https://doi.org/10.1016/j.atmosres.2019.104629

A modified regional L-moment method for regional extreme precipitation frequency analysis in the Songliao River Basin of China

Chen Hu[1,2], Jun Xia[1,2,3*], Dunxian She[1,2*], Chongyu Xu[1,4], Liping Zhang[1,2], Zhihong, Song[1,2], Lin Zhao[5]

1.State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan 430072, P.R. China

2. Hubei Provincial Key Laboratory of Water System Science for Sponge City Construction, Wuhan University, Wuhan, 430072, P.R. China

3.Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, P.R. China

4.Department of Geosciences, University of Oslo, N-0316 Oslo, Norway

5.School of Resource and Environmental Science, Wuhan University, Wuhan 430072, P.R. China

Corresponding authors:

Prof. Jun Xia; email: xiajun666@whu.edu.cn

Dr. Dunxian She; Email: shedunxian@whu.edu.cn

Abstract

The regional frequency analysis (RFA) is a widely used method in analyzing the changes of extreme precipitation (EP). The uncertainties in the identification of homogeneous subregions and the selection of optimal regional frequency distributions can largely influence the results of the RFA. In this study, the fuzzy c-means method combined with the extended Xie-Benn index (FCXB) is used to help determine the optimal division of subregions. In addition, we introduce a new comprehensive index (CI) to overcome the shortcomings of present measures and reduce the uncertainty in regional frequency distribution selection. The changes of EP at 93 meteorological stations in the Songliao River Basin (SRB) during 1960 to 2016 is analyzed. The main results show that: (1) FCXB can effectively identify the optimum number of homogeneous subregions automatically, and the corresponding subregion division is proven to be reasonable and reliable; (2) compared with the single goodness-of-fit measure, the developed CI can reduce the uncertainties in distribution selection and determine the optimal regional distribution in a reliable way; (3) the estimated EP under different return periods both decrease from the south to the north of the SRB, which indicates the risk of high-intensity EP events in the southern SRB is relatively higher. These findings can provide technical support for local policymakers to formulate effective measures to lessen the damages of the EP on ecosystem and society.

Keywords: regional frequency analysis; L-moments method; a comprehensive index; Xie-Benn index; the Songliao River Basin

1 Introduction

According to the fifth IPCC report (2013), the frequency of extreme precipitation (EP) events has significantly increased in many land regions since the 1950s, causing huge economic losses (Jun et al., 2017; Su et al., 2008). Moreover, the increasing trend of EP is projected to continue in the 21st century in many regions across the world, especially in the high latitudes and tropical regions (Alexander and Arblaster, 2009; Hao et al., 2013). The frequent occurrence of EP, which is mainly attributed to climate change and human activities (Konisky et al., 2016; Tao et al., 2011), can probably alter the balance of local ecosystem, and will likely trigger other natural hazards such as floods (Ashfaq et al., 2010; Mishra et al., 2012), drought (Mukherjee et al., 2018), and landslides (Wu et al., 2017). Therefore, understanding the changes in EP (frequency, trends, etc.) is of great significance and can help water resources managers formulate effective adaptation strategies to minimize catastrophic losses.

Basically, there are two ways to carry out an EP frequency analysis: the at site frequency analysis (ASF) method and the regional frequency analysis (RFA) method (Hosking and Wallis, 1997). Compared to the ASF method, which cannot be applied at ungauged stations, the main advantage of the RFA method is that it can use data from multiple gauged stations to predict the meteorological characteristics in an ungauged region (Cunnane, 1988; She et al., 2014). The most well-known RFA method is the regional L-moment method (denoted as LM), which has been widely used in the RFA of EP in many regions around the world (Chen et al., 2014; Fowler and Kilsby, 2003;

Ngongondo et al., 2011; Yang et al., 2010a). Generally, the procedure of LM consists of five steps: (a) identification of homogeneous subregions; (b) screening of the data using discordancy measures; (c) the homogeneity test using heterogeneous measures; (d) selection of the optimal regional distributions; and (e) quantiles estimation and accuracy assessment.

Determining the number of subregions plays an important role in the work of LM since an inappropriate number of subregions will lead to an unreasonable division pattern that may not satisfy the homogeneity assumption of the RFA (Forestieri et al., 2018; Rao and Srinivas, 2006). Many previous studies determined the number of subregions by setting an initial number based on the regional geographical characteristics and then adjusting the number back and forth until all divided subregions could pass the homogeneity test (Du et al., 2014; Yang et al., 2010b). However, such an approach is subjective, and the determined number of subregions may not be the most appropriate. Some objective cluster methods such as the K-means method and the fuzzy c-means method (denoted as FC) with some cluster validity indices have also been employed for determining the optimum number of subregions in some studies (Forestieri et al., 2018; Halkidi et al., 2001). However, selecting a reliable cluster validity index is important for the determination of the optimum subregion number as different indices usually lead to different results. In this study, the FC with a reliable index, the extended Xie-Benn index (denoted as FCXB), is employed to identify the homogeneous subregions (Basu and Srinivas, 2015; Rao and Srinivas, 2006; Xie and

Beni, 1991). As confirmed by Xie and Beni (1991) and Hamerly and Elkan (2002), FCXB performs better than the K-means method in producing hard clustering solutions. In addition, the FCXB presents a direct connection to the properties of the input data so that the results of the FCXB do not exhibit monotonic increasing or decreasing trends with the increase in the subregion number. However, the connection to the input data is lacking in some other validity indices (such as the fuzzy partition coefficient or the fuzzy partition entropy), which will lead to the monotonic trends in their results and unreasonable decisions for the subregion number (Halkidi et al., 2001; Hall and Minns, 1999).

The determination of an appropriate regional frequency distributions is also important in the RFA. An inappropriate selection of distributions may result in a large overestimation or underestimation of the risk of EP. Several goodness-of-fit measures have been proposed to select the optimal regional distribution, and the goodness-of-fit measure proposed by Hosking and Wallis (1997) (denoted as HWGOF) and the graphical measure using the L-diagram (denoted as GMLM) (Peel et al., 2001; Vogel et al., 1993) are the two most widely used. However, the results of GMLM may not be reliable because the selection of the optimal regional distributions mainly depends on subjective judgments. In addition, Kjeldsen and Prosdocimi (2015) claimed that the distributions suggested by the HWGOF are not robust enough due to the inadequate consideration of the variability of the L-skewness, and they proposed a new bivariate extension goodness-of-fit measure (denoted as KPGOF), which considers the variability of both the L-skewness and L-kurtosis as well as their correlation. The KPGOF can improve the ability and stability in selecting the true population distribution in most situations except when the true distribution is a generalized logistic distribution (Kjeldsen et al., 2017; Kjeldsen and Prosdocimi, 2015). In this study, we will first compare the efficiency of the HWGOF, GMLM, and KPGOF in determining the optimal regional distribution and then develop a comprehensive index (denoted as CI) based on the joint consideration of the three previously described measures. We aim to provide a more reliable way based on CI to help choose the most appropriate regional distributions, especially when the results of HWGOF, GMLM, and KPGOF are inconsistent.

In this study, the Songliao River Basin (SRB) is selected as the study area, and the RFA of the EP in this region is investigated through the LM with FCXB and CI. The objectives of this study are to (1) reduce the uncertainty in the identification of the homogeneous subregions through applying the FCXB to determine the optimum number of subregions and the corresponding division of homogeneous subregions and (2) construct a comprehensive index with the joint considerations of three different goodness-of-fit measures to determine the optimal regional distribution in a more reliable way.

2 Study area and methodology

2.1 Study area and data

The Songliao River Basin (115°31'E–135°9'E, 38°35'N–53°35'N) is one of the most important agricultural and industrial regions in northeastern China (Liang et al., 2011; Song et al., 2014). However, the frequency of EP in the SRB has been largely increased in recent years, causing significant economic losses and huge damage to local infrastructures (Ma et al., 2004; Wang et al., 2013). Thus, it is of great necessity to assess the EP over the whole SRB. The SRB covers a drainage area of approximately 1.24×10^6 km^2, including the northeast of the Inner Mongolia province, Liaoning province, Jilin province, and Heilongjiang province. The Amur River, Liao River, Yalu River, Tumen River, Suifen River, Daling River, and Ergun River are the seven major rivers in the SRB. Of these seven rivers, the largest is the Amur River, with a length of approximately 4440 km and a drainage area of approximately 8.88×10^5 km^2 in the territory of China. The largest tributary of the Amur River in China is the Songhua River, with a length of 1927 km and a drainage area of 5.568×10^5 km^2, which is formed by the confluence of the southern Second Songhua River tributary and the Northern Nenjiang River tributary (Song et al., 2015). The SRB belongs to temperate and cold temperate zones and has a continental monsoon climate. The long-term annual mean precipitation (denoted as LAMP) of the SRB shows a south-north gradient, varying from more than 1000 mm in the southern SRB to less than 350 mm on the northern edge of the SRB (Qi, 2006). The east, west, and north of the SRB are surrounded by mountains, and the highest mountain with an elevation of 2439 m is located in the west of the SRB. The long-term annual mean air temperature varies from 1°C to 5°C, and the annual range of air temperature can reach up to 40°C. Fig. 1 shows the location and the geographic information of the study area.

