

Research Article

Estimation of Areal Mean Rainfall in Remote Areas Using B-SHADE Model

Tao Zhang,^{1,2} Baolin Li,^{1,3} Jinfeng Wang,¹ Maogui Hu,¹ and Lili Xu^{1,2}

¹State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

²University of Chinese Academy of Sciences, Beijing 100049, China

³Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China

Correspondence should be addressed to Baolin Li; libl@lreis.ac.cn

Received 3 August 2015; Revised 2 November 2015; Accepted 1 December 2015

Academic Editor: Brian R. Nelson

Copyright © 2016 Tao Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This study presented a method to estimate areal mean rainfall (AMR) using a Biased Sentinel Hospital Based Area Disease Estimation (B-SHADE) model, together with biased rain gauge observations and Tropical Rainfall Measuring Mission (TRMM) data, for remote areas with a sparse and uneven distribution of rain gauges. Based on the B-SHADE model, the best linear unbiased estimation of AMR could be obtained. A case study was conducted for the Three-River Headwaters region in the Tibetan Plateau of China, and its performance was compared with traditional methods. The results indicated that B-SHADE obtained the least estimation biases, with a mean error and root mean square error of -0.63 and 3.48 mm, respectively. For the traditional methods including arithmetic average, Thiessen polygon, and ordinary kriging, the mean errors were 7.11 , -1.43 , and 2.89 mm, which were up to 1027.1% , 127.0% , and 358.3% , respectively, greater than for the B-SHADE model. The root mean square errors were 10.31 , 4.02 , and 6.27 mm, which were up to 196.1% , 15.5% , and 80.0% , respectively, higher than for the B-SHADE model. The proposed technique can be used to extend the AMR record to the presatellite observation period, when only the gauge data are available.

1. Introduction

Accurate estimation of areal mean rainfall (AMR) is necessary for water balance calculations, climate studies, flood modeling, and runoff forecasting [1]. Currently, rainfall data are derived mainly from three principal sources: (1) rain gauges; (2) weather radar; (3) and satellites [2–5]. However, the series of rainfall data based on weather radar and satellites are short, and therefore, to obtain the longest data series possible, AMR is mainly determined using rain gauge data.

Methods commonly used for the estimation of AMR can be classified into two categories. The first involves the calculation of the average or weighted average of rain gauge data, for example, using an arithmetic average method or Thiessen polygon technique [1], and the second requires the acquisition of the average precipitation for all pixels of a grid map using spatial interpolation methods such as one of the kriging techniques [5].

The arithmetic average method assumes that the rainfall field is homogeneous and that the rain gauge observations are independent; thus, it gives equal weight to all rain gauges [1, 6]. The Thiessen polygon technique sets a fixed weighting factor for the areal coverage of a station. It also assumes that the rainfall field is homogeneous within the area of coverage of each station [1, 7]. These assumptions are really only reasonable in primarily low-lying or flat terrain and/or where rainfall events are typically extensive and homogeneous (e.g., widespread frontal storms) [5].

Kriging methods (e.g., ordinary kriging, universal kriging, and cokriging) are known as geostatistical approaches or multivariate geostatistical techniques with spatial correlation patterns [6, 8–11]. Previous case studies have confirmed their effectiveness in the spatial interpolation of rainfall within mountainous terrain [11–14]. However, their estimation accuracy is influenced considerably by the density and distribution of the rain gauge network [15–17]. Generally, when

TABLE 1: Details of meteorological stations within the Three-River Headwaters region.

Station ID	Station names	Latitude (°)	Longitude (°)	Elevation (m)	Annual mean precipitation for 1998–2011 (mm)
52908	Wudaoliang	35.217	93.083	4612.2	332.6
52943	Xinghai	35.583	99.983	3323.2	392.4
56004	Tuotuohe	34.217	92.433	4533.1	339.5
56018	Zaduo	32.900	95.300	4066.4	543.9
56021	Qumalai	34.133	95.783	4175.0	439.7
56029	Yushu	33.017	97.017	3681.2	485.5
56033	Maduo	34.917	98.217	4272.3	344.5
56034	Qingshuihe	33.800	97.133	4415.4	520.0
56043	Guoluo	34.467	100.250	3719.0	513.0
56046	Dari	33.750	99.650	3967.5	564.4
56065	Henan	34.733	101.600	3519.0	574.5
56067	Jiuzhi	33.433	101.483	3628.5	728.2
56125	Nangqian	32.200	96.483	3643.7	574.6
56151	Banma	32.933	100.750	3612.0	655.9

the number of stations is sparse and/or their distribution is uneven, kriging methods may obtain biased estimations when lacking additional information from covariates (e.g., elevation).

Covariates, for example, latitude, longitude, and elevation are commonly used in kriging techniques [9–11]. They are also used in other interpolation methods, for example, regression techniques [18]. Many published high resolution precipitation climatologies, such as Parameter Regression on Independent Slopes Model (PRISM) and WorldClim, were developed by integrating explanation variables in their algorithms [18, 19]. However, uncertainty may also be the highest in mountainous and in poorly sampled areas [19]. For example, in a simple linear climate-elevation regression technique [18], available rain gauges often do not span the whole range of elevations in a region, especially in mountainous areas where most of the stations are located in river valleys or plains, which may introduce much uncertainty for spatial extrapolation. Satellite products, such as radar and Tropical Rainfall Measuring Mission (TRMM) data, are increasingly used as covariates in interpolation methods [20–23]. Satellite data are very promising, as they provide detailed spatial distribution information of rainfall, and can be transferred to different regions, especially in data scarce areas [23, 24].

In western China, difficulty of access and low population densities result in poor spatial coverage of rain gauges, particularly in the Gobi Desert or high mountains and plateaus. Thus, the distribution of rain gauge stations is sparse and uneven. Most gauges are located in river basins or valleys; however, rainfall varies considerably with elevation within mountainous areas [25]. Therefore, the estimated AMR will be biased when using the arithmetic average, Thiessen polygon, or kriging techniques. From the viewpoint of spatial sampling and statistical inference, the problem when estimating AMR using any of the abovementioned methods is that the biased samples cannot indicate the total population precisely. Thus, a proper estimator that is capable

of obtaining an unbiased estimation of the population from the biased samples is required.

The Biased Sentinel Hospital based Area Disease Estimation (B-SHADE) model, which was developed to correct biased samples, held the ability to obtain the best linear unbiased estimation of the total (or mean) population [26–29]. It considers both the correlation between each sample and the population and the dependence between each pair of samples based on historical satellite data. Thus, the accurate AMR may be estimated using biased rain gauge observations in combination with satellite data based on this model. A case study was implemented by comparison with traditional methods including the arithmetic average, Thiessen polygon, and ordinary kriging. Such a technique can improve the accuracies of AMR estimation in remote areas where the rain gauge distribution is sparse and uneven and also extend the AMR record to the presatellite observation period, when only the gauge data are available.

2. Study Area and Data

2.1. Study Area. The Three-River Headwaters region was selected as the study area. This is located in the middle of the Tibetan Plateau, south of Qinghai Province in China, and it encompasses an area of about 360,000 km². It belongs to the plateau continental climate region. From northwest to southeast, the annual average temperature increases from -5.6 to 3.8°C ; annual average precipitation increases from 262.2 to 772.8 mm, and the climatic zones change from semiarid to humid subtropical [30]. In addition, there are many huge mountains in the northwest of the region, where glaciers cover some areas [31].

