Flood Frequency Analysis of Annual Maximum Stream Flows in Pakistan Using L-Moments and TL-Moments

By

Muhammad Fawad

Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan
2015

- Frequency distribution
- Natural disasters
- Probability Theory.

Flood Frequency Analysis of Annual Maximum Stream Flows in Pakistan Using L-Moments and TL-Moments

By

Muhammad Fawad

Supervised by

Dr. Ishfaq Ahmad

Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan

2015

بسم الله الرحمن الرحيم

Flood Frequency Analysis of Annual Maximum Stream Flows in Pakistan Using L-Moments and TL-Moments

By

Muhammad Fawad

A Dissertation
Submitted in the Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
IN
STATISTICS

Supervised by

Dr. Ishfaq Ahmad

Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan

2015

Certificate

Flood Frequency Analysis of Annual Maximum Stream Flows in Pakistan Using L-Moments and TL-Moments

By

Muhammad Fawad

A DISSERTATION SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF THE ***MS in STATISTICS***

We accept this dissertation as conforming to the required standard.

1. ____________
Dr. Javid Shabbir
(External Examiner)

2. ____________
Dr. Muhammad Akbar Awan
(Internal Examiner)

3. ____________
Dr. Ishfaq Ahmad
(Supervisor)

4. ____________ 4/8/15
Dr. Rahmat Ellahi
(Chairman)

Department of Mathematics & Statistics
Faculty of Basic and Applied Sciences
International Islamic University, Islamabad
Pakistan
2015

Dedication

To my family,
For the endless support and patience.

To my Teachers,
For the constant source of Knowledge and
Inspiration.

To my friends,
The ones that are close and the ones that are far.

Forwarding Sheet by Research Supervisor

The thesis entitled **"Flood Frequency Analysis of Annual Maximum Stream Flows in Pakistan Using L-Moments and TL-Moments"** submitted by **Muhammad Fawad** (Registration # 10-FBAS/MSST/F12) in partial fulfillment of M.S degree in Statistics has been completed under my guidance and supervision. I am satisfied with the quality of his research work and allow him to submit this thesis for further process to graduate with Master of Science degree from Department of Mathematics and Statistics, as per IIU Islamabad rules and regulations.

Dated: ______________________

Dr. Ishfaq Ahmad,
Assistant Professor
Department of Maths & Statistics,
International Islamic University,
Islamabad.

Acknowledgements

All the praises to Almighty ALLAH the most gracious, the most merciful and the creator of all the creature in the creature. Thanks to ALLAH and his Prophet (P.B.U.H). There are few people who made this journey easier for me with their encouragement and ideas.

I express my gratitude to my supervisor **Dr. Ishfaq Ahmad** for his regardless and inspirational efforts and moral support throughout my research carrier. His sound advices and lots of good ideas were very helpful to me. I would have been lost without him. May ALLAH bless him with all kinds of happiness and success in his life and may all his wishes come true.

I am also thankful to Chairman Department of Mathematics and Statistics **Dr. Rahmat Ellahi** for providing such necessary facilities.

I pay my full regards and heartiest thoughts of appreciation to all my teachers of Departments of Mathematics and Statistics **Dr. Zahid Iqbal, Dr. Babur Wasim Arif** and **Dr. M Akbar Awan** for their cooperation and valuable suggestions throughout this research work.

I would also acknowledge the support and facilitation provided by Federal Flood Commission, Indus River System Authority and WAPDA for providing the appropriate Annual Maximum flows data for analysis.

I am thankful to **J. R. M. Hosking** who guided me step by step in this study.

My Deepest gratitude to my family who are the real pillars of my life. They always encouraged me and showed their everlasting love, care and support throughout my life. Their continuous encouragement and, humble prayers, support (both financially and moral) from my Father, Uncle and mother is unforgettable.

I am also thankful to my uncle Prof. Mohammad Ayaz for his love, care and support in my life, which has been directly encouraging me for my study. I am also thankful to my friends Aamir Abbas, Waqas Mehmood, Muhammad Yasin, Qasim Jan, Shah Suhail, Sayyad Anwar, Mahboob Elahi, and Adnan Ali khan. I am grateful for all the emotional support, entertainment, love and care they provided.

Finally, I express my sincere appreciation to all the helpful staff of the Mathematics Department IIUI. I would like to thanks to everybody who was important to successful realization of this thesis as well as expressing my apology to those that I could not mention.

Muhammad Fawad

DECLARATION

I hereby declare that this thesis, neither as a whole nor a part of it, has been copied out from any source. It is further declared that I have prepared this dissertation entirely on the basis of my personal efforts made under the supervision of my supervisor **Dr. Ishfaq Ahmad.** No portion of the work, presented in this dissertation, has been submitted in the support of any application for any degree or qualification of this or any other learning institute.

Signature: ____________________

Muhammad Fawad
MS (Statistics)
Reg. No 10-FBAS/MSST/F12
Department of Mathematics and Statistics,
Faculty of Basic and Applied Sciences,
International Islamic University Islamabad, Pakistan.

TABLE OF CONTENTS

List of Figures

Figure No.	Title	Page

List of Tables

Table No. **Title** **Page**

List of Abbreviations

AM	Annual Maximum
AD	Anderson Darling
C-S	Chi-Square
CPA	Change Point Analysis
CUSUM	Cumulative Sum
CDF	Cumulative Distribution Function
E	Exponential
EV1	Extreme Value Type 1
MENT	Maximum Entropy
FFA	Flood Frequency Analysis
FA	Frequency Analysis
FFC	Federal Flood Commission
GLO	Generalized Logistic
GEV	Generalized Extreme Value
GPA	Generalized Pareto
GU	Gumbel
GNO	Generalized Normal
G	Gamma
GMM	Generalized Method of Moments
HCDN	Hydro-Climate Data Network
IID	Identical Independent Distributed
ICM	Incomplete Mean Method
K-S	Kolmogorov-Smirnov
L-Moments	Linear Moments
L-CV	Linear Coefficient Variation
L-Skewness	Linear Skewness
L-Kurtosis	Linear Kurtosis
TL-Moments	Trimmed Linear Moments

TL-CV	Trimmed Linear Coefficient Variation
TL-Skewness	Trimmed Linear Skewness
TL-Kurtosis	Trimmed Linear Kurtosis
LS	Least Square
LP3	Log Pearson Type III
LN	Log Normal
LH-Moments	Linear Higher order Moments
LN2	Log Normal with two parameter
L3	Logistic with three parameter
LBQ	Ljung-Box Q
MOM	Method of Moments
MLE	Maximum likelihood Estimation
MADI	Mean Absolute Deviation Index
MSDI	Mean Square Deviation Index
M-W	Mann-Whitney
MK	Mann-Kendall test
MIX	Mixed Moments
N2	Normal distribution with two parameter
NDMA	National Disaster Management Authority
PPCC	Probability Plot Correlation Coefficient
PWM	Probability Weighted Moments
PDF	Probability Distribution Function
RMED	Root Mean Error Deviation
RMSE	Root Mean Square Error
RMAE	Relative Mean Absolute Error
RRMSE	Relative Root Mean Square Error
T	Return Period
W3	Weibull Type III
WAPDA	Water and Power Development Authority

Abstract

L-Moments and TL- Moments techniques are being used in this study to estimate the parameters of the best fitted distribution for Annual Maximum flows (AM) of each of 18 gauging sites of Pakistan separately. The three selected distributions Generalized Logistic (GLO) Distribution, Generalized Extreme Value (GEV) Distribution, Generalized Pareto (GPA) Distribution are fitted for each of 18 sites, which are commonly used in Flood Frequency Analysis (FFA). The performance of these distributions is compared by using goodness-of-fit methods, such as Mean Absolute Deviation Index (MADI), Relative Root Mean Square Error (RRMSE) and Probability Plot Correlation Coefficient (PPCC). Further L-Moments Ratio diagram and TL-Moments Ratio diagram are also being used to confirm the goodness-of-fit for the above three distributions. Finally the results show that GPA distribution is the most suitable distribution for the AM flows for the majority of the sites in Pakistan followed by GLO and GEV distributions respectively, when the parameters are estimated by using L-moments and TL- Moments techniques. The Return Periods from 2 to 1000 years and corresponding quantiles are provided for best fitted distribution. It is noted that for shorter and high return periods the quantile estimates are larger for TL-moment as compare to L-moment. The difference in quantile estimates occurs because of trimming, as we used trimming $(t_1 = 1, t_2 = 0)$ in this study which mean trimming is on the smallest order statistics.

CHAPTER 1

1.1 Floods in Pakistan

Environmental events have always played a vital role in human history. Floods, rainstorms, droughts and high winds are some of the manifestations of these events that cause tremendous destruction and bring misery to human existence. Flood, is one of the natural disasters that could wrought havoc to the properties, infrastructures, animals, plants and even human lives.

Pakistan faced the floods for so many years. These are repeated in almost every monsoon season when there is heavy rainfall and snowmelt in summer. The financial loss for Pakistan in the last six decades has estimated to be US$ 37.554 billion. During the period 1950 to 2013, there were around 11,572 human casualties, 188,531 villages were partially or completely damaged and approximately total area of 603,942 Sq.km was affected due to 21 major flood events.

Floods are perennial problem in Pakistan, however, according to the Federal Flood Commission (FFC) report there are 22 major floods i-e 1950, 1955, 1956, 1957, 1959, 1973, 1975, 1976, 1977, 1978, 1981, 1983, 1984, 1988, 1992, 1994, 1995, 2010, 2011, 2012, 2013 and 2014. These floods have greatly affected the economic and social conditions of the people.

Floods are a natural disaster that not only happen in Pakistan, but also in other parts of the world. Flood is also known as spate, a natural disaster that could destroy human properties and lives. Therefore, it is important that if we cannot stop them completely atleast we can minimize destructions due to these floods.


Fig. 1.1 Show Damages caused by the Floods of 2010


Fig.1.2 Flood of 2014


Fig.1.3 Show Damages caused of Houses by the Floods of 2011


Fig.1.4 Show the Flooded Area of 2014

According to FFC report that the 1950 disaster was the first deadliest disaster in term of human losses in the country. Which was spreaded over 17,920 kilometres area. In which 10,000 villages were flooded and 21, 920 lives were lost. In 1977 the flood damaged 2,185 villages and killed 848 people. The 1992 flood sunken 13,208 villages and eliminated 1,000 lives.

The second deadliest flood was recorded in 2010 in which 1,985 lives were killed and 21 million people were affected. That flood had destroyed 17,533 villages over 160,000 square kilometre area.

According to FFC report in 2011, 516 people died and 571 died in 2012. From 2010 to 2012 total human loses are 3,072 and $16 billion of economic losses.

According to National Disaster Management Authority (NDMA) report, 2013 Monsoon floods have killed 69 people: 18 in Baluchistan, 22 in Sindh, 14 in Khyber Pakhtunkhwa and 15 in Punjab. The authority reported that 81,674 people and 333 villages were affected; 2,533 houses were completely destroyed, 1,782 partially and 135,076 acres of crops are affected.

Details are given in the following table 1.1.

Table 1.1 Historical Flood Events Experienced in Pakistan

Sr.No.	Year	Direct losses (US$ million) @ 1US$=PKR 86	Lost lives (No)	Affected Villages (No)	Flooded area (Sq.km)
1	1950	488	2,190	10,000	17,920
2	1955	378	679	6,945	20,480
3	1956	318	160	11,609	74,406
4	1957	301	83	4,498	16,003
5	1959	234	88	3,902	10,424
6	1973	5134	474	9,719	41,472
7	1975	684	126	8,628	34,931
8	1976	3485	425	18,390	81,920
9	1977	338	848	2,185	4,657
10	1978	2227	393	9,199	30,597
11	1981	299	82	2,071	4,191
12	1983	135	39	643	1,882
13	1984	75	42	251	1,093
14	1988	858	508	100	6,144
15	1992	3010	1,008	13,208	38,758
16	1994	843	431	1,622	5,568
17	1995	376	591	6,852	16,686
18	2010	10,000 @ 1US$=PKR 86	1,985	17,553	160,000
19	2011	3730* @ 1US$=PKR 94	516	38,700	27,581
20	2012	2640** @ 1US$=PKR 95	571	14,159	4,746
21	2013	2,000^ @ 1US$=PKR 98	333	8,297	4,483
Total		**37,554**	**11,572**	**188,531**	**603,942**

The NDMA assessment has proved that floods in the last five years have caused more damages than the floods in the last six decades together. Due to the absence of effective disaster management system, Pakistan suffers loss of almost $800 million each year. The effects of these floods are visible in every aspect of human life. These causes' unbearable problems in housing, health, education, irrigation, transport, communication, water supply, and sanitation and energy sector. More importantly, agriculture sector, which is the backbone of Pakistan's economy is severely affected by these floods. Consequently, it also affected private sector and industries.

The flood of September 2014, has killed around 360 people, injured 646, affected 2,523,681 people, damaged 56,644 houses, inundated 4,065 villages, affected 2,416,558 acres of crops, and total of 8,957 cattle were perished.

1.2 Introduction to Flood Frequency Analysis (FFA)

It is important to analyse the AM flows data to extract the probability distribution for floods and other associated phenomena. Through probability distribution, it is easy to predict flood events and it also facilitates to determine their characteristics. Additionally, preventive measures can be adopted and models for flash flood warning can be also designed.

Before we study the characteristics of floods events it is pertinent to know the science of water. The movement, distribution, resources, hydrologic cycle and quality of water is called hydrology. On the basis of these characteristics modelling is adopted for water on the whole earth. Moreover, on the basis of these hydrological data like rainfall or river flow and their statistical analysis, hydrologists are able to predict future hydrological phenomena. Frequency of flood is considered to be a very active area of investigation in Statistical Hydrology Rao *et al* (2000).

FFA is a statistical method that is commonly used for forecasting future events at different return periods. Frequency Analysis (FA) is the estimation of how often a specified event will happen Hosking and Wallis (1997).The main objective of FFA is to relate the magnitude of extreme events to their frequency of happening through the use of probability distributions assuming that the data are independent and identically distributed Chow *et al* (1988).

Statistical analysis of hydrological data at different time scales generally met in water resources planning studies are established on a set of vital assumptions, these are that the series is stationary, homogenous and independent. The violation of these assumptions can lead us in surprising inaccurate effects on the results. Stationarity implies that the future will be statistically not different from past. Operationally, the notion of stationarity is qualified. Strict stationarity means that the probability laws that governed the past will govern the future in same way. A less strict form of stationarity that is weak stationarity implies that for the probability law to change with time, but in a manner that does not affect specific statistical characteristics, e.g., the lower order moments or the degree of importance of the absolute times of events over that of specific lags in the time of events. In water management we assume strict stationarity while in Hydrologic processes we assume weak stationarity, such that the covariances of events separated in time depend only on the time differences and not on the absolute times of the events. Further it is also assumed that the processes are stationary up to at least the third moment and perhaps the fourth moment, such that the mean, variance, skewness, and perhaps the kurtosis of the past will be those of the future. (Nicholas C. Matalas). As if the time series data is nonstationary, we can analyse its performance simply for the time period under consideration. Individually set of time series data will be tested consequently for a specific period. As a result, it is not feasible to generalize it to other time periods. Thus for the purpose of predicting such (nonstationary) time series might be of little practical value. The concept of Independence of a time series is more related to strict stationarity as compared to weak stationarity. In hydrological data trends are generally presented through natural or artificial changes. Hydrologic Nonstaionarity in the data which may be due to climate change cannot be dismissed during analysis of the data. Homogeneity in a series implies that that the data in the series belongs to one population and subsequently have a time-invariant mean. Here it is to be noted that we are dealing only one variable that is the water flow of a particular location for different years. We have a time series and time series of flow. We are not concerned about neither two different variables nor their independence. In this sense homogeneity or identically distributed means time homogeneity; that means, you have a time series and then this should be homogeneous among time. The space homogeneity are mainly related to regional frequency analysis. In the present study we are not interested in space homogeneity. Non-homogeneity in a series may occurs due to variety of sources e.g. changes in data collection method, climate change, changes in land use in the catchment and changes in abstractions and river regulations Fernando & Jayawardena (1994).

1.3 Introduction Linear –Moments (L-Moments) and Trimmed Linear-Moments (TL-Moments)

The Uncertainty events such as floods are uncommon and happen in a short amount of time. Therefore, it is crucial for probability distribution to study its characteristics. FA is used to find the probability distribution for uncertainty of events through replacing space for time. Outlier values are generally to be found in the flood data which are logically danger.

The standard estimation methods such as Least Square (LS), Method of Moments (MOM) and Maximum Likelihood Estimation (MLE) are extremely effected in the presence of outlier. For example, by using these three methods of estimation to estimate the mean of a normal population by the sample mean $\bar{X}$ which is unique minimum variance unbiased estimator, however is not robust to outlier and exits from normality. Hence, as the outlier observation has extreme effect on these methods, we need a strong method of estimation, to reduce the effect of outliers on the estimates.

