International Journal of Business & Management Study – IJBMS 2018
Copyright © Institute of Research Engineers and Doctors , SEEK Digital Library
Volume 5 : Issue 1 - [ISSN : 2372-3955] - Publication Date: 25 June, 2018

The Quantiles Estimation in the Monsoon Season of the Seven Unguaged Sites of Punjab Pakistan with Regional Rainfall Frequency Analysis

Uzma Nawaz*, Tooba Nihal, Aisha Ejaz

Abstract— **Pakistan is located between latitude 25^o and 36^o with strong spells of monsoon rains showering from June to September with somewhat variation in showers according to location. Heavy rains in monsoon season make dams, rivers and lakes overflow causing destructive floods. The scenario provoked the study on the estimation of regional rainfall quantiles of Annual Maximum Monsoon Rainfall Totals (AMMRT) of the seven unguaged sites of the Punjab province in Pakistan [1]. The estimates are obtained from L-moment based Regional Rainfall Frequency Analysis (RRFA) [2]. The monsoon data series of the seven sites of northern Punjab were found random and identically distributed with no serial correlation. These results were the output of the run test, lag-1 correlation and Mann-Whitney U test. The L-moment based discordancy measure exposed no site discordant in the group of seven. The heterogeneity measure concluded a homogenous region of the seven sites. The L-moment Ratio diagrams exposed the Generalized Normal as the best choice of the regional distribution for the quantile estimation. The estimates of the study may be used for the estimation of the rainfall quantiles of the seven sites for different return periods. The estimates will be provocative to design future preventive measures for the harmful impact of hydro meteorological events at these sites in Punjab Pakistan [3].**

Keywords— L- moments, Regional Rainfall Frequency Analysis, Discordancy Measures, Heterogeneity Measure.

1. Introduction

The climate of Pakistan is a segregation of four weathers. December to February is a cool dry weather that runs through the hot, dry spring from March through May. The monsoon period spell in from June pulling it up to September. The monsoons draw back in October and November. The climatic segmentation is more or less also dependent on the location. Pakistan has extreme weather spiking with high and low temperatures, heaviest rainfall and flooding. The capital city of Islamabad in the upper Punjab has a heat wave that drops to 2 ° Celsius as a mean temperature per day in January and the mean spikes up to of 38 °C in June. Fifty percent of the yearly rainfall storms in July and August, averaging about 255 millimeters in each of these two months. In the remaining months the average rainfall is about fifty millimeters per month and spring comes with hailstorms. Monsoon rainfall has its strong impact on the hydrological perspective of a location or a site specifically the dams, open water storages, agriculture and sewers.

* Dr Uzma Nawaz (Author)
Department of Statistics
The Women University Multan Pakistan

Tooba Nihal and Aisha Eijaz (Co Authors)
Department of Statistics
The Women University Multan
Pakistan

This necessitates to find the estimates of parameters for finding flood risks, optimal designing for the management of water related issues and the like [4]. The study comprises of the AMMRT data of seven sites of upper Punjab (also called the northern Punjab) in the monsoon season [5].

The following four steps foundation for the application of RRFA.

1) A random check of the AMMRT data series concluded with application of Run Test, Mann-Whitney U test and Lag-1 correlation coefficient.
2) L-Moments, Discordancy Measure and the homogeneity check of the region of seven sites of northern Punjab.
3) L-Moment Ratio Diagram and Z^{DIST} criteria figuring the most appropriate distribution for the homogeneous region formed.
4) Quantile estimation and Inference for different return periods. Step 3) and 4) are the component of RRFA and complete the analysis [6].

1).a Run Test on the AMMRT data

Table 1. Output of Run Test

S. No.	Sites	N	ONR	P-values
1	Faisalabad	36	15	0.850
2	Islamabad	36	18	0.968
3	Jhelum	36	22	0.300
4	Lahore	36	13	0.299
5	Murree	36	17	0.543
6	Sargodha	36	15	0.820
7	Sialkot	36	17	1.000

The p-values of observed number of runs (ONR) give insignificant results providing evidence of randomness in the data series of the sites under study.

1).b Mann-Whitney U Test

Table 2. Application output of Mann-Whitney U Test

S.No.	Sites	N	Groups	W	P-values
1	Faisalabad	36	18,18	281.0	0.1032
2	Islamabad	36	18,18	329.0	0.9118
3	Jhelum	36	18,18	330.0	0.9370
4	Lahore	36	18,18	340.0	0.8371
5	Murree	36	18,18	322.0	0.7397
6	Sargodha	36	18,18	305.0	0.3843
7	Sialkot	36	18,18	350.0	0.6016

The p-values of the test statistic (W) signify that the two groups are identically distributed at each site [7].

