ORIGINAL PAPER

Precipitation extremes in the Yangtze River Basin, China: regional frequency and spatial–temporal patterns

Yongqin David Chen · Qiang Zhang · Mingzhong Xiao · Vijay P. Singh · Yee Leung · Luguang Jiang

Abstract Regional frequency analysis and spatial–temporal patterns of precipitation extremes are investigated based on daily precipitation data covering 1960–2009 using the index-flood L-moments method together with some advanced statistical tests and spatial analysis techniques. The results indicate that: (1) the entire Yangtze River basin can be divided into six homogeneous regions in terms of extreme daily precipitation index. Goodness-of-fit test indicates that Pearson type III (PE3, three parameters), general extreme-value (GEV, three parameters), and general normal (GNO, three parameters) perform well in fitting regional precipitation extremes; (2) the regional growth curves for each homogeneous region with 99 % error bands show that the quantile estimates are reliable enough and can be used when return periods are less than 100 years, and the results indicate that extreme precipitation events are highly probable to occur in regions V and VI, and hence higher risk of floods and droughts; and (3) spatial patterns of annual extreme daily precipitation with return period of 20 years indicate that precipitation amount increases gradually from the upper to the lower Yangtze River basin, showing higher risks of floods and droughts in the middle and lower Yangtze River basin, and this result is in good agreement with those derived from regional growth curves.

Y. D. Chen · Y. Leung
Department of Geography and Resource Management, and
Institute of Environment, Energy and Sustainability, The Chinese
University of Hong Kong, Hong Kong, China

Q. Zhang (✉) · M. Xiao
Department of Water Resources and Environment, Sun Yat-sen
University, Guangzhou 510275, China
e-mail: zhangq68@mail.sysu.edu.cn

Q. Zhang · M. Xiao
Key Laboratory of Water Cycle and Water Security in Southern
China of Guangdong High Education Institute, Sun Yat-sen University, Guangzhou 510275, China

Q. Zhang · M. Xiao
School of Geography and Planning, and Guangdong Key
Laboratory for Urbanization and Geo-simulation, Sun Yat-sen
University, Guangzhou, People's Republic of China

V. P. Singh
Department of Biological and Agricultural Engineering and
Department of Civil and Environmental Engineering, Texas A & M
University, College Station, TX 77843-2117, USA

L. Jiang
Institute of Geographic Sciences and Natural Resources Research,
Chinese Academy of Sciences, Beijing 100101, China

1 Introduction

Changes in precipitation extremes in both time and space have a strong influence on human society. Higher frequencies of weather extremes (e.g., Easterling et al. 2000; Alan et al. 2003) pose a potential threat to the at-risk populations. Tropical storms, floods, and droughts affect human welfare directly through catastrophic damage or indirectly through adverse effects on crop productivity (e.g., Huntington 2006). In recent years, climate changes and related impacts on human society, particularly on agriculture and food security, have received an unprecedented attention (e.g., Schmidhuber and Tubiello 2007; MacDonald 2010; Piao et al. 2010) because of the significant role of the availability, accessibility, and security of food and water in the stability and sustainability of human society. Moreover, investigation of precipitation changes could be the first step into thorough understanding of variations of hydrological cycles in the backdrop of global climate changes (e.g., Allen and Ingram 2002; Alan et al. 2003; Dore 2005). All these could be the reasons that the precipitation changes have drawn over-whelmingly increasing concerns from researchers around the world in recent decades (Fatichi and Caporali 2009), and it is particularly true for the precipitation extremes

(Easterling et al. 2000; Emori and Brown 2005; Zhang et al. 2011, 2012).

The Yangtze River (Fig. 1), being the longest river in China and the third longest river in the world, plays an important role in the sustainable development in China. However, frequent floods in the Yangtze River basin inflicted considerable loss on economy and human life. In 1998, disastrous floods occurred to large parts of the basin, which was the largest flood since 1954. The economic loss was 166 billion RMB (nearly 28 billion US dollars). This flood hazard was the direct result of unusually high precipitation that occurred between June and August (670 mm) due to a strong El Niño event (Yin and Li 2001). Therefore, more and more investigations have been conducted on floods and the possible mechanisms behind (e.g., Zhang et al. 2006, 2007). Extreme precipitation events are the main causes for the flood hazards in the Yangtze River basin (Zhang et al. 2005).

The southern part of the Yangtze River basin is climatically close to tropical climate and the northern part to the temperate zone. The annual mean temperature in the southern and northern parts of the middle and lower Yangtze River basin is roughly 19 and 15 °C, respectively; and the mean annual precipitation in the basin varies from 300 to 500 mm in the western region to 1,600–1,900 mm in the southeastern region and the precipitation is mostly concentrated in the summer season (from June to August), accounting for nearly half of the annual total precipitation amount (Guo et al. 2012). In this regard, uneven distribution of precipitation changes is one of the major causes behind the frequent floods and droughts across the Yangtze River basin. Therefore, regionalization of precipitation extremes and the related statistical behaviors do definitely merit further study. There are some studies pertaining to the precipitation extremes in

the Yangtze River basin (e.g., Su et al. 2005). Zhang et al. (2008) analyzed the changing characteristics of precipitation maxima using Mann–Kendall trend test and explored the possible causes for the changes by using NCEP/NCAR reanalysis dataset. The research results indicate changes of precipitation maxima from relatively stable patterns to the significant increasing/decreasing trend in the mid-1970s. With respect to the intra-annual variability, the rainy days are decreasing and precipitation intensity is increasing, and significant increasing trend of precipitation intensity was detected in the middle and lower Yangtze River basin.

However, the regional frequency analysis of precipitation extremes with the state-of-art L-moments techniques has not been conducted in the whole Yangtze River Basin, except similar study in the Pearl River basin (Yang et al. 2010), as a highly complicated river system that encompasses a large area with various random and systematic variations so far. This is definitely instrumental in terms of sound understanding of spatial and temporal variations of precipitation extremes, and is also greatly helpful for understanding of occurrence of floods and droughts in both space and time. Therefore, efforts should be made to carry out the regional frequency analysis of extreme precipitation in the basin and this is the major motivation of this study.

The major objectives of this study are to: (1) identify and delineate the hydrological homogeneous regions for precipitation extremes defined in this study; (2) determine the best probability distribution for rainfall extremes, conduct regional frequency analysis with uncertainty assessment including the corresponding error bounds and root mean squared error using the L-moments; and (3) characterize the spatial–temporal patterns of extreme precipitation events in order to reveal the underlying impacts of climate variations dominated in the Yangtze River basin. The study results will
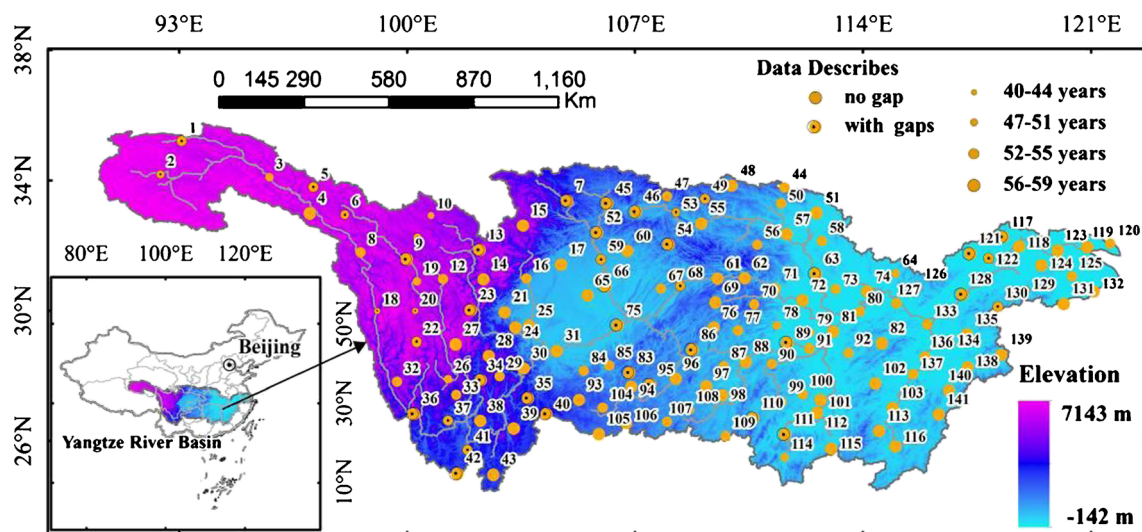

Fig. 1 Locations of the Yangtze River basin and selected meteorological stations

definitely benefit further development of understanding of the unique and complex features of extreme rainfall in the Yangtze River basin, the largest river basin in China, and hence the basin-scale water resources management in the backdrop of changing climate.