There are only 14 available meteorological stations within this area, of which just two are located in the west. Most of the stations are located in river valleys (Figure 1), and their elevations range from 3323.2 to 4612.2 m a.s.l.; there are no stations higher than 5000 m a.s.l. (Table 1). Such a sparse and

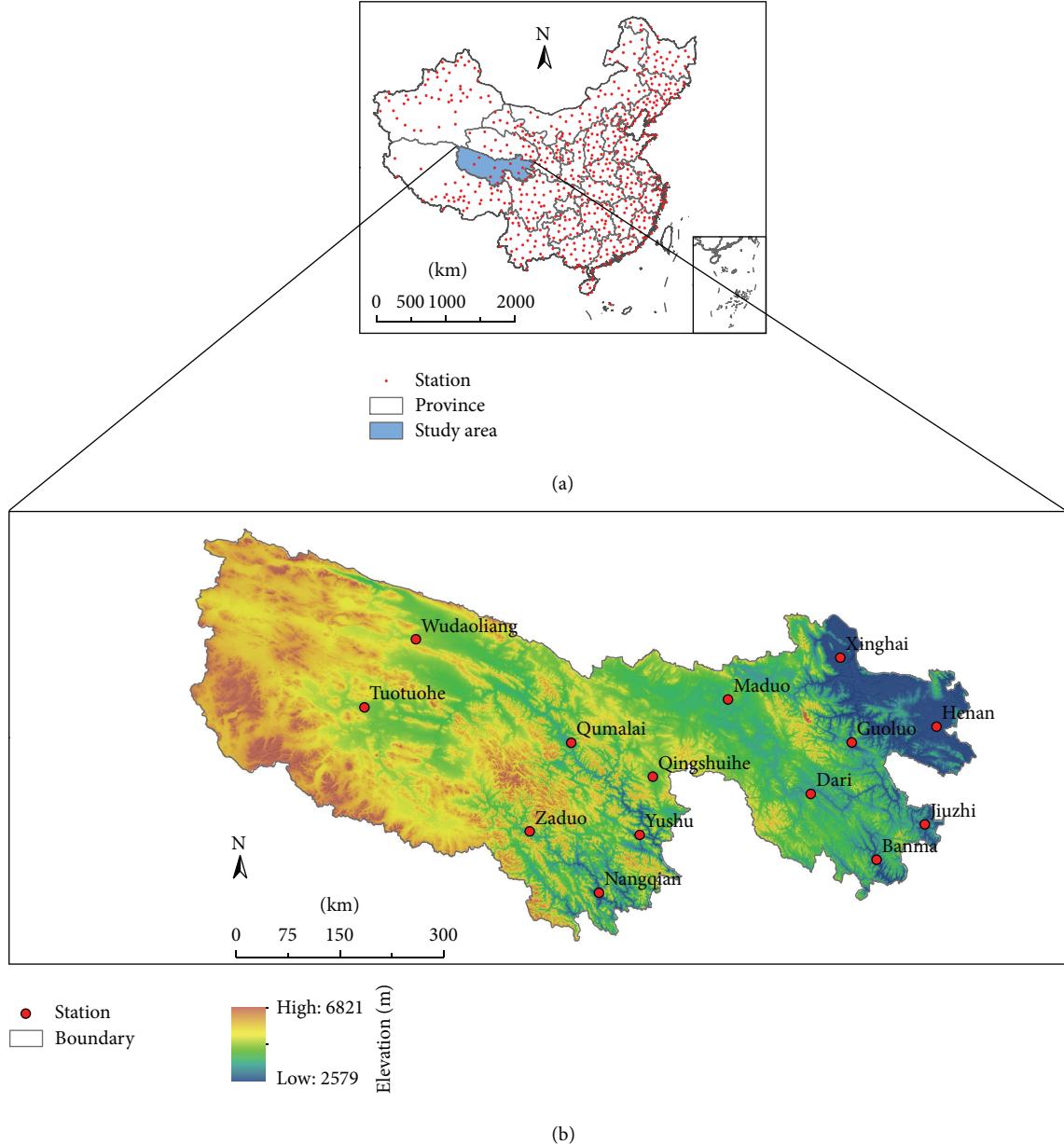


FIGURE 1: (a) Location of study area and positions of all meteorological stations in China; (b) topography and distribution of meteorological stations available within the Three-River Headwaters region.

uneven distribution of meteorological stations is typical of western China (Figure 1).

2.2. Data

2.2.1. Rain Gauge Observations. Rain gauge data were downloaded from the China Meteorological Data Sharing Service System (<http://cdc.nmic.cn>). For the study period (January 1998 to December 2011), complete monthly rainfall datasets were acquired for 649 stations (Figure 1), 14 of which were located within the Three-River Headwaters region. The annual mean precipitation within the study area from 1998 to 2011 was between 332.6 and 728.2 mm (Table 1).

2.2.2. TRMM Rainfall Estimates. For this study, we choose TRMM 3B43 (Version 7) monthly precipitation data from 1998 to 2011. This dataset was acquired from the National Aeronautics and Space Administration (NASA) website (<http://trmm.gsfc.nasa.gov/>). It has spatial resolution of $0.25^\circ \times 0.25^\circ$ [32, 33], which is selected as the research scale of this study; namely, the bias and comparison statistics used in the development, application, and validation of the techniques are calculated at the TRMM resolution of $0.25\text{ deg lat/lon grid}$. It has been validated that the Version 7 TRMM 3B43 data have better accuracy than previous versions [34, 35]. Figure 2 shows the spatial distribution of TRMM-estimated annual mean precipitation from 1998 to

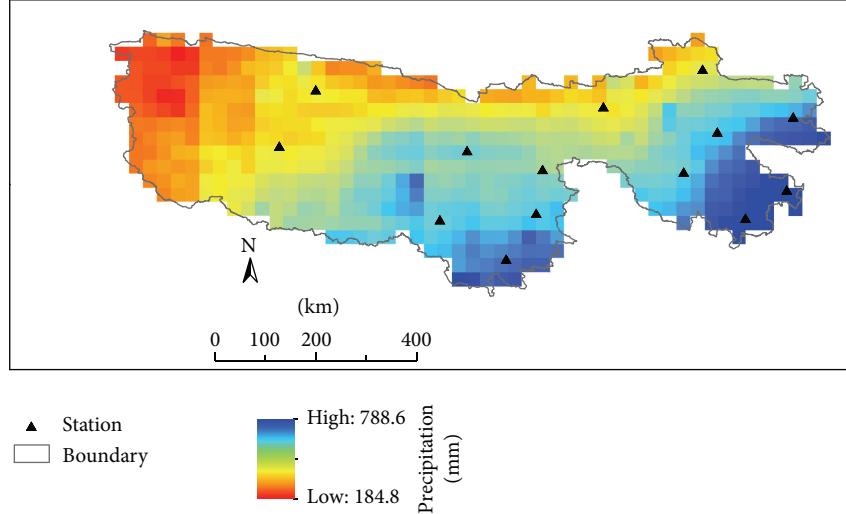


FIGURE 2: Spatial distribution of TRMM-estimated annual mean precipitation from 1998 to 2011 in the Three-River Headwaters region.

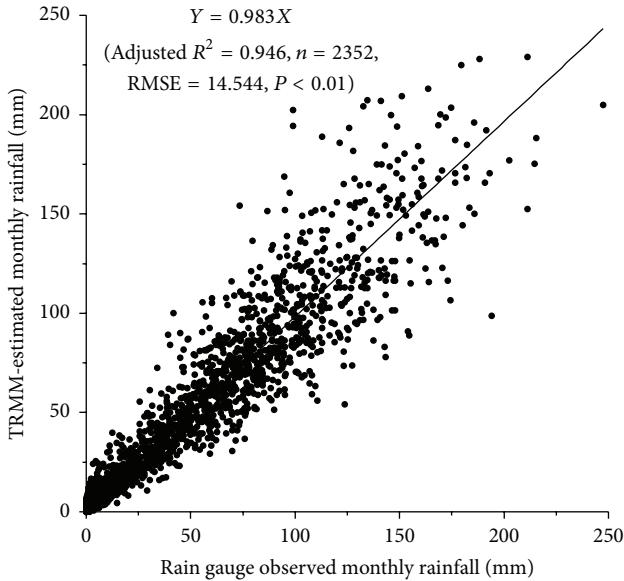


FIGURE 3: Linear regression of monthly precipitation from TRMM 3B43 and rain gauges for 1998–2011.

