fig. 2(a)), indicating the effective correction of the BMA model. As for the precipitation threshold of 10.0 mm (Fig. 2(b)), the BMA method is effective on the calibration of all schemes for almost the whole period, except for Scheme A on August 12 and Scheme B on August 14. Importantly, it is emphasized that before the BMA calibration, Scheme C is relatively worse compared with the other two schemes at the threshold of 10.0 mm, whereas it is greatly improved and shows better performance than Scheme A and B for 5 days after the calibration. For the probabilistic precipitation forecast at the 25.0 mm threshold (Fig. 2(c)), the BMA method makes positive calibrations for 4 days in both Scheme A and B, while for 6 days in Scheme C. With respect to the threshold of 50.0 mm (Fig. 2(d)), the BMA model is effective in the bias correction for only 2 days in Scheme A and B, but for 4 days in Scheme C with 3 days of remarkable improvement. It is indicated that the improvement of probabilistic precipitation forecasts induced by the BMA model is more remarkable in Scheme C than in Scheme A and B at the thresholds of 25.0 mm

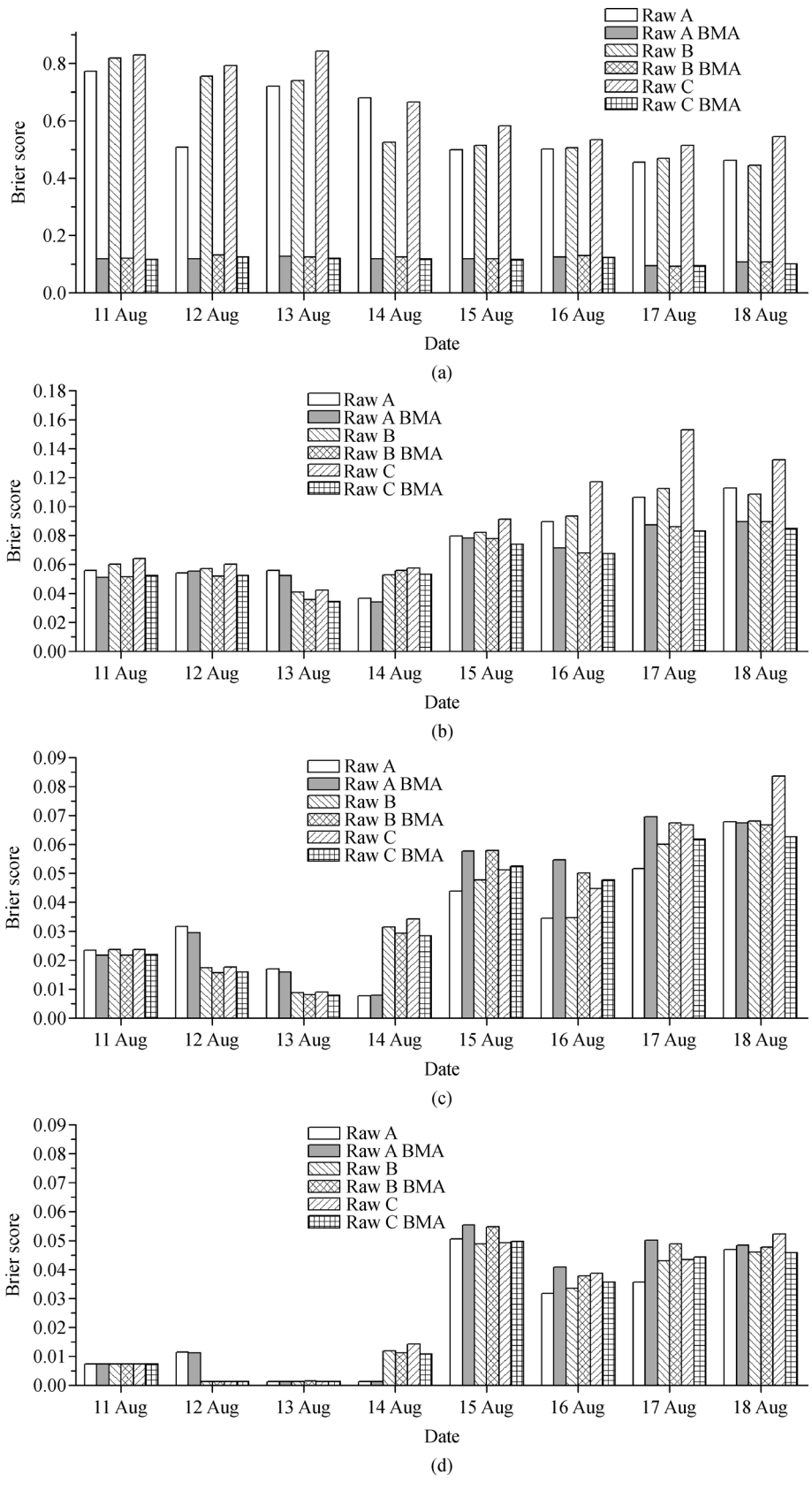

Fig. 2 Brier scores of ensemble forecasts based on different schemes and the BMA during Typhoon "Utor" (a: precipitation threshold of 0.1 mm; b: precipitation threshold of 10.0 mm; c: precipitation threshold of 25.0 mm; d: precipitation threshold of 50.0 mm).

and 50.0 mm.

During the "Trami" influence period of 7 days, the BMA model is considerably effective in the bias correction of the ensemble prediction systems based on all the three schemes at the precipitation threshold of 0.1 mm (Fig. 3(a)). For the probabilistic precipitation forecast at the 10.0 mm threshold (Fig. 3(b)), the BMA model has manifest calibration on the ensemble schemes for almost the entire period except for Scheme A and B on August 24. The BMA model makes positive impacts on probabilistic forecasts from ensemble Schemes A, B and C for 1 day, 3 days and 6 days, respectively, at the 25.0 mm threshold (Fig. 3(c)). As for the 50.0 mm threshold (Fig. 3(d)), the well-calibrated results occur in 2 days for Scheme A, and 3 days for Scheme B and C after the BMA processing.

Moreover, it can be revealed that the BMA model makes similarly pronounced calibration for the probabilistic precipitation forecast at thresholds of 0.1 mm and 10.0 mm during the period influenced by Typhoon "Usagi" (Figs. 4(a) and 4(b)), similar as those of "Utor" and "Trami." For the 25.0 mm precipitation threshold (Fig. 4(c)), the BMA model effectively improves the ensemble forecast for 2 days, 3 days and 4 days, respectively for the three schemes. On the other hand, the calibration for the probabilistic precipitation forecast at the 50.0 mm threshold (Fig. 4(d)) works in 2 days, 3 days and 4 days for Scheme A, B and C, respectively, which are more than the improved days during the other two typhoons.

Based on the above illustrations of the BMA model calibrations on the probabilistic precipitation forecast during the periods influenced by three typhoons, the results can be concluded as follows. For the 0.1 mm threshold, the BMA model shows relatively consistent improvements on the three ensemble forecast schemes. However, it differs a lot for the calibration effects at the higher precipitation thresholds. For the 10.0 mm threshold, the BMA model is characterized by similar effects between Scheme A and B and the largest improvement on Scheme C. For the precipitation thresholds of 25.0 mm and 50 mm, it is also the best for the forecast correction of Scheme C, followed by Scheme B. According to investigations on the forecast skills for different precipitation thresholds in a specific ensemble scheme, the best calibration occurs in the probabilistic precipitation forecast at the 10.0 mm threshold for all the three schemes. In Scheme A, it is followed by the precipitation threshold of 50.0 mm and then 25.0 mm, while they are reversed in both Scheme B and C.