In this study, the daily precipitation observations (from 1960 to 2016 without gaps) of 93 national meteorological stations in the SRB are used. These data were obtained from the National Meteorological Administration of China (http://data.cma.cn/), which is the official institute of China providing high-quality meteorological data. The specific locations of the 93 stations in SRB are given in Fig. 1. In this study, the EP time series at each station is obtained as the annual maximum daily precipitation (denoted as AMP, mm) derived from the daily rainfall observations.

2.2 Methodology

In this study, the LM with FCXB and CI is applied for the RFA of the EP in the SRB. More details about the LM can be found in the study by Hosking and Wallis (1997). The major steps of the LM are briefly introduced below.

2.2.1 The identification and delineation of homogeneous subregions

FCXB is employed in this study for the identification of the homogeneous subregions, as it can determine the optimum number of subregions automatically in a reliable way. The procedure of FCXB consists of three parts: First, select and use several geographical or meteorological characteristics (such as longitude, LAMP, etc.,) of all stations to form the input data matrix. Second, apply the input data matrix to calculate the partition membership matrixes with different numbers of subregions; more details on the process of obtaining the partition membership matrix are provided in the study by Rao and Srinivas (2006). Finally, based on these partition membership matrixes, the values of the extended Xie-Benn index corresponding to different numbers of subregions can be computed as:

$$V_{XB} = (\sum_{j=1}^{c}\sum_{k=1}^{K}(\mu_{jk})^{r}\left\|V_j - W_k\right\|^2) \Big/ (K \min_{j,j\neq k}\left\|V_j - V_k\right\|^2) \tag{1}$$

where μ_{jk} denotes the member of partition membership matrix U at row j and column k. K and c denote the numbers of column and row of U, respectively. W_k denotes the vector of input data matrix W at row k. V_j denotes a fuzzy centroid vector that can be calculated as $V_j = \sum_{k=1}^{K}\left(\mu_{jk}\right)^{r} W_k \Big/ \sum_{k=1}^{K}\left(\mu_{jk}\right)^{r}$, where r denotes the fuzzifier, which controls the extent of membership shared among fuzzy clusters. For most data sets, $1.25 \le r \le 2.5$ gives good results for FC (Srinivas et al., 2008; Zhang and Hall, 2004). The minimum V_{XB} indicates the optimum number of compact and well-separated subregions. After determining the optimum number of subregions and the corresponding partition membership matrix, each station can be assigned to a specific subregion according to its maximum membership in the partition membership matrix.

2.2.2 Screening of data using the discordancy measure

Assume that there are N stations in the study region, let $t^{(i)}, t_3^{(i)}, t_4^{(i)}$ denote the coefficient of variation (L-CV), L-skewness, and L-kurtosis at station i, respectively, and $u_i = \left[t^{(i)}, t_3^{(i)}, t_4^{(i)}\right]^T$. The discordancy measure for station i, D_i, can be calculated as:

$$D_i = N\left(u_i - \bar{u}\right)^T S^{-1}\left(u_i - \bar{u}\right) \Big/ 3 \tag{2}$$

where $u_i = \left[t^{(i)}, t_3^{(i)}, t_4^{(i)}\right]^T$, $\bar{u} = \sum_{i=1}^{N} u_i \Big/ N$ and $S = \sum_{i=1}^{N}(u_i - \bar{u})(u_i - \bar{u})^T$ (Neykov et al., 2007). The stations with D_i values larger than the critical value that is related to the number of sites in the study region are considered to be discordant with the other stations.

2.2.3 Testing of regional homogeneity using the heterogeneity measure

Three heterogeneity measures (H), namely H_1, H_2, and H_3, are used to examine the assumption of the RFA that each divided subregion is a homogeneous region (She et al., 2014). H_1, H_2, and H_3 can be calculated as:

$$H_1 = (V_1 - \mu_v)/\sigma_v \, , H_2 = (V_2 - \mu_{v2})/\sigma_{v2} \, , H_3 = (V_3 - \mu_{v3})/\sigma_{v3} \tag{3}$$

where V_1, V_2, and V_3 are the weighted standard deviations of the at-site sample L-moment ratios that can be calculated as:

$$\begin{cases} V_1 = \left[\sum_{i=1}^{N} n_i (t^{(i)} - t^R)^2 \Big/ \sum_{i=1}^{N} n_i \right]^{1/2} \\ V_2 = \sum_{i=1}^{N} n_i \left[(t^{(i)} - t^R)^2 + (t_3^{(i)} - t_3^R)^2 \right]^{1/2} \Big/ \sum_{i=1}^{N} n_i \\ V_3 = \sum_{i=1}^{N} n_i \left[(t_3^{(i)} - t_3^R)^2 + (t_4^{(i)} - t_4^R)^2 \right]^{1/2} \Big/ \sum_{i=1}^{N} n_i \end{cases} \tag{4}$$

where n_i denotes the recording length at station i and t^R, t_3^R, and t_4^R denote the regional average L-moments ratios ($t_j{}^R = \sum_{i=1}^{N} n_i t_j{}^{(i)} \Big/ \sum_{i=1}^{N} n_i$, $j = 1,3,4$) (Yang et al., 2010a). μ and σ are the mean and standard deviation of V derived from a large number (N_m), which is set as 500 here, of simulated realizations of the study region by Monte Carlo simulation. Each simulated realization contains N stations with the same record length as their real-world counterpart and has a four parameter Kappa distribution fitted by the regional average L-moments ratios $(1, t^R, t_3^R, t_4^R)$ as the frequency distribution. More details about Kappa distribution can be found in the study by Kjeldsen et al. (2017). The study region can be regarded as being "acceptably homogeneous" if $H_i < 1 (i = 1,2,3)$, "possibly homogeneous" if $1 \le H_i < 2 (i = 1,2,3)$ and "definitely heterogeneous" if $H_i \ge 2 (i = 1,2,3)$.

2.2.4 Choice of optimal regional frequency distribution

In this study, four candidate probability distributions that are frequently used in the LM, i.e., the Pearson-Ⅲ distribution (denoted as PE3), generalized normal distribution (denoted as GNO), generalized logistic distribution (denoted as GLO), and generalized extreme value distribution (denoted as GEV), are considered (Fowler and Kilsby, 2003; Yang et al., 2010a). More detail about these frequency distributions can be seen in the study by Hosking and Wallis (1997). Three measures (HWGOF, GMLM, KPGOF) and a newly constructed comprehensive index CI will be separately introduced below.

2.2.4.1 HWGOF

HWGOF (Z^{DIST}), using the L-kurtosis for the distribution selection (Hosking and Wallis, 1997), can be calculated for each candidate distribution as:

$$Z^{DIST} = (\tau_4^{DIST} - t_4^R + \beta_4) / \sigma_4 \tag{5}$$

where τ_4^{DIST} is the L-kurtosis of the fitted candidate distribution to the data. β_4 and σ_4 denote the bias and standard deviation of t_4^R, respectively, which can be computed as:

$$\beta_4 = \sum_{m=1}^{N_m} (t_4^{(m)} - t_4^R) \Big/ N_m \tag{6}$$

$$\sigma_4 = \left\{ \frac{1}{N_m - 1} \left[\sum_{m=1}^{N_m} (t_4^{(m)} - t_4^R)^2 - N_m \beta_4^2 \right] \right\}^{1/2} \tag{7}$$

where $t_4^{(m)}$ denotes the sample regional L-kurtosis derived from the *m*th simulation. If $|Z^{DIST}| \le 1.64$, the candidate distribution can be considered acceptable. The distribution with the minimum $|Z^{DIST}|$ among all the acceptable distributions is the best regional distribution.

2.2.4.2 GMLM

GMLM is a widely-used graphical measure based on L-diagrams, which is a simple but useful visual comparison method. With the verification of several studies (Kjeldsen and Prosdocimi, 2015; Vogel et al., 1993; Vogel and Wilson, 1996), the distribution whose theoretical curve is the closest to the regional mean L-moment ratios point can be suggested to be the most appropriate regional distribution.

2.2.4.3 KPGOF

The procedure of KPGOF consists of two steps (Kjeldsen and Prosdocimi, 2015): first, using the L-diagrams with the $(1-\alpha)100\%$ confidence ellipse, where α is the given significance level, to determine the acceptable distributions. The distributions whose theoretical curves are located within the area of the ellipse are considered as acceptable distributions. Second, for each acceptable distribution, D^{DIST} can be calculated as:

$$D^{DIST} = \left(\tau^{DIST} - t^{R}\right)^{T} \Omega^{-1} \left(\tau^{DIST} - t^{R}\right) \tag{8}$$

where $t_B^R = (t_3^R - \beta_3, t_4^R - \beta_4)$ denotes the vector of the bias-corrected regional L-skewness and L-kurtosis. Ω is a covariance matrix as $\begin{bmatrix} \sigma_3^2 & \sigma_{34} \\ \sigma_{43} & \sigma_4^2 \end{bmatrix}$, where σ_{34} is the covariance between the L-skewness and L-kurtosis, which can be estimated as:

$$\sigma_{34} = (N_m - 1)^{-1} \left\{ \sum_{m=1}^{N_m} (t_3^{(m)} - t_3^R)(t_4^{(m)} - t_4^R) - N_m \beta_3 \beta_4 \right\} \tag{9}$$

where $t_3^{(m)}$ is the regional average L-skewness derived from the *m*th simulation. β_3 and σ_3 denote the bias and standard deviation of t_3^R, respectively. If $\left|D^{DIST}\right| \le 4.61$, the candidate distribution can be viewed as acceptable. The distribution with the minimum $\left|D^{DIST}\right|$ can be accepted as the most appropriate one among all the acceptable distributions.

