JIMMA UNIVERSITY

JIMMA INSTITUTE OF TECHNOLOGY

SCHOOL OF GRADUATE STUDIES

FACULTY OF CIVIL AND ENVIRONMENTAL ENGINEERING
HYDROLOGY AND HYDRAULIC ENGINEERING CHAIR

REGIONAL LOW FLOW ANALYSIS: THE CASE OF UPPER OMO GIBE RIVER BASIN, ETHIOPIA

A Thesis Submitted to School of Graduate Studies of Jimma University, Jimma institute of technology in partial fulfillment of the requirements for the degree of Master in Hydraulic Engineering.

By DARARA DABTARA

October, 2019
Jimma, Ethiopia

JIMMA UNIVERSITY

JIMMA INSTITUTE OF TECHNOLOGY

SCHOOL OF GRADUATE STUDIES

FACULTY OF CIVIL AND ENVIRONMENTAL ENGINEERING
HYDROLOGY AND HYDRAULIC ENGINEERING CHAIR

REGIONAL LOW FLOW ANALYSIS: THE CASE OF UPPER OMO GIBE RIVER BASIN, ETHIOPIA

A Thesis Submitted to School of Graduate Studies of Jimma University, Jimma institute of technology in partial fulfillment of the requirements for the degree of Master in Hydraulic Engineering.

By DARARA DABTARA
Main Advisor: Zeinu Ahmed (PhD)
Co- Advisor: Tolera Abdisa (MSc)

ACKNOWLEDGEMENTS

First and foremost, I want to thanks the Almighty God for giving me the chance, strength and courage to continue my study. Secondly I would like to express my gratitude appreciation to my main advisor Zeinu Ahmed (PhD) and co-advisor Tolera Abdisa (MSc) for their excellent guidance, encouragement and enthusiastic support.

Lastly my sincere appreciation also goes to instructor Chala Hailu and JIT instructors for sharing their experience, and materials.

DECLARETION

I, the undersigned declare that the thesis entitled as “Regional low flow Analysis: the Case of upper Omo gibe River Basin, Ethiopia” is my own original work and has not been presented for any degree in Jimma institute of technology and any other university or institute. All the sources of materials used in this study have been duly acknowledged.

Name	Signature	Date
Darara Dabtara	____________	___________

This thesis has been submitted for examination to my approval as Jimma university institute of technology supervisor.

Advisors: Name	Signature	Date
Zeinu Ahmed (PhD)	____________	___________
Tolera Abdisa (MSc.)	____________	___________

APPROVAL

The thesis entitled “Regional low flow Analysis: the Case of upper Omo gibe River Basin, Ethiopia” was submitted by Darara Dabtara Bayana was approved and accepted as a Partial Fulfillment of the Requirements for the Degree of Masters of Science in Hydraulic Engineering at Jimma Institute of Technology.

Name	Signature	Date
Advisor: Dr. Zeinu Ahmed	___________	__________
Co-Advisor: Tolera Abdisa (MSc.)	___________	__________

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

Name	Signature	Date
Chairman Fayera Gudu (MSc.)	_________	__________
External Examiner Dr.	_________	_________
Internal Examiner	_________	_________

ABSTRUCT

Assessment of low flow of a river in magnitude as well as in frequency is crucial for planning and design of water resource projects such as planning of water supplies, water quality management issuing, hydropower project, irrigation project and the impact prolonged drought on aquatic ecosystems. The objectives of this study were to quantify the characteristics of low flows in rivers of upper Omo gibe, and to estimate the magnitude, frequency, flow regionalization, and to fit best fit statistical distribution. Twenty hydrometric stations in the upper Omo gibe which have more than 16 years of complete data were selected for the current low flow study. L-moment based approach and geographical proximity location of the station were applied for regional frequency analysis of annual minimum 7-day low flows and four separate homogeneous regions were identified.

The most frequently used distributions in the analysis of hydrologic extreme variables are: Generalized Extreme Value (GEV), Lognormal (LN), Generalized Pareto distribution (GP) and generalized logistic (GLO). For selection of best-fit distributions L-MRD, XLSTAT Statistical computer software, EASY FIT Statistical computer software and Matlab were employed. XLSTAT Statistical computer software were used to select methods of parameters estimation for at-site low flow Frequency Analysis and Matlab software were selected for parameter estimation depend on RMSE. Using the goodness-of- fit tests (Chi test and Kolmogorov-Smirnov test), for most of the stations, the selected probability distributions were GEV and GP distributions. Method of Moments (MOM) was selected for all distribution used in the study for estimation of parameter. The result of ($Z^{dist)}$) *indicated that Generalized Extreme Value and Generalized Pareto distributions are most appropriate probability distribution for Region-1, Region-2 and Region-3, Region-4 respectively. Finally, the growth curves developed using the estimated at site and regional quantiles for all stations and identified regions.*

***Key words**: FDC; low flow; L-moments; quantile; regionalization.*

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONOMY/ AND ABBREVIATION

AFDC	Annual flow duration curve
AMF	Annual minimum flow
BFI	Base flow index
C-CV	Combined coefficient of variation
Cdf	Cumulative distribution function
DEM	Digital elevation model
Di	Discordance
DMC	Double mass curve
FFA	Flow frequency analysis
FDC	Flow duration curve
GEV	General extreme value
GIS	Geographical information system
GLO	Generalized logistics
GOF	Goodness of fit
GPA	Generalized Pareto
IHR	Identification of homogeneous region
IWMI	International water management institute
L-CKu	L- Co-efficient kurtosis
L-Cs	L-Coefficient of skewness
L-CV	L- Co-efficient variation
L-MRD	L-moment ratio diagram
LLO	Log-logistic
LFFA	Low flow frequency analysis
MAM	Mean Annual minimum
MWIE	Ministry of water irrigation and electricity
OGRB	Omo gibe river basin
PDF	Probability distribution function
P-P	Probability plot
PWM	Probability weighted moment

Q-Q	Quantile -quantile plot
RLFFA	Regional low flow frequency analysis
R-1, R-2, R-3, R-4	Region-1, 2, 3, 4
RMSE	Root mean square error
SPSS	Statistical software package for social science
SMADA	Storm water management aided and design analysis
ZFI	Zero flow index day ratio
K	Shape parameter
µ (mu)	Location parameter
δ (sigma)	Scale parameter
Ʈ 2, Ʈ3, Ʈ4	Coefficient of (variation, skewness and kurtosis)
7Q10	seven -day consecutive low flow with ten -year return period

1. INTRODUCTION

Stream flows naturally vary both during a year and from year to year. In the face of these variabilities, water management decisions can only be made with predicted estimates of stream flows. More importantly, design and planning of water resource projects requires the assessment of the probability of extreme hydrological events such as low or high flows. A low flow condition can be defined as a period during which the average stream flow is a minimum for the year. The characterization and estimation of low flows are important issues in hydrologic studies such as the determination of minimum downstream flow requirement of hydropower station, estimation of available water supply for municipal and industrial uses, water quality management, determination of potential capacity for effluent dilution, assessing the impact of low flows on aquatic ecosystem, and in general for environment impact assessment studies (Zadeh, 2012), (Mélanie , et al., 2017),and (Hamadi, et al., 2013).

Low flow values are expressed in terms of their averaging period (for example, a 4-day average flow or a 7-day average flow) and their recurrence frequency (generally once in ‘n’ years) (EPA, 2018). Now, low flows and droughts are more and more recognized as risk situations due to the huge consequences of water shortage. Furthermore, climate change constitutes a new threat, even though the uncertainty about the evolution of low-flows remains high. (Winlar & Khin, 2014).

Numerous indices can be obtained from LFFC. Among the most commonly used ones are the quintiles of the lowest mean discharge over a continuous period of 7- days corresponding to a recurrence interval of *n*-years. From the perspective of LFFA, the available flow records are generally insufficient for reliable quantification of extreme low flow events and as a result frequency analysis relies on different types of theoretical distribution functions to extrapolate beyond the limits of observed values and to ameliorate the accuracy of low-flow estimation. The true probability distributions of low flows are unknown and the practical problem is to identify a reasonable functional distribution and estimate its parameters. In flow frequency analysis, different distributions are used to determine magnitude of extreme flow events (Winlar & Khin, 2014).

According to (Smakhtin, 2001) a universally accepted distribution for low flow analysis is unlikely to be identified. A number of commercial software packages available for selecting best fit distribution are; (easyfit, Matlab, xlstat, Minitab, statistics, smada) (Joshi & André , 2013).

Planning and designing of hydraulic structures need adequate hydrological information of the specific area or the region at large. This includes observed stream flows at many sites. But many sites do not have adequate number of gauging station or that are recently established. There may not be gauging stations in the catchments at all. In such cases, transfer of required information from gauged sites to un-gauged sites becomes very essential. If an interest site has no record of flow, regionalization must be forwarded (Cunnane, 1985) and (Benyahya, et al., 2011).

Analysis of low flow magnitude is needed in three situational cases 1) for water resource development, 2) to execute the daily decisions on managing the water resource development during operational phases, and 3) when there is a current need to decide on the operations based on the estimations of the future stream flows (Kidist, et al., 2018).

There are many methods in modern-day hydrology that could be used to calculate hydrological characteristics at gauged and ungauged sites. Some of the methods used are hydrological regionalization (growth curve method), drainage area ratio methods, and regional regression equation. The said parameters can be estimated by different software such as easyfit, xlstat, Matlab, sigma magic and so on and ArcGIS is used to delineate homogeneous region (Ramakar & Smakhtin, 2008).

Unless quite management is done on stream flow, every project fails fully or partially in functional as well as destruction of aquatic animals, environmental pollution will face urban area and people would be faced by disease due to excess water disposed to urban area (McMahon , et al., 1982).

Generally, Ethiopia has excess river basin and less management system, a quite analysis for that resources were needed. Among those river basin Omo Gibe is the popular river basin contain flow of almost 16.6 BM^3 in volume. By using minimum river flow data that were recorded for many year and analyzing the data and using statistical distribution methods estimation of minimum flow was possible and used to overcome problem related to allocation of discharge.

1.2 Statement of the problem

Even if many projects have been developed and proposed on upper Omo gibe river basin, there is no research done on low flow analysis before, which shows magnitude of minimum flow available within a year on it. Because of the extra demand on the river may create undesirable environmental as well as upstream and downstream conflict. Since understanding the low flow regimes and evaluating the magnitudes for incorporating in water resources management is vital for the countries like Ethiopia (Kidist, et al., 2018).

Due to huge demand of land and favorable land for agriculture available around the site, development of irrigation project that will be depend on firm discharge will be needed to increase agricultural productivity throughout the year by the farmers and by the government at all level to alleviate poverty. Most of all, the availability and quality of information is not adequate so further development of any water resource project within the basin are difficult and unreliable unless the low flow characteristics is well known (Tatek Worku, 2015).

Moreover, this study aimed to identify watershed variables best explaining the variation in the hydrological regime, with a special focus on low flows that may be susceptible to management policies for developing and securing water resources in dry periods. In general variability of river discharge in Ethiopia to be under difficult condition due to unknown estimation magnitude of minimum flow for instance reduction or shifting follow of electric power in the country requires a great attention toward the study of low flow (Gebrehiwot, et al., 2011) and (Tseday, 2007). However, many annual flow series are too short to allow for a reliable estimation of extreme events or there is no flow record available at the site of interest. These are typical of the case in Upper Omo-Gibe sub-basin, where many of rivers are un-gauged and some of gauged stations in sub-basin face problems, such as shortness of records and incomplete records, among others.

Therefore, this study would be used to overcome problem related to minimum flow and characterize low flow of upper Omo gibe river basin and provide the necessary information about the low flow of the basin.

1.3 Objective

1.3.1 General objective

The general objective of this research is to determine a regional low flow of upper Omo Gibe river basin.

1.3.2 Specific objective

1. To estimate the magnitude corresponding to return period of at site gauged upper Omo Gibe river basin.
2. To regionalize and develop low flow duration curves of upper Omo gibe river basin.
3. To determine the best fit distribution for upper Omo gibe river basin.

1.4 Research questions

The study of the research will answer the following questions:

1. What is the magnitude and frequency of low flow of upper Omo Gibe river basin?
2. What seems regionalization and low flow duration curve of upper Omo gibe river basin?
3. Which distribution is best fit for upper Omo Gibe river basin?

1.5 Significance of the Study

The importance of low flow analysis is to achieve environmental sustainability during consumption of water, water management during construction and after construction of water resource projects, to avoid unfunctionality of project on that basin, to allocate discharge for different purpose at the downstream and to show trends of low flow of upper Omo gibe river to government body or private company that has a power to develop any project on that area.

1.6 Scope of the study

Generally, the study address issues related to estimation of low flow and its magnitude, regionalization of site under study and estimation of best fit distribution for upper Omo gibe river basin that might take place depending on the river basin hydrology. And also the thesis contain the procedure of regionalization as well as how to determine statistical distribution by L-MRD for specific site depend on stream flow data.

The study is limited mainly on regionalization of stream flow data on the upper Omo gibe River Basin, Ethiopia.

2. LITERATURE REVIEW

Low flows have been investigated only in the recent past few decades. This includes low flow frequency analysis, base flow separation, recession analysis, flow spell analysis, and low flow estimation at gaged and ungauged sites. Although there is a high interest in low flow studies, the mass of literature has still been relatively less compared with flood or precipitation studies. The characteristics and estimation of low flows are important issues in many hydrologic studies and in general for environmental impact assessment studies. Such studies often require that the hydrologists estimate the magnitude, frequency, duration, and spells of low flow events as different aspects of low flow analysis (Zadeh, 2012).

Information on the characteristics of low flows for streams and rivers is important for planning, design and operation of water-related projects and water resource systems. Such information is used in designing wastewater treatment and storage facilities to ensure that releases do not exceed the assimilative capacity of receiving waterways, reservoir storage design for multi-purpose systems and the allocation of water for various purposes such as industrial, agricultural, domestic and in-stream ecological needs (WMO, 2009).

Low-flow frequency analysis and flow-duration curves are the two most commonly used analytical tools to assess the low-flow characteristics of streams. Both approaches typically require at-site continuous stream flow data, unless regional approaches are used to estimate at site characteristics. Other characteristics that are sometimes useful include the amount of time or frequency for which flows might be below a certain threshold during a season and the volume of water or deficit that might arise during the period in which flows are below a threshold. Low flows within a year or season may result from different mechanisms forcing the hydrological response. It is important to understand the processes producing the low flows, as these may determine the analytical approaches taken to analyze their characteristics and results (WMO, 2009).

Anthropogenic intervention can greatly alter the natural low-flow regime. For example, increased extraction from surface water for irrigation may occur during periods of prolonged absence of rainfall, resulting in artificially suppressed flow values, compared with what naturally would have occurred. Significant extraction of groundwater for agricultural, industrial and human uses can reduce water-table levels and result in reduced stream flow. A variety of other anthropogenic interventions can occur within a basin and should be known prior to proceeding with analyses of data. Such operations may cause increases or decreases in flow rates (WMO, 2009).

For a large number of applications, the minimum annual flow of 7 days duration and 10-year recurrence or an estimation of the flow of 95% of permanence are sufficient. These quantiles can be estimated directly from the empirical distribution, without resorting to a theoretical model adjustment. On the other hand, really severe droughts, such as occurred exceptionally during the rainy season of year, where the estimated return period was over 100 years, require the application of theoretical statistical models that present a good fit to the hydrological variables (Eloy, 2018).

There are still many details to be discovered and clarified for the distribution function properties of minima. Increasing recognition of the importance of minimum flows for ecosystem viability, economic sustainability and as a climate change alert, makes the study of small extreme minima more and more important (Eloy, 2018).

River low flows have always been an important parameter in hydrological studies. Low flow conditions are mainly due to local climate, soils, topography, vegetation, as well as by lakes and swamps. Human activities can also influence low flow conditions such as irrigation, water withdrawals and climate change. These conditions and factors all need to be considered during the planning, design, construction and the maintenance of different hydraulic structures and water resource systems. River low flows can also impact fish habitat and instream water toxicity by reducing the dilution capacity and increasing water temperatures (Benyahya, et al., 2011).

Navigation and power supply sectors can also be affected by low flows. Furthermore, as pressures on rivers become more important during low flows, some conflicts between the different water users can arise, especially between instream water use and water abstraction demand. Low flows can have different meanings depending on the definitions of authors. In this study, low flows are considered as the lowest discharge values observed in a river. The index chosen to characterize low flows is MAM7 which stands for mean annual minimum flow on a 7-day average basis (Grandry & Gailliez, 2013).

Low-flow calculation and frequency analysis are easy to handle for long-time gauged catchments. For ungauged catchments, however, low-flow index has to be inferred using neighboring gauged catchment data (Smakhtin, 2001).

The other method uses time-series simulation, regional prediction curves, or spatial interpolation. Regional regression consists in delineating hydrologically homogeneous regions based on catchment characteristics and developing, for each region, a regression model relating the low flow index to these characteristics. As this approach is based on physical parameters of the basin, it allows a better understanding of low flows from a physical point of view (Laaha & Bl¨oschl, 2006), (Vezza, et al., 2010) and (Tsakiris, et al., 2011). However, only a few of them included the return period in their analysis (Grandry & Gailliez, 2013).

Low flow regime is tightly dependent on the catchment hydrogeological feature and a detailed surface and groundwater catchment analysis is necessary for an accurate characterization. However, on a practical perspective, although scientifically proven, statistical analysis is often applied to derive indices to characterize low flow regimes and as a measure for low flows. Low flow indices can be easily evaluated at gauged sites from observed stream flow time series, but their reliability can be affected by poor and not accurate stream flow data. Another approach to estimate low flow statistics in ungauged sites is the regional analysis, widely used since long time and in different disciplines. It is the most widely used technique in flow estimation in ungauged sites or where few data are available (Rossi & Caporali , 2010).

2.1 Low flow frequency analysis

Daily low flows of the stations would be extracted from stream flow data. From those data, the minimum annual discharge could be calculated for a 1-day, 7-day and 14-day duration. Many statistical distributions were used to represent low flows. However, the 3 parameter distribution were the most often used distribution for low flows and was chosen for fitting the annual minimum discharge in the present study. The 3 parameter cumulative distribution function (*Cdf*) is referred by equation 2.1 (Benyahya, et al., 2011):

$$F(x) = 1 - e^{-((x-t)/\eta)^{\beta}};\ x \geq t \qquad 2.1$$

Where x represents discharge, t is a threshold parameter, $\eta > 0$ is a scale parameter, and β is a shape parameter.

Many methods have been used for the estimation of the distribution’s parameters, and two commonly used methods are the method of moments and maximum likelihood method and selection of those method of parameter estimation was done by xlstat software.

With the distribution's parameters and the *cdf*, the low flow estimates *QT* were then calculated for different recurrence intervals. For instance the 3 parameter Weibull quantiles *QT* are obtained from equation 2.2 (Benyahya, et al., 2011):

$$Q_T = \eta\ (-ln\ (1-F(x))^{\ 1/\ \beta)} + t \qquad 2.2$$

Where the relation between *F(x)* and recurrence interval, *T*, is given by equation 2.3

$$F(x) = 1/T \qquad 2.3$$

The coefficient of determination (R^2) and the root mean square error (RMSE) are the best criterion for evaluating the good fit of a regression equation. (Benyahya, et al., 2011).

Unlike the flow duration curve which shows the proportion of time during which a flow is exceeded, a low flow frequency curve shows the proportion of years when a flow is exceeded or equivalently the average interval in years (return period or recurrence interval) that the stream flow falls below a given discharge (Zadeh, 2012).

2.2 Regional flow frequency analysis

Regionalization of stream flow characteristics is based on the premise that catchments with similar geology, topography, climate, vegetation, and soils would have similar stream flow responses (Rossi & Caporali , 2010). Regional flow frequency analysis (RFFA) involves two major steps: (1) Grouping of sites into homogeneous regions, and (2) Regional estimation of flow quantiles at the site of interest. The performance of any regional estimation method strongly depends on the grouping of sites into homogeneous regions and geographically contiguous regions have been used for a long time in hydrology (Malekinezhad, et al., 2011).

In order to carry out, the regional low flow analysis initially the homogeneous group of stations will be identified and categorized by using the L-MRD and coefficient of variation (C-C) test. The station year method was used for estimating the standardized long-term quantiles for developing the regional frequency curve and pull the standardized low flow values as one station for each homogenous group. The best probability distribution for each homogenous group will be fitted and the long-term standardized quantiles will estimated for various return periods. The regional growth curve was established as the relationship between the standardized quantiles and return periods for each discharge. Hence the estimated standardized quantile will used to compute the normal low flow quantiles for both gaged and ungauged stations using equation 2.4.

$$X_T = Q_T/\bar{Q} \qquad 2.4$$

Where; $\bar{Q}$ - is the mean annual low flow (m^3/s) is the index flow, Q_T - is the quantile (m^3/s) function of fitted distribution at site I and X_T - regional quantile which can be obtained from regional growth curve; this defines the frequency distribution common to all the sites in a homogenous region. (Kidist, et al., 2018).

Regional flow frequency analysis dealing the following subheadings that constitute the general procedure for the analysis: Data screening, Delineation of homogeneous regions, Regional homogeneity test, Selection and estimation of regional frequency distribution, Estimation of flow magnitudes; and regionalization (Zadeh, 2012).

In order to estimate low flows of different return periods for ungauged areas, it is necessary to establish a relationship between the annual 7-day low flow and the pertinent physiographic and climatic characteristics at gauged catchments. The established relationship can be used to obtain the estimation for the ungauged catchments which are located together with gauged catchments in a homogeneous region. Since the climatic factors can be considered to be identical throughout the study region, catchment size is an important factor in determining the magnitude of discharge. The index flood procedure used in the previous section assumes that, for a homogeneous region, the frequency distributions for all sites are identical except a site-specific scale factor. Note that the heterogeneity can be seen from the area-*Q*7 plot and can be fixed by log-transformation. Therefore, the relationship between the mean 7-day low-flow *Q*7 (m^3/s) and the catchment area *A* (km^2) is estimated as:

$$Q_U = Q_g \, (A_U/A_G)^{a} \qquad 2.5$$

If a target study site exist between upstream gauged and downstream gauged site the exponent is calculated as equation 2.6 and ungauged discharge was calculated by equation 2.5. (Nancy , et al., 2007).

$$a = \log(Qgu/Qgd)/\log(Agu/Agd) \qquad 2.6$$

where Q_U mean annual flow (volume units) for the ungauged site, Q_g mean annual flow (volume units) for the gauged site, Qgu/Qgd gauged discharge at upstream and downstream, Agu/Agd area at upstream and area at downstream and Au and Ag are the areas of the ungauged and gauged catchments, respectively and value of ‘a’ is b/n 0.5 and 1 and this method is best if the ungauged sites is near to the gauged site (Gordon, 2004) and (Pamela, 1992).

2.3 Flow duration analysis

Flow-duration curves shows the percentage of days that the flow of a stream is greater than or equal to given amounts over a given period of record. Some of the most common uses of flow-duration curves are in computing hydroelectric power potential for firm power and secondary power, water-supply and irrigation planning and other water-quality management problems. It will be illustrating flow characteristics from flood to low flows for a basin. The shape of the curve reflecting physiography and climatology condition of that basin. If the stream flow data are stationary, the derived flow-duration curve should provide the long-term exceedance probabilities for the entire range of flows, which is a useful planning tool (Guastard, 2009) and (Nancy , et al., 2007).

A flow duration curve (FDC) is one of the most informative methods of displaying the complete range of river discharges from low flows to flood events (Nancy , et al., 2007). It displays the relationship between stream flow and the percentage (probability) of time it is exceeded (Zadeh, 2012).

2.3.1 Flow duration curve construction

In general, a FDC is constructed by reassembling the flow time series values in decreasing order of magnitude, assigning flow values to class intervals and counting the number of occurrences (time steps) within each class interval. Cumulative class frequencies are then calculated and expressed as a percentage of the total number of time steps in the record period. Finally, all ranked flows are plotted against their rank which is again expressed as a percentage of the total number of time steps in the record (Zadeh, 2012) and (Nancy , et al., 2007). The most convenient way of constructing a FDC is using the log-normal probability plot. This allows FDCs in some cases to be linearized and low- and high-flow ends of the curve to be more clearly displayed (Smakhtin, 2001).

2.3.2 Application FDC

Various low-flow indices may be estimated from this part of the FDC. The flows within the range of 70–99% time exceedance are usually most widely used as design low flows. Some common example indices are: one- or *n*-day discharges exceeded 75, 90, and 95% of the time, i.e. Q75 (7), Q75 (10), Q90 (1), Q95 (1), Q95 (10). Some less conventional indices include the percentage of time that 25% average flow is exceeded (Smakhtin, 2001). Flow duration curve analysis, which can be used as general indicators of hydrologic conditions (i.e., wet versus dry and severe). Flow duration curve intervals were grouped into broad categories, or zones which provide additional insight about conditions and patterns associated with the impairment. The duration curve is dividing into five zones, representing high flows (0-10%), moist conditions (10-40%), mid - range flows (40-60%), dry conditions (60-90%), and low flows (90-100%) (Richard, et al., 2016). FDC has many significant application including stream water quality study and for estimation of optimal release schedule from reservoirs, in design of flow diversions (Zadeh, 2012).

2.3.3 Interpretation and indices

The slope of a FDC reflects the catchment's response to precipitation. If groundwater contributions are significant, the slope of the curve at the lower end tends to be flattened whereas a steep curve indicates low base flows. Streams draining the same geologic formations will tend to have similar FDC at the low flow end (Nancy , et al., 2007). Flow duration curves can provide a number of indices to characterize the stream for classification and regionalization purposes. Of most interest for low flow studies is the low flow section of FDCs, which may be arbitrarily defined, for example, as part of the curve with flows below the median which corresponds to the discharge equaled or exceeded 50% of time or Q50 to Q99 (Zadeh, 2012).

2.3.4 Base flow index

The base-flow index is the total volume of base flow divided by the total volume of runoff for a period. The base flow component of river flow is commonly expressed as a proportion of the total river flow, termed the Base Flow Index (BFI). A catchment's with BFI approaching 1 is highly dominated by base flow and considered as highly permeable, while a catchments with BFI approaching 0 receives little base flow contribution and considered as impermeable. From the main hydrograph, base flow hydrograph and base flow index is calculated. The obtained result of BFI is used to determine the catchment behavior at all and to compute the regression analysis between catchment and its discharge in order to predict the volume of ungauged catchment in a provided region. Generally, station having almost the same BFI have the same hydrological characteristics and geographical behavior to precipitation response (Tatek Worku, 2015) and (WMO, 2009).There are a significant relationship between the shape of flow duration curve and the BFI coefficient (Cheng, et al., 2012) and also it is used as a measure of the base flow characteristics of catchments and site having nearly the same BFI considered as homogeneous and it provides a systematic way of assessing the proportion of base flow in the total runoff of a catchment (Adane & Gerd, 2006). BFI was found to be a good indicator of the effects of geology on low-flows and for that reason is widely used in many regional low-flow studies (Smakhtin, 2001).

2.4 Statistical tools used in this study

Statistical tools or models are characterizations of the real-world system. A model is physical or mathematical description of a physical system, including the interaction with its outside world, which can be used to simulate the effects of changes in the system itself. A watershed model simulates hydrologic processes in a more holistic approach compared to many other models which primarily focus on individual processes or multiple processes at relatively small-or field-scale without full incorporation of a watershed.

Deterministic models are mathematical models in which outcomes are obtained through known relationships among states and events. Stochastic models will have most, if not all, of their inputs or parameters represented by statistical distributions which determine a range of outputs. Even though most models are deterministic in nature, stochastic models provide two important advantages. First, their conceptually simple framework makes it possible to describe heterogeneity when there are limited spatial or temporal details. Second, they provide decision makers with the ability to determine uncertainty-associated with prediction. Empirical models consist of functions used to approximate or fit available data.

2.4.1 Selection of statistical tools

Each model type serves a purpose, and a particular model type may not categorically be considered more appropriate than others in all situations. Choice of a suitable model structure relies heavily on the function that the model needs to serve. There are various criteria which can be used for choosing the right statistical models for a specific problem. These criteria are always project dependent, since every project has its own specific requirements and needs. Further, some criteria are also user-depended. Among the various project-dependent selection criteria, there are: Required model outputs important to the project and therefore to be estimated by the model and as well Hydrologic processes that need to be modeled to estimate the desired outputs adequately and Availability of input data. The selection of model was by considering the above criteria with inclusive of availability of data, level of application, purpose, required accuracy, space and time scale, catchment area, simplicity, previous trends (studies) in the surrounding area & Ethiopia as a whole. Considering all the criteria's set above data driven model xlstat, Matlab, SPSS, EASYFIT and physical based models SMADA model is adopted for this study.

2.4.1.1 Matlab software package

Is type of software program used to compute the best fit distribution for hydrological time series data and compute its statistical parameter and used to calculate RMSE of the distribution selected. In this study, Matlab2018a were used to find the best-fit distribution and its estimation parameters. During the importation of data into Matlab the data must be saved in terms of numeric matrix form rather than table form during curve fitting and estimation of the distribution for the provided hydrological data. In addition to this Matlab is used to check the heterogeneity of station in a homogeneous region by providing the discordance between the stations depend on critical value of discordance provided by (Hosking and Wallis, 1997).

2.4.1.2 Easyfit software

Easyfit is a data analysis and simulation software which enables us to fit and simulate statistical distributions with sample data, choose the best model, and use the obtained result of analysis to take better decisions. This software can function as a stand-alone windows application or as an add-in for Excel spread sheet. Easy Fit combines the classical statistical analysis methods and innovative data analysis techniques, making it a tool of choice for anyone dealing with probability data. Prominent features of this program are: Ability to test performed operations, integrated help system, Interactive graphs, Goodness of fit tests, Easy to use interface (http://www.mathwave.com/help/easyfit).

2.4.1.3 Sigma magic software

These methods are similar to EASYFIT it is preferred especially where there is little or no information about the base distribution pattern in data and the need to find the best distribution type. In order to determine whether the distribution model could fit the data properly, chi-square goodness-of-fit tests were used. In this study sigma magic Statistical Software Package, trial version 11.3 was used to find the best-fit distribution and its estimation parameters.

2.4.1.4 XLSTAT

The other statistical software used to estimate the missed data and best fit distribution for data to be used in this study by applying different goodness of fit test such as Kolmogorov simirnov and chi-square test method. The other use of this software package is used to determine which parameter estimation method is used with the selected distribution parent. Therefore in this study xlstat2018 trial version were used for different computation

2.5 Previous low flow study in Ethiopia

Low flow is an important part of the natural flow regime of rivers where the water resource planning and design consider its spatial and temporal variability. The spatiotemporal variability of a stream flow due to the complex interaction of catchment attributes and rainfall induce complexity in hydrology. Researchers have been trying to address this complexity with a number of approaches; river flow regime is one of them. The flow regime can be quantified by means of hydrological indices characterizing five components: magnitude, frequency, duration, timing, and rate of change of flow. Similarly, the writer of this research aimed to understand the flow variability of Ethiopian Rivers using the observed daily flow data from 208 gauging stations in the country. With this process, the Hierarchical Ward Clustering method was implemented to group the streams into three flow regimes (1) ephemeral, (2) intermittent, and (3) perennial. The mean flow per unit catchment area and Base flow index (BFI) show an incremental trend with ephemeral, intermittent and perennial streams. Whereas the number of mean zero flow days ratio (ZFI) and coefficient of variation (CV) show a decreasing trend with ephemeral to perennial flow regimes (Belete, et al., 2015).

