Original Research

Regional Frequency Analysis of Annual Peak Flows in Pakistan Using Linear Combination of Order Statistics

Ishfaq Ahmad[1]*, Muhammad Fawad[1], Muhammad Akbar[1], Aamar Abbas[2], Hafiz Zafar[3]

[1]Department of Mathematics and Stats, International Islamic University, Islamabad, Pakistan
[2]Department of Mathematics, The University of Poonch, AJK, Rawalakot, Pakistan
[3]Department of Statistics, University of Sargodha Pakistan

Abstract

For this paper we conducted a regional analysis (RA) of annual peak flows using linear combination of order statistics, i.e., linear-moments (LM) and trimmed linear moments (TLM). Design flood estimates are calculated and compared at different return periods, which are useful for water resources management, including hydrological structures and basin management. The main objective of our study was to compare regional design flood estimates for untrimmed and trimmed samples. LM is the special case of TLM, when we have no trimming from either side. First, regional flood frequency analysis is performed for LM and then for TLM. After initial screening of the annual peak flow series, a discordancy measure was used to diagnose the discordant sites. No site was found to be discordant. For homogeneity of the region, the homogeneity measure "H" was employed using simulation study based on Kappa distribution, and found that the nine sites on the Indus Basin included in the study constitute a single homogeneous region. In this study we used TLM with trimming values $(\gamma, 0)$, where $\gamma = 1, 2, 3, 4$. In order to determine the most appropriate probability distribution for regional quantile estimates, different probability distributions are used, namely: generalized extreme value (GEV), generalized pareto (GPA), generalized logistic (GLO), Pearson type three (PE3), and generalized normal (GNO). L-moments ratio diagram and Z-test as goodness of fit are engaged to identify the most suitable probability distribution. A comparison revealed that GNO is the best distribution for first three cases as $(0, 0)$, $(1, 0)$, and $(2, 0)$, while for the last two cases of $(3, 0)$ and $(4, 0)$ the most appropriate choice is GEV. A simulation study was also carried out to evaluate the performance and robustness of the best fit probability distribution using relative bias (RB) and relative root mean square error (RRMSE).

*e-mail: ishfaq.ahmad@iiu.edu.pk

Introduction

Pakistan, being a developing country, has long faced the problem of flooding. Floods are natural disasters that cause tremendous destruction in the form of human lives, economy, health, infrastructure, industry, and agriculture. Estimating the frequency and magnitude of flooding is of great importance in order to make policies and measure disaster prevention and mitigation. Punjab Province in Pakistan faces more severe floods than other provinces. Heavy rainfall in the monsoon season burdens the water level in the basin and results in floods. Floods cannot be eliminated completely but several preventive measures could be adopted to mitigate losses due to these floods. Among others, some preventive measures include reliable knowledge on quantity and frequency of the discharge flows. In view of such disastrous floods, it is of dire need to estimate the magnitude and frequency of these peak flows for a given return period with high accuracy for future hydrological projects such as successful operation and economical design of different hydrologic structures such as water reservoirs, dams, bridges, small culverts, spillways, urban drainage systems, etc.

The exclusive objective of any on-site or regional flood frequency analysis using index flood procedure is to investigate reliable quantile estimates (design flood estimates) for each site [1]. The selection of the suitable probability model for flood frequency analysis (FFA) has received extensive attention [2]. In Pakistan, data of annual block maxima/peak flows series (one value with maximum flow in the whole year) are often of shorter length and sometimes unavailable at some (ungauged) stations, which put a hurdle in front of hydrologists when evaluating reliable flood design estimates at a particular location. In such a scenario, on-site FFA is of limited use and, ultimately, regional analysis of peak flows is more appropriate and applicable for providing the information on design flood estimates that are useful for water resource management [3]. During RA, information from different sites is combined and this prudent way of pooling information is known as the "index flood method."

In environmental sciences, use of the same kind of data at many different sites by forming homogeneous regions of these sites, keeping in view some common hydrological and geographical characteristics and to estimate the suitable frequency distribution of these homogenous regions, is known as regional frequency analysis [4]. A selection of different methods for estimating parameters for reliable estimates always remained a matter of great concern. In this study we compared the design flood estimates using ML and TLM. The method of trimmed L-moments is useful when we have small samples and are interested in design flood estimates of larger time periods. We will implement the same methodology as proposed by [4] with the addition of trimmed L-moments.

Regional analysis based on L-moments has been applied effectively in different countries across the world: Parida et al. analyzed regional flood frequency analysis in India of subzones of Zone 3 based on L-moments [5], Lim and Lye in Malaysia evaluated regional flood estimation for ungauged basins in Sarawak [6], Hussain and Pasha in Pakistan carried out regional flood frequency analysis of seven sites of Punjab using L-moments [1], Shahzadi (also in Pakistan) carried out regional analysis on annual maximum daily rainfall series across different stations [7], Hassan and Ping [8] implemented the same methodology for flood data in China's Luanhe Basin with the help of cluster analysis for regions, Saf [9] evaluated regional studies for rainfall data in Turkey using L-moments, and the same kind of analysis was carried out in the UK by Fowler and Kilsby [10]. And in Italy Noto and La Logia [11] and Cannarozo et al. [12] and in Canada Glaves and Waylen [13], Yue and Wang [14], and Abida and Ellouze [15] all carried out regional FFA using L-moments.

When our concern is outliers in the data, we can use a generalized form of LM, i.e., TLM developed by Elamir and Seheult [16], by assigning zero weight to extreme observations. TLM moments are more robust as compared to LM. TLM show certain advantages over LM for example TLM exist even the mean of probability distribution does not exist such as Cauchy distribution [16]. Some studies on TLM with application in hydrology and meteorology are available in the literature, for example [17-22]. In the present study, regional flood frequency analysis using LM and TLM are presented and compared using different values of trimming $(\gamma, 0)$, where $\gamma = 1, 2, 3, 4$. LM is a special case of TLM, when we make zero trimming from both sides. In this study we considered different probability distributions such as GLO, GEV, GPA, PE3, and GNO, which are short-listed based on previous studies such as [20-22], in which these distributions are more acceptable for tropical regions for modeling precipitation extremes [23-24].

