Jimma University
School of Graduate Studies
Jimma Institute of Technology
Faculty of Civil and Environmental Engineering
Hydrology and Hydraulic Engineering Chair
Masters of Science Program in Hydraulic Engineering

Evaluation of Predictive Accuracy of Regional Flood Frequency Estimations: A Case Study of Tekeze River Basin, Ethiopia

A Thesis Submitted to the School of Graduate Studies of Jimma University Jimma Institute of Technology in Partial fulfillment of the requirements for the Degree of Masters of Science in Hydraulic Engineering

By: Tarekegn Zeleke Gela

December, 2019

Jimma, Ethiopia

Jimma University
School of Graduate Studies
Jimma Institute of Technology
Faculty of Civil and Environmental Engineering
Hydrology and Hydraulic Engineering Chair
Masters of Science Program in Hydraulic Engineering

Evaluation of Predictive Accuracy of Regional Flood Frequency Estimations: A Case Study of Tekeze River Basin, Ethiopia

A Thesis Submitted to the School of Graduate Studies of Jimma University Jimma Institute of Technology in Partial fulfillment of the requirements for the Degree of Masters of Science in Hydraulic Engineering

By: Tarekegn Zeleke Gela

Main Advisor: Prof., Dr. - Ing. Esayas Alemayehu
Co-advisor: Mr. Kiyya Tesfa (MSc.)

December, 2019

Jimma, Ethiopia

DECLARATION

I hereby declare that the research entitled “**Evaluation of Predictive Accuracy of Regional Flood Frequency Estimations: A Case Study of Tekeze River Basin, Ethiopia**” is my work which I submit for partial fulfillment of the degree of Master of Science in Hydraulic Engineering to Jimma University, school of graduate studies, Jimma Institute of Technology, and Hydrology and Hydraulic Engineering Chair. The research was conducted under the guidance of main advisor Prof., Dr. - Ing Esayas Alemayehu and co-advisor Mr. Kiyya Tesfa (MSc.).

Tarekegn Zeleke Gela ______________ _____________
Candidate Signature Date

Prof., Dr.- Ing. Esayas Alemayehu ________________ _____________
(Main Advisor) Signature Date

Mr. Kiyya Tesfa (MSc.) ________________ _____________
(Co-Advisor) Signature Date

APPROVAL

The thesis entitled "**Evaluation of Predictive Accuracy of Regional Flood Frequency Estimations: A Case Study of Tekeze River Basin, Ethiopia**" submitted by Tarekegn Zeleke Gela is approved and accepted as a Partial Fulfillment of the Requirements for the Degree of Masters of Science in Hydraulic Engineering at Jimma Institute of Technology.

Name	Signature	Date
Main Advisor: Prof., Dr.-Ing Esayas Alemayehu		
Co-Advisor: Kiyya Tesfa (MSc.)		

As members of the examining board of MSc. thesis, we certify that we have read and evaluated the thesis prepared by Tarekegn Zeleke Gela. We recommend that the thesis could be accepted as a Partial Fulfillment of the Requirements for the Degree of Masters of Science in Hydraulic Engineering.

	Name	Signature	Date
Chairman:	Birhan Tekuame (MSc.)		31/12/2019
External Examiner:	Dr. Zelalem Biru		31/12/2019
Internal Examiner:	Tolera Abdisa (MSc.)		31/12/2019

ABSTRACT

Floods are among the most recurring and devastating natural disasters and are responsible for significant loss of life and property throughout the world. It causes physical suffering, economic losses, limit the efficiency of drainage, and disturb existence of life in the study area. An evaluation of predictive accuracy of regional flood frequency estimation methods has been the backbone of water resources project planning, design of any structures and the economic analysis of flood control projects. It is due to the fact that floods represent the most disastrous natural event causing several damages to enormous economic and life losses in the study area. The aim of this study was to evaluate the predictive fit of probability distributions to annual maximum flood data, and in particular to evaluate which combination of distribution and estimation method gives the best fit of Tekeze River Basin. Subsequently, the probability distribution fits were evaluated according to several goodness-of-fit measures and to the variability of the predicted flood quantiles. To achieve this, based on data from eleven stream gauged sites, three hydrological homogeneous sub regions were defined and delineated based on L-moment homogeneity tests, namely Region-A, Region-B and Region-C. Delineation of homogeneous regions were accomplished using ArcGIS10.4.1. Discordancy of regional data of the L-moment statistics was identified using Matlab2018a. All regions have shown satisfactory results for discordance measures and homogeneity tests. For the regions, best-fit distributions were selected. L-moment ratio diagrams and Easy Fit statistical software was used to select best-fit probability distributions. The performances of the distributions were evaluated using Kolmogorov Smirnov, Anderson-Darling and Chi-Squared goodness-of-tests. After three goodness of fit tests was carried out, generalized extreme value (GEV) with MOM for Region-A and generalized Pareto (GPA) with PWM for Region-B and C were identified as suitable distributions for analyzing accurate annual maximum flows in the basin. Based on best-fit distributions for the three regions, regional flood frequency curves were constructed. In this study, the flood magnitude is estimated for 2, 5, 10, 15, 20, 25, 50, 75, 100, 200, 500 and 1000 years return period. The derived flood frequency curves at a given return period suggested that how important engineering decisions and actions such as design and operation of the water resources project have to be undertaken. Consequently, statistical analysis of gauged sites was revealed an acceptable method of regionalization. Finally, the study can be further extended into flood hazard, risk and inundation mapping of identified regions of the study area.

Keywords: *Best-fit statistical distribution; Homogeneity; L-moment; Parameter estimation methods; Regionalization.*

ACKNOWLEDGEMENTS

Above all I thank the Almighty God for his mercy, and give me patience, audacity, wisdom and who made it possible, to begin and finish this thesis successfully. Without his support and blessings this piece of work would never have been accomplished and I know that all are in his name.

I would like to express my deep to my advisors Prof., Dr.-Ing Esayas Alemayehu and Kiyya Tesfa (MSc.) for their persistent, support, constructive guidance and advice throughout the work.

I express my gratitude to all staff members of Ministry of Water Irrigation and Electricity, particularly department of GIS and Hydrology, for their appreciable support in providing me relevant data and other related reference materials from their library.

I would like to acknowledge Ethiopia Roads Authority the financial support made through postgraduate program in Jimma University. I also thank to Jimma Institute of Technology, Faculty of Civil and Environmental Engineering and Jimma University community for their support.

I am very grateful to Tarekegn Dejene and Chala Hailu for their support, valuable advice and follow up with their research experience.

Last but not least, I would like to thanks to all my families especially to my sister Adey Zeleke and friends who helped me a lot in carrying out my thesis through remarkable encouragement, advise, endless support and comments in several ways for the realization of this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

AMF	Annual Maximum Flow
CC	Combined Coefficient of variation
DEM	Digital Elevation Model
FFA	Flood Frequency Analysis
FFC	Flood Frequency Curve
GEV	Generalized Extreme Value distribution
GIS	Geographical Information System
GPA	Generalized Pareto distribution
IHR	Identification of Homogeneous Regions
LCk	Linear Coefficient of Kurtosis
LCs	Linear Coefficient of Skewness
LCV	Linear Coefficient of Variation
Matlab	Mathematics Laboratory
MML	Method of Maximum Likelihood
MOM	Method of Moment
PWM	Probability Weighted Method
Q	Discharge
Q-T	Discharge-Return Period
RFFA	Regional Flood Frequency Analysis
SEE	Standard Error of Estimate
T	Return Period
XT	Estimated flow quantiles value

1. INTRODUCTION

1.1. Background

Floods are among the catastrophic natural events that cause severe consequences for human society and are responsible for significant loss of life and property throughout the world. It causes damages to properties, agricultural lands, economic losses, limit the efficiency of drainage, and disturb existence of life (Yucel and Keskin, 2011).

The flood prone damages are common in different parts in Ethiopian regions; this is due to lack of adequately studied information and prevention mechanism. The frequency and magnitude of floods have increased, affecting large parts of the country and causing damage to property, loss of life, and the health of the populations (Akirso, 2017).

Flood frequency estimation is essential for flood management. It is used to map floodplain areas, design hydraulic structures (dams, retaining basins, storm water systems) and infrastructures (roads, bridges), and define the frequency of flood events for natural disaster assessments and alert methods (Javelle *et al*., 2010).

The study of hydrological hazard uses flood frequency analysis (FFA) and has led to the development of various methods ranging from purely statistical approaches to simulation approaches. The development of these methods is often influenced by the availability of observation data and by the objectives to be met (Boughton and Droop, 2003; Castellarin *et al*., 2011; Pathiraja *et al*., 2012).

Flood frequency analysis is an important factor in flood risk assessment studies and for the design of various hydraulic structures. Flood quantiles estimates are required at locations where stream flow series are very short or where no data are available, making a direct flood frequency analysis impossible. Regional flood frequency analysis such as the index flood method (Jingyi and Hall, 2004; Kjeldsen and Jones, 2007; Das and Cunnane, 2011; Malekinezhad *et al*, 2011; Zaman *et al*, 2012) offers a solution to this problem and is widely used to estimate flood quantiles in these situations.

The accurate estimation of design floods remains one of the major challenges for many engineers and planners who are involved in project design where hydrological data and information are limited. Flood frequency analysis involves estimation of a flood magnitude corresponding to a required return period or probability of exceedence. Flood estimation is important for design and safety assessments, flood risk management schemes and spatial planning (Lim and Lye, 2003). To reduce flood damage and save human lives, flood modeling is generally undertaken to estimate floods associated with return periods of interest, which is called design flood. Design flood estimation is needed for various purposes including design of hydraulic structures, flood plain management, development and planning controls, and flood insurance studies (Zaman, 2013).

A flood frequency analysis consists of a study of past records of flow discharge and an estimate of frequencies of future floods. If adequate records are available the common methods give acceptably uniform results within the range of data (Badreldin and Fengo, 2012). Flood frequency analysis is based on the analysis of observed historical flood events and estimates the magnitudes of floods with a given return period. Flood frequency analysis is an important factor in flood risk assessment studies and for the design of various hydraulic structures. The primary objectives of flood frequency analysis are to determine the return periods and then to estimate the magnitudes of events for design return periods beyond the recorded range (Mustapha and Yusuf, 2012).Flood frequency analysis is usually carried out to estimate flood quantities, at a project location for a return period. The analysis primarily uses annual maximum flood data observed at a desired project location or gauging station to estimate flood quintiles.

Regional flood frequency analysis enables estimation of flood magnitude in different return periods at any stream location within a region (Solana and Solana, 2001; Atiem and Harmancioglu, 2006) to improve the at-site estimates by using the available flood data within a region and attempt to respond to the need of flood estimation in ungauged basins. Thus, it allows flood quantile estimation for any site in a region to be expressed in terms of flood data recorded at all gauging sites in the same region, including those at the specific site (Harmancioglu, 2006).

To estimate the frequency and magnitude of floods for design purposes, the availability of stream flow data is a fundamental requirement. Flood frequency analysis is often used by practitioners to support the design of river engineering works, flood mitigation measures and civil protection strategies. It is generally carried out by fitting peak flow observations to a suitable probability distribution (Baratti, 2012).

Flood frequency analysis provides vital information for design and economic appraisal of a variety of engineering and water resources planning and development projects. Frequency analysis of flood is a very active area of investigation in statistical hydrology. Various distributions, methods of estimation of parameters, problems related to regionalization and other related topics are being investigated. The analysis involves estimation of a flood magnitude corresponding to a required return period or probability of exceedance (Mengistu, 2008).

The aim of this study was to evaluate the predictive fit of probability distributions to annual maximum flood data, and in particular to evaluate which combination of distribution and estimation method gives the best fit of Tekeze River Basin.

1.2. Statement of the Problem

The main feature of a flood, from the water management point of view, is its interference with human activities. The interference is measured in terms of actual and potential economic losses and danger to human life. The purpose of flood analysis is to assess the magnitude and frequency of this interference. This flood frequency analysis provides vital information for the planning and design of many hydraulic structures and for risk assessment in flood plain uses (Mengistu and Sivakumar, 2018).

The estimation of flood quantiles is complicated because of both lack of a physical basis for determining the form of the underlying flood frequency distribution and the necessity of evaluating flood risk for return periods that exceed the length of the observed record. Flood quantile estimates are strongly dependent on the form of a portion of the underlying flood frequency distribution and that is difficult to estimate from observed data. Regional flood frequency analysis is becoming an important subject because of most structures are constructed in areas where recorded flood data are either missing or inadequate (Demissie, 2008).

Tekeze River Basins have sparse network of observation sites with short record length of observed flow that makes the use of single site analysis to estimate design parameters at many potential project sites unreliable due to lack of fund and qualified person, density of gauging station is low, and the operation and maintenance of stream gauging networks are difficult, so reliable estimation of the flow quantile is difficult for design of hydraulic structures such as culvert, spillway, bridge, reservoir and dikes and for integrated water resources management such as water supply, irrigation and hydropower and reducing flood induced losses.

1.3. Objective of the Study

1.3.1. General Objective

The general objective of this study is evaluating predictive accuracy of regional flood frequency estimations on Tekeze River Basin of Ethiopia.

1.3.2. Specific Objectives

The specific objectives of the study are:

1. To identify and delineate hydrologically homogeneous regions of the basin;
2. To identify the best fit statistical distributions and quantile estimation for the basin;
3. To develop regional frequency curves for the delineated homogeneous region.

1.4. Research Questions

1. How hydrologically homogeneous regions of the basin for regional flood estimation is identified and delineated?
2. What are the best-fit probability distributions for prediction of hydrological events of gauging stations of the basin?

3. How regionalization method is used for regional flood frequency analysis in the future?

1.5. Significance of the Study

This study will be used for design and safety assessments, flood risk management schemes and spatial planning, economic evaluation of flood control projects, proper planning, and design of water resources management options on the study area. In addition to this, the study can be used for policy and decision makers (higher government bodies), water resource planner, hazard management bodies, Disaster Management and Food Security Sectors (DMFSS) and for researchers on flood modeling.

1.6. Scope of the Study

This study investigates the flood problem which was occurred in the selected basin. Based on the available data a possible flood mitigation alternative measure is worked out. The study is limited mainly on regionalization of stream flow data on the Tekeze River Basin, Ethiopia.

1.7. Limitation of the Study

The problem faced through this study was lack of sufficient and reliable data regarding the evaluation of predictive accuracy of regional flood frequency estimations on the river basin.

2. LITERATURE REVIEW

2.1. Flood Frequency Analysis

Extreme flow quantiles estimated from stream flow data provides vital information for engineering design of any project and economic appraisal of a variety of engineering and water resources planning and development projects, when there is sparse of observed flow data. Extreme flow quantiles are estimated from observed flow data using flood frequency analysis. Flood frequency analysis is a hydrologic field dealing with estimation of a flood magnitude corresponding to any required return period of occurrence (Rao and Srinivas, 2008; Bhagat, 2017; Kanti *et al*., 2017).

Flood frequency analyses are used to predict design floods for sites along a river. The technique involves using observed annual peak flow discharge data to calculate statistical information such as mean values, standard deviations, skewness, and recurrence intervals. These statistical data are then used to construct frequency distributions, which are graphs and tables that tell the likelihood of various discharges as a function of recurrence interval or exceedence probability (Jos, 2017).

Flood frequency analysis uses historical records of peak flows to produce guidance about the expected behavior of future flooding. Primary applications of flood frequency analyses are to predict the possible flood magnitude over a certain time period and to estimate the frequency with which floods of a certain magnitude may occur (Sah and Pradas, 2017).

Flood frequency analysis involves the fitting of a probability model to the sample of annual flood peaks recorded over a period of observation, for a catchment of a given region. The model parameters established can then be used to predict the extreme events of large recurrence interval (Pegram and Parak, 2004) reliable flood frequency estimates are vital for floodplain management; to protect the public, minimize flood related costs to government and private enterprises, for designing and locating hydraulic structures and assessing hazards related to the development of flood plains (Tumbare, 2000).

Flood Frequencies are highly affected by the physical and climatic characteristics of the catchments such as storm duration, intensity, and magnitude, catchment size, shape, relief, drainage density, morphology, land cover, presence or absence of storage, soil type, and land use (Baguis, 2008).

2.2. Flood Estimation Techniques

A significant difficulty in hydrology is the evaluation of flood magnitudes, mainly as planning and design of water resource projects and flood plain management depend on the frequency and magnitude of peak discharges (Saf, 2009). Floods have become more frequent and severe due to effects of global climate change and human alterations of the natural environment. All flood forecasting systems serve specific purposes and in most cases they are designed to prevent, minimize, or mitigate people's suffering and to limit economic losses. The forecast of flooding would benefit greatly from the use of hydrological models, which are designed to simulate flow processes of surface or subsurface water.

The literature identified two comprehensive methods for flood frequency analysis, statistical and derived. Statistical flood frequency analysis is the modern method of determining the frequency of peak stream flows. This method of frequency analysis involves fitting extreme value probability distribution functions to the historical record of annual maximum floods. This method is reliant upon the availability of observed stream flow to fit suitable probability distributions relevant to gauged sites (Kumar and Chatterjee, 2011; Vivekanandan, 2015). The derived techniques of flood frequency analysis involve the quantification of the processes that govern flood behavior which is less dependent upon historical data (Badreldin and Fengo, 2012).

In the statistical analysis of floods extreme value probability distribution are fitted to calculated peak flows. This technique is data serious and is relevant just to gauge station. Choice of possibility distribution is commonly random; no physical source is accessible to reduce the use of any particular distribution (Yue and Wong, 2004).

2.3. Flood Frequency Models

The purpose of flood frequency analysis is to determine a Q-T link at any essential site along a river. At any river site it is frequently assumed that character provides an exclusive Q-T bond and that Q is a monotonically increasing function of T (Haberldin and Radtke, 2014).

In flood frequency modeling the problems related to the following points have to point out: Choice of model type, choice of distribution to be used and choice of method of parameter and quantile estimation. It should be noted that two separate features are important. These are the descriptive and predictive properties of the method. The descriptive property relates to the requirements that the chosen distribution shape resembles the observed sample distribution of floods and that random samples drawn from the chosen model distribution must be statistically similar to the properties of real flood series, the predictive properties relates to the requirement that quantile estimates are robust with small bias and standard error (Murphy *et al*., 2014).

At any river site it is usually assumed that nature provides a unique Q-T relationship and that Q is a monotonically increasing function of T (Desalegn *et al*., 2016; Das and Simonovic, 2012) and the following two models was available for this purpose.

2.3.1. Annual Maximum Series Model

In the annual maximum flow (AMF) series, only the peak flow in each year of record is considered. (Ketsela *et al*., 2017) has stated that a series of annual maximum flood is assumed to form a random sample from stationary population in which Q is a random variable with parameter regression distribution. The variate values with exceedence probability 1/T is said to have return period T.

2.3.2. Partial Duration Series Model

In this model, most of the flow hydrograph is disregarded and the hydrograph is viewed as a series of randomly spaced flood peaks of random magnitude (Vivekanandan, 2015). For case of statistical modeling and also for case of identification of the values, which form the series, only the series of peak exceeding an arbitrary threshold are considered. In partial duration series, all peaks above a certain base value are considered. The base is usually selected low enough to include at least one event each year.

Thus, to avoid the problem of data dependency, the annual maximum flow series model was selected. In addition to this, AMF series is widely and universally used model by different researchers for the purpose of flood frequency analysis (Desalegn *et al*., 2016). As a result, to keep the concern of requirement on data, AMF series model was chosen.

2.4. Regionalization

Regionalization, in the context of an evaluation of predictive accuracy of regional flood frequency estimation, refers to identification of homogeneous regions through homogeneity test and selection of appropriate frequency estimation for the identified region and stations. A more specific definition of a homogeneous region is that the region consists of sites having the same standardized frequency distributional form and parameters (Chebana and Ouarda, 2008; Gottschalk and Krasovskaia, 2001; Kachroo *et al*., 2000). In regional flood frequency analysis (RFFA), the established curve of flood variate versus return period can be used for estimating flood quantiles at any site within the region.

Regionalization is done to reduce prediction error by identifying areas with similar characteristics and creating the ability to prediction methods based on catchment similarity (Ding and Haberlandt, 2017), used regionalization techniques to predict design peak flows and found that further effort needs to be made to determine the best predictor variables for parameter regionalization.

The use of regional in formation to estimate flood magnitudes at sites with little or no observed data has become increasingly important since many projects which require design flood information are located in areas where observed flood data are either missing or inadequate. Regional analysis consists of analyzing the record of all gauged sites in a hydrologically homogeneous region, in order to be able to use or transfer information contained in the record of many sites to estimate quintiles at any individual gauged or ungauged catchments in the region (Willems, 2003).

2.4.1. Identification and Delineation of Homogeneous Regions

The identification of homogenous regions is an elementary step in RFFA. The application typically involves the allocation of an ungauged catchment to an appropriate homogenous group and the prediction of flood quantiles using developed models based on catchment characteristics. That is, the RFFA based on homogenous regions can transfer the information from similar gauged catchments to ungauged catchments to allow for flood prediction (Haddad, 2013).

Due to the complexity in understanding the factors that have direct and indirect effect on the generation of flood, there are no simple guidelines for identifying homogeneous regions (Kachroo *et al*., 2000). Meanwhile, experience, prior information and personal judgments can provide possible guidelines to delineate regions with similar hydrological features.

There were several attempts made by different authors to identify hydrologically homogeneous regions and their emphasis were either on geographical considerations or on hydrological characteristics or a combination of both (Kachroo *et al*., 2000).

2.4.2. Statistical Homogeneity Tests

Regional flood estimation methods are based on the premises that standardized flood variate has the same distribution at every site in the chosen region. The importance of homogeneity has been demonstrated by (Demissie, 2008). Homogeneity implies that region have similar flood generating mechanism A more specific definition of a homogeneous region is that region which consists of sites having the same standardized frequency distribution form and parameter.

Homogeneity tests based on Cv and LCv are applied to verify if the preliminary identified and delineated region is homogeneous. In this case, the hydrological data have to be used and the region is confirmed to be homogeneous if it satisfies both criteria of homogeneity tests (Nobert *et al*., 2014).

The discordance measure is intended to identify those sites that are grossly discordant within the group as a whole. It estimates how far a given site is from the center of the group. It is also helpful to screen out the data from unusual sites to look for the appropriate datasets for regionalization. These sites were due to the presence of inaccuracies in data or some other local conditions (Rao and Hamed, 2000; Noto and Loggia, 2009; Guru and Jha, 2016; Kanti *et al*., 2017). The delineation of the homogeneous region is important for site characteristics to be truly representative of the observed discharge data used to estimate hydrologic design values (Irwin *et al*., 2014).

2.5. Statistical Distributions

The best choice of distribution is the one which is robust or capable of giving good quantile estimates even though future values may come from a distribution somewhat different from the fitted distribution, when several distributions fit the data adequately (Hailegeorgis and Alfredsen, 2017).The choice of frequency distributions is determined based on goodness-of-fit measures, which indicate how much the considered distributions fit the available data. A shape parameter describes the weight of the distribution tail of the random variable.

Selection of distribution for AM series has expected broad extend concentration (Haberldin and Radtke, 2014).The selection of distribution is inclined by many factors, such as methods of difference between distributions, methods of estimation parameters, the availability of data.

2.5.1. Best fit Probability Distributions

Probability distribution fitting or simply distribution fitting is the fitting of a probability distribution to a series of data concerning the repeated measurement of a variable phenomenon. The aim of distribution fitting is to predict the probability or to forecast the frequency of occurrence of the magnitude of the phenomenon in a certain interval. Probability distributions can be fitted more closely to the observed frequency of the data than others, depending on the characteristics of the phenomenon and of the distribution. The distribution giving a close fit is supposed to lead to good predictions (Athulya and James, 2017).

In FFA accurate estimation of maximum flood are obtained by fitting probability distribution for a specified return period. The objective is to predict the frequency of occurrence of the magnitude of phenomenon in a certain interval. This can lead to a good prediction of flood. The probability distributions most closely fitted to the observed data depends on the nature of the occurrence and the distribution (Vivekanandan, 2015; Athulya and James, 2017). In flood frequency analysis, an assumed probability distribution is fitted to the available data to estimate the flood magnitude for a specified return period.

A group of goodness-of-fit tests have been conducted such as Kolmogorov-Smirnov test, Anderson-Darling test along with the chi-square test to assess the reasonability and check the adequacy of best-fitting probability distributions to the recorded data. These are statistical tests, which provide a probabilistic framework to evaluate the adequacy of a distribution. The selection of a distribution for flood frequency analysis goes with the selection of the method of parameter estimation (Das and Simonovic, 2012).

2.5.2. Goodness of Fit Tests

The goodness of fit test is a statistical model computes the well-matched of a random sample with a theoretical probability distribution function. In other words, these tests show how fine the selected distribution fits to the data. Similar to in a linear regression, in fundamental nature, the goodness of fit test compares the observed values to the expected (fitted or predicted) values. The performances of the distributions was evaluated using Kolmogorov Smirnov, Anderson-Darling and Chi-Squared goodness-of-tests (Rao and Hamed, 2000).

2.5.3. Method of L-Moment Ratio Diagram

L-moments ration diagram developed by Hosking (1990) is a graphical plot between L-skewness and L-kurtosis by comparing visually sample L-moment ratios to theoretical values. LMRD can be used as a guide tool in selecting an appropriate distribution (Vogel and Wilson, 1996; Peel *et. al*., 2001). The distribution with theoretical value visually close to sample values can be considered as the most suitable PDF that can represent the sample data well. This evaluation test is used as a supportive visual evaluation to ensure that the selected overall best distribution fits the observed data well.

