Jimma University

School of Graduate studies

Jimma Institute of Technology

Faculty of Civil and Environmental Engineering

Chair of Hydrology and Hydraulic Engineering

Masters of Science Program in Hydraulic Engineering

Assessment of Predictive Accuracy for Selected Regional Flood Frequency Distribution Estimation Methods. A case of Upper Omo-Gibe River Basin, Ethiopia

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

BY: Tesfaye Yismaw Deneke

January 2022

Jimma, Ethiopia

Jimma University

School of Graduate studies

Jimma Institute of Technology

Faculty of Civil and Environmental Engineering

Chair of Hydrology and Hydraulic Engineering

Masters of Science Program in Hydraulic Engineering

Assessment of Predictive Accuracy for Selected Regional Flood Frequency Distribution Estimation Methods. A case of upper Omo-Gibe River Basin, Ethiopia

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

BY: Tesfaye Yismaw Deneke

Main advisor Dr.-ing Tamene Adugna (Ph.D.)

Co-advisor: Mr. Kiyya Tesfa (MSc.)

January 2022

Jimma, Ethiopia

DECLARATION

I hereby declare that the research entitled "**Assessment of predictive accuracy for regional flood frequency distribution estimation methods on Upper Omo-Gibe River Basin, Ethiopia**" is my work which I submit for partial fulfillment of the degree of Master of Science in Hydraulic Engineering to Jimma University, school of graduate studies, Jimma Institute of Technology, and Hydrology and Hydraulic Engineering Chair. The research was conducted under the guidance of the main advisor Dr.-Ing Tamene Adugna (PhD.) and co-advisor Mr. Kiyya Tesfa (MSc.).

Mr. Tesfaye Yismaw ______________ ______________

Candidate Signature Date

This research has been submitted for presentation with my approval as a university supervisor.

Dr.-Ing Tamene Adugna (PhD.) ______________ ______________

Main-advisor Signature Date

Mr. Kiyya Tesfa (MSc.) ______________ ______________

Co-advisor Signature Date

APPROVAL

The thesis entitled "**assessment of predictive accuracy for regional flood frequency distribution estimation methods on Upper Omo-Gibe River Basin, Ethiopia**" submitted by Tesfaye Yismaw Deneke is approved and accepted as a Partial Fulfillment of the Requirements for the Degree of Masters of Science in Hydraulic Engineering at Jimma Institute of Technology.

	Name	Signature	Date
Main Advisor	Dr.-Ing Tamene Adugna (PhD.)	___________	___________
Co-Advisor:	Kiyya Tesfa (MSc.)	___________	___________

As members of the examining board of the MSc. thesis, we certify that we have read and evaluated the thesis prepared by Tesfaye Yismaw Deneke. 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:		_____________	_____________
External Examiner:		_____________	_____________
Internal Examiner:		_____________	_____________

ABSTRACT

Floods are one of the most common and destructive natural disasters, causing substantial loss of life and property all over the world. Assessment of predictive accuracy for regional flood frequency distribution estimation method has been the backbone of water resources project planning, design of any structures, and the economic analysis of flood control projects. The goal of this study was to test the predictive fit of probability distributions to yearly maximum flood data and to determine which distribution and estimation method provide the best match for the Upper Omo-Gibe River Basin. Using a river basin as a case study, the performance of nine probability distributions, three fitting tests, evaluation processes, and selection procedures was examined. To achieve this, data from eleven stream gauged sites, three hydrological homogeneous sub-regions were defined and delineated based on L-moment homogeneity tests, namely Region-A, Region-B, and Region-C. Delineation of homogeneous regions was accomplished using ArcGIS10.4.1. Discordancy of regional data of the L-moment statistics was identified using Matlab2018a. The R programming language was used in conjunction with a collection of the most recent computer statistical programs in an integrated development environment. The most relevant distribution models were identified using maximum likelihood estimation, goodness-of-fit tests-based analysis, and information criteria-based selection techniques. The performances of the distributions were evaluated using Kolmogorov Smirnov, Anderson-Darling, and Chi-Squared goodness-of-tests. After three goodness of fit tests were carried out, the results showed that the lognormal and gamma distribution models were the best-fit functions for Region-A. Because they had the lowest Akaike Information Criterion (AIC) values of -46.251 *and* -45.802, *and Bayesian Information Criterion (BIC) values corresponding to* -43.320 *and* -42.870, *respectively. Similarly, for both Region-B and Region-C the lognormal and gamma functions were the best-fit distribution functions and identified as suitable distributions for analyzing accurate annual maximum flows in the basin. Based on best-fit distributions for the three regions, regional flood frequency curves were constructed. In this study, the flood magnitude is estimated for 2, 5, 10, 15, 20, 25, 50, 75, 100, 200, 500, and 1000 years return period, and their respective extreme event for Region-A became 427.65, 547.24, 626.42, 671.1, 702.37, 726.46, 800.6, 800.68, 843.82, 874.35, 917.31, and 947.75. 1044.59. The derived flood frequency curves at a given return period suggested that how important engineering decisions and actions such as design and operation of the water resources project have to be undertaken carefully.*

Keywords: *Best-fit distribution, Flood frequency analysis, Homogeneity, Parameter estimati on, and Regionalization.*

ACKNOWLEDGMENT

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

I would like to express my deep to my advisors Dr.-Ing Tamene Adugna (PhD.) and Kiyya Tesfa (MSc.) for their persistence, support, constructive guidance, and advice throughout the work.

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

I would like to acknowledge the Mizan Tepi University, MTU, for its kind help, support, and sponsorship of my study program for the M.Sc.

I am very grateful to Dessyie Asaye and yeshigeta Nurieye for their support, valuable advice, and follow-up with their research experience. I also thank Jimma Institute of Technology, Faculty of Civil and Environmental Engineering, and the Jimma University community for their support.

Last but not least, I would like to thank all my families especially my Mather werkinesh Anagaw and friends who helped me a lot in carrying out my thesis through remarkable encouragement, advice, endless support, and comments in several ways for the realization of this work

TABLE OF CONTENTS

Contents

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATION

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

1. INTRODUCTION

1.1 Background

Floods are one of the most devastating natural disasters that have serious ramifications for human society and result in enormous loss of life and property all over the world. It ruins properties, agricultural lands, causes economic losses, reduces drainage efficiency, and disrupts life (Yucel and Keskin, 2011). It is a natural occurrence, but human encroachment on natural streams, as well as the effects of land use and climate change, have the potential to alter catchment runoff responses, resulting in the occurrence of catastrophic floods, increasing susceptibility and risk (Hailegeorgis & Alfredsen, 2017).

The calculation of flood quantiles corresponding to return periods (T) of interest is commonly done using statistical methods for flood frequency analysis based on systematic streamflow data. Prevalence of severe floods or increasing trends in one or more flood characteristics (e.g., flood frequency, magnitude, and timing) in different parts of the world substantiate the need for more reliable prediction of flood quantiles for design and management of water and transportation infrastructure such as spillways, culverts, bridges, sewers, etc. (Teklu & Knut, 2017).

Estimating flood frequency is critical for flood management. It's used to map floodplain areas, develop hydraulic structures (dams, retaining basins, stormwater systems), and infrastructure (roads, bridges) for natural disaster assessments and alert mechanisms, as well as quantify the frequency of flood episodes (Javelle *et al*., 2010). The application of flood frequency analysis (FFA) in the assessment of hydrological hazards has led to the development of a variety of methods ranging from purely statistical approaches to simulation approaches. The availability of observation data and the objectives to be attained often influence the development of these methods (Castellarin, *et al*., 2011; Pathiraja *et al*., 2012).

The most important statistical tool for determining the nature and size of a river's discharge is flood frequency analysis. It aims to relate the magnitude of events to their frequency of occurrence through probability distribution (Bhagat, 2017; Ganamala & Kumar, 2017). If adequate records are available, the common methods give acceptably uniform results within the range of data. However, the location of the gauging station occasionally coincides with the sites of interest, or the available records become too short to make important statistical implications (Badreldin & Fengo, 2012). Hence, the estimation of design floods for a site has been a common problem particularly for ungagged basins or for sites of a short record length (Hailegeorgis & Alfredsen, 2017).

Flood-prone damages are frequent in several locations of Ethiopia, owing to a lack of properly researched knowledge and prevention mechanisms. The frequency and magnitude of floods have increased, affecting large parts of the country and causing damage to property, loss of life, and the health of the population (Akirso, 2017). Using flood data from neighboring sites within a homogenous region, regional flood frequency analysis has been an effective technique for calculating flood quantiles at ungauged sites or with insufficient streamflow data (Dubey, 2014; Lu, 2016; Wu, *et al.*, 2018). It is a data-driven approach, which attempts to transfer flood information from a group of gauged catchments to the catchment location of interest. This technique is expected to be simple so that design flood estimates can be obtained from readily available input data and the region is considered homogeneous (Rahmana, *et al.,* 2015).

The established flood versus return period curve is utilized in regional flood frequency analysis to estimate flood quantiles for every site within the region. These regional relations can alleviate the effects of outliers from time-series data (Mishra, *et al*., 2009). It is recommended to use observed time-series data to develop flood frequency estimates for a site with limited time-series data. Since they are the bases for regional information (Wilson, *et al*., 2011). Therefore, the use of regional information derived from data at gauged sites and regionalized for use at any location within the basin has practiced major setbacks due to the absence of tools and methods.

However, most catchments in impoverished nations such as Ethiopia are inadequately gauged or ungauged, making water supply management and flood prediction challenging (Rabba, *et al*., 2018). The low density of gauging stations, the complexity of operating and maintaining gauging networks, and the lack of infrastructure required to collect sufficient hydrologic data are all contributing factors (Gedefa & Seleshi, 2009). The availability of such technologies would help flood risk estimation, water management, and engineering decisions in the basin (Share Bale Eco-Region, 2017). As a result, the primary goal of this research is to use annual maximum series (AMF) estimation modeling of stream gauging data to undertake regional flood frequency analysis on the Upper Omo-Gibe River basin in Ethiopia.

1.2 Statement of the problem

Floods will continue to wreak havoc on the economy and the environment. Flood disasters are account for roughly one-third of all-natural disasters in terms of both number and economic losses. Flood and droughts are the world's costliest natural disasters, causing an average of $6–$8 billion in global damages annually and collectively affecting more people than any other form of natural disaster (Lampros, 2009). For example, in 2006, a flood in the Omo basin overwhelmed the Dasenech and Nyangatom Weredas, killing 364 people and displacing 6000 to 10,000 people in Kuraz District, South Omo. More than 3000 livestock are also reported to have been claimed due to the flooding (OCHA UN office of the coordination of humanitarian affairs, 2006).

The Omo-Gibe river basin is one of the flood-affected areas in Ethiopia flooding (OCHA UN office of the coordination of humanitarian affairs, 2006). It regularly floods settlements in the basin's lower reaches, reducing community benefits from flood and recession agriculture. The necessity to conduct this research is the frequent happening of floods in the area. Flood management bodies require information about flooding characteristics and their effects to make decisions about flood management strategies such as the construction of flood protection structures (engineering structures), the development of flood emergency plans, and human settlement planning. As the region is known as a crop productive area of the country, most of the crops may be suffered. Aside from that, the Upper Omo-Gibe river basin is one of the country's water resource potential zones, suited for a variety of water resource projects such as hydropower projects, water supply, recreation, and small and large-scale irrigation projects.

Deferent researcher conducts a number of research on the study area by different tittle like, Surface water potential Assessment and demand scenarios analysis In Omo-Gibe River Basin by Dereje Atinafu(2016), The effects of land use land cover change on hydrological process of Gilgel Gibe, Omo Gibe river basin conducted by (Wakjira Takalaa, Tamene Adugna(PhD) DawudTamam), Hydro Meteorological Trends in the Upper Omo-Ghibe River Basin, Ethiopia by Dessalegn Jaweso, Application of a Satellite Based Rainfall-Runoff Estimation: in Upper Omo-Gibe Basin to simulate the extreme flood event at Omorate conducted by Samuel Bekele feb (2020), but there is research gap in flood frequency analysis in the Upper Omo- Gibe River basin which is very important for proper planning and design of water resources management options and flood risk management in the study area.

The design and successful operation of hydraulic and drainage facilities such as dams, spillways, bridges, culverts channels, and flood protection schemes rely heavily on this data, both in quantity and quality (Saf, 2009). (Tanaka, *et al*., 2017). Unfortunately, these powerful inputs are frequently insufficient and, in the majority of situations, completely unavailable at points of interest (Rabba, *et al*., 2018). This is also true in this study area; there is insufficient data to do flood frequency analysis that is why regional flood frequency analysis was recommended. The project was providing regional growth curves to depict regional frequency curves, which is critical for estimating flood quantile magnitude Q (T) for the river system.

1.3 Objective of the study

1.3.1 General objective

The general objective of this study is to assess the predictive accuracy of selected regional flood frequency distribution estimations on the Upper Omo River basin of Ethiopia.

1.3.2 Specific objectives

The specific objectives of the study are:

i. To identify the best-fit parameter estimation approach and statistical distributions to the data of each gauge.
ii. To develop a suitable parameter estimation method for each station in the study area.
iii. To develop regional flood frequency curves for delineated homogeneous regions corresponding to the required return periods on the basin.

1.4 Research Questions

i. What is the best-fit probability distributions for the prediction of hydrological events of gauging stations of the basin?
ii. What is the suitable parameter estimation method for each station in the study area? and
iii. How regionalization method is used for regional flood frequency analysis in the future?

1.5 Significance of the study

This research is expected to provide useful information for flood risk estimates, economic evaluation of flood control projects, effective planning, and design of water resource management alternatives in the study area. The study will also serve as a point of reference for policymakers and decision-makers, as well as any future research on the Upper Omo-Gibe River watershed. This research addressed several issues, including the effects of underestimating flood quantiles, which can result in future flood risks and hazards, as well as the effects of overestimating design flood or flood quantiles during the design of various hydraulic structures, which will aid in analyzing and estimating the exact or appropriate budget determination for various water resource development.

This research can be used as a foundation for future research in the Upper Omo-Gibe river basin. For example, effective design of various water resources development projects such as large and small-scale irrigation projects, water supply, recreation areas, and hydropower development projects.

1.6 Scope of the study

In general, the study addresses issues related to the likelihood of flooding and the magnitude of flooding that may occur depending on the hydrological response of the chosen basin. It is important to obtain an estimate of flood quantile magnitude Q (T) for locations on the river system for proper planning and design of water resources management options on the study area. The research is primarily focused on regionalizing stream flow data from the Upper Omo-Gibe river basin.

1.7. Limitation of the Study

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

2. LITERATURE REVIEW

2.1 Flood Frequency Analysis

When there is rarely observed flow data, extreme flow quantiles computed from stream flow data give crucial information for the design of any project and economic appraisal of a variety of engineering and water resources planning and development projects. Flood frequency analysis is used to estimate extreme flow quantiles from observed flow data. Flood frequency analysis is a hydrologic field concerned with estimating the magnitude of a flood matching to any specified return time of occurrence (Rao & Srinivas, 2008; Bhagat, 2017 Kanti, *et al*., 2017).

Professionals are interested in determining the appropriate estimation of extreme occurrences with defined return periods when planning and designing water resources projects (Rahmana, *et al*., 2015). These extreme events are required for the design of various flow control structures such as levees, culverts, bridges, barrages, and dams, reservoir management, economic evaluation of flood protection projects, land use planning and management, and flood risk assessment (Rao & Srinivas, 2008; Noto & Loggia, 2009; Bhagat, 2017; Kanti, *et al*., 2017).

Design floods for places near a river are predicted using flood frequency studies. Statistical information such as mean values, standard deviations, skewness, and recurrence intervals are calculated using recorded yearly peak flow discharge data. These statistics have been used to create frequency distributions, which are graphs and tables that show the probability of particular discharges as a function of recurrence interval or exceedance probability (Jos, 2017). Flood frequency analysis is a method of predicting future flooding behavior based on previous records of peak flows.

Flood frequency assessments are primarily used to forecast prospective flood magnitudes throughout time and to determine the frequency with which floods of a given magnitude may occur (Sah & Prasad, 2017). The fitting of a probability model to a sample of yearly flood peaks observed during the observation for a catchment in a certain region is called flood frequency analysis. The developed model parameters can then be utilized to predict extreme events with long recurrence intervals (Pegram & d Parak, 2004). Floodplain management requires accurate flood frequency estimations to protect the public, minimize flood-related costs to government and private enterprises, for designing and locating hydraulic structures, and assessing hazards related to the development of flood plains (Tumbare, 2000).

Various research on the regionalization of basin hydrology has been conducted in Ethiopian River Basins. However, most studies (Gebeyehu, 1989; Sine & Ayalew, 2004; Demissie, 2008); Gedefa & Seleshi, 2009; Mekoya & Seleshi, 2010; Hussein & Wagesho, 2016; Ketsela, *et al*., 2017). However, research was carried out by (Share Bale Eco-Region, 2017) argued different drivers of hydrological dynamics in the research area are vulnerable to flooding, according to the findings. To overcome this problem, the study recommended a regional flood frequency analysis by grouping stations into homogenous regions for the Omo gibe River Basin.

Ethiopia's water resources are inadequately gauged or ungauged, making flood prediction and management difficult (Rabba, *et al*., 2018). This is due to the low density of gauging stations, the difficulty of operating and maintaining gauging networks, and the absence of infrastructure needed to collect acceptable hydrologic data (Gedefa & Seleshi, 2009). The design and successful operation of hydraulic and drainage facilities such as dams, spillways, bridges, culverts channels, and flood protection schemes rely heavily on this data, both in quantity and quality (Saf, 2009; Tanaka, *et al*., 2017). Unfortunately, these vigorous inputs are usually inadequate, in most cases incredibly unavailable at points of interest.

