



# Improving quantitative precipitation estimates by radar-rain gauge merging and an integration algorithm in the Yishu River catchment, China

Yuehong Shao<sup>1</sup> · Aolin Fu<sup>1</sup> · Jun Zhao<sup>1</sup> · Jinchao Xu<sup>1</sup> · Junmei Wu<sup>2</sup>

Received: 12 August 2020 / Accepted: 4 January 2021 / Published online: 21 February 2021  
© The Author(s), under exclusive licence to Springer-Verlag GmbH, AT part of Springer Nature 2021

## Abstract

High-precision areal rainfall is crucial for hydrometeorological coupled forecasts. The accuracy of quantitative precipitation estimates (QPE) is improved by merging radar-rain gauge data with an integration approach based on a statistical weight matrix in the Yishu River catchment, China. First, a local Z-R relationship ( $Z = 85R^{1.82}$ ) is reconstructed using a genetic optimization algorithm to minimize the error from different precipitation patterns and climate zones. Next, based on the local Z-R relationship, six methods of merging radar-rain gauge data are respectively adapted to improve the accuracy of QPE, as follows: mean field bias (MFB), Kalman filter (KLM), optimum interpolation (OPT), variation method (VAR), two-step calibration of KLM and OPT (KOP), and two-step calibration of KLM and VAR (KVR). The results indicate that QPE accuracy is clearly improved, and is in good agreement with rain gauge observations, after the six merging methods are applied. Among these methods, KOP performs the best, reducing the mean relative error from 55.2 to 15.1%. An innovative aspect of this work is the inclusion of an integrated ideology based on a statistical weight matrix, which further improves the accuracy of QPE by incorporating the advantages of each estimation mode. The results further show that the accuracy of QPE derived from the integration approach is higher than that obtained by any individual method; QPE values are similar to those obtained the automatic rain gauge network in both the spatial distribution and location of the intense precipitation centers, and better reflects the precipitation status over the ground surface. This approach could serve as a promising conventional method for QPE in the study region.

## 1 Introduction

Precipitation is a key variable of water cycle, and has profound impacts on hydrological and meteorological processes,

including the potential to cause flash floods, debris flow, and other natural disasters (Jonkman 2005; Maggioni and Massari 2018). Accurate and timely rainfall data is therefore crucial for hydrometeorological forecasting and early flash-flood warnings (Sharifi et al. 2018). Rain gauges are a conventional observation method that can provide direct and fairly accurate precipitation measurements at a single point (Cong and Liu 2011). However, because precipitation is a weather phenomenon with substantial spatial and temporal fluctuations, it is also associated with a high degree of error and uncertainty. Regional rainfall estimates interpolated by rain gauges often produce large errors because of sparse rain gauge networks, and/or complex and irregular terrain—especially for heavy or local convective storms (Lee et al. 2013; Ku et al. 2015; Ochoa-Rodriguez et al. 2019).

With progress being made in remote sensing technology, in recent decades, weather radar has been widely used for estimating precipitation because it can measure rainfall in real time with high spatial resolution and temporal continuity (Germann et al. 2006; Berne and Krajewski 2013). The relationship between radar reflectivity ( $Z$ ) and surface rainfall rate

✉ Yuehong Shao  
syh@nuist.edu.cn

Aolin Fu  
2252758417@qq.com

Jun Zhao  
zsmzyg@126.com

Jinchao Xu  
xujinchao301@foxmail.com

Junmei Wu  
wjmqingkong@126.com

<sup>1</sup> School of Hydrology and Water Resources, Nanjing University of Information Science and Technology, Nanjing 210044, China

<sup>2</sup> Kunshan Meteorological Bureau, Kunshan 215300, China

( $R$ ) has been previously studied; in these papers, an empirical equation of  $Z = 200R^{1.6}$  has commonly been used, regardless of climate region (Marshall and Palmer 1948). However, rainfall intensity can be affected by season, region, and rainfall type, which can lead to a large deviation between rainfall values estimated by an empirical Z-R relationship and observed values (Chapon et al. 2008). Many studies have shown that the accuracy of quantitative precipitation estimates (QPE) can be improved by reconstructing the local Z-R relationship to some extent (Wang et al. 2012; Gou et al. 2015; Zhang et al. 2016). However, precipitation values estimated by the Z-R relationship may still be associated with large errors due to distance attenuation, non-meteorological echoes, and bright-band contamination, among other phenomena (Lafont and Guillemet 2004; Jacobi and Heistermann 2016).

To improve the accuracy of QPE, in recent years, efforts have been made to merge radar and rain gauge data using a number of methods (Goudenhoofdt and Delobbe 2009; Martens et al. 2013; Berndt et al. 2014; Rabiei and Haberlandt 2015; Wang et al. 2015; Hasan et al. 2016). Early research focused on bias correction in radar precipitation estimates using rain gauge observations. Mean field bias (MFB) is broadly used to reduce the mean radar-gauge error, calculated by dividing the gauge amount by the radar amount (Smith and Krajewski 1991; Seo et al. 1999). However, due to a uniform multiplicative adjustment factor in the precipitation field, MFB may lead to an obvious overestimation or underestimation at rainfall stations where the reflectivity factor values are too high or low (Li et al. 2014; Wang et al. 2013). Alternatively, some scholars have applied a Kalman filter (KLM) to calibrate the error from the temporal domain. The results have shown that KLM performs well when the error field is stable over time. However, estimation reliability is significantly reduced in cases of drastic changes in intensity (Chumchuan et al. 2006; Kim and Yoo 2014a, b).

Subsequent research focused mainly on spatial variability in radar and rain gauge data to improve the accuracy of precipitation estimates. A variety of methods, including as optimum interpolation (OPT), variation method (VAR), kriging, co-kriging, kriging with external drift, and conditional merging, have been applied to correct radar and rain gauge biases (Krajewski 1987; Sinclair and Pegram 2005; Haberlandt 2007; Shao et al. 2008; Sideris et al. 2014; Cantet 2017). For example, VAR takes into account the spatial distribution field of precipitation, and corrects the radar observation in space and time; this method is more accurate and reliable when the precipitation is more uniform, and less affected by terrain, or station density (Zhang et al. 1992; Bianchi et al. 2013; Li et al. 2015a, b). The OPT makes full use of the high precision of rain gauge at point and objective and spatial precipitation field measured by radar, which has an advantage over the VAR in cases of heterogeneous precipitation caused by convective precipitation and complex terrain. However, precipitation recovery

ability cannot be achieved, as in the VAR method, when ground objects are blocked (Victor and Alvarez 2010). Co-kriging, kriging with external drift, and conditional merging are all popular geo-statistical techniques used for merging radar and gauge data that are suitable for merging spatially continuous grid-based measurements and rain gauge data as a primary source. However, many studies have shown that the results are inaccurate when the covariance function does not match the actual spatial distribution of rainfall (Sinclair and Pegram 2005; Yoo and Park 2008; Martens et al. 2013). In recent years, two-step calibration methods like two-step calibration of KLM and OPT (KOP) and KLM and VAR (KVR) have been used in radar QPE from the temporal and spatial domains. Research has shown that KOP and KVR are better than KLM, OPT, and VAR alone, but they can make the precipitation echo center smooth when it is far away from the correction rain gauge, and may fail to reflect the true precipitation spatial structure (Sun et al. 1993; Zhao et al. 2001; Gao et al. 2004; Li et al. 2009).

Due to the specific applicability of each mode of estimation, the most effective use of data is unlikely to be achieved using a deterministic mode (Villarini and Krajewski 2010; He et al. 2013; Li et al. 2014). An integrated method is beneficial, not only because it reflects the inherent chaos and stochasticity of the precipitation system but also because it effectively reduces the uncertainty in estimation caused by errors in observation and analysis (Wu et al. 2016; Cecinati et al. 2017). The method of average weight integration is commonly used for this purpose, wherein the arithmetic average is obtained from several estimation results of integration, and a deterministic result is obtained from the perspective of probability (Chumchuan et al. 2006b; Chao et al. 2018). However, the average integration fails to make full use of the advantages of each method, nor does it faithfully reflect the spatial structure of rainfall. Li et al. (2014) proposed an integration method based on critical probability, and selected different thresholds to output results according to practical application needs, so as to make reasonable recommendations. However, this method has not been tested or evaluated with data and a study area. To better reflect the spatial structure of rainfall and make full use of the advantages of each estimation method, in this paper, an integration method based on a statistical weight matrix is used to improve the accuracy of QPE.