The L-moment ratio diagram is usually used as the first visual inspection tool for selecting a regional frequency distribution from sample data of a region. The L-moment ratio diagram has the ability to provide an elementary visual judgment of a regional frequency distribution by plotting the sample L-moment ratios and average sample L-moment ratios (τ_3 and τ_4) or record length weighted average L-moment ratios as a scatterplot with theoretical curves of several candidate distributions in a L-skewness-L-kurtosis space. The selected distribution should give the closest approximation to the regional data (Lu, 2016; Hosking and Wallis, 1997).

2.5.4. Parameter Estimation

In regional flood frequency analysis (RFFA), the established relationship between the flow variation and the return period can be used for estimating flow quantiles at any site within the region. The parameters are estimated from the sample data; the estimates are subject to sampling errors. A method of fitting must be chosen to minimize these errors. A method suitable to estimate the parameters of one distribution might not necessarily be as efficient for another distribution (Ahmad *et al*., 2011; Badreldin and Fengo, 2012). Several methods can be used for parameter estimation. In this study, the method of moments (MOM), the maximum likelihood method (MLM) and the probability weighted moment method (PWM) are used for parameter estimation.

A. Method of Moments

It is one of the most commonly used methods of estimating parameters of a probability distribution. The estimates of the parameters of a probability distribution function are obtained by equating the moments of the sample with the moments of the probability distribution function. Method of moments (MOM) is a relatively easy parameter estimation method. Unfortunately, MOM estimates are usually inferior in quality and generally not as efficient as the MLM estimates especially in the case where the distributions have a large number of parameters. This is due to the fact that higher order moments are more likely to be highly biased in relatively small samples. The most popularized method to frequency analysis in recent time is that L-moment approach introduced by Hosking and Wallis (1997).

B. Maximum Likelihood Method

The maximum likelihood method (MLM) is considered to be the most accurate method. This is because it provides the smallest sampling variance of the estimated parameters which leads to the smallest sampling variance of the estimated quantiles compared to other methods. Estimation by the Maximum Likelihood Method (MLM) involves the choice of parameter estimates that produce a maximum probability of occurrence of the observations (Vivekanandan, 2015).

In general, the PWM and MOM are better for estimating the parameters for three and two parameter distributions respectively of the underlying distribution from which the data are sampled. They are less sensitive than others are to sampling variability (outliers), and therefore, they yield more accurate and robust estimates of the characteristics or parameters of the underlying probability distribution (Rao and Hamed, 2000).

C. Probability Weighted Moments

The method of probability weighted moments (PWM) is widely used in practice and research (Yurtal, 2010). This method makes use of the analytical relationships among the parameters and the so-called Probability-Weighted Moments of a probability distribution in calculating magnitudes for the parameters. Parameter estimates are obtained in PWM method by equating moments of the distributions with the corresponding sample moments of observed data. Probability-weighted moments (PWM) are useful in the deriving expression for the parameters of distributions can be explicitly defined.

Parameter estimation by PWM, which is relatively new is as easy to apply as ordinary moments (MOM) is usually unbiased and is almost as efficient as MLM. Indeed, in small samples PWM may be as efficient as MLM; with a suitable choice of distribution PWM estimation also contributes to robustness and is attractive from that point of view. Another attraction of the PWM method is that it can be easily used in regional estimation schemes (Rao and Hamed, 2000).

D. L-Moment Method

L-Moments are analogous to method of moments but are estimated by linear combinations of an ordered data set, namely L-statistics (Rao and Hamed, 2000).It is a powerful and efficient method to compute statistical parameters, because such methods can give an unbiased estimate of sample parameters, and cannot easily influence with the presence of outliers.

Compared to the method of moments and maximum likelihood, L-moments can characterize a wide range of distributions. Sample estimates of L-moments are so forceful, may not be affected by the presence of an outlier in the dataset and less subjected to bias in estimation. L-moments can yield accurate estimates of the parameters of a fitted distribution. Even some times parameter estimated form samples are more accurate than maximum likelihood (Cunnane, 1989).

2.5.5. Quantile Estimation

Quantile estimation is the main focus of hydrologic frequency analysis and estimated by applying a distribution function. The selected quantile of under or over design criterion concerning with hydraulic structures is exposed to risk as the return period is determined according to cost and economic strategic significance of the structure. Selecting a reliable design quantile, are necessary for the delineation of floodplains, the development of floodplain management and flood warning systems, which effects on design, operation, and management of a hydraulic structure, considerably depends on statistical methods used in parameter estimation belonging to the probability distribution (Amalina *et al*., 2016).

After the parameters of a distribution are estimated, quantile estimates ($\mathrm{X_T}$) which correspond to different return periods T may be computed. The return period is related to the probability of non-exceedence (F) by the relation, F=1- $\frac{1}{T}$ (where) (F=F (X_T) is the probability of having a flood of magnitude $\mathrm{X_T}$ or smaller. The problem then reduces to evaluating $\mathrm{X_T}$ for a given value of F (Kumar and chatterjee, 2005).

2.6. Derivation of Flood Frequency Curves

A flood frequency curve plots the peak annual flow of a particular stream at a specific location against how often that flow is exceeded. Flood frequency curves provide the annual probability of exceeding a specific flood flow. Regional flood frequency curves have the ability to considering the spatial pattern of variation of hydrologic phenomena across many gauging sites; can be used for estimating flood quantiles at any ungauged site within the region (Ergish, 2010).

Flood frequency curves (FFC) describe the relationship between the magnitude of river peak flows and the recurrence interval or return period. Flood frequency curve, the estimation of flood for various return periods is needed when analyzing flood risk. Developing FFC for different return period helps to estimate flood quantiles (Das and Simonovic, 2012).

In every RFFA, the main goal of the analysis is to develop a regional curve that can represent the averagely weighted distribution of the homogeneous regions. It is the final process of flood frequency analysis to estimate the normalized regional quantile floods (X_T); FFC for a give return period (Tadesse *et al*., 2011).

For a given region, the model parameters derived from the best-fitted distribution to the observed data are used. This helps to compute standardized quantile estimates and then used to construct a regional flood frequency curve for the homogeneous region. These curves are plots of quantiles representing for all sites of a homogeneous region (Hailegeorgis and Alfredsen, 2017).

2.7. Previous Studies on RFFA in Ethiopian River Basins

Investigation of regional flood frequency analysis based on monthly rainfall pattern and geographical proximity was conducted by Gebeyehu (1989) for the Blue Nile River Basin. The study had some limitation about the way that it does not delineate homogeneous regions accurately because the responses of the statistical approach in similar rainfall regions are different consequences of changes in basin topography. In his conclusion, Gebeyehu (1989) point out the following information. The regionalization approach provides useful information about the flood frequency of gauged and ungauged catchments, a small amount of site data greatly improves the estimate of the mean annual flood that can be used with a regionally based estimate of X_T relationship and the results of regional flood frequency analysis should always be updated as more relevant information becomes available.

Blue Nile River Basin has also been regionalized into similar flood producing characteristics based on statistics of at site data (Sine and Ayalew, 2004). The author defined a homogeneous region found have to be with geographical proximity and it performs mainly for carrying out regional frequency analysis for estimation of flood magnitude for water resources project planning and design. Identification and delineation of homogeneous regions for all stations of the respective regions satisfy homogeneity criteria. The types of distribution most likely to fit data of each region were identified from the regional average statistical value of L-Moment ratio. The study recommended that selection of best-fit single distribution and dynamic parameter estimation method require further investigation.

Demissie and Michael (2008), Mekoya and Seleshi (2010) established regional flood frequency analysis for Upper Awash sub-basin using the application of index flood method. The former regionalizes the sub basin into two as upper and lower regions and the later delineated the sub basin into five homogeneous regions and log Pearson type-III as best fit distribution for quantile estimations. The former recommended that additional testing of stations for homogeneity should be done considering geographical factors are a good method in regional flood frequency analysis of the basin and the later to extend the method of regional flood frequency analysis for the other Ethiopian river basins.

Gedefa and Seleshi (2009) investigated Upper Omo-Gibe sub-basin using index flood estimation based on the observed annual maximum flow. L-moment based statistical homogeneity tests were used to identify homogeneous regions. The study concluded that regionalization provides valuable information even in possibly heterogeneous regions, and regional analysis is more accurate and flexible than single-site analysis.

According to Hussein and Wagesho (2016), regionalization of Abaya-Chamo sub-basin was performed based on site characteristics such as elevation, soil type, soil texture, slope, land use land cover and mean annual rainfall. Site statistics were used for testing of homogeneity of the proposed region. The authors concluded that to get reliable quantile estimate more gauging stations should be installed in the basin to infer something for ungagged sites.

Ketsela *et al*. (2017) performed regional flood frequency analysis on Awash River Basin using statistical distribution technique. The Easy Fit Software was employed for selection of best-fit distributions and estimation of parameters for stations. Kolmogorov–Smirnov test was used for the choice of a suitable distribution for estimation of maximum flood discharge. According to this study, Awash basin was delineated into five satisfactory homogeneous regions and recommended software-based techniques like Easy Fit and Hyfran plus software to get accurate and reliable flood estimation results.

2.8. Parameter Estimation Model

Data fitting plays an important role in many natural sciences, engineering and other disciplines. The key idea is to estimate unknown parameters in a mathematical model that describes a real life situation, by minimizing the distance of some known experimental data from the theoretical model function values. Easy Fit is a data analyzer and simulation software which allows to fit probabilistic distributions to given data samples, simulate them, choose the best fitting sample, and implement the results of analysis to take better decisions (Pakgohar, 2014).

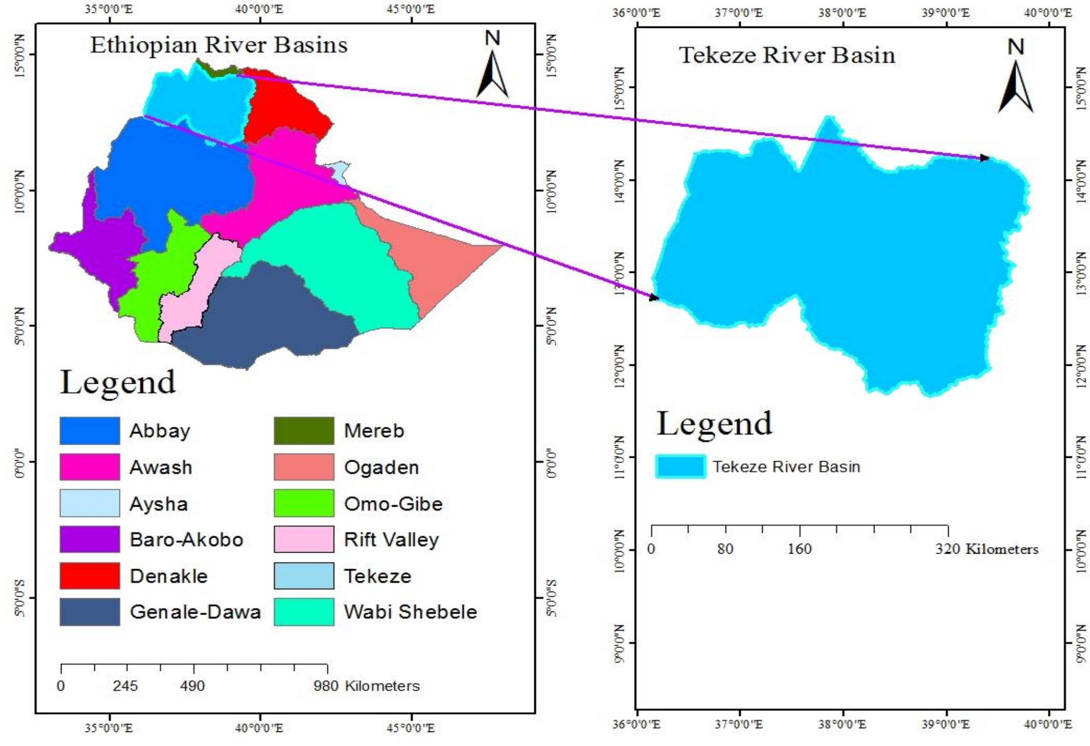
In this study selection of best-fit probability distribution and its method of parameter estimation suitable for each distribution within the interface was conducted using Easy Fit software due to the results of analysis leads to taking a better decision (Romani and Yusop, 2017).

According to Irwin *et al*. (2014), watersheds are delineated using ArcGIS with DEM data and subsequently, several flood generation characteristics are assigned to each watershed. The outcome of this procedure can be directly applied in regionalization to group watersheds into hydrologically homogeneous regions based on the similarity of their attributes, and hydrologic variables are estimated from the regions. Hence, to delineate and characterize watersheds for regionalization ArcGIS10.4.1 environment was used for this study using the procedure of Abdulla (2011) and Irwin *et al*. (2014).

3. MATERIALS AND METHODS

3.1. Description of the Study Area

Tekeze River basin is situated in the North West part of Ethiopia between 11°40' to 15°12' N, and 36°30' to 39°50'E begins at the springs near Lalibela in the central Ethiopian Highlands near Mount Qachen within Lasta, Wollo and shared with Ethiopia and Eritrea after entering northeastern Sudan joins the Atbarah River a tributary of the Nile. The basin has a total drainage area of 83,530km²; of which 79,513 km² (95.19%) in Ethiopia covering parts of the Amhara and Tigray regional states and relatively small part of the basin 4,017 km² (4.81%) is situated in Eritrea. The River commences from the highlands of Wollo and Gonder in the South and drains central, southern and small portion of the western Tigray and Northern Gonder westward to the Nile. The river basin has a lowest elevation of 536 m in low lands of Metema area and a highest elevation of 4517 m at Semen Mountains (Fentaw *et al*., 2017; MoWR, 1998).



(a). Ethiopian River Basins (b). Tekeze River Basin

Figure 3.1: Location map of the study area

3.1.1. Climate and Hydrology

The climate of the basin can be divided into two: the west region of the Semen Mountains with wet season and the east region with dry (small rainy) and wet (main rainy) seasons. The mean temperatures in the basin vary from 10^0C in the Semen Mountains, to 22^0C in the highlands and to 26^0C in the lowlands. Minimum and maximum temperature ranges are $3\text{-}21^0$C and $19\text{-}43^0$C respectively. Rainfall decreases from south to north from 1,200 to 600mm.The mean annual rainfall is 600mm in the lowlands and 1,300mm in the Semen Mountains (Fentaw *et al*., 2017).

3.1.2. Land Use Land Cover

The land use and land cover of the basin includes 27% of cultivated land, 35.1% of shrub land, 0.3% of wooded grassland, and 32.5% of bushy/open wood land; shrub by grass land and sparsely vegetated shrub land/exposed rock/soil. Most of the climax vegetation of the basin has disappeared and only little of the original vegetation is evident while only little of the lowland woodlands and bush lands in the western and northern parts of the basin are nearer to climax. However, the Afro-alpine and sub afro-alpine heath vegetation lies above 3700 to 3900m.a.s.l around Semen Mountains (Fentaw *et al*., 2017).

3.2. Tools Used

For the proper execution of this study, materials and software used was based on the capability to work on achieving the predetermined objectives. Microsoft excel spread sheet and XLSTAT 2018 was used for data arrangement, filling missed data, calculate the statistical parameters of hydrological data used in the flood frequency analysis and frequency curves, return period and quantiles are also plotted. Arc-GIS 10.4.1 software was used to generate the study area map representing geographical location of gauging stations and delineate hydrologically homogeneous regions. Easy Fit 5.6 statistical software (trial version) was used to select the best fit probability distribution with its method of parameter estimations, a goodness of fit tests and to check the estimation accuracy of each of data of stations. Matlab2018a to execute discordancy of sites from the identified regions.

3.3. Data Collection and Analysis

Defining a clear and efficient methodology is vital for the quality of the findings of the study. The procedures of data analysis in this study includes from the preliminary screening of data to develop a regional flood frequency curve depending on AMF series data. Screening the data was carried out to check for gross errors and make sure the continuity of data. After relevant data which were useful for the regional analysis identified from the study basin, checking of data for its quality was performed.

Identifying homogeneous region was done to decide on which sub basins can be grouped together which might have similar flood producing nature. This was performed based on the L-moment ratio diagram and site characteristics of stations. The regional frequency distribution by the average L-moment ratios and a goodness of fit test with help of Easy Fit software was then used to confirm how well the selected distribution fit the data in the region. Estimation of the frequency distribution is then designed to compute the flood quantiles for certain return periods at ungauged sites derived from the regional growth curve. In general, to achieve the evaluation of predictive accuracy of regional flood frequency estimations of this study, the following procedures were employed. Homogenous region is identified, standardized data from different sites within the region can be pooled together and a single frequency curve applicable to the region can be derived. In cases where adequate rainfall or river flow records are not available at or near the site of interest, it is difficult for hydrologists and engineers to derive reliable flood estimates directly and regional studies can be useful.

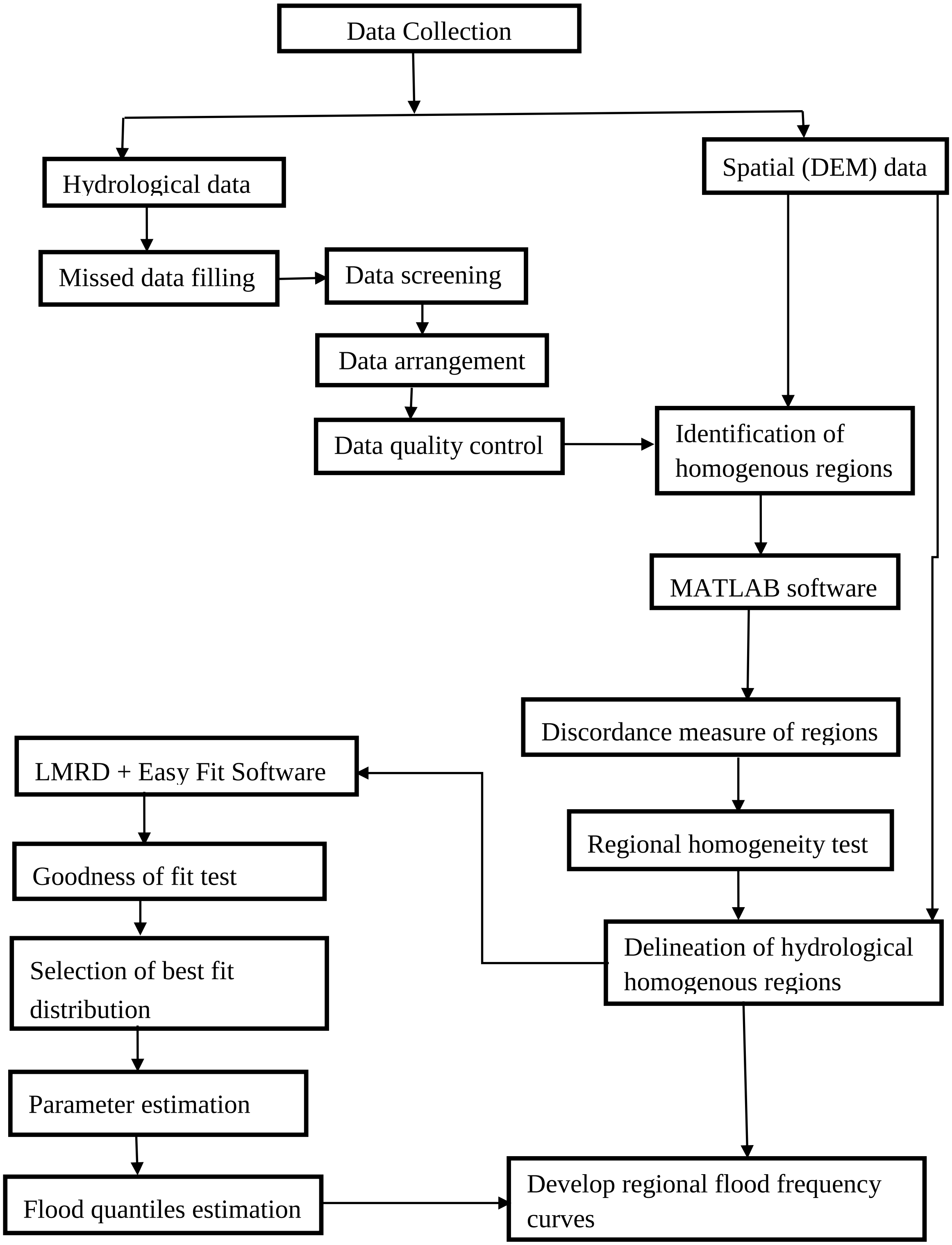


Figure 3.2: Flow chart of the methodology

3.3.1. Sources and Availability of Data

Flood frequency estimation primarily uses observed annual maximum flood data at gauging stations to estimate flood magnitude. Hydrological and DEM (digital elevation model) data of Tekeze River Basin were collected from Ministry of Water Irrigation and Electricity, department of hydrology and GIS. DEM data was employed as basic input for delineation and specifying the location of the gauging stations in the basin. In the study area, there are about twenty gauging stations, out of these only eleven gauging stations were selected for the proper regional flood frequency estimation. The selected stations by themselves have no fully recorded data; they have a number of years of record having missing data that needs to be filled before analysis. Accordingly, eleven gauging sites which satisfied the minimum record length were selected.

The site characteristics of stations for this study includes the code of the stations, the name of the river and their gauging sites, the locations (latitude and longitude) and catchment area (km^2).

Table 3.1: The site characteristics of stations used in detail analysis

Station code	River name	Location of gauging station	Coordinate		Area (km2)	Record period	Record length
			Latitude	Longitude			
121004	Gheba	nr.Mekele	13°6'N	39°38'E	5000	2001-2015	15
121006	Tekeze	nr.Embamadre	13°74'N	38°2'E	18425	1994-2015	22
121007	Sulluh	nr.Hawzen	13°85'N	39°51'E	2951	1991-2013	23
121008	Gheba	nr.Adikumsi	13°46'N	39°02'E	6893	1998-2017	20
121010	Genfel	at Wukro	13°8'N	39°6'E	3032	1992-2015	24
121012	Metere	nr.Aynalem	13°46'N	39°49'E	2621	1991-2006	16
121014	Maydungur	nr.Adwa	14°18'N	38°88'E	6124	1991-2007	17
121023	Tekeze	nr.Kulmesk	12°6'N	39°19'E	7909	1996-2015	20
122002	Tekeze	at Humera	13°84'N	36°88'E	22624	1990-2004	15
122003	Buya	nr.Maitsemry	13°59'N	38°15'E	2951	2000-2014	15
123049	Mekezo	nr.Dansha	13°57'N	36°97'E	5000	2000-2014	15

3.3.2. Data Screening

Data screening is the first task in which employed methods that the unwanted observation from the data series as well as the sites from the analysis can be filtered. It is used to check the data are appropriate for performing the regional flood frequency (Kumar and Chatterjee, 2011; Kachroo *et al.*, 2000). This allows visual detection of whether the observations have been consistently or accidentally credited to the wrong day, or whether they contain misplaced decimal points. Visual observation of the daily flow records implied on errors such as overstated numbers, misplaced decimal points, and very high flow records during dry months and/or very low flow records during rainy months. In this study, stream flow data was used from gauging stations in the Tekeze River Basin.

The minimum and maximum lengths of the at-site AMF records respectively are 15 and 24 years. For all the stations listed in Table 3.1 and shown in Figure 3.3, the AMF data were selected and later subjected for investigative data analysis in order to choose representative stations for the study area.

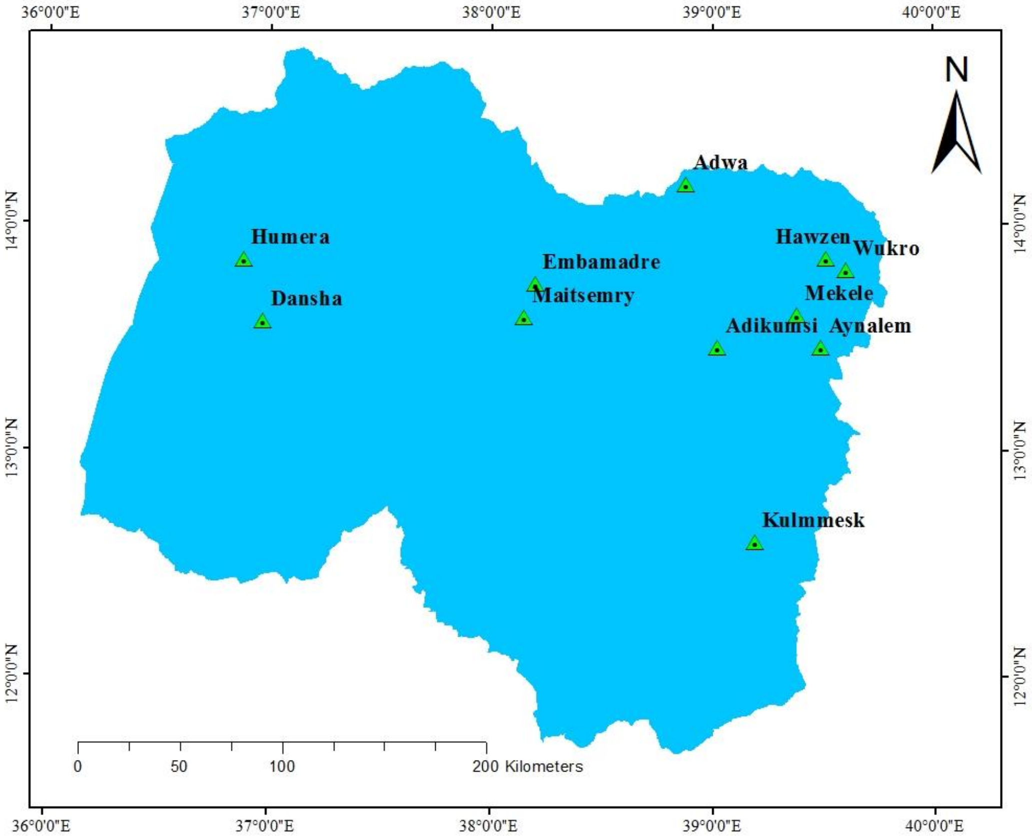


Figure 3.3: The spatial distribution of gauging stations in Tekeze River Basin

3.3.3. Missed Data Filling

When undertaking an analysis of stream flow data from gauges where observations are made, it is often to find times where no observations are recorded at one or more gauges. The continuity of the record may be broken with missing data due to many reasons such as the absence of recorder, carelessness of the observer, break or failure of instruments (Sine, 2004). Different methods used for filling of missing flow data for a given gauging station.