4.2 Longer lead time skill of the PQPF by BMA

The BMA model calibrations on the deviation of the probabilistic precipitation forecast and the ensemble mean forecast from the three ensemble schemes for the typhoon precipitation with the lead time of 24-h have already been detailed analyzed in Section 4.1, which indicates the significant improvement made by the BMA model on the ensemble precipitation forecast during the typhoon influence period. On the other hand, the calibration effects of the BMA method on precipitation forecast for different lead time in the same ensemble scheme still need to be further investigated. In the current section, Scheme C is employed for the following analysis on forecast calibrations for different lead time. The CRPS of probabilistic precipitation forecasts in different lead time are displayed in Table 6, Table 7 and Table 8 for ensemble forecasts before and after the BMA calibration on Scheme C during the three-typhoon influence periods. It is detected that the BMA model forecast is superior to the raw ensemble for the whole period, with the CRPS reduced by 26.2% on average. After the BMA calibration, the CRPS scores are lower than the raw ensemble for lead time of 24-h, 48-h and 72-h in Scheme C, that is, the BMA model can significantly improve the probabilistic precipitation forecast of Scheme C for the three lead time. Largest improvement features the 24-h precipitation forecast by the BMA model, followed by the lead time of 48-h. It is also indicated that the calibration capability of the BMA model is closely related to forecasts from the raw ensemble.

4.3 Typhoon cases study of the PQPF by BMA

Based on the analysis in Sections 4.1 and 4.2, the typhoon "Utor" is chosen for the case study to further investigate the influence of the BMA model calibration on the 24-h precipitation forecast from the ensemble system based on Scheme C. Since the BMA outputs are represented by probabilistic forecasts at a specific percentile, the 50th percentile precipitation forecast is employed as the "ensemble mean" of BMA calibration. Afterwards, the BMA calibrated results are compared with the corresponding observations as follows.

4.3.1 Spatial distribution of PQPF

The distributions of daily precipitation during the "Utor" landfall period of 11–18 August 2013 are shown in Fig. 5. It is shown that the heavy rainfall induced by the typhoon starts on August 14, and is detected in the south-western part of Fujian, most of Guangdong, south-eastern Guangxi and northern Hainan. According to the comparison between the 50th percentile precipitation forecast of Scheme C before (Fig. 6) and after (Fig. 7) the BMA calibration as well as the observations (Fig. 5), the BMA model has a significant deviation correction on false predictions in the precipitation forecast from Scheme C. And it's indicated that the results of the 50th percentile precipitation forecast are obviously much closer to the observations after the BMA calibration, making the main

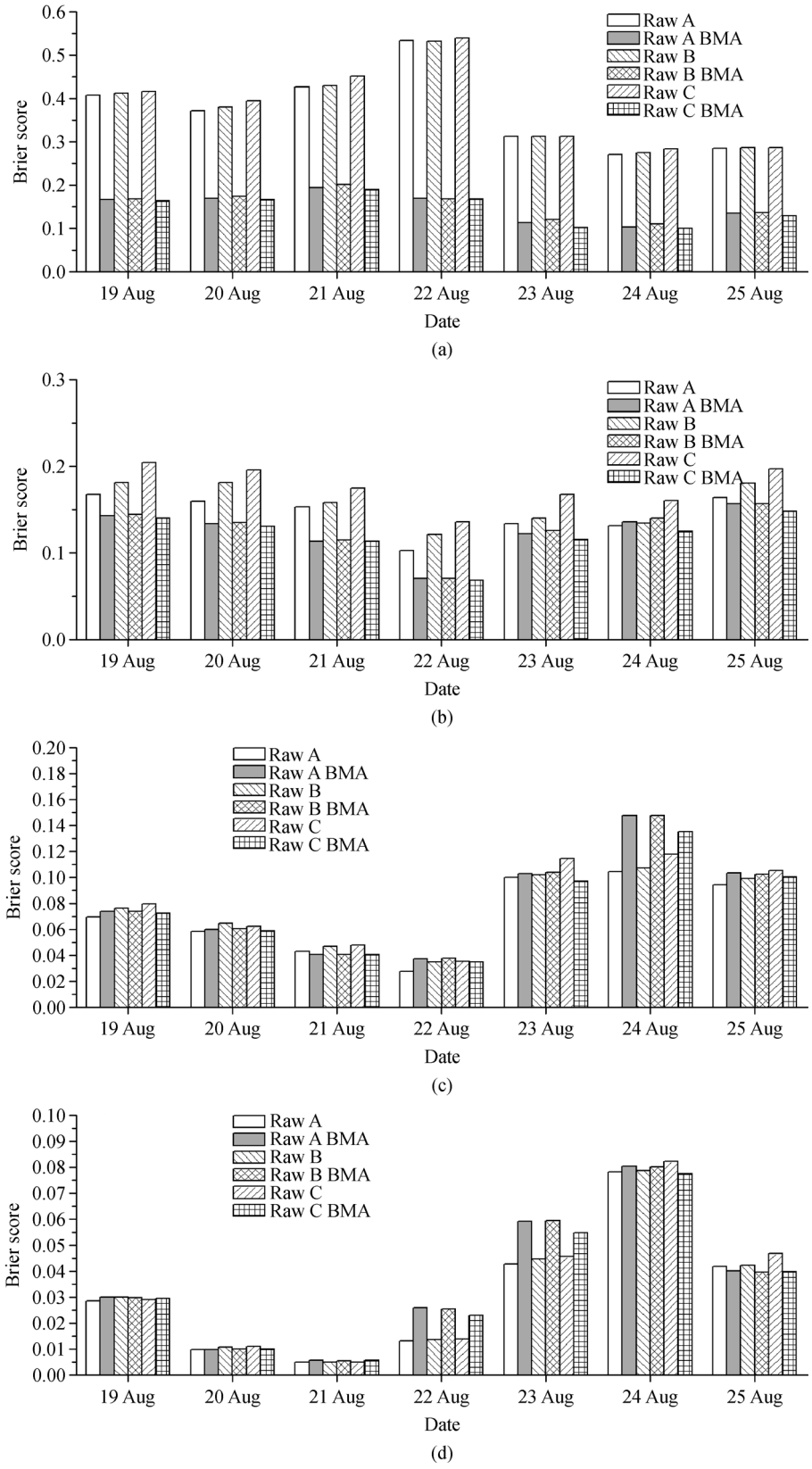

Fig. 3 The same as Fig. 2, but for Typhoon "Trami" (a: precipitation threshold of 0.1 mm; b: precipitation threshold of 10.0 mm; c: precipitation threshold of 25.0 mm; d: precipitation threshold of 50.0 mm).