2011 in the Three-River Headwaters region. The annual mean precipitation for 1998–2011 ranged from 184.8 to 788.6 mm for TRMM data (Figure 2), while it ranged from 332.6 to 728.2 mm for rain gauges (Table 1); that is, precipitation range of TRMM is up to 52.6% higher than that from rain gauge observations.

2.2.3. Relationship between TRMM Data and Rain Gauge Records. The TRMM data indicated high accuracy with respect to individual rain gauge observations. A scatter diagram of monthly TRMM-estimated rainfall against rain gauge observations from January 1998 to December 2011 demonstrates good linear correlation (Figure 3). Setting rain gauge precipitation as the independent variable and

TRMM-estimated rainfall as the dependent variable, the linear regression through the origin was performed using SPSS 19 software. The results show that the slope of their linear regression function is 0.983 and the adjusted determination coefficient (R^2) is 0.946 ($P < 0.01, n = 2352$).

For each station, the TRMM-estimated rainfall showed a good linear relationship with the rain gauge observations. For each of the 14 stations within the study area, the linear regression through the origin was also performed between the monthly TRMM-estimated rainfall and rain gauge observations from 1998 and 2011. It is shown in Table 2 that the slopes of the regression functions range from 0.809 to 1.141; the adjusted determination coefficients (R^2) are all >0.92 , and all of the regression functions passed the significance test at the 0.01 level.

3. Methodology

3.1. The Principle of the B-SHADE Model. The B-SHADE model is a model-assisted and data-driven method used for deviation correction of population estimation using biased samples [26]. Hu et al. [28] extended the original technique from total population estimation to mean population evaluation. They also clarified that historical unbiased samples (instead of the historical population) were effective when estimating the population mean. To achieve an unbiased estimation of a population using biased samples, the model takes into account both the “vertical” correlation and the “horizontal” dependence between the samples and the population. The “vertical” relevance between each sample and the population, which is represented by their ratio, is used to correct the sampling bias. The “horizontal” interaction between each pair of samples, which is the covariance between them, is used to increase the effectiveness of the sample size and reduce estimation variance [28]. Moreover, to guarantee a best linear unbiased estimation, each sample is allocated a unique weight. Then, the mathematical expectation of

TABLE 2: Linear regression of precipitation values of stations from TRMM 3B43 (Y) and rain gauges (X) for 1998–2011 in the Three-River Headwaters region.

Station	Regression equation	Adjusted coefficient of determination (adjusted R^2)	RMSE (mm)	P-level
Wudaoliang	$Y = 0.998X$	0.961	9.104	0.01
Xinghai	$Y = 0.862X$	0.921	12.713	0.01
Tuotuohe	$Y = 1.029X$	0.940	12.009	0.01
Zaduo	$Y = 0.908X$	0.963	11.775	0.01
Qumalai	$Y = 1.141X$	0.940	15.739	0.01
Yushu	$Y = 1.065X$	0.949	14.279	0.01
Maduo	$Y = 0.933X$	0.953	8.844	0.01
Qingshuihe	$Y = 0.809X$	0.936	13.090	0.01
Guoluo	$Y = 1.001X$	0.952	14.067	0.01
Dari	$Y = 0.911X$	0.966	11.455	0.01
Henan	$Y = 1.087X$	0.963	14.432	0.01
Jiuzhi	$Y = 1.015X$	0.968	15.195	0.01
Nangqian	$Y = 0.957X$	0.932	18.755	0.01
Banma	$Y = 1.020X$	0.959	15.820	0.01

the population estimation should be equal to the true value of population, and the variance of the estimator should be minimized [26, 28].

As important input data for driving the B-SHADE model, a time series of historical unbiased sample set or population (historical data) is used for model construction and weights calculation. The estimated population value can be finally calculated by weighting sentinel samples. Being driven by the historical data, the B-SHADE model can be used to choose the best combination of sentinel sites according to the minimum estimation variance of the population, when the number of samples is given, or to calculate the smallest number of samples needed for a given estimation variance [28].

3.2. Objective. The objective is to calculate the best linear unbiased estimation of AMR based on rain gauge observations. The discrete format true value of precipitation is expressed as

$$\bar{Y} = \frac{1}{N} \sum_{k=1}^N Y_k, \quad (1)$$

where Y_k is the value of the k th grid of TRMM gridded product and N is the total number of grids within the area. The estimated AMR is

$$\bar{y} = \sum_{i=1}^n w_i y_i, \quad (2)$$

where y_i is the gauge (or alternatively corresponding TRMM grid) value of the i th station, w_i is the weight of the i th station, and n is the number of stations ($n < N$). Accurate estimation of areal average precipitation in (2) should address two terms: the unbiased estimation condition

$$E(\bar{Y}) = E(\bar{y}) \quad (3)$$

and the minimum estimation variance

$$\min \left[\sigma_y^2 = E(\bar{y} - \bar{Y})^2 \right], \quad (4)$$

where E is the statistical expectation and σ_y^2 is the variance of \bar{y} . Considering (2), (3) can be written as

$$E(\bar{Y}) = E\left(\sum_{i=1}^n w_i y_i\right). \quad (5)$$

The relationship between the i th station and AMR is expressed as

$$b_i E(\bar{Y}) = E(y_i), \quad (6)$$

where b_i is the bias of the i th station with respect to the areal mean \bar{Y} . Considering (5), (6) can be written as

$$\sum_{i=1}^n w_i b_i = 1. \quad (7)$$

3.3. Estimation of Weights. The objective can be finally converted into solving the combined equations (8) and (7) using the least-squares method:

$$\mathbf{K} \cdot \mathbf{W} = \mathbf{M}, \quad (8)$$

where

$$\mathbf{K} = \begin{bmatrix} C(y_1, y_1) & \cdots & C(y_1, y_n) & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ C(y_n, y_1) & \cdots & C(y_n, y_n) & b_n \\ b_1 & \cdots & b_n & 0 \end{bmatrix},$$

$$\mathbf{W} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \\ \mu \end{bmatrix}, \quad (9)$$

$$\mathbf{M} = \begin{bmatrix} C(y_1, \bar{Y}) \\ \vdots \\ C(y_n, \bar{Y}) \\ 1 \end{bmatrix};$$

$C(y_i, \bar{Y}) = (1/N') \sum_{j=1}^{N'} C(y_i, y_j)$ ($i = 1, \dots, n$) is the covariance between the i th station and the areal average value. N' is the number of unbiased historical samples of the population, and $C(y_i, y_j)$ is the covariance between the i th and j th stations, both \bar{Y} and $C(y_i, y_j)$ are acquired based on historical TRMM-derived AMR data because of the high accuracy of the TRMM data, as discussed in Section 2.2.3, and μ is a Lagrange multiplier.

The weights can be calculated by

$$\mathbf{W} = \mathbf{K}^{-1} \cdot \mathbf{M}. \quad (10)$$

The theoretical variance can be calculated using

$$\sigma_y^2 = (r_n - 1) \sum_{i=1}^n \sum_{j=1}^n w_i w_j C(y_i, y_j) - 2\mu, \quad (11)$$

$$r_n = \frac{(1/N'^2) \sum_{i=1}^{N'} \sum_{j=1}^{N'} C(y_i, y_j)}{\sum_{i=1}^n \sum_{j=1}^n w_i w_j C(y_i, y_j)},$$

where r_n is also evaluated using historical data [28].