For this study, any missing data was filled by the method of linear regression. Simple linear regression was applied to fill missing stream flow values using nearby flow gauging station observations. The equation for linear regression is given as:

$$y = ax + b \quad \text{……………………………………………………………………3.1}$$

Where x and y are instantaneous daily stream flows ($m^3/_{sec}$) and, a and b-constants.

This method is selected due to the following reasons: i). It is the most widely used method when compared to other method for large data. ii). Estimation of significant missing observations as accurate as possible. iii). It is applied by creating a correlation with the nearby station.

3.4. Data Quality Control

Some errors may exist in the stream flow observation that were collected, such as misplaced decimal numbers, very huge unrealistic numbers and negative flow records in some cases. The following approaches were considered to check stream flow data quality.

3.4.1. Test for Randomness and Independence

It is usually assumed that all peak magnitude in annual maximum series is usually mutually independent in the statistical sense. It is known that FFA is carried out when the at-site data are independent and identically distributed conditions satisfied. This provides that the extreme events might appear randomly and all might have the same frequency distribution.

The requirement of RFFA is that the AMF at different stations in a homogeneous region should be spatially independent. Hailegeorgis and Alfredsen (2017) noted that independence of data series is one of the main assumptions in frequency analysis and the intersite correlation has a considerable effect on the variance of regional parameters and flood quantiles and reduces the effective length of records.

It is assumed that all the peak magnitudes in the AM series are mutually independent in the statistical sense. In this study, the correlation coefficient was applied to verify the independence of the data of the selected hydrological stations. According to Dahmen and Hall (1990), the lag-1 serial correction coefficient, R_1, defined as follows:

$$R_1 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(x_i + 1 - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.2$$

Where Xi is an observation,

Xi+1 is the following observation and

n is the number of data.

After computing R_1, the test hypothesis is that H_0: R_1= zero (that there is no correlation between two consecutive observations) against the alternative hypothesis, H_1: R_1<> 0.

Anderson (1942) defines the critical region, R_1 at the 5% level of significance as: (−1, (LCL) R_1 (UCL), 1) and equation 3.2 gives:

The upper confidence limit, UCL, for R_1 as:

$$\mathrm{UCL}(R_1) = \frac{(-1 + 1.96(N-2)^{0.5}}{N-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.3$$

The lower confidence limits, LCL, for R_1 as:

$$\mathrm{LCL}(R_1) = \frac{(-1 - 1.96(N-2)^{0.5}}{N-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.4$$

To accept the hypothesis H_0: R_1=0, the value of R_1 should fall between the UCL and LCL. Applying this condition to the time series, the condition: LCL (R_1) <R_1< UCL (R_1) is satisfied for the all stations.

Table 3.2: Result of test for independence of stations time series data

Station name	LCL (R_1)	R_1	UCL (R_1)	Station name	LCL (R_1)	R_1	UCL (R_1)
Embamadre	-0.465	0.214	0.370	Mekele	-0.576	0.330	0.433
Humera	-0.576	0.354	0.433	Adwa	-0.537	0.319	0.412
Kulmesk	-0.490	-0.253	0.385	Maitsemry	-0.576	0.127	0.433
Hawzen	-0.454	-0.057	0.363	Dansha	-0.576	0.390	0.433
Wukro	-0.443	0.102	0.356	Aynalem	-0.556	-0.219	0.422
Adikumsi	-0.490	-0.050	0.385				

Thus, no correlation exists between successive observations. The data are independent and there is no persistence in the time series. The summarized result of the test for annual maximum flow series for example for Embamadre station -0.465<0.214<0.370 and the other stations are given in Table 3.2 and the results show that the annual maximum flow series for all stations were independent**.**

3.4.2. Test for Consistency and Stationary

A time series is stationary if in the long term it is invariant with respect to time. The two tests were adopted to check stream flow observations stationary and consistency. According to Rao and Hamed (2000), F-test for the stability of variance and t-test for the stability of mean verify not the stationary of time series, but also its absolute consistency and homogeneity. According to this, if F-test shows stable variance and t-test shows stable mean, then we can say that the time series is stationary, consistent and homogenous. Thus, the two tests were adopted to check stream flow observations stationarity and consistency.

a. F-test for Stability of Variance

The test statistic is the ratio of the variances of two split, non-overlapping, sub-sets of the series (Dahmen and Hall, 1990). The annual maximum stream flow observations during are divided into equal or nearly equal time series. Then, the variance of both time series is calculated for all gauging stations.

The test statistic (Ft) is calculated as:

$$\text{Ft} = \frac{\text{variance of series 1}}{\text{variance of series 2}} \quad \ldots\ldots 3.5$$

According to this method, the variance of the time series is stable if and only if: F (V_1, V_2, 2.5%) <Ft < F (V_1, V_2, 97.5%), where $V_1 = n_1 - 1$, $V_2 = n_2 - 2$, and $n_1 = n_2$ -the number of observation point in each subset.

b. Test for the Stability of Mean

The test for stability of the mean involves computing and then comparing the mean of non-overlapping subsets of the time series (Dahmen and Hall, 1990). The same subsets from the F-test are used for calculations of the t-test values.

The statistic t-test (Tt) is given as:

$$Tt = \frac{\bar{X}series1 - \bar{X}series2}{\left[(n_1-1)S_1^2 + (n_2-1)S_2^2 * \frac{1}{n_1+n_2-2} * \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\right]^{0.5}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.6$$

Where $\bar{X}$: is the mean of the series

n: is the number of monthly stream flow records

S: is the standard deviation of the two series

According to this test, the mean of the time series is stable if and only if: t (V,2.5%) < Tt< t (V,97.5%), Where the value of V is different for each station and values are read from Appendix-D using percentile columns (2.5% and 97.5%).

Noting that both F {V_1, V_2, 2.5%} and F {V_1, V_2, 97.5%} values for 5% significance level as Appendix-B. For the station having year are listed using V_1, V_2 and percentile row 2.5 % or 97.5 % Appendix-C. The results of observations of data of gauging stations T-test and F-test, are presented in Appendix-E and shows that mean and variance of the time series was stable.

3.4.3. Check for Data Adequacy and Reliability

The accuracy of statistical the mean is a function of the sample size. The data taken for analysis were checked for its adequacy and reliability. Accuracy and adequacy of data were checked and defined in (McCuen, 1998) using the equation 3.7.

$$De = \frac{C_v}{N^{0.5}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.7$$

Where, De- Standard error

Cv-Coefficient of variation and

N-number of yearly data in the series

The data series could be regarded as reliable and adequate if De is less than 10% significance level. Hence, the data of stations are found accurate, adequate and reliable as De value for most of the stations are less than 10% significant level (McCuen, 1998).

Table 3.3: Results of test for adequacy and reliability of AMF data

Site location	Cv	N	De	Site location	Cv	N	De
Embamadre	0.2176	22	0.0571	Mekele	0.2542	15	0.0656
Humera	0.3831	15	0.0989	Adwa	0.3912	17	0.0949
Kulmesk	0.4351	20	0.0973	Maitsemry	0.1567	15	0.0405
Hawzen	0.3451	23	0.0720	Dansha	0.3751	15	0.0969
Wukro	0.1978	24	0.0404	Aynalem	0.3540	16	0.0885
Adikumsi	0.4089	20	0.0914				

3.4.4. Check for Outliers of the Data Series

Outliers are data points that depart significantly from the trend of the remaining data. Outliers may come due to personal error during recording and inadequacy of measuring device or really due to very extreme condition of natural phenomenon that is important information for flood frequency analysis. The presence of outliers in the data causes difficulties when fitting a distribution to the data. Low and high outliers are both possible and have different effects on the analysis. For statistical tests of out-lying observation, it is generally recommended that a low significant level such as 1% be used and that significance level greater than 5% should not be common practice (Dahmen and Hall, 1990; Ketsela *et al*., 2017). To minimize or avoid the effect of outliers in this study L-moment an efficient parameter estimation technique was employed.

3.5. Regionalization of Tekeze River Basin

In this study, the index flood L-moment approach of regionalization was applied depending on the data homogeneity of the stations. The statistical values have been checked for the stations whether they can be classified under one or more regions. Flood statistics of Tekeze River Basin stations were computed using L-moment methods. Due to the fact that such methods can give a balanced estimation of sample parameters and cannot be easily influenced by the presence of outliers (Rao and Hamed, 2000).

3.5.1. Identification of Homogeneous Regions

Identification of homogeneous regions is the significant step in regional frequency analysis. To identification of homogeneous regions, the specification of variables characterizing this similarity has been made. The identification of homogeneous regions is usually the most difficult stage and requires the greatest amount of personal judgment (Amalina *et al*., 2016).

The regionalization process is statistically verified by using discordant measures and homogeneity tests to assess the degree of variability within the pool. This is proven by comparing the measure of scale and dispersion value of both the L-moment (LCv) and conventional moment (Cv) of gauging stations that belong to different regions. The statistical nature of annual maximum flow variation within the group of stations is best explained by LCv and Cv. The stream gauging stations were grouped into geographically continuous sites such that the response of streams to physiographic variables should be similar. DEM size of 30mx30m the basin was used to identify site characteristics. This enables stream flow records to be transferred from gauged basins to ungauged basins within a region (Sine *et al*., 2013).

3.5.1.1. Site Characteristics

In this study, preliminary IHRs of stations into a certain category is achieved by looking at stations site characteristics. The following site characteristics were used as a preliminary IHR; latitude and longitude, AMF, station area and altitude of the flow gauging station. Then stations having nearly same kind of site characteristics are clustered on the same region.

3.5.1.2. Method of L-Moment Ratio Diagram

L-moments ration diagram developed by Hosking (1990) is a graphical plot between L-skewness and L-kurtosis by comparing visually sample L-moment ratios to theoretical values. LMRD can be used as a guide tool in selecting an appropriate distribution (Vogel and Wilson, 1996; Peel *et al*., 2001). The distribution with theoretical value visually close to sample values can be considered as the most suitable PDF that can represent the sample data well. This evaluation test is used as a supportive visual evaluation to ensure that the selected overall best distribution fits the observed data well.

3.5.2. Test for Homogeneity of Stations and Regions

The preliminary identified regions have to be checked by various homogeneity tests. The main advantage of L-moments is that being a linear combination of data, they are less influenced by outliers, and the bias of their small sample estimates remains fairly small. Unbiased sample estimators of the first four PWMs are given as (Hosking and Wallis, 1997) and suggested a homogeneity test based on L-moments which proved to be efficient. The tests used in this study were discordance measure tests, measure of scale, dispersion based tests (Cv-based homogeneity test and LCv-based homogeneity test), and statistical comparison.

3.5.2.1. Discordancy Measure of Regions

The discordance measure is used to identify those sites that are grossly discordant with the group as a whole. The discordance measure Di estimates how far a given site is from the center of the group based on statistical properties (Rao and Hamed, 2000). A suitable criterion to classify a station as discordant is when Di is greater than or equal to 3 (Hosking and Wallis, 1993). If a vector, $\mathrm{Ui} = \left(\tau^{\mathrm{i}}{}_{2,}\ \tau^{\mathrm{i}}{}_{3,}\ \tau^{\mathrm{i}}{}_{4,}\right)^{\mathrm{T}}$, which controlled the L-moment ratios for site i, T is the transpose of the vector Ui (Hosking and Wallis, 1997), then the discordancy measure may be defined as:

$$\mathrm{Di} = \frac{1}{3}(\mathrm{Ui} - \overline{\mathrm{U}}\mathrm{i})\mathrm{S}^{-1}(\mathrm{Ui} - \overline{\mathrm{U}}\mathrm{i})^{\mathrm{T}} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.8$$

$$\overline{\mathrm{U}}\mathrm{i} = \frac{1}{\mathrm{N}} * \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{Ui} \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.9$$

$$S = \frac{1}{(N-1)} * \sum_{i=1}^{N} (Ui - \overline{U}i)(Ui - \overline{U}i)^{T} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.10$$

Where N -is the total number of sites

Di -discordancy measure

Ui -is defined as a vector containing the L-moment ratios for site i,

$\overline{U}i$ -is the group averages Ui,

S -sample covariance matrix of Ui.

Hosking and Wallis (1997) tabulated critical values of the discordancy statistic Di for various numbers of sites in a region at a significance level of 10%. These were used to assess each of the study sites and identify whether they should be analyzed further to ensure homogeneity. The identified regions have tested for discordancy using equation 3.8. However, to determine the value of Di using simple matrix multiplication was difficult and quite cumbersome.

Due to this, Hosking and Wallis (1997) recommended using Fortran, Matlab and other computer programs to simplify the work and get acceptable accuracy results. For this study, following this recommendation Matlab2018a programming code was employed to simplify the numerical calculations of discordancy index (Di). The programming code used to calculate the covariance matrix and Di were given on Appendix-F.

Table 3.4 : Critical values of discordancy measure with N sites

Number of sites in a region	Critical value	Number of sites in a region	Critical value
5	1.333	6	1.648
7	1.917	8	2.14
9	2.329	10	2.491
11	2.632	12	2.757
13	2.869	14	2.971
>15	3		

(Source: Hosking and Wallis, 1997)

3.5.2.2. Adjustment of Regions

If the regions formed are not statistically homogeneous, they are adjusted to improve their homogeneity. This step is justified because regions are not generally likely to be homogeneous based on the homogeneity assessment and discordant sites may also exist.

Rao and Srinivas (2008) point out the following options for revising regions that are grossly discordant with respect to other sites within the region. i). eliminating one or more sites from the data set; ii) transferring (or moving) one or more sites from a region to other regions; iii) dividing a region to form two or more new regions; iv) allowing a site to be shared by two or more regions; v) dissolving regions by transferring their sites to other regions; vi) merging a region with another or others; vii) merging two or more regions and redefining groups; and viii) obtaining more data and redefining regions. Among these, the first three options are useful in reducing the values of heterogeneity measures of a region, whereas the options (iv) to (vii) help in ensuring that each region is sufficiently large.

3.5.2.3. Conventional Homogeneity Test

Stations in a region can be tested for homogeneity that is fall in a region; homogeneity can be taken as a base for many criteria of the basin. The criterion used to check for regional homogeneity was based on the value of CC. According to some researchers, the higher the value of Cv and CC, the lower will be the performance of the index-flood method for the region under consideration. This is due to the dominance of the flood quantile estimation variance by the variance of the at-site sample mean. Hence, for better performance of the index flood method, CC should be kept low. In this study both conventional and L-moment has been used to calculate CV, L-Cv and their respective CC, values. The procedures are described below.

For each site in the delineated regions; the mean $\bar{Q}$, standard deviation (δ) and coefficient of variation (Cv) were given and calculated by Guru and Jha (2016) and Sine and Ayalew (2004) equation (3.11-3.16).

The mean of AMF of the station:

$$\overline{\mathrm{Q}}\mathrm{i} = \frac{1}{\mathrm{n}} * \sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{Qi} \quad \ldots\ldots\ldots 3.11$$

The standard deviation of AMF of the station;

$$\delta i = \sqrt{\frac{\sum_{i=1}^{n}(\mathrm{Qi}-\bar{\mathrm{Q}}\mathrm{i})^2}{\mathrm{n}}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.12$$

$$\mathrm{Cv_i} = \frac{\delta \mathrm{i}}{\bar{\mathrm{Q}}_\mathrm{i}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.13$$

Where: Qi= the flow rate of the station in the region (m^3/s), at site i

$\bar{\mathrm{Q}}$i=the mean flow rate for the region (m^3/s), at site i

δi= standard deviation for the region, at site i

n = number of a record year

$\mathrm{Cv_i}$= Coefficient of variation of a region, at site i

For each region, using the statistic calculated Cv above, the regional mean, Cvi and finally the corresponding CC value using the following relation:

Regional mean; $\overline{\mathrm{Cv}}_\mathrm{i} = \frac{1}{\mathrm{N}} * \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{Cv_i}$ ………………………..…..……….……....3.14

Regional standard deviation, $\delta\mathrm{cv} = \sqrt{\frac{\sum_{i=1}^{N}(\mathrm{Cv_i}-\overline{\mathrm{Cv}}_\mathrm{i})^2}{\mathrm{N}}}$ …………………………....3.15

$$\mathrm{CC} = \frac{\delta_{\mathrm{cv}}}{\overline{\mathrm{Cv}}_\mathrm{i}} < 0.3 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.16$$

Where: N=number of the site in a region

$\overline{\mathrm{Cv}}$i = the mean coefficient of at site Cvi values

δ_{cv} = standard deviation of at site Cvi values

3.5.2.4. L-moment Based Homogeneity Test

L-Moment method is a powerful and efficient method to compute statistical distribution and parameters, because such methods can give unbiased estimate of sample parameters and also cannot be easily influenced with the presence of outliers. Effective way of testing the homogeneity of the site based LCV homogeneity test is more accurate when compared with that of the Cv-based homogeneity test. The procedural calculation is the same as that of the Cv. The following are advantage of LCv (Cunnane, 1989; Karchroo *et al*., 2000): Compared to Cv, LCv can characterize a wide range of distribution, sample estimates are so strong that they are not affected by the presence of outliers in the data set, they are less matter to bias in estimation, yields more accurate estimate of the parameter of a fitted distribution.

According to Rao and Hamed (2000), L-moments have the following advantages: i). characterize most of probability distributions than conventional moments, ii). less sensitive to outliers in the data, iii). Approximate their asymptotic normal distribution more closely, iv). Nearly unbiased for all combinations of sample sizes and populations.

Hosking and Wallis (1993) gave the unbiased estimators of β_0, $\beta_{1,}$ β_2 and β_3, defined as;

$$\beta_0 = \frac{1}{n}\sum_{i=1}^{n} Q_i \quad \text{……………………………………………………………….........3.17}$$

$$\beta_1 = \sum_{i=1}^{n-1} \frac{(j-1)(Q_i)}{n(n-1)} \quad \text{……………………….....……………………………...…...3.18}$$

$$\beta_2 = \sum_{i=1}^{n-2} \frac{(j-1)(j-2)(Q_i)}{n(n-1)(n-2)} \quad \text{……………………………...…...……....………...3.19}$$

$$\beta_3 = \sum_{i=1}^{n-3} \frac{(j-1)(j-2)(j-3)(Q_i)}{n(n-1)(n-2)(n-3)} \quad \text{……………………………........…...…………3.20}$$

Where Qi- annual maximum flow (m^3/s) from stations dataset

n - the number of years, j-rank

β_0, $\beta_{1,}$ β_2 and β_3 -are L-moments estimator.

The first few moments are:

$\lambda_1 = \beta_0$; $\lambda_2 = 2\beta_1 - \beta_0$; $\lambda_3 = 6\beta_2 - 6\beta_1 + \beta_0$; $\lambda_4 = 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0$ ……………........3.21

In specific, λ_1 is the mean of the distribution or measure of location; λ_2 is a measure of scale; τ_3 is a measure of skewness, and τ_4 is a measure of kurtosis. L-skewness and L-kurtosis are both defined relative to the L-scale, λ_2; and sample estimates of L-moment ratios can be written as L-Cv, L-Cs, and L-Ck.

L-moment ratios are independent of units of measurement and are given by Hosking and Wallis (1997) as follows:

$$\tau_2 = \frac{\lambda_2}{\lambda_1}, \tau_3 = \frac{\lambda_3}{\lambda_2}, \tau_4 = \frac{\lambda_4}{\lambda_2}, \ldots\ldots\ldots 3.22$$

Using the above procedural formula,

$$\overline{\mathrm{Lcv}_i} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{Lcv}_i \ldots\ldots\ldots 3.23$$

$$\delta cv = \sqrt{\frac{\sum_{i=1}^{n}(\mathrm{Lcv}_i - \overline{\mathrm{Lcv}_i})^2}{n-1}} \ldots\ldots\ldots 3.24$$

The weighted regional LCvi, of all the sites, CC is defined as follows:

$$CC = \frac{\delta_{Lcv}}{\overline{Cvi}} < 0.3 \ldots\ldots\ldots 3.25$$

A region that confidently satisfies all criteria for being hydrologically homogeneous can be derived.

3.5.3. Delineation of Homogeneous Regions

The performance of any regional estimation method highly depends on the grouping of sites into homogeneous regions (Karchroo *et al.*, 2000). In this study, the geographical proximity and LMRD were used in order to cluster preliminary regions which then tested for hydrologic similarity. The delineation of homogeneous regions is closely related to the identification of the common regional distributions that apply within each region. A region can only be considered homogeneous if sufficient evidence can be established that at different sites in the region are drawn from the same parent distribution.

In this study, the Digital Elevation Model (DEM) size of 30m×30m Tekeze River Basin was used and the delineation of homogeneous regions was performed by taking in to account the drainage boundaries of the subbasin with ArcGIS 10.4.1 environment. The preliminarily identified regions have to be checked by various homogeneity tests. All sample stations are located on a digitized map by latitude and longitude. For each station, the statistical values (LCs, LCk) were computed. It was assumed that the LCs and LCk values of one station vary linearly with the neighboring stations.

According to Abdulla (2011) and Irwin et al. (2014), the procedures in the delineation of the boundary of the region are as follows: i). Compute the (LCs, LCk) value of each station, ii). Identify the location of stations along the distributions of LMRD for the defined regions statistical comparison of observed flood data, iii). Identify the group based on step (ii), iv). Each region that was identified in step-i was checked for statistical homogeneity using the proposed test.

Finally, the drainage boundaries of each sub-region the delineation was carried out using ArcGIS10.4.1 environment.

3.6. Selection of Regional Frequency Distribution

Selection of regional frequency distribution is one of the important elements of the regional flood frequency analysis (RFFA). Presence of adequate hydrometric stations is essential in each of the hydrologic regions for reliable selection of regional frequency distributions. The choice of frequency distributions is determined based on goodness-of-fit measures, which indicates how much the considered distributions fit the available data (Hailegeorgis and Alfredsen, 2017; Mishra *et al.*, 2009). In flood frequency analysis, the annual maximum flow corresponding to a given T can be estimated from the annual flood series using varies theoretical distributions.

3.6.1. L-Moment Ratio Diagram

L-moment ratio diagram is a graphical tool which can be used as goodness of fit measure for selection of best fit distribution. It is a graph of the L-skewness and L-kurtosis which compares the fit of several distributions on the same graph. LMRD can be used as a guide tool in selecting an appropriate distribution (Hosking and Wallis., 1997; Vogel and Wilson, 1996; Peel *et al.*, 2001). The distribution with theoretical value visually close to sample values can be considered as the most suitable PDF that can represent the sample data well.

Therefore, some acceptable design procedures are essentially required to choose a model that minimizes uncertainties. Generalized extreme value (GEV), Generalized logistic (GLO), Generalized Pareto (GPA), Logistic, Log-Normal (LN), Log-Pearson type 3 (LPIII) and Normal distributions are among the employed distributions in this study.

The GEV distribution is a family of continuous probability distributions. GEV makes use of 3 parameters: location, scale and shape. The location parameter describes the shift of a distribution in a given direction on the horizontal axis. The scale parameter describes how spread out the distribution is, and defines where the bulk of the distribution lies. As the scale parameter increases, the distribution will become more spread out. The 3rd parameter in the GEV family is the shape parameter, which strictly affects the shape of the distribution, and governs the tail of each distribution. The shape parameter is derived from skewness, as it represents where the majority of the data lies, which creates the tail(s) of the distribution. The GEV is probably the most widely used distribution when measuring AM series of river flow.

The generalized logistic (GLO) distribution is evaluated for flood frequency analysis. The performance of the GLO distribution is compared with those of the generalized extreme value (GEV), three parameters log-normal (LN3) and three parameter Pearson (P3) distributions. The results are reported in terms of empirical distribution function (EDF) tests of goodness of fit, on both individual and regional flood series through the application of these distributions to a set of reasonably long annual maximum series.

The LPIII distribution is a member of the family of Pearson Type 3 distributions, and is also referred to as the Gamma distribution. The LPIII distribution is complicated, as it has two interacting shape parameters. Similar to GEV it uses three parameters, shape (k), scale (σ) and location (μ).

Hence, these distributions were considered for the evaluation of the possible distributions that can represent the average frequency distribution of the regional data of the basin.

