



Regional precipitation frequency analysis for 24-h duration using GPM and L-moments approach in South China

Xiaoyang Chen¹ · Dehua Zhu^{2,3} · Min Wang^{1,4} · Yifan Liao⁵

Received: 31 March 2022 / Accepted: 16 February 2023 / Published online: 1 March 2023
© The Author(s) 2023

Abstract

There are little researchers that combined satellite-based products and L-moments methods in frequency analysis. The objective of this study was to explore and evaluate the satellite precipitation data applied in univariate L-moments method. The Shaoguan City of South China was selected as the study area. Annual maximum precipitation (AMP) respectively from observational rainfall and Integrated Multi-satellitE Retrievals for Global Precipitation Mission (IMERG) gridded rainfall was used for the calculation of estimates using regional L-moments analysis (RLMA) method. Then, the quantiles at grids were corrected by mean bias correction (MBC) method with the estimates at sites. The results showed that the number of homogeneous regions as well as the spatial distribution of parameters (C_v , C_s , and C_k) for two data sets are different. When the return periods are larger than 40 years, the estimates at sites are also larger than those at grids and the difference will increase as return period increases. After being modified, the quantiles from IMERG are more reasonable and combine the spatial distribution characteristic of estimates from two cases, which indicates that the estimation calculated by satellite products with L-moments approach is applicable and more reasonable. This study can also provide reference for solving the issues regarding the density and coverage of rain gauges that affect the accuracy of frequency analysis, especially in remote areas of China or most parts of the world.

1 Introduction

The rapid development of global climate change and urbanization has caused the frequent occurrences of extreme events in recent years (Wu 2017). From 18:00

on July 18 to 00:00 on July 21, 2021 (Beijing time, the same below), influenced by the combined effects of terrain and water vapor which is transported by typhoon “fireworks,” heavy rains across large parts of Henan Province, especially torrential rains, hit Zhengzhou, Xuchang, and Xinxiang cities (Central Meteorological Observatory 2021a, b). The daily rainfall of 10 meteorological stations in Henan Province has been breaking the historical extreme value since the record began. At Zhengzhou City, the maximum 24-h accumulated rainfall reached 624.1 mm, exceeding the annual accumulated rainfall of 2019 (509.5 mm) (Central Meteorological Observatory 2021a, b). Shaoguan is a typical city in South China, in which mountains and hills are the most widely distributed and average annual rainfall is over 1500 mm. During the flood season, it is often attacked by continuous rainstorms and the flooding, flush flood, and debris which occurred frequently. These disasters have brought huge losses to the safety of people’s lives and the social economy every year. From 08:00 on June 8 to 08:00 on June 8, 2020, there were heavy rains across central and southern Shaoguan and the local parts were attacked by torrential rains. In Qujiang County, the daily rainfall reached 460.7 mm

✉ Dehua Zhu
d.zhu@nuist.edu.cn
Xiaoyang Chen
cxy921020@126.com

¹ Meteorological Bureau of Shaoguan City, Shaoguan 512028, China
² School of Hydrology and Water Resources, Nanjing University of Information Science and Technology, Nanjing 210044, China
³ Key Laboratory of Hydrometeorological Disaster Mechanism and Warning of Ministry of Water Resources, Nanjing University of Information Science and Technology, Nanjing 210044, China
⁴ Tianjin Meteorological Information Center, Tianjin 30074, China
⁵ Guangxi Vocational College of Water Resources and Electric Power, Nanning 530000, China

and had refreshed the historic record of daily rainfall in Shaoguan City (China Meteorological Administration 2020). The extreme precipitation caused road collapses, houses flooded, and walls collapsed in part of this city, and traffic was disrupted. Therefore, it has great practical significance to perform long-duration 24-h precipitation frequency analysis in Shaoguan, with compiling a 24-h rainfall frequency atlas which can provide scientific suggestions for city management in carrying out classification prevention and rescues of extreme precipitation disaster events.

There are two core issues in hydrological frequency analysis: the precise estimate of parameters and the accuracy of quantiles. Simultaneously, two major difficulties are existent: it is impossible to derive the optimal frequency distribution function of the rainfall data through theoretical analysis, and the true value of the quantiles is unknown (Lin 2010a, b; Lin et al. 2012). At present, two approaches are frequently used to perform the frequency analysis of precipitation, at-site and regional. Since at-site frequency analysis needs a collection of long periods of records, the reliability of estimates is usually hampered by the insufficient length of recorded data, especially when estimating rainfall of large return periods (Bernard et al. 1995). On the contrary, regional frequency analysis can improve the accuracy of quantile estimates by incorporating spatial data from several sites having similar characteristics in a homogeneous region (Hosking and Wallis 1998). The index-flood method proposed by Hosking and Wallis (1993) is widely used with L-moments approach to undertake regional frequency analysis. It is reported that L-moments have the theoretical advantages over conventional moments of being more robust to the presence of outliers in the data and being less subjective to biasness in the estimation of parameters (Hosking and Wallis 1998; Lin et al. 2006; Lin 2010b; Liang et al. 2013). Later, by applying the concept of regional analysis, Lin et al. (2006) and Lin (2010b) separated a rainfall into common component and local component, in the practices of precipitation frequency analysis for the NOAA Atlas 14 in the USA (Lin et al. 2006; Lin 2010b; Bonnin et al. 2012) as well as for the Taihu Lake Basin in China (Lin 2010b; Wu et al. 2015). Other studies abroad have also been done using the regional L-moments approach which combines the regional analysis with the L-moments for rainfall frequency analysis (e.g., UK (Fowler and Kilsby 2010), Italy (Norbiato et al. 2007), Norway (Hailegeorgis et al. 2013)) and many regions of China (e.g., Guangxi (Chen et al. 2014), Yangtze River Delta region (Yin et al. 2016), Huaihe River Basin (Shao et al. 2016), Jiangxi (Ding et al. 2018; Liu et al. 2018), Sichuan (Li et al. 2019), and Xiamen (Shang et al. 2019)). The research results indicated that the regional L-moments method is able to solve the problems related to identification of homogeneous regions, testing and selection of the

best-fitting distribution, as well as estimations of parameters and quantiles at places of interest.

In general, L-moments methods are usually used with observational rainfall data from stations, but the gauges in Shaoguan are distributed unevenly and sparsely. Most of the satellite-based precipitation products appeared after 2000, as a new type of precipitation data, there are little researchers which combined satellite precipitation data and L-moments methods in frequency analysis. In this study, Integrated Multi-satellitE Retrievals for Global Precipitation Mission (IMERG) gridded rainfall data is used as an example to explore and evaluate the applicability of satellite precipitation data in the univariate L-moments method which has been used in many studies and thus relatively mature (Jing et al. 2014; Modarres 2010; Sarmadi and Shokoochi 2014). The GPM products can provide precipitation data with higher accuracy, larger coverage, and higher spatial and temporal resolution, as well as improve the detection accuracy of precipitation; therefore, it is also expected the results can complement the shortage of rainfall stations in the study area.

The contents of this paper are as follows: Section 2 introduces the study area and data sets used. Following that, the theoretical method of regional L-moments analysis (RLMA) and mean bias correction (MBC) are described in Section 3. Later, the results and discussion are given in Section 4, and the main conclusions are summarized in Section 5.

2 Study area and data

2.1 Study area

The target area, Shaoguan City, lies off the north of Guangdong Province with the geographic coordinates between 112°50' and 114°45'E longitude and 23°5' and 25°3'N latitude and covers a total area of about 18,400 km². According to the statistics, the average 24-h annual rainfall of Shaoguan City is 1695.1 mm over the past 30 years (from 1991 to 2020), while the air temperature is 20.3 °C. As the elevation map shows (Fig. 1), two-thirds of the city is occupied by mountains which are mostly located in the western, while plains are mainly distributed in the central part. Three main mountain ranges run approximately in parallel from northwest to southeast, including Yaoling, Huashi and Qingyun Mountain Range, with a maximum elevation of 1902 m.