After the weights of the rain gauges have been determined using (10), the AMR can be computed using (2). Even though the rain gauge observations are accurate, the discrepancy in scale between a gauge (point) and its corresponding TRMM grid (0.25 deg lat/lon) still exists, which will introduce potential uncertainties if the gauge values are used in calculating covariances. At the same time, as the TRMM data indicated high accuracy with respect to individual rain gauge observations, the TRMM values are used for covariance calculation to keep the consistency of data scale.

3.4. Validation. To obtain the AMR estimation, both \bar{Y} and $C(y_i, y_j)$ in (8) should be determined based on historical TRMM-derived AMR data. In this study, we classified the monthly TRMM data into two intervals: January 1998 to

December 2004 and January 2005 to December 2011. The data from the first interval were used to compute the weights that the B-SHADE model required, and the data from the second interval were used to compute the AMR as validation.

To evaluate the effectiveness of the AMR estimation using the B-SHADE model, three commonly used methods (arithmetic average, Thiessen polygon, and ordinary kriging) were chosen for comparison. The B-SHADE model was implemented using a software package which was developed by Hu et al. [28], and Thiessen polygon and ordinary kriging were carried out in ArcGIS 10.0 platform. The mean error (ME), root mean square error (RMSE), and mean absolute percent error (MAPE) were calculated using

$$\text{ME} = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i),$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \quad (12)$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100,$$

where y_i and \hat{y}_i are the true and estimated AMR of the i th month, respectively, and n represents the number of months in the second 7-year study interval.

Two AMRs at two temporal scales were validated: the first was over a year and the second was over a season. Accuracies of all the four methods were also compared for optimum combinations of 2, 3, 4, ..., 14 stations. For a given number of stations, the best combination of sentinel sites, which held the minimum estimation variance of the population (calculated based on historical TRMM data from 1998 to 2004), were selected from all possible combinations to compute the ME, RMSE, and MAPE for each technique.

4. Results

4.1. Multiyear Estimations of Monthly AMR

4.1.1. Accuracies Based on Different Estimators. Table 3 shows that for different number of station combinations, the ME of the arithmetic average is always >3 mm, except when there are two stations. For the other methods, the MEs are smaller, especially for the B-SHADE model, for which the worst ME is only -2.30 mm. This indicates that the AMR is generally biased when estimated using the arithmetic average technique but that the other methods can correct the bias to different degrees. Among them, the B-SHADE model always has the least biased estimation.

Figure 4 shows the RMSEs of the arithmetic average, Thiessen polygon, and ordinary kriging fluctuate for the different numbers of stations. The RMSE of the arithmetic average is always the highest, followed by the kriging and Thiessen polygon techniques. The RMSE of the B-SHADE model is always the smallest in comparison with the other three methods. This indicates that the B-SHADE model

TABLE 3: MEs for the arithmetic average, Thiessen polygon, ordinary kriging, and B-SHADE methods by counting all the months from January 2005 to December 2011.

Number of stations	Arithmetic average (mm)	Thiessen polygon (mm)	OK (mm)	B-SHADE (mm)
2	0.56	-0.56	—	-1.59
3	4.57	2.32	3.43	-2.30
4	5.58	2.78	4.84	-1.55
5	9.84	4.20	6.78	-1.43
6	5.74	2.17	4.52	-1.43
7	4.06	2.33	2.51	-1.33
8	3.11	1.35	1.67	-1.00
9	4.16	1.58	2.73	-1.07
10	4.58	1.59	2.70	-1.07
11	7.13	1.89	3.10	-0.81
12	6.67	1.81	2.79	-0.61
13	7.01	2.03	2.93	-0.57
14	7.11	-1.43	2.89	-0.63

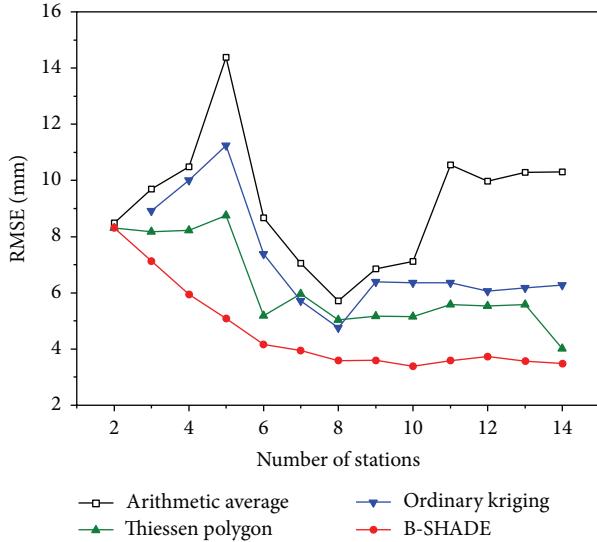


FIGURE 4: RMSEs of different methods by counting all the months from January 2005 to December 2011.

obtains minimum estimation biases and the highest accuracies.

Table 3 and Figure 4 show that the accuracies of the AMRs estimated by the B-SHADE model are the highest of all four methods when all the rain gauge data are used. The ME and RMSE of the B-SHADE model are -0.63 and 3.48 mm, respectively. For arithmetic average, Thiessen polygon, and ordinary kriging, the mean errors were 7.11, -1.43, and 2.89 mm, which were up to 1027.1%, 127.0%, and 358.3%, respectively, greater than for the B-SHADE model; the root mean square errors were 10.31, 4.02, and 6.27 mm, which were up to 196.1%, 15.5%, and 80.0%, respectively, higher than for the B-SHADE model.

4.1.2. Accuracy Change with an Increase of Rain Gauge Numbers. Accuracies of the estimations for different numbers

of stations, by counting all the monthly precipitation from January 2005 to December 2011, show that when the number of stations increases, both the MEs and the RMSEs tend to decrease, except for the arithmetic average method (Table 3 and Figure 4). The ME and RMSE estimated by the arithmetic average model indicate no trend, varying along the averages of 5.4 and 9.2 mm, respectively. When the number of stations increases, the MEs and RMSEs of the arithmetic average in particular, but also of the Thiessen polygon and ordinary kriging methods, fluctuate severely (Table 3 and Figure 4). The RMSE of the arithmetic average method reaches its highest value of 14.39 mm when there are five stations and attains its lowest value of 5.71 mm when the number of stations is eight. Only the ME and RMSE of the B-SHADE method decrease gradually and reach their lowest values when all the rain gauge data are used. This indicates that the B-SHADE method has the most robust performance of the four methods.

4.2. Estimation of Monthly AMR Variation with Season

4.2.1. Accuracies Based on Different Estimators. Accuracies of estimations for different numbers of stations, by counting all the months every season from 2005 to 2011, show that both the ME and the RMSE estimated by the B-SHADE model generally indicate the lowest values, while those of the arithmetic average method have the largest values (Tables 4–7 and Figure 5). In summer, the accuracies of the estimations for the different numbers of stations using the B-SHADE model are the highest (Table 5 and Figure 5). In autumn, they are the highest except when only 2, 4, or 14 rain gauges are considered (Table 6 and Figure 5). In winter, the accuracies of the estimations using the B-SHADE model are similar to those based on the Thiessen polygon, but higher than the other two methods (Table 7 and Figure 5). In spring, the accuracies of the estimations using the B-SHADE model are similar to those based on ordinary kriging, lower than those using the Thiessen polygon method, and higher than those using the arithmetic average method (Table 4 and Figure 5).

TABLE 4: MEs for the arithmetic average, Thiessen polygon, ordinary kriging, and B-SHADE methods by counting the spring months from 2005 to 2011.