This paper is structured as follows: First, the local Z-R relationship is established using a genetic optimization algorithm to reduce errors caused by an empirical Z-R relationship. Second, a radar-rain gauge merging method is adopted to improve the accuracy of QPE, and evaluated from the Z-R relationship and different merging methods based on a contingency table approach. Finally, an integration method based on a statistical weight matrix is used to further improve the accuracy of QPE by making full use of the advantage of each deterministic mode of estimation. Considering the spatiotemporal heterogeneity of rainfall, this integration method can reduce the uncertainty caused by observation and analysis errors, which is a major

advancement of this research compared to previous studies. The results of this study are expected to provide a stable and reliable method for radar QPE in the Yishu River basin.

## 2 Study area and data

### 2.1 Study area

The study area (34° 26'~36° 09' N, 117° 30'~119° 08' E) is located in the Yishu River basin of eastern China and covers drainage area of approximately  $2.6 \times 10^4 \text{ km}^2$ . The altitude ranges from 48 to 545 m. The study area consists of plains located mainly in the southeast, and mountains and hills concentrated in the north. The area corresponds to a typical, warm temperate continental monsoon climate zone, and is characterized by a hot and rainy season due to the influence of subtropical, high pressures in summer. The annual average temperature and rainfall are 12 °C and 830 mm, respectively. Precipitation varies widely and the spatial distribution of precipitation is uneven, with 50–80% of rain falling between June and September. The elevation and meteorological stations in the study area are shown in Fig. 1.

### 2.2 Precipitation data

Precipitation data are obtained from rain gauge observations and weather radar inversion. First, according to the collected data, hourly rainfall data in 2006 are used to estimate QPE from 10 conventional weather stations (CWS) and 127 intensified automatic weather stations (IAWS) in the study area. Areal rainfalls from the rain gauge network are interpolated to  $1 \text{ km} \times 1 \text{ km}$  via

an inverse distance-weighted (IDW) method, which corresponds to the spatial resolution of radar rainfall field. Second, the radar data are received from an S-band single polarization Doppler radar at Linyi station that has an effective range of 230 km. The radar performs approximately nine different elevation scans between 0.5 and 19.5° above the horizontal axis. The radar completes a volume scan in approximately 6.0 min. The constant altitude plan position indicators (CAPPI) of the compound plane at a height of 1.0 km are obtained by the radar reflectivity factors of the four lowest elevation angles, which correspond to PPI data from 3.4, 2.4, 1.5, to 0.5° elevation angles within 0–20 km, 20–35 km, 35–50 km, and 50–230 km, respectively. The CAPPI is converted to rain rate using the local Z-R relationship.

Considering the consistency of rainfall type and weather system, six strong precipitation events (20060629, 20060703, 20060806, 20060816, 20060826, and 20060829) are selected to evaluate the radar-rain gauge merging and integration methods. Eighty-eight samples with a 1-h time resolution are collected from six processes. According to the principle of uniform sampling, 44 samples are used for multiple regression to generate a weight matrix for integration algorithm, and the remaining 44 samples are used to test the algorithm.

## 3 Methodology

### 3.1 The Z-R relationship based on a genetic optimization algorithm

In mathematics, a genetic optimization algorithm is based on the idea of natural selection and includes the selection of an

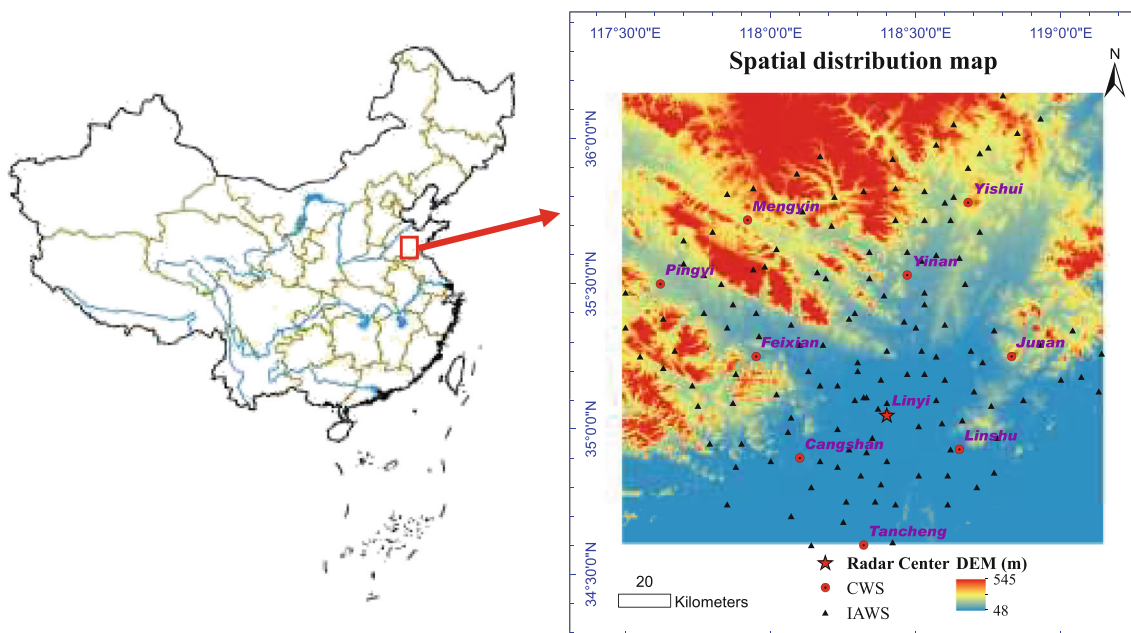


Fig. 1 Spatial map showing the elevation of the study area, and meteorological stations, and radar center

objective function, coding parameters, and configuration of the genetic operator. The procedure occurs as follows: first, the amount of rainfall per hour is measured by each rain gauge. Second, the corresponding radar echo reflectivity ( $Z$ ) in space-time measured by the radar above these rain gauges is converted into rain intensity ( $R$ ) based on an assumed Z-R relationship ( $Z = AR^b$ ), and the radar estimation of rainfall per hour is obtained by accumulated rain intensity over time. The formula can be written as follows:

$$H_n = \sum_{m=1}^M (W_{nm} \times I_{nm}) \tag{1}$$

where  $H_n$  is the radar estimation value at the  $n^{\text{th}}$  hour;  $W$  and  $I$  are the time-weighted factor and rain intensity at the  $m^{\text{th}}$  volume scan during the  $n^{\text{th}}$  hour, respectively; and  $m$  is the number of radar volume scans per hour.

Finally, a criterion function (CTF) is selected according to the principle of minimum error. It is defined as follows:

$$\text{CTF} = \min \left\{ \sum_{n=1}^N (H_n - G_n)^2 + (H_n - G_n) \right\} \tag{2}$$

where  $G_n$  is the measured value of the rain gauge at the  $n^{\text{th}}$  hour.

Two-dimensional floating-point encoding and a Gaussian equality crossover operator are respectively used to improve local search ability and enhance the stability of optimization.

$$x_i^{c1} = x_i^{p1} + \frac{d_i}{3}u(0, 1) \quad x_i^{c2} = x_i^{p2} + \frac{d_i}{3}u(0, 1) \tag{3}$$

where  $x_i^{p1}$  and  $x_i^{p2}$  are the  $i^{\text{th}}$  gene of the two parent individuals undergoing crossover operation, respectively;  $x_i^{c1}$  and  $x_i^{c2}$  are the  $i^{\text{th}}$  gene of the offspring generated by the crossover operator;  $d_i$  is the distance between the two parent individuals at the  $i^{\text{th}}$  gene. The Gaussian mutation operator is defined as:

$$x_i^c = x_i^p + f_{\text{dev}} \times u(0, 1) \tag{4}$$

where  $f_{\text{dev}}$  is a constant that controls the range of mutation. When seeking the optimal solution, the random initial population size is 120; the crossover probability and mutation probability are 0.85 and 0.05, respectively; and the maximum evolutionary algebra is selected as 300.

The parameters  $A$  and  $b$  are continuously adjusted via the genetic optimization algorithm until the CTF reaches the minimum value. At this time, the optimal  $A$  and  $b$  are determined in the Z-R relationship.

### 3.2 Merging of radar and rain gauge data and evaluation of QPE

In this study, radar and rain gauge merging is adopted to improve the accuracy of QPE. Considering the source of error

in the spatial-temporal domain, six common merging approaches—that is, AVG, KLM, OPT, VAR, KOP, and KVR—are used.

