

RAINFALL FREQUENCY ANALYSIS USING FRECHET AND LOG-LOGISTIC DISTRIBUTIONS OF SITES OF AZAD JAMMU AND KASHMIR, PAKISTAN

HUSSAIN, Z.^{1*} – ABBAS, K.²

¹*Research Center for Modelling and Simulation (RCMS), National University of Sciences and Technology (NUST), H-12, Islamabad, Pakistan*

²*Department of Statistics, University of Azad Jammu and Kashmir, Muzaffarabad, Pakistan
(phone: +92-51-9085-5739)*

**Corresponding author*

e-mail: zami_cr@yahoo.com, zamir@rcms.nust.edu.pk

(Received 10th Apr 2019; accepted 4th Jul 2019)

Abstract. The paper illustrates the application of at-site frequency analysis using relatively non-traditional probability distributions adopting four methods of parameter estimation using annual maximum rainfall series (AMRS) (from 1980 to 2015) of three sites; Muzaffarabad, Garhi Dupatta and Kotli of Azad Jammu and Kashmir, Pakistan. Critical assumptions related to AMRS for frequency analysis have been validated through run test (to check the randomness of AMRS at each site) and Mann Whitney test (to check the distribution of AMRS at given sites as identical or not). Two probability distributions; Frechet and Log-logistic have been used as candidate distributions to model AMRS at given sites. The parameters of these distributions have been estimated using four methods; Probability Weighted Moments Estimation (PWME), Maximum Likelihood Estimation (MLE), Bayesian Estimation (BE) and Moments Estimation (ME) methods. The Kolmogorov Smirnov test has been used as goodness-of-fit criteria. The results show suitability of the selected probability distributions for the observed AMRS series at given sites with BE and MLE methods as adequate for estimation. Moreover, the results indicate that ME method is not a preferred choice for the estimation of parameters and quantiles for various return periods.

Keywords: *at-site frequency analysis, Bayesian estimation, maximum likelihood estimation, method of moments, probability weighted moments*

Introduction

Frequency analysis is a technique for estimation of a specific event that occurs frequently. Estimation of a return period of geophysical events, for instance, extreme rainfall or floods of a site or many sites (together, assuming a homogeneous region) which occurs rarely is a familiar problem in environmental engineering. For rainfall frequency analysis (RFA), usually, at site and regional approaches with various typical probability distributions are in practice for the estimation of quantiles. Success depends on the availability and span of the observed data along with nature of the associated variables (site characteristics) having strong correlations with the observed data series at given sites. With the introduction of several new class of probability distributions in the literature, it is very important to examine the suitability of these distributions with optimum method of estimation for analysing the extremes of geophysical, hydrological and meteorological events like earthquakes, floods, rainfall, etc. It will bring variability in the choice of selecting best-fit distribution for the estimation of extreme events. The current study has used two atypical probability distributions in the family of extreme value distributions namely Frechet distribution (FD) and Log-logistic distribution

(LLD), with four methods of estimation of parameters to estimate rainfall quantiles for T-years return period.

Climate is a composition of temperature, sunshine, pressure, wind speed, humidity, rainfall and clouds, etc. Rainfall is an important characteristic as it can cause disasters like land slide or floods if extremely befallen. Scarcity of rainfall also affects the water management systems hence impacts agricultural productions. Pakistan is a populated country, economy based on agriculture and high vulnerable index to ecological disasters, like rainfall and floods. Therefore, it is also important to have estimates of rainfall using suitable methods in numerous climate zones of the country. This study illustrates the results of at-site frequency analysis using two relatively new probability distributions with four methods of estimation using rainfall data of three sites of Azad Jammu and Kashmir, Pakistan. The results of this study will be useful in terms of introducing new probability distributions, to bring variety and flexibility in the choice of probability distributions, for modelling of extreme events with various methods of estimation. The estimated quantiles with best suitable estimation method(s) would also be useful for officials dealing with water resources management and disasters management of the study area.

Based on the aforementioned details, the major objectives of the study are:

- Assessing the suitability of FD and LLLD for modelling extremes of rainfall using four methods of estimation namely Maximum Likelihood Estimation (MLE), Bayesian Estimation (BE), Moments Estimation (ME) and Probability Weighted Moments Estimation (PWME)
- Estimating rainfall quantiles for various return periods of the most suitable method(s) of estimation

Moreover, to establish the usefulness of the adopted methodology, in general, the quantile estimates obtained through this study has been compared with the results of the previously published study. The comparison reveals that the results are comparable in terms of forecasting the magnitude of extreme events at given sites for shorter as well as longer return periods.

Literature review

Various probability distributions with different methods of estimations using regional as well as at-site approach have been applied in the literature for modelling rainfall data all over the world including Pakistan. Some of the published international and indigenous studies are summarized below.

International studies

Zalina et al. (2002) performed rainfall frequency analysis using eight probability distributions including Gamma, Generalized Normal, Generalized Pareto, Generalized Extreme Value, Gumbel, Log Pearson Type III, Pearson Type III and Wakeby distribution to estimate extreme rainfall in Malaysia. The variable of analysis was annual maximum rainfall series of seventeen stations located in Peninsular Malaysia. The study concluded that the Generalized Extreme Value distribution is the most appropriate distribution for the study area.

Mohamed and Ibrahim (2016) analyzed annual rainfall data for fourteen rainfall stations in Sudan. Five distributions namely Normal, Log Normal, Gamma, Weibull and

Exponential have been used and the study concluded that Normal and Gamma distributions are suitable probability distributions to model rainfall data.

Bahrawi (2018) investigates annual rainfall data of more than 20 stations in Makkah Al-Mukarramah, Saudi Arabia using Lognormal, Normal, Extreme Value Type I, Pearson Type III and Log-Pearson Type III distributions. Different distributions have been identified as best-fitted distribution for different regions with suitability of Log-normal for 36% of the rainfall stations.

Few other studies include the work of Mo et al. (2019) in South China using Generalized Extreme Value and Generalized Pareto distributions with MLE method for analyzing rainfall data of eight gauging stations and concluded that Generalized Pareto distribution performed better relative to Generalized Extreme Value distribution. Liang et al. (2017) performed L-moments based regional frequency analysis and compared the results with at-site analysis using Generalized Extreme Value distribution analyzing rainfall data in the Taihu basin China; Wang et al. (2017) applied regional approach using fuzzy c-means and L-moments to analyze precipitation extremes in Mainland China; and Alam et al. (2018) performed at-site rainfall frequency analysis using Generalized Extreme Value, Pearson Type III and Log Pearson Type III probability distributions for various sites of Bangladesh and concluded that different probability distributions are suitable for different sites with Generalized Extreme Value distribution suitable for 36% of the stations.