When we found outliers in our data set, commonly we reject the observation that are greater than some number times of the sample standard deviation away from sample mean. Therefore, in the existence of outliers the sample standard deviation will be completely overestimated and the method will lack power.

Hosking (1990) introduced L-moments that are summary statistics for probability distributions and data samples. They are similar to conventional moments, they deliver measures of location, dispersion, skewness, kurtosis, and further characteristics of the shape of probability distributions or data samples, and however they are calculated through the linear arrangement of the sample ordered statistics. Since the Sample L-moments are unbiased estimates of population L-moments. L-moments take the following general favor over usual moments:

- L-moments exist for any probability distribution that has specific mean, no requirement of high order moments to be finite.
- We require only that the distribution have finite variance for standard errors of L-moments to be finite, no need of higher-order moments to be finite.
- Kirby (1974) and Dalén, J. (1987) noted that when conventional ratio diagram is calculated for finite samples they are bound and cannot get the full array of values available to the population skewness and kurtosis. While Hosking (1986) show that the sample L-moment

ratios (t_3, t_4) computed from a finite sample can be take all of the possible values of the population L-moment ratios (τ_3, τ_4).

- L-moments compared with conventional moments are better for asymptotic approximation to the sampling distributions.
- As L-moments are linear moments of the ordered statistic, do not arise data to powers of 2, 3, and 4 as necessary for variance, skewness and kurtosis respectively, and therefore superior parameter estimates for data having outlying values. Royston & Thompson (1992) and Vogel & Fennessey (1993).
- L-moments provide superior identification of the parent distribution that produced a specific data sample

TL-moments are derived by Elamir and Seheult in 2003 from L-moments and have additional robust properties compared to L-moments. In other words, TL-moments are claimed to be more robust than the L-moment. The original L-moments use all of the order statistics. Trimmed L-moments with degree of trimming (t_1, t_2) do not use the t_1 smallest and the t_2 largest order statistics. However, with reference to Cunnane (1987), Wang (1990) and Shabri *et al* (2011) censoring the data from lower side might be useful as the small peak flows are less significant to large ones, therefore the use of small peak flows can sometimes be only a nuisance value in the estimation of large return period event and also in model formation testing and verification. This option may also be useful in the presence of lower outliers where some undesirable effects are likely to occur especially in arid or semiarid areas. Based on this idea, in this study we use trimmed L-moments with degree of trimming (1,0) which mean do not use one smallest order statistic. But it mean that this smallest one value does not contribute to the (1,0)-trimmed sample L-moment, because this smallest data value is removed from every subsample that includes it. Hence, for extreme data, TL-moments are also considered for estimating the parameters of the selected probability distributions.

1.4 Objectives of the Study

- To estimate L-moments and TL-moments for these selected distributions e.g. Generalized Logistic (GLO), Generalized Extreme Value (GEV) and Generalized Pareto (GPA) Distributions for AM flows. These have been commonly used for flood estimation in different countries.
- To find the most suitable distribution among three commonly used distributions for AM flows data by using the Mean Absolute Deviation Index (MADI), Relative Root Mean Square Error (RRMSE), Probability Plot Correlation Coefficient (PPCC), L-moment ratio and TL-moment ratio diagrams.
- To compare the results obtained from TL-moments and L-moments for best fitted distributions.
- To find the quantile estimates for best fitted distribution with different return periods.
- To identify the sites likely to receive maximum or minimum flows in future.
- To give some solutions to mitigate these extreme events for policy implications in the country.

CHAPTER 2

Literature Review

Literature review is based on finding concepts and collecting data from other's research works, which is done on the basis of certain theories and approaches. These findings may be theoretical and empirical studies, which have been conducted in a particular region or area. Researchers of different countries including USA, UK, Malaysia, India, Turkey, and China worked on FFA of AM flows. To build up this study, we reviewed the work of different researchers. In present chapter, we included the references, starting from 1987 to 2014, which explain the methodology of FFA based on the methods of L-moments and TL-moments.

2.1 Review of Methodology

Tasker (1987) applied frequency analysis based on different methods of estimation to annual 7-days minimum flow series across 20 rivers of Virginia. Various distributions were fitted and it was concluded that Weibull type III distribution (W3) and Log Pearson type III (LP3) distribution were the most suitable probability distributions for mentioned data.

Karim and Chowdhury (1995) performed FFA based on L-moment to annual maximum discharge using data across 31 gauging sites of Bangladesh. Four distributions namely Lognormal (LN), Gumbel (GU), LP3 and GEV were used in the study. The performance of distributions were compared by using Root Mean Error Deviation (RMED), PPCC and L-moments ratio diagrams. The results indicated that GEV distribution was the best fit for the modeling of annual peak discharges in Bangladesh.

Vogel and Wilson (1996) applied different probability distributions to annual maximum flood flows, annual mean daily stream flows and annual minimum low flows across United States.

They used the method of L-moment for estimation of parameters and L-moment ratio diagram for goodness-of-fit.

Önöz and Bayazit (1999) applied various probability distributions to low flows across 16 European rivers and their performances were compared by PPCC test. The results concluded that GEV distribution was best fit for these rivers.

Rao and Hamed (2000) analyzed the basic methods of FFA by applying Method of L-Moment, MOM, MLE and Method of Probability Weighted Moments (PWM).

Zaidman *et al* (2002) applied low FA based on L-moments to 25 UK Rivers. GLO, GEV, Pearson type III (P3) and GPA distributions were used in that study. The performance of distributions was compared by Root Mean Square Error (RMSE).

Kroll and Vogel (2002) applied certain probability distribution to 1-day, 7-days and 30-days annual minimum stream flow data of Hydro-Climate Data Network (HCDN) river sites in United States. The parameter of the distributions were estimated by L-moment method. The performance of distribution was compared by L-ratio diagram.

Lee and Maeng (2003) fitted GLO, GEV and GPA distributions based on LH-moment to annual maximum flood of six watersheds in Korea. The suitability of distributions were compared by using Kolmogorov-Smirnov (K-S) test and LH-ratio diagrams.

Kjeldsen and Jones (2004) recommended L-moment for the estimation of FFA across UK using GLO distribution. They calculated quantile estimates for GLO distribution with different return periods.

Bhattara (2005) performed FFA based on three variants of L-moments, namely the simple L-moments, the LH-moments and the Partial L-moments using river flow data of Ireland. According to their finding, GEV distribution was fitted to AM flows and then the results were compared. The results showed that the methods of LH-moment and partial L-moment were best for FFA at Irish flood data.

Yurekli *et al* (2005) carried out FA of low flow series using data over Cekerek stream Basin in Turkey. The aim of study was to derive suitable probability distribution for the frequency analysis of 7-days annual low flows data at three gauging stations of Cekerek Stream Basin. Several distributions were used in the study for fitting seven days annual low flows data including two parameters Normal (N2), Lognormal (LN2) distributions and three parameters LN, Logistic (L3), GLO, Extreme Value Type 1(EV1), GEV, GPA and LP3 distributions. L-moments were used for estimation of parameters of these distributions. MADI and Mean Square Deviation Index (MSDI) were used to compare the distributions. The results based on MADI and MSDI showed that GPA distribution was the best fitted distribution for 7-days annual low flows data at three gauging stations of Cekerek Stream.

Abida and Ellouze (2008) applied FFA based on L-moment to AM flows data of 42 stations across different zones of Tunisia. In their research five distributions namely, GLO, Generalized Normal (GNO), GEV, GPA and P3 were fitted to AM flows and L-moment ratio diagram was used as goodness-of-fits test. The results showed that GNO distribution was best fitted to Northern Tunisia while GEV distribution was the most appropriate in central/southern Tunisia.

Zin *et al* (2009) used L-moment and LQ-moment to analyze annual maximum rainfall for 50 gauge stations in Peninsular Malaysia. Three parameters distributions were fitted to annual maximum rainfall including GLO, GNO, GEV, GPA and P3 distributions. The performances of these distributions were compared by goodness-of-fits test. The results showed that GLO distribution is best fitted for the most of the stations. Further it has found that L-moments performances on the average were better than LQ-moments.

Deka *et al* (2009) used L-moment and LQ-moment approaches to analyze the annual maximum daily rainfall of 9 stations in north east India. The extreme value distributions were fitted, such as GLO, GEV, GPA, LN and P3. The distributions parameters were estimated by L-moment and LQ-moment. Three goodness-of-fits tests Relative Root Mean Square (RRMS), Relative Mean Absolute Error (RMAE) and PPCC and also L-moment ratio diagrams were used for the evaluation of the five probability distribution. The goodness-of-fits tests results indicated that GLO was the most suitable distribution for annual maximum daily rainfall data.

Shabri and Ariff (2009) applied FA based on L-moments to maximum daily rainfall data of stations in Selangor and Kula Lumpur in Malaysia. Several distributions were used in their research containing two parameters N2, LN2 distributions and three parameters LN, L3, GLO, EV1, GEV and GPA distributions. MADI, MSDI and L-moment ratio diagram was used to measure the performance of the distributions. The results indicated that GLO distribution was the most appropriate distribution for the data of maximum daily rainfalls to stations in Selangor and Kula Lumpur in Malaysia.

Gubareva and Gartsman (2010) used L-moments and MLE to estimate the parameters of different distributions that are used in Hydrology and Meteorology namely GEV, GPA, P3 and LN distributions. Their work showed that L-moments method gives more stable results as compared to MLE method.

Ahmad *et al* (2011) performed FFA of the AM flows using data over stations in Negeri Sembilan Malaysia. Three selected distribution GLO, GEV and GPA distributions were used in their research. The estimations of these distributions were performed by using the methods of L-moments and TL-moment with trimming (1, 0). The best fitted distribution was determined by the use of MADI, L- moment ratios and TL-moment ratios diagrams. The

S) tests. According to AD test results indicated that GPA distribution was best for AM flows, Galoie *et al* (2013) used L-moment to examine the best fit probability distribution for rainfall data in Schoeckelbach basin across the northern Graz in Austria. GEV, GU, LP3 and 3-parameter LN distributions were fitted and their performances were compared by using goodness-of-fit tests such as C-S, K-S and RMSE. Goodness-of -fit tests showed GU is the best fitted distribution which furtherly used for forecasting of rainfall data.

Ahmad *et al* (2013) applied Kappa probability distributions based on L-moment to monsoon rainfall in Pakistan at 27 meteorological stations from to 1960-2006. Using these estimates they calculated quantiles for different return periods from 2, 5, 10, 20, 50,100, 200 and 500 years.

Izinyon and Ehiorobo (2014) used the FFA based on L-moments method to analyze AM flow series data across site in Benin Owena River in Nigeria. The three selected distributions GLO, GEV and GPA were used, Parameters were estimated by using the method of L-moments. The three distributions were compared by using four goodness-of -fit tests RMSE, RRMSE, MADI and PPCC. The results showed that GPA is the best fitted distribution for examining the AM series at site study for return periods.

Gocic *et al* (2014) used L-moment to analyze the annual monthly precipitation data of 29 stations across Serbia. Three distributions GLO, GEV and GPA were fitted to estimate the data. Further goodness-of-fits tests namely RRMSE, RMAE and PPCC and also L-moment ratio diagrams were used to confirm the best fit distribution. The results showed that GEV is the most suitable distribution for annual precipitation data in Serbia, also return levels were calculated for the GEV distribution.

Bílková (2014) used different methods of estimation including TL-moments, L-moments and MLE. In his research exponential (E), GU, L2, N2, GPA, GEV, GLO, LN and G distributions were fitted to wages data. Results of almost all cases showed that Method of TL-moments provided the most accurate results compared to Method of L-moments and MLE. Method of L-moment as the second best choice and so on.

Proposed Work:

- To extend Ahmad *et al* (2011) by using data of 18 sites of Pakistan.
- To use MADI as used by Ahmad *et al* (2011) augmented by RRMSE, PPCC, L-moment ratio and TL-moment ratio Diagrams as goodness-of- fits for FFA of AM flows for 18 sites of Pakistan using L-Moments and TL-Moments.
- To apply Trimmed L-Moments for the first time in Pakistan.
- To compute quantile estimates for different return periods by using Trimmed L-Moments for the first time in Pakistan.

CHAPTER 3

Materials and Methods

3.1 Study Area and Data

The AM flows data for 18 sites of Pakistan, located on the five rivers namely Indus, Jhelum, Chenab, Ravi, and Sutlej have been used in this study. The geographical location of these river's sites are given in fig. 3.1. The data of these sites have been collected from the Hydrology Department of Water and Power Development Authority (WAPDA) Lahore and Federal Flood Commission (FFC) Islamabad. The AM flows at given sites have been recorded in the months of monsoon (from July to September).

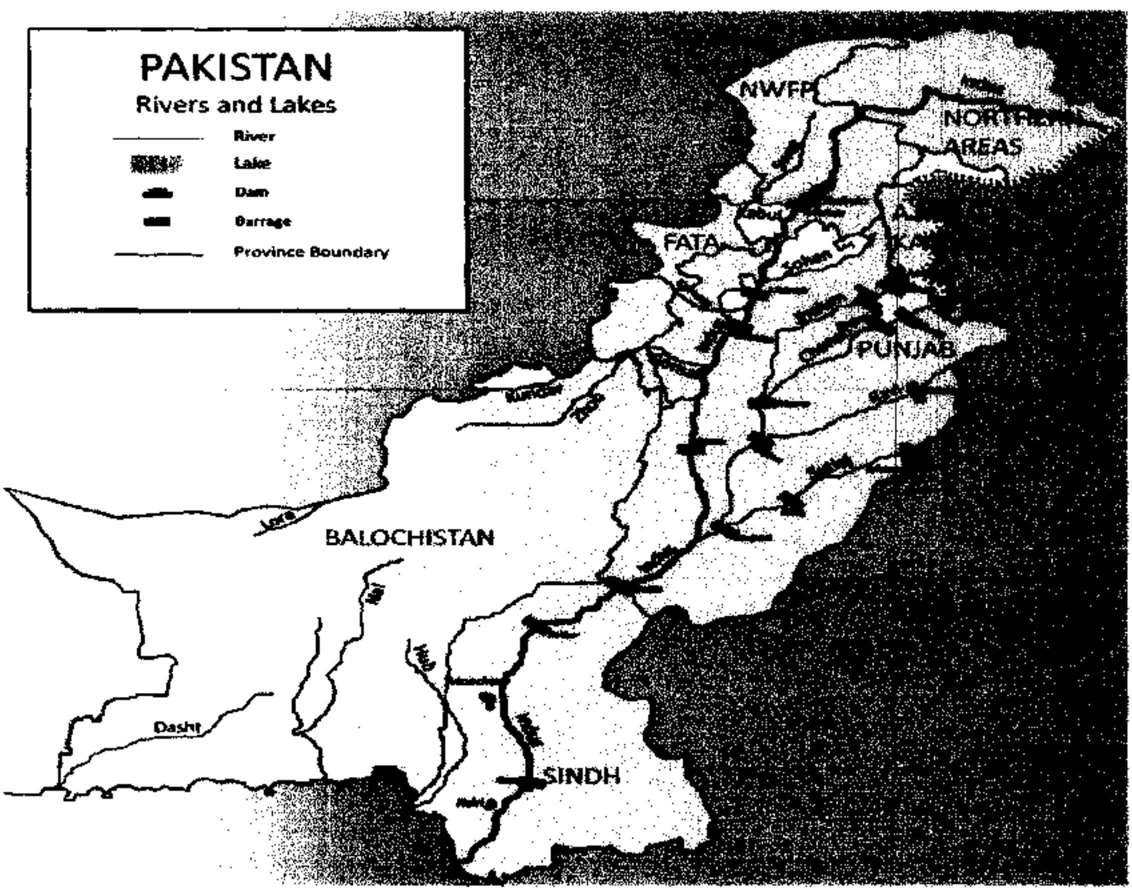


Fig. 3.1: Geographical Locations of the 18 Sites of Pakistan used in the study

Himalaya and Karakorum is the origination of Pakistani rivers system. There are five rivers, Chenab, Indus, Sutlej, Ravi, and Jhelum. They are mostly pass through Punjab province. The word 'panj' means five and 'aab' means water (in Persian language) that's why Punjab is called the land of five rivers. The irrigation system of Pakistan is the widest system of irrigation in the world. This system has a great importance in the agriculture field of the country. These river's water is the biggest source of water in Pakistan.

Indus River is originating from Himalayan region (Tibetan plateau near Lake Mansarovar) in China. In Asia the Indus River is a major river which flows through Pakistan and also the longest river in Pakistan. Length of Indus River is 3180 Kilometers. Balram River, Gilgit River, Tanubal River, Astor River, Kabul River, and Zanskar River are the tributaries of Indus River. Indus River consist of 16 Barrages, 3 reservoirs, 2 Siphons across major river 12 inter link canals, 2 head works, 44 canal system: 23 in Punjab, 14 in Sindh, 5 in KPK, and 2 in Baluchistan.