For the AVG method, MFB is used to calibrate the deviation of radar estimation to obtain ground precipitation estimated by radar, which is calculated by dividing the gauge amount by the radar amount. KLM is primarily used to eliminate interference of random noise on radar QPE. The basic idea is to obtain the  $f_1$  and  $f_2$  values of deviation estimation from the state equation and measurement equation, which are weighted to obtain the best estimation of  $f$  with the smallest variance. The OPT method makes use of the deviation between the radar estimation and the value measured by  $n$  rain gauges within a certain radius around the radar grid point. Radar estimation at the grid point is corrected by solving the optimal weight coefficient. The VAR method is an objective analysis method that considers the spatiotemporal distribution of precipitation. From the standpoint of extreme values, VAR attempts to find an optimal analysis field between the ground rain gauge field and the radar initial value field to minimize the analysis error. The KOP and KVR methods are both two-step calibration methods. First, the precipitation field estimated by the radar is calibrated via the KLM method in the time domain. Second, the spatial distribution field of precipitation is calibrated by means of the OPT and VAR methods in the spatial domain. Additional details and procedures for the six methods can be found in Shao et al. (2008) and Shao (2010).

The contingency table approach is used to evaluate the precipitation accuracy based on the six methods outlined above, represented by a  $2 \times 2$  matrix. Each cell of the matrix represents whether rainfall gauge observations and radar estimates have reached or exceeded a certain threshold over a certain period of time. Four statistic parameters, the bias score ( $B_S$ ), threat score ( $T_S$ ), bias percentage ( $B_P$ ), and root mean square error (RMSE) are used to assess the precision. These parameters are calculated as follows:

$$\begin{aligned} B_S &= \frac{A + B}{A + C}, \\ T_S &= \frac{A}{A + B + C}, \\ B_P &= \frac{\sum_{n=1}^{N_{\text{est}}} P_n}{\sum_{n=1}^{N_{\text{est}}} X_n}, \end{aligned} \tag{5}$$

$$\text{RMSE} = \left[ \frac{\sum_{n=1}^{N_{\text{obs}}} (P_n - X_n)^2}{N_{\text{obs}}} \right]^{\frac{1}{2}}$$

where  $X_n$  and  $P_n$  are the rain gauge observation value and the radar-estimated precipitation value at the  $n^{\text{th}}$  station;  $A$  is the number of radar estimations and rain gauge observations at each station that are greater than or equal to a given threshold;  $B$  is the number of radar-estimated values only that reach or exceed a given threshold;  $C$  is the number of rain gauge values only that reach or exceed a given threshold; and  $D$  is the number of values for which neither the rain gauge nor radar-

**Table 1** Evaluation of time points and the number of rain gauge stations reaching threshold values

Time	The number of rain gauges exceeding the threshold value						
	0.1 mm	1.0 mm	2.5 mm	5.0 mm	7.5 mm	10.0 mm	12.5 mm
0703T00	11	8	7	5	4	0	0
0703T01	18	11	8	6	3	0	0
0703T02	28	20	13	7	5	3	1
0703T03	33	22	12	7	2	1	1
0703T04	46	33	16	5	5	4	3
0703T05	57	44	31	17	12	10	6
0703T06	56	43	36	31	23	18	13
0703T07	59	47	41	32	25	22	20
0703T08	61	53	41	27	15	11	10
0703T09	61	55	28	12	7	6	5
0703T10	63	42	25	13	10	9	8
0703T11	62	39	21	8	7	3	1
0703T12	61	30	14	10	7	5	3
0703T13	60	34	19	13	8	8	6
0703T14	61	39	16	6	6	6	5
0703T15	62	45	19	13	9	5	5
0703T16	63	44	32	22	11	8	6
0703T17	63	60	53	43	31	25	14
0703T18	63	63	54	44	33	28	19
0703T19	62	51	37	31	25	19	17
0703T20	54	40	31	23	15	12	11

estimated values reach a certain threshold.  $N_{tot}$  is the sum of  $A$  from all sites, and  $N_{obs}$  is the sum of  $C$  from all sites.

### 3.3 Integration approach based on a statistical weight matrix

The optimal precipitation field is obtained from the weighted summation of different kinds of precipitation distribution fields; that is, the sum of the products of the estimation values of various methods and the weight coefficient at each grid. The basic formulation is defined as follows:

$$P_k = \sum_{i=1}^m R_{ik} W_{ik} \tag{6}$$

where  $P_k$  is the integrated optimal estimation value of different methods at the  $k^{th}$  grid point, and  $R_{ik}$  and  $W_{ik}$  denote the radar estimation value and the weight coefficient of grid  $k$  from the  $i^{th}$  approach, respectively.

In order to minimize arbitrariness in the general weighted scheme, the weighting size is determined according to the relative importance of multiple independent variables in the dependent variable. Assuming that there is a multivariate linear correlation between the observed rain gauge value and the estimated radar value, the equation set of  $n$  times can be written as follows:

$$\begin{cases} w_{1k}R_{11} + w_{2k}R_{12} + \dots + w_{mk}R_{1m} = G_1 \\ w_{1k}R_{21} + w_{2k}R_{22} + \dots + w_{mk}R_{2m} = G_2 \\ \vdots \\ w_{1k}R_{n1} + w_{2k}R_{n2} + \dots + w_{mk}R_{nm} = G_n \end{cases} \tag{7}$$

where  $w_{mk}$  is the weight coefficient of the grid  $k$  in the  $m^{th}$  approach;  $R_{nm}$  is the radar estimation value of the  $n^{th}$  time in the  $m^{th}$  approach;  $G_n$  is the observation value of the  $n^{th}$  time;  $w_{mk}$  can be calculated using Eq. (7) according to the least squares principle when the residual sum of squares between observation and the regression values reaches the minimum.

$$Q = \sum_{j=1}^n (G_j - \hat{G}_j)^2 = \sum_{j=1}^n [G_j - (w_{1k}R_{j1} + w_{2k}R_{j2} + \dots + w_{mk}R_{jm})]^2 \tag{8}$$

where because  $Q$  is a non-negative quadratic form of the weight coefficient, its minimum value must exist. According to the extremum principle,  $w_{mk}$  should satisfy the condition of  $\partial Q / \partial w_{ik} = 0$  when  $Q$  gets the extremum.

In order to examine the quality of the integrated data, the mean error ( $e$ ), the root mean square error of the mean error ( $\sigma_e$ ), correlation coefficient ( $\rho$ ), and mean relative error (RE) are used to evaluate the degree of relative influence that the

various methods have in the statistical weight matrix integration.

$$\begin{aligned}
 e &= \frac{1}{M} \sum_{j=1}^M \left[ \frac{1}{N} \sum_{i=1}^N (R_i - G_i) \right] \\
 \sigma_e &= \frac{1}{M} \sum_{j=1}^M = 1 \left\{ \left[ \frac{1}{N} \sum_{i=1}^N (e_i - e)^2 \right]^{\frac{1}{2}} \right\} \\
 \rho &= \frac{\frac{1}{N} \sum_{i=1}^N G_i R_i - \frac{1}{N} \sum_{i=1}^N G_i \frac{1}{N} \sum_{i=1}^N R_i}{\left\{ \left[ \frac{1}{N} \sum_{i=1}^N G_i^2 - \left( \frac{1}{N} \sum_{i=1}^N G_i \right)^2 \right] \left[ \frac{1}{N} \sum_{i=1}^N R_i^2 - \left( \frac{1}{N} \sum_{i=1}^N R_i \right)^2 \right] \right\}^{\frac{1}{2}}} \\
 RE &= \frac{1}{M} \sum_{j=1}^M \left( \left| \bar{R}_j - \bar{G}_j \right| / \bar{G}_j \right)
 \end{aligned}
 \tag{9}$$

where  $R_i$  and  $G_i$  are radar-estimated precipitation and rain gauge observations at the  $i^{\text{th}}$  station or grid point, respectively;  $N$  is the number of the rain gauges or grid points;  $M$  is the count of the durations;  $\bar{R}_j$  and  $\bar{G}_j$  are the mean areal precipitation values from radar estimations and gauge observations, respectively.