Number of stations	Arithmetic average (mm)	Thiessen polygon (mm)	OK (mm)	B-SHADE (mm)
2	-0.94	-2.05	—	-1.43
3	2.95	0.74	1.87	-1.27
4	2.76	0.11	1.72	-1.22
5	6.83	1.36	3.19	-0.79
6	3.98	0.73	2.66	-0.87
7	2.06	0.30	0.47	-0.97
8	1.55	-0.51	-0.14	-0.80
9	2.57	-0.32	0.37	-0.76
10	3.00	-0.42	0.45	-0.77
11	5.11	-0.18	0.96	-0.59
12	4.89	-0.12	0.70	-0.42
13	4.85	-0.35	0.55	-0.52
14	4.99	-1.93	0.59	-0.53

TABLE 5: MEs for the arithmetic average, Thiessen polygon, ordinary kriging, and B-SHADE methods by counting the summer months from 2005 to 2011.

Number of stations	Arithmetic average (mm)	Thiessen polygon (mm)	OK (mm)	B-SHADE (mm)
2	4.55	2.56	—	-2.21
3	11.13	7.26	9.17	-5.49
4	14.85	9.98	13.65	-3.03
5	22.48	12.62	17.23	-3.60
6	12.42	5.73	10.25	-4.20
7	9.97	6.79	6.87	-3.76
8	7.91	5.28	5.38	-2.70
9	9.73	5.69	7.99	-3.05
10	10.61	6.11	8.14	-2.86
11	14.94	6.17	8.51	-2.65
12	13.50	5.66	7.46	-2.41
13	14.50	6.56	8.12	-2.19
14	14.79	-2.70	8.06	-2.27

TABLE 6: MEs for the arithmetic average, Thiessen polygon, ordinary kriging, and B-SHADE methods by counting the autumn months from 2005 to 2011.

Number of stations	Arithmetic average (mm)	Thiessen polygon (mm)	OK (mm)	B-SHADE (mm)
2	-1.66	-2.74	—	-3.01
3	3.82	1.20	2.47	-2.41
4	4.44	1.00	3.77	-1.83
5	9.41	2.79	6.45	-1.05
6	5.79	2.00	4.50	-0.71
7	3.79	2.08	2.38	-0.62
8	2.74	0.74	1.17	-0.50
9	4.08	1.09	2.39	-0.41
10	4.27	0.68	2.04	-0.59
11	7.98	1.65	2.78	0.02
12	7.74	1.75	2.81	0.42
13	8.07	1.94	2.88	0.43
14	7.92	-0.87	2.69	0.29

TABLE 7: MEs for the arithmetic average, Thiessen polygon, ordinary kriging, and B-SHADE methods by counting the winter months from 2005 to 2011.

Number of stations	Arithmetic average (mm)	Thiessen polygon (mm)	OK (mm)	B-SHADE (mm)
2	0.28	0.01	—	0.31
3	0.37	0.07	0.20	-0.04
4	0.26	0.03	0.20	-0.11
5	0.62	0.02	0.27	-0.29
6	0.77	0.21	0.67	0.06
7	0.42	0.14	0.33	0.03
8	0.23	-0.09	0.25	0.00
9	0.26	-0.13	0.16	-0.04
10	0.45	-0.02	0.19	-0.04
11	0.49	-0.09	0.13	-0.02
12	0.55	-0.06	0.20	-0.01
13	0.60	-0.02	0.19	-0.01
14	0.71	-0.20	0.22	-0.01

4.2.2. Accuracy Change with an Increase of Rain Gauge Numbers. Accuracies of the estimations for different numbers of stations, by counting all the months in one of the four seasons from 2005 to 2011, show that when the number of stations increases, both the MEs and RMSEs tend to decrease, except for the arithmetic average method (Figure 5). The ME and RMSE estimated by the arithmetic average model indicate no trend. The RMSEs of the arithmetic average vary along the averages of 5.57, 15.14, 8.68, and 0.84 mm in spring, summer, autumn, and winter, respectively, and the MEs vary along the averages of 6.0, 15, 9, and 0.8 mm in spring, summer, autumn, and winter, respectively.

When the number of stations increases, the MEs and RMSEs of the arithmetic average in particular, but also of the Thiessen polygon and ordinary kriging methods, fluctuate severely in every season. Only the MEs and RMSEs of the B-SHADE method decrease gradually and attain their lowest values when all the stations are considered. This is further evidence that the B-SHADE method has the most robust performance of the four methods at the season scale.

4.2.3. Seasonal Variation of Accuracy Based on the B-SHADE Model. Figure 6 illustrates the MAPEs of the B-SHADE model in the different seasons. As the number of stations increases, the MAPE decreases gradually. When the number of stations is >10 , the MAPE becomes relatively stable for all seasons. Among the four seasons, the MAPE displays the best accuracy in summer and poorest accuracy in winter. When the number of stations is >6 , the accuracy in autumn is higher than in spring, whereas it is the opposite for smaller numbers of station combinations.

5. Discussion

5.1. Accuracy Improvement of AMR Estimation Combined with TRMM Data Based on B-SHADE Model. Based on the historical TRMM data, the population of the areal precipitation can be determined. Thus, through the historical TRMM data, the ratio of rainfall between each station and the population

can be considered in calculating the weights for correcting the bias of the stations in the B-SHADE model. All the other methods take weighted averages to estimate the AMR; however, their weights are based simply on samples (rain gauges), which means that rainfall in areas where there is no rain gauge coverage cannot be reflected in the AMR.

Because the AMR bias correction using the B-SHADE model is based on both spatial and temporal distribution information of precipitation in the historical TRMM data, the B-SHADE model can be taken in AMR estimation in the pre-TRMM data period. If the spatial pattern information from the TRMM data can also be used by introducing covariates, the other three methods can perform better. However, the annual spatial patterns of precipitation in an area vary greatly. Thus, AMR estimation through introducing covariates cannot be extrapolated and be taken in the pre-TRMM data period.

5.2. Robustness of AMR Estimation Based on B-SHADE Model. The B-SHADE model can select the best station combination that presents the smallest theoretical variance of precipitation from all the stations [26, 28]. The spatial pattern of rainfall based on the selected rain gauges can be a closer representation of the “real” one based on the TRMM data than other station combinations. Thus, the AMR will be the best linear unbiased estimation. Conversely, for the other three methods, there are no similar regimes for selecting the best station combination. In this study, the selected stations for the other three methods are the same as those based on the B-SHADE model. The bias of the AMR with the three methods cannot be corrected as it is in the B-SHADE model using historical TRMM data. If the bias of rainfall of the selected stations is high, the accuracy of the estimation will be low and vice versa. Thus, it is easy to understand that the accuracies of the estimated AMR can vary greatly with these three methods.

5.3. Seasonal Variation of AMR Estimation Accuracy. Of all the four methods examined, the B-SHADE model does not illustrate the best accuracy in winter and spring, as it does in

