JIMMA UNIVERSITY
SCHOOL OF GRADUATE STUDIES
JIMMA INSTITUTE OF TECHNOLOGY
FACULTY OF CIVIL AND ENVIRONMENTAL ENGINEERING

LOW FLOW ANALYSIS AND REGIONALIZATION FOR GHEBA RIVER, 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

ADHANOM GEBREMICHAEL BERHE

FEBRUARY, 2019

JIMMA, ETHIOPIA

JIMMA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

JIMMA INSTITUTE OF TECHNOLOGY

FACULTY OF CIVIL AND ENVIRONMENTAL ENGINEERING

LOW FLOW ANALYSIS AND REGIONALIZATION FOR GHEBA RIVER, 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

ADHANOM GEBREMICHAEL BERHE

ADVISOR: Dr.-Ing. Tamene Adugna (PhD)

CO-ADVISOR: Mr. Berhan Tekuame (MSc)

FEBRUARY, 2019

JIMMA, ETHIOPIA

DECLARATION

This thesis entitled **"Low flow analysis and regionalization for Gheba River, Ethiopia"** has been approved by the following advisor(s), and department for the commencement of the required research in Hydraulic Engineering.

……………………… …………….. ………………..

Candidate Signature Date

CERTIFICATION

I, the undersigned, certify that I read and hereby recommend for the acceptance by the Jimma University a research entitled: -**Low flow analysis and regionalization for Gheba River, Ethiopia"** In partial fulfillment of a degree of Masters of Science in Hydraulic Engineering.

Dr.-Ing. Tamene Adugna	----------------------	-------------------
(Advisor)	Signature	Date
Mr. Berhan Tekuame	----------------------	-------------------
(Co-Advisor)	Signature	Date

APPROVAL PAGE

A thesis submitted to the School of Graduate Studies of Jimma University in Partial fulfillment of the requirements for the Degree of Masters of Science in Hydraulic Engineering.

BY

Adhanom Gebremichael Berhe

APPROVED BY BOARD OF EXAMINERS

Dr.-Ing. Tamene Adugna	---------------------	----------------------
(Advisor)	Signature	Date
Mr. Berhan Tekuame	-------------------	---------------------
(Co-Advisor)	Signature	Date
Dr.-	-------------------	-------------------
(External Examiner)	Signature	Date
Mr.	---------------------	--------------------
(Internal Examiner)	Signature	Date
Mr.	----------------------	--------------------
(Chairman)	Signature	Date

DEDICATION

I dedicate this research to my mother she has paved the road and eliminated the darkness of my life.

ABSTRACT

Analyze the low flow in Gheba river and identify their low flow characteristic for better management and operation of reservoir & water resource as general were estimated by low flow frequency analysis and flow duration curve approaches to the selected 7 days annual minimum flow observed at four stations with an average of 17 years record length. Goodness-of-fit tests, including L-moment methods for regional data, were used to determine the best distributions. Gheba river watershed is the sub-basin of the Tekeze river basin and situated in the northern part of Ethiopia, Tigray Regional State. The total drainage area of Gheba up to the junction to Tekeze is about 5163 km^2 and it is characterized by a minimum elevation of 1311 m, with a maximum of 3065m a.s.l. The main problem in the Gheba river basin is that the watershed is ungauged and also water resources in the basin are scarce due to continuously climate change and deforestation. The aim of this study was to estimate the low flow quantile and the low flow frequency growth curve using the statistical parameters were computed using EASY FIT computer program. Accordingly Gheba river sub basin has been divided in to two homogeneous regions. The region one of the basin includes the stations Gheba and Suluh and region two encloses the stations of Agulae and Ilala. The Generalized Logic distribution provides a good fit to low flows in the region one, while the generalized extreme value distribution fitted well in the region two. And the frequency curve for a return period of T was developed for both regions of low flow. On the basis of flow duration curve the flow duration curve indices for low flow of the basin Q60, Q70, Q80, Q90, Q95 and Q100, were quantified. For more certain on the results of this study, further studies are required on the points which are not assessed due to scarce of data, like effect of regionalization by topographic and climatic parameters and the ground water contributions of the regions.

Generally this study contains the information required for the development of water resource projects on the sub basin.

***Key words**: Low flow, L-moment, homogeneous regions, low flow frequency, regionalization and FDC*

ACKNOWLEDGEMENT

First of all I will like to thankful my almighty God in all aspects of my life he has presented what a priceless gift.

It is difficult in a work like this to make suitable acknowledgement to all the many people from whom I have learned. I hope silence will not be taken to imply lack of gratitude.

I would like to express my sincere gratitude to my advisor, Dr.-Ing. Tamene Adugna for his enthusiasm and dedication to producing the best product possible.

I also extend my special thank my Co-advisor Mr. Berhan Tekuame for making this work possible. I am very grateful for his guidance, support and for supplying me relevant materials to carry out the research.

I would like as well to acknowledge Jimma university/JU for the financial support made through the postgraduate program of the university, which is available to me during my research time.

I wish to extend my appreciation to all staff members of water resource and environmental engineering for their cooperation in all aspects.

I express my gratitude to all staff members of ministry of water resource, particularly hydrology department for their appreciable support in providing me hydrological data and other reference materials.

Finally I take this opportunity to deliver my heartfelt acknowledgement to my family for their well willingness and support to my success, and also I need to thank all my friends who helped me in different ways to accomplish my research.

Table of Contents

Contents

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AMS	Annual Maximum / Minimum Series
CC	Combined Coefficient
CDF	Cumulative Distribution Function
CV	Coefficient of variance
DEM	Digital Elevation Model
DMC	Double Mass Curve
ECDF	Empirical Cumulative Distribution Function
FDC	Flow Duration Curve
GIS	Geographical Information System
L-Cs	Linear Coefficient of skewness
L-Ck	Linear Coefficient of kurtosis
LCL	Lower confidence level
LFFA	Low flow Frequency Analysis
LMRD	L- Moment Ratio Diagram
ML	Maximum Likelihood
MOM	Method of Moment
PWM	Probability Weighted Moment
UCL	Upper confidence level
WMO	World Metrological Organization

1. INTRODUCTION

1.1 Background

Climatic change, population growth and the related development in domestic, industrial and agricultural use of water have sited an increasing demand on water resources throughout the world.

Hydrological analysis is an essential prerequisite for any project involving the implementation of works in a river or stream. The hydrology of the catchment defines boundary conditions which are the source of risks such as low flow and flood which affect the functions of river. Low flow can be the flow in a river during dry periods of the year (Tennant, 1976).

International glossary of hydrology (WMO, 2008) defines low flow as "flow of water in a stream during prolonged dry weather". This definition does not make a clear distinction between low flows and droughts. Low flows are a seasonal phenomenon and an integral component of a flow regime of any river, where as drought on the other hand is a natural event resulting from a less than the normal precipitation for an extended period of time.

Low flow hydrological regimes are crucial for an efficient development of water resources management tools, especially in those areas stressed by the combination of a dry climate and an excessive water demand (Dahmani, 2009).

There are three scenarios when low flow information is required. The first is when a water resource projects being developed. The second is during the operational stage of water resource projects, when it is constructed and includes decision on how to manage the project on day to day basis. The third scenario is when it is necessary to make operational decisions to day based on estimates of future river flows which look days, weeks or sometimes months ahead. These forecasts can increase the efficiency of water use and the economic importance in terms of reducing the operational costs of water resource projects (WMO, 2008).

Low flow characteristics are best estimated from observed stream flow data but for sites where these data are unavailable hydrological regionalization techniques can be used to gather them from other catchments where stream flow data have been collected (Tharme, 2003).

The estimation of flow regimes at ungauged sites may be achieved by transfer of statistics derived from gauged catchments using regionalization procedures. The term regionalization in hydrology refers to grouping catchments in to homogenous regions (Stall, 1971).

Low flow information can be quantified in a variety of ways depending on the type of data available and the output information desired. Low flow frequency analysis and flow duration curve are the techniques use to quantify the low flow (Chang, 1977).

Low flow frequency analysis is a stochastic approach for characterizing low flow events. And it shows the percentage of a particular time that a flow is exceeded during a given interval of time on the average (Tharme, 2003).

A flow duration curve is one of the most use full ways of displaying the complete range of river discharge from low to floods. It gives a relationship between a discharge value and percentage of time that the discharge is equaled or exceeded (Foster, 1934).

1.2 Statement of the problem

Due to increased food insecurity and recent food price hikes in the country, it is envisaged that there would be huge demand for irrigated agriculture by the farmers themselves and as a matter of policy priority by the government at all level. This will undoubtedly create huge demand on the water resources particularly during the lean flow season. The extra demand on the river may create undesirable environmental as well as upstream downstream conflict. Therefore, for informed decision making and appropriate policy recommendations on the minimum environmental and downstream requirements, estimation of low flow is required that enables generation of flows at any required reach of the river system, low flow estimates are vital for planning water supplies, water quality management issuing and renewing of waste disposal permits, hydropower, and the impact prolonged drought on aquatic ecosystems. Now a day's extreme low flow events are more diligently analyzed in the emerging field of Eco-Hydrology. Estimation of low flow is also important for small scale irrigation projects that contribute significantly to poverty alleviation by means of increased crop production and generation of rural employment (Abebe, 2003). The main problem in the Gheba river basin is that the watershed is ungauged and also water resources in the basin are scarce due to continuously climate change and deforestation (Behailu, 2009). So analysis of low flows would provide an accurate understanding of the demand that may safely be placed on a stream flow.

Regional low flow frequency analysis is an important subject because, most of hydraulic structures are constructed in areas where recorded flow data either missing or inadequate.

Moreover, most of hydrological regions have scarce of observation sites with short record length of observed flow data that makes the use of single site analysis to estimate design parameters at many potential project sites unreliable.

Generally this study tries to analysis the low flow of Gheba river sub basin and provides the necessary information concerned on the low flow of the basin.

1.3 Objective of the study

1.3.1 General objective

The main objective of this study is to analyze the low flow in Gheba river and identify their low flow characteristic for better management and operation of reservoir & water resource as general.

1.3.2 Specific objectives

The specific objectives of this study are;

i. To identify and delineate the basin in to hydrological homogeneous low flow regions.
ii. To determine low flow quantiles and regional frequency curve.
iii. To develop and analysis the flow duration indices of the basin.

1.4 Research question

i. Why low flow statistics of Gheba River are needed to study?
ii. What is the purpose of selecting the best statistical parent distribution and parameters?
iii. The need of analysis of flow duration curve?

1.5 Significant of the study

Most of the time in the design of hydraulic structures does not consider the low flow of a stream. This study has a general theme of providing information on low flow characteristics, for the purpose of water resource development projects in Gheba river sub basin. To this concern analysis of flow duration curve and low flow frequency analysis has been carried out.

The major objective of low flow estimation is to evaluate the consequence of flows in conjunction with future water use demand, the development of restriction for public water supplies, reservoir operation, irrigation and hydropower scheduling.

This study is useful for designer and operation policy makers to make appropriate measures or decisions on water resource management.

1.6 Scope of the study

This study summarizes streamflow data collected in the Gheba River sub basin, including data from five gaging stations funded by the Ministry of Water Resources Irrigation and Electricity of the GIS Department, since the stations of the basins are very minimum the EASY FIT COMPUTER software were used to fit the selected distributions for the purpose of quantile estimation at site and developing regional regression equations to estimate low flow frequency statistics at ungauged sites. Low-flow frequency statistics, such as theMAM7 (mean annual minimum 7 day) were calculated for 17year of the streamflow-gaging stations in the basin. Sub basin drainage areas were then delineated using catchment physiographic and climatic parameters obtained from maps and hydro meteorological data (as 7 days mean annual minimum discharge) were determined using Geographic Information System (GIS) methods.

2. LITERATURE REVIEW

2.1 Low flow analysis

Characterization of the magnitude, frequency and duration of low stream flows and droughts is essential for assessing the consistency of flows for all in-stream and withdrawal uses and for defining resource shortages and drought. Hydropower capacity design is highly dependent on the low flow of the river. Since, the water used is directly returned back to the river, the impact on the quality and quantity of the flow is minimum, even though, the release might be regulated (Tallaksen, 2004). Some of the impacts of human activities on the low flow generating processes are; construction of dams and regulation of a river flow regime, irrigation return flows from agricultural fields, industrial and municipal waste disposed directly to the river and river abstraction for industrial, agricultural or municipal purposes.

2.2 Methods of low flow analysis

2.2.1 Low flow frequency analysis

Frequency analysis is one of the most common and earliest applications of statistics with in hydrology. And it assess ; definition of hydrological events and extreme characteristics to be studied, selection of the extreme events and probability distribution to describe the observed data, estimation of the parameters of the distribution and estimation of extreme events or design values for a given problem (Tallaksen, 2004). Low flow frequency analysis is used to evaluate the ability of stream to meet specified flow requirement at a particular location. The analysis can indicate the amount of storage that would be required to meet a given demand again with a stated probability of being shortage. The design of hydroelectric power plants, determination of minimum flow requirement for different cases and design of water storage projects can benefits from low flow frequency analysis (Hickox, 1933). Low flow frequency curves are a graphical means of understanding the characteristics (frequency, duration and magnitude) of low flow events. Low flow frequency curve is normally constructed on the basis of a series of annual flow minimum (daily or monthly minimum discharges or flow volumes), which are extracted from the available original continuous flow series (one value from every of record) (Poff, 1989).

2.2.2 Flow duration curve

The flow duration curve is a graph of river discharge plotted against exceedance frequency and is normally derived from complete time series of recorded river flows. It is simple to construct and used in many different water resource applications over the entire range of river flows. Flow duration curves (FDC) give a relationship between magnitude and frequency of stream flow discharges. The construction of FDC is based on ranking the data (normally daily discharge) and calculating the frequency of exceedance for each value.

It effectively re-orders the observed hydrograph from one ordered by time to one ordered by magnitude. The percentage of time that only particular discharge is exceeded can be estimated from the plot (Smakhtin, 2001). To construct the FDC the data values are first ordered by size, the largest value is given rank 1, the second largest a rank 2 and so on until the lowest has a rank equal to N, which is the total number of data points. If the two values are equal, they should be assigned different ranks. A plotting position is assigned to each data using a plotting formula such as Weibull, California, Hazen, Chegodayev and Gringorten. A flow duration curve is one of the most informative methods of displaying the complete range of river discharges from low flows to flood events. It is a relationship between any given discharge value and the percentage of time that this discharge is equaled or discharge is exceeded, or in other words, the relationship between magnitude and frequency of stream flow discharge (calculation of the cumulative frequency of streamflow discharge) (Searcy, 1959).

2.3 Probability distribution

The primary objective of frequency analysis is to relate the magnitude of extreme events to their frequency of occurrence through the use of probability distributions (Chow, 1988).

Probability distribution is a function representing the probability of occurrence of random variable which satisfies the properties of probability distribution function. Some of the probability distributions commonly applied in hydrological analysis are; normal family (normal distribution, log normal distribution and log normal type III distribution), generalized extreme value family (extreme value type I (Gumbel), extreme value type II and extreme value type III (Weibull)), and exponential or Pearson type family (exponential distribution, Pearson type III distribution and log pearson type III distribution).

2.4 Fitting data to probability distribution

Fitting a probability distribution means estimating the parameters of the distribution. By fitting a distribution to a set of hydrologic data, a great deal of the probabilistic information in the sample can be completely summarized in the function and its associated parameters.

The two methods used for fitting a probability distribution are; method of moment and method of maximum likelihood. The procedure includes trying to fit several theoretical distributions to the observed low flow data and selecting an appropriate distribution by using statistical tests.

For low flows, the recommended distributions include Weibull, Gumbel, Pearson Type-III and lognormal distribution.

Many studies have attempted to ascertain suitable distributions for annual minima and those occurring at different averaging intervals. Despite several attempts, no fixed probability distribution for low flows has been agreed on. One of the crucial issues that most of low flow frequency analysis and distribution fitting studies confront is the occurrence of zero values. Hydrological data sets for example, stream flow and precipitation, often have zero as a lower limit. Ignoring zero values may lead to an unreliable estimation of the concerned variable. However, distributions fit to zero values assign positive probabilities to negative values of the variable. In such cases, the distributions can be restricted to have a lower limit, which may give physically meaningless results along with challenging the flexibility of the distribution (Smakhtin, 2001).

In another approach, (Haan, 1977) used a conditional probability approach to account for zero values of low flows. Using the theorem of total probability, for non-negative values of flows, denoted by x, this approach can be expressed as: $p(x=0)+p(x>0)=1$ (if all $x \geq 0$) Proportion of values equating to zero are accounted for by primarily analyzing all non-zero observations and then multiplying the resulting probabilities by the fraction of non-zero values in the data. That is; $P(x) = cp'(x)$ Where p(x) is the probability of exceedance of all values, c is the probability that x is not zero and $p'(x)$ is the probability of exceedance for the non-zero values.

Probability distributions are characterized by their parameters. To fit a distribution to a dataset, true parameter values of the same must be ascertained using the sample data series. There are three dominant parameter estimation techniques exist: Method of maximum likelihood, Method of Moments and probability weighted moments.

2.4.1 Method of moment (mom)

This method is one of the most commonly and simply used method for estimating the parameters of a probability distribution. However, it is well known that some moment estimates are biased, especially for the cases of short sample series or higher order moments (Greenwood, 1979).

It is confirmed theoretically that coefficient of variation (CV) and skewness (Cs) in random samples are bounded and that the bounds are a function of sample size. Much research work has been done in deriving factors or functions by which the bias can be corrected.

2.4.2 Maximum likelihood (ml)

In this method the parameters are determined by maximizing the sample log likelihood function. The unknown parameters may be obtained by setting each of the partial derivatives with respect to each parameter equal to zero and solving the resulting equations simultaneously. These equations unfortunately do not often take a simple closed form and the numerical solutions have to be used. Even though the maximum likelihood estimator is asymptotically unbiased, sufficient and consistent, sometimes there is a failure to obtain proper solutions due to the complexity of log likelihood functions, particularly when sample size is small or when the distribution has more than two parameters. This method is statistically the most efficient one in large samples.

2.4.3 Probability weighted moments

Probability weighted moments (PWM) were introduced by (Greenwood, 1979) and further analyzed by (Hosking, 1986). PWMs were proposed for the purpose of derivation of expressions for the parameter of distributions whose inverse forms x = x (F) can be explicitly defined, such as Extreme Value type 1, General Extreme Value, Log Logistic and Wakeby distributions. The use of PWMs was later extended to be used to estimate P3 parameters and other distributions whose inverse forms cannot be explicitly defined Hosking. The PWM method has recently come to be regarded as the best method by research hydrologists Hosking. This procedure leads to relatively unbiased quantile estimates due to the fact that the PWMs being linear functions of the data, suffer less from the effects of sampling variability than do the conventional moments, because the latter involve higher powers of the data.

2.5 Selection of distribution function

The procedures that have been used to assess the goodness of fit of distribution for observed data series contains different ways of testing like, special statistical tests. Statistical goodness of fit tests such as the chi squared test and Kolmogorov Smirnov tests are designed to check the best distribution. The chi squared statistical test is based on a comparison of the number of observed and expected events in each class interval of the sample histogram. And the Kolmogorov Smirnov test is based on the absolute deviation between the empirical and theoretical cumulative distribution function.

The conclusion that can be drawn from these tests depends on the level of confidence adopted commonly at 95% level is chosen (Silesh, 2005).