## 4 Results and discussion

### 4.1 The results of six radar and rain gauge merging methods

A local Z-R relationship is reconstructed using a genetic optimization algorithm, which is further used to estimate precipitation. On the basis of  $Z = 85R^{1.82}$ , the accuracy of QPE is evaluated via a contingency table of the six methods of radar-rain gauge merging. According to the principle of uniform sampling, 70 stations are used for radar and rain gauge merging, and the remaining 67 stations are used to evaluate the results. Seven rainfall intensity thresholds corresponding to 0.1, 1.0, 2.5, 5.0, 7.5, 10.0, and 12.5 mm/h are selected for the study region. For the threshold value of 0.1 mm/h, rainfall is evaluated when the number of stations is greater than 10. For example, consider the rainfall process on July 3, 2006: different time points and the number of stations exceeding the threshold from 00:00 to 20:00, Beijing time (BJT, namely UTC+8), are shown in Table 1. The results of  $A$ ,  $B$ ,  $C$ , and  $D$  can be found in Table 2. The statistical indices of  $B_S$ ,  $T_S$ ,  $B_p$ , and RMSE can be calculated from Table 1 and Table 2.

#### 4.1.1 Evaluation of the accuracy of radar quantitative precipitation estimates at site

$B_S$  and  $T_S$  reflect only the degree of deviation above a certain threshold. In order to quantitatively access the error, the  $B_p$  and RMSE are used to further evaluate the accuracy of radar-

**Table 2** The results of  $A$ ,  $B$ ,  $C$ , and  $D$  from the contingency table at 7:00 BJT, July 3, 2006

Threshold (mm)	Method	A	B	C	D	Method	A	B	C	D
0.1	Z-R	43	3	8	12	MFB	48	3	11	4
1.0		38	2	9	17		43	4	4	15
2.5		24	0	17	25		37	5	4	20
5.0		13	1	19	33		27	8	6	25
7.5		4	0	21	41		23	7	3	33
10.0		0	0	22	44		16	6	7	37
12.5		0	0	20	46		13	4	8	41
0.1	KLM	50	3	9	4	OPT	52	7	2	5
1.0		44	4	3	15		45	5	2	14
2.5		38	5	3	20		37	4	4	21
5.0		28	8	5	25		28	4	4	30
7.5		24	7	2	33		23	5	2	36
10.0		17	6	6	37		18	4	4	40
12.5		14	4	7	41		15	5	5	41
0.1	VAR	49	5	4	8	KOP	53	6	2	6
1.0		44	5	3	14		44	6	3	13
2.5		38	5	3	20		37	4	4	21
5.0		29	3	3	31		29	4	4	29
7.5		23	3	2	38		24	5	2	35
10.0		17	5	5	39		18	5	5	38
12.5		14	4	6	42		16	5	5	40
0.1	KVR	50	4	5	8					
1.0		43	6	4	13					
2.5		38	5	3	20					
5.0		29	4	3	30					
7.5		23	3	2	38					
10.0		18	5	4	39					
12.5		15	5	5	41					

estimated precipitation. The results of  $B_S$ ,  $T_S$ ,  $B_p$ , and RMSE for different thresholds are shown for the 20060703 rainfall process (Fig. 2).

Figure 2 shows that, of the seven methods of estimation, radar QPE based on the Z-R relationship has the poorest accuracy. Values from this method are clearly underestimated, especially for high-intensity rainfalls.  $T$ -evident errors can be attributed to the unchanged coefficient in the Z-R relationship. The accuracy of QPE is improved dramatically by radar-rain gauge merging. Figure 2 a and c show that  $B_S$  and  $B_p$  are approximately 1.0 as the threshold increases, which indicates that radar estimations from six merging methods are in good agreement with the rain gauge observations. Figure 2 b and d show that  $T_S$  and RMSE from all methods clearly decrease and increase with an increase of threshold value, which indicates that the radar estimation of the heavy precipitation center is greatly deviated, and the accuracy worsens with increasing threshold values—this is especially true for estimations from

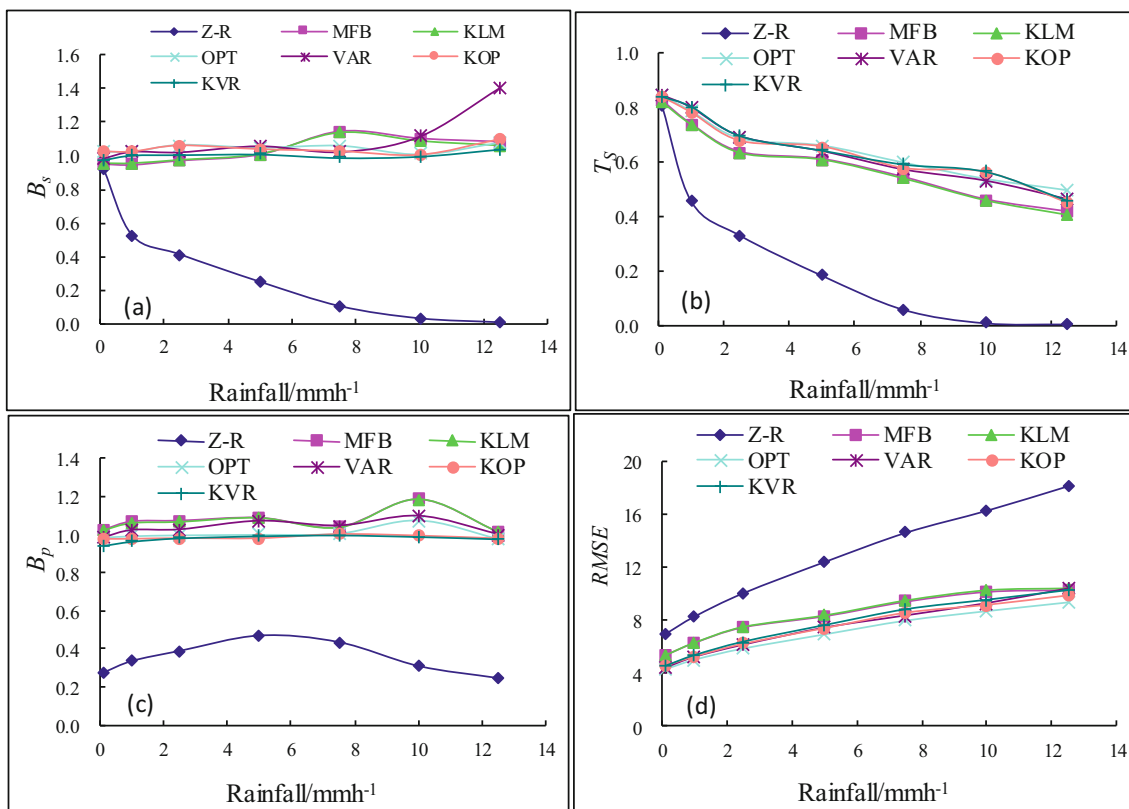


Fig. 2  $B_s$ ,  $T_s$ ,  $B_p$ , and RMSE averages over all time levels for different thresholds on July 3, 2006

the Z-R relationship. Among the six merging methods, the accuracies of KOP and KVR are slightly higher than those of OPT, VAR, KLM, and MFB. The results obtained from these different radar-rain gauge merging methods are consistent with the results of Li et al. (2014, 2015a, b). The main reason for this finding is that KOP and KVR fully consider the random characteristics of the spatial and temporal distributions of precipitation, which not only eliminate measurement noise but also highlight the precipitation structure; accordingly, precipitation accuracy is highest with these methods.

4.1.2 Evaluation of regional rainfall precision

Figure 3 shows the mean areal rainfall and relative errors from rain gauge fields and radar-estimated precipitation for six strong precipitation processes that occurred in 2006.