2.2 Data and sources

2.2.1 Observational precipitation data

The observational precipitation data from 8 meteorological stations and 62 hydrological stations provided by

Guangdong Hydrology Bureau and Shaoguan Meteorological Bureau was used in this study. The distribution of these stations is shown in Fig. 1, where 39 stations are distributed in the “buffer zone” within 30 km around Shaoguan City in order to improve the reliability of frequency estimation in the edge of the study area. The data includes hydrological 24-h annual maximum precipitation and meteorological hourly precipitation. The effective data length was selected as 20 years (from 1980 to 2020) to ensure the reliability of the calculation results. After review, the length of data of each station is greater than 20 years, meaning that 70 stations are available.

2.2.2 IMERG rainfall data

One of the major problems in precipitation frequency analysis in Shaoguan area is the stations which have long-term records of reliable rainfall data are distributed unevenly and sparsely. The in situ-based gridded rainfall data can be used in the region where long-term reliable precipitation observation is not available because of their spatial and temporal continuity (Ahmed et al. 2019). In this paper, the IMERG satellite gridded rainfall products provided by the National Aeronautics and Space Administration (NASA) were used for research and analysis. The data was selected from June 1, 2000, to December 31, 2020, with the coordinates between 112° 36' and 115° E longitude and 23° 42' and 25° 48' N latitude. Besides that, the time resolution of data is 1 day while spatial resolution is $0.1^\circ \times 0.1^\circ$. Existing studies have shown that this data set has excellent adaptability in China, with high spatial and temporal resolution and global coverage

(Chen et al. 2017; Wu et al. 2019). The experimental results also showed that in Shaoguan and its surrounding areas, the correlation coefficient between precipitation observation and corresponding IMERG rainfall on daily and annual scales is 0.7~0.8, especially exceeding 0.9 on seasonal scale. Therefore, the IMERG data is reasonable and can be applied to this study.

3 Methodology

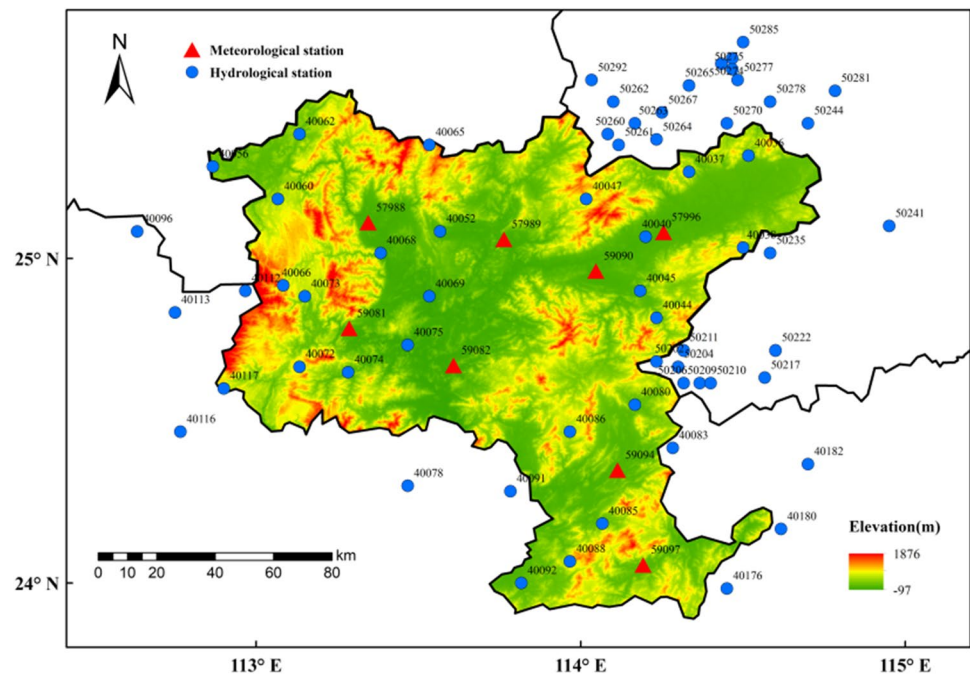
3.1 Regional L-moments analysis

The RLMA includes the following steps which mainly use an index-flood procedure: identification of homogeneous regions, parameter estimation, choice of an appropriate probability distribution, and estimation of quantiles. The theory of RLMA and the methods used in these steps are briefed in this section.

3.1.1 L-moments approach

L-moments are an alternative system to summarize the statistical properties of the hydrologic data and to describe the shape of the probability distributions. L-moments arose as modifications of the probability-weighted moments (PWMs) of Greenwood et al. (1979) and are defined by Hosking (1990) as expectations of certain linear combinations of order statistics. Letting $X_{1:n} \leq X_{2:n} \leq \dots \leq X_{n:n}$ be the order statistics of a random sample of size n drawn from the

Fig. 1 Elevation map of Shaoguan City, and locations of the rain gauge stations



distribution of X , the L-moments of a probability distribution are defined by Hosking et al. (1998):

$$\begin{aligned} \lambda_1 &= E(X_{1:1}) \\ \lambda_2 &= \frac{1}{2}E(X_{2:2} - X_{1:2}) \\ \lambda_3 &= \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3}) \\ \lambda_4 &= \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) \end{aligned} \tag{1}$$

In general, the r th L-moment of variable X is

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E(X_{r-k:r}), r = 1, 2, \dots \tag{2}$$

where $E(X_{r-k:r})$ is the $(r - k)$ th order statistics from a sample size of r .

Hosking and Wallis (1998) defined the L-moment ratios as follows:

$$\begin{aligned} \text{Coefficient of } L - \text{variation } (L - C_v) : \tau_2 &= \lambda_2/\lambda_1 \\ \text{Coefficient of } L - \text{skewness } (L - C_s) : \tau_3 &= \lambda_3/\lambda_2 \\ \text{Coefficient of } L - \text{kurtosis } (L - C_k) : \tau_4 &= \lambda_4/\lambda_2 \end{aligned} \tag{3}$$

3.1.2 The index-flood procedure

The key assumption of an index-flood procedure is that the N sites form a homogeneous region and the frequency distributions of the N sites are identical apart from a site-specific scaling factor (Hosking and Wallis 1993). The index-flood actually is a location estimator which is commonly the at-site sample mean \bar{x}_i . The frequency values (q_{T_j}) at several desired return periods (T_j) of the dimensionless regional distribution are called regional growth factors (RGFs). The quantiles for return periods (T_j) at site i can be written as (Hosking and Wallis 1993):

$$Q_{T_j,i} = \bar{x}_i q_{T_j}, j = 2\text{years}, 5\text{years}, \dots, 100\text{years}, \dots, 1000\text{years} \tag{4}$$

The q_{T_j} can be determined by a set of regional parameters for a selected distribution. The regional parameters are weighted average values over N sites, with the parameter estimate of site i given weight proportional to the length of record, n_i . The at-site parameter estimate is obtained via the L-moments method based on rescaling data at site i to its mean \bar{x}_i .

3.1.3 Identification of homogeneous regions

The aim of identifying homogeneous regions is to form groups of sites that approximately satisfy the homogeneity condition. In the homogeneous region, the statistical parameters ($L - C_v$, $L - C_s$, and $L - C_k$) of the sites are consistent

within a tolerance. The measure to judge the degree of homogeneity of the region is (Hosking and Wallis 1998):

$$H_1 = \frac{V_{L-C_v} - \bar{V}_{L-C_v}}{\sigma_{L-C_v}} \tag{5}$$

where V_{L-C_v} is the variation of the $L - C_v$ over stations in a region. So, H_1 is the standardized \bar{V}_{L-C_v} . Hosking and Wallis (1993) suggested the region can be considered as ‘‘acceptably homogeneous’’ if $H_1 < 1$. The effect of $L - C_s$ on the formation of homogeneous regions is also regarded.

3.1.4 Selection of an appropriate probability distribution

In each region, only the best distribution is applied to undertake regional frequency analysis for better and more accurate calculation result. Three goodness-of-fit tests are used to select the best distribution from five commonly used three-parameter frequency distributions: generalized logistic (GLO), generalized extreme value (GEV), generalized normal (GNO), generalized Pareto (GPA), and Pearson type III (PE3).

Test 1: the Monte-Carlo simulation test

The goodness of fit is judged by the deviation from the regional average of $L - C_s$ and $L - C_k$ of the observed data to the $L - C_s$ and $L - C_k$ of the fitted distribution. For each distribution, the measure is (Hosking and Wallis 1998):

$$Z^{\text{DIST}} = (\tau_4^{\text{DIST}} - t_4^R + \beta_4)/\sigma_4 \tag{6}$$

where DIST refers to the candidate distribution; t_4^R is the average $L - C_k$ value computed from the data of a given region; τ_4^{DIST} is the $L - C_k$ of the fitted distribution to the data using the candidate distribution; β_4 and σ_4 are the bias and standard deviation of the regional average $L - C_k$, respectively; and t_4^R is the Monte-Carlo simulation samples by kappa distribution.