Case studies of Pakistan

There are few published studies performing rainfall frequency analysis of different sites/areas of Pakistan using at-site or regional approach with various probability distributions using different methods of estimation of parameters and quantiles. Some of them are briefly summarized below:

Amin et al. (2016) performed at-site frequency analysis using four probability distributions; Log Pearson Type-III, Normal, Gumbel Maximum and Log-Normal distributions to model annual maximum rainfall for sites in the northern areas of Pakistan. The study concluded that the Normal and Log Pearson Type-III distributions are best-fit distributions for the study area. Iqbal and Ali (2013) performed at-site frequency analysis using extreme annual rainfall of four cities of Punjab and concluded that the General Extreme Value distribution is the best-fitted distribution for the study area. Shahzadi (2013) and Shahzadi et al. (2013) performed L-moments based regional frequency analysis using annual maximum rainfall series of 23 stations of Pakistan. The study area was divided into three homogenous regions. The study concluded that Generalized Normal, Generalized Extreme Value and Generalized Logistic distributions are suitable to model the observed data series for the regions under study. The study area of Shahzadi et al. (2013) includes the sites used in this study with similar variable of analysis, i.e. annual maximum rainfall series. Therefore, a comparison of the estimates provided by Shahzadi et al. (2013) with the results of current study has been provided in the later section of the paper, to establish the usefulness of the current methodology. In another study, Hussain et al. (2017) illustrated the application of regional rainfall frequency analysis using rainfall data of seven sites of Sindh, Pakistan with annual maximum monthly rainfall totals as variable of analysis. The study concluded that three distributions namely: Generalized Normal, Pearson Type III and Generalized Pareto are suitable for the observed data series of the study region. Ahmad et al. (2017) performed L-moments based regional frequency analysis using annual total

rainfall of 30 met sites of Pakistan. The study area (includes sites from Sindh, Punjab, Balochistan and Northern Areas) was divided into four homogeneous regions with different three-parameters probability distributions identified as best-fitted probability distributions. Ahmad et al. (2016) performed at site frequency analysis using annual daily maximum rainfall series. Different probability distributions have been used with methods of estimations as L-moments, TL moments and LH moments. The study area includes 9 sites from Sindh, 7 sites from Balochistan, 6 sites from Punjab, 3 sites from Khyber Pakhtunkhwa and 3 sites from Northern Areas of Pakistan. Hussain et al. (2010) applied Box-cox transformed hierarchical Bayesian multivariate spatio-temporal interpolation method to estimate space time monthly precipitation in the monsoon periods during 1974 to 2000, using 27 years monthly average rainfall data of 51 stations in Pakistan. Khan et al. (2017) performed regional frequency analysis using monthly precipitation data from 1999 to 2012 at 17 stations of Northern areas and Khyber Pakhtunkhwa. L-moments and partial L-moments have been used as estimation methods with four three-parameter probability distributions. The study concluded the suitability of Generalized Pareto distribution using L-moments and Generalized Normal distribution using partial L-moments for the study region.

The aforementioned details show that both at-site and regional approaches have been adopted to analyse rainfall frequency analysis in Pakistan using some typical probability distributions with various methods of estimation of parameters but non of the studies have used FD and LLD to model the extremes of rainfall in Pakistan.

Materials and methods

Study area and data utilized

The study area consists of three sites of Azad Jammu and Kashmir (AJ&K) namely Muzaffarabad, Kotli and Garhi Dupatta. AJ&K commonly known as Pakistan administered Kashmir borders Pakistan's Punjab province to the south and Khyber Pakhtunkhwa province to the west. It has a total area of 13,297 km² with a total population of 4,045,366 as per the 2017 census of Pakistan. The northern part of AJ&K encompasses the lower area of the Himalayas. The region receives rainfall in both the winter and the summer. However, in the summer season, monsoon floods of the rivers Jhelum and Leepa are common due to extreme rainfall and snow melting. The southern parts of AJ&K including Kotli usually have extremely hot weather in summers and moderate cold weather in winters. This area receives rainfall mostly in monsoon (from June to September). The central and the northern parts of AJ&K including Muzaffarabad have moderate hot weather in summers and very cold (chilly) in winter with snowfall occurs in December and January.

Annual maximum rainfall series (AMRS) in millimeters, i.e. maximum value from the daily rainfall record of each year has been selected for the analysis, which usually considered as an important variable for rainfall frequency analysis as used in a number of case studies mentioned in the literature review section. Data for the years 1980 to 2015 have been used, provided by Pakistan Meteorological Department, head office Karachi, Pakistan. The site characteristics along with few summary statistics and geographical locations of the three sites are presented in *Table 1* and *Figure 1*, respectively. The descriptive statistics provided in *Table 1* are useful to understand the general trends and tendencies of AMRS at given sites.

Among the three sites, Garhi Dupatta has the highest average value of observed AMRS as 75.72 mm whereas the average of observed AMRS at Kotli and Muzaffarabad is 59.91 mm and 74.46 mm, respectively. To observe the dispersion around mean of the given data series, standard deviation has been calculated which shows that there exist variations in the observed data series at different sites. The highest variation has been observed in AMRS of site Muzaffarabad.