To understand the causes and take remedial action for the sustainable utilization of the low flows the dynamics in low flows in a river system should be evaluated, which could include quantifying the trend of low flow quantiles, and developing regional curves (ungauged catchments) is a very important approach for proper management of the water resources. The doer of this paper concluded the results of study have indicate mainly a decrease in low flows values for the selected stations in the Blue Nile Basin. The L-Moment ratio diagram provides a practical method to identify the underlying distribution for a given station. The use of the L-Moment ratio diagram was very convenient that one can compare the fit of several distributions using which are superior to conventional moment ratios because L-moments are less biased than ordinary moments. Land use management plans should recognize that woodland, dense wet forest and savannah grassland can promote higher low flows, while grazing land diminishes low flows (Gebrehiwot, et al., 2011) and (Kidist, et al., 2018).

The distributions used, which represent four of the most frequently used distributions in the analysis of hydrologic extreme variables are: (I) Generalized Extreme Value (GEV), (ii)Lognormal (LN), (iii) Lognormal with three parameters (LN3) and (iv) Log-Pearson three(LP3). According to this research the best best-fit distributions and methods of parameters estimation for at-site Flood Frequency Analysis General extreme value, lognormal and EASY FIT Statistical computer software were employed. According to this study upper Omo-Gibe sub-basin was grouped in to two homogeneous regions (Tamiru, 2009).

L-moments have been used for parameter estimation in both cases. Finally, the growth curves developed using the estimated at site and regional quantiles for all stations indicate that as recurrence interval increases the magnitude of flood increase which shows amount of infiltration decreases that push surrounding to drought and Estimation of low flow for ungaged sites (below 50% of exceedance) can be obtained using regional regression method and subsequently Flow Duration Curve for each regions (Tsedey, 2007) and (Tamiru, 2009).

In Ethiopia the population is rapidly expanding and a consequence the landscape is rapidly changing. The hydrological effects of the changing landscape on river (low) flows have not been well documented and therefore the amount of water available in the future might be over optimistic. The researcher found a statistically significant decreasing trend (P < 0.00001) of low flow in the Gilgel Abay. From 1980’s to 1990’s the low flow decreased by 25% and from 1990’s to 2000’s the low flow was reduced by 46%. The deterministic analysis with the Parameter Efficient Distributed (PED) model supported the statistical findings and indicated that in the middle of the nineteen nineties, after irrigation projects and eucalyptus plantations increased greatly, the low flows decreased more rapidly (Temesgen , et al., 2014).

3 MATERIAL AND METHODS

3.1 Description of Study Area

Omo-Gibe River basin is almost 79,000 km^2 in area and is situated in the southwest of Ethiopia, between 4°30’ and 9°30’N and 35° and 38°E with an average altitude of 2800masl. It flows from the northern highlands through the lowland zone to discharge into Lake Turkana at the Ethiopia/ Kenya border in the south and is fed along its course by some important tributaries. The key characteristic of the Omo-Gibe river basin is its complex topography. Thus the basin is divided sharply into highlands in the northern half of the area and lowlands in the southern half. This division is reflected in almost all other aspects of the basin. The northern highlands are deeply dissected with steep slopes and drained by the Gibe and Gojeb systems which merge to form the Omo in a deeply entrenched gorge which slices into the highlands. The Gibe River is called Omo River in its lower reach, south and south westwards from its confluence with the Gojeb River. The northern part of the catchment has a number of tributaries emanating from the north-east, of which the largest are the Walga and Wabe rivers. Another two tributaries are the Tunjo and Gilgel Gibe rivers which drain mainly cultivated lands with less permeable soils in the south-west. The Gojeb River is a major tributary to the Omo River, draining the uplands that have been less intensively cultivated than the other parts of the basin. To the south of the Gojeb River the catchments of the Sherma, Guma and Denchiya rivers, which are tapering streams that join the Omo at the northern end of the flood plain.

Except in the driest years, these rivers usually maintain some flow throughout the year (Chaemiso, et al., 2016). Generally, this study would be cover almost the upper part of OGRB and it cover around 44,837.7 km^2 out of 79,000km^2. Detail of the study area was located on Figure 3.1.

3.1.1 Climate and hydrology

The climate of Omo-Gibe River basin varies from a hot arid climate in the southern part of the flood plain to a tropical humid one in the highlands that include the extreme north and northwestern part of the Basin. Intermediate between these extremes and for the greatest part of the basin the climate is tropical sub-humid. Rainfall of Omo-Gibe river basin varies from 2000 mm per annum to 300mm. The amount of rainfall decreases throughout the Omo-Gibe catchments with a decrease in elevation. The mean annual temperature in Omo-Gibe basin varies from 16^{0}C in the highlands of the north to over 30^{0}C in the lowlands of the south.

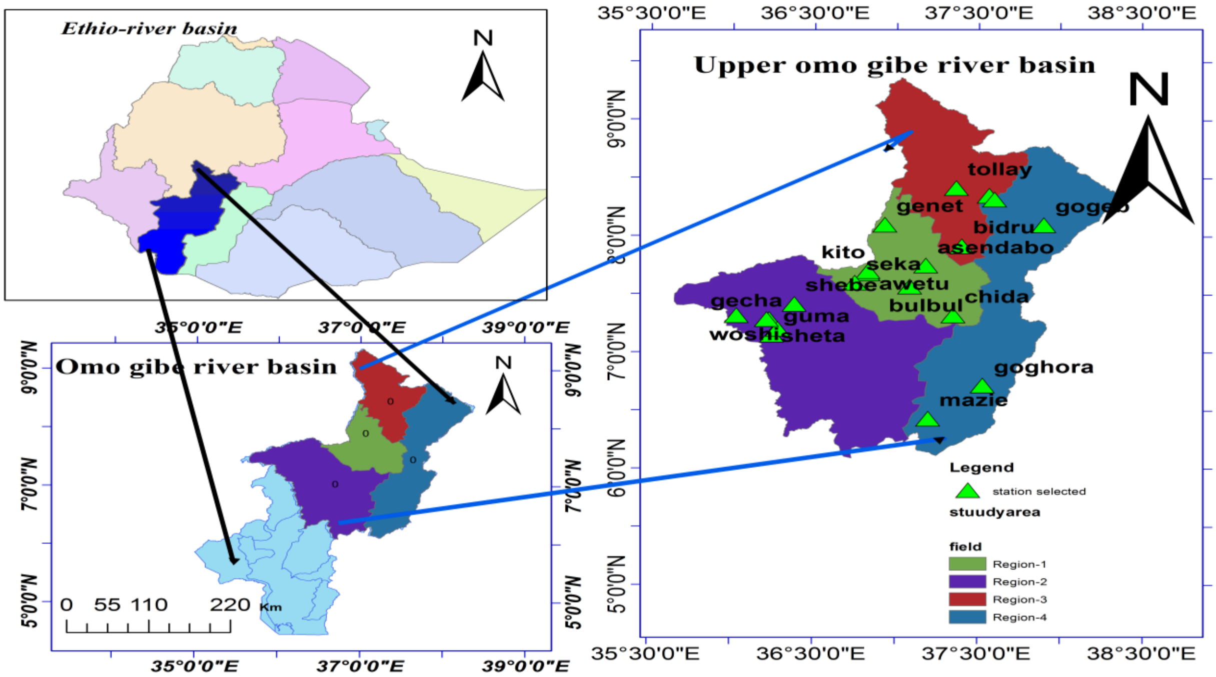


Figure 3.1 Location map of Omo gibe and upper Omo gibe river basin.

3.2 Software

For proper execution of this study, Software were used based on the capability to work on meeting the predetermined objectives of the study. ArcGIS10.4.1 Software was used to generate the study area map representing geographical location of gauging stations and delineate hydrologically homogeneous regions. Easy Fit 5.6 Statistical Software, Matlab2018a, XLSTAT 2018 and l-moment ratio diagram were used to select the best fit probability distribution with its method of parameter estimations and a goodness of fit tests for each station. As well as SMADA online package was used to compute frequency and return period of low flow. XLSTAT2018 with Microsoft Excel were used in filling of missed hydrological data and estimate the method used for parameter estimation for selected statistical distribution function for hydrological data used in low flow frequency analysis.

3.3 Sources, collection and analysis of data

Low Flow frequency analysis primarily use observed annual minimum flow data at gauging stations to estimate low flow magnitude. Hydrological and DEM (digital elevation model) data of Omo gibe River Basin were collected from Ministry of Water Irrigation and Electricity (MWIE), and department of hydrology and GIS. DEM data was employed as basic input for delineation and specifying the location of the gauging stations in the basin. The site characteristics of stations for this study includes the code of the stations, the name of the river and their gauging sites, the locations (latitude and longitude) and catchment area (km^2). Table 3.1 indicate the site characteristics of stations used in this study.

Defining a clear and efficient methodology is crucial tools for the quality of the findings of the study. The procedures of data analysis in this study includes from the preliminary screening of data to develop a regional low flow frequency curve depending on 7day MAM flow series of data. The selection of 7-day AMF was depend on the following evidences;

- The 7-day low-flow (7Q) is the most widely used index in the USA, UK and many other countries. The minimum 7-day average flow is known as “dry weather flow” or “mean annual 7-day minimum flow” (MAM7) (Smakhtin, 2001). The 7-day period covered by MAM eliminates the day-to-day variations of the river flow.
- Previous studies, as reviewed by Smakhtin (2001), have shown that, compared with 1-day low flow, an analysis based on a time series of 7-day average flows is less sensitive to measurement errors and outliers.

Practically, the 7-day low flow better represents the drought conditions of concern and can be used more effectively in water management. The 30-day low flow is considered to be a more suitable index in arid and semiarid climate regions in that it avoids too many zeros which may appear in minimum 1-day or 7-day low flow series (Yongkin, et al., 2010)

Screening of data was carried out to check the gross errors and make sure the continuity of data. After relevant data that was used for the regional low flow analysis identified from the study basin, checking of data for its quality was performed. Identifying homogeneous region was done to decide on which sub-basins can be grouped together which might have similar low flow in nature. This was performed based on the geographical location of the station, L-moment ratio diagram and site characteristics of stations. The regional frequency distribution by the average L-moment ratios and a goodness-of-fit test with help of Easy Fit Software was then used to confirm how well the selected distribution fit the data in the region. Estimation of the frequency distribution is then designed to compute the low flow quantiles for certain return periods at ungauged sites derived from the regional growth curve. In general, to achieve the regional low flow frequency analysis of this study, detail of the procedures were described on figure 3.3.

3.3.1 Screening of Data

It is the first task in which unwanted observation from the data series as well as the sites from the analysis can be filtered. It is used to check the data are appropriate for performing the regional flow frequency analysis. In this study, stream flow data were used from gauging stations in the upper Omo gibe River Basin. 20 stations were decided according to the guideline for FFA which allows a minimum of 10 years' historical flow data and no consecutive gap exist. In the study area, there are about 36 gauging stations, out of these only 22 gauging stations were selected for the proper RLFFA. The selected stations by themselves have no fully recorded data; they have a number of years of record having missing data that needs to be filled before analysis. Out of 22 selected gauging stations almost 20 stations have less missed data and two stations have less than 10 years of record data that is less than the guideline for FFA exist (Hosking & Wallis, 1997) and (Gadefa, 2009). Accordingly, 20 gauging station which satisfied the minimum record length were selected. The minimum and maximum length of the at-site AMF records were 16 and 27 years respectively. For all the stations listed on Table 3.1 the AMF data were selected and later subjected for investigative data analysis in order to choose representative stations for the study area.

3.3.2 Filling of Missed data

When undertaking analysis of stream flow data from gauges; where observations are made, it is often to find times where no observations are recorded at one or more gauges. The continuity of the record may be broken with missing data due to many reasons such as the absence of recorder, carelessness of the observer, break or failure of instruments. Therefore, it is often necessary to estimate these missing data. The missing data can be estimated by using the data of the neighboring station. There are different methods used for filling the missed flow data records of a given gauging station. For this study, any missing data were filled by the method of linear regression by multiple imputation. Reference variables were the same type i.e. flow vs. flow. Simple linear regression has been applied to fill missing stream flow values using nearby flow gauging station observations. The equation for linear regression is given as:

$$y = ax + b \qquad 3.1$$

Where x, y are daily stream flows (m^3/sec) and a, b are constants. In this study, regression with correlated stations by scatter plot was checked and used to obtain missing daily flow data, by comparing nearby station by deriving a common equation using a scatter plot.

Table 3.1 Site characteristics of the station

Station code	River name	Location of station (near)	Latitude coordinate	Longitude coordinate	Area (km^2)	Record period	Record length
091012	Gojeb	Shebe	7^{0}25’0”N	36^{0}23’0” E	3577.0	1990-2016	27
092005	Dincha	At Bonga	7°12’0”N	36^{0}17’0” E	443.8	1990-2016	27
092002	Gecha	Nr Bonga	7^{0}17’0”N	36^{0}13’0” E	175	1990-2016	27
092004	Guma	Andaracha	7^{0}9’0”N	36^{0}15’0” E	231.3	1998-2013	16
091025	Woshi	Nr Dimbira	7^{0}19’0”N	36^{0}2’0” E	47.5	1999-2014	16
200000	Gojeb	Chida	7^{0}19’60’’N	37^{0}21’0’’E	234	2001-2016	16
092003	Sheta	At Bonga	7^{0}17’0”N	36^{0}14’0” E	190.6	1990-2016	27
091008	Gilgel gibe	Asendabo	7^{0}45’0”N	37^{0}11’0” E	2966.	1990-2016	27
091007	Gogeb	Nr. Endeber	8^{0}5’60’’N	37^{0}54’0’’E	109.	1990-2010	21
091014	Gilgel gibe	Limmu genet	8^{0}6’0”N	36^{0}56’0”E	533	1990-2010	21
091017	Gibe	Nr Seka	7^{0}36’0”N	36^{0}45’0”E	280.4	1990-2010	21
091023	Kito	Nr Jimma	7^{0}42”0”N	36^{0}50’0”E	85	1990-2010	21
091032	Bulbul	Nr serbo	7^{0}34’0”N	37^{0}5’0” E	526	1990-2010	21
091019	Bidru	Sokoru	7^{0}55’0”N	37^{0}24’0”E	41	1990-2010	21
061015	Gilgel gibe	At Abelti	8^{0}13’48’’N	37^{0}34’48’’E	15746.	1990-2010	21
0111111	Gilgel Gibe	Nr Tollay	8^{0}25’12’’N	37^{0}22’12’’E	6580.8	2000-2015	16
091010	WALGA	Nr. wolkite	8^{0}19’48’’N	37^{0}36’0E	1792	1990-2005	16
092008	Mazie	Nr. Morka	6^{0}26’0”N	37^{0}12’0”E	937	1990-2010	21
092013	Goghora	Nr. Dana1	6^{0}43’0” N	37^{0}32’0”E	266	1990-2010	21
091024	Awetu	At Jimma	7^{0}41’0” N	36^{0}50’0” E	72.0	1990-2010	21

The model performance can be good if the correlation coefficient (R) between 0.6 and 1 (Ketsela , et al., 2017). They discussed that method was selected and commonly used due to the following reasons: It is the most widely used method when compared to other methods for large data, estimation of significant missing observations as accurate as possible, it is applied by creating a correlation with the nearby station. As a result, linear regression analysis is used to fill the missing instantaneous daily flow data with satisfactory correlation coefficients.

3.4 Data quality control

Some errors may exist in the stream flow observation that were collected, such as misplaced decimal numbers, very huge unrealistic numbers and negative flow records in some cases. Performing observation quality before using it for necessary purposes is a vital step. The following approaches were considered to check stream flow data quality.

3.4.1 Test for randomness and independence

By principle, it is known that FFA is carried out when at-site data are independent and identically distributed conditions satisfied (Hosking & Wallis, 1997).This provides that the extreme events might appear randomly and all might have the same frequency distribution. The requirement of RLFFA is that the AMF at different stations in a homogeneous region should be spatially independent. However, (Hosking & Wallis, 1997) noted that a small amount of serial dependence in annual data series has little effect on the quality of quantile estimates. It is assumed that all the low magnitudes in the AM series are mutually independent in the statistical sense. In this study, the correlation coefficient was applied to verify the independence of the data of the selected hydrological stations. According to (Dahmen & Hall, 1990), the lag-1 serial correction coefficient, R, defined as follows:

$$R=\sum ((Xi-X_m)*(Xi+1-X_m))/\sum (Xi-X_m)^2 \qquad 3.2$$

Where Xi is an observation, Xi+1 is the following observation. After computing R, the test hypothesis is that H0: R= zero (that there is no correlation between two consecutive observations) against the alternative hypothesis, H1: R< or > 0. In other case it is simple to use Xlstat software package to compute randomness.

The computed R by equation 3.2 should be between the upper and lower confidence interval that were written on equation 3.3 and 3.4.

$$UCL = (-1+1.96(N-2)^{0.5})/\ (N-1) \qquad 3.3$$

$$LCL = (-1-1.96(N-2)^{0.5})/\ (N-1) \qquad 3.4$$

To accept the hypothesis H0: the value of R should fall between the UCL and LCL. Applying this condition to the time series, the condition: LCL (R) <R< UCL (R) is satisfied for the all stations. The data are independent and there is no persistence in the time series. The summarized result of the test for mean annual minimum of 7 day flow series for instance for Shebe station -0.540 <0.382<0.44 and the other stations were given on Table 3.2.and the results show that the mean annual minimum of 7 day flow series for all stations were independent.

3.4.2 Test for homogeneity

By using Xlstat software package 2018 trial version it is simple to test the homogeneity of the time series hydrological data by comparing the null hypothesis with confidence interval that could allowed as standard by the software. If the calculated P value is greater than the null hypothesis it is considered as homogeneous. According to this study the selected station is homogeneous after some computation is made by Xlstat. And RAINBOW software package were also independent software used to test homogeneity of hydrological data. In both cases the hydrological data of station were homogeneous and detail of the description is represented on table 3.3.

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

Station Name	UCL	R	LCL	Station Name	UCL	R	LCL
Gojeb Shebe	0.440	0.382	-0.54	Gibe Nr Seka	0.377	-0.019	-0.477
Dincha Nr Bonga	0.440	0.249	-0.54	Kito Nr Jimma	0.377	0.186	-0.477
Sheta At Bonga	0.440	0.390	-0.54	Bulbul Nr serbo	0.377	-0.041	-0.477
Guma Andaracha	0.432	0.431	-0.556	Bidru Nr Sokoru	0.377	.365	-0.477
Woshi Nr Dimbira	0.422	0.359	-0.556	Gibe At Abelti	0.377	0.368	-0.477
Gojeb Chida	0.470	0.400	-0.56	Gibe Nr Tollay	0.422	0.266	-0.556
Gecha At Bonga	0.440	0.396	-0.540	Walga Nr. wolkite	0.422	0.181	-0.556
Gibe Nr Asendabo	0.440	0.116	-0.540	Mazie Nr. Morka	0.377	-0.238	-0.477
Gogeb Nr. Endeber	0.377	0.211	-0.477	Goghora Nr. Dana1	0.454	0.446	-0.477
Gibe Nr Limmu	0.384	0.377	-0.477	Awetu At Jimma	0.377	0.093	-0.477

As indicated on table 3.3 the null hypothesis H_0 is greater than confidence interval (5%) taken as standard by Xlstat software and the computed R^2 by RAIN BOW software is acceptable for all station depend on chi-square goodness of fit test.

Table 3.3 Homogeneity test by XLSTAT and RAINBOW

River name	H_0 by XLSTAT	R^2 by RAINBOW	River name	H_0 by XLSTAT	R^2 by RAINBOW
Gojeb Shebe	6.31%.	0.85	Gibe Nr Seka	29.69%.	0.89
Gecha nr Bonga	5.16%	0.88	Kito Nr Jimma	20.58%.	84.66
Dincha At Bonga	11.27%	0.92	Bulbul Nr serbo	47.24%.	0.95
Guma Andaracha	42.91%.	0.97	Bidru Sokoru	12.12%	0.84
Woshi Nr Dimbira	9.43%	0.91	Gilgel At Abelti	5.39%.	0.93
Gojeb Chida	10.23%.	0.98	Gibe Nr Tollay	12.13%.	0.87
Sheta At Bonga	27.06%.	0.92	Walga Nr. wolkite	13.91%.	99.82
Gibe Asendabo	15.38%.	0.92	Mazie Nr. Morka	49.02%.	87.30
Gogeb Nr. Endeber	5.52%.	0.92	Goghora Nr. Dana1	7.85%.	0.97
Gibe Limmu genet	64.76%.	0.86	Awetu At Jimma	7.57%.	0.80

3.4.3 Test for consistency and stationarity

A time series hydrological data is relatively consistent if the periodic data are proportional to an appropriate simultaneous time series. Therefore in order to obtain the accurate result checking of consistency and stationarity of data is the obligation of the researcher. This can be done through double mass curve (DMC) methods. In this study the observed DMC from the whole station indicate that the correlation coefficient R^2 indicate 0.96 to 0.99 those indicate consistency and stationarity of the site. Thus, this tests were adopted to check stream flow observations stationarity and consistency (McCuen, 1998), (Searcy & Hardison, 1966). For instance figure 3.3 indicate the consistency of Asendabo catchment thus R^2 (0.99> 0.6) that is acceptable and for the other station is indicated on appendix A.

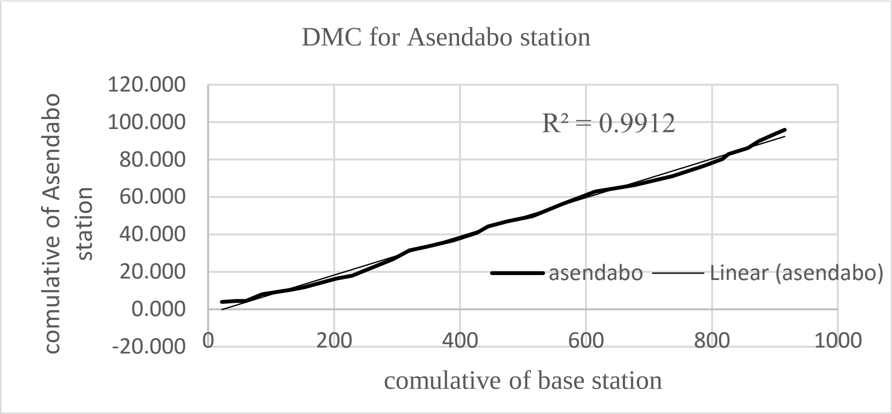


Figure 3.2 Double mass curve for Asendabo station

Note: Base station is a group of nearby station containing five to ten station and its correlation is best if R^2 is greater than 0.87 (Goyal, 2016).

3.4.4 Check for outliers of the data series

An outlier is an observation that deviates a lot from the bulk of the data. This may be due to errors in data collection, misplaced decimal points, very high flow records during dry months and or low flow record during rainy months or due to natural causes. For statistical tests of outlying observation, it is generally recommended that a low significant level such as 1% is used and that significance level greater than 5% should not be common practice (Ketsela, et al., 2017) . However, to minimize or avoid the effect of outliers in this study L-Moments were efficient parameter estimation technique was used (Hosking & Wallis, 1997).

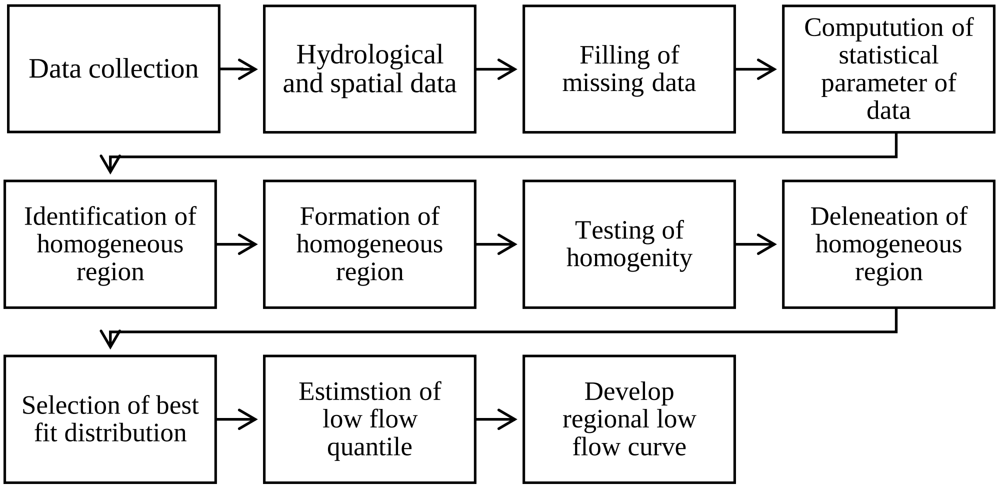


Figure 3.3 Procedural steps of methods

3.5 Regionalization of upper Omo Gibe River Basin

In this study, the L-moment approach of regionalization for geographically proximity was applied depending on homogeneity of the stations. The statistical values have been checked for the stations whether they can be classified under one or more regions. Flow statistics of upper Omo Gibe River Basin were computed using L-moment methods. Due to the fact that such methods can give a balanced estimation of sample parameters and cannot be easily influenced by the presence of outliers (Rao, 2008).

3.5.1 Identification of homogeneous regions

Identification of homogeneous regions (IHR) is the significant step in regional flow frequency analysis (Amalina, et al., 2016). To identify HRs the specification of variables characterizing this similarity has been made. The IHR is usually the most difficult stage and requires the greatest amount of personal judgment. Consequently, the clustering of sites into homogeneous regions was carried out by applying the hierarchical geographic regionalization technique with the method of L-moments as a guideline for regionalization. The stream gauging stations were grouped into geographically continuous sites such that the response of streams to physiographic variables should be similar. DEM size of 30mx30m of the Basin were used to identify site characteristics by ArcGIS software. This enables stream flow records to be transferred from gauged basins to ungauged basins within a region. To IHR one should follow the following procedure;

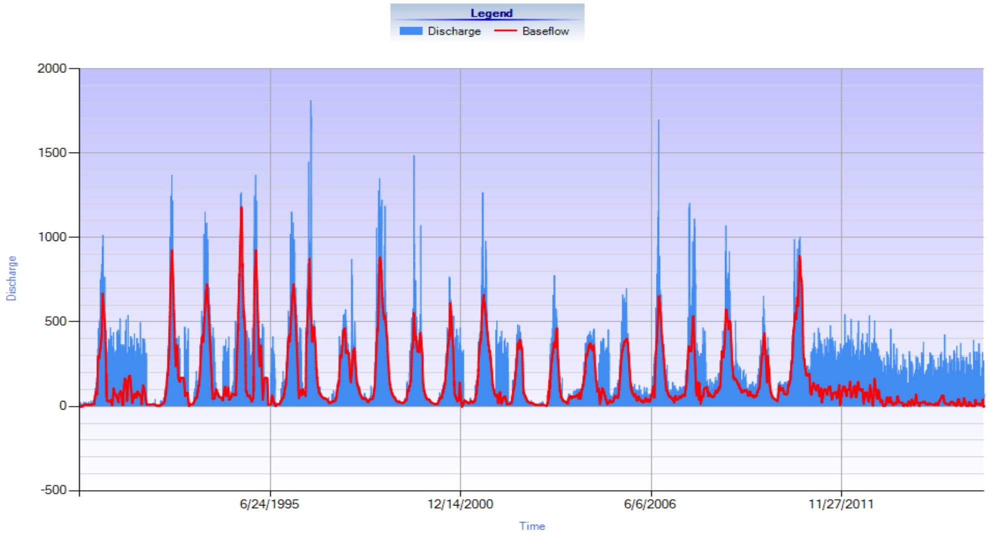
3.5.1.1 Site characteristics

Preliminary IHRs of stations into a certain category is achieved by looking at stations site characteristics. The following site characteristics were used as a preliminary IHR; latitude and longitude, AMF, station area and altitude of the flow gauging station. Then stations having nearly same kind of site characteristics are clustered on the same region (WMO, 2018).

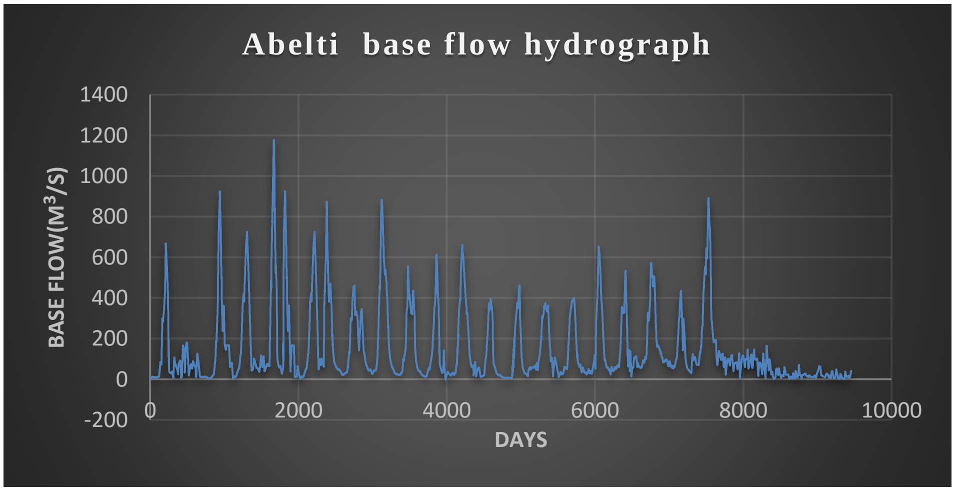
3.5.1.2 Base flow and base flow index

The base-flow index is the total volume of base flow divided by the total volume of runoff for a period and this index is the best parameter that shows the geology of the site (Wahl, 1995). In order to determine the contribution from overland flow in a watershed to the streams in the watershed it is necessary to separate out the base flow from stream gage data. BFI+ Software package were also used to separate base flow from total hydrograph. For all stations base flow hydrograph is separated and they are attached in appendix- B. For illustration typical separated base flow hydrograph of Abelti station presented on figure 3.4 (a) and (b).

A catchment's with BFI approaching 1 is highly dominated by base flow, while a catchments with BFI approaching 0 receives little base flow contribution. From the main hydrograph, base flow hydrograph and base flow index is calculated. The obtained BFI result is used as one of the input parameter for the regression analysis for ungauged site and the result of BFI were presented on table 3.4.



(A)Hydrograph of Abelti station with base flow



(B) Base flow separated for Abelti station

Figure 3.4 Hydrograph of Abelti station with its base flow

Table 3.4 Base flow index for all station

Station	BFI	Station	BFI	Station	BFI	Station	BFI
Abelti	0.67	Bulbul	0.53	Dincha	0.56	Guma	0.4
Bidru	0.64	Limmu	0.46	Woshi	0.51	Walga	0.34
Tollay	0.67	Seka	0.56	Gecha	0.65	Mazie	0.34
Awetu	0.43	Kito	0.72	Sheta	0.49	Gogeb	0.39
Asendabo	0.7	Chida	0.73	Shebe	0.72	Goghora	0.37

Those stations have a BFI of 0.34-0.73 which means 34% to 73% of the flow contributes from delayed storage. And as a number of base flow index increased the contribution to base flow is also increased. Therefore the flow grouped under region 1,2,3 have more contribution to base flow were as station grouped under region 4 indicate flow is mostly dominated by surface runoff.