2.2 Flood estimation techniques

The evaluation of flood frequency is critical for flood management. It deals with flood risk assessment, which is required in flood zoning and spatial planning, as well as the arrangement of flow values for the design of flood mitigation and control works (Murphy, *et al*., 2014; Engeland, 2015). The accurate estimation of flood magnitude with the corresponding frequency of occurrence is a challenge for hydrologists because the frequency and magnitude of maximum floods (Chavoshi & Azmin, 2009; Saf, 2009; Javelle, *et al*., 2010; Dubey, 2014; Alam, *et al*., 2016) on the sites of interest are dependent on the planning, management, and design of water resource projects.

Statistical and derived methods for flood frequency analysis are identified in the literature. The modern approach of determining the frequency of peak stream flows is statistical flood frequency analysis. Fitting extreme value probability distribution functions to the historical record of yearly maximum floods is a method of frequency analysis. This method relies on the availability of measured streamflow to fit appropriate probability distributions for gauged sites (Kumar & Chatterjee, 2011; Vivekanandan, 2015).

The flood frequency analysis methodologies that have been developed entail the quantification of the factors that regulate flood behavior and are less reliant on historical data (Badreldin & Fengo, 2012). The computed peak flows are fitted to an extreme value probability distribution in flood statistical analysis. This method is data-driven and only applies to gauge stations. The choice of possibility distribution is frequently haphazard; there is no physical source available to limit the usage of any certain distribution (Wu, *et al*., 2018).

2.3 Flood frequency models

The main goal of flood frequency analysis is to find a Q-T link at every critical location along a river. It is commonly thought that character offers an exclusive Q-T link and that Q is a monotonically rising function of T at any river point (Haberldin & Radtke, 2014). It is necessary to use a statistical or stochastic model of the continuous hydrograph to calculate this natural Q-T connection from a good quality constant hydrometric record of N years' duration, which keeps and discards information in the hydrograph relevant to the Q-T relation (Dessalegn, *et al*., 2016). For this idea, three distinct models are offered. Annual Maximum Series Model, Annual Maximum Partial Duration Series Model, PD or Peak over a threshold, (POT) Time Series Model and TS are a few examples. The challenges relating to the following points must be answered in flood frequency modeling (Dessalegn, *et al*., 2016). Selection of model type (AM or PD), distribution to be employed in the chosen model, and parameter and quintile estimation technique. The significance of two-part components of such an alternative is thought to be illustrious. The descriptive and predictive features of the chosen approach are as follows (Haberldin & Radtke, 2014).

The descriptive property' refers to the requirement that the chosen distribution form closely resembles the practical sample distribution of floods and that arbitrary samples drawn from the chosen model distribution must be statistically similar to the properties of actual flood series. The analytical property' refers to the requirement that quintiles estimates are strong with low bias and standard deviation (Dessalegn, *et al*., 2016) and the following two models were available for this purpose.

2.3.1. Annual maximum series model

Only the peak flow in each year of the record is considered in the annual maximum flow (AMF) dataset. (Desalegn *et al*. 2016) proposed that a series of AMF floods is required to generate a random sample from a stationary population in which the accidental variable has a distribution. Only the peak flow in each year of data is used in the AMF flow series, which may result in some information loss (Chow *et al*.,1988). An AMF is a flood frequency analysis model that is universally employed by different investigators (Badreldin and Fengo, 2012).

2.3.2. Partial duration series model

The majority of the flow hydrograph is ignored in this paradigm, and the hydrograph is interpreted as a sequence of randomly spaced flood peaks of varying amplitude. Only the series of peaks exceeding an arbitrary threshold is evaluated in statistical modeling and identification of the values that make up the series. (Desalegn *et al*., 2016). All peaks over a particular base value are taken into account in partial duration series. The base is normally set low enough to accommodate at least one annual event (Rao and Hammed, 2000). As a result, the yearly maximum flow series model was chosen to overcome the problem of data dependency. Furthermore, the AMF series is a widely and globally utilized model for flood frequency analysis by various scholars (Desalegn *et al*., 2016). As a result, the AMF series model was adopted to avoid the worry about data requirements.

2.4 Regionalization

Regionalization refers to the identification of homogeneous regions through a homogeneity test and the selection of an appropriate frequency distribution for the identified region and stations in the context of evaluating the predictive accuracy of regional flood frequency distribution estimation methods. There is no objective method of regionalization that is universally acknowledged (Sine, 2004). A common major stage in any RFFA is the delineation of hydrologically homogenous regions or zones. The hydrologic characteristics of gauged basins are transferred to ungauged basins through regionalization.

There is no commonly acknowledged impartial way of regionalization because of the complexity of components that influence flood generation (Kachroo, *et al*., 2000); (Mishra, *et al*., 2000). (2009). Expectations in regionalization are based on the statistical similarity of the locations in an area. The values of the coefficient of variation and the site-to-site coefficient of variation must be employed in this analysis. The mean, standard deviation, and coefficient of variance of each site in a region must be computed for the homogeneity test (Nobert, *et al*., 2014). Depending on the availability of gauging stations in the area, hydrologic data can be used in a variety of ways.

Different writers have attempted to identify hydrologically similar locations based on geographical factors, flood data characteristics, or a mix of both (Kachroo, *et al*., 2000). After that, the collection of defined homogeneous catchments can be pooled and described using statistical features. As a result, regional flood approaches are extensively employed, and they give a realistic way to predict discharge in areas with limited data. (Zaman and colleagues, 2012).

2.4.1 Identification and Delineation of Homogeneous Regions

A basic stage in regional flood frequency analysis is the identification and definition of homogeneous regions (RFFA). In recent years, there has been a lot of interest in identifying flood-producing natures in data-poor areas (Smith, *et al*., 2015). The application usually entails assigning an ungauged watershed to a suitable homogeneous group and predicting flood quantiles with created models based on catchment features. To put it another way, an RFFA based on homogeneous areas can transfer information from similar gauged catchments to ungauged catchments, allowing for flood prediction (Haddad, 2013).

(Hosking & Wallis, 1997) Mentioned all the stages in RFFA involving many sites. The authors pointed out that identifying homogenous areas (IHRs) is often the most challenging and demands the most subjective assessment. The identification of spatially contiguous regions can be used to divide the basin into homogeneous regions. Geographical proximity, on the other hand, does not imply hydrological similarities (Patil & Stieglitz, 2012).

There are no clear rules for selecting homogeneous zones due to the complexity of comprehending the components that have a direct and indirect influence on flood formation (Kachroo, *et al*., 2000). Meanwhile, prior knowledge, experience, and personal judgments might help to define regions with similar hydrological characteristics. Several authors attempted to identify hydrologically homogeneous locations, focusing on either geographical considerations or hydrological traits, or a mix of the two (Kachroo, *et al*., 2000).

2.4.2. Statistical Homogeneity Tests

The assumption behind regional flood estimating methods is that a standardized flood variate has the same distribution at every location in the designated region. The significance of homogeneity has been proven by (Demissie, 2008). Homogeneity means that the flood-generating mechanisms in different areas are similar. A more precise definition of a homogenous zone is a group of locations that share the same standardized frequency distribution form and parameter. To see if the preliminary defined and demarcated region is homogeneous, homogeneity tests based on Cv and LCv are used. In this situation, hydrological data must be used, and the region is considered homogeneous if it meets both homogeneity test conditions (Nobert *et al*., 2014). The discordance metric is used to detect sites that are significantly out of sync with the rest of the group. It calculates how far a given location is from the group's center. It's also a good idea to filter out data from atypical locations when looking for relevant datasets for regionalization. These locations were chosen because of data flaws or other local factors (Rao and Hamed, 2000; Noto and Loggia, 2009; Guru and Jha, 2016; Kanti *et al*., 2017) The delineation of the homogeneous region is important for site characteristics to be truly representative of the observed discharge data used to estimate hydrologic design values (Irwin *et al*., 2014).

2.5 Statistical distributions for flood frequency analysis

The fundamental goal of regional flood frequency analysis is to find a distribution that produces as precise quantile values for each site as possible. When numerous distributions fit the data well, the optimal distribution is the one that can give good quantile estimates (Hosking & Wallis, 1997). Because there is no sound physical foundation to explain the selection of a distribution for a certain application, it is usually made arbitrarily (Rahmana, *et al*., 2015). The distribution models were chosen based on past research, with the majority of them having been applied and recommended in various nations. Many factors influence this, including the methods used to discriminate across distributions, the methods used to estimate parameters, and the availability of data (Kumar & Chatterjee, 2011).

2.5.1 Best fit probability distributions

The process of determining a suitable probability distribution for a given dataset is known as probability distribution fitting. Flood frequency analysis provides a precise estimate of the greatest flood by fitting a probability distribution for a given return time (Vivekanandan, 2015). The goal is to anticipate the frequency of occurrence of the phenomenon's magnitude in a certain interval. This can lead to an accurate flood forecast. The most closely fitting probability distributions to the observed data are determined by the nature of the occurrence and the distribution (Athulya & James, 2012).

As a result, the most critical component in frequency analysis is selecting the best statistical distribution. As a result, several distributions must be used, and then the most suited data distribution should be chosen (Amirataee, *et al*., 2014). In flood frequency analysis, the available data is fitted to an assumed probability distribution to estimate the flood amplitude for a certain return time. In this document, you'll find details on some of the most widely utilized distributions in flood data (Rao & Hamed, 2000). The first error is connected with the incorrect assumption of a particular distribution for the given data, which can be checked to some extent using goodness-of-fit tests (Millington, *et al*., 2011). To analyze the reasonability and check the appropriateness of best-fitting probability distributions to the recorded data, several goodness-of-fit tests were performed, including the Kolmogorov-Smirnov test, the Anderson-Darling test, and the chi-square test at the significance level (0.05).

In the section about flood data distributions, you'll find information on the most widely used distributions (Rao & Hamed, 2000). The first error, which is connected with the incorrect assumption of a specific distribution for the given data, can be checked to some extent using goodness-of-fit tests (Millington, *et al*., 2011). To analyze the reasonability and check the appropriateness of best-fitting probability distributions to the recorded data, a variety of goodness-of-fit tests were performed, including the Kolmogorov-Smirnov test, the Anderson-Darling test, and the chi-square test at the significance level (=0.05).

2.5.2 Goodness of fit tests

The goodness of fit test is a statistical model that uses a theoretical probability distribution function to produce a well-matched random sample. To put it another way, these tests demonstrate how well the chosen distribution matches the data. The goodness of fit test compares observed values to expected (fitted or predicted) values in the same way that linear regression does. Kolmogorov Smirnov, Anderson-Darling, and Chi-Squared goodness-of-fit tests were used to assess the distributions' performance (Rao & Hamed, 2000).

2.5.3 Method of L-moment ratio diagram

L-moment ratio diagrams can be used to evaluate the adequacy of a probability distribution. This strategy is effective in selecting a distribution for a region in regional frequency analysis (Das & Simonovic, 2012). The L-moment ratio diagrams (LMRD) are a trusted diagnostic tool for determining a probability distribution. For a goodness-of-fit test, this is always superior to a product-moment ratio diagram because it allows a visual comparison of the sample estimates with the population values of L-moments (Hosking & Wallis, 1997). (Amalina, *et al*., 2016). The ability to compare the fit of several distributions using a single graphical instrument is a benefit of LMRD (Chavoshi & Azmin, 2009).

The L-moment ratio diagram is commonly used as the first visual examination tool for picking a regional frequency distribution from sample data. By plotting the sample L-moment ratios and average sample L-moment ratios (3 and 4) or record length weighted average L-moment ratios as a scatterplot with theoretical curves of several candidate distributions in an L-Skewness-L-kurtosis space, the L-moment ratio diagram can provide an elementary visual judgment of a regional frequency distribution. The chosen distribution should be as near to the regional data as possible (Lu, 2016). (Hosking & Wallis, 1997).

2.5.4 Parameter Estimation

For parameter estimation, only the ordinary methods of moments (MOM) were previously described. MOM, PWM, and ML are the most efficient methods of parameter estimate known because of their lesser inaccuracy in quintile estimation, according to several parameters estimating approaches that have been suggested and researched (Badreldin & Fengo, 2012). The Maximum Likelihood (ML), Approach of the Moment (MOM), Probability Weighted Moments, and L-Moment method are all estimate methods. The L-Moment approach is used to estimate the parameters for this study since it is less affected by outliers and severe data series. (1997, Hosking & Wallis). Some of the parameter estimation methods may not yield good estimates. Hence, some guidance is needed for estimation methods.

i. Method of maximum likelihood

The maximum likelihood technique (MML) is regarded as the most accurate method, especially for big datasets. Because it leads to efficient parameter estimators with Gaussian asymptotic distributions, when compared to other approaches, it delivers the least variance in the estimated parameters, and hence the estimated quintiles. However, with small samples, the results may not converge. This method involves the choice of parameter estimates that produce a maximum probability of occurrence of the observations (Cunnane, 1989). In general, the PWM and MOM are better for estimating the parameters for three and two-parameter distributions respectively of the underlying distribution from which the data are sampled. They are less susceptible to sampling variability (outliers) than others, resulting in more accurate and robust estimates of the underlying probability distribution's features or parameters (Rao and Srinivas, 2008).

ii. Method of moment

The method of moment (MOM) is a simple and widely used method of determining probability distribution parameters. It can also be used to get beginning values for numerical techniques used in ML estimates. MOM estimates, on the other hand, are generally less efficient than ML estimates. Higher-order moments are more likely to be utilized to establish starting values for numerical procedures involved in ML estimation and to be strongly biased for relatively small samples, especially for distributions with a large number of parameters. In recent years, the L-moment approach has become the most preferred method for frequency analysis established by (Hosking and Wallis, 1997)

L-Moments (LMM) is a system of expressing the shapes of probability distributions that is similar to the method of moments and linear functions of the expectations of order statistics. It is a powerful and efficient way for computing statistical parameters because such methods can provide an unbiased estimate of sample parameters that is unaffected by outliers (Ghosh *et al*., 2016; Rao and Hamed, 2000). The L-moments technique illustrates reliable forecasts of all types of statistical analysis; therefore, it can be recommended for policies and decision-making in hydrological catchment planning (Kanti *et al*., 2017). L-moments can characterize a wide range of distributions when compared to the method of moments and maximum likelihood. L-moment sample estimates are sufficiently powerful that they may not be affected by the presence of an outlier in the dataset and are less prone to estimation bias. L-moments can provide precise parameter estimates for a fitted distribution (Cunnane, 1989).

iii. Probability-weighted moments

Probability-Weighted Moments (PWMs) are useful in deriving expressions for distribution parameters that can be explicitly described. This approach obtains parameter estimation by equating the distribution's moment with the relevant sample moment of observed data. The initial sample moments are set equal to the corresponding population moments for a distribution with a parameter. The unknown parameters were then solved simultaneously using the equation that resulted. PWM parameter estimation is a relatively recent technique that is as simple to use as regular moments, is usually unbiased, and is nearly as efficient as MML. PWM may be equally efficient as MML in small samples; with larger sample size, however, MML may be more efficient.

The data fitting procedure entails the application of statistical techniques that allow for the estimation of fitness parameters in line with the data sample. One advantage of utilizing software to fit data and understand probability data is that it can automatically fit data with many known distribution patterns at the same time. R studio programming Software is a data analyzer and simulation program capable of fitting and simulating statistical distributions with sample data, selecting the best model, and then using the analysis results to make better judgments. For several distributions, R studio used for the selection parameter estimation approach for fitting distribution to a dataset maximum likelihood method moment matching method, quantile matching approach, and maximizing goodness of fitting approach.

R studio Software is an interactive software system to identify parameters, allows the most flexible input of the underlying model in form of FORTRAN code, and is executable independently from the interface. It consists of a database containing models, data, and results, and of underlying numerical algorithms for solving the parameter estimation problem depending on the mathematical structure (Schittkowski, 2002). The selections of the distribution models are based on the previous studies where most of these have been used and recommended in various countries. In this study selection of best-fit probability distribution and its method of parameter estimation suitable for each distribution within the interface were conducted using R studio software due to the results of analysis leading to taking a better decision (Romani and Yusop, 2017). According to Irwin et al. (2014), watersheds are delineated using ArcGIS with DEM data and subsequently, several flood generation characteristics are assigned to each watershed. The outcome of this procedure can be directly applied in regionalization to group watersheds into hydrologically homogeneous regions based on the similarity of their attributes, and hydrologic variables are estimated from the regions. Hence, to delineate and characterize watersheds for regionalization ArcGIS10.4.1 environment was used for this study using the procedure of (Abdulla, 2011) and (Irwin *et al.*, 2014)

2.6 Quantile Estimation and Derivation of the Flood Frequency Curve

After estimating the parameters of a distribution, quintile estimates (X_T) for different return periods T can be derived. The chance of non-exceedance (F) is related to the return period by the relation F=1-1/T, where F= F (X_T) is the likelihood of having a flood of magnitude X_T or smaller. There are two types of distribution functions that are found in practice. The first is that which may be represented in the inverse form $X_T = \varphi$ (F). In this situation, X_T is calculated by substituting its value from the above equation for φ (F). The distribution of the second kind cannot be directly represented in the inverse form X_T = (F). Frequency analysis and is estimated by applying a distribution function once the parameters of distribution have been estimated.