International Journal of Business & Management Study – IJBMS 2018
Copyright © Institute of Research Engineers and Doctors , SEEK Digital Library
Volume 5 : Issue 1 - [ISSN : 2372-3955] - Publication Date: 25 June, 2018

1).c Lag-1 Correlation Coefficient

Table 3. Results of Ljung-Box Q (LBQ) Statistics

S. No.	Sites	r_1	LBQ Statistics	P-values
1	Faisalabad	0.096	0.358	0.549
2	Islamabad	-0.046	0.084	0.772
3	Jhelum	-0.242	2.280	0.131
4	Lahore	0	0	0.998
5	Murree	0.114	0.512	0.474
6	Sargodha	0.070	0.194	0.660
7	Sialkot	-0.072	0.201	0.654

r_1 is the lag-1 correlation coefficient.

The column of p-values supports the hypothesis of no serial correlation in the data

2).a L-Moments and Discordancy Measure

L-moments are the linear combination of the probability weighted moments (PWM), introduced by Hosking (1990) [8]. The PWM were worked out by Greenwood *et al.* (1979), mathematically defined as,

$$\beta_r = E[x\{F(x)\}^r]$$

The r^{th} L-moments by Hosking (1990) are,

$$\lambda_{r+1} = \sum_{l=0}^{r} \beta_l (-1)^{r-k} \binom{r}{l}\binom{r+l}{l}$$

Concluding the first four L-moment ratios as,

Location: λ_1 = *Mean of each site*

Scale, L-Cv (τ): $\tau = \lambda_2/\lambda_1$

L-Skewness (τ_3): $\tau_3 = \lambda_3/\lambda_2$

L-Kurtosis (τ_4): $\tau_4 = \lambda_4/\lambda_2$

Discordancy measurement (Di) of a site is calculated by,

$$D_i = \frac{N}{3(N-1)}\left(\mu_i - \bar{\mu}\right)^T M^{-1}\left(\mu_i - \bar{\mu}\right)$$

Where N is the total no. of sites,

$$\mu_i = \left[\tau^i \quad \tau_3^i \quad \tau_4^i\right]^T,$$

$$\bar{\mu} = N^{-1}\sum_{i=1}^{N}\mu_i ,$$

and M is the variance-covariance matrix of μ_i.

If any of the D_i value of a site exceeds the critical value of discordancy statistics as determined by Hosking and Wallis (1997) the site is considered as discordant [9].

Table 4. Summary Statistics of L-Moments and Discordancy Measure of each Site

S.No.	Sites	n	λ_1	τ	τ_3	τ_4	D_i
1	Faisalabad	36	60.68	0.34	0.38	0.23	0.83
2	Islamabad	36	109.4	0.30	0.22	0.10	1.03
3	Jhelum	36	81.08	0.29	0.18	0.12	1.79
4	Lahore	36	51.36	0.26	0.33	0.20	1.45
5	Murree	36	139.0	0.27	0.17	0.07	0.65
6	Sargodha	36	69.31	0.34	0.46	0.31	1.00
7	Sialkot	36	81.05	0.31	0.39	0.24	0.25

Table 4 contains the Summary Statistics $(\lambda_1, \tau, \tau_3, \tau_4)$ and discordancy measure D_i. The D_i value of seven sites is 1.917 (Hosking 1997). Therefore it is observed that none of the seven sites of northern Punjab are discordant to one another [10].

2).b Homogeneity Check of the Region of Seven Sites

Hosking and Wallis (1993) presented a heterogeneity measure that compares the observed dispersion of L-moments with the simulated dispersion of L-moments calculated through simulations. The four parameter Kappa distribution is used to generate 500 simulated data regions (Hosking and Wallis 2000). These regions are homogeneous have no autocorrelation, and sites have the equal record lengths as real data.

The heterogeneity measure is given by,.

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

where V for each simulated sample with μ_v and σ_v.

The weighted standard deviation of L CVs is,

$$V = \frac{\sum_{i=1}^{N} v_i (t^i - \bar{t})^2}{\sum_{i=1}^{N} v_i}$$

The region is

"Acceptably homogeneous" if H < 1 ,

"Possibly heterogeneous" if 1 < H < 2 ,

"Absolutely heterogeneous" if H ≥ 2 .