For a confidence level of 90%, a distribution is acceptable if $|Z^{\text{DIST}}| \leq 1.64$. Among accepted distributions, the distribution with the smallest $|Z^{\text{DIST}}|$ is the most appropriate $L - C_k$ distribution.

Test 2: root mean square error of the sample L-moments

In this test, a weighted root mean square error (RMSE) calculated for each of the candidate distributions serves as an index. It is expressed by Lin et al. (2006):

$$\text{RESE} = \left(\sum_{i=1}^N n_i (S_{i,L-C_k} - D_{i,L=C_k})^2 / \sum_{i=1}^N n_i \right)^{\frac{1}{2}}, i = 1, 2, \dots, N \tag{7}$$

where n_i is the data length at site i , and $S_{i,L-C_k}$ and $D_{i,L=C_k}$ are the sample $L - C_k$ and the distribution’s $L - C_k$ at site i ,

respectively. The distribution with the smallest RMSE is the most appropriate distribution.

Test 3: real data check test

Having had a quantile estimate from a fitted distribution at a given return period (T_j) based on the real data series, an empirical exceedance frequency (F_{i,T_j}) to the quantiles can be calculated. Then, F_{i,T_j} is compared with its corresponding theoretical probability (P_{T_j}). A relative error (RE) can be calculated over several return periods to reflect the degree of the match between the empirical frequencies and the theoretical probabilities at site i . The smaller the RE, the better the fitting will be. Then, the regional average RE calculated over N sites in a homogeneous region can be used as an index for the goodness of fit (Lin et al. 2006):

$$RE = \sum_j \sum_{i=1}^N \frac{(F_{i,T_j} - P_{T_j})}{P_{T_j}} / \sum_j \sum_{i=1}^N 1, i = 1, 2, \dots, N; j = 2\text{years}, 5\text{years}, \dots, 100\text{years} \quad (8)$$

For each return period, REs over the five three-parameter frequency distributions mentioned above are ranked from large to small, and the total ranking number over 2 years, 5 years, 10 years, 25 years, and 50 years for each distribution is taken as an index, called RE score, to evaluate the goodness of fit of the distribution. The higher the RE score, the better the fitted distribution will be.

The above three goodness-of-fit tests are performed for each homogeneous region, and then, a final decision to choose the most appropriate distribution can be made based on a summary of the test results.

3.2 Mean bias correction

The fundamental idea of the MBC method is to use the total error of the estimates for desired return period which are calculated respectively based on IMERG rainfall and observational rainfall, to correct the former. This method assumes that in the same homogeneous region, the quantiles obtained from IMERG data have uniform deviation (Gjertsen et al. 2003). For each homogeneous region, the correction factor B can be calculated by the following formula:

$$B = \frac{\sum_{i=1}^n P_{si}}{\sum_{i=1}^n P_{gi}} \quad (9)$$

where n is the number of rain gauge stations in the homogeneous region, and P_{si} and P_{gi} is the estimate for desired return period at station and at adjacent grid, respectively.

4 Results and discussion

4.1 Consistency test of hydrological and meteorological precipitation data

In this study, the annual maximum series (AMS) approach was used to construct 24-h annual maximum precipitation (AMP) at stations and grids. For IMERG data, daily rainfall should be converted to 24-h precipitation first. By comparing the annual maximum value of daily and 24-h precipitation at 8 meteorological stations, the average amplification factor of IMERG daily rainfall was determined to be 1.17.

For precipitation observation, after getting 24-h AMP at each station, data quality control should be carried out for examining whether precipitation data from meteorological and hydrological stations have same distribution and can be used together in frequency analysis. Therefore, 6 groups of hydrological and meteorological rainfall stations with close distances were selected (59,776–40,046, 57,988–40,068, 59,082–40,075, 59,090–40,045, 59,081–40,074, 59,094–40,083), and the statistical consistency of precipitation data between members of each group was tested by using t test method ($\alpha = 0.05$). The result showed that 6 groups all have the same distribution at the significance level of 5%.

4.2 Parameter estimation

Based on the 24-h AMP at stations and grids, the statistical parameters such as mean, C_v , C_s , and C_k are calculated with the L-moments method, and their spatial distribution map is drawn in Fig. 2. The range of mean of AMP at sites is 93.32~155.63 mm, while that at grids is 101.22~169.15 mm. The maximums of mean are mainly distributed in the southern plains, while the extreme rainfall in the northwestern mountainous region is relatively small. The $L - C_v$ at sites showed that the AMP from observational rainfall has a large degree of dispersion in western Shaoguan, and the maximum zone of $L - C_v$ at grids is mainly located in the northwest corner, southwest corner, and the east-central mountainous region; that is, the AMP from IMERG rainfall in these areas has more maximum or minimum precipitation. The $L - C_s$ can reflect the frequency distribution characteristics of the AMP flag values while $L - C_k$ describes the steepness of the peak of AMP distribution. Domestic and overseas studies have presented that there is a good positive correlation between $L - C_s$ and $L - C_k$ (Hosking and Wallis 1998). The spatial distributions of $L - C_s$ and $L - C_k$ at sites are generally similar, indicating that the western mountainous region of Shaoguan has

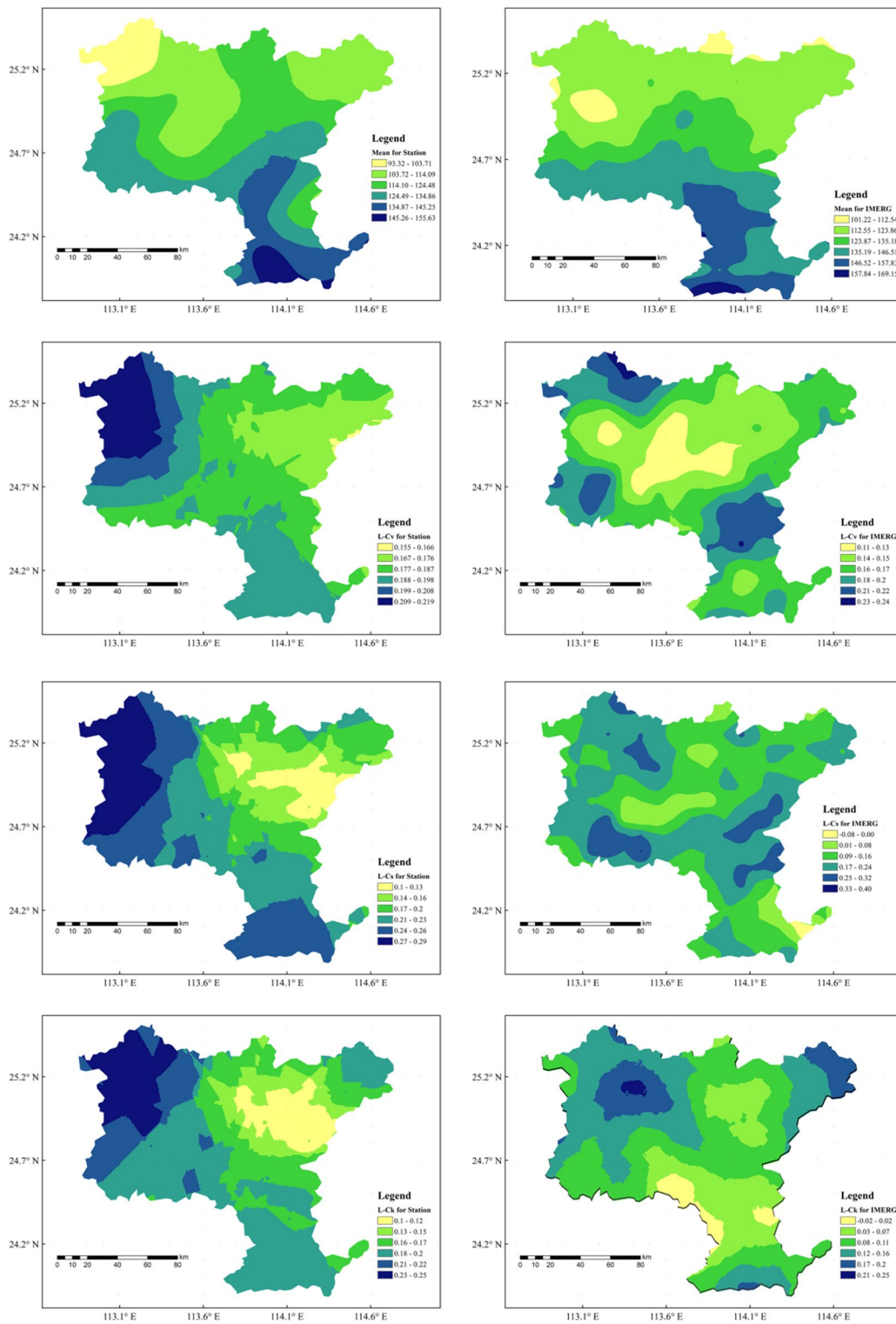


Fig. 2 Spatial distribution of mean, $L - C_v$, $L - C_s$, and $L - C_k$ at sites or grids for 24-h AMP in Shaoguan area