The chosen quantile of under or over design criterion for hydraulic structures is at risk since the return period is dictated by the structure's cost and economic-strategic importance. After estimating the parameters of a distribution, selecting a reliable design quantile is required for the delineation of floodplains, the development of floodplain management, and flood warning systems. The effects of statistical methods used in parameter estimation belonging to the probability distribution on the design, operation, and management of a hydraulic structure are highly dependent on statistical methods used in parameter estimation belonging to the probability distribution. (Amalina and colleagues, 2016) The parameter estimates that maximize the likelihood function are generated by partial differentiation for each parameter, setting the partial derivatives to zero, and then solving the resulting set of equations all at once. As a result of this problem, the equations are frequently complex, and the solution set may not be properly determined (Cunnane, 1989). Although using these parameters provides less skewed estimates than using two, there is no broad consensus on the ones to employ (Parida, *et al*., 1998).

When quantiles must be computed for sites where no observations have been recorded or where observations have been obtained for only a short length of time, frequency analysis estimates are neither possible nor accurate. RFFA is one of the methods for overcoming these issues while accurately quantifying flood forecasts at appropriate frequencies for series inside a more or less hydrologically homogeneous region (Dubey, 2014).

2.7 Derivation of Flood Frequency Curves

The peak annual flow of a given stream is plotted on a flood frequency curve at a specified point in the recurrence interval or return period. The annual likelihood of exceeding a certain flood flow is calculated using flood frequency curves. Regional flood frequency curves can be used to estimate flood quantiles at an ungauged site within the region by taking into account the spatial pattern of change of hydrologic phenomena over numerous gauging sites (Ergish, 2010).

The relationship between the amplitude of river peak flows and the recurrence interval or return time is described by flood frequency curves (FFC). When calculating flood risk, a flood frequency curve and flood estimation for various return times are required. The development of FFC for various return durations aids in the estimation of flood quantiles (Das & Simonovic, 2012). The primary purpose of every RFFA is to create a regional curve that can represent the averagely weighted distribution of homogeneous regions. The normalized regional quantile floods (XT); FFC for a specific return period are estimated using the final phase of flood frequency analysis (Tadesse, *et al*., 2011).

The model parameters produced from the best-fitted distribution to the observed data are used for a given region. This aids in the calculation of standardized quantile estimates, which are then utilized to create a regional flood frequency curve for the homogenous region. These graphs are quantile plots that depict all locations in a homogeneous region (Hailegeorgis & Alfredsen, 2017).

2.8. Previous Studies on Regional flood frequency analysis in Ethiopian River Basins

Gebeyehu (1989) investigated regional flood frequency analysis based on monthly rainfall patterns and geographical closeness for the Blue Nile River Basin. Because the responses of the statistical technique in similar rainfall regions are diverse outcomes of changes in basin topography, the study has some limitations in that it does not correctly designate homogeneous regions. (Gebeyehu, 1989) points out the following information in his conclusion. The regionalization approach provides useful information about the flood frequency of gauged and ungauged catchments; a small amount of site data greatly improves the estimate of the mean annual flood that can be used with a regionally based estimate of the XT relationship, and the results of regional flood frequency analysis should be updated as more relevant information becomes available.

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

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

According to (Hussein and Wagesho 2016), the regionalization of the Abaya-Chamo sub-basin was performed based on site characteristics such as elevation, soil type, soil texture, slope, land use land cover, and mean annual rainfall. Site statistics were used for testing of homogeneity of the proposed region. The authors concluded that to get a reliable quantile estimate more gauging stations should be installed in the basin to infer something for ungagged sites. (Ketsela *et al*. 2017) performed FFA on Awash River Basin using statistical distribution technique. The Easy Fit Software was employed for the selection of best-fit distributions and estimation of parameters for stations. Kolmogorov–Smirnov test was used for the choice of a suitable distribution for estimation of maximum flood discharge. According to this study, the awash basin was delineated into five satisfactory homogeneous regions and recommended software-based techniques like Easy Fit and other alternative statistical software packages to get accurate and reliable flood estimation results.

3. MATERIALS AND METHODS

3.1 Study area Discretion

3.1.1 Location

The Upper Omo-Gibe basin is one of the major river basins in Ethiopia and is situated in the south western part of the country covering parts of SNNPR and Oromia region. The basin lies between 5°40' to 9°40'N latitude and 35°40' to 38°20'E longitude. The basin is characterized by diverse topographic features with elevations ranging from 746 m asl in the southern part to 3522 m asl in the northern highlands. The rainfall in the basin has a mono-modal pattern. The mean annual rainfall during the study area is 1425 mm and 92% of the annual rainfall occurred during the wet season (March–October), while the dry season (November–February) receives only 8% of the annual rainfall (Adnew and Woldeamlak, 2013). The mean annual temperature is 19.2 ◦C. The area is drained by some of the major rivers of the country, such as the Omo, Gilgel Gibe, Gojeb, and their numerous tributaries. These have created the dissected terrain. Nearly half of the country's remaining natural forests are found in this region.

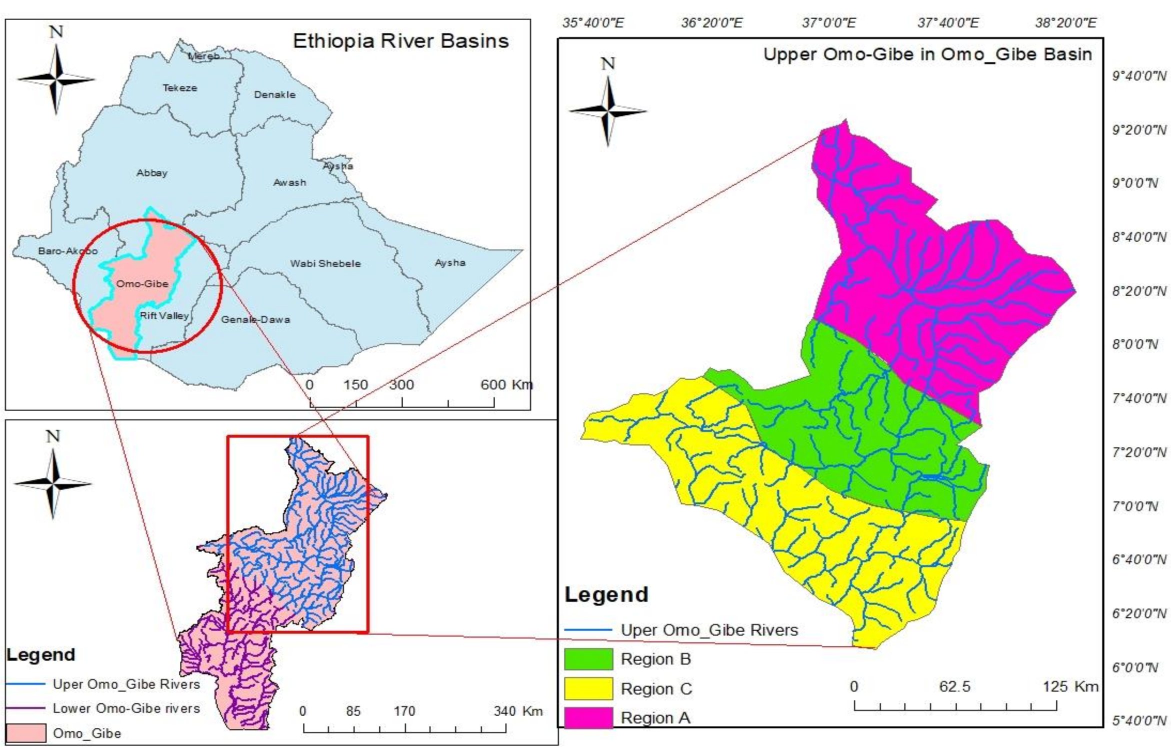


Figure 3. 1 Study area map

3.1.2 Climate

The Upper Omo-Gibe River Basin's climate ranges from a hot, arid environment in the floodplain's southern reaches to a tropical, humid climate in the highlands of the Basin's extreme north and north-western reaches. The climate is tropical sub-humid for most of the basin, which lies halfway between these extremes. In the study area, annual rainfall ranges from over 1900 mm in the north-central parts to less than 300 mm in the south. Furthermore, the rainfall regime in the northern and central regions of the basin is unimodal. The average annual temperature in the Upper Omo-Gibe Basin ranges from 160°C in the northern highlands to over 290°C in the southern lowlands. (Richard Woodroof and Associates, 1996).

3.1.3 Topography

Physical variation characterizes the terrain of the Omo-Gibe basin as a whole. The Omo, Gojeb, and Gilgel-Gibe Rivers gorges pass through mountainous to hilly terrain in the northern two-thirds of the basin, while the southern one-third is a flat alluvial plain broken by steep sections. The northern and central halves of the basin are over 1500 m a.s.l., with a maximum elevation of 3522 m a.s.l. (between the Gilgel-Gibe and Gojeb rivers), and the lower Omo plains are between 400 and 500 m a.s.l.

The Great-Gibe River's headwaters are around 2200 meters above sea level. The Gibe River flows southwards, towards the Omo River and subsequently to Lake Turkana, a fault feature filled with alluvial and lacustrine materials of recent origin connected with the Great Rift Valley, despite the presence of some significant tributaries from various directions. In its lower sections, southwest of the confluence with the Gojeb River, the Gibe River is known as the Omo River. This is how the Omo-Gibe River Basin got its name. (Richard Woodroof and Associates, 1996).

3.1.4 Land Use

In a general sense, much of the Omo-Gibe Basin's northern catchments are under intensive cultivation with increased land pressure, which means that cultivated areas are expanding into increasingly marginal regions at the expense of forest lands. Deforested regions are currently restricted to places that are too steep and difficult to agriculture. The northern catchments' flatter, poorer-drained bottomlands are normally not cultivated but are used for dry-season grazing and eucalyptus tree plantations. The main gorges of the basin are relatively unpopulated and support a cover of open woodland and bush land with grasses, the eastern part of the basin has some of the most densely populated and intensively farmed areas in the country, let alone the basin. The south of the basin is more sparsely populated with a greater population of natural vegetation, though even here the forest is decimated at an alarming rate. (Richard Woodroof and Associates, 1996)

3.2 Materials and Tools/Software's used

The materials indispensably used in conducting this study were; kinds of literature, the internet, etc. In addition, the tools used to undertake the research were software. The general descriptions of the Tools/Software's are described herein Table 3-1.

Table 3. 1 Description of Tools/Software's used

NO	Tools used	Function
1	Microsoft excel spreadsheet and XLSTAT2018	For data arrangement, filling missed data and calculate the statistical parameters of hydrological data. Used to fill in missing streamflow data of stations
2	SPSS Software	Used for data quality control.ie for F-tests, and T-tests
3	Arc-GIS	For the delineation of the study area map, delineation of the hydrologically homogeneous region.
4	R studio software	For the selection of suitable probability distribution for each selected station, selection of parameter estimation methods and estimation of the goodness of fit.
5	Origin plot 2019	Plot curves and graphs that are huge, create and share diagrams and templates and resize shapes curves and diagrams
6	Matlab2018a	For the measures of discordancy of the site from identified regions.

3.3 Data Collection and Analysis

It is critical to define a clear and efficient approach for the study's findings to be of high quality. The data analysis processes in this work range from preliminary data screening to developing a regional flood frequency curve based on AMF series data. The data were screened to look for major errors and to ensure that the data was consistent. After identifying pertinent data from the study basin that would be beneficial for the regional analysis, the data was checked for quality. The goal of identifying homogenous regions was to determine whether sub-basins may be grouped based on their flood-producing nature. This was done using the L-moment ratio diagram and station site parameters. The regional frequency distribution was calculated using average L-moment ratios, and the goodness of fit test was performed using R studio software to ensure that the chosen distribution fits the data in the region. The flood quantiles for specified return periods at ungauged locations are then computed using the estimated frequency distribution derived from the regional growth curve.

The following approaches were used in general to assess the predictive accuracy of regional flood frequency estimation in this study. A homogeneous region is defined, and standardized data from various sites within the region can be pooled together to produce a single frequency curve appropriate to the region. It is difficult for hydrologists and engineers to generate credible flood estimates directly when enough rainfall or river flow records are not available at or near the place of interest, thus regional studies might be valuable.

3.3.1 Sources and Availability of Data

To estimate flood magnitude, flood frequency estimation generally uses recorded yearly maximum flood data at gauging stations. The Ministry of Water, Irrigation, and Electricity's department of hydrology and GIS provided hydrological and DEM (digital elevation model) data for the Omo-Gibe River Basin. DEM data was used as a starting point for defining the basin's boundaries and pinpointing the locations of the gauging stations. There are around twenty gauging stations in the research area, but only eleven gauging stations were chosen for adequate regional flood frequency calculation. The chosen stations do not have fully recorded data on their own; they have several years of records with missing data that must be filled in before analysis. As a result, eleven gauging stations were chosen. The site characteristics of stations for this study include the code of the stations, the name of the river and their gauging sites, the locations (latitude and longitude), and the catchment area in km^2.

Table 3. 2 The site characteristics of stations used in detail analysis

Code	River name	Location of gauging station	Latitude	Longitude	Area (km2)	Record period	Recor d length
111111	Ghibe	Tollay	8'25"10.9'N	37:2:58E	2,572	2000-2019	20
61015	Gibe	Abelt	8'13.47"N	37:34:44E	15,746	1985-2017	33
91004	Wabi	Wolkite	8'14"53.7N	37:45:34E	1,866	1975-2009	35
91010	Walga	Wolkite	8'19"32.1N	37:35:25E	1,792	1975-2005	31
91007	Gogob	Endeber	7°50′0N	37°40′0″	109	1990-2014	25
91008	Gilgelghbe	Asendabo	7:45: 0 N	37:11: 0 E	2,966	1982-2016	35
91032	Bulbul	Serbo	7:34: 0 N	37: 5: 0 E	526	2000-2017	18
91023	Kito	Jimma	7:42: 0 N	36:50: 0 E	85	1990-2013	24
92002	Gecha	Bonga	7:17: 0 N	36:13: 0 E	175	1991-2019	29
91012	Gojeb	Shebe	7:25: 0 N	36:23: 0 E	3,577	1974-2008	35
92004	Guma	Andaracha	7: 9: 0 N	36:15: 0 E	231.3	1990-2015	26

3.3.2 Data Screening

Data screening is the initial task in which procedures are used to filter out undesired observations from the data series as well as analysis sites. It is used to ensure that the data used in the regional flood frequency calculation are correct (Kumar and Chatterjee, 2011; Kachroo et al., 2000). This allows for visual assessment of whether observations have been routinely or unintentionally ascribed to the wrong day, or if decimal points have been misplaced. Visual inspection of daily flow data revealed problems such as inflated numbers, missed decimal points, and extremely high and/or extremely low flow records during dry months and/or rainy months. Streamflow data from gauging stations in the Upper Omo-Gibe River Basin was used in this investigation.

The at-site AMF records have a minimum and maximum length of 18 and 35 years, respectively. The AMF data were selected and afterward submitted to investigative data analysis for all of the stations indicated in Table 3.2 and presented in figure 3.3 to choose representative stations for the study area.

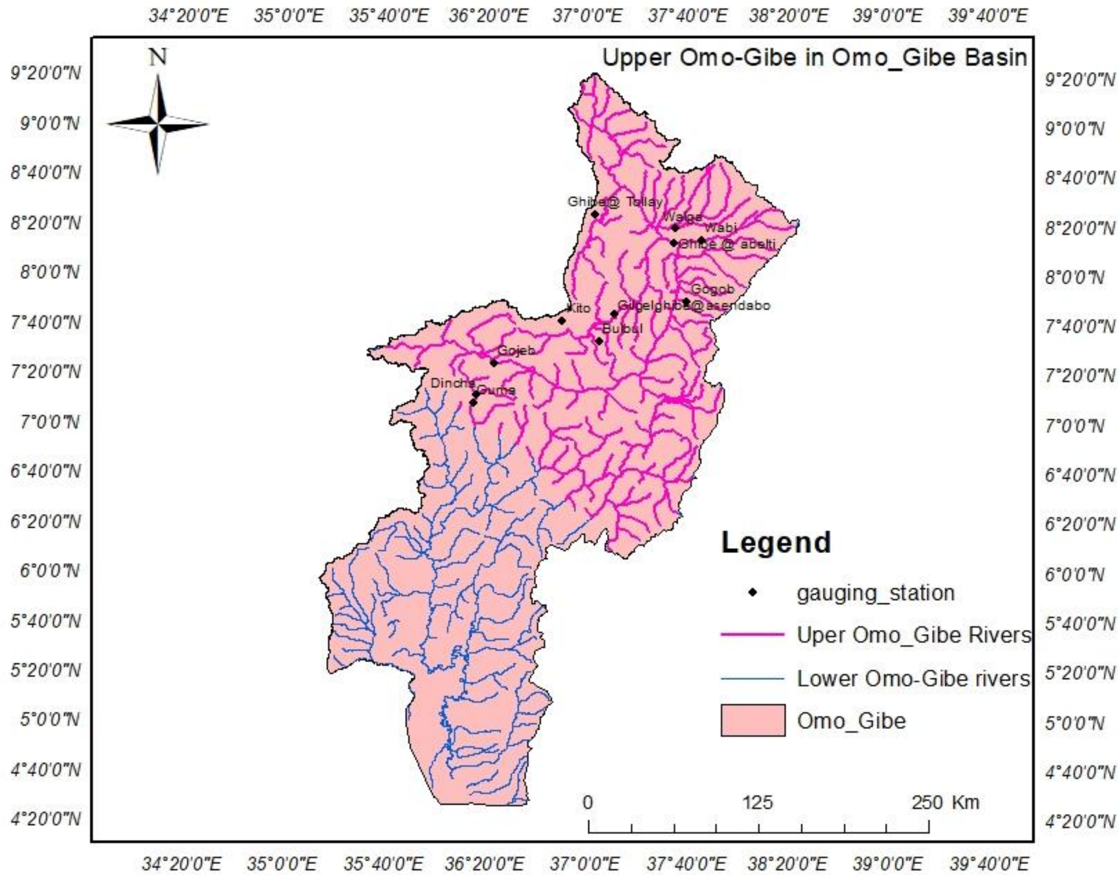


Figure 3. 2 The spatial distribution of gauging stations in Upper Omo-Gibe River Basin

3.3.2 Missed data filling

When analyzing runoff data from gauges that take monthly observations, it's common to find months where no observations are taken at one or more gauges. The record's stability can be harmed by missing data for a variety of reasons, including the absence of a recorder, the observer's carelessness, and the break or failure of instruments. It's common to have to make educated guesses about these missing records. Data from nearby stations can be used to estimate missing data (Sine, 2004). For a given gauging station, different methods are employed to fill in missing flow data. Arithmetic Average Method, Graphical Correlation Method, Normal Ratio Method, and Linear Regression Method are a few examples.