River Jhelum originates from the south-eastern part of valley Kashmir. River Jhelum is the tributary of River Chenab and 774 kilometer long. Mangla is one of the world largest dam which is constructed on Jhelum in 1967. The storage capacity of Mangla dam is 5.9 million acre-feet. Many other dams and Barrages are also constructed on Jhelum like Rasul Barrage and Trimmu Barrage.

Chenab River is joined by Chandra and Bhaga rivers in the upper Himalayas. It flows through Jammu and Kashmir and at Trimmu Barrage Chenab rivers joined by the Jhelum River. Chenab River is 960 Kilometers long. Under the Indus waters treaty the water of Chenab is allocated to Pakistan and after Indus this is the strongest river of Pakistan.

The Ravi River also known as Parushani or Iravati, Hydraotes in Indian Vedic, and an ancient Greeks. River Ravi originates from Himalayas near Chamba in Himachal Pradesh State, northern India. It flows in Indian Punjab south-west region and flows beside with indo-Pak border and enters into Pakistan and merges with Chenab. The length of River Ravi is nearly 720 kilometers long and in Pakistan its length is 675 kilometer. On Ravi River Important Engineering's head-works are constructed e.g. Balloki and Sidhani.

Sutlej River is also known as Red River and it is 550 kilometers long. It flows through northern Punjab area of India and Pakistan. The location of Sutlej is east of the central Suleiman range in Pakistan, north of the Vindhya Range, and south of the Hindu Kush segment of the Himalayas. On Sutlej Important Engineering Barrages are constructed e.g. Islam Barrage and Sulemanki Barrage.

3.2 Basic Assumptions of Flood Frequency Analysis (FFA)

There are various basic assumptions which are necessary in FFA. The purpose of these assumptions is to test observations at each site for stationary, independence and identical distribution. These assumptions are checked for different kinds of data practically for AM flows, rainfalls, droughts etc. The AM flows as used in our study different plots and tests are applied which are given below.

3.2.1 Time Series Plots

The time series plots are important tool of statistics which are used to observe the pattern of variability in a time series data. The graphical inspection is always useful to provide an initial clue about the likely nature of the series. When the data of the same variable over a long period of time is recorded, then it is hard to determine any trend or pattern. However, the graphically display of the same data points make easier to spot trends. The trends are very significant as they can be used to plan into the future.

3.2.2 Ljung-Box Q-Statistics (LBQ)

The LBQ –statistic which is developed by Ljung-Box (1978). It is used to check whether a sequence of observations over time is random/independent or not. If the sequence of observations is not random/ independent then one observation can be connected with other observation k time units later, a relationship called autocorrelation. The autocorrelation can also reduce the accuracy of a time-series model, such as time series plot which can lead to misunderstanding of the data.

The LBQ statistics is a modified version of the Box-Pierce Q statistics, which is defined as.

$$LB = n(n+2)\sum_{k=1}^{m}\left[\frac{\hat{\rho}_k^2}{n-k}\right] \sim \chi_m^2 \qquad (3.2.2.1)$$

Where n is sample size, m is lag length and $\hat{\rho}_k^2$ is sample autocorrelation at lag k. if $LB > \chi_{1-\alpha,m}^2$ then we can reject the null hypothesis at chosen level of significance .For large sample size both Q and LBQ statistics follow chi-square distribution with m degree of freedom. But for small and large sample LBQ statistics is found better than Q statistics.

3.2.3 Mann-Whitney U Test (M-W)

The Mann-Whitney (M-W) U test (1947) a non-parametric test is used to check the null hypothesis whether the two samples are same. Let P and Q represent the two sample respectively and $P \leq Q$. Then total of these samples size is $N=P+Q$, ranked in ascending order. The (M-W) U test considers the measures W and V in the following equations.

$$W = PQ - V \tag{3.2.1.1}$$

$$V = R - \left\{P\left(\frac{P+1}{2}\right)\right\} \tag{3.2.1.2}$$

In the above equation R is the total of the ranks of the observations for first sample size P in the combined series (N). The two measures W and V are calculated from R, P and Q. V denotes the number of times observation in sample 1 followed by observation in sample 2 in the ranking. Also, W can be calculated for sample 2 following sample 1.

For small sample size we use M-W statistic U as defined by V and W. But when $N > 20$ and P, $Q > 3$, then U follows the normal distribution with mean $\bar{U} = \frac{PQ}{2}$ and variance $Var(U)$ given below under the null hypothesis that the two samples came from the same population.

$$Var(U) = \left[\frac{PQ}{N(N-1)}\right]\left[\frac{N^3 - N}{12} - \sum T\right] \tag{3.2.1.3}$$

$$T = \left(\frac{J^3 - J}{12}\right) \tag{3.2.1.4}$$

Here J is the number of tied observation at a given rank and T is total overall of groups tied elements in both sample sizes P and Q. The U statistic is used to test the null hypothesis of homogeneity at α level of significance.

$$u = \left[\frac{U - \bar{U}}{\sqrt{var(U)}}\right] \tag{3.2.1.5}$$

3.2.4 Mann-Kendall Trend Test (MK)

The MK test used to check the monotonic pattern of increasing or decreasing trend of the variable of interest over time Mann (1945), Kendall (1975) & Gilbert (1987). The MK test is a non-parametric test and it is a special case of Kendall's tau test. The MK statistic gives a clue of whether a trend occurs, (is it significant?) and whether the trend is positive or negative. Hamed (2008) show that MK test is robust to the influence of extremes and suitable for application with skewed variables. The MK test is rank-based procedure for time series $X = (x_1, x_2, x_3, \dots, x_n)$ the statistic test "*S*" as given by.

$$S = \sum_{i<j} a_{ij} \qquad (3.2.4.1)$$

Where

$$a_{ij} = sign(x_j - x_i) = sign(R_j - R_i) = \begin{cases} 1 & x_i < x_j \\ 0 & x_i = x_j \\ -1 & x_i > x_j \end{cases} \qquad (3.2.4.2)$$

In above equation R_j and R_i are the ranks of the observations x_j and x_i of the time series respectively. From equation (3.2.4.2) as we can see that the test statistics only depends on the ranks of the observations. Kendall (1975) give the mean and variance of the *S* Statistic in equation (3.2.4.1) under the assumption of identical independent distributed (IID).

$$E(S) = 0 \qquad (3.2.4.3)$$

$$V_0(S) = \frac{n(n-1)(2n+5)}{18} \qquad (3.2.4.4)$$

In above equation "*n*" is the sample size. In the case of tied ranks in the data the variance of *S* becomes.

$$V_0^*(S) = \frac{n(n-1)(2n+5) - \sum_{j=1}^{m} t_j(t_j-1)(2t_j+5)}{18} \tag{3.2.4.5}$$

M represents the number of groups of tied ranks and t_j shows tied observation. As the number of observation become large, the *S* statistic follows normal distribution. For sample size greater than *10*, Kendall (1975) shows that *S* statistic follows normal distribution and can be calculated as.

$$U = \begin{cases} \dfrac{(S-1)}{\sqrt{V_0\ (S)}} & S > 0 \\ 0 & S = 0 \\ \dfrac{(S+1)}{\sqrt{V_0\ (S)}} & S < 0 \end{cases} \tag{3.2.4.6}$$

Now trend can be tested by normal variate at the chosen significance level α. In above equation the subtracting or adding of unity is a continuity correction.

3.2.5 Change Point Analysis (CPA)

In historical process data we try to answer the following questions. Is there any change occurs? Is there more than one changes occur? If the changes occurred then when? The changes occurred with which confidences? To answer all these question we apply CPA. CPA is an effective tool to recognize changes in historical data. The aim of CPA is to detect any changes in the mean of process. It is designed to use for long period of historical data Taylor (2011). The CPA assumes that time series (data) are IID. The properties of CPA are that, no specific distribution is assumed, no need that data should be normally distributed. It can handle all type of time series data. It will give results that are robust to outlier when CPA is applied on the ranks. The CPA can identify smaller continuous changes which cannot identify by control charts. CPA also provides confidence level and confidence interval for the identification of changes.

The CPA is combination of two powerful methods: Cumulative sum charts (CUSUM) and bootstrapping to identify changes Taylor (2000a). The CPA is procedure starts with the construction of CUSUM charts. The construction of CUSUM charts are based on the data by calculating and plotting a cumulative sum. Let $X_1, X_2, X_3, \ldots, X_N$ be the time series data and the cumulative sums are $S_0, S_1, S_2, \ldots, S_N$ calculated for this. The following steps are adopted:

1. calculate the average

$$\bar{X} = \frac{X_1 + X_2 + X_3 \ldots + X_N}{N} \tag{3.2.5.1}$$

2. Begin with the cumulative sum at zero setting by$S_0 = 0$.
3. We determine the further cumulative sums by adding the difference between present value and the average to the previous sum, i.e.,

$$S_j = S_{j-1} + (X_j - \bar{X}) \qquad j = 1,2,3 \ldots, N. \tag{3.2.5.2}$$

4. The CUSUM we calculate the estimators of the magnitude of changes is required.

$$S_{diff} = S_{max} - S_{min} \tag{3.2.5.3}$$

Where

$$S_{max} = max_{j=1,2,..,N}\ S_j \qquad S_{min} = min_{j=1,2,..,N}\ S_j$$

Following this, now we are performing the bootstrap analysis to fix the estimator of change. A single bootstrap is performed by the following steps:

1. First generate N units of bootstrap sample and denote by $X_1^0, X_2^0, \ldots, X_N^0$. then randomly rearranging to the N original data values.
2. Next calculated bootstrap CUSUM from the bootstrap sample and denoted by $S_0^0, S_1^0, \ldots, S_N^0$.
3. The difference of the bootstrap CUSUM is computed by taking maximum, minimum and denoted by S_{max}^0, S_{min}^0 and S_{diff}^0.
4. In the final step check if the bootstrap difference S_{diff}^0 is less or more than the original difference S_{diff}.

   The idea behind bootstrapping is to check the behavior of CUSUM for all random bootstrap samples. If no change has occurred than how much S_{diff} vary from the bootstrap sample. The bootstrap CUSUM chart can also be compared with the original CUSUM. Further we calculate the confidence level for bootstrap sample. Let N be the number of bootstrap sample and X be the number of bootstraps for which $S_{diff}^0 < S_{diff}$. Then percentage change occurred is given as.

$$Confidence\ Level = 100 \times \frac{X}{N} \qquad (3.2.5.4)$$

3.3 Selection of Parameter's Estimation Methods

For the estimation of parameters, there are several techniques including the MLE, MOM, L-moments, TL-moments, LS, Maximum Entropy (MENT), Generalized Method of Moments (GMM), Mixed Moments (MIXM), and Incomplete Means (ICM) Rao *et al* (2000).

The MLE is generally considered as the most efficient technique, because it provides the minimum sampling variance of the estimated parameters as compared to other techniques. But in the presence of outlier, small sample size and large number of parameter MLE gives inappropriate results and also biased estimates.

For the parameter estimation another method is MOM which is relatively easy for computation. The MOM is not efficient as compared to MLE method in the case of large number of parameters of the distributions. In the presence of small sample size the higher order moments are more likely to be extremely biased.

Hence, as the outlier observation and small sample size has extreme effect on these methods, we need a robust method of estimation to reduce the effect of outliers on the estimates.

The L-moment method Hosking (1986) are more robust in the presences of outlier observation and small sample size as compared to other methods. The TL-moment method is modified version of L-moments Elamir & Seheult (2003) is the most appropriate for the outlier and small sample size as compare to L-moment and other methods of estimation. However our data contain extreme observation and small sample size therefore we use the L-moment and TL-moment methods for estimation.

3.3.1 Method of Linear-Moments (L-Moments)

As L-moment have been defined by Hosking (1990), L-moments are expectations of certain linear combination/arrangements of order statistic. They can be explained for any random variable that has finite mean. Let $X_1, X_2, \ldots, X_r$ be the random sample of magnitude r, with cumulative distribution Function $F(X)$ and quantile function $X(F)$. Let $X_{1:r} \leq X_{2:r} \leq \ldots \leq X_{r:r}$ be the order statistic of random sample. For the random variable X, the rth population L-moment as explained by Hosking (1990) is:

$$\lambda_r = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} \mathrm{E}(X_{r-k:r}) \qquad r = 1, 2 \ldots \qquad (3.3.1.1)$$

In L-moments L highlight that λ_r is a linear function of the expected order statistics. The first four L-moments are given by:

$$\lambda_1 = E(X_{1:1}) \qquad (3.3.1.2)$$

$$\lambda_2 = \frac{1}{2}\mathrm{E}(X_{2:2} - X_{1:2}) \qquad (3.3.1..3)$$

$$\lambda_3 = \frac{1}{3}E(X_{3:3} - 2X_{2:3} + X_{1:3}) \qquad (3.3.1.4)$$

$$\lambda_4 = \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} + X_{1:4}) \qquad (3.3.1.5)$$

The L-moments ratio has defined as:

$$\tau = {\lambda_2}/{\lambda_1} \qquad (3.3.1.6)$$

$$\tau_3 = {\lambda_3}/{\lambda_2} \qquad (3.3.1.7)$$

$$\tau_4 = {\lambda_4}/{\lambda_2} \qquad (3.3.1.8)$$

In the above equation λ_1 is the measure of location, τ is measure of Linear-coefficient of variation (L-CV), τ_3 and τ_4 are Linear- Skewness (L-Skewness) and Linear-Kurtosis (L-Kurtosis) respectively.

3.3.2 Estimation of Linear-Moments (L-Moments)

In practice, L-moments need commonly be estimated after a random sample drawn from an anonymous distribution. As λ_r is a meaning of the expected order statistics of a sample of magnitude r. Let $x_1, x_2, \ldots, x_n$ be the sample and $x_{1:n} \leq x_{2:n} \leq \cdots \leq x_{n:n}$ is the order statistics of the samples, then we can define the rth sample L-moments as by Asquith (2007).

$$l_r = \frac{1}{r}\sum_{i=1}^{n}\left[\sum_{j=0}^{r-1}\frac{(-1)^j\binom{r-1}{j}\binom{i-1}{r-1-j}\binom{n-1}{j}}{\binom{n}{r}}\right]X_{i:n} \qquad r = 1, 2 \ldots \tag{3.3.2.1}$$

By using Wang (1996) direct estimators method of L-moments the first four sample L-moments are defined as:

$$l_1 = \sum_{i=1}^{n}\left[\frac{x_{i:n}}{\binom{n}{1}}\right] \tag{3.3.2.2}$$

$$l_2 = \frac{1}{2}\sum_{i=1}^{n}\left[\frac{\binom{i-1}{1} - \binom{n-i}{1}}{\binom{n}{2}}\right]x_{i:n} \tag{3.3.2.3}$$

$$l_3 = \frac{1}{3}\sum_{i=1}^{n}\left[\frac{\binom{i-1}{2} - 2\binom{i-1}{1}\binom{n-i}{1} + \binom{n-i}{2}}{\binom{n}{3}}\right]x_{i:n} \tag{3.3.2.4}$$

$$l_4 = \frac{1}{4}\sum_{i=1}^{n}\left[\frac{\binom{i-1}{3} - 3\binom{i-1}{2}\binom{n-i}{1} + 3\binom{i-1}{1}\binom{n-i}{2} - \binom{n-i}{3}}{\binom{n}{4}}\right]x_{i:n} \tag{3.3.2.5}$$

The sample L-ratios are defined as:

$$t = \frac{\ell_2}{\ell_1} \tag{3.3.2.6}$$

$$t_3 = \frac{\ell_3}{\ell_2} \tag{3.3.2.7}$$

$$t_4 = \frac{\ell_4}{\ell_2} \tag{3.3.2.8}$$

l_1 is the measure of location, t is L-CV, t_3 and t_4 are the samples L- Skewness and L-Kurtosis respectively. In summary statistics these measures l_1,l_2, t, t_3 and t_4 are very useful for observed data. These measures are used to classify the distribution from which sample is drawn. Furthermore these measure are used in the estimation of parameters when distribution is fitted to sample data by equating the sample L-moments and population L-moments.