3.6.2. Easy Fit Software for Distribution Fitting

Easy Fit is a data analyzer and simulation software which allows to fit probabilistic distributions to given data samples, simulate them, choose the best fitting sample, and implement the results of analysis to take better decisions. In order to determine whether the distribution model could fit the data properly, goodness-of-fit tests were used. In this study Easy Fit 5.6 Statistical Software Package, trial version 5.6 was used to find the best-fit distribution and its estimation parameters (Pakgohar, 2014).

3.6.3. Goodness of Fit Tests

The goodness fit measure involves identifying a distribution that fits the observed data. When computing the magnitudes of extreme events, such as flood flows, it is required that the probability distribution function be invertible, so that a given value of recurrence interval (T) and the corresponding value of frequency factor (K) can be determined.

In this study, to test the statistical hypothesis whether a particular distribution provides an adequate fit to the observed AMF series data three goodness of fit tests were applied. The reason for selecting three different tests is that there is no single test that can give conclusive results and a particular test emphasizes a particular aspect of the goodness-of-fit. All test statistics were defined and carried out at 5% significance level (Ashraful *et al.*, 2018).

i. Kolmogorov-Smirnov Test (KS)

The test statistic in the Kolmogorov-Smirnov test is extremely simple. The KS test was used to check whether the sample came from hypothesized continuous distribution. It was based on the empirical distribution function (Di Baldassarre *et al*, 2009). In this method, the hypotheses take dependability of a specified distributions data of stations. Kolmogorov-Smirnov (KS) test is a different and commonly used goodness-of-fit moreover Chi-squares test. A statistic based on the deviations of the sample distribution function FN (X) is use in this test. The test statistic DN is defined in equation 3.26.

The test statistic DN is defined as:

$$\mathrm{DN} = \underset{1 \le i \le n}{\max} \left| F_n(x_i) - F_{O(x_i)} \right| \quad \text{...3.26}$$

The values of FN (x) are predictable as Nj/N where Nj is the cumulative number of sample events in class j. Fo(x) is then 1/K, 2/k…...etc., Similar to the chi-square test. The value of DN must be less than a tabulated value of DN at the specified confidence level for the distribution to be received.

ii. Chi-Squared Test (X^2)

The chi- square goodness of fit test is one of the most commonly used tests for testing the goodness of fit of probability distribution functions to empirical frequency distribution.

In X^2 goodness of fit test, sample data is separated into intervals. Then the numbers of points that drop into the interval are compared, with the predictable numbers of points in every interval. The null hypothesis assumes that there is no notable variation between the observed and the expected value. The degree of freedom depends on the distribution of the data sample (Ghosh *et al.*, 2016).

In Chi-Square goodness of fit test, the alternative hypothesis assumes that there is an essential variation between the observed and the expected value.

$$X^2 = \frac{(O-E)^2}{E} \quad \text{..3.27}$$

Where X^2 =chi-Square goodness of fit test

O = observed value

E = expected value

The considered value of Chi-Square goodness of fit test is compared with the table value. If the considered value of Chi-Square goodness of fit test is less than the table value, will admit the null hypothesis and conclude that there is no important differentiation between the observed and expected value.

iii. Anderson-Darling Test (AD)

The AD test was used to check whether the given sample came from a particular probability distribution at hand. The null hypothesis at chosen level of significance would be rejected if calculated value of above statistic exceeds the critical value given in the table (Onoz and Bayazit, 2001; Ahmad *et al*, 2015).

AD test can be used in RFFA studies to assess the fitness of the candidate regional frequency distributions. This method is based on statistical frequency distribution behavior of the observed value (Viglione *et al*, 2007).

3.6.4. Evaluation the Performance of Frequency Distributions

The results obtained from statistical analysis can be uncertain, and to be trustful methods of uncertainty assessments should be applied (Hosking and Wallis, 1997). Assessment of the accuracy of the estimates should, therefore, take into account the possibility of heterogeneity in the region, misspecification of the frequency distribution and statistical dependence between observations at different sites, to an existent that is consistent with the data. Analytical goodness-to-fit criteria are helpful as an approval for whether a particular elimination of the data from the model is statistically significant or not.

The distribution that has the most number of points nearby to the line signifies the best-fitted distribution model. This implies that the frequency distributions that were chosen as the best distribution could be fitting regional flood models for the basin. Hence, for this analysis, two methods of uncertainty assessments were achieved. Thus are probability-probability (P-P) and quantile-quantile (Q-Q) plots. The performance of the best distribution model identified for the respective regions was evaluated by comparing observed with simulated values by employing the P-P and Q-Q plot techniques with Easy Fit software.

i. Probability-Probability (P-P) Plots

Probability plots are generally used to decide whether the distribution of a variable matches a given distribution. P-P Plots is the variable's cumulative magnitude in opposition to the cumulative magnitude of any of a number of trial distributions.

If the selected variable matches the test distribution, the points come together approximately a straight line. The following fundamental issues should arise when selecting a distribution: (1). It is true and reliable with the distribution for which the observations are drawn, (2). It should be used to obtain reasonably perfect and strong estimations of design quantiles and hydrologic risk (Desalegn *et al*., 2016).

ii. Quantile-Quantile (Q-Q) Plots

The Quantile- Quantile plot is a graphical technique for determining if two data sets come from populations with a common distribution. Quantile-quantile(Q-Q) plots are plots of two quantiles against each other. A quantile is a small part where certain values fall below that quantile. The purpose of Q-Q plots is to get out if two sets of data come from the same distribution. It is the graph of the input observed and analysis data values plotted against their theoretical or fitted distribution. These are produced by plotting the data values against the x-axis, and the following values against the y-axis. Q-Q plots were used to compare the estimated quantiles and the observed flood values and to check the validity of the estimates provided by a fitted theoretical distribution. The best frequency distribution was subjected to randomly simulate the same size as observed series (Desalegn *et al*., 2016).

3.6.5. Parameter and Quantile Estimation

In flood frequency analysis, the probability distribution is fitted to the available data to estimate the flood magnitude for a specified return period. The choice of an appropriate probability distribution is quite arbitrary, as no physical basis is available to rationalize the use of any particular distribution (Saf, 2009; Rao and Hamed, 2000).

In the present study, the parameter estimation was done by using the Easy Fit statistical software. Based on the selected distributions for each station, the quantile can be calculated according to the formula of the selected distributions. For regions with a computed value of scale, location and shape parameter, then it is possible to determine the quantile with different return periods using different equations for different distributions.

For GEV distribution the flow quantile can be estimated as;

$X_T = \mu + \frac{\sigma}{K}\left(1 - (-\ln\left(1 - \frac{1}{T}\right)^{K}\right)$, for k≠0 …………………………….…………....3.28

$X_T = \mu + \sigma\left(\ln(-\ln\left(1 - \frac{1}{T}\right)\right)$, for k=0 ……………………………………......3.29

For GPA distribution the flow quantile can be estimated as;

$X_T = \mu + \frac{\sigma}{K}\left(1 - \left(\frac{1}{T}\right)^{K}\right)$, for k≠0 ……………………...…………………….…….3.30

$X_T = \mu + \sigma\left(\ln\left(\frac{1}{T}\right)\right)$, for k=0 ………………………………...........…………...3.31

Where, σ = scale parameter, T= return period, μ= location parameter and k = shape parameter

In this study, estimation of parameters and calculation of the magnitude of flood for 1000 years return period were executed. Comparing the result of the flood events of 1000 years return period is significant. The reason that dam safety risk analysis, sizing of emergency spillways, the design of dam crest level and any other hydraulic structures, the critical flood peaks are mostly based on the criterion of 1000 years return period flood. This may help to make balanced engineering decisions on the choice of design floods used to ensure a satisfactory and reliable standard in the planning and design of flood control structures (Donnelly *et al*., 2008; Haktanier *et al*., 2010).

3.6.6. Standard Error of Parameter Estimation

The standard error of the estimate is a measure of the accuracy of predictions. The development of the relationship between the mean annual flood or index flood and the catchment characteristics was a necessary step in predicting flood magnitudes at any point in a region where the frequency curve has been derived and error in quantitative terms. Different researchers use different measures of error. The most common measures are standard errors. From the various source of error only sampling error can be evaluated theoretically a consensus seems to be emerging that at least sampling error should normally be reported in quantitative terms. Error in flood frequency estimates should normally be reported either numerically or graphically. The standard error of a given quantile due to sampling error should generally be computed (Rao and Hammed, 2000).

$$\bar{Q} = \frac{\sum \bar{Q}_T}{Q_T} \quad \ldots\ldots 3.32$$

$$SEE = \left[\frac{\sum(\bar{Q} - Q_T)^2}{\bar{Q}_T}\right]^{0.5} \quad \ldots\ldots 3.33$$

Where; SEE – standard error of estimate

$\bar{Q}_T$ - is the estimated value of standard quantile

$\bar{Q}$ - is the mean annual flood (m^3/s) is the index flood

Q_T - is the quantile (m^3/s) function of fitted distribution at site i

3.7. Derivation of the Regional Flood Frequency Curves

In every regional flood frequency analysis, the main goal of the analysis is to develop regional frequency curve that can represent the average weighted distribution of the homogenous regions. It is the final procedure of flood frequency analysis to estimate the normalized regional quantile floods (X_T); flood frequency curve (X_T vs. T); and at-site flood quantiles, for a give return period, T. For a given region, the model parameters derived from the best fitted distribution to the observed data are the most essential one. Because, these values are used to compute standardized quantile estimates, X_T for the return periods T, and then used to construct regional frequency curves for the homogenous region (a curve showing X_T against return period, T) (Kachroo *et al*., 2000; Mkhandi *et al*., 2000; Rosbjerg, 2007; Yang *et al*., 2010).

3.7.1. Estimation of Index-Flood

The fundamental assumption of the index flood method is that data at different sites in a region follow the same distribution consisting of identification of homogeneous regions, determination of best-fit distribution and derivation of the regional flood frequency curve. In this study, the index flood L-moment approach of regionalization is applied depending on the homogeneity of the stations by testing for the homogeneity using different techniques. Derivation of the mean annual flood ($\bar{Q}$) for each station was obtained by relating the annual flood data from each station (Qi) and dividing it by the number of record years.

Flood quantiles estimation in flood frequency analysis were corresponding to the required return periods. The model parameters for the distributions estimated for each station were used to compute standardized flow estimates conforming to the return periods 2, 5, 10, 15, 20, 25, 50, 75, 100, 200, 500 and 1000 years. Plots of Q/Qm against the Gumble reduced variate (-ln(-ln(1-1/T)) known as growth curves, were generated for each station and used in the derivation of the regional growth curves.

To do this, the following stages were employed. Select best fitted distributions the parameter values such as shape (k), location (μ) and scale (σ) which were estimated using Easy Fit software, the model parameters estimated for a given region were used to compute the standardized quintiles estimates for the return periods, the growth curves for each station were developed.

After the regional frequency distribution is determined, the flood quantiles having a return period of T year within a homogeneous region can be estimated based on the equation (3.34) proposed by Hosking and Wallis (1997). The dimensionless regional growth curves used to estimate X_T. The common practice is to get the dimensionless data by dividing the values by an estimate of the at-site mean.

$$X_T = \frac{Q_T}{\bar{Q}} \quad \text{.....} 3.34$$

Where; $\bar{Q}$ - is the mean annual flood (m^3/s) is the index flood

Q_T - is the quantile (m^3/s) function of fitted distribution at site i

X_T - Regional quantile of which can be obtained from regional growth curve; this defines the frequency distribution common to all the sites in a homogenous region.

4. RESULTS AND DISCUSSIONS

4.1. Identification and Delineation of Homogeneous Region

4.1.1. Identification of Homogeneous Region

The identification of homogenous regions is usually the most difficult stage and requires greatest amount of subjective judgment.The aim is to form group of sites that approximately satisfy homogeneity condition that the sites frequency distributions are identical. The homogeneity of the region is evaluated using homogeneity measures which are based on site characteristics and L-moment ratio diagram (LMRD) of flood statistics. This method considers the stations that were geographically continuous (the spatial proximity of network of gauging stations as indicated in Figure 3.3) and in clustering, the annual maximum flow of sites in the region should satisfy the homogeneity test criteria (Hosking and Wallis, 1997; Tallaksen *et al.*, 2004).

The LMRD shown on Figure 4.1 was used to identify homogeneous regions with site characteristics of gauging stations described in Table 3.1. As indicated in Table 4.1, the accentuated distributions were designated to the same group since stations lie close to the identical distribution. Hence, based on L-moment statistics and suitability of gauging site networks, three homogeneous subregions were identified. Namely Region-A, Region-B and Region-C as shown in Table 4.1.

Table 4.1: Preliminary identified homogeneous regions

Group name	Station name	Possible distributions from LMRD	
Region-A	Adikumsi	GEV	LN/LPIII
	Adwa	LN	LPIII
	Aynalem	GEV	LN
	Hawzen	LN/LPIII	GEV
	Kulmesk	GLO	LN
	Mekele	GLO	GEV
	Wukro	LN/LPIII	GLO
Region-B	Embamadre	GEV	GPA/LPIII
	Maitsemry	GPA	GEV/LPIII
Region-C	Dansha	GPA	GEV/LPIII
	Humera	GPA	GEV/LPIII

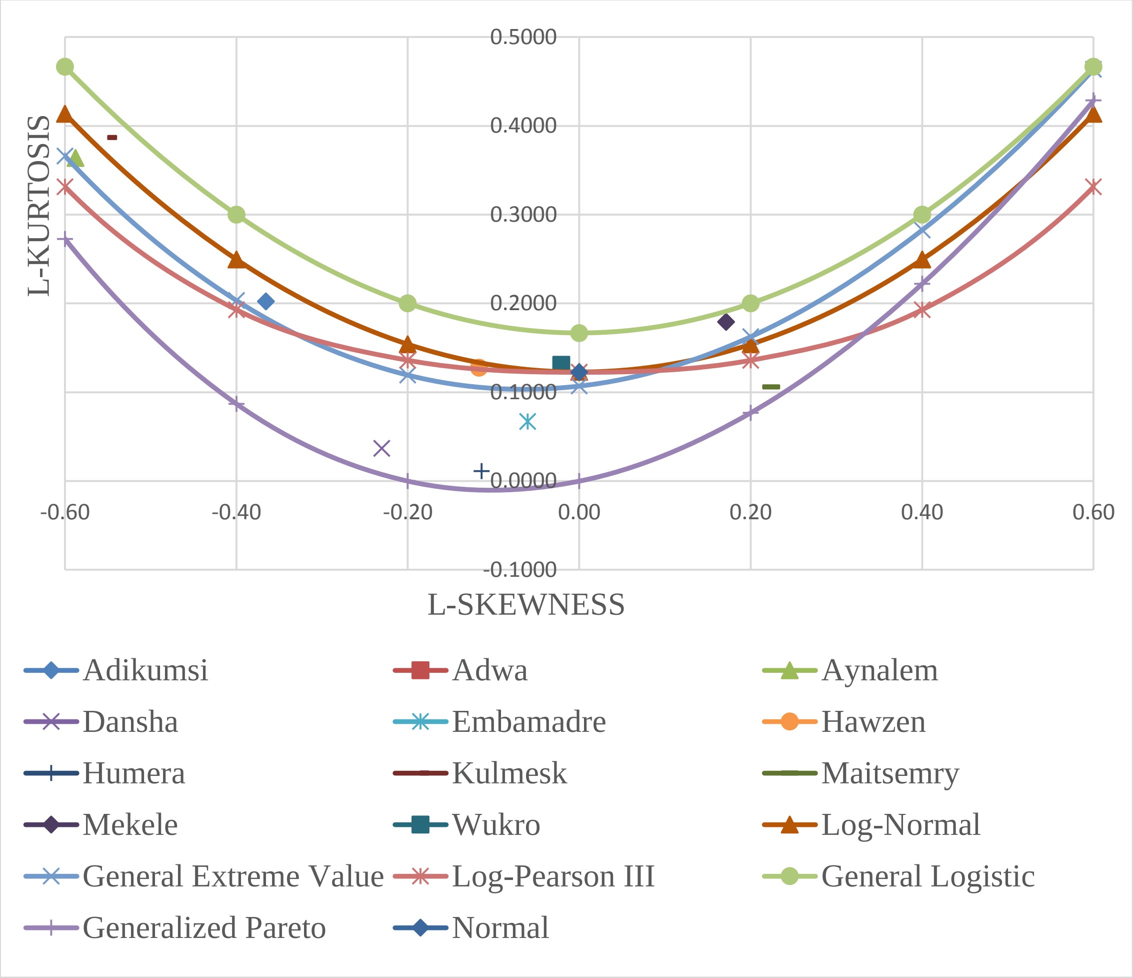


Figure 4.1: L-moment ratio diagram for identification of homogeneous regions

4.1.2. Test for Regional Homogeneity

The identified homogeneous regions from statistical values have to be statistically homogenous to verify the acceptability of regions.

4.1.2.1. Discordancy Measure of Regions

The discordancy measure (Di) proposed by Hosking and Wallis (1993) aims to screen out the unusual sites from other sites in a group by comparing their L-moment ratios. Values of discordancy of L-moment statistics have been calculated for all the eleven gauging sites of the basin. Using Equation (3.8) with Matlab program code presented in Appendix-F, the values of discordance index (Di) measure for different sites within the regions were presented in Table 4.2, 4.3 and 4.4 for Region-A, Region-B and Region-C respectively. The critical values of the discordancy index Di for various numbers of sites in a region at a significance level of 10% were obtained from Table 3.4. It was observed that the Di values for all eleven sites vary from 0.3120 to 1.8212.

According to Hussen and Wagesho (2016), Kanti *et al*. (2017), Lim (2007), Nobert *et al.* (2014), and the region on their study under investigation, has been declared homogeneous if Di is less than 3. In this condition, a site is declared to be unusual if Di is large. This would be considered as grossly discordant and would justify elimination from the defined regions and can be redefined as a single site or merged into other regions. Hence, all of the stations grouped as a homogeneous in Region-A, Region-B, and Region-C were satisfied the discordance test criteria. As shown in Table 4.2, 4.3 and 4.4, the result of all the Di was below the critical value which implies that all the regions are homogeneous. So, none of the identified regions was found to reveal Di greater than the critical value. This indicated that all sites do not reflect any outlier and discordancy.Thus, data of all gauging sites could be considered for further regional flood frequency analysis.

Table 4.2: Results of major statistics and discordant measures in Region-A

Station Name	LCv	LCs	LCk	Di	Remark
Adikumsi	0.9694	-0.3176	0.1421	0.6720	Homogenous
Adwa	1.0867	-0.5787	0.2753	0.6907	Homogenous
Aynalem	1.082	-0.4766	0.1922	0.4186	Homogenous
Hawzen	0.6496	-0.1028	0.114	1.6640	Homogenous
Kulmesk	1.0584	-0.5044	0.2012	0.3120	Homogenous
Mekele	0.5192	0.1045	0.3478	1.8212	Homogenous
Wukro	0.5918	-0.2815	0.3621	1.4216	Homogenous

Table 4.3: Results of major statistics and discordant measures in Region-B

Station Name	LCv	LCs	LCk	Di	Remark
Embamadre	0.2176	0.0075	0.0003	0.9999	Homogenous
Maitsemry	0.1567	0.2676	0.0961	0.9999	Homogenous

Table 4.4: Results of major statistics and discordant measures in Region-C

Station Name	LCv	LCs	LCk	Di	Remark
Dansha	0.6707	-0.2305	0.0366	0.9999	Homogeneous
Humera	0.6971	-0.0750	-0.0125	0.9999	Homogeneous

4.1.2.2. CC- based Regional Homogeneity Test

In this test the site-to-site coefficient of variation of the coefficient of variation (CC) of both conventional and L-moments of the proposed region are used. The (L-Cs, L-Ck) of standardized flow values at each station have been plotted on the LMRD together with various theoretical distribution function. Those stations close to a particular theoretical distribution linear considered to be homogeneous stations and grouped together. The combined coefficients of variation for the region (CC) values were calculated and the results in sites of each region were summarized as shown in Table 4.5, 4.6 and 4.7.

The value of CC varies from region to region depending on L-moment statistics of flow data. From Cv-based homogeneity test, the CC values were 0.2524, 0.2138 and 0.0149 for Region-A, Region-B and Region-C respectively. On the other way, from LCv-based homogeneity test, the CC values were 0.2972, 0.2301 and 0.0273 for Region-A, Region-B and Region-C respectively.

According to Sine and Ayalew (2004), Guru and Jha (2016) and Nobert *et al*. (2014) noted that for the study regions under their consideration, a region is declared to be homogeneous if CC values were less than 0.3.Therefore, from the results in Table 4.5, 4.6 and 4.7, it can be concluded that all regions were hydrologically homogeneous for both Cv and LCv based homogeneity tests since the CC values were less than 0.3. The results obtained below, all stations grouped as homogeneous were satisfied the stated homogeneity test criteria. As a result, it can be concluded that all regions were reasonably homogeneous.

Table 4.5: Results of Cv and LCv-based homogeneity test for Region-A

Station Name	LCv	LCs	LCk	Cv	Cs	Ck
Adikumsi	0.9694	-0.3176	0.1421	0.4089	1.4361	2.2821
Adwa	1.0867	-0.5787	0.2753	0.3942	2.2874	6.0170
Aynalem	1.0820	-0.4766	0.1922	0.3540	1.9630	4.3191
Hawzen	0.6496	-0.1028	0.1140	0.3451	1.3020	3.6676
Kulmesk	1.0584	-0.5044	0.2012	0.4351	2.3013	6.5411
Mekele	0.5192	0.1045	0.3478	0.2542	0.3346	1.7202
Wukro	0.5918	-0.2815	0.3621	0.1978	3.3472	14.2998
Mean	0.8510	-0.3082	0.2335	0.3413	1.8531	5.5496
Std.dev	0.2529	0.2427	0.0972	0.0861	0.9502	4.2459
CC	0.2972			0.2524		

Table 4.6: Results of Cv and LCv-based homogeneity test for Region-B

Station Name	LCv	LCs	LCk	Cv	Cs	Ck
Embamadre	0.2176	0.0075	0.0003	0.4386	-0.0003	-1.0227
Maitsemry	0.1567	0.2676	0.0961	0.3234	-1.416	1.6754
Mean	0.1872	0.1376	0.0482	0.3810	-0.7082	0.3264
Std.dev	0.0431	0.1839	0.0677	0.0815	1.0011	1.9078
CC	0.2301			0.2138		

Table 4.7: Results of Cv and LCv-based homogeneity test for Region-C

Station Name	LCv	LCs	LCk	Cv	Cs	Ck
Dansha	0.6707	-0.2305	0.0366	0.3751	-0.0003	-1.0227
Humera	0.6971	-0.075	-0.0125	0.3831	-1.4160	1.6754
Mean	0.6839	-0.1528	0.0121	0.3791	-0.7082	0.3264
Std.dev	0.0187	0.1100	0.0347	0.0057	1.0011	1.9078
CC	0.0273			0.0149		

4.1.3. Delineation of Homogeneous Regions

The delineation of homogeneous regions is closely related to the identification of the common regional distributions that apply within each region. The preliminary identified regions have to be checked by various homogeneity tests. The tests used in this study are dispersion based tests (Cv based homogeneity test and L-Cv based homogeneity test) and statistical comparison. The regions have covered an area of 34,530, 21,376 and 27,624km^2 for Region-A, Region-B and Region-C respectively.

Accordingly, the first region which includes most of gauging stations in the Gheba, Tserare and upper Tekeze sub-river basins including Adikumsi, Adwa, Aynalem, Hawzen, Kulmesk, Mekele and Wukro stations were delineated under Region-A. The second region, which includes the gauging stations in Sibta, Zarema, Belesa and Middle Tekeze sub-river basins including Embamadre and Maitsemry stations were delineated under Region-B. The third region, which is most of gauging stations in Angereb, Goang and Lower Tekeze sub-river basins including Dansha and Humera stations were delineated under Region-C. Having proven to be statistically homogeneous, the delineated homogenous regions shown in Figure 4.2 could be used to generate a regional growth curve at any site located in the study area.