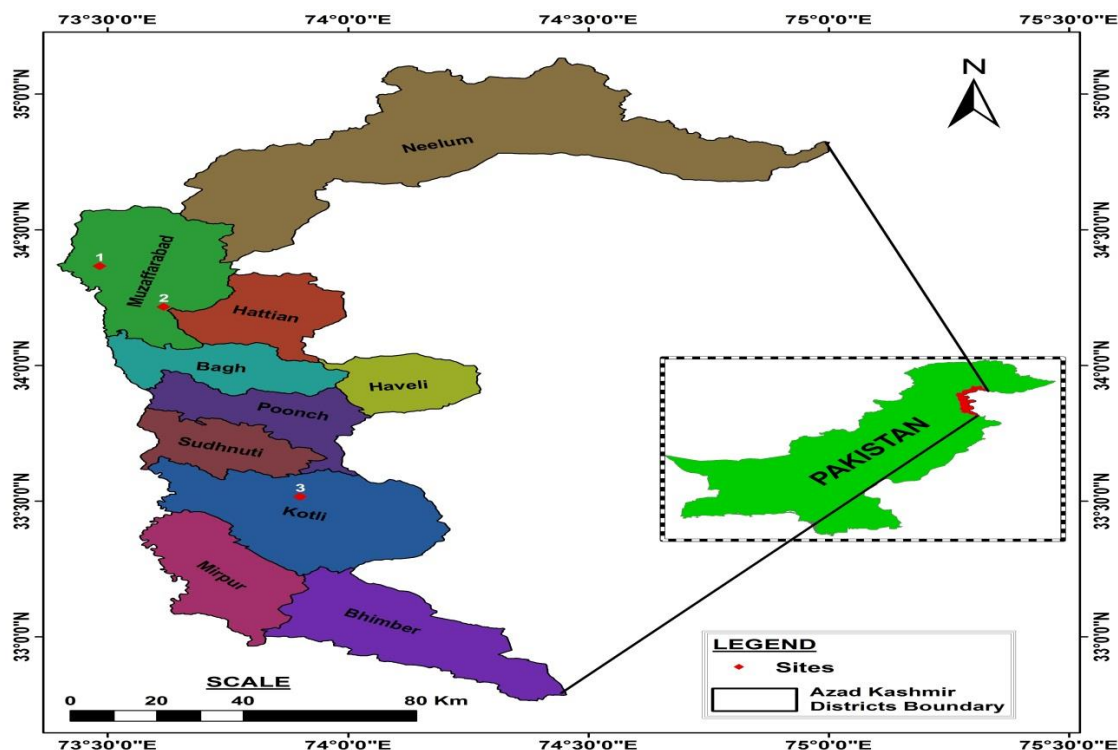


Figure 1. Study area and geographical locations of the three sites of AJ&K, Pakistan

Table 1. Site characteristics and descriptive statistics of the data series at three sites

Site name	Lat	Lon	Ele (m)	n	Mean (mm)	SD (mm)	CV	Min (mm)	Max (mm)	Skew
Garhi Dupatta	34°13'	73°37'	813.50	36	75.72	11.32	14.95	32.15	184.27	1.12
Kotli	33°31'	73°54'	614.00	36	59.91	24.55	40.98	28.72	129.28	1.29
Muzaffarabad	34°22'	73°29'	702.04	36	74.46	31.43	42.21	30.55	153.98	1.09

Lat, Lon, and Ele are the Latitude, Longitude and Elevation of the sites, mm denotes millimeters while n, SD, CV, Min, Max and Skew are the number of observations, standard deviation, coefficient of variation, minimum value, maximum value and moments measure of skewness of the observed data series at each site, respectively

Coefficient of variation (CV) has been calculated as a relative measure of dispersion. The values of CV show that the least consistency exists in AMRS of Muzaffarabad, followed by Kotli and Garhi Dupatta. The coefficient of skewness ranges from 1.09 to 1.29 indicates that the distribution associated to the observed data series at these sites is positively skewed (non-normal). Therefore, it would be appropriate to select positively skewed distribution(s) as candidate(s) for the observed data series at given sites. The

maximum of the observed AMRS at respective sites so far is 184.27 mm at Garhi Dupatta, 129.28 mm at Kotli and 153.98 mm at Muzaffarabad.

For rainfall frequency analysis, there exists certain commonly used probability distributions with variations in the methods of estimation as discussed in the literature review. However, this study has introduced FD and LLD to model the rainfall data at given sites using different methods of estimation. The illustration of the results of this application provides flexibility to the researchers performing inferential statistics of extreme events using probability distribution(s).

Basic assumptions of the observed data series

The procedure of rainfall frequency analysis is usually based on few assumptions related to the observed data series at given sites. For instance, the observed data series at given sites is random and identically distributed. Therefore, to check these assumptions Run test for randomness (Bradley, 1968) and Maan-Whitney test (Mann and Whitney, 1947) for identical distribution of the data series at given site have been utilized. Furthermore, Kolmogorov Smirnov (KS) test has been used as goodness-of-fit measure.

Estimation methods

For the two selected probability distributions, i.e. FD and LLD, four estimation methods, i.e. MLE, PWME, ME and BE methods have been used for the estimation of parameters and quantiles. Few details related to the procedures are:

FD was introduced by Maurice Frechet in 1927 for largest extremes (Frechet, 1927). It is a special case of Extreme value distribution (EVDs) and used to model the extreme events like floods, maximum rainfall, river discharge in hydrology, etc. The reason for selecting this probability distribution for the analysis is that various published studies related to rainfall frequency analysis found Generalized Extreme Value distribution as appropriate to model extremes of rainfall in Pakistan. Therefore, the idea is to provide a suitable alternative to bring flexibility in the choice of available probability distributions for modeling rainfall data. The probability density function (PDF) of FD is

$$f(x, \alpha, \beta) = \frac{\alpha}{\beta} \left(\frac{\beta}{x}\right)^{\alpha+1} \exp\left[-\left(\frac{\beta}{x}\right)^{\alpha}\right], x > 0, \alpha, \beta > 0, \quad (\text{Eq.1})$$

where α is the shape parameter and β is the scale parameter.

The LLD is also a type of extreme value distribution. It can be used for model the extreme values like maximum daily one day rainfall, river discharge per day or per month, etc. The PDF of LLD is:

$$f(x, \alpha, \beta) = \frac{\left(\frac{\beta}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\beta-1}}{\left[1+\left(\frac{x}{\alpha}\right)^{\beta}\right]^2}, x > 0, \alpha, \beta > 0, \quad (\text{Eq.2})$$

where α is scale and β is shape parameter.