3.3.3 Method of Trimmed Linear-Moments (TL-Moments)

TL-Moments are defined by the same way as L-moments. Assume that $X_1, X_2, \ldots, X_r$ are the random sample of size r from continues distribution with quantile function $Q(F)$ and let $X_{1:r} \leq X_{2:r} \leq \ldots \leq X_{r:r}$ indicates the order statistics. The TL-moments are defined by Elamir & Scheult (2003), in the above equation (3.3.1.1) the expectation term $E(Xr - k{:}r)$ is substituted by $E(Xr + t_1 - k{:}r + t_1 + t_2)$. Therefore, for every r, we increased the theoretical sample size from r to $(r + t_1 + t_2)$ and work simply with the expectation of the r order statistic $X_{t_1+1:r+t_1+t_2}, \ldots, X_{t_1+r:r+t_1+t_2}$ by means of trimming t_1 the lowest and t_2 the biggest from the random sample. The rth TL-moment is denoted by $\lambda_r^{(t_1,t_2)}$ and defined as:

$$\lambda_r^{(t_1,t_2)} = \frac{1}{r}\sum_{k=0}^{r-1}(-1)^k \binom{r-1}{k} E\left(X_{r+t_1-k:r+t_1+t_2}\right) \qquad r = 1,2,\ldots \tag{3.3.3.1}$$

The connection among L-moments and TL-moments is that when $(t_1 = t_2 = 0)$ then TL-moment reduces to L-moment. In this study, we use unequal trimming value where just only one lowest value will be trimmed from our random sample ($t_1 = 1, t_2 = 0$). The first four TL-moments when ($t_1 = 1, t_2 = 0$) are given by:

$$\lambda_1^{(1,0)} = E(X_{2:2}) \quad (3.3.3.2)$$

$$\lambda_2^{(1,0)} = \frac{1}{2} E(X_{3:3} - X_{2:3}) \quad (3.3.3.3)$$

$$\lambda_3^{(1,0)} = \frac{1}{3} E(X_{4:4} - 2X_{3:4} + X_{2:4}) \quad (3.3.3.4)$$

$$\lambda_4^{(1,0)} = \frac{1}{4} E(X_{5:5} - 3X_{4:5} + 3X_{3:5} - X_{2:5}) \quad (3.3.3.5)$$

The population TL-ratio are defined as:

$$\tau^{(1,0)} = \frac{\lambda_2^{(1,0)}}{\lambda_1^{(1,0)}} \quad (3.3.3.6)$$

$$\tau_3^{(1,0)} = \frac{\lambda_3^{(1,0)}}{\lambda_2^{(1,0)}} \quad (3.3.3.7)$$

$$\tau_4^{(1,0)} = \frac{\lambda_4^{(1,0)}}{\lambda_2^{(1,0)}} \quad (3.3.3.8)$$

$\lambda_1^{(1,0)}$ is the measure of location, $\tau^{(1,0)}$ is measure of Trimmed Linear-Coefficient of Variation (TL-CV), $\tau_3^{(1,0)}$ and $\tau_4^{(1,0)}$ are Trimmed Linear- Skewness (TL-Skewness) and Trimmed Linear-Kurtosis (TL-Kurtosis) respectively. It must be noted that TL- Skewness and TL-Kurtosis are location and scale invariant, and exist for Cauchy distribution. Elamir (2010) provided two methods for the selection of appropriate trimming, quantile function approach and simulation approach.

In quantile function approach the error between quantile function and its TL-moments are calculated, further to choose optimal trimming values of t_1 and t_2 with less sum of absolute error. We use two methods for choosing the value of absolute error are, exact method and approximation method.

In the simulation approach select the known distribution and find its TL-moments estimators functions. Next from the distribution with knowing parameters simulate data. The sample size of the data should be large at least 100 and the large number of replication at least 10000. Select the approximate optimal trimming with is smaller variance.

3.3.4 Estimation of Trimmed Linear-Moments (TL-Moments)

We study the estimators of population TL-moments that are the linear combinations of the order statistics $x_{1:n} \leq x_{2:n} \dots \leq x_{n:n}$ of the random sample $x_1, x_2, \dots, x_n$ of size n from the population. These sample TL-moments are the unbiased estimator of population TL-moments. Hosking (2007) define the rth sample TL-moments as:

$$l_r^{(t_1,t_2)} = \frac{1}{r} \sum_{j=t_{1+1}}^{n-t_2} \left[\sum_{k=0}^{r-1} \frac{(-1)^k \binom{r-1}{k} \binom{j-1}{r+t_1-k-1} \binom{n-j}{t_2+k}}{\binom{n}{r+t_1+t_2}} \right] x_{j:n} \qquad (3.3.4.1)$$

As in this study we use trimming of ($t_1 = 1, t_2 = 0$), In this case we can write the above as given below.

$$l_r^{(1,0)} = \frac{1}{r}\sum_{j=2}^{n}\left[\sum_{k=0}^{r-1} \frac{(-1)^k \binom{r-1}{k}\binom{j-1}{r-k}\binom{n-j}{k}}{\binom{n}{r+1}}\right] x_{j:n} \tag{3.3.4.2}$$

The first four sample TL-moments can be defined as:

$$l_1^{(1,0)} = \sum_{j=2}^{n}\left[\frac{\binom{j-1}{1}}{\binom{n}{2}}\right] x_{j:n} \tag{3.3.4.3}$$

$$l_2^{(1,0)} = \frac{1}{2}\sum_{j=2}^{n}\left[\frac{\binom{j-1}{2} - \binom{j-1}{1}\binom{n-j}{1}}{\binom{n}{3}}\right] x_{j:n} \tag{3.3.4.4}$$

$$l_3^{(1,0)} = \frac{1}{3}\sum_{j=2}^{n}\left[\frac{\binom{j-1}{3} - 2\binom{j-i}{2}\binom{n-j}{1} + \binom{j-1}{1}\binom{n-j}{2}}{\binom{n}{4}}\right] x_{j:n} \tag{3.3.4.5}$$

$$l_4^{(1,0)} = \frac{1}{4}\sum_{j=2}^{n}\left[\frac{\binom{j-1}{4} - 3\binom{j-1}{3}\binom{n-j}{1} + 3\binom{j-1}{2}\binom{n-j}{2} - \binom{j-1}{1}\binom{n-j}{3}}{\binom{n}{5}}\right] x_{j:n} \tag{3.3.4.6}$$

The sample TL-skewness and TL-kurtosis are calculated as:

$$t_3^{(1,0)} = \frac{l_3^{(1,0)}}{l_2^{(1,0)}} \tag{3.3.4.7}$$

$$t_4^{(1,0)} = \frac{l_4^{(1,0)}}{l_2^{(1,0)}} \quad (3.3.4.8)$$

The TL-skewness and TL-kurtosis are dimensionless compute of the shape of a data set, and also the reliable estimators of the population TL-moments.

3.4 Selection of Probability Distributions

Probability distribution are widely used in hydrological studies, studies of Floods and droughts, rainfall quantities, and reservoir volume. For FFA many probability models have been proposed due to which scientists carried considerable research work throughout the world. Below table show the previous study of probability models in hydrology based on methods of L-moments and TL-moments.

Table 3.4 Previous flood frequency Model based L-moment and TL-moment.

Location	Number of stations	Recommended distribution	Reference
Eastern United States	55	GEV	Wallis (1988)
Central Victoria, Australia	53	GEV	Nathan and Weinmann (1991)
South Island , New Zealand	275	EV1,EV2, and GEV	Pearson (1992)
New Brunswick, Canada	53	GEV	Gingras and Adamowski (1992)
Nova Scotia, Canada	25	GEV	Pilon and Adamowski (1992)
Australia	61	GEV and GPA	Vogel et al .(1993)
Bangladesh	31	GEV	Karim and Chowdhury (1993)
The world	19	GEV	Onoz and Bayazit (1995)
Negeri Sembilan, Malaysia	9	GLO,GPA and GEV	Ahmad et al.(2011)
North Brahmaputra region, India	10	GEV and GPA	Deka and Borah (2012)
Benin Owena river basin, Nigeria	1	GPA	Izinyon and Ehiorobo (2014)

We observed that GLO distribution, GEV distribution and GPA distribution has been used in all over the world for FFA. However, based on previous studies we will use these models for FFA in Pakistan.

3.4.1 Generalized Logistic Distribution (GLO)

In recent years GLO distribution has been frequently used for the estimation of extreme value event containing flood, rainfall and droughts. Hosking and Wallis (1997) introduced GLO distribution for the first time in hydrological literature. Reed & Robson (1999) and Ahmad *et al* (2011) recommended GLO distribution for extreme floods estimation in UK and Malaysia respectively. Further it is also used in UK for the estimation of return periods in the range of 2 to 200 years by Chadwick *et al* (2004). The probability density function (PDF) of GLO distribution is given by:

$$f(x) = \frac{1}{\alpha}\left[1 - k\left(\frac{x-\xi}{\alpha}\right)\right]^{\left(\frac{1}{k}-1\right)} \left[1 + \left\{1 - k\left(\frac{x-\xi}{\alpha}\right)\right\}^{\left(\frac{1}{k}\right)}\right]^{-2} \qquad (3.4.1.1)$$

Where k is a shape, ξ is a location and α is scale parameter. The range of x is:

$-\infty < x \le \xi + \frac{\alpha}{\kappa}$ $\quad if\, k > 0\,;\; -\infty < x < \infty \quad if\, k = 0\,;\; \xi + \frac{\alpha}{\kappa} \le x < \infty \quad if \;\; k < 0.$

The distribution function (DF) of GLO distribution is:

$$F(x) = \left[1 + \left\{1 - k\left(\frac{x-\xi}{\alpha}\right)\right\}^{\left(\frac{1}{k}\right)}\right]^{-1} \qquad (3.4.1.2)$$

And the corresponding quantile function of GLO distribution is:

$$x(F) = \xi + \frac{\alpha}{\mathrm{k}}\left[\left(\frac{1-F}{F}\right)^{k}\right] \tag{3.4.1.3}$$

The L-moments, L-moment ratios and parameters estimates of GLO distribution are given by Hosking (1986).

The GLO distribution first four population L-moment with relation to estimates of parameters are as:

$$\lambda_1 = \xi + \frac{\alpha}{k}\,\{1 - \Gamma(1-k)\,\Gamma(1+k)\} \tag{3.4.1.4}$$

$$\lambda_2 = \alpha\,\{\Gamma(1-k)\,\Gamma(1+k)\} \tag{3.4.1.5}$$

$$\lambda_3 = -\alpha k\,\{\Gamma(1-k)\,\Gamma(1+k)\} \tag{3.4.1.6}$$

$$\lambda_4 = \frac{\pi\alpha k(5k^2)+1}{6\sin(\pi k)} \tag{3.4.1.7}$$

The GLO distribution L-skewness τ_3 and L-kurtosis τ_4 can be defined as:

$$\tau_3 = \frac{\lambda_3}{\lambda_2} = -k \tag{3.4.1.8}$$

$$\tau_4 = \frac{\lambda_4}{\lambda_2} = \frac{1}{6} + \frac{5k^2}{6} \tag{3.4.1.9}$$

Parameter estimation of GLO distribution by using L-moments method.

$$\hat{\alpha} = \frac{l_2}{\Gamma(1+\hat{k})\Gamma(1-\hat{k})} \tag{3.4.1.10}$$

$$\hat{\xi} = l_1 + \frac{(l_2 - \hat{\alpha})}{\hat{k}} \tag{3.4.1.11}$$

$$\hat{k} = -t_3 \tag{3.4.1.12}$$

If $k=0$ then this is the special case of log-logistic distribution. But this generalization of logistic distribution changes from others that have been defined in literature. This is a parameterized form of the log-logistic distribution of Ahmad *et al* (1988).

The TL-Moments of GLO distribution with trimming *(1, 0)* is given by Shabri *et al* (2011) and compared the L-moment and TL-moment with *(1, 0)*. Results showed that TL-moment with *(1, 0)* is better than the L-moment and TL-moment also for modeling.

The TL-moment of GLO distribution are:

$$\lambda_1^{(1,0)} = \xi - \frac{\alpha[\Gamma(1-k)\,\Gamma(k+1)-1]}{k} + \alpha[\Gamma(1-k)\Gamma(k+1)] \tag{3.4.1.13}$$

$$\lambda_2^{(1,0)} = \frac{-3\alpha[\Gamma(1-k)\,\Gamma(k+1)]}{2} - \frac{3\alpha[\Gamma(1-k)\,\Gamma(k+1)(k-3)]}{4} \tag{3.4.1.14}$$

$$\lambda_3^{(1,0)} = \frac{10\alpha}{3k} + \frac{10\alpha[\Gamma(1-k)\,\Gamma(k+1)-1]}{3k} + \frac{8\alpha[\Gamma(1-k)\,\Gamma(k+1)(k-3)]}{3} - \frac{5\alpha[\Gamma(4-k)\,\Gamma(k+1)]}{9k} + 2\alpha[\Gamma(1-k)\Gamma(k+1)] \tag{3.4.1.15}$$

$$\lambda_4^{(1,0)} = \frac{25\alpha\Gamma(4-k)}{\Gamma(k+1)}8k - \frac{10\alpha[\Gamma(1-k)\,\Gamma(k+1)-1]}{k} - \frac{5\alpha[\Gamma(1-k)\,\Gamma(k+1)]}{2} - \frac{25\alpha[\Gamma(1-k)\,\Gamma(k+1)\Gamma(k-3)]}{4} - \frac{10\alpha}{k} - \frac{35\alpha[\Gamma(5-k)(k+1)]}{96k} \tag{3.4.1.16}$$

The TL-skewness and TL-kurtosis for GLO distribution are given below:

$$\tau_3^{(1,0)} = \frac{\lambda_3^{(1,0)}}{\lambda_2^{(1,0)}} = \frac{4}{27} - \frac{20k}{27} \tag{3.4.1.16}$$

$$\tau_4^{(1,0)} = \frac{\lambda_4^{(1,0)}}{\lambda_2^{(1,0)}} = \frac{35k^2}{72} - \frac{5k}{24} - \frac{5}{36} \tag{3.4.1.17}$$

The scale α, location ξ and shape k can be estimated by using TL-Moment method with trimming *(1, 0)* are:

$$\hat{\alpha} = \frac{4l_2^{(1,0)}}{-6\Gamma(1-\hat{k})\Gamma(\hat{k}+1) - 3\Gamma(1-\hat{k})\Gamma(\hat{k}+1)(\hat{k}-3)} \tag{3.4.1.18}$$

$$\hat{\xi} = l_1^{(1,0)} + \frac{\hat{k}[\Gamma(1-\hat{k})\Gamma(\hat{k}+1) - 1]}{\mathrm{k}} - \hat{\alpha}[\Gamma(1-\hat{k})\Gamma(\hat{k}+1)] \tag{3.4.1.19}$$

$$\hat{k} = \frac{4 - 27t_3^{(1,0)}}{20} \tag{3.4.1.20}$$

3.4.2 Generalized Extreme Value Distribution (GEV)

Due to extreme or rare events, extreme floods and snowfalls, high wind speeds, extreme temperatures, and large fluctuations in exchange rates, and market crashes. The problem of modeling arises in business analyst and engineers are using GEV model for assessing the risks caused by these events. GEV distribution, are widely used in many areas, risk management, finance, insurance, economics, hydrology, material sciences etc.

The GEV distribution is the combination of three types of extreme value distributions i.e. type I, type II and type III which are Gumbel, Fréchet and Weibull distribution respectively. In other word if $k=0$ it is the Gumbel distribution, $k>0$ is Fréchet distribution and $k<0$ is Weibull distribution.

The GEV distribution is developed through the extreme value theory. As these extreme value distribution are the limiting case of above three continuous probability distributions when "*n*" tend to infinity. That's why here the name 'extreme value' is attached with these distribution. Keep in mind as these distribution are tag "extreme values" they do not represent distributions of all kinds of "extreme values" but they can be used empirically like other distributions by Johnson *et al* (1970). The PDF and CDF of GEV distribution are given below.

$$f(x) = \frac{1}{\alpha}\left[1 - k\left(\frac{x-\xi}{\alpha}\right)\right]^{\left(\frac{1}{k}-1\right)} e^{-\left[1-k\left(\frac{x-\xi}{\alpha}\right)\right]^{\left(\frac{1}{k}\right)}} \quad (3.4.2.1)$$

$$F(x) = exp\left\{-\left[1 - k\left(\frac{x-\xi}{\alpha}\right)\right]^{\frac{1}{k}}\right\} \quad (3.4.2.2)$$

The range of X: $-\infty < x \leq \xi + \frac{\alpha}{k}$ $if\ k > 0$; $-\infty < x < \infty\ if\ k = 0$; $\xi + \frac{\alpha}{k} \leq x < \infty\ if\ k < 0$.