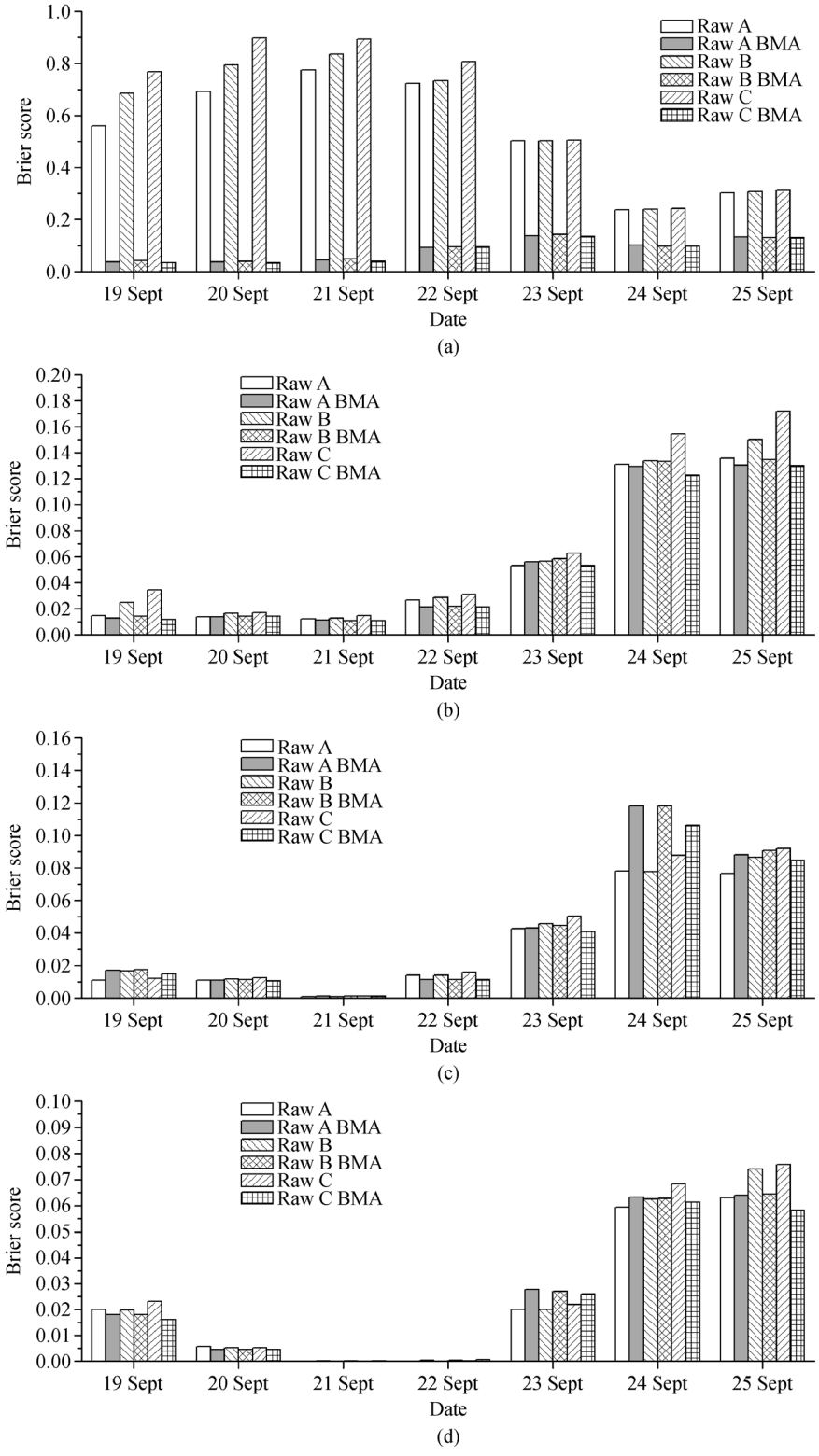

Fig. 4 The same as Fig. 2, but for Typhoon "Usagi" (a: precipitation threshold of 0.1 mm; b: precipitation threshold of 10.0 mm; c: precipitation threshold of 25.0 mm; d: precipitation threshold of 50.0 mm).

Table 6 CRPS verifications for BMA models and raw ensemble for Scheme C with three days lead times for precipitation during the Typhoon "Utor" (unit: mm)

Date	24-h lead time forecasts		48-h lead time forecasts		72-h lead time forecasts	
	Scheme C	BMA base on scheme C	Scheme C	BMA base on scheme C	Scheme C	BMA base on scheme C
11 Aug	2.63	1.68	3.5	2.32	5.4	3.38
12 Aug	2.21	1.46	4.05	2.84	4.76	3.38
13 Aug	1.93	1.02	3.42	2.19	5.14	3.51
14 Aug	3.19	2.32	4.06	3.19	5.04	4.23
15 Aug	6.09	5.3	7.77	6.37	9.25	8.03
16 Aug	6.43	5.01	11.9	9.77	14.28	11.7
17 Aug	7.75	5.87	12.32	10.05	18.58	15.64
18 Aug	7.72	6.01	13.05	10.74	18.05	14.83

Table 7 The same as Table 6, but for Typhoon "Trami"

Date	24-h lead time forecasts		48-h lead time forecasts		72-h lead time forecasts	
	Scheme C	BMA base on scheme C	Scheme C	BMA base on scheme C	Scheme C	BMA base on scheme C
19 Aug	6.81	5.15	12.69	10.22	17.28	14.42
20 Aug	5.82	3.89	10.44	7.81	15.85	12.51
21 Aug	4.91	3.12	8.52	5.7	12.83	9.23
22 Aug	5.62	3.75	8.58	6.02	12.75	8.92
23 Aug	8.68	7.22	12.58	10.14	14.94	12.05
24 Aug	9.22	7.96	15.77	13.43	18.39	16.06
25 Aug	8.14	6.27	14.39	11.89	20.11	17.61

Table 8 The same as Table 6, but for Typhoon "Usagi"

Date	24-h lead time forecasts		48-h lead time forecasts		72-h lead time forecasts	
	Scheme C	BMA base on scheme C	Scheme C	BMA base on scheme C	Scheme C	BMA base on scheme C
19 Sept	2.57	1.39	4.07	2.9	5.55	3.83
20 Sept	1.34	0.72	3.44	2.16	4.81	3.63
21 Sept	0.75	0.29	1.85	0.91	3.7	2.33
22 Sept	1.34	0.77	2.4	1.14	3.16	1.75
23 Sept	4.86	4.33	6.5	4.97	7.2	4.95
24 Sept	7.89	7.26	12.88	11.35	15.79	12.94
25 Sept	8.95	6.91	15.03	13.7	20.26	18.25

precipitation areas more prominent.

It is suggested from Figs. 6 and 7 that the BMA model effectively calibrates the ensemble precipitation forecast of Scheme C for the thresholds from 0.1 mm to 10.0 mm. When the precipitation is larger than 50.0 mm, even exceeding 100.0 mm, the 50th percentile precipitation forecast mainly locates in the range from 25.0 mm to 50.0 mm after the BMA calibration. It is demonstrated that the BMA model significantly improves the ensemble forecast for the precipitation thresholds of 0.1 mm and 10.0 mm, which corresponds well with the analysis on the BS examinations. That is, plenty of false predictions can be eliminated by the BMA model. However, simultaneously, the magnitudes of the heavy rainfall intend to be underestimated.