Estimates of parameters for FD using MLE

Let $x_1, x_2, x_3, \dots, x_n$ be a random sample of size 'n' from two parameters FD, then the likelihood of function of Equation 1 is:

$$L(x, \alpha, \beta) = \alpha^n \beta^{n\alpha} \prod_{i=1}^n x_i^{-(\alpha+1)} \exp \left[-\sum_{i=1}^n \left(\frac{\beta}{x_i}\right)^\alpha \right], \quad (\text{Eq.3})$$

The log-likelihood function is:

$$\log L(x, \alpha, \beta) = n \log \alpha + n \log \beta - (\alpha + 1) \sum_{i=1}^n \log x_i - \sum_{i=1}^n \left(\frac{\beta}{x_i}\right)^\alpha, \quad (\text{Eq.4})$$

The score equations are

$$\frac{\partial \log L}{\partial \alpha} = \frac{n}{\alpha} + n \log \beta - \sum_{i=1}^n \log x_i - \sum_{i=1}^n \left(\frac{\beta}{x_i}\right)^\alpha \log \left(\frac{\beta}{x_i}\right) = 0, \quad (\text{Eq.5})$$

$$\frac{\partial \log L}{\partial \beta} = \frac{n\alpha}{\beta} - \alpha \beta^{(\alpha-1)} \sum_{i=1}^n \left(\frac{1}{x_i}\right)^\alpha = 0. \quad (\text{Eq.6})$$

The above equations cannot be written in closed form, therefore, the BFGS (Broyden Fletcher GoldFarb Shanno), available in the R software, have been used to estimate α and β with MLE.

Estimates of parameters for FD using PWME

The Probability weighted Moments proposed by Greenwood et al. (1979) has the following form

$$M_{p,r,s} = \int X^p \{F(x)^r\} \{1 - F(x)\}^s dF(x), \quad (\text{Eq.7})$$

Hosking et al. (1985) presented the special case

$$B_r = E[X\{F(X)\}^r], r = 0, 1, 2, 3, \dots \quad (\text{Eq.8})$$

For any population, the r^{th} non-central moments can be expressed as

$$\alpha_r = M_{1,0,r} = E[X\{1 - F(X)\}^r], \quad (\text{Eq.9})$$

by putting $r = 1, 0$ in Equation 9,

$$\alpha_0 = \beta \tau \left(1 - \frac{1}{\alpha}\right), \alpha > 1, \quad (\text{Eq.10})$$

$$\alpha_1 = \frac{\beta \tau (1 - \frac{1}{\alpha}) (2 - 2\frac{1}{\alpha})}{2}, \quad (\text{Eq.11})$$

from observed sample, the corresponding r^{th} moments are

$$\alpha_r = n^{-1} \frac{\sum_{i=1}^n \binom{n-i}{r} X_i}{\binom{n-1}{r}}, \quad (\text{Eq.12})$$

by substituting $r = 1, 0$ in Equation 12,

$$a_0 = \bar{X}, \quad (\text{Eq.13})$$

$$a_1 = \frac{\sum_{i=1}^n (n-1)x_i}{n(n-1)}, \quad (\text{Eq.14})$$

Then equating *Equations 10 and 11* to *Equation 13 and 14*, estimates of α and β for FD using PWM are:

$$\hat{\alpha}_{PWM} = \frac{\log 2}{\log 2 + \log a_0 - \log a_1}, \quad (\text{Eq.15})$$

$$\hat{\beta}_{PWM} = \frac{\bar{X}}{\tau(1-\frac{1}{\alpha})}, \alpha > 1 \quad (\text{Eq.16})$$

Estimates of parameters for FD using ME

The r^{th} moment for FD is

$$E(X^r) = \beta^r \tau \left(1 - \frac{r}{\alpha}\right), \quad (\text{Eq.17})$$

by substituting $r = 1, 2$ respectively in *Equation 17*, the first and second moments for FD, i.e.

$$m_1^* = \bar{X}, \quad (\text{Eq.18})$$

$$m_2^* = E(X^2) = \beta^2 \tau \left(1 - \frac{2}{\alpha}\right), \quad (\text{Eq.19})$$

By solving *Equations 18 and 19* using iterative method, the estimates of α and β for FD using ME can be obtained.

Estimates of parameters for FD using BE

The estimates of α and β using BE can be obtained by considering the reference prior, developed by Abbas and Tang (2015), i.e.

$$\pi_R(\alpha, \beta) = \frac{1}{\alpha\beta}, \quad (\text{Eq.20})$$

Then the joint posterior distribution for α and β is

$$\pi(\alpha, \beta | x) = \frac{\alpha^{n-1} \beta^{n\alpha-1} \pi_{i=1}^n x_i^{-(\alpha+1)} \exp[-\sum_{i=1}^n (\frac{\beta}{x_i})^\alpha]}{\int_0^\infty \int_0^\infty \alpha^{n-1} \beta^{n\alpha-1} \pi_{i=1}^n x_i^{-(\alpha+1)} \exp[-\sum_{i=1}^n (\frac{\beta}{x_i})^\alpha] d\alpha d\beta}, \quad (\text{Eq.21})$$

Laplace approximation can be used to obtain the estimates of α and β using BE.

Estimates of parameters for LLD using MLE

Let $x_1, x_2, x_3, \dots, x_n$ be a random sample of size ' n ' from LLD, then the likelihood of function of *Equation 2* is:

$$L(x_i; \alpha, \beta) = \beta^n \alpha^{-n\beta} \prod_{i=1}^n x_i^{\beta-1} \prod_{i=1}^n [1 + \left(\frac{x_i}{\alpha}\right)^\beta]^{-2}, \quad (\text{Eq.22})$$

Then the log-likelihood function can be written as:

$$\log L(x_i; \alpha, \beta) = n \log \beta - n\beta \log \alpha + (\beta - 1) \sum_{i=1}^n \log x_i - 2 \sum_{i=1}^n \log \left[1 + \left[\frac{x_i}{\alpha} \right]^\beta \right], \quad (\text{Eq.23})$$

Derivatives of Equation 23 with respect to α and β are:

$$\frac{\partial \log L}{\partial \alpha} = -\frac{n\beta}{\alpha} + \frac{2\beta}{\alpha} \sum_{i=1}^n \frac{\left(\frac{x_i}{\alpha}\right)^\beta}{\left(1 + \left(\frac{x_i}{\alpha}\right)^\beta\right)} = 0, \quad (\text{Eq.24})$$

$$\frac{\partial \log L}{\partial \beta} = \frac{n}{\beta} - n \log \alpha + \sum_{i=1}^n \log x_i - 2 \sum_{i=1}^n \frac{\left(\frac{x_i}{\alpha}\right)^\beta \log \left(\frac{x_i}{\alpha}\right)}{\left[1 + \left(\frac{x_i}{\alpha}\right)^\beta\right]} = 0. \quad (\text{Eq.25})$$

Equations 24 and 25 do not have a closed form solution. Therefore, these may be solved to obtain the estimates of parameters of LLD using BFGS method which is available in R software.