2.2.4.4 CI

The CI is constructed based on the results of the three foregoing introduced measures to help determine the optimal regional distribution, especially when the best distributions suggested by the three measures are different. The theory of CI is that the distribution accepted by the majority of goodness-of-fit measures should be considered as the most appropriate regional distribution. For each candidate distribution, the value of CI (T^{DIST}) can be calculated as:

$$\begin{cases} T^{DIST} = A * \mathrm{D}^{DIST} / D^{Th} + B * \left|Z^{DIST}\right| / Z^{Th} + C * \mathrm{G}^{DIST} \\ A+B+C=1 \end{cases} \tag{10}$$

where the values of G^{DIST} are set as 0 for all possible optimal distributions suggested by GMLM and 1 for the other distributions. D^{Th} and Z^{Th} denote the critical values of KPGOF and HWGOF, which are set as 4.61 and 1.64 under the significance level of 5%, respectively. *A*, *B* and *C* are the weights of KPGOF, HWGOF, and GMLM, respectively. The values of *A*, *B*, and *C* are related to the performance of corresponding measures, with a smaller value indicating a more robust performance. In this study, the initial values of weights are all set as 1/3 on the condition that comparisons of the performance of the three measures have never been conducted for the study region. For those unacceptable distributions determined by any of the three measures, the values of T^{DIST} will be set as its upper limit of 1. The distribution with the minimum value of T^{DIST} is considered the optimal regional distribution.

3 Results and discussions

3.1 Trend analysis

The Mann Kendall (MK) test (Romanić et al., 2015) is used to examine the trends of the AMP time series at 93 stations in the SRB from 1960 to 2016. A boxplot of the MK test results of 93 sites is given in Fig. 2 (a). The 5% level is used to determine the significance of the trends. It can be seen from Fig. 2 (a) that all the MK statistics are in the range of -1.96–1.96, which means that the AMP time series at the 93 stations do not have significant changes in their trends. The spatial distribution of the MK results of the 93 stations is presented in Fig. 2 (b). It can be found that the AMP time series at 39 stations show increasing trends, while the remaining 54 stations show decreasing trends. The AMP time series at the 93 stations can be used for the RFA of the EP without consistency correction, as there are no significant changes in the trends of any of them.

3.2 Identification of homogeneous subregions

In this study, four factors, including the longitude, latitude, elevation, and LAMP of each station in the SRB, are employed in FCXB to obtain the optimal division of the study area, and the result is presented in Fig. 3. We can observe that the values of the extended Xie-Benn index do not present an overall monotonic increasing or decreasing tendency when the subregion number increases. This change pattern is consistent with the results in the study by Rao and Srinivas (2006) and Srinivas et al. (2008). The optimum number of subregions of the SRB can be determined to be six because the extended Xie-Benn index obtains the minimum value under this case (Fig. 3). We depict the boundaries of the six subregions in Fig. 4 and list the stations in each subregion in Tab. 1. Generally, subregion I, with the highest elevation and the lowest LAMP of the entire SRB, is located in the west of the SRB, and the whole Ergun River is located in this subregion. Subregion II mainly represents the central areas of the SRB, which is characterized by a large plain area with a low LAMP of approximately 450 mm. This subregion includes the Nenjiang River and the west of the Liao River. Subregion III with an elevation of less than 500 m, is located in the northeast of the SRB. There is not much precipitation in this subregion, and the LAMP of most areas in this subregion is approximately 550 mm. Most areas with low elevation in the west and the center of Liaoning province are associated with subregion IV, and the LAMP over this subregion is 616 mm, which is larger than that of the three previous subregions. The Daling River and the mainstream of Liao River are both located in this subregion, and the southern part of subregion IV is near Bohai Bay of China. Most of the eastern areas of Liaoning province are located in subregion V. The elevation of subregion V is higher than that of subregion IV, and this subregion has the highest LAMP of the whole SRB, as the LAMP in most areas exceeds 750 mm. Subregion VI contains the high-elevation areas in the southeastern SRB. The LAMP of this subregion is almost the same as that of subregion IV, and both are higher than the LAMPs of subregions I and III.

To verify the rationality of the current division of the SRB, we compare it with the previous division pattern in the study by Zhang et al. (2012), which mainly considers the hydrological and geographical characteristics of the SRB. It can be found that these two divisions show a large similarity in the northern SRB but present some discrepancy in the southern SRB with the complicated river system. Such a result shows that the division in this study is not only based on the geographical and hydrological characteristics of the SRB but also on the spatial distribution of the LAMP.

To further verify the reliability and robustness of the results of FCXB, we also apply other validity indices, the fuzzy partition coefficient and fuzzy partition entropy, for the determination of the optimum subregions number of the SRB. The values of the two indices present an overall monotonic increasing or decreasing trend with an increase in the subregion number, suggesting an optimum number of two, but the two subregions have been proven by us to not be capable of passing the homogeneity test. This result is similar to the conclusions in some previous studies (Güler and Thyne, 2004; Hamerly and Elkan, 2002; Srinivas et al., 2008). In addition, considering the large spatial extent of the SRB and the spatial variability of the LAMP in the SRB (Zhang et al., 2012), it is inappropriate to divide the whole SRB into only 2 subregions. Thus, our division of subregions can be considered reasonable and reliable for the RFA of the EP in the SRB. The better performance of the extended Xie-Benn index is related to its strong connection with the input data (Rao and Srinivas, 2006; Xie and Beni, 1991).

After separating the SRB into 6 subregions, the discordancy measure is used to find the grossly discordant stations in each subregion. The critical values ($D_{critical}$) depend on the number of stations in each subregion, which are 2.632, 3, 2.971, 3, 2.869 and 2.971 for subregions I to VI, respectively. The discordancy measure values of all stations are presented in Tab. 1. The results show that the discordancy measure values of all stations are smaller than the regional critical values, which means that all the stations in the SRB have passed the discordancy test.

Then, the homogeneity of each subregion is tested using three heterogeneity measures (H_1, H_2, and H_3), and the results are shown in Tab. 1. It is obvious that the values of the three heterogeneity measures H_1, H_2, and H_3 for the six subregions are less than 1, which demonstrates that all subregions without subjective adjustments can be accepted as homogeneous regions. The results of our study are consistent with the conclusion drawn in the study by Rao and Srinivas (2006), which is that the subregions separated by FCXB are close to being homogeneous.

Tab. 1. The results of discordancy and heterogeneity tests for 93 meteorological stations in SRB.

Homogeneous subregion	Containing stations **Station number** (D_i)	Discordancy measure $D_{critical}$	Heterogeneity Test H$_1$	Heterogeneity Test H$_2$	Heterogeneity Test H$_3$
I (11 sites)	**1** (0.162), **3** (0.964), **5** (0.172), **6** (0.226), **10** (2.464), **11** (1.731), **14** (0.313), **42** (0.576), **48** (2.597), **56** (0.884), **57** (0.912)	2.632	0.837	-1.863	-1.139
II (24 sites)	**7** (1.444), **8** (2.467), **12** (0.270), **13** (1.091), **15** (0.704), **16** (0.292), **17** (0.046), **21** (1.098), **22** (0.796), **23** (2.692), **25** (0.214), **30** (2.582), **31** (0.079), **32** (1.511), **33** (0.270), **34** (0.918), **35** (0.842), **41** (1.092), **43** (0.428), **44** (0.520), **45** (2.702), **46** (0.687), **49** (0.816), **50** (0.440)	3	0.875	0.035	-0.585
III (14 sites)	**2** (0.933), **4** (0.966), **9** (0.370), **18** (0.471), **20** (0.626), **24** (0.845), **26** (1.532), **27** (1.531), **28** (1.341), **29** (0.131), **36** (2.302), **38** (0.404), **39** (1.663), **40** (0.885)	2.971	-0.812	-0.471	-0.508
IV (17 sites)	**58** (2.177), **71** (1.367), **72** (1.201), **73** (0.375), **74** (0.685), **75** (1.211), **76** (1.287), **77** (0.376), **85** (0.755), **86** (0.807), **87** (1.115), **88** (0.535), **89** (1.582), **90** (0.808), **93** (0.137), **94** (1.259), **96** (1.323)	3	-1.293	-1.213	-0.777
V (13 sites)	**61** (0.811), **63** (0.568), **64** (1.725), **65** (0.437), **66** (0.425), **78** (1.198), **79** (1.086), **80** (1.120), **81** (1.563), **82** (0.275), **83** (1.935), **91** (0.477), **92** (1.380)	2.869	0.649	-0.170	-1.196
VI (14 sites)	**37** (0.330), **47** (0.375), **51** (0.565), **52** (1.072), **53** (0.813), **54** (0.605), **55** (1.299), **59** (0.131), **60** (1.720), **62** (1.255), **67** (0.557), **68** (2.819), **69** (0.957), **70** (1.500)	2.971	-0.474	-0.744	-0.920

3.3 Choice of the optimal regional distribution

We first use three different goodness-of-fit measures (HWGOF, GMLM, and KPGOF) to determine the optimal regional distributions in each subregion, and their performance are then compared. Then, the CI is used to give the final decision of the choice of distribution, as it can balance the discrepancy and bias of different measures and provide a more reliable suggestion. The results of HWGOF for the four types of distributions (PE3, GEV, GNO, and GLO) are given in Tab. 2. It can be seen that GEV, PE3, and GNO can be considered as acceptable distributions for subregions I, III and VI since their $\left|Z^{DIST}\right|$ values are no more than the critical value 1.64. However, GNO, with the minimum $\left|Z^{DIST}\right|$ value, is the optimal distribution for these three subregions. Similarly, for subregions II, IV and V, GNO and GEV are both considered acceptable regional models but GEV is determined to be the best model for these subregions.