more and stronger extreme precipitation. However, the spatial distributions of $L - C_s$ and $L - C_k$ at grids are quite different, but both have a maximum zone in the northwestern plain, meaning that the distributions of AMP from IMERG

have the characteristics of steep peak and thick tail in this area. In addition, the maximum value of $L - C_s$ at grids is also located in the southwest corner and the east-central mountainous region, which corresponds to the maximum

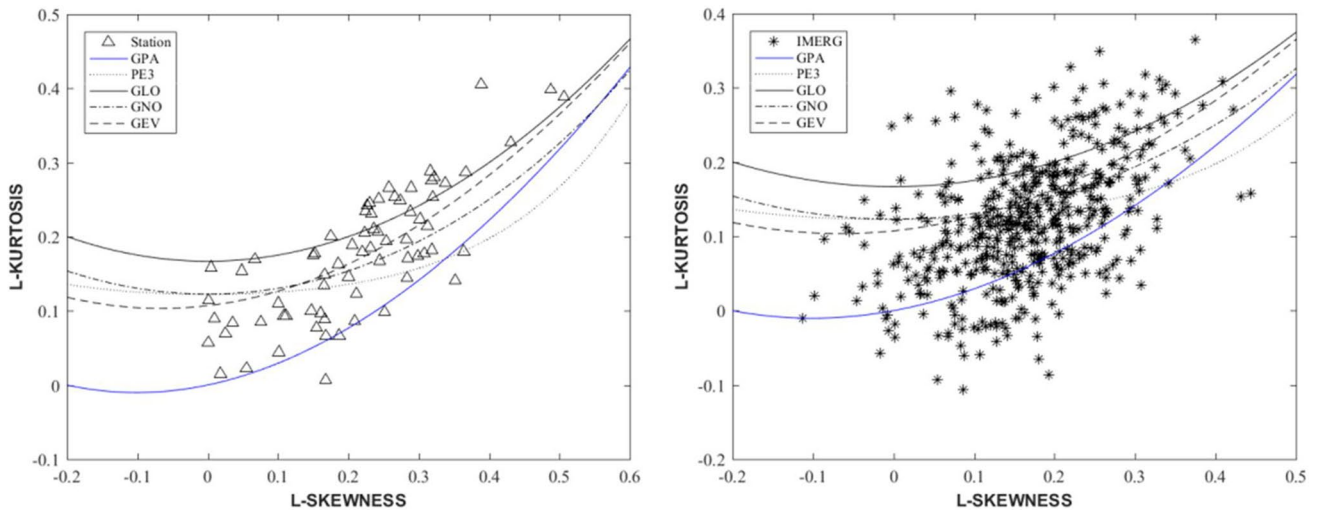


Fig. 3 The theoretical curves of $L - C_s$ and $L - C_k$ of five commonly three-parameter distributions and the point distance of parameters at 70 stations and at 504 grids

of $L - C_v$ at grids, indicating that there has extreme precipitation which is far exceeding the average of AMP from IMERG.

The five curves in Fig. 3 represent the theoretical curves of $L - C_s$ and $L - C_k$ of five commonly three-parameter distributions, and the point distance of these two parameters at 70 stations or 504 grids is also plotted in Fig. 3. It can be seen that when considering the Shaoguan area as one homogeneous region, the spot of stations or IMERG is scattered, indicating that the same distribution cannot be chosen for the 24-h extreme rainfall of all stations or grids. Therefore,

Shaoguan should be divided into several homogeneous regions in the next step.

4.3 Homogeneous region of Shaoguan

4.3.1 According to 24-h AMP from observational rainfall

Based on the criteria for identifying homogeneous regions mentioned above, Shaoguan area was delineated into four subregions for 24-h AMP from precipitation observation (Fig. 4). Table 1 presents the homogeneity measure (H_1)

Fig. 4 Sketch of 24-h homogeneous regions of Shaoguan and the best-fit distributions for homogeneous regions according to 24-h AMP from observational rainfall

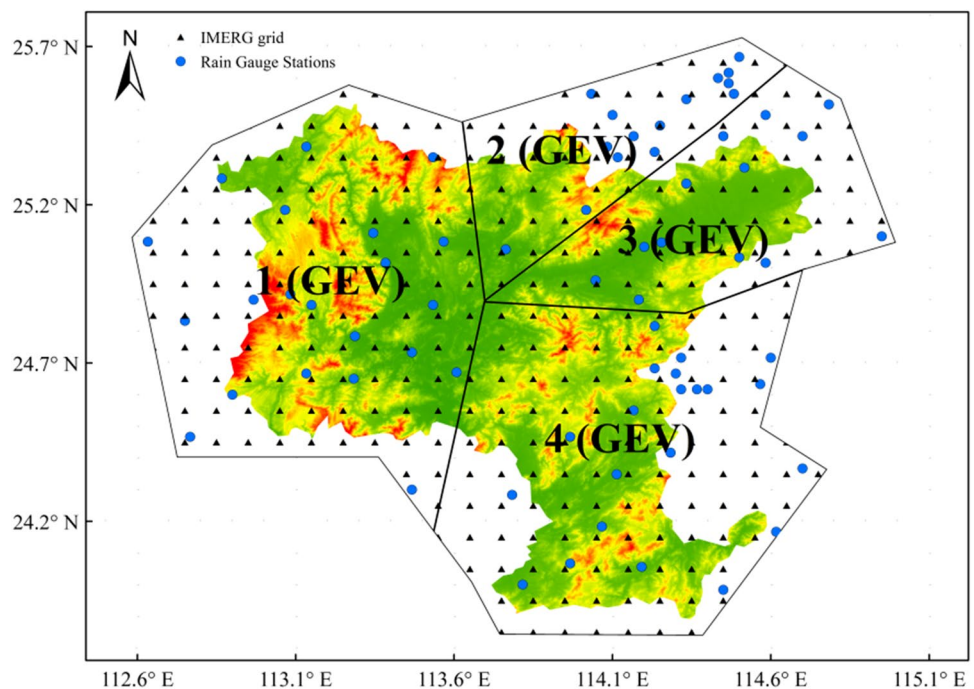


Table 1 Numbers of sites, values of homogeneity measure (H_1), and goodness-of-fit test results, as well as the best-fit distributions for 24-h homogeneous regions of Shaoguan according to observational rainfall data

Homogeneous regions	Number of site	H_1	Z^{DIST}			$\text{RMSE} \times 10^{-5}$						RE score						Best-fit distribution
			GLO	GEV	GNO	GPA	PE3	GLO	GEV	GNO	GPA	PE3	GLO	GEV	GNO	GPA	PE3	
1	21	0.46	0.64	-1.14	-	-	6780	5731	6554	10,080	8642	17	20.5	18.5	7	12	GEV	
2	21	0.97	-	-0.62	-1.17	-	6119	4365	4815	9385	6161	15.5	17	17	9	16.5	GEV	
3	13	0.24	-	-0.05	-0.12	-	8222	5937	6589	10,017	7519	18.5	17	15.5	10	14	GEV	
4	15	0.11	1.61	0.07	-0.62	-	7954	5953	6232	8038	7984	15.5	16	15.5	12	16	GEV	

The best-fitting results of each test are presented in *italic*, and an dash indicates the value of $|Z^{\text{DIST}}|$ exceeds 1.64

and the goodness-of-fit test results of each subregion. The grouping is reasonable in terms of homogeneity because all the values of H_1 are less than 1. In each test, the distributions ranked at the top place in performance of the fit can be considered as the appropriate ones. Taking the three goodness-of-fit test results into consideration, the best-fit distributions for each homogeneous region are determined. Ultimately, the distribution which best fitted for four homogeneous regions is GEV.