The Quantile function of GEV distribution is given below.

$$x(F) = \xi + \frac{\alpha}{k}[1 - (-lnF)^k] \quad (3.4.2.3)$$

The L-moments for GEV distribution

$$\lambda_1 = \xi + \frac{\alpha}{k}\{1 - \Gamma(1+k)\} \quad (3.4.2.4)$$

$$\lambda_2 = \frac{\alpha}{k}\{(1-2^{-k})\,\Gamma(1+k)\} \quad (3.4.2.5)$$

$$\tau_3 = \frac{2(1-3^{-k})}{(1-2^{-k})} - 3 \quad (3.4.2.6)$$

$$\tau_4 = \frac{\{5(1-4^{-k}) - 10(1-3^{-k}) + 6(1-2^{-k})\}}{(1-2^{-k})} \quad (3.4.2.7)$$

Using L-moment to estimate the parameters of GEV distribution

For equation (3.4.2.6), it is important to estimate k.

$$\hat{\alpha} = \frac{l_2\hat{k}}{\Gamma(1+\hat{k})\Gamma(1-2^{-\hat{k}})} \quad (3.4.2.8)$$

$$\hat{\xi} = l_1 + \frac{\hat{\alpha}[\Gamma(1+\hat{k})-1]}{\hat{k}} \quad (3.4.2.9)$$

$$\hat{k} = 7.8590\hat{C} + 2.9554\hat{C}^2 \quad (3.4.2.10)$$

$$\hat{C} = \frac{2}{3+t_3} - \frac{ln2}{ln3} \quad (3.4.2.11)$$

Using TL-moment to estimates the parameters of GEV distribution

$$\hat{\alpha} = \frac{-2(6)^{\hat{k}}\, l_2}{3\Gamma(1+\hat{k})\,(2^{\hat{k}} - 3^{\hat{k}})} \quad (3.4.2.12)$$

$$\hat{\xi} = l_1 + \frac{\hat{\alpha}}{\hat{k}} + \frac{\alpha\Gamma(k)}{2^k} \quad (3.4.2.13)$$

$$\hat{k} = 0.49 - 2.08t_3 + 0.61(t_3)^2 - 0.60(t_3)^3 + 0.48(t_3)^4 \quad (3.4.2.14)$$

3.4.3 Generalized Pareto Distribution (GPA)

The GPA distribution is also very suitable for modeling events like flood and rainfall and particularly for flood flow data it provides the best estimates. In many countries including Turkey, Malaysia, Nigeria and many most the GPA distribution is used for FFA Saf (2009) and Izinyon & Ehiorobo (2014). Rao *et al* (2000) derived the scale and shape parameter of GPA distribution and it is right-skewed distribution. Special cases arise when $k=0$ and $k=1$, it becomes exponential and uniform distributions respectively and it is also the special case of wakeby distribution. The PDF of GPA distribution is.

$$f(x) = \frac{1}{\alpha}\left[1 - \frac{k}{\alpha}(x - \xi)\right]^{\left(\frac{1}{k}-1\right)} \quad (3.4.3.1)$$

Range of x: $\xi \leq x \leq \xi + \frac{\alpha}{k}$ if $k > 0$; $\xi \leq x < \infty$ if $k \leq 0$.

Where ξ, α, and k are location, scale and shape parameters, respectively. The corresponding distribution function is given below:

$$F(x) = 1 - \left[1 - \frac{k}{\alpha}(x - \xi)\right]^{\frac{1}{k}} \quad (3.4.3.2)$$

Quantile function of GPA distribution:

$$x(F) = \xi + \frac{\alpha}{k}\left[1 - (1 - F)^{k}\right] \quad (3.4.3.3)$$

The L-moment of GPA distribution are given below:

$$\lambda_1 = \xi + \frac{\alpha}{(1 + k)} \quad (3.4.3.4)$$

$$\lambda_2 = \frac{\alpha}{(1 + k)(2 + k)} \quad (3.4.3.5)$$

$$\tau_3 = \frac{(1 - k)}{(3 + k)} \quad (3.4.3.6)$$

$$\tau_4 = \frac{(1 - k)(2 - k)}{(3 + k)(4 + k)} \quad (3.4.3.7)$$

The relationship among τ_3 and τ_4 is given by.

$$\tau_4 = \frac{\tau_3(1 + 5\tau_3)}{(5 + \tau_3)} \quad (3.4.3.8)$$

Using L-moments to estimate the parameters of GPA distribution

$$\hat{\alpha} = l_2\left[(\hat{k}+1)(\hat{k}+2)\right] \qquad (3.4.3.9)$$

$$\hat{\xi} = l_1 - l_2\left(\hat{k}+2\right) \qquad (3.4.3.10)$$

$$\hat{k} = \frac{(1-3t_3)}{(1+t_3)} \qquad (3.4.3.11)$$

Ahmad *et al* (2012) estimates the first four sample TL-moment and the parameters of GPA distribution using TL-moments with trimming *(1, 0)*:

$$\lambda_1^{(1,0)} = \xi + \alpha\left[\frac{(k+3)}{(k+1)(k+2)}\right] \qquad (3.4.3.12)$$

$$\lambda_2^{(1,0)} = \frac{3\alpha}{(k+1)(k+2)(k+3)} \qquad (3.4.3.13)$$

$$\lambda_3^{(1,0)} = \frac{4\alpha(1-k)}{(k+1)(k+2)(k+3)(k+4)} \qquad (3.4.3.14)$$

$$\lambda_4^{(1,0)} = \frac{5\alpha(k-1)(k-2)}{(k+1)(k+2)(k+3)(k+4)(k+5)} \qquad (3.4.3.15)$$

TL-skewness and TL-kurtosis

$$\tau_3^{(1,0)} = \frac{\lambda_3^{(1,0)}}{\lambda_2^{(1,0)}} = \frac{4(1-k)}{3(k+4)} \qquad (3.4.3.16)$$

$$\tau_4^{(1,0)} = \frac{\lambda_4^{(1,0)}}{\lambda_2^{(1,0)}} = \frac{5(1-k)(k-2)}{3(k+4)(k+5)} \qquad (3.4.3.17)$$

Parameters estimation:

$$\hat{\alpha} = \left[\frac{l_2^{(1,0)}(\hat{k}+1)(\hat{k}+2)(\hat{k}+3)}{3}\right] \quad (3.4.3.18)$$

$$\hat{\xi} = l_1^{(1,0)} - \hat{\alpha}\left[\frac{(\hat{k}+3)}{(\hat{k}+1)(\hat{k}+2)}\right] \quad (3.4.3.19)$$

$$\hat{k} = \left[\frac{4 - 12t_3^{(1,0)}}{3t_3^{(1,0)} + 4}\right] \quad (3.4.3.20)$$

3.5 Comparison of the Probability Distributions Using Goodness-of-Fit Criteria:

The Selected distribution models are fitted to the observed AM flows data and their performances are compared by goodness-of-fit tests, Mean Absolute Deviation Index (MADI), Relative Root Mean Square Error (RRMSE) and Probability Plot Correlation Coefficient (PPCC). The MADI and RRMSE are used to measure the deviation between current flow and predicted flow while the last one is used to measure correlation between current and predicted flow. On the basis of the results of these tests we will be able to determine which distribution is best fitted to the current data among the three selected distributions. We are also using L- ratio diagram and TL-ratio diagram to confirm the distribution that is best fitted to the current data.

3.5.1 Mean Absolute Deviation Index (MADI)

This method has been applied for the first time by Jain and Singh (1987) as goodness-of-fit. The goal of this method is to check whether a given distribution fits the data closely by selecting from a number of candidate distributions, the one that gives the best fit to the data. The MADI is calculated by:

$$MADI = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - y_i}{x_i}\right| \quad (3.5.1.1)$$

In the above equation x_i's are the observed flows, "y_i" the predicted flows and N is the number of observation at any site. The $y_{i,s}$ predicted flows are calculated by using Hosking (1990) plotting position formula. Which gives better results for GEV, GPA and GLO distributions Hosking *et al* (1985b) and Hosking & Wallis (1987a).

$$F = \frac{(j - 0.35)}{n}, \qquad j = 1,2,\ldots n \quad (3.5.1.2)$$

Where'j' is the observation in ascending order and n is sample size. The MADI of smaller value obtained for given distribution shows that the distribution is more closely fitted to the actual data. Therefore, the distribution with the smallest value of MADI indicates that the particular distribution is best fitted to the current observed data.

3.5.2 Relative Root Mean Square Error (RRMSE)

For the judgment of probability distributions RRMSE is applied to AM flows data for the three selected distributions. About the distribution of overall fit RRMSE gives better result because it calculates every single error in proportion to the size of the observation. It also reduces the effect of outlier which are commonly found in flood data Tao *et al* (2002). It is calculated as:

$$RRMSE = \left[\frac{\sum_{i=1}^{n}\left(\frac{x_i - y_i}{x_i}\right)^2}{(n-m)}\right]^{\frac{1}{2}} \tag{3.5.2.1}$$

Where x_i the current values of flows are, y_i is calculated from the assumed probability distributions, *m* is the number of parameters estimated for the distributions and *n* is the number of observation at any site. The RRMSE of smaller value obtained for given distribution shows that the distribution is more closely fitted to the actual data. Therefore, the distribution with the smallest value of RRMSE indicates that the particular distribution is best fitted to the current observed data.

3.5.3 Probability Plot Correlation Coefficient (PPCC)

Basically it is the relationship between current observations and corresponding predicted plotting position Karim *et al* (1995). PPCC between the current and predicted flows are given by:

$$r = \frac{\sum_{i=1}^{n}[(x_i - \bar{x})\,(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{3.5.3.1}$$

Where x_i are the observed flows and y_i are the predicted flows .The range of correlation *r* is -1 to +1. When a value of correlation of fitted distribution is closer to 1 or -1 it suggests that the current data have been drawn from the fitted distribution and when a value of correlation of fitted distribution is closer to '0' it suggests that the current data is not drawn from the fitted distribution.

3.5.4 Linear-Moments and Trimmed Linear-Moments Ratio Diagrams

The simplest method to determine the best distribution to fit the actual data is by using L-moment and TL-moment ratio diagrams, which based on the relationships between L-moment and TL-moment ratios respectively. The L-moment and TL-moment ratio diagrams provide a graphical indication of which distribution can be expected to give a better fit to the sample data or samples. Therefore, this allows better discrimination between the distributions.

Hence, from the L-moment and TL-moment ratio diagrams the identification of parent distribution can be also achieved. The associations among τ_3 and τ_4 are used for the three selected distribution containing GLO, GEV and GPA used in this study. For each distribution the sample L-moment and TL-moment ratios take the range $-1 \leq \tau_3 \leq 1$. For this interval, τ_4 is computed for the three selected distribution using their relationships with τ_3. Then the sample L-moment and TL-moment ratios are plotted in the diagram as (τ_3, τ_4) and $\left(\tau_3^{(1,0)}, \tau_4^{(1,0)}\right)$, the distribution for which their L-moment or TL-moment ratios are close to the values of sample ratios are considered to be the best distribution for fitting the observed data.

3.6 Quantiles of Best Fit Distribution

One of the objectives of FFA is to get handy estimates of the quantiles for return period of scientific significance. After estimating the parameters of best fitted distribution we need to find out the Quantiles' estimates corresponding to different return periods (T). AM flows do not occur with any fixed pattern in time or magnitude. Large floods naturally have large return periods with less probability and vice versa. The relationship between return periods and occurrence of an extreme event (e.g. amount of flow greater than or equal to some threshold value in a dam) may be established through notion of geometric random variable as: $P = \frac{1}{T}$ where P is probability of occurrence of T year return period event i.e. $X \geq = x_t$. $P = \frac{1}{T}$ implies that T year return period event will occurs on average once in T years. The above equation is the basis for estimating the magnitude of the AM flows (estimates of the quantiles), x_t with exceedance probability P. Further, probability of non-exceedance can be derived as below.

$$F(Q_T) = P(Q_T \leq q) = 1 - P(Q_T > q) = 1 - \frac{1}{T} \tag{3.6.1}$$

Where, q is some threshold value, under which flood will occur. The above equation is the basis for estimating the magnitude of a flood. Substituting $P = \frac{1}{T}$ or $F = 1 - \frac{1}{T}$ in a known statistical distribution function, one can solve for the estimate of Q_T $i.e$ "q". Putting equation (3.6.1) in the quantile function of GLO, GEV and GPA distributions we can get quantile estimates for different return periods. The following equations (3.6.2 to 3.6.4) represents the corresponding relationships for GLO, GEV and GPA distributions, respectively.

$$\hat{x}_T = \hat{\xi} + \frac{\hat{\alpha}}{\hat{k}}\left[1 - (T-1)^{-\hat{k}}\right] \tag{3.6.2}$$

$$\hat{x}_T = \hat{\xi} + \frac{\hat{\alpha}}{\hat{k}}\left[1 - \left\{-log\left(1 - \frac{1}{T}\right)\right\}^{-\hat{k}}\right] \tag{3.6.3}$$

$$\hat{x}_T = \hat{\xi} + \frac{\hat{\alpha}}{\hat{k}}\left[1 - T^{-\hat{k}}\right] \tag{3.6.4}$$

The above equations (3.6.2 to 3.6.4) are used for the determination of quantile estimates for different return period of the GLO, GEV and GPA distributions for methods of the L-moments and TL-moments.

The quantile function of the probability distribution shows the magnitude of an event in terms of non-exceedance probability as well as exceedance probability (whatever we prefer because total probability is unity, and random variable has two categories "occur not occur"). For example, a 5 years return period event yields a probability of exceedance (occur of flood) equal to 0.2 and the probability of non- exceedance is 0.8 and the corresponding quantile value is based on "P" probability and the selected distributions (GLO, GEV and GPA). We can say that 5 years return period event will occur (some threshold value on average once in 5 years) with probability of 0.2. Sometimes we are interested to find out the probability of a T year return period event to occur at least once in N years. The answer of this equation can be achieved by using the relationship given by $1-\left(1-\frac{1}{T}\right)^{N}$. In case of 5 years return period event (as discussed above), the probability of this 5 years return period event to occur at least once in next 5 year is: 0.672.

CHAPTER 4

Results and Discussion

4.1 Basic Assumption:

Before we analyze the data, initially we check the basic assumptions of FFA which are stationary, independence, and homogeneity of data. First we apply time series plot to identify the patterns in the AM flows. Generally the AM flow of all 18 sites showed that there is no systematic jumps or trend. So finally we conclude that there is no constant increasing/decreasing in the data of 18 sites of Pakistan. Time series plots of Kalabagh and Chashma sites are shown in the fig 4.1. Next for Stationarity, independency and homogeneity we apply Ljung-Box Q-statistics, lag-one correlation and Mann- Whitney U tests respectively. The results of all the 18 sites show Stationarity, independency and homogeneity. Further to check trend we apply Mann Kendall test to all sites. All the sites show insignificant trends. Here we also apply Change Point Analysis (CPA) developed by Wayne Taylor (2011) an effective procedure to detect changes in historical data (Flood data). We apply CPA to our series of AM flows which detect changes, for each change we provide confidence interval, estimate the time of the change where CPA is robust to outliers. In our data the changes occur according to Wayne Taylor (2011) but these changes are acceptable.

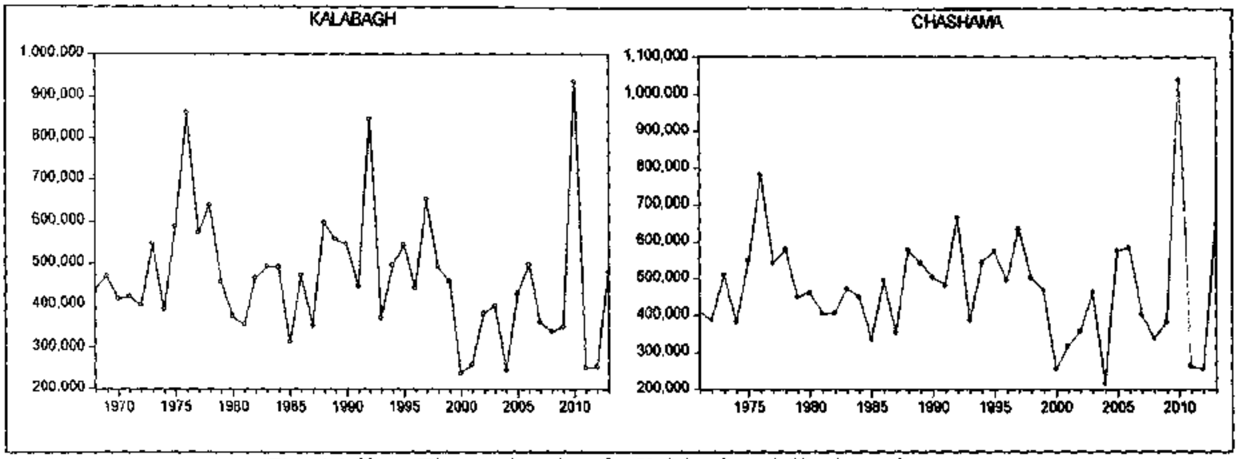


Fig. 4.1 Show Time series plots for Kalabagh and Chashma sites

4.2 Sample Linear-Moments and Linear-Moments Ratios:

L-moments and L-moments ratios of 18 sites are analyzed by using the R language programming. The results are tabulated in table no 4.2. Further these values are used for the calculation of quantile function for the three selected distributions by using Method of L-moments.