Fig. 5 presents the results of GMLM for each subregion. It indicates that GNO is the best model for subregions I, III and VI and GEV is the optimal distribution for subregion IV because their theoretical curves are the closest to the regional average L-moment ratios points compared with those of the other distributions. GNO and GEV are both considered acceptable distributions for subregions II and V since their corresponding theoretical curves are both close to the regional average L-moment ratios points, but the optimal distribution cannot be determined by GMLM.

The 95% confidence ellipses used in KPGOF for identifying the acceptable regional distributions are also shown in Fig. 5. It shows that GNO and GEV are acceptable distributions for all subregions since their theoretical curves intersect the confidence ellipses of all subregions. PE3 can be considered an acceptable distributions of subregions I and III for the same reason. The difference between the results of HWGOF and KPGOF is that PE3 is not accepted for subregion VI by KPGOF. Then, the optimal regional distributions are determined from the identified candidate distributions of each subregion, and the results are shown in Tab. 2. It can be found from Tab. 2 that the results of KPGOF are consistent with those of HWGOF.

Tab. 2. The results of HWGOF, KPGOF, and CI for 6 subregions in SRB

('NA' represents the distribution is not accepted as a candidate distribution for this region, which also means the values of these distributions are larger than the critical value of KPGOF).

Measures / Subregion	HWGOF					KPGOF					CI				
	GNO	PE3	GEV	GLO	Best model	GNO	PE3	GEV	GLO	Best model	GNO	PE3	GEV	GLO	Best model
I	**0.608**	-1.104	1.541	3.239	GNO	**0.377**	2.349	4.398	NA	GNO	**0.150**	0.726	0.963	1.000	GNO
II	-0.524	-2.381	**0.466**	3.298	GEV	0.888	NA	**0.158**	NA	GEV	0.503	1.000	**0.106**	1.000	GEV
III	**0.21**	-1.271	0.994	3.25	GNO	**0.001**	3.25	1.162	NA	GNO	**0.042**	0.825	0.618	1.000	GNO
IV	-0.756	-2.554	**0.197**	2.108	GEV	2.424	NA	**0.023**	NA	GEV	0.661	1.000	**0.041**	1.000	GEV
V	-0.413	-1.827	**0.36**	2.275	GEV	0.593	NA	**0.127**	NA	GEV	0.126	1.000	**0.082**	1.000	GEV
VI	**-0.081**	-1.583	0.686	2.767	GNO	**0.14**	NA	0.736	NA	GNO	**0.026**	0.986	0.525	1.000	GNO

With the results of the three measures, the final choices of the most appropriate distributions for each subregion are determined by the CI. The results of CI are also presented in Tab. 2, and the minimum values in the results of CI for each subregion indicate that the most appropriate frequency distributions for subregions I, III and VI and for subregions II, IV and V are still determined to be GNO and GEV, respectively.

To examine the reasonability and reliability of the results of CI, a performance comparison is conducted between CI and the three other measures. Based on previous studies (Hosking and Wallis, 1997; Peel et al., 2001; Vogel and Wilson, 1996) and the results of three measures in this study, it can be accepted that HWGOF, KPGOF, and CI are generally more reliable than GMLM, as GMLM is a subjective measure and may fail to determine the optimal regional distributions in some subregions like subregions II and V. The performance comparison between HWGOF, KPGOF, and CI is slightly adapted from the evaluations in the studies of Hosking and Wallis (1997) and Kjeldsen and Prosdocimi (2015). For each subregion, Monte Carlo simulations are used to generate 500 replicas of this region according to the corresponding optimal regional frequency distribution. Each replica contains the same number of stations and records length as their real-world counterpart. The time series at each site in this subregion is randomly generated according to the real L-moment ratios of this site and the optimal regional distribution. Then, for each replica, the HWGOF, KPGOF, and CI are applied to select the best distribution among GLO, GEV, GNO, and PE3. The percentages of each distribution selected as the optimal one by the different measures are recorded in

Tab. 3. The performance comparison between HWGOF and KPGOF shows that the HWGOF selects the true distributions of subregions II, IV and V more often than KPGOF, while KPGOF performs better in subregions I, III and VI. In addition, it can be generally accepted that both KPGOF and CI perform better than HWGOF in selecting the true regional distributions, as HWGOF selects GEV more often than the true distribution GNO for subregions I, III and VI, while KPGOF and CI always choose the true distribution more often than the other distributions. Moreover, the performance comparison between CI and KPGOF indicates that the correct regional distribution is chosen more often by CI than KPGOF for all subregions except subregion I, while the selected percentage difference of subregion I is not large. Furthermore, for subregions II, IV and V, where HWGOF performs better than KPGOF, CI selects the true distribution GEV more frequently than HWGOF. Thus, it can be concluded that CI performs better than KPGOF and HWGOF in selecting the true distributions, which also means that the results of CI are more reliable.

The advantage of CI is that it can effectively reduce the uncertainty of a single measure. There exist situations in which it is hard for a single measure to identify the most appropriate distribution from several alternatives with similar results (Chen et al., 2014; Du et al., 2014; Hosking and Wallis, 1997). For example, in this study, it is difficult to identify the optimal distributions of GNO and GEV using GMLM for subregions II and V. In addition, there remains a large uncertainty in the process of selecting the optimal distribution for subregion II by HWGOF, as the difference between the values of GNO and GEV is quite small and may be caused by statistical errors. A similar problem was found in the study by Chen et al. (2014), in which the values of GNO and GEV were close when HWGOF was applied to select the optimal distribution for subregion III of the Yangtze River Basin. Therefore, compared with a single measure, CI can identify the differences between all candidate distributions and find the most appropriate distribution accepted by the majority of robust measures, which can largely reduce the uncertainties in the process of choosing the optimal regional distribution. As shown in Tab. 2, GMLM or HWGOF cannot distinguish the better performance of GNO and GEV in subregion II. However, CI definitely suggests GEV as the best regional model since the T^{DIST} value of GEV is much smaller than that of the other distributions.

Tab. 3. The percentages of simulations where a particular selected distribution is chosen by HWGOF, KPGOF, and CI (the bold values represent the percentages of the true regional distributions selected by the three measures).

	HWGOF				KPGOF				CI			
Region (best distribution)	GLO	GEV	GNO	PE3	GLO	GEV	GNO	PE3	GLO	GEV	GNO	PE3
Subregion I (GNO)	15%	36%	**30%**	19%	4%	24%	**44%**	28%	5%	28%	**40%**	27%
Subregion II (GEV)	13%	**66%**	19%	2%	4%	**62%**	32%	3%	5%	**67%**	26%	2%
Subregion III (GNO)	9%	36%	**32%**	23%	1%	27%	**41%**	31%	2%	31%	**41%**	26%
Subregion IV (GEV)	30%	**54%**	14%	2%	13%	**53%**	31%	3%	12%	**56%**	30%	3%
Subregion V (GEV)	29%	**49%**	14%	8%	14%	**47%**	28%	11%	17%	**51%**	23%	10%
Subregion VI (GNO)	9%	39%	**33%**	19%	2%	29%	**41%**	28%	3%	28%	**42%**	27%

3.4 Accuracy analysis of extreme precipitation estimations

To assess the accuracy of EP estimations with different return periods in the SRB, the relative RMSE (Hosking and Wallis, 1997; Yang et al., 2010a) of the estimated quantiles for each station and each subregion is calculated. Tab. 4 lists the values of the regional average relative RMSEs of the estimated quantiles for the six subregions. The relative RMSE values of the EP estimates for the 6 subregions range from 0.054 to 0.160 when the return period is no more than 100 years, which indicates that these quantile estimates of the EP are reliable and can be used with confidence. Similar conclusions were drawn from the accuracy analysis of the EP estimations in other regions of China (Chen et al., 2014; Du et al., 2014; Yang et al., 2010a). Such a phenomenon implies that the estimations of the EP by RFA in most regions of China can be considered accurate when the return level is less than 100 years. Meanwhile, the estimated regional growth curves together with the 95% error bounds for each subregion are presented in Fig. 6. Fig. 6 also shows that the quantile estimates can be valid to use when the return period is less than 100 years. In addition, the estimated regional growth curves of all subregions are concave, but the estimated regional growth curves in most humid regions of China are convex (Chen et al., 2014; Du et al., 2014). Such a result shows that the estimated EP in the SRB has a rapidly increasing tendency with the increase in the return levels. In addition, the regional growth curve of subregion IV is steeper than that of the other 5 regional growth curves, which indicates that the EP increments in subregion IV are larger than those in the other subregions and implies a high risk of EP occurrence in this subregion.