2.6 L-moments and frequency analysis

In the last century, probably one of the most significant scientific contributions to statistical hydrology is, the L-moments (Hosking, 1990). L-moments have the advantage of providing parameter estimates that are nearly unbiased and highly efficient. They are not much influenced by outliers (extreme observations) in the data because they are linear combinations of the observed data values, unlike the ordinary statistical moments that need the data to be squared, cubed in their computations (Stedinger, 1993). Although they lead to quite efficient estimates of the parameters of a distribution, this may not be used for the estimates of extreme quantiles in some cases. Because very large or very small sample values are given little weight in the estimation. Several researchers illustrated that compared to the product moment ratio diagram, the L-moment ratio diagrams posse a better ability to discriminate between distributions.

2.7 Regionalization

The availability of data is an important aspect in frequency analysis. The estimation of probability of occurrence of extreme floods/low flow is an extrapolation based on limited data. Thus the larger the database, the more accurate the estimates will be. From a statistical point of view, estimation from small samples may give unreasonable or physically unrealistic parameter estimates, especially for distributions with a large number of parameters (three or more). Large variations associated with small sample sizes cause the estimates to be unrealistic. In practice, however, data may be limited or in some cases may not be available for a site. In such cases, regional analysis is most useful. Regional analysis is based on the concept of regional homogeneity which assumes that annual minimum flow populations at several sites in a region are similar in statistical characteristics and are not dependent on catchment size (Cunnane, 1989). Although this assumption may not be strictly valid, it is convenient and effective.

2.8 Delineation of hydrological homogeneous regions

An important aspect in regional frequency analysis concerns the delineation of homogeneous regions. The sites should be grouped to satisfy the homogeneity assumption of the index method that is sites within the group should follow the same distribution. In practice, the homogeneity of regions formed using a regionalization approach is tested by using flow statistics. Hence, these statistics are not supposed to be used as attributes to form regions. The regionalization of stream flow characteristics in general is based on the evidence that catchments with similar climate, geology, topography, vegetation and soils would normally have similar stream flow responses, for example, in terms of unit runoff from the catchment area, average monthly flow distribution, duration of certain flow periods, frequency and magnitude of high and low-flow events in similar sized catchments.

The delineation of regions may be accomplished using convenient boundaries based on geographic, administrative or physiographic considerations. The regions that result using such an approach may not always appear to be sufficiently homogeneous. However, this pragmatic approach may appear to be suitable in conditions of limited data availability. Geographically nearby regions may be established on the basis of residuals from a regional regression model developed to estimate flow characteristics at ungauged catchments. A homogeneous region may be viewed as a collection of catchments, which are similar in terms of catchment hydrological response, but not necessarily geographically contiguous. Development of a separate regression model for each of the identified homogeneous regions is likely to improve the predictive ability of the final prediction equation. Classification of catchments into groups may be based on standardized flow characteristics estimated from the available observed or simulated stream flow records (Wiltshire, 1986).

Alternatively the regions are delineated using catchment physiographic and climatic parameters obtained from maps and hydro meteorological data (rainfall, evaporation) (Mkhandi, 2000). This last approach has more relevance to the estimation of flow characteristics at ungauged sites since an ungauged catchment, for which, for example, low flow estimation is undertaken, should first be assigned to one of the identified regions and this can only be done on the basis of catchment physiographic information. A drawback in using flow statistics as attributes to form regions is that the resulting regions may appear homogeneous but are not necessarily effective for frequency analysis studies.

2.9 Low flow estimation for ungauged catchments

For many practical problems of surface water hydrology, such as water resources management or Planning and designing hydraulic structures, reliable estimates of low flow statistics (e.g. mean annual low flow and low flow quantiles) are needed. However, historical data, that are needed to estimate these statistics, are not always available at the site of interest or available data may not be representative of the basin flow because of the changes in the watershed characteristics. The spatial variations in low flow statistics are closely related with the variations in regional physiographic, geologic and climatic factors. Making use of this observation, regional regression models are often used to make estimates of low flow statistics for ungauged sites (Sivapalan, 2003).

2.10 Previous studies in the country and global experiences

The case which is entitled the low flow analysis and regionalization were not studied in the Gheba river basin as the desired amount. Concerning the low flow analysis and regionalization most possibly studies made in Ethiopia is in Blue Nile river basin by (Tegenu, 2007), for the master thesis submitted to Arbaminch University. This study was provided the low flow analysis and regionalization for 38 stations with an average of 23 years record length in the basin. And this study applied three methods to accomplish the study i.e. low flow frequency analysis, flow duration curve and base flow index. The overall methodology and results obtained from this study are; the GEV and Wakeby distributions were used for the identified regions and the PWM method of parameter is used to estimate the parameters of the selected distributions and to fit the selected distributions, the MATLAB software program were used. On the basis of FDC the 1, 7, 10 and 30 day FDCs were developed for the 38 stations of the basin. Finally the third approach which is the base flow index, were used here to know the sub surface contribution of the regions using the catchment characteristics.

In the same case but out of the country, the low flow analysis were study at three rivers in eastern Canada by (Deepti Joshi, 2013) and here the flow duration and the low flow frequency analysis were used as a methods of analysis. The HYFRAN- PLUS and MATLAB software's are used to fit the statistical distributions. On the basis of the FDC the 70-99% exceedance were used as FDC indices for the low flow of the rivers. Generally from the observed low flow data the 7 and 10 days annual minimum data are used in the Russian and Eastern Europe for analysis of low flow and the 1 and 30 day are widely used in US (Smakhtin, 2001). (Matalas, 1963) analyzed for 34 stations in the united states using the pearson III, the pearson V, the Gumbel type III, which is also known as the three parameter Weibull and the three parameter log normal distributions. He concluded that the Gumbel type III and Pearson type III distributions performed equally well and tended to outperform the other two distributions. (Condie, 1975) Performed a similar analysis of data from 38 Canadian rivers using the same distributions as (Matalas, 1963). (Zaidman, 2003) and others performed an analysis of 25 natural streams within the United Kingdom having more than 30 years of record. They derived data times series for durations of 1, 7, 30 days for each of the basins. In turn, four three parameter distributions, namely the generalized extreme value distribution, generalized logistic distribution, Pearson type III and generalized Pareto distribution were used. Starting from the above experiences this case this study followed the methods; low flow frequency analysis, flow duration curve and base flow index. But here the base flow index is not applied because it needs full information about the basin catchment characteristics but these raw data are not available and it is difficult to observe these data by single person. And the other thing which makes differ from the above researchers is the developing of FDCs for the basin; here the only 7 day FDCs for the stations were developed because the Gheba river contains the artificial components or structures which makes a variation on the flow, the 7 day period annual minimum flow reduces the day- to day variation of the basin. Since the stations of the basin are very minimum the simplest method of distribution fitting which is the EASY FIT COMPUTER software were used to fit the selected distributions.

3. MATERIALS AND METHODS

3.1 Study area description

3.1.1 Location and topography

Gheba river watershed is the sub-basin of the Tekeze river Basin and situated in the northern part of Ethiopia, Tigray Regional State. The Gheba sub basin lies between latitudes $13^0$17'46" and $14^0$15'00" N and longitudes $38^0$37'37"and $39^0$47'47''E. The total drainage area of the Gheba up to the junction to Tekeze is about 5163 km^2; the length of the main watercourse is 236.4km. The principal tributaries of this river are: Suluh, Genfel, Agulae, and Ilala. The elevation of Gheba watershed ranges between 3065m to 1311m a.s.l with a mean elevation of 2178.32m a.s.l (Setegn, 2015).

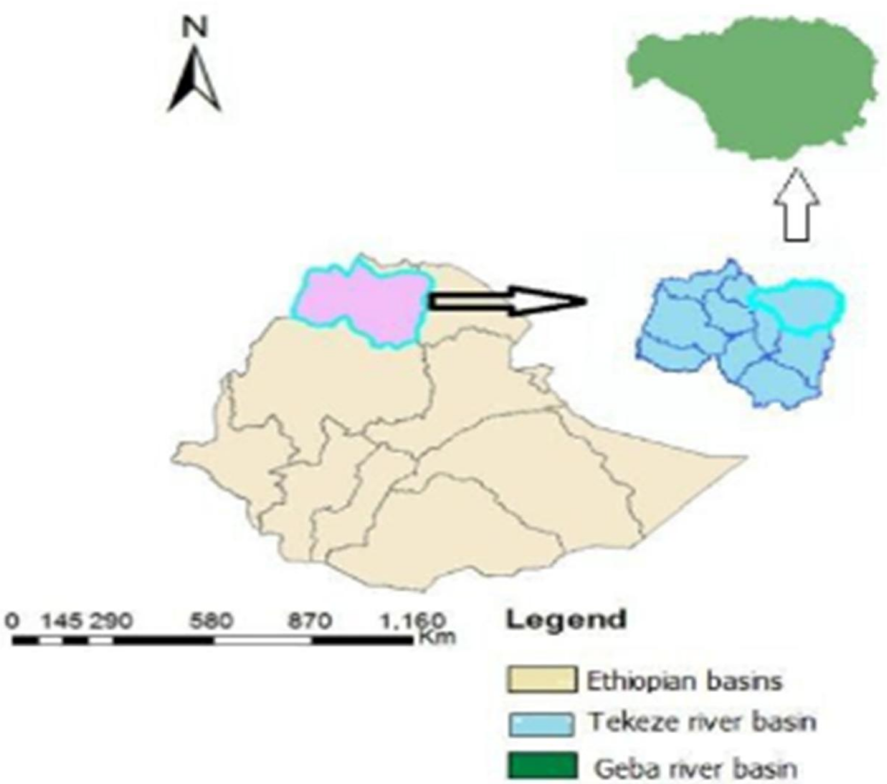


Figure 1 location of study area

3.1.2 Climate

According to the definition given by the world metrological organization (WMO), climate is defined as the synthesis of weather condition in a given area characterized by long term statistics (mean, variance, probability of extremes) of the metrological elements in an area. The WMO usually accepts 30 years of statistical data series to define climate. The climate of Tekeze basin in general comes under the influence of the Inter Tropical Convergence Zone (ITCZ) (NEDECO, 1997). The metrological elements include rainfall, temperature, wind, radiation, humidity and sunshine hour. The rainfall of the Gheba basin is highly seasonal with an annual rainfall ranging from over 500mm in the northern part to about 700mm in the central and west of the basin. The average over the whole basin is about 666mm, with more than 60% of the rainfall occurring between July and August, is unimodal pattern. The distribution of rainfall indicates that the rainfall is concentrated in the central part of the basin.

The annual temperature ranges from between 25°C in the area of lowlands to about 22°C on the high plateaus. The temperature of the coldest month average less than 6°C on the high plateau and reaches 11°C near the lowland area (Gonfa, 1996).

The minimum mean monthly relative humidity of the basin occurs in the dry periods (October – March). And the maximum values are occurring in the rainy season (July – August). The mean monthly sunshine hour for different stations in the basin lied between 6.5 and 8.5 hr. /day. The maximum monthly sunshine hour values occurred during dry period and the minimum value occurred during rainy season are varied up to 10hr/day and 4hr/day respectively. The mean monthly wind speed of the basin are up to 300km/day occur during the rainy season and the minimum values are occur during dry season sometimes below 100km/day (NEDECO, 1997).

3.2 Source and availability of data

In order to accomplish the goals of the study various data from different agencies and individuals were collected including time series data, topographical data and a digitized map of the study area.

3.2.1 Hydrological data

The most important data needed for this research is the long term average daily stream flow record.

The numbers of hydrological stations should include in the Gheba sub basin is 10, but 5 gauging stations with an average record length of 17 years were used for analysis.

The stations daily flow series were used to produce different durations of low flows and flow duration curves. Moreover, they were used for low flow frequency analysis for the whole region.

The summery of hydrological station and hydrological data used for this study are listed in the [appendix B & C].

3.2.2 Digitized map

A digitized contour map of the basin was collected from the Ministry of Water Resources Irrigation & Electricity of the GIS Department. These data were used as basic input for developing a digital elevation model (DEM), flow accumulation and flow direction. They were also employed in specifying the exact location of the gauging stations, all of which are necessary for regionalizing the basin into a homogeneous region.

3.3 Methodology

In general the study comprised the following methods:

• Collection of relevant data such as hydrological data, topographic and a digitized map of the basin

• Data analysis

• Application of low flow analysis techniques:

i. Low flow frequency analysis

ii. Flow duration curve

Low flow frequency analysis were computed by taking the statistical parameters of the selected stations, from the statistical parameters using the LMRD select frequency distribution for the delineated regions with in the basin for grouping the basin in to hydrological homogeneous region based on statistical value using the arc GIS and the regional homogeneity test describing the delineated region as region one and region two. Then finally obtain the quantile and deriving the frequency curves using the standardized flow data for all regions (Chow, 1988).

A flow duration curve of a stream is a plot of discharge against the percent of time flow was equaled or exceeded. And the tasks perform here are:

• Identification of FDC indices of the basin

• Grouping of the FDCs based upon the right-end tail portion of the curve

3.4 Hydrological data analysis

Annual minima data were derived from a daily flow series by selecting the lowest flow every year and the mean of the minima calculated. Minima of different durations were determined, with 1, 7, 10, 30 and 90 days being commonly used. The annual minima can be used to determine a distribution function for estimating the frequency or return period of low flows. The period or the calendar year is used in defining an annual minimum value. Only stations without long term zero data in their series are analyzed. The 7 day moving average annual minimum flow series are used in frequency analysis of this study, as it is usually necessary to nullify the effects of minor river regulations and an analysis based on a time series of 7 day average flows is less sensitive to measurement errors (Riggs, 1980). At the same time, in the majority of cases there are no big differences between 1 day and 7 day low flows (Singh, 1947). 7 day low flow data is defined as the smallest average of seven consecutive daily discharges (Hidley, 1973).

Before extracting annual minimum series from daily data, missing data were filled by using the following procedure:

1. Missing periods of consecutive days up to 30 days with in the record period are filled using interpolation and daily and seasonal mean values (annex A).

2. To fill missing stream flow records the regression correlation has been carried out between days and months of the record periods for those who have correlation coefficient (r) greater than or equal to 0.6 (annex A).

3.4.1Test for independency and stationery

Independent means that there should be no serial correlation or short-range dependence in the time series and the hydrologic regime remains stationary during the period of record or no long term trend in the time series.

So the following test was used to identify the serial correlation of the data:

i. W-W test (Wald-Wolfowitz test)
ii. Lag –one serial correlation coefficient test

Lag – one serial correlation coefficient test

The serial correlation coefficient can help to verify the independence of a time series.

If a time series is completely random, the population auto correlation function will be zero for all lags other than zero (when its value is unity all data sets are perfectly correlated with themselves).

For this purpose it sufficient to compute the lag-one serial correlation coefficient i.e. the correlation between adjacent observations in a time series. Lag- one serial correlation coefficient r1 defined as (Hall, 1990):

$$r_1 = \frac{\sum_{i=1}^{n-1}(x_i-\bar{x})*(x_{i+1}-\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \quad \text{……………………………………………………………….Equation 1}$$

After computing r1 check whether the value is fall between the upper confidence level and the lower confidence level at 5% of confidence level.

The upper and lower confidence level at 5% is calculated using:

$$UCL(r1) = \frac{(-1+1.96(N-2)^{0.5})}{N-1} \quad \text{......Equation 2}$$

$$LCL(r1) = \frac{(-1-1.96(N-2)^{0.5})}{N-1} \quad \text{......Equation 3}$$

Where: N is the number of sample size.

To accept the hypothesis the value of r1 must fall between UCL and LCL of the data set.

In any time series, data outliers may or may not exist. These outliers may be due to personal errors during recording and to the inadequacy of measuring devices as well as to extreme conditions of natural phenomena. Unless the source of the outliers is clearly identified, it is difficult to completely remove outliers from the analysis. Outliers can be excluded from the estimation procedure only if it is certain that annual minimum flows can be adequately modeled by a single distribution form (Cunnane, 1989). Thus, an outliers test was not done in this study; but to avoid the effect of outliers an efficient method of parameter estimation like PWM was used. Even if outliers are retained in the analysis, they have only a small effect if an efficient method of parameter estimation such as ML or PWM is used (Cunnane, 1989).

3.4.2Checking consistency of the data

Before using stream flow data it should be checked for consistency. If the conditions relevant to the recording of rain gauge station have undergone a significant change during the period of record, inconsistency would arise. Since the study area contains neighboring stations of ones another station in the basin the Double mass curve (DMC) is a desirable method for checking inconsistency of recorded data. The curve is a plot on arithmetic graph paper, of cumulative stream flow collected at a gauge where measurement condition may have changed significantly against the average of the cumulative stream flow for the same period of record collected at several gauges in the same region. A change in the proportionality between the measurements at the suspect station and those in the region is reflected in a change in the slope of the trend of the plotted points. If a double mass curve shows a change in slope significantly, the values of the earlier period should be adjusted to be consistent with the later period records. A correction factor K is applied on the records that needs adjustment at the place where the slope changes (Chow, 1988). And it is given by the following relationship:

$$K = \frac{slope\ for\ the\ period\ after\ slope\ change}{slope\ for\ the\ period\ before\ slope\ change} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots .. \textbf{Equation 4}$$

3.5 Regionalization

Regionalization refers to grouping of basins in to homogeneous regions. In other words regionalization means identification of homogeneous regions, which contain stations of similar flow producing characteristics. Several studies have shown that delineation of regions in the past has often relied on physiographic, political or administrative boundaries. The resulting regions were assumed to be homogeneous in terms of hydrologic response. This assumption actually is not true as it may have very different relief and stations with in the same geographical region. Before the low flow frequency procedure was applied homogeneous regions has been identified using statistical moments. Since the regionalization is done in the case of the hydrological response of the basin, the statistical moment method is the best option to identify homogeneous regions. Thus conventional moments, L-moments and L-moment ratios had to be computed. Final grouping of similar regions was validated using tests such as homogeneity tests (Stall, 1971).

3.5.1 Identification of homogeneous regions

The L-Cs and L-Ck moment ratio of standardized flow of the stations was plotted with the L-moment ratio diagram (LMRD) of various distribution functions. Furthermore, stations which lie in the same distribution were preliminary grouped as a homogeneous region. This was done on the origin of the fact that stations from the same homogenous region will be distributed along the same distribution of LMRD (Dalrymple, 1960).

In this study the statistical values of the basin are considered to group stations in to homogeneous region.

3.5.2 Delineation of homogenous regions

The performance of any regional estimation method strongly depends on the grouping of sites in to homogeneous regions. 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. The tool used for delineation of homogenous regions of the basin is the GIS software. All stations under analysis were identified according to their geographical location (latitude and longitude) on the digitized map of the basins. So from GIS screen the distance between one station and its neighboring station was determined and (L-Cs, L-Ck) values where interpolated to fix the boundary between two stations of different regions (Tucci, 1995).

The procedures followed in the delineation of the area boundary are as follows:

For delineating the area boundary compute the (L-Cs, L-Ck) values of each station and identifying the location of each stations along the distributions LMRD based on the skewness and kurtosis, the values more close to the distribution functions can identify the same group. Interpolate between L-Cs and L-Ck values of two stations of different groups to fix two boundaries one from the LCs and the other from L-Ck values the finally the boundary of the region is fixed between the mid ways of the two boundaries.

3.6 Homogeneity tests

Homogeneity can be taken as base for many criteria of the basin. Different tests are available to examine regional homogeneity in terms of the hydrologic response of the stations in regions.

Stations in a region can be tested for homogeneity that fall in a region; there are various homogeneity tests. The testes used in this study are Cv-based homogeneity test and L-Cv-based homogeneity test.

3.6.1 Cv- based homogeneity test

In regionalization assumption must be made about the statistical similarity of the site in a region. To investigate whether those has been meet or not many researches (Cunnane, 1989) have used the values of mean coefficient of variation (Cv) and the site-to-site coefficient of variation (Cv) of both convention and L-moment of the proposed region.