The calculated heterogeneity measure (H=-0.59<1) form an acceptably homogenous region for the monsoon spell in the northern Punjab.

3).a L-moment Ratio Diagram

L-moments ratio diagram is a graph of L-skewness against L-kurtosis. The curve shows the imaginary relation among L-skewness and L-kurtosis, which is used for selection of suitable frequency distribution. L-moments ratio diagram are widely used for the choice of a probability

International Journal of Business & Management Study – IJBMS 2018
Copyright © Institute of Research Engineers and Doctors , SEEK Digital Library
Volume 5 : Issue 1 - [ISSN : 2372-3955] - Publication Date: 25 June, 2018

distribution in RFA. L-moment ratios are also called standardized L-moments. The diagram illustrates the most appropriate distribution for the estimation of regional quantiles.

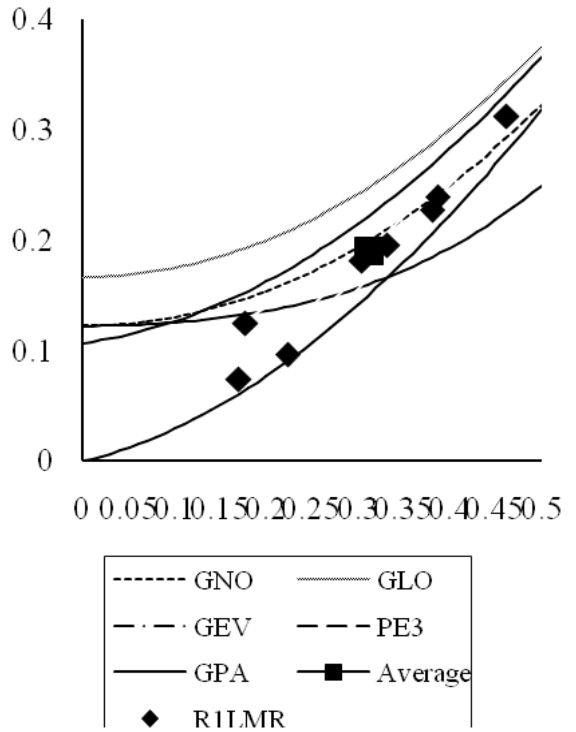

Fig.1. L-Moment Ratio Diagram

In Fig.1 the average (τ_3, τ_4) lies exactly on the curve of the Generalized Normal Distribution (GNO). The GNO is visualized as the best fitted regional distribution on the data of the seven sites as homogeneous region of northern Punjab.

3).b Z^{DIST} Criteria

It is used to check the goodness-of-fit of all the proposed distribution in the L-Moment Ratio diagram. Hosking and Wallis (1997) give a valuable measure that is defined as,

$$Z^{DIST} = \frac{(\tau_4^{DIST} - t_4^R + B_4)}{\sigma_4}$$

where τ_4^{DIST} is the L-kurtosos of the fitted Kappa distribution, t_4^R s the L-kurtosis, B_4 is the bias of t_4^R and σ_4 is the simulated standard deviation of t_4^R.

This fit is good if $\left|Z^{DIST}\right| \leq 1.64$, the most suited distribution is selected which has small $|Z^{DIST}|$ measure. Simulation and calculations are covered using FORTRAN routines of Hosking (2000) available in international mathematical and statistical libraries (IMSL).

Table 5. Results of Goodness-of-Fit Test

S. No.	Distributions	$\|Z^{DIST}\|$
1	GNO	0.29
2	PE3	0.89
3	GEV	0.98
4	GPA	1.30
5	GLO	1.82

GNO has the smallest value of the Z^{DIST} criteria computed as 5% level of significance far less than the Zα value; 1.64, validating the appropriate graphical appearance of the GNO from the L-Moment Ratio Diagram.

4. Statistics Estimation from GNO and Statistical Inference

The following are the parameters estimates of the best fitted frequency distributions the GNO.

Table 6. Parameter Estimates

	μ	α	Λ
GNO	0.840	0.452	-0.637

Where μ is the location parameter, α is the scale and λ is the shape of GNO distribution.