Materials and Methods

Theory of L-Moments and TL-Moments

L-moments are summary statistics for probability distributions as well as for data samples. They are tantamount to conventional moments and are considered superior to them. They also provide measures of location, dispersion, skewness, kurtosis, and other aspects of the shape of probability distributions and data samples. In their computation we use linear combinations of the ordered data values. They are more robust in small and moderate samples and are superior to methods of moments and maximum likelihood method [25-27]. Hosking [27] defined L-moments as a linear function of probability weighted moments earlier determined by Greenwood et al. [28] as:

$$\lambda_{r+1} = \sum_{k=0}^{r} p_{r,k}^{\star} \beta_k \qquad (1)$$

$$where ; \sum_{k=0}^{r} p_{r,k}^{\star} = \sum_{k=0}^{r} (-1)^{r-k} \binom{r}{k} \binom{r+k}{k} \beta_k$$

The first four L-moments and probability weighted moments are related as follows:

$\lambda_1 = \beta_0; \quad \lambda_2 = 2\beta_1 - \beta_0; \quad \lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0;$
$\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$

… where λ_1 is measure of location, λ_2 is measure of scale,

$\tau = \dfrac{\lambda_2}{\lambda_1}$ is $L_{c.v.}$ (coefficient of variation based on LM)

$$For \ r = 3, \quad \tau_3 = \frac{\lambda_3}{\lambda_2} \ is \ L - skewness \tag{2}$$

$$For \ r = 4, \quad \tau_4 = \frac{\lambda_4}{\lambda_2} \ is \ L - kurtosis \tag{3}$$

In practice, L-moments are estimated from sample observations. It would be appropriate to use an unbiased estimator b_r of PWMs i.e. β_r as:

$$b_r = n^{-1} \sum_{j=r+1}^{n} \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} x_{j:n}$$

The relationship between sample L-moments and PWMs is given by

$l_1 = b_0; \quad l_2 = 2b_1 - b_0; \quad l_3 = 6b_2 - 6b_1 + b_0; \quad l_4 = 20b_3 - 30b_2 + 12b_1 - b_0$

If our concern is extreme events inserting undue influence, a more robust estimation method could be used [16]. The sample trimmed linear moments are unbiased estimators to population TLM and relatively more robust to outliers. In TLM, the expectations of the order statistics are substituted by expectations of the order statistics of a higher size. The size is augmented equal to the total amount of trimming.

The r_{th}, TL-moments ($t_1 = \gamma$, $t_2 = 0$, *wher*, $\gamma = 0$, 1, 2, 3, 4) could be written as:

$$\lambda_r^{(\gamma,0)} = \frac{1}{r} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E\left(X_{r+\gamma-k:r+\gamma}\right) \qquad r = 1,2,\dots$$

The population measure of location, scale, skewness and kurtosis can be defined as in case of LM. In this study we trim from the lower side only, and such TLMs are also known as LHMs (higher order linear moments). The TLMs are useful for illustrating the upper part of the distributions and larger events in the data samples. The LMs are oversensitive to the lower part of distributions and give inadequate weight to larger values in the data having meaningful information about the upper tail of the distribution. The TLMs (trimming from lower side only) alleviate the unwanted disturbances due to small samples during estimation of larger events. A comprehensive discussion on LM and TLM may be found in [16 and 29]. The PWMs and TLM at different trimming are related as:

$\lambda_{r+1}^{(\gamma,0)} = \sum_{k=0}^{r} p_{r,k}^{\star} \beta_k^{(\gamma,0)}$, where $\gamma = 1,2,3,4$
$\lambda_1(\gamma,0) = \beta_0(\gamma, 0)$
$\lambda_2(\gamma,0) = 2\beta_1(\gamma,0) - \beta_0(\gamma,0)$
$\lambda_3(\gamma,0) = 6\beta_2(\gamma,0) - 6\beta_1(\gamma,0) + \beta_0(\gamma,0)$

$\lambda_4(\gamma,0) = 20\beta_3(\gamma,0) - 30\beta_2(\gamma,0) + 12\beta_1(\gamma,0) - \beta_0(\gamma,0)$

The corresponding CV based on TLM and TLM ratios for skewness, kurtosis, and further sample quantities could be defined as in the case of LM.

Regional frequency analysis based on L-Moments

Regional frequency analysis methodology based on L-moments as reported by [4] will be implemented in this study for Pakistan peak flow data. In addition, we also carried out this study for trimmed L-moments. The following four steps for regional frequency analysis are:

1) ***Screening the data for regional analysis:*** To discard inconsistent sites, D_i (discordancy measure), D_i is defined as:

$$D_i = \frac{1}{3} \ N \ (u_i - \bar{u})^T C^{-1} (u_i - \bar{u})$$

…where u_i is the vector of sample LMs ratios and C is a variance-covariance matrix. To screen out site i from the data, if the value of statistic (D_i) for site i exceeds the critical value, we would discard site i from analysis. It could be useful as it provided initial guidance to formulate homogeneous regions.