2 Data

Daily precipitation data covering 1960–2009 were collected at 141 precipitation stations and the precipitation dataset has been provided by the National Climate Center of the China Meteorological Administration. The locations of the stations are shown in Fig. 1. There are 42 precipitation stations containing missing daily precipitation data, as illustrated in circles with dots (Fig. 1). We excluded the stations with missing values for a month or longer and the largest missing data rate is only 1.09 %. The missing data were filled by the long-term average of the same days of other years. Given the objectives of this study, the gap-filling method, which was similarly employed by Zhang et al. (2011), does not significantly affect the final results. As for precipitation extremes, Alexander et al. (2006) has set an exact definition and formula for 27 extreme daily precipitation and temperature indices and there are 11 extreme daily precipitation indices, involving five different categories that are percentile-based indices, absolute indices, threshold indices, duration indices, and so on. In this study, four indices are selected: (1) the annual number of heavy precipitation days with precipitation >10 mm (R10); (2) the annual amount of rainfall exceeding the 95th (R95P) percentiles, and the 95th percentile of precipitation was defined on the wet days (defined as the rainy day with daily precipitation larger than 1 mm) in a 30-year period, and is selected during the period of 1971–2000; (3) maximum 1-day precipitation amount (RX1DAY); and (4) maximum 5-day precipitation amount (RX5DAY).

3 Methodology

Regional frequency analysis based on L-moment method has been used to characterize the spatial pattern of precipitation extremes within the Yangtze River basin. The methods are presented in this section, including the stationarity test, serial correlation check, L-moments approach, regional frequency analysis based on the L-moment, and spatial analysis.

3.1 Stationarity test and serial correlation check

Both stationarity and independence are important underlying assumptions inherent in frequency analysis, thus the analysis without stationarity and serial correlation tests may lead to incorrect results and conclusions (Yang et al. 2010).

Therefore, it is beneficial to conduct stationarity and serial correlation tests prior to the regional frequency analysis. A stationary series is usually defined as constant mean, variance and autocorrelation, etc., and in this paper we mean a flat-looking series. The trend test is one of the most popular methods for examining the stationarity in hydrological series. With the advantage of not requiring any distribution assumptions in the data while having the same power as its parametric counterparts, the rank-based Mann–Kendall method (Mann 1945; Kendall 1975) is recommended by the World Meteorological Organization for assessing the significance of monotonic trends in hydrological series (Mitchell et al. 1966). The effect of the serial correlation on the Mann–Kendall (MK) test was eliminated using the pre-whitening technique developed by Yue et al. (2002) in this study.

The serial correlation check was carried out mainly by examining the autocorrelation coefficients of the time series. When the autocorrelation coefficients of different lag times, calculated for a time series, are within the 95 % confidence level, the observations in this time series can be accepted as being independent from each other. In this study according to the calculated autocorrelation coefficients of lag-1, for each annual series, the observations in that series can be accepted as being independent at the 95 % significance level. In the paper, for the stations that are not stationary or independent are excluded in the dataset for the regional frequency analysis of extreme daily precipitation index.

3.2 L-moments theory

L-moments have the theoretical advantages over conventional moments for being able to characterize a wider range of distributions and being more robust to the presence of outliers in the data. Details about the L-moments approach can be found in Hosking and Wallis (1997). In brief, it is a modification of the probability weighted moments method with the advantage of offering a description of the shape of a probability distribution by L-skewness and L-kurtosis (Yang et al. 2010). The L-moment is a linear combination of the probability weighted moments. The sample L-moment ratios are defined as

$$t = l_2/l_1 \text{ and } t_r = l_r/l_2, \ r = 3, 4, \dots \quad (1)$$

where l_r being the unbiased rth L-moments, being analogue to the traditional ratios, that is, t is the coefficient of variation (L-CV), t_3 the L-skewness, and t_4 the L-kurtosis. The L-moment ratios will be used for homogeneity analysis in the regional frequency analysis.

Compared with conventional moments, L-moments have less bias in estimation and their asymptotic are closer to the normal distribution in finite samples. And the L-moments

approach covers the characterization of probability distributions, the summary of observed data samples, the fitting of probability distributions to data, and the testing of the distributional form (e.g., Yang et al. 2010).

3.3 Regional frequency analysis based on L-moments method

Suppose that there are N sites in a region with sample size n_1, $n_2, \ldots, n_N$, respectively. And the sample L-moment ratios (L-CV, L-skewness, and L-kurtosis) at site i are denoted by $t^{(i)}$, $t_3^{(i)}$, and $t_4^{(i)}$, respectively. Then the regional weighted average L-moment ratios are given by:

$$\bar{t} = \sum_{i=1}^{N} n_i t^{(i)} \bigg/ \sum_{i=1}^{N} n_i \text{ and } \bar{t}_r = \sum_{i=1}^{N} n_i t_r^{(i)} \bigg/ \sum_{i=1}^{N} n_i \, \mathrm{r} = 3, 4, \ldots \ (2)$$

The regional frequency analysis using L-moments consists of five steps (Hosking and Wallis 1993, 1997): (1) identification of homogenous regions by cluster analysis; (2) screening of the data using the discordancy measure D_i; (3) homogeneity testing using the heterogeneity measure H; (4) distribution selection using the goodness-of-fit measure Z; and (5) regional estimation of precipitation quantiles using the L-moment approach. These five steps were followed to conduct a regional frequency analysis for the Yangtze River basin and the statistical methods employed are introduced below.

3.3.1 Identification of homogenous regions by cluster analysis

In this study, five variables are adopted to characterize the precipitation changes, i.e., latitude, longitude, elevation, the mean annual precipitation, and the standard deviation of annual precipitation. And then the Fuzzy c-mean based on the five variables was used to cluster the stations in the Yangtze River basin with the number of clusters identified based on the cluster validity index for the fuzzy c-mean as proposed by Wu and Yang (2005).

The output from the cluster analysis is not the final results. Subjective adjustments can often be found to improve the physical coherence of regions and to reduce the heterogeneity of regions as measured by the heterogeneity measure. Several adjustments of regions may be recommended (Hosking and Wallis 1997): (1) move a site or a few sites from one region to another; (2) delete a site or a few sites from the dataset; (3) subdivide the region; (4) break up the region by reassigning its sites to other regions; (5) merge the region with another or others; (6) merge two or more regions and redefine groups; and (7) obtain more data and redefine groups.

3.3.2 Screening the data using the discordancy measure

Let $u_i = [t^{(i)}, t_3^{(i)}, t_4^{(i)}]^T$ be the vector containing the t, t_3, and t_4 values for site i where the superscript T denotes transposition of a vector or matrix. And the (unweighted) regional average is defined as (Hosking and Wallis 1993):

$$\bar{u} = \sum_{i=1}^{N} u_i \bigg/ N \tag{3}$$

Then the discordancy measure for site i is defined as (Hosking and Wallis 1993):

$$D_i = \frac{1}{3} N \left(u_i - \bar{u} \right)^T A^{-1} \left(u_i - \bar{u} \right) \tag{4}$$

where $A = \sum_{i=1}^{N} \left(u_i - \bar{u} \right)\left(u_i - \bar{u} \right)^T$.

Obviously, a large value of D_i indicates the discordancy of site i with other sites. Hosking and Wallis (1997) found that there was no single fixed number which can be considered to be a "large" D_i value and suggested some critical values for discordancy test which are dependent on the number of sites in the study region.

3.3.3 Homogeneity testing using the heterogeneity measure

The regional average L-CV, L-skewness, and L-kurtosis, represented by t^R, t_3^R, and t_4^R, respectively, are computed as (Hosking and Wallis 1993; Yang et al. 2010):

$$t^R = \sum_{i=1}^{N} n_i t^{(i)} \bigg/ \sum_{i=1}^{N} n_i t_3^R = \sum_{i=1}^{N} n_i t_3^{(i)} \bigg/ \sum_{i=1}^{N} n_i t_4^R$$

$$= \sum_{i=1}^{N} n_i t_4^{(i)} \bigg/ \sum_{i=1}^{N} n_i \tag{5}$$

where $n_i / \sum_{i=1}^{N} n_i$ denotes the weight applied to sample L-moment ratios at site i. And the regional average mean l_1^R is, without loss of generality, set to be 1. As Hosking and Wallis (1997) found that H_2 is a weaker test of heterogeneity than H_1, the measures of dispersion in this study is defined as (Hosking and Wallis 1997; Yang et al. 2010):

$$V_1 = \left\{ \sum_{i=1}^{N} n_i \left[t^{(i)} - t^{(R)} \right]^2 \bigg/ \sum_{i=1}^{N} n_i \right\}^{1/2} \tag{6}$$

Let μ_v and σ_v denote the mean and standard deviation values of V_1 computed on the basis of a large number of

simulated homogeneous regions. Hence the statistic that measures the heterogeneity of a set of sites is given by:

$$H_1 = \frac{(V_1 - \mu_v)}{\sigma_v} \tag{7}$$

In order to obtain reliable values of μ_v and σ_v, the number N_{sim} of simulations needs to be large and $N_{sim} = 1,000$ was used in this study, and these were calculated using the R package lmomRFA maintained by Hosking (2012a). The region is regarded to be "acceptably homogeneous" if $H_1 < 1$, "possibly heterogeneous" if $1 \leq H_1 < 2$, and "definitely heterogeneous" if $H_1 > 2$. In this study a region has been accepted as homogeneous if $H_1 < 2$.