4.3.2 According to 24-h AMP from IMERG rainfall

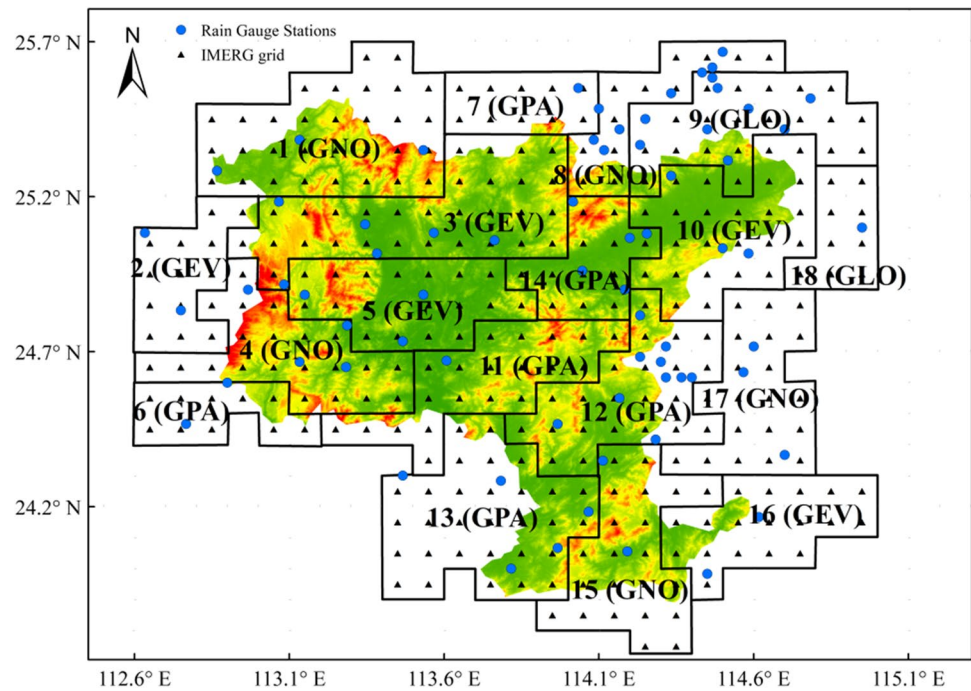
Firstly, referring to the results in Section 4.3.1, same homogeneous regions were used for IMERG. The homogeneity measure (H_1) of each subregion is presented in Table 2. Although the PE3 distribution of region 1 and the GEV, GNO, and PE3 distribution of region 3 past the Monte-Carlo simulation test (the value of $|Z^{\text{DIST}}|$ of these distributions does not exceed 1.64), but the value of H_1 of region 1 was greater than 1 while that of region 3 was less than -3, indicating that region 1 did not pass homogeneous test and region 3 had a strong correlation between the extreme precipitations at grids. The homogeneity of region 2 and region 4 was reasonable, but five commonly used frequency distributions all did not pass the Monte-Carlo simulation test, meaning that it is impossible to select a suitable frequency distribution for these two regions. In addition, because of the sparse distribution of stations, there are fewer stations at the boundaries of four regions. It is uncertain whether to move the grids near the boundary into the region, which affects the final frequency estimation results. Therefore, the homogeneous region of Shaoguan should be re-divided according to the statistical parameters from IMERG rainfall.

Except the criteria mentioned above, it is also necessary to ensure that the rain gauge stations exist in each subregion delineated, for correcting the quantiles at grids with the calculation result at sites later. Finally, 344 IMERG grids were used in this process and Shaoguan area was delineated into 18 subregions for 24-h AMP from IMERG rainfall (Fig. 5). Table 3 also presents the homogeneity measure (H_1) and the best-fit distributions were selected for each homogeneous region. All the values of H_1 are less than 1, most of which are between -2 and -3 because of the strong correlation of gridded rainfall data. Furthermore, the best-fit distributions used for 18 regions were different, such as region 14 and adjacent region 5. In Section 4.2, $L - C_s$ and $L - C_k$ at grids have indicated that the frequency distributions of extreme precipitation in northwestern plain have the characteristics of steep peak and thick tail, which makes region 5 select GEV distribution. For region 14, the mean of AMP at grids of this region is close to that of region 5, but $L - C_v$ and $L - C_s$ are larger while $L - C_k$ is smaller according to Fig. 2,

Table 2 Numbers of grid, values of homogeneity measure (H_1), and Monte-Carlo simulation test results

Homogeneous regions	Number of grid	H_1	Z^{DIST}				
			GLO	GEV	GNO	GPA	PE3
1	110	2.06	–	–	–	–	0.20
2	36	–1.44	–	–	–	–	–
3	48	–3.24	–	–0.49	–0.73	–	–1.61
4	108	–0.90	–	–	–	–	–

The en dash indicates the value of $|Z^{\text{DIST}}|$ exceeds 1.64

Fig. 5 Sketch of the 24-h homogeneous regions of Shaoguan and the best-fit distributions for the homogeneous regions according to 24-h AMP from IMERG rainfall

indicating that the AMP at grids of region 14 is scattered and has larger extreme precipitation; moreover, the peak of distribution is gentle. Thus, the distribution of extreme precipitation in region 14 has a longer and thinner tail, which makes region 14 select GPA distribution. Similarly, the best-fit distribution of other regions was also related to the characteristics of their statistical parameters.

4.4 Rainfall quantiles in Shaoguan area

The rainfall quantiles over 24 h at recurrence intervals of 1 to 10,000 years at sites or grids in Shaoguan area were obtained by using the index-flood procedure. Two calculation results were plotted and compared to each other (Fig. 6). On the whole, the quantiles at grids are slightly higher than those at sites if the return period is less than 40 years, but the situation is opposite over 40 years and the difference between them will increase as return period increases.

According to statistics, the mean of 24-h AMP from observational rainfall or IMERG data is close, but the $L - C_s$ and $L - C_k$ at stations (0.22 and 0.18) are larger than those that at grids (0.17 and 0.12), indicating that the extreme precipitation from observation rainfall is more and stronger than that from IMERG data, possibly because the GPM satellites have strong ability to capture weak precipitation but are poor to detect heavy rainfall. The peak of distribution of stations is steeper, and the tail is thicker than that of grids, which results in larger estimates of stations at higher return periods. Therefore, quantiles at grids should be corrected for reducing their estimation error.

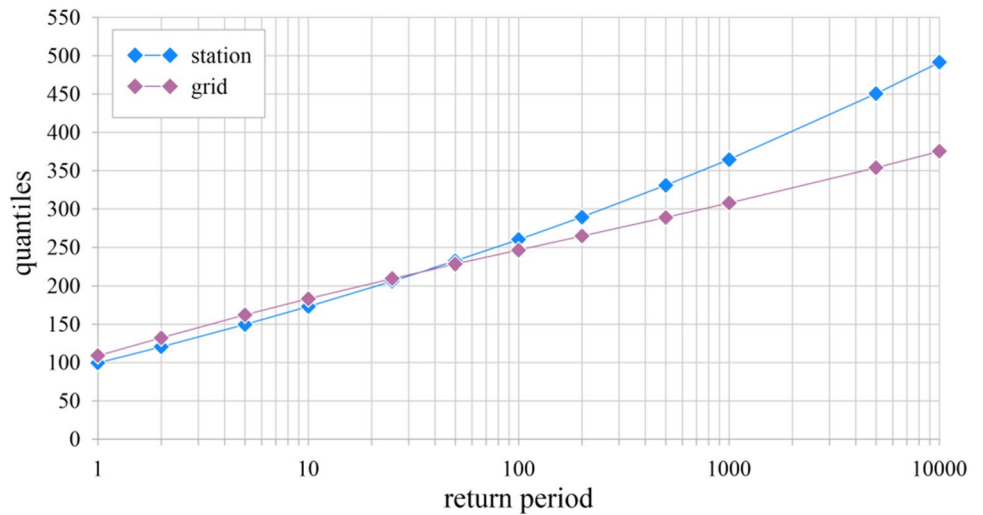
In this paper, the quantiles at grids at recurrence intervals of 2 years were corrected via the MBC method, and the revised results were checked by two approaches. The first is to perform actual data inspection and analysis on the adjusted frequency estimation. For site 59,097 with the longest observation record in the study area, the length of 24-h AMP is 41 a and the maximum observed