Table 4.2: Sample L-Moments and L- Moments Ratio for all sites:

Sites	ℓ_1	ℓ_2	ℓ_3	ℓ_4	t	t_3	t_4
Tarbela	386962.963	42836.478	10548.191	9095.715	0.111	0.246	0.212
Kalabagh	464719.96	81158.68	14085.12	18309.99	0.175	0.174	0.226
Chashma	475333.047	79990.796	9055.046	16374.089	0.168	0.113	0.205
Taunsa	452791.554	76857.484	4325.554	13131.199	0.170	0.056	0.171
Guddu	609909.420	161117.76	21164.22	14568.30	0.264	0.131	0.090
Sukkur	546609.594	176548.689	30461.464	8293.265	0.323	0.173	0.047
Kotri	395262.07	161959.43	65544.90	52419.16	0.410	0.405	0.324
Mangla	132481.78	54023.67	24245.62	19729.55	0.408	0.449	0.365
Rasul	134418.39	66402.58	32458.78	23316.78	0.494	0.489	0.351
Marala	308572.41	106247.39	31321.98	10904.27	0.344	0.295	0.103
Khanki	351963.19	123768.11	45550.74	22860.59	0.352	0.368	0.185
Qadirabad	356547.70	135642.60	36453.86	12931.54	0.380	0.269	0.095
Trimmu	261376.22	104959.78	30819.02	12010.62	0.402	0.294	0.114
Panjnad	260134.72	106855.13	23005.86	12780.08	0.411	0.215	0.120
Balloki	87914.728	30921.621	12070.471	7577.139	0.352	0.390	0.245
Sidhani	64143.427	27038.629	10954.863	6896.196	0.422	0.405	0.255
Sulemanki	70254.92	40419.47	18280.68	10021.65	0.575	0.452	0.248
Islam	49089.450	29493.988	14184.388	7808.931	0.601	0.481	0.265

4.3 Sample Trimmed Linear-Moments and Trimmed Linear - Moments Ratios with Trimming($t_1 = 1, t_2 = 0$):

By using the R language programming, TL-moments and TL-moments ratios with smallest trimming ($t_1 = 1, t_2 = 0$) of 18 sites are examined. The results are shown in the table no 4.3. Further, to calculate the quantile function for three selected distributions the below mentioned values in table no.4.3 are applied.

Table 4.3: Sample TL-Moments and TL- Moments Ratio for all sites :

Sites	ℓ_1	ℓ_2	ℓ_3	ℓ_4	t	t_3	t_4
Tarbela	429799.441	40038.502	13095.938	8948.614	0.093	0.327	0.224
Kalabagh	545878.64	71432.85	21596.74	16806.35	0.131	0.302	0.235
Chashma	555323.84	66784.38	16952.76	15741.76	0.120	0.254	0.236
Taunsa	529649.04	60887.28	11637.84	13776.24	0.115	0.191	0.226
Guddu	771027.18	136711.48	23821.68	10681.65	0.177	0.174	0.078
Sukkur	723158.282	155257.614	25836.486	6021.071	0.215	0.166	0.038
Kotri	557221.50	170628.25	78642.71	54803.28	0.306	0.461	0.321
Mangla	186505.45	58701.97	29316.78	21549.97	0.315	0.499	0.367
Rasul	200820.97	74146.02	37183.71	25819.28	0.369	0.501	0.348
Marala	414819.802	103177.027	28150.830	7263.726	0.249	0.273	0.070
Khanki	475731.30	126989.14	45607.56	17665.41	0.267	0.359	0.139
Qadirabad	492190.302	129072.347	32923.601	9572.133	0.262	0.255	0.074
Trimmu	366335.997	101834.098	28553.093	7246.251	0.278	0.280	0.071
Panjnad	366989.86	97395.75	23857.30	11919.84	0.265	0.245	0.122
Balloki	118836.349	32244.069	13098.406	6456.135	0.271	0.406	0.200
Sidhani	91182.056	28495.119	11900.706	6610.139	0.313	0.418	0.232
Sulemanki	110674.39	44025.11	18868.22	10080.48	0.398	0.429	0.229
Islam	78583.438	32758.782	14662.212	7604.301	0.417	0.448	0.232

4.4 Comparison of Sample Linear-Moments and Sample Trimmed Linear-Moments:

It is observed from both table 4.2 and table 4.3 that Guddu site has the largest average AM flows, while Sukkur and Kotri has the largest variation. Now comparing the results of L-cv and TL-cv from table 4.1 and table 4.2 we found that the TL-cv are smaller than L-cv for all sites. Its mean that TL-cv are consistent than L-cv. It is also observed from both tables that the flows of all sites follow a positive extreme with a degree of skewness ranging from 0.056 to 0.501. That is enough by itself to characterize the distribution as non-normal. Some of the sites are highly skewed i.e. Islam, Rasul, Mangla, Sulemanki.

4.5 MADI, RRMSE and PPCC for Linear-Moments and Trimmed Linear-Moments:

The MADI, RRMSE and PPCC of 18 sites are analyzed by using Microsoft Excel for two cases (L-Moments & TL-Moments) and therefore the results are found and tabulated. For every site, their MADI and RRMSE are then ranked starting with the minimum value the distribution which best fits the data. The PPCC for every site is then ranked with the value closer to one for best fit distribution.

4.5.1 Results of MADI, RRMSE and PPCC for Linear-Moments:

The MADI, RRMSE and PPCC are obtained by using L-moment technique for three selected distributions which are GLO, GEV and GPA distributions as shown in Table 4.5.1.1. The best fitted distribution for each site are given in the table no. 4.5.1.2.

Table 4.5.1.1: MADI, RRMSE and PPCC using L-Moments:

NAME OF SITES	MADI			RRMSE			PPCC		
	GLO	GEV	GPA	GLO	GEV	GPA	GLO	GEV	GPA
Tarbela	0.033	0.043	0.040	0.055	0.064	0.063	0.918	0.900	0.895
Kalabagh	0.159	0.080	0.099	0.199	0.092	0.134	0.975	0.971	0.942
Chashma	0.071	0.061	0.084	0.087	0.078	0.108	0.973	0.955	0.917
Taunsa	0.040	0.060	0.068	0.061	0.077	0.096	0.986	0.970	0.948
Guddu	0.121	0.077	0.177	0.212	0.151	0.204	0.980	0.992	0.982
Sukkur	0.195	0.136	0.121	0.254	0.177	0.163	0.986	0.959	0.991
Kotri	0.169	0.253	0.226	0.217	0.330	0.327	0.992	0.930	0.935
Mangla	0.111	0.155	0.188	0.153	0.194	0.255	0.929	0.955	0.913
Rasul	0.112	0.200	0.225	0.161	0.269	0.447	0.987	0.903	0.959
Marala	0.378	0.119	0.134	0.411	0.175	0.156	0.986	0.959	0.987
Khanki	0.161	0.107	0.090	0.186	0.211	0.108	0.933	0.822	0.973
Qadirabad	0.222	0.144	0.116	0.344	0.266	0.137	0.896	0.964	0.987
Trimmu	0.214	0.207	0.259	0.309	0.306	0.333	0.971	0.885	0.967
Panjnad	0.406	0.194	0.141	0.953	0.324	0.227	0.974	0.965	0.990
Balloki	0.149	0.119	0.097	0.223	0.172	0.207	0.929	0.975	0.981
Sidhani	0.113	0.114	0.081	0.137	0.153	0.147	0.965	0.969	0.987
Sulemanki	0.779	1.298	0.481	1.066	2.516	0.620	0.979	0.941	0.982
Islam	0.633	0.530	0.282	1.508	0.916	0.492	0.910	0.952	0.956

Table 4.5.1.2: Best fitting distributions based on MADI, RRMSE, PPCC and L-moment ratio diagram using L-Moment methods for all sites:

Sites name	MADI	RRMSE	PPCC	L-moment ratio diagram	Best
Tarbela	GLO	GLO	GLO	GLO	GLO
Kalabagh	GEV	GEV	GLO	GLO	GEV
Chashma	GEV	GEV	GLO	GLO	GEV
Taunsa	GLO	GLO	GLO	GLO	GLO
Guddu	GEV	GEV	GEV	GEV	GEV
Sukkur	GPA	GPA	GPA	GPA	GPA
Kotri	GLO	GLO	GLO	GLO	GLO
Mangla	GLO	GLO	GEV	GLO	GLO
Rasul	GLO	GLO	GLO	GLO	GLO
Marala	GEV	GPA	GPA	GPA	GPA
Khanki	GPA	GPA	GPA	GPA	GPA
Qadirabad	GPA	GPA	GPA	GPA	GPA
Trimmu	GEV	GEV	GLO	GPA	GEV
Panjnad	GPA	GPA	GPA	GPA	GPA
Balloki	GPA	GEV	GPA	GPA	GPA
Sidhani	GPA	GLO	GPA	GPA	GPA
Sulemanki	GPA	GPA	GPA	GPA	GPA
Islam	GPA	GPA	GPA	GPA	GPA

Table 4.5.1.2 shows the best fitted distribution for each of 18 sites constructed on the value of MADI, RRMSE and PPCC for L-moment technique. From the MADI, RRMSE, PPCC and L-moments ratio diagram results as given in the table 4.5.1.2, it is clear that GPA distribution has the best fit, followed by GLO and GEV distributions. The overall results as given in the percentages of the best fitting distributions with respect to different sites based on the L-moment ratio diagram are, for GPA 50%, for GLO 27.77% and GEV 22.22%. Therefore, the results obtained from goodness-of-fit tests and L-Moment ratio diagram are agreed on the best fitting distribution for 15 out of 18 sites examined in this study (83.33%).

We also compared the performances of the distributions under individual goodness-of-fit tests.

Case1: for MADI in table 4.5.1.2 it is clear that GPA has the most appropriate distribution followed by GLO and GEV distributions. We can also describe these results in percentages as: for GPA 44.44%, for GLO 27.77% and for GEV 27.77%. So in this case GPA is best fitted distribution.

Case2: when using RRMSE goodness-of-fit test to compare the performances of distributions, we noted from table 4.5.1.2 that GPA is best fitted distribution. The results of distributions are in percentages as, GPA 38.88%, for GLO 33.33% and for GEV is 27.77%. Similarly in this case also GPA is the most appropriate distribution followed by GLO distribution.

Case3: The results for PPCC in percentages are, for GPA 50%, GLO 38.88%, and GEV 11.11%.

In general, the most frequently selected best fitted distributions are GPA and GLO for both goodness-of-fit tests and L-moment ratio diagram.

4.5.2 Results of MADI, RRMSE and PPCC for Trimmed Linear-Moments:

Table 4.5.2.1 shows that MADI, RRMSE and PPCC are obtained in the case of TL-moment with smallest trimming ($t_1 = 1, t_2 = 0$) for all 18 sites .Table 4.5.2.2 presents the sum of each ranking for each distribution for all 18 sites separately.

Table 4.5.2.1: MADI, RRMSE and PPCC using TL-Moments:

NAME OF SITES	MADI			RRMSE			PPCC		
	GLO	GEV	GPA	GLO	GEV	GPA	GLO	GEV	GPA
Tarbela	0.271	0.236	0.255	0.363	0.255	0.297	0.914	0.916	0.894
Kalabagh	0.277	0.398	0.419	0.326	0.473	0.547	0.980	0.973	0.967
Chashma	0.225	0.406	0.374	0.293	0.492	0.535	0.969	0.950	0.935
Taunsa	0.313	0.385	0.394	0.366	0.482	0.572	0.935	0.965	0.969
Guddu	0.461	0.427	0.414	0.518	0.563	0.549	0.984	0.985	0.973
Sukkur	0.772	0.379	0.690	0.975	0.554	0.844	0.956	0.994	0.986
Kotri	0.895	0.749	0.668	1.148	1.118	1.097	0.894	0.907	0.971
Mangla	0.676	0.547	0.907	0.900	0.761	1.091	0.947	0.872	0.937
Rasul	0.809	0.751	1.069	1.601	1.057	1.236	0.956	0.931	0.929
Marala	0.769	1.022	0.485	0.845	1.206	0.560	0.948	0.967	0.990
Khanki	0.942	1.299	0.528	1.001	1.587	0.594	0.496	0.481	0.756
Qadirabad	0.848	0.717	0.650	1.069	0.959	0.801	0.965	0.987	0.979
Trimmu	3.224	0.786	0.778	5.125	1.038	1.039	0.928	0.957	0.911
Panjnad	0.985	0.986	0.965	1.612	1.570	1.988	0.987	0.980	0.983
Balloki	0.937	1.144	0.527	1.009	1.403	0.207	0.472	0.837	0.938
Sidhani	0.994	1.752	0.745	1.134	2.052	0.992	0.599	0.859	0.917
Sulemanki	3.596	4.357	2.767	5.245	8.310	4.904	0.960	0.959	0.979
Islam	2.799	2.784	2.784	4.608	4.939	5.151	0.904	0.954	0.971

Table 4.5.2.2: Best fitting distributions based on MADI, RRMSE, PPCC and TL-moment ratio diagram using TL-Moment methods for all sites:

Sites name	MADI	RRMSE	PPCC	TL-moment ratio diagram	Best
Tarbela	GEV	GEV	GEV	GEV	GEV
Kalabagh	GLO	GLO	GLO	GLO	GLO
Chashma	GLO	GLO	GLO	GLO	GLO
Taunsa	GLO	GLO	GPA	GLO	GLO
Guddu	GPA	GLO	GEV	GPA	GPA
Sukkur	GEV	GEV	GEV	GPA	GEV
Kotri	GPA	GPA	GPA	GEV	GPA
Mangla	GEV	GEV	GLO	GEV	GEV
Rasul	GEV	GEV	GLO	GEV	GEV
Marala	GPA	GPA	GPA	GPA	GPA
Khanki	GPA	GPA	GPA	GPA	GPA
Qadirabad	GPA	GPA	GEV	GPA	GPA
Trimmu	GPA	GEV	GEV	GPA	GEV
Panjnad	GPA	GEV	GLO	GPA	GPA
Balloki	GPA	GPA	GPA	GPA	GPA
Sidhani	GPA	GPA	GPA	GPA	GPA
Sulemanki	GPA	GPA	GPA	GPA	GPA
Islam	GPA	GLO	GPA	GPA	GPA

The table 4.5.2.2 presents the best fitted distribution for each of 18 sites constructed on the value of MADI, RRMSE and PPCC based on TL-moment technique. The results obtain from the MADI, RRMSE, PPCC and TL-moments ratio diagram are shown in the table 4.5.2.2, indicates that GPA distribution obtained the first rank frequently, GEV distribution is second ranked and GLO distribution is last ranked. We describe the overall best fitted distributions results in the percentages for all site as: GPA 55.55%, GEV 27.77% and GLO 16.66%. It is noted that the results obtained from MADI, RRMSE, PPCC and TL-Moment ratio diagram agree on the best fitting distribution for 15 out of 18 sites examined in this study. Hence, our results of graphical method TL-ratio diagrams 83.33% coincide with mathematical results for goodness-of-fit tests, given in table 4.5.2.2. The evaluations of distributions are also individually compared under goodness-of-fit tests.

Case1: Taking in consideration the column of MADI in table 4.5.2.2, it is observed that GPA is best fitted distribution for most sites as compare to GLO and GEV distributions. The results obtained from MADI for best fitted distributions in percentages are, GPA 61.11%, GEV 22.22% and GLO 16.66%. So in this case most of the sites follow GPA distribution.

Case2: The results obtained from RRMSE is presented in the table 4.5.2.2. The GPA distribution ranked first most of the time compared with the other two distributions for the 18 sites .The results of RRMSE for three selected distribution in percentages for GPA 38.88%, GEV 33.33% and GLO 27.77%. Similarly in this case also GPA has the most appropriate distribution followed by GEV distribution.

Case3: GPA distribution is rank first the most from table 4.5.2.2. The GLO distribution and GEV distribution are obtained second rank for the calculations of including all the 18 sites. The results for PPCC in percentages are, for GPA 44.44%, for GEV 27.77% and for GLO 27.77%. In general GPA and GEV are the best fitted distributions from both goodness-of-fit tests and TL-moment ratio diagram.

4.6 Results of Linear-Moment Ratio Diagram and Trimmed Linear-Moment Ratio Diagram:

From L-moment and TL-moment ratio diagrams as shown in the fig.4.6.1 & fig.4.6.2 it is clear that, most of the sites follow GPA distribution curve, followed by GLO and the GEV distributions. Here we noted from L-moment ratio diagram that the sites with high skewness and high kurtosis follow GLO distribution. The sites with moderate skewness and small kurtosis follow GPA distribution and the sites with moderate skewness and kurtosis follow GEV distribution. Similarly from TL-moment ratio diagram the sites with high skewness and kurtosis follow GEV distribution, the site that have moderate skewness and kurtosis follow GLO distribution and the sites that have high skewness and small kurtosis follow GPA distribution. We also noted from both diagrams that most of the sites have equal skewness and kurtosis follow GPA distribution. It is observed that if we compare the L-moments and TL-moments, GPA distribution describes the TL-moments-ratio diagram better than the L-moment ratio diagram. Therefore from MADI, RRMSE, PPCC, L-ratio diagram and TL- ratio diagram it is clear that for the majority of sites GPA distribution is best fitted, followed by GLO and GEV distributions. Our results of graphical method L-ratio and TL-ratio diagrams coincide with results for other goodness-of-fit tests, given in table 4.5.1.2 and 4.5.2.2.