A simple test based on the variability of at-site CV values was used. In a region of m number of stations and n_j years of records at the j^{th} station having a standardized low flow discharge (qi,i= 1,2,3,. n_j), the coefficient of variation is defined as follows:

$$\boldsymbol{\sigma i} = \sqrt{\frac{\sum_i^n (Qi - \overline{Qi})^2}{n-1}} \quad \text{......Equation 5}$$

$$\boldsymbol{CVi} = \frac{\sigma i}{\bar{Q}i} \quad \text{......Equation 6}$$

Where; Qi = the flow rate of station i

$\bar{Q}i$ = the mean flow rate for station i σi = standard deviation of Qi for station i

n = number of years

CVi = coefficient of variation of station i

For each region using the statistics calculated using the above equations the regional mean Cv and finally the corresponding CC values were computed using the following relation:

$$\overline{CV}\boldsymbol{i} = \sum_{i=1}^{N} \frac{CVi}{Ni} \ldots \ldots \text{Equation 7}$$

$$\sigma_{CV} = \sqrt{\frac{\sum_{i=1}^{N}(CVi - \overline{CVi})^2}{Ni-1}} \ldots \ldots \text{Equation 8}$$

$$CC = \frac{\sigma cv}{\overline{Cv}} \ldots \ldots \text{Equation 9}$$

Where; N = number of station in a region

$\overline{CV}i$= Mean coefficient of variation

σ_{cv}= standard deviation of at site CV values

CC = combined coefficient of variation for the region. And the region declared to be homogeneous if

$CC < 0.3$

Stations which have approximately close Cv values and which lie within the same geographical proximity were first grouped together in one region and the internal homogeneity of the regions expressed numerically in terms of the flow statistics CC was determined. The criterion used to check for regional homogeneity was based on the value of CC. For homogeneous regions a value of CC was allowed to be 0.3 or less (Sine, 2004).

3.6.2 L-cv based homogeneity test

L- Moments are defined as linear combinations of probability weighted moments (PWMs) that have simple interpretations as measures of the location, dispersion and shape of the sample data (Hosking, 1990). He demonstrated that L-moment methods perform competitively with the best available statistical techniques. L-Cv-based homogeneity test is more accurate and effective way of testing the homogeneity of the station when compare with that of the Cv-based homogeneity test.

The procedural calculation is the same as that of the Cv, but it has its own formula and calculation;

The unbiased PWM estimators are:

$$b_0 = \frac{1}{n}\sum_{j=1}^{n} Xj \quad \text{...Equation 10}$$

$$b_1 = \frac{1}{n}\sum_{j=2}^{n} {(j-1)x_j}/{n-1} \quad \text{...Equation 11}$$

$$b_2 = \frac{1}{n}\sum_{j=3}^{n} {(j-1)(j-2)x_j}/{(n-1)(n-2)} \quad \text{...Equation 12}$$

$$b_3 = \frac{1}{n}\sum_{j=3}^{n} {(j-1)(j-2)(j-3)x_j}/{(n-1)(n-2)(n-3)} \quad \text{...Equation 13}$$

Where n is the sample size and Xj denotes the jth element of a sample of size n sorted in to ascending order.

The first few moments are:

L1 = b0

L2 = 2b1 – b0

L3 = 6b2 – 6b1 + b0

L4 = 20b3-30b2 +12b1 –b0

Like the conventional moments L- moment can be used to specify and summarize probability distribution. In particular L_1, the first L- moment is the mean of the statistical distribution and identical to the first conventional moment. And L_2 is a linear measure of spread or dispersion analogous to standard deviation.

The L- moment ratios are used as estimation procedures for obtaining growth curves and are defined by:

$$\tau_r = {\lambda_r}/{\lambda_2} \quad \text{...Equation 14}$$

r = 3, 4

L- Skewness; $\tau_3 = {\lambda_3}/{\lambda_2}$ ...Equation 15

L- Kurtosis; $\tau_4 = {\lambda_4}/{\lambda_2}$ ...Equation 16

L- Coefficient of variation; $\tau = {\lambda_2}/{\lambda_1}$ ...Equation 17

In particular 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. The L-Cv, $\tau = \lambda_2/\lambda_1$ is analogous to the usual coefficient of variation. 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.

Using the above procedural formulas we have:

$$\boldsymbol{LC_V = \frac{L_2}{L_1}} \quad \text{……} \quad \textbf{Equation 18}$$

And

$$\boldsymbol{\overline{LC}_V = \sum_{i=1}^{N} \frac{LC_V}{N}} \quad \text{……} \quad \textbf{Equation 19}$$

$$\boldsymbol{\sigma_{LCv} = \sqrt{\frac{\sum_{i=1}^{N}(LC_V - \overline{LC}_V)^2}{N-1}}} \quad \text{……} \quad \textbf{Equation 20}$$

$$\boldsymbol{CC = \frac{\sigma_{LCv}}{\overline{LC}_V}} \quad \text{……} \quad \textbf{Equation 21}$$

Where, N = number of station in a region

$\overline{CV}$ = mean coefficient of variation

$\sigma_{LCv} =$ Standard deviation of at station Cv values

CC = combined coefficient of variation for the region, the stations and regions declared to be homogeneous if CC< 0.3.

3.7 Selection of best fit statistical parent distributions

According to (Hosking, 1990) the L- moment ratio diagram are based on the relationship between the L- moment ratios and it is best way for selection of distributions for the homogeneous regions. To assess the reasonability of the selected distribution statistical tests like Chi-squared test and Kolmogorov- Smirnov tests are used. Depending on the results of goodness of fit tests by Kolmogorov-Smirnov Test and Chi-square, appropriate distributions were identified for each at site data and regions. The Chi square test and Kolmogorov-Smirnov tests are discussed below;

3.7.1 Kolmogorov – smirnov test

This test is used to decide if a sample comes from a hypothesized continuous distribution. It is based on the empirical cumulative distribution function (ECDF). Assume that we have a random sample X1, X2 ……Xn from some distribution with CDF F(x). The empirical CDF is denoted by: Fn(x) = 1/n (number of observations ≤ x).

Definition

The Kolmogorov- Smirnov statistic (D) is based on the largest vertical difference between the theoretical and the empirical cumulative distribution function:

$D_N = max|F_N(x) - F_0(x)|$

Hypothesis testing

The null and the alternative hypothesis are:

H0 = the data follow the specified distribution

Ha = the data do not follow the specified distribution

The hypothesis regarding the distributional form is rejected at the chosen significance level (α) if the test statistic, D, is greater than the critical value obtained from a table. The fixed values of α (0.01, 0.05 etc.) are generally used to evaluate the null hypothesis (H0) at various significance levels. A value of 0.05 is typically used for most applications.

The standard tables of critical values used for this test are only valid when testing whether a data set is from a completely specified distribution. If one or more distribution parameters are estimated, the results will be conservative; the actual significance level will be smaller than that given by the standard tables and the probability that the fit will be rejected in error will be lower.

P- Value

The P- value in contrast to fixed α value is calculated based on the test statistic and denotes the threshold value of the significance level in the sense that the null hypothesis (H0) will be accepted for all values of α less than the P-value.

For example, if P=0.025, the null hypothesis will be accepted at all significance levels less than P (i.e. 0.01 and 0.02), and rejected at higher levels, including 0.05 and 0.1.

The P- value can be useful when the null hypothesis is rejected at all predefined significance level, and you need to know at which level it could be accepted.

3.7.2 Chi- squared test

The Chi- squared test is used to determine if a sample comes from a population with a specific distribution.

The steps follow to use chi squared test are:

By estimating each of the unknown parameters of the assumed distribution need to divide the data into K classes or cells and determine the probability of a random value falling within each class using the fitted distribution then multiply each of the cell probabilities by the sample size n, let $P_i = \frac{n_A}{n}$ determining the expected number of observations (Ei) for each cell (i=1, 2, 3,…k), according to the fitted distribution $Ei = Pi * n$

From the given data, count the number of observed values in each cell (Mi) to compute the Chi – square Test statistic as follows,

$$\boldsymbol{X_C^2 = \sum_{i=1}^{k} \frac{(Mi-Ei)^2}{Ei} = X_C^2 = \sum_{i=1}^{n} n \frac{[f_s(x_i)-P(x_i)]^2}{P(x_i)}} \quad \text{…………………………………………………Equation 22}$$

Compare the computed value of Chi – square with the tabulated value of Chi – Square using (k-p-1) degrees of freedom (v) and a certain significance level (α) where p = number of parameters estimated in step II, k = number of classed, v = (k-p-1) = degrees of freedom and α = significance level (5% or 10%).

Finally if; $X_{v,a}^2 < X_C^2$ is not good fit but if $X_{v,a}^2 > X_C^2$is good fit.

3.8 Selection of parameter estimation methods

After a distribution or a number of distributions are selected to fit the data, their parameters must be estimated. A number of methods can be used for parameter estimation. The most commonly used methods are the Method of Moments (MOM), Method of Maximum Likelihood (MLM) and Method of probability weighted Moments (PWM).

The PWM method gives more accurate results than ML for parameter estimation from small samples. The parameter estimation procedures are less complicated compared to ML. Another attraction of the PWM method is that it can be easily used in regional estimation schemes (Rao, 2000). In this study, Method of Maximum Likelihood (ML) and Method of Moments (MOM) are not applied for estimation of parameters of selected probability distribution both at-site and regionally. But the PWM were applied to estimate the parameters because to avoid the effect of outliers in the recorded row data PWM is an efficient alternative for estimating the parameters and due to accuracy compared to the others using EASY FIT COMPUTER PROGRAM.

3.9 Quantile estimation

After the parameters of a distribution are estimated, quantile estimates (X_T) which correspond to different return periods T may be computed. The return period is related to the probability of non exceedance (F) by the relation, F=1-1/T where F= F (XT) is the probability of having a flow of magnitude X_T or smaller. The problem then reduces to evaluating X_T for a given value of F. In practice, two types of distribution functions are encountered.

The first type is that which can be expressed in the inverse form $X_T = \varphi$ (F). In this case, X_T is evaluated by replacing ф (F) by its value from above equation. In the second type the distribution cannot be expressed directly in the inverse form $X_T = \varphi$ (F). In this case numerical methods are used to evaluate X_T corresponding to a given value of φ (F).

Region – I: $\boldsymbol{X_T = \frac{Q_T}{Q_{mean}}}$ … ..Equation 23

Region – II: $\boldsymbol{X_T = \frac{Q_T}{Q_{mean}}}$ ….Equation 24

Where, Q_T is the low flow quantile,

Q_{mean} is the at-site mean annual low flow,

And T is the return period.

The low flow quantiles for the selected distributions of both regions are computed depending on the parameters of the basin as shown below:

$\boldsymbol{Q_T = \mu + \sigma log(T - 1)}$ ….Equation 25, **for generalized logic distribution**

$\boldsymbol{Q_T = \mu + \frac{\sigma}{k}[1 - (-logF)^k]}$ ….Equation 26, **for generalized extreme value distribution**

Where; σ, μ and k are parameters of the stations in the basin.

3.10 Regional frequency curve for homogeneous regions

The regional low flow frequency curve has important implications for hydrological processes. Each homogeneous region has its own fitted regional frequency distribution. Once the delineated region have been shown to be homogeneous and suitable distributions have been identified for the respective regions, regional growth curves can be developed based on the distributions. But the low flow event had follow the extreme event type III distribution and the standardized annual minimum discharge data for the sites in the regions were plotted against the set of plotting positions derived for the site. The plotting positions are related to the EV III reduced variant via the return period, T.

The graph is than called Regional growth curve. And it is possible to develop the regional frequency curve for each region using the standardized annual minimum flow with the return periods.

3.11 Flow duration curve

3.11.1 Characteristics of flow duration curve in low flow regimes

The cumulative frequency distribution of daily mean flow shows the percentage of time during which specified discharge is equaled or exceeded during the period of record. The relationship is normally referred to as the flow duration curve and although it does not convey any information about the sequencing properties of flows it is one of the most informative methods of displaying the complete range of river discharges from low to flood flows. Flow duration curve is a relationship between any given discharge value and the percentage of time that this discharge is equaled or exceeded, or in other words, the relationship between magnitude and frequency of stream flow discharge (Smakhtin, 2001). The flow duration curve has been developed by ranking all low flow data in descending order regardless of the sequence in which they occurred. The plotting position/probability of exceedance of each data has been calculated using the most commonly used formula known as Weibull plotting position:

$$\boldsymbol{P = \frac{m}{(N+1)} * 100} \quad \text{Equation 27}$$

Where; P = the probability that a given flow will be equaled or exceeded (% of time exceedance)

m = the ranked position on the listing (dimensionless)

N = the number of events for period of record (dimensionless)

In this study flow duration curves of 7, day have been derived for four stations. Low flow indices derived from FDC on the basis of daily flow are the percentage which indicates a high frequency of exceedance and the low flow period of regime. Common practice used as the low flow is 60, 70, 80, 90, 95 and 100 percentile Q60, Q70 Q80 Q90 Q95. Q100 (Flies, 2004)

4. RESULT AND DISCUSSION

4.1 Data analysis

4.1.1 Test for independency and stationery

This analysis of observed data is done to check whether the data have no serial correlation or short range dependence in the time series and the hydrological regime remains stationary during the period of record or no long term trend in the time series.

Based on the above description the results of the test of the lag one correlation coefficients for the stations are tabulated below:

Table 1: Result of independence test using serial correlation coefficient

S. No	Station name	Confidence level at 5%		Lag-one/r1	Remark
		UCL	LCL		
1	Gheba	0.4119	-0.5369	0.4093	Independent
2	Ilala	0.4119	-0.5369	0.2223	Independent
3	Agulae	0.4119	-0.5369	0.000835	Independent
4	Genfel	0.4119	-0.5369	0.6819	Not independent
5	Suluh	0.4119	-0.5369	0.1890	Independent

Generally the values of lag one correlation coefficients of the Gheba, Ilala, Agulae and Sulu stations are fall between the UCL (r1) and LCL (r1) and the hypothesis is acceptable and the data are independent and have no correlation between each observed data or no observation in the data series has any influence on any following observation. But station Genfel is not in the range of the lag one correlation coefficient and they are not acceptable for the further analysis.

4.1.2 Checking consistency of the data

As shown in the graph below the DMC analysis of the data in Gheba river sub basin all stations are approximately straight and consistent. But in the above analysis, station of Genfel is already rejected and no need for further analysis.

Generally the data for these consistent stations are generated by the same mechanism that generated similar or related data at other stations or neighboring stations.

The DMC of Gheba station with the nearby stations is shown below;

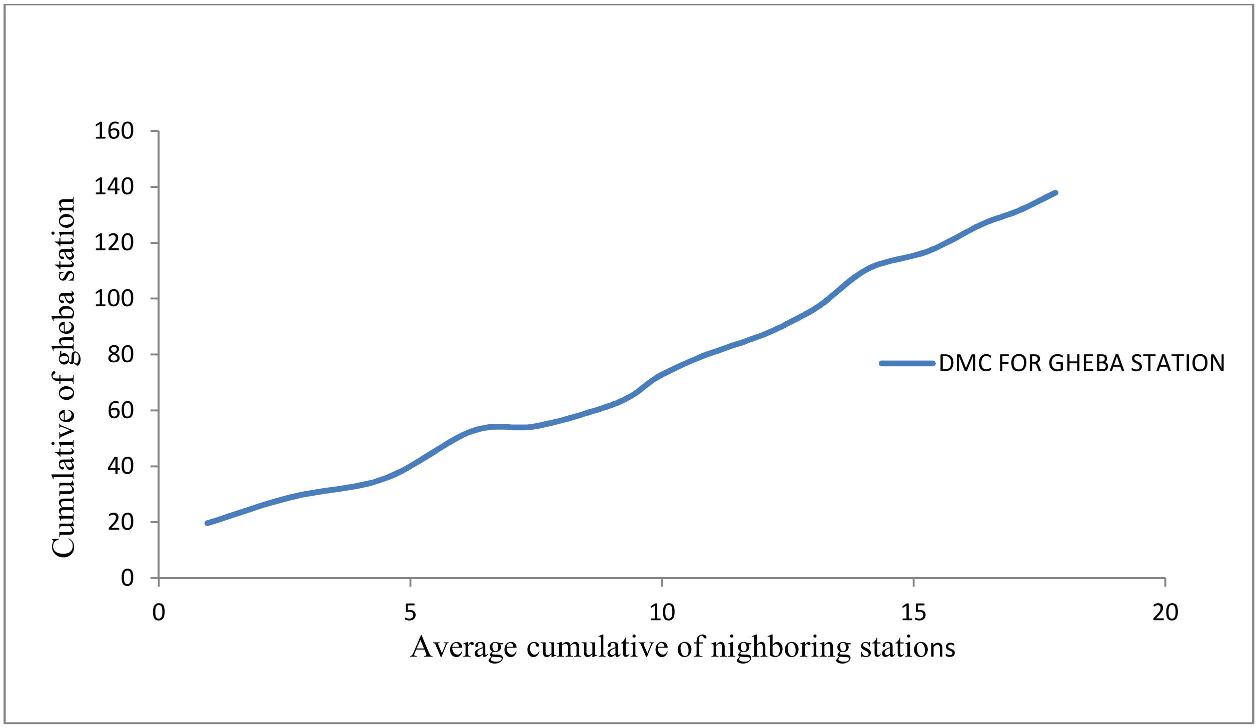


Figure 2 Double mass curve for Gheba station

The DMC of Ilala station with the nearby stations is as shown below;

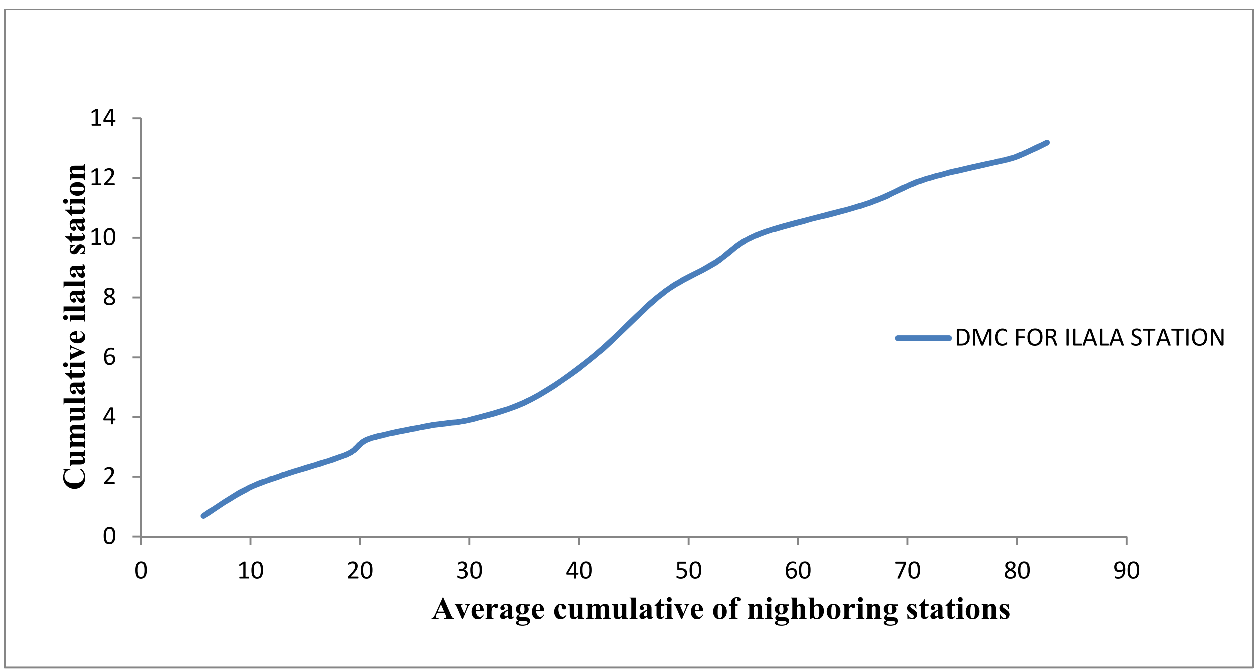


Figure 3 Double mass curve for Ilala station

The DMC of Agulae station with the nearby stations is as shown below;