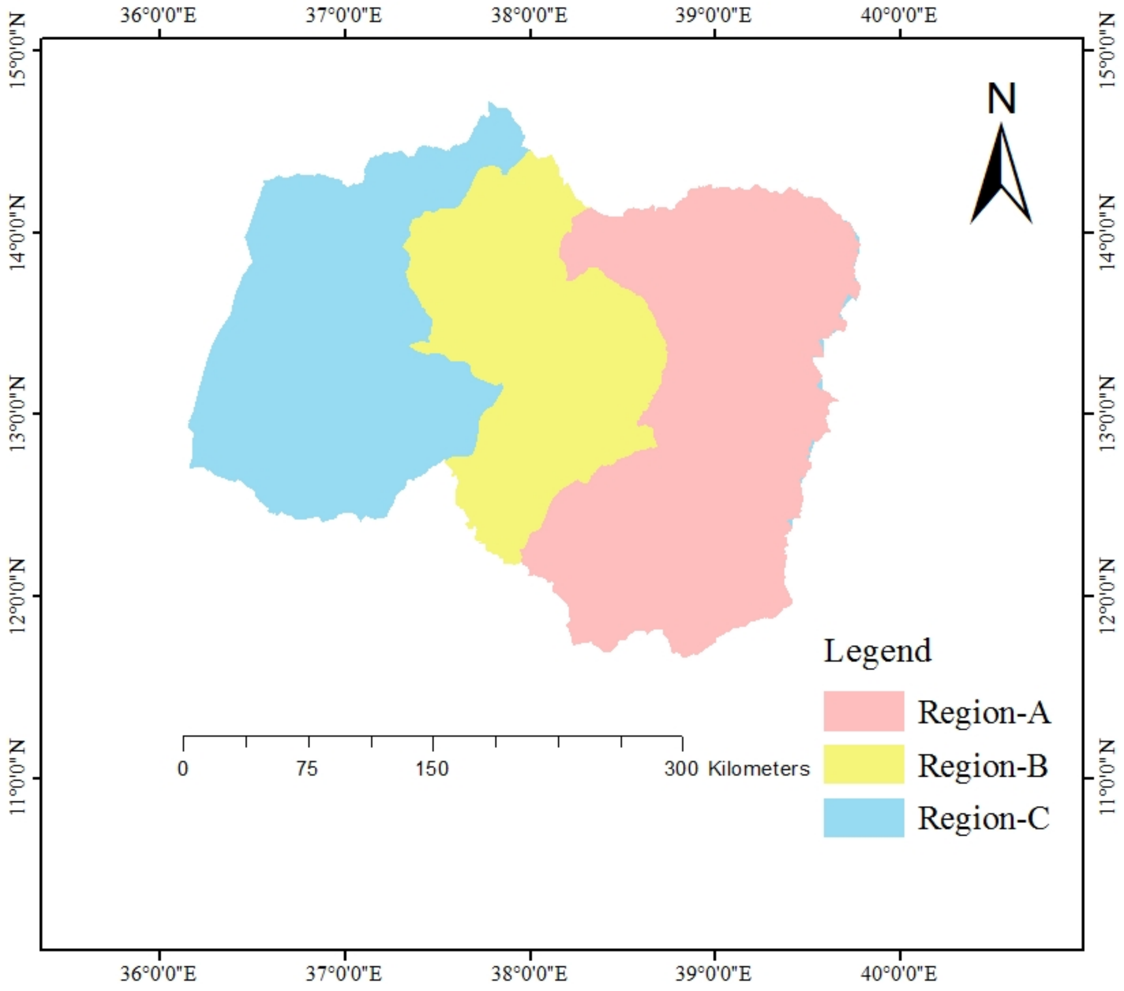


Figure 4.2: Spatial distribution of delineated homogeneous regions

4.2. Determination of Suitable Regional Probability Distribution

In this study, the annual maximum series model was adopted where only the maximum flow in each year is considered.

4.2.1. Goodness of Fit Tests

The purpose of the goodness-of-fit test is to determine the best fitting frequency distribution by computing the difference of the L-kurtosis between the sample data and using Easy Fit statistical software.

In this study, the goodness of fit tests was performed for all distributions using Kolmogorov Smirnov, Anderson-Darling and Chi-Squared methods for the data of gauging stations. They were applied to determine whether the distribution to be fitted to the data or not. The best-fit result of each station was taken as the distribution with the lowest sum of the rank orders from each of the three test statistics. This GOFs at 5% level of significance was used to define the best-fit ranking using Easy Fit statistical software.

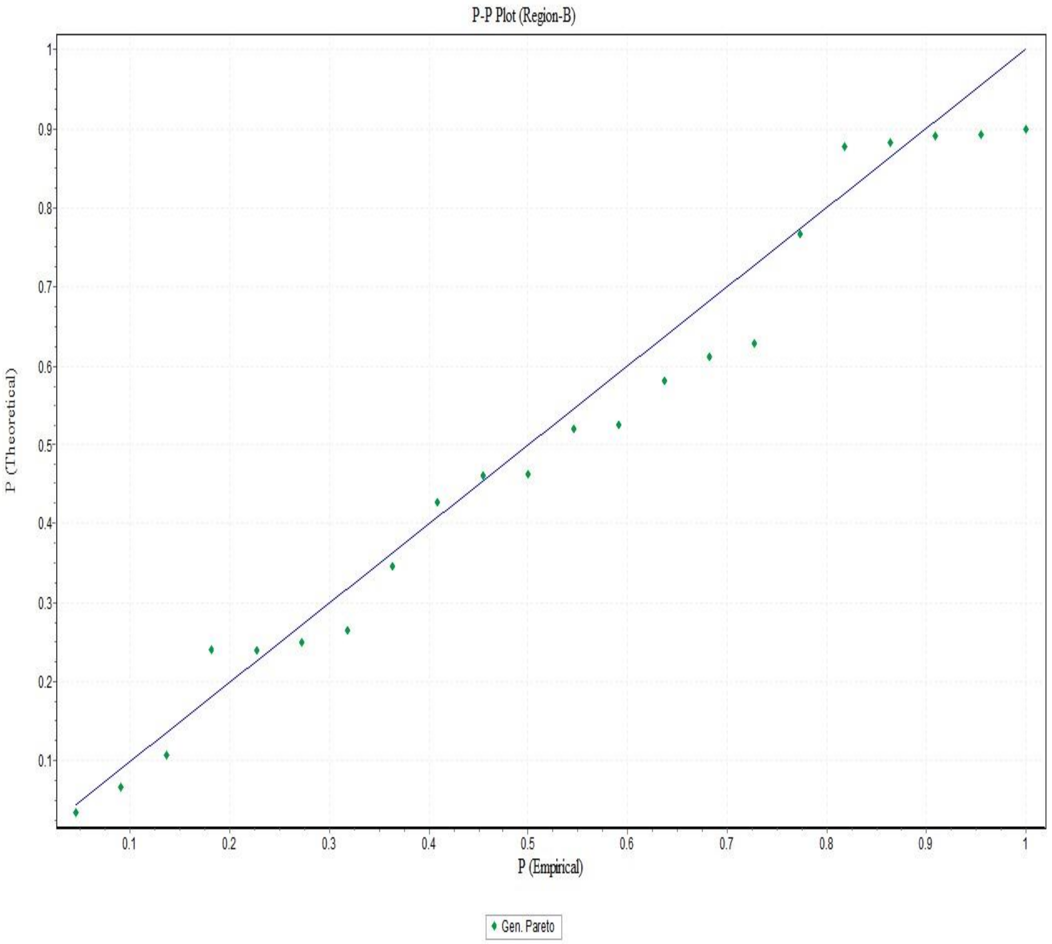
The probability distribution having the first rank along with their test statistic was presented in Table 4.8 and Appendix-H. The justification of results was summarized in Table 4.8 for Region-A and Appendix-H for other Regions were presented depending on the ranking of the goodness of fit tests. Using the three tests from Table 4.8, it was detected that generalized extreme value distribution for Region-A provides the best fit to the AMF data and generalized pareto distribution for Region-B and Region-C. Comparing the results of goodness-of-fit tests, the generalized extreme value distribution affords a good fit for the recorded data of stations.

Table 4.8: Goodness of fit test values for selected distributions of Region-A

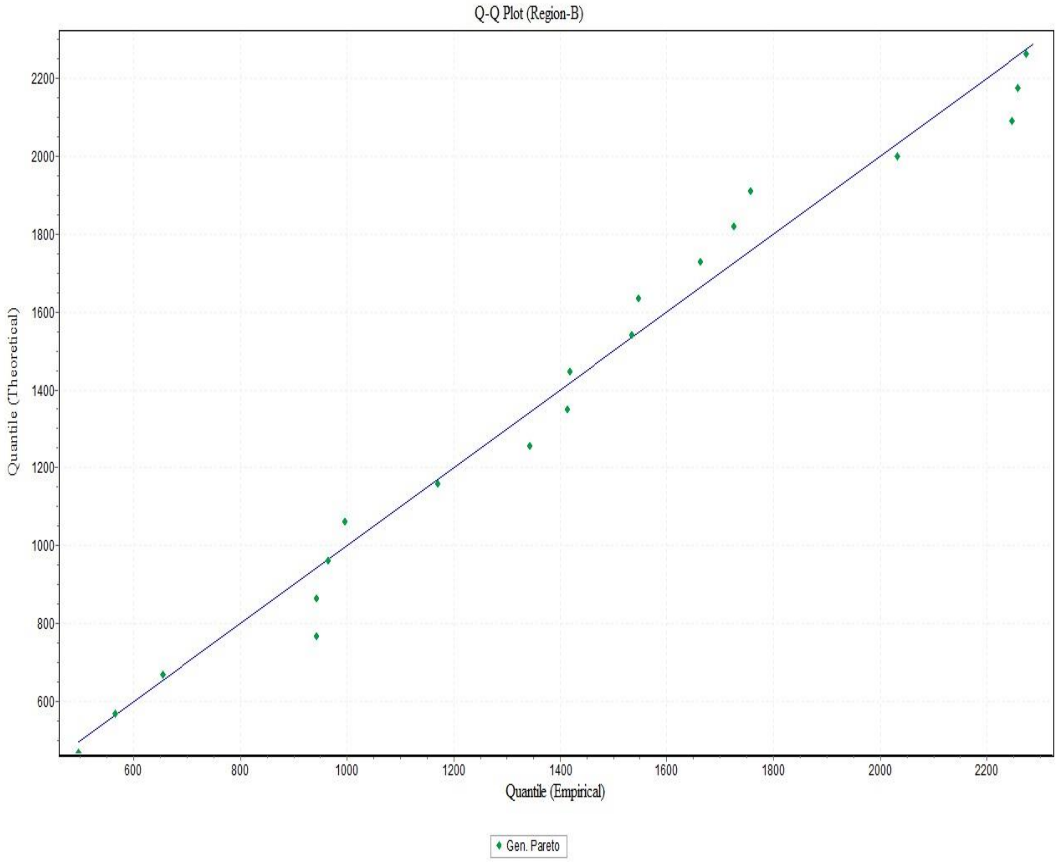
Distribution	Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
	Statistics	Rank	Statistics	Rank	Statistics	Rank
Generalized Extreme Value	0.0937	1	0.3588	2	0.1368	1
Generalized Pareto	0.0966	2	0.3282	1	0.1382	2
Log-Pearson 3	0.1152	3	0.4080	3	1.4312	5
Lognormal (3P)	0.1282	4	0.4778	4	2.6691	6
Logistic	0.1469	5	0.8280	5	1.3896	4
Normal	0.1536	6	0.8636	6	1.2001	3
Lognormal	0.1814	7	1.0900	7	6.1031	7

4.2.2. Evaluating Estimation Accuracy of Selected Distribution

The performance of the chosen distribution as a best fitted regional model was assessed using probability-probability plot and quantile-quantile plot.The probability-probability (P-P) plot is a graph of the empirical values plotted against the theoretical values. It is used to determine how well a specific distribution fits to the observed data. The quantile-quantile (Q-Q) plot is a graph of the input (observed) data values plotted against the theoretical (fitted) distribution quintiles. Both axes of this graph are in units of the input data set. The P-P and Q-Q plot have to be more or less linear if the particular theoretical distribution is the correct model. It was observed that from the results shown in Figure 4.3 for Region-B and Appendix-I and J for the rest of the regions, indicated that almost all plots were well fitted to the line. The study reveals that GEV and GPA distributions performed well for most of the stations in the basin. Therefore, results from both methods validated that the flood frequencies of the regions were well addressed. Hence, using these distributions and annual maximum flow modeling could have a wide range of applications in agriculture, hydrology, engineering design and future climate evaluation in the study area.



(a)



(b)

Figure 4.3: Performance evaluation of frequency distributions

4.2.3. Method of L-Moment Ratio Diagram

The L-moment ratio diagram is usually used as the first visual inspection tool for selecting a regional frequency distribution from sample data of a region. The corresponding average weighted value of L-moment statistics results were obtained from regional data as presented in Table 4.5,4.6 and 4.7 plotted along with the theoretical lines for some distributions on LMRD to determine a regional probability distribution.

As shown in Figure 4.4, the points representing the regional average values of L-kurtosis versus L-skewness were fitted with GPA and GEV distributions. The choice of a suitable standard frequency distribution is often uncertain and LMRD might not guarantee that the distribution is the actual representative of flood statistics in the given region.

The results between the goodness-of-fit test with Easy Fit and LMRD indicated that due to the common acceptance of GEV and GPA distributions, could be used as a best-fit distribution for the study area. Therefore, GEV and GPA distributions could be adopted as the regional distribution, while GLO, LN, LPIII, Normal and Logistic distributions should not be considered. As a result, this justified that the two distributions would be acceptable and the dominate probability distributions in the Tekeze River Basin for estimation of regional flood frequency.

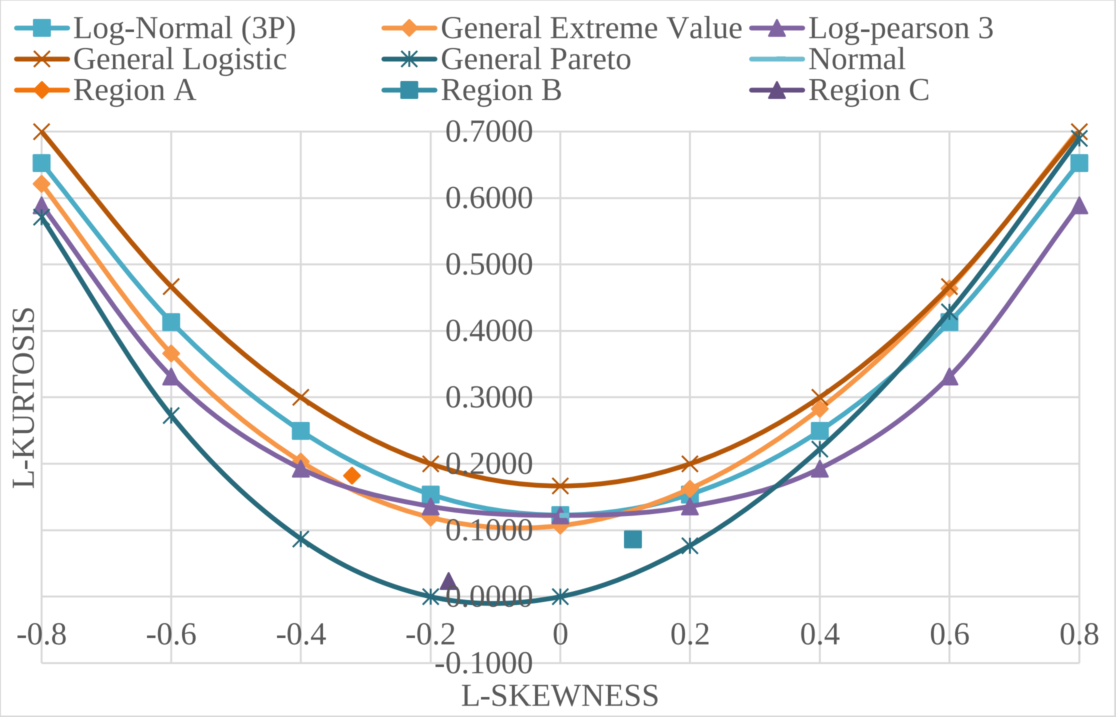


Figure 4.4: Regional weighted L-moment ratio diagram for the selected regions

4.3. Estimation of Regional Flood Frequency Curves

After regions have been accepted as homogeneous, suitable distributions were identified for the regions. The flood frequency curves were established for each region based on suitable distribution to calculate the deviations in the standardized flow of various return periods.

4.3.1. Parameter and Quantile Estimations

Estimation by the MML involves the choice of parameter estimates that produce a maximum probability of occurrence of the observations. The best parameter estimates from Easy Fit for selected distribution models were displayed as shown in Table 4.9. These results were generated according to the ranks and descriptive statistics of the goodness fit tests shown in Table 4.8 and Appendix-H. As a result, these distributions could be adopted as the appropriate and found to be the dominating distribution in the Tekeze River Basin for accurate evaluation and estimation of floods.

Estimation of flood quantiles was applied for 2, 5, 10, 15, 20, 25, 50, 75,100, 200, 500 and 1000 years return period and flood frequency curves for regions were developed. Flood frequency curves were estimated using equation 3.28 and 3.31. This estimation of the flood can be utilized in the designing of vital hydraulic structures in the river reach.

Table 4.9: Results of estimation parameters for fitted distributions in the region

Name of Regions	Best-fitted distribution	Value of parameters		
		K	σ	μ
Region-A	Generalized Extreme Value	0.06119	29.33	29.176
Region-B	Generalized Pareto	-1.0817	2206	417.86
Region-C	Generalized Pareto	-0.6755	298.4	80.132

4.3.2. Standard Error of Parameter Estimation

The standard error of a flood estimate indicates the reliability of that estimate. The development of the relationship between the mean annual flood or index flood and the catchment characteristics was a necessary step in predicting flood magnitudes at any point in a region where the flood frequency curve has been derived and error in quantitative terms.

Standard error measure is the most common measure of estimation. From the various source of error only sampling error can be evaluated theoretically a consensus seems to be emerging that at least sampling error should normally be reported in quantitative terms. The least estimator values are the most robust flood estimation for a given region and station. Selection of the most efficient method that gives the smallest standard error of estimate. Standard error of estimate (SEE) were estimated using equation 3.33. Depending on the result of SEE the best fit parameter estimation and distributions were selected.

Table 4.10: Standard error for the selected distribution of Region -A

T	Distribution		
	GEV/MLM	GEV/MOM	GEV/PWM
2	0.003	0.026	0.183
5	0.026	0.146	0.515
10	0.070	0.255	0.606
15	0.111	0.315	0.635
20	0.150	0.355	0.650
25	0.188	0.386	0.659
50	0.366	0.480	0.677
75	0.534	0.532	0.683
100	0.694	0.569	0.686
200	1.297	0.654	0.691
500	2.931	0.763	0.695
1000	5.410	0.841	0.696
Avg. SEE	0.982	0.444	0.615

Table 4.11: Standard error for the selected distribution of Region -B

T	Distribution		
	GPA/MLM	GPA/MOM	GPA/PWM
2	0.002	0.072	0.145
5	0.015	0.175	0.169
10	0.038	0.256	0.299
15	0.062	0.326	0.353
20	0.087	0.348	0.385
25	0.113	0.368	0.407
50	0.244	0.479	0.460
75	0.382	0.536	0.483
100	0.523	0.564	0.497
200	1.114	0.674	0.523
500	3.010	0.763	0.546
1000	6.377	0.876	0.558
Avg. SEE	0.997	0.453	0.402

Table 4.12: Standard error for the selected distribution of Region -C

T	Distribution		
	GPA/MLM	GPA/MOM	GPA/PWM
2	0.008	0.296	0.071
5	0.066	0.405	0.276
10	0.140	0.506	0.285
15	0.203	0.673	0.397
20	0.259	0.707	0.399
25	0.311	0.728	0.486
50	0.531	0.732	0.495
75	0.717	0.788	0.586
100	0.883	0.796	0.596
200	1.446	0.809	0.675
500	2.736	0.818	0.677
1000	4.405	0.821	0.758
Avg. SEE	0.975	0.673	0.475

According to the result presented on the above table the selected distributions and method of parameter estimation gives less value of standard error of estimation, Therefore, it can be selected for estimation of quantiles.

4.3.3. Estimation of Index-Flood for Standardization

In this case, the average of the growth curves was determined to represent the flood frequency curves of regions. The results of Table 4.13 show that the standardized quantiles for regions using the selected distribution and parameters with their corresponding return periods. It was observed that the magnitude of flood increases as the return period increases for selected distribution parameter for all stations. This may be due to the variability of the flood regimes of hydrological phenomena generating the flood events. This can significantly help in risk assessment works, water resources management, and engineering decisions and actions in the study area.

Table 4.13: Estimated standardize flood quantiles of Regions

Gumbel reduced variate	RGC-A	RGC-B	RGC-C
0.367	0.953	0.767	0.189
0.910	1.261	1.272	2.111
1.363	1.455	1.480	2.662
1.643	1.561	1.568	2.846
1.847	1.634	1.619	2.941
2.009	1.689	1.654	2.999
2.528	1.855	1.739	3.119
2.841	1.949	1.777	3.162
3.068	2.014	1.799	3.185
3.626	2.167	1.841	3.220
4.386	2.360	1.878	3.244
4.974	2.500	1.897	3.253

(RGC: Regional Growth Curve)

Depending on selected distributions, regional growth curves were derived as indicated in Figure 4.5. Based on figure 4.5 revealed that lower elevation catchments have lower flood values but higher extreme flood variability than higher elevation catchments.

The constructed regional frequency curves from three regions reflect that all curves have different flood characteristics. This could be due to the fact that the flood in different regions has different flood statistics. As indicated in Figure 4.5, the derived regional growth curve of Region-C was revealed higher quantile estimates than Region-A and B, for the same return periods. This high flood within the region might cause tremendous damages and disruptions to local communities. This could be attributed to the variability in their flood regimes and the corresponding contributing areas.

The higher variations of regional curves may be due to the considerable spatial fluctuations of elevations with their spatially undulating mountainous topography of regional boundaries, which causes uncertainties in flood prediction.

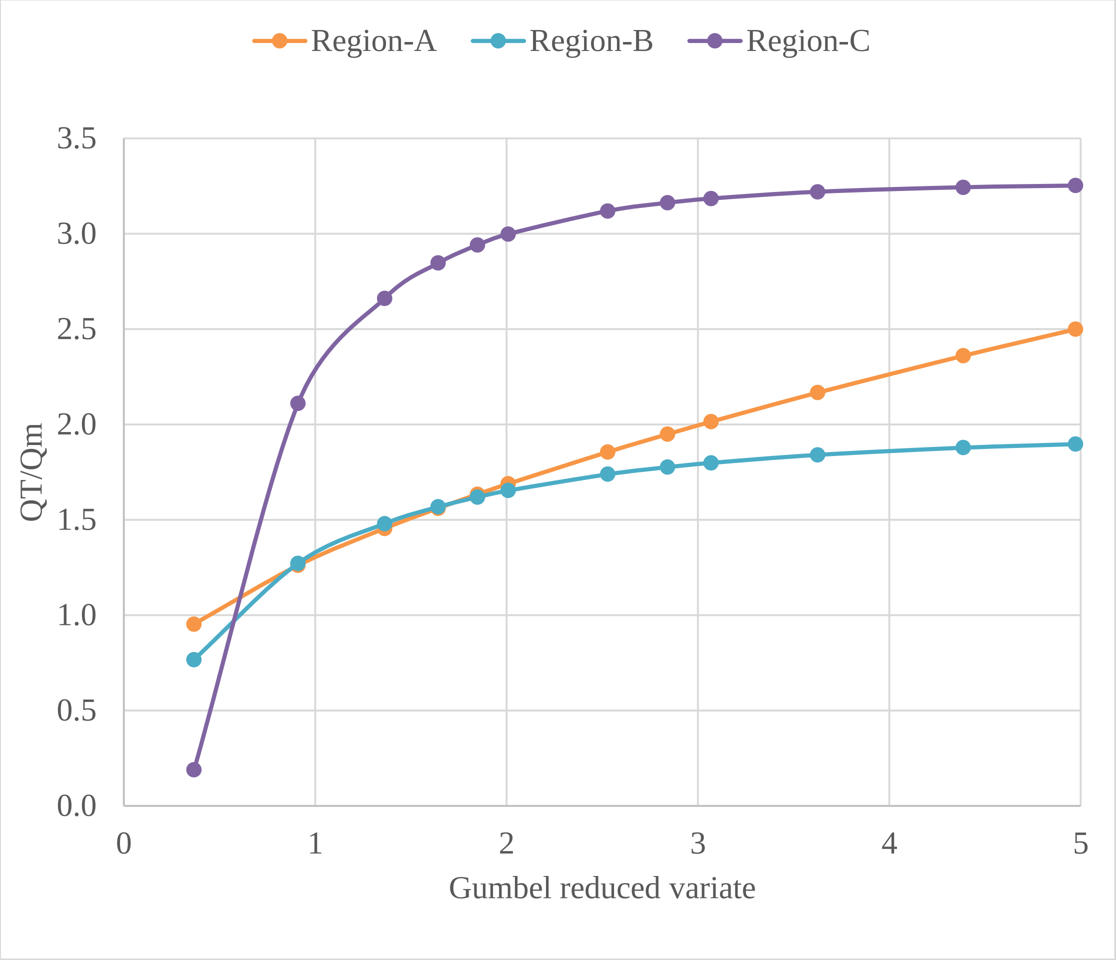


Figure 4.5: Regional growth curves for delineated homogeneous regions

4.3.4. Flood Frequency Curves

In this study, Flood frequency curves were plotted on the bases of different return periods versus the estimated flood quantiles values (X_T) as shown in Figure 4.6.

Estimated flood quantiles help to develop policies, which will reduce risk and damage from extreme flood events in both short and long-term planning which might happen in the study area. The estimated flood frequency curves at a given return period advances the accuracy and reliability of flood risk estimations for regions.