Fig.4.6.1 present L-moment ratio diagram for each of 18 sites in Pakistan

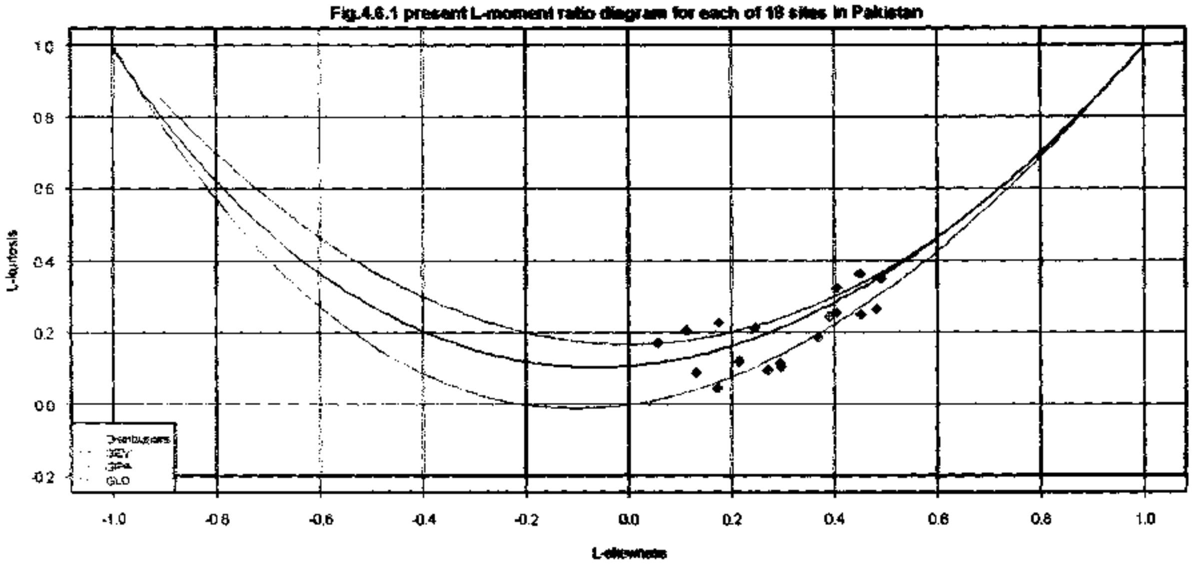
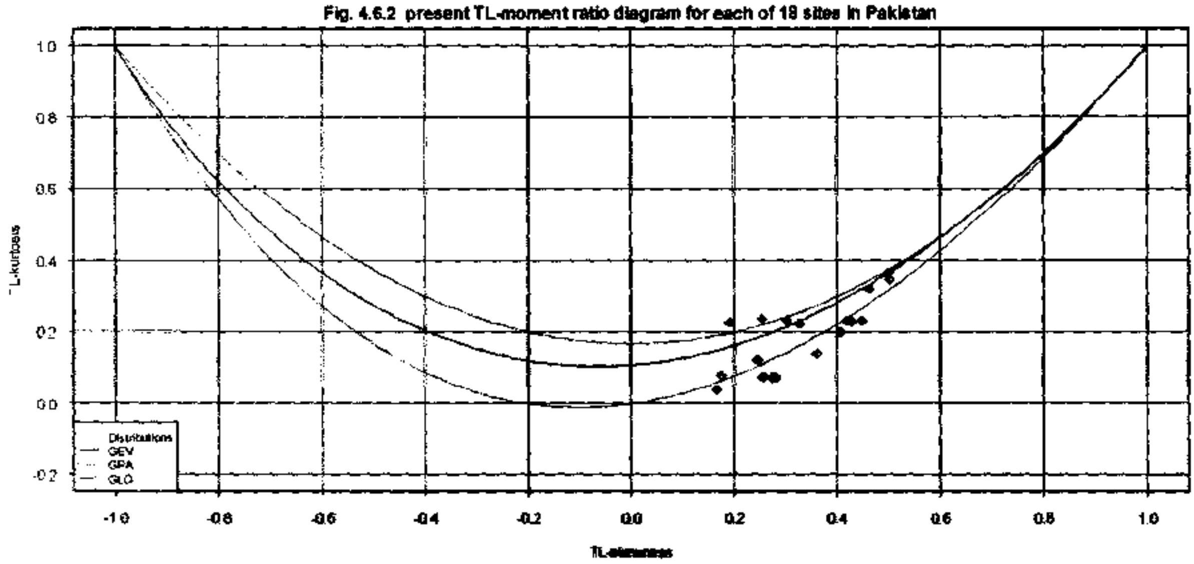


Fig. 4.6.2 present TL-moment ratio diagram for each of 18 sites in Pakistan



4.7 Results of Quantile Estimates for selected Return Periods using Methods of Linear-Moments and Trimmed Linear-Moments:

The quantile estimates $\hat{q}(F)$ for each sites with non-exceedance probabilities and return periods are given respectively in the following table 4.7.1 and table 4.7.2 for L-moment and TL-moments. We find quantile estimates for those distribution which are best fitted for that site. For Tarbela site GLO is best fitted distribution and we can interpret as, a 100 year return period event yields a probability of occurrence equal to 0.01 and the corresponding quantile value based on this probability and best fit distribution selected is 700,849.147 cusec. We can say that 100 years return period event will occur (greater or equal to 700,849.147 cusec) on an average once in 100 years with probability 0.01.

Another way to interpret that the probability of occurrence at least once of a T return period event in next N year can be given as $p_N = 1-(1-p)^N$. The probability that 100 year flood will at least once occur within next 50 year period, $p_N = 1-(1-1/100)^{50} = 0.395$. There is a 39.5% chance of the 100 year flood occurring in the next 50 years. Larger extreme events (Larger time period events) have less probability and vice versa. As we can see in both tables 4.7.1 & table 4.7.2 most of the sites including Tarbela, Guddu, Taunsa, Marala, and Mangla etc. have high quantile for large return period as compared to the other sites. Because that sites have huge design capacity of water as compared to the other sites.

We noted from both table 4.7.1 and table 4.7.2 that for shorter return period quantile estimates for TL-moment is larger than L-moment. We also note that for high return period the quantile estimates are larger for TL-moment as compare to L-moment. The difference in quantile estimates occurs because of trimming.

Table 4.7.1: Quantile estimates for each Sites for selected return periods using that distribution which is best for that site. (L-Moments)

Sites Name	Best Distribution	0.500 2	0.800 5	0.900 10	0.950 20	0.980 50	0.990 100	0.998 500
Tarbela	GLO	370,093.687	434,196.502	483,173.747	537,822.565	623,241.434	700,849.147	939,573.703
Kalabagh	GEV	439,458.483	572,445.627	661,021.222	746,381.408	857,450.643	941,111.905	1,135,860.898
Chashma	GEV	458,672.751	588,334.823	667,140.008	737,859.630	822,809.717	881,928.704	1,005,048.794
Taunsa	GLO	445,711.906	556,101.160	624,760.857	690,846.241	778,710.236	847,018.979	1,014,713.265
Guddu	GEV	571,380.508	833,758.752	997,678.599	1,147,961.937	1,332,864.528	1,464,640.529	1,748,351.612
Sukkur	GPA	483,104.035	828,009.585	1,015,117.387	1,155,933.267	1,290,098.660	1,362,882.040	1,469,847.471
Kotri	GLO	295,709.114	522,116.107	727,002.354	985,651.369	1,448,912.443	1,928,001.511	3,716,497.052
Mangla	GLO	96,348.508	169,097.689	238,056.774	328,132.863	495,683.123	675,253.276	1,383,050.346
Rasul	GLO	86,965.836	172,632.289	257,317.727	371,419.800	591,093.095	834,275.297	1,841,644.643
Marala	GPA	249,137.657	448,900.141	589,575.977	721,853.822	884,674.980	999,336.163	1,239,863.617
Khanki	GPA	270,353.870	490,663.494	667,820.002	854,562.187	1,117,018.401	1,328,066.043	1,863,198.868
Qadirabad	GPA	285,960.094	544,535.730	717,548.006	873,249.863	1,055,429.295	1,177,325.072	1,415,379.292
Trimmu	GEV	209,501.913	376,582.998	508,044.272	652,433.656	870,164.499	1,059,666.107	1,603,173.150
Panjnad	GPA	213,767.329	421,665.868	545,824.359	647,220.438	753,309.441	816,666.551	922,544.569
Balloki	GPA	66,700.085	120,525.407	165,443.818	214,336.178	285,668.900	345,197.620	504,527.505
Sidhani	GPA	45,112.590	91,425.039	131,036.210	175,079.806	240,955.113	297,298.453	453,648.241
Sulemanki	GPA	39,793.243	104,949.643	164,935.000	236,032.038	350,619.247	456,112.410	782,665.744
Islam	GPA	25,998.518	71,615.964	115,457.015	169,398.850	260,233.840	347,531.614	635,816.096

Table 4.7.2: Quantile estimates for each Sites for selected return periods using that distribution which is best for that site. (TL-Moments)

Sites Name	Best Distribution	0.500 2	0.800 5	0.900 10	0.950 20	0.980 50	0.990 100	0.998 500
Tarbela	GEV	408,234.755	470,679.870	521,973.154	580,259.336	671,602.969	754,136.053	1,004,284.583
Kalabagh	GLO	511,919.527	617,375.535	702,923.093	802,621.085	966,097.151	1,121,570.008	1,633,300.960
Chashma	GLO	528,354.317	627,706.367	704,269.517	790,231.377	925,515.611	1,049,234.684	1,433,484.403
Taunsa	GLO	510,853.934	601,829.457	667,345.678	737,389.416	841,850.923	932,584.524	1,193,874.374
Guddu	GPA	721,690.352	988,290.532	1,133,262.039	1,242,586.358	1,346,987.539	1,403,758.681	1,487,454.028
Sukkur	GEV	676,893.954	931,109.499	1,098,455.166	1,258,256.951	1,464,061.903	1,617,516.774	1,969,630.468
Kotri	GPA	427,093.088	698,663.097	952,085.450	1,256,005.741	1,752,633.534	2,216,074.192	3,680,059.219
Mangla	GEV	142,827.216	220,208.092	298,453.985	403,374.718	601,917.721	816,989.611	1,674,586.163
Rasul	GEV	145,578.187	242,903.768	341,540.943	474,045.506	725,308.429	998,030.477	2,088,846.046
Marala	GPA	360,374.836	557,118.179	689,817.820	810,063.133	951,876.193	1,047,526.404	1,236,418.690
Khanki	GPA	393,503.629	621,276.618	801,609.575	989,165.701	1,248,680.142	1,454,143.514	1,963,515.889
Qadirabad	GPA	427,817.425	675,454.622	836,496.510	977,936.001	1,138,839.276	1,243,476.897	1,439,925.338
Trimmu	GEV	318,099.065	481,043.953	606,888.041	743,095.279	945,108.040	1,118,110.158	1,602,676.271
Panjnad	GPA	319,991.077	507,471.329	626,888.022	729,924.352	844,767.245	917,917.127	1,051,381.212
Balloki	GPA	96,157.620	151,195.021	198,344.996	250,844.609	329,496.235	396,876.274	584,298.971
Sidhani	GPA	70,697.311	118,791.641	160,794.541	208,350.116	280,997.403	344,443.485	526,011.915
Sulemanki	GPA	78,526.262	151,758.076	216,834.011	291,630.155	407,919.398	511,257.426	814,693.509
Islam	GPA	54,008.070	107,152.122	155,782.439	213,117.796	304,956.073	388,994.161	646,781.230

We have also drawn extreme value plots for best fitted distributions. The extrem value plot for four sites are shown in the following figures. It is observed from these plots that quantiles based on best fitted distribution are in close agreement with observed flows.