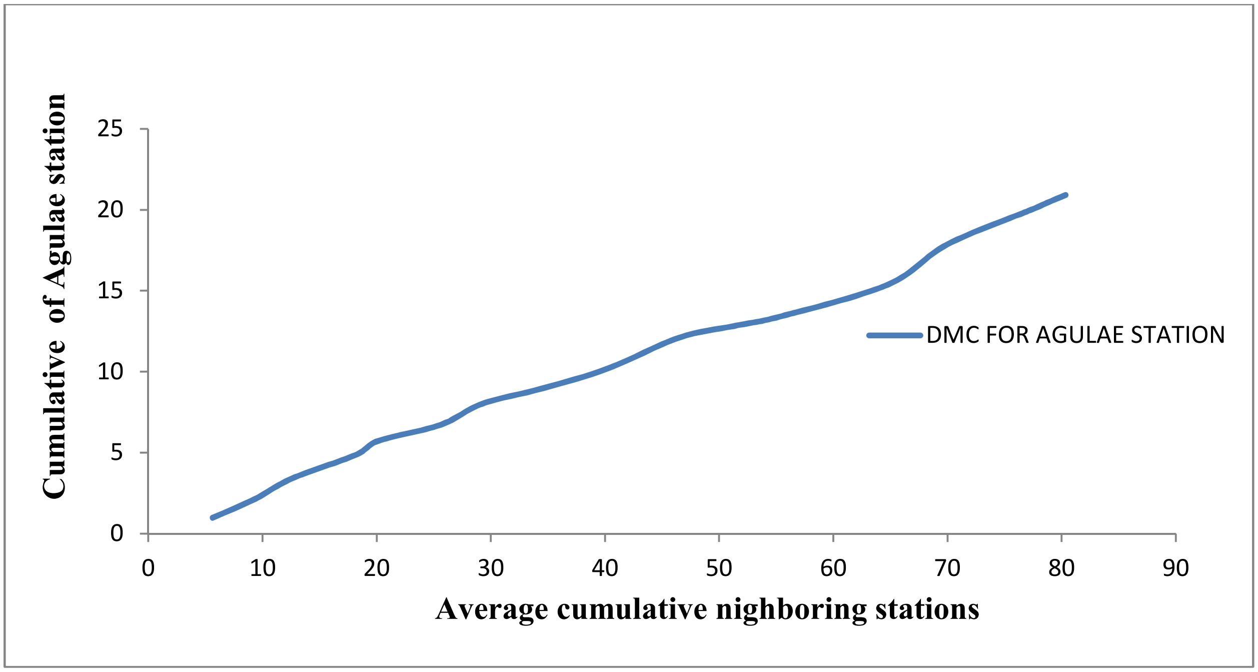


Figure 4 Double mass curve for Agulae station

The DMC of Suluh station with the nearby stations is as shown below;

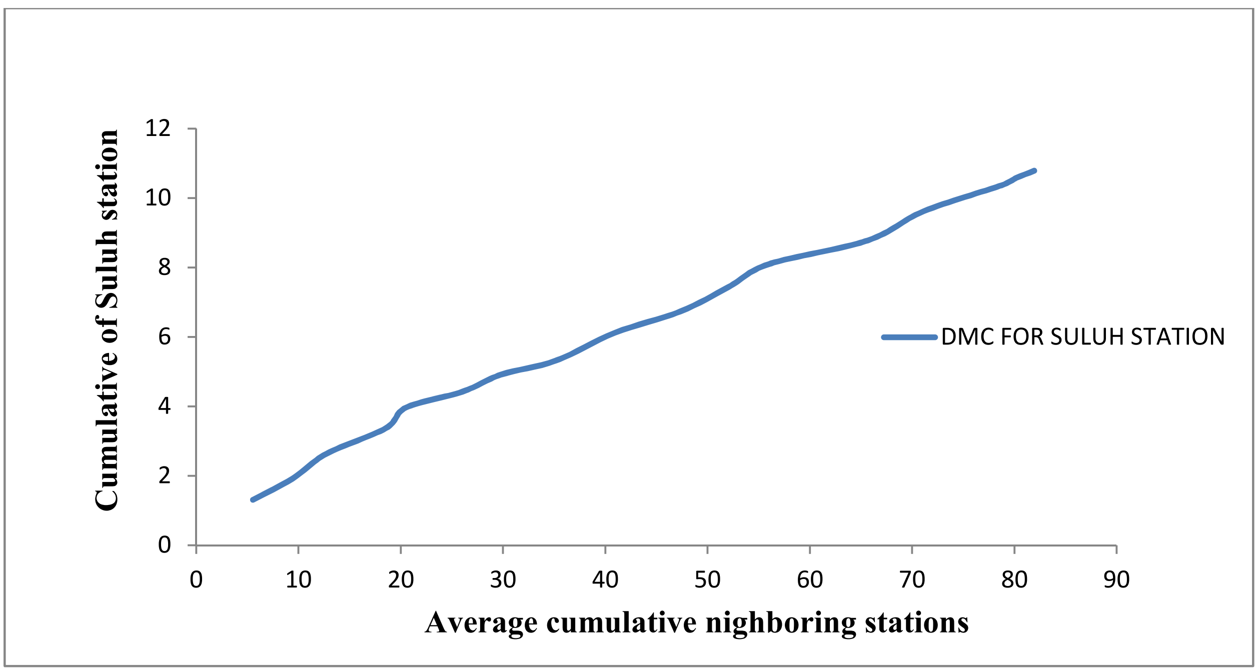


Figure 5 Double mass curve for Suluh station

Generally these four stations of the study area are with good quality of data and they are used for further analysis of the study.

4.2 Low flow statistics of the basin

Low flow statistics of Gheba sub basin stations were computed using both conventional moment and L-moment methods using procedure indicated in the above section 3.5. The low flow statistics computed for the basin for selected stations are used for the low flow frequency analysis technique as an input for identification of homogeneous regions, testing of the selected homogeneous region and for LMRD and the result is shown in table 2 below:

Table 2 Low flow statistics of the basin

S.No	Station name	Mean flow	No. year	Area (km^2)	Moment values					
					Cv	Cs	Ck	L-Cv	LCs	L-Ck
1	Gheba	15.7232	17	4342	0.9263	0.3895	3.8884	0.8319	0.0277	0.1723
2	Agulae	1.0092	17	692	0.5742	-0.3299	2.7056	0.6176	-0.0844	0.0640
3	Ilala	0.4517	17	190	0.773	1.7421	6.0811	0.8256	0.4460	0.2725
4	Suluh	0.6364	17	399	0.8041	2.3829	10.553	0.6705	0.3313	0.3535

4.3 Identification of homogeneous regions

The L-Cs vs. L-Ck moment ratio of standardized flow of the stations was plotted with the L-moment ratio diagram (LMRD) of various distribution functions as shown below in Figure 5. Furthermore, stations which lie in the same distribution were preliminary grouped as a homogeneous region.

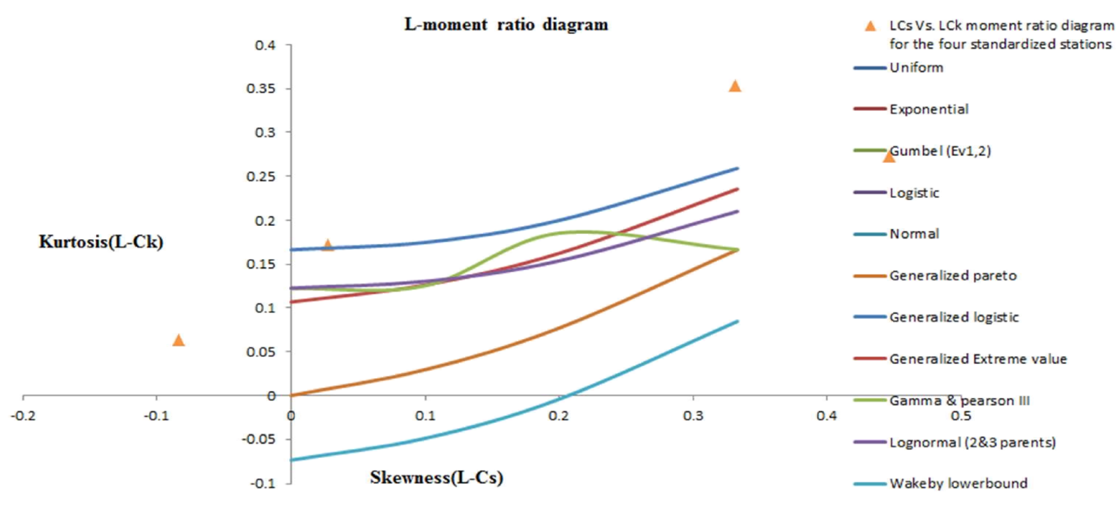
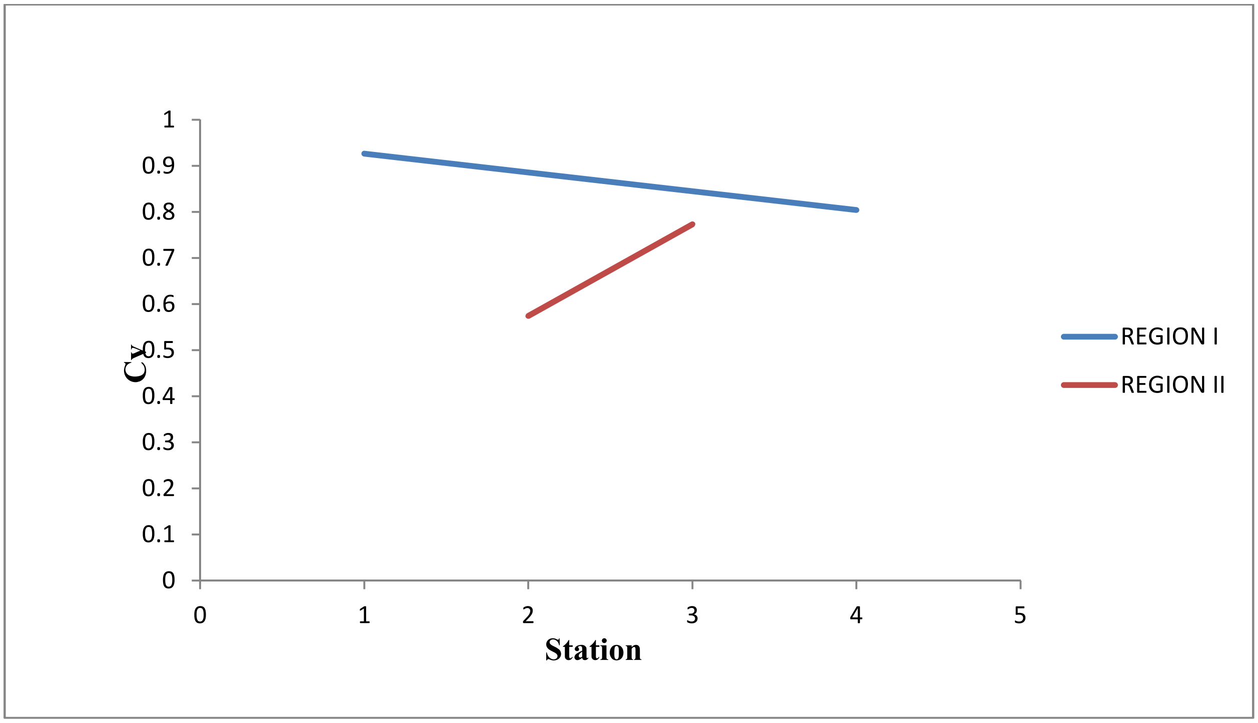


Figure 6 L-Cs – L-Ck moment ratio diagram for the four standardized stations

This was done on the origin of the fact that stations from the same homogenous region are distributed along the same distribution of LMRD.

According to the LMRD of the stations of the basin the basin was grouped in to two regions.

Those stations in the same region show statistical parameters within specified range as indicated in Table 3. Particularly, when these statistical values are plotted as in Figures 7 and figure 8, they show layers or group of regions that indicate different regions. Moreover, coefficient of variation was found to be a good indicator of homogeneity of the stations.



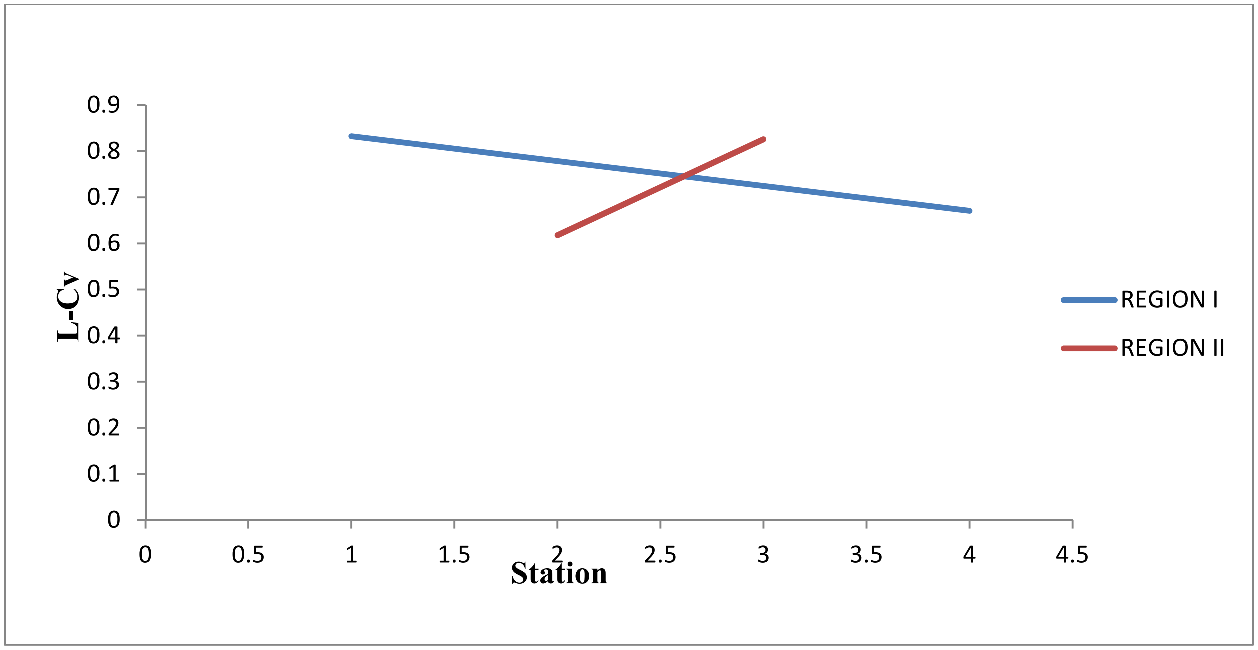
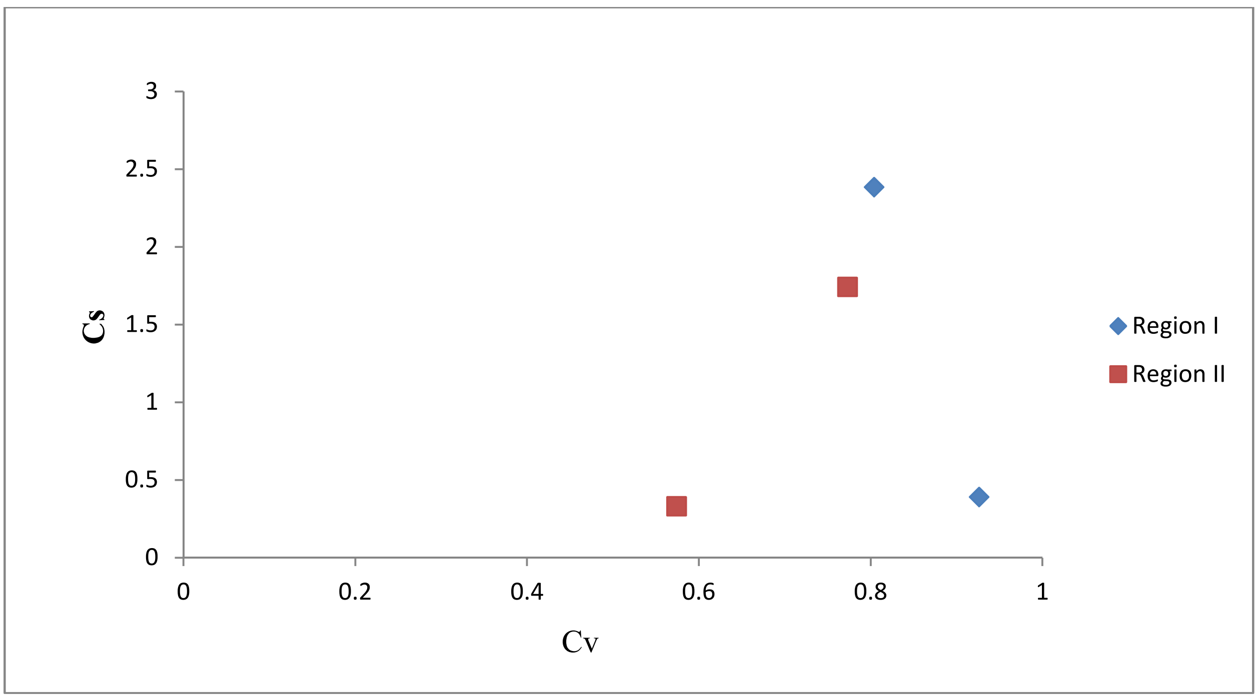


Figure 7 Homogeneity test by comparison of statistical values of both conventional and L-moments



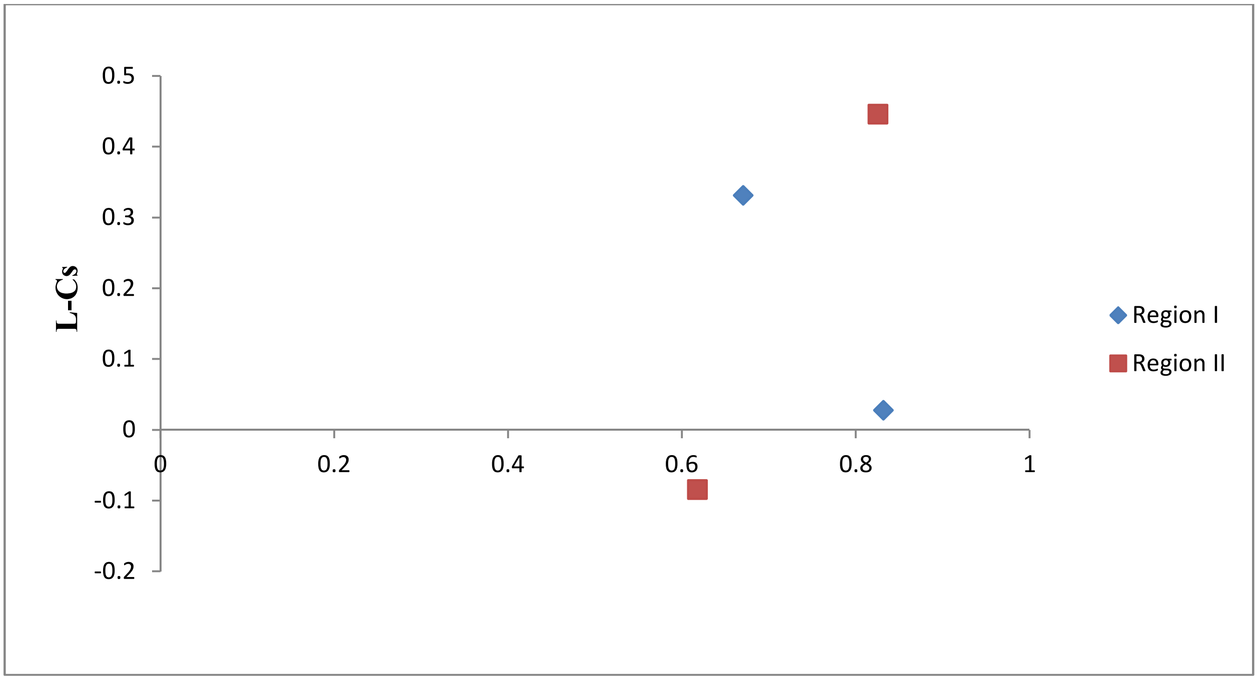


Figure 8 plots of conventional and L-moment ratio for two regions

From the above Figures 7 and 8 the layers of the conventional moment coefficients and L-moment coefficients are indicated the homogeneous regions of the basin.