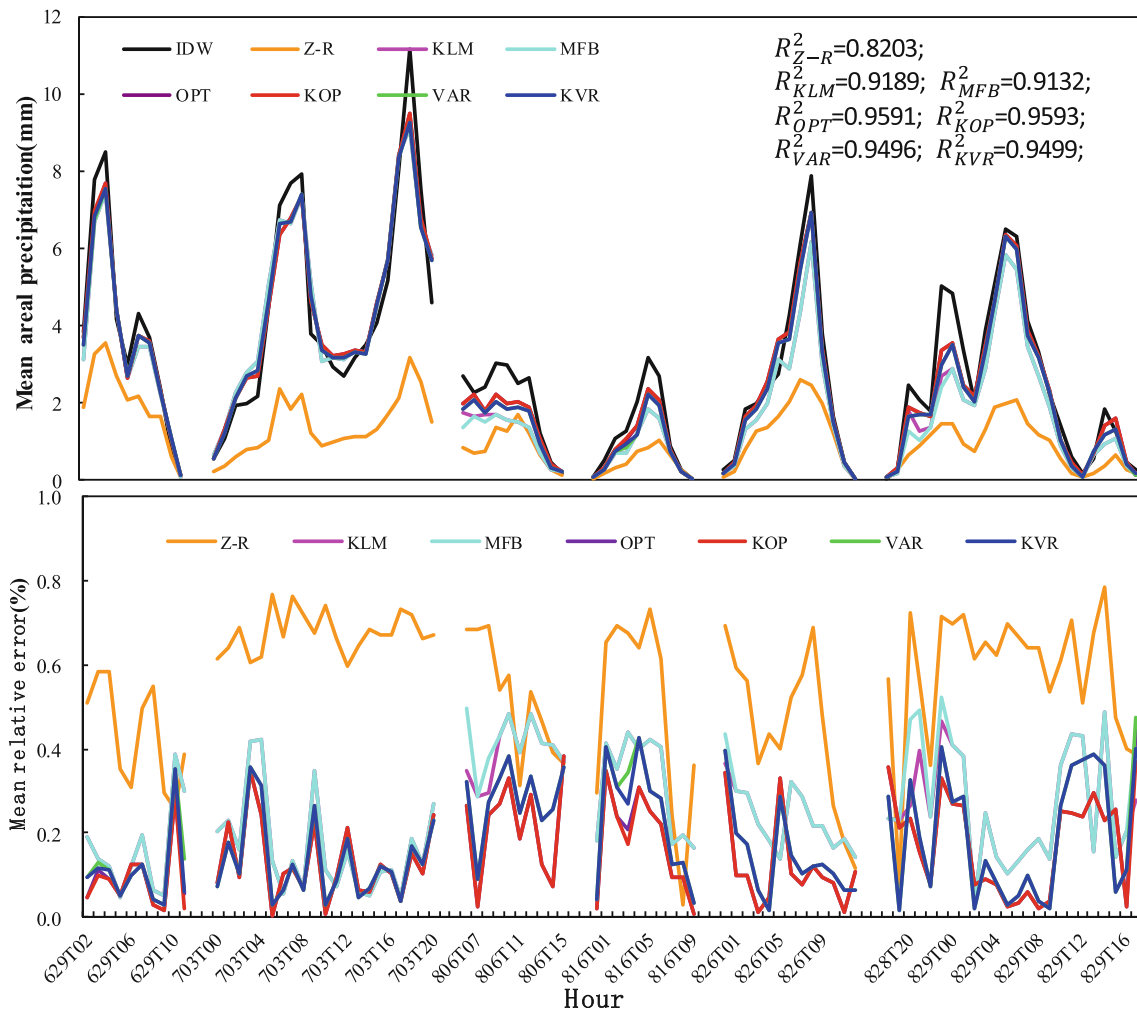
Figure 3 shows that the estimated mean regional rainfall based on the Z-R relationship is substantially lower than that from gauge observations, and the mean relative error reaches 55.2%, although there is a high correlation of  $R^2 = 0.8203$  for precipitation from Z-R relationship and IDW. In this case, underestimation may be attributed to an unchanged Z-R relationship being applied to all rainfall processes, radar parameters, distance attenuation, and disturbance from non-meteorological echoes from the ground over rugged terrain, etc. After merging radar and rain gauge data, the accuracy of

regional precipitation from radar estimation is improved dramatically and is in good agreement with the gauge network. This is because merging data take advantages of high temporal and spatial resolution of radar, and accurate, single-point gauge observations. The mean relative errors of MFB, KLM, OPT, VAR, KOP, and KVR are approximately 25.3%, 24.4%, 15.4%, 17.9%, 15.1%, and 17.7%, respectively. KOP performs the best and has the smallest mean relative error, generally around 10%, except for a few values exceeding 30%, which is in agreement with previous research (Chen et al. 2008; Li et al. 2009, 2015a, b). The accuracy and stability of radar QPE based on KOP method can be attributed to two main factors. One is the fact that KOP uses two-step calibration to eliminate the error source from the time and space domains. The other is that OPT has an advantage over VAR in the case of complex terrain and convective precipitation system. Therefore, KOP can accurately reflect precipitation status over the ground surface to some extent.

4.2 Results of integration based on the statistical weight matrix algorithm

4.2.1 Comparison of QPE before and after integration

The statistical features of  $e$ ,  $\sigma_e$ , and  $\rho$  from radar QPE are shown in Table 3. It can be seen that  $e$  and  $\sigma_e$  from the Z-R



**Fig. 3** Mean areal precipitation and relative error from six rainfall processes

relationship are the largest in both the statistical sample and test sample, which is clearly underestimated and is the worst-integrated analysis data among the methods considered. The precipitation estimates from the six merging methods are close to the gauge observations, but all of them are slightly higher

**Table 3** Statistical error features for rainfall estimation modes at gauge sites

Method	Statistical samples			Test samples		
	$e$	$\sigma_e$	$\rho$	$e$	$\sigma_e$	$\rho$
Z-R	-2.435	4.369	0.615	-2.746	4.478	0.618
MFB	0.302	4.292	0.664	0.334	4.375	0.646
KLM	0.280	4.290	0.662	0.328	4.370	0.644
OPT	0.106	4.121	0.706	0.144	4.156	0.682
VAR	0.175	4.184	0.682	0.185	4.258	0.678
KOP	0.088	4.126	0.698	0.128	4.137	0.692
KVR	0.143	4.168	0.692	0.169	4.269	0.687
INT	-0.128	4.085	0.712	-0.130	4.124	0.704

than the observation values. The error statistics indicate that OPT, VAR, KOP, and KVR have higher weights than MFB and KLM in the integration analysis. OPT and KOP have the lowest errors, and highest correlation coefficients, and are considered to be the best-integrated analysis data. After integration, the error is further reduced; however, no obvious improvement is observed. The main reason for this finding is that the integration method emphasizes the spatial distribution field of precipitation based on the weight coefficient matrix obtained by fitting of multiple linear regression equations.

Statistical characteristics of spatial rainfall fields from different estimation algorithms are presented in Table 4. Although the results of  $e$  and  $\sigma_e$  are similar to those in Table 3 for both the statistical and test samples, the correlation coefficient is clearly improved after integration analysis; this finding is further evident in Fig. 4, which shows the scatter plot of correlation coefficients. Correlations between rainfall from radar estimations and gauge observations between 06:00 and 07:00 on July 3, 2006, are shown in Fig. 4. Considering the MFB method with a high correlation coefficient, for



**Table 4** Statistical error features for rainfall estimate models at grid points

Method	Statistical samples				Test samples			
	$e$	$\sigma_e$	$\rho$	RE (%)	$e$	$\sigma_e$	$\rho$	RE (%)
Z-R	- 2.863	4.545	0.417	54.6	- 2.974	4.664	0.356	55.2
MFB	0.202	4.394	0.624	23.9	0.296	4.549	0.580	26.6
KLM	0.248	4.346	0.487	22.9	0.242	4.485	0.486	25.5
OPT	- 0.187	4.160	0.428	14.5	0.211	4.362	0.369	16.0
VAR	- 0.193	4.176	0.424	17.6	0.257	4.389	0.370	18.7
KOP	- 0.172	4.143	0.448	13.8	0.204	4.276	0.392	15.4
KVR	- 0.200	4.172	0.433	16.8	0.224	4.294	0.388	17.9
INT	0.256	2.171	0.824	12.6	0.321	2.470	0.805	14.4