3.3.4 Distribution selection using the goodness-of-fit measure

For each candidate distribution, the goodness-of-fit measure:

$$Z^{DIST} = \left(\tau_4^{DIST} - \bar{t}_4 + \beta_4 \right) \Big/ \sigma_4 \tag{8}$$

was used, as suggested by Hosking and Wallis (1993) using the L-kurtosis, where τ_4^{DIST} is the L-kurtosis of the fitted distribution to the data using the candidate distribution, and:

$$\beta_4 = \sum_{m=1}^{N_{sim}} \left(\bar{t}_4^m - \bar{t}_4 \right) \Big/ N_{sim} \tag{9}$$

is the bias of $\bar{t}_4$ estimated using the simulation technique as before with $\bar{t}_4^m$ being the sample L-kurtosis of the mth simulation, and:

$$\sigma_4 = \left\{ (N_{sim} - 1)^{-1} \left[\sum_{m=1}^{N_{sim}} \left(\bar{t}_4^m - \bar{t}_4 \right) - N_{sim} \beta_4^2 \right] \right\}^{\frac{1}{2}} \tag{10}$$

is the standard deviation of $\bar{t}_4$. The fit is considered to be adequate if $|Z^{DIST}|$ is sufficiently close to zero, and a reasonable criterion being $|Z^{DIST}| \leq 1.64$.

As the goodness-of-fit test allows only to decide to keep a distribution or not but does not allow to select among the distributions kept, then if more than one candidate distribution is acceptable, the L-moment ratio diagram has been used to identify the distribution by comparing its closeness to the L-skewness and L-kurtosis combination in the L-moment ratio diagram furthermore.

3.4 Spatial interpolation

To understand the spatial patterns of statistical characteristics of extreme precipitation regimes across the Yangtze River basin, the geostatistical or stochastic methods were used as they capitalize on the spatial correlation between neighboring observations to predict attributed values at unsampled locations (Goovaerts 1999). And Goovaerts (1999) indicated that the major advantage of the Kriging method over other simpler interpolation methods is that sparsely sampled observations of the primary attribute can be complemented by secondary attributes that are more densely sampled. Therefore, the Kriging interpolation method has been used in this study to characterize the spatial patterns of the extreme precipitation for the study region.

4 Results and discussion

4.1 Regionalization of precipitation extremes based on L-moment technique

The topographical features and the climate types of the Yangtze River basin are complex. The elevation of the Yangtze River basin is gradually decreasing from the northwest to the southeast and the river basin is generally affected by the southeast monsoon with the southern part of the basin is climatically close to tropical climate and the northern part to the temperate zone. The precipitation changes are far from being spatially homogeneous, so different precipitation climate patterns were identified. In this case, the homogeneous regions were first differentiated using the fuzzy c-means clusters in terms of the latitude, longitude, elevation, mean annual precipitation, and the standard deviation of annual precipitation. Generally, six homogeneous regions were identified based on the cluster validity index for the fuzzy c-mean proposed by Wu and Yang (2005) and the results of the cluster validity index are not shown in the paper.

As suggested by Yang et al. (2010), it is beneficial to make stationarity and serial correlation tests prior to the regional frequency analysis. And the Mann–Kendall test was conducted on the observations of precipitation extremes (R10, R95P, RX1DAY, and RX5DAY) for all of the stations in the study basin. For the stations with trends being significant at >95 % confidence level are considered as not stationary and excluded from the dataset analyzed in this study and the results of stations with significant trend at >95 % confidence level are shown in Fig. 2. Besides, the stations are considered as not independent if the autocorrelation coefficients of lag-1 are beyond the 95 % confidence level and the results of the stations, considered as not independent are also shown in Fig. 2. Similarly, for the stations considered as not independent will also be removed from the dataset for analysis.

Then the discordance test was done for each homogeneous region of R10, R95P, RX1DAY, and RX5DAY, respectively. Results of the initial cluster show that there are some sites being discordant with other sites within a homogeneous region. For these discordant stations, subjective adjustments are often necessary to improve the physical coherence of regions and to reduce the heterogeneity of
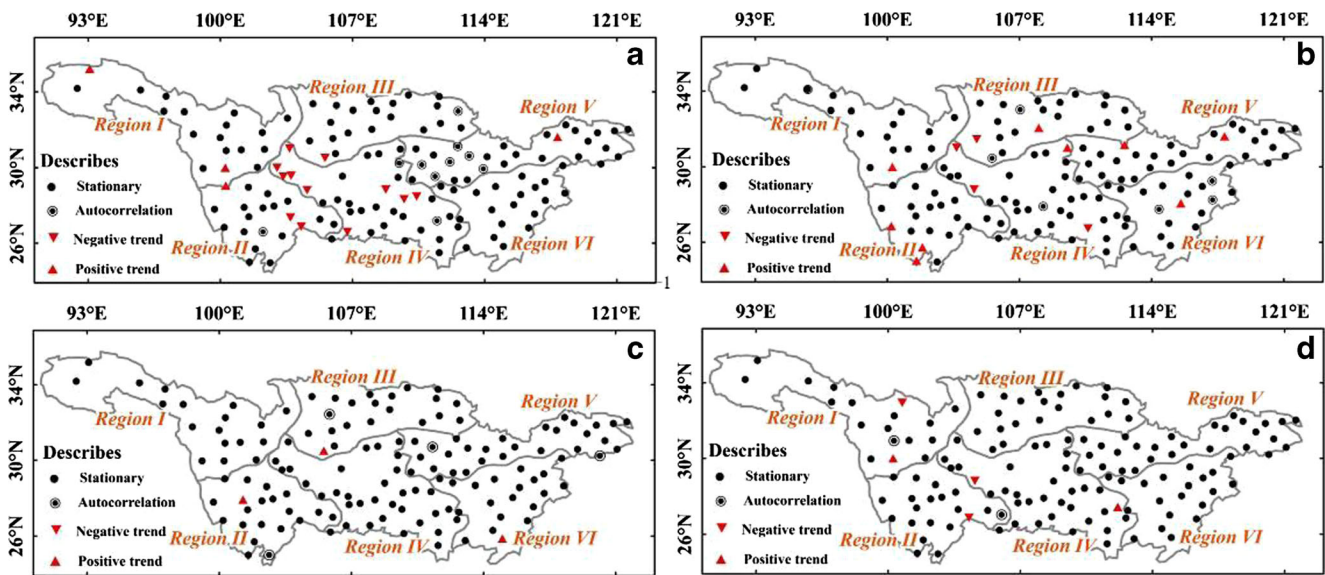
Fig. 2 Homogeneous regions clustered by the fuzzy *c*-means and the stations being not stationary or independent, where **a** for R10; **b** for R95P; **c** for RX1DAY, and **d** for RX5DAY

regions as measured by the heterogeneity measure, and these can be done as illustrated in Section 3.3.1. After several adjustments, the final results of the identified homogeneous regions for each extreme precipitation index are obtained and illustrated in Fig. 3, which shows that the entire Yangtze River basin can be categorized into six homogeneous regions for each extreme daily precipitation index. The homogeneity measure H_1 defined in Eq. 7 was calculated for the detected homogeneous regions and the results (Fig. 3) indicate that the homogeneous regions identified for each extreme daily precipitation index are corroborated statistically to be homogeneous with $H_1 < 2$. And for the sites discordant with the other sites in a homogenous region are also illustrated in Fig. 3.