Table 3 summary of moment values of the regions

Moment value		Region I	Region II
Cv	Range	0.8041-0.9263	0.5742-0.773
	Average	0.8652	0.5632
Cs	Range	0.3895-2.3829	-0.3299-1.7421
	Average	1.3862	0.7061
Ck	Range	3.8884-10.5529	2.7056-6.0811
	Average	7.2207	4.3934
L-Cv	Range	0.6705-0.8319	0.6176-0.8256
	Average	0.7512	0.7216
L-Cs	Range	0.0277-0.3313	-0.0844-0.4460
	Average	0.1795	0.1808
L-Ck	Range	0.1723-0.3535	0.0640-0.2725
	Average	0.2629	0.1683

Starting with the above LMRD and the graph for the layer of statistical values, the basin has divided into two regions. The first region identified includes the Gheba and Suluh stations and the second region includes the Agulae and Ilala stations.

The results obtained from homogeneity tests have strengthened the preliminary division of regions. Table 4 and Table 5 below shows the results of coefficient of variation based homogeneity test for each regions.

Table 4 coefficient of variation based homogeneity test for region I

S.No	Station	Area (km^2)	No. years	Mean flow	Cv	Cs	Ck	L-Cv	L-Cs	L-Ck
1	Gheba	4342	17	15.7232	0.9263	0.3895	3.8884	0.8319	0.0277	0.1723
2	Suluh	399	17	0.6364	0.8041	2.3829	10.5529	0.6705	0.3313	0.3535
Average Coefficient of variation					0.8652			0.7512		
Average Standard deviation					0.0864			0.1141		
Combined coefficient					0.1			0.1519		

The combined coefficient of variation for region I was computed as mentioned in section 3.5. From the above Table 4 has been concluded that the region is homogeneous for both convectional Cv based and L-moment L- Cv based homogeneity tests since the CC values are less than or equal to 0.3 for both case.

Table 5 coefficient of variation based homogeneity test for region II

S.No	Station	Area (km^2)	No. years	Mean flow	Cv	Cs	Ck	L-Cv	L-Cs	L-Ck
1	Agulae	692	17	1.0092	0.5247	-0.3299	2.7056	0.6176	-0.0844	0.0640
2	Ilala	190	17	0.4517	0.773	1.7421	6.0811	0.8256	0.4460	0.2725
Average Coefficient of variation					0.6736			0.7216		
Average Standard deviation					0.1406			0.1471		
Combined coefficient					0.2087			0.2038		

Similarly, From the above Table 5 it can concluded that region II is also homogeneous for both convectional Cv based and L-moment L- Cv based homogeneity test; since the CC values are less than or equal to 0.3.

Delineation of homogeneous regions was done using Arc GIS by taking in to account the drainage boundaries of the sub-catchments in each region. See Figure 9 below which shows the map of the established homogenous regions.

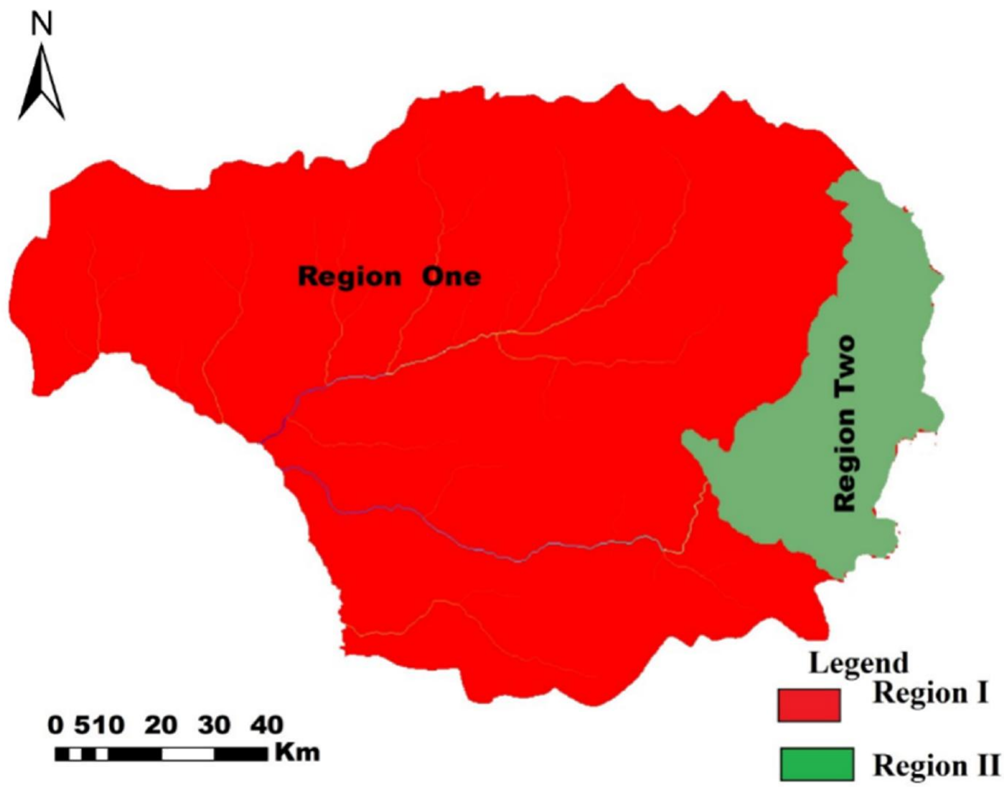


Figure 9 map for homogeneous region

4.4 Selection of best fit statistical parent distributions

Even the commonly used frequency distribution for low flow event is known i.e. EV III, but it is necessary to select appropriate distributions for the study area. After confirming the homogeneity of the study area an appropriate statistical parent distribution needs to be selected for the regional low flow frequency analysis. The selection was carried out by plotting the sample average L-skewness (L-Cs) and L- kurtosis (L-Ck) of the regions on L- moment ratio diagram (LMRD) for the theoretical distribution.

A diagram based on L-Cs versus L-Ck can be used to identify appropriate distributions. The best parent distribution is that one for which the average value of the point (L-Cs, L-Ck) of all stations within the region gets close to one of the drawn LMRD of the parent distribution.

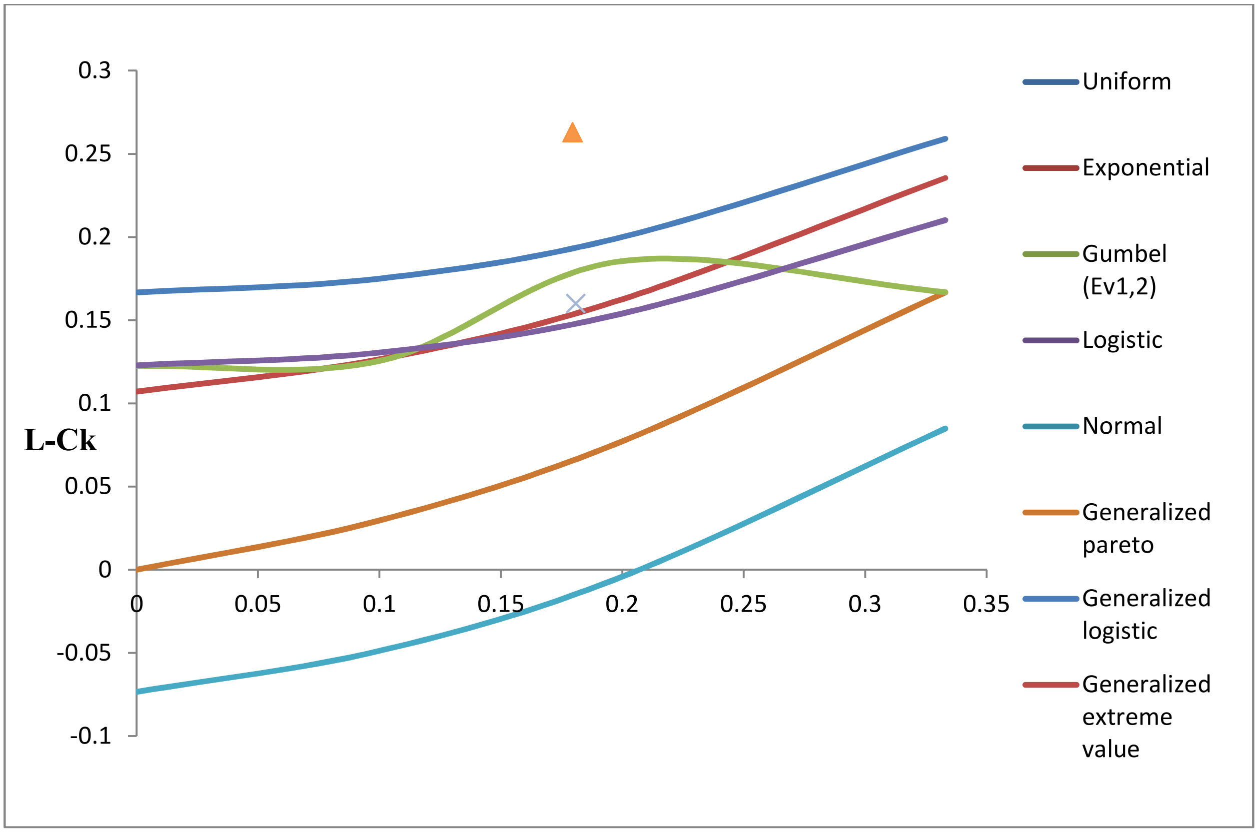


Figure 10 L-Cs - L-Ck moment ratio diagram for regionalized stations

Accordingly the most possible underlying candidate distributions from Figure 10 are summarized in the Table 6 below:

Table 6 selected candidate distributions for the regions of the basin

Region	Average regional L-moments (L-Cv, L-Cs , L-Ck)	Selected candidate distributions
Region I	(0.7512, 0.1795, 0.2629)	Generalized logic
Region II	(0.7216, 0.1808, 0.1683)	Generalized extreme value

These distributions may or may not represent the region, to assess the reasonability of the selected distribution statistical tests like Chi-squared test and Kolmogorov- Smirnov tests are used.

4.5 Goodness of fit test for candidate distributions

4.5.1 Chi-squared test

Using the procedures discussed in the above chapter three the chi-square value of each stations were computed. And the computed value was compared with the critical value of chi- square at a given degree of freedom, at 5% significant level, finally as indicated below in Table 7 for Xv^2,α>X^2c is good fit but if Xv^2,α<Xc^2 is not good fit. But here the parameter value (P-value) can be useful when the hypothesis are rejected at all predefined significance level and you need to know at which level it could be accepted.

Table 7 result of chi square test for selected stations

Region	Station name	Distribution	$X^2_{v,\alpha}$	X^2_c	Remark
Region one	Gheba	Generalized logistic	5.99	5.059	Accepted
	Suluh	Generalized logistic	5.99	0.941	Accepted
Region two	Agulae	Generalized extreme value	3.84	0.941	Accepted
	Ilala	Generalized extreme value	3.84	2.118	Accepted

In the above Table 7 all the distributions are acceptable for all stations in both regions.

4.5.2 Kolmogorov- smirnov tests

The Kolmogorov- Smirnov statistic (D) is based on the largest vertical difference between the theoretical and the empirical cumulative distribution function:

$$DN = max|FN(x) - FO(x)|$$

If the computed DN is greater than the critical value obtained from table at chosen confidence level the distribution will be rejected, if not the distribution is best fit and accepted.

Table 8 result of kolmogorov-smirnov test for selected stations

Region	Station name	Distribution	D_m	Critical value @ α = 5%	Remark
Region one	Gheba	Generalized logistic	0.188	0.318	Accepted
	Suluh	Generalized logistic	0.106	0.318	Accepted
Region two	Agulae	Generalized extreme value	0.106	0.318	Accepted
	Ilala	Generalized extreme value	0.094	0.318	Accepted

Using the above criteria for selecting the best fit distribution; the results for Kolmogorov-smirnov test are indicated in the above Table 8. And generally the common distributions which pass the criteria were selected from both tests for each regions of the basin.

4.6 Selection of parameter estimation methods

For the distributions which were chosen in the previous section 4.5 the following candidate methods of parameter estimation were applied; for Generalized Logistic distribution (ML, MOM, PWM) and for Generalized extreme value distribution (ML, MOM, PWM), are applied to estimate the parameters.

From the candidate methods of parameter estimation the probability weighted moment method was selected and applied to estimate the parameters, because it is the best method to avoid the effect of outliers in the recorded row data. Due to this reason, PWM is an efficient alternative for estimating the parameters.

Based on the results obtained in the sections 4.4 and 4.5 above the following distributions were found to be the best procedures for describing the annual minimum low flow and for predicting acceptable low flow estimates for the delineated regions:

4.6.1 At-site and regional parameters

The parameters of each station in a region were computed by applying the idea discussed in the above section 4.5. At site parameter estimation, for each station was done by using EASY FIT computer software and the result are shown in Table 9 below. This program makes the parameter estimation easy and simplest way for computing parameters of each station.

Table 9 estimated parameters for selected distributions

Distribution	Parameters	Station name			
		Gheba	Suluh	Agulae	Ilala
Generalized logistic	σ	4.6864	0.11209		
	μ	15.723	0.63645		
Generalized extreme value	K			-0.44387	0.38961
	σ			0.39032	0.15326
	μ			0.90875	0.26841

The regional parameters of the selected distributions for region one and region two are simply computed as the average of site parameters as shown in Table 10.

Table 10 average regional parameters for selected distributions for both regions

Distribution	Parameter	Region	
		ONE	TWO
GENERALIZED LOGIC	σ	2.399	
	μ	8.180	
GENERALIZED EXTREME VALUE	k		-0.027
	σ		0.272
	μ		0.589

4.7 Low flow quantiles and regional low flow frequency growth curve

The results of estimated dimensional quantiles for selected four stations in the basin are shown in Table 11 and for recurrence intervals: 2, 5, 10, 25, 50, 100, 200, 500, 1000 and 10000years.

Using the computed regional parameters in the above Table 10 for selected distributions i.e. Generalized Logic for Region-1 and Generalized extreme value for Region-2, the standardized regional low flow quantiles were estimated as shown in Table 12 for recurrence intervals of 2, 5, 10, 25, 50, 100, 200, 500, 1000 and 10000 years. Figure 11 and the regional growth curves for both region-1 and region-2 is developed using the variant of the distribution applied for low flow events; which is the extreme event type III versus the regional standardized low flow quantile for both regions.

This is computed using the relationship:

Table 11 estimated quantiles at site for the four stations in the basin

T	GENERALIZED LOGIC				GENERALIZED EXTREME VALUE					
	GHEBA		SULUH		AGULAE			ILALA		
	σ = 4.6864	μ = 15.723	σ = 0.11209	μ = 0.63645	K = -0.44387	σ = 0.39032	μ = 0.90875	K = 0.38961	σ = 0.15326	μ = 0.26841
2	15.723		0.63645		1.527			0.4154		
5	18.544		0.704		2.507			0.5034		
10	20.195		0.743		3.486			0.5435		
25	22.191		0.791		5.294			0.5801		
50	23.644		0.826		7.223			0.5997		
100	25.075		0.860		9.837			0.6145		
200	26.496		0.894		13.385			0.6257		
500	28.367		0.9389		20.1005			0.6366		
1000	29.780		0.973		27.337			0.6425		
10000	34.468		1.085		75.929			0.6540		

The above Table 11 shows the overall at site quantiles for generalized logic and generalized extreme value distribution for four stations in the basin and they are used for computing the regional low flow quantiles for the homogeneous regions and the standardized low flow for developing the low flow frequency curve.

Table 12 estimated quantiles at the basin for both regions

T	REGION I		REGION II		
	GENERALIZED LOGIC		GENERALIZED EXTREME VALUE		
	σ = 2.399	μ = 8.180	K = -0.027	σ = 0.272	μ = 0.589
2	8.18		0.9290		
5	9.6243		1.2443		
10	10.4692		1.4639		
25	11.4911		1.7478		
50	12.2348		1.9632		
100	12.9675		2.1810		
200	13.6949		2.4022		
500	14.6527		2.7005		
1000	15.3760		2.9309		
10000	17.7759		3.7274		

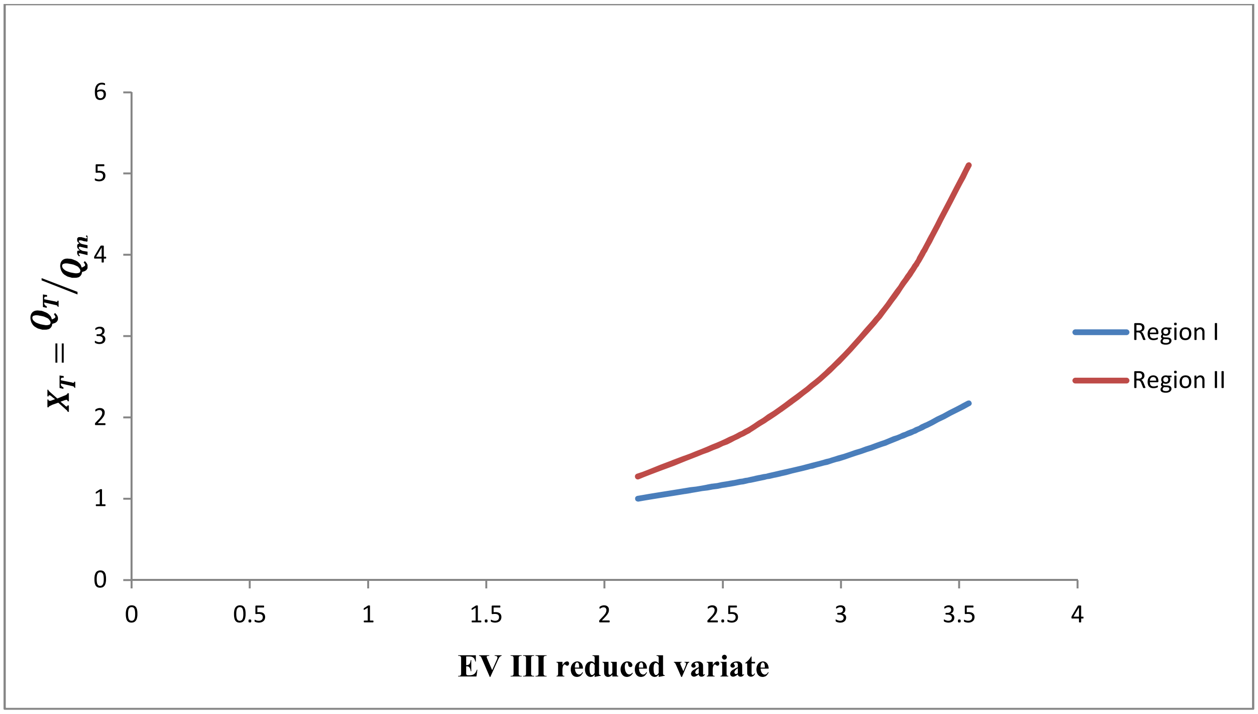
The regional growth curve is used by designers and hydrologists during planning and prefeasibility studies in areas where data is scarce. Since the low flow event follow the EV III the varait of EV III is used to develop the low flow frequency of both region one and region two. And finally using the standardized regional low flow for region one (Generalized logic distribution) and for region two (Generalized extreme value) versus the EV III the growth curve were developed as shown in figure 10.