4.3.2 Temporal evolution of PQPF

Aiming at further analyzing the calibration of the BMA model on the heavy rainfall forecast, two stations on the track of the "Utor" after landfall are selected as the specific representations, which are the near landfall point "Yang-

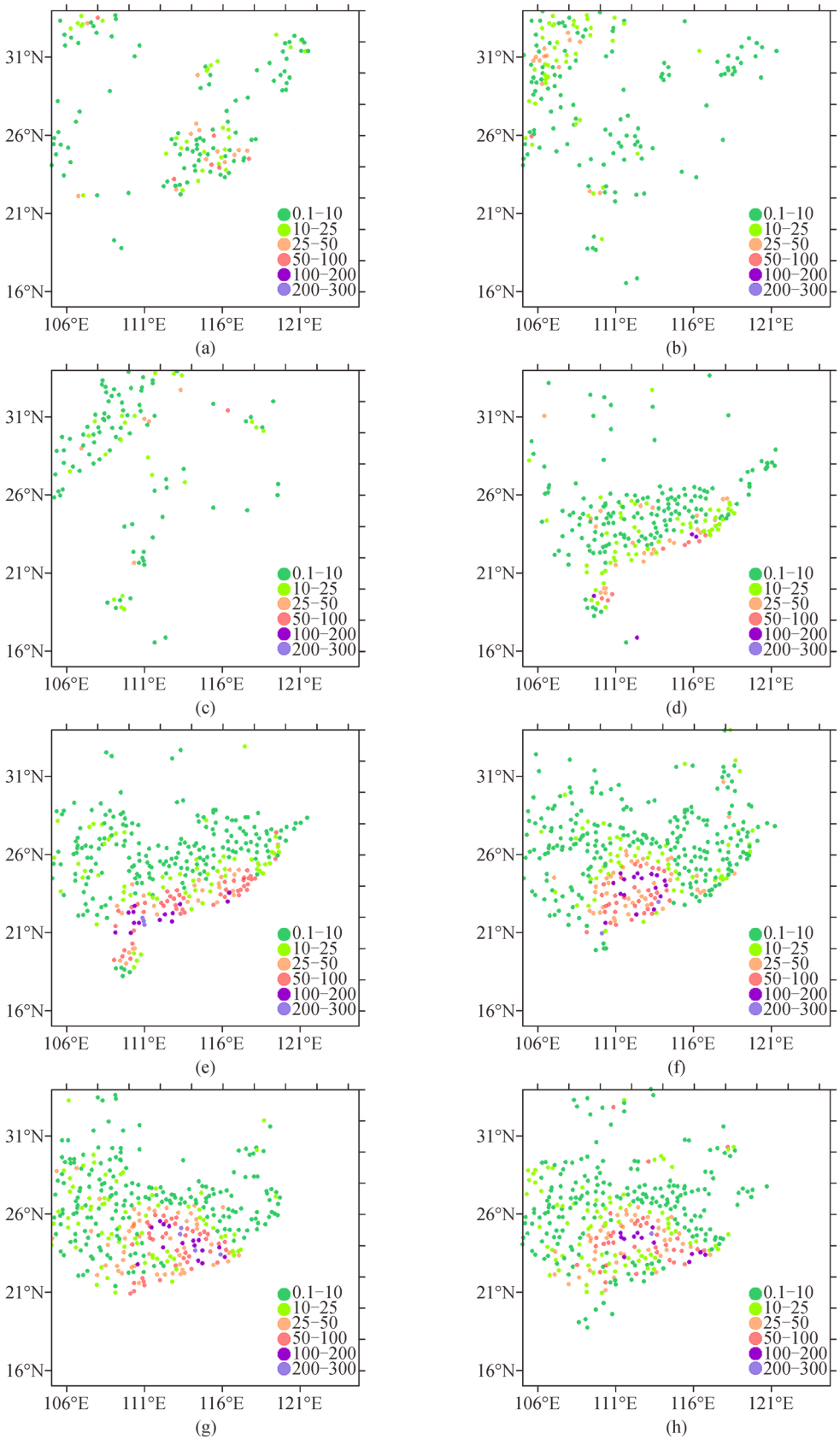

Fig. 5 Observations of the precipitation (unit: mm) for the Typhoon "Utor" during August 11–18, 2013 (The figures of a, b, …, h represent the specific dates of August 11, 12, …, 18, respectively).

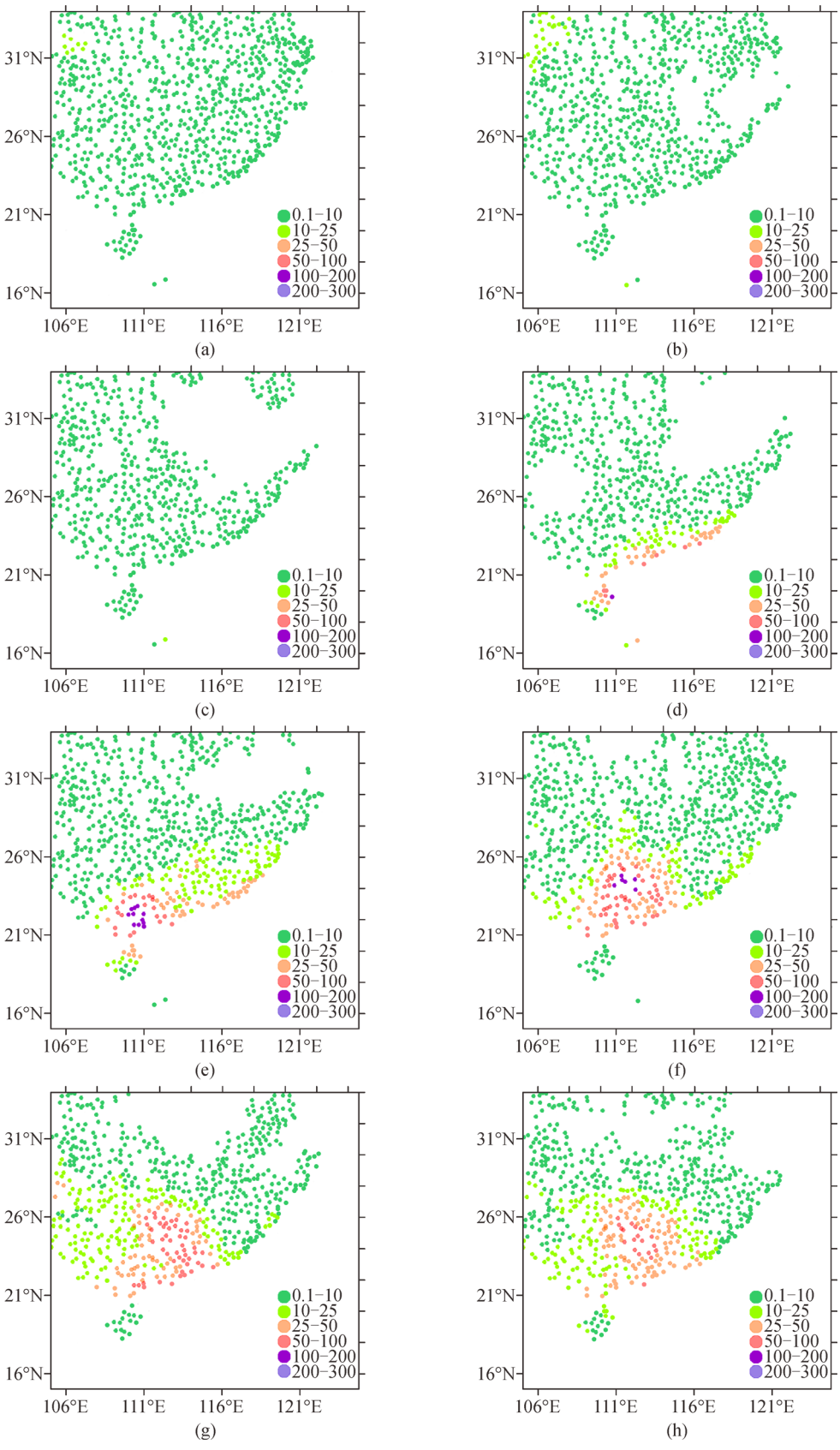

Fig. 6 The 50th percentile forecast of the precipitation (unit: mm) using the Scheme C for the Typhoon "Utor" during August 11–18, 2013 (The figures of a, b, …, h represent the specific dates of August 11, 12, …, 18, respectively).