Estimates of parameters for LLD using PWME

The estimates of α and β can be obtained as:

$$\beta_r = \frac{\alpha \Gamma\left(\frac{1}{\beta} + r + 1\right) \Gamma\left(1 - \frac{1}{\beta}\right)}{\Gamma(r + 2)}, \quad (\text{Eq.26})$$

by putting $r = 0, 1$ respectively, in Equation 26, we get

$$\beta_0 = \alpha \Gamma\left(\frac{1}{\beta} + 1\right) \Gamma\left(1 - \frac{1}{\beta}\right), \quad (\text{Eq.27})$$

$$\beta_1 = \frac{\alpha}{2} \Gamma\left(\frac{1}{\beta} + 2\right) \Gamma\left(1 - \frac{1}{\beta}\right), \quad (\text{Eq.28})$$

Their corresponding sample moments are

$$\widehat{\beta}_r = \frac{\sum_{i=1}^n \binom{i-1}{r} x_i}{n \binom{n-1}{r}}, \quad (\text{Eq.29})$$

put $r = 0, 1$ respectively, in Equation 29, we have

$$\widehat{\beta}_0 = \bar{X}, \quad (\text{Eq.30})$$

$$\widehat{\beta}_1 = \sum_{i=1}^n \frac{x_i(i-1)}{n(n-1)}. \quad (\text{Eq.31})$$

By comparing Equation 27 and Equation 28 to Equation 30 and Equation 31, estimates of α and β through PWME can be obtained.

Estimates of parameters for LLD using ME method

The r^{th} moment for LLD is:

$$E(X^r) = \int_0^{\infty} x^r f(x; \alpha, \beta) dx, \quad (\text{Eq.32})$$

$$E(X^r) = \alpha^r \Gamma\left(\frac{r}{\beta} + 1\right) \Gamma\left(1 - \frac{r}{\beta}\right), \quad (\text{Eq. 33})$$

By putting $r = 1, 2$ respectively in Equation33, the first and second moments for LLD are:

$$m_1^* = E(X) = \alpha \Gamma\left(\frac{1}{\beta} + 1\right) \Gamma\left(1 - \frac{1}{\beta}\right) = \bar{X}, \quad (\text{Eq.34})$$

$$m_2^* = E(X^2) = \alpha^2 \Gamma\left(\frac{2}{\beta} + 1\right) \Gamma\left(1 - \frac{2}{\beta}\right). \quad (\text{Eq.35})$$

By solving Equation34 and Equation35 by iterative method, the estimates of α and β through ME can be obtained.

Estimates for LLD using BE

Estimates of parameters for LLD using BE are obtained by considering the reference prior developed by Abbas and Tang (2016). The joint posterior distribution of α and β is:

$$\pi(\alpha, \beta | x) = \frac{\alpha^{-n\beta-1} \beta^{n-1} \pi_{i=1}^n x_i^{\beta-1} \pi_{i=1}^n [1 + (\frac{x_i}{\alpha})^\beta]^{-2}}{\int_0^{\infty} \int_0^{\infty} \alpha^{-n\beta-1} \beta^{n-1} \pi_{i=1}^n x_i^{\beta-1} \pi_{i=1}^n [1 + (\frac{x_i}{\alpha})^\beta]^{-2} d\alpha d\beta}. \quad (\text{Eq.36})$$

Here Metropolis within Gibbs sampling algorithm has been used to obtain the Bayesian estimates.

Results and discussion

Initial data screening

To check for the randomness and identical distribution of the observed data series at given sites, Table 2 illustrates the results of Run test and Maan-Whitney statistics (W). The corresponding p-values of observed number of runs indicates that the AMRS at given sites is random. For the calculation of W, the data series at each site has been divided into two groups. The corresponding p-values of W provide sufficient evidence that the observed data series at each site is identically distributed. Both the tests (Run test and Maan-Whitney) have been performed using 5% level of significance.

Estimation of probability distributions

The estimates of parameters of the two selected probability distributions along with p-values of Kolmogorov Smirnov (KS) test (used as a goodness of fit measure) are illustrated in Table 3. Based on the corresponding p-values, in general, all the methods of estimation have passed the goodness of fit criteria for the three sites considering both

the selected probability distributions at 5 percent level of significance. However, observing closely, the values of ME have the highest value of KS statistic so as the lowest p-values for both the probability distributions for the three sites. It indicates that the method of moments may not be a preferable choice for the estimation of parameters for the given data series at three sites. These results are consistent with the general practice of the application of extreme value probability distributions that the methods of moments is not an optimal choice for the data series having skewed distribution.

In particular, for FD the corresponding p-values of KS test are approximately similar with MLE and BE methods for Garhi Dupatta and Muzaffarabad, while PWME method have near similar tendency to MLE and BE as the skewness in the observed data series increases for the site Kotli. Similar tendencies hold for LLD (with PWME having similar tendency comparable to MLE and BE for Garhi Dupatta and Muzaffarabad); while MLE method outperforms for the site Kotli. Therefore, the results suggest that both the distributions (FD and LLD) using MLE and BE methods are suitable (with PWME method as a good alternative, if the observed data series exhibit large skewness) for modeling extremes of rainfall.

To graphically visualize the performance of both the probability distributions with four methods of estimation, the plots of probability density functions of FD and LLD have been constructed and presented in *Figure 2*. These graphs also suggest that, in general; the selected distributions are suitable for the observed data series at given sites. Particularly, for LLD, the behavior of pdf with BE method is relatively different for the sites Kotli and Muzaffarabad; while for the site Garhi Dupatta with PWME the curve seems to fit the observed data adequately and the performance of MLE, BE and ME methods is approximately similar. On the other hand, for FD, the graph using BE and MLE are comparable for all the three sites. These results are consistent with the findings of KS test.

Based on the aforementioned details both the distributions are suitable to model the extremes of rainfall with MLE or BE as preferred methods of estimation.

The quantile estimates for various return periods for each site can be obtained using the quantile function of the selected probability distribution(s). These quantiles have been calculated using MLE, BE and PWME methods (as ME is not a preferred choice of estimation). The results are illustrated in *Table 4*. The results of *Table 4* show that the estimated quantiles for $F = 0.90$ (10 year return period) and so on are greater than the average values of AMRS at given site. Moreover, the magnitude of 100 year return period is greater the maximum value of the observed AMRS (see *Table 1*) for the three sites (for both the probability distributions). These results are useful for the officials and researchers dealing with disasters and water resources management of the study for proper planning and effective management.