Table 13 standardized regional low flow quantiles using generalized logic distribution for region I

REGION I			
Return period, T	Q_T	EV III reduced variate	$X_T = \frac{Q_T}{Q_m}$
			$Q_m = 8.1798$
2	8.18	2.1408	1
5	9.6243	2.5152	1.1766
10	10.4692	2.6959	1.2799
25	11.4911	2.8778	1.4048
50	12.2348	2.9898	1.4957
100	12.9675	3.0846	1.5853
200	13.6949	3.1736	1.6742
500	14.6527	3.2750	1.7913
1000	15.3760	3.3442	1.8798
10000	17.7759	3.5405	2.1731

Table 14 standardized regional low flow quantiles using generalized extreme value distribution for region II

REGION II			
Return period, T	Q_T	EV III reduced variate	X_T
			$Q_m = 0.7305$
2	0.9290	2.1408	1.2717
5	1.2443	2.5152	1.7034
10	1.4639	2.6959	2.0040
25	1.7478	2.8778	2.3926
50	1.9632	2.9898	2.6875
100	2.1810	3.0846	2.9856
200	2.4022	3.1736	3.2884
500	2.7005	3.2750	3.6968
1000	2.9309	3.3442	4.0122
10000	3.7274	3.5405	5.1025



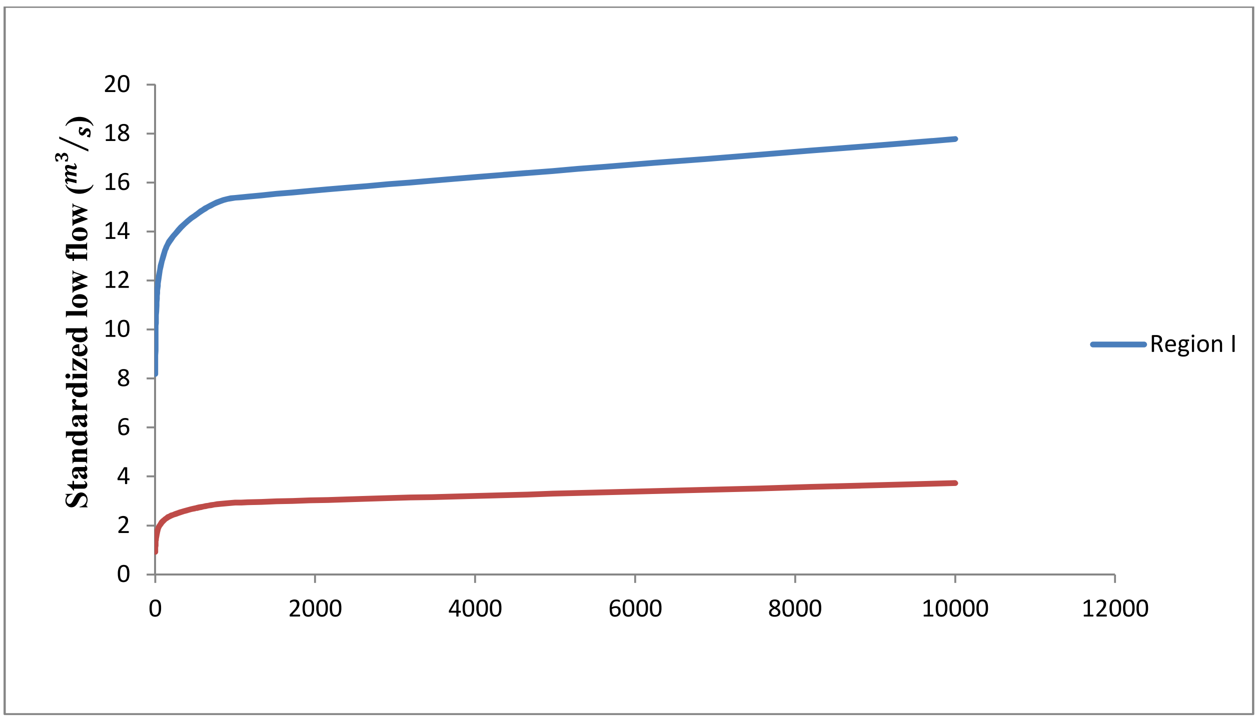


Figure 11 Regional growth curves

The slope of a frequency curve graphically represents the standard deviation of the low flow frequency distribution, and the higher the slope, the greater the standard deviations in low flow discharges. And also the steep slope of the curve represents maximum variation on the observed low flow but, the flat slope represents the observed data have minimum variation.

4.8 Characteristics of flow duration curve in low flow regimes

Low flow indices derived from FDC on the basis of 7- day annual minimum daily flows are the percentage which indicates a high frequency of exceedance and the low flow period of regime. Common practices used as the low flow are 60, 70, 80, 90, 95 and 100 percentile Q60, Q70, Q80, Q90, Q95& Q100 (Flies, 2004).

The low flow indices of the stations in the basin were taken from the flow duration curve plotted using the above section 3.11 procedures of developing FDC and the results for standard or common low flow indices are listed below in the Table 15.

Table 15 Low flow indices for the basin

Station name	60%	70%	80%	90%	95%	100%
GHEBA	0.5278	0.3669	0.1962	0.086	0.0155	0.0028
AGULAE	0.0561	0.0513	0.0407	0.0219	0.0039	0.0007
ILALA	0.0058	0.005	0.005	0.005	0.0009	0.0002
SULUH	0.0076	0.0054	0.0034	0.0029	0.0005	0.0001

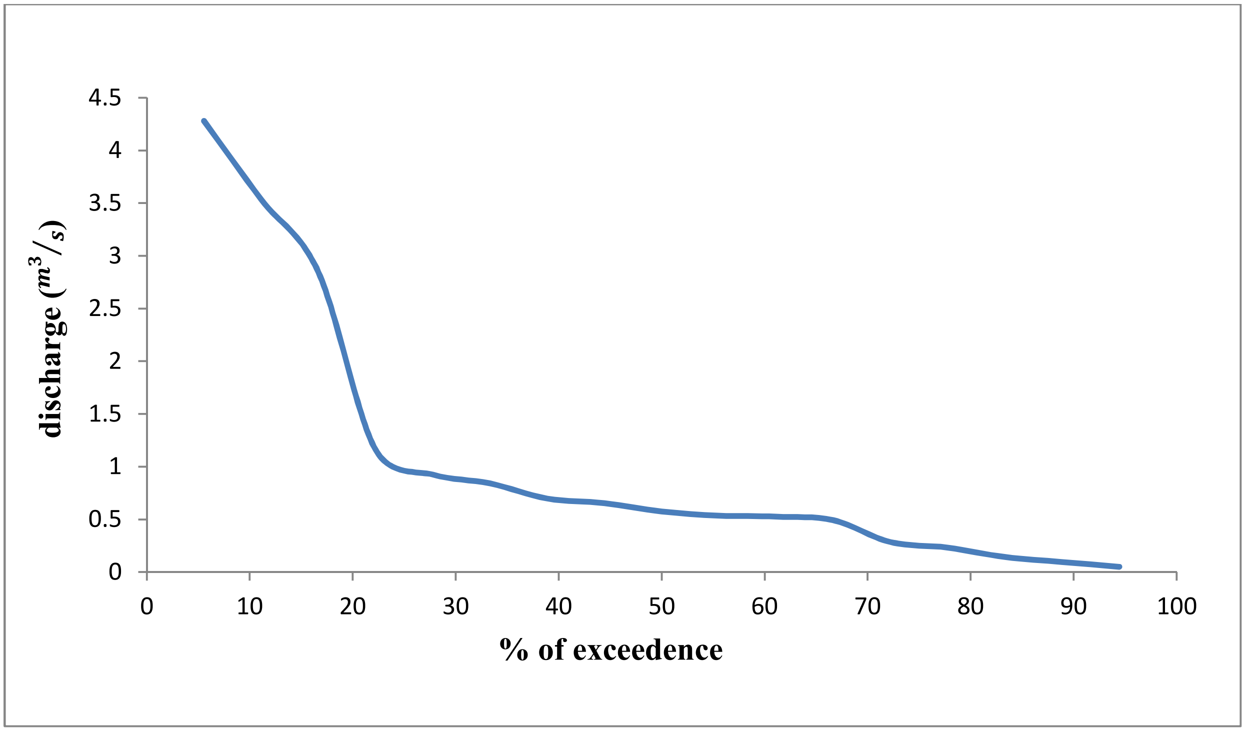


Figure 12 Flow duration curve for Gheba station

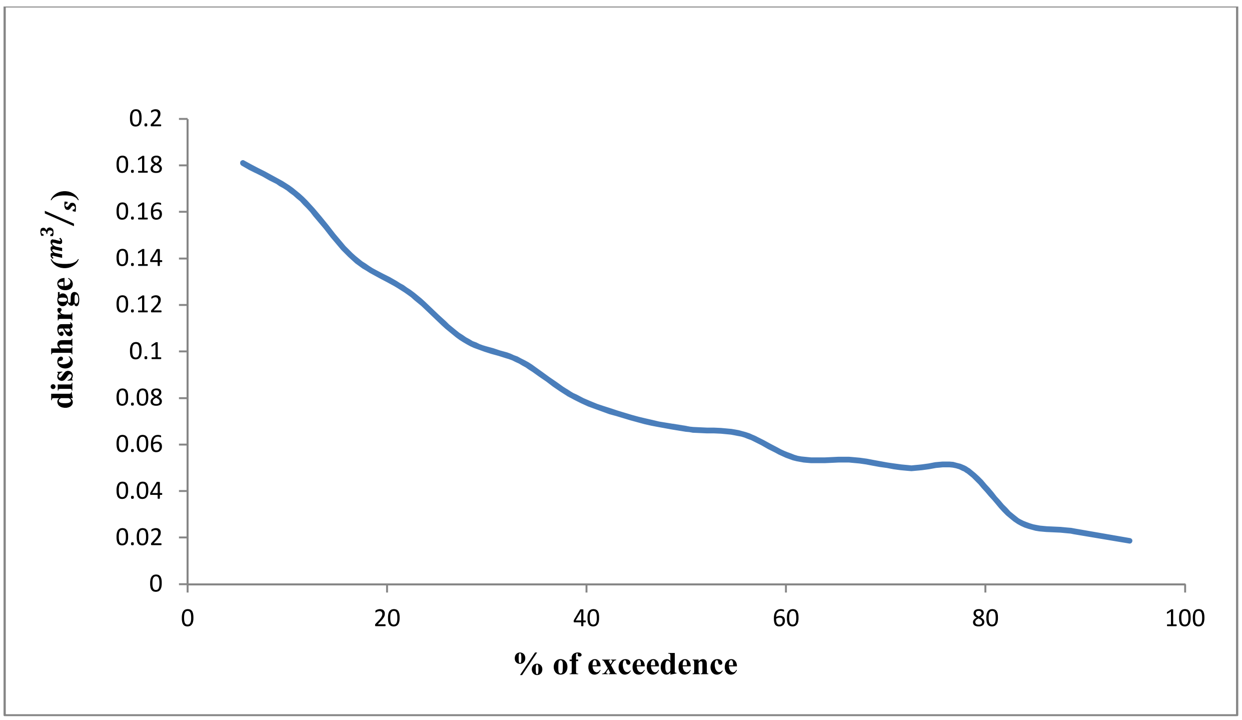


Figure 13 Flow duration curve for Agulae station

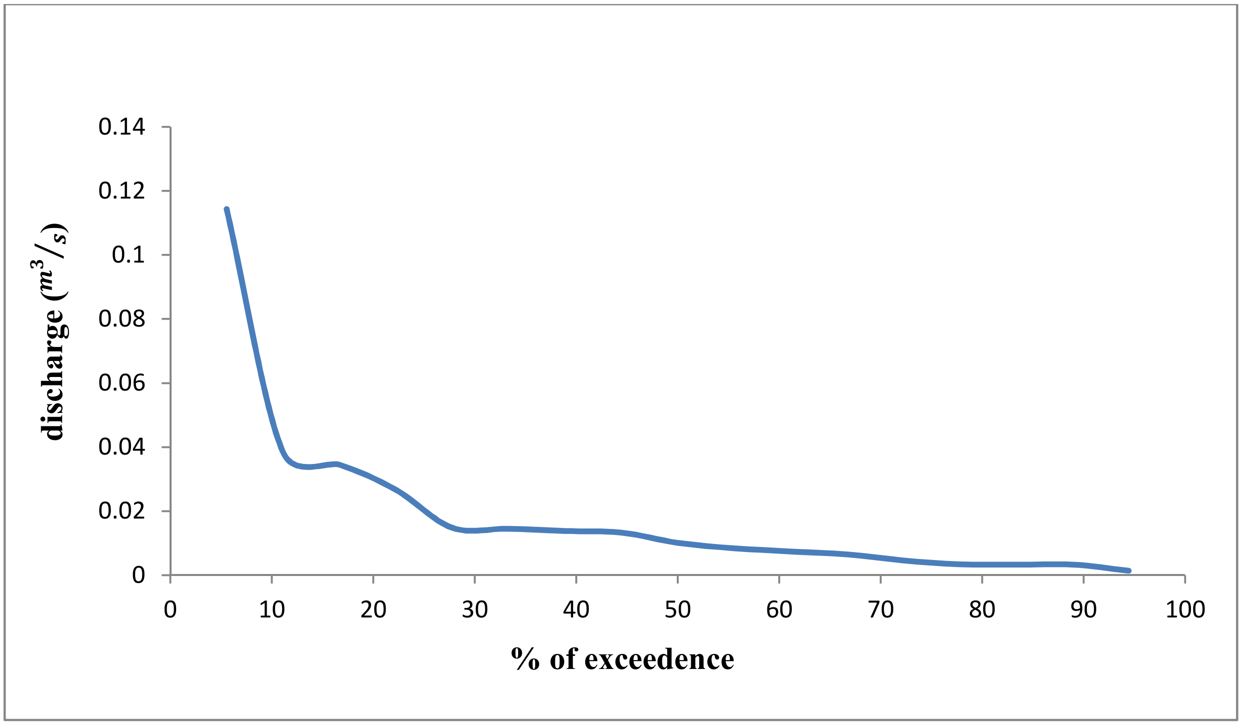


Figure 14 Flow duration curve for Suluh station

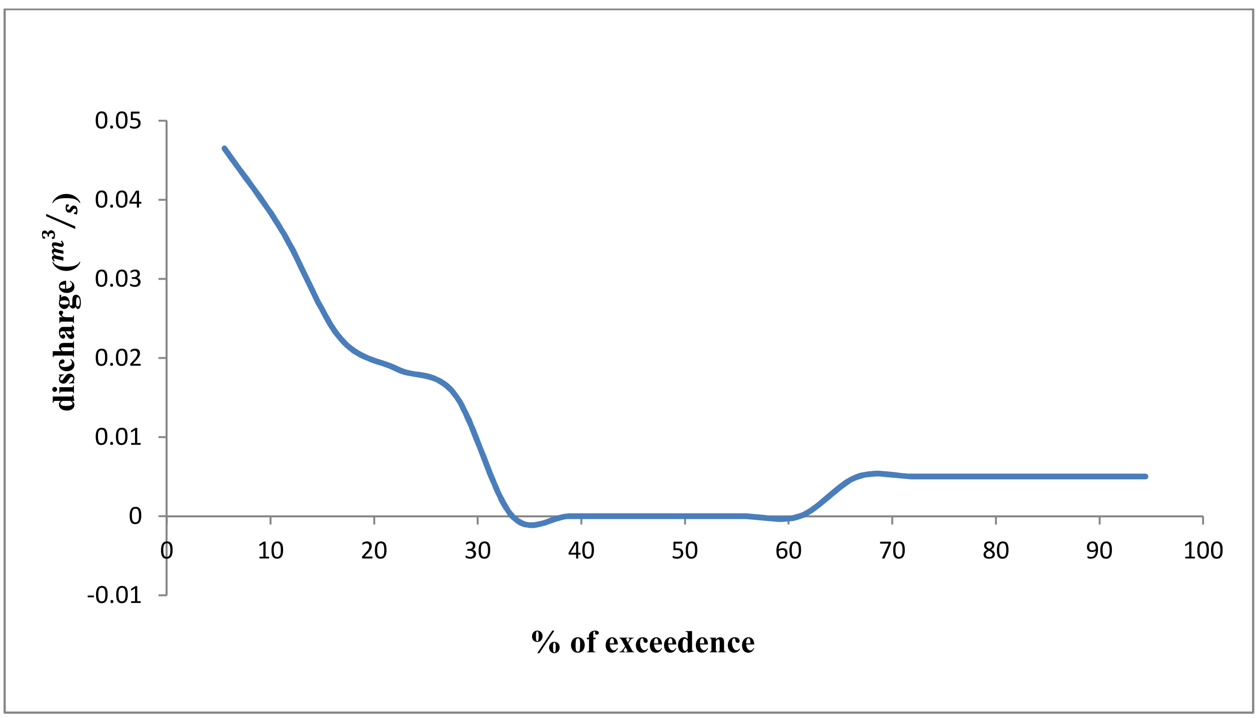


Figure 15 Flow duration curve for Ilala station

Based on the characteristics shown in the right tailed end portion of FDC and the slope of the FDC the basin FDC were categorized in to two groups. And this was done by simply observing by visual inspection. In the group one FDC indices the stations Gheba and Suluh are grouped and in group two stations Agulae and Ilala are grouped as shown in the Figure 18 and Figure 19 below.