Tab. 4. The values of regional average RMSE of quantile estimates for 6 subregions in SRB.

Return period \ RMSE	Sub-region I	Sub-region II	Sub-region III	Sub-region IV	Sub-region V	Sub-region VI
1 year	0.097	0.104	0.076	0.102	0.100	0.085
2 years	0.061	0.056	0.054	0.063	0.056	0.056
10 years	0.093	0.086	0.079	0.093	0.087	0.083
50 years	0.155	0.147	0.130	0.160	0.152	0.135
100 years	0.187	0.182	0.156	0.209	0.185	0.165

3.5 Return period analysis

The spatial patterns of the EP under different return periods, which can serve as an important indicator for the risk analysis, are investigated with the estimated EP of each station and the spatial interpolation method. In this study, the Inverse Distance Weighting method is adopted to obtain spatial maps of the EP in the SRB under different return periods (T=1, 5, 10, 50 and 100 years), and these spatial maps are presented in Fig. 7. It can be found that the estimated precipitation extremes of the different return periods in the SRB present similar spatial variabilities. The values of the estimated EP decrease from the southern SRB to northern SRB, which means that the values of the estimated EP in subregions IV and V are larger than those in the other subregions. Moreover, the maximum values of the EP can be found in the south of subregions IV and V near the China Yellow Sea and China Bohai Sea, while the minimum extreme values are usually located in the north of subregion I and the northeast of subregion III.

Therefore, the spatial patterns reveal that the estimated EP in the southern subregions (subregions IV and V) of the SRB is much larger than that in the other subregions under the same return period, which means that the risk of high-intensity floods is higher in the southern SRB. The main reasons behind the high values of the EP in the southern SRB can be illustrated from two aspects. First, the increase in precipitation in the southern SRB is related to the influence of the strong variations of the East Asian monsoon, while the impact of the monsoon is not obvious in the northern SRB. In addition, the water vapor pressure is higher in the southern SRB in the summer because most of the water vapor in the SRB comes from the south or, more specifically, from the Bay of Bengal and the western equatorial Pacific Ocean (Wu et al., 2017).

4 Conclusions

In this study, a modified L-moments method is used for a regional extreme precipitation frequency analysis for the Songliao River Basin (SRB). The uncertainties in the identification of homogeneous subregions and in the selection of optimal regional frequency distributions can largely influence the results of the RFA and should be carefully addressed. Based on the original regional L-moments method, the fuzzy c-means method with the extended Xie-Benn index (FCXB) is applied to help determine the optimum number of subregions in the process of identifying homogeneous subregions. Moreover, we develop a new comprehensive index (CI), which gives an integrated consideration to three different goodness-of-fit measures, to reduce the uncertainties in regional frequency distribution. Moreover, the accuracy of the estimated quantiles and the spatial distributions of the estimated precipitation extremes in the SRB are also calculated and analyzed. The main conclusions of this study can be given as follows:

1) No significant change trends can be detected in the AMP series at any stations in the study area. The FCXB suggests that the whole SRB can be divided into six homogenous subregions, and this division is more reliable compared with the results of other cluster methods.

2) The results of CI suggest that GNO is the optimal distribution for subregions I, III and VI and GEV is the best distribution for subregions II, IV and V. The performance comparisons between HWGOF, GMLM, KPGOF, and CI showed that the CI is the most reliable measure of all, as the objective measure CI can select the true distribution more often than the other measures. The results also show that KPGOF performs better than HWGOF since HWGOF selects GEV more often than the true distribution GNO for subregions I, III and VI. In addition, the CI can effectively reduce the uncertainty in the selection of optimal distributions when the distributions suggested by different single measures show differences.

3) The values of the relative RMSE of the estimated precipitation extremes for the 6 subregions range from 0.054 to 0.160 when the return period is no more than 100 years, which indicates a high confidence for the quantile estimates of extreme precipitation.

4) The spatial distributions of the precipitation extremes in the SRB with different return periods (T=1, 5, 10, 50 and 100 years) all show similar decreasing trends from the southern SRB to the northern SRB. Thus, the southern subregions in the SRB (subregions IV and V) have a higher risk of high-intensity floods than the

northern subregions.

The results of this study prove that FCXB and CI are two reliable and effective methods that can help obtain a more robust and reliable result of RFA. This study can also be beneficial for finding the regions in the SRB prone to suffering from EP events and can provide scientific support for local policymakers to determine corresponding measures to reduce losses to the minimum level.

Acknowledgment

This study was financially supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (grant no. XDA23040304), the Research Council of Norway (FRINATEK Project 274310) and the Fundamental Research Funds for the Central Universities (grant no. 2042018kf0222). The meteorological data used in this study were collected from the National Meteorological Administration of China, which is highly appreciated.

References

Alexander, L.V., Arblaster, J.M., 2009. Assessing trends in observed and modelled climate extremes over Australia in relation to future projections. Int. J. Climatol. 29, 417-435. doi: https://doi.org/10.1002/joc.1730

Ashfaq, M., Bowling, L.C., Cherkauer, K., Pal, J.S., Diffenbaugh, N.S., 2010. Influence of climate model biases and daily-scale temperature and precipitation events on hydrological impacts assessment: A case study of the United States. J. Geophys. Res.-Atmos. 115. doi: https://doi.org/10.1029/2009jd012965

Basu, B., Srinivas, V.V., 2015. Analytical approach to quantile estimation in regional frequency analysis based on fuzzy framework. J. Hydrol. 524, 30-43. doi: https://doi.org/10.1016/j.jhydrol.2015.02.026

Chen, Y.D., Zhang, Q., Xiao, M., Singh, V.P., Leung, Y., Jiang, L., 2014. Precipitation extremes in the Yangtze River Basin, China: regional frequency and spatial-temporal patterns. Theor. Appl. Climatol. 116, 447-461. doi: https://doi.org/10.1007/s00704-013-0964-3

Cunnane, C., 1988. Methods and merits of regional flood frequency analysis. J. Hydrol. 100, 269-290. doi: https://doi.org/10.1016/0022-1694(88)90188-6

Du, H., Xia, J., Zeng, S., 2014. Regional frequency analysis of extreme precipitation and its spatio-temporal characteristics in the Huai River Basin, China. Nat. Hazards 70, 195-215. doi:

https://doi.org/10.1007/s11069-013-0808-6

Forestieri, A., Conti, F.L., Blenkinsop, S., Cannarozzo, M., Fowler, H.J., Noto, L., 2018. Regional frequency analysis of extreme rainfall in Sicily (Italy). Int. J. Climatol. 38. doi: https://doi.org/10.1002/joc.5400

Fowler, H.J., Kilsby, C.G., 2003. A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. International Journal of Climatology 23, 1313-1334. doi: https://doi.org/10.1002/joc.943

Güler, C., Thyne, G.D., 2004. Delineation of hydrochemical facies distribution in a regional groundwater system by means of fuzzy c-means clustering. Water Resour. Res. 40. doi: https://doi.org/10.1029/2004WR003299

Halkidi, M., Batistakis, Y., Vazirgiannis, M., 2001. On Clustering Validation Techniques. intelligent information systems 17, 107-145. doi: https://doi.org/10.1023/a:1012801612483

Hall, M.J., Minns, A.W., 1999. The classification of hydrologically homogeneous regions. Hydrol. Sci. J.-J. Sci. Hydrol. 44, 693-704. doi: https://doi.org/10.1080/02626669909492268

Hamerly, G., Elkan, C., 2002. Alternatives to the k-means algorithm that find better clusterings, conference on information and knowledge management, pp. 600-607.

Hao, Z., Aghakouchak, A., Phillips, T.J., 2013. Changes in concurrent monthly precipitation and temperature extremes. Environ Res Lett 8, 034014. doi: https://doi.org/10.1088/1748-9326/8/3/034014

Hosking, J.R.M., Wallis, J.R., 1997. Regional Frequency Analysis: An Approach Based on L-Moments. Cambridge University Press, New York, United States of America.

IPCC, 2013. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, in: Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (Ed.), Cambridge, United Kingdom and New York, NY, USA, p. p. 1535 pp.