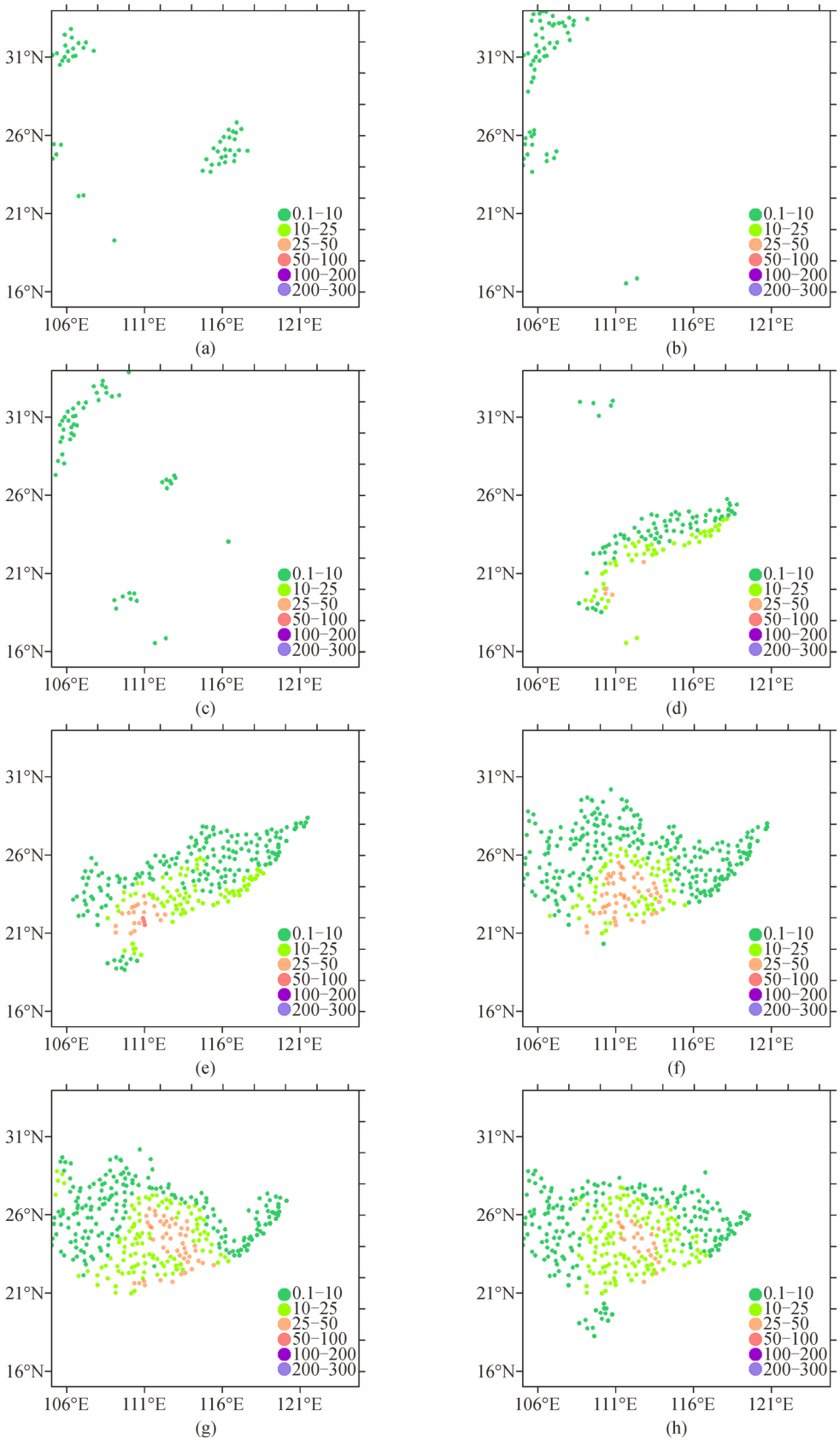

Fig. 7 The same as Fig. 6 (The figures of a, b, ..., h represent the specific dates of August 11, 12, ..., 18, respectively), but for the BMA-calibrated Scheme C.

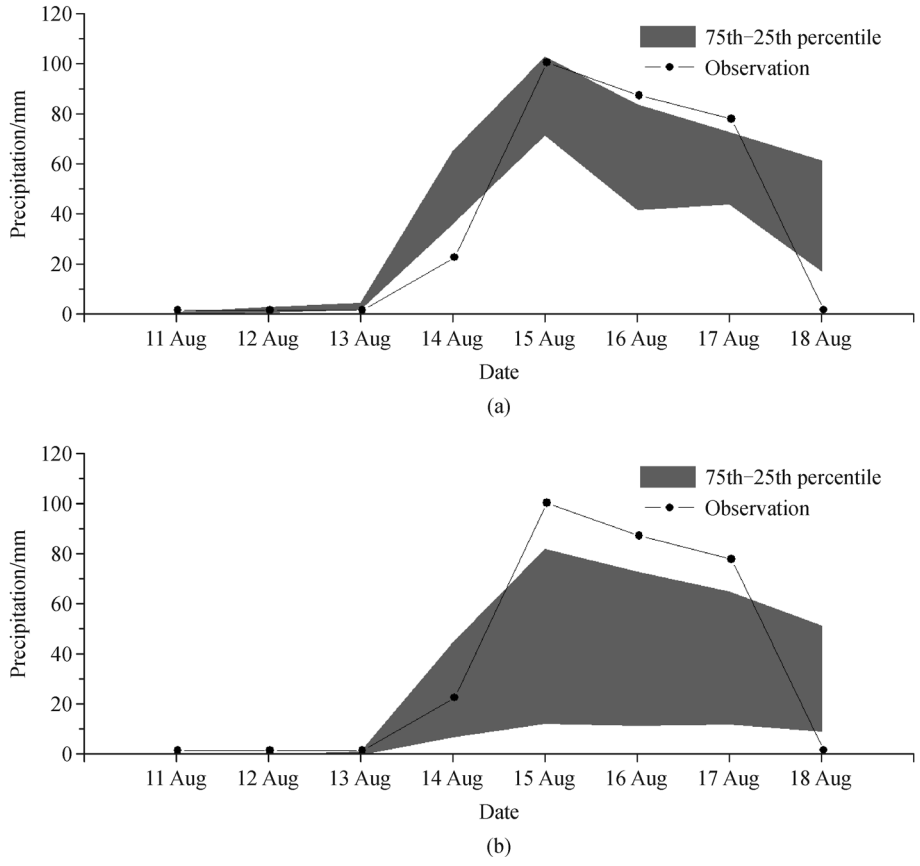

Fig. 8 Percentile forecasts of 24-h accumulated precipitation at Guangdong Yangjiang during August 11–18, 2013 (The gray zone is between 25th and 75th quantile, and the black circles indicate observed precipitation at Guangdong Yangjiang station (a: forecasts based on Scheme C; b: forecasts from the BMA model).