Table 7. Regional Quantiles Estimates from GNO

Quantiles Estimates of GNO Distribution for different non-exceedance Probabilities (F)							
F	0.0100	0.0500	0.1000	0.2000	0.5000	0.9000	0.9500
GNO	0.2916	0.3793	0.4441	0.5455	0.8400	1.7356	2.1536
F	0.9800	0.9900	0.9950	0.9970	0.9980	0.9987	0.9990
GNO	2.7556	3.2534	3.7913	4.3722	4.5689	4.9989	5.2108

Table 7 provides the quantiles estimates from GNO of different non-exceedance probabilities with varying return periods. Return period is also a synonym to recurrence interval. It is the estimation of the likelihood of the occurrence of an event like rainfall, flood and other extreme events. The return period T is the inverse of the probability of exceedance (P), $\left(T = \frac{1}{P}\right)$. P is the probability of an event over a given time period. Given T, P is defined as $\left(P = \frac{1}{T}\right)$. The non-exceedance probability F is the complement of P i.e $(F = 1 - P)$. For T=10 years return period P=0.1 complementing the on-exceedance probability as F=1-0.1=0.900. The values of quantiles given in the Table 7 can be interpreted as for GNO (0.0900) = 1.7356 is the magnitude of rainfall which will happen once on the average in 10 years, for GNO (0.9900) = 3.2534 is the magnitude of rainfall which will happen once on the average in 100 years and similarly for GNO (0.9980) = 4.5689 is the magnitude of rainfall which will happen once on the average in 500 years. All these values can be interpreted in the similar manner.

5. Results

The quantile estimates resting on the best regional distribution present the expected occurrence of rainfall in the varying years to come. Therefore quantiles estimates may be used for hydrological planning and management of the water resources. In particular in the monsoons spells in the upper Punjab in the years to come. Further the utility lies in the development of the preventive measures in estimated anticipation for natural disasters to come. The city of Multan which does not fall in the upper Punjab yet forms a homogeneous with the other seven sites of the province. This may be attributed to the particular time duration of the monsoons that strike the whole region.

6. Recommendations

The RRFA was first developed as a statistical series of applications for the data of extreme events in 1993 [11]. In Pakistan some statistical work is contributed focusing on the sites of Sindh and Punjab on the data series of the whole year [12]. The novelty of the study lies here in particular of the quantile estimation in the monsoons seasons [13]. The monsoon winds are the main generation of rainfall in the season so are the major cause of water related issues.

The workouts for the regional quantiles in the rainy season of the meteorological sites in the provinces of Pakistan has now become a necessity. It would be advisably useful for hydrological engineers and scientist.

7. References

[1]. Koutsoyiannis, D., & Baloutsos, G. (2000). Analysis of a long record of annual maximum rainfall in Athens, Greece, and design rainfall inferences. *Natural Hazards, 22*(1), 29-48.

[2]. Salma, S., Rehman, S., & Shah, M. (2012) Rainfall trends in different climate zones of Pakistan. *Pakistan Journal of Meteorology, 9*.

[3]. Mekonnen, S. (2013). *Rainfall Intensity-Duration-Frequency Analysis for Selected Areas of Wollega Zone, Ethiopia.* Haramaya University.

[4]. Zahid, M., & Rasul, G. (2011). Frequency of extreme temperature and precipitation events in Pakistan 1965–2009. *Sci. Int.(Lahore), 23*(4), 313-319.

[5]. Van Gelder, P., & Neykov, N. M. (1998). *Regional frequency analysis of extreme water levels along the North-Sea Coasts.* Paper presented at the European Geophysical Society, XXIV General Assembly, The Hague, The Netherlands.

[6]. Hussain, Z., Shahzad, M. N., & Abbas, K. (2016). Application of regional rainfall frequency analysis on seven sites of Sindh, Pakistan. *KSCE Journal of Civil Engineering*, 1-8.

[7]. Hirsch, R. M., & Slack, J. R. (1984). A nonparametric trend test for seasonal data with serial dependence. *Water Resources Research, 20*(6), 727-732.

[8]. Hosking, J. R. M., & Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. *Water Resources Research, 29*(2), 271-281. doi: 10.1029/92wr01980.

[9]. 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).

[10]. Vivekanandan, N. (2014). Rainfall Frequency Analysis Using L-moments of Probability Distributions. *International Journal of Computer Application and Engineering Technology, 3*(3), 248-256.

[11]. Khalifa, A. H. N., & Al-Omran, A. T. *Consumption Adaptive Voltage Controlled Oscillator Using Fpaa.*

[12]. Khattak, M. S., & Ali, S. (2015). Assessment of temperature and rainfall trends in Punjab province of Pakistan for the period 1961-2014. *Journal of Himalayan Earth Science, 48*(2).

[13]. Lee, C. (2005). Application of rainfall frequency analysis on studying rainfall distribution characteristics of Chia-Nan plain area in Southern Taiwan. *Crop Environ. Bioinf, 2*, 31-38.