Table 2. Results of run test and Maan-Whitney (W) statistics

Site name	n	Runs	P-values	Groups	W	P-values
Garhi Dupatta	36	17	0.6920	18, 18	339	0.8619
Kotli	36	19	0.7510	18, 18	340	0.8371
Muzaffarabad	36	20	0.6750	18, 18	346	0.6927

n is the number of observations at each site, Runs are the observed number of runs, W is the calculated test statistic for Maan-Whitney test. These tests are performed at 5% level of significance

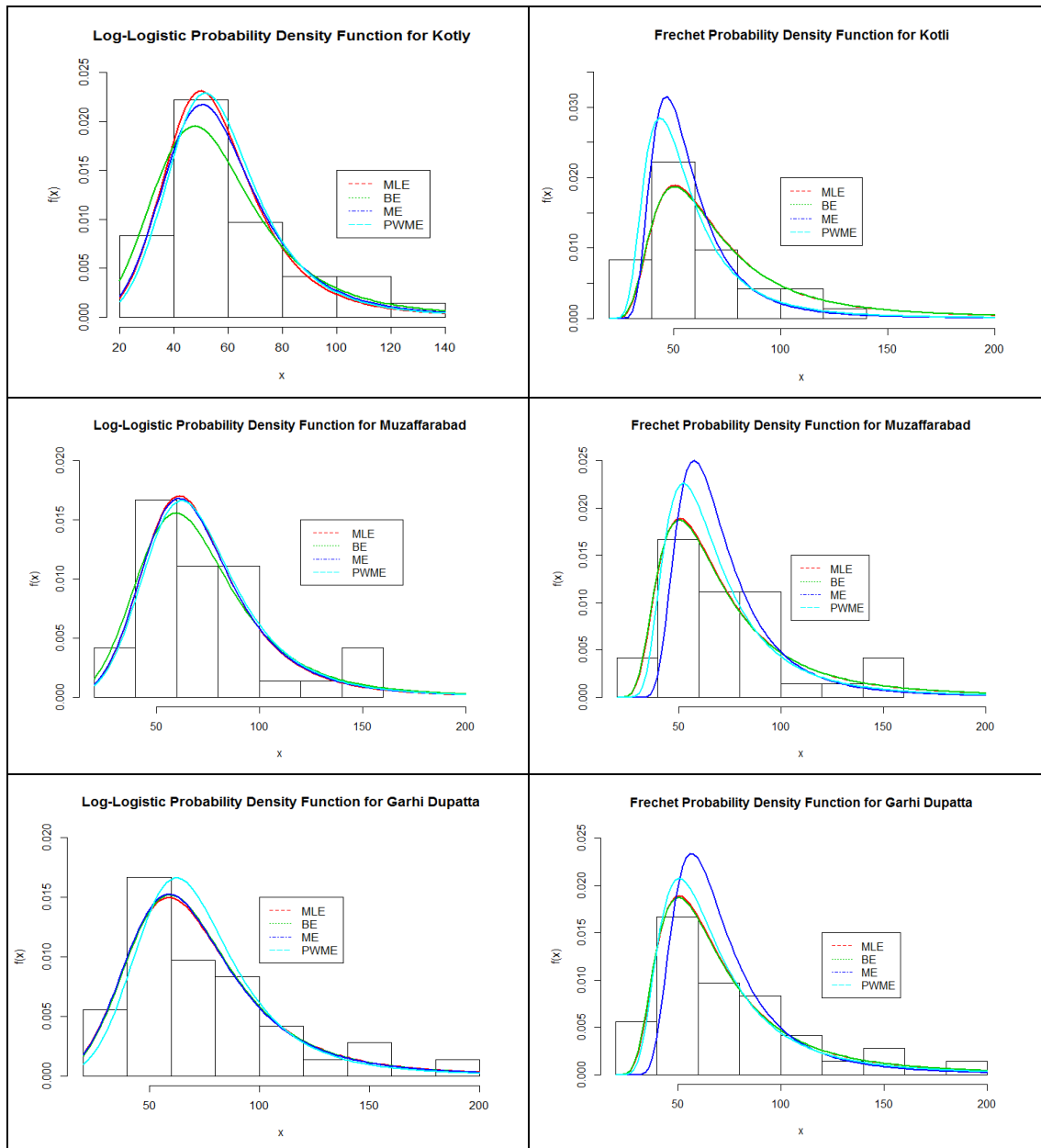


Figure 2. Plots of probability density function of Log-Logistic and Frechet distribution for three sites using four estimation methods

Table 3. Estimates of parameters and goodness of fit measure for FD and LLD

For FD				
Muzaffarabad				
Methods	A	B	KS	P-values
MLE	2.7462	56.6804	0.0971	0.8539
BE	2.7103	56.5792	0.0995	0.8336
ME	4.0172	60.8367	0.1344	0.4912
PWME	3.3454	57.4418	0.1269	0.5645

Garhi Dupatta				
MLE	2.5047	55.0170	0.1026	0.8056
BE	2.4705	54.8953	0.1014	0.8162
ME	3.7203	60.5172	0.1778	0.1817
PWME	3.0015	55.9328	0.1281	0.5523
Kotli				
MLE	3.2386	46.9789	0.0924	0.8899
BE	3.1941	46.9301	0.0929	0.8868
ME	4.1011	49.2182	0.1258	0.5749
PWME	3.5121	47.0087	0.0965	0.8587
For LLD				
Muzaffarabad				
Methods	A	B	KS	P-values
MLE	68.0016	4.3984	0.0855	0.9346
BE	67.9053	3.9635	0.0687	0.9913
ME	69.2406	4.7908	0.1183	0.6507
PWME	68.1396	4.3437	0.0852	0.9363
Garhi Dupatta				
MLE	67.9686	3.7960	0.0795	0.9631
BE	68.0521	3.8644	0.0825	0.9502
ME	69.3718	4.3719	0.11589	0.6761
PWME	67.5925	3.8496	0.0758	0.9757
Kotli				
MLE	54.3723	4.8122	0.0906	0.9036
BE	54.0521	3.9565	0.1097	0.7381
ME	55.9075	4.9094	0.1243	0.5908
PWME	55.3301	4.5832	0.1115	0.7200