3.5.1.3 Method of L-moment ratio diagram

Method of L-moment ratio diagram (L-MRD) is used as a tool to give priority for IHRs and distributions based on the statistical principles. The main hypothesis of the study is that if the annual minimum flows of different stations come from a single distribution model and the kurtosis and skew ness determined by l-moment are nearly the same, then these stations belong to the same group and form a homogeneous region. This is a useful way of representing the moments of different distributions depending on the statistical nature of data. L-moment statistics are used to group stations comparing with geographical proximity and continuity of gauging stations. To use the statistical parameters L-Cs and L-Cku are first computed and those stations that has nearly same coefficient calculated are considered to be in the same region. The formed regions and stations included in the group are then tested by different homogeneity test methods.

3.6 Homogeneity test for identified regions

Once a homogeneous region has been preliminary identified, the degree of homogeneity of the candidate region with respect to flow statistics has to be tested. The necessity is that the region is satisfactorily homogenous that no further division of the region into individual sites would improve the accuracy of low flow estimates. Unbiased sample estimators of the first four PWMs are given as and suggested a homogeneity test based on L-moments which proved to be efficient. Stations in a region can be tested for homogeneity if then it form a region. Different tests are available to check regional homogeneity in terms of the hydrologic response of the stations. In this study, to verify the acceptability of clustering techniques; discordance measure, C-Cv and L-Cv-based statistical homogeneity tests were applied.

3.6.1 Discordancy test of regions

To estimate discordancy values for sites in a region, the sites are considered as points in three dimensional space of sample L-moment ratios (L-CV, L-Cs, and L-Cku). If a vector, Ui = (τ_2^i, τ_3i, τ_4^i) ^{T}, which controlled the L-moment ratios for site i, T is the transpose of the vector Ui (Hosking and Wallis, 1997), then the discordancy measure may be defined as:

$$Di=\frac{1}{3}\ (Ui-\overline{Ui})\ S^{-1}(Ui-\overline{Ui})^T \qquad 3.5$$

$$\overline{Ui}=\frac{1}{N}\sum_{i=1}^{n}Ui \qquad 3.6$$

$$S=\frac{1}{n-1}*\sum_{i=1}^{n}(Ui-\overline{Ui})(Ui-\overline{Ui})^T \qquad 3.7$$

Where N = is the total number of sites

D_i = discordancy measure

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

$\overline{Ui}$ = is the group averages Ui,

S = sample covariance matrix of Ui.

(Hosking and Wallis, 1997) tabulated critical values of the discordancy statistic Di for various numbers of sites in a region at a significance level of 10%. These were used to assess each of the study sites and identify whether they should be analyzed further to ensure homogeneity. The identified regions have tested for discordancy using equation 3.5. However, to determine the value of D_i using simple matrix multiplication was difficult. Due to this, (Hosking and Wallis, 1997) recommended using FORTRAN, Matlab and other computer programs to simplify the work and get acceptable accurate results. For this study, Matlab2018a, R package programming code was employed to simplify the numerical calculations of discordancy index (D_i). The programming code used to calculate the covariance matrix and D_i were given on Appendix-C.

Table 3.5 Critical values of discordancy measure with N sites

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

(Source: Hosking and Wallis, 1997)

3.6.2 Conventional homogeneity test

The criteria used to check regional homogeneity was based on the value of combined variation coefficient (C-Cv). According to some researchers, the lower value of C-Cv will be the performance of the index-flow method for the region under consideration. This is due to the dominance of the flow quantile estimation variance by the variance of at-site sample mean. Hence, for better performance of the index flow method, C-Cv should be kept low. In this method to calculate C-Cv values, the procedures are described below. For each site in the delineated regions; the mean discharge$\overline{Q\imath}$, standard deviation (σ) and coefficient of variation (C_V) were calculated by equation 3.5, 3.6 and 3.7 depend on the mean of AMF of the station (Murphy, et al., 2014).

$$\bar{Q}_{i=} (\frac{1}{N}\textstyle\sum_{i=1}^{N} Qi \tag{3.8}$$

The standard deviation of AMF of the station;

$$\delta_{i=} = \sum_{i=1}^{N} (\sqrt{(Qi - \bar{Q}_i)2}\ /N) \tag{3.9}$$

$$\mathrm{C}V_{i=}\delta_{i/\bar{Q}_i} \tag{3.10}$$

Where: Qi= the flow rate of the station in the region (m3/s), at site I

$\bar{Q}_i$ =The mean flow rate for the region (m3/s), at site i

δ_i= Standard deviation for the region, at site i

N = number of a record year

Cvi = Coefficient of variation of a region, at site I

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

Regional mean; $\overline{CV}_{i_{i=}} \frac{1}{N}(\sum_{i=1}^{N} \mathrm{CV}_i)$ 3.11

Regional standard deviation, $\delta C_i = \sum_{i=1}^{N} \sqrt{(CVi - \overline{CV}_i)2}\ /N$ 3.12

C-Cv =$\delta Cv_i/\overline{CV}_i$ < 0.3 3.13

Where: N=Number of the site in a homogeneous region

$\overline{CV}_i$ = the mean coefficient of at site CV values

δCv_i= Standard deviation of at site CV values

According to (Murphy, et al., 2014) if the C-Cv value of station in a region is less than 0.3 the station in a region is homogeneous.

3.6.3 L-moment based homogeneity test

L-CV-based homogeneity test is more accurate and effective way of testing the homogeneity of the site when compared with that of the Cv-based homogeneity test. The procedural calculation is the same as that of the Cv. The following are advantage of L-Cv (Cunnane, 1989): Compared to Cv. L-Cv can characterize a wide range of distribution, sample estimates are so strong that they are not affected by the presence of outliers in the data set, they are less matter to bias in estimation, yields more accurate estimate of the parameter of a fitted distribution. According to the Central Water Commission (2010), L-moments has the following advantages: I). characterize most of probability distributions than conventional moments, ii). Less sensitive to outliers in the data, iii). Approximate their asymptotic normal distribution more closely, IV). Nearly unbiased for all combinations of sample sizes and populations.

(Hosking and Wallis. 1997) gave the unbiased estimators of β0, β1, β2 and β3 as: defined as;

$$\beta 0= (\frac{1}{N}\sum Qi) \qquad 3.14$$

$$\beta 1=\sum_{i=1}^{N}(\frac{J-1}{N(N-1)})Qi \qquad 3.15$$

$$\beta 2=\sum_{i=1}^{N}(\frac{(J-1)(J-2)}{N(N-1)(N-2)})Qi \qquad 3.16$$

$$\beta 3=\sum_{i=1}^{N}(\frac{(J-1)(J-2)(J-3)}{N(N-1)(N-2)(N-3)})Qi \qquad 3.17$$

Where Qi - annual minimum flow (m3/s) from stations data set n - the number of years, j-rank βo, β1, β2, and β3- are L-moments estimator. The first few moments are:

$\lambda 1= \beta O; \lambda 2 = 2\beta 1-\beta O; \lambda 3 = 6\beta 2-6\beta 1+ \beta O; \lambda 4 = 20\beta 3-30\beta 2 + 12\beta 1-\beta O$ 3.18

in specific, λ1 is the mean of the distribution or measure of location; λ2 is a measure of scale; τ3 is a measure of skewness, and τ4 is a measure of kurtosis. L-skewness and L-kurtosis are both defined relative to the L-scale, λ2; and sample estimates of L-moment ratios can be written as L-Cv, L-Cs, and L-Ck. L-moment ratios are independent of units of measurement and are given by (Hosking and Wallis, 1997) as follows:

$\tau 2= \lambda 2/ \lambda 1, \tau 3= \lambda 3/ \lambda 2, \tau 4= \lambda 4/ \lambda 2$ 3.19

Using the above procedural formula,

Regional mean; $L\overline{CV}_{i_{i=}}\frac{1}{N}(\sum_{i=1}^{N}\mathrm{LCV}_i)$ 3.20

Regional standard deviation,

$$\delta LCv_i = \sum_{i=1}^{N}(\sqrt{(LCVi-\overline{LCV}_i)2}\ /N) \qquad 3.21$$

The weighted regional L-Cv of all the sites, CC is defined as follows:

CC $=\delta LCv_i/L\overline{CV}_i < 0.3$ 3.22

Where: N=Number of the site in a region

$\overline{CV}_i$ = the mean coefficient of at site CV values

L-Cv = Standard deviation of at site CV values

Generally, in both techniques the value of CC is less than 0.3 that indicate the region is quite homogenous and no need of forming another region (Murphy, et al., 2014).

3.7 Delineation of homogeneous regions

The performance of any regional estimation method highly depends on the grouping of sites into homogeneous regions. In this study, the geographical proximity and L-MRD were used in order to classify preliminary regions which then tested for hydrologic similarity. The delineation of homogeneous regions is closely related to the identification of the common regional distributions and closeness of L-Cs and L-C_{ku} that apply within each region. In this study, the DEM of upper Omo gibe River Basin (OGRB) were used and the delineation of homogeneous regions were performed by taking in to account the drainage boundaries of the sub basin with ArcGIS 10.4.1. The preliminarily IHR have to be checked by various homogeneity tests. All sample stations are located on a digitized map by latitude and longitude. For each station, the statistical values (L-Cs, L-Cku) were computed. It was assumed that the L-Cs and L-Cku values of one station vary linearly with the neighboring stations.

The procedures followed in the delineation of the regions are as follows: 1). Compute the (L-Cs, L-Cku) value of each station, 2). Identify the location of stations along the distributions of L-MRD for the defined regions statistical comparison of observed flow data, 3). Identify the group located nearby based on step (2), 4). Each region that was identified on step-3 was checked for statistical homogeneity using the proposed test. In this particular study, (Irwin, et al., 2014) procedures were used in delineating the defined homogeneous regions. According to those authors, the methodology used gives efficient and consistent watershed delineation on DEMs of any size.

Finally taking into consideration the drainage boundaries of each sub-region the delineation was carried out accordingly with the ArcGIS10.4.1 environment.

3.8 Selection of regional frequency distribution

The choice of frequency distributions is determined based on goodness-of-fit measures, which indicates how much the considered distributions fit the available region. In flow event analysis, the annual minimum flow corresponding to a given return period T can be estimated from the annual low flow series using varies theoretical distributions.in this study goodness of fit is determined by L-MRD, xlstat2018, matlab2018a and by EASYFIT software for individual station and Z^{th} distribution test and average L-MRD for delineated region.

3.8.1 L-moment ratio diagram

Regional frequency distribution were fitted by using L-MRD which highly depends on a regional average weighted L-moment statistical value of L-Cs and L-Cku of all sites for the defined homogeneous regions. This shows that grouping of the sample data sets around the theoretical relationships between L-Cs and L-Cku of different probability distributions. Thus, some acceptable design procedures are essentially required to choose a model that minimize uncertainties. Those distribution are generalized extreme value (GEV), generalized logistic (GLO), Generalized Pareto (GPA), Normal, Log Pearson type 3 (LPIII) and Lognormal (LN) distributions, exponential distribution, Gumbel distribution, uniform distribution are among the employed distributions in this study. Many flow frequency distributions have been practiced for flow modeling, but none has been accepted as universal. Hence, these distributions were considered for the evaluation of the possible distributions that can represent the average frequency distribution of the regional data of the basin. During the computation of moment statistic every distribution has its own formula to calculate the skewness and kurtosis of their distribution and the formula for some distribution are expressed below (Hosking, 1990 and 1991a, b):

a) Uniform distribution: $Z_3=0$ $Z_4=0$
b) Exponential distribution: $Z_3=1/3$ $Z_4=1/6$
c) Normal distribution: $Z_3=0$ $Z_4=0.1226$
d) Gumbel distribution: $Z_3=0.1699$ $Z_4=0.1504$
e) Logistic: $Z_3=0$ $Z_4=1/6$
f) Lognormal (two and three parameters):

Z_4=0.12282+0.77578 $(Z_3)^2$ +0.12279 $(Z_3)^4$ - 0.13638$(Z_3)^6$ +0.113638$(Z_3)^8$ 3.23

g) General Extreme Value (GEV)

Z_4=0.1070+0.1109 $(Z_3)^2$ -0.0669 $(Z_3)^3$ + 0.60567$(Z_3)^4$ - 0.04208$(Z_3)^5$ +0.03763$(Z_3)^6$ 3.24

h) Gamma and Pearson type III

Z_4=0.1224+0.30115 $(Z_3)^2$ +0.95812 $(Z_3)^4$ - 0.57488$(Z_3)^6$ + 0.19383$(Z_3)^8$ 3.25

Generalized Logistic:

$Z_4= (1 + 5Z^2{}_3)/6$

Or $Z_4= 0.16667 + 0.83333\ Z^2{}_3$ 3.26

H. Generalized Pareto:

$Z_4 = 0.20196\ Z_3 + 0.95924\ Z^2{}_3 – 0.20096\ Z^3{}_3 + 0.04061\ Z^4{}_3$ 3.27

or Z4 = $Z_3\ (1 + 5Z_3) / (5 + Z_3)$

3.8.2 EasyFit Software for distribution fitting

These methods are preferred especially in cases where there is little or no information about the base distribution pattern in data and the need to find the best distribution type. In order to determine whether the distribution model could fit the data properly, goodness-of-fit tests were used. In the present study Easy Fit 5.6 Statistical Software Package, trial version 5.6 was used to find the best-fit distribution and its estimation parameters.

3.8.3 Sigma magic software package

These methods are similar to EASYFIT it is preferred especially where there is little or no information about the base distribution pattern in data and the need to find the best distribution type. In order to determine whether the distribution model could fit the data properly, chi-square goodness-of-fit tests were used. In this study sigma magic Statistical Software Package, trial version 11.3 was used to find the best-fit distribution and its estimation parameters.

3.8.4 Matlab

Is another software program used to compute the best fit distribution for hydrological time series data and compute its statistical parameter and used to calculate RMSE of the distribution selected. In this study, Matlab2018a were used to find the best-fit distribution and its estimation parameters.

3.8.5 XLSTAT

Xlstat is the other statistical software used to estimate the best fit distribution for data to be used in this study by applying different goodness of fit test such as Kolmogorov simirnov and chi-square test methods. The other use of this software package is used to determine which parameter estimation method is used with the selected distribution parent. Therefore in this study xlstat2018 trial version were used for different computation.

3.9 Performance evaluation of probability distributions

Assessment of the accuracy of the estimates should, therefore, take into account the possibility of heterogeneity in the region, miss-specification of the frequency distribution and statistical dependence between observations at different sites, to an existent that is consistent with the data. Analytical goodness-to-fit criteria are helpful as an approval for whether a particular elimination of the data from the model is statistically significant or not.

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

I. Probability-probability plots

Probability plots are generally used to decide whether the distribution of a variable matches a given distribution. P-P plots show that the observed values together with the simulated from the regional values may reveal a systematic regional bias in the estimation of the quantile events. This is for visually informative the character of a data set and to determine if fitted distribution seems reliable with the data.

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

ii. Quantile-quantile plots

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

iii. Zth-statistics test methods: in this method if the residual error calculated from recorded flow data and predicted value of minimum flow that has different return interval calculated by recommended distribution is less than 1.64 the distribution selected taken as the best fit distribution for the delineated homogeneous region (Hosking and Wallis, 1997). Several methods are available for testing the goodness of fit of a distribution to data from a single sample. These include quantile-quantile plots, chi-squared, Kolmogorov-Smirnov, and other general goodness-of-fit tests and tests based on moment or L-moment statistics. Some of these methods can be adapted for using in the regional network. The fit of a postulated regional frequency distribution to each site's data can be assessed by goodness-of-fit statistics calculated at each site, and the resulting statistics then combined into a regional goodness-of-fit statistic. The goodness-of-fit criterion for each of the various distributions is defined in terms of L-moments and is termed the *Z*-statistic (Hosking and Wallis, 1997):

$$Z^{Dist} = (\tau_4^{Dist} - \tau_4^{R} + \beta_4) / \delta_4 \tag{3.28}$$

Where β_4 is the bias of τ_4^{R}, and δ_4 is the standard deviation of τ_4^{R} defined as follow:

$$\beta_4 = \sum(_{m=1}^{N})(\tau_4^{Dist} - \tau_4^{R)}/N \tag{3.29}$$

$$\delta_4 = \sum \{(\tau_4^{Dist} - \tau_4^{R})^{2} - N^{*} \beta_4^{2})/N-1\}^{\wedge 1/2} \tag{3.30}$$

The smallest value of Z^{dist} will represent the more fit distribution to the data under the study (Malekinezhad, 2011).

iv. **RMSE** methods this is done by calculating the root mean square error for identified distribution and the distribution having minimum value of RMSE is the best fit distribution for homogeneous region and calculated as:

Residual = (Observed value-predicted value)

$$RMSE = \Sigma\ ((Residual^{2}/n))^{0.5} \tag{3.31}$$

Table 3.6 RMSE for selection of parameter

Station name	RMSE by xlstat	RMSE by Easyfit	RMSE by Matlab	Station name	RMSE by xlstat	RMSE by EASYFIT	RMSE by Matlab
Abelti	9.1125	10.463	8.136	Dincha	0.368	1.4643	0.254
Bidru	0.00228	0.00416	0.0008	Woshi	1.349	1.14917	0.028556
Tollay	0.815	0.14	0.135	Gecha	0.0931	0.0706	0.0573
Awetu	0.01328	0.019	0.027	Sheta	0.1066	0.10588	0.0896
Asendabo	2.48	2.26	1.2788	Shebe	2.727	2.315	0.089
Bulbul	0.10744	0.22	0.0185	Guma	0.222	0.412	0.149
Limmu	0.1292	0.091344	0.0505	Walga	0.039	0.0017	0.0006
Seka	0.368	1.4643	0.22654	Mazie	0.0246	0.0098	0.00007
Kito	0.031	0.0103	0.0083	Gogeb	0.014	0.039	0.005
Chida	3.69	1.77	1.42	Goghora	0.086	0.32	0.03

3.10 Parameter and quantile estimation

The method used for parameter estimation for quantile determination of low flow with different return period is determined by the help of Xlstat software. According to this software method of moment is used for parameter estimation with the help of Xlstat depend on most likelihood probability of the distribution to fit the data. These parameters are used to calculate the quantiles related to return periods. The method used for regionalization is the index flow method which comprises the standardized AMF series of each station divided by site averaged AMF values. The frequency distribution procedure of AMF data in a homogeneous region consists of similar quantile distribution. After the parameters of a distribution are estimated, low flow quantile estimates (X_T) which correspond to different return periods can be computed. In the present study, the parameter estimation was done by using the Easy Fit, XLSTAT, and Matlab. Based on the selected distributions for each station, the quantile can be calculated according to the formula of the selected distributions. For stations with a computed value of scale, location and shape parameter, then it is possible to determine the quantile with different return periods using different equations for different distributions and in other case it is simple to estimate quantile by using Xlstat, SMADA on line and statistics software package (www.smadaonline.com or SMADA2013.App-code.SMADA.distribution).

For GEV distribution the flow quantile can be estimated as;

$$X_T = \mu + \frac{\delta}{K}(1- (-\ln(1- \frac{1}{T})^{K}) \text{ for } K \neq 0 \tag{3.32}$$

$$X_T = \mu + \delta(\ln(-\ln((1- \frac{1}{T})^{K}) \text{ for } K = 0 \tag{3.33}$$

For GPA distribution the flow quantile can be estimated as;

$X_T = \mu + \delta(\ln(\frac{1}{T})^K)$ for $K = 0$ 3.34

$X_T = \mu + \delta/K(1-(\frac{1}{T})^K)$ for $K \neq 0$ 3.35

Where σ= Scale parameter,

T= return period

μ= Location parameter; and

k = Shape Parameter

In general estimation of parameters and calculation of the magnitude of low flow for T years return period were executed.

3.11 Derivation of the regional low flow frequency curves

The average regional growth curves was determined to represent the frequency curves of regions. Index flow method employs data of the gauged catchments to evaluate a regional correlation from the flow magnitudes of various return periods for ungauged catchments to be evaluated (Modi & Mitra, 2017).In this study, the index flow method was used to determine the magnitude and frequency of low flow quantiles for sites located within a homogeneous region.

3.12 Estimation of flow- index

Based on the index flow procedure, for estimating a T-year return period flow quantile at ungauged sites, an estimate of the index flow or sample mean of annual low flow data is required. Since observed flow data are not available at ungauged sites, the at-site mean cannot be computed. In such a situation, it is necessary to establish a relationship between the mean annual low flow of the gauged catchments within the homogeneous region, and their pertinent physiographic and climatic characteristics to obtain an estimate of the mean annual low flow.

The model parameters for the distributions estimated for each station were used to compute standardized flow estimates conforming to the return periods 5, 10, 20, 25, 50, 75 100, and 150 years. Plots of regional estimated flow with its recurrence interval is known as growth curves, were generated for each station and used in the derivation of the regional growth curves. To do this, the following stages were employed. Select best fitted distributions with parameter values such as shape (k), location (μ) and scale (σ) which were estimated using Matlab Software, the model parameters estimated for a given region were then used to compute the standardized quintiles estimates for the return periods. In this method, the dimensionless regional growth curves used to estimate XT. After the regional frequency distribution is determined, the flow quantiles having a return period of T year within a homogeneous region can be estimated based on the equation 2.4. The common practice is to get the dimensionless data by dividing the values by an estimate of the at-site mean.

3.13 Estimation of Regional Growth Curve

The parameters of the regional growth curve which is identical to the selected regional distribution can be estimated by pooling the information available from all the sites within the homogeneous region. (Hosking and Wallis, 1997) suggested the following procedure to estimate the parameters of regional growth curve.

1. The first four unbiased L-moments and their ratios should be computed separately for each site within the homogeneous region.
2. The average L-moment ratio weighted proportionally to the record length of each site should be obtained.
3. The parameters of selected regional distribution for homogeneous region should be estimated using the regional average L-moment ratios. These estimations can be performed by using the provided relationships between the L-moments and the parameters of some distributions.
4. Plot the quantile function q (f) of the regional frequency distribution estimated in step (3) versus the return period. The resulting curve is the regional low flow growth curve for the region. Figure 4.11 (a-d) presents a typical regional growth curve.

4 RESULT AND DISCUSSION

4.1 Assessment of magnitude and return period of at site upper Omo gibe river basin

The magnitude and return period of low flow of at site upper Omo gibe river basin were determined for 5, 10, 15, 25, 50, 75, 100, 150, return period were calculated by SMADA (storm water management and design aid on line software) and by theoretical formula based on selected distribution for further analysis selected by L-MRD and different software packages. For the gauged upper Omo gibe river basin it is tabulated on table 4.1 for Woshi and Gecha station and appendix-D for the other station.

Table 4.1 Magnitude and return period of at site station

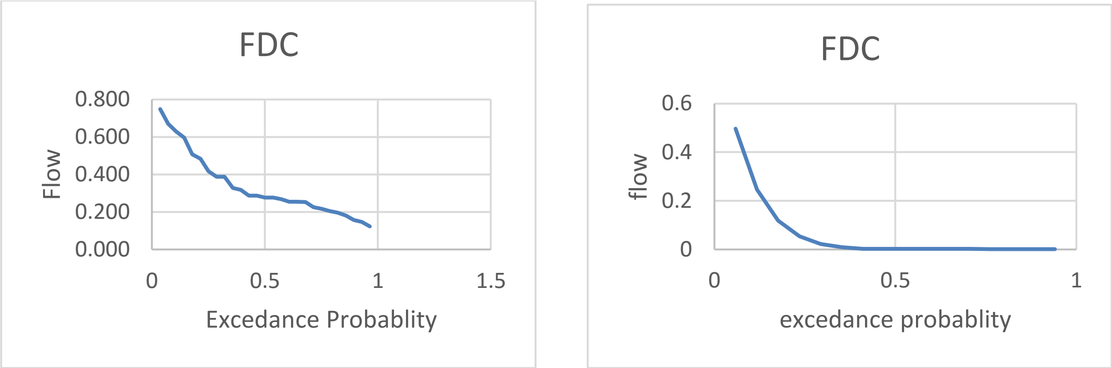
station	Return P.(year)	Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	station	Return P.(year)	Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab
Woshi	5	0.508815	1.19369	1.266746	Gecha	5	0.513201	0.471773	0.447801
	10	0.53119	1.838367	2.023041		10	0.58476	0.527112	0.493254
	15	0.539831	2.337432	2.632182		15	0.622825	0.555585	0.515952
	25	0.547769	3.140171	3.645238		25	0.667824	0.588361	0.541463
	50	0.555	4.649938	5.635396		50	0.724552	0.62828	0.571591
	75	0.557937	5.834726	7.257134		75	0.755737	0.649544	0.587191
	100	0.559599	6.848713	8.67888		100	0.77705	0.663793	0.597462
	150	0.561477	8.576663	11.16194		150	0.806013	0.682782	0.610911

Remark the 7MAM flow of Gecha and Woshi station that has 10 year recurrence interval are 2.023 and 0.493 m^3/s respectively.

4.2 Flow duration curves

It displays the relationship between stream flow and the percentage (probability) of time it is exceeded (Zadeh, 2012). According to this study the 7-day annual minimum flow duration curves (LFDC) were computed and for demonstration LFDC of Chida and Woshi station were presented on figure 4.1 and for the other station were presented on appendix -E.

Remark: low flow duration curve indicate that the exceedance probability increasing with decreasing value of 7MAM and flow with greater than 50% exceedance probability considered as low flow of (Zadeh, 2012).



(a) Gecha station (b) Woshi station

Figure 4.1 Flow duration curve of Chida and Woshi station

4.3 Identification of homogeneous region

The homogeneity of a proposed region was preliminarily judged based on site characteristics and L-moment ratio diagram (L-MRD) of low flow statistics. The classification of sites was carried out by hierarchical geographic regionalization procedure. This method considers the stations that were geographically continuous (Contiguous) as indicated on Figure 4.2 and the annual minimum flow of sites in the region should satisfy the homogeneity test criteria.

L-MRD shown on Figure 4.2 were used to identify homogeneous regions with site characteristics of gauging stations described in Table 3.1. As indicated on Table 4.2, the fitted distributions were designated to the same group since stations lie close to the identical distribution. Hence, based on L-moment statistics and suitability of gauging site networks, four homogeneous regions were identified. Namely Region-1, Region-2, Region-3 and Region-4 as shown on Table 4.2 and lastly the delineated region were situated on figure 4.3.

Table 4.2. Distribution selected for regions by L-MRD

Region name	Station name	Possible distribution	Region name	Station name	Possible distribution
R-1	Awetu Limmu genet Asendabo Gojeb Chida Seka Kito Bulbul	Lognormal/GLO GEV/lognormal GLO/lognormal lognormal GLO/lognormal Lognormal/GLO GEV/lognormal	R-2	Dincha Woshi Gecha Shebe Sheta Guma	Lognormal/GEV GEV Pearson/GEV GP/GEV GEV/Pearson GEV
R-3	Abelti Bidru Tollay	GEV/GP GP/GEV GEV/GP	R-4	Gogeb Goghora Walga Mazie	GEV/GP GEV/GPA GEV GEV

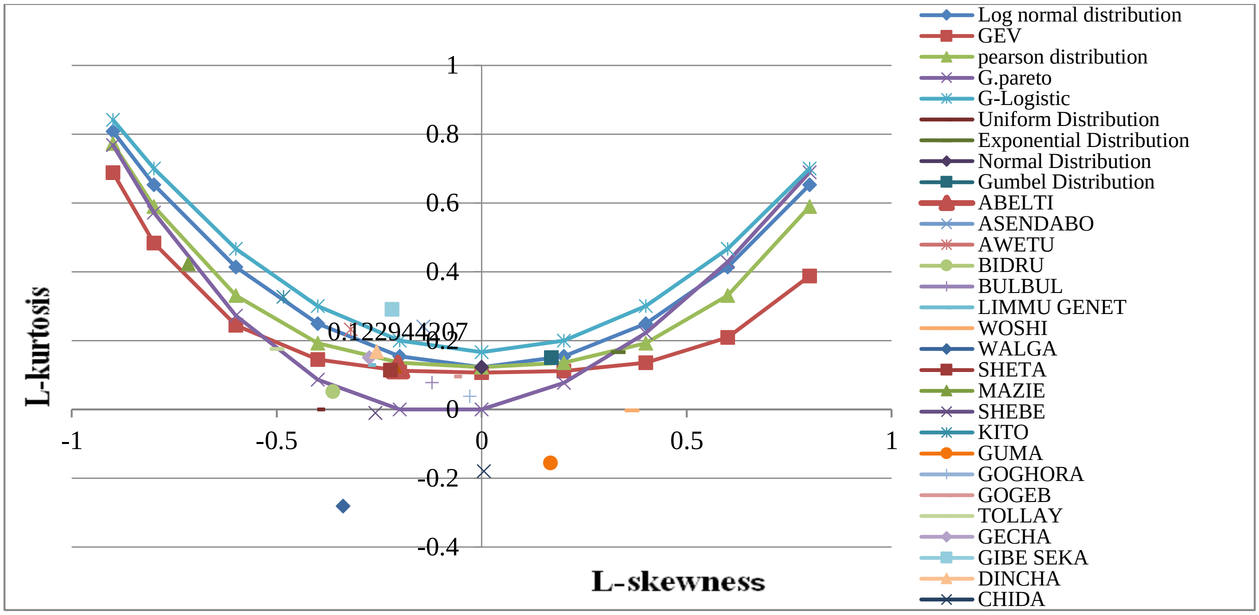


Figure 4.2 L-moment ratio diagram for identification of homogeneous regions

Table 4.3 Distribution selected for station by EASYFIT

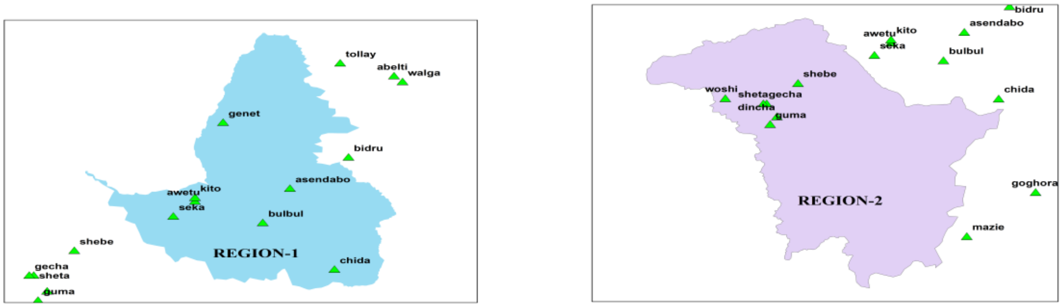
Region name	Station name	Possible distribution	Region name	Station name	Possible distribution
R-1	Awetu Limmu genet Asendabo Gojeb Chida Seka Kito Bulbul	Lognormal/GEV GEV/lognormal GEV/lognormal GEV/lognormal GEV/lognormal Lognormal/GLO GEV/lognormal	R-2	Dincha Woshi Gecha Shebe Sheta Guma	Lognormal/GEV GEV/GP GEV/lognormal Weibull /GEV GEV/Pearson GEV
R-3	Abelti Bidru Tollay	GEV/lognormal lognormal/GEV GEV/Weibull	R-4	Gogeb Goghora Walga Mazie	GEV/GP GEV/GP GEV/GP GP/GEV

Table 4.4 Distribution selected for a region by matlab

Region name	Station name	Possible distribution	Region name	Station name	Possible distribution
R-1	Awetu Limmu genet Asendabo Gojeb Chida Seka Kito Bulbul	Lognormal/GEV GEV/lognormal GEV/lognormal GEV/lognormal GEV/lognormal GEV GEV/lognormal	R-2	Dincha Woshi Gecha Shebe Sheta Guma	GEV GEV GEV/lognormal Weibull /GEV GEV/lognormal GEV
R-3	Abelti Bidru Tollay	GEV/lognormal Log-logistic/GEV GEV/Weibull	R-4	Gogeb Goghora Walga Mazie	GEV/GP GEV/GP GEV/GP GLO/GEV

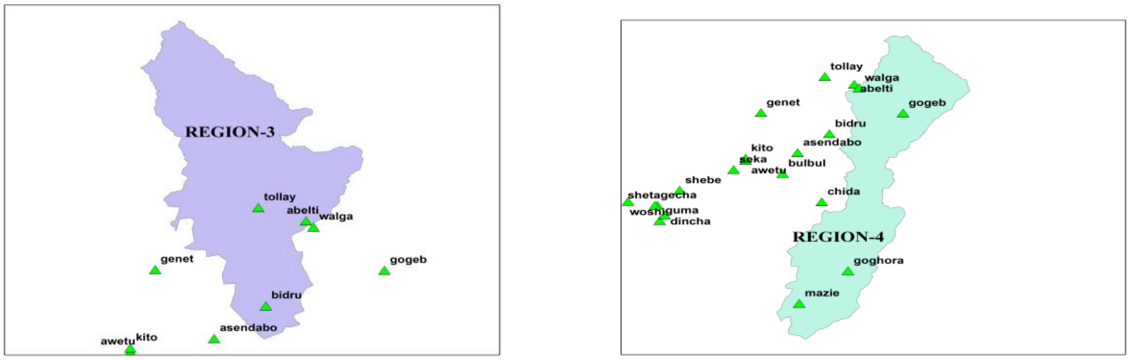
Table 4.5 Distribution selected by xlstat

Region name	Station name	Possible distribution	Region name	Station name	Possible distribution
R-1	Awetu Limmu genet Asendabo Gojeb Chida Seka Kito Bulbul	Lognormal/GEV GEV/lognormal lognormal/GEV GEV/lognormal Lognormal/GEV GEV GEV/lognormal	R-2	Dincha Woshi Gecha Shebe Sheta Guma	GEV GEV GEV/lognormal Weibull /GEV GEV/lognormal GEV
R-3	Abelti Bidru Tollay	GEV/lognormal Lognormal/GEV GEV	R-4	Gogeb Goghora Walga Mazie	GEV/GP GEV/GP GEV/GP GLO/GEV



(a) Region-1

(b) Region-2



c) Region-3

d) Region-4

Figure 4.3 Identified homogeneous region

4.4 Test for regional homogeneity

Once a homogeneous region has been preliminary identified, the degree of homogeneity of the candidate region with respect to flow statistics has to be tested. The necessity is that the region is satisfactorily homogenous that no further division of the region into individual sites would improve the accuracy of low flow estimates. Stations in a region can be tested for homogeneity if then it form a region. Different tests are available to test regional homogeneity in terms of the hydrologic response of the stations. In this study, to verify the acceptability of clustering techniques; discordance measure, C-Cv and L-Cv-based statistical homogeneity tests were applied. Detail of this test were discussed below.