In this study, any missing data were filled in using the linear regression method by using the excel state software. Using adjacent flow gauging station observations, simple linear regression was used to fill in missing streamflow estimates. The following are the reasons why this strategy was chosen: i) It is the most extensively utilized method for massive data when compared to other methods. ii) As accurate estimation of significant missing observations as possible. iii). It is used by establishing a link with a nearby station.

The equation for linear regression is given as:

$$Y = ax + b \qquad (3.1)$$

3.4 Data quality control

In some circumstances, errors in the streamflow observation to be collected may exist, such as misplaced decimal values, extremely large unrealistic numbers, and negative flow records. A critical step is to test observation quality before employing it. To check the quality of streamflow data for this investigation, the following procedures were examined.

3.4.1 Test for randomness and independence

Flood frequency analysis FFA is known to be carried out in principle when the at-site data are independent and identically distributed requirements are met (Hosking & Wallis, 1997). This implies that extreme occurrences could occur at any time and have the same frequency distribution. Independence is one of the most important assumptions in frequency analysis, and the interstice correlation has a significant impact on the variance of regional parameters and flood quantiles, as well as reducing the effective length of records. The randomness test is required to determine independent annual maximum (AM) series from all data set values at each station, according to (Guru & Jha, 2016). It is assumed that all the peak magnitudes in the AM series are mutually independent in the statistical sense.

The AMF at multiple sites in a homogenous region must be spatially independent, according to the criteria of regional flood frequency analysis (RFFA) (Hailegeorgis & Alfredsen, 2017). It should be noted that the correlation coefficient was used in this study to validate the independent data from the chosen hydrological stations. The lag-1 serial correction coefficient R is defined as follows by Dahmenand Hall (1990)

$$R = \sum_{n=1}^{\infty} \frac{(xi-\bar{x})(xi+1-\bar{x})}{(xi-\bar{x})} \quad \text{……………………………………………………………………………3.2}$$

Where Xi is an observation,

Xi+1 is the following observation and

n is the amount of data.

After computing R1, the test hypothesis is H0: R1= 0 (no connection between two consecutive observations) against H1: R1> 0. At the 5% level of significance,

Anderson (1942) defines the crucial area, R1, as (-1, (LCL) R1 (UCL), 1), and equation 3.2 gives: R1's upper confidence limit (UCL) is calculated as follows:

$$UCL(R1) = \frac{(-1+1.96(N-2)\ 0.5)}{N-1} \quad \text{……………………………………………………………......3.3}$$

The lower confidence limits, LCL, for R1 as:

$$LCL(R1) = \frac{(-1-1.96(N-2)\ 0.5)}{N-2} \quad \text{………………………………………………………………3.4}$$

To accept the hypothesis Ho: R1=0, the value of R1 should fall between the UCL and LCL. Applying this condition to the time series, the condition: LCL (R1) <R1< UCL (R1) is satisfied for all stations.

Table 3. 3 Result of test for independence of stations time-series data

Station	LCL (R1)	R1	UCL (R1)	Station	LCL (R1)	R1	UCL (R1)
Tollay	-0.556	-0.251	0.422	Serbo	-0.52	0.402	0.456
Abelti	-0.361	0.309	0.302	Jimma	-0.537	0.016	0.412
Wolkite	-0.361	0.042	0.302	Bonga	-0.385	0.290	0.318
Wolkite	-0.372	-0.174	0.309	Shebe	-0.372	0.242	0.309
Endeber	-0.052	0..012	0.456	Andaracha	-0.477	0.344	0.377
Asendabo	-0.361	0.126	0.32				

As a result, there is no correlation between successive observations. The data are unrelated, and the time series has no consistency. The summarized result of the test for annual maximum flow series for example for Tollay station -0.556<R1< 0.422

3.4.2 Test for Consistency and Stationary

If the periodic data are proportional to an adequate contemporaneous time series, a time series of hydrological data is relatively consistent (Dahmen & Hall, 1990). The F-test for variance stability and the t-test for mean stability, according to (Dahmen & Hall, 1990), evaluate not only the stationary of time series but also their absolute consistency and homogeneity. If the F-test reveals stable variance and the t-test shows a stable mean, then the time series is steady, consistent, and homogeneous, according to this. The following are the two criteria used to ensure the consistency and stationary of streamflow observations:

a. F-test for Stability of Variance

The ratio of the variances of two split, non-overlapping subsets of the series is the test statistic (Dahmen and Hall, 1990). The annual maximum streamflow observations are separated into time series that are equal or nearly equivalent. Then, for each gauging station, the variation of each time series is determined.

The test statistic (Ft) is calculated as:

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

According to this method, the variance of the time series is stable if and only if:

F (V1, V2, 2.5%) < Ft<F (V1, V2, 97.5%)

Where, $V_1 = n_1 - 1$ and $V_2 = n_2 - 2$, and

n_1 and n_2 -the number of observation points in each subset.

b. Test for the Stability of Mean

The mean stability test entails calculating and comparing the mean of non-overlapping subsets of the time series (Dahmen and Hall, 1990). The t-test values are calculated using the same subgroups as the F-test.

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

$$Tt = \frac{\bar{X}series1 - \bar{X}series2}{[(n1-1)S1^2 + (n2-1)S2^2 * \frac{1}{n1+n2-2*\left(\frac{1}{n1}+\frac{1}{n2}\right)}]\ 0.5} \quad \ldots\ldots 3.6$$

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

n: is the number of monthly streamflow records

S: is the standard deviation of the two series

The mean of the time series is stable according to the Stability of Mean test if and only if: t (V, 2.5 percent) < Tt < (V, 97.5 percent). The value of V varies per station, and the values are taken from Appendix-D using percentile columns (2.5 percent and 97.5 percent).

Noting that both F {V1, V2, 2.5%}, and F {V1, V2, 97.5%} values for 5% significance level as Appendix-B. For the station having years are listed using V1, V2 and percentile row 2.5 % or 97.5 % Appendix-C.

For this research, the T-test and F-test were conducted by using SPSS statistical software. SPSS is software for editing and analyzing all sorts of data. These data may come from basically any source: scientific research, a customer database, Google Analytics, or even the server log files of a website. SPSS can open all file formats that are commonly used for structured data such as spreadsheets from MS Excel or Open Office, plain text files (.txt or .csv), relational (SQL) databases, Stata and SAS. An SPSS data file always has a second sheet called variable view. It shows the metadata associated with the data. Metadata is information about the meaning of variables and data values. This is generally known as the "codebook" but in SPSS it's called the dictionary.

The results of observations of data of gauging stations T-test and F-test, are presented in Appendix-E, and show that mean and variance of the time series was stable. The reason why SPSS software is selected, i) Quick and easy to learn ii) Can handle large amounts of data iii) Great user interface iv) SPSS can take data from almost any type of file and use them to generate tabulated reports, charts, and plots of distributions and trends, descriptive statistics and conduct complex statistical analyses.

3.4.3 Check for data adequacy and reliability

The sample size affects the accuracy of statistical the mean. The data used for analysis was double-checked for accuracy and consistency. The data's accuracy and sufficiency were evaluated and specified.

Using equation 3.7, the accuracy and appropriateness of data were assessed and defined in (McCuen, 1998).

$$\mathrm{De} = \frac{\mathrm{Cv}}{N^{0.5}} \quad \text{.......3.7}$$

Where, De- Standard error

Cv-Coefficient of variation and

N-number of yearly data in the series

Table 3. 4 Results of test for adequacy and reliability of AMF data

Station	Cv	N	De	Station	Cv	N	De
Tollay	0.291	16	0.073	Serbo	0.556	18	0.093
Abelti	0.173	35	0.029	Jimma	0.488	17	0.077
Wolkite	0.615	35	0.09	Bonga	0.598	31	0.091
Wolkite	0.626	33	0.087	Shebe	0.353	33	0.061
Endeber	0.708	18	0.067	Andaracha	0.796	21	0.099
Asendabo	0.305	35	0.052				

If De is less than the 10% significance threshold, the data series can be considered dependable and acceptable. As a result, the data from the stations are judged to be accurate, adequate, and dependable, as the Devalue for most of the stations is less than 10% significant. (McCuen, 1998).

3.4.4 Check for Outliers of the data series

An outlier is a data point that deviates significantly from the rest of the data. This could be due to data entry errors, decimal point losses, abnormally high and/or extremely low flow records during dry months and/or extremely low flow records during rainy months, or natural causes. For statistical tests of outlying observations, it is usually recommended that a moderate significance level, such as 1%, be employed, and that significance values greater than 5% should not be standard practice. (Grubbs, 14969) as quoted in (Dahmen & Hall, 1990; Ketsela, *et al*., 2017) However, L-Moment will apply an effective parameter estimate technique to minimize or eliminate the impact of outliers in this study.

3.5 Regionalization of Upper Omo-Gibe River basin

The index flood L-moment approach of regionalization was utilized in this study based on the data homogeneity of the stations. The statistical values for the stations must be examined to determine whether they may be grouped into one or more categories. Flood statistics for Upper Omo-Gibe River basin stations were calculated using L-moment methods. These approaches are widely utilized because they give a balanced estimation of sample parameters and are not easily altered by the presence of outliers (Rao & Hamed, 2000).

3.5.1 Identification of homogeneous regions

The first stage in regional frequency analysis is to identify homogeneous regions. The specification of variables indicating this similarity has been made to identify homogeneous zones. The most challenging stage is frequently identifying homogeneous zones, which necessitates the most personal judgment (Amalina *et al*., 2016). To determine the degree of heterogeneity within the pool, discordant measures and homogeneity tests are used to statistically verify the regionalization process. This is demonstrated by comparing the scale and dispersion values of the L-moment (LCv) and conventional moment (Cv) of gauging stations from various regions. LCv and Cv best explain the statistical character of yearly maximum flow variation among a collection of stations.

The stream gauging stations were grouped into spatially continuous sites to ensure those stream responses to physiographic variables were comparable. The basin was utilized to identify site characteristics using a DEM with a size of 30mx30m. This allows for the transfer of streamflow records from gauged basins to ungauged basins within a region (Sine *et al*., 2013). Check for station and region homogeneity

3.5.1.1. Site Characteristics

Preliminary IHRs of stations into a specific group are determined in this study by examining station site features. As a preliminary IHR, the following site variables were used: latitude and longitude, AMF, station area, and flow gauging station altitude. Stations with nearly identical site features are then clustered together in the same region.

3.5.1.2. Method of L-Moment Ratio Diagram

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

3.5.2. Test for Homogeneity of Stations and Regions

Various homogeneity tests must be performed on the initially selected regions. Because L-moments are a linear combination of data, they are less influenced by outliers, and the bias of their small sample estimates is kept to a minimum. (Hosking and Wallis, 1997). provided unbiased sample estimators for the first four PWMs and proposed a homogeneity test based on L-moments, which proved to be effective. Discordance measure tests, a measure of scale, dispersion-based tests (Cv-based homogeneity test and LCv-based homogeneity test), and statistical comparison were employed in this work.

3.5.2.1. Discordancy Measure of Regions

The discordance metric is used to detect sites that are significantly out of sync with the rest of the group. Based on statistical features, the discordance measure Di calculates how far a specific location is from the group's center (Rao and Hamed, 2000). When Di is larger than or equal to 3, it is a good criterion to define a station as discordant (Hosking and Wallis, 1993). If a vector Ui = I $_2$, I $_3$, I $_4$,) T is the transpose of the vector Ui (Hosking and Wallis, 1997), then the discordancy measure can be defined as

$$Di = \frac{1}{3}(Ui-\bar{U}i)\, S^{-1}(Ui-\bar{U}i)^{T} \qquad 3.8$$

$$\bar{U}I = \frac{1}{N} * \sum_{i=1}^{n} Ui \qquad 3.9$$

$$S = \frac{1}{N-1} * \sum_{i=1}^{n}(Ui-\bar{U}i)\,(Ui-\bar{U}i)^{T} \qquad 3.10$$

Where, N -is the total number of stations, Di -discordancy measure

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

$\bar{U}i$ -is the group averages Ui and S -sample covariance matrix of Ui.

At a significance level of 10%, Hosking and Wallis (1997) tabulated critical values of the discordancy statistic Di for varying numbers of sites in an area. These were used to evaluate each of the study locations and determine whether they needed to be further investigated to assure homogeneity. Equation 3.8 was used to check for discordancy in the selected locations. However, calculating Di's value using simple matrix multiplication was complicated and time-consuming.

As a result, (Hosking and Wallis (997) advocated for the use of Fortran, Matlab, and other computer tools to streamline the process and achieve acceptable accuracy. Following this guideline, the Matlab2018a programming code was used to ease the numerical calculations of the discordancy index in this work (Di). Appendix-F contains the programming code for calculating the covariance matrix and Di.

Table 3. 5 Critical values of discordancy measure with N sites.

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

(Source: Hosking and Wallis, 1997)

3.5.2.2. Adjustment of regions

If the generated zones aren't statistically homogeneous, they're tweaked to make them more so. This step is necessary because, based on the homogeneity assessment, regions are unlikely to be homogeneous and discordant sites may exist. (Rao and Srinivas, 2008), provide the following strategies for amending regions that are significantly discordant with other sites in the region. I) removing one or more sites from the data set; ii) transferring (or moving) one or more sites from a region to another region; iii) dividing a region to form two or more new regions; iv) allowing a site to be shared by two or more regions; v) dissolving regions by transferring their sites to other regions; vi) merging two or more regions and redefining groups; and vii) obtaining more data and redefining regions. The first three options help in lowering the values of a region's heterogeneity measures; whereas options (iv) to (vii) assist in ensuring that each region is large enough.

3.5.2.3. Conventional homogeneity test

The value of CC was employed as a criterion for determining regional homogeneity. According to some studies, the larger the value of Cv and CC, the worse the index-flood approach will perform for the region in question. This is owing to the at-site sample mean-variance dominating the flood quantile estimation variance. As a result, CC should be kept low to improve the index flood method's performance. The processes for calculating CC values using this method are outlined below. For each site in the delineated regions; the mean $\overline{Q}$, standard deviation (σ) and coefficient of variation (Cv) were given and calculated by (Sine and Ayalew, 2004; Nobert *et al*., 2014; Guru and Jha, 2016)

The mean of AMF of the summation:

$$\text{Qi} = \frac{1}{N} * \sum_{i=1}^{n} Qi \quad \text{...3.11}$$

The standard deviation of AMF of the station; values, the procedures are described below. For each site in the delineated regions; the mean $\overline{Q}$, standard deviation (σ) and coefficient of variation (Cv) were given and calculated by Sine and (Ayale, 2004; Nobert *et al*., 2014; Guru and Jha, 2016) equation (3.12-3.16).

The standard deviation of AMF of the station;

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

$$\text{Cvi} = \frac{\delta\text{I}}{\text{Qi}} \quad \text{...3.13}$$

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

Qi =the mean flow rate for the region (m3 /s), at site i

δi = 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, Cvi, and finally the corresponding CC value using the following relation:

Regional mean; $c\bar{v}i = \frac{1}{N} * \sum_{i=1}^{n} vi$……………………………………………………3.14

Regional standard deviation, $\delta cv = \frac{\sqrt{\sum_{i=1}^{n}(cvi - c\bar{v}i)^2}}{n}$……………………………………3.15

The weighted regional Cvi of all the sites, Cc is defined as follows:

$CC = \frac{\delta Cv}{cvi} < 0.3$ ……………………………………………………………………………3.16

Where: N=number of the site in a region

Cvi = the mean coefficient of at site Cvi values

δCv = Standard deviation of at site Cvi values

3.5.2.4. L-moment based homogeneity test

LCV-based homogeneity test is a 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 an advantage of LCv (Cunnane, 1989), Compared to Cv, LCv can characterize a wide range of distribution, sample estimates are so strong that they are not affected by the presence of outliers in the data set, they are less matter to bias in estimation, yields more accurate estimate of the parameter of a fitted distribution.

According to the (Central Water Commission 2010), L-moments have the following advantages: i). Characterize most of the probability distributions than conventional moments, ii). Less sensitive to outliers in the data, iii). Approximate their asymptotic normal distribution more closely, iv). Nearly unbiased for all combinations of sample sizes and populations. (Hosking and Wallis, 1993) gave the unbiased estimators of β0, β1, β2and β3as: defined as

$\beta_0 = \frac{1}{N} * \sum_{i=1}^{n} Qi$……………………………………………………………………3.17

$$\beta 1 = \frac{\sqrt{\sum_{i=1}^{n} (j-1)(Qi)}}{n(n-1)} \qquad (3.18)$$

$$\beta 2 = \frac{\sqrt{\sum_{i=1}^{n} (j-)(j-2)(Qi)}}{n(n-1)(n-2)} \qquad (3.19)$$

$$\beta 3 = \frac{\sqrt{\sum_{i=1}^{n} (j-1)(j-2)(j-3)(Qi)}}{n(n-1)(n-2)(n-3)} \qquad (3.20)$$

Where Qi - annual maximum flow (m3 /s) from stations dataset

n - the number of years, j-rank

βo, β1, β2, and β3- are L-moments estimator.