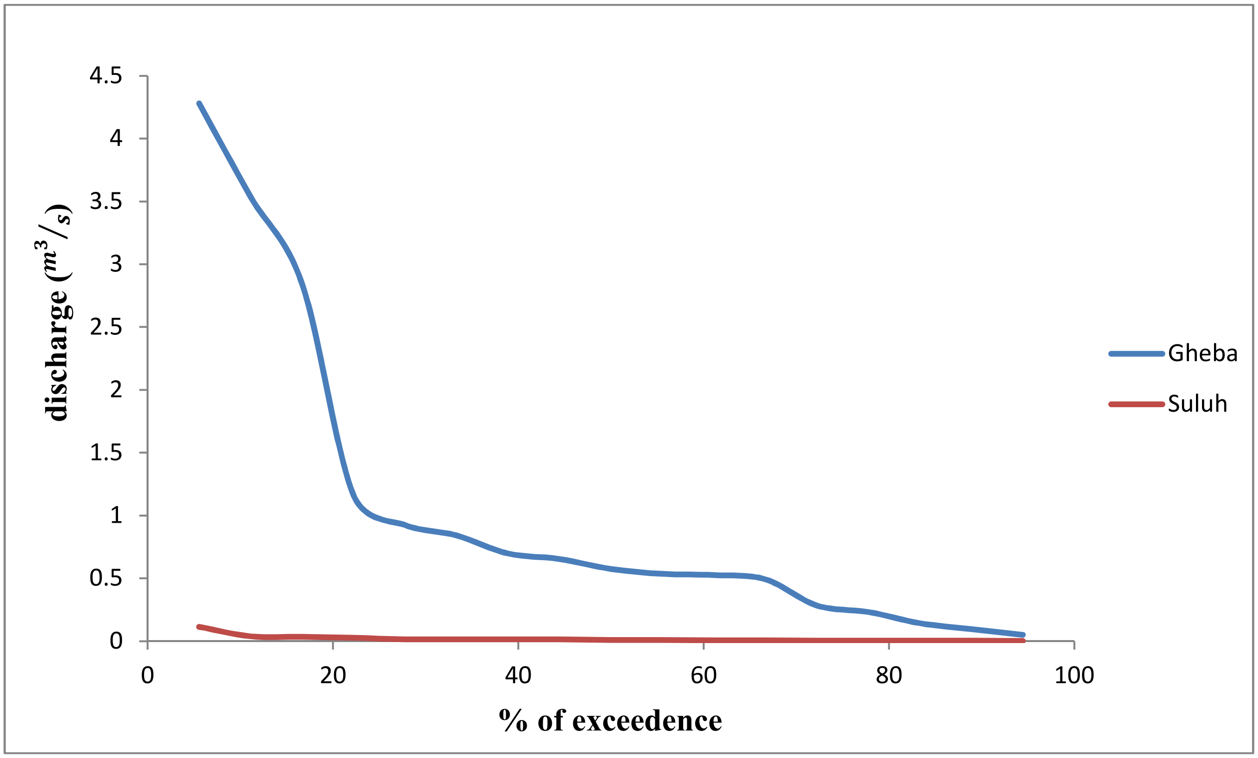


Figure 16 FDCs for group one stations

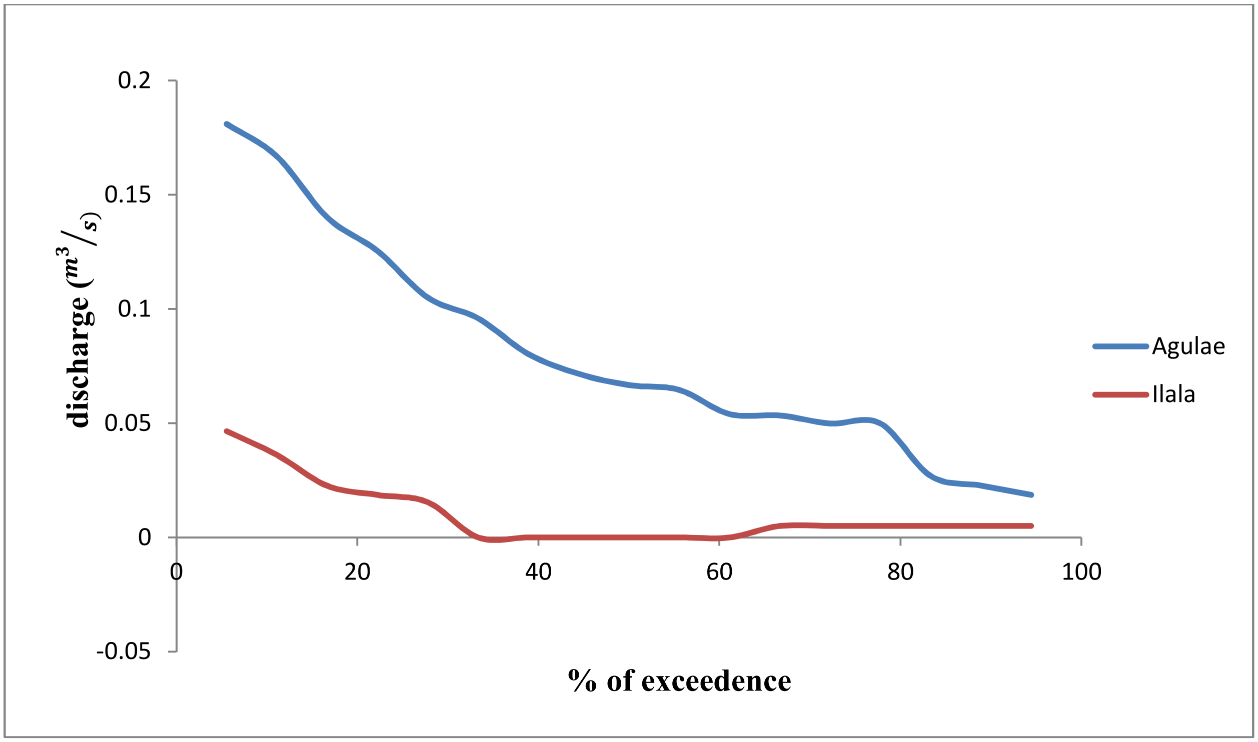


Figure 17 FDCs for group two stations

CONCLUSION AND RECOMMENDATION

1. Conclusion

Gheba River is one of the sub basins of Tekeze utilized for different purposes. The main objective of the study was to delineate Gheba River in to hydrological homogeneous regions on the basis of their low flow characteristics. The two techniques were applied; the low flow frequency analysis and the flow duration curve. The typical annual minimum mean 7- day low flow was used in the analysis. Out of five stations in the study area, four stations are used for the analysis of the study.

On the basis of low flow frequency analysis two homogeneous regions were identified and finally delineated using Arc GIS. Region one of the basin includes the Gheba and Suluh stations, and region two encloses the stations of Agulae and Ilala. The Homogeneity of regions was validated using the L-CV and CV based tests. The L-Moment ratio diagram provided a practical means to group stations into different regions and also identified the underlying distribution for a given region. Suitable distributions were fitted on the basis of the goodness of fit tests. Accordingly the Generalized Logic distribution provided a good fit to low flows in region one while the Generalized extreme value distribution fitted well in the region two. The method of probability weighted moments was considered the best parameter estimation procedure compared with the method of moments and maximum likelihood. And the regional and the station parameters of the basin were computed using EASY FIT computer program.

Regression analysis was also applied to develop regression models for prediction of mean annual minimum flow from ungauged catchments using catchment area.

Flow duration curves of different durations were established, and the flow duration curve indices for low flow of the basin i.e. Q60, Q70, Q80, Q90, Q95 and Q100, were quantified.

2. Recommendation

The hydrological recorded data for the analysis of this study were used from four gauging stations for 17 years record length. But it is recommended to have better number of gauging stations with more years of recorded length to increase the regions of the basin.

The most common methods used for the analysis of low flow and regionalization are the low flow frequency analysis, flow duration curve and the base flow index methods. But due to insufficient availability of data in the study area the base flow index method were not applied here. And it is recommended to use this method to know the mean ground water contribution of the regions and to classify the basin on the basis of the base flow index method.

The regionalization of homogeneous regions of the basin was identified by using the low flow statistics of the basin. But it is not sufficient for this case and it is preferable to assess the effect of regionalization by different topographic and climatic parameters like slope, elevation, soil type, geology and hydrological indexes of the basin which are not included here due to the scope of the research and several obstacles.

It is also recommended that similar work can be done for all basins in Ethiopia to facilitate the development in hydropower, irrigation, water supply and others that the country is in need at this stage of development.

Generally, it is recommended that further studies are required on the following points of this research that are not assessed due to the scarce of data and other obstacles:

- The effect of regionalization by different topographic and climatic parameters like slope, elevation, soil type, geology and hydrological indexes of the basin.
- To know the mean ground water contribution of the regions the basin also classify on the basis of the base flow index.
- To have better number of gauging stations with more years of recorded length to increase the regions of the basin.

REFERENCES

ABEBE, A. (2003) Ethiopian journal of water science and technology.

BEHAILU, M. (2009) Monitoring productivity of water in agriculture and integrated systems in tekeze river basin MSc. thesis report. Mekeele, Ethiopia.

CHANG, M. A. B., D.G., (1977) Estimation of low flows using water shed and climatic parameters. *Water Resources Research,,* 13 : 997-1001.

CHOW, V. T. M. A. D. R. A. M., L.W. (1988) Applied Hydrology. Singapore, McGRAW.

CONDIE, R., AND G.A, NIX. (1975) modelling of low flow frequency distributions and parameter estimation, presented at the international water resource symposium, water for arid lands, Tehran,Iran, Dec 8-9.

CUNNANE, C. (1989) "Statistical distributions for flood frequency analysis". *World Meteorological Organization Operational Hydrology,,* Report No.33, WMO-No.718, Genwva,Switzerland.

DAHMANI, A., MEDDI, M., (2009) Climate variability and its impact on water resources in the catchment area of Wadi Fekan Wilaya of Mascara (West Algeria). . *European journal of Scientific Research,* 36 (3), 458-472.

DALRYMPLE, T. (1960) Flood frequency ananlysis. Reston, VA.: US Geology survey on water supply paper 1543-A.

DEEPTI JOSHI, A. S.-H. (2013) Low flow frequency analysis. INRS-ETE.

FLIES, A. (2004) Methods for estimating flow duration curve and an application to the Burdekin river catchment.

FOSTER, H. A. (1934) "Duration curves". Amer. soc. Civil Engineers Trens. 99: 1213-1267.

GONFA (1996) Climate classification of Ethiopia. Addis Ababa, Ethiopia.

GREENWOOD, J. A. A. L. J. (1979) Probability weighted moments, definition and relation to parameters of several distributions expressable in invers form. water resource RES,. 15, 1049-1054.

HAAN, C. (1977) statistical methods in hydrology..Iowa state university press, Ames, IA.

HALL, D. (1990) screening of hydrological data: Test for stationarity relative consistency. Wageningen, the Netherlands. Internationl Institute for Land Reclamation and Improvement / ILRI.

HICKOX, K. P., AND G.O. WESSENAUER. (1933) "Application of duration curve to hydro-electric studies". . *Amer. Soc. Civil Engineers Trens. 98: 1276-1308.*

HIDLEY (1973) The definition of dry weather flow in river flow measurment. *J. Inst. water engng 27, 438-440.*

HOSKING, J. (1986) Theory of probability weighted moments, Res, Rc 12210. New york Heights, NY 10598: research division.

HOSKING, J. R. M. (1990) L-Moments; analysis and estimation of distribution using linear combination of orderd of statistics.

MATALAS, N. (1963) Probability distribution of low flows,. *statistical studies in hydrology. U.S. ,* Geological survey professional paper 434-A.

MKHANDI, E., AL. (2000) Flood frequency analysis of south Africa. II Identification of regional distributions. *Hydrology science Joiurnal. ,* 45(3), 449-466.

NEDECO (1997) Tekeze river basin master plan study report, Adiss Ababa, Ethiopia.

POFF, N. L. A. W., J.V., (1989) Implications of streamflow variability and predictability for lotic community structure: a regional analysis of streamflow pattern. *canadian Journal of Fisheries and Aquatic Sciences, ,* 46: 1805-1818.

RAO, A. A. (2000) flood frequency analysis. Florida: CRCP press LLC.

RIGGS, H. C. (1980) "Characteristics of low flow". *Journal of Hydraulics Div., Proc. ASCE 106(HY5): 717-731.*

SEARCY, J. K. (1959) "Flow duration curves, manual of hydrology: part 2, Lowflow techniques." U.S. Geol. Survey Water-supply Paper 1542-A.

SETEGN, A. (2015) assesment of surface water potential and demand in tekeze basin Msc. thesis report. AA: AAU.

SILESH (2005) In D .R.Ven Te chow, Applied Hydrology and Engieering Hydrology, principles and practices. *Engineering Hydrology*.

SINE, A. (2004) Regional flood frequency model for Blue Nile River Basin (Thesis report, Arbaminch University).

SINGH, K. P. (1947) Hydrology of 7-day 10-yr low flows. *journal of the Hydraulics Division, ,* HY12: 1753-1771.

SIVAPALAN, M. (2003) Prediction in ungauged basin: a grand challenge for theoretical hydrology. hydrological prosses, 17: 3163-3170.

SMAKHTIN, V. U. (2001) low flow hydrology. *review, journal of hydrology,* vol.240, 147-186.

STALL, J. B. (1971) "Derivation of base flow recession curves and parameters". *Water Resource Res. 7(2): 292-303.*

STEDINGER, J. R. (1993) frequency analysis of extreme events in "Hand book of hydrology", ed. D.R. Maidment. MC Grew Hill. Newyork, NY, pp 18.1-18.66. .

TALLAKSEN, L., AND VAN LANEN, H. (2004) hydrological drought processes and estimation of methods for stream flow and ground water.

TEGENU, Z. (2007) Low flow analysis and regionalization of Abbay river basin, Msc. thesis report to Arbaminch university. Arbaminch, Ethiopia.

TENNANT, D. L. (1976) Instream flow regimes for fish, wildlife, recreation and related environmental resources. Fisheries, 1 : 6-10.

THARME, R. E. (2003) Aglobal perspective on environmental flow assessment: emerging trends in the development and application of environmental flow methodologies for rivers. . *River Research and Application,,* 19 : 397-441.

TUCCI (1995) Flow regionalization in the upper Paraguay basin, Brazil. Hydrol. Sci. *J. 40 (4), 485-497*.

WILTSHIRE, S. (1986) Identification of homogeneous regions for flood frequency analysis, Journal of hydrology 84; 287-302.

WMO (2008) mannual on low flow estimation. NO. 1029 and pridiction operational hydrology report NO. 50.

ZAIDMAN, M. V. (2003) Flow duration frequency behaviour of Britsh Rivers based on annual minimuma data,. *journal of hydrology, 277,pp. 195-213*.

APPENDIX A: TYPICAL EXAMPLE ON FILLING MISSING DATA

Seasonal & daily mean

Interpolation

y = -2.6671x +3.36 R = 0.71

Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0.624	0.322	2.5012	1.7411	0.981	0.632	0.283	4.816	5.516	0.158	0.133	0.112
0.624	0.322	2.5012	1.7411	0.981	0.5945	0.208	3.063	2.438	0.158	0.133	0.092
0.624	0.322	2.5012	1.7411	0.981	0.557	0.133	3.063	1.9	0.133	0.133	0.092
0.688	0.322	2.5012	3.6586	4.816	3.0305	1.245	2.739	1.551	0.283	0.112	0.092
0.688	0.826	1.157	0.95599	0.755	0.9095	1.064	6.544	1.662	0.247	0.092	0.133
0.688	0.826	1.157	0.92249	0.688	0.9213	1.1545	2.739	1.551	0.215	0.092	0.133
0.688	0.826	1.157	0.83299	0.509	0.877	1.245	2.438	1.444	0.185	0.092	0.133
0.688	0.826	1.157	0.83299	0.509	1.4183	2.3275	2.438	1.245	0.158	0.092	0.133
0.688	0.755	1.3463	0.92767	0.509	1.9595	3.41	7.685	1.245	0.158	0.092	0.133
0.688	0.688	1.525	1.01702	0.509	2.776	5.043	3.144	1.245	0.158	0.158	0.133
0.688	0.565	1.8531	2.09779	2.3425	2.7028	3.063	1.982	0.901	0.133	0.158	0.133
0.688	0.565	1.8531	3.01454	4.176	3.978	3.78	2.0945	0.409	0.133	0.158	0.133
0.688	0.565	1.8531	2.54304	3.233	4.8885	6.544	1.9	0.247	0.112	0.158	0.133
0.688	0.565	1.8531	1.59754	1.342	1.4465	1.551	6.276	0.112	0.092	0.133	0.133
0.688	0.565	1.8531	1.30404	0.755	0.755	0.755	4.7545	0.036	0.06	0.133	0.133
0.624	0.565	1.8531	1.30404	0.755	0.7215	0.688	3.233	0.009	0.047	0.133	0.133
0.565	0.565	1.8531	1.30404	0.755	1.3275	1.9	5.276	0.092	0.036	0.133	0.133
0.409	0.565	1.8531	1.30404	0.755	0.519	0.283	18.359	0.215	0.019	0.133	0.133
0.363	0.565	1.662	1.2085	0.755	0.519	0.283	15.552	0.185	0.019	0.133	0.133
0.322	0.565	1.885	2.158	0.755	0.836	0.917	15.552	0.185	0.019	0.133	0.133
0.322	0.565	1.885	1.9	0.755	1.153	1.551	10.534	0.185	0.019	0.133	0.133
0.322	0.565	1.885	0.755	0.755	1.994	3.233	5.516	0.158	0.019	0.112	0.133
0.322	0.565	1.885	1.32	0.755	2.4655	4.176	4.816	0.158	0.019	0.112	0.133
0.322	0.565	1.885	1.32	0.755	0.868	0.981	3.592	0.185	0.019	0.112	0.133
0.322	0.565	1.885	1.32	0.755	1.9418	3.1285	8.3	0.185	0.247	0.112	0.133
0.322	0.509	1.9555	1.35525	0.755	3.0155	5.276	2.7305	0.185	0.185	0.112	0.133
0.322	0.509	2.026	1.3905	0.755	1.909	3.063	1.624	0.185	0.185	0.112	0.133
0.322	0.509	2.026	1.3905	0.755	1.747	2.739	1.462	0.185	0.133	0.112	0.133
0.322		2.026	1.3905	0.755	1.747	2.739	1.462	0.185	0.133	0.112	0.133
0.322		2.026	1.3905	0.755	0.133	2.739	5.043	0.185	0.133	0.112	0.133
0.322		2.026		0.755		2.739	7.1		0.133		0.133