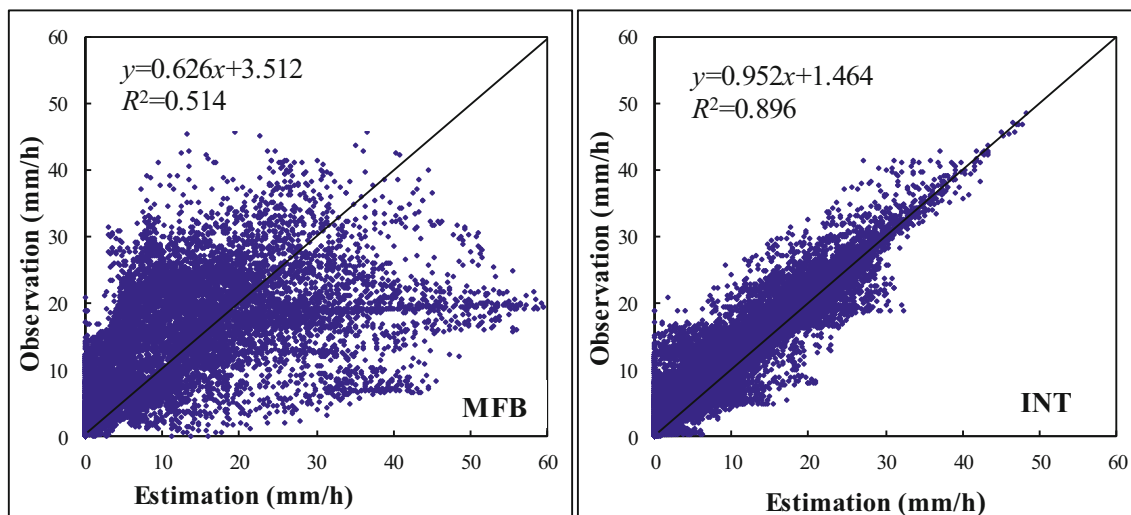
example, the scattered data are presented on both sides of the line  $y = x$ , which is indicative of a large degree of dispersion and a poor correlation. After statistical weight integration, the data points are uniformly distributed on both sides of the line  $y = x$ , and  $\rho$  is increased from 0.717 to 0.947, which indicates that the estimation is in good agreement with observation. The accuracy of precipitation estimated by the integration algorithm is higher than that noted before integration, which is consistent with or similar to previous research (Guan et al. 2004). However, Guan et al. (2004) failed to take into account the spatial heterogeneity of precipitation; that is, the weight coefficients across all grid points were identical for a certain mode.

**4.2.2 Analysis of regional precipitation using different methods**

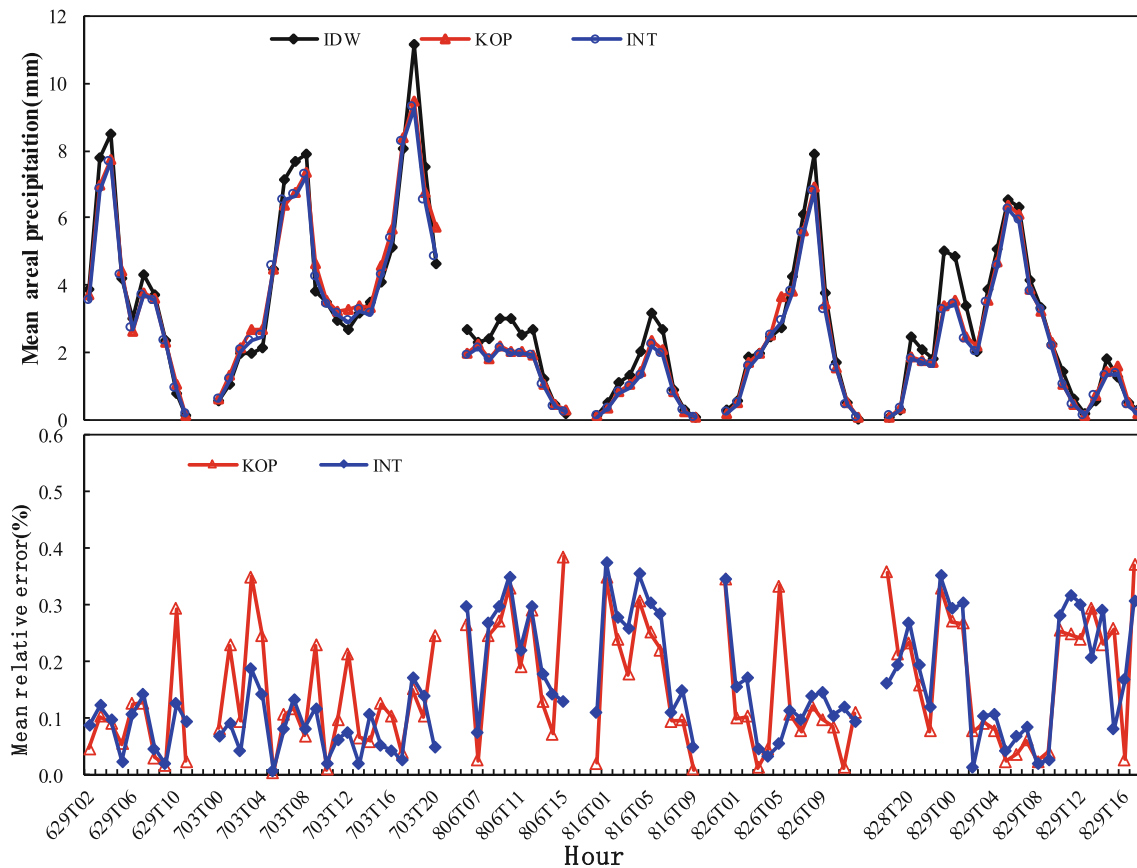
The mean areal rainfall and the relative error from the integration method are further shown in Fig. 5. The mean regional

rainfall matches well with the gauge observation values for all precipitation processes. The accuracy of regional precipitation is further improved after weight integration, and the mean relative error is reduced to 13.5%. In general, the relative error is larger in moderate rainfall processes (approximately 20%), than in the process of heavy rainfall (approximately 10%). Overall, the extremely large error values for all times are less than before integration; namely, RE volatility becomes smaller and more stable over time. The accuracy of precipitation estimates is improved by weight integration; that is, the spatial distribution field reflects the precipitation situation on the ground, and is in good agreement with the actual rain gauge network.

Spatial distribution fields based on the rain gauge network, radar estimation fields, and weight integration between 06:00 and 07:00 on July 3, 2006, are shown in Fig. 6. It is found that the spatial distributions of precipitation from all methods are largely in good agreement with those interpolated by precipitation measured by the rain gauge network. However,



**Fig. 4** Correlations between radar estimations and rain gauge observations made from 06:00 to 07:00 BJT using the MFB and INT methods, respectively