4.4.1 C-CV-based regional homogeneity test

The criterion used to check regional homogeneity was based on the value of combined coefficient variation (C-Cv). According to some researchers, the higher the value of CV and C-Cv, the lower will be the performance of the index-flow method for the region under consideration. In this study the value of C-Cv of region was less than 0.3 that indicate the region formed were homogeneous as stated on 3.6.2 of previous section and detail on table 4.6.

4.4.2 L-CV-based regional homogeneity test

The criteria used to check regional homogeneity was based on the value of combined variation coefficient (L-Cv). According to some researchers, the higher the value of L-CV and CC, the lower will be the performance of the index-flow method for the region under consideration. In this study the value of CC of the all-region was less than 0.3 that indicate the region formed is homogeneous as stated on 3.6.3 of previous section and on table 4.6.

4.4.3 Discordancy test

(Hosking and Wallis, 1997) tabulated critical values of the discordancy statistic Di for various numbers of sites in a region at a significance level of 10%. These were used to assess each of the study sites and identify whether they should be analyzed further to ensure homogeneity. The identified regions have tested for discordancy using equation 3.5. However, to determine the value of Di using simple matrix multiplication was difficult. Due to this, (Hosking and Wallis, 1997) recommended that using Matlab and other computer programs to simplify the work and get acceptable accuracy results. Depend on these the computed value of Di for the proposed region was less than the critical value of discordancy stated on table 3.5 of previous discussion and detail of these region is discussed on table 4.6.

Table 4. 6 Di, C-CV and L-CV test for homogeneity

Region name	Station name	C-CV test	L-CCV test	Discordance test	Region name	Station name	C-CV test	L-CCV test	Discordance test
R-1	Awetu	0.24	-0.57	0.57	R-2	Dincha	0.28	-0.49	0.44
	Asendabo			1.08		Shebe			1.56
	Limmu genet			0.94		Gecha			0.69
	Kito			1.32		Sheta			0.23
	Seka			1.01					
	Bulbul			1.53		Woshi			1.59
						Guma			1.51
	Gojeb Chida			1.18	R- 4	Gogeb	0.26	-0.51	1
						Walga			1
R- 3	Abelti	0.28	-0.22	0.38		Goghora			1
	Bidru			0.67					
	Tollay			0.402		Mazie			1

4.5 Delineation of homogeneous regions

After organizing and assembling the data set, important statistical parameters have been computed and interpolation of these statistical values (L-Cs, L-Cku) in collaboration with site characteristics are then used to come up with the following results of delineation. Delineation of regions were done depending on the fact that the statistical homogeneity tests satisfied. The regions have covered an area of 4696.4, 4665.2, 22367.8 and 3104.3 km^2 for Region-1, 2, 3 and 4 respectively. Having proven statistically homogeneous region, the delineated homogenous regions shown on Figure 4.3 could be used to generate a regional growth curve at any site located in the study area.

4.6 Determination of suitable regional probability distribution

In this study, the annual minimum 7-day series model was adopted where only the minimum flow in each water year is considered. Depend on Z-distribution selection of candidate distribution function were done for proposed region. In this thesis Generalized extreme distribution and generalized Pareto distribution were selected for region -1, ragion-2, region-3 and region-4 respectively.

4.6.1 Goodness of fit tests

The goodness of fit tests was performed for all distributions using Kolmogorov Smirnov, Anderson-Darling and chi-square methods for the data of gauging stations. They were applied to determine whether the distribution to be fitted to the data or not. The best-fit result of each station was taken as the distribution with the lowest sum of the rank orders from each of the three test statistics. This GOFs at 5% level of significance was used to define the best-fit using Easy Fit Statistical Software and Matlab programing depend on high ranks by easy fit and most likely probability to fit the data by Matlab. The probability distribution having the first rank along with their test statistic were presented on Table 4.6 and on Appendix-G for the other. Using the three tests from Table 4.6 and Appendix-G, it was detected that generalized extreme value distribution for region 1 and 2, and generalized Pareto distribution for region 3 and 4 provides the best fit to the AMF data. Comparing the results of goodness-of-fit tests, the generalized extreme value and generalized Pareto distributions afford a good fit for the recorded data of stations. It was also observed that most of the probability distributions have the first rank in both Kolmogorov Smirnov and Anderson Darling tests. This indicates that the two goodness-of-fit tests lead to a reasonable estimation of low flow in the upper Omo gibe River Basin.

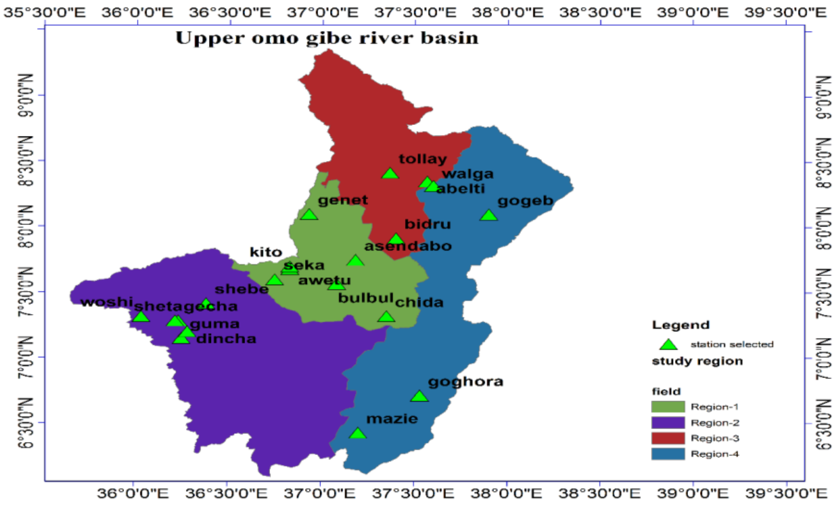


Figure 4. 4 Delineated homogeneous region.

4.7 Evaluating estimation accuracy of selected distribution

The P-P and Q-Q plot have to be more or less linear if the particular theoretical distribution is the correct model or not. It was observed that from the results shown in Figure 4.5 for delineated region and Appendix-H for the rest of the stations, indicated that almost all plots were well fitted to the line. Through all the patterns, the study reveals that lognormal and GEV distributions performed well for most of the stations in the basin. Therefore, results from both methods validated that the flow frequencies of the regions were well addressed. Hence, using these distributions and annual minimum flow modeling could have a wide range of applications in agriculture, hydrology, engineering design and future climate evaluation in the study area.

Table 4.7 Affordable selected distributions for Bulbul station and their ranks

#	Distribution	Kolmogorov Smirnov		Anderson Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Exponential	0.16993	11	3.7222	6	0.25887	3
2	Gamma	0.13919	7	4.0462	9	1.3344	10
3	Gen. Extreme Value	0.08241	1	0.20907	1	0.09486	1
4	Gen. Pareto	0.08416	2	0.27002	2	0.1335	2
5	Lognormal	0.14203	9	4.7523	12	0.51075	5

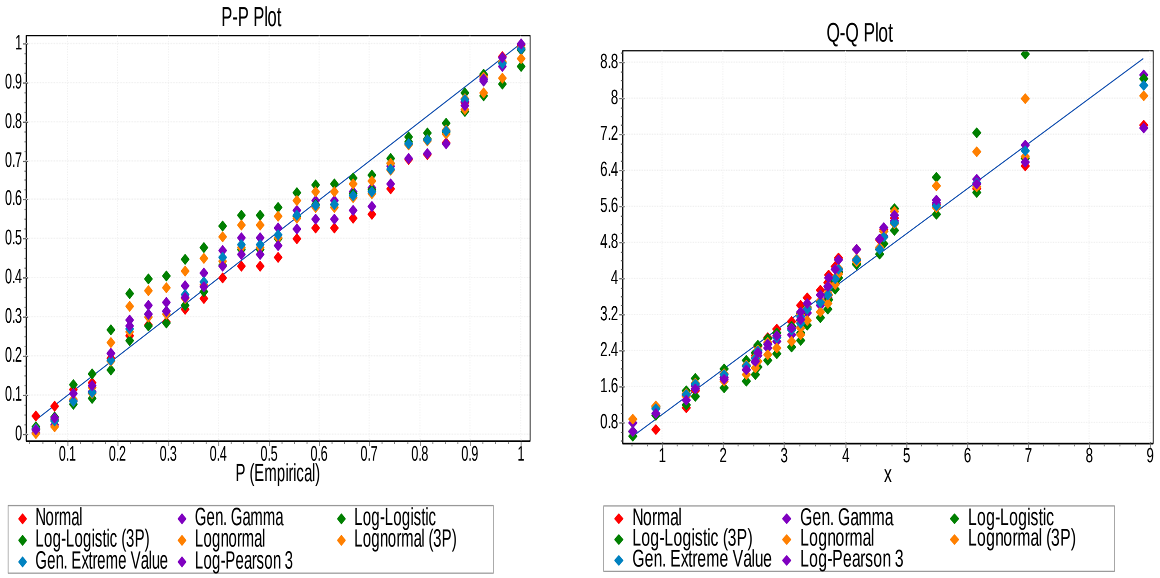


Figure 4.5 P-P plot and Q-Q plot for Asendabo station

4.7.1 Method of L-moment ratio diagram

This method is used for assessing the performance of the average values of the point (L-Cs, L-Cku) of all stations within the region close to L-MRD of the selected parent distribution. The corresponding average weighted value of L-moment statistics results were obtained from regional data as presented on Table 4.8. As shown on Figure 4.6, the points representing the regional average values of L-Kurtosis versus L-

Skewness were fitted with General Pareto and GEV distributions. Therefore, it appears that the GEV and GPA distributions would be suitable distributions for the regions. The choice of a suitable standard frequency distribution is often uncertain and L-MRD might not guarantee that the distribution is the actual representative of flow statistics in the given region. For this reason, a confirmation of candidate distributions is needed. Hence, the results between the goodness-of-fit test with Easy Fit and L-MRD indicated that due to the common acceptance of GEV and GPA distributions, could be used as a best-fit distribution for the study area. Therefore, GEV and GPA distributions could be adopted as the regional distribution, while the other distribution such as LPIII, Normal, LLO and LN distributions should not be selected. As a result, this justified that the two distributions would be acceptable and the dominate probability distributions in the Upper Omo gibe River Basin for estimation of regional low flow frequency.

Table 4.8 (a and b) Regional average weight of L-moment statistics

Region	station name	L-Cs	L-Cku	Average L-CS	Average L-Cku
R-1	Awetu	-0.320	0.232	0.057	0.160
	Asendabo	-0.142	0.241		
	Bulbul	-0.122	0.078		
	Kito	0.483	0.327		
	limmu.G	0.276	0.129		
	Chida	0.005	-0.179		
	gibe Seka	0.219	0.291		
R-4	Walga	1.102	0.523	0.478	0.270
	Mazie	0.716	0.424		
	Goghora	0.029	0.038		
	Gogeb	0.066	0.096		

a) Regional average weight of l-moment statistics for region-1 and region-4

R-2	Dincha	-0.257	0.168	0.155	0.070
	Woshi	0.384	-0.474		
	Gecha	-0.274	0.151		
	Shebe	0.259	-0.010		
	Sheta	0.223	0.115		
	Guma	0.595	-0.372		
R-3	Abelti	-0.204	0.123	-0.356	0.117
	Bidru	-0.364	0.052		
	Tollay	-0.499	0.176		

(b) Regional average weight of l-moment statistics for region-2 and region-3

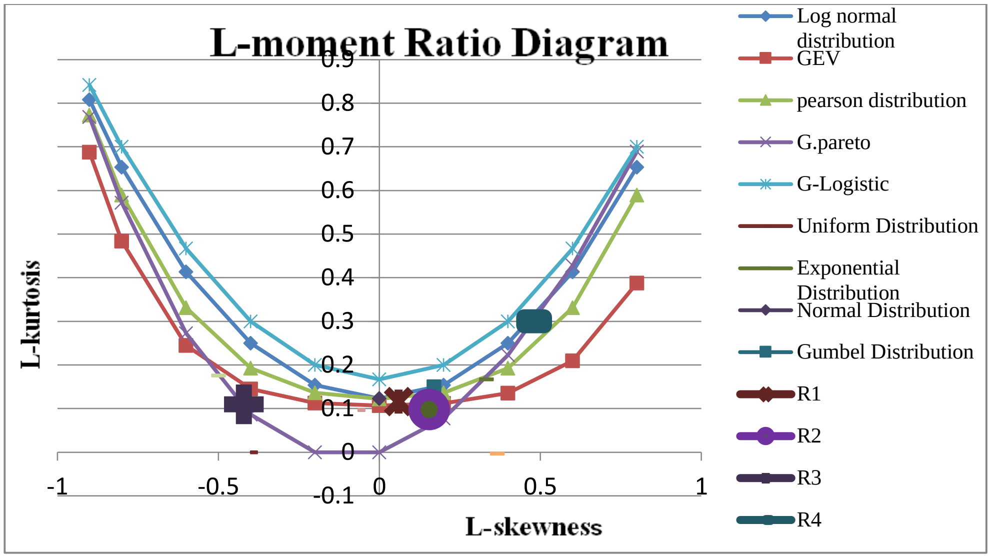


Figure 4.6 Average regional L-MRD for provided regions

4.7.2 Hosking and Wallis goodness of fit test (Z^{th} -statistics test)

The candidate distribution is declared as adequate fit if Z^{Dist} is sufficiently close to zero. (Hosking and Wallis, 1997) suggested a reasonable criterion being $|Z^{Dist}| \leq 1.64$. Then depend on the above formula the best fit distribution for identified homogeneous region is then selected as easyfit and L-MRD and detail of the discussion were illustrated on table 4.9.

Table 4.9 Z^{th} distribution computed for delineated regions

Region name	Z^{th} for GEV	Z^{th} for GPA	Z^{th} for GLO	Z^{th} for LN
R_1	0.201499469	0.607758145	0.444930338	0.2170801464
R_2	0.780669754	0.94125966	0.917853384	0.876376763
R_3	0.507979431	0.383177417	0.761961675	0.724913469
R_4	0.469215591	0.2479143996	0.682342546	0.622532026

As indicated on table 4.9 the smallest value of Z^{Dist} for region-1 and region-2 were obtained by GEV distribution whereas for region-3 and region-4 GPA is the best fit distribution.

4.8 Estimation of regional low flow frequency curves

After regions have been accepted as homogeneous, suitable distributions were identified for the regions. The low flow frequency curves were established for each station based on suitable distribution to calculate the deviations in the standardized flow of various return periods. The IHR to be fit GEV and GPA distribution for region-1, 2 and region-3, 4 respectively.

4.8.1 Parameter and quantile estimations

MOM were selected for parameter estimation by xlstat statistical software. Parameter were estimated by Matlab for selected distribution and they displayed on Table 4.10. These results were generated according to the ranks and descriptive statistics of the goodness fit tests and the probability of fitting the flow data shown on Table 4.10. As a result, these distributions could be adopted as the dominating distribution in the upper Omo gibe River Basin for accurate estimation of low flow quintiles. Estimation of low flow quintiles were applied for 5, 10, 20, 25, 50, 100 years return period and low flow frequency curves for stations were developed. Low flow frequency curves (LFFC) were estimated using equation 2.4. This estimation of flow can be utilized in the designing of vital hydraulic structures in the river reach.

Table 4.10 Results of estimated parameters for fitted distribution for station

Region	station name	distribution	K (shape)	Sigma (scale)	Mu (location)
R-1	Awetu Asendabo Bulbul Kito limmu.G Chida gibe Seka	GEV	0.3 -0.12 -0.05 0.501 0.31 0.296 -0.0755	0.02 1.656 0.154 0.028 0.126 3.42 0.175	0.03 2.785 0.162 0.075 0.511 6.96 0.3145
R-2	Dincha Woshi Gecha Shebe Sheta Guma	GEV	0.157 -0.617 0.206 0.125 0.114 -0.1	0.406 0.588 0.106 3.57 0.125 0.3	0.616 0.286 0.252 2.48 0.234 0.6
R-3	Abelti Bidru Tollay	GPA	0.105 0.994 0.115	11.12 0.016 1.095	15.38 0.13 0.376
R-4	Walga Mazie Goghora Gogeb	GPA	0.011 0.1 -0.2352 -0.25247	0.052 0.033 0.1377 0.02224	0.001 0.0012 0.1715 0.057346

4.8.2 Estimation of index-flow for standardization or regional growth curve

In this case, the average growth curves were determined to represent the low flow frequency curves of regions. The results of Appendix-I and Table 4.11 show that the standardized quantiles for stations using the selected distribution and parameters with their corresponding return periods. It was observed that the magnitude of low flow increases as the return period increases for selected distribution parameter for all stations. This may be due to the variability of the flow regimes of hydrological phenomena generating the low flow events. This can significantly help in risk assessment works, water resources management, and engineering decisions and actions in the study area**.** For more discussion the regional growth curve developed on basis of average l-moment statistics was presented on figure 4.7 (a-d) that is depend on table 4.11 (a), (b) and (c).

Table 4.11 Estimated standardized parameter for estimation of flow quantiles.

Region	Station name	Distribution	k-Average	Sigma-Average	Mu-Average
R-1	Awetu Asendabo Bulbul Kito limmu.G Chida gibe Seka	GEV	0.166	0.798	1.549
R-2	Dincha Woshi Gecha Shebe Sheta Guma	GEV	0.159	0.482	0.532
R-3	Abelti Bidru Tollay	GPA	0.422	3.90	5.199
R-4	Walga Mazie Goghora Gogeb	GPA	0.195	0.0458	0.058

Depending on selected distributions, regional growth curves were derived as indicated on Figure 4.7 (a-d). Figure 4.7 (a and c) indicated that the growth curves of Region-1and region-3 which represents the main reaches of most of the rivers Awetu and Kito near Jimma, Gibe near Asendabo, Bulbul, Gojeb Chida, Gibe near Seka, Gilgel gibe near Limmu genet, and Gilgel gibe at Abelti, Bidru near Sokoru, and gibe near Tollay cause extensive high flow in their lower reaches and causes extensive floods in their lower reaches. Therefore, the lower reaches of homogeneous Region-1 might be affected by the occurrence of low flow at those station.

Table 4.12 Standardized reduced Gumbel variate for regional growth curve.

Station		Awetu	Asendabo	Bulbul	LGenet	Seka	Chida	Kito	
R.P.	G.Variate	G.Variate	G.Variate	G.Variate	G.Variate	G.Variate	G.Variate	G.Variate	Reg-1
5	0.786	0.831	0.570	0.586	0.903	0.613	0.852	0.929	0.755
10	0.883	0.920	0.706	0.729	0.954	0.743	0.929	0.974	0.851
15	0.932	0.961	0.788	0.813	0.978	0.820	0.966	0.993	0.903
25	0.989	1.006	0.896	0.919	1.004	0.918	1.005	1.012	0.966
50	1.058	1.056	1.051	1.065	1.033	1.056	1.049	1.030	1.049
75	1.095	1.081	1.147	1.153	1.047	1.140	1.071	1.038	1.097
100	1.119	1.097	1.219	1.216	1.055	1.200	1.084	1.042	1.131
150	1.151	1.116	1.323	1.307	1.067	1.288	1.102	1.048	1.179

Station		Dincha	Woshi	Gecha	Shebe	Sheta	Guma	
Return P	G.Variate	G.Variate	G.Variate	G.Variate	G.Variate	G.Variate	G.Variate	Reg-2
5	0.74	0.76	0.183	0.823	0.687	0.752	0.596	0.634
10	0.85	0.868	0.293	0.907	0.822	0.857	0.726	0.745
15	0.92	0.923	0.381	0.95	0.89	0.914	0.804	0.811
25	0.986	0.987	0.527	0.996	0.976	0.98	0.906	0.895
50	1.07	1.064	0.815	1.05	1.078	1.061	1.051	1.02
75	1.115	1.105	1.05	1.08	1.13	1.106	1.141	1.103
100	1.145	1.13	1.256	1.099	1.17	1.136	1.207	1.167
150	1.185	1.169	1.615	1.12	1.223	1.177	1.303	1.268

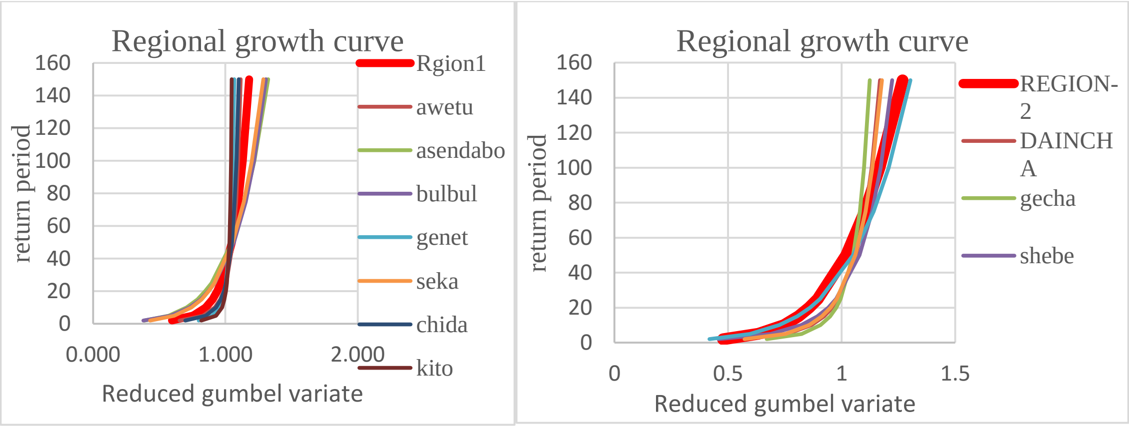
a) Standardized reduced Gumbel variate of region-1 and region-2

STATION		Abelti	Bidru	Tollay	
return P	G.Variate	G.Variate	G.Variate	G.Variate	Reg-3
5	0.808	0.670	0.933	0.561	0.721
10	0.912	0.802	0.991	0.735	0.843
15	0.960	0.875	1.010	0.832	0.906
25	1.012	0.962	1.026	0.949	0.979
50	1.066	1.074	1.037	1.097	1.069
75	1.092	1.135	1.041	1.180	1.119
100	1.107	1.177	1.043	1.236	1.152
150	1.127	1.235	1.045	1.313	1.198

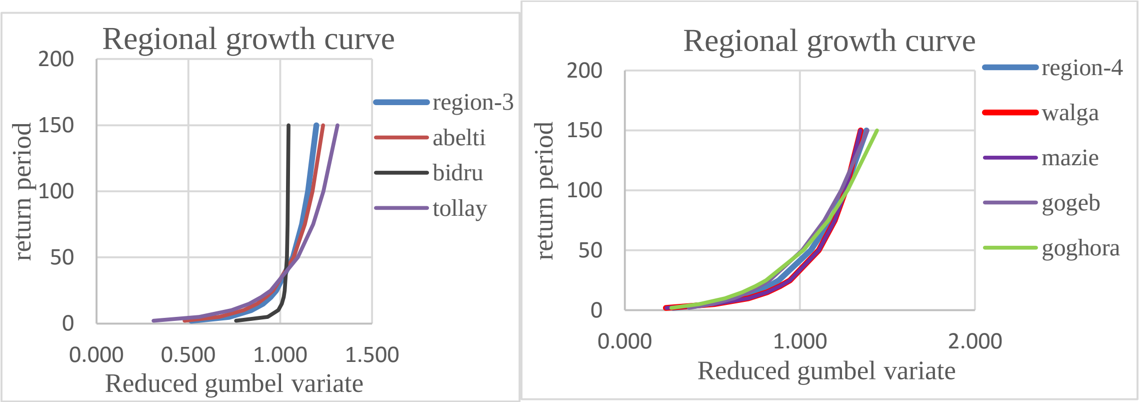
B) Standardized reduced Gumbel variate of region-3

STATION		Walga	Mazie	Gogeb	Goghora	
return P	G.Variate	G.Variate	G.Variate	G.Variate	G.Variate	Reg-4
5	0.516	0.512	0.514	0.500	0.427	0.488
10	0.672	0.706	0.707	0.624	0.573	0.653
15	0.767	0.814	0.814	0.708	0.671	0.752
25	0.890	0.943	0.943	0.826	0.807	0.880
50	1.065	1.108	1.108	1.013	1.021	1.062
75	1.171	1.199	1.199	1.139	1.163	1.175
100	1.248	1.262	1.261	1.236	1.273	1.258
150	1.360	1.347	1.346	1.385	1.440	1.380

c) Standardized reduced Gumbel variate of region-4



(a) Regional growth curve for region-1 (b) Regional growth curve for region-2



(c) Regional growth curve for region-3 (d) Regional growth curve for region-4

Figure 4. 7 Regional growth curve for different return period

Figure 4.7 (b and d) indicated that, the growth curves of Region-2 and region-4, which represents the main reaches of most of the rivers Dincha, Gecha and Sheta near Bonga, Woshi near Dimbira, Gojeb near Shebe, Guma near Andaracha and Walga near wolkite, Mazie near Morka, Gogeb near

Endeber and Goghora near Dana that causes extensive low floods in their lower reaches experiencing low flow generation from the highlands of gibe sub-watershed. The flood that comes from the highlands might inundate low-lying areas in their outfall reaches. Hence, the middle and lower reaches of this region were might be susceptible to the risk of drought. Generally, Figure 4.7 (a), (b), (c) and (d) revealed that lower elevation catchments have lower flood values but higher extreme flood variability than higher elevation catchments. Due to the fact that the flow in different regions has different flow statistics, all curves have different flow characteristics. As indicated in Figure 4.7 (c), the derived regional growth curve of Region-3 was revealed higher quantile estimates than Region-1 and 2, 4 for the same return periods.

4.8.3 Confidence limits of estimated low flow

Confidence limits indicated that the uncertainty of a given estimation of frequency curves. The results of the confidence limit of the study areas at 95% of quantile values for the distribution models were determined as shown for Guma, Woshi, Chida and Limmu genet station on Table 4.13 and Appendix-J for the other station and figure 4.8 for Asendabo and Goghora station.