2) ***Identification of homogeneous regions***: The most prudent step in RFFA is formulating homogeneous regions. A region is considered to be homogeneous if all sites included in the region have some common characteristics. There are different grouping methods available in literature used for this purpose, e.g., geographical convenience, subjective partitioning, objective partitioning, and cluster analysis. In geographical convenience usually regions constitute contiguous sites, with the help of administrative areas or main physical groupings [4]. In subjective partitioning, regions may be defined subjectively on the basis of site characteristics (mean annual precipitation, latitude, longitude, drainage area, and time of occurrence of largest flood in the year, etc.). Although these methods are subjective in nature and once the regions formed on the basis of these approaches, they must be tested for homogeneity using some standard criteria such as heterogeneity measure H. In objective partitioning regions are formed on the basis of some threshold value of site characteristics of each site, and sites are allocated to one of two group using this threshold value. Cluster analysis is considered another multivariate technique used to form homogenous regions in RFFA. Sites are divided into different groups (also called clusters) on the basis of different on-site characteristics, site characteristics, or a combination of both. The number of clusters is normally a subjective choice. Cluster analysis is an effective method if we have a large number of sites in each cluster while the range of drainage area of these sites in homogenous regions should not be too large, as reported in [30-36]. The tests to check the homogeneity of the regions suggested by [4], are based on L-moments and Kappa

distribution in its simulation study for the reason that Kappa distribution is less restrictive than Gumbel or GEV. Initially, by considering one region of all sites included in the plan, it is suggested to calculate the heterogeneity measure H as:

$$H = \frac{V - \mu_v}{\sigma_v},$$

…where V = weighted standard deviation of sample l_{cv} based on all N sites. Mathematically,

$$V = \left\{ \sum_{i=1}^{N} n_i \ (t^i - t^R)^2 / \sum_{i=1}^{N} n_i \right\}^{\frac{1}{2}} :$$

…where t^R= Regional Average *of sample* l_{cv} defined as:

$$t^R = \frac{\sum_{i=1}^{N} n_i \ t^{(i)}}{\sum_{i=1}^{N} n_i}$$

For simulation study we find mean and standard deviation from a large number N_{sim} of realizations of a region with "N" sites, each having Kappa distribution. In other words, we calculated μ_v and σ_v given in the above expression by fitting Kappa distribution using average L-moment ratios (L-CV is more preferable as compared to L-skewness and L-kurtosis, as it captures more variation, although these two measures can also be used). If H statistic shows a sufficient large value, state the region to be heterogeneous, i.e., if H < 1 distribution in the region is possibly homogeneous; if $1 \leq$ H < 2, the region is possibly heterogeneous; and if H $\geq$ 2, the region is definitely heterogeneous.

3) ***Choice of frequency distribution for homogenous regions:*** After formulating homogeneous regions, the next step is to choose the most robust frequency distribution for each homogenous region. Hosking and Wallis [4] argued that such frequency distribution should be capable of giving reasonable quantile estimates even though future data values come from some different ones than the fitted distribution. Hence the objective should not be to find a distribution, which gave only best fit for a particular homogeneous region, but also a robust distribution with good quantile estimates from which future data values would arise. A distribution with large parameters (3 to 5) is considered reasonable as it gave less bias and is appropriate for estimates of the quantiles in the tails of distribution. It is more reliable to use such distribution in RFFA as compared to on-site frequency analysis. To use Z^{Dist}, through regional L-moment statistics let σ_4 be the standard deviation of t_4^R, which can be obtained by repeated simulation of a homogeneous region whose sites supposedly had a particular frequency distribution. Initially it is defined as:

$$Z^{Dist} = (\tau_4^{Dist} - t_4^R)/\sigma_4 \qquad (4)$$

The smaller the value of Z^{Dist}, the better the indication of good fit, meaning that the true distribution is the same as the distribution we are fitting to the data, where:

$$\sigma_4 = \left[(N_{sim})^{-1} \{ \sum_{m=1}^{Nsim} (t_4^{(m)} - t_4^R)^2 \} \right]^{1/2}, \text{ and } t_4^{(m)},$$

is the regional average L-kurtosis of m times simulated region. The fit is considered to be good if $|Z^{Dist}|$ had a smaller value sufficiently close to zero or a reasonable criteria might be adopted as if $|Z^{Dist}| \leq 1.64$, i.e., the fit is declared to be adequate.

4) ***Estimation of Frequency Distribution and accuracy of estimated quantile:*** The final step of RFA is the estimation of parameters for selected distribution and further to determine its robustness in producing reasonably reliable quantile estimates for all stations in the homogenous region. We use index flood procedure that used the summary statistic of the data at each site and combined them to form regional estimates. The resulting procedure as suggested by [4] is known as the regional L-moments algorithm, which is more robust even if some of the basic assumptions of the index flood method could not be fulfilled. The advantage of this algorithm is also that it took into account the possibility of heterogeneity in the region, error in specification of correct frequency distribution, and statistical dependence of the data at different sites. In the simulation procedure, quantile estimates are calculated for various non-exceedance probabilities. The quantile function of the fitted regional frequency distribution could be denoted as Q (.). Furthermore, the quantile estimates "design flood estimates" at each site could be obtained by combining the estimates of μ_i and Q (.) with nonexceedance probability F using the relationship as:

$$\hat{q}_i(F) = l_1^{(i)} \hat{Q}(F) \qquad (5)$$

The robustness of the designated regional frequency is further explored with estimation of design flood quantiles. To achieve this objective, Meshgi and Khalili [29] suggested a Monte Carlo simulation to calculate errors between simulated quantiles and calculated regional quantile estimates. For this, relative bias (RB) and relative root mean square error (RRMSE) are evaluated as:

$$RB(F) = M^{-1} \sum_{m=1}^{M} \frac{\{ \hat{Q}_i \, [\, m](F) - Q_i(F) \}}{Q_i(F)} \qquad (6)$$

$$RRMSE(F) = \sqrt{M^{-1} \sum_{m=1}^{M} \left(\frac{\{ \hat{Q}_i \, [\, m](F) - Q_i(F) \}}{Q_i(F)} \right)^2} \qquad (7)$$

…where M is the sample size, and $\hat{Q}_i \, [m](F)$ and $Q_i(F)$ *are* simulated and calculated regional quantiles.