Table 4.14: Estimated flood quantiles of Regions

T (year)	FFC-A (m3/s)	FFC-B (m3/s)	FFC-C (m3/s)
2	154	120	105
5	371	233	354
10	482	439	569
15	588	641	873
20	792	842	976
25	895	943	1077
50	1204	1046	1480
75	1410	1150	1683
100	1513	1254	1886
200	1822	1461	2193
500	2133	1670	2604
1000	2241	1883	2914

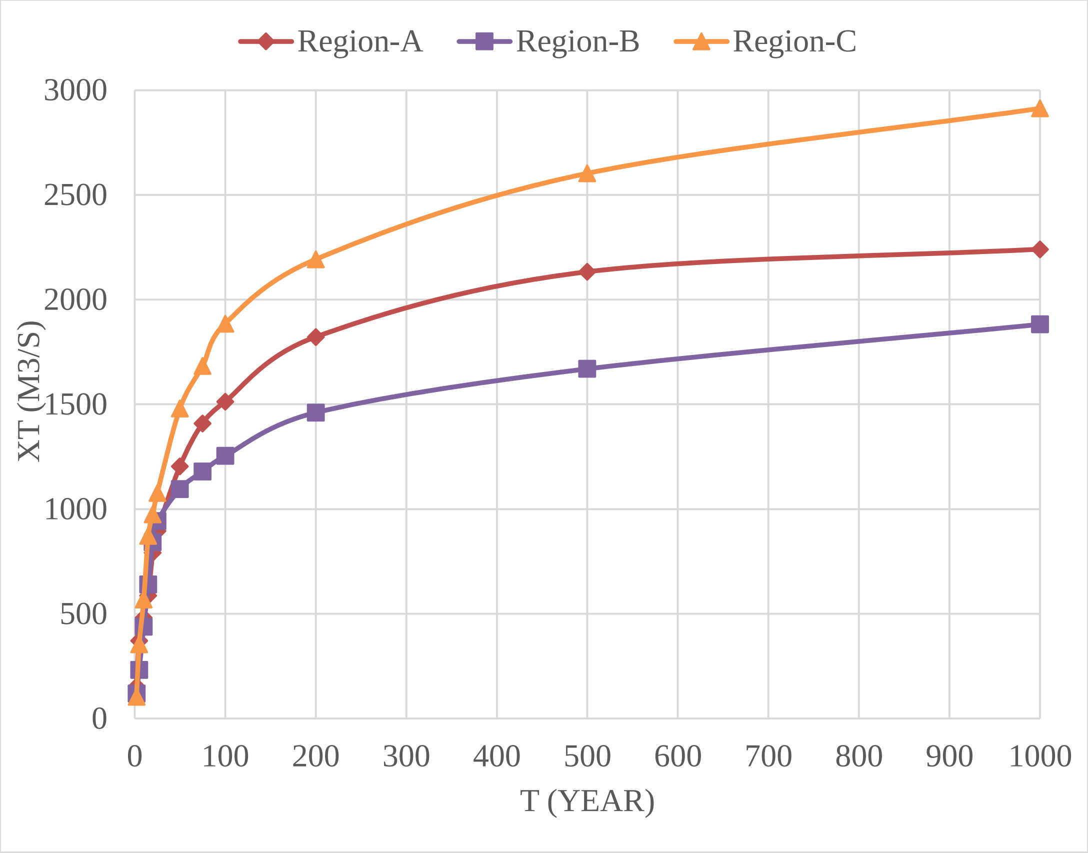


Figure 4.6: Flood frequency curves of regions

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

In this study, regional flood frequency analysis was performed using the data of eleven stream gauging stations so as to ensure reliable estimation of flood in Tekeze River Basin. The main objective of the study was to delineate the Tekeze river basin into hydrological homogeneous regions which would form the basic units to form and develop frequency curves for each region. The basin has defined and delineated into three hydrologically homogeneous regions using AMF frequency model. The regions were named as Region-A, Region-B and Region-C comprising seven, two and two gauging sites respectively. The delineation of the regions was done with ArcGIS10.4.1. LMRD and Easy fit software were used to check weather all stations in the same region are found to lie on the same type of distribution. Further, a discordance measure using Matlab2018a and CC test was conducted to check their homogeneity.

Regional average values of LCs and LCk were used to select the best fit statistical distribution of each region and goodness of fit test by using the Easy Fit software was used to approve the best fit distribution. Standard error estimation was conducted to select a superlative method of parameter estimation for the selected distribution. For Region-A Generalized Extreme Value (GEV) with MOM was selected, for Region-B Generalized Pareto(GPA) with PWM was selected and finally Generalized Pareto(GPA) with PWM was selected for Region-C. And these distributions with method of parameter estimations are finally used to develop a regional growth curve of each homogeneous region.

The regional growth curve can be used to safely and feasibly design hydrologic projects under prediction in both gauged and ungauged catchments. The derived results can be useful as a reference in any hydrological considerations like flood risk management, proper planning, and designing of pivotal hydraulic structures such as dams, spillways, bridges, culverts, and urban drainage systems in the study area.

5.2. Recommendations

Based on the result obtained in this study, the following recommendations are made for further work in the area.

Delineation of homogenous regions based on statistical parameter of gauged site could be one of an alternative method of regionalization to identify stations of similar flood producing characteristics. Due to the adequacy of best-fit distributions and acceptability of results, Easy Fit statistical software can be used for other related studies. Matlab and other programming should be used to simplify and get the accurate and reasonable results of any statistical analysis.

Testing of stations for homogeneity using statistical methods to form homogeneous regions considering other geographical, topographical and altitude factors is a good method in regional flood frequency analysis of the basin. Stations having different distribution with the same number of parameter and method of parameter estimation are statistically similar and depending on other external factors they can be categorized under the same region.

Flood frequency curves should be developed using varies types of catchment characteristics such as elevation, slope, area, precipitation, soil type, land use land cover and shape factor to compare the results and get a more reasonable flood estimation for ungauged catchments.

In order to get a reliable estimate of regional flood quantile more hydrometric stations should be installed in the basin.

Usually flood frequency analysis is done by using statistical distribution technique (Easy fit software). But, this method is not always the best and efficient method, so try to use another soft wares are important in some cases. For example: Hyfran plus software.

It is advisable to extend this approach of regional flood frequency analysis for other Ethiopian river basins to establish the homogeneous regions so that problems related to absence of sufficient discharge data for water resources project planning and design could be reduced.

REFERENCES

Abdulla, M. (2011). Catchment Area Delineation Using GIS Technique for Bekhma Dam. Spatial Information Processing II, Paper No. 5335.

Ahmad, I., Fawad, M. and Mahmood, I. (2015). At-site flood frequency analysis of annual maximum stream flows in Pakistan using robust estimation methods. Polish Journal of Environmental Studies, 24(06): 2345-2353.

Ahmad, U.N., Shabri, A. and Zakaria, Z.A. (2011). Flood Frequency Analysis of Annual Maximum Streamflow Using L-Moments and TL-Moments. Applied Mathematical Sciences, Vol. 5, 243-253.

Akirso, B. (2017). Flood Frequency under Changing Climate in the Upper Omo-Gibe River Basin, Southwestern Ethiopia: Case Study Abelti-Catchment.

Amalina, N.M, Shabri, A. and Badyalina, B. (2016). Selecting Probability Distribution for Regions of Peninsular Malaysia Streamflow. AIP Conference Proceedings 1750, 060014.The American Institute of Physics. doi.org.: 10.1063/1.4954619.

Ashraful A. M., Emura, K., Farnham, C. and Yuan. (2018). Best-Fit Probability Distributions and Return Periods for Maximum Monthly Rainfall in Bangladesh. The Journal of climate. Vol.6 (9).

Athulya, P. S. And James, K.C. (2017). Best Fit Probability Distributions for Monthly Weather Data. International Journal of Advances in Management, Technology and Engineering Sciences. Vol.7 (12).

Atiem, IA., Harmancioglu, N. (2006). Assessment of regional floods using L-moments approach: the case of the River Nile. Water Resource Management 20:723-747.

Badreldin, G. and Fengo, P. (2012). Regional Rainfall Frequency Analysis for the Luanhe Basin using L-moment and Cluster Techniques.

Baguis, P., Boukhris, O., Ntegeka, V., Roulin, E., Willems, P., Demaree, G. (2008). Climate change impact on flood frequency along rivers and urban drainage systems. Technical report CCIHYDR project by K.U. Leuven- Hydraulics Section & Royal Meteorological Institute of Belgium. May 2008. 57 p.

Baratti, E., Montanari, A., Castellarin, A., Salinas, J.L., Viglione, A., Bezzi, A. (2012). Estimating the Flood Frequency Distribution at seasonal and Annual Time Scale. Hydrol. Earth Syst. Sci. 9, 7947-7967.

Bhagat, N. (2017). Flood Frequency Analysis Using Gumbel's Distribution Method: A Case Study of Lower Mahi Basin. Journal of Water Resources and Ocean Science.Vol.6(4),5154.ISSN:23287969;ISSN:23287993.doi:10.11648/j.wros .20170604.11.

Boughton, W., Droop, O. (2003). Continuous simulation for design flood estimations. Environ Modell Softw,18(4):309-318.DOI:http://dx.doi.org/10.1016/S1364-8152(03)00004-5.

Castellarin, A., Kohnová, S., Gaál, L., Fleig, A.K., Salinas, J.L., Toumazis, A., Kjeldsen, T.R. and Macdonald, N. (2011). Review of applied statistical methods for flood frequency analysis in Europe. WG2 of COST Action ES0901.

Chebana, F., Ouarda, T.B.M.J. (2008). Depth and Homogeneity in Regional Flood Frequency Analysis. Water Resources Research, 44(11), W11422, doi:10.1029/2007WR006771.

Cunnane, C. (1989). Statistical distribution for flood frequency analysis. World Metrological Organization, Operational Hydrology Report No.33, WMO-NO.718. Geneva, Switzerland.

Dahmen, E.R. and Hall, M.J. (1990). Screening of Hydrological Data: Tests for Stationarity and Relative Consistency. Wageningen, The Netherlands. International Institute for Land Reclamation and Improvement.

Das, S. and Cunnane, C. (2011). Examination of homogeneity of selected Irish pooling groups. Hydrol. Earth Sys. Sci., 15, 819–830.

Das, S. and Simonovic, S.P. (2012). Assessment of Uncertainty in Flood Flows Under Climate Change-the Upper Thames River basin (Ontario, Canada). Water Resources Research Report.

Demissie, M. and Michael, Y. (2008). Regional Flood Frequency Analysis for Upper Awash Sub-basin (Upstream of Koka). MSc. Thesis Submitted to Addis Ababa University.

Demissie, M. (2008) Regional Flood Frequency Analysis for Upper Awash Sub-Basin (Upstream of Koka).

Dessalegn, B., Ayalew, S. and Hailu, D. (2016). Evaluation of Extreme Flow Quantiles Estimated from Global Reanalysis Runoff data. A Case Study on Blue Nile River Basin. MSc. Thesis Submitted to Addis Ababa University.

Di Baldassarre, G., Laio, F. and Montanari, A. (2009). Design flood estimation using model selection criteria. Physics and Chemistry of the Earth, Parts A/B/C, 34(10):606–611.

Ding, J., and Haberlandt, U. (2017). Estimation of instantaneous peak flow from maximum mean daily flow by regionalization of catchment model parameters.HydrologicalProcesses,31(3),612626.https://doi.org/10.1002/hyp.11053.

Donnelly, C.R., Zhou, R.D. and Judge, D.G. (2008). On the Relationship Between the 10,000 Year Flood and Probable Maximum Flood. Conference Paper No. 131.

Ergish, N. J. (2010). Flood Frequency Analysis for Regulated Watersheds.

Fentaw, F, Hailu, D. and Nigussie, A. (2017). Trend and Variability Analysis of Rainfall and Stream Flow Series at Tekeze River Basin, Ethiopia. International Journal of Scientific and Engineering Research. Vol. 8 (11), 2229-5518.

Gebeyehu, A. (1989). Regional Flood Frequency Analysis, PhD. Dissertation Submitted to Stockholm University. Hydraulics Laboratory, the Royal Institute of Technology. Florida: CRC Press LLC.

Gedefa, T.W. and Seleshi, Y. (2009). Regional Flood Frequency Analysis for Upper Omo-Gibe Sub-basin; MSc. Thesis Submitted to Addis Ababa University.

Ghosh, S., Roy, M.K., and Biswas, S.C. (2016). Determination of the Best Fit Probability Distribution for Monthly Rainfall Data in Bangladesh. American Journal of Mathematics and Statistics.Vol.6 (4), 170-174.DOI: 10.5923/j.ajms.20160604.05.

Gottschalk, L. and Krasovskaia, I. (2001). Regional flood frequency analysis, a theoretical background. Lecture notes in Stochastic Hydrology. Department of Geophysics. University of Oslo.

Guru, N. and Jha, R. (2016). Flood Frequency Analysis of Partial Duration Series Using Soft Computing Techniques for Mahanadi River Basin. Aquatic Proceedings. Vol. 4, 427-434.

Haberldin, U and Radtke,I. (2014). Hydrological Model Calibration for Derived Flood Frequency Analysis using Stochastic Rainfall and Probability Distribution of Peak Flow.

Haddad, K. (2013). Regional Flood Frequency Analysis in the Range of Small to Large Floods: Development and Testing of Bayesian Regression-based Approaches.

Hailegeorgis, T.T and Alfredsen, K. (2017). Regional Flood Frequency Analysis and Prediction in Ungauged Basins including Estimation of Major Uncertainties for Mid-Norway.

Haktanier, T., Cobaner, M., and Kisi, O. (2010). Frequency Analysis of Annual Extreme Rainfall. Journal of Hydrological Process.

Harmancioglu, N. (2006). Assessment of regional flood using L-moments approach.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society. Series B, Methodological 52, 105-124.

Hosking, J. R. M., and Wallis, J. R. (1993). Some statistics useful in regional frequency analysis. Water Resour. Res., 29(2), 271–281.

Hosking, J.R.M. and Wallis, J.R. (1997). Regional frequency analysis. An approach based on L-moments. Cambridge University Press, New York.

Hussein, B. and Wagesho, N. (2016). Regional Flood Frequency Analysis for Abaya-Chamo Sub-Basin, Rift-Valley River Basin, Ethiopia: Journal of Resources Development and Management. An International Peer-Reviewed Journal: Vol.24: ISSN 2422-8397.

Irwin, S.K, Srivastav, R. and Simonovic, S. P. (2014). Instruction for Watershed Delineation in an ArcGIS Environment for Regionalization Studies: Water Resources Research Report No: 87.

Javelle, P., Fouchier, C., Arnaud, P., Lavabre, J. (2010). Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations. Journal of Hydrology, 394(1-2): 267-274. DOI: DOI 10.1016/j.jhydrol.2010.03.032.

Jingyi, Z. & Hall, M.J. (2004). Regional flood frequency analysis for Gan-Ming river basin in China. J. Hydrol., 296, 98–117.

Jos, K. (2017). Flood Frequency Analysis of the Yewa River at Ijaka-Oke, Ogun state, Southwest Nigeria. Nigerian Journal of Hydrological Sciences, Volume 5, pp.112-122.

Kachroo, R. K., Mkhandi, S. H., & Parida, B. P. (2000). Flood frequency analysis of southern Africa: I. Delineation of homogeneous regions. Hydrological Sciences Journal- Journal Des Sciences Hydrologiques, 45(3), 437-447.

Kanti, K. K., Sung-Kee Y., Jun-Ho, L. and Khan, K. (2017). Regional Flood Frequency Analysis for Using L-moments Approach in Jeju Island, Korea. Geo-environmental Disasters. Vol.4(18). DOI 10.1186/s40677-017-0082-0.

Ketsela, H., Tadele, K. and Temam, D. (2017). An Assessment of Predictive Accuracy for Regional Flood Frequency Distributions and Analysis

Estimation Methods on Awash River basin. MSc. Thesis Submitted to Jimma University.

Kjeldsen, T.R. and Jones, D. (2007). Estimation of an index flood using data transfer in the UK. Hydrol. Sci. J., 52(1), 86–98.

Kumar, R. and Chatterjee, C. (2005). Regional Flood Frequency Analysis using L-Moments for North Brahmaputra Region of India.

Kumar, R. and Chatterjee, C. (2011). Development of Regional Flood Frequency Relationships for Gauged and Ungauged Catchments using L-Moments. Journal of Hydrology, in Extremis. Descriptive events and trends in Climate and Hydrology. DOI 10.1007/978-3-642-14863-7-5.

Lim, Y.H. (2007). Regional Flood Frequency Analysis of the Red River Basin Using L-Moments Approach. World Environmental and Water Resources Congress: Restoring Our Natural Habitat.

Lim, Y.H. and Lye, L.M. (2003). Regional flood estimation for ungauged basins in Sarawak, Malaysia. Hydrological Sciences Journal 48(1), 79-94.

Lu, Y. (2016). Regional Flood Frequency Analysis for Newfoundland and Labrador Using theL-Moments Index-Flood Method.MSc. Thesis Submitted to the Memorial University of Newfoundland.

Malekinezhad, H., Nachtnebel, H.P. & Klik, A. (2011). Comparing the index flood and multiple regression methods using L-moments. J. Phys. Chem. Earth, 36, 54–60.

McCuen, R.H. (1998). Hydrologic Analysis and Design. Second Edition. Library of Congress Cataloging-in-Publication Data.

Mekoya, S.Y. and Seleshi, Y. (2010). Regional Flood Frequency Analysis Upstream of Awash at the Confluence of Kesem River. MSc. Thesis submitted to Addis Ababa University.

Mengistu, D. (2008). Regional flood frequency analysis for upper awash sub- basin (Upstream of koka), MSc thesis submitted to Addis Ababa University.

Mengistu, K.H and Sivakumar, K. (2018). An Assessment of Predictive Accuracy for Regional Flood Frequency Distribution Estimation Methods on Awash River Basin. International Journal of Latest Trends in Engineering and Technology.Vol.11(1),30039.ISSN:2278621X:DOI:http://dx.doi.org/10.21172/1.111.06.

Mishra, B.K., Takara, K., Yamashiki, Y. and Tachikawa, Y. (2009). Selection of Regional Distribution Using Simulated Flood Data. Annual Disaster Prevention Research Institute, Kyoto University, No.52B.

Mkhandi, S. H., Kachroo, R. K., & Gunasekara, T. A. G. (2000). Flood frequency analysis of southern Africa: II. Identification of regional distributions. Hydrological Sciences Journal-Journal Des Sciences Hydrologiques, 45(3), 449-464.

MoWR. (1998). Tekeze River Basin Integrated Resources Development Master Plan Project.

Murphy, C., Cunnane, C., Das, S. and Mandal, U. (2014). Flood Frequency Estimation. Flood Studies Update: Technical Research Report. Volume II.

Mustafa, S. and Yusuf, M.I.Y. (2012). A Textbook of Hydrology and Water Resources. Topsmerit Page Publishers, Abuja, Nigeria.

Nobert, J., Mugob, M. and Gadain, H. (2014). Estimation of Design Floods in Ungauged Catchments Using a Regional Index Flood Method. A Case Study of Lake Victoria Basin in Kenya: Journal of Physics and Chemistry of the Earth. 30(30),1474-7065.http://dx.doi.org/10.1016/j.pce.2014.02.001

Noto, L.V. and Loggia, G.L., (2009). Use of L-Moments Approach for Regional Flood Frequency Analysis in Sicily, Italy. Journal of Water Resources Management.Vol.3, 2207–2229. DOI 10.1007/s11269-008-9378-x.

Onoz, B. and Bayazit, M. (2001). Power distribution for low stream flows. Journal of Hydrologic Engineering, 6(05):429-435.

Pakgohar, A. (2014). Using Easy Fit Software for Goodness-of-Fit Test and Data Generation. International Journal of Mathematical Archive-5(1), 7(118-124).

Pathiraja, S., Westra, S., Sharma, A. (2012). Why continuous simulation? The role of antecedent moisture in design flood estimation. Water Resour. Res., 48(W06534): 1-15. DOI: doi:10.1029/2011WR010997.

Pegram, G and Parak, M. (2004). A review of the regional maximum flood and rational formula using geomorphological information and observed floods, Water SA 30 (3) 377-388.

Peel, M. C., Wang, Q. J., Vogel, R. M., and McMAHON, T. A. (2001). The utility of L-moment ratio diagrams for selecting a regional probability distribution. Hydrological sciences journal. 46(1), 147 -155.

Rao, A.R. and Hamed, K.H. (2000). Flood Frequency Analysis. CRC Press, Washington.

Rao, A.R. and Srinivas, V.V. (2008). Regionalization of Watersheds by Hybrid Cluster Analysis. Journal of Hydrology. ISBN: 978-1-4020-6851-5/ISBN: 978-1-4020-6852-2.

Rosbjerg, D. (2007). Regional flood frequency analysis. Extreme Hydrological Events: New Concepts for Security, 78, 151-171.

Saf, B. (2009).Regional Flood Frequency Analysis using L-Moment for the West Mediterranean Region of Turkey.Journal of Water Resources Management.Vol.23(3),531-551. DOI 10.1007/s11269-008-9287-z.

Sah, S. and Prasad,J. (2017). FloodFrequency Analysis of Kumaon Region Rivers Uttarakhand,India, International Journal of Engineering Researches and Management Studies, http://www.ijerms.com,4(2)ISSN: 2394-7659.

Sine, A. (2004). Regional Flood Frequency Analysis Model for Blue Nile Thesis Report.

Sine, A. and Ayalew, S. (2004). Identification and Delineation of Hydrological Homogeneous Regions. The Case of Blue Nile River Basin. Lake Abaya Research Symposium Proceedings. Vol.4.

Sine, A. G., Ayalew, S. M., and Bekele, S.A. (2013). Basin Regionalization for the Purpose of Water Resource Development in Limited Data Situation: Case of Blue Nile River Basin, Ethiopia. Journal of Hydrologic Engineering. Vol. 18(10),1084-0699.

Solana, AO., Solana, V. (2001). Entropy-based inference of simple physical models for regional flood analysis. Stoch Environ Res Risk A 15:415–446.

Tadesse, A.H., Gottschalk, L. and Tallaksen, L.M. (2011). Regional Flood Frequency Analysis in Southern Africa, MSc. Thesis Submitted to the University of Oslo.

Tallaksen, LM., Madsen, H., Hisdal, H. (2004). Frequency Analysis, Hydrological Drought–Processes and Estimation Methods for Streamflow and Ground water, Developments in Water Sciences 48. Elsevier Science Publisher: The Netherlands.

Tumbare, M. J. 2000. Mitigating floods in Southern Africa. Paper presented at the1st WARSFA/Water Net Symposium: Sustainable Use of Water Resources, 1-2 November, Maputo.

Viglione, A., Laio, F., & Claps, P. (2007). A comparison of homogeneity tests for regional frequency analysis. Water Resources Research, 43(3), W03428.

Vivekanandan, N. (2015). Flood Frequency Analysis using Method of Moment and L-Moment of Probability Distribution.

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

Willems, P.K., Abdo, G., Sonbol, M. (2003). Flood Frequency Analysis of the Eastern Nile Rivers.

Wilson, D., Fleig, A. K., Lawrence, D., Hisdal, H., Pettersson, L.E, Holmqvist, E. (2011). A review of NVE's flood frequency estimation procedures in Norway, Norwegian Water Resources and Energy Directorate.

Yang, T., Xu, C. Y., Shao, Q. X., & Chen, X. (2010). Regional flood frequency and spatial patterns analysis in the Pearl River Delta region using L-moments approach. Stochastic Environmental Research and Risk Assessment, 24(2), 165-182.

Yurtal, R. (2010). Comparison of Probability Weighted Moments and Maximum Likelihood Methods used in Flood Frequency Analysis for Ceyhan River Basin.

Yucel, I. and Keskin, F. (2011). Assessment of Flash Flood Events using Remote Sensing and Atmospheric Model-Derived Precipitation in A Hydrological Model, in Melbourne, Australia.

Yue, S. and Wong, C. (2004). Possible regional probability distribution type of Canadian annual stream flow in the united states.

Zaman, M. A. (2013). A Study on Selection of Probability Distributions for at-Site Flood Frequency Analysis in Australia.Melbourne, VIC, Australia.

Zaman, M.A., Rahman, A. & Haddad, K. (2012). Regional flood frequency analysis in arid regions. A case study for Australia. J. Hydrol., 475(0): 74-83.