4.2 Selection of probability functions

Within each homogeneous region for R10, R95P, RX1DAY, and RX5DAY, the regional distribution of each extreme daily precipitation index was fitted by the general normal (GNO, three parameters); Pearson type III (PE3, three parameters), general extreme-value (GEV, three parameters), generalized logistic (GLO, three parameters), and generalized Pareto (GPA, three parameters) distributions and these distributions
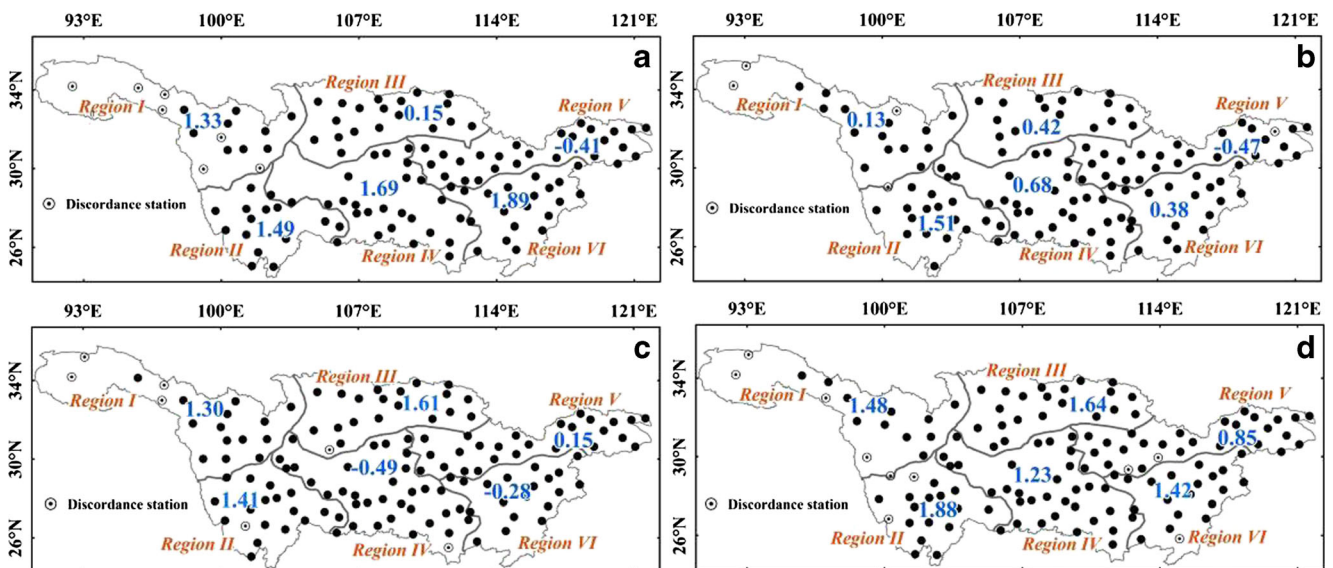

Fig. 3 Homogeneous regions identified for each extreme daily precipitation index and the sites discordant with the other sites in a homogenous region are also illustrated by *circles with dots*, where **a** for R10, **b** for R95P, **c** for RX1DAY, and **d** for RX5DAY

Table 1 Goodness-of-fit test of regional distribution for each homogeneous region

Regions	GLO R10	GEV	GNO	PE3	GPA	GLO R95P	GEV	GNO	PE3	GPA
Region I	3.88	0.12	0.82	0.75	−7.02	6.02	1.39	2.10	1.92	−7.57
Region II	5.32	0.23	1.46	1.46	−9.05	3.60	−0.36	−0.22	−0.78	−8.53
Region III	6.12	0.58	1.51	1.35	−10.06	5.88	2.13	1.92	1.02	−5.95
Region IV	4.53	−2.02	−0.57	−0.60	−14.15	6.44	0.92	1.09	0.30	−10.51
Region V	6.06	−0.40	0.58	0.32	−12.93	6.65	2.08	1.80	0.65	−7.79
Region VI	5.33	−0.46	0.60	0.48	−11.47	4.99	1.41	0.99	−0.17	−6.52
	RX1DAY					RX5DAY				
Region I	3.07	−0.19	−0.27	−0.95	−7.12	3.93	−0.42	0.55	0.53	−8.49
Region II	1.67	−0.59	−1.31	−2.69	−6.00	3.94	−1.29	−0.11	−0.14	−10.96
Region III	3.20	0.39	−0.37	−1.89	−6.20	6.88	0.81	1.78	1.58	−10.91
Region IV	3.63	0.47	−0.66	−2.78	−7.18	5.47	−1.58	−0.16	−0.24	−14.82
Region V	2.02	−0.32	−1.48	−3.58	−6.24	5.66	−0.39	0.61	0.42	−12.03
Region VI	2.60	0.29	−0.67	−2.43	−5.42	4.87	−0.79	0.27	0.16	−11.53

are introduced in the Appendix. The goodness-of-fit test was carried out to generate the results as shown in Table 1, where the criterion value was 1.64 for the absolute value of goodness-of-fit measure statistic Z (Hosking and Wallis 1997), corresponding to acceptance of the hypothesized distribution at a confidence level of 90 %. As the criterion that the absolute value of Z less than 1.64 is used only to decide whether or not a distribution should be accepted but does not allow to decide the probability distributions with the highest goodness-of-fit value. In this case, the L-moment ratio diagram is also used to scrutinize the distribution by comparing its closeness to the L-skewness and L-kurtosis

combination in the L-moment ratio diagram. The L-moment ratio plot for R10, R95P, RX1DAY, and RX5DAY at six homogeneous regions for extreme daily precipitation index R10, R95P, RX1DAY, and RX5DAY are illustrated in Figs. 4, 5, 6, and 7, respectively.

Table 1 and Figs. 4, 5, 6, and 7 indicate that the regional distributions with the highest goodness-of-fit performance for R10, R95P, RX1DAY, and RX5DAY for each homogeneous region are displayed in Table 2. It can be observed that PE3, GEV, and GNO perform well in fitting regional precipitation extremes. And then the regional growth curves of R10, R95P, RX1DAY, and RX5DAY for each homogeneous

Fig. 4 L-moment ratio plot for R10 at six homogeneous regions and **a**, **b**, **c**, **d**, **e**, and **f** are for regions I to VI, respectively
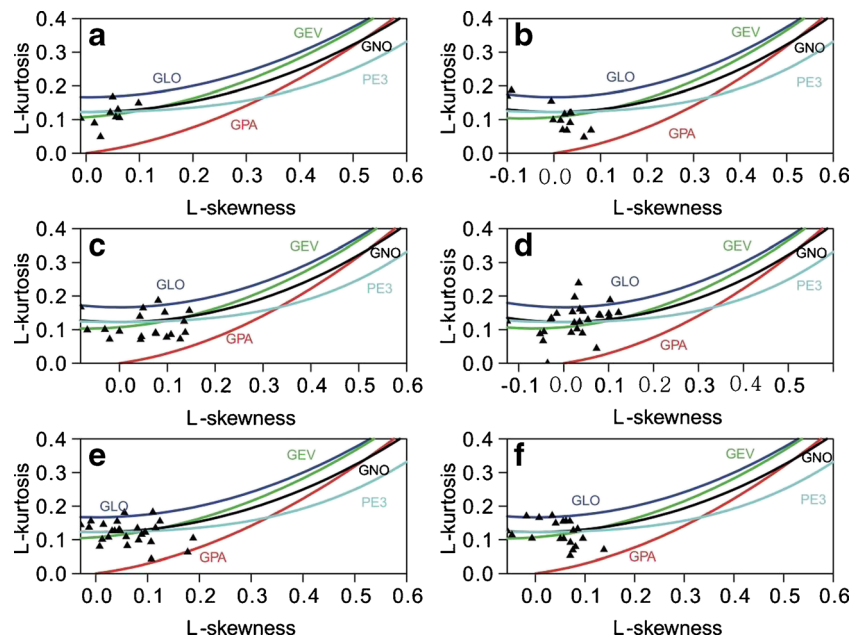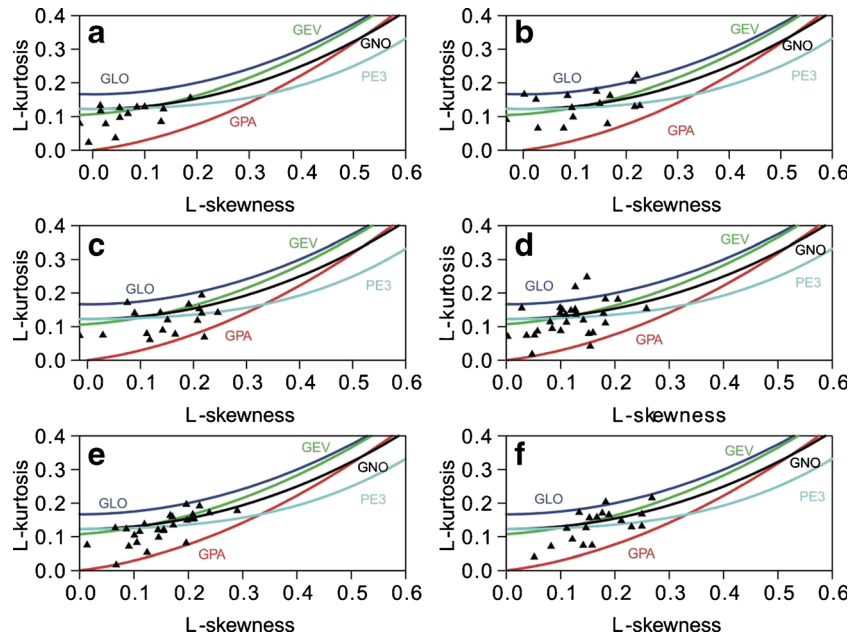
Fig. 5 L-moment ratio plot for
R95P at six homogeneous
regions and **a**, **b**, **c**, **d**, **e**, and **f** are
for regions I to VI, respectively