Table 1. Basic Information about all of the sites used in the study.

Sites	Trim. level	l_1	l_2	t	t_3	t_4	Latitude	Longitude	Annual mean peak flows	Elevation
Mangla ($n = 54$)	(0,0)	132,481.78	54,023.67	0.408	0.449	0.365	33.15	73.65	132,480	14,902.7
Rasul ($n = 44$)	(0,0)	34,418.39	66,402.58	0.494	0.489	0.351	32.68	73.50	134,420	8,828.66
Marala ($n = 54$)	(0,0)	308,572.41	106,247.39	0.344	0.295	0.103	32.68	74.43	308,570	9,358.38
Khanki ($n = 89$)	0,0)	351,963.19	123,768.11	0.352	0.368	0.185	32.40	73.92	361,160	7,839.85
Qadirab ad ($n = 44$)	(0,0)	356,547.70	135,642.60	0.380	0.269	0.095	32.33	73.73	356,550	7,522.02
Trimmu ($n = 46$)	(0,0)	261,376.22	104,959.78	0.402	0.294	0.114	31.14	72.15	261,380	5,297.19
Panjnad ($n = 54$)	(0,0)	260,134.72	106,855.13	0.411	0.215	0.120	29.33	71.00	260,130	4,838.10
Balloki (n=54)	(0,0)	87,914.728	30,921.621	0.352	0.390	0.245	31.22	73.86	91,272.0	6,497.89
Sidhani ($n = 89$)	(0,0)	64,143.427	27,038.629	0.422	0.405	0.255	30.58	72.07	70,600.0	4,944.05

RFFA using Trimmed L-Moments (LH-Moments)

The same procedure would be repeated for each level of trimming, i.e., (1, 0), (2, 0), (3, 0), and (4, 0). All measures for skewness, kurtosis, and CV would be replaced by the corresponding measures based on TLM for discordancy and the homogeneity test. Selection of an adequately fitted distribution would be carried out using the TLM-based Z-test. The distribution with absolute Z-test value less than 1.64 would be declared as the possible candidate distribution, and furthermore – in the case of more than one distribution qualifying this criteria – the distribution would be best with the smallest value. Estimation of the design floods for a specific return period would be calculated for each trim level.

Results and Discussion

Study Area and Data

This study investigated the RFFA using the peak flow series of nine sites at four tributaries of the Indus Basin in

Fig. 1. Selected sites of the Indus Basin.

Table 2. Results of different tests for basic assumptions.

Sites	Mann-Whitney U test		Ljung-Box Q Statistics test	
	Statistic-Value	P-value	Statistic-Value	P-value
Mangla	209.500	0.890	10.926	0.814
Rasul	196.2	0.56	9.541	0.56
Marala	99.500	0.125	5.663	0.773
Khanki	320.500	13.091	0.519	0.62
Qadirabad	318.000	26.653	0.056	0.35
Trimmu	334.500	23.149	0.058	0.25
Panjnad	212.3	13.606	0.754	0.64
Balloki	142.8	19.252	0.376	0.41
Sidhani	102.3	18.055	0.800	0.26

Pakistan. The sites for the proposed study are selected at the Indus Basin using standard criteria of area, record length, quality, urbanization, regulation, climate variability, and change. These sites are located at its tributaries, namely the Jhelum, Chenab, Ravi, and Sutlej rivers. The data are retrieved from the Federal Flood Commission, the Water and Power Development Authority, the Indus River System Authority, and the Provincial Irrigation Departments. Record lengths range from 44 to 92 years. All sites are located at the Indus Basin with a latitude ranging from 29.33 to 33.15 and longitude ranging from to 71 to 74.43 (Table 1 and Fig. 1).

Before conducting RFFA, the basic assumptions of stationarity, homogeneity, and independence of peak flows series are tested. The Ljung test is employed for stationarity and independence, while the Mann-Whitney test is used for homogeneity and consistency. All tests indicated that the series are suitable for RFFA. (Table 2).

In RFFA, the first step is screening out the discordant sites using discordancy measure D_i. The discordancy statistic D_i is computed for individual sites using different levels of trim. The calculated values for all sites are less than the critical value of 3. The same methodology has been adopted by many studies around the world [30-32, 34-36]. In the present study we declared no one site to be discordant (Table 3).

Although the maximum calculated value for discordancy statistic D_i is 2.15 at the Panjnad site using LM, it is also less than the critical value and indicates that all sites do not reflect any outlier and discordancy and could be considered for further RFFA at all trimming levels. The second step is the formulation of homogeneous regions. By homogeneous regions we meant to group the sites into mutually exclusive regions on the basis of their common characteristics, which might be geographical, hydrological, or other characteristics. In the present study we initially assume a single region on the basis of geographical convenience and subjective partining. All of these sites lie in the middle of the Indus Basin. Seven of these nine sites have already been used in a previous study [1], where these sites formed a single homogenous region. The two extra sites named Panjnad and Trimmu are also in the vicinity of the remaining seven sites. Although it is tedious work to formulate a homogeneous region, a useful measure "*H*" is based on the degree of heterogeneity used to identify such homogeneous regions. For this purpose regional average ratios of LM and TLM are calculated and the results of various heterogeneity statistic *H*, based on 500 simulations using four parameters of Kappa distribution, are given in Table 3.

Heterogeneity measure "H_1" (preferable because it captures more variations using sample L.CV) at all levels of trim) is calculated. Results showed that the said region is homogenous, consisting of all nine sites, as the calculated value "H_1" is less than 1 (critical value). There is no need

Table 3. Discordancy measure D, homogeneity measure H, and goodness of fit test.