The first few moments are:

$\lambda_1 = \beta_O$; $\lambda_2 = 2\beta_1\ \beta_O$; $\lambda_3 = 6\beta_2\ 6\beta_1 + \beta_O$; $\lambda 4 = 20\beta_3\ 30\beta_2 + 12\beta_1\ \beta_O$ ……………………………3.21

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

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

$$\tau_{2=} \frac{\lambda 2}{\lambda 1} = \tau_3\ \frac{\lambda 3}{\lambda 2} = \tau_{4=} \frac{\lambda 4}{\lambda 2} \qquad (3.22)$$

Using the above procedural formula,

$$Lc\bar{v}I = \frac{1}{N} * \sum_{i=1}^{n} Lcvi \qquad (3.23)$$

$$\Delta cv = \frac{\sqrt{\sum_{=1}^{n} (Lcvi - Lc\bar{v}i)^2}}{n-1} \qquad (3.24)$$

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

$$CC = \frac{\delta lCv}{cvi} < 0.3 \qquad (3.25)$$

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

3.5.3. Delineation of Homogeneous Regions

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

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

According to Abdulla (2011) and Irwin et al. (2014), the procedures in the delineation of the boundary of the region are as follows: i). Compute the (LCs, LCk) value of each station, ii). Identify the location of stations along with the distributions of LMRD for the defined regions statistical comparison of observed flood data, iii). Identify the group based on steps iv). Each region that was identified in step-i was checked for statistical homogeneity using the proposed test. Finally, the drainage boundaries of each sub-region the delineation was carried out using ArcGIS10.4.1 environment

3.6. Selection of Regional Frequency Distribution.

The Selection of regional frequency distribution is one of the important elements of the regional flood frequency analysis (RFFA). The Presence of adequate hydrometric stations is essential in each of the hydrologic regions for reliable selection of regional frequency distributions. The choice of frequency distributions is determined based on goodness-of-fit measures, which indicate how much the considered distributions fit the available data

(Hailegeorgis and Alfredsen, 2017); (Mishra *et al*., 2009). In flood frequency analysis, the annual maximum flow corresponding to a given T can be estimated from the annual flood series using various theoretical distributions.

3.6.1. The Cullen and Frey graph

The Cullen and Frey diagram is a graphical tool that may be used to determine the goodness of fit of a distribution. It's a graph that compares the fit of multiple distributions on the same graph and indicates possible model distribution candidates using the L-skewness and L-kurtosis. The Cullen and Frey diagram can be used to help you choose the right distribution. The distribution with theoretical values that are visually similar to sample values can be deemed the best appropriate PDF for accurately representing sample data. To choose a model that reduces uncertainty, several accepted design approaches are generally required. Among the distributions used in this study are the generalized extreme value (GEV), generalized logistic (GLO), Generalized Pareto (GPA), Logistic, Log-Normal (LN), Log-Pearson type 3 (LPIII), and Normal, Gamma, Weibull, Cauchy and beta distributions are among the employed distributions in this study. A family of continuous probability distributions is known as the GEV distribution. GEV relies on three variables: location, scale, and shape. The location parameter describes the shift of a distribution on the horizontal axis in a certain direction. The scale parameter determines how spread out the distribution is and where the majority of it is located. The distribution will get more spread out if the scale parameter is increased. The shape parameter, which determines the shape of the distribution and governs the tail of each distribution, is the third parameter in the GEV family.

The shape parameter is obtained from skewness since it represents where the majority of the data is located, resulting in the distribution's tail(s). When quantifying AM series river flow, the GEV is perhaps the most often employed distribution.

3.6.2. R studio Software for Distribution Fitting

R Studio is an Integrated Development Environment (IDE) for R, a programming language for statistical computing and graphics. It is available in two formats: R Studio Desktop is a regular desktop application while R Studio Server runs on a remote server and allows accessing R Studio using a web browser. It includes a console, syntax-highlighting editor that supports direct code execution, as well as tools for plotting, history, and debugging, and workspace management. R Studio is a data analyzer and simulation software that allows to fit probabilistic distributions to given data samples, simulate them, choose the best fitting sample, and implement the results of the analysis to make better decisions. To determine whether the distribution model could fit the data properly, goodness-of-fit tests were used. In this study R Studio is (version 4.1.1.2021) Statistical Software was used to find the best-fit distribution and its estimation parameters (Pakgohar, 2014).

3.6.3. Goodness of Fit Tests

The goodness fit measure involves identifying a distribution that fits the observed data. When computing the magnitudes of extreme events, such as flood flows, it is required that the probability distribution function be invertible, so that a given value of recurrence interval (T) and the corresponding value of frequency factor (K) can be determined. In this study, to test the statistical hypothesis of whether a particular distribution provides an adequate fit to the observed AMF series data three goodness of fit tests were applied. The reason for selecting three different tests is that there is no single test that can give conclusive results and a particular test emphasizes a particular aspect of the goodness-of-fit. All test statistics were defined and carried out at a 5% significance level (Ashraful *et al.,* 2018).

i. Kolmogorov-Smirnov Test (KS)

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

$$DN = \max_{1 \leq i \leq n}\left(\mathrm{Fn}\ (\mathrm{xi}\,)\quad \mathrm{FO}(\mathrm{xi}\,)\right) \quad \ldots\ldots 2.26$$

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

ii. Chi-Squared Test ($\boldsymbol{X}^2$)

The chi-square goodness of fit test is one of the most commonly used tests for testing the goodness of fit of probability distribution functions to empirical frequency distribution. In the X^2 goodness of fit test, sample data is separated into intervals. Then the numbers of points that drop into the interval are compared, with the predictable numbers of points in every interval. The null hypothesis assumes that there is no notable variation between the observed and the expected value. The degree of freedom depends on the distribution of the data sample (Ghosh *et al*., 2016). In the Chi-Square goodness of fit test, the alternative hypothesis assumes that there is an essential variation between the observed and the expected value.

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

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

O = observed value

E = expected value

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

iii. Anderson-Darling Test (AD)

The AD test was used to check whether the given sample came from a particular probability distribution at hand. The null hypothesis at the chosen level of significance would be rejected if the calculated value of the above statistic exceeds the critical value given in the table (Onoz and Bayazit, 2001; Ahmad *et al*, 2015). AD test can be used in RFFA studies to assess the fitness of the candidate regional frequency distributions. This method is based on the statistical frequency distribution behavior of the observed value (Viglione *et al*, 2007).

3.6.4. Evaluation of the Performance of Frequency Distributions

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

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

i. Probability-Probability (P-P) Plots

Probability plots are generally used to decide whether the distribution of a variable matches a given distribution. P-P Plots are the variable's cumulative magnitude in opposition to the cumulative magnitude of any of several trial distributions. If the selected variable matches the test distribution, the points come together approximately a straight line. The following fundamental issues should arise when selecting a distribution: (1). It is true and reliable with the distribution for which the observations are drawn, (2). It should be used to obtain reasonably perfect and strong estimations of design quantiles and hydrologic risk (Desalegn *et al*., 2016).

ii. Quantile-Quantile (Q-Q) Plots

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

3.6.5 Selection of best fit probability distribution approach to a data set

The analysis involves the estimation of the parameter values' distribution or distributions selected and the identification of the best distribution model suited for the estimation of extreme hydrological events. There are many methods for the selection of best-fit probability distribution to a data set to the estimation of time-dependent hydro-climatic flood parameters. This research only discusses the most common and major ones.

3.6.5.1 The method of moments (MOM)

The method of moments estimates population parameters by taking known facts about the population sample as the first-moment condition and extending the same concepts to derive higher moments. Moments such as the skewness (s), coefficient of variations (σ 2), kurtosis (k), expected moment (μ), and the parameters (θ) are presumed to be related to a distribution function—ó = g (μ, α 2, s, k)—and are considered members of the underlying distribution (û, s, ˆ k, ˇ ᾷ) for providing parameters ô = g (û, s, ˆ k, ˇ ᾷ) (21). This method has the advantages of being simple to derive, consistent in providing estimators for continuous function, and providing starting estimates in search of maximum likelihood values. However, the estimators may not be unique in a given dataset, and thus can provide multiple solutions to a set of equations; furthermore, sometimes parameter estimates may suffer from inaccurate and insufficient statistics, especially for smaller population sizes.

Moment matching estimation consists of equalizing theoretical and empirical moments. Estimated values of the distribution parameters are computed in R studio software by a closed-form formula for the following distributions: "norm", "lnorm", "pois", "exp", "gamma", "nbinom", "geom", "beta", "unif" and "logis". Otherwise, the theoretical and the empirical moments are matched numerically, by minimization of the sum of squared differences between observed and theoretical moments. In this last case, further arguments are needed in the call to fitdist.

3.6.5.2 Quantile matching estimation

Quantile matching estimation consists of equalizing theoretical and empirical quantile. Numerical optimization is carried out in qmedist via optim to minimize the sum of squared differences between observed and theoretical quantiles. The use of this method requires additional argument probs, defined as the numeric vector of the probabilities for which the quantile(s) is(are) to be matched.

3.6.5.3 Maximum goodness-of-fit estimation

Maximum goodness-of-fit estimation consists in maximizing goodness-of-fit statistics. Numerical optimization is carried out in mgedist via optim to minimize the goodness-of-fit distance. The use of this method requires an additional argument GoF coding for the goodness-of-fit distance chosen. One can use the classical Cramer-von Mises distance ("CvM"), the classical Kolmogorov-Smirnov distance ("KS"), the classical Anderson-Darling distance ("AD") which gives more weight to the tails of the distribution, or one of the variants of this last distance proposed by Luceno (2006). This method is not suitable for discrete distributions.

3.6.5.4 Maximum Likelihood Estimation

The MLE method is an approach that is used to determine values for the parameters of a model, the maximum likelihood Estimation Theory. The MLE method is an approach that is used to determine values for the parameters of a model, which are calculated in such a way that they maximize the likelihood that the model process description produces the data that were observed. In an observed sample series, the probability of any random variable to occur can be obtained by the multiplication of the probability density functions of each observed data of that series by each other with the assumption that the events of the random variable are independent of each other, which results in what is known as the likelihood function (LF). The parameter values that give the maximum likelihood function among so many other possible sample series of the population are considered the most suitable ones for that sample series. It is analytically convenient to use the derivative of the logarithm of the likelihood function (LLF) (summation of logarithms of the probability density function (PDF)), which is also called the cumulative distribution function (CDF), because the maximum values of the likelihood function and the logarithm of the likelihood function result in the same magnitudes of the distribution parameters. In this study, the selection of the best-fit probability distribution to the data set was done by using a programing language of R studio software.

The method used for the selection of the best-fit probability distribution was R Studio (2014 version), which is an integrated development environment for the R programming language. Flood flow extreme value analysis was done by fitting a parametric distribution to the extremal data using fitdistrplus, which is the freely available packages. It of univariate distributions to non-censored data by maximum likelihood (MLE), moment matching (mme), quantile matching (qme), or maximizing goodness-of-fit estimation (mge). The latter is also known as minimizing distance estimation.

A character string "name" naming a distribution for which the corresponding density function dname, the corresponding distribution function pname, and the corresponding quantile function qname must be defined, or directly the density function. A character string coding for the fitting method: "mle" for 'maximum likelihood estimation, "mme" for 'moment matching estimation', "qme" for 'quantile matching estimation', "mge" for 'maximum goodness-of-fit estimation, and "mse" for 'maximum spacing estimation. A named list gives the initial values of parameters of the named distribution or a function of data computing initial values and returning a named list.

In this study, the best-fit probability distribution approach for the data set was selected by checking the likelihood values (LL), the Akaike information criteria values (AIC), and the Bayesian information criteria (BIC) values for several distributions in each method. likelihood values show the strength or accuracy of the model and the AIC and BIC values shows the measure of the error of the model.

3.6.6. Parameter and Quantile Estimation

In flood frequency analysis, the probability distribution is fitted to the available data to estimate the flood magnitude for a specified return period. The choice of an appropriate probability distribution is quite arbitrary, as no physical basis is available to rationalize the use of any particular distribution (Saf, 2009; Rao and Hamed, 2000). In the present study, the parameter estimation was done by using the R studio statistical software. Based on the selected distributions for each station, the quantile can be calculated according to the formula of the selected distributions. For regions with a computed value of scale, location, and shape parameter, then it is possible to determine the quantile with different return periods using different equations for different distributions.

For GEV distribution the flow quantile can be estimated as;

$XT = \mu + \sigma/K\ (1 - (-\ln (1 - 1/T)\ ^{K})$, fork≠0 ………………………………….………….....3.28

$XT = \mu + \sigma\ (\ln (-\ln (1 - 1/T))$, for k=0 ………………………………………….………......3.29

For GPA distribution the flow quantile can be estimated as;

$XT = \mu + \sigma\ K\ (1 - (1/T)\ ^{K})$, fork≠0……………………..……………………….…….…….3.30

$XT= \mu +\sigma\ (\ln (1/T))$, for k=0 ………………………….……........................…………...3.31

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

3.6.7. Standard Error of Parameter Estimation

The standard error of the estimate is a metric for determining how accurate predictions are. The creation of a link between the mean annual flood or index flood and watershed parameters was an essential step in predicting flood magnitudes at any point in a region where the frequency curve had been developed and inaccuracy in quantitative terms was quantified. Various measurements of error are used by different researchers. Standard errors are the most frequent metric. Only sampling error may be evaluated from the numerous sources of error. Theoretically, a consensus appears to be forming that at the very least, sampling error should be stated quantitatively. In most cases, errors in flood frequency estimations should be indicated numerically or graphically. In most cases, the standard error of a particular quantile attributable to sampling error should be calculated (Rao and Hammed, 2000).

$$\bar{Q} = \frac{\sum \bar{Q}T}{QT} \qquad \text{3.32}$$

$$SEE= [\frac{\sum(\bar{Q}-QT)\ 2}{\bar{Q}T}]^{0.5} \qquad \text{3.33}$$

Where; SEE – standard error of estimate

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

$\bar{Q}$ - is the mean annual flood (m3 /s) is the index flood

QT - is the quantile (m3 /s) function of fitted distribution at site i

3.7. Derivation of the Regional Flood Frequency Curves

The basic purpose of any regional flood frequency analysis is to create a regional frequency curve that can represent the average weighted distribution of homogeneous regions. For a given return period, T is the final step of flood frequency analysis to determine normalized regional quantile floods (XT); flood frequency curve (XT vs. T); and at-site flood quantiles. The model parameters produced from the best fitting distribution to the observed data are the most important for a given region. Because these values are used to produce standardized quantile estimates, XT, for the return times T, which are then utilized to create regional frequency curves for the homogeneous region. (a curve showing XT against return period, T) (Kachroo *et al*., 2000; Mkhandi *et al*., 2000; Rosbjerg, 2007; Yang *et al*., 2010).

3.7.1. Estimation of Index-Flood

The index flood approach is based on the premise that data from different locations in a region follow the same distribution. It involves identifying homogenous regions, determining the best-fit distribution, and deriving the regional flood frequency curve. The index flood L-moment strategy of regionalization is used in this study, which is based on the homogeneity of the stations, which is tested using various techniques. The mean annual flood (Q) for each station was calculated by multiplying the annual flood data (Qi) from each station by the number of record years.