jiang" station and the near the following track of "Xinyi" station in Guangdong province, China (Fig. 1). Taking the 25th to 75th percentiles of PQPF as effective forecast ranges, it is indicated from Fig. 8 that, in the BMA model, observations of 4 days are included in the effective ranges of forecast during the "Utor" influence period (8 days) in the Yangjiang station, whereas it is no days are included in the effective ranges for the raw ensemble of Scheme C. During the 24-h before August 14 (The observation are 0.0, 0.0, 0.0 and 22.0, respectively), the BMA model (The 25th percentiles of BMA are 0.0, 0.0, 0.0, respectively) well corrected the deviations of exceeding precipitation forecast from the raw Scheme C (The 25th percentiles of raw Scheme C are 0.14, 0.89, 1.87 and 36.13, respectively). However, for the 24-h before August 15, the effective forecast ranges from BMA model are lower than those from the raw Scheme C, even showing larger biases. Figure. 9 shows that observations of 5 days are included in the effective ranges of forecast from the BMA model during the 8 day influence period in the Xinyi station, while only 3 days for the raw ensemble forecast of Scheme C. Additionally, the BMA model effectively overcomes the great deficiency of overestimation precipitation amount in the raw ensemble of Scheme C for 14, 15 and 17 August 2013. Hence, the results suggest that although the magnitudes of the precipitation forecast are reduced, the probability of including the observations in the effective forecast ranges is increased after the BMA calibration, that is, more information, such as the prediction uncertainty, can be provided by the probabilistic precipitation forecast from the BMA model, implying its relatively higher capability than the raw ensemble of Scheme C. Furthermore, during the influence period of Typhoon "Utor," the effective forecast ranges of the BMA model are generally lower than the observations at the Yangjiang station (Fig. 8(b)), while at the Xinyi station, most observations are covered (Fig. 9(b)), which could be attributed to the model deficiency in dealing with the regional underlying surface on sea-land changes. Therefore, the BMA model is definitely superior to the raw ensemble Scheme C, but still not perfect and needs further improvements.

5 Summary and discussion

In this study, the ensemble precipitation results from CMA, ECMWF and NCEP are employed in the paper to design three multi model ensemble schemes with lead time of

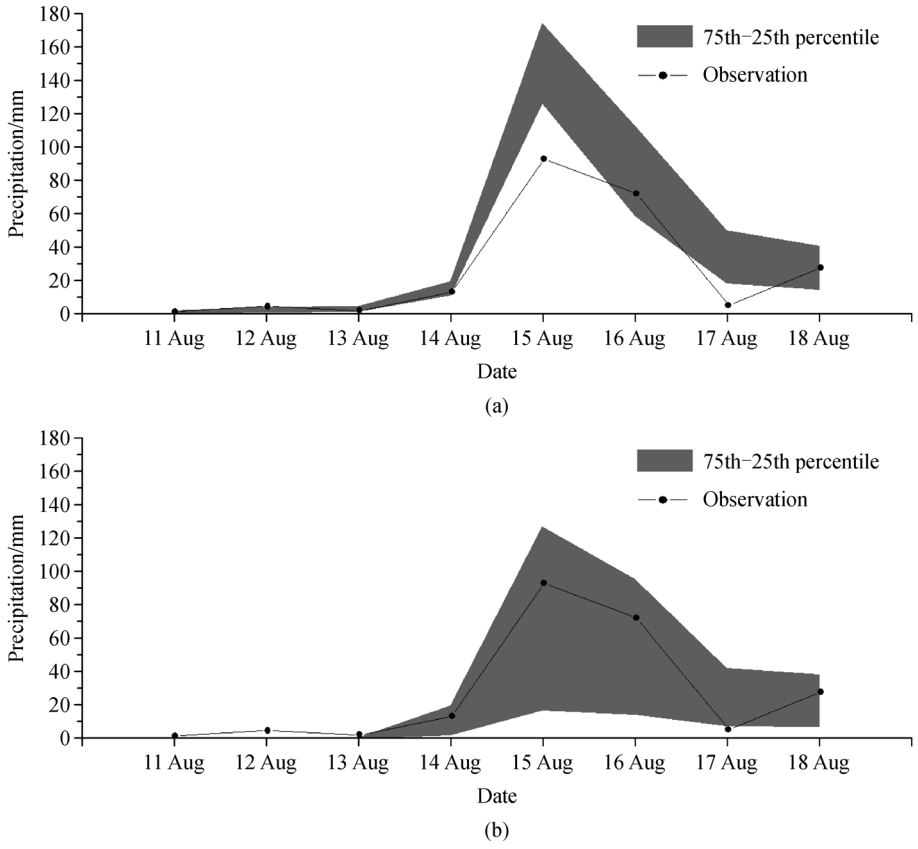

Fig. 9 The same as Fig. 8, but for the Xinyi station (a: forecasts based on Scheme C; b: forecasts from the BMA model).

24-h, 48-h and 72-h. The BMA model is then established based on the three schemes to carry out the experiments of PQPF during the period influenced by three typhoons from July 1 to 30 September 2013. The prediction skills of the BMA calibrated forecast are evaluated by the hourly precipitation observations extracted from the 818 national surface meteorological stations. To a great extent, the BMA model improves the precipitation forecast during the three typhoon influence periods in this study.

As discussed above, in general, the BMA model is superior to the raw ensemble in the probabilistic precipitation forecast during the typhoon influence periods. The MAE is reduced by 12.4% after the BMA calibration. It shows the most significant improvement on Scheme C, followed by Scheme B. Before the typhoon landfall, the MAE scores of three schemes are fairly close after the bias correction by the BMA model. After landfall, the BMA calibrated Scheme C shows obviously lower MAE values than those of non-calibrated or calibrated Schemes A and B. On the other hand, according to the BSs at different precipitation thresholds, the BMA model correction is considerably effective for the light and moderate precipitation, and has some certain improvement on the heavy rainfall prediction. For precipitation at the 50.0 mm threshold, the forecast capability enhancement by the BMA model is relatively limited. Resulted from the BMA calibration, the prediction accuracy of the ensemble precipitation forecast is improved for all lead time.

Comparisons among different ensemble schemes show that, for the 0.1 mm precipitation threshold, the BMA model shows relatively consistent improvements on the three ensemble forecast schemes. However, it differs a lot for the calibration effects at the thresholds over moderate rainfall. For the 10.0 mm threshold, the BMA model is characterized by similar effects between Schemes A and B and largest improvement on Scheme C. For precipitation thresholds of 25.0 mm and 50 mm, it is also the best for the forecast correction of Scheme C, followed by Scheme B.

According to investigations on the forecast skills for different precipitation thresholds in a specific ensemble scheme, the best calibration of the BMA model occurs in the probabilistic precipitation forecast at the 10.0 mm threshold for all the three schemes. In Scheme A, it is followed by the precipitation threshold of 50.0 mm and then 25.0 mm, while they are reversed in both Schemes B and C.

The BMA model forecast is superior to the raw ensemble for the whole period, with the CRPS reduced by 26.2%. The BMA model can significantly improve the probabilistic precipitation forecast of Scheme C for the three lead time. Greatest improvement features the 24-h precipitation forecast by the BMA model, followed by the

lead time of 48-h. It is indicated that the calibration capability of the BMA model is closely related to forecasts from the raw ensemble.

The calibrated probabilistic precipitation forecast from the BMA model for the typhoon case of "Utor" shows significant deviation correction on false predictions in Scheme C, making the main precipitation areas more prominent. When the precipitation is larger than 50.0 mm, even exceeding 100.0 mm, the 50th percentile precipitation forecast locates in the range from 25.0 mm to 50.0 mm after the BMA calibration, that is, to some extent, the magnitudes of the heavy rainfall are underestimated by the BMA model.