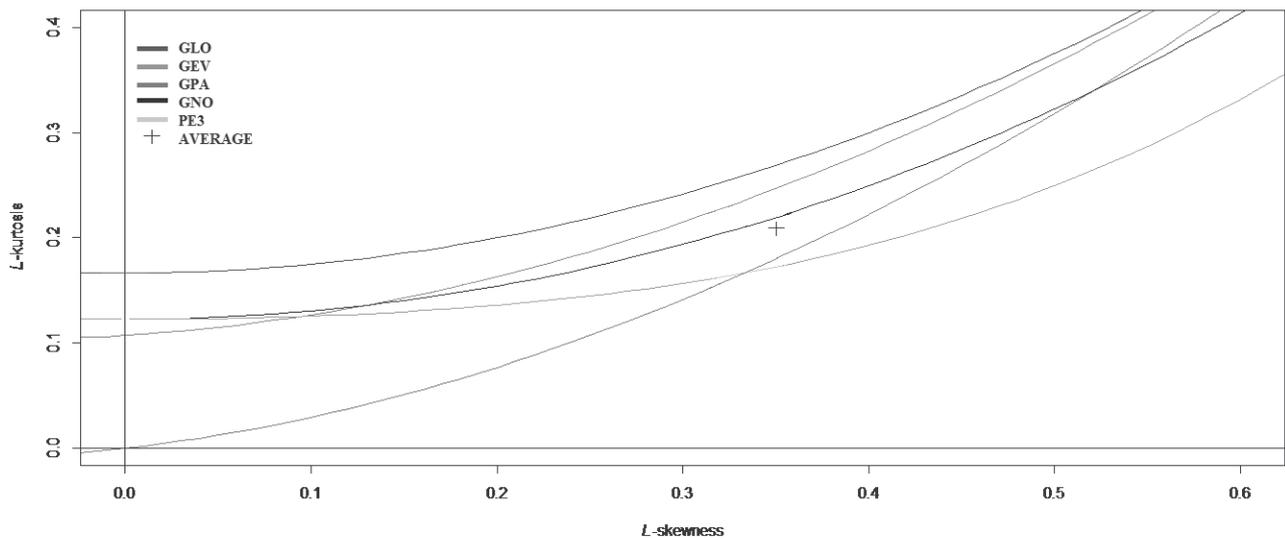| Sr. no	Site Name	$D_i^{(0,0)}$	$D_i^{(1,0)}$	$D_i^{(2,0)}$	$D_i^{(3,0)}$	$D_i^{(4,0)}$	Heterogeneity Statistics with Trimming (0, 0)	Heterogeneity Statistics with Trimming (1, 0)	Heterogeneity Statistics with Trimming (2, 0)	Heterogeneity Statistics with Trimming (3, 0)	Heterogeneity Statistics with Trimming (4, 0)
1	Mangla	1.59	1.83	2.02	2.14	2.17	H_1= 0.94	H_1= 0.78	H_1= 0.67	H_1= 0.51	H_1= 0.43
2	Rasul	2.00	1.96	1.94	1.93	1.92	H_2= 1.04	H_2= 0.94	H_2= 0.88	H_2= 0.79	H_2= 0.68
3	Marala	0.65	0.50	0.43	0.42	0.45	H_3= 1.25	H_3= 0.99	H_3= 0.83	H_3= 0.81	H_3= 0.68
4	Khanki	0.62	0.70	0.75	0.77	0.79					
5	Qadirabad	0.43	0.41	0.42	0.46	0.51					
6	Trimmu	0.56	0.59	0.61	0.60	0.61					
7	Panjnad	2.15	1.89	1.56	1.22	0.96					
8	Balloki	0.79	0.81	0.88	0.96	1.01					
9	Sidhani	0.20	0.30	0.40	0.51	0.58					

Fig. 2. Ratio diagram based on L-moments.

to make more regions. All heterogeneity measures are sufficient evidence to declare the region as homogenous. After successful formulation of a homogeneous region the next step is the selection of robust and suitable statistical distribution for the homogenous region. Suitable distribution does not mean only the distribution provided as a best fit to the selected region but also the distribution with accurate quantile estimates for different return periods. The ratio diagram for graphical inspection and a goodness of fit measure Z^{DIST} statistic suggested by [4] is calculated for this purpose. It is reasonable to use distributions with three to five parameters for regional frequency analysis as compared to the distributions having two parameters, which might cause bias in tail quantile estimates. Hence, for this study the appropriate distributions are selected from among a family of three parameter distributions, namely: generalized logistic (GLO), generalized Pearson type III (PE3), generalized extreme value (GEV), generalized normal (GNO), and generalized Pareto (GPA). These five distributions are shortlisted based on previous studies such as [17-23]. Ratio diagrams (Fig. 2) display LM ratios, i.e. LM-Skewness and LM-Kurtosis of different regional distributions and data samples. That distribution is considered as appropriate distribution to whom the curve of population skewness and kurtosis is in close agreement with ratios of regional sample skewness and kurtosis.

The results of the ratio diagram coincided with the Z^{DIST} statistic for the (0, 0) trimming level. For other trimming levels, Z-test is evaluated as goodness of fit measure. Using 10% significance level if $|Z^{DIST}| \leq 1.64$, the distribution is considered an appropriate candidate distribution for the said homogenous region. If there is more than one candidate, probability distributions meet the said criteria and then the distribution with a value close to zero will be considered best among others. After careful examination of Table 4, it is revealed that for LM ("0, 0" trimming), GNO and GPA are designated appropriate

distributions at first and second positions. While for trimming of (1, 0) and (2, 0), GNO and GEV are declared as appropriate distributions at first and second positions, respectively. While for (3, 0) GEV and GNO and finally, for trimming of (4, 0) GEV and GLO are the best among others at first and second rank, respectively. Overall, GNO remained at the first position for the first three cases and GEV for the last two cases (Table 4).