In flood frequency analysis, the estimated flood quantiles corresponded to the needed return periods. The standardized flow estimates for the return periods 2, 5, 10, 15, 20, 25, 50, 75, 100, 200, 500, and 1000 years were computed using the model parameters for the distributions determined for each station. For each station, growth curves (Q/Qm plotted against the Gumble reduced variate (-ln(-ln(1-1/T)) were generated and used in the construction of the regional growth curves. The stages below were used to accomplish this. The model parameters calculated for a given region were utilized to construct the standardized quintiles estimates for the return periods, the growth curves for each station were generated, and the parameter values such as shape (k), location (), and scale () were estimated using R studio software. Following the determination of the regional frequency distribution, the flood quantiles with a return period of T year within a homogeneous region can be approximated using Hosking and Wallis' equation (3.34). (1997). XT was estimated using dimensionless regional growth curves. The dimensionless data is often obtained by dividing the values by an estimate of the at-site mean.

$$XT = \frac{QT}{\bar{Q}} \quad \text{......3.34}$$

Where; $\bar{Q}$ - is the mean annual flood (m3 /s) is the index flood

QT - is the quantile (m3 /s) function of fitted distribution at site i

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

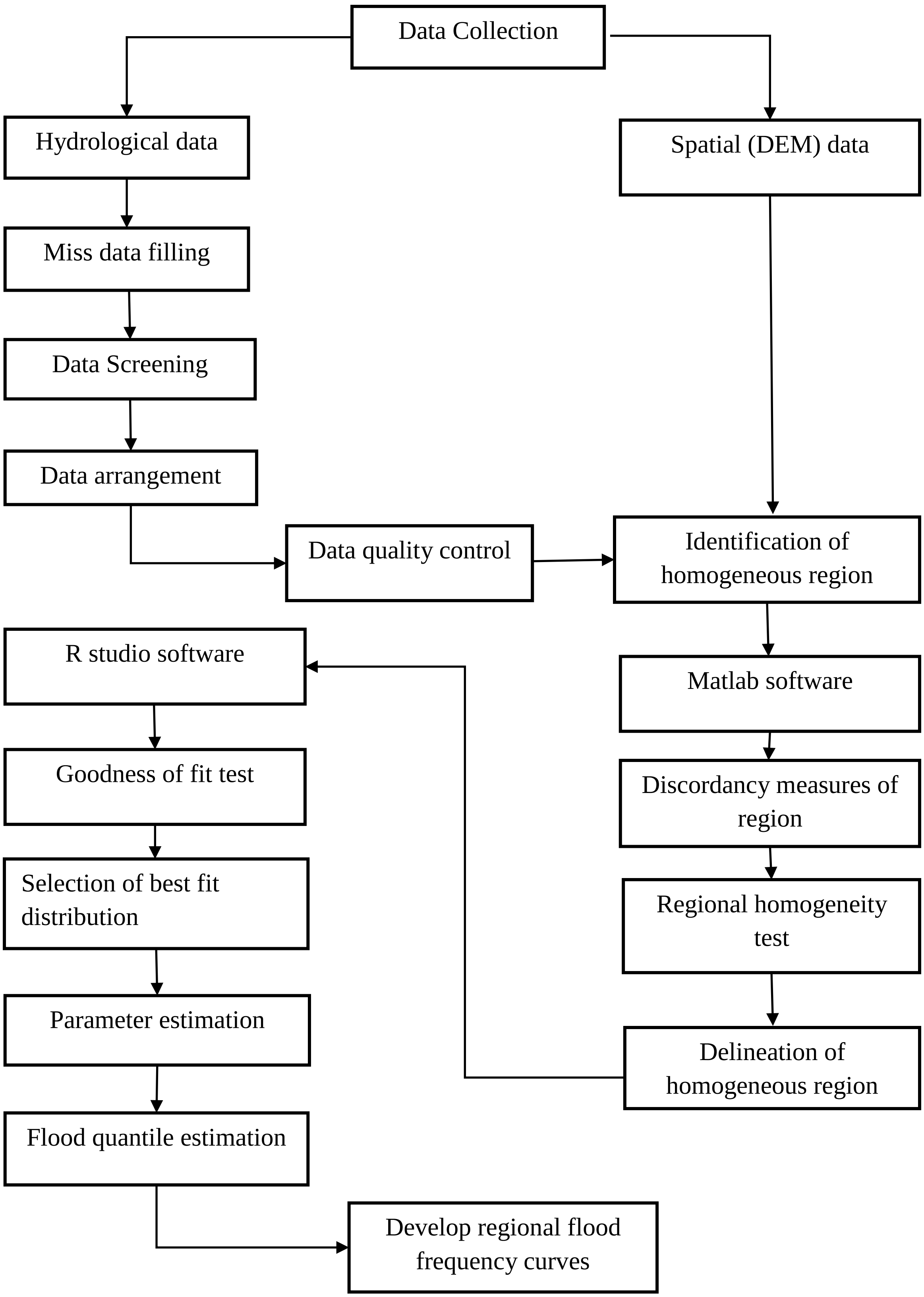


Figure 3. 3 Flow chart of the methodology

4. RESULT AND DISCUSSION

4.1 Identification of Homogeneous Region

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

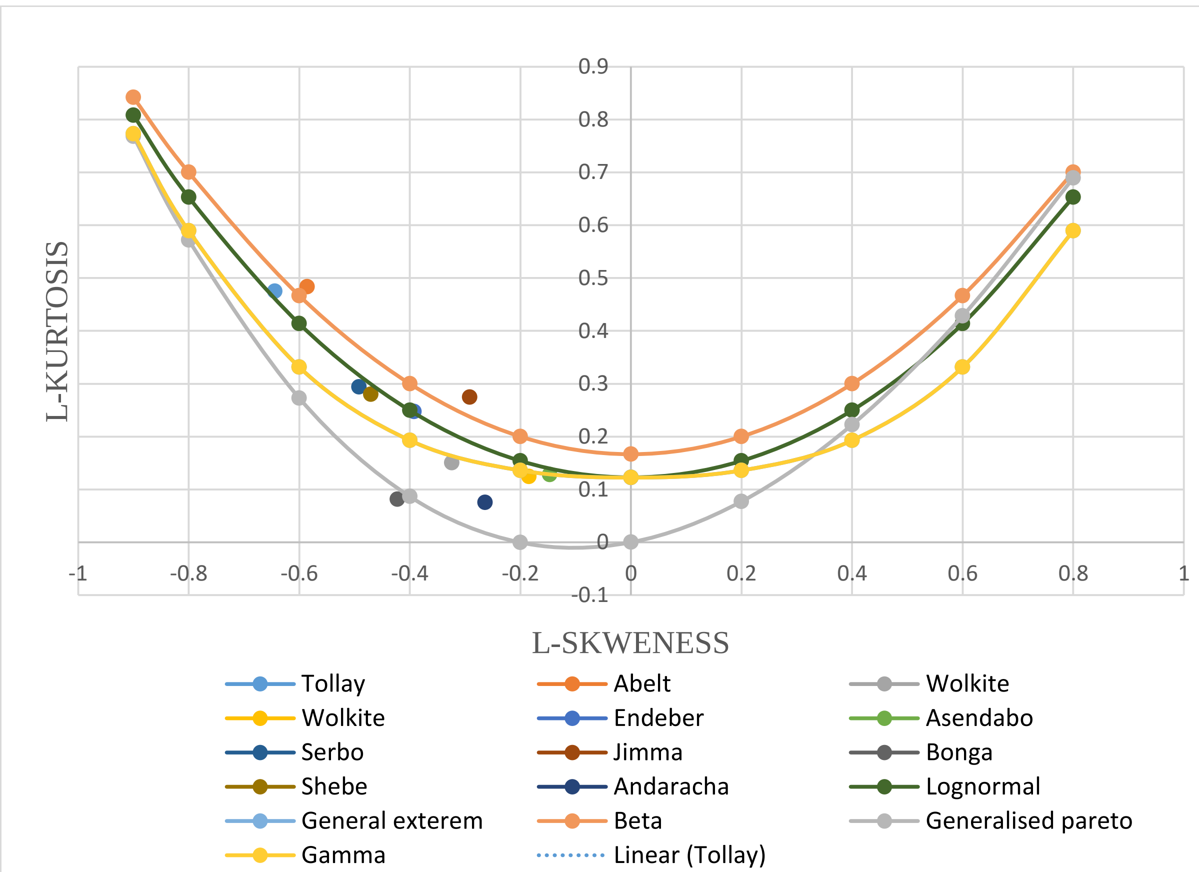


Figure 4. 1: L-moment ratio diagram for identification of homogeneous regions

The LMRD shown on Figure 4.1 was used to identify homogeneous regions with site characteristics of gauging stations described in Table 3.2. As indicated in Table 4.1, the accentuated distributions were designated to the same group since stations lie close to the identical distribution. Hence, based on L-moment statistics and suitability of gauging site networks, for further homogeneity tests three Preliminary identified homogeneous sub regions were identified. Namely Region-A, Region-B and Region-C as shown in Table 4.1.

Table 4. 1: Preliminary identified homogeneous regions

Group name	Station name	Possible distributions from R-Studio
Region-A	Tollay	Weibull
	Abelt	Gamma
	Wolkite	Lognormal
	Wolkite	Beta
	Endeber	Lognormal
Region-B	Asendabo	Lognormal
	Serbo	beta
	Jimma	Lognormal
Region-C	Bonga	Lognormal
	Shebe	Gamma
	Andaracha	Gamma

Table 4. 2: Classical descriptive statistics for the primarily delineated homogeneous region located Upper Omo river basin

Statistic	Extreme Streamflow Datasets		
	Region-A	Region-B	Region-C
Minimum (m3 /s)	235.4	51.75	101.9
Maximum (m3 /s)	779.8	269.54	295.3
Median (m3 /s)	434.5	96.30	158.3
Mean (m3 /s)	449.3	123.15	169.2
Estimated standard deviation (m3 /s)	118.45	64.61	41.805
Estimated skewness	0.982	0.966	1.1373
Estimated kurtosis	1.5889	-0.393	2.0266
Coefficient of variance	0.2657	0.53221	0.2603
1st Quartile	370.5	73.72	131.0
3rd Quartile	504.6	175.37	169.24

4.1.1 Test for Regional Homogeneity

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

4.1.1.1 Discordancy Measure of Regions

Hosking and Wallis (1993) proposed the discordancy measure (Di) to distinguish odd sites from other sites in a group by comparing their L-moment ratios. Values of discordancy of L-moment statistics have been calculated for all the eleven gauging sites of the basin. Using Equation (3.8) with Matlab program code presented in Appendix-F, the values of discordance index (Di) measure for different sites within the regions were presented in Table 4.3, 4.4, and 4.5 for Region-A, Region-B, and Region-C respectively. The critical values of the discordancy index Di for various numbers of sites in a region at a significance level of 10% were obtained from Table 3.5. It was observed that the Di values for all eleven sites vary from 0.6682 to 1.4216.

According to (Hussen and Wagesho, 2016; Kanti et al., 2017; Lim, 2007; Nobert *et al.*, 2014), and the region on their study under investigation, has been declared homogeneous if Di is less than 3. If Di is large, a place is termed exceptional in this scenario. This would be considered as grossly discordant and would justify elimination from the defined regions and can be redefined as a single site or merged into other regions. Hence, all of the stations grouped as homogeneous in Region-A, Region-B, and Region-C were satisfied the discordance test criteria. As shown in Table 4.3, 4.4, and 4.5, the result of the entire Di was below the critical value which implies that all the regions are homogeneous. So, none of the identified regions was found to reveal Di greater than the critical value. This indicated that all sites do not reflect any outlier and discordancy. Thus, data of all gauging sites could be considered for further regional flood frequency analysis.

Table 4. 3 Results of major statistics and discordant measures in Region-A

Station Name	LCv	LCs	LCk	Di	Remark
Tollay	0.79097	-0.6442	0.47234	0.6770	Homogeneous
Abelt	-0.0595	-0.5855	0.4833	0.6682	Homogeneous
Wolkite	-0.3196	-0.3239	0.1506	1.3216	Homogeneous
Wolkite	-0.3571	-0.1845	0.1244	0.8565	Homogeneous
Endeber	-0.2662	-0.3923	0.3671	0.6961	Homogeneous

Table 4. 4 Results of major statistics and discordant measures in Region-B

Station Name	LCv	LCs	LCk	Di	Remark
Asendabo	-0.1714	-0.1469	0.2983	0.9999	Homogeneous
Serbo	-0.2708	-0.4920	0.2938	0.9999	Homogeneous
Jimma	-0.1792	-0.2914	0.2743	0.9999	Homogeneous

Table 4. 5 Results of major statistics and discordant measures in Region-C

Station Name	LCv	LCs	LCk	Di	Remark
Shebe	-0.1877	-0.4705	0.2798	0.9999	Homogeneous
Bonga	-0.4971	-0.6223	0.0817	0.9999	Homogeneous
Andaracha	-0.1877	-0.2637	0.0755	0.9999	Homogeneous

4.1.1.2. CC-based Regional Homogeneity Test

In this test, the site-to-site coefficient of variation of the coefficient of variation (CC) of both conventional and L-moments of the proposed regions are used. The (L-Cs, L-Ck) of standardized flow values at each station have been plotted on the LMRD together with various theoretical distribution functions. Those stations close to a particular theoretical distribution linear considered to be homogeneous stations and grouped together. The combined coefficients of variation for the region (CC) values were calculated, and the results in sites of each region were summarized as shown in Table 4.6, 4.7, and 4.8.

The value of CC varies from region to region depending on L-moment statistics of flow data. From Cv-based homogeneity test, the CC values were 0.1806, 0.0837and 0.11231for Region-A, Region-B and Region-C respectively. On the other way, from the LCv-based homogeneity test, the CC values were 0.2358, -0.2179, and -0.5015 for Region-A, Region-B, and Region-C respectively. According to (Sine and Ayalew, 2004; Guru and Jha, 2016; Nobert et al. 2014) noted that for the study regions under their consideration, a region is recognized to be homogeneous if CC values were less than 0.3. Therefore, from the results in Table 4.6, 4.7, and 4.8, it can be concluded that all regions were hydrologically homogeneous for both Cv and LCv based homogeneity tests since the CC values were less than 0.3. The results obtained below; all stations grouped as homogeneous were satisfied the stated homogeneity test criteria. As a result, it can be concluded that all regions were reasonably homogeneous.

Table 4. 6 Results of Cv and LCv-based homogeneity test for Region-A

Station Name	LCv	LCs	LCk	Cv	Cs	Ck
Tollay	0.79097	-0.6442	0.47234	0.1733	-1.3832	1.7213
Abelt	0.6595	-0.5855	0.4833	0.6148	-0.5431	8.6065
Wolkite	0.3619	-0.3239	0.4506	0.6263	1.6885	3.6899
Wolkite	0.6571	-0.4845	0.3244	0.6495	0.5649	-0.9714
Endeber	0.7662	-0.3923	0.3671	0.3053	1.9776	4.2283
Mean	0.6471	-0.4861	0.41955	0.4128	0.4609	3.4549
Std.dev	0.1526	0.11831	0.06264	0.0745		
CC	0.2358			0.1806		

Table 4. 7 Results of Cv and LCv-based homogeneity test for Region-B

Station Name	LCv	LCs	LCk	Cv	Cs	Ck
Asendabo	-0.1714	-0.1469	0.2983	0.5562	0.8269	0.772
Serbo	-0.2708	-0.492	0.2938	0.4875	1.4905	0.6580
Jimma	-0.1792	-0.2914	0.2743	0.5988	1.0511	0.4259
Mean	-0.2071	-0.3101	0.2888	0.5475	1.1228	0.6186
Std.dev	0.0451	0.1415	0.0104	0.0458	0.2756	0.1440
CC	-0.2179			0.0837		

Table 4. 8 Results of Cv and LCv-based homogeneity test for Region-C

Station Name	LCv	LCs	LCk	Cv	Cs	Ck
Shebe	-0.1877	-0.4705	0.2798	0.6528	0.7519	1.9502
Bonga	-0.4971	-0.3223	0.2817	0.8593	1.0528	1.3838
Andaracha	-0.1877	-0.2637	0.1755	0.7965	0.7373	2.6724
Mean	-0.2908	-0.3521	0.2456	0.76953	0.8473	2.0021
Std.dev	0.1458	0.0870	0.0496	0.08643	0.1454	0.5273
CC	-0.5015			0.11231		

4.2 Delineation of Homogeneous Regions

The delineation of homogeneous regions is highly related to the identification of the similar regional distributions that apply within each region. The preliminarily identified regions have to be checked by various homogeneity tests. The tests used in this study are dispersion-based tests (Cv-based homogeneity test and L-Cv-based homogeneity test) and statistical comparison. The regions have covered an area of 22,085, 3,577, and 3,983.3 km^2, for Region-A, Region-B, and Region-C respectively.

Accordingly, the first region which includes Five gauging stations in the Upper Omo-Gibe sub-river basins including Tollay, Abelt, Endeber, and two gauging stations at Wolkite were delineated under Region-A. The second region, which includes the gauging stations in Asendabo, Serbo, and Jimma stations were delineated under Region-B. The third region, which is the rest of the gauging stations selected for this study including Shebe, Bonga, and Andaracha stations were delineated under Region-C. Having proven to be statistically homogeneous, the delineated homogenous regions shown in Figure 4.2 could be used to generate a regional growth curve at any site located in the study area.