region are demonstrated in Figs. 8, 9, 10, and 11, respectively, and the 99 % error bounds are also shown in these figures. It can be identified from these figures that the quantile estimates are valid and justified to be used when return periods are less than 100 years. Meanwhile, Fig. 8 also shows that the regional growth curve of R10 for each homogeneous region is convex, i.e., the precipitation increments in all of the homogeneous region are generally decreasing when occurrence frequencies increase for precipitation extremes with return periods of 5, 10, 50, and 200 years. This means that the extreme values of R10 deviate little from the mean or the median in the Yangtze River basin. As the skewness is a

measure of the asymmetry of the probability distribution of a real world random variable; and kurtosis is a measure of the "peakedness" of the probability distribution of a real world random variable, it can be seen from the L-moment ratio diagram of R10 (Fig. 4) that the regional average L-skewness and L-kurtosis are also small, which is in good agreement with the results of regional growth curves.

It can be seen from Fig. 9 that the regional growth curve of RX95P for homogeneous regions I to IV are evidently convex; but for regions V and VI, the regional growth curves are nearly the straight lines, especially for region VI. Similarly, this can also be discerned from the L-moment ratio diagram

Fig. 6 L-moment ratio plot for
RX1DAY at six homogeneous
regions and **a**, **b**, **c**, **d**, **e**, and **f** are
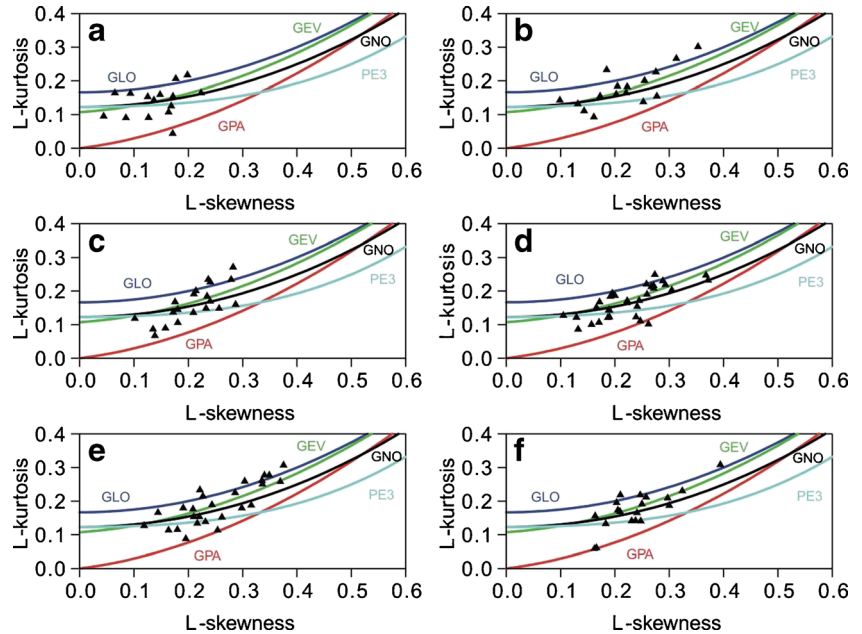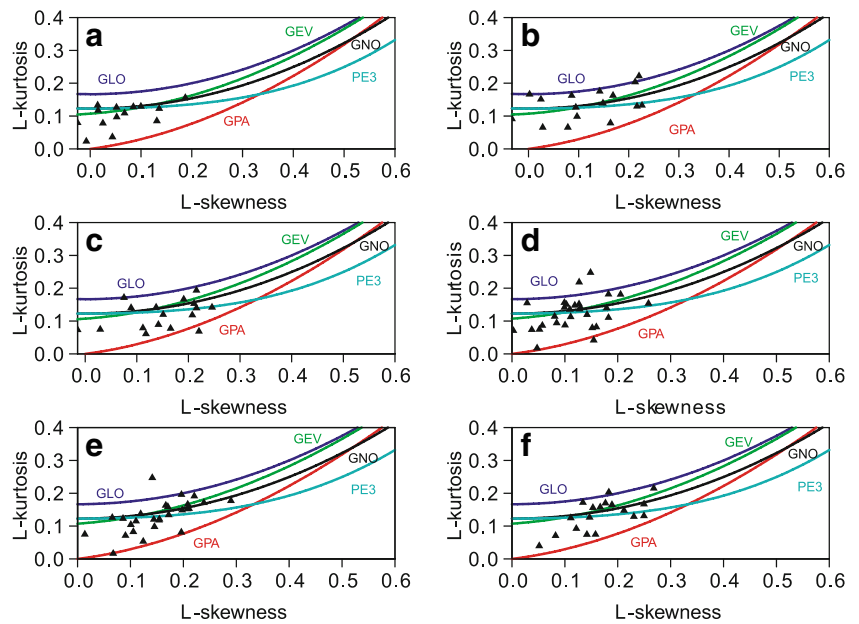for regions I to VI, respectively

Fig. 7 L-moment ratio plot for RX5DAY at six homogeneous regions and **a**, **b**, **c**, **d**, **e**, and **f** are for regions I to VI, respectively

for R95P (Fig. 5), which shows that the regional average L-skewness and L-kurtosis are small in regions I to IV, while larger in region V and region VI. This means that the values of R95P in region V are in more asymmetrical distribution and are prone to extreme value events, especially for region VI. It can be observed from Fig. 10 that the regional growth curves of RX1DAY for each homogeneous region are concave, opposite of convex, except for region I which is nearly a straight line, and this can also be expected from the L-moment ratio diagram of RX1DAY (Fig. 6). It can be seen from Fig. 6 that the regional average L-skewness and L-kurtosis are large in each homogeneous region and this means that it is prone to extreme value events in the Yangtze River basin for 1-day maximum precipitation. Figure 11 indicates that the regional growth curves of RX5DAY for homogeneous regions V and VI are evidently concave which can be identified from the L-moment ratio diagram of RX5DAY (Fig. 7) that the regional average L-skewness and L-kurtosis in regions V and VI are larger than other regions. And this means that, taking account of consecutive 5 days precipitation amount, it is prone to extreme value events only in regions V and VI, implying that regions V and

VI are more likely to suffer from extreme precipitation events.

4.3 Spatial patterns of precipitation extremes of different return periods

Spatial patterns of the precipitation extremes can serve as one of the most important environmental indicator for integrated basin-scale water resources management, thus it is instrumental to quantify the spatial associations of precipitation extremes between sites and to map precipitation extremes with different return periods across the Yangtze River basin using the Kriging interpolation technique. To improve the accuracy of spatial interpolation of precipitation extremes, the stations which could not be grouped into any differentiated homogeneous region are not excluded in the dataset. It can be seen from the above that the performance of PE3, GEV, and GNO are good in fitting the regional distribution for R10, RX95P, RX1DAY, and RX5DAY. The discordant stations of R10, RX95P, RX1DAY, and RX5DAY are fitted by the three distributions, the goodness-of-fit tests by Kolmogorov–Smirnov method show that for all of the

Table 2 The best-fitted regional distributions of R10, R95P, RX1DAY, and RX5DAY for each homogeneous region

	Region I	Region II	Region III	Region IV	Region V	Region VI
R10	GEV	GNO	GEV	GNO	PE3	PE3
R95P	GEV	GEV	PE3	PE3	PE3	GNO
RX1DAY	GNO	GEV	GEV	GNO	GEV	GEV
RX5DAY	GEV	GEV	PE3	PE3	GNO	GNO