APPENDIXES

Appendix-A: Results for correlation of gauging stations used for analysis

Code	Gauging station(Y)	Nearby station(X)	Regression equation	R^2	Remark
121004	Mekele	Aynalem	$y = 0.4098x + 2.5875$	0.7812	WC
121006	Embamadre	Maitsemry	$y = 35.06x + 529.94$	0.9193	WC
121007	Hawzen	Wukro	$y = 0.1173x - 0.4129$	0.7954	WC
121008	Adikumsi	Mekele	$y = 7.5865x - 124.39$	0.8246	WC
121010	Wukro	Hawzen	$y = 0.1721x + 6.5064$	0.8853	WC
121012	Aynalem	Mekele	$y = 0.3987x - 8.0479$	0.6236	WC
121014	Adwa	Hawzen	$y = 0.1332x -3.3624$	0.7726	WC
121023	Kulmesk	Adikumsi	$y = 2.825x - 50.553$	0.8408	WC
122002	Humera	Dansha	$y = 14.741x - 91.083$	0.9068	WC
122003	Maitsemry	Embamadre	$y = 5.1657x - 52.693$	0.9137	WC
123049	Dansha	Humera	$y = 1.2194x - 13.172$	0.6000	WC

(WC: Well Correlated)

Appendix–B: Critical values of the Grubbs T Test Statistic as a function of the number of Observations and Significance level

N	5%	2.50%	1%	N	5%	2.50%	1%
3	1.15	1.15	1.15	20	2.56	2.71	2.88
4	1.46	1.48	1.49	21	2.58	2.73	2.91
5	1.67	1.71	1.75	22	2.6	2.76	2.94
6	1.82	1.89	1.94	23	2.62	2.78	2.96
7	1.94	2.02	2.1	24	2.64	2.8	2.99
8	2.03	2.13	2.22	25	2.66	2.82	3.01
9	2.11	2.21	2.32	30	2.75	2.91	
10	2.18	2.29	2.41	35	2.82	2.98	
11	2.23	2.36	2.48	40	2.87	3.04	
12	2.29	2.41	2.55	45	2.92	3.09	
13	2.33	2.46	2.61	50	2.96	3.13	
14	2.37	2.51	2.66	60	3.03	3.2	
15	2.41	2.55	2.71	70	3.09	3.26	
16	2.44	2.59	2.75	80	3.14	3.31	
17	2.47	2.62	2.79	90	3.18	3.35	
18	2.5	2.65	2.82	100	3.21	3.38	
19	2.53	2.68	2.85				

(source: Grubbs,1969)

Appendix–C: Percentile Points of the F-Distribution F {V1, V2, P} for the 5 % level of Significance (Two-Tailed)

P=P(F<FP)		V1:4	5	6	7	8	9	10	11	12	14	16
0.025 0.975	V2:5	.107 .739	.140 7.15	.169 6.98								
0.025 0.975	6		.143 5.99	.172 5.82	.195 5.70							
0.025 0.975	7			.176 5.12	.200 4.99	.221 4.90						
0.025 0.975	8				.204 4.53	.226 4.43	.244 4.36					
0.025 0.975	9					.230 4.10	.248 4.03	.265 3.96				
0.025 0.975	10						.252 3.78	.269 3.72	.284 3.66			
0.025 0.975	11							.273 3.53	.288 3.47	.301 3.43		
0.025 0.975	12								.292 3.32	.305 3.28	.328 3.21	
0.025 0.975	14									.312 3.05	.336 2.98	.355 2.92
		V1:14	16	18	20	24	30	40	60	100	160	∞
0.025 0.975	V2:16	.342 2.82	.362 2.76	.379 2.71								
0.025 0.975	18		.368 2.64	.385 2.60	.400 2.56							
0.025 0.975	20			.391 2.50	.406 2.46	.430 2.41						
0.025 0.975	24				.415 2.33	.441 2.27	.468 2.21					
0.025 0.975	30					.453 2.14	.482 2.07	.515 2.01				
0.025 0.975	40						.498 1.94	.533 1.88	.573 1.80			
0.025 0.975	60							.555 1.74	.600 1.67	.642 1.60		
0.025 0.975	100								.625 1.56	.674 1.48	.706 1.44	
0.025 0.975	160									.696 1.42	.733 1.36	
0.025 0.975	∞											1.00 1.00

(Source: Dahmen and Hall, 1990)

Appendix-D: Percentile Points of the t-distribution t {V, p for the 5% level of Significance (Two-Tailed)}

P = P(t < = tp)	0.025	0.975	P = P(t < = tp)	0.025	0.975
4	-2.78	2.78	16	-2.12	2.12
5	-2.57	2.57	18	-2.1	2.1
6	-2.54	2.54	20	-2.09	2.09
7	-2.36	2.36	24	-2.06	2.06
8	-2.31	2.31	30	-2.04	2.04
9	-2.26	2.26	40	-2.02	2.02
10	-2.23	2.23	60	-2	2
11	-2.2	2.2	100	-1.98	1.98
12	-2.18	2.18	160	-1.97	1.97
14	-2.14	2.14	∞	-1.96	1.96

(Source: Dahmen and Hall, 1990)

Appendix-E: Result of hydrological data quality test for stationarity of stations time series data

Station name	Subset-1	Subset-2	V1,V2	Ft2.5%	Ft	Ft97.5%	V	Tt2.5%	Tt	Tt97.5%
Adikumsi	1998-2007	2008-2017	10,10	0.269	0.775	3.720	20	-2.1	1.174	2.1
Adwa	1991-1999	2000-2007	9,8	0.244	0.571	4.360	17	-2.131	-1.511	2.131
Aynalem	1991-1998	1999-2006	8,8	0.226	0.646	4.430	16	-2.145	-0.469	2.145
Hawzen	1991-2002	2003-2013	12,11	0.301	0.838	3.430	23	-2.08	-1.379	2.08
Kulmesk	1996-2005	2006-2015	10,10	0.269	0.496	3.720	20	-2.1	-1.753	2.1
Mekele	2001-2008	2009-2015	8,7	0.204	0.967	4.530	15	-2.16	-0.087	2.16
Wukro	1992-2003	2004-2015	12,12	0.305	0.574	3.280	24	-2.074	1.348	2.074
Embamadre	1994-2004	2005-2015	11,11	0.288	0.994	3.470	22	-2.086	1.016	2.086
Maitsemry	2000-2007	2008-2014	8,7	0.204	4.207	4.530	15	-2.16	1.247	2.16
Dansha	2000-2007	2008-2014	8,7	0.204	0.688	4.530	15	-2.16	-1.204	2.16
Humera	1981-1988	1989-1995	8,7	0.204	0.984	4.530	15	-2.16	0.805	2.16

Appendix-F: (Translated Matlab code for Discordancy Measure as provided by Hosking and Wallis, 1997)

```
U=xls. read ('c:\users\name of group\desktop\U.xls'); % File
% ratios ($\tau_2^i$, $\tau_3^i$, $\tau_4^i$,) of the gauging sites in the region
U= number of gauging sites in the region (Enter the matrix of test statistics);
n=; % input ('enter the number of gauging sites in the group:');
Ubar= [0;0;0];
for i=1: n
Ubar=Ubar+1/n*(U(i,1:3)');
end
S=zeros (3);
for i=1: n
S=S+(U(i,1:3)'-Ubar) *(U(i,1:3)'-Ubar)';
end
for i=1: n
Di(i)=1/3*(U(i,1:3)'-Ubar)'*inv(S)*(U(i,1:3)'-Ubar);
End
disp ('The Di of U Statistics');
disp ('Di, Di+1,…. Dn');
```

Appendix-G: Candidate probability distributions of AMF for this study

To select the type of distribution which fit to the given data the following equations were used and obtained from Mishra et al. (2009).

- Normal distribution,

$$\tau_3 = 0, \tau_4 = 0.1226$$

- Logistic,

$$\tau_3 = 0,\ \tau_4 = \frac{1}{6}$$

- Generalized Pareto (GPA),

$$\tau_4 = 0.20196(\tau_3) + 0.95924(\tau_3)^2 - 0.20096(\tau_3)^3 + 0.04061(\tau_3)^4$$

- Log Normal Distribution,

$$\tau_4 = 0.12282 + 0.77518(\tau_3)^2 + 0.12279(\tau_3)^4 - 0.13638(\tau_3)^6 + 0.113638(\tau_3)^8$$

- Generalized Extreme Value (GEV),

$$\tau_4 = 0.10701 + 0.1109(\tau_3) + 0.84838(\tau_3)^2 - 0.06669(\tau_3)^3 + 0.00567(\tau_3)^4 - 0.04208(\tau_3)^5 + 0.03763\,(\tau_3)^6$$

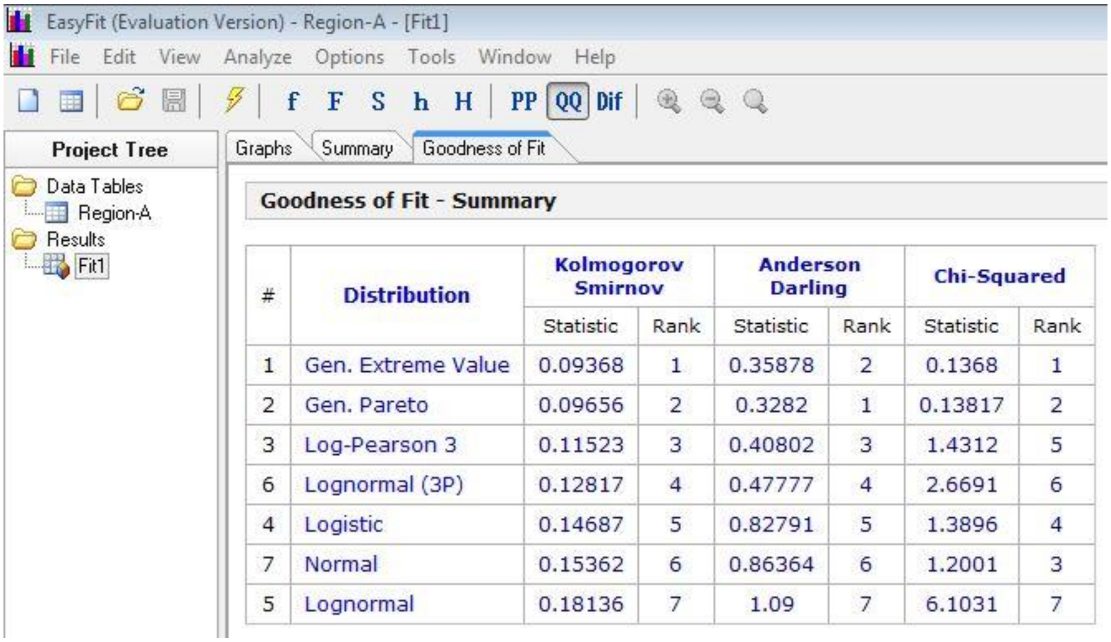
- Log-Pearson Type III,

$$\tau_4 = 0.1224 + 0.30115(\tau_3)^2 + 0.95812(\tau_3)^4 - 0.57488(\tau_3)^6 + 0.19383(\tau_3)^8$$

- Generalized Logistic,

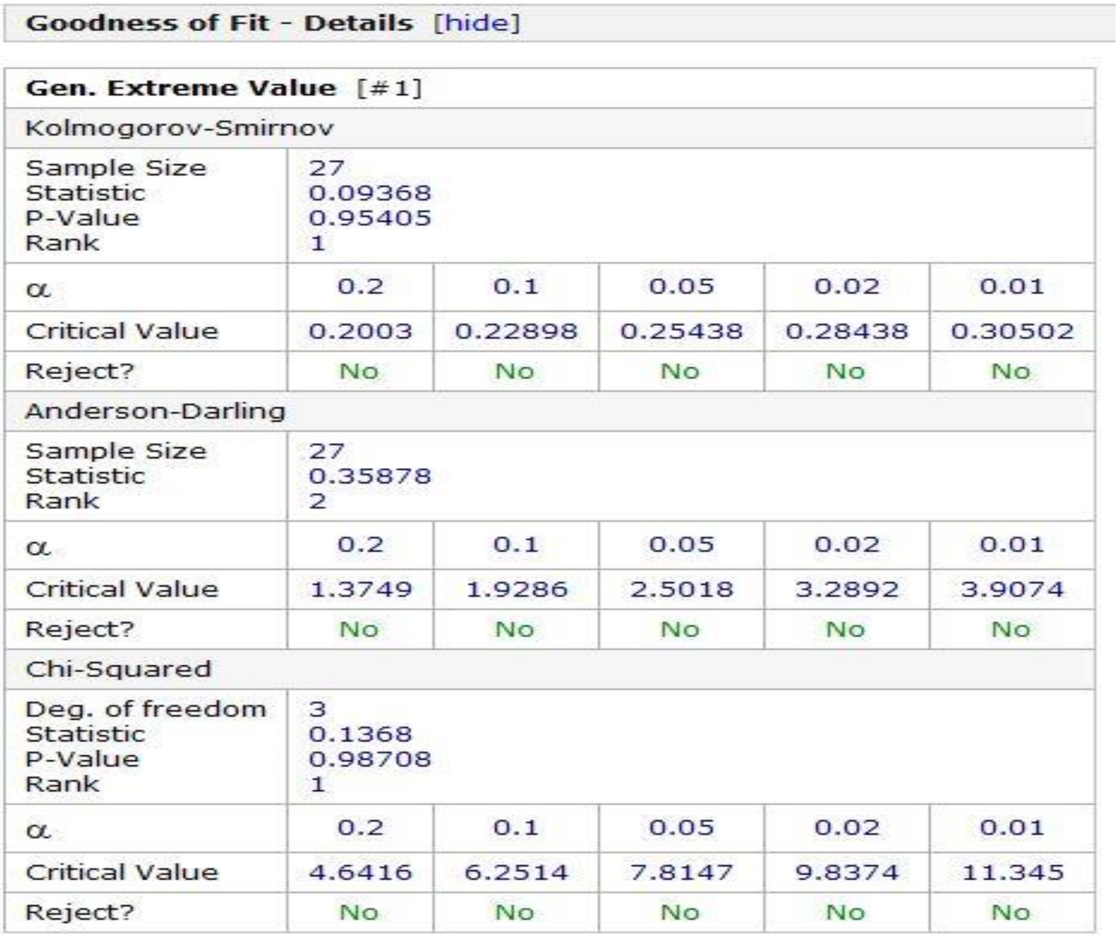
$$\tau 4 = 0.16667 + 0.83333(\tau_3)^2$$

Appendix-H: Goodness of fit test results and descriptive statistics for selected distribution of Regions



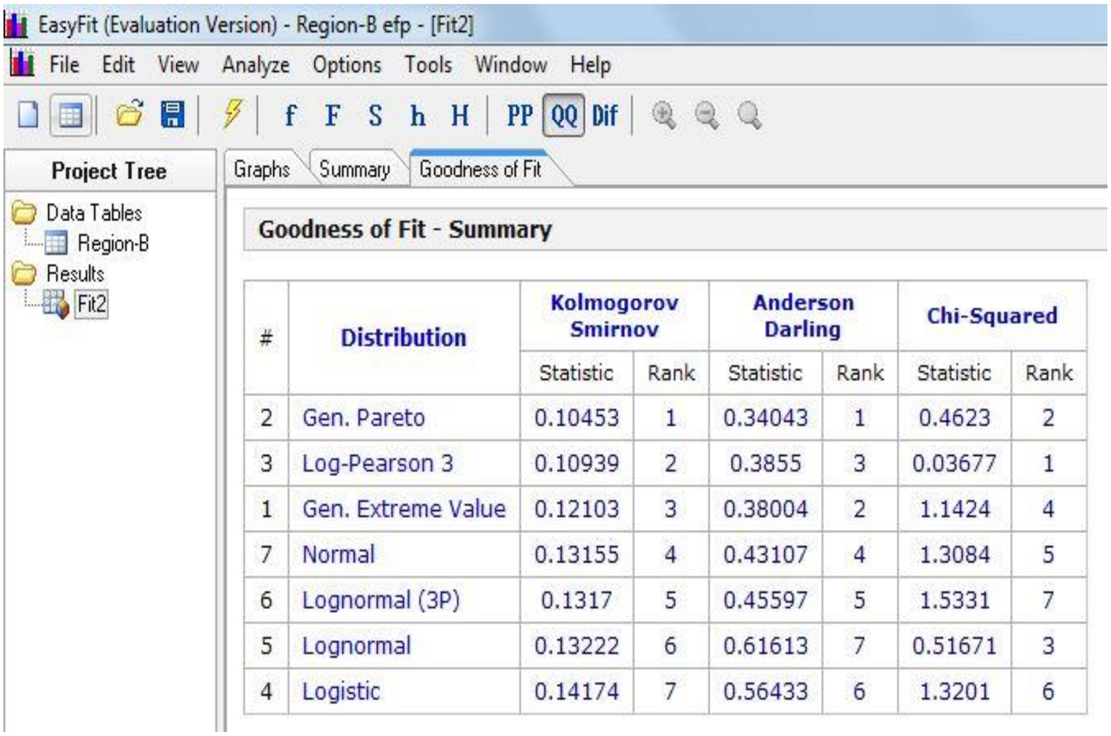
Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Gen. Extreme Value	0.09368	1	0.35878	2	0.1368	1
2	Gen. Pareto	0.09656	2	0.3282	1	0.13817	2
3	Log-Pearson 3	0.11523	3	0.40802	3	1.4312	5
6	Lognormal (3P)	0.12817	4	0.47777	4	2.6691	6
4	Logistic	0.14687	5	0.82791	5	1.3896	4
7	Normal	0.15362	6	0.86364	6	1.2001	3
5	Lognormal	0.18136	7	1.09	7	6.1031	7



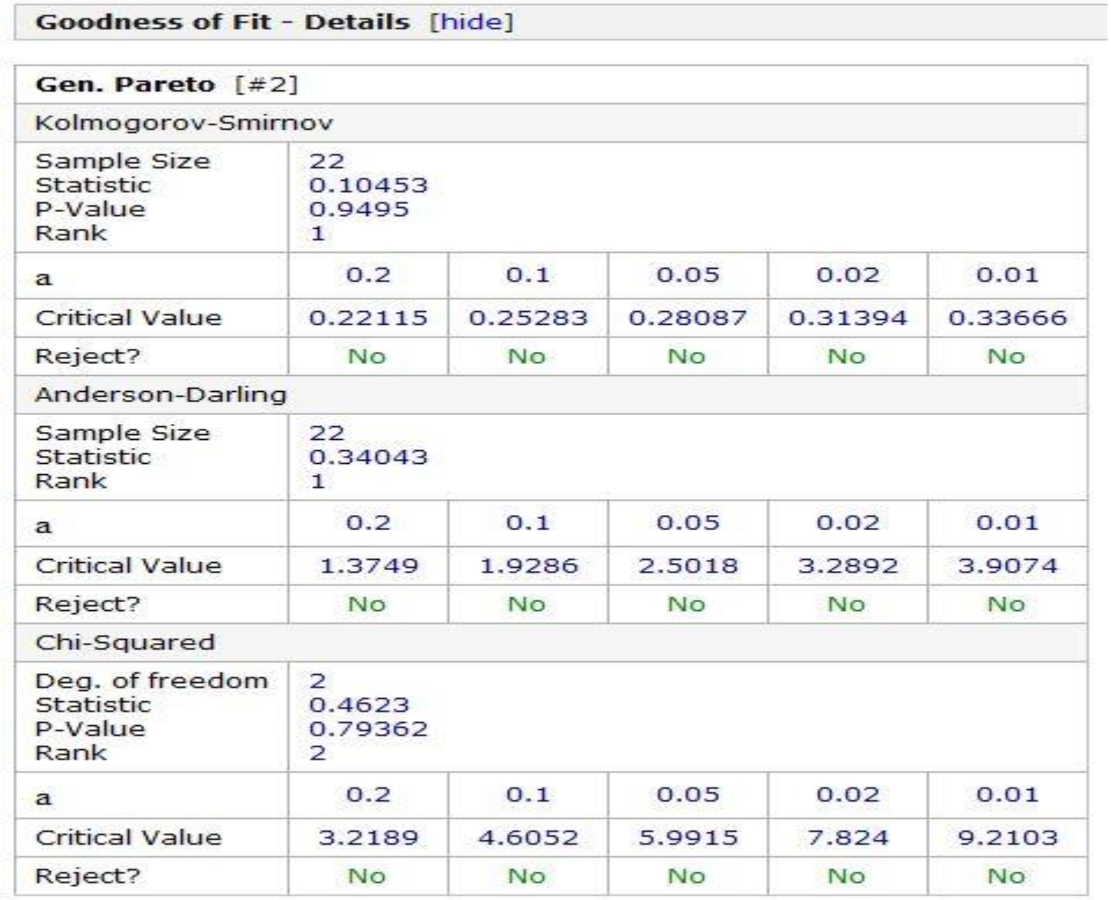
Goodness of Fit - Details [hide]

Gen. Extreme Value [#1]					
Kolmogorov-Smirnov					
Sample Size Statistic P-Value Rank	27 0.09368 0.95405 1				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	0.2003	0.22898	0.25438	0.28438	0.30502
Reject?	No	No	No	No	No
Anderson-Darling					
Sample Size Statistic Rank	27 0.35878 2				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	No	No	No	No	No
Chi-Squared					
Deg. of freedom Statistic P-Value Rank	3 0.1368 0.98708 1				
α	0.2	0.1	0.05	0.02	0.01
Critical Value	4.6416	6.2514	7.8147	9.8374	11.345
Reject?	No	No	No	No	No



Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
2	Gen. Pareto	0.10453	1	0.34043	1	0.4623	2
3	Log-Pearson 3	0.10939	2	0.3855	3	0.03677	1
1	Gen. Extreme Value	0.12103	3	0.38004	2	1.1424	4
7	Normal	0.13155	4	0.43107	4	1.3084	5
6	Lognormal (3P)	0.1317	5	0.45597	5	1.5331	7
5	Lognormal	0.13222	6	0.61613	7	0.51671	3
4	Logistic	0.14174	7	0.56433	6	1.3201	6



Goodness of Fit - Details [hide]

Gen. Pareto [#2]

Kolmogorov-Smirnov

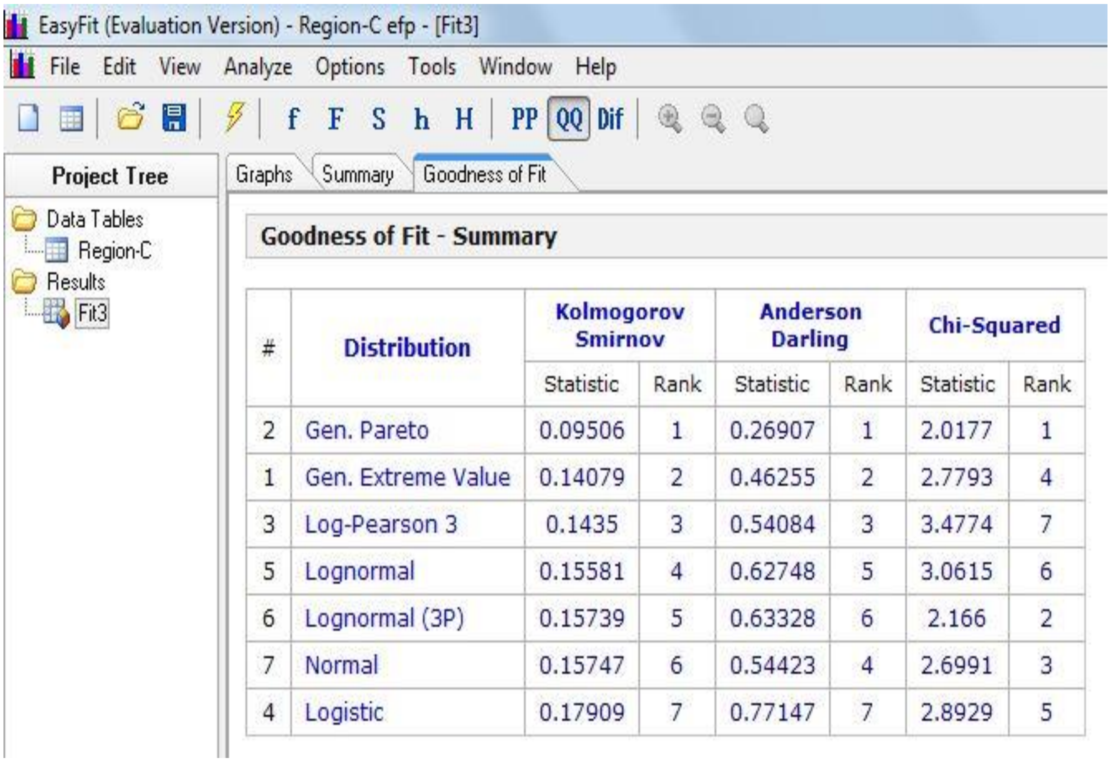
Sample Size Statistic P-Value Rank	22 0.10453 0.9495 1				
a	0.2	0.1	0.05	0.02	0.01
Critical Value	0.22115	0.25283	0.28087	0.31394	0.33666
Reject?	No	No	No	No	No

Anderson-Darling

Sample Size Statistic Rank	22 0.34043 1				
a	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	No	No	No	No	No

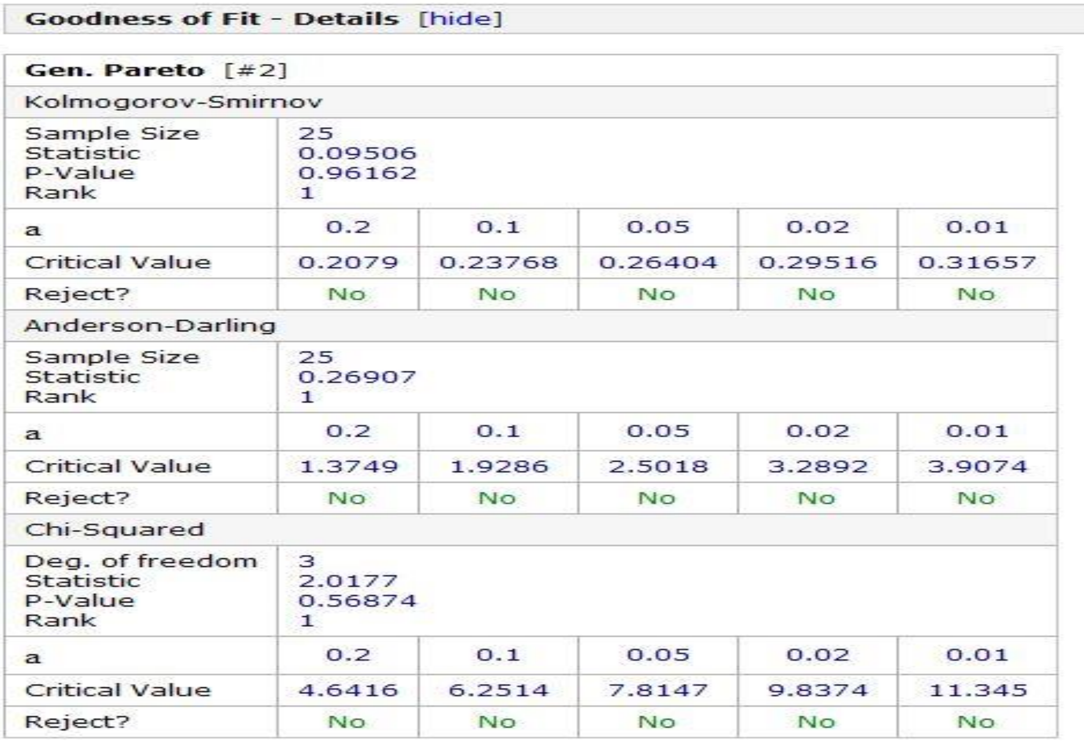
Chi-Squared

Deg. of freedom Statistic P-Value Rank	2 0.4623 0.79362 2				
a	0.2	0.1	0.05	0.02	0.01
Critical Value	3.2189	4.6052	5.9915	7.824	9.2103
Reject?	No	No	No	No	No