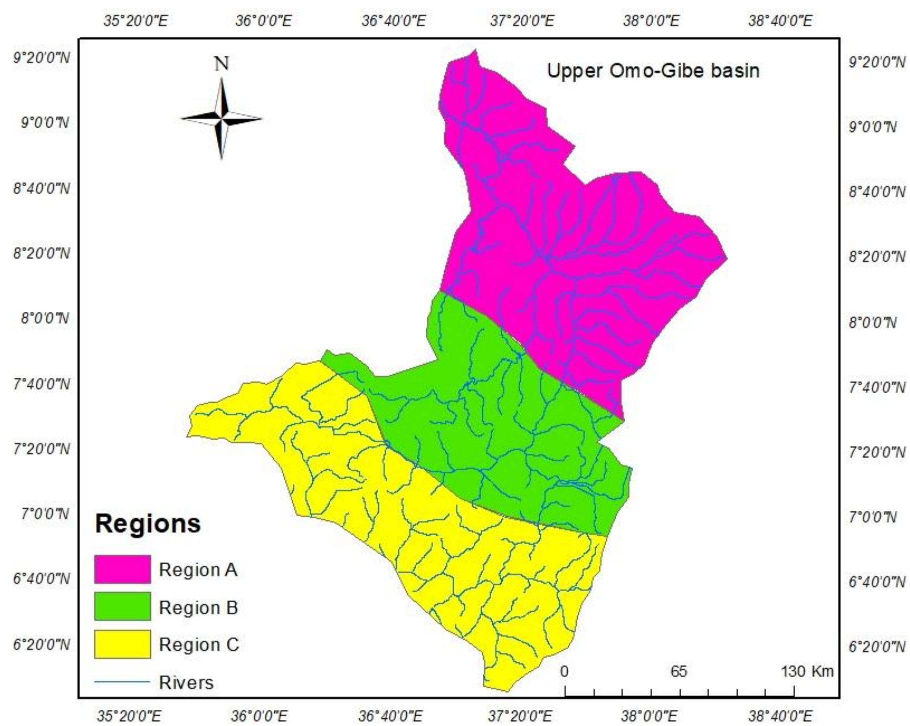


Figure. 2: Spatial distribution of delineated homogeneous regions

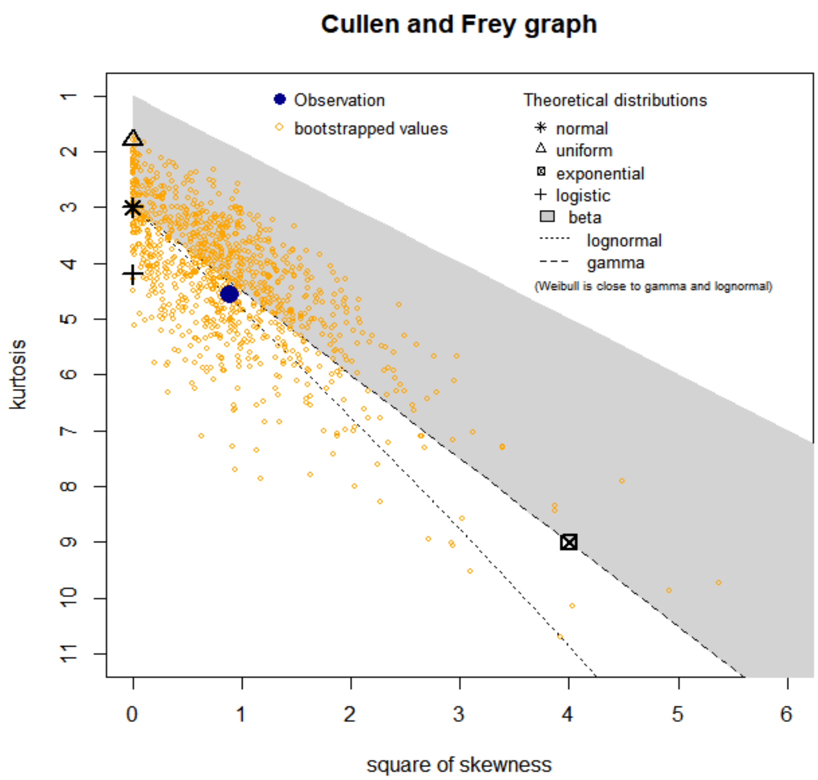
4.3 Determination of Suitable Regional Probability Distribution

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

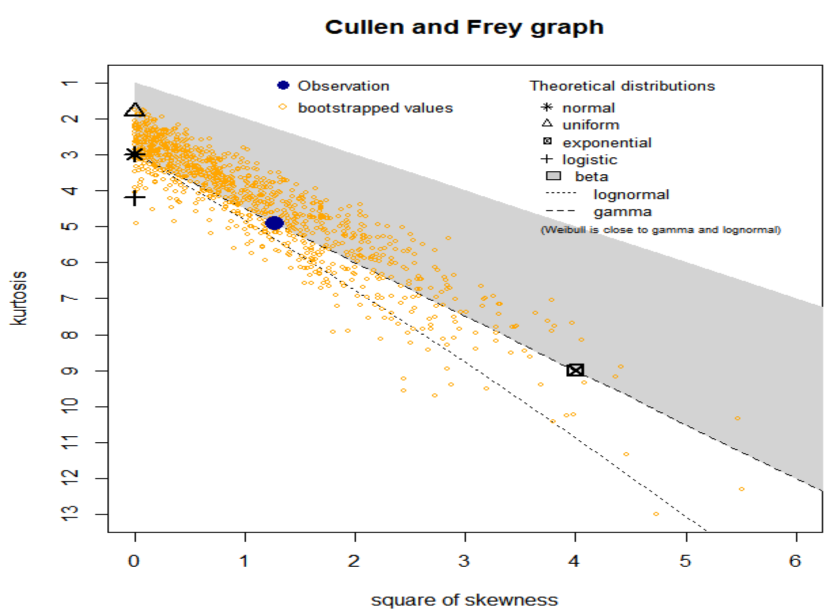
4.4 Preliminary Assessment and Visualization

The best-fit probability distribution was calculated using R-Studio (Version4.1.1, 2021), an integrated development environment for the R programming language. Flood flow extreme value analysis was performed by fitting a parametric distribution to the extreme data with fitdistrplus, freely available packages. The package uses the maximum likelihood estimation, moment matching (Flunn, *et al*., 2006), quantile matching (QME), and maximum goodness-of-fit estimation (MGE) methods.

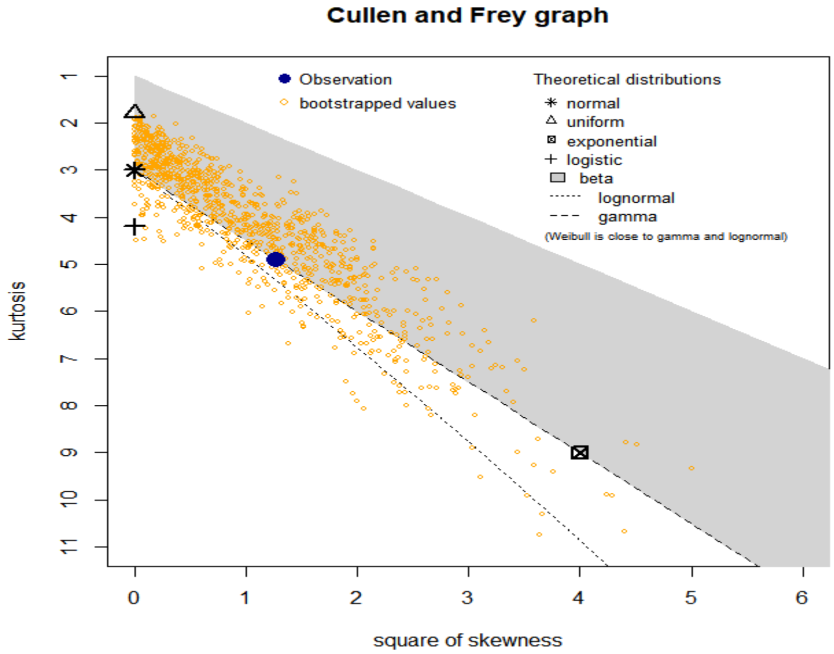
There are different methods for selecting an appropriate regional frequency distribution that describes the features of the sample data. The Cullen and Frey graph show a possible model distribution candidate for a given data set. Figure 4.3 illustrates an initial skewness-kurtosis graph of the unbiased distribution of the extreme data to aid in visualization and model selection. For the skewness, and kurtosis, uniform, exponential, normal, and logistic models have only one potential distribution value, whereas the possible lognormal, Weibull and gamma areas are depicted bylines, and the possible Beta areas are represented by larger areas. In the Cullen and Frey graph, the kurtosis and squared skewness of extreme datasets are represented as a blue point denoting "observation." The empirical distribution has zero skewness, indicating that it has symmetry, but the kurtosis measures the weight of tails in comparison to the normal distribution. A kurtosis value of three indicates normal distribution. From figure 4.3, lognormal, gamma, normal, and Weibull being common right-skewed distributions, are indicated as possible model distribution candidates because of the positive skewness and a kurtosis value that is close to three. However, the skewness and kurtosis exhibited high variations for all the distributions, and therefore, this inference can only be taken as indicative



(a) Region-A



(b) Region-B



(c) Region-C

Figure 4. 3: Description of regional streamflow samples from a normal distribution with uncertainty on skewness and kurtosis estimated by bootstrap.

The MLE approach optimizes the scale and shape parameters and offers parameter estimates for the fitted distribution. The standard errors were determined using the Hessian matrix estimate and the correlation matrix between parameter estimations from the maximum likelihood simulation solution. The standard error (Std error) reflects the consistency of a sample dataset's mean, and that the sample mean is a more accurate representation of the actual dataset. A small standard error indicates that the sample dataset's mean is reasonably close to the true mean of the entire dataset. Table 4.9, Table 4.10, Table 4.11 and Appendix-G show the probability plots simulated related with the parameter estimations.

Table 4.9 Distribution Sample Estimates Shape and Scale values estimated using the Moment Matching Estimation (MME) Method

		Region-A	Region-B	Region-C
Distribution	parameter	Sample Est	Sample Est	Sample Est
Normal	Mean	0.41	0.123	1.862
	Sd	0.117	0.0637	0.256
Lognormal	mean log	0.833	0.832	1.862
	sg log	0.255	0.255	0.256
Gamma	Shape	14.799	14.799	14.745
	Rate	32.94	32.94	91.860
Beta	Shape	7.7	7.701	12.217
	Scale	9.44	9.43	63.897
Cauchy	Shape	-	-	-
	Scale	-	-	-
Weibull	Shape	-	-	-
	Scale	-	-	-
Exponential	Shape	-	-	-
	Scale	2.226	2.225	6.229
Uniform	Min	0.246	0.012	0.088
	Max	0.651	0.233	0.232
Logistic	Shape	0.45	0.449	0.160
	Scale	0.064	0.064	0.02

(Est: Estimate)

Table 4. 10 Distribution Sample Estimates Shape and Scale values estimated using the Quantile Matching Estimation (QME) Method.

		Region-A	Region-B	Region-C
Distribution	parameter	Sample Est	Sample Est	Sample Est
Normal	mean	0.48	0.152	0.189
	sd	0.113	0.056	0.037
Lognormal	mean log	0.807	-2.174	1.84
	sg log	0.247	0.642	0.233
Gamma	Shape	16.379	2.712	18.755
	rate	35.794	20.37	113.568
Beta	Shape	8.916	2.424	15.755
	Scale	9.568	16.009	79.083
Cauchy	Shape	0.483	0.1245	0.170
	Scale	0.0295	0.0508	0.0098
Weibull	Shape	4.994	1.814	5.293
	Scale	-0.537	0.146	0.188
Exponential	Shape	-	-	-
	Scale	-	-	-
Uniform	min	0.276	0.022	0.100
	max	0.69	0.226	0.239
Logistic	Shape	0.483	0.1245	0.170
	Scale	0.063	0.0462	0.021

(Est: Estimate)

Table 4. 11 Distribution Sample Estimates Shape and Scale values estimated using the Maximizing Goodness of fit Estimation (MGE) Method.

		Region-A	Region-B	Region-C
Distribution	parameter	Sample Est	Sample Est	Sample Est
Normal	Mean	0.436	0.122	-1.866
	sd	0.061	0.066	0.252
Lognormal	mean log	0.833	2.237	1.866
	sg log	0.255	0.556	0.252
Gamma	Shape	19.8	3.759	15.65
	Rate	37.69	32.862	83.761
Beta	Shape	10.265	3.359	13.361
	Scale	13.158	26.214	70.874
Cauchy	Shape	0.436	0.0996	0.156
	Scale	0.0611	0.288	0.018
Weibull	Shape	4.27	1.869	4.164
	Scale	0.483	0.137	0.172
Exponential	Shape	1.652		4.623
	Scale	2.226		
Uniform	Min	0.246	0.0184	
	Max	0.651	0.217	
Logistic	Shape	0.45	0.11	0.157
	Scale	0.064	0.038	0.022

(Est: Estimate)

4.4.1 Selection of best fit probability distribution approach

In this study, the Maximum Likelihood, Moment Matching, Quantile Matching, and Maximizing Goodness of fit techniques or models were tested by considering nine different distributions, namely Normal, Lognormal, Gamma, Beta, Weibull, Cauchy, Exponential, Uniform, and Logistic). According to the Likelihood value (LL), Akaike information criterion (AIC), and Bayesian information criterion (BIC), the Maximum Likelihood technique provides a better parameter estimation procedure, since it provides greater LL values, and lower AIC and BIC values in all regions, Region-A, Region-B, and Region-C as shown in Table 4.12 up to Table 4.20. Because the LL value indicates the model's strength or correctness, while the AIC and BIC values indicate the model's measure of error. Now is the time the second and third best parameter estimation procedures are maximizing goodness of fit and moment matching, respectively, while the quantile matching estimation process is the last.

Table 4.12 Region-A Goodness-of-fit information criterion. Likelihood (LL).

		Maximum likelihood estimation (MLE) method.		Moment matching estimation (MME) method.		Quantile matching estimation (QME) method	
Distribution		Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	LL	23.310	4	23.318	4	21.81	4
Lognormal	LL	25.126	1	25.123	1	24.944	1
Gamma	LL	24.901	2	24.876	2	24.438	2
Beta	LL	23.208	5	23.201	5	21.914	5
Cauchy	LL	20.489	7	-	7	10.353	7
Weibull	LL	22.183	6	-	6	18.968	6
Exponential	LL	-6.396	8	-6.396	8	-	8
Uniform	LL	19.457	9	-Inf	9	-Inf	9
Logistic	LL	24.236	3	24.166	3	21.745	3

Table 4. 13 Region-A Maximizing Goodness-of-fit information (MGFE). LL: Likelihood Criterion

	Kolmogorov–Smirnov		Cramer–von Mises		Anderson–Darling	
Distribution	Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	22.414	4	22.316	4	23.071	4
Lognormal	25.035	1	24.884	1	25.089	1
Gamma	24.769	2	24.562	2	24.842	2
Beta	21.967	5	21.714	5	22.831	5
Cauchy	20.443	7	20.438	7	20.362	7
Weibull	16.100	6	16.973	6	21.479	6
Exponential	-8.328	8	-7.758	8	-7.684	8
Uniform	Inf	9	-Inf	9	Inf	9
Logistic	24.213	3	24.214	3	24.235	3

Table 4. 14 Region-A Goodness-of-fit information criterion. AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion

Distribution		Maximum likelihood estimation (MLE) Method.		Moment matching estimation (MME) Method.		Quantile matching estimation(MME) Method	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	AIC	-42.622	4	-42.62	3	-39.66	3
	BIC	-39.69		-39.69		-36.728	
Lognormal	AIC	-46.251	1	-46.24	1	-45.887	1
	BIC	-43.32		-43.31		-42.956	
Gamma	AIC	-45.802	2	-45.75	2	-44.876	2
	BIC	-42.87		-42.819		-41.945	
Beta	AIC	-42.415	5	-42.402	4	-39.82	4
	BIC	-39.484		-39.47		-36.898	
Cauchy	AIC	-36.979	7	-		-16.707	7
	BIC	-34.048		-		-13.775	
Weibull	AIC	-40.366	6	-		-33.936	6
	BIC	-37.434		-		-31.005	
Exponential	AIC	14.792	9	14.79	6	-	
	BIC	16.258		16.258		-	
Uniform	AIC	-34.915	8	Inf		Inf	
	BIC	-37.983		Inf		Inf	
Logistic	AIC	-44.437	3	-44.23	5	-39.49	5
	BIC	-41.54		-41.3		-36.558	

Table 4. 15 Region-B Goodness-of-fit information criteria. Likelihood (LL)

Distribution		Maximum likelihood estimation (MLE) Method		Moment matching estimation (MME) Method.		Quantile matching estimation (MME) Method	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	LL	45.34	5	24.13	4	39.997	6
Lognormal	LL	51.8	1	25.123	1	49.447	1
Gamma	LL	50.394	2	24.875	2	48.499	2
Beta	LL	50.053	3	23.201	3	48.37	3
Cauchy	LL	43.236	7	-		35.967	7
Weibull	LL	48.599	4	-		47.738	4
Exponential	LL	37.206	8	-6.393	6	-	
Uniform	LL	31.823	9	-Inf		-Inf	
Logistic	LL	44.69	6	24.116	5	43.343	5

Table 4. 16 Region-B Maximizing Goodness-of-fit information criteria (MGFE). LL: Likelihood Criterion

Distribution	Kolmogorov–Smirnov		Cramer–von Mises		Anderson–Darling	
	Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	44.825	5	43.707	5	44.688	5
Lognormal	51.073	1	51.305	1	51.26	1
Gamma	49.75	2	49.903	2	49.8	2
Beta	49.451	3	49.512	3	49.468	3
Cauchy	41.52	7	39.978	7	-	
Weibull	48.212	4	47.633	4	48.246	4
Exponential	-		-		-	
Uniform	-Inf		-Inf		--Inf	
Logistic	44.541	6	44.353	6	44.169	6

Table 4. 17 Region-B Goodness-of-fit information criterion. AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion

Distribution		Maximum likelihood estimation (MLE) method.		Moment matching Estimation (MME) method.		Quantile matching estimation (MME) method	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	AIC	-86.34	5	-86.68	1	-75.99	6
	BIC	-83.68		-83.628		-72.94	
Lognormal	AIC	-99.61	1	-46.247	2	-94.86	1
	BIC	-96.56		-43.315		-91.84	
Gamma	AIC	-96.78	2	-45.875	3	-92.99	2
	BIC	-93.73		-42.819		-89.94	
Beta	AIC	-96.11	3	-42.201	5	-92.74	3
	BIC	-93.05		-39.47		-89.68	
Cauchy	AIC	-82.24	7	-		-67.93	7
	BIC	-79.42		-		-64.88	
Weibull	AIC	-93.2	4	-		-91.48	4
	BIC	-90.15		-		-88.42	
Exponential	AIC	-72.41	9	14.792	6	-	
	BIC	-70		16.25		Inf	
Uniform	AIC	-79.64	8	Inf		Inf	
	BIC	-71.59		Inf		-82.69	5
Logistic	AIC	-85.3	6	-44.23	4	-79.63	
	BIC	-82.3		-41.3		-	

Table 4. 18 Region-C Goodness-of-fit information criterion. Likelihood

Distribution		Maximum Likelihood Estimation (MLE) Method.		Moment Matching Estimation (MME) Method		Quantile Matching Estimation (MME) Method	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	LL	59.699	5	60.528	5	58.25	5
Lognormal	LL	62.64	1	62.589	1	62.445	1
Gamma	LL	61.998	2	61.925	2	61.546	2
Beta	LL	61.641	3	61.575	3	61.045	3
Cauchy	LL	58.491	6	-		47.447	7
Weibull	LL	58.35	7	-Inf		53.685	6
Exponential	LL	28.197	9	28.197	6	-	-
Uniform	LL	55.849	8	-Inf		-Inf	-
Logistic	LL	60.72	4	60.545	4	58.352	4