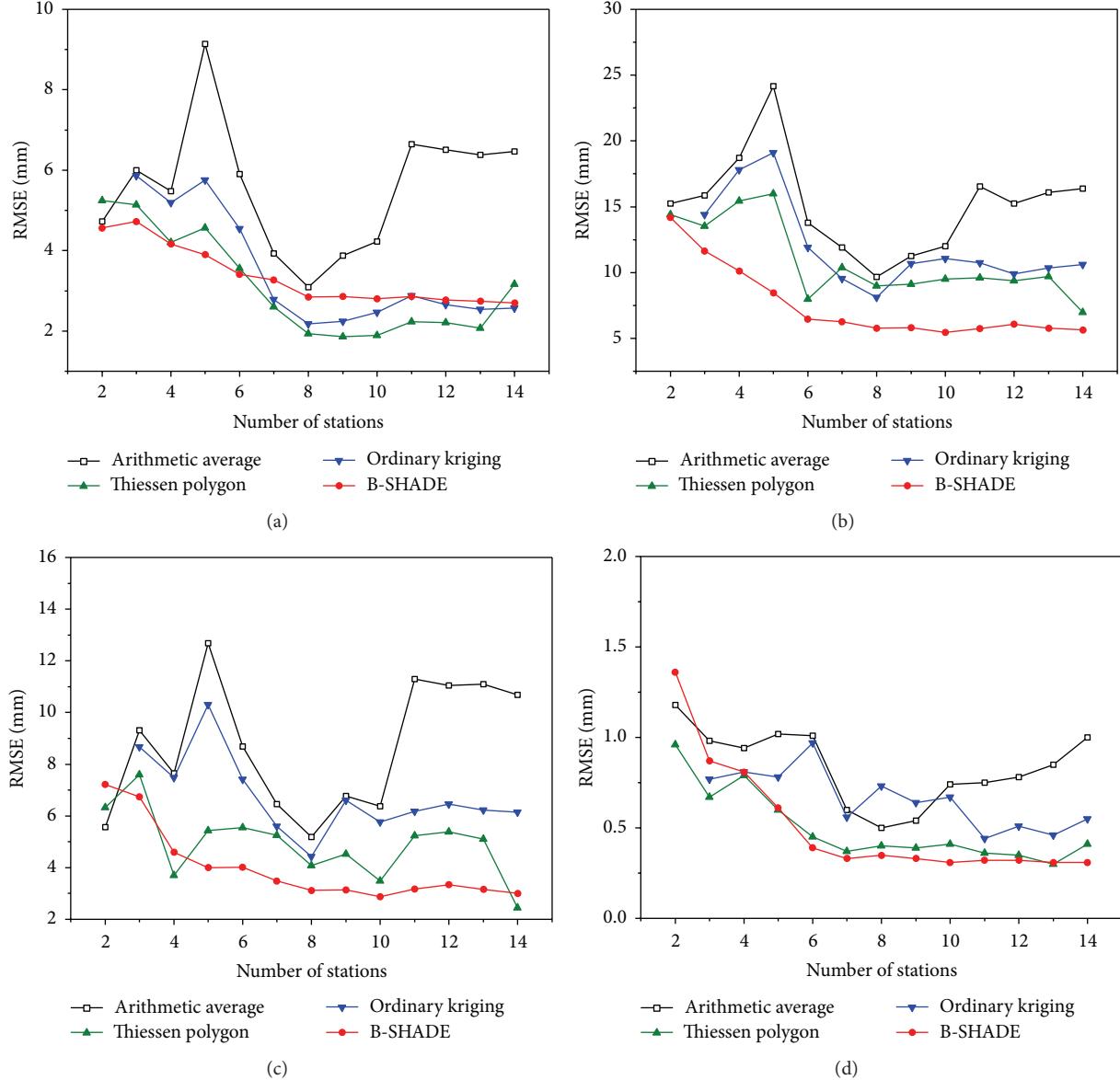


FIGURE 5: RMSEs of the different methods by counting the months from 2005 to 2011 in (a) spring, (b) summer, (c) autumn, and (d) winter.

the other seasons. This might be due to the different spatial distribution of precipitation in these two seasons compared with that of the entire study period. In this study, the length of the TRMM time series is limited for each season (21 months), and it is insufficient to ensure stable estimations for the B-SHADE model. Therefore, we derived the weights for the AMR estimation on the season scale based on the monthly estimations of all four seasons over the first 7 years of the study period (84 months).

Strictly, there are only two seasons in the Three-River Headwaters region [36, 37]: the wet season from May to September and the dry season throughout the rest of the year. The wet season is relatively warm and brings about 80% of the annual precipitation. In contrast, the dry season is quite cold and brings little rainfall. Different atmospheric circulations and sources of moisture between the wet and dry

seasons collectively lead to different distribution patterns of precipitation within the study area [38]. In the wet season, the southwest monsoon controls the weather within the study area, and moisture from the Indian Ocean is transported to the region by the southwesterly wind, which brings heavy rainfall. In contrast, during the dry season, the entire Tibetan Plateau is under the influence of a westerly circulation and, thus, long distances and topographic barriers restrict the transportation of moisture into the region, leading to little precipitation.

As the spatial distribution of precipitation in winter and spring (November to April) is different from that in summer and autumn (June to September) and the wet season brings about 80% of the annual precipitation, the weighting of the stations used in this study is largely decided by the spatial distribution of rainfall during this period, which conversely

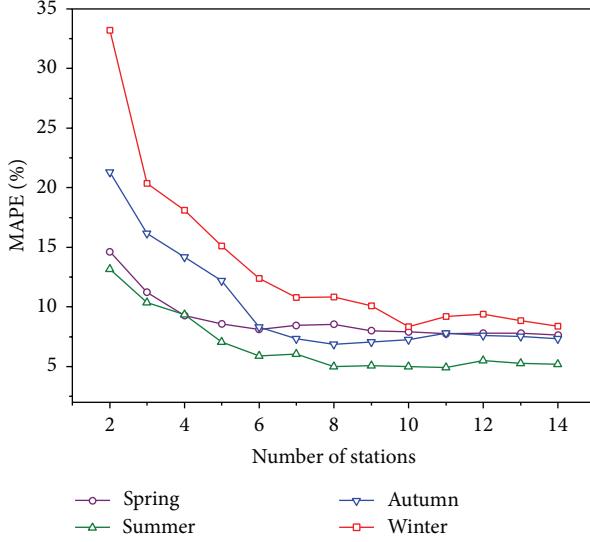


FIGURE 6: MAPEs of the B-SHADE model by counting months in spring, summer, autumn, and winter from 2005 to 2011.

results in relatively poor performance of the B-SHADE model in winter and spring.

In addition, the poor performance of the B-SHADE model in winter and spring is also caused by the lower accuracies of TRMM satellite precipitation product in these seasons. During this period, snowfall is the primary precipitation form, and it is covered by snow and ice in the study area. However, as a representative satellite precipitation product, the TRMM data presents limited skills in capturing and quantifying precipitation in this period for its intrinsic inability of sensors [39–41]. The TRMM combines two different types of satellite sensors, that is, microwave (e.g., TRMM Microwave Imager) and infrared. Microwave sensors show weak ability to capture short rainfall events, because of the anomalous sampling of low earth-orbiting and the limited number of overpasses [39, 42]. Meanwhile, snow and ice surfaces may be interpreted as rainy clouds by microwave sensors, which may introduce errors in mountainous regions [41]. Besides, discrimination of rain/no rain category based on cloud top infrared temperature may also fail in mountainous regions [41, 42]. Therefore, when inaccurate TRMM data is used in the proposed model, the accuracy of the AMR estimation will be limited (see Section 5.4).

5.4. Limitations of the AMR Estimation Based on the B-SHADE Model. As stated in Section 5.3, because the length of TRMM time series is limited for each season, we have to determine the weights for estimating the AMR based on the monthly estimations of all four seasons over the first 7 years of the study period. Furthermore, it has been validated that the B-SHADE model is effective in estimating reliable AMR on the multiyear scale. To estimate seasonal AMR for multiyear or annual scales, it is better to calculate the weights at the corresponding time scale, and historical data for a longer time series are necessary to obtain sufficient samples to ensure stable estimations by the B-SHADE model.

The B-SHADE model assumes that the rainfall features of an area are consistent within the time series. Because of the limitations of TRMM data, we used the data from January 1998 to December 2004 for model construction, and the data from January 2005 to December 2011 for validation. The validation period is very close to that of the model construction and may be not enough to validate the effectiveness of AMR estimation in the past decades. If the spatiotemporal patterns changed significantly during the study period, the estimated error might become larger. However, we believe the likelihood of substantial change in the spatiotemporal pattern within one century is not high. Therefore, as the records of most rainfall gauges are less than this time scale, historical AMR could be determined further using the B-SHADE method proposed in this study.

The performance of the B-SHADE model also relies on the accuracy of the satellite data, for it provides spatiotemporal distribution information of precipitation which is the basis of parameters calculation. Over- and underestimation of the TRMM data can both introduce errors to the AMR estimation. These errors which indicate inaccurate rainfall relationships between stations and the population can lead to incorrect ratios for bias correction and false weights for AMR estimation. Moreover, it might also mislead the determination of the best station combination. Therefore, prevalidation and potentially correction for a good linear relationship between TRMM data and rain gauge records are inevitable in the proposed method of this study.