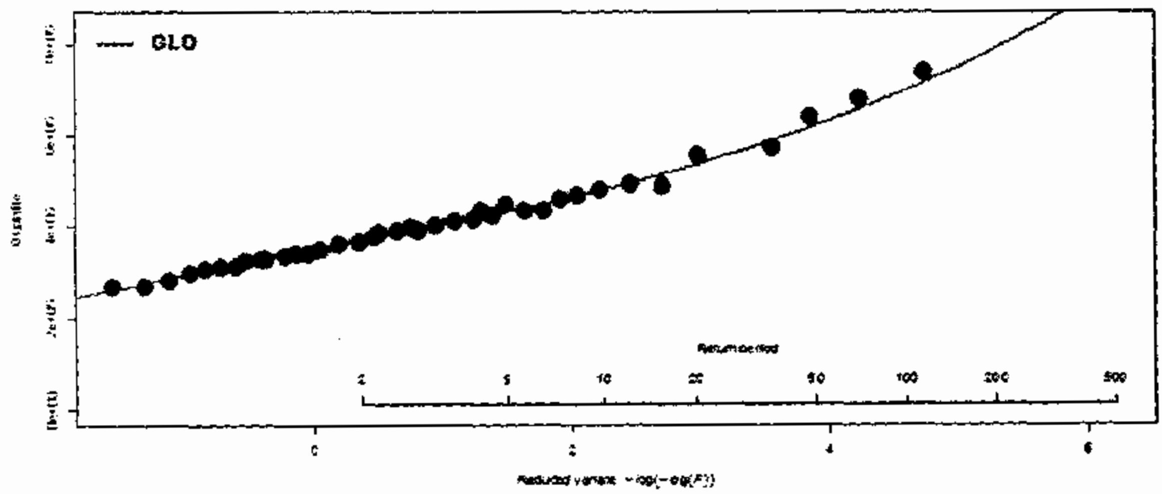


Fig 4.7.1. Extreme value plot for Tarbela site using L-Moment

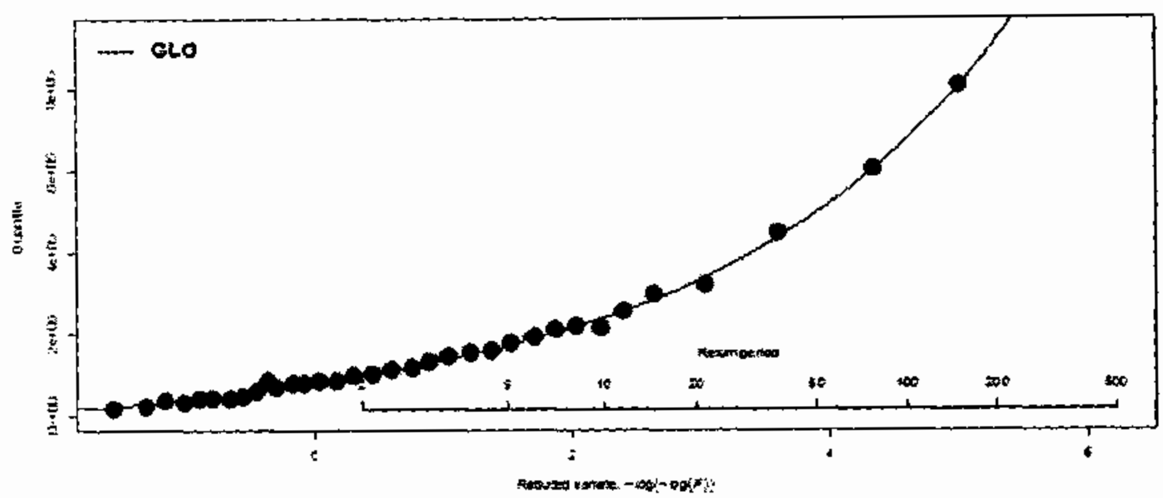


Fig 4.7.2. Extreme value plot for Mangla site using L-Moment

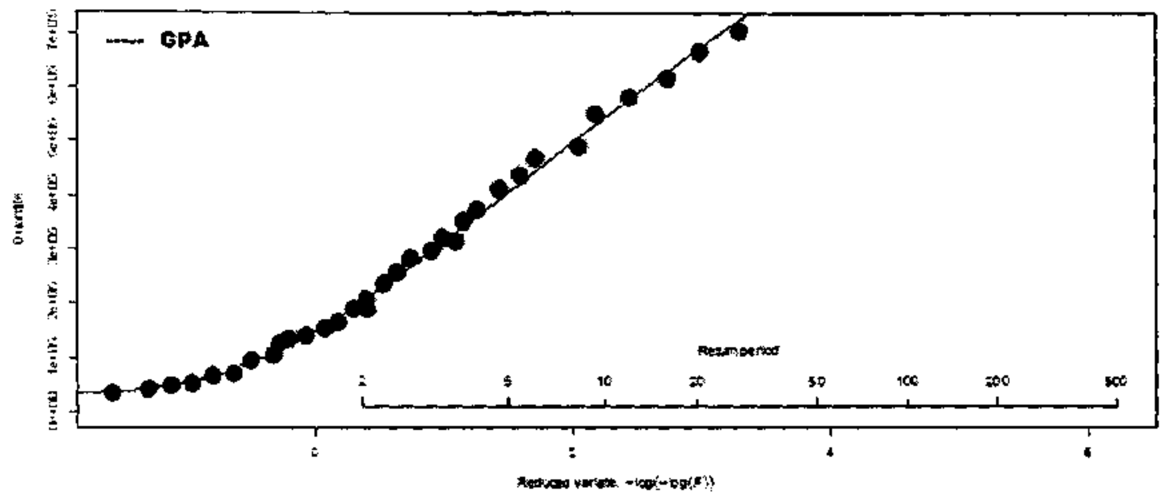


Fig 4.7.3. Extreme value plot for Trimmu site using TL-Moment

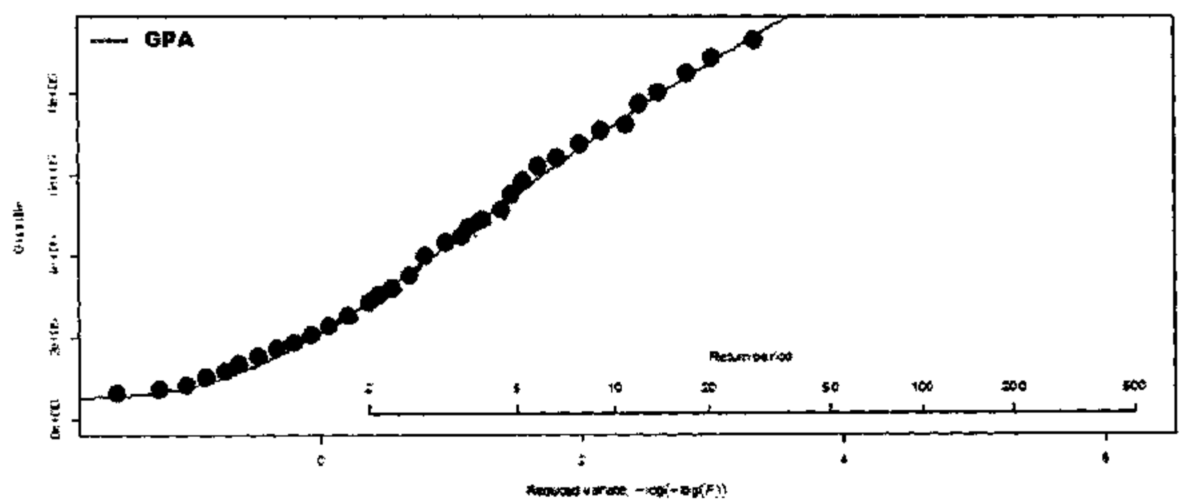


Fig 4.7.4. Extreme value plot for Qadirabad site using TL-Moments

CHAPTER 5

Summary and Conclusion

FFA has been performed on 18 sites of Pakistan. The AM flows data was taken from Water and Power Development Authority (WAPDA) and Federal Flood Commission (FFC) of Pakistan. The AM flows was being measured in cusec. The record length of AM flows varies from 32 to 92 year. (Sukkur 32 & Balloki 92).

The AM flows of 18 sites of Pakistan being studied, are located on the five rivers namely Indus, Jhelum, Chenab, Ravi, and Sutlej. Initially the basic assumptions of FFA are tested by different tests that are time series plots, Mann-Whitney test, Mann-Kendall's tau test, Ljung-Box-Q-Statistics and CPA. All sites satisfied these tests, which means that observations at any site are independent, stationary and identically distributed. Therefore data of all the 18 sites were used for further analysis.

The sample L-moments, TL-moments, L-ratios and the TL-ratios are very useful for summarizing statistical properties of hydrological data and can be used for parameters estimation of distribution and choice of best fit distribution. The three selected distributions GLO, GEV, and GPA are being applied to the AM flows of 18 sites.

Then the distributions are compared by using goodness-of-fit tests as MADI, RRMSE, and PPCC. While for graphical representation of best fit distributions, the L-Moment and TL-Moment ratio diagrams are being used. In the case of L-moments the result showed that for the most of the sites GPA are best fitted distribution, followed by GLO distribution and GEV distribution. The overall results in the percentages of the best fitted distributions with respect to different sites based on the

L-moments are, for GPA 50%, GLO 27.77% and GEV 22.22%. In the case of TL-moments the overall results in percentage of best fitted distributions are, GPA 55.55%, GLO 16.66% and GEV 27.77%. It is observed that if we compare the L-moments and TL- moments on the basis of Ratio diagrams, TL-moments-ratio diagram give better fit than the L-moment ratio diagram. Our results of graphical method of L-ratio and TL-ratio diagrams coincide with results for other goodness-of-fit tests. For policy implication the GPA distribution can be used for AM flows at these sites. It can be used for estimates of flood magnitude.

The Quantile for any return periods has the probability of occurrence/exceedance equal $(\frac{1}{T})$ for any given return period "T". The quantile estimates $\hat{q}(F)$ for each site with exceedance probability and return periods are calculated for three selected distributions. In the study quantile estimates are found for those distribution which are best fitted for that site. We noted that for shorter return period and high return period quantile estimates for TL-moment is larger than L-moment. It can be suggested that gauging networking system should be applied and increased in the country to improve national water resources planning and development.

These return periods and subsequently estimates of quantiles are very significant in the design of dams, bridges, culverts, flood controlling devices.

Recommendations for the Future Study

In this study there are also some recommendations and suggestions which are given below for future research and developments of this study.

1. For future study the method of L-moment and TL-moment may be compared with other estimation methods such as Maximum likelihood method or the Method of Moments.
2. The study may be conducted using other estimation methods like LQ-moments, LH-moments and partial L-moments, then estimated quantiles can be compared with Trimmed L-moments.
3. The Regional Flood Frequency Analysis may be conducted for TL-moments and then compare with the regional study of L-moments.

References

1. Abida, H., & Ellouze, M. (2008). Probability distribution of flood flows in Tunisia. *Hydrology and Earth System Sciences, 12*(3), 703-714.
2. Afreen, S., & Muhammad, F. (2012). Flood frequency analysis of various dams and barrages in Pakistan. *Irrigation and Drainage, 61*(1), 116-128.
3. Ahmad, I., Shah, S. F., Mahmood, I., & Ahmad+, Z. (2013). MODELING OF MONSOON RAINFALL IN PAKISTAN BASED ON KAPPA DISTRIBUTION. *Science International, 25*(2), 333-336.
4. Ahmad, M. I., Sinclair, C. D., & Werritty, A. (1988). Log-logistic flood frequency analysis. *Journal of Hydrology, 98*(3), 205-224.
5. Ahmad, U. N., Shabri, A., & Zakaria, Z. A. (2011). Flood frequency analysis of annual maximum stream flows using L-moments and TL-moments approach. *Applied Mathematical Sciences, 5*(5), 243-253.
6. Ahmad, U. N., Shabri, A., & Zakaria, Z. A. (2011). TL-moments and L-moments estimation of the generalized logistic distribution. *Journal of Mathematics Research, 3*(1), p97.
7. Ahmad, U. N., Shabri, A., & Zakaria, Z. A. (2011). Trimmed L-moments (1, 0) for the generalized Pareto distribution. *Hydrological Sciences Journal, 56*(6), 1053-1060.
8. Asquith, W. H. (2007). L-moments and TL-moments of the generalized lambda distribution. *Computational Statistics & Data Analysis, 51*(9), 4484-4496.

9. Bhattara, K. P. (2005). Flood frequency analysis of Irish river flow data using variants of L-Moment. In *National Hydrology Seminar.*

10. Bílková, D. (2014). Trimmed L-Moments: Analogy of Classical L-Moments. *American Journal of Mathematics and Statistics*, *4*(2), 80-106.

11. Chadwick, A., Morfett, J. & Borthwick, M. (2004). *Hydraulics in Civil and Environmental Engineering,* Spon Press, London.

12. Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). *Applied hydrology.*

13. Cunnane, C. (1987). Review of statistical models for flood frequency estimation. In *Hydrologic frequency modeling* (pp. 49-95). Springer Netherlands.

14. Cunnane, C. (1989). Statistical distributions for flood frequency analysis. *Operational Hydrology Report (WMO).*

15. Dalén, J. (1987). Algebraic bounds on standardized sample moments. *Statistics & probability letters*, *5*(5), 329-331.

16. Deka, S., & Borah, M. (2012) Statistical Analysis of Flood Peak data of North Brahmaputra Region of India based on the methods of TL-moment [online]. http://interstat.statjournals.net/YEAR/2012/articles/1206002.pdf[accessed4 September 2014].

17. Deka, S., Borah, M., & Kakaty, S. C. (2009). Distributions of Annual Maximum Rainfall Series of North-East India. *European Water*, *27*(28), 3-14.

18. Elamir, E. A. H. (2010). Optimal choices for trimming in trimmed L-moment method. *Applied Mathematical Sciences*, *4*(58), 2881-2890.

19. Elamir, E. A., & Seheult, A. H. (2003). Trimmed L-moments. *Computational Statistics & Data Analysis*, *43*(3), 299-314.

20. Galoie, M., Zenz, G., & Motamedi, A (2013). Rainfall Analysis for the Schoeckelbach Basin (Austria) and Determining its Best-Fit Probability Distribution Model. *Water and Environmental Dynamics*, 6th International Conference on Water Resources and Environmental Research was held at Koblenz (Germany).

21. Gingras, D., & Adamowski, K. (1992). COUPLING OF NONPARAMETRIC FREQUENCY AND L-MOMENT ANALYSES FOR MIXED DISTRIBUTION IDENTIFICATION 1. *JAWRA Journal of the American Water Resources Association*, *28*(2), 263-272.

22. Gocic, M., Velimirovic, L., & Stankovic, M. (2014). Determining the best fitting distribution of annual precipitation data in Serbia using L-moments method. *ArXiv preprint arXiv: 1410.5309.*

23. Gubareva, T. S., & Gartsman, B. I. (2010). Estimating distribution parameters of extreme hydrometeorological characteristics by L-moments method. *Water resources*, *37*(4), 437-445.

24. Hamed, K. H. (2008). Trend detection in hydrologic data: the Mann–Kendall trend test under the scaling hypothesis. *Journal of Hydrology*, *349*(3), 350-363.

25. Hosking, J. R. M. (1986). *The theory of probability weighted moments.* IBM Research Division, TJ Watson Research Center.

26. Hosking, J. R. M. (2007). Some theory and practical uses of trimmed L-moments. *Journal of Statistical Planning and Inference*, *137*(9), 3024-3039.

27. Hosking, J. R. M., & Wallis, J. R. (1997). *Regional frequency analysis: an approach based on L-moments.* Cambridge University Press, United Kingdom.

28. Hosking, J.R.M. (1990). L-Moments, Analysis and estimation of distributions using linear combinations of order statistics. *Journal Royal Statistics Society. Series B*, *52*(1), 105-124.

29. Izinyon, O. C., & Ehiorobo, J. O. (2014). L-moments approach for flood frequency analysis of river Okhuwan in Benin-Owena River basin in Nigeria. *Nigerian Journal of Technology*, *33*(1), 10-18.

30. Jain, D., & Singh, V.P. (1986). A comparison of Transformation Methods for Flood Frequency Analysis. *Water Resources Bulletin*, *22*(6), 903-912.

31. Johnson, N. L., & Kotz, S. (1970). Distributions in Statistics: Continuous Univariate Distributions. *Wiley Series in Probability and Mathematical Statistics.*

32. Karim, M. A., & Chowdhury, J. U. (1995). A comparison of four distributions used in flood frequency analysis in Bangladesh. *Hydrological Sciences Journal*, *40*(1), 55-66.

33. Karim, M.A., & Chowdhury, J.U. (1993). A comparison of five distributions used in flood frequency analysis in Bangladesh. In *Proc. Int. Conf. on Hydro. and Water Resour.*

34. Kjeldsen, T.R., & Jones, D.A. (2004). Sampling Variance of flood quantiles from the generalised logistic distribution estimated using method of L-moment. *Hydrology and Earth System Sciences*, 8(2), 183-190.

35. Kroll, C. N., & Vogel, R. M. (2002). Probability distribution of low streamflow series in the United States. *Journal of Hydrologic Engineering*, *7*(2), 137-146.

36. Lee, S. H., & Maeng, S. J. (2003). Comparison and analysis of design floods by the change in the order of LH-moment methods. *Irrigation and drainage*, *52*(3), 231-245.

37. Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*(2), 297-303.

38. Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The annals of mathematical statistics*, 50-60.

39. Mann, H. B., (1945) Nonparametric tests against trend: *Econometrica: Journal of the Econometric Society*, 245-259.

40. Nathan, R. J., & Weinmann, P. E. (1991). Application of at-site and regional flood frequency analyses. In *National Conference Publication- Institute of Engineers. Australia.*

41. Önöz, B., & Bayazit, M. (1995). Best-fit distributions of largest available flood samples. *Journal of Hydrology, 167*(1), 195-208.

42. Önöz, B., & Bayazit, M. (1999). GEV-PWM model for distribution of minimum flows. *Journal of Hydrologic Engineering, 4*(3), 289-292.

43. Pearson, C. (1992). New Zealand regional flood frequency analysis using L-moments. *Journal of Hydrology.*

44. Pilon, P. J., & Adamowski, K. (1992). The value of regional information to flood frequency analysis using the method of L-moments. *Canadian Journal of Civil Engineering, 19*(1), 137-147.

45. Rao, A.R., & Hamed, K.H. (2000). *Flood Frequency Analysis*, CRC Press, Boca Raton, London, New York, Washington, D.C.

46. Reed, D. W., & Robson, A. J. (1999). Flood Estimation Handbook Volume 3: Statistical procedures for flood frequency estimation. *Institute of Hydrology: Wallingford.*

47. Royston, P., & Thompson, S. G. (1992). Model-based screening by risk with application to down's syndrome. *Statistics in medicine, 11*(2), 257-268.

48. Saf. B. (2009). Regional flood frequency analysis using L-moments for the West Mediterranean region of Turkey. *Water resources management, 23*(3), 531-551.

49. Salarpour, M., Yusop, Z., & Yusop, F. (2013). Comparison of distribution Models for Peak flow, Flood Volume and Flood Duration. *Research Journal of Applied Sciences, Engineering and Technology, 6*(4), 733-738.

50. Shabri, A., & Ariff, N. (2009). Frequency analysis of maximum daily rainfalls via l-moment approach. *Sains Malaysiana, 38*(2), 149-158.

51. Tao, D. Q., Nguyen, V. T., & Bourque, A. (2002, June). On selection of probability distributions for representing extreme precipitations in Southern Quebec. In *Annual Conference of the Canadian Society for Civil Engineering. 5th* (pp. 1-8).

52. Tasker, G. D. (1987). A comparison of methods for estimating low flow characteristics of streams. *Water Resources Bulletin, 23* (6), 1077-1083.

53. Taylor .W. (2011). Change-point analysis: a powerful new tool for detecting changes. Taylor Enterprises, Libertyville, Illinois. *Web: http://www.variation.com/cpa.*

54. Taylor, W. (2000a). Change-Point Analyzer 2.0 shareware program, Taylor Enterprises, Libertyville, Illinois. *Web: http://www. Variation. Com/cpa.*

55. Vogel, R. M., & Fennessey, N. M. (1993). L moment diagrams should replace product moment diagrams. *Water Resources Research, 29*(6), 1745-1752.

56. Vogel, R. M., & Wilson, I. (1996). Probability distribution of annual maximum, mean, and minimum streamflows in the United States. *Journal of hydrologic Engineering, 1*(2), 69-76.

57. Vogel, R. M., McMahon, T. A., & Chiew, F. H. (1993). Flood flow frequency model selection in Australia. *Journal of Hydrology, 146*, 421-449.

58. Wallis, J. R. (1988). *Catastrophes, computing and containment: Living with our restless habitat.* IBM Thomas J. Watson Research Division.

59. Wang, Q. J. (1990). Estimation of the GEV distribution from censored samples by method of partial probability weighted moments. *Journal of Hydrology*, *120*(1), 103-114.

60. Wang, Q. J. (1996). Direct sample estimators of L moments. *Water resources research*, *32*(12), 3617-3619.

61. Yurekli, K., Kurunc, A., & Gul, S. (2005). Frequency Analysis of Low Flow Series from Cekerek Stream Basin. *Journal of Agricultural Sciences*, *11*(1), 72-77.

62. Zaidman, M.D., Keller, V., and Young, A.R., (2002) Probability distributions for x-day daily mean flow events: final technical Report. *Center for Ecology and Hydrology*, Wallingford, UK 247pp.

63. Zin, W. Z. W., Jemain, A. A., & Ibrahim, K. (2009). The best fitting distribution of annual maximum rainfall in Peninsular Malaysia based on methods of L-moment and LQ-moment. *Theoretical and applied climatology*, *96*(3-4), 337-344.