Table 4. 19 Region-C Maximizing Goodness-of-fit information criteria (MGFE). LL: Likelihood Criterion

Distribution	Kolmogorov–Smirnov		Cramer–von Mises		Anderson–Darling	
	Statistic	Rank	Statistic	Rank	Statistic	Rank
Normal	55.752	5	59.51	5	62.613	5
Lognormal	61.291	1	62.598	1	62.613	1
Gamma	59.811	2	61.779	2	61.968	2
Beta	59.613	3	61.267	3	61.597	3
Cauchy	39.978	6	58.49	6	39.978	6
Weibull	47.547	7	47.772	7	57.708	7
Exponential	25.704	9	26.739	9	26.825	9
Uniform	-	8	-	8	-	8
Logistic	59.541	4	60.563	4	60.716	4

Table 4. 20 Region-C Goodness-of-fit information criterion. AIC: Akaike Information Criterion, BIC: Bayesian Information Criterion

Distribution		Maximum Likelihood Estimation (MLE) Method. Statistic	Rank	Moment Matching Estimation (MME) Method. Statistic	Rank	Quantile Matching Estimation (MME) Method Statistic	Rank
Normal	AIC	-115.4	5	-151.18	5	-112.51	5
	BIC	-112.35		-112.13		-109.45	
Lognormal	AIC	-121.28	1	-121.18	1	-120.89	1
	BIC	-118.23		-118.13		-117.83	
Gamma	AIC	-120	2	-119.85	2	-119.09	2
	BIC	-116.95		-116.8		-116.04	
Beta	AIC	-119.28	3	-119.15	3	-118.09	3
	BIC	-116.23		-116.1		-115.03	
Cauchy	AIC	-112.98	6	-		-90.822	6
	BIC	-109.93		-		-87.776	
Weibull	AIC	-112	7	-		-103.37	7
	BIC	-109		-		100.317	
Exponential	AIC	-54.39	9	-54.395	6	-	
	BIC	-52.869		-52.869		-	
Uniform	AIC	-107.7	8	Ibf		Inf	
	BIC	-104.64		Inf		Inf	
Logistic	AIC	-117.45	4	-117	4	-112.7	4
	BIC	-114.39		-114.03		-109.65	

Sample Parameter Estimation using R-Studio

R version 4.1.1 (2021-08-10) -- "Kick Things"

[Workspace loaded from ~/.RData]

> region_one_r_data <- read.csv("C:/Users/Amane/Desktop/r regional data/region one r data.csv")

> View(region_one_r_data)

> summary(region_one_r_data)

X335.758

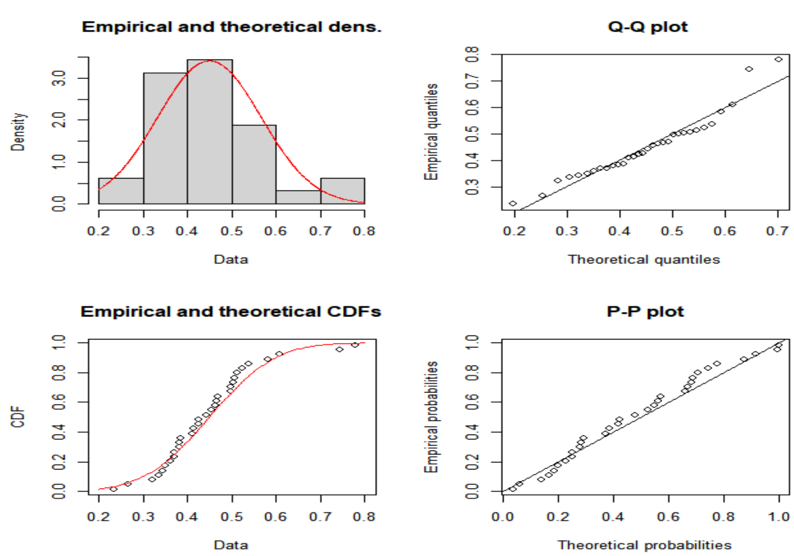
Min. :235.4

1st Qu.:370.5

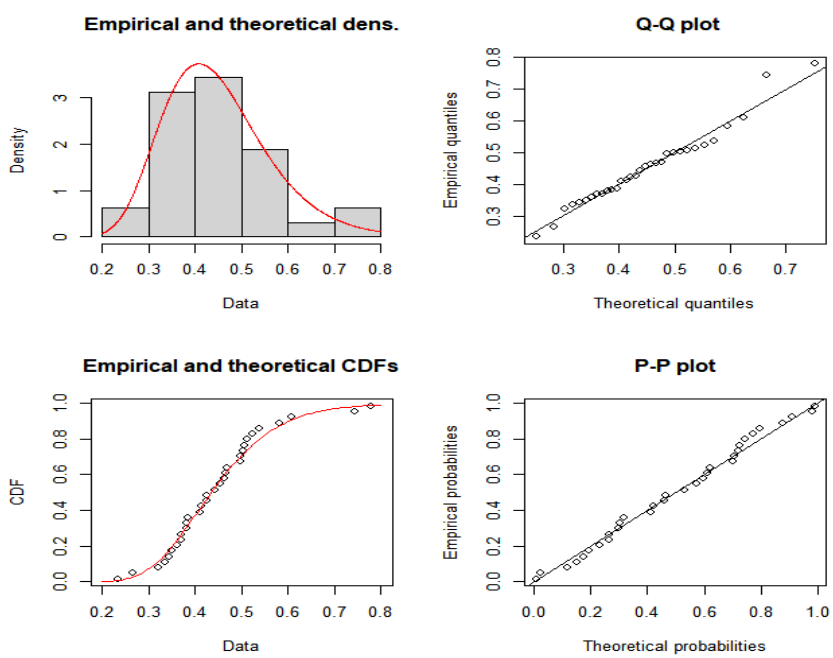
Median :434.5

Mean :449.3

```
3rd Qu.:504.6
Max.   :779.8
> View(region_one_r_data)
> x<-region_one_r_data$X335.758/1000
> summary(x)
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
0.2354  0.3705  0.4345  0.4493  0.5046  0.7798
> norMLE<-fitdist(x,"norm",method = "mle")
> summary(norMLE)
Fitting of the distribution ' norm ' by maximum likelihood
Parameters :
      estimate      Std. Error
mean 0.4492776      0.02064522
sd   0.1167870      0.01459356
Loglikelihood:  23.31088   AIC:  -42.62176   BIC:  -39.69028
Correlation matrix:
              mean            sd
mean  1.000000e+00  -2.675968e-13
sd   -2.675968e-13   1.000000e+00
> plot(norMLE)
```



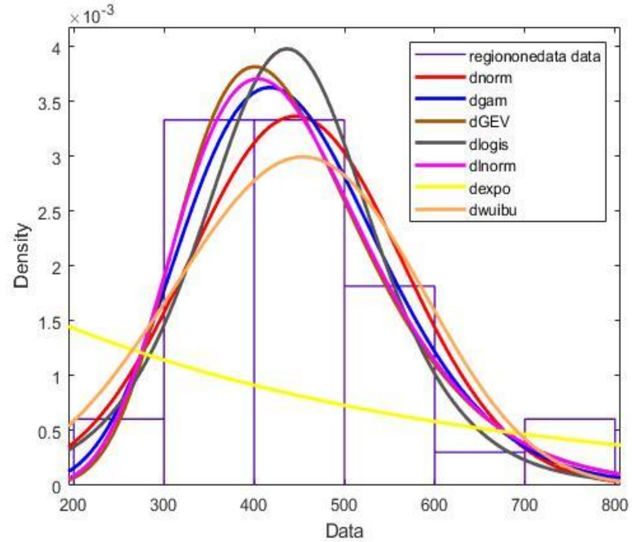
```
> lnorMLE<-fitdist(x,"lnorm",method = "mle")
> summary(lnorMLE)
Fitting of the distribution ' lnorm ' by maximum likelihood
Parameters :
          estimate     Std. Error
meanlog 0.8323966     0.04484289
sdlog    0.2536697     0.03170649
Loglikelihood:  25.12577   AIC:  -46.25154   BIC:  -43.32007
Correlation matrix:
        meanlog   sdlog
meanlog      1      0
sdlog        0      1
> plot(lnorMLE)
```



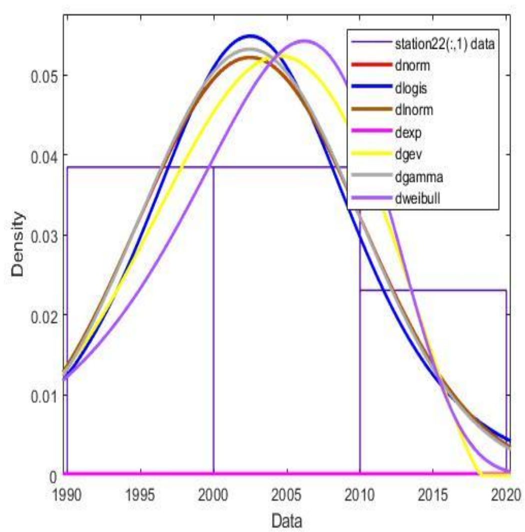
4.4.2 Goodness of Fit Using Assessment-Based Graphs

Various graphical functions can be used to investigate the goodness-of-fit of fitted distributions. For selecting the best-fit parameter estimation methods utilizing the MLE approaches illustrated above the MLE approach was the best parameter estimation approach for the regions. Figure 4.4 shows the output for the best and worst fitted distributions to the regions by the MLE approach. It shows the histogram of the empirical distribution (data) generated according to (Blom, 1958) superimposed on the PDF of the theoretically suited distributions. The results of fitting the chosen distribution functions to the streamflow datasets are shown. The plots of the lognormal, gamma, and logistic models appear to fit the region-A flow dataset series, and may thus be the recommended models for this dataset. The log-likelihood values of the lognormal and gamma models are substantially skewed to the left for Region-B, therefore models lognormal, and gamma models appear to be best suited for this dataset series. For Region-C, the lognormal, gamma, beta, and logistic models are the best options.

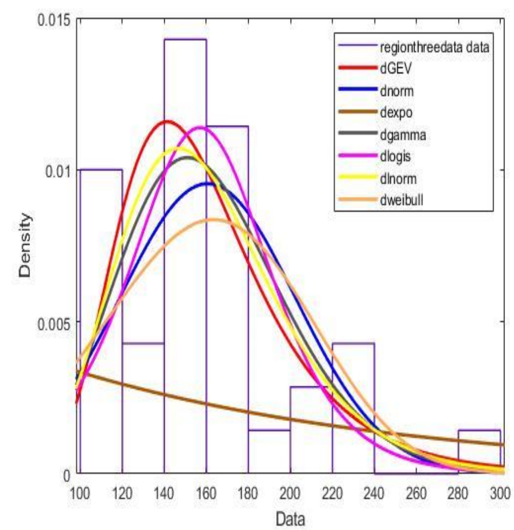
Quantile–quantile (Q–Q) plots were created to analyze and visualize the goodness of fit of the selected model distributions graphically, as well as to determine if the dataset series were generated from the nine theoretical distributions. A probability–probability (P–P) plot is a simple graphical approach for evaluating the quality of forecast prediction and forecast uncertainty. It compares the observed stream-flow probability values in the 0 to 1.0 range against a uniform distribution within the stream-flow ensemble. The P–P plot would be 1:1 if the data set series were perfectly regularly distributed. This holds true for the Q as well. The P–P and Q–Q charts for the data series related to the Region-A, Region-B, and Region-C for the best and the worst theoretical distributions are shown in fig.4. 5. The P–P plots for Region-A data series are regularly distributed for lognormal, gamma, and logistic distribution whereas both regions are unfairly distributed for exponential and uniform distribution. Both Region-B and Region-C data series are regularly distributed for lognormal, gamma, and beta distribution. Similarly, those two regions are unfairly distributed for Exponential, Weibull, and uniform distribution, as seen in Figure 4.6. As a result, the observed data series were presumed to be the real distribution, and the quantiles computed from the observed data were thought to be the theoretical distribution's true quantiles as well. From the chosen distribution models lognormal, gamma, logistic, normal, Weibull, and beta fit the data well in region-one and region-two however, for region-three lognormal, gamma, beta and logistic fit the data well during the computation process.



(a) Region one

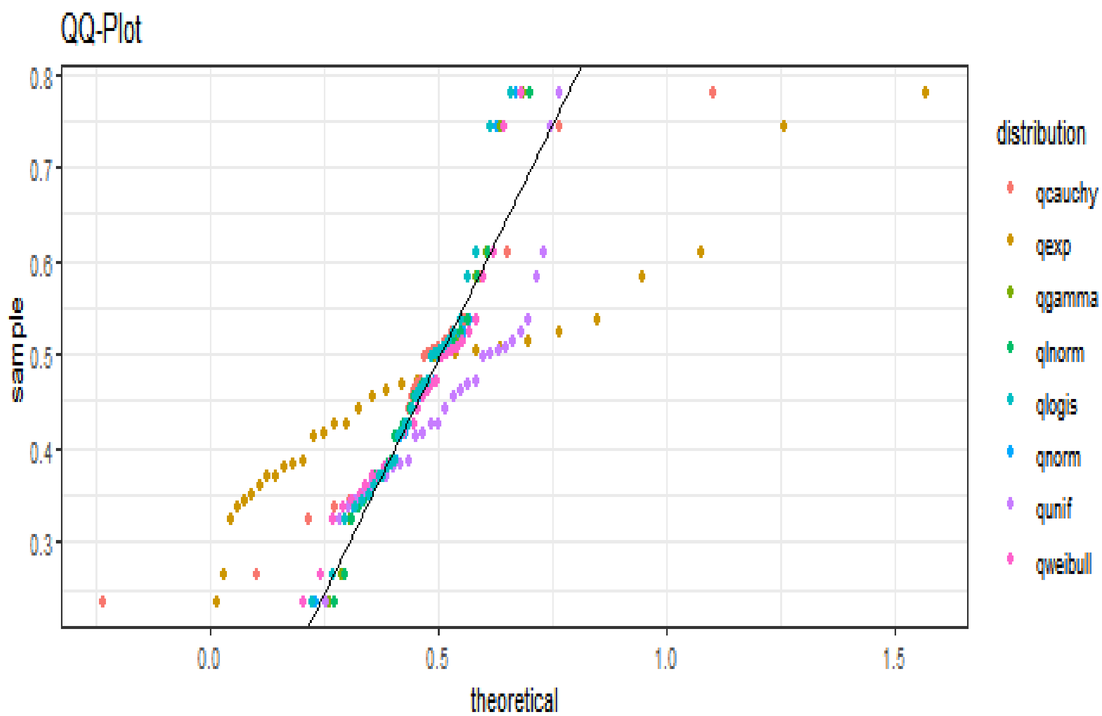


(b) Region two

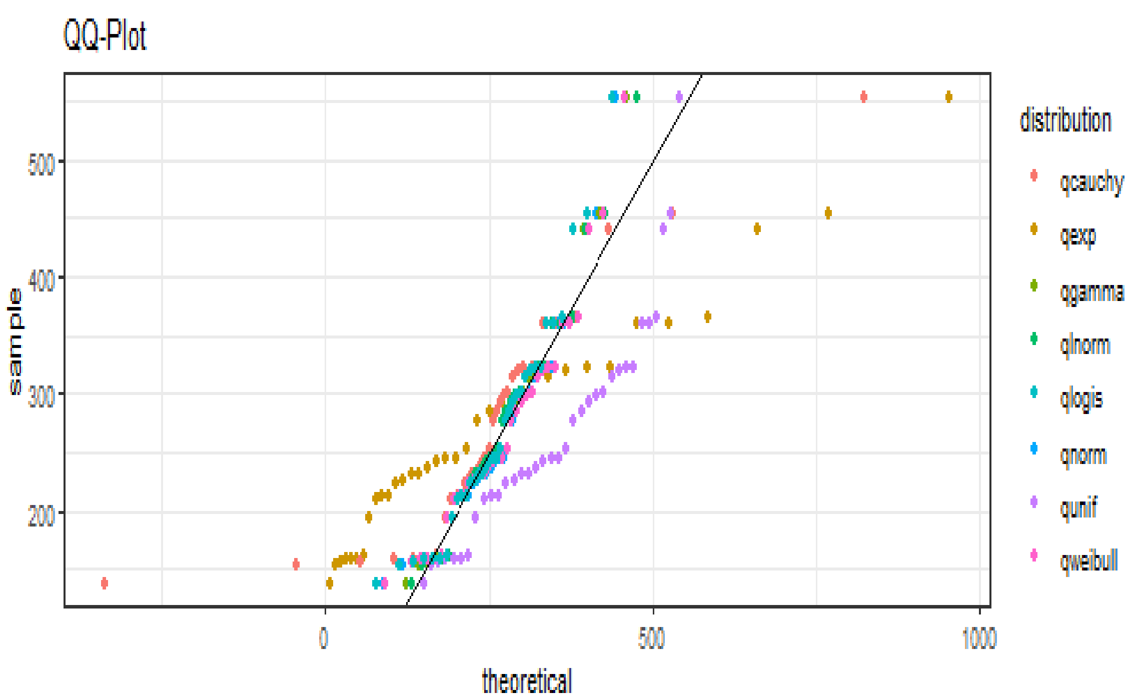


(c) Region three