Table 4.13 Estimated quantiles and Confidence limits of stations (m3/s)

Station				Station			
Guma	UCL	LCI	Flow(m^3/s)	Woshi	UCL	LCI	Flow(m^3/s)
5	1.946325	0.939643	1.442984		4.206978	0.26749	1.266746
10	2.260421	1.253738	1.75708		4.963273	0.49171	2.023041
15	2.449037	1.442355	1.945696		5.572414	0.50805	2.632182
25	2.694955	1.688273	2.191614		6.58547	0.70506	3.645238
50	3.046912	2.04023	2.543571		8.575628	2.695164	5.635396
75	3.263943	2.257261	2.760602		10.19737	4.316902	7.257134
100	3.423367	2.416685	2.920026		11.61911	5.738648	8.67888
150	3.656173	2.64949	3.152831		14.10217	8.221708	11.16194
Station				Station			
Chida	UCL	LCI	Flow(m^3/s)	L. Genet	UCL	LCI	Flow(m^3/s)
5	13.80134	11.63766	12.7195		0.75786	0.682295	0.720077
10	14.95259	12.78891	13.87075		0.798715	0.723149	0.760932
15	15.49757	13.33389	14.41573		0.817891	0.742326	0.780109
25	16.08466	13.92098	15.00282		0.838412	0.762846	0.800629
50	16.74106	14.57738	15.65922		0.861157	0.785591	0.823374
75	17.06418	14.9005	15.98234		0.872264	0.796698	0.834481
100	17.27031	15.10663	16.18847		0.879314	0.803749	0.841531
150	17.53195	15.36828	16.45011		0.888219	0.812654	0.850436

5. CONCLUSION AND RECOMMENDATION

5.1 Conclusions

In this study, regional low flow frequency analysis was performed using the data of 20 stream gauging stations so as to ensure reliable estimation of flow in upper Omo gibe River Basin. The basin has defined and delineated into four hydrologically homogeneous regions using AMF frequency model. The regions were named as Region-1, Region-2, Region-3 and Region-4. The delineation of the regions was done with ArcGIS10.4.1. The discordancy of sites from the region was estimated using Matlab2018a. Further, regional homogeneity tests were conducted to verify the homogeneity of regions. All regions were shown acceptable results for discordancy index and statistical homogeneity tests. Thus, a method of L-moment has found suitable for regional frequency analysis of the study area. Kolmogorov-Smirnov and Anderson-Darling of goodness-of-fit tests were applied and found suitable for checking the adequacy of fitting a suitable distribution for the recorded data of the each catchment. As a result, GEV and GPA were identified as the best fit distributions in the study area with the help of Easy Fit, Matlab and L-MRD. Using the model parameters of the distributions of each station, low flow quantiles were estimated corresponding to different return periods. The study concluded that L-MRD and Matlab Software were acceptable methods for selecting best-fit distribution in upper Omo gibe River Basin. The regional flow frequency curves were significantly different for each regions, which confirmed that the heterogeneity of regions. This variation of curves may be due to the variability of hydrological phenomena of low flow events. This information can be used to design feasible hydrologic projects under prediction uncertainty in both gauged and ungauged catchments. The derived results can be useful as a reference in any hydrological considerations like drought risk management, proper planning, and designing of hydraulic structures such as dams and reservoir during rainy season in the study area.

5.2 Recommendation

1. For the regional estimation of the T-year return period of low flows 7day at the gauged or ungauged locations in upper Omo gibe, the provided equations 2.4 and 2.5 are recommended.
2. Homogeneity test that consider not only higher moments of low flow but also some important basin characteristics should be investigated.
3. Delineation of hydrological homogeneous regions based on statistical parameter of gauged sites could be one of the possible alternative methods of regionalization.
4. Different photographic and climatic parameters like elevation, soil type, geology and hydrogeology indexes, etc. of each region contributed in identifying homogeneous region and the estimates of the mean annual minimum flow will be accurately estimated.
5. Low flow frequency analysis of stations which have zero flows have to be studied further.
6. It is advisable to extend this approach of LFFA for other parts of Ethiopian river basins to establish the homogeneous regions of the country so that problems related of absence of sufficient discharge data for water resources project planning and design could be reduced.
7. Generally for good quality of regional analysis of low flow having data with long record length and no missing or little missing value were recommended.

References

Adane, . A. & Gerd, . F., 2006. Catchment characteristics as predictors of base flow index in Wabi shebele river basin, east Africa. *Research Institute for Water and Environment, Civil Engineering Department University of Siegen, 57068 Siegen, Germany,* Issue For fulfilment of degree of masters.

Amalina, N. M., Shabri, A. & Badyalina,, . B., 2016. Selecting Probability Distribution for Streamflow. *AIP Conference Proceedings 1750, 060014. the American Institute of Physics.*

Arbeláez, A. C., 2009. *Low Flow Discharges Regional Analysis.* Colombia: s.n.

Belete, . B., Yilma , S., Solomon , S. D. & Assefa , M. M., 2015. Flow Regime Classification and Hydrological Characterization: a case study of Ethiopian river. *www.mdpi.com/journal/water,* ISSN 2073-4441(for partial fulfilment of degree of masters).

Benyahya, Daniel, C. & L., 2011. Multivariate analysis of the low-flow regimes in eastern Canadian rivers. *Hydrological Sciences Journal,* ISSN 0706-6457(56), pp. 51-67.

Chaemiso, S. E., Abebe, A. & Murlidhar , S. P., 2016. Assessment of the impact of climate change on surface hydrology. *Model. Earth Syst. Environ.,* (2), p. 205.

Chen D, Y., Huang, G., Shao, . Q. & Xu, C., 2006. Regional analysis of low flow using L-moments for dongjiang basin, south china. *Hydrological Sciences Journal,* pp. 51(6),1051-1064.

Cheng, . L., Yaeger, . M., Viglione, . M. A. & Coopersmith , E., 2012. Exploring the physical controls of regional patterns of flow duration curves Insights from statistical analyses. *Hydrol. Earth Syst. Sci.,* Issue 116, p. 4435–4446.

Cunnane, C., 1985. Unbiased plotting positions review. *Journal of Hydrology,,* (3-4), (37), pp. 205-222.

Dahmen, E. R. & Hall, . M. J., 1990. Screening of Hydrological Data: Tests for Stationarity and Relative Consistency.. *Wageningen, The Netherlands. International Institute for land reclamation and improvment.*

Dessalegn, B., Ayalew, . S. & Hailu, D., 2016. Evaluation of Extreme Flow QuantEstimated from Global Reanalysis Runoff data. A Case Study on Blue Nile River. Issue MSc. Thesis Submitted to Addis Ababa University..

Desta, A. G., 2017. Evaluation of the impact of climate change on omo gibe river basin. Issue Addis Ababa Institute Technology for Msc.

Eloy, K., 2018. Frequency analysis of minimum flows. *Brazilian Journal of Water Resources,* Volume https://doi.org/10.1590/2318-0331.0318170080: Versão On-line ISSN 2318-0331, p. 2.

EPA, 2018. *Low flow statistics,* USA: EPA-833-B-18-001.

Fleming, G. W. &. M. J., 2010. *Manual of HEC-SSP,* USA: USA Corps of Engineers.

Gebrehiwot, S. G., Ilstedt , U. & Gardenas, A. . I., 2011. Hydrological characterization of watersheds in the Blue Nile Basin, Ethiopia. *Hydrol. Earth Syst. Sci.,* Issue 15, p. 11–20.

Gedefa, T., 2009. Regional Flood Frequency analysis for upper omo gibe river basin. Issue Addis Ababa University for degree of masters.

Gordon, N. D., 2004. *stream hydrology book.* Second Edition ed. British: British Library Cataloguing in Publication Data.

Goyal, K. M., 2016. Engineering hydrology. *PHI learning private limited,* Volume ISBN-978-81-203-5243-8.

Grandry , M. & Gailliez, . S., 2013. A method for low-flow estimation at ungauged sites. *Hydrol. Earth Syst. Hydrol. Earth Syst. Sci.,* doi:10.5194/hess-17-1319-2013(17), p. 1319–1330.

Guastard, A., 2009. *Fedral Institute of hydrology,* Geneva: World Metrological Organization (WMO).

Hamadi, k., Samaneh, . A. & Zohrab, . N., 2013. Frequency Analysis of Low-Flows. *International Research Journal of Applied and Basic Sciences,* ISSN 2251-838X / Vol, 4 (1), pp. 146-151.

Hosking, M. J. R. & Wallis, J. R., 1997. *RegionaFl requencyA nalysis,* Newyork: ISBN10-0-0-521-01940-0.

Irwin, . S. K., Srivastav, . R. & Simonovic, . S. P., 2014. *Instruction for Watershed Delineation in an ArcGIS Environment for Regionalization ,* s.l.: s.n.

Joshi, D. & André , S.-H., 2013. Low Flow Frequency analysis. *Statistical Hydroclimatology Research Group.*

Ketsela , H., Tadele , K. & Temam, D., 2017. An Assessment of Predictive Accuracy for Regional flood frequency analysis the case of awash catchment. Issue MSc. Thesis Submitted to Jimma University for degree of masters.

Kidist, . A., Mamaru , . A. & Moges, 2018. Low Flow Trends and Frequency Analysis. *Journal of Water Resource and Protection,* ISSN Print: 1945-3094, ISSN Online: 1945-3108(10), pp. 182-203.

Laaha, G. & Bl¨oschl, . G., 2006. A comparison of low flow regionalisation methods. *J. Hydrol.,* doi:10.1016/j.jhydrol.2005.09.001(323), p. 193–214.

Larwin, N. & Khin Maung, w., 2014. Low Flow Frequency Analysis on Selected Rivers in Malaysia. *International Journ. of earth science,* Volume Volume 07, No. 04, pp. P.P.1315-1321.

Malekinezhad, H., Nachtnebel, . H. P. & Klik, . A., 2011. Regionalization Approach for Extreme Flood Analysis Using L-moments. *J. Agr. Sci. Tech.,* Vol. 13: 1183-1196(13), pp. 1183-1196.

McCuen, R. H., 1998. *Hydrologic analysis and design.* second edition ed. United States of America (USA): s.n.

McMahon , . T. A., Diaz , A. & Arenas, 1982. Methods of computation stream flow. I S B N 92-3-102 013-7(7).

Mélanie , T., Pierre, L., Doucet, G. & Robert , L., 2017. Assessing River Low-Flow Uncertainties Related to Hydrological Model Calibration. *www.mdpi.com/journal/climate,* doi:10.3390/cli5010019(5), p. 19.

Modi, . M. & Mitra, A., 2017. Regional Flood Frequency Analysis:. *International Journal of Engineering Science Invention.,* ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726(6(7)), pp. 55-65.

Murphy, C., Cunnane, . C., Das, S. & Mandal, . U., 2014. Flood Frequency Estimation.. *Flood Studies Update:,* Volume Volume II.

Nancy , D. . G., Thomas , . A. M. & Brian, . L. . F., 2007. *Stream Hydrology to An Introduction for Ecologists.* second edition ed. British: The University of Melbourne.

Paix, a. d. l., 2009. *Guide to hydrological practice.* sixth ed. Geneva: World Meteorological Organization.

Pamela, 1992. Techneque of estimating 7-day low flow,. *U.S. Geological Survey,* pp. Water-Resources Investigations Report 91-4130.

Ramakar, . J. & Smakhtin, V., 2008. A Review of Methods of hydrological estimation at ungauged site. *(International Water Management) InstituteIWMI,* ISBN 978-92-9090-693-3(3).

Rao, A. a. S. V., 2008. Regionalization of Watersheds by Hybrid Cluster Analysis.. *Journal of Hydrology.,* Volume ISBN: 978-1-4020-6851-5/ ISBN: 978-1-4020-6852-2.

Richard, . K., Wilson, . A. & Nicholas , K., 2016. Estimation of Flow-Duration and Low-Flow Frequency parameter. *American Scientific Research Journal for Engineering, Technology, and Sciences (ASRJETS),* ISSN (Print) 2313-4410, ISSN (Online) 2313-4402(1), pp. 62-75.

Rossi , G. & Caporali , E., 2010. Regional analysis of low flow in Tuscany (Italy). *Global Change: Facing Risks and Threats to Water Resources,* Volume giuseppe.rossi@dicea.unifi.it, p. 135:141.

Searcy, k. J. & Hardison, H. C., 1966. Manual of hydrology. *Method and practices of geology survey,* 3(3), pp. 31-40.

Smakhtin, V., 2001. Low flow hydrology: a review. *Journal of Hydrology,* Issue 240, p. 147–186.

Tatek, W., 2015. Estimation of low flow quantiles in ungauged river chatchment. p. 18.

Temesgen , E., Adugnaw , T. & Debebe, . L. Y., 2014. *Biohydrology of low flows in the humid Ethiopian highlands:a case of Gilgel abay cathment,* Adis Ababa: School of Civil and Water Resources Engineering, Bahir Dar Institute of Technology, Bahir Dar University, Ethiopia.

Tsakiris, G., Nalbantis, . I. & Cavadias , G., 2011. Regionalization of low flows based on Canonical Correlation Analysis,. *Adv. Water Resour.,* Issue 34, p. 865–872.

Tsedey, T., 2007. Regionalization of flow duration curve for for Abay river basin. Issue AAit for partial fulfilment masters degree.

USA department of agriculture, 2007. *Stream Hydrology.* united state: U.S. Department of Agriculture.

Vezza, P., Comoglio, C., Rosso, . M. & V., 2010. Low flows regionalization in North-Western Italy. *Water Resour. Manag.,* Issue 24, p. 4049–4074.

Winlar & Khin, M., 2014. Low Flow Frequency Analysis. *international journal of earth science and engineering,* ISSN 0974-5904, Volume 07(No. 04), pp. 1315-1321.

WMO, 2009. Management of Water Resources and Application of Hydrological Practices. *Guide To Hydrology Practice,* ISBN 978-92-63-11029-9(183).

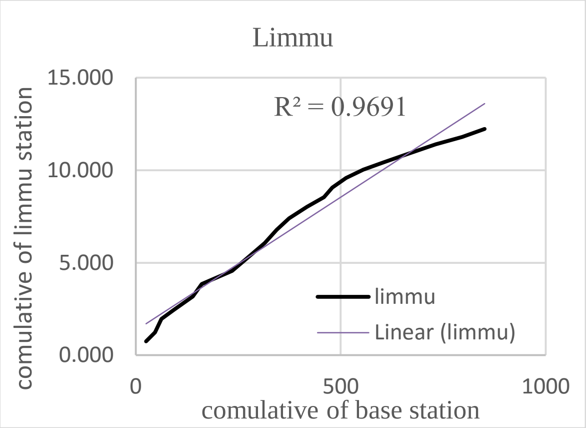
Yongkin, D. C., Guoru, H., Quanxi, S. & Chong-Yu, X., 2010. Regional analysis of low flow using L-moments. *Hydrological Sciences Journal,* ISSN: 0262-6667 (Print) 2150-3435 (Online) : DOI: 10.1623/hysj.51.6.1051(51), pp. 1051-1064.

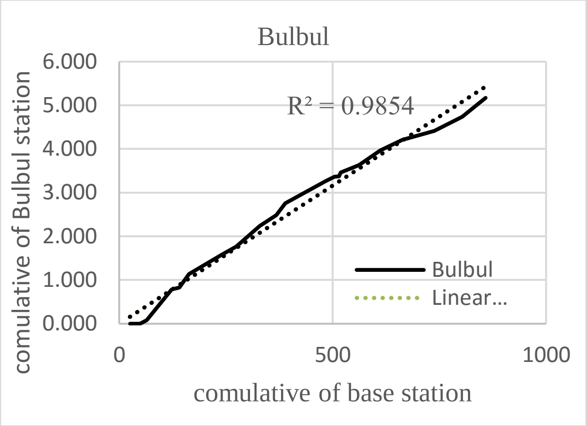
Zadeh, S. M., 2012. Low Flow Frequency Study for Newfoundland and Labrador. *Faculty of Engineering and Applied Science, Memorial University of Newfoundland,* Issue A Thesis submitted to the School of Graduate Studies for degree of masters.

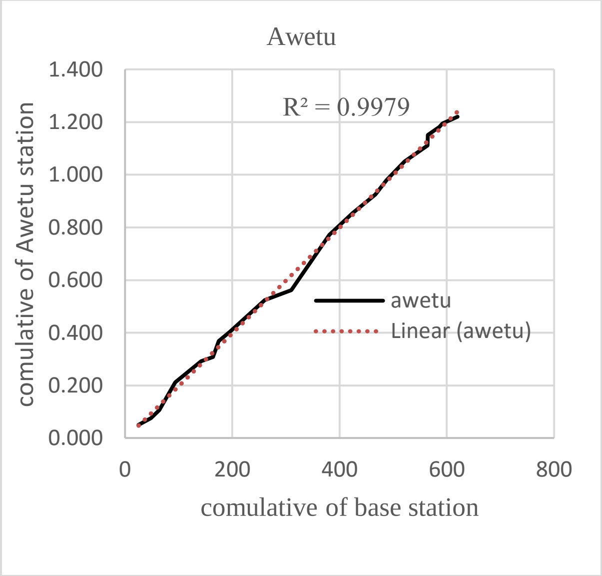
Appendices

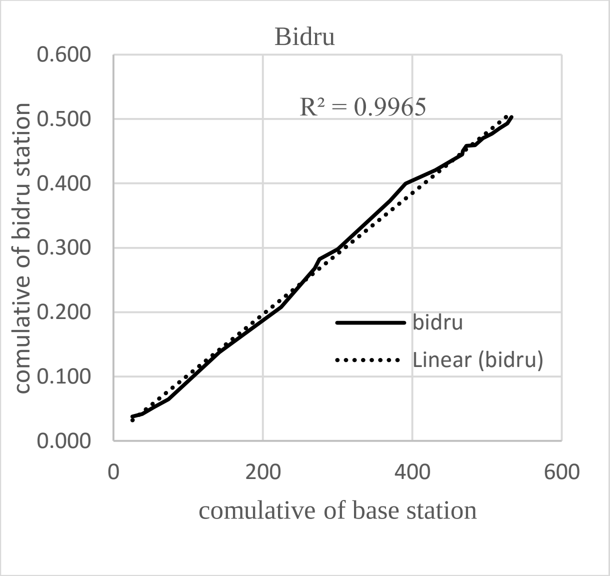
Appendix-A

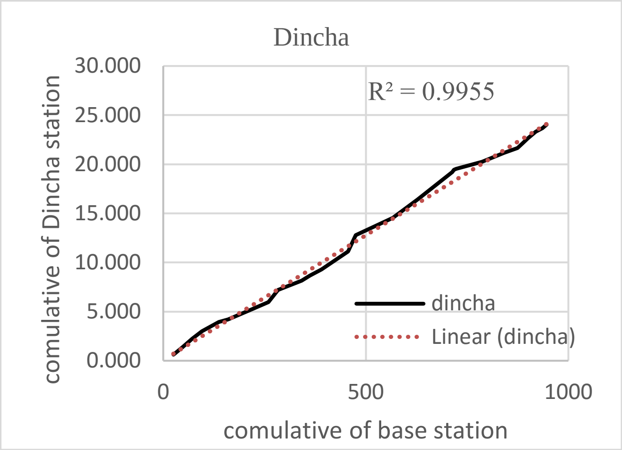
This appendix A represents the double mass curve of the representative station selected for the analysis. In addition all the graph shows that the same property: as stream flow of one station increase stream flow of the other station were also increased to show consistency and stationarity.

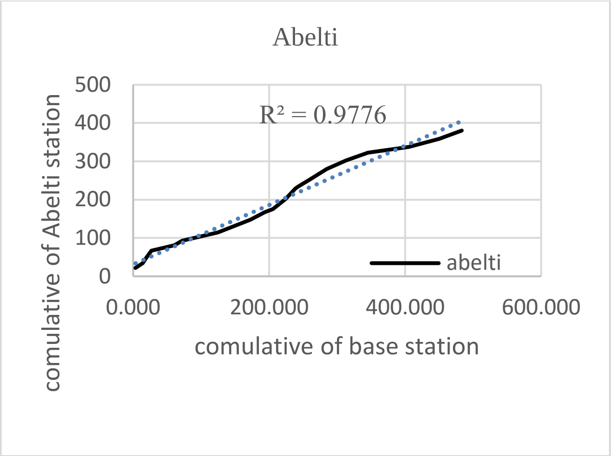


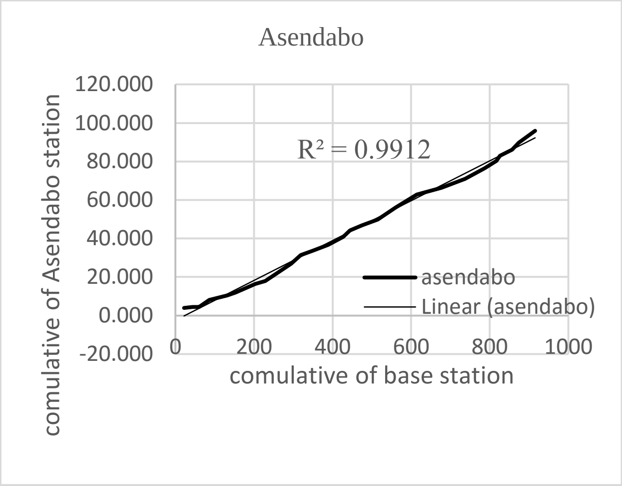


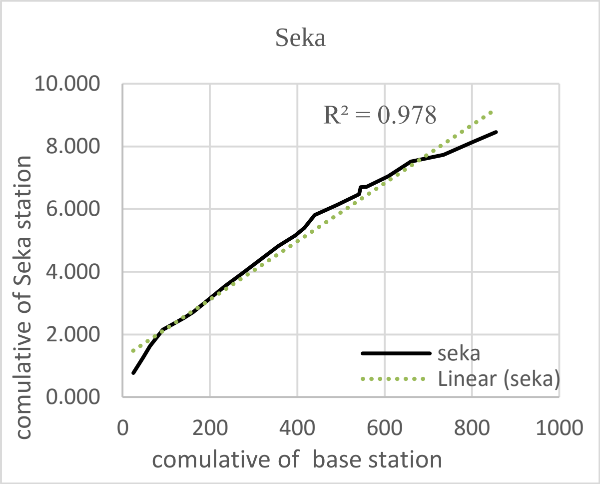


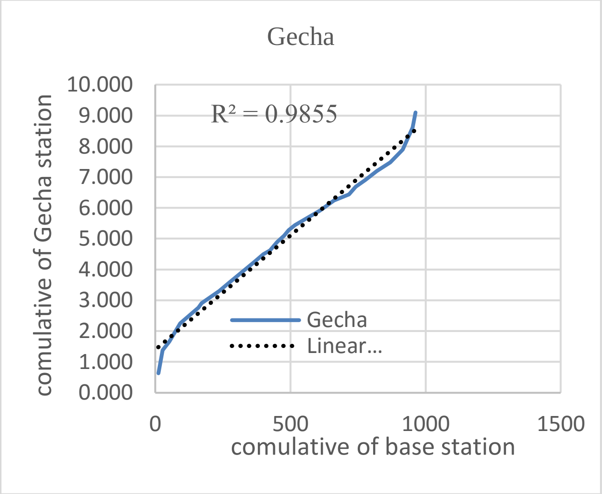


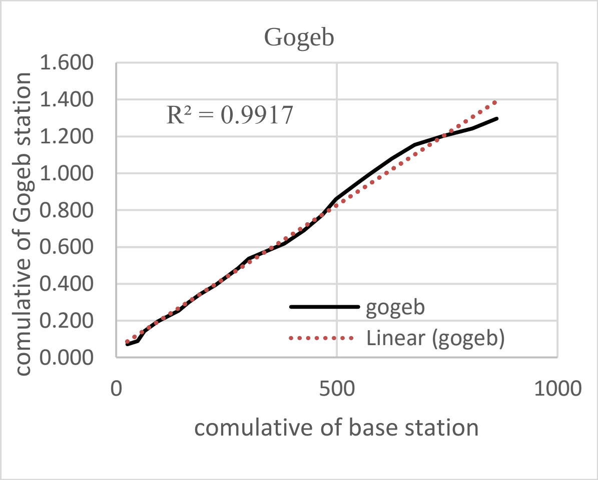

Dincha
R² = 0.9955
comulative of Dincha station
30.000
25.000
20.000
15.000
10.000
5.000
0.000
0
500
1000
comulative of base station
dincha
Linear (dincha)

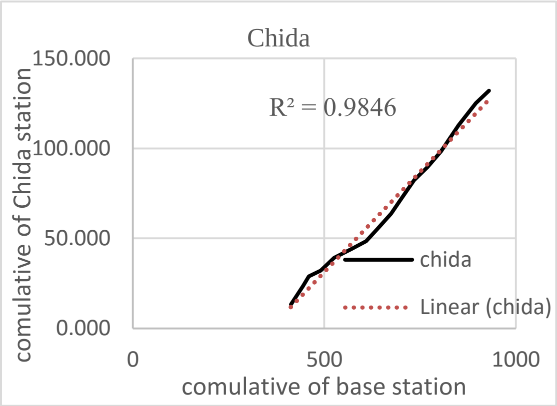

Abelti
R² = 0.9776
comulative of Abelti station
500
400
300
200
100
0
0.000
200.000
400.000
600.000
comulative of base station
abelti

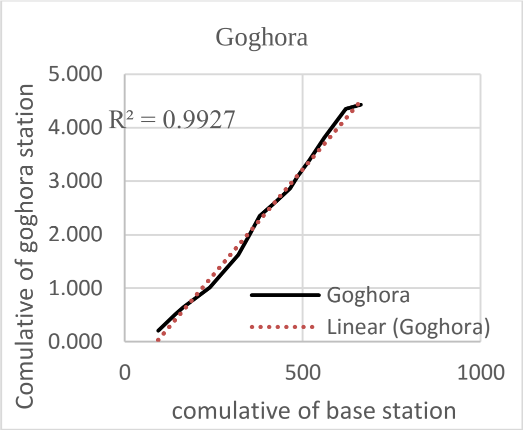

Asendabo
R² = 0.9912
comulative of Asendabo station
120.000
100.000
80.000
60.000
40.000
20.000
0.000
-20.000
0
200
400
600
800
1000
comulative of base station
asendabo
Linear (asendabo)

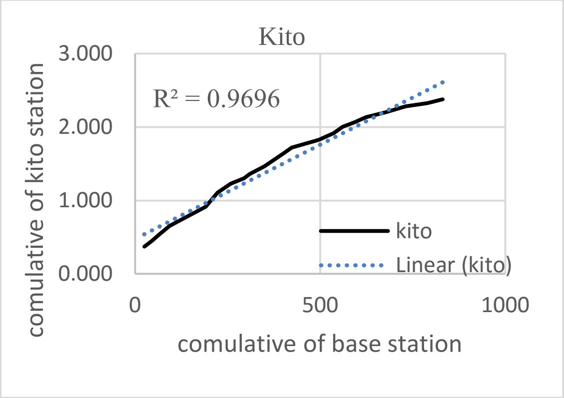

Seka
R² = 0.978
comulative of Seka station
10.000
8.000
6.000
4.000
2.000
0.000
0
200
400
600
800
1000
comulative of base station
seka
Linear (seka)

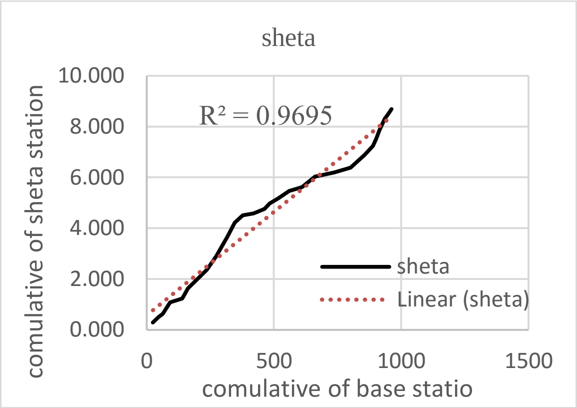

Gecha
R² = 0.9855
comulative of Gecha station
10.000
9.000
8.000
7.000
6.000
5.000
4.000
3.000
2.000
1.000
0.000
0
500
1000
1500
comulative of base station
Gecha
Linear...

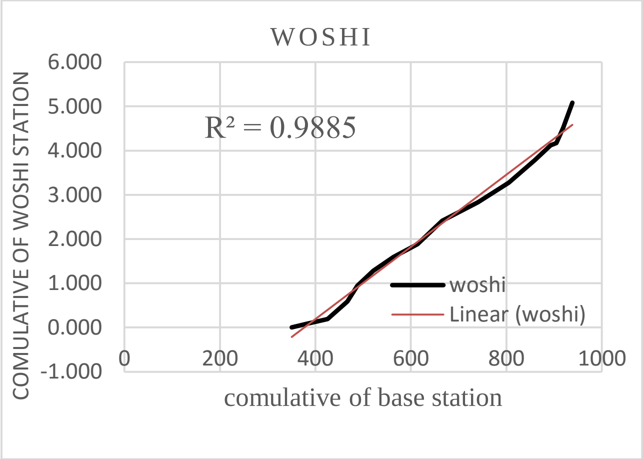

Gogeb
R² = 0.9917
comulative of Gogeb station
1.600
1.400
1.200
1.000
0.800
0.600
0.400
0.200
0.000
0
500
1000
comulative of base station
gogeb
Linear (gogeb)

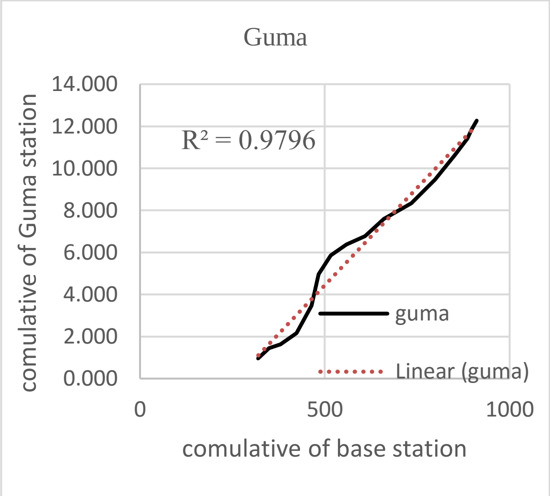

Chida
150.000
100.000
50.000
0.000
comulative of Chida station
R² = 0.9846
chida
Linear (chida)
0
500
1000
comulative of base station


Goghora
5.000
4.000
3.000
2.000
1.000
0.000
Comulative of goghora station
R² = 0.9927
Goghora
Linear (Goghora)
0
500
1000
comulative of base station


Kito
3.000
2.000
1.000
0.000
comulative of kito station
R² = 0.9696
kito
Linear (kito)
0
500
1000
comulative of base station

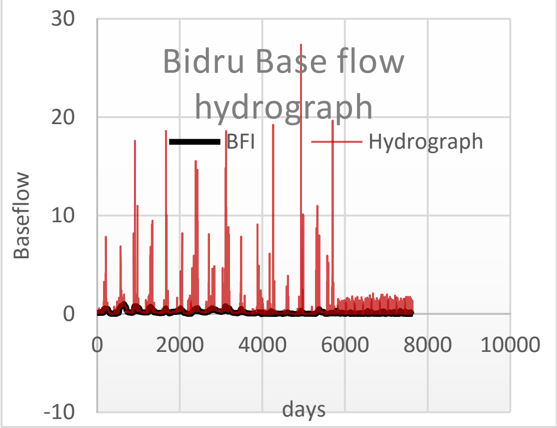

sheta
10.000
8.000
6.000
4.000
2.000
0.000
comulative of sheta station
R² = 0.9695
sheta
Linear (sheta)
0
500
1000
1500
comulative of base statio

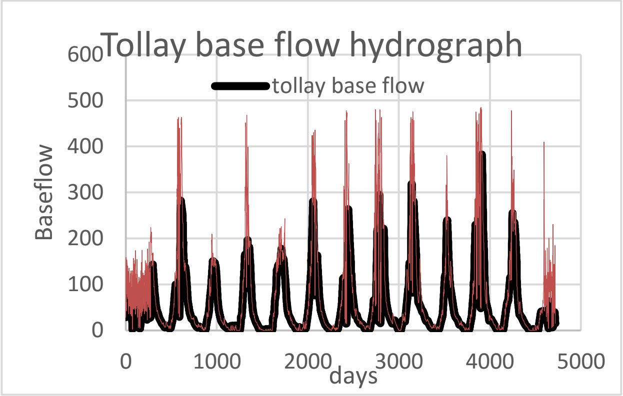

WOSHI
6.000
5.000
4.000
3.000
2.000
1.000
0.000
-1.000
COMULATIVE OF WOSHI STATION
R² = 0.9885
woshi
Linear (woshi)
0
200
400
600
800
1000
comulative of base station

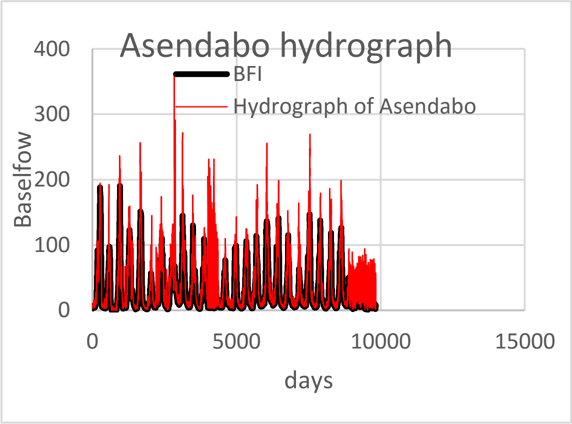

Guma
14.000
12.000
10.000
8.000
6.000
4.000
2.000
0.000
comulative of Guma station
R² = 0.9796
guma
Linear (guma)
0
500
1000
comulative of base station

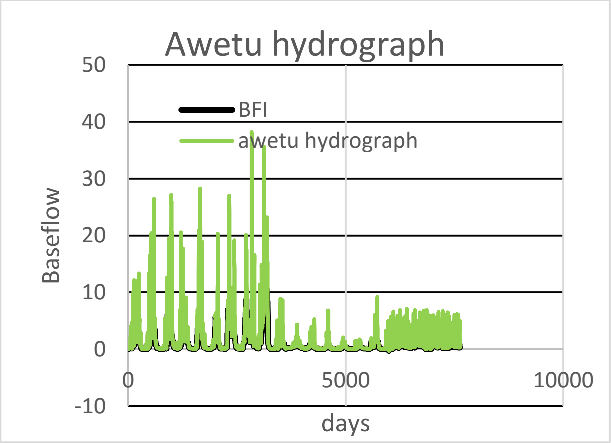
Appendix- B

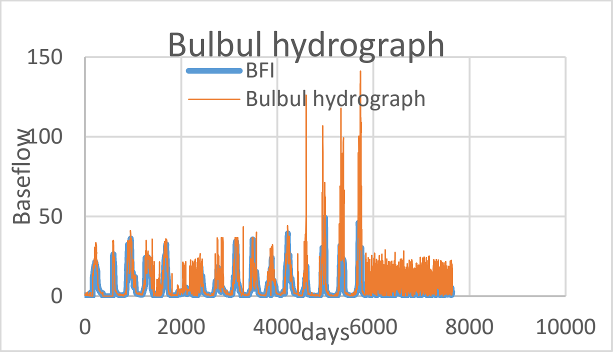
This appendix B (a-t) shows that hydrograph and base flow of all station contributed in this thesis. As shown on figure the property of the hydrograph indicated were shows the same property.