Table 3 Numbers of grid, values of homogeneity measure (H_1), and goodness-of-fit test results, as well as the best-fit distributions for 24-h homogeneous regions of Shaoguan according to IMERG rainfall data

Homogeneous regions	Number of grid	H_1	Z^{DIST}			RMSE $\times 10^{-5}$						RE score			Best-fit distribution			
			GLO	GEV	GNO	GPA	PE3	GLO	GEV	GNO	GPA	PE3	GLO	GEV		GNO	GPA	PE3
1	31	-2.86	-	1.43	0.79	-	-0.46	7433	4527	4084	6323	4183	16.5	15.5	14.5	15	13.5	GNO
2	12	-2.83	0.33	-0.87	-1.23	-	-	6407	5759	6506	11,008	8221	17.5	18.5	15	11	13	GEV
3	30	-1.62	-	0.37	0.24	-	-0.41	8111	5821	6258	9774	6777	13.5	18	17	9	17.5	GEV
4	21	-1.99	-	1.07	0.48	-	-0.61	9009	7298	6626	8852	6215	13.5	15.5	18	11.5	16.5	GNO
5	19	-1.74	-	1.02	1.29	-	1.03	10,686	7893	8095	10,126	7954	10	17.5	16	15	16.5	GEV
6	10	-1.44	-	-	-	-1.49	-	12,691	8614	8874	6914	8443	12	14	14	21	14	GPA
7	10	0.05	-	-	-	-0.67	-	15,145	10,363	11,157	6065	10,942	7.5	17.5	17.5	15	17.5	GPA
8	12	-1.07	-	0.54	0.23	-	-0.45	7526	4947	5121	7907	5771	13	12.5	13.5	20	16	GNO
9	21	-2.36	-0.95	-	-	-	-	6174	8309	7984	16,542	8320	20.5	14.5	19	5	16	GLO
10	24	-2.95	-	-0.25	-0.59	-	-1.39	6096	4658	4665	9898	5181	12	16	18.5	13	15.5	GEV
11	20	-0.57	-	-	-	1.34	-	15,536	12,232	11,368	6086	9787	12.5	13	15	18	16.5	GPA
12	22	-2.19	-	-	-	1.38	-	13,397	10,555	9314	5847	7331	13	12	14.5	20.5	15	GPA
13	30	-1.39	-	-	-	-0.14	-	15,152	11,045	11,272	6236	10,777	8	15.5	12.5	22	17	GPA
14	12	-2.49	-	-	-	-1.64	-	15,045	11,087	10,894	5194	9977	8	13	14	22	17.5	GPA
15	22	-2.72	-	0.11	-0.04	-	-0.64	10,418	8622	9224	11,565	9979	16	15	17.5	13.5	13	GNO
16	18	-2.22	-	1.11	-	-	-	9871	6302	7078	10,220	7046	15.5	15.5	15.5	13	15.5	GEV
17	20	-2.85	-	0.79	0.24	-	-0.8	9334	7072	7223	8175	8115	10.5	12	13.5	21	17.5	GNO
18	9	-1.9	-0.24	-0.94	1.4	-	-	6159	6666	7222	9161	11,690	11.5	12.5	16.5	17	17.5	GLO

The best-fitting results of each test are presented in italic, and en dash indicates the value of $|Z^{DIST}|$ exceeds 1.64

Fig. 6 Comparison of quantiles at sites and at grid



rainfall is 306.5 mm. The 2-year estimate at adjacent grid before and after correction is 149.78 mm and 131.93 mm, respectively, with the empirical frequency of 0.31 and 0.45 according to factual data of site 59,097. Obviously, the

corrected empirical frequency is closer to 0.5 which is the theoretical probability of 2-year return period, indicating that the adjusted frequency estimation is more reasonable. Similarly, the empirical frequency at other grids has also

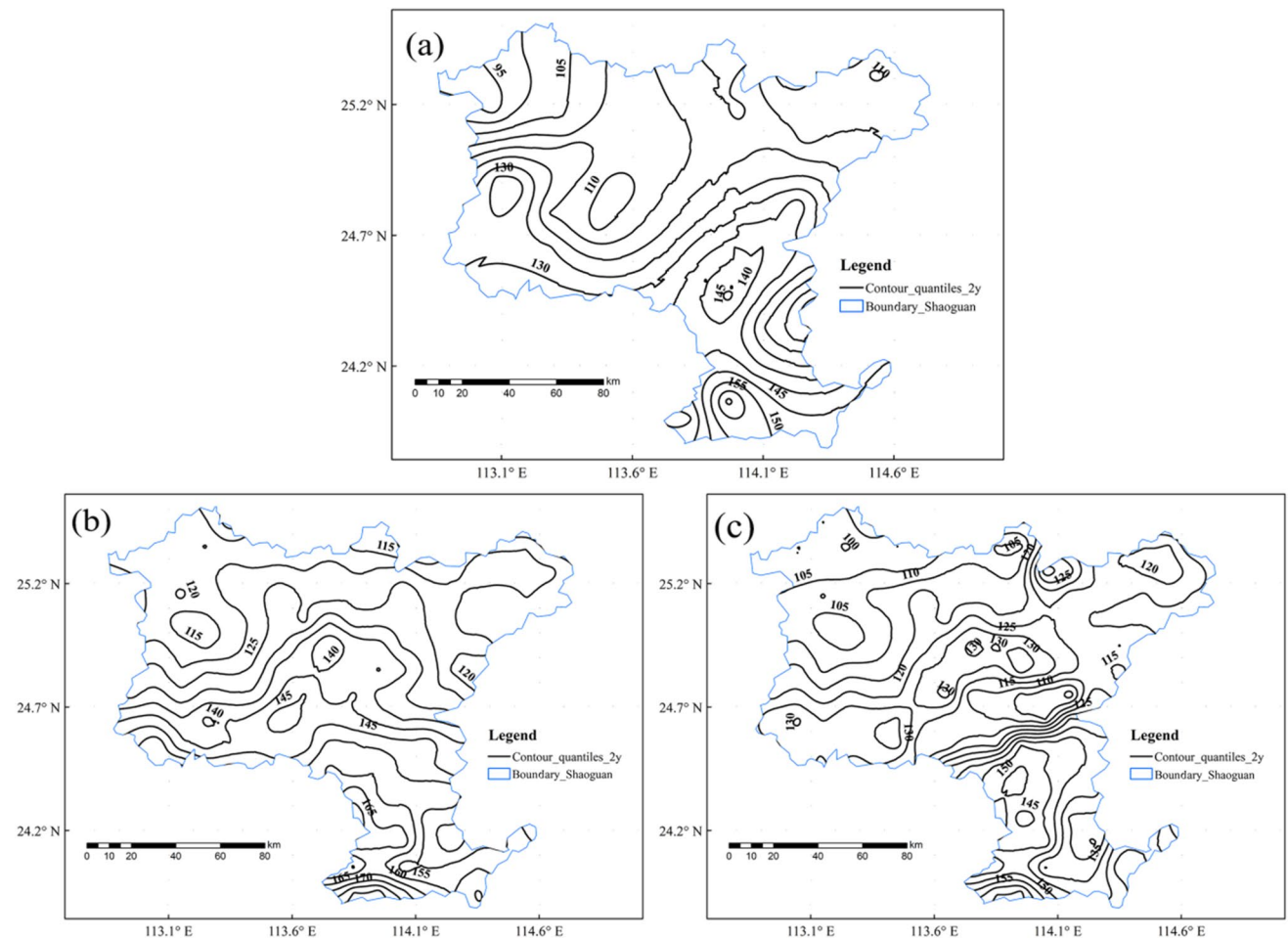


Fig. 7 The contour plots of 2-year estimates for 24 h at sites (a) and at grids (before correction (b) and after correction (c))