Jun, C., Qin, X., Tung, Y., De Michele, C., 2017. Storm event-based frequency analysis method. Hydrol. Res. 49, 700-710. doi: https://doi.org/10.2166/nh.2017.175

Kjeldsen, T.R., Ahn, H., Prosdocimi, I., 2017. On the use of a four-parameter kappa distribution in regional frequency analysis. Hydrol. Sci. J.-J. Sci. Hydrol. 62, 1354-1363. doi: https://doi.org/10.1080/02626667.2017.1335400

Kjeldsen, T.R., Prosdocimi, I., 2015. A bivariate extension of the Hosking and Wallis goodness-of-fit measure for regional distributions. Water Resour. Res. 51, 896-907. doi: https://doi.org/10.1002/2014WR015912

Konisky, D.M., Hughes, L., Kaylor, C.H., 2016. Extreme Weather Events and Climate Change Concern. Clim. Change 134, 533-547. doi: https://doi.org/10.1007/s10584-015-1555-3

Liang, L., Li, L., Liu, Q., 2011. Precipitation variability in Northeast China from 1961 to 2008. J. Hydrol. 404, 67-76. doi: https://doi.org/10.1016/j.jhydrol.2011.04.020

Ma, Z., Li, D., Hu, Y., 2004. The Extreme Dry/Wet Events in Northern China During Recent 100 Years. J. Geogr. Sci. 14, 275-281. doi: https://doi.org/10.1007/bf02837407

Mishra, V., Wallace, J.M., Lettenmaier, D.P., 2012. Relationship between hourly extreme precipitation and local air temperature in the United States. Geophys Res Lett 39, n/a-n/a. doi: https://doi.org/10.1029/2012gl052790

Mukherjee, S., Mishra, A., Trenberth, K.E., 2018. Climate Change and Drought: a Perspective on Drought Indices. Current Climate Change Reports, 1-19. doi: https://doi.org/10.1007/s40641-018-0098-x

Neykov, N.M., Neytchev, P.N., Van Gelder, P.H.A.J.M., Todorov, V.K., 2007. Robust detection of discordant sites in regional frequency analysis. Water Resour. Res. 43. doi: https://doi.org/10.1029/2006WR005322

Ngongondo, C., Xu, C., Tallaksen, L.M., Alemaw, B.F., Chirwa, T., 2011. Regional frequency analysis of rainfall extremes in Southern Malawi using the index rainfall and L-moments approaches. Stoch Environ Res Risk Assess 25, 939-955. doi: https://doi.org/10.1007/s00477-011-0480-x

Peel, M., Wang, Q.J., Vogel, R., Mcmahon, T., 2001. The utility of L-moment ratio diagrams for selecting a regional probability distribution. International Association of Scientific Hydrology Bulletin 46, 147-155. doi: https://doi.org/10.1080/02626660109492806

Qi, J., 2006. Characteristics of temperature and precipitation in Northeast China from 1951 to 2000 (in Chinese). Journal of Meteorology and Environment. doi:

Rao, A.R., Srinivas, V.V., 2006. Regionalization of watersheds by fuzzy cluster analysis. J. Hydrol. 318, 57-79. doi: https://doi.org/10.1016/j.jhydrol.2005.06.004

Romanić, D., Ćurić, M., Jovičić, I., Lompar, M., 2015. Long-term trends of the 'Koshava' wind during the period 1949-2010. Int. J. Climatol. 35, 288-302. doi: https://doi.org/10.1002/joc.3981

She, D., Xia, J., Zhang, D., Ye, A., Sood, A., 2014. Regional extreme-dry-spell frequency analysis using the L-moments method in the middle reaches of the Yellow River Basin, China. Hydrol. Process. 28, 4694-4707. doi: https://doi.org/10.1002/hyp.9930

Song, X., Li, L., Fu, G., Li, J., Zhang, A., Liu, W., Zhang, K., 2014. Spatial–temporal variations of spring drought based on spring-composite index values for the Songnen Plain, Northeast China. Theor. Appl. Climatol. 116, 371-384. doi: https://doi.org/10.1007/s00704-013-0957-2

Song, X., Song, S., Sun, W., Mu, X., Wang, S., Li, J., Li, Y., 2015. Recent changes in extreme precipitation and drought over the Songhua River Basin, China, during 1960–2013. Atmos. Res. 157, 137-152. doi: https://doi.org/10.1016/j.atmosres.2015.01.022

Srinivas, V.V., Tripathi, S., Rao, A.R., Govindaraju, R.S., 2008. Regional flood frequency analysis by combining self-organizing feature map and fuzzy clustering. J. Hydrol. 348, 148-166. doi: https://doi.org/10.1016/j.jhydrol.2007.09.046

Su, B., Gemmer, M., Jiang, T., 2008. Spatial and temporal variation of extreme precipitation over the Yangtze River Basin. Quat. Int. 186, 22-31. doi: https://doi.org/10.1016/j.quaint.2007.09.001

Tao, H., Gemmer, M., Bai, Y., Su, B., Mao, W., 2011. Trends of streamflow in the Tarim River Basin during the past 50 years: Human impact or climate change? J. Hydrol. 400, 1-9. doi: https://doi.org/10.1016/j.jhydrol.2011.01.016

Vogel, R.M., Mcmahon, T.A., Chiew, F.H.S., 1993. Floodflow frequency model selection in Australia. J. Hydrol. 146, 421-449. doi: https://doi.org/10.1016/0022-1694(93)90288-K

Vogel, R.M., Wilson, I., 1996. Probability Distribution of Annual Maximum, Mean, and Minimum Streamflows in the United States. J. Hydrol. Eng. 1, 69-76. doi: https://doi.org/10.1061/(ASCE)1084-0699(1996)1:2(69)

Wang, B., Zhang, M., Wei, J., Wang, S., Li, X., Li, S., Zhao, A., Li, X., Fan, J., 2013. Changes in extreme precipitation over Northeast China, 1960–2011. Quat. Int. 298, 177-186. doi:

https://doi.org/10.1016/j.quaint.2013.01.025
Wu, Q., Zhao, C., Niu, D., Cao, H., Zhang, Z., Wang, D., Xu, F., 2017. Climate characteristics of flood season rainstorms from 1960 to 2014 in Liaoning province (in Chinese). Journal of Meteorology and Environment 33, 37-43. doi:
Xie, X.L., Beni, G., 1991. A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Mach. Intell. 13, 841-847. doi: https://doi.org/10.1109/34.85677
Yang, T., Shao, Q., Hao, Z., Chen, X., Zhang, Z., Xu, C., Sun, L., 2010a. Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River Basin, China. J. Hydrol. 380, 386-405. doi: https://doi.org/10.1016/j.jhydrol.2009.11.013
Yang, T., Xu, C., Shao, Q., Chen, X., 2010b. Regional flood frequency and spatial patterns analysis in the Pearl River Delta region using L-moments approach. Stoch Environ Res Risk Assess 24, 165-182. doi: https://doi.org/10.1007/s00477-009-0308-0
Zhang, J., Hall, M.J., 2004. Regional flood frequency analysis for the Gan-Ming River basin in China. J. Hydrol. 296, 98-117. doi: https://doi.org/10.1016/j.jhydrol.2004.03.018
Zhang, X., Xu, K., Zhang, D., 2012. Risk assessment of water resources utilization in Songliao Basin of Northeast China. Environ. Earth Sci. 67, 1319-1329. doi: https://doi.org/10.1007/s12665-012-1575-5

List of Fig. Captions

points and blue circle points indicate the stations with increasing and decreasing trends, respectively.

Fig. 3. The values of the extended Xie-Benn index with different numbers of subregions.

Fig. 4. The division of 6 homogeneous subregions in SRB (a) and the long-term annual mean precipitation of each subregion (b).

Fig. 5. The L-diagrams for AMP at 6 subregions. The black circle points represent L-skewness and L-kurtosis of each station, the plus signal points represent the regional average L-skewness and L-kurtosis. The black ellipses represent the confidence regions with the 5% significance level.

Fig. 6. The estimated regional growth curves of AMP with 95% error bounds for six subregions. The red lines represent the regional growth curves and the black lines represent the 95% error bounds, and the grey shadow regions denote the confidence intervals.

Fig. 7. The spatial distributions of estimated precipitation extremes in SRB when the return period equals to 1, 5, 10, 50 and 100 years.