The regional distribution parameters and the quantile estimates are found for best distribution against different return periods from 2 to 1,000 years. Specifically, (0, 0), (1, 0), (2, 0) GNO and (3, 0) and (4, 0) GEV are considered the most appropriate distribution for estimating design flood quantiles (Table 5).

The robustness of best regional frequency distribution (as determined by Z-criteria) is evaluated by the Monte Carlo simulation proposed by Meshgi and Khalili [29] for design flood estimates. The evaluation is based on RRMSE and RB by calculating the differences between simulated and calculated quantiles. RRMSE and RB for different sample sizes and return periods up to 100 years are presented in Table 6.

The values for RRMSE and RB are generally observed to be increased with smaller sample sizes and larger return periods as compared to larger samples. At each trimming level, the best distribution (Ranked 1st distribution

Table 4. Z^{-DIST} statistics for various distributions.

Trimming	ZGLO	ZGEV	ZGNO	ZPE3	ZGPA
(0, 0)	2.37	1.54 *	**0.44 ***	-1.48 *	-1.03 *
(1, 0)	1.98	1.06 *	-0.08 *	-2.06	-1.76
(2, 0)	1.48 *	0.57 *	**-0.46 ***	-2.26	-2.14
(3, 0)	0.98 *	**0.07 ***	-0.89 *	-2.54	-2.61
(4, 0)	0.67 *	**-0.23 ***	-1.09 *	-2.59	-2.82

Table 5. Regional quantile estimates for different levels of trimming.

Trimming		Parameters			Regional quantile estimates with nonexceedance probability F							
	Dist.	ε	α	k	0.500* 2**	0.800 5	0.900 10	0.950 20	0.980 50	0.990 100	0.998 500	0.999 1000
(0, 0)	GNO	0.761	0.544	-0.759	0.761	1.402	1.940	2.542	3.450	4.234	6.413	7.525
(1, 0)	GNO	0.764	0.557	-0.737	0.822	1.290	1.687	2.134	2.813	3.402	5.052	5.899
	GNO	0.768	0.570	-0.716	0.851	1.255	1.592	1.967	2.532	3.016	4.359	5.042
(3, 0)	GEV	0.611	0.445	-0.234	0.877	1.220	1.507	1.837	2.362	2.843	4.332	5.179
(4, 0)	GEV	0.608	0.459	-0.220	0.890	1.211	1.473	1.772	2.238	2.658	3.931	4.639

* Nonexceedance probability F;   ** return periods

determined by Z-test) performed better than others. For example, for trimming levels of (0, 0), (1, 0), and (2, 0), GNO is the most robust distribution, while for levels of (3, 0) and (4, 0), GEV is the most appropriate choice. Overall, TLM outperformed LM. The impact of trimming in the sense of robustness is shown in Fig. 3 (a and b), where we plotted RRMSE and RB of top distribution (most robust in Table 6) with respective trimming level (for n = 20 and n = 80) versus return periods. On the basis of RB (Fig. 3a), it is shown that for a sample size of n = 20 and up to return period of 20 years, GNO with trimming level of (1, 0) and (2, 0), GEV with trimming level (3, 0) and (4, 0) superseded GNO with trimming (0, 0). In other words, TLM performed better in the form of reduced

bias as compared to ML for a return period up to 20 years. Interestingly, after a return periods of 20 years, GNO (0, 0) superseded GNO (1, 0) and GNO (2, 0), but not than GEV (3, 0) and GEV (4, 0). Overall, GEV (3, 0) is better for a sample size of n = 20. By increasing sample size up to n = 80, significant changes appeared to be observed in the results. Overall performance and robustness of GNO (1, 0) and GNO (0, 0) is better. This implied that if the sample size is increased, the robustness of frequency distribution moved towards less trimmed values. For example, GNO (0, 0) outperforms GEV (3, 0) and GEV (4, 0), although for n = 80 the overall robust distribution is GNO (1, 0). Similar results are observed for RRMSE Table 6 and Fig. 3b.

Table 6. RB and RRMSE for best fitted distribution at different trimming levels.