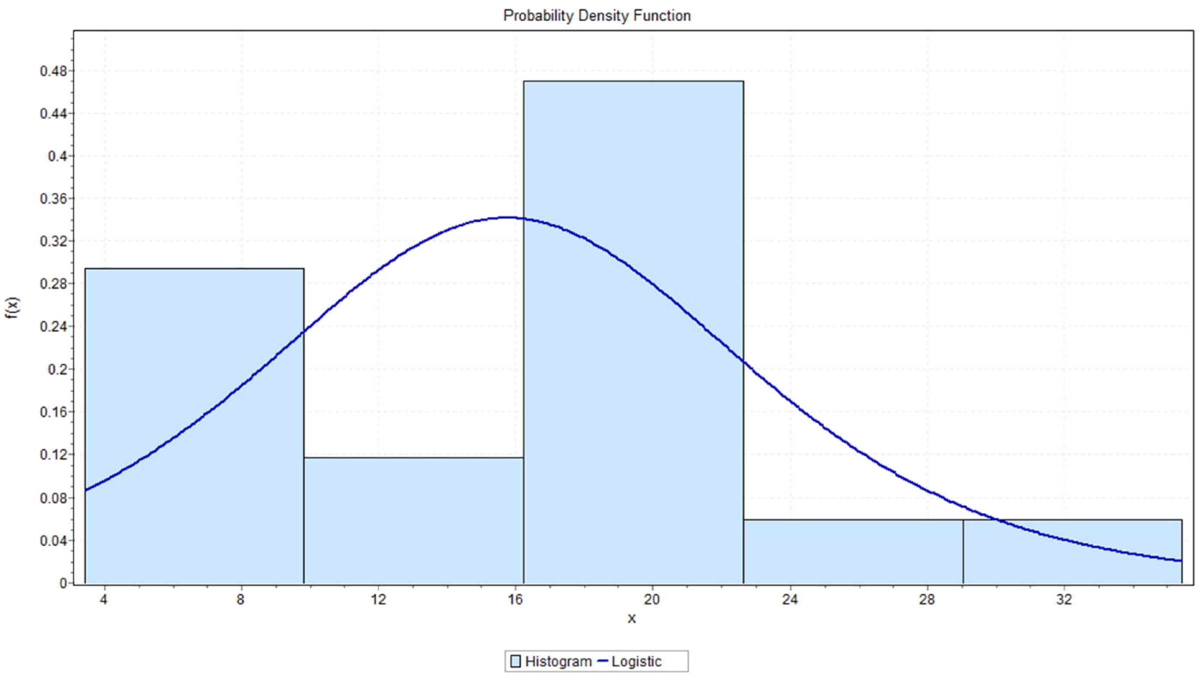
APPENDIX B: SUMMARY OF HYDROLOGICAL STATIONS

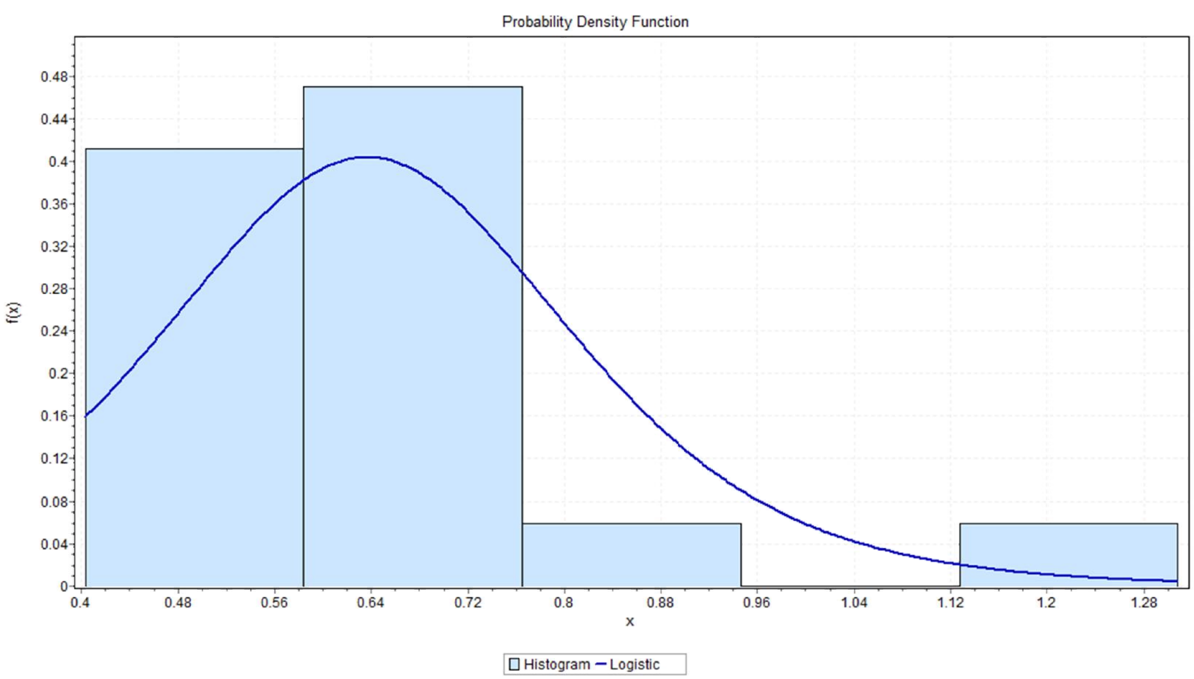
S. no.	River	Station	Duration	Number of year	Catchment area(km^2)
1	Gheba	Adikumsi	1998-2014	17	4342
2	Agulae	Agulae	1992-2014	23	692
3	Genfel	Wukro	1992-2014	23	481
4	Ilala	Mekelle	1992-2014	23	190
5	Gheba	Mekelle	1992-2014	23	2449
6	Suluh	Hawsien	1992-2014	23	399
7	Metere	Aynalem	1991-2006	16	69

APPENDIX C: SUMMARY OF HYDROLOGICAL DATA USED FOR THE ANALYSIS OF THE STUDY

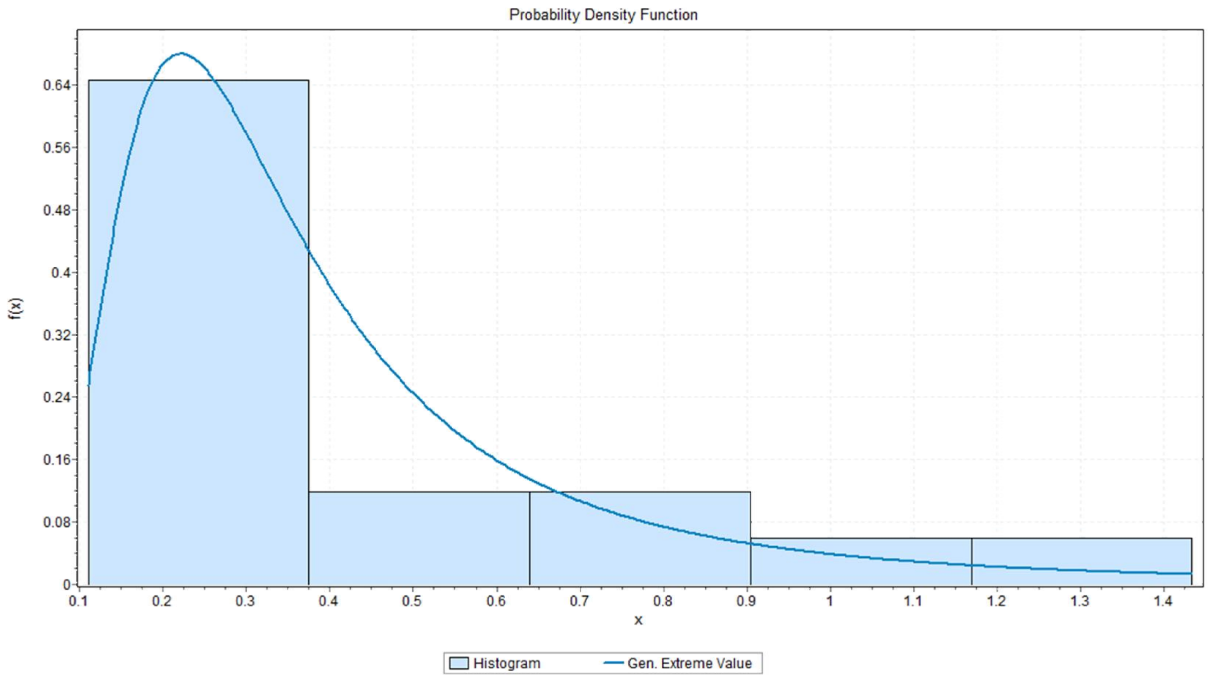
year	Gheba	Agulae	Ilala	Suluh
	Area = 4243km^2	Area = 692km^2	Area = 190km^2	Area = 399km^2
1998	0.9244	0.0227	0.0083	0.1143
1999	0.2836	0.181	0.0361	0.0267
2000	0.2319	0.1668	0.0226	0.0148
2001	0.0503	0.0534	0.0186	0.0345
2002	0.0949	0.1254	0.015411	0.0384
2003	0.5747	0.054	0.0465	0.0074
2004	0.5356	0.0499	0.005	0.0133
2005	0.4919	0.0499	0.005	0.0033
2006	0.654	0.0803	0.0066	0.0084
2007	0.6974	0.0959	0.0066	0.0065
2008	0.8411	0.0269	0.005	0.0138
2009	1.1601	0.0186	0.005	0.0145
2010	2.8445	0.0716	0.0058	0.0101
2011	3.5358	0.105	0.0074	0.0034
2012	4.2803	0.1401	0.0058	0.0033
2013	0.5259	0.0646	0.005	0.0014
2014	0.1427	0.0667	0.005	0.0046

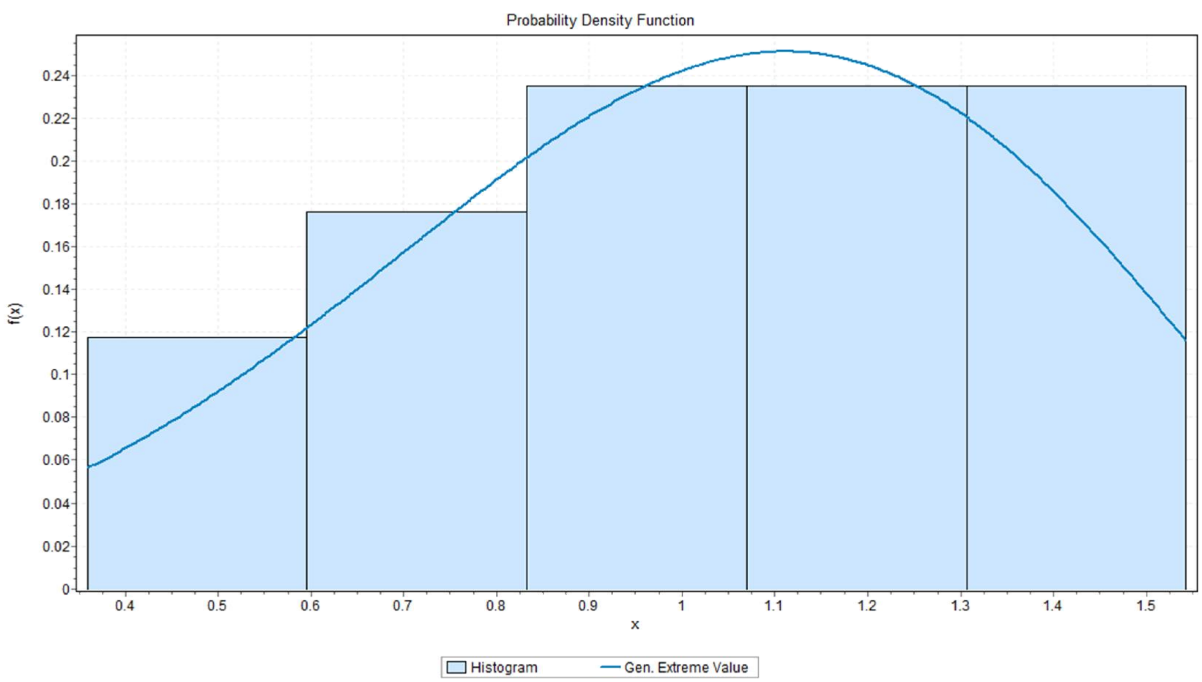
APPENDIX D: GRAPHICAL REPRESENTATION FOR PROBABILITY DENSITY FUNCTION OF REGION I (GENERALIZED LOGISTIC DISTRIBUTION) STATIONS GHEBA AND SULUH FROM EASY FIT SOFT WARE





APPENDIX E: GRAPHICAL REPRESENTATION FOR PROBABILITY DENSITY FUNCTION OF REGION II (GENERALIZED EXTREME VALUE DISTRIBUTION) STATIONS 121013 AND 121023 FROM EASY FIT SOFT WARE





APPENDIX F: L- MOMENT RATIO DIAGRAM FOR COMMON THEORETICAL DISTRIBUTIONS (HOSKING, 1990)

1. Uniform: $\tau_3 = 0$, $\tau_4 = 0$
2. Exponential: $\tau_3 = 1/3$, $\tau_4 = 1/6$
3. Gumbel (EV 1(2)): $\tau_3 = 0.1699$, $\tau_4 = 0.1504$
4. Logistic: $\tau_3 = 0$, $\tau_4 = 1/6$
5. Normal: $\tau_3 = 0$, $\tau_4 = 0.1226$
6. Generalized Pareto: $\tau_4 = \tau_3(1+5\tau3)/(5+ \tau_3)$ or
   $\tau_4 = 0.20196\ \tau_3 + 0.95924\ \tau^2{}_3 + 0.04061\ \tau^4{}_3$
7. Generalized logistic: $\tau_4 = (1 + 5\ \tau^2{}_3)/6$ or $0.16667 + 0.83333\ \tau^2{}_3$
8. Generalized Extreme value:

$$\tau_4 = 0.10701 + 0.11090\ \tau_3 + 0.84838\tau^2{}_3 - 0.06669\tau^3{}_3 + 0.00567\tau^4{}_3 - 0.04208\tau^5{}_3 + 0.03763\tau^6{}_3$$

9. Gamma and Pearson III:
   $\tau4 = 0.1224 + 0.30115\tau^2{}_3 + 0.95812\tau^4{}_3 - 0.\ 57488\tau^5{}_3 + 0.19383\tau^6{}_3$
10. Lognormal (two and three parameters)
   $\tau4 = 0.12282 + 0.77518\tau^2{}_3 + 0.12279\tau^4{}_3 - 0.\ 13638\tau^6{}_3 + 0.11368\ \tau^8{}_3$
11. Wakeby lower bound
   $\tau4 = -0.07347 + 0.14443\tau_3 + 1.03879\tau^2{}_3 - 0.14602\tau^3{}_3 + 0.03357\tau^4{}_3$

APPENDIX G: CRITICAL VALUES OF THE CHI SQUARED TEST (SOURCE: ANGEWANDTE, SPRINGER, 1997, 427-431)

TABLE 2–7 Critical Values of the Chi-Square Distribution

	α							
ν	0.995	0.990	0.975	0.95	0.050	0.025	0.010	0.005
1	0.0^4393	0.0^3157	0.0^3982	0.0^3393	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	5.99	7.38	9.21	10.60
3	0.0717	0.115	0.216	0.352	7.81	9.35	11.34	12.84
4	0.207	0.297	0.484	0.711	9.49	11.14	13.28	14.86
5	0.412	0.554	0.831	1.145	11.07	12.83	15.09	16.75
6	0.676	0.872	1.237	1.635	12.59	14.45	16.81	18.55
7	0.989	1.239	1.690	2.167	14.07	16.01	18.48	20.28
8	1.344	1.646	2.180	2.733	15.51	17.53	20.09	21.96
9	1.735	2.088	2.700	3.325	16.92	19.02	21.67	23.59
10	2.156	2.558	3.247	3.940	18.31	20.48	23.21	25.19
11	2.603	3.053	3.816	4.575	19.68	21.92	24.72	26.76
12	3.074	3.571	4.404	5.226	21.03	23.34	26.22	28.30
13	3.565	4.107	5.009	5.892	22.36	24.74	27.69	29.82
14	4.075	4.660	5.629	6.571	23.68	26.12	29.14	31.32
15	4.601	5.229	6.262	7.261	25.00	27.49	30.58	32.80
16	5.142	5.812	6.908	7.962	26.30	28.85	32.00	34.27
17	5.697	6.408	7.564	8.672	27.59	30.19	33.41	35.72
18	6.265	7.015	8.231	9.390	28.87	31.53	34.81	37.16
19	6.844	7.633	8.907	10.117	30.14	32.85	36.19	38.58
19	6.844	7.633	8.907	10.117	30.14	32.85	36.19	38.58
20	7.434	8.260	9.591	10.851	31.41	34.17	37.57	40.00
21	8.034	8.897	10.283	11.591	32.67	35.48	38.93	41.40
22	8.643	9.542	10.982	12.338	33.92	36.78	40.29	42.80
23	9.260	10.196	11.689	13.091	35.17	38.08	41.64	44.18
24	9.886	10.856	12.401	13.848	36.42	39.36	42.98	45.56
25	10.520	11.524	13.120	14.611	37.65	40.65	44.31	46.93
26	11.160	12.198	13.844	15.379	38.89	41.92	45.64	48.29
27	11.808	12.879	14.573	16.151	40.11	43.19	46.96	49.64
28	12.461	13.565	15.308	16.928	41.34	44.46	48.28	50.99
29	13.121	14.256	16.047	17.708	42.56	45.72	49.59	52.34
30	13.787	14.953	16.791	18.493	43.77	46.98	50.89	53.67

APPENDIX H: CRITICAL VALUES OF THE KOLMOGOROV TEST (SOURCE: ANGEWANDTE, SPRINGER, 1997, 427-431)

Critical values for the Kolmogorov-Smirnov Test for goodness of fit

For completely specified continuous distributions:

$1-\alpha$ / n	0.9	0.95	0.99
1	0.950	0.975	0.995
2	0.776	0.842	0.929
3	0.636	0.708	0.829
4	0.565	0.624	0.734
5	0.510	0.563	0.669
6	0.468	0.520	0.617
7	0.436	0.483	0.576
8	0.410	0.454	0.542
9	0.387	0.430	0.513
10	0.369	0.409	0.489
11	0.352	0.391	0.468
12	0.338	0.375	0.450
13	0.325	0.361	0.432
14	0.314	0.349	0.418
15	0.304	0.338	0.404
16	0.295	0.327	0.392
17	0.286	0.318	0.381
18	0.279	0.309	0.371
19	0.271	0.301	0.361
20	0.265	0.294	0.352
21	0.259	0.287	0.344
22	0.253	0.281	0.337
23	0.247	0.275	0.330
24	0.242	0.269	0.323
25	0.238	0.264	0.317
26	0.233	0.259	0.311
27	0.229	0.254	0.305
28	0.225	0.250	0.300
29	0.221	0.246	0.295
30	0.218	0.242	0.290
31	0.214	0.238	0.285
32	0.211	0.234	0.281
33	0.208	0.231	0.277
34	0.205	0.227	0.273
35	0.202	0.224	0.269
> 35	$\frac{1.224}{\sqrt{n}}$	$\frac{1.358}{\sqrt{n}}$	$\frac{1.628}{\sqrt{n}}$

for $n > 35$: $\frac{\sqrt{-0.5 \cdot \ln(\frac{\alpha}{2})}}{\sqrt{n}}$

For a partially specified normal disitribution: (Lilliefors Modification)

n / $1-\alpha$	5	8	10	12	15	20	25	30	40
0.90	0.319	0.265	0.241	0.222	0.201	0.176	0.159	0.146	0.128
0.95	0.343	0.288	0.262	0.242	0.219	0.190	0.173	0.159	0.139
0.99	0.397	0.333	0.304	0.281	0.254	0.223	0.201	0.185	0.162

$n > 30$:

$1-\alpha = 0.90$: $\frac{0.819}{\sqrt{n} - 0.01 + \frac{0.83}{\sqrt{n}}}$ $\quad 1-\alpha = 0.95$: $\frac{0.895}{\sqrt{n} - 0.01 + \frac{0.83}{\sqrt{n}}}$

$1-\alpha = 0.99$: $\frac{1.035}{\sqrt{n} - 0.01 + \frac{0.83}{\sqrt{n}}}$

For a partially specified Poisson distribution:

n	$1-\alpha$	$\overline{X} \leq 1$	$1 < \overline{X} \leq 2$	$2 < \overline{X} \leq 3$	$3 < \overline{X} \leq 5$	$5 < \overline{X} \leq 10$
6	0.90	0.202	0.214	0.226	0.237	0.254
	0.95	0.234	0.242	0.254	0.265	0.281
	0.99	0.290	0.300	0.310	0.324	0.334
12	0.90	0.152	0.166	0.172	0.179	0.185
	0.95	0.180	0.188	0.194	0.199	0.206
	0.99	0.223	0.234	0.236	0.243	0.249
20	0.90	0.120	0.132	0.140	0.144	0.149
	0.95	0.141	0.151	0.156	0.160	0.165
	0.99	0.176	0.185	0.188	0.195	0.197
30	0.90	0.100	0.112	0.116	0.120	0.124
	0.95	0.116	0.125	0.129	0.134	0.140
	0.99	0.149	0.154	0.158	0.160	0.168
40	0.90	0.087	0.097	0.102	0.106	0.110
	0.95	0,101	0,108	0,113	0,118	0,122
	0.99	0.130	0.135	0.137	0.143	0.146
> 40	0.90	$\frac{0.55}{\sqrt{n}}$	$\frac{0.61}{\sqrt{n}}$	$\frac{0.65}{\sqrt{n}}$	$\frac{0.67}{\sqrt{n}}$	$\frac{0.70}{\sqrt{n}}$
	0.95	$\frac{0.64}{\sqrt{n}}$	$\frac{0.69}{\sqrt{n}}$	$\frac{0.72}{\sqrt{n}}$	$\frac{0.75}{\sqrt{n}}$	$\frac{0.77}{\sqrt{n}}$
	0.99	$\frac{0.82}{\sqrt{n}}$	$\frac{0.86}{\sqrt{n}}$	$\frac{0.87}{\sqrt{n}}$	$\frac{0.90}{\sqrt{n}}$	$\frac{0.93}{\sqrt{n}}$

(from: L. Sachs, Angewandte Statistik, Springer 1997, 427 - 431)