Fig. 8 Estimated regional growth curves of R10, with 99 % error bands for six homogeneous regions and A, B, C, D, E and F are for regions I to VI, respectively
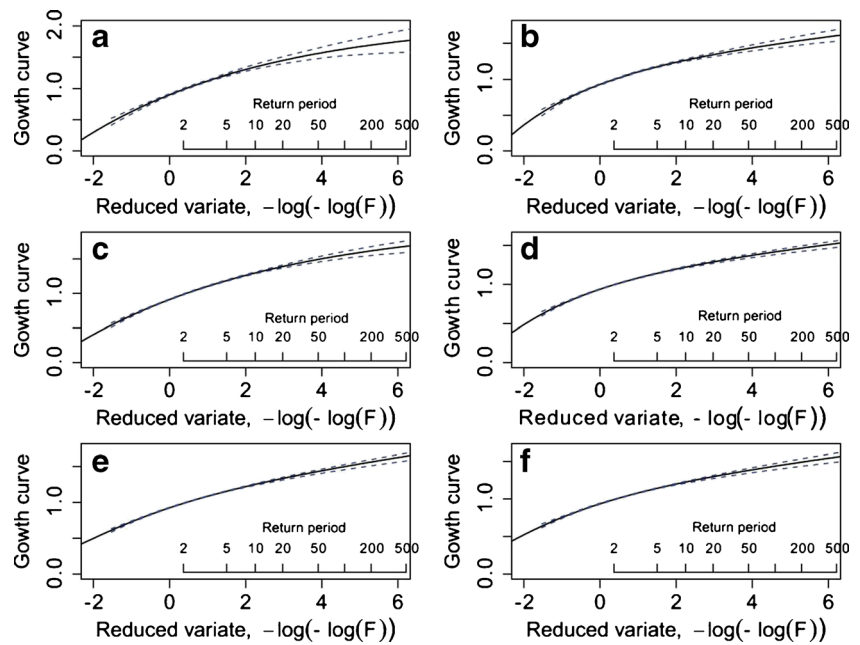

stations these three distributions are all significant at 95 % level (the results not shown here). Then the best-fitted distributions are selected by plotting position figure, here taking the station 114 which is discordant for the regional frequency analysis of RX1DAY as an example, the plotting position figure of station 114 is illustrated in Fig. 12. It can be seen from Fig. 12 that the GEV fits well the RX1DAY. And then the best-fitted distributions for all of the discordant stations are selected and the results are shown in Table 3.

Then the spatial patterns of annual extreme daily precipitation index of R10, RX95P, RX1DAY, and RX5DAY with

return period of 20 years are illustrated in Fig. 13. It can be seen from Fig. 13 that precipitation amount increases gradually from the upper to the lower Yangtze River basin. Meanwhile it can also be seen that there are two regions with the highest precipitation extremes in the eastern part of the Yangtze River basin and the regions covered by the stations 21, 24, and 25 (Fig. 1), specifically the regions east of the 107° E. However, there is also another region with higher precipitation extremes in the central parts of the basin, specifically the region between 100° E–107° E, and it is particularly the case for R95P, RX1DAY, and RX5DAY (Fig. 13b–

Fig. 9 Estimated regional growth curves of R95P, with 99 % error bands for six homogeneous regions and **a**, **b**, **c**, **d**, **e**, and **f** are for regions I to VI, respectively
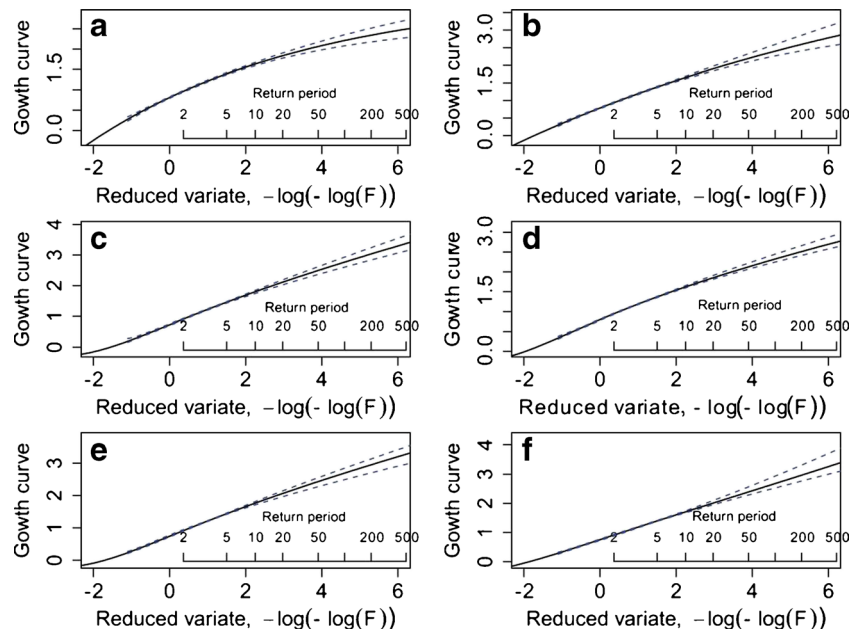
Fig. 10 Estimated regional growth curves of RX1DAY, with 99 % error bounds for six homogeneous regions and **a**, **b**, **c**, **d**, **e**, and **f** are for regions I to VI, respectively
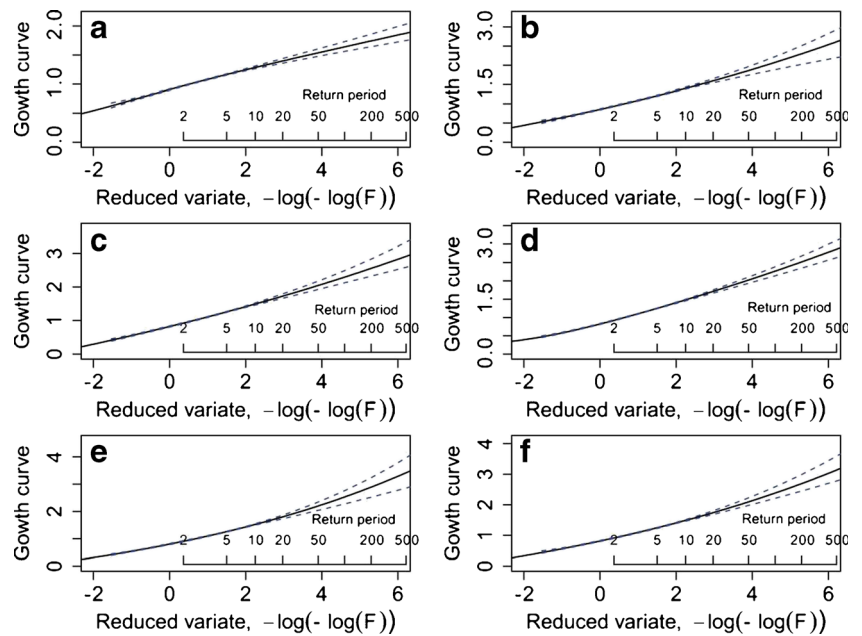
d). Figure 1 shows that the locations of stations 21, 24, and 25 are in the boundary of a basin, which is also called as the Szechwan Basin. The elevation at the southeastern part of the Szechwan Basin is relatively lower, and the water vapor of southeast monsoon is conducive to input to the Szechwan Basin while the water vapor is blocked in the other boundary of the basin, causing the convection of water vapor. Then the precipitation in the boundary of the Szechwan Basin will be more than other regions in the vicinity of the precipitation center. It can be seen from Figs. 13 and 1 that the precipitation in the boundary of the Szechwan Basin, except the

southeastern part, is more than the ambient regions. And there is an obvious precipitation center at the west part of the Szechwan Basin. Meanwhile it can be seen from Fig. 13a that precipitation in the center part of the Yangtze River basin is large in amount and this is due to the fact that stations 21, 24, and 25 with significant trends of R10 are not included in the spatial interpolation analysis.

The result that higher precipitation amount of precipitation events with return periods of 20 years is identified in the middle and southeastern parts of the Yangtze River basin, or mainly in the lower Yangtze River basin is consistent with our