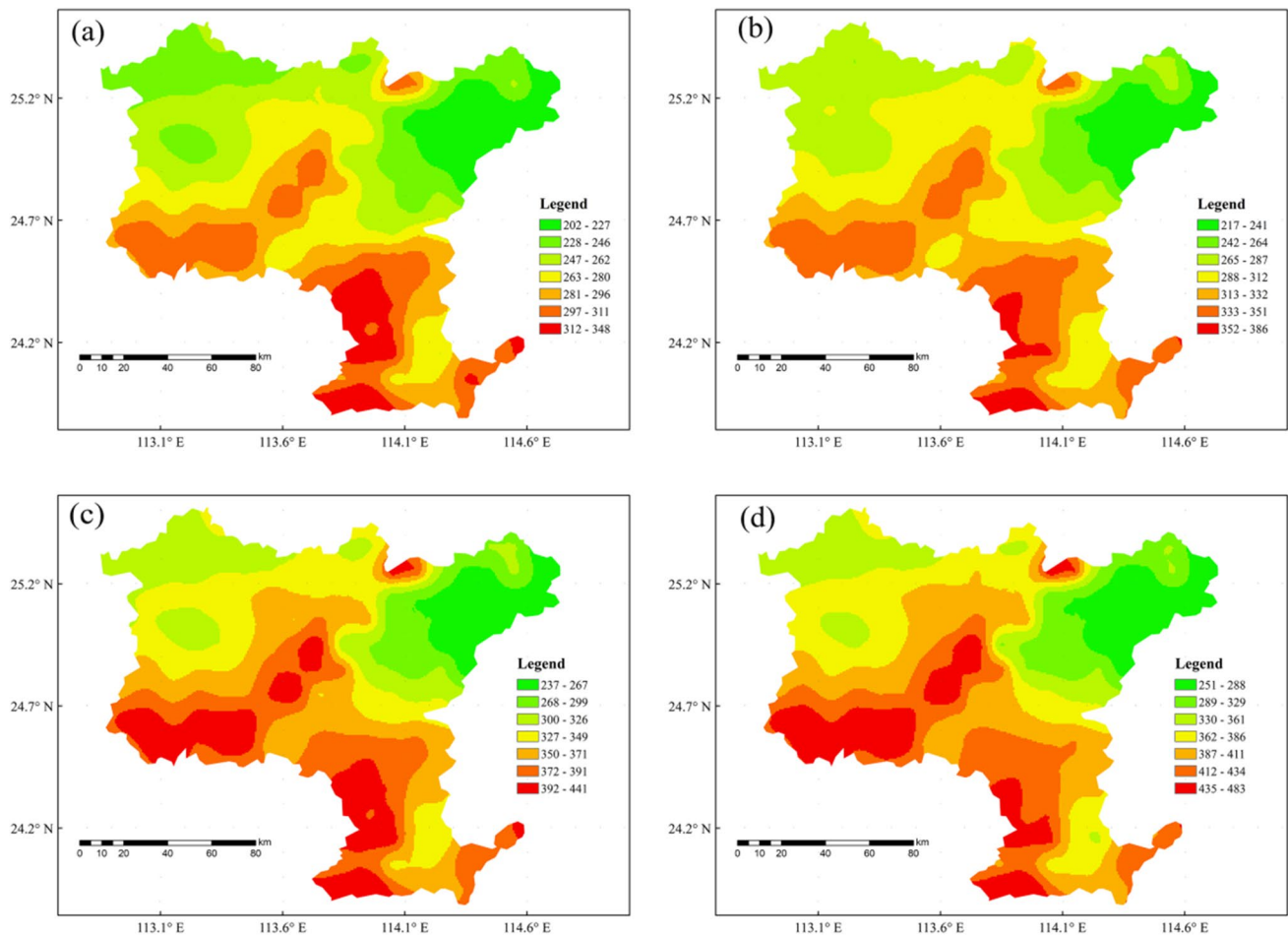


Fig. 8 The distribution map of rainfall quantiles over 24 h at recurrence intervals of 100 years (a), 200 years (b), 500 years (c), and 1000 years (d) after correction

been improved. The second approach is to compare the contour map of 2-year estimates before and after correction. From Fig. 7, the frequency estimation in the central and southern Shaoguan had been significantly decreased after the revision but the precipitation maximum center has not changed. Meanwhile, the precipitation minimum center in east-central and southeastern is more prominent, in which the former did not appear at sites. The corrected results combine the characteristics of quantiles at sites and at grids, meaning that the spatial distribution of frequency estimation had been appropriately adjusted.

Finally, the distribution map of 100-year, 200-year, 500-year, and 1000-year rainfall estimates over 24 h which have been corrected is presented in Fig. 8. It is noticeable that the quantiles increase with the increase of return period, and the spatial distribution of estimates for different return periods is approximately the same, showing a characteristic that estimates in the south were larger than those in the north. The maximum rainfall estimates occur in the southwest to the central, as well as in the

west-central, and the minimum value is distributed in the northeast. Actually, the orientation of prevailing moisture jet in Shaoguan is from south to north; thus, heavy rain is often caused in the southwest to central because of the local topographic forcing. The estimation results are proved to be consistent with reality and are conducive to providing a scientific basis for flood control decisions in the future.

5 Summary and conclusions

In this study, 24-h AMP respectively from observational precipitation and IMERG gridded rainfall was used to estimate 1- to 10,000-year rainfall quantiles for 24 h in Shaoguan area via RLMA. Then, the estimates at grids were compared with that at sites and corrected based on the latter by the MBC method. The findings as well as some critical questions that should be analyzed and/or discussed in further studies are summarized as follows:

1. By using the t test method, the precipitation observation from meteorological stations and hydrological stations is proved to have the same distribution and can be used together in frequency analysis.
2. The mean of AMP at grids is slightly larger than that at stations, the maximums of them are mainly distributed in the southern plains. The $L - C_v$ at stations shows that the AMP from observational rainfall has a large degree of dispersion in western Shaoguan, while the maximum zone of $L - C_v$ at grids is also located in the northwest corner, southwest corner, and the east-central mountainous region. The $L - C_s$ and $L - C_k$ at sites or grids both indicate that northwestern Shaoguan has more and stronger extreme precipitation.
3. For 24-h AMP from observational precipitation, Shaoguan was delineated into 4 regions because the stations used were distributed unevenly and sparsely. If the same regions are used for 24-h AMP from IMERG rainfall, one of the regions has strong correlation, and two of the regions cannot select suitable distribution. Based on the statistical parameters at grids, the number of homogeneous regions for IMERG is determined to be 18 finally.
4. For stations, all regions select GEV as the best-fit distribution. But for IMERG, the frequency distributions chosen for regions are different. For example, region 5 selects GEV and region 14 selects GPA because $L - C_v$ and $L - C_s$ of region 14 are larger while $L - C_k$ is smaller, indicating that the distribution for region 14 has thinner tail.
5. Compared with the estimates at sites, quantiles calculated from IMERG rainfall were underestimated if return period is greater than 100 years, and the difference will increase as return period increases. The reason is that the $L - C_s$ and $L - C_k$ at sites are larger than those at grids, which means that the extreme precipitation from observation rainfall is more and stronger than that from IMERG data. Thus, GPM satellite may be poor to detect heavy rainfall.
6. Taking 2-year estimates at grids as an example, the MBC method was proved to combine the characteristics of the quantiles at sites and at grids well, with the adjusted frequency estimation that is more reasonable and the precipitation center having no change.
7. The spatial distribution of 100-year, 200-year, 500-year, and 1000-year revised rainfall quantiles over 24 h at grids is similar. The rainstorm high-risk area of Shaoguan is in the southwest to the central, as well as in the west-central, which should be paid special attention in disaster prevention and reduction.

At present, there are still many issues regarding the density and coverage of rain gauges that affect the accuracy of univariate L-moments method, especially in

remote areas of China or most parts of the world. Therefore, the main purpose of this study is to explore and evaluate the satellite precipitation data applied in the univariate L-moments method. The result shows that the frequency estimation calculated by satellite products with the univariate L-moments method is applicable and more reasonable. In the future, carrying out more further studies on multivariate L-moments which had replaced the univariate one and attracted the attention of many researchers all over the world today should be considered (Chebana and Ouarda 2007; Chebana et al. 2009; Wu et al. 2017; Zhang et al. 2015), using satellite precipitation data, as well as other parameters that affect the rainfall, such as the temperature and relative humidity.