6. Conclusions

This study presented a method to estimate the AMR using the B-SHADE model with biased rain gauge observations and TRMM data, for remote areas where the distribution of rain gauges is sparse and uneven. Based on the B-SHADE model, the AMR estimation addresses two terms: the unbiased estimation condition between the rainfall from each gauge and the AMR, and the minimum estimation variances between two rain gauges. TRMM data were used to determine the population of the AMR for model construction and validation. The objectives were finally solved using the least-squares method.

A case study was conducted for the Three-River Headwaters region in the middle of the Tibetan Plateau in the south of Qinghai Province, China, which covers an area of about 360,000 km². The results indicated that the B-SHADE model obtained the least estimation biases, that is, the highest accuracy for the AMR estimation; when all the stations were used, the ME and RMSE were -0.63 and 3.48 mm, respectively. For arithmetic average, Thiessen polygon, and ordinary kriging, the mean errors were 7.11, -1.43, and 2.89 mm, which were up to 1027.1%, 127.0%, and 358.3%, respectively, greater than for the B-SHADE model; the root mean square errors were 10.31, 4.02, and 6.27 mm, which were up to 196.1%, 15.5%, and 80.0%, respectively, higher than for the B-SHADE model. When the number of stations increased, the MEs and RMSEs of the arithmetic average in particular, but also of the Thiessen polygon and ordinary kriging methods, fluctuated

severely; only the B-SHADE method demonstrated robust performance.

Because the length of TRMM time series is limited for each season, the weights for the AMR estimation were determined based on the monthly estimations of all four seasons over the first 7 years of the study period. To estimate seasonal AMR over multiyear or annual scales, it would be better to calculate the weights at the corresponding time scale, and historical data for longer time series should be provided to obtain sufficient samples to ensure stable estimations by the B-SHADE model. The proposed technique can be to extend the AMR record to the presatellite observation period, when only the gauge data are available.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This study was supported by the National Program on Key Basic Research Project of China (nos. 2015CB954103 and 2015CB954101) and the National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2012BAH33B01). The constructive criticisms and comments from the anonymous referees are also acknowledged.

References

- [1] P. R. Rakhecha and V. P. Singh, *Applied Hydrometeorology*, Springer Science & Business Media, 2009.
- [2] D. I. F. Grimes, E. Pardo-Igúzquiza, and R. Bonifacio, "Optimal areal rainfall estimation using raingauges and satellite data," *Journal of Hydrology*, vol. 222, no. 1–4, pp. 93–108, 1999.
- [3] D.-J. Seo and J. P. Breidenbach, "Real-time correction of spatially nonuniform bias in radar rainfall data using rain gauge measurements," *Journal of Hydrometeorology*, vol. 3, no. 2, pp. 93–111, 2002.
- [4] M. Li and Q. Shao, "An improved statistical approach to merge satellite rainfall estimates and raingauge data," *Journal of Hydrology*, vol. 385, no. 1–4, pp. 51–64, 2010.
- [5] K. Sene, *Flash Floods Forecasting and Warning*, Springer Science & Business Media, 2012.
- [6] C. Yoo, U. Kim, K. Kim, and K. Kim, "Inter-station correlation and estimation errors of areal average rain rate," *Stochastic Environmental Research and Risk Assessment*, vol. 22, no. 2, pp. 143–157, 2008.
- [7] A. H. Thiessen, "Precipitation Averages for large areas," *Monthly Weather Review*, vol. 39, no. 7, pp. 1082–1089, 1911.
- [8] J. D. Creutin and C. Obled, "Objective analyses and mapping techniques for rainfall fields: an objective comparison," *Water Resources Research*, vol. 18, no. 2, pp. 413–431, 1982.
- [9] G. Q. Tabios and J. D. Salas, "A comparative analysis of techniques for spatial interpolation of precipitation," *Journal of the American Water Resources Association*, vol. 21, no. 3, pp. 365–380, 1985.
- [10] D. L. Phillips, J. Dolph, and D. Marks, "A comparison of geostatistical procedures for spatial analysis of precipitation in mountainous terrain," *Agricultural and Forest Meteorology*, vol. 58, no. 1–2, pp. 119–141, 1992.
- [11] C. Daly, "Guidelines for assessing the suitability of spatial climate data sets," *International Journal of Climatology*, vol. 26, no. 6, pp. 707–721, 2006.
- [12] G. Matheron, "Principles of geostatistics," *Economic Geology*, vol. 58, no. 8, pp. 1246–1266, 1963.
- [13] J. A. Hevesi, J. D. Istok, and A. L. Flint, "Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: structural analysis," *Journal of Applied Meteorology*, vol. 31, no. 7, pp. 661–676, 1992.
- [14] E. Pardo-Igúzquiza, "Comparison of geostatistical methods for estimating the areal average climatological rainfall mean using data on precipitation and topography," *International Journal of Climatology*, vol. 18, no. 9, pp. 1031–1047, 1998.
- [15] H. L. Xu, C.-Y. Xu, H. Chen, Z. X. Zhang, and L. Li, "Assessing the influence of rain gauge density and distribution on hydrological model performance in a humid region of China," *Journal of Hydrology*, vol. 505, pp. 1–12, 2013.
- [16] A. K. Mishra, "Effect of rain gauge density over the accuracy of rainfall: a case study over Bangalore, India," *SpringerPlus*, vol. 2, no. 1, pp. 1–7, 2013.
- [17] C. L. Shope and G. R. Maharjan, "Modeling spatiotemporal precipitation: effects of density, interpolation, and land use distribution," *Advances in Meteorology*, vol. 2015, Article ID 174196, 16 pages, 2015.
- [18] C. Daly, R. P. Neilson, and D. L. Phillips, "A statistical-topographic model for mapping climatological precipitation over mountainous terrain," *Journal of Applied Meteorology*, vol. 33, no. 2, pp. 140–158, 1994.
- [19] R. J. Hijmans, S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis, "Very high resolution interpolated climate surfaces for global land areas," *International Journal of Climatology*, vol. 25, no. 15, pp. 1965–1978, 2005.
- [20] U. Haberlandt, "Geostatistical interpolation of hourly precipitation from rain gauges and radar for a large-scale extreme rainfall event," *Journal of Hydrology*, vol. 332, no. 1–2, pp. 144–157, 2007.
- [21] R. Schiemann, R. Erdin, M. Willi, C. Frei, M. Berenguer, and D. Sempere-Torres, "Geostatistical radar-raingauge combination with nonparametric correlograms: methodological considerations and application in Switzerland," *Hydrology and Earth System Sciences*, vol. 15, no. 5, pp. 1515–1536, 2011.
- [22] A. Verworn and U. Haberlandt, "Spatial interpolation of hourly rainfall-effect of additional information, variogram inference and storm properties," *Hydrology and Earth System Sciences*, vol. 15, no. 2, pp. 569–584, 2011.
- [23] P. D. Wagner, P. Fiener, F. Wilken, S. Kumar, and K. Schneider, "Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions," *Journal of Hydrology*, vol. 464–465, pp. 388–400, 2012.
- [24] P. D. A. Borges, J. Franke, Y. M. T. da Anunciação, H. Weiss, and C. Bernhofer, "Comparison of spatial interpolation methods for the estimation of precipitation distribution in Distrito Federal, Brazil," *Theoretical and Applied Climatology*, pp. 1–14, 2015.
- [25] Y. H. Ding, S. W. Wang, J. Y. Zheng, H. J. Wang, and X. Q. Yang, Eds., *Climate of China*, Meteorological Press, Beijing, China, 2013.
- [26] J.-F. Wang, B. Y. Reis, M.-G. Hu et al., "Area disease estimation based on sentinel hospital records," *PLoS ONE*, vol. 6, no. 8, Article ID e23428, 2011.