**Fig. 5** Mean areal precipitation and relative error from the KOP and INT methods

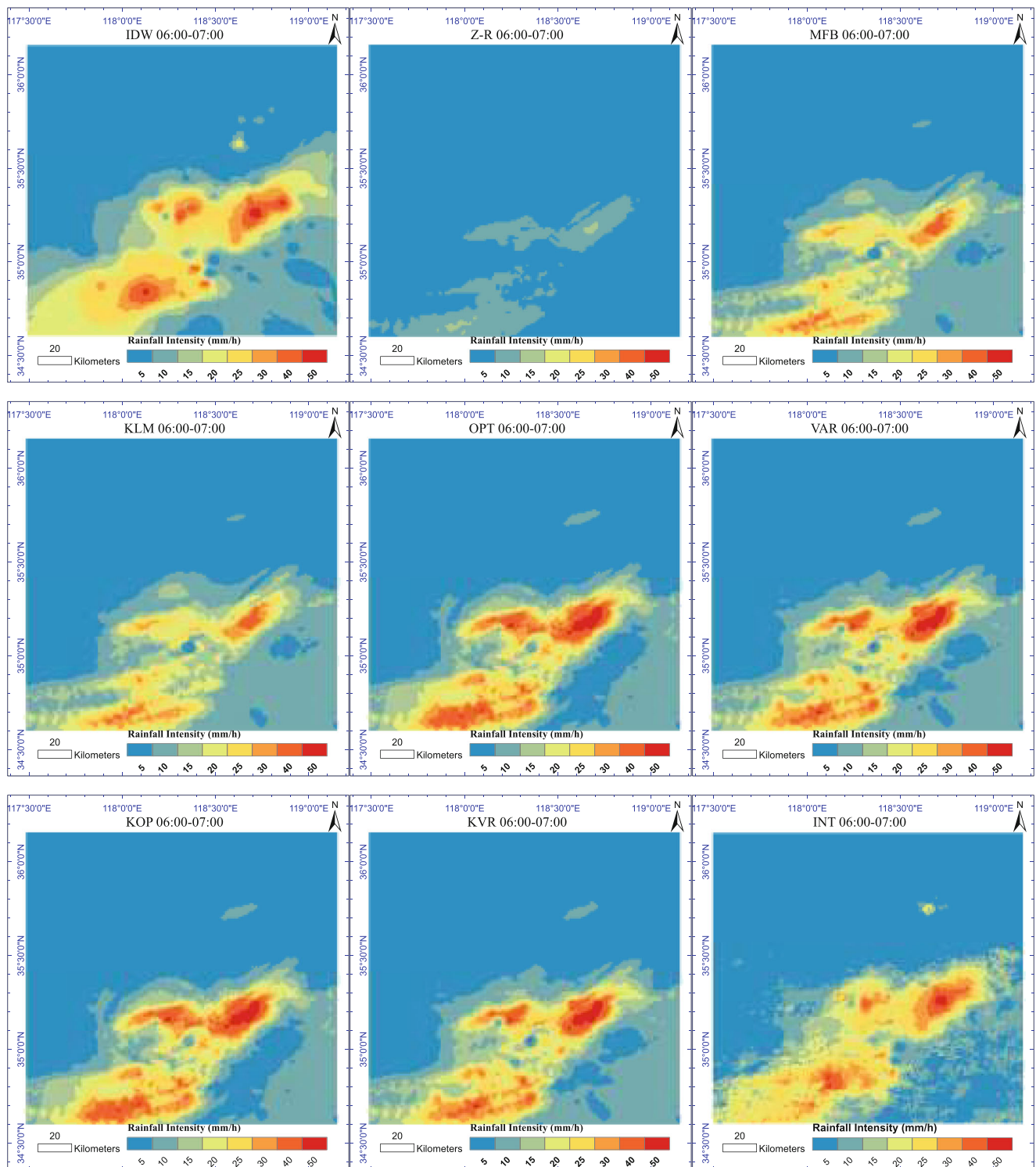
locations with centers of intense rainfall exhibit different patterns; in particular, precipitation estimated by the Z-R relationship is clearly underestimated. After merging radar and rain gauge data, the accuracy of radar QPE is improved dramatically, both in the spatial distribution and the location of intense precipitation centers, which shows good performance when compared with previous studies (Shao et al. 2008; Li et al. 2015a, b). The results indicate that it is reasonable to use radar to estimate precipitation patterns; however, calculations of regional precipitation may be subject to error to some extent. Generally, the OPT and KOP methods perform better than MFB, KLM, VAR, and KVR, both in terms of spatial patterns and intense precipitation centers. It can be also seen from Fig. 6 that the precipitation accuracy derived by integrating each of the abovementioned modes is higher than that obtained by any individual method, and is close to the observed values from the automatic rain gauge network in the spatial distribution and location of the intense precipitation centers. These results verify and validate the integrated ideology and theory outlined by Li et al. (2014), who proposed an integration algorithm based on critical probability. Regional precipitation estimated by an integration approach better represents precipitation status over the ground surface because it effectively reflects the inherent chaos and randomness of a rainfall system, and reduces the estimated uncertainty from observation and analysis

error. Therefore, it shows promise as a conventional method for estimating regional rainfall in the study region

## 5 Conclusions

The main objective of this research was to improve quantitative precipitation estimates by radar-rain gauge merging and integration based on a statistical weight matrix. The main conclusions from the results are as follows:

Results from analyses using a contingency table approach show that the accuracy of QPE based on the Z-R relationship is poor; the QPE is clearly underestimated, especially for high-intensity rainfalls. After merging the radar and rain gauge data, the QPE is largely in agreement with rain gauge observations. Among the six merging methods, the accuracies of the KOP, KVR, OPT, and VAR methods are slightly superior to those of KLM and MFB. The performance of the KOP method is maximal, because it fully considers random characteristics inherent to the spatial and temporal distributions of precipitation. For areal mean precipitation, the Z-R relationship yields the largest mean relative error of 55.2%, and is clearly underestimated. The accuracy of QPE is improved dramatically after radar-rain gauge merging as a result of the fact that the mean relative errors from MFB, KLM, OPT,



**Fig. 6** The spatial distribution of the rainfall field between 06:00 and 07:00 on July 3, 2006, as obtained from different methods

VAR, KOP, and KVR are reduced to 25.3%, 24.4%, 15.3%, 17.9%, 15.1%, and 17.7%, respectively. The KOP performs the best and the mean relative error is generally around 10%, which is more reflective of the true precipitation status over the ground surface.

Consideration of the statistical error features from each individual mode at sites and grid points reveals that the Z-R relationship and KOP are the worst- and best-integrated analysis data, respectively. The OPT, VAR, KOP, and KVR methods have higher weights

than MFB and KLM in the integration analysis. After integration, accuracy is further improved, especially the correlation coefficient of spatial grids, which is distinctly increased. For areal mean precipitation, the accuracy of QPE estimated by the integration approach is further improved and the mean relative error is reduced to 13.5%. At all times, the high errors are less than those before integration, which indicates that error fluctuations become small and stable over time. As a whole, the accuracy of QPE derived from integration of the statistical weight matrix is higher than that obtained by any individual mode.

The spatial distribution fields of precipitation from all methods are in good agreement with those interpolated by the rain gauge network. However, rainfall intensity in the center of rainstorm exhibited different patterns compared with the interpolated rain gauge network, such as an obvious underestimation from the Z-R relationship. The OPT and KOP methods perform better than the MFB, KLM, VAR, and KVR methods, both in terms of spatial patterns and intense precipitation centers. However, they are slightly overestimated in comparison with rain gauge observations. After undertaking the integration approach, the QPE is similar to values obtained from the automatic rain gauge network, both in terms of spatial distribution and in the location of intense precipitation centers. Thus, it better reflects the precipitation status over the ground surface and is a promising conventional method for QPE in the study region.

Accurately estimating quantitative precipitation directly effects forecasting precision and disaster assessment. The results of this research provide insight into methods for improving QPE in different basins in China. The main take-away from this study is that the integration method effectively reduces uncertainty in precipitation estimates caused by observation and analysis errors. The results of this study are valuable for obtaining high-precision QPE using radar-rain gauge data and an integration approach, and should be useful for further research focused on improving hydrometeorological coupling forecasts and early warnings of flash floods and debris flow.

**Acknowledgments** The authors are grateful to Linyi Meteorological Bureau of China for providing Doppler radar data and rain gauge data at sites. The authors would like to acknowledge the anonymous reviewers and the editor for their thoughtful comments and suggestions, which have greatly improved the presentation of this paper.

**Authors' contributions** All authors contributed substantially towards the success of this study. Dr. Shao plays a guiding role in the whole process as first author and corresponding author. She mainly takes charge of experiment design, data analysis, and manuscript writing of this research. Under the guidance of Dr. Shao, Aolin Fu revised the manuscript according to the reviewer's suggestions. For manuscript improvement, Dr. Zhao corrected errors in spelling, grammar, consistency, word choice, and sentence

clarity. Dr. Xu mainly finished the drawing of some figures in the manuscript. Junmei Wu mainly did the preliminary collation of the data.

**Funding** This work is financially supported by the Special Fund for Natural Science Foundation of Jiangsu province (BK20141001), by the Meteorological Open Research Fund in Huaihe River basin (HRM201702).

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interests.

## References

- Berndt C, Rabiei E, Haberlandt U (2014) Geostatistical merging of rain gauge and radar data for high temporal resolutions and various station density scenarios. *J Hydrol* 508:88–101
- Berne A, Krajewski WF (2013) Radar for hydrology: unfulfilled promise or unrecognized potential? *Adv Water Resour* 51:357–366
- Bianchi B, Jan van Leeuwen P, Hogan RJ, Berne A (2013) A variational approach to retrieve rain rate by combining information from rain gauges, radars, and microwave links. *J Hydrometeorol* 14:1897–1909
- Cantet P (2017) Mapping the mean monthly precipitation of a small island using kriging with external drifts. *Theoret Appl Climatol* 127(1–2):31–44
- Cecinati F, Moreno Ródenas M, Rico-Ramirez MA (2017) Integration of rain gauge errors in radar-rain gauge merging techniques. In: 10th World Congress on Water Resources and Environment, Athens, pp 279–285
- Chao L, Zhang K, Li Z, Zhu Y, Wang J, Yu Z (2018) Geographically weighted regression based methods for merging satellite and gauge precipitation. *J Hydrol* 558:275–289
- Chapon B, Delrieu G, Gosset M, Boudevillain B (2008) Variability of rain drop size distribution and its effect on the Z-R relationship: a case study for intense Mediterranean rainfall. *Atmos Res* 87:52–65
- Chen QP, Liu JX, Yu JH, Yang LZ, Xia WM (2008) Quantitative estimate of different sorts of precipitation with radar. *Meteorol Sci Technol* 36(2):233–236
- Chumchean S, Seed A, Sharma A (2006) Correcting of real-time radar rainfall bias using a Kalman filtering approach. *J Hydrol* 317:123–137
- Chumchean S, Sharma A, Seed A (2006b) An integrated approach to error correction for real-time radar-rainfall estimation. *J Atmos Ocean Technol* 23:67–79
- Cong F, Liu L (2011) A comprehensive analysis of data from the CINRAD and the ground rainfall station. *Meteorol Monogr* 37(5): 532–539
- Gao J, Xue M, Droegemeier KK (2004) A three-dimensional variational data analysis method with recursive filter for Doppler radars. *J Atmos Ocean Technol* 21:457–469
- Germann U, Galli G, Boscacci M, Bolliger M (2006) Radar precipitation measurement in a mountainous region. *Q J R Meteorol Soc* 132(618):1669–1692
- Gou YB, Liu L, Wang D, Zhong L, Chen C (2015) Evaluation and analysis of the Z-R storm-grouping relationships fitting scheme based on storm identification. *Torrential Rain Disasters* 34(01):1–8
- Goudenhoofd E, Delobbe L (2009) Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrol Earth Syst Sci* 13(2):195–203