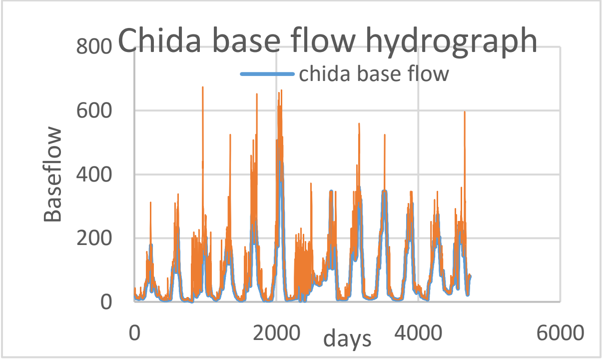


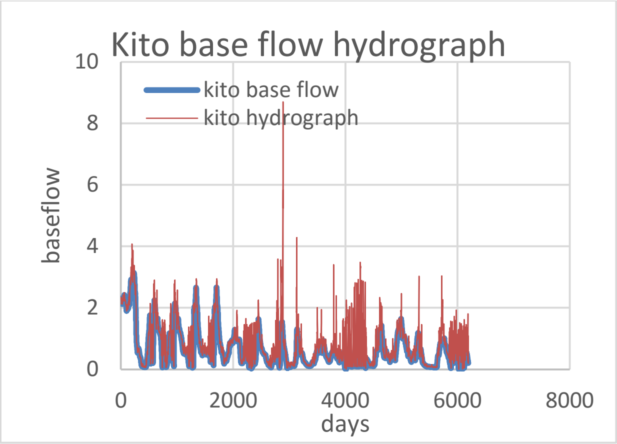


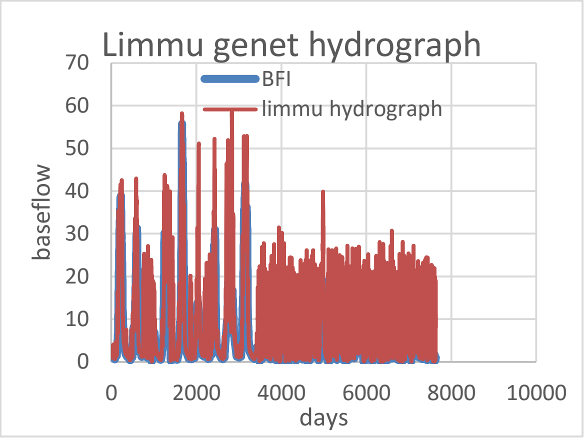


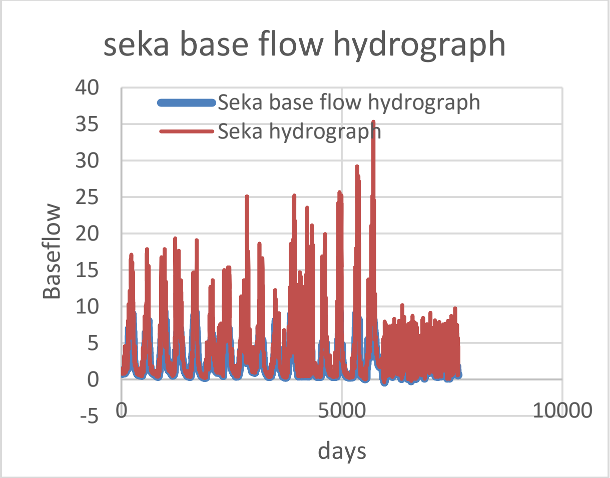


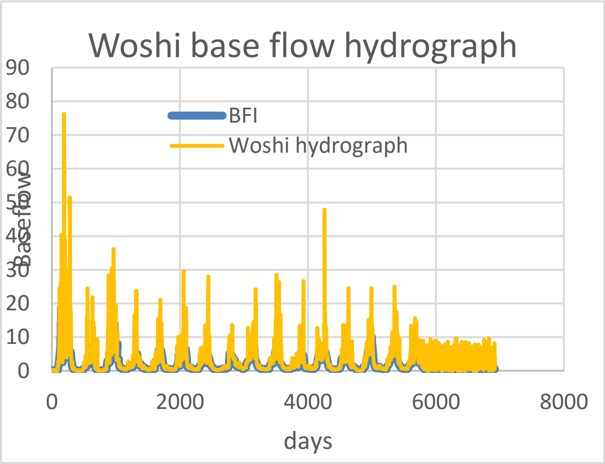


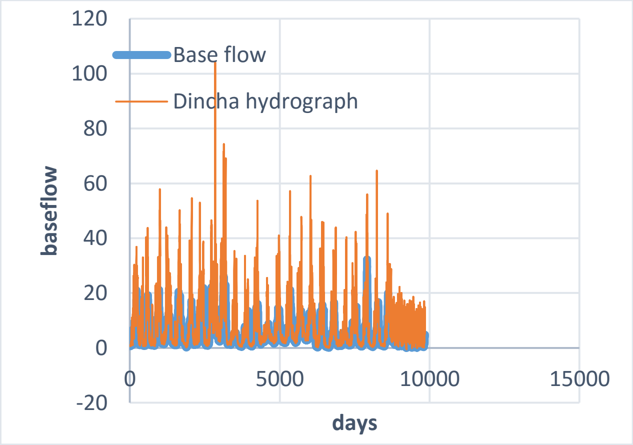


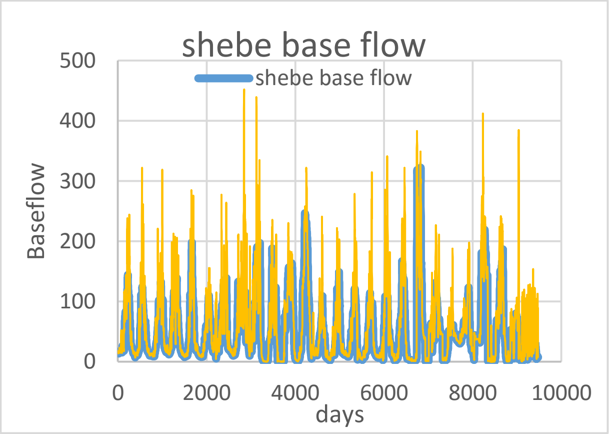

Kito base flow hydrograph
kito base flow
kito hydrograph
baseflow
10
8
6
4
2
0
0
2000
4000
6000
8000
days

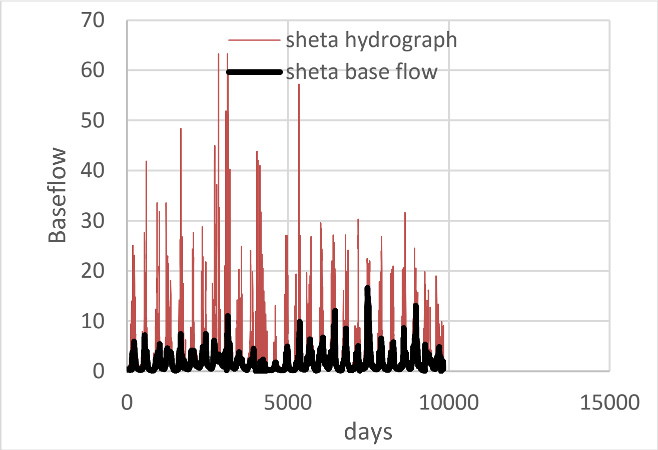

Limmu genet hydrograph
BFI
limmu hydrograph
baseflow
70
60
50
40
30
20
10
0
0
2000
4000
6000
8000
10000
days

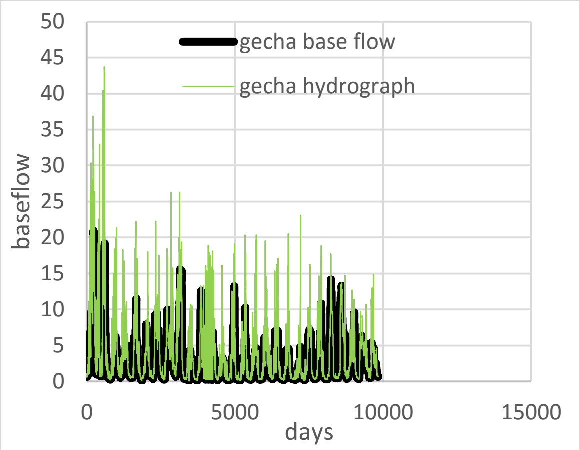

seka base flow hydrograph
Seka base flow hydrograph
Seka hydrograph
Baseflow
40
35
30
25
20
15
10
5
0
-5
0
5000
10000
days

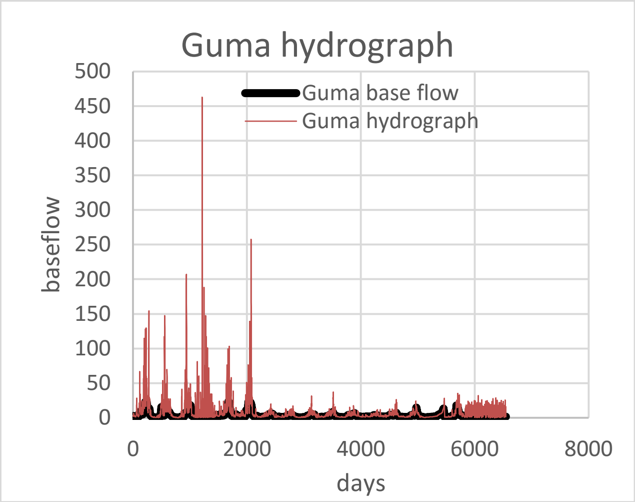

Woshi base flow hydrograph
BFI
Woshi hydrograph
Baseflow
90
80
70
60
50
40
30
20
10
0
0
2000
4000
6000
8000
days

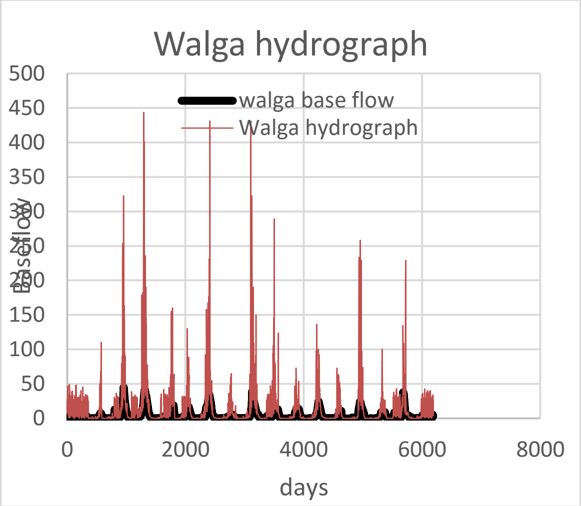

Base flow
Dincha hydrograph
baseflow
120
100
80
60
40
20
0
-20
0
5000
10000
15000
days

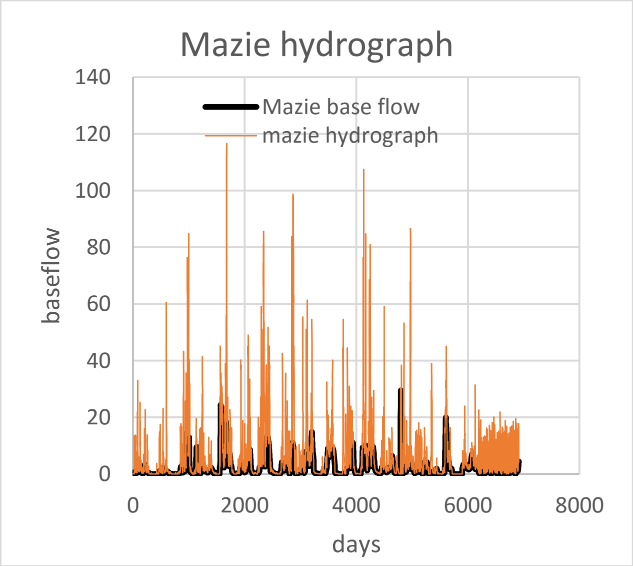

shebe base flow
shebe base flow
Baseflow
500
400
300
200
100
0
0
2000
4000
6000
8000
10000
days

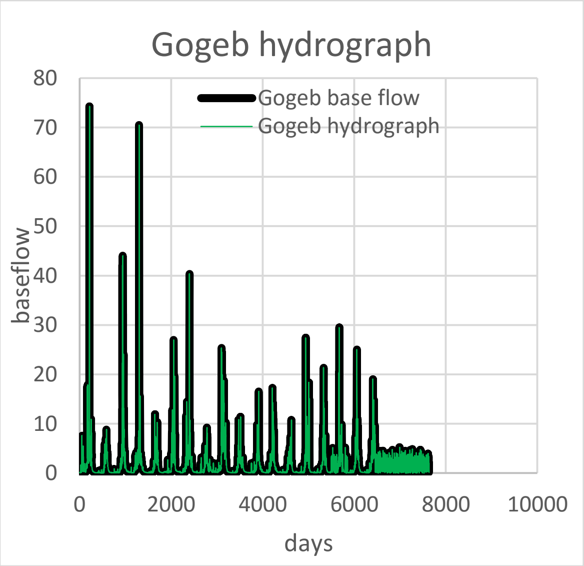

sheta hydrograph
sheta base flow
Baseflow
70
60
50
40
30
20
10
0
0
5000
10000
15000
days


gecha base flow
gecha hydrograph
baseflow
50
45
40
35
30
25
20
15
10
5
0
0
5000
10000
15000
days


Guma hydrograph
Guma base flow
Guma hydrograph
baseflow
500
450
400
350
300
250
200
150
100
50
0
0
2000
4000
6000
8000
days


Walga hydrograph
walga base flow
Walga hydrograph
Baseflow
500
450
400
350
300
250
200
150
100
50
0
0
2000
4000
6000
8000
days


Mazie hydrograph
Mazie base flow
mazie hydrograph
baseflow
140
120
100
80
60
40
20
0
0
2000
4000
6000
8000
days


Gogeb hydrograph
Gogeb base flow
Gogeb hydrograph
baseflow
80
70
60
50
40
30
20
10
0
0
2000
4000
6000
8000
10000
days

Appendix –C

This appendix was the Matlab program that is used to test the heterogeneity or discordancy of the delineated regions provided by Hosking and Wallis, 1997).

```
U=BULBULS5 (:,:); file name contains (CV, Cs and Cku) for delineated site

n=7; %input ('enter the number of sites in the group :');

ubar =[0;0;0];

for I=1:n

Ubar=ubar+1/n*(U (I, 1:3)');

End

S=zeros (3);

For I=1: n

S=S+ (U (I, 1:3)'-ubar)*(U (I, 1:3)'-ubar)';

End

For I=1: n

Di (I) =1/3*n*(U (I, 1:3)'-ubar)'*inv(S)*(U (I, 1:3)'-ubar);

End
```

Appendix- D Quantile for all station by GEV and GPA by selected distribution

Bulbul and Guma Station

Q(m^3/s) XLSTST	Q(m^3/s) EASYFIT	Q(m^3/s) Matlab	Q(m^3/s) xlstat	Q(m^3/s) EASYFIT	Q(m^3/s) Matlab	Return P.(year)
0.338295	0.372164	0.35451	12.66875	10.67914	10.41413	2
0.475319	0.590475	0.546419	17.38499	13.37497	12.7195	5
0.557861	0.746378	0.679924	20.22601	14.792	13.87075	10
0.601769	0.838585	0.757606	21.73728	15.4845	14.41573	15
0.631425	0.905043	0.813036	22.75798	15.92865	14.75856	20
0.653675	0.957326	0.856325	23.52382	16.24959	15.00282	25
0.71911	1.124563	0.992994	25.77604	17.13334	15.65922	50
0.755081	1.226212	1.074786	27.01413	17.58168	15.98234	75
0.779666	1.300205	1.133749	27.86031	17.87312	16.18847	100
0.813075	1.407278	1.218246	29.01019	18.24997	16.45011	150

Kito and Awetu Station

Q(m^3/s) xlstat	Q(m^3/s EASYFIT	Q(m^3/s Matlab	Return P (year)	Q(m^3/s) XLSTST	Q(m^3/s) EASYFIT	Q(m^3/s) Matlab
0.123321	0.101733	0.100721	2	0.078902	0.068902	0.059969
0.162718	0.117265	0.114266	5	0.113844	0.133844	0.076331
0.18645	0.124045	0.119813	10	0.134893	0.154893	0.084463
0.199075	0.126998	0.122134	15	0.14609	0.15609	0.088302
0.207601	0.128763	0.123489	20	0.153652	0.163652	0.090712
0.213998	0.129976	0.124404	25	0.159326	0.169326	0.092427
0.232812	0.133035	0.126645	50	0.176012	0.176012	0.097026
0.243154	0.134427	0.127626	75	0.185185	0.195185	0.099283
0.250223	0.135272	0.128209	100	0.191454	0.21454	0.100721
0.259828	0.136295	0.128898	150	0.199974	0.29974	0.102542

Limmu Genet and Sheta station

Q(m^3/s) xlstat	Q(m^3/s) EASYFIT	Q(m^3/s) Matlab	Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	Return P. (year)
0.717695	0.676045	0.637141	0.402655	0.386208	0.375111	2
0.88272	0.801163	0.720077	0.538491	0.517642	0.491559	5
0.982129	0.872512	0.760932	0.620317	0.5983	0.560793	10
1.03501	0.909193	0.780109	0.663844	0.641705	0.597322	15
1.070725	0.933457	0.79211	0.693242	0.671228	0.621871	20
1.097522	0.951388	0.800629	0.715299	0.693494	0.640224	25
1.176329	1.002736	0.823374	0.780166	0.759571	0.693853	50
1.21965	1.030065	0.834481	0.815825	0.796294	0.723108	75
1.249259	1.04837	0.841531	0.840197	0.821566	0.743005	100
1.289494	1.072751	0.850436	0.873315	0.856142	0.769915	150

Asendabo and Shebe station

Q(m^3/s) XLSTAT	Q(m^3/s) EASYFIT	Q(m^3/s) Matlab	Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	Return P. (year)
6.172663	7.016302	7.247662	10.12856	9.354925	9.706885	5
7.090587	8.604267	8.969768	12.22601	11.15645	11.61701	10
7.574495	9.552524	10.01212	13.34174	12.0956	12.61821	15
8.142431	10.78324	11.37994	14.6607	13.188	13.78786	25
8.851746	12.53426	13.35399	16.32345	14.53661	15.23999	50
9.238361	13.60829	14.58044	17.2375	15.26392	16.02717	75
9.501188	14.39461	15.48562	17.86221	15.75509	16.56049	100
9.856458	15.53902	16.81373	18.71114	16.41467	17.27896	150

Dincha and Gecha station

Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	Return P.(year)
1.205983	3.125216	1.06087	0.39441	0.374861	0.364279	2
1.676397	5.083861	1.409929	0.513201	0.471773	0.447801	5
1.959769	6.224098	1.609048	0.58476	0.527112	0.493254	10
2.110508	6.817882	1.711438	0.622825	0.555585	0.515952	15
2.212317	7.213746	1.779174	0.648535	0.57443	0.530702	20
2.288704	7.507968	1.829237	0.667824	0.588361	0.541463	25
2.513347	8.358972	1.972619	0.724552	0.62828	0.571591	50
2.636839	8.817461	2.048948	0.755737	0.649544	0.587191	75
2.721239	9.126895	2.100079	0.77705	0.663793	0.597462	100
2.835933	9.542163	2.168191	0.806013	0.682782	0.610911	150

Woshi and Seka station

Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	Return P. (year)
0.508815	1.19369	1.266746	0.620813	0.637473	0.760469	5
0.53119	1.838367	2.023041	0.71069	0.721221	0.921375	10
0.539831	2.337432	2.632182	0.758501	0.76473	1.016261	15
0.544637	2.762435	3.163946	0.790791	0.793695	1.084526	20
0.547769	3.140171	3.645238	0.815019	0.8152	1.13816	25
0.555	4.649938	5.635396	0.88627	0.877287	1.309312	50
0.557937	5.834726	7.257134	0.925438	0.910662	1.41305	75
0.559599	6.848713	8.67888	0.952207	0.933156	1.488431	100
0.561477	8.576663	11.16194	0.988585	0.963301	1.597322	150

Quantile by Generalized Pareto

Abelti and Bidru station

Return P.(year)	Q(m^3/s) XLSTST	Q(m^3/s) EASYFIT	Q(m^3/s) Matlab	Q(m^3/s) XLSTST	Q(m^3/s) EASYFIT	Q(m^3/s) Matlab
2	24.14644	23.60586	22.80621	0.050227	0.036777	0.020853
5	34.14894	34.0896	31.82552	0.082349	0.054988	0.025593
10	41.12945	41.69063	38.0938	0.104766	0.065966	0.027181
15	44.99403	46.00992	41.55424	0.117177	0.071465	0.027712
20	47.6425	49.01901	43.92152	0.125682	0.075004	0.027978
25	49.64498	51.32183	45.70904	0.132113	0.077559	0.028138
50	55.58777	58.30474	51.00152	0.151197	0.084548	0.028459
75	58.87783	62.27279	53.92326	0.161763	0.088048	0.028566
100	61.13258	65.03718	55.92201	0.169004	0.090301	0.02862
150	64.20233	68.8629	58.6384	0.178862	0.093182	0.028673

Tollay and Mazie station

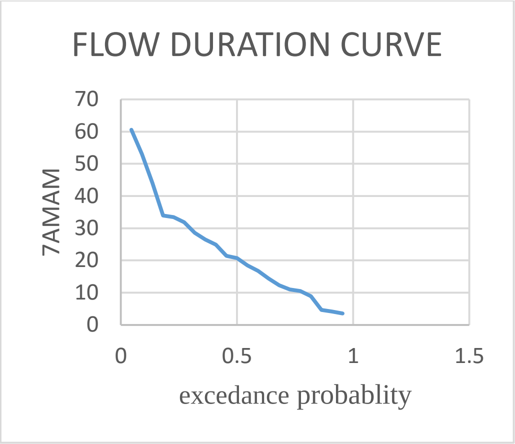
Return P.(year)	Q(m^3/s) XLST	Q(m^3/s) EASY	Q(m^3/s) Matlab	Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab
2	1.109904	0.550236	0.582578	0.023099	0.067145	0.056006
5	2.004646	0.832411	1.021267	0.050058	0.098477	0.085761
10	2.629066	0.979773	1.310949	0.068872	0.112182	0.102991
15	2.974759	1.046669	1.465375	0.079287	0.11769	0.111388
20	3.21167	1.087157	1.568722	0.086426	0.120781	0.116699
25	3.390795	1.115097	1.645507	0.091823	0.122798	0.120485
50	3.922389	1.185769	1.866428	0.10784	0.127442	0.130616
75	4.216691	1.21785	1.984198	0.116707	0.129308	0.135553
100	4.418383	1.237267	2.063013	0.122784	0.130356	0.138675
150	4.692977	1.260645	2.167811	0.131058	0.131527	0.142603

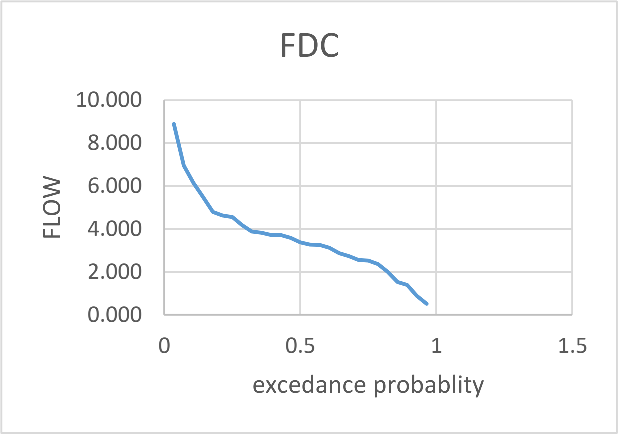
Walga and Gogeb station

Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	Q(m^3/s) xlstat	Q(m^3/s) easyfit	Q(m^3/s) Matlab	Return P.(year)
0.035823	0.137963	0.117856	0.067724	0.074196	0.070639	2
0.078303	0.201091	0.148042	0.083245	0.101515	0.087371	5
0.107949	0.226906	0.159337	0.094078	0.126812	0.09841	10
0.124362	0.236812	0.163414	0.100075	0.143797	0.104292	15
0.13561	0.242213	0.165555	0.104184	0.156947	0.108227	20
0.144115	0.245667	0.166886	0.107292	0.167825	0.11115	25
0.169353	0.253338	0.169711	0.116514	0.205804	0.119555	50
0.183326	0.256282	0.170731	0.121619	0.231304	0.124033	75
0.192902	0.257887	0.171266	0.125118	0.251046	0.127029	100
0.205939	0.259633	0.171828	0.129882	0.281422	0.131011	150

Appendix-E

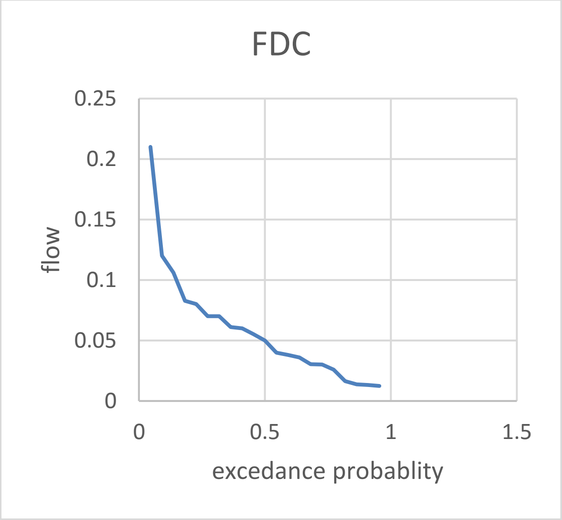
This appendix shows the property of flow duration curve that were extracted from 7-days AMF of stations to shows the probability of flows to occur. For all section the flow duration curve were represented on figure below (a-r) where as for Chida and Woshi station were discussed on section 4.2.