- [27] J. Wang, C. Xu, M. Hu et al., "A new estimate of the China temperature anomaly series and uncertainty assessment in 1900–2006," *Journal of Geophysical Research: Atmospheres*, vol. 119, no. 1, pp. 1–9, 2014.
- [28] M.-G. Hu, J.-F. Wang, Y. Zhao, and L. Jia, "A B-SHADE based best linear unbiased estimation tool for biased samples," *Environmental Modelling & Software*, vol. 48, pp. 93–97, 2013.
- [29] C.-D. Xu, J.-F. Wang, M.-G. Hu, and Q.-X. Li, "Interpolation of missing temperature data at meteorological stations using P-BSHADE," *Journal of Climate*, vol. 26, no. 19, pp. 7452–7463, 2013.
- [30] L. Tong, X. Xu, Y. Fu, and S. Li, "Wetland changes and their responses to climate change in the 'Three-River Headwaters' region of China since the 1990s," *Energies*, vol. 7, no. 4, pp. 2515–2534, 2014.
- [31] Z. C. Xie and C. H. Liu, *An Introduction to Glaciology*, Shanghai Popular Science Press, Shanghai, China, 2010.
- [32] C. Kummerow, W. Barnes, T. Kozu, J. Shiue, and J. Simpson, "The Tropical Rainfall Measuring Mission (TRMM) sensor package," *Journal of Atmospheric and Oceanic Technology*, vol. 15, no. 3, pp. 809–817, 1998.
- [33] C. Kummerow, J. Simpson, O. Thiele et al., "The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit," *Journal of Applied Meteorology*, vol. 39, no. 12, pp. 1965–1982, 2000.
- [34] Z. Duan, W. G. M. Bastiaanssen, and J. Z. Liu, "Monthly and annual validation of TRMM Multisatellite Precipitation Analysis (TMPA) products in the Caspian Sea Region for the period 1999–2003," in *Proceedings of the 32nd IEEE International Geoscience and Remote Sensing Symposium (IGARSS '12)*, pp. 3696–3699, IEEE, Munich, Germany, July 2012.
- [35] K. Fleming and J. L. Awange, "Comparing the version 7 TRMM 3B43 monthly precipitation product with the TRMM 3B43 version 6/6A and Bureau of Meteorology datasets for Australia," *Australian Meteorological and Oceanographic Journal*, vol. 63, no. 3, pp. 421–426, 2013.
- [36] J. X. Dai, *Climate in Qinghai-Tibet Plateau*, China Meteorological Press, Beijing, China, 1990.
- [37] G. X. Wang, *Land Surface Processes and Environmental Changes in River Headwater Regions of Qinghai-Tibet Plateau*, Science Press, Beijing, China, 2010.
- [38] L. Cao and S. Pan, "Changes in precipitation extremes over the 'Three-River Headwaters' region, hinterland of the Tibetan Plateau, during 1960–2012," *Quaternary International*, vol. 321, pp. 105–115, 2014.
- [39] Y. Tian, C. D. Peters-Lidard, J. B. R. Eylander et al., "Component analysis of errors in satellite-based precipitation estimates," *Journal of Geophysical Research: Atmospheres*, vol. 114, no. 24, Article ID D24101, 2009.
- [40] G. J. Huffman, R. F. Adler, D. T. Bolvin, and E. J. Nelkin, "The TRMM multi-satellite precipitation analysis (TMPA)," in *Satellite Rainfall Applications for Surface Hydrology*, pp. 3–22, Springer, Dordrecht, The Netherlands, 2010.
- [41] T. Dinku, S. J. Connor, and P. Ceccato, "Comparison of CMORPH and TRMM-3B42 over mountainous regions of Africa and South America," in *Satellite Rainfall Applications for Surface Hydrology*, pp. 193–204, Springer, Dordrecht, The Netherlands, 2010.
- [42] A. S. Gebregiorgis and F. Hossain, "Understanding the dependence of satellite rainfall uncertainty on topography and climate for hydrologic model simulation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 1, pp. 704–718, 2013.



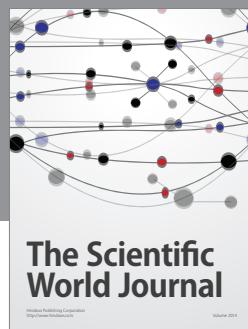
Journal of
Geochemistry



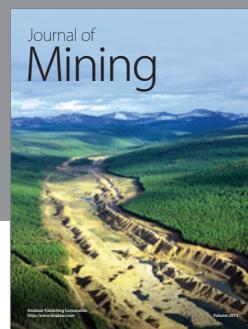
Scientifica



International Journal of
Ecology



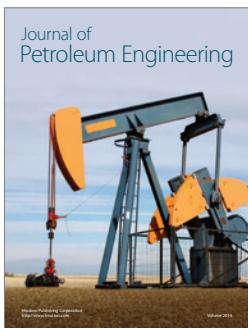
The Scientific
World Journal



Journal of
Mining



Journal of
Earthquakes

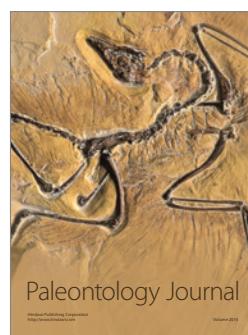


Journal of
Petroleum Engineering

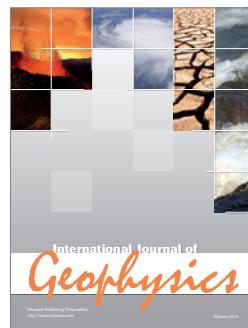


Hindawi

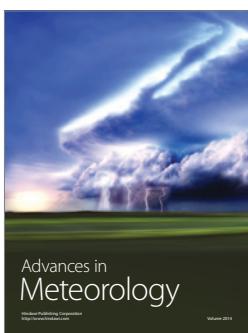
Submit your manuscripts at
<http://www.hindawi.com>



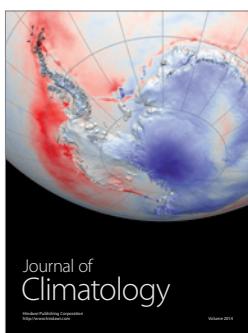
Paleontology Journal



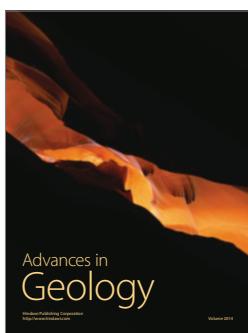
International Journal of
Geophysics



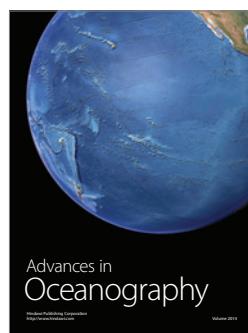
Advances in
Meteorology



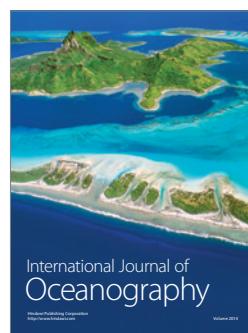
Journal of
Climatology



Advances in
Geology



Advances in
Oceanography



International Journal of
Oceanography



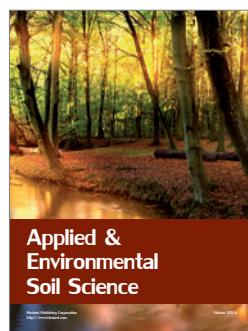
International Journal of
Mineralogy



Journal of
Geological Research



International Journal of
Atmospheric Sciences



Applied &
Environmental
Soil Science



Journal of
Computational
Environmental Sciences