Fig. 11 Estimated regional growth curves of RX5DAY, with 99 % error bands for six homogeneous regions and **a**, **b**, **c**, **d**, **e**, and **f** are for regions I to VI, respectively
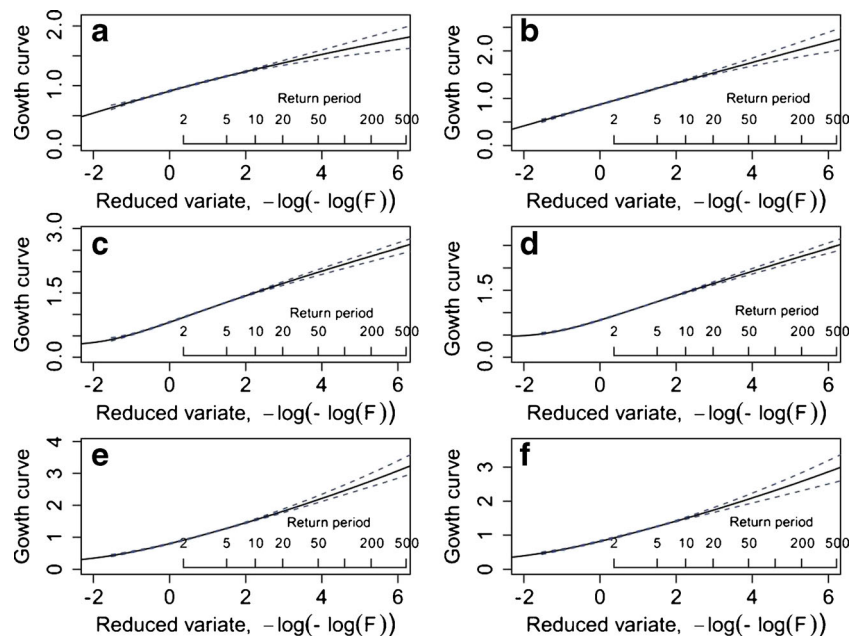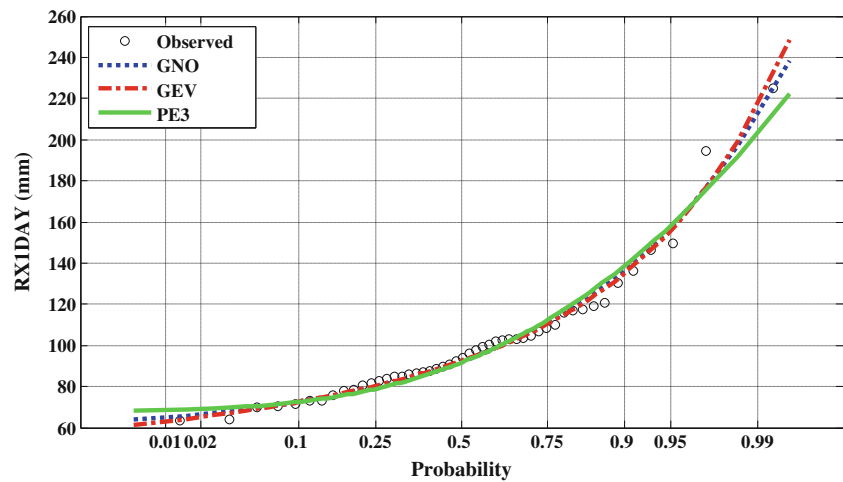
Fig. 12 Fitted probability
functions for RX1DAY at station
114

previous studies (e.g., Zhang et al. 2008). Large-scale atmospheric circulation analysis indicates decreasing strength of East Asian summer monsoon during 1975–2005 as compared to that during 1961–1974 and increasing geopotential height in the north China, South China Sea, and west Pacific regions, all of which combine to negatively impact the northward propagation of the vapor flux. This circulation pattern will be beneficial for the longer stay of the Meiyu front in the Yangtze River basin, leading to more precipitation in the middle and lower basin in summer months (Zhang et al. 2008). Larger precipitation amount of precipitation events with return periods of 20- years indicate higher risk of floods and droughts. Moreover, Zhang et al. (2008) also demonstrated that the significant increasing summer precipitation intensity and changing frequency in the rain/no rain days in the middle and lower Yangtze River basin have potential to result in higher occurrence probability of flood and drought hazards in the region. Therefore, under the influences of changing climate, the risk of droughts and floods will be significantly amplified.

5 Conclusions

Regional frequency analysis based on extreme daily precipitation index has scientific and practical value in the context of basin-scale water resource and flood risk management. This paper presents a regional frequency analysis of precipitation extremes and characterization of the spatial pattern of rainfall extremes variations in the Yangtze River basin using the well-known index-flood L-moments approach together with some advanced statistical tests and spatial analysis methods.

The Yangtze River basin can be categorized into six homogeneous regions in terms of extreme daily precipitation index. For each homogeneous region of R10, R95P, RX1DAY, and RX5DAY, the regional distribution of each extreme daily precipitation index was fitted by the general normal, Pearson type III, general extreme-value, generalized logistic, and generalized Pareto distribution and the goodness-of-fit test results indicated that PE3, GEV, and GNO performed well in fitting regional distributions.

Table 3 The best fitted distributions for all of the discordant stations of each extreme daily precipitation index

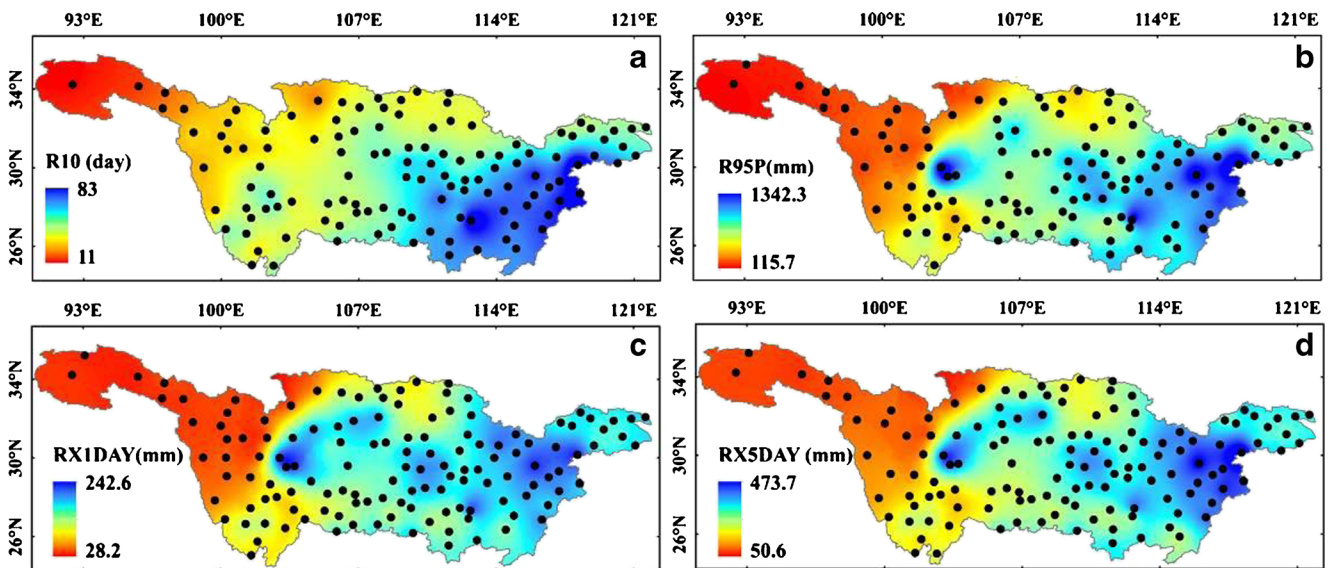| R10		R95P		RX1DAY		RX5DAY	
Station	Distribution	Station	Distribution	Station	Distribution	Station	Distribution
2	PE3	1	PE3	1	PE3	1	PE3
3	GEV	2	PE3	2	GEV	2	GEV
4	GEV	10	GEV	4	PE3	4	PE3
5	GEV	22	PE3	5	GEV	18	PE3
9	PE3	123	GEV	37	GNO	22	PE3
18	PE3			65	GNO	27	GNO
23	GEV			114	GEV	36	GEV
						79	GEV
						80	GEV
						116	PE3

Fig. 13 Spatial patterns of annual extreme daily precipitation index with return period of 20 years: **a** for R10; **b** for R95P; **c** for RX1DAY, and **d** for RX5DAY

Additionally, this study also presents the regional growth curves of R10, R95P, RX1DAY, and RX5DAY for each homogeneous region with the 99 % error bands. The results indicate that the quantile estimates are reliable enough and can be used when return periods are less than 100 years. Furthermore, analysis results of the shape of the regional growth curves and the regional average L-skewness and L-kurtosis of R10, R95P, RX1DAY, and RX5DAY for each homogeneous region indicate that, for R95P, RX1DAY, and RX5DAY, the regional average L-skewness and L-kurtosis in regions V and VI are larger than other regions, implying that extreme precipitation events are highly probable to occur in regions V and VI, and hence higher risk of floods and droughts in regions V and VI.

Spatial patterns of annual extreme daily precipitation index of R10, RX95P, RX1DAY, and RX5DAY with return period of 20 years indicate that precipitation amount increases gradually from the upper to the lower Yangtze River basin. Moreover, there are two precipitation centers with larger precipitation amount in the eastern part of the Yangtze River basin and the western part of the Szechwan Basin boundary. As the Yangtze River basin is dominated by the southeast monsoon, it is obvious that the precipitation gradually decreases from southeast to northwest, forming an obvious precipitation center with higher precipitation amount. The precipitation center with higher precipitation amount in the western part of the Szechwan Basin boundary should be attributed to the effects of topography because the southeast monsoon is blocked in the boundary of the basin, causing the convection of water vapor. Furthermore, decreasing strength of East Asian summer monsoon during 1975–2005 as compared to that during 1961–1974 and increasing geopotential height in the north China, South China

Sea, and west Pacific regions combine to negatively impact the northward propagation of the vapor flux, and which should contribute much to the higher precipitation amount in the middle and lower Yangtze River basin (Zhang et al. 2008).