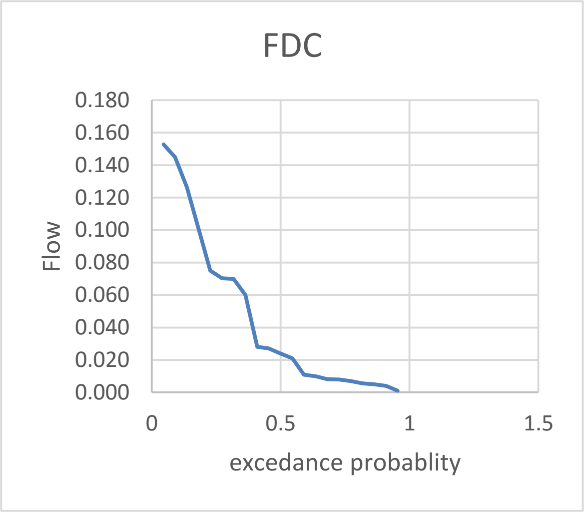




(a) Abelti station

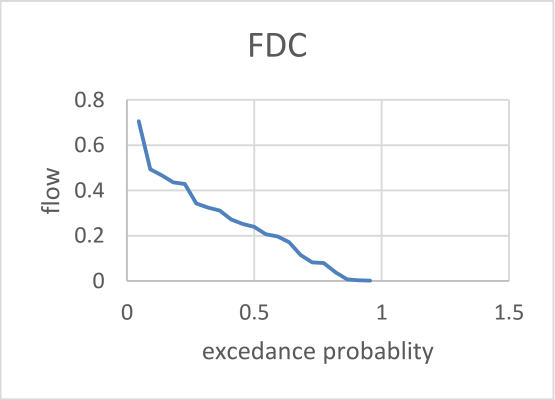
(b) Asendabo station

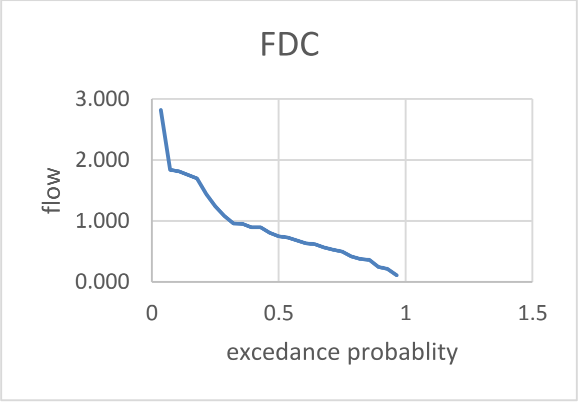




(c) Awetu station

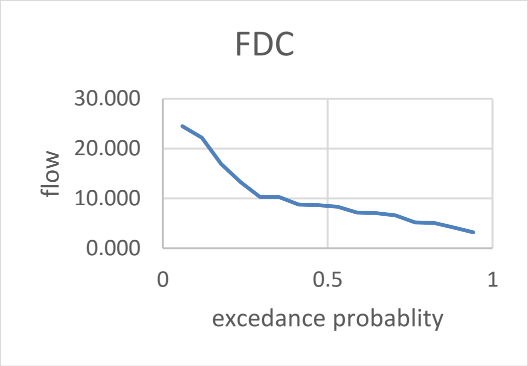
(d) Bidru station



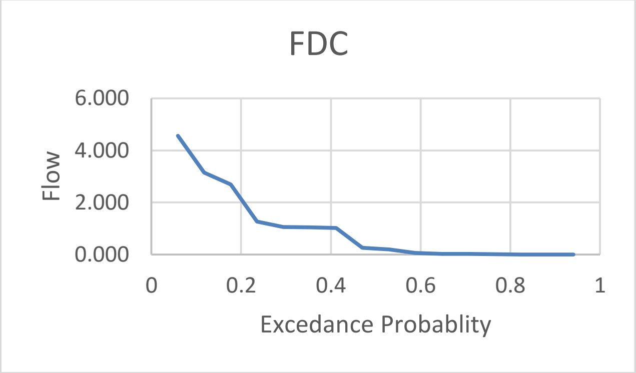


(e) Bulbul station

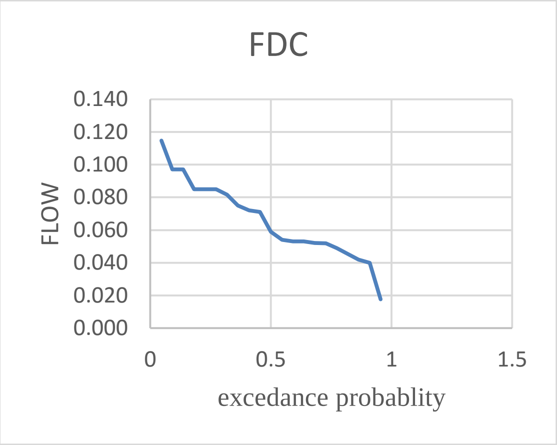
(f) Dincha station



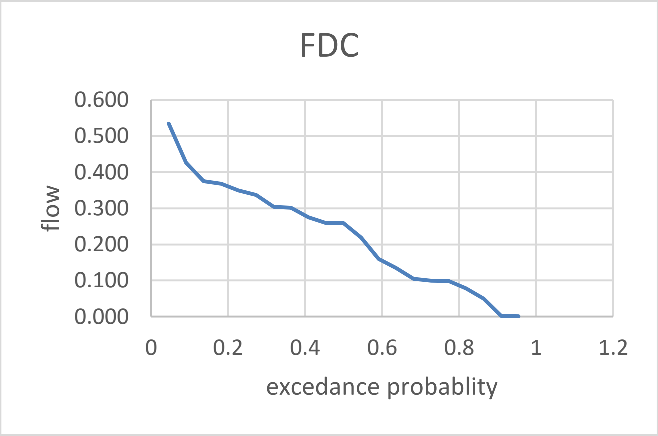
(g) Chida station



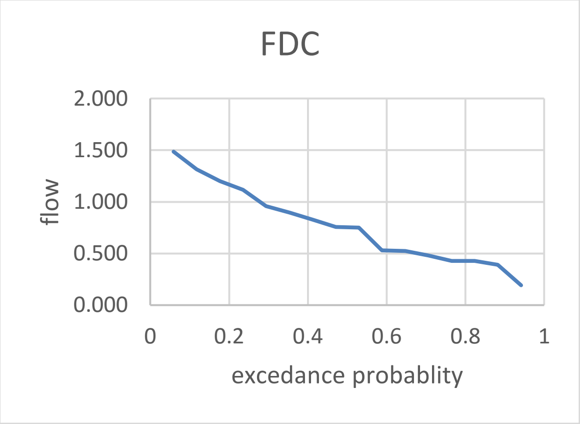
(h) Tollay station



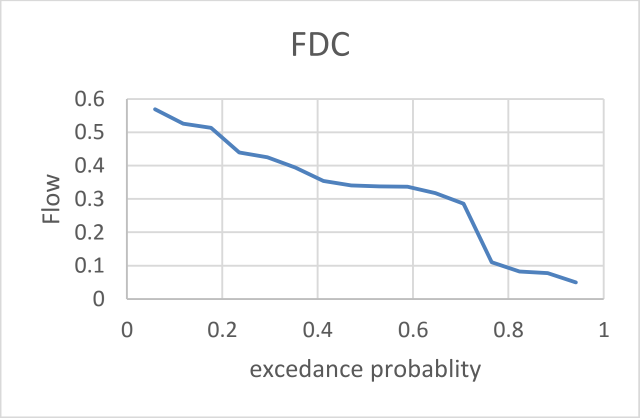
(I) Gogeb station



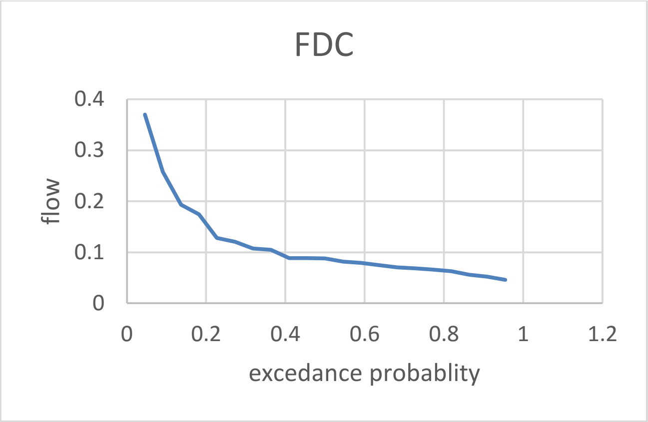
(j) Goghora station

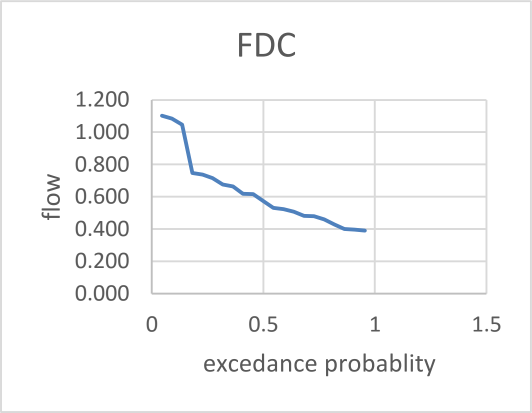


(k) Guma



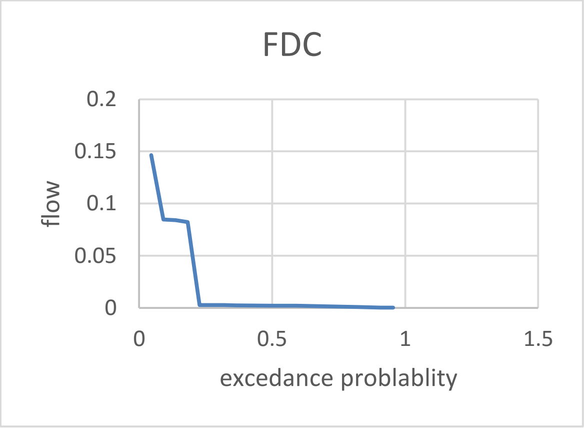
(l) Walga

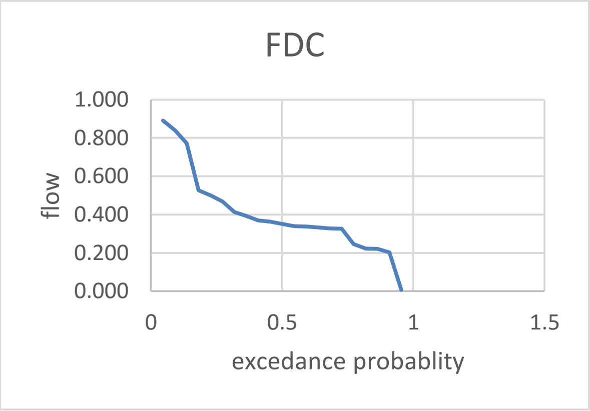




(m) Kito station

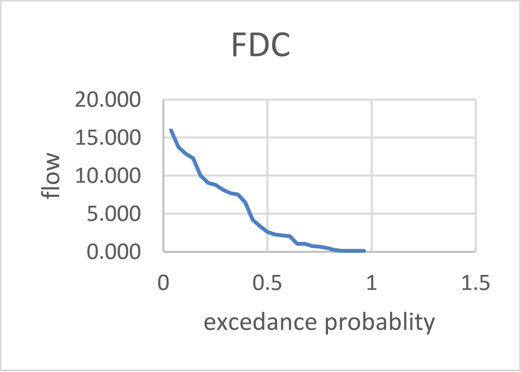
(n) Limmu genet station

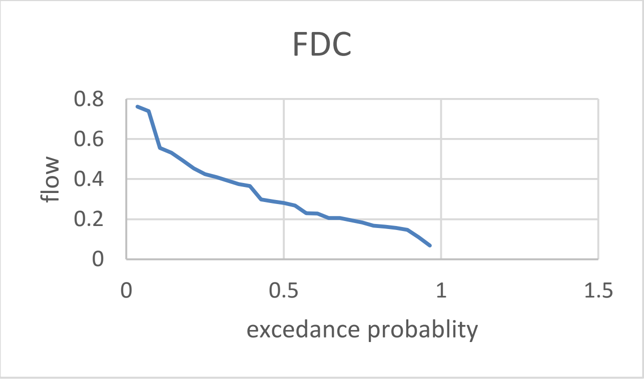




(o) Mazie station

(p) Seka station

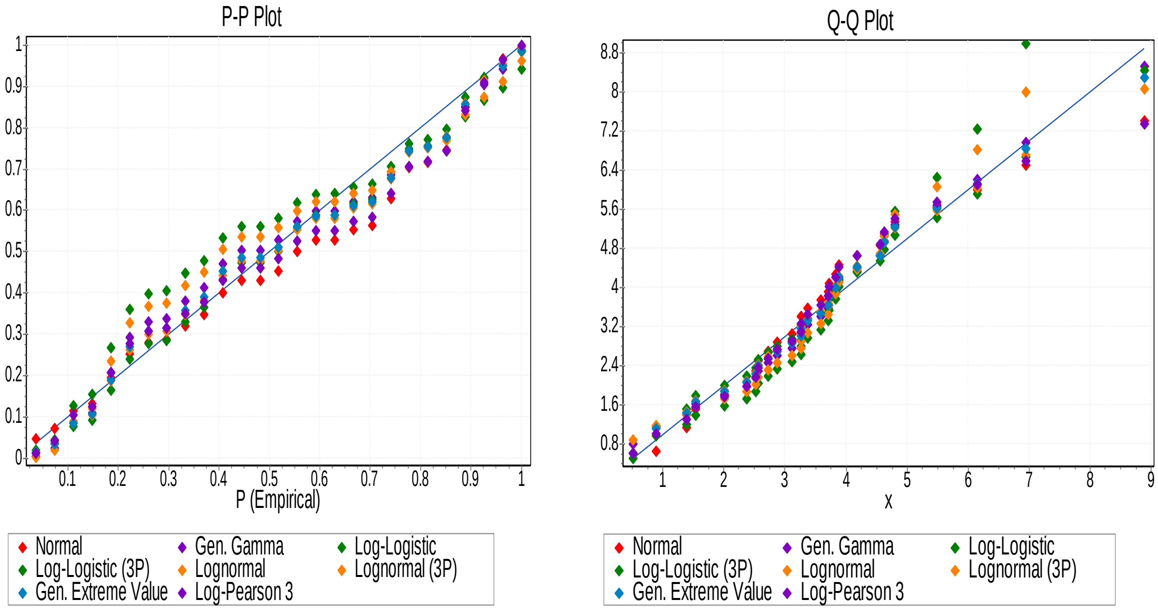




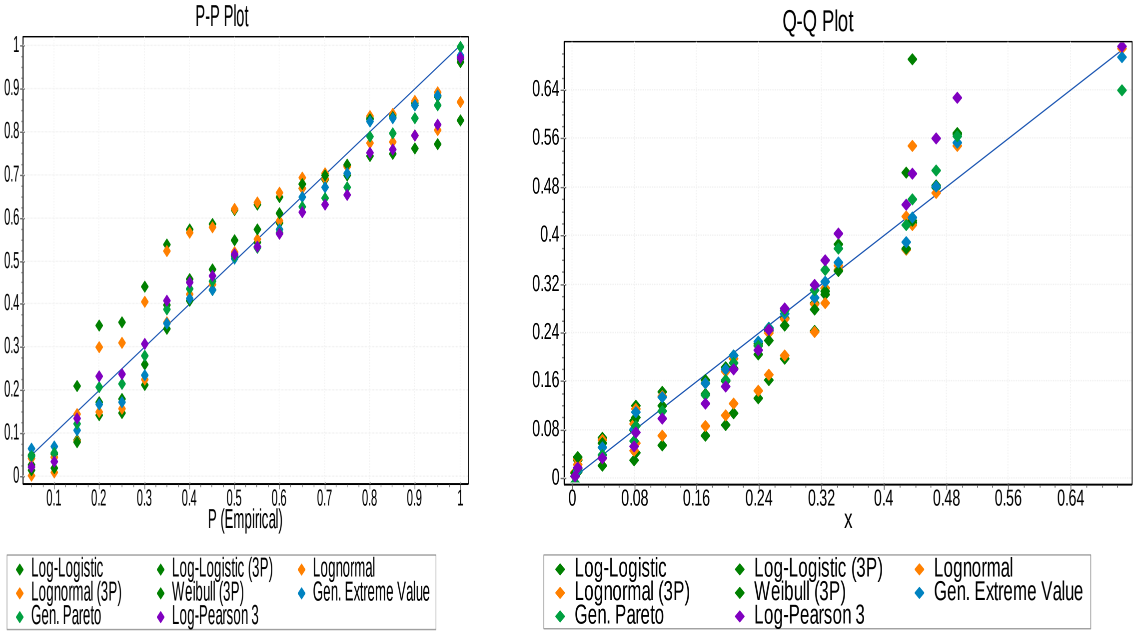
(q) Shebe station

(r) Sheta station

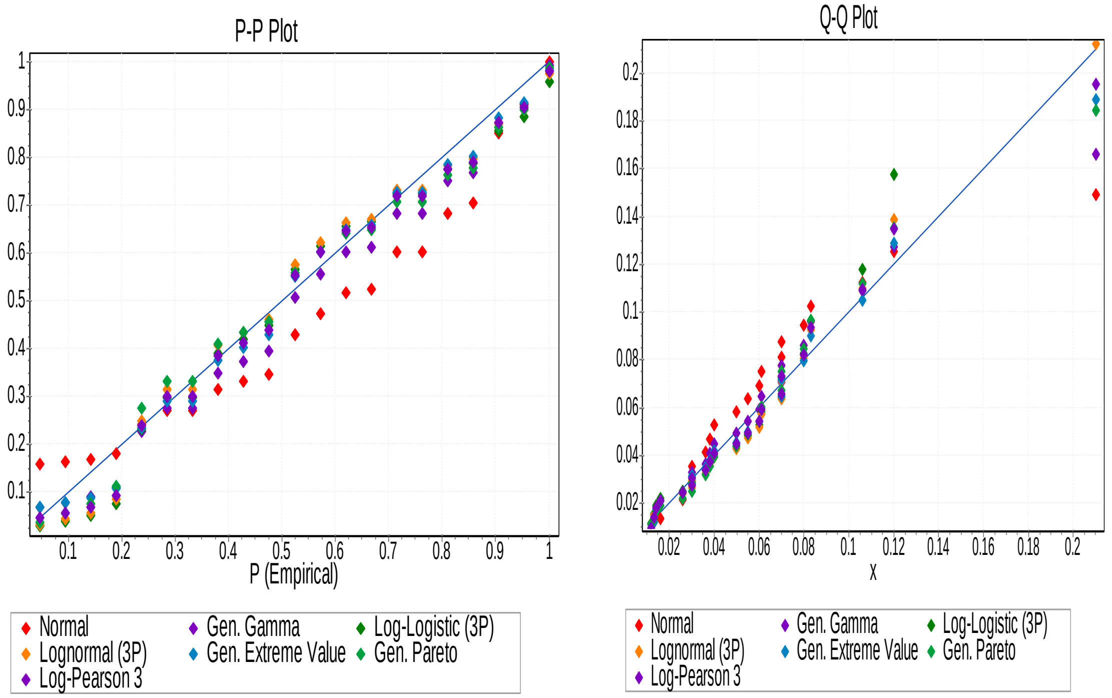
Appendix- F. This appendix indicate the P-P plot and Q-Q Plot to show goodness of fit test.