Explorations imply that in view of probabilistic precipitation forecasts, the BMA can effectively improve the ensemble forecast, and the probability of including the observations in the effective forecast ranges are increased after the BMA calibration. The BMA model is characterized by relatively higher capability of probabilistic precipitation forecast than the raw ensemble of Scheme C.

Ensemble forecast is an effective method to quantify the forecast uncertainties, which can capture the extreme weather events well. As a post-processing method of the ensemble forecast, BMA model can make full use of the ensemble information. The members' weights are determined by the posterior probabilities of corresponding model, representing the relative forecast skill of each member in the training period, and they are obtained from the training tests on various ensemble prediction systems combined with the observations. Additionally, aiming at the optimal forecast, it is of great importance to select the training data and determine the optimal length of training period from continuous training experiments.

It should be noted that the statistical distributions of precipitation amounts are distinctly asymmetric, and skewed to the right, which are physically constrained to be nonnegative. The gamma distribution model is applied in the precipitation distribution, which describes the nonzero precipitation in BMA. The gamma distribution (i.e., Type III from the Pearsonian system) is a limiting case of Type I distribution of Pearsonian system. It is next to the normal distribution in simplicity, and, at the same time, it covers a wide range of skewness. And the gamma distribution is positively skewed, meaning that it has an extended tail to the right of the distribution. This is advantageous because it models actual rainfall distributions for many areas, where there is a nonzero probability of extremely high rainfall amounts, even though the typical rainfall may not be very large (Ananthakrishnan and Soman, 1989). Because the PDF of the gamma distribution takes on a wide variety of shapes depending on the value of the shape parameter, α. The role of the scale parameter, β, effectively is to stretch or squeeze (i.e., to scale) the gamma density function to the right or left, depending on the overall magnitudes of the data values represented. As the distribution is stretched to the right by larger values of β. Conversely, as the density is squeezed to the left, its height must rise (Wilks, 2006). This flexibility allows for the gamma distribution to be fit to any number of rainfall regimes with reasonable accuracy, while other distributions may fit only a single, specific rainfall regime.

However, it is more difficult to work with the Gaussian distribution, because obtaining good parameter estimates from particular batches of data are not as straightforward. The simplest (although certainly not best) approach to fitting a gamma distribution is to use the method of moments. Even here, however, there is a complication, because the two parameters for the gamma distribution do not correspond exactly to moments of the distribution, as was the case for the Gaussian. The mean of the gamma distribution is given by the product $\alpha\beta$, and the variance is $\alpha\beta^2$ (Wilks, 2006). Equating these expressions with the corresponding sample quantities yields a set of two equations in two unknowns, which can be solved to yield the moments estimators. The moments estimators for the gamma distribution are not too bad for large values of the shape parameter, perhaps $\alpha > 10$, but can give very bad results for small values of α (Thom, 1958; Wilks, 1990). Some results of research show that since the occurrence of such an extreme value is very rare and the historic records are not so long, it is difficult to fit tail of gamma distribution function with observed records, which is also likely due to the observed data containing many periods in which successive values are highly correlated, whereas the synthetic data vary more randomly (Semenov and Porter, 1995).

Moreover, it must be emphasized that the application of the BMA method also has some limitations because it fairly relies on the observations. Additionally, in the paper, the BMA model median and the raw ensemble mean are applied as deterministic results in the MAE comparison between forecasts before and after the BMA calibration. Thus, the information of extreme events, such as the super Typhoon "Utor," may not be well captured. How to extract the valuable information of typhoon precipitation remains to be investigated more comprehensively in the PQPF.

Acknowledgements This research was funded by the National Key R&D Program of China (No. 2017YFC1502000), and the Chinese Ministry of Science and Technology Project (No. 2015CB452806), the National Natural Science Foundation of China (Grant No. 41475044) and National Key Technology Research and Development Program of the Ministry of Science and Technology of China (Grant No. 2015BAK10B03). We gratefully acknowledge the anonymous reviewers for spending their valuable time and providing constructive comments and suggestions on this manuscript.

References

Ajami N K, Duan Q Y, Gao X G, Sorooshian S (2006). Multimodel combination techniques for analysis of hydrological simulations: application to distributed model intercomparison project results. J Hydrometeorol, 7(4): 755–768

Ananthakrishnan R, Soman M K (1989). Statistical distribution of daily rainfall and its association with the coefficient of variation of rainfall series. Int J Climatol, 9(5): 485–500

Bankman I N (2009). Handbook of Medical Image Processing and Analysis. Elsevier

Buizza R, Hollingsworth A, Lalaurette F, Ghelli A (1999). Probabilistic predictions of precipitation using ECMWF ensemble prediction system. Weather Forecast, 14(2): 168–189

Chen B Y, Guo Y Q, Dai K, Qian Q F (2016). Research for the ensemble member optimization correction technique on typhoon rainstorm forecast and its application experiment. Meteorol Monogr, 42(12): 1465–1475 (in Chinese)

Chen P Y, Lei X T, Ying M (2013). Introduction and application of a new comprehensive assessment index for damage caused by tropical cyclones. Trop Cyclone Res Rev, 2(3): 176–183

Cheng Z Q, Lin L X, Sha T Y, Yang G J (2014). Analysis of atmosphere stratification in extremely heavy rainfall event associated with severe typhoon "Utor". Meteorological Monthly, 40(12): 1507–1512 (in Chinese)

Cui B, Toth Z, Zhu Y, Hou D (2012). Bias correction for global ensemble forecast. Weather Forecast, 27(2): 396–410

ECMWF (2001).Grid point to grid point interpolation. Available at ECMWF Website

Fraley C, Raftery A E, Gneiting T (2010). Calibrating multimodel forecast ensembles with exchangeable and missing members using bayesian model averaging. Mon Weather Rev, 138(1): 190–202

Glahn H R, Lowry D A (1972). The use of model output statistics (MOS) in objective weather forecasting. J Appl Meteorol, 11(8): 1203–1211

Gu H, Qian C H, Gao S Z, Xiang C Y (2017). The impact of tropical cyclones on China in 2016. Trop Cyclone Res Rev, 6(1–2): 1–12

Guo D F, Bao H M, Zheng J G, Wu J (2014). Typhoon "Trami" track and the precipitation diagnostic analysis during its impacting on Jiangxi. Meteorology Disaster, 37(1): 29–38 (in Chinese)

Hamill T M, Colucci S J (1998). Evaluation of Eta-RSM ensemble probabilistic precipitation forecasts. Mon Weather Rev, 126(3): 711–724

Hamill T M, Snyder C, Morss R E (2000). A comparison of probabilistic forecasts from bred, singular-vector, and perturbed observation ensembles. Mon Weather Rev, 128(6): 1835–1851

Hamill T M, Snyder C, Whitaker J S (2003). Ensemble forecasts and the properties of flow-dependent analysis-error covariance singular vectors. Mon Weather Rev, 131(8): 1741–1758

Hamill T M, Whitaker J S, Wei X (2004). Ensemble reforecasting improving medium-range forecast skill using retrospective forecasts. Mon Weather Rev, 132(6): 1434–1447

Hurricane Research Division (2015). Which countries have had the most tropical cyclones hits? Frequently Asked Questions NOAA AOML. (Available at NOAA Website)

Krzysztofowicz R, Drzal W J, Drake T R, Weyman J C, Giordano L A (1993). Probabilistic quantitative precipitation forecasts for river basins. Weather Forecast, 8(4): 424–439