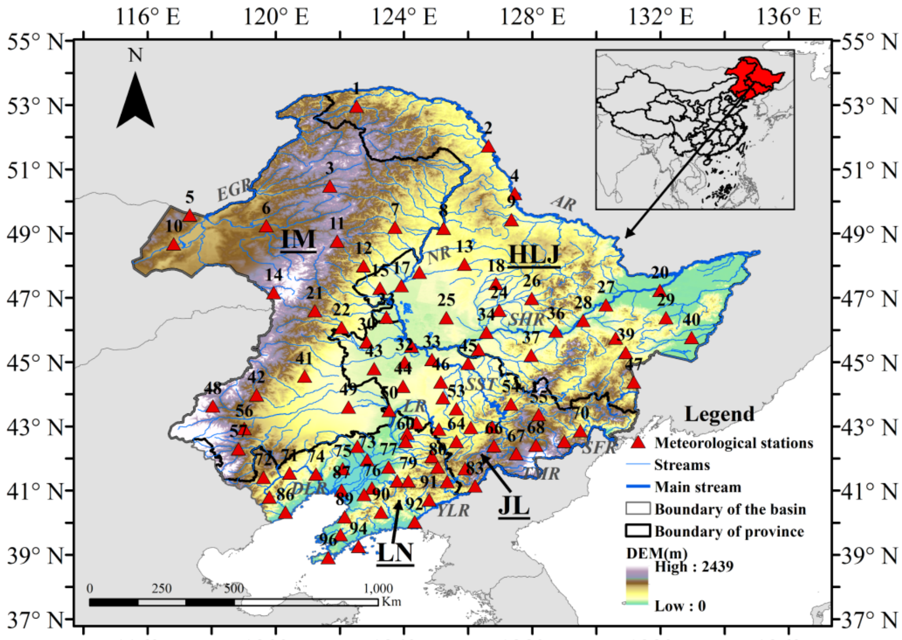


Fig.1. The location of the Songliao River Basin (SRB) in the northeastern China. The meteorological stations are represented by the red triangle points. The underlined words in black color represent the names of provinces, and IN, LN, JL, and HLJ represent Inner Mongolia province, Liaoning province, Jilin province, and Heilongjiang province, respectively. The gray italic words are the abbreviated names of main rivers. From the top to the bottom of this Fig., EGR, AR, NR, SHR, SST, LR, SFR, DLR, TMR, YLR represent Ergun River, Amur River, Nenjiang River, Songhua River, Second Songhua Tributary, Liao River, Suifen River, Daling River, Tumen River, Yalu River, respectively.

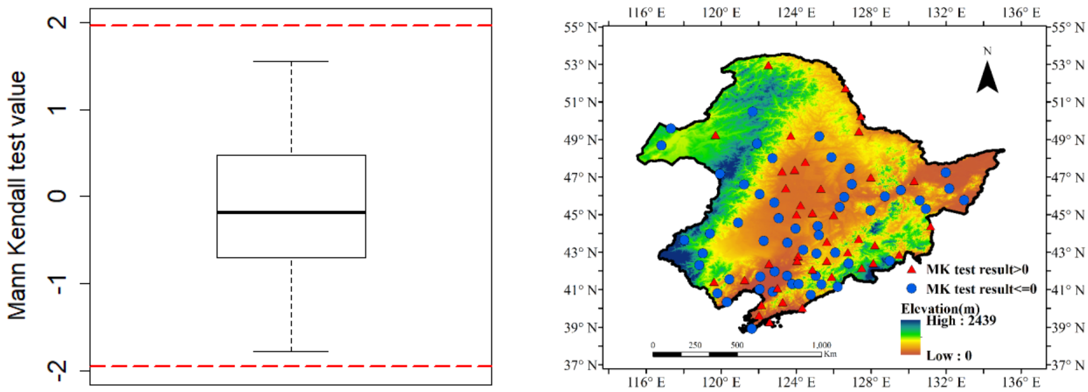


Fig.2. The boxplot of the Mann Kendall test results of 93 stations (a). The red horizontal dash lines represent the critical value (±1.96) for the MK test at 5% significance level. The spatial pattern of the Mann-Kendall test result for the SRB over the period 1960–2016 (b). The red triangle points and blue circle points indicate the stations with increasing and decreasing trends, respectively.

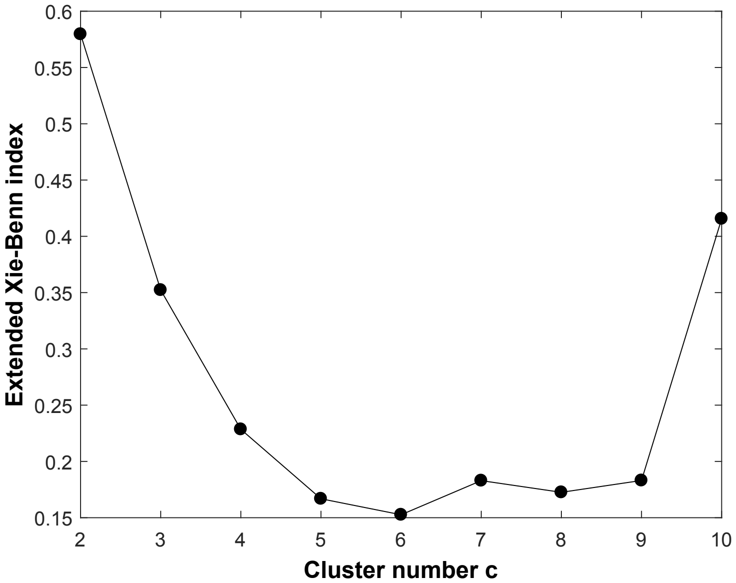
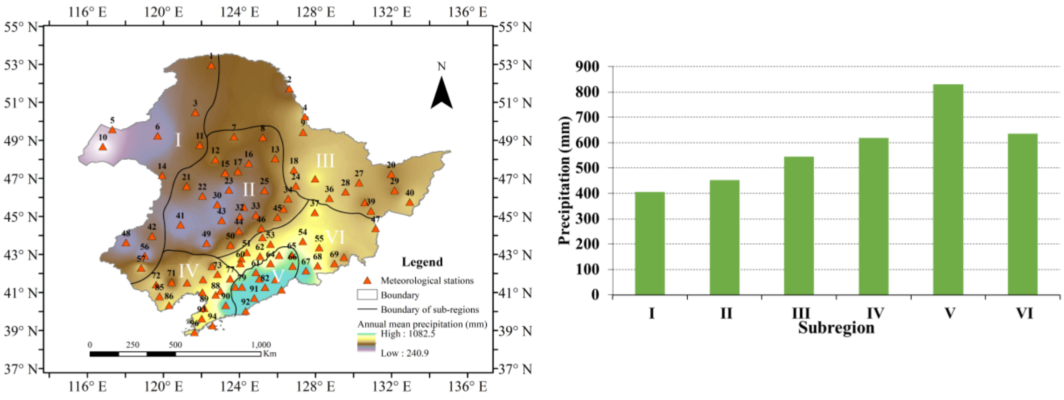


Fig.3. The values of the extended Xie-Benn index with different numbers of subregions.



(a) (b)

Fig.4. The division of 6 homogeneous subregions in SRB (a) and the long-term annual mean precipitation of each subregion (b).

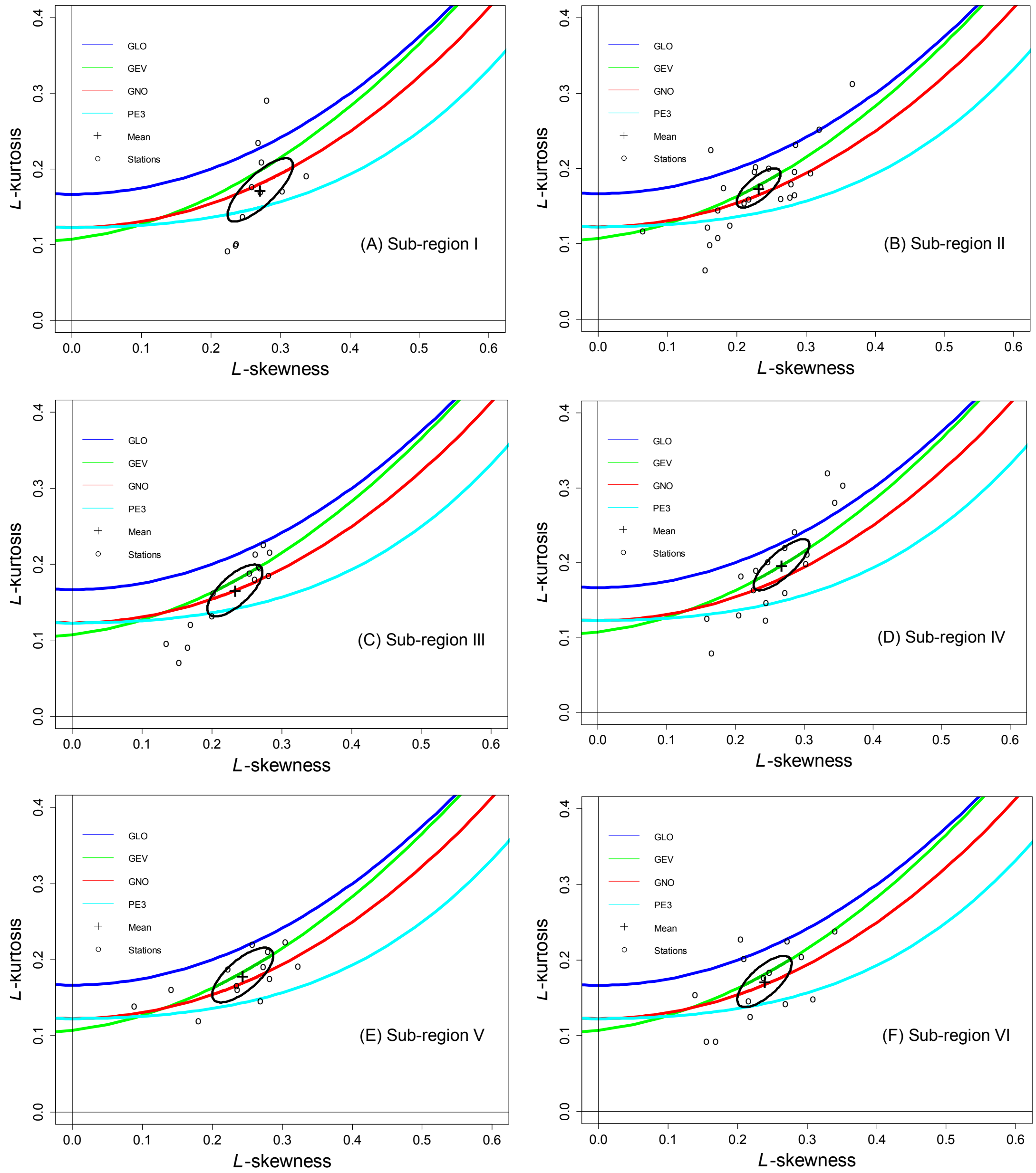


Fig.5. The L-diagrams for AMP at 6 subregions. The black circle points represent L-skewness and L-kurtosis of each station, the plus signal points represent the regional average L-skewness and L-kurtosis. The black ellipses represent the confidence regions with the 5% significance level.