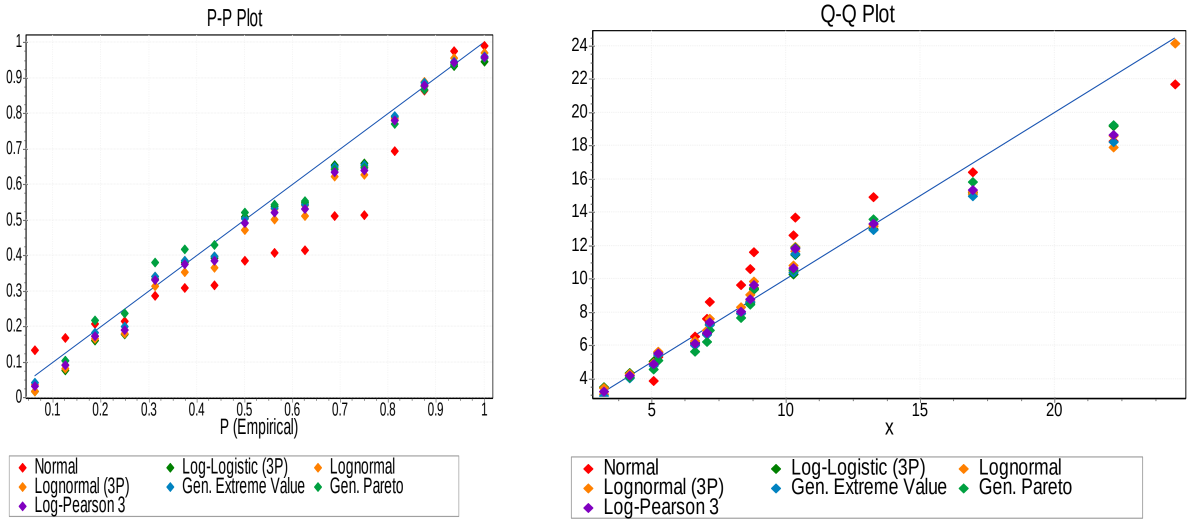
Asendabo station



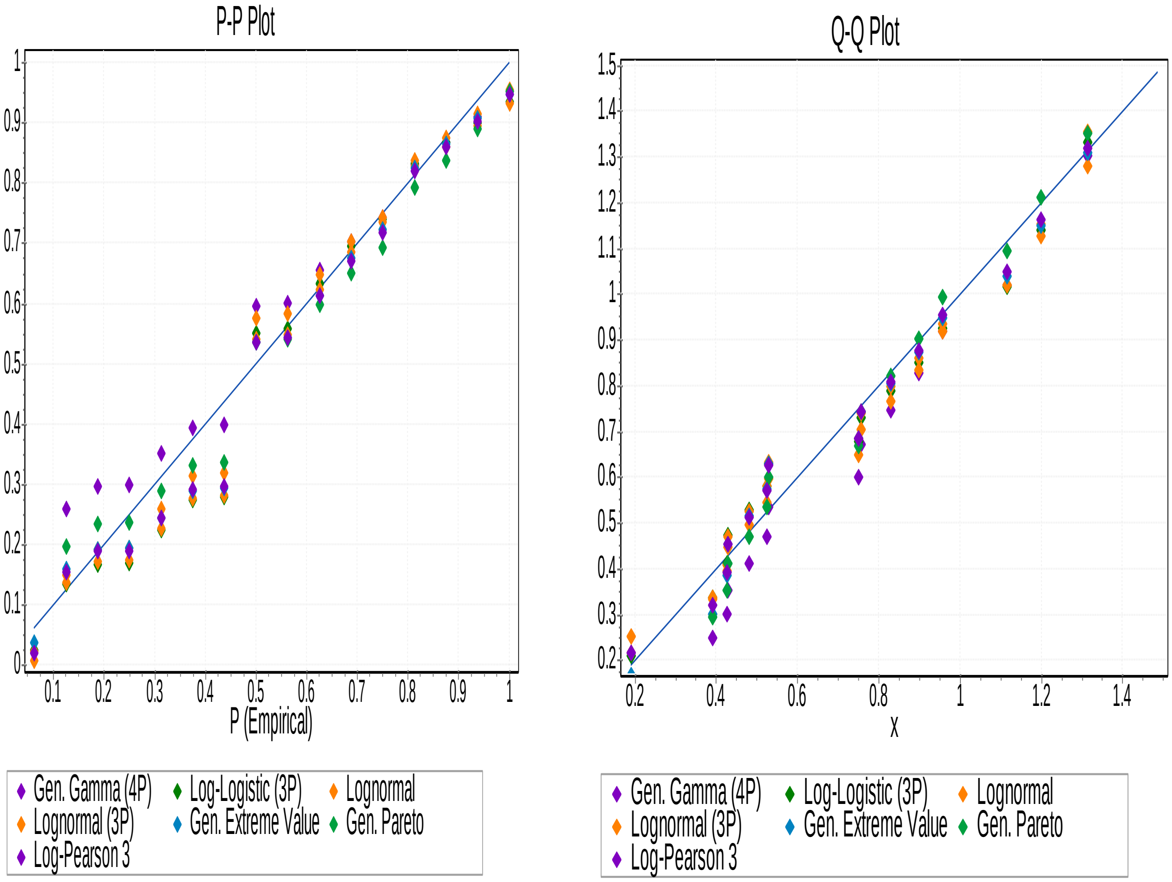
Bulbul station



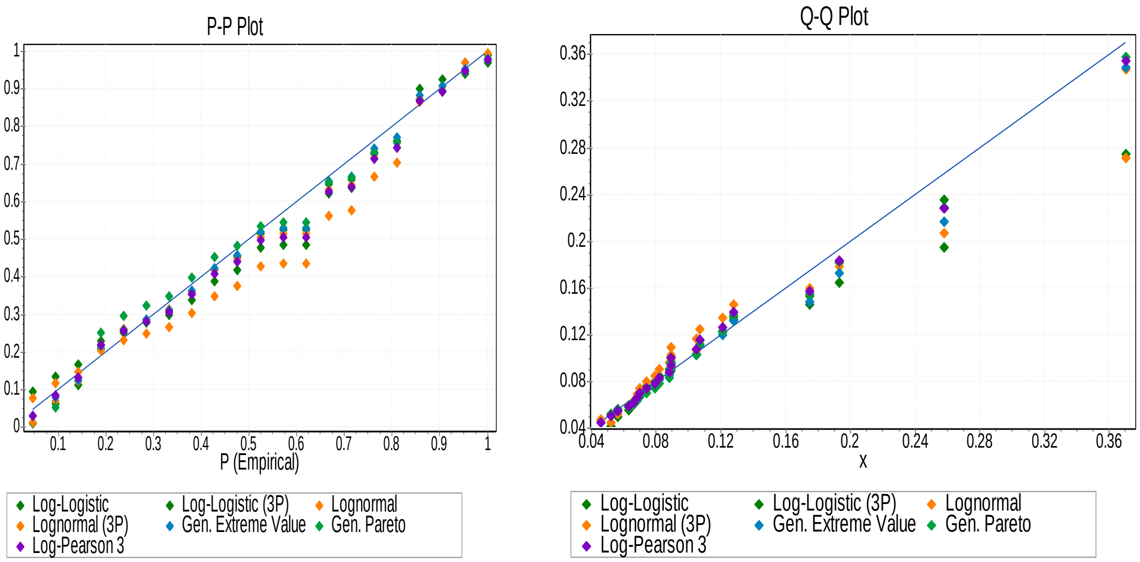
Awetu station



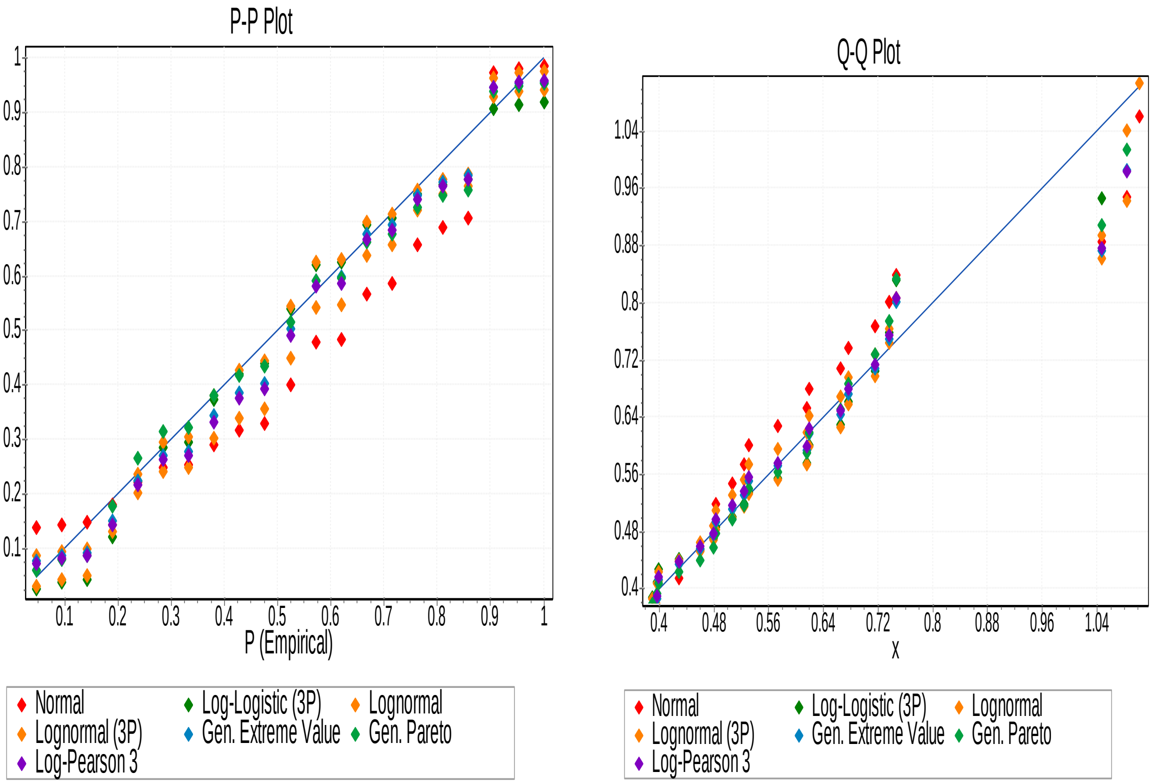
Chida station



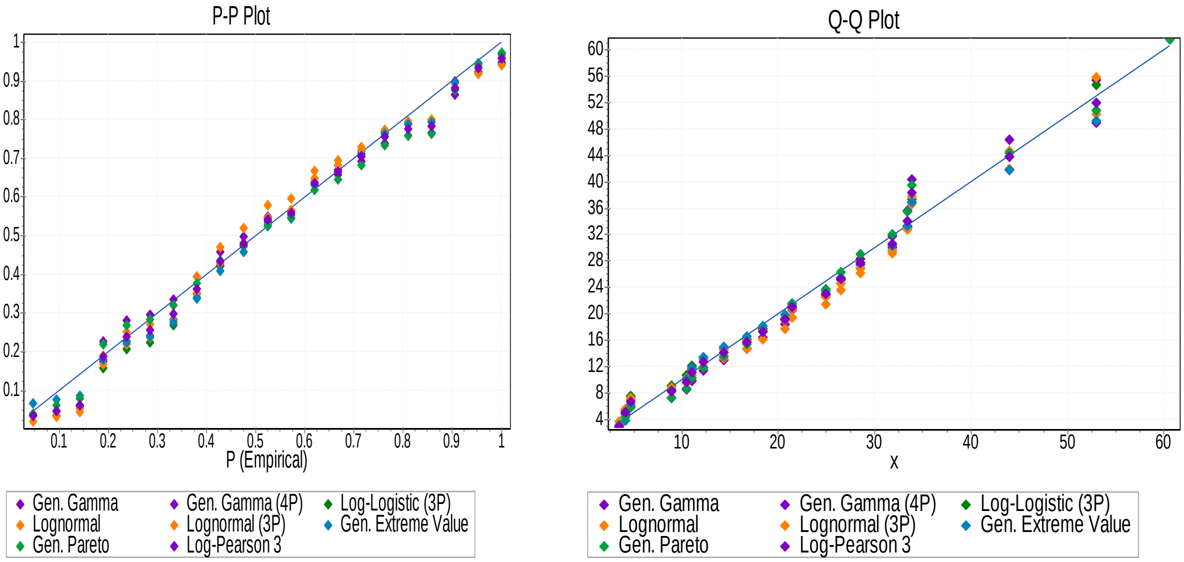
Guma station



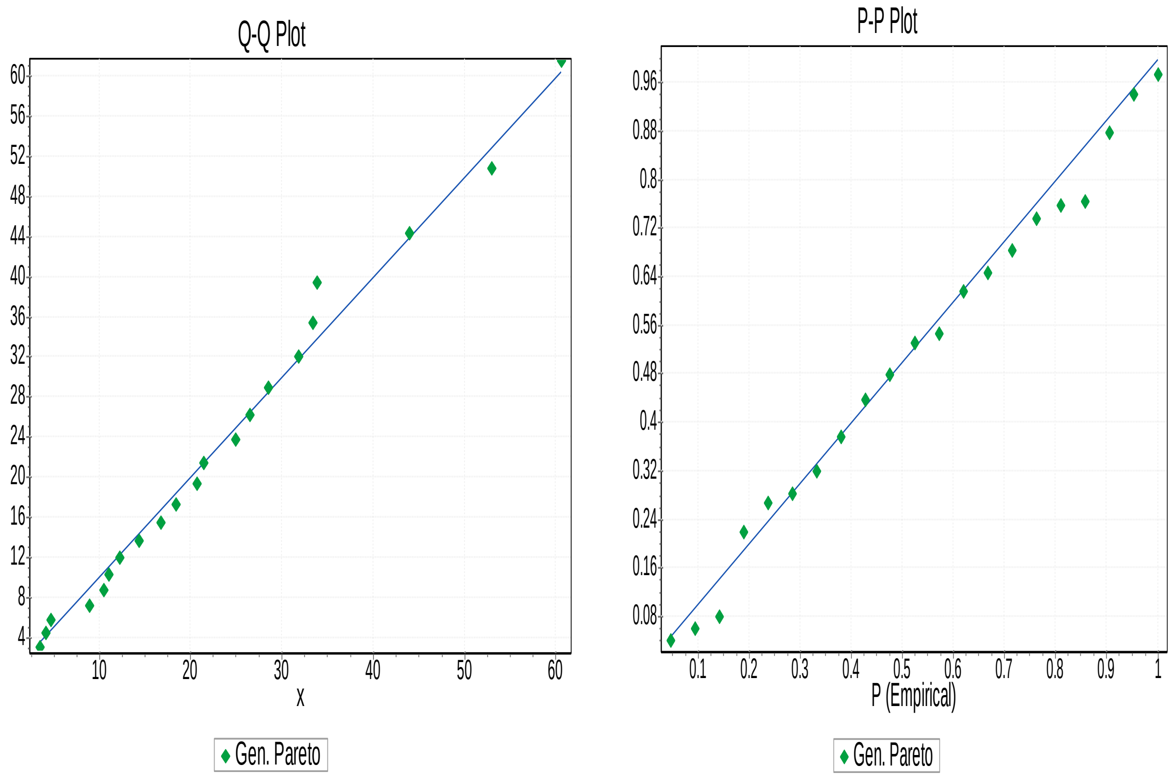
Kito station



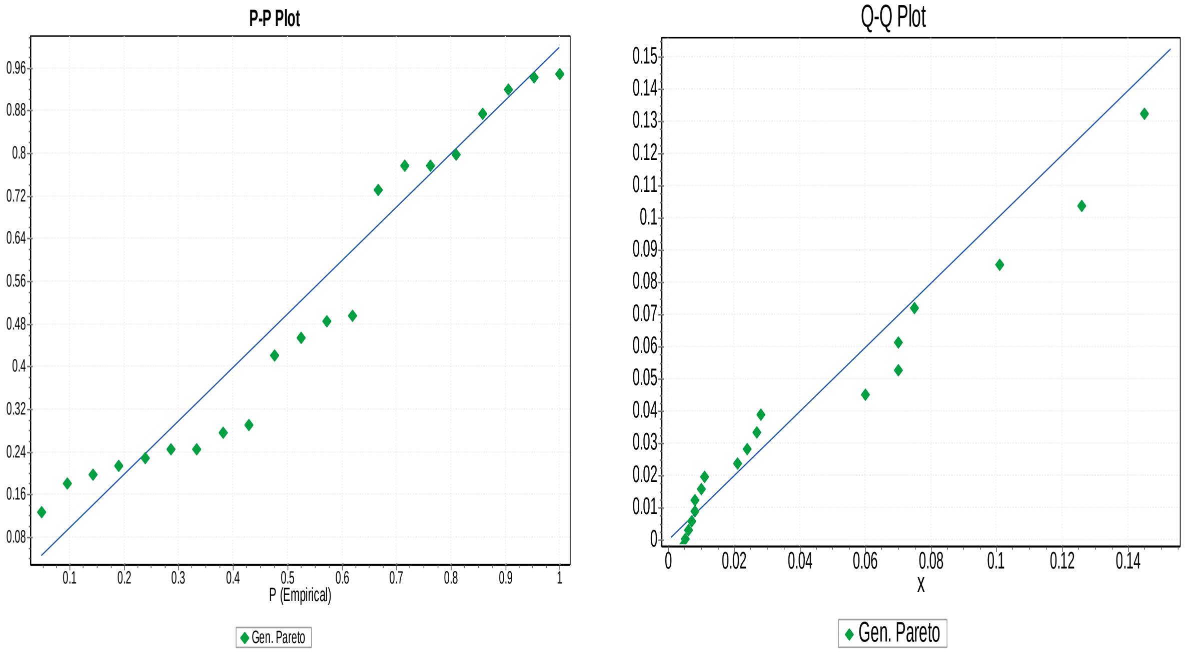
Limmu station



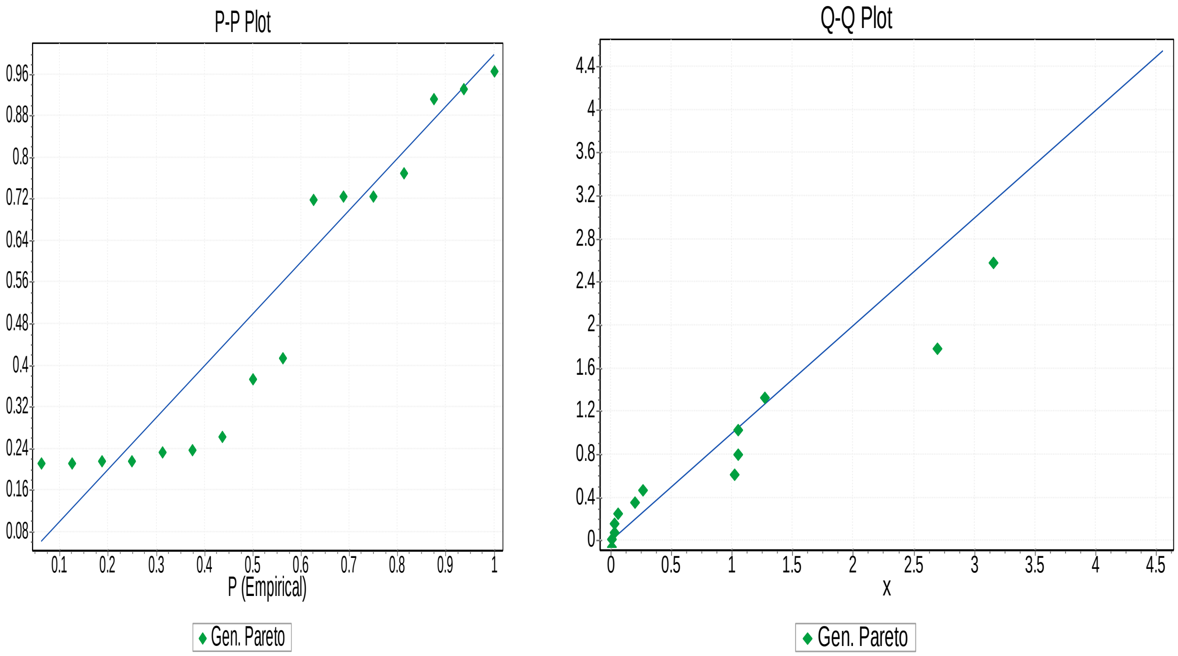
Seka station



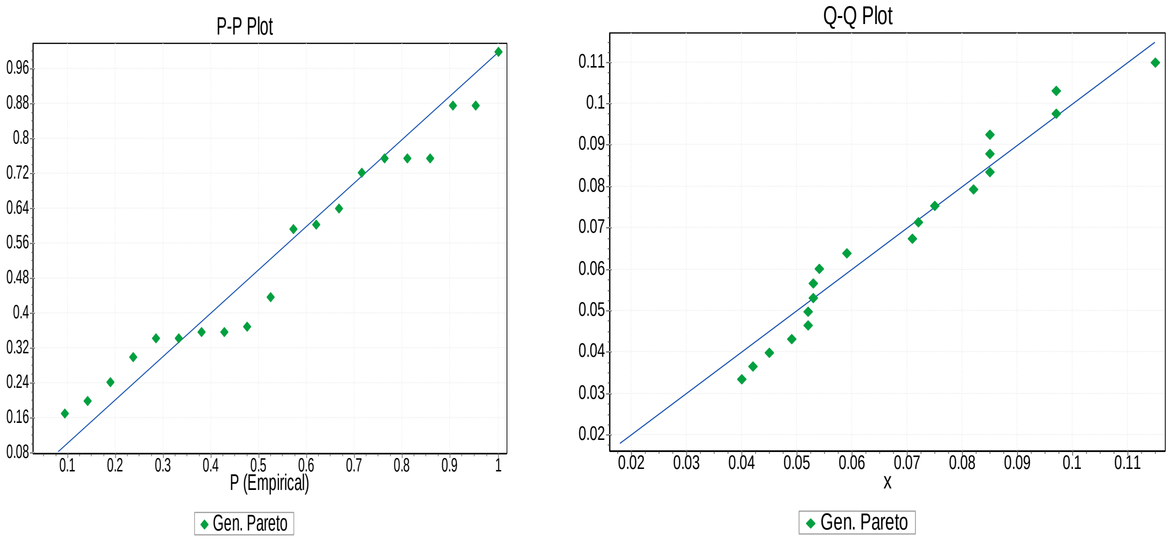
Abelti station



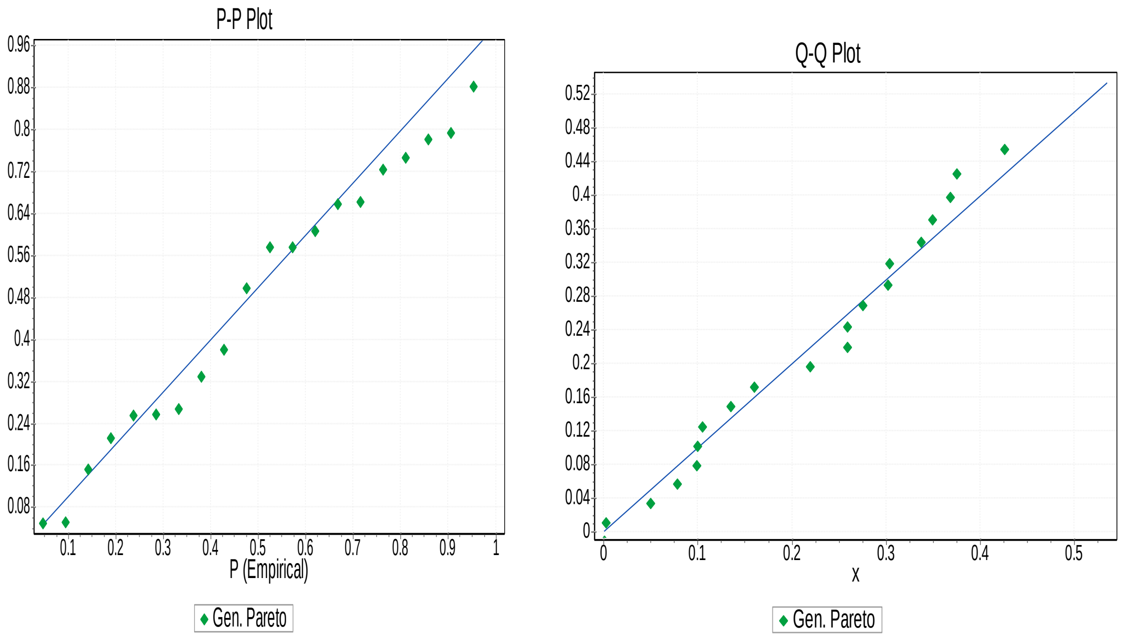
Bidru station



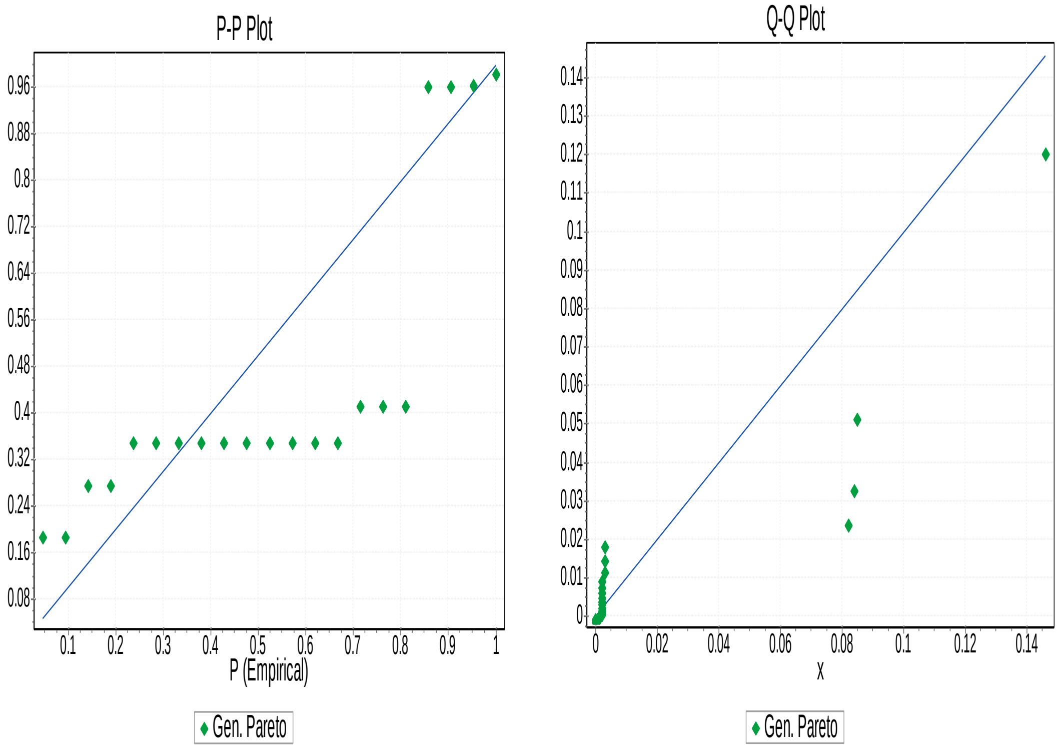
Tollay station



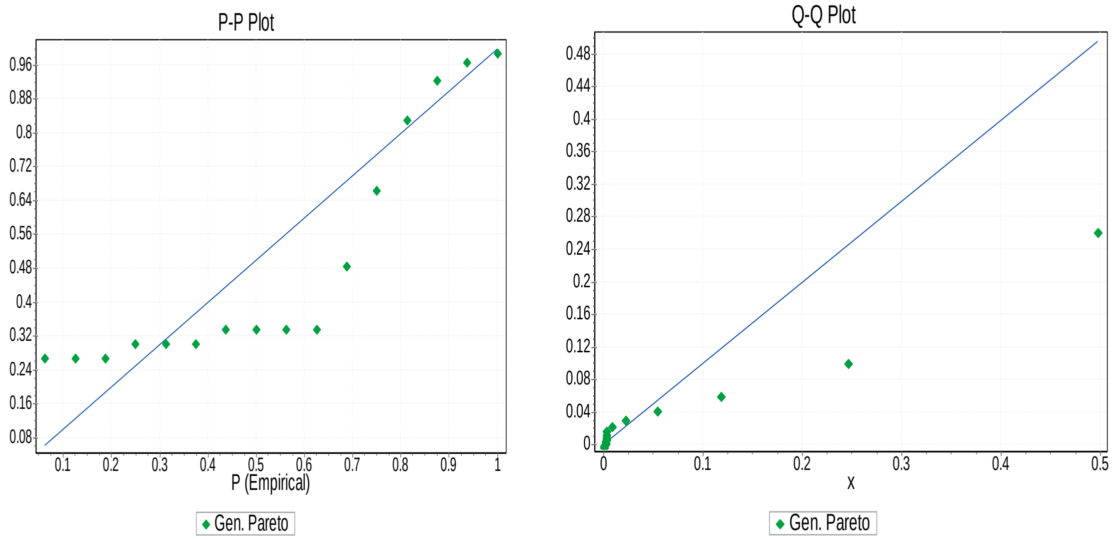
Gogeb station



Goghora station



Mazie station



Walga station

Appendix-G

To select the best fit distribution observing the most probable distribution that fit hydrological data extracted were selected by Matlab as follows (a and b).

Station name	GEV	GPA	Lognormal	GLO
Awetu	41.2056	34.0285	40.6513	41.1868
Asendabo	58.6432	56.5949	54.8655	53.7019
Seka	4.44225	2.02278	4.08872	1.22815
Limmu	6.66369	3.91231	5.50648	5.11038
Kito	35.4505	16.1359	32.9825	33.2182
Bulbul	6.67683	6.18307	3.50409	4.48042
Chida	57.6424	50.5734	47.7049	47.9919
Dincha	23.7777	22.7908	21.1024	21.054
Woshi	6.63648	6.09255	3.43811	3.44192
Gecha	14.7365	14.1093	14.6928	14.0801
Shebe	76.1122	75.0261	71.2157	72.0895
Sheta	11.5521	4.08516	11.4675	11.1794
Guma	6.88569	4.53935	6.2497	6.49128

(a)

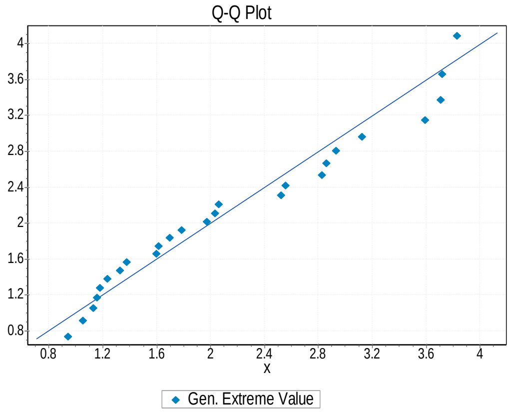
Station name	GEV	GPA	Lognormal	GLO
Abelti	85.0411	85.6854	85.0616	85.4216
Bidru	42.9408	46.2041	44.0007	43.2644
Tollay	6.94282	7.90681	6.77523	7.59494
Walga	44.47999	44.5686	44.0703	43.8228
Mazie	77.8664	78.4189	77.7175	78.2347
Goghora	11.1729	13.1419	3.35544	6.21684
Gogeb	49.7147	50.8198	47.8473	48.5566

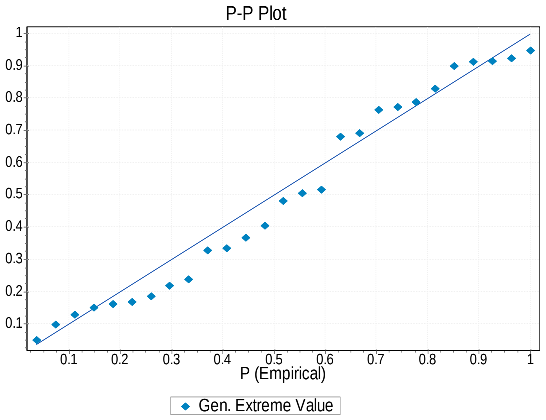
(b)

Appendix -H

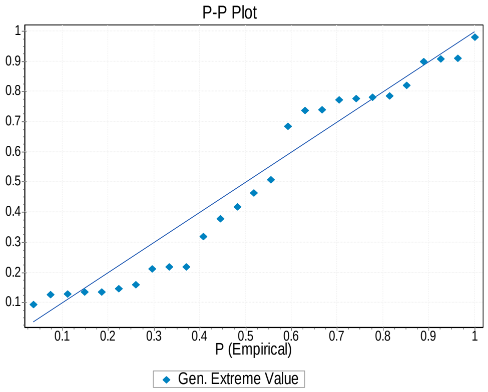
Representation of P-P and Q-Q plot by easyfit to selected distribution for delineated region

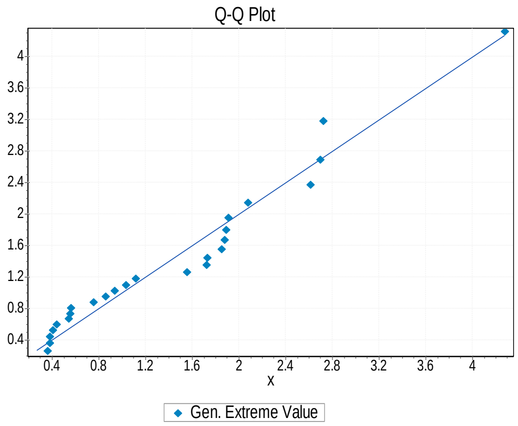
Region -1



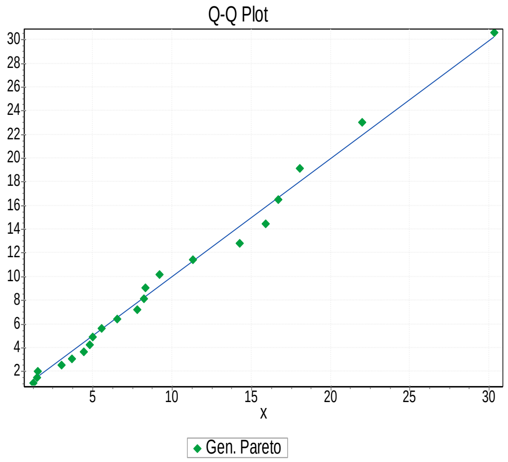


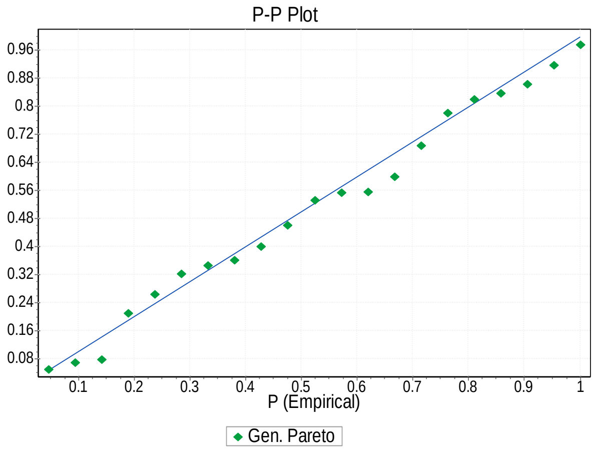
Region -2



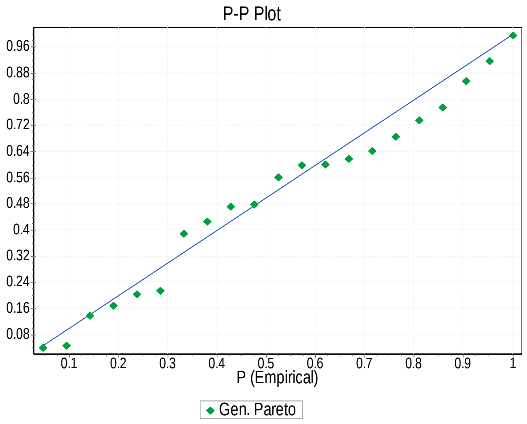


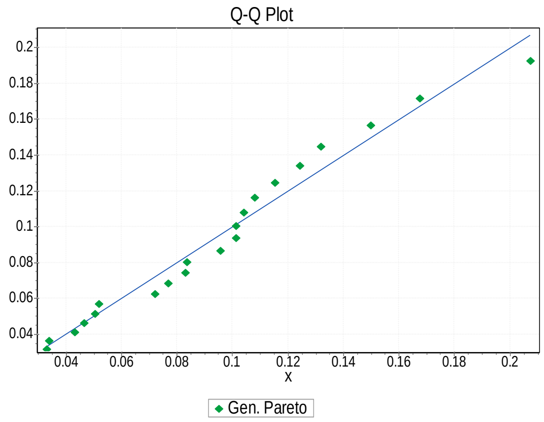
Region -3





Region -4


P-P Plot


Q-Q Plot

Appendix- I. Parameter used in quantile estimation by different tools.

Station	Xlstat			Easyfit			Matlab		
	k	δ	μ	κ	δ	μ	κ	δ	μ
Awetu	0.1	0.037	0.037	0.221	0.027	0.035	0.302	0.0245	0.035
Asendabo	0.13	1.67	2.753	-0.097	1.626	2.755	-0.120	1.656	2.784
Bulbul	0.1	0.144	0.175	-0.081	0.168	0.162	-0.053	0.154	0.163
Kito	0.1	0.041	0.076	0.436	0.029	0.074	0.502	0.029	0.075
limmu.G	0.1	0.174	0.521	0.158	0.146	0.517	0.311	0.126	0.511
Chida	0.10	4.963	7.055	0.243	3.643	6.885	0.297	3.419	6.960
gibe Seka	0.1	0.157	0.294	0.14	0.163	0.313	0.075	0.175	0.315
Dincha	0.1	0.495	0.646	0.136	0.693	1.195	0.157	0.406	0.616
Woshi	0.14	0.165	0.308	-0.547	0.194	0.278	0.617	0.188	0.287
Gecha	0.1	0.125	0.253	0.156	0.113	0.252	0.182	0.106	0.252
Shebe	0.1	3.664	2.502	0.134	3.452	2.438	0.125	3.578	2.482
Sheta	0.1	0.143	0.241	0.080	0.134	0.233	0.114	0.126	0.235
Guma	0.12	0.298	0.583	-0.063	0.33	0.595	-0.106	0.313	0.612
Abelti	0.1	12.244	15.947	0.053	12.2	15.33	0.105	11.106	15.38
Bidru	0.1	0.039	0.024	0.281	0.03	0.020	0.994	0.016	0.01
Tollay	0.1	1.095	0.376	0.456	0.517	0.243	3.1683	0.058	0.020
Walga	0.1	0.052	0.001	0.754	0.0161	-0.005	0.086	0.086	0.073
Mazie	0.1	0.033	0.001	0.672	0.007	-0.003	0.712	0.006	0.002
Goghora	0.10	0.116	0.165	-0.233	0.148	0.167	-0.235	0.138	0.171
Gogeb	0.1	0.019	0.055	-0.169	0.022	0.056	-0.252	0.022	0.057

Appendix-J. Confidence limit for estimated low flow magnitude of different return periods.

Station				Station			
Abelti	UCL	LCI	Flow(m^3/s)	Bidru	UCL	LCI	Flow(m^3/s)
5	45.81272	30.92123	38.36698		0.026475	0.02471	0.025593
10	52.08257	37.19108	44.63683		0.028063	0.026298	0.027181
15	55.4077	40.51621	47.96195		0.028594	0.02683	0.027712
25	59.32894	44.43745	51.88319		0.02902	0.027256	0.028138
50	64.25673	49.36524	56.81099		0.029341	0.027576	0.028459
75	66.95788	52.06639	59.51213		0.029448	0.027684	0.028566
100	68.80067	53.90919	61.35493		0.029502	0.027737	0.02862
150	71.3004	56.40892	63.85466		0.029556	0.027791	0.028673
Station				Station			
Tollay	UCL	LCI	Flow(m^3/s)	Gogeb	UCL	LCI	Flow(m^3/s)
5	1.62581	1.020753	1.323281		0.182748	0.073383	0.128065
10	1.902105	1.297047	1.599576		0.215879	0.106514	0.161196
15	2.043257	1.438199	1.740728		0.237534	0.128169	0.182852
25	2.20481	1.599752	1.902281		0.254135	0.14477	0.199453
50	2.400131	1.795073	2.097602		0.267797	0.158432	0.213114
75	2.503471	1.898413	2.200942		0.31519	0.205825	0.260507
100	2.572429	1.967371	2.2699		0.346862	0.237496	0.292179
150	2.663939	2.058881	2.36141		0.371337	0.261972	0.316655
Station				Station			
Goghora	UCL	LCI	Flow(m^3/s)	Walga	UCL	LCI	Flow(m^3/s)
5	0.983515	0.215778	0.599647		0.00357	0.003561	0.003566
10	1.179161	0.411425	0.795293		0.003581	0.003573	0.003577
15	1.305769	0.538033	0.921901		0.003583	0.003574	0.003579
25	1.481248	0.713511	1.09738		0.003583	0.003575	0.003579
50	1.753138	0.985401	1.369269		0.003584	0.003576	0.00358
75	1.933117	1.16538	1.549248		0.003584	0.003576	0.00358
100	2.071372	1.303635	1.687504		0.003584	0.003576	0.00358
150	2.282544	1.514808	1.898676		0.003584	0.003576	0.00358

Station				Station			
Mazie	UCL	LCI	Flow(m^3/s)	Asendabo	UCL	LCI	Flow(m^3/s)
5	0.092777	0.047569	0.070173		10.23296	4.262361	7.2476622
10	0.111683	0.066475	0.089079		11.95507	5.984467	8.969767965
15	0.12174	0.076532	0.099136		12.99742	7.026823	10.01212374
25	0.133629	0.088421	0.111025		14.36524	8.394636	11.37993684
50	0.148616	0.103408	0.126012		16.33929	10.36869	13.35398998
75	0.156856	0.111647	0.134251		17.56574	11.59514	14.58044262
100	0.162487	0.117278	0.139882		18.47092	12.50032	15.48562346
150	0.170139	0.12493	0.147534		19.79903	13.82843	16.81373025
Station				Station			
Awetu	UCL	LCI	Flow(m^3/s)	Kito	UCL	LCI	Flow(m^3/s)
5	0.083931	0.068731	0.076331		0.118479	0.110053	0.114266
10	0.092063	0.076864	0.084463		0.124026	0.115599	0.119813
15	0.095901	0.080702	0.088302		0.126347	0.117921	0.122134
25	0.100027	0.084827	0.092427		0.128617	0.120191	0.124404
50	0.104625	0.089426	0.097026		0.130858	0.122432	0.126645
75	0.106883	0.091683	0.099283		0.13184	0.123413	0.127626
100	0.10832	0.093121	0.100721		0.132422	0.123996	0.128209
150	0.110142	0.094943	0.102542		0.133111	0.124685	0.128898
Station				Station			
Bulbul	UCL	LCI	Flow(m^3/s)	Seka	UCL	LCI	Flow(m^3/s)
5	0.743829	0.349009	0.546419		1.006564	0.514374	0.760469
10	0.877334	0.482514	0.679924		1.16747	0.675279	0.921375
15	0.955016	0.560196	0.757606		1.262356	0.770165	1.016261
25	1.053735	0.658915	0.856325		1.384255	0.892064	1.13816
50	1.190403	0.795584	0.992994		1.555407	1.063217	1.309312
75	1.272196	0.877376	1.074786		1.659145	1.166955	1.41305
100	1.331159	0.936339	1.133749		1.734526	1.242336	1.488431
150	1.415656	1.020836	1.218246		1.843417	1.351227	1.597322

Station				Station			
Dincha	UCL	LCI	Flow(m^3/s)	Gecha	UCL	LCI	Flow(m^3/s)
5	1.630994	1.188864	1.409929		0.495264	0.400339	0.447801
10	1.830114	1.387983	1.609048		0.540717	0.445792	0.493254
15	1.932503	1.490373	1.711438		0.563415	0.46849	0.515952
25	2.050302	1.608172	1.829237		0.588925	0.494001	0.541463
50	2.193684	1.751554	1.972619		0.619054	0.524129	0.571591
75	2.270014	1.827883	2.048948		0.634654	0.539729	0.587191
100	2.321144	1.879014	2.100079		0.644924	0.55	0.597462
150	2.389256	1.947125	2.168191		0.658373	0.563448	0.610911
Station				Station			
Shebe	UCL	LCI	Flow(m^3/s)	Sheta	UCL	LCI	Flow(m^3/s)
5	11.91723	7.496539	9.706885		0.572848	0.410269	0.491559
10	13.82736	9.406669	11.61701		0.642082	0.479503	0.560793
15	14.82856	10.40787	12.61821		0.678611	0.516033	0.597322
25	15.9982	11.57751	13.78786		0.721513	0.558935	0.640224
50	17.45033	13.02964	15.23999		0.775143	0.612564	0.693853
75	18.23752	13.81683	16.02717		0.804397	0.641818	0.723108
100	18.77084	14.35015	16.56049		0.824294	0.661716	0.743005
150	19.4893	15.06861	17.27896		0.851204	0.688625	0.769915