| n	Trimming	Best Distr.	Q_2	Q_5	Q_{10}	Q_{20}	Q_{50}	Q_{100}
20	**(0, 0)**	GNO (RB)	0.004	0.016	0.02	0.023	0.029	0.032
		GNO (RRMSE)	0.075	0.081	0.086	0.101	0.126	0.142
	(1, 0)	GNO (RB)	-0.008	0.006	0.016	0.025	0.036	0.044
		GNO (RRMSE)	0.065	0.069	0.073	0.099	0.139	0.158
	(2, 0)	GNO (RB)	-0.007	0.007	0.016	0.024	0.034	0.041
		GNO (RRMSE)	0.06	0.059	0.065	0.093	0.126	0.15
	(3, 0)	GEV (RB)	-0.007	-0.004	-0.002	-0.001	0.002	0.006
		GEV (RRMSE)	0.043	0.044	0.048	0.09	0.11	0.133
	(4, 0)	GEV (RB)	-0.002	0.004	0.009	0.015	0.024	0.032
		GEV (RRMSE)	0.046	0.049	0.052	0.091	0.119	0.136
80	**(0, 0)**	GNO (RB)	0.003	0.0003	0.0001	0.0005	0.002	0.003
		GNO (RRMSE)	0.029	0.053	0.081	0.103	0.128	0.146
	(1, 0)	GNO (RB)	0.01	0.002	-0.002	-0.006	-0.009	-0.01
		GNO (RRMSE)	0.027	0.043	0.064	0.083	0.104	0.12
	(2, 0)	GNO (RB)	0.003	0	0	0.001	0.006	0.011
		GNO (RRMSE)	0.051	0.036	0.068	0.101	0.143	0.174
	(3, 0)	GEV (RB)	-0.001	0.005	0.01	0.015	0.021	0.026
		GEV (RRMSE)	0.047	0.039	0.068	0.103	0.156	0.203
	(4, 0)	GEV (RB)	0.003	0.002	0.003	0.004	0.008	0.013
		GEV (RRMSE)	0.043	0.041	0.073	0.109	0.16	0.205
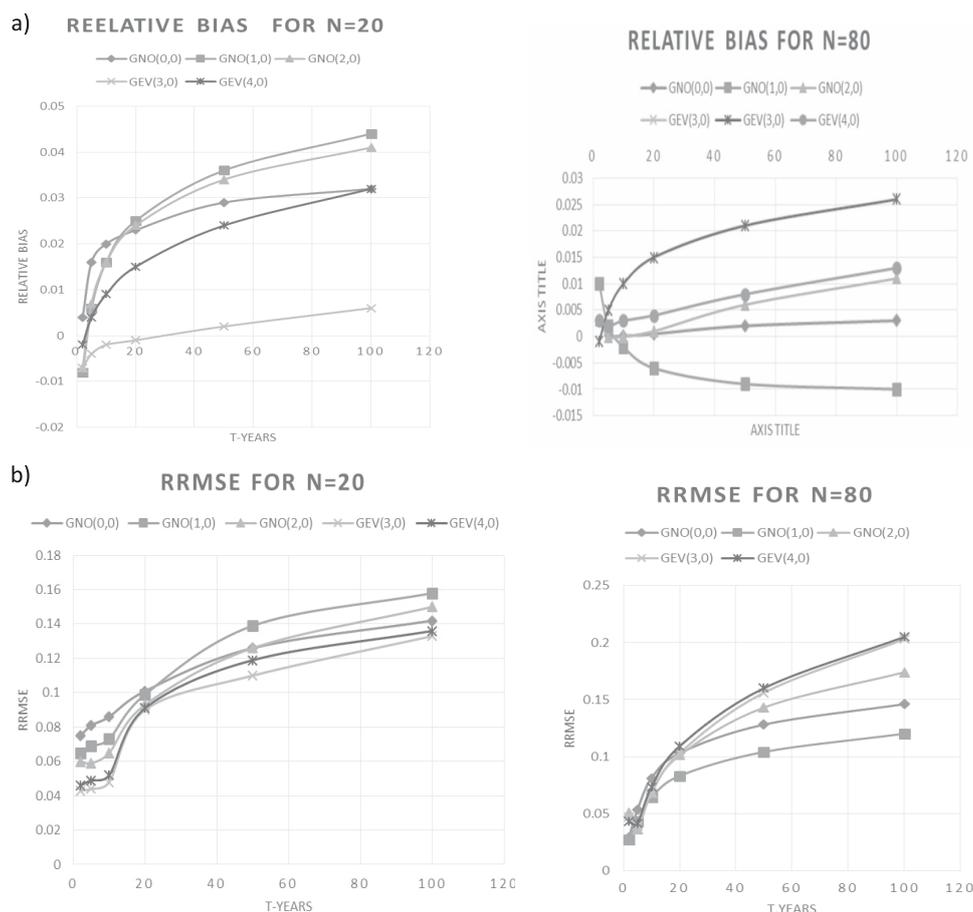
Fig. 3. Best fitted distribution at different trimming levels for n = 20 and n = 80; a) relative bias, b) relative root mean square error.

Conclusions

Regional Analysis of annual peak flows of nine sites is conducted at the Indus Basin in Pakistan. Data of all sites exhibited random behavior and the observations are found to be consistent, identical, and independently distributed. The discordancy measure is used for screening the data. No site is found to be discordant. Initially all sites are considered as one homogenous region and the regional homogeneity measure is calculated, which assures that these nine constitute a single homogenous region. For the selection of the robust frequency distribution a "Z" statistic criteria and ratio diagram (for (0, 0)) are used. Generalized normal (GNO) is found to be the most suitable distribution at trimming of (0, 0), (1, 0), and (2, 0), while GEV is found to be the most appropriate distribution at (3, 0) and (4, 0) trimming levels.

After this, design flood estimates are evaluated at different return periods. Estimating the frequency and magnitude of the floods is of great importance to make policies and take measures for disaster prevention and mitigation. For example, design flood estimates could be of great importance in water resources management for future hydrological projects such as successful operation and economical design of different hydrological structures such as water reservoirs, dams, bridges, small culverts,

spillways, urban drainage systems, etc. To evaluate their robustness, we calculated the RRMSE and RB for calculated quantiles and simulated quantiles through a Monte Carlo simulation performing 10,000 simulations for n = 20 and n = 80. The simulation indicated that the values of RRMSE and RB are generally increased with smaller sample sizes and larger return periods. Overall, TLM outperformed MLM for lower and higher quantile estimations.

Acknowledgements

References

1. HUSSAIN Z., PASHA G.R. Regional flood frequency analysis of the seven stations of Punjab, Pakistan, using L-moments. Water Resour Manage. **23**, 1917, **2009**.

2. RAHMAN A.S., RAHMAN A., ZAMAN M.A., HADDAD K., AHSAN A. AND IMTEAZ M. A study on selection of

probability distributions for At-site flood frequency analysis in Australia. *Natural* hazards, **69**, 1803. **2013**.

3. JINGYI Z., HALL M.J. Regional flood frequency analysis for the Gan-Ming River basin in China. J Hydrol. **296**, 98, **2004**.

4. HOSKING J.R.M., WALLIS J.R. Regional frequency analysis: An approach based on L.Moments. University press, Cambridge, **1997**.

5. PARIDA B.P., KACHROO R.K., SHRESTHA D.B. Regional flood frequency analysis of MahiSabarmati Basin (Subzone 3-a) using index flood procedure with L-moments. Water Resour Manage. **12**, 1, **1998**.

6. LIM Y.H., LYE L.M. Regional flood estimation for ungauged basins in Sarawak, Malaysia. Hydrol Sci J. **48**, 79, **2003**.