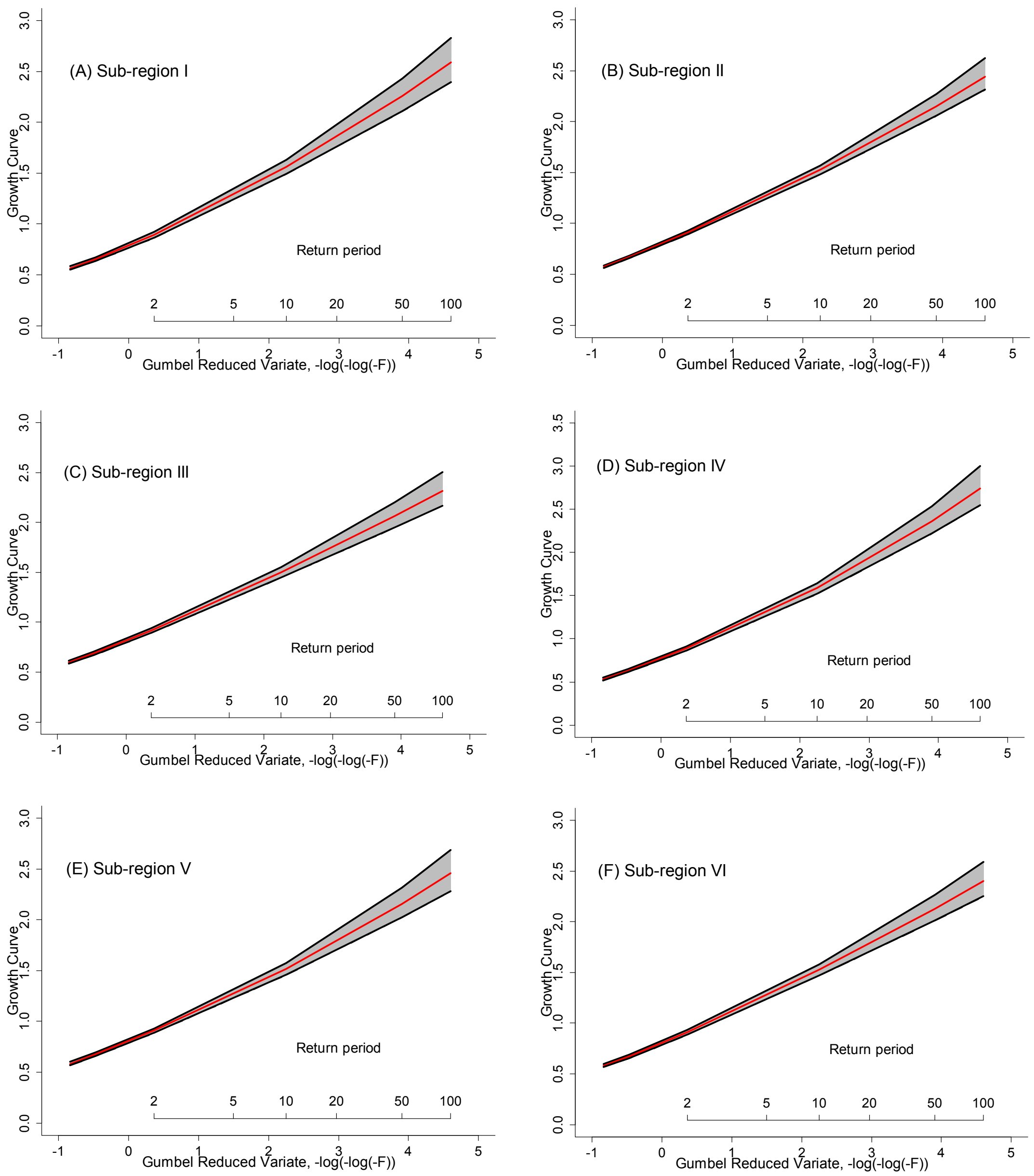


Fig.6. The estimated regional growth curves of AMP with 95% error bounds for six subregions. The red lines represent the regional growth curves and the black lines represent the 95% error bounds, and the grey shadow regions denote the confidence intervals.

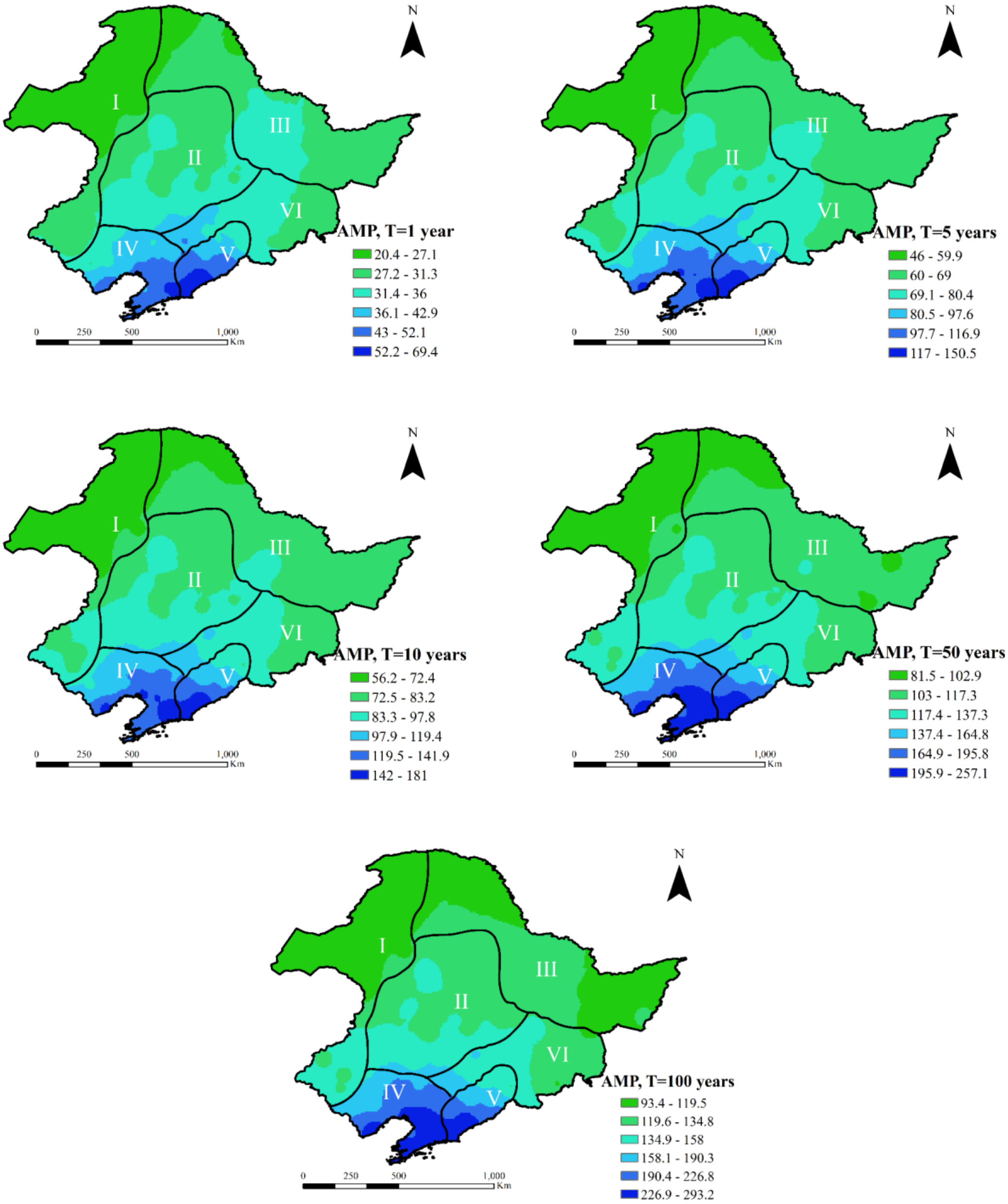


Fig.7. The spatial distributions of estimated precipitation extremes in SRB when the return period equals to 1, 5, 10, 50 and 100 years.