Krzysztofowicz R (1998). Probabilistic hydrometeorological forecasts: toward a new era in operational forecasting. Bull Am Meteorol Soc, 79(2): 243–251

Li J, Du J, Liu Y (2015). A comparison of initial condition-, multi-physics- and stochastic physics-based ensembles in predicting Beijing "7.21" excessive storm rain event. J Meteorol Resh, 73(1): 50–71 (in Chinese)

Li W T, Duan Q Y, Miao C Y, Ye A Z, Gong W, Di Z H (2017). A review on statistical postprocessing methods for hydrometeorological ensemble forecasting: statistical postprocessing methods. WIRes Water, 4(6): e1246

Li Y Z, Wu D Y, Liu B F, Wang X Y, Lin X Q, Zhao H W (2013). Consideration on the service for forecasting and warning of wind disaster caused by the severe typhoon "Usagi" landing on the Shanwei. Guangdong Meteorol, 35(6): 60–64 (in Chinese)

Liang L, Zhao L N, Qi D, Wang C X, Bao H J, Zhang Y J (2013). The experiment of hydrologic probabilistic forecast based on the precipitation forecast calibrated by Bayesian Model Averaging. J Appl Meteorolog Sci, 24(4): 416–424 (in Chinese)

Liu J, Xie Z (2014). BMA probabilistic quantitative precipitation forecasting over the Huaihe Basin using TIGGE multimodel ensemble forecasts. Mon Weather Rev, 142(4): 1542–1555

Raftery A E, Gneiting T, Balabdaoui F, Polakowski M (2005). Using Bayesian Model Averaging to calibrate forecast ensembles. Mon Weather Rev, 133(5): 1155–1174

Ren F M, Wu G X, Wang X L, Wang Y M, Dong W J, Liang J, Bai L N (2011). Tropical Cyclones Affecting China over the Last 60 years. Beijing: Meteorological Press (in Chinese)

Scheuerer M, Hamill T M (2015). Statistical postprocessing of ensemble precipitation forecasts by fitting censored, shifted gamma distributions. Mon Weather Rev, 143(11): 4578–4596

Shrestha D L, Robertson D E, Bennett J C, Wang Q J (2015). Improving precipitation forecasts by generating ensembles through postprocessing. Mon Weather Rev, 143(9): 3642–3663

Semenov M A, Porter J R (1995). Climatic variability and the modelling of crop yields. Agric Meteorol, 73(3-4): 265–283

Sloughter J M, Raftery A E, Gneiting T, Fraley C (2007). Probabilistic quantitative precipitation forecasting using Bayesian Model Averaging. Mon Weather Rev, 135(9): 3209–3220

Smith R L, Tebaldi C, Nychka D, Mearns L O (2009). Bayesian modeling of uncertainty in ensembles of climate models. J Am Stat Assoc, 104(485): 97–116

Theis S E, Hense A, Damrath U (2005). Probabilistic precipitation forecasts from a deterministic model: a pragmatic approach. Meteorol Appl, 12(3): 257–268

Thom H C S (1958). A note on the gamma distribution. Mon Weather Rev, 86(4): 117–122

Nakazawa T, Swinbank R, Toth Z, Ebert B (2010). THORPEX/TIGGE applications to TC motion and forecasting. The World Meteorological Organisation (WMO) International Workshop on Tropical Cyclones (IWTC-VII) Report. La Réunion

Wang Y J, Wen S S, Li X C, Fischer T, Su B D, Wang R, Jiang T (2016). Spatiotemporal distributions of influential tropical cyclones and associated economic losses in china in 1984–2015. Nat Hazards, 84 (3): 2009–2030

Wilks D S (1990). Maximum likelihood estimation for the gamma distribution using data containing zeros. J Clim, 3(12): 1495–1501

Wilks D S (2006). Statistical Methods in the Atmospheric Sciences. 2nd ed. Academic Press, 627

Wilson L J, Beauregard S, Raftery A E, Verret R (2007). Calibrated surface temperature forecasts from the Canadian ensemble prediction

system using Bayesian model averaging. Mon Weather Rev, 135(4): 1364–1385

Yang C, Yan Z W, Shao Y H (2012). Probabilistic precipitation forecasting based on ensemble output using generalized additive models and Bayesian Model Averaging. Acta Meteorol Sin, 26(1): 1–12

Ying M, Bai L N, Zhan R F (2014a). Tropical cyclone activity over the western North Pacific in 2013. Trop Cyclone Res Rev, 3(3): 131–144

Ying M, Zhang W, Yu H, Lu X Q, Feng J X, Fan Y X, Zhu Y T, Chen D Q (2014b). An overview of the China meteorological administration tropical cyclone database. J Atmos Ocean Technol, 31(2): 287–301

Zhao L N, Bai X M, Xing C, Wang B Y, Li Y T, Li X M, Yang R W (2017). Probability distribution of hourly precipitation in typhoons during summer seasons in the southeast China. Torrential Rain Disa, 36(2): 97–107 (in Chinese)

Zhao L N, Tian F Y, Wu H, Qi D, Di J Y, Wang Z (2011). Verification and comparison of probabilistic precipitation forecasts using the TIGGE data in the upriver of Huaihe basin. Adv Geosci, 29: 95–102

Zhu J S, Kong F Y, Ran L K, Lei H C (2015). Bayesian Model Averaging with stratified sampling for probabilistic quantitative precipitation forecasting in northern China during summer 2010. Mon Weather Rev, 143(9): 3628–3641

AUTHOR BIOGRAPHIES

Linna Zhao received the Ph.D. degree in meteorology from the university of Chinese Academy of Sciences, China in 2003.

She is currently a Professor of Centre for Analysis and Applications of Meteorological Data, Chinese Academy of Meteorological Sciences. Her research interests include comprehensive study of ensemble prediction, post-processing of numerical weather prediction and probability quantitative precipitation forecast. Her recent research projects have been funded by National Key R&D Program of China, the Chinese Ministry of Science and Technology project (973 Program), and the National Natural Science Foundation of China.

Prof. Zhao is the Expert of the CHy-XIII in WMO, and Vice-chairman of the Commission of Hydrometeorology of the Chinese Meteorological Society, Member of Chinese Meteorological Society, National Planning Environmental Impact Assessment Expert, Member of the National Advisory Council of Environmental and Health Expert. She won the 2011' Outstanding Science and Technology Achievement Prize of the Chinese Academy of Sciences (Major Contributor) Award for Meteorology.
E-mail: zhaoln@cma.gov.cn

Xuemei Bai received the M. S. degree in meteorology from Chengdu University of Information Technology, China in 2016. She is currently a weather forecaster of Heilongjiang Meteorological Observatory, China. Her research interests focus on probability quantitative precipitation forecast and prediction of landing typhoon precipitation.
E-mail: b_xmei@163.com

Dan Qi received the M. S. degree in meteorology from Nanjing University of Information Science & Technology, China in 2005. She is a weather forecaster in National Meteorological Center of China and constructed several real-time operating systems as Senior Engineer. Her research interests include quantitative precipitation forecast and quantitative precipitation estimation, and the application in the coupling of meteorological and hydrological models. Ms. Qi is the Member of the CHy-XIV in WMO and published lots of papers in journals.
E-mail: qidan@cma.gov.cn

Cheng XING received the B. S. degree in meteorology from Chengdu University of Information Technology, China in 2013. His research interests focus on numerical weather prediction.
E-mail: xingc1010@163.com