MLE, BE, ME and PWME are estimates using Maximum Likelihood Estimation, Bayesian Estimation, Moments Estimation and Probability Weighted Moments Estimation methods. KS are the calculated values of Kolmogorov Smirnov test

Table 4. Quantiles estimates for various return periods for each site using selected probability distributions and three methods of estimation

For Muzaffarabad						
Distributions	Estimation methods	Quantile estimates (in millimeters) with non-exceedance probability F/return period (years)				
		0.10/(1)	0.50/(2)	0.90/(10)	0.99/(100)	0.999/(1000)
Frechet	MLE	41.83	64.77	128.62	302.63	701.09
	BE	41.59	64.77	129.79	308.87	723.54
	PWME	49.43	66.65	106.52	191.20	339.55
Log-Logistic	MLE	41.26	68.00	112.07	193.30	326.96
	BE	39.01	67.91	118.21	216.47	387.88
	PWME	41.75	69.24	109.43	114.83	339.56

For Garhi Dupatta

Distributions	Estimation methods	Quantile estimates (in millimeters) with non-exceedance probability F/return period (years)				
		0.10/(1)	0.50/(2)	0.90/(10)	0.99/(100)	0.999/(1000)
Frechet	MLE	39.44	63.69	135.11	345.24	867.28
	BE	39.17	63.67	136.50	353.35	899.03
	PWME	42.36	63.20	118.38	258.98	558.59
Log-Logistic	MLE	38.10	67.97	121.25	228.05	419.27
	BE	38.54	68.05	120.16	223.49	406.49
	PWME	38.20	67.59	119.61	223.00	406.53

For Kotli

Distributions	Estimation methods	Quantile estimates (in millimeters) with non-exceedance probability F/return period (years)				
		0.10/(1)	0.50/(2)	0.90/(10)	0.99/(100)	0.999/(1000)
Frechet	MLE	36.31	52.61	94.11	194.43	396.43
	BE	36.15	52.64	94.93	198.11	407.96
	PWME	37.07	52.18	89.22	174.19	335.99
Log-Logistic	MLE	34.44	54.37	85.83	141.28	228.40
	BE	31.01	54.05	94.19	172.67	309.71
	PWME	34.26	55.33	89.37	150.80	249.71

Comparison with previously published study

Though the illustrated results of this study indicates that the selected probability distributions are suitable to model the rainfall data at given sites so as the estimated quantiles are useful but the precedence through practical comparison of the adopted methodology (as an acceptable alternative to the conventional methods) is still requisite. To do so, the quantile estimates for 1, 10 and 100 year return period of this study have been compared with the quantiles provided by Shahzadi et al. (2013). Few details of the comparison are: the study of Shahzadi et al. (2013) has been selected as it is the only study with similar variable for analysis, i.e. annual maximum rainfall series and the same set of sites. The sites Muzaffarabad, Garhi Dupatta and Kotli were included in Region 2 by Shahzadi et al. (2013) and Generalized Extreme Value, Generalized Normal and Generalized Logistic distributions identified as regional distributions (for more details and regional quantiles see *Table 5* of Shahzadi et al., 2013). Those estimated regional quantiles are utilized to estimate the quantiles for the specific sites (for comparison) using index flood method, i.e. those estimates of regional quantiles (for 1, 10 and 100 year return period) are multiplied by the mean of the respective sites provided in *Table 1* of this study. The results are illustrated in *Table 5*. The comparison (of *Tables 4* and *5*) reveals that, for all the three sites, the estimates of FD and LLD distributions with BE, MLE and PWME are quite close to the results of L-moments based regional frequency analysis with Generalized Extreme Value, Generalized Normal and Generalized Logistic distributions for shorter return period, i.e. 1 and 10 year. Moreover, for longer return period, i.e. 100 year, LD distribution with BE method and FD with PWME method provides comparative results to the results of Shahzadi et al. (2013), hence can be a preferred choice to model the extremes of rainfall.

Table 5. Quantiles for 1, 10 and 100 year return period of Shahzadi et al. (2013)

For Muzaffarabad				
Distributions	Estimation methods	Return period (years)		
		1	10	100
Generalized extreme value	L-moments based regional frequency analysis	40.66	117.05	206.55
Generalized normal	L-moments based regional frequency analysis	40.51	118.47	201.94
Generalized logistic	L-moments based regional frequency analysis	40.51	201.94	213.40
For Garhi Dupatta				
Distributions	Estimation methods	Return period (years)		
		1	10	100
Generalized extreme value	L-moments based regional frequency analysis	41.34	119.03	210.05
Generalized normal	L-moments based regional frequency analysis	41.19	120.47	205.35
Generalized logistic	L-moments based regional frequency analysis	41.19	116.38	217.01
For Kotli				
Distributions	Estimation methods	Return period (years)		
		1	10	100
Generalized extreme value	L-moments based regional frequency analysis	32.71	94.18	166.19
Generalized normal	L-moments based regional frequency analysis	32.59	95.32	162.48
Generalized logistic	L-moments based regional frequency analysis	32.59	92.08	171.70

Conclusions

The study illustrates application of two atypical probability distributions with four methods of estimation of parameters to model the behavior of AMRS at three sites of AJ&K, Pakistan. Some of the major conclusions are:

- i. The observed data series at given sites is random and identically distributed as validated by Run test and Man Whitney test.
- ii. Summary statistics have shown that there exists variations in the observed data series and the distribution related to the observed data series at given sites is positively skewed.
- iii. In general, both the distributions with suggested methods for the estimation of parameters have passed the goodness of fit criteria of KS test. However, the tendency of the KS values along with corresponding p-values have suggested that ME method is no a preferred choice for estimation.
- iv. The graphs of probability density functions for FD and LLD have also illustrated that both the probability distributions are suitable for the observed data series at given sites. Moreover, the trends and tendencies of probability density functions are dissimilar using various methods of estimation. However, the results are in agreement with the results of KS test.
- v. The study suggest that both the probability distributions using MLE and BE methods are preferred choice for modeling the extremes of rainfall of the study area. These results are encouraging to adopt relatively atypical probability distributions for the estimation of extreme events, like floods, rainfall, etc. This will bring diversity and flexibility in the choice of selection of a probability distribution for the modeling of extremes of geophysical events.
- vi. The quantiles for various return periods have been calculated using the quantiles functions of the selected probability distributions. The estimated

quantiles for 10 year return period and above are greater than the average values of the observed data series at given sites. Moreover, the estimates of 100 year return period are greater than the maximum values of the observed data series for all the three sites. Therefore, the concerned officials have to adopt preemptive measures in advance to deal with these extreme events of such magnitude.