Acknowledgements We acknowledge the Guangdong Hydrology Bureau and Shaoguan Meteorological Bureau for providing the precipitation observations. We acknowledge the National Aeronautics and Space Administration (NASA) for providing the Integrated Multi-satellite Retrievals for GPM (IMERG) gridded rainfall products.

Author contribution XYZ was the primary designer of this study, produced the figures, and wrote the paper. DHZ contributed to the conceptualization, methodology, and validation. MW and YFL provided input to the figure and text editing.

Funding This research was supported by the National Natural Science Foundation of China (Grant No. 42075190) and the Science Research Project of Guangdong Meteorological Bureau (Grant No. GRMC2019M40).

Data availability The observational precipitation sharing is not applicable to this article. The IMERG rainfall products are available on <https://disc.gsfc.nasa.gov/datasets?keywords=GPM&page=1>. The frequency estimates generated during the current study are available from the corresponding author on reasonable request.

Code availability Datasets were analyzed using MATLAB and FORTRAN. The figures were displayed with ArcGIS.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication All authors have approved the manuscript and gave their consent for submission and publication.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahmed K, Shahid S, Ali RO et al (2019) Evaluation of the performance of gridded precipitation products over Balochistan Province, Pakistan. *Desalin Water Treat* 79:73–86
- Bobée B, Rasmussen PF (1995) Recent advances in flood frequency analysis. *Rev Geophys* 33:1111–1116
- Bonnin GM, Martin DL, Lin B, et al (2012) NOAA Atlas 14 volume 1 version 4.0, Precipitation-Frequency Atlas of the United States, Alaska Central Meteorological Observatory (2021a) Heavy precipitation typhoon “Chapaka” in Henan, Hebei and other places continues to affect the coast of South China. http://www.cma.gov.cn/2011xwzx/2011xqxw/2011xzytq/202107/t20210721_581154.html. Accessed 18 Jan 2022
- Central Meteorological Observatory (2021b) The timeline of rainstorm in Zhenzhou: 28 early warning within 4 days, emergency response level 3. https://view.inews.qq.com/k/20210722A0758500?web_channel=wap&openApp=false. Accessed 18 Jan 2022
- Chebana F, Ouarda T (2007) Multivariate L-moment homogeneity test. *Water Resour Res* 43(8):199–212
- Chebana F, Ouarda TBMJ, Bruneau P (2009) Multivariate homogeneity testing in a northern case study in the province of Quebec, Canada. *Hydrol Process* 23(12):1690–1700
- Chen X, Lin B, Wu J et al (2014) Application of hydrometeorological regional L-moments method to storm frequency analysis in Guangxi. *Water Resour Power* 32:5–9 (In Chinese)
- Chen X, Zhong R, Wang Z et al (2017) Evaluation on the accuracy and hydrological performance of the latest-generation GPM IMERG product over South China. *J Hydraul Eng* 48(10):1147–1156 (In Chinese)
- China Meteorological Administration (2020) Shaoguan: heavy rain breaks records and classes are suspended in many places. http://www.cma.gov.cn/2011xwzx/2011xqxw/2011xjctz/202006/t20200610_555641.html. Accessed 18 Jan 2022
- Ding H, Liao Y, Lin B (2018) High risk flash flood rainstorm area mapping and its application in Jiangxi Province, China. *MATEC Web Conf* 246(2):01087
- Fowler HJ, Kilsby CG (2010) A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. *Int J Climatol* 23(11):1313–1334
- Gjertsen U, Šálek M, Michelson DB (2003) Gauge adjustment of radar-based precipitation estimates in Europe. *Proceedings of Erad*
- Greenwood JA, Landwehr JM, Matalas NC et al (1979) Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water Resour Res* 15:1049–1054
- Hailegeorgis TT, Thorolfsson ST, Alfredsen K (2013) Regional frequency analysis of extreme precipitation with consideration of uncertainties to update IDF curves for the city of Trondheim. *J Hydrol* 498(498):305–318
- Hosking JRM (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J R Stat Soc Ser B Methodol* 52(1):105–124
- Hosking JRM, Wallis JR (1993) Some statistics useful in regional frequency analysis. *Water Resour Res* 29:271–281
- Hosking JRM, Wallis JR (1998) *Regional frequency analysis: an approach based on L-moments*. Cambridge University Press, New York
- Jing F, Yan D, Li C, et al (2014) Regional frequency analysis of extreme precipitation after drought events in the Heihe River Basin, Northwest China. *J Hydrol Eng* 19(6):1101–1112
- Li M, Ao T, Li X (2019) Regional frequency analysis in hydrological frequency analysis of Sichuan Province. *Southwest China. J Agric Sci* 32:1938–1943 (In Chinese)
- Liang Y, Liu S, Zhong G et al (2013) Comparison between conventional moments and L-moments in rainfall frequency analysis for Taihu Lake Basin. *J China Hydrol* 33:16–21 (In Chinese)
- Lin B (2010a) Research and application of hydrometeorological regional L-moment in flood control design standards. *Proceedings of 2010a CHES Annual Conference*, Guiyang. (In Chinese)
- Lin B (2010b) Study and application of hydrometeorological regional L-moments analysis method on flood control standards. *Proceedings of the 2010b Chinese Hydraulic Engineering Society Annual Conference*. (In Chinese)
- Lin B, Bonnin GM, Martin DL et al (2006) Regional frequency studies of annual extreme precipitation in the united states based on regional L-moments analysis. *World Environ Water Resour Congress*
- Lin B, Shao Y, Yan G et al (2012) Development of hydrometeorology in promoting research on the core project of engineering hydrological calculation. *Proceedings of 2012 Chinese Symposium on Hydrology*, Nanjing. (In Chinese)
- Liu M, Yin Y, Han C et al (2018) Frequency analysis of extreme precipitation in Jiangxi Province based on regional L-moments method. *Water Resour Power* 36:1–5 (In Chinese)
- Modarres R (2010) Regional dry spells frequency analysis by L-moment and multivariate analysis. *Water Resour Manag* 24(10):2365–2380
- Norbiato D, Borga M, Sangati M et al (2007) Regional frequency analysis of extreme precipitation in the eastern Italian Alps and the August 29, 2003 flash flood. *J Hydrol Amsterdam* 345(3–4):149–166
- Sarmadi F, Shokoochi A (2014) Regionalizing precipitation in Iran using GPCP gridded data via multivariate analysis and L-moment methods. *Theor Appl Climatol* 118(1–2):1–8
- Shang H, Lin B, Xu L (2019) Visualization of precipitation frequency atlas & rainstorm high-risk regionalization atlas. *China Water Wastew* 35(05):131–138 (In Chinese)
- Shao Y, Wu J, Li M (2016) Frequency analysis of extreme precipitation in Huaihe River basin based on hydrometeorological regional L-moments method. *J China Hydrol* 36:16–23 (In Chinese)
- Wu P (2017) The intensity of extreme precipitation will increase in most regions of the world. http://www.cma.gov.cn/2011xwzx/2011xqxk/qxkjgqy/201705/t20170519_412596.html. Accessed 20 Jan 2022
- Wu J, Lin B, Shao Y (2015) Application of regional L-moments analysis method in precipitation frequency analysis for Taihu Lake Basin. *J China Hydrol* 35:15–22 (In Chinese)
- Wu S, Zhang X, She D (2017) Joint pattern of seasonal hydrological droughts and floods alternation in China’s Huai River Basin using the multivariate L-moments. *J Earth Syst Sci* 126(4)
- Wu Y, Zhang Z, Jin Q et al (2019) Accuracy evaluation of GPM precipitation product in Yangtze River basin. *Yangtze River* 50(9):77–85,152. (In Chinese)
- Yin Y, Chen H, Xu C et al (2016) Spatio-temporal characteristics of the extreme precipitation by L-moment-based index-flood method in the Yangtze River Delta region, China. *Theor Appl Climatol* 124(3–4):1005–1022.ta
- Zhang Q, Qi T, Singh VP et al (2015) Regional frequency analysis of droughts in China: a multivariate perspective. *Water Resour Manag* 29(6):1767–1787

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.