7. SHAHZADI A. 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**, 97, **2013**.

8. HASSAN B.G.H., PING F. Formation of homogenous regions for Luanhe Basin – by using L-Moments and cluster techniques. Int J Environ Sci Dev **3**, 205, **2012**.

9. SAF B. Regional flood frequency analysis using L-moments for the West Mediterranean region of Turke. Water Resour Manage. **23**, 531, **2010**.

10. FOWLER H.J., KILSBY C.G. A regional frequency analysis of United Kingdom extreme rainfall from 1961 to 2000. Int J Climatol **23**, 1313, **2003**.

11. NOTO L.V., LA LOGGIA G. Use of L-moments approach for regional frequency analysis in Sicily Italy. Water Resour Manag **1**, 23, **2009**.

12. CANNAROZZO M., NOTO L.V., LA LOGGIA G. Annual runoff regional frequency analysis in Sicily. Phys Chem Earth **34**, 679, **2009**.

13. GLAVES R., WAYLEN P.R. Regional flood frequency analysis in Southern Ontario using L-moments. Can Geogr **41**, 178, **1997**.

14. YUE S., WANG C. Determination of regional probability distributions of Canadian flood flows using L-moments. J Hydrol NZ **43**, 59, **2004**.

15. ABIDA H., ELLOUZE M. Probability distribution of flood flows in Tunisia. Hydrol Earth Syst Sci **12**, 703, **2008**.

16. ELAMIR E.A., SEHEULT A.H .Trimmed L-moments. Comp Stat Data Anal **43**, 299, **2003**.

17. SHABRI A., AHMAD U.N., ZAKARIA Z.A. TL-moments and L-moments estimation of the generalized logistic distribution. J Math Res **3**, 97, **2011**.

18. JAN NA, SHABRI A. Estimating distribution parameters of annual maximum streamflows in Johor, Malaysia using TL-moments approach. Theoretical and Applied Climatology. DOI No: 10-1007/s00704-015-1623-7. **2015**.

19. SHABRI A., DAUD Z.M., ARIFF N.M. Regional analysis of annual maximum rainfall using TL-moments method. Theor Appl Climatology **104**, 561, **2011**.

20. AHMAD I., ABBAS A., ASLAM M., AHMAD I. Total Annual Rainfall Frequency Analysis in Pakistan using Methods of L-Moments and TL-Moments. Sci. Int. (Lahore), **27**, 233, **2015**.

21. AHMAD I., FAWAD M., MAHMOOD I. At-Site Flood: Frequency Analysis of Annual Maximum Stream Flows in

Pakistan Using Robust Estimation Methods. Polish Journal of Environmental Studies, **24**, 2345, **2015**.

22. AHMAD U.N., SHABRI A., ZAKARIA Z.A. Trimmed L-moments (1, 0) for the generalized Pareto distribution. Hydrolog Sci J **56**, 1053, **2011**.

23. ZIN W.Z.W., JEMAIN A.A., IBRAHIM K. The best fitting distribution of annual maximum rainfall in Peninsular Malaysia based on methods of L-moment and LQ-moment. Theor Appl Climatol **96**, 337, **2009**.

24. KYSELÝ J. Trends in heavy precipitation in the Czech Republic over 1961-2005. Int J Climatol **29**, 1745, **2009**.

25. HOSKING J.R.M., WALLIS J.R. Parameter and quantile estimation for the generalized Pareto distribution. Technometrics, **29**, 339, **1987a**.

26. ALAM J., MUZZAMMIL M., KHAN M.K., Regional flood frequency analysis: comparison of L-moment and conventional approaches for an Indian catchment. ISH Journal of Hydraulic Engineering, **7**, 1, **2016**.

27. HOSKING J.R.M. "L-moments: Analysis and estimation of distributions using linear combinations or order statistics," Journal of Royal Statistical Society, Series B Methodology, **52**, 105, **1990**.

28. GREENWOOD J.A., LANDWEHR J.M., MATALAS N.C., WALLIS J.R. Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. Water Resour Res **15**, 1049, **1979**.

29. MESHGI A., KHALILI D. Comprehensive evaluation of regional flood frequency analysis by L- and LH-moments. II. Development of LH-moments parameters for the generalized Pareto and generalized logistic distributions. Stoch Environ Res Risk Assess **23**, 137, **2009**.

30. MODARRES R., SARHADI A. Statistically-based regionalization of rainfall climates of Iran. Global and Planetary Change, **75**, 67, **2011**.

31. YANG T., XU C.Y., SHAO Q.X., CHEN X., Regional flood frequency and spatial patterns analysis in the Pearl River Delta region using L-moments approach. Stochastic Environmental Research and Risk Assessment, **24**, 165. **2010**.

32. SHIAU J.T., WU F.C. Regionalization of natural flow regime: Application to environmental flow optimization at ungauged sites. River research and applications, **25,** 1071, **2009**.

33. KUMAR R., GOEL N.K., CHATTERJEE C., NAYAK P.C. Regional Flood Frequency Analysis using Soft Computing Techniques. Water Resources Management, **29**, 1965, **2015**.

34. KOMI K., AMISIGO B.A., DIEKKRÜGER B., HOUNTONDJI F.C. Regional Flood Frequency Analysis in the Volta River Basin, West Africa. *Hydrology*, **3**, .5, **2016**.

35. NGONGONDO C.S., XU C.Y., TALLAKSEN L.M., ALEMAW B., CHIRWA T. Regional frequency analysis of rainfall extremes in Southern Malawi using the index rainfall and L-moments approaches. Stochastic Environmental Research and Risk Assessment, **25**, 939, **2011**.

36. AYDOĞAN D., KANKAL M., ÖNSOY, H. „Regional flood frequency analysis for Çoruh Basin of Turkey with L-moments approach." Journal of Flood Risk management, **6**, 69, **2016**.