- vii. To establish the usefulness of the selected distributions along with the adopted methods of estimation, the quantiles for 1, 10 and 100 year return period are compared with the previously published studies related to this area of study and variable of analysis. The comparison reveals that LD distribution with BE method and FD distribution with PWME method provides comparative results relative to L-moments based rainfall frequency analysis using popular three-parameter distributions, especially for longer return period, i.e. 100 year.

The study provides useful guidelines for the concerned officials of study area, home-grown farmers, meteorologists studying precipitation and rainwater management planning. Moreover, the study provides illustration of some new probability distributions with different methods of parameter estimation for at site frequency analysis of extreme events. The procedure can be adopted in general to bring diversity and flexibility in the choice of probability distributions for modeling of extremes of events like floods, rainfall, etc.

Acknowledgements. We are thankful to the editorial office of the journal and anonymous reviewer(s) for their constructive comments and suggestions to improve the quality of the paper. We are also grateful to Pakistan Meteorological Department (CDPC, Head Office Karachi) for providing the rainfall data for the study.

REFERENCES

- [1] Abbas, K., Tang, Y. (2015): Analysis of Frechet distribution using reference priors. – Communications in Statistics - Theory and Methods 44: 14: 2945-2956.
- [2] Abbas, K., Tang, Y. (2016): Objective Bayesian analysis for log-logistic distribution. – Communications in Statistics-Simulation and Computation 45(8): 2782-2791.
- [3] Ahmad, I., Abbas, A., Saghir, A., Fawad, M. (2016): Finding probability distributions for annual daily maximum rainfall in Pakistan using linear moments and variants. – Polish Journal of Environmental Studies 25(3): 925-937.
- [4] Ahmad, I., Abbas, A., Fawad, M. (2017): A study on regional frequency analysis of annual total rainfall in Pakistan using method of linear moments. – NUST Journal of Engineering Sciences 10(1). <http://journals.nust.edu.pk/index.php/njes/article/view/157>.
- [5] Alam, M. A., Emura, K., Farnham, C., Yuan, J. (2018): Best-fit probability distributions and return periods for maximum monthly rainfall in Bangladesh. – Climate 6(9). DOI: 10.3390/cli6010009.
- [6] Amin, M. T., Rizwan, M., Alazba, A. A. (2016): A best-fit probability distribution for the estimation of rainfall in northern regions of Pakistan. – Open Life Sciences 11(1): 432-440.
- [7] Bahrawi, J. A. (2018): Rainfall distribution and its characteristics in Makkah Al-Mukarramah Region, Saudi Arabia. – Applied Ecology and Environmental Research 16(4): 4129-4144.

- [8] Bradley, J. V. (1968): *Distribution-Free Statistical Tests*. – Prentice Hall, Upper Saddle River, NJ.
- [9] Frechet, M. (1927): Sur la loi de probabilité de l'écart maximum. – *Annales de la Société Polonaise de Mathématique* 6: 93-116.
- [10] Greenwood, J. A., Landwehr, J. M., Matalas, N. C., Wallis, J. R. (1979): Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. – *Water Resour Res* 15: 1049-1054.
- [11] Hosking, J. R., Wallis, J. R., Wood, E. F. (1985): Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. – *Technometrics* 27(3): 251-261.
- [12] Hussain, I., Spöck, G., Pilz, J., Yu, H.-L. (2010): Spatio-temporal interpolation of precipitation during monsoon periods in Pakistan. – *Advances in Water Resources* 33(8): 880-886.
- [13] Hussain, Z., Shahzad, M. N., Abbas, K. (2017): Application of regional rainfall frequency analysis on seven sites of Sindh, Pakistan. – *KSCE Journal of Civil Engineering* 21(5): 1812-1819.
- [14] Iqbal, M. J., Ali, M. (2013): A probabilistic approach for estimating return period of extreme annual rainfall in different cities of Punjab. – *Arabian Journal of Geosciences* 6(7): 2599-2606.
- [15] Khan, S. A., Hussain, I., Hussain, T., Faisal, M., Muhammad, Y. S., Mohamd Shoukry, A. (2017): Regional frequency analysis of extremes precipitation using L-moments and partial L-moments. – *Advances in Meteorology*. DOI: 10.1155/2017/6954902.
- [16] Liang, Y., Liu, S., Guo, Y., Hua, H. (2017): L-moment-based regional frequency analysis of annual extreme precipitation and its uncertainty analysis. – *Water Resources Management* 31(12): 3899-3919.
- [17] Mann, H. B., Whitney, D. R. (1947): On a test of whether one of two random variables is stochastically larger than the other. – *Annals of Mathematical Statistics* 18(1): 50-60.
- [18] Mo, C., Ruan, Y., He, J., Jin, J., Liu, P., Sun, G. (2019): Frequency analysis of precipitation extremes under climate change. – *International Journal of Climatology* 39(3): 1373-1387.
- [19] Mohamed, M. T., Ibrahim, A. A. (2016): Fitting probability distributions of annual rainfall in Sudan. – *SUST J. Eng. Comput. Sci.* 17(2): 34-39.
- [20] Shahzadi, A. (2013): A review: regional frequency analysis of annual maximum rainfall in monsoon region of Pakistan using L-moments. – *International Journal of Advanced Statistics and Probability* 1(3): 97-101.
- [21] Shahzadi, A., Akhter, A. S., Saf, B. (2013): Regional frequency analysis of annual maximum rainfall in monsoon region of Pakistan using L-moments. – *Pakistan Journal of Statistics and Operation Research* 9(1): 111-136.
- [22] Wang, Z., Zeng, Z., Lai, C., Lin, W., Wu, X., Chen, X. (2017): A regional frequency analysis of precipitation extremes in Mainland China with fuzzy c-means and L-moments approaches. – *International Journal of Climatology* DOI: 10.1002/joc.5013.
- [23] Zalina, M. D., Desa, M. N. M., Nguyen, V. T. A., Kassim, A. H. M. (2002): Selecting a probability distribution for extreme rainfall series in Malaysia. – *Water Science and Technology* 45(2): 63-68.