Appendix

The univariate distributions used in this study are introduced as follow, referring to Hosking (2012b):

1. GEV

    The generalized extreme-value distribution with location parameter ξ, scale parameter α, and shape parameter κ has distribution function

    $$F(x) = \exp(-\exp(-y)) \qquad (1)$$

    where $y=-\kappa^{-1}\log(1-\kappa(x-\xi)/\alpha)$, with x bounded by $\xi+\alpha/k$ from below if $\kappa<0$ and from above if $\kappa>0$. Extreme-value distribution types I, II, and III (Gumbel, Frechet, and Weibull) correspond to shape parameter values $\kappa=0$, $\kappa<0$, and $\kappa>0$, respectively.

2. GNO

    The generalized normal distribution with l location parameter ξ, scale parameter α, and shape parameter κ has distribution function

    $$F(x) = \Phi(y) \qquad (2)$$

where $y=-\kappa^{-1}\log(1-\kappa(x-\xi)/\alpha)$, and $\Phi(y)$ is the distribution function of the standard normal distribution, with x bounded by $\xi+\alpha/k$ from below if $\kappa<0$ and from above if $\kappa>0$.

3. GLO

The generalized logistic distribution with location parameter ξ, scale parameter α, and shape parameter κ has distribution function

$$F(x) = 1/(1 + \exp(-y)) \tag{3}$$

where $y=-\kappa^{-1}\log(1-\kappa(x-\xi)/\alpha)$, with x bounded by $\xi+\alpha/k$ from below if $\kappa<0$ and from above if $\kappa>0$.

4. GPA

The generalized Pareto distribution with location parameter ξ, scale parameter α, and shape parameter κ has distribution function

$$F(x) = 1-\exp(-y) \tag{4}$$

where $y=-\kappa^{-1}\log(1-\kappa(x-\xi)/\alpha)$, with x bounded by $\xi+\alpha/k$ from below if $\kappa<0$ and from above if $\kappa>0$. And the exponential distribution is the special case $\kappa=0$. The uniform distribution is the special case $\kappa=1$.

5. PE3

The Pearson type III distribution contains as special cases the usual three-parameter gamma distribution (a shifted version of the gamma distribution) with a finite lower bound and positive skewness; the normal distribution, and the reverse three-parameter gamma distribution, with a finite upper bound and negative skewness. The distribution's parameters are the first three (ordinary) moment ratios: μ (the mean, a location parameter), σ (the standard deviation, a scale parameter), and γ (the skewness, a shape parameter).

If $\gamma\neq0$, let $\alpha=4/\gamma^2$, $\beta=\frac{1}{2}\sigma|\gamma|$, $\xi=\mu-2\sigma/\gamma$. The probability density function is

$$f(x) = \frac{|x-\xi|^{\alpha-1}\exp\left(-|x-\xi|\big/\beta\right)}{\beta^{\alpha}\Gamma(\alpha)} \tag{5}$$

with x bounded by ξ from below if $\gamma>0$ and from above if $\gamma<0$. If $\gamma=0$, the distribution is a normal distribution with mean μ and standard deviation σ.

References

Alan DZ, Justin S, Edwin PM, Bart N, Eric FW, Dennis PL (2003) Detection of intensification in global- and continental-scale hydrological cycles: temporal scale of evaluation. J Clim 16:535–547

Alexander L, Zhang X, Peterson T, Caesar J, Gleason B, Klein Tank A, Haylock M, Collins D, Trewin B, Rahimzadeh F (2006) Global observed changes in daily climate extremes of temperature and precipitation. J Geophys Res 111, D05109. doi:10.1029/2005JD006290

Allen M, Ingram WJ (2002) Constraints on future changes in climate and the hydrologic cycle. Nature 419:224–232

Dore HIM (2005) Climate change and changes in global precipitation patterns: what do we know? Environ Int 31:1167–1181

Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, Mearns LO (2000) Climate extremes: observations, modeling, and impacts. Science 289:2068–2074

Emori S, Brown SJ (2005) Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. Geophys Res Lett 32, L17706. doi:10.1029/2005GL023272

Fatichi S, Caporali E (2009) A comprehensive analysis of changes in precipitation regime in Tuscany. Int J Climatol. doi:10.1002/joc.1921

Goovaerts, P., 1999. Performance comparison of geostatistical algorithms for incorporating elevation into the mapping of precipitation. In: The IV International Conference on GeoComputation was hosted by Mary Washington College in Fredericksburg, VA, USA, 25–28 July 1999.

Guo J, Chen H, Xu CY, Guo S (2012) Prediction of variability of precipitation in the Yangtze River Basin under the climate change conditions based on automated statistical downscaling. Stoch Env Res Risk A 26(2):157–176

Hosking J, Wallis J (1993) Some statistics useful in regional frequency analysis. Water Resour Res 29(2):271–281

Hosking J, Wallis J (1997) Regional frequency analysis: an approach based on L-moments. Cambridge University Press, Cambridge

Hosking, J., 2012a. Regional frequency analysis using L-moments. R package, version 2.4. URL:http://CRAN.R-project.org/package=lmomRFA.

Hosking, J., 2012b. L-moments. R package, version 1.6. URL:http://CRAN.R-project.org/package=lmom.

Huntington GT (2006) Evidence for intensification of the global water cycle: review and synthesis. J Hydrol 319:83–95

Kendall, M., 1975. Rank Correlation Methods: Charles Griffin, edited, London

MacDonald MG (2010) Water, climate change, and sustainability in the southwest. Proc Nat Acad Sci U S A 107(50):21256–21262

Mann HB (1945) Nonparametric tests against trend. Econometrica J Econ Soc 13(3):245–259

Mitchell, J.M., Dzerdzeevskii, B., Flohn, H., Hofmeyr, W.L., Lamb, H.H., Rao, K.N., Wallen, C.C., 1966. Climate Change, WMO Technical Note No. 79, World Meteorological Organization.

Piao SL, Ciais P, Huang Y, Shen Z, Peng S, Li J, Zhou L, Liu H, Ma Y, Ding Y, Friedlingstein P, Liu C, Tan K, Yu Y, Zhang T, Fang J (2010) The impacts of climate change on water resources and agriculture in China. Nature 467:43–51

Schmidhuber J, Tubiello NF (2007) Global food security under climate change. Proc Nat Acad Sci U S A 104(50):19703–19708

Su BD, Xiao B, Zhu DM, Jiang T (2005) Trends in frequency of precipitation extremes in the Yangtze River basin, China: 1960–2003. Hydrol Sci J 50(3):479–492

Wu KL, Yang MS (2005) A cluster validity index for fuzzy clustering. Pattern Recognit Lett 26(9):1275–1291

Yang T, Shao Q, Hao ZC, Chen X, Zhang Z, Xu CY, Sun L (2010) Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River Basin, China. J Hydrol 380(4):386–405

Yin HF, Li CA (2001) Human impact on floods and flood disasters on the Yangtze River. Geomorphology 41:105–109

Yue S, Pilon P, Phinney B, Cavadias G (2002) The influence of autocorrelation on the ability to detect trend in hydrological series. Hydrol Process 16(9):1807–1829

Zhang Q, Jiang T, Gemmer M, Becker S (2005) Precipitation, temperature and discharge analysis from 1951–2002 in the Yangtze Catchment, China. Hydrol Sci J 50(1):65–80

Zhang Q, Liu C-L, Xu C-Y, Xu Y-P, Jiang T (2006) Observed trends of annual maximum water level and streamflow during past 130 years in the Yangtze River basin, China. J Hydrol 324:255–265

Zhang Q, Xu C-Y, Jiang T, Wu YJ (2007) Possible influence of ENSO on annual maximum streamflow of the Yangtze River, China. J Hydrol 333:265–274

Zhang Q, Xu C-Y, Zhang Z, Chen YD, Liu C-L (2008) Spatial and temporal variability of precipitation maxima during 1960–2005 in the Yangtze River basin and possible association with large-scale circulation. J Hydrol 353:215–227

Zhang Q, Singh VP, Li JF, Chen XH (2011) Analysis of the periods of maximum consecutive wet days in China. J Geophys Res 116, D23106. doi:10.1029/2011JD016088

Zhang Q, Zhou Y, Singh VP, Li JF (2012) Scaling and clustering effects of extreme precipitation distributions. Journal of Hydrology. doi:10.1016/j.jhydrol.2012.06.015