- Guan L, Wang ZH, Pei XF (2004) The consensus methods and effect of estimating rainfall using radar. *J Meteorol Sci* 24(1):104–111
- Haberlandt U (2007) Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event. *J Hydrol* 332:144–157
- Hasan MM, Sharma A, Mariethoz G, Johnson F, Seed A (2016) Improving radar rainfall estimation by merging point rainfall measurements within a model combination framework. *Adv Water Resour* 97:205–218
- He X, Sonnenborg TO, Refsgaard JC, Vejen F, Jensen KH (2013) Evaluation of the value of radar QPE data and rain gauge data for hydrological modeling. *Water Resour Res* 49(9):5989–6005
- Jacobi S, Heistermann M (2016) Benchmarking attenuation correction procedures for six years of single-polarized C-band weather radar observations in South-West Germany. *Geomatics, Nat. Hazards Risk* 7:1785–1799
- Jonkman SN (2005) Global perspectives on loss of human life caused by floods. *Nat Hazards* 34:151–175
- Kim J, Yoo C (2014a) Using extended Kalman filter for real-time decision of parameters of Z-R relationship. *J Korea Water Resour Assoc* 47(2):119–133
- Kim J, Yoo C (2014b) Use of a dual Kalman filter for real-time correction of mean field bias of radar rain rate. *J Hydrol* 519:2785–2796
- Krajewski WF (1987) Cokriging radar-rainfall and rain gauge data. *J Geophys Res:Atmos* 92:9571–9580
- Ku JM, Ro Y, Kim K, Yoo C (2015) Analysis on characteristics of orographic effect about the rainfall using radar data: a case study on Chungju Dam basin. *J Korea Water Resour Assoc* 48(5):393–407
- Lafont D, Guillemet B (2004) Subpixel fractional cloud cover and inhomogeneity effects on microwave beam-filling error. *Atmos Res* 72(1):149–168
- Lee J, Byun H, Kim H, Jun H (2013) Evaluation of a raingauge network considering the spatial distribution characteristics and entropy: a case study of Imha dam basin. *J Korean Soc Hazard Mitig* 13(2):217–226
- Li JT, Gao ST, Guo L, Liu XY, Yang HP, Cai YY (2009) The two-step calibrate technique of estimating areal rainfall. *Chin J Atmos Sci* 33(3):501–512
- Li JT, Li B, Yang HP, Liu XY, Zhang L, Guo L (2014) A study of regional rainfall estimation by using radar and rain gauge: proposal of model integration method. *Meteorol Sci Technol* 42(4):556–562
- Li JT, Li B, Yang HP, Liu XY, Zhang L, Guo L (2015a) Verification and assessment of regional rainfall estimation by using radar and rain-gauge. *Meteorol Mon* 42(2):200–211
- Li H, Hong Y, Xie PP, Gao JD, Niu Z, Kirstetter P, Yong B (2015b) Variational merged of hourly gauge-satellite precipitation in China: preliminary results. *J Geophys Res Atmos* 120:9897–9915
- Maggioni V, Massari C (2018) On the performance of satellite precipitation products in riverine flood modeling: a review. *J Hydrol* 558:214–224
- Marshall JS, Palmer WMK (1948) The distribution of raindrops with size. *J Meteorol* 5:165–166
- Martens B, Cabus P, De Jongh I, Verhoest NEC (2013) Merging weather radar observations with ground-based measurements of rainfall using an adaptive multi-quadric surface fitting algorithm. *J Hydrol* 500(3–4):84–96
- Ochoa-Rodriguez S, Wang LP, Willems P, Onof C (2019) A review of radar-rain gauge data merging methods and their potential for urban hydrological applications. *Water Resour Res* 55(8):6356–6391
- Rabiei E, Haberlandt U (2015) Applying bias correction for merging rain gauge and radar data. *J Hydrol* 522:544–557
- Seo DJ, Breidenbach JP, Johnson ER (1999) Real-time estimation of mean field bias in radar rainfall data. *J Hydrol* 223:131–147
- Shao YH (2010) Precipitation retrieved by Doppler radar and its assimilation study with the improved regional climate model RIEMS. Nanjing University, Nanjing
- Shao YH, Zhang WC, Liu YH (2008) Analysis of quantitative precipitation estimation with different methods by using Doppler radar data. *Int Workshop Geosci Remote Sens Symp* 2:21–22
- Sharifi E, Steinacker R, Saghafian B (2018) Multi time-scale evaluation of high resolution satellite-based precipitation products over north-east of Austria. *Atmos Res* 206:46–63
- Sideris IV, Gabella M, Erdin R, Germann U (2014) Real-time radar-rain gauge merging using spatio-temporal co-kriging with external drift in the alpine terrain of Switzerland. *Q J R Meteorol Soc* 140:1097–1111
- Sinclair S, Pegram G (2005) Combining radar and rain gauge rainfall estimates using conditional merging. *Atmos Sci Lett* 6:19–22
- Smith JA, Krajewski WF (1991) Estimation of the mean field bias of radar rainfall estimates. *J Appl Meteorol* 30:397–412
- Sun SX, Liu GQ, Ge WZ (1993) A method of variational analysis combined with Kalman filter for radar rainfall field correction. 26<sup>th</sup> international conference on radar meteorology, Orman, Amer, Meteor. Soc. 755–757
- Victor H, Alvarez VH, Aznar M (2010) An efficient approach to optimal interpolation of experimental data. *J Taiwan Inst Chem Eng* 41(2):184–189
- Villarini G, Krajewski WF (2010) Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surv Geophys* 31:107–129
- Wang GL, Liu LP, Ding YY (2012) Improvement of radar quantitative precipitation estimation based on real-time adjustments to Z-R relationships and inverse distance weighting correction schemes. *Adv Atmos Sci* 29(3):575–584
- Wang LP, Ochoa-Rodriguez S, Simões N, Onof C, Maksimovic Č (2013) Radar-rain gauge data combination techniques: a revision and analysis of their suitability for urban hydrology. *Water Sci Technol* 68:737–747
- Wang HY, Wang GL, Liu LP, Jiang Y, Wang D, Li F (2015) Development of a real-time quality control method for automatic rain gauge data using radar quantitative precipitation estimation. *Chin J Atmos Sci* 39(1):59–67
- Wu MC, Lin GF, Wang LR (2016) Optimal integration of the ensemble forecasts from an ensemble quantitative precipitation forecast experiment. *Procedia Eng* 154:1291–1297
- Yoo C, Park J (2008) Combining radar and rain gauge observations utilizing Gaussian process based regression and support vector learning. *J Korean Inst Intel Syst* 18(3):297–305
- Zhang PC, Dai TP, Fu DS, Wu ZF (1992) Principle and accuracy of adjusting the area precipitation from digital weather radar through variational method. *Chin J Atmos Sci* 16(2):248–256
- Zhang J, Howard K, Langston C (2016) Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: initial operating capabilities. *Bull Am Meteorol Soc* 97:621–638
- Zhao K, Liu GQ, Ge WZ (2001) Precipitation calibration by using Kalman filter to determine the coefficients of the variational equation. *Clim Environ Res* 6(2):180–185