Goodness of Fit - Summary

#	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
2	Gen. Pareto	0.09506	1	0.26907	1	2.0177	1
1	Gen. Extreme Value	0.14079	2	0.46255	2	2.7793	4
3	Log-Pearson 3	0.1435	3	0.54084	3	3.4774	7
5	Lognormal	0.15581	4	0.62748	5	3.0615	6
6	Lognormal (3P)	0.15739	5	0.63328	6	2.166	2
7	Normal	0.15747	6	0.54423	4	2.6991	3
4	Logistic	0.17909	7	0.77147	7	2.8929	5



Goodness of Fit - Details [hide]

Gen. Pareto [#2]

Kolmogorov-Smirnov

Sample Size Statistic P-Value Rank	25 0.09506 0.96162 1				
a	0.2	0.1	0.05	0.02	0.01
Critical Value	0.2079	0.23768	0.26404	0.29516	0.31657
Reject?	No	No	No	No	No

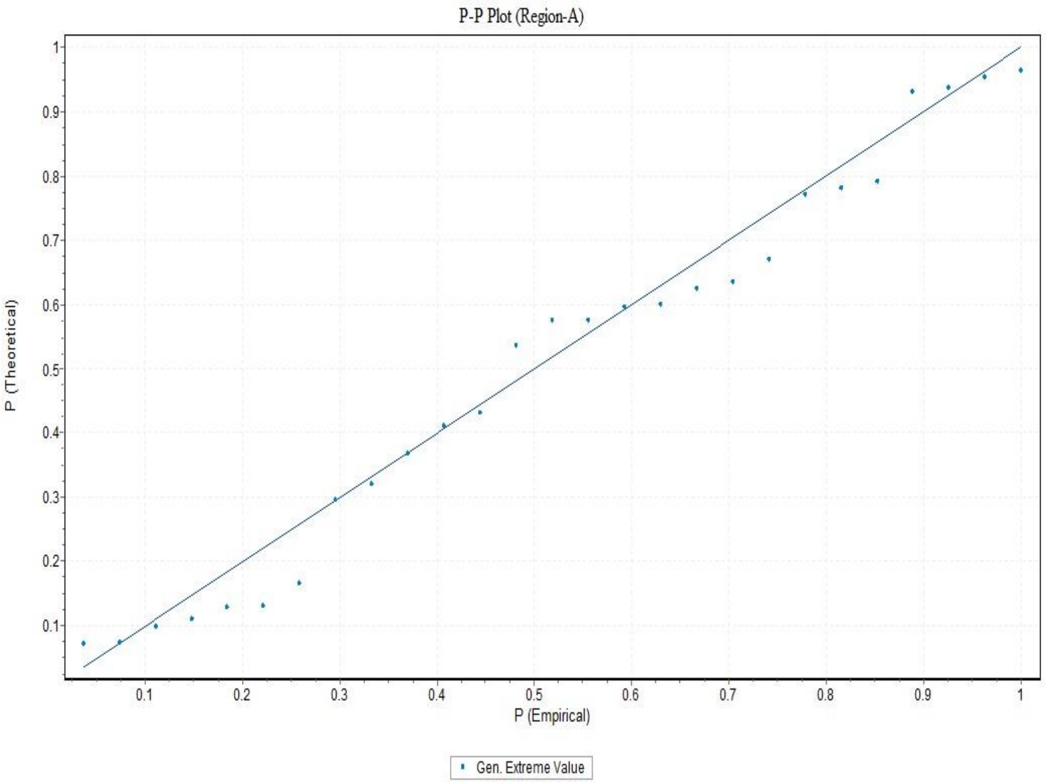
Anderson-Darling

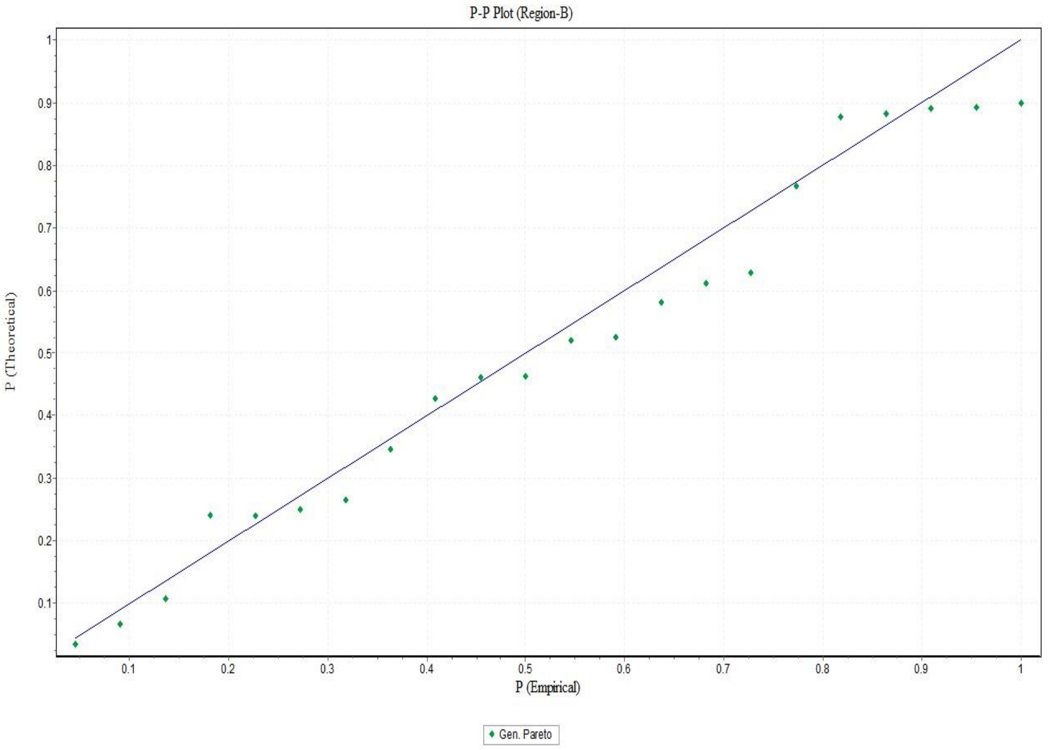
Sample Size Statistic Rank	25 0.26907 1				
a	0.2	0.1	0.05	0.02	0.01
Critical Value	1.3749	1.9286	2.5018	3.2892	3.9074
Reject?	No	No	No	No	No

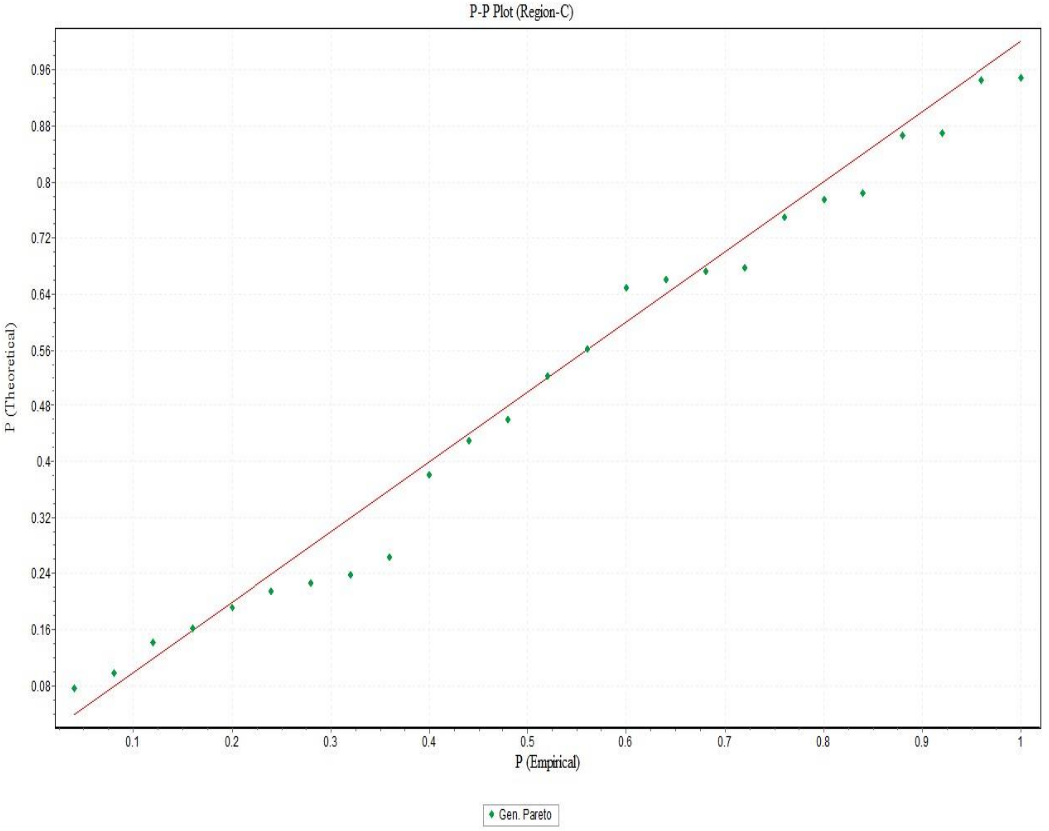
Chi-Squared

Deg. of freedom Statistic P-Value Rank	3 2.0177 0.56874 1				
a	0.2	0.1	0.05	0.02	0.01
Critical Value	4.6416	6.2514	7.8147	9.8374	11.345
Reject?	No	No	No	No	No

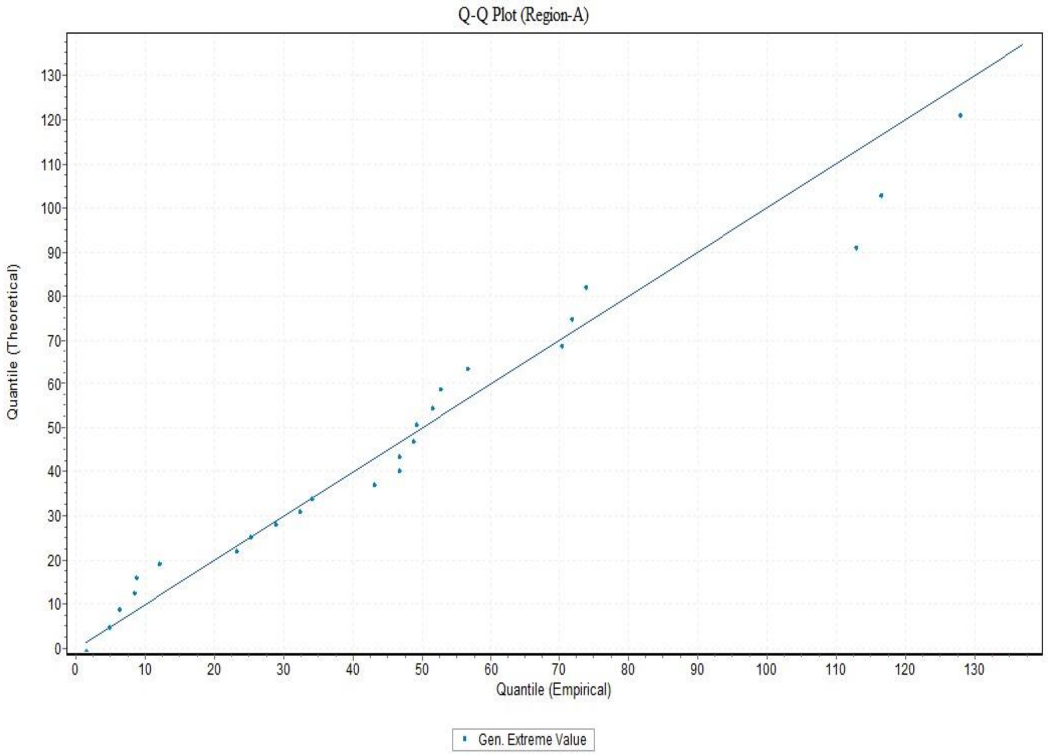
Appendix-I: Probability-probability plots of stations

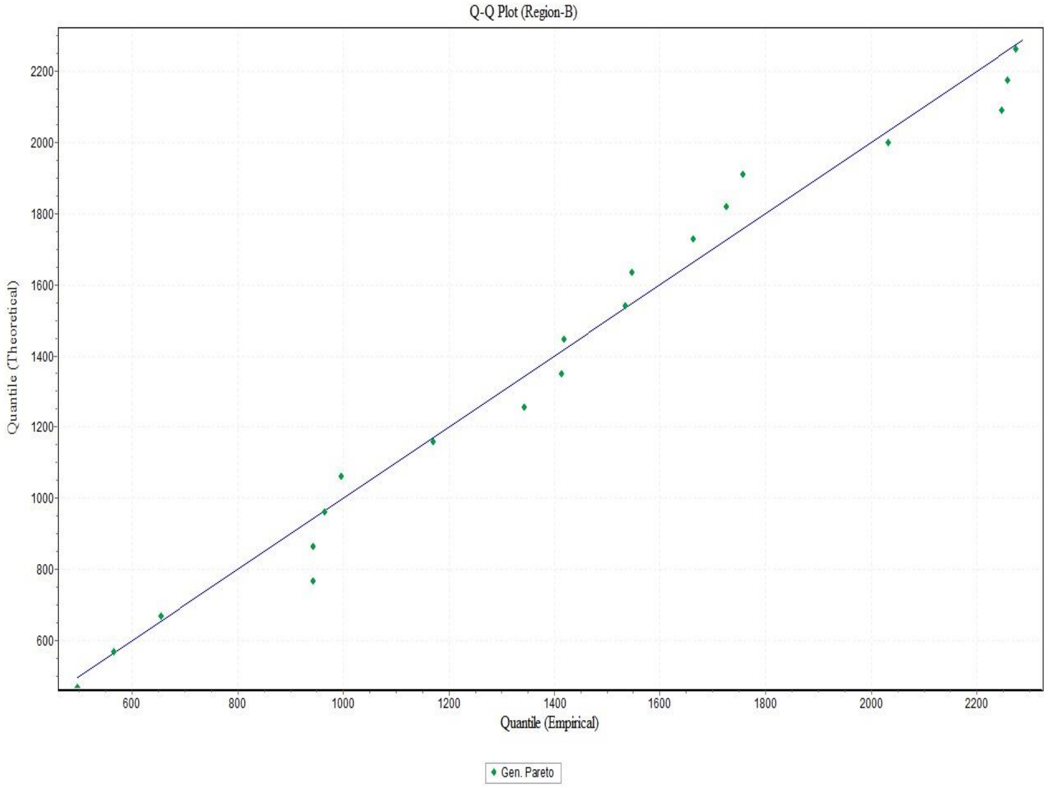


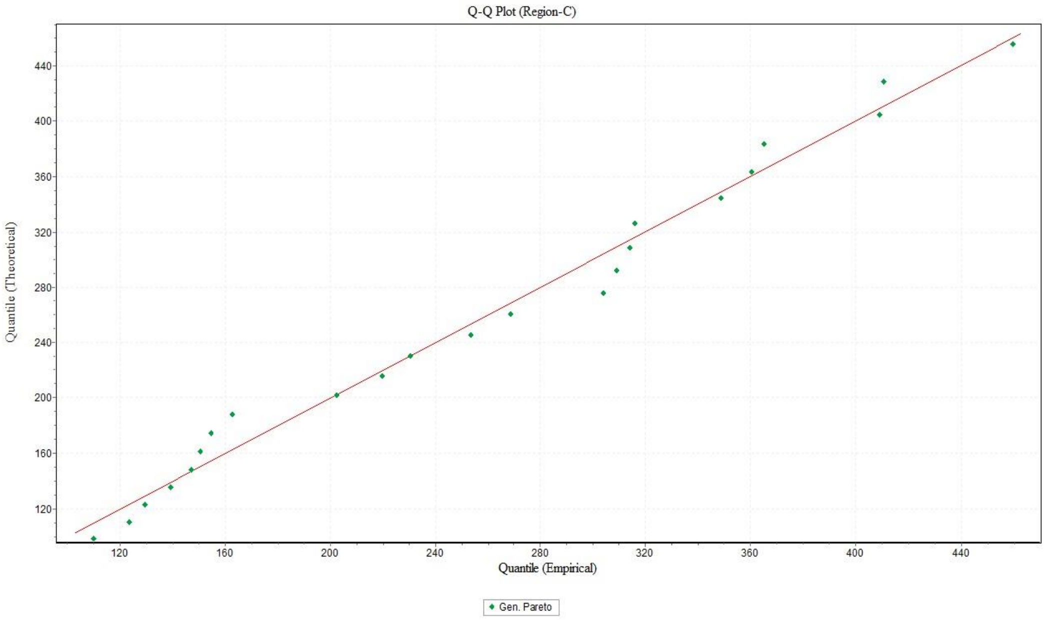

P-P Plot (Region-B)
P (Theoretical)
P (Empirical)
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Gen. Pareto


P-P Plot (Region-C)
P (Theoretical)
P (Empirical)
0.08
0.16
0.24
0.32
0.4
0.48
0.56
0.64
0.72
0.8
0.88
0.96
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Gen. Pareto

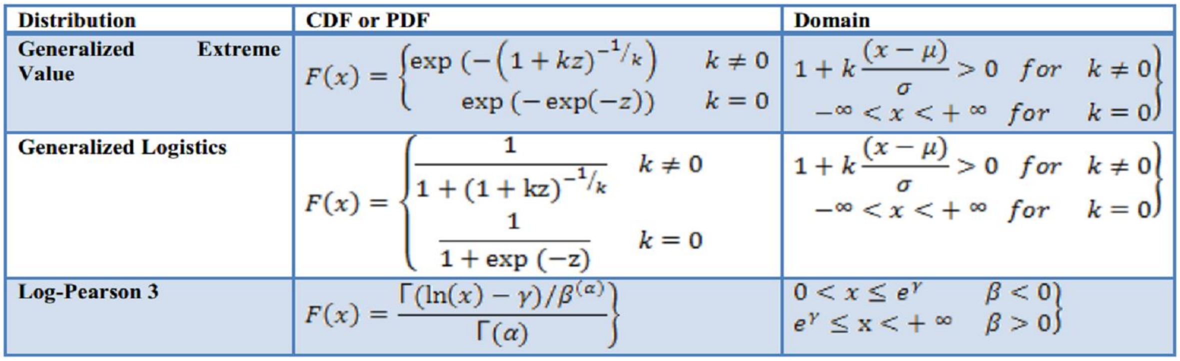
Appendix-J: Quantile-Quantile plots of stations



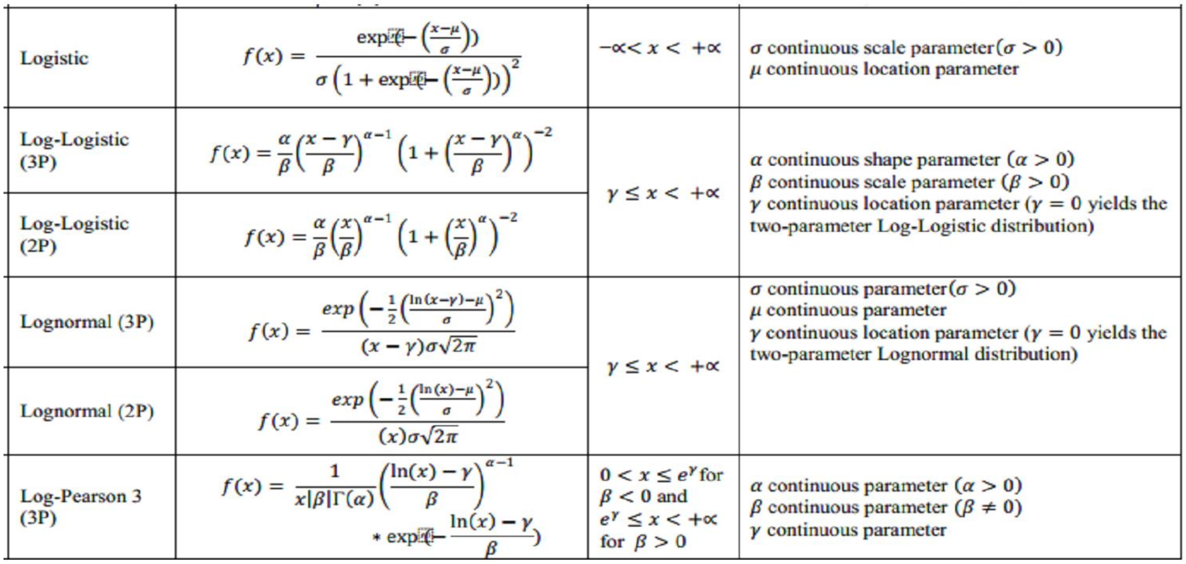




Appendix-K: Probability Density functions for selected distributions (Chow, 1964)



Distribution	**CDF or PDF**	**Domain**
Generalized Extreme Value	$F(x) = \begin{cases} \exp\left(-\left(1+kz\right)^{-1/k}\right) & k \neq 0 \\ \exp\left(-\exp(-z)\right) & k = 0 \end{cases}$	$\left. \begin{matrix} 1 + k\dfrac{(x-\mu)}{\sigma} > 0 & for & k \neq 0 \\ -\infty < x < +\infty & for & k = 0 \end{matrix} \right\}$
Generalized Logistics	$F(x) = \begin{cases} \dfrac{1}{1+(1+\mathrm{kz})^{-1/k}} & k \neq 0 \\ \dfrac{1}{1+\exp(-\mathrm{z})} & k = 0 \end{cases}$	$\left. \begin{matrix} 1 + k\dfrac{(x-\mu)}{\sigma} > 0 & for & k \neq 0 \\ -\infty < x < +\infty & for & k = 0 \end{matrix} \right\}$
Log-Pearson 3	$\left. F(x) = \dfrac{\Gamma(\ln(x)-\gamma)/\beta^{(\alpha)}}{\Gamma(\alpha)} \right\}$	$\left. \begin{matrix} 0 < x \leq e^{\gamma} & \beta < 0 \\ e^{\gamma} \leq \mathrm{x} < +\infty & \beta > 0 \end{matrix} \right\}$



Logistic	$f(x) = \dfrac{\exp(-\left(\frac{x-\mu}{\sigma}\right))}{\sigma\left(1+\exp(-\left(\frac{x-\mu}{\sigma}\right))\right)^2}$	$-\propto < x < +\propto$	σ continuous scale parameter($\sigma > 0$) μ continuous location parameter
Log-Logistic (3P)	$f(x) = \dfrac{\alpha}{\beta}\left(\dfrac{x-\gamma}{\beta}\right)^{\alpha-1}\left(1+\left(\dfrac{x-\gamma}{\beta}\right)^{\alpha}\right)^{-2}$	$\gamma \leq x < +\propto$	α continuous shape parameter ($\alpha > 0$) β continuous scale parameter ($\beta > 0$) γ continuous location parameter ($\gamma = 0$ yields the two-parameter Log-Logistic distribution)
Log-Logistic (2P)	$f(x) = \dfrac{\alpha}{\beta}\left(\dfrac{x}{\beta}\right)^{\alpha-1}\left(1+\left(\dfrac{x}{\beta}\right)^{\alpha}\right)^{-2}$		
Lognormal (3P)	$f(x) = \dfrac{exp\left(-\frac{1}{2}\left(\frac{\ln(x-\gamma)-\mu}{\sigma}\right)^2\right)}{(x-\gamma)\sigma\sqrt{2\pi}}$	$\gamma \leq x < +\propto$	σ continuous parameter($\sigma > 0$) μ continuous parameter γ continuous location parameter ($\gamma = 0$ yields the two-parameter Lognormal distribution)
Lognormal (2P)	$f(x) = \dfrac{exp\left(-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2\right)}{(x)\sigma\sqrt{2\pi}}$		
Log-Pearson 3 (3P)	$f(x) = \dfrac{1}{x\lvert\beta\rvert\Gamma(\alpha)}\left(\dfrac{\ln(x)-\gamma}{\beta}\right)^{\alpha-1} * \exp(-\dfrac{\ln(x)-\gamma}{\beta})$	$0 < x \leq e^{\gamma}$ for $\beta < 0$ and $e^{\gamma} \leq x < +\propto$ for $\beta > 0$	α continuous parameter ($\alpha > 0$) β continuous parameter ($\beta \neq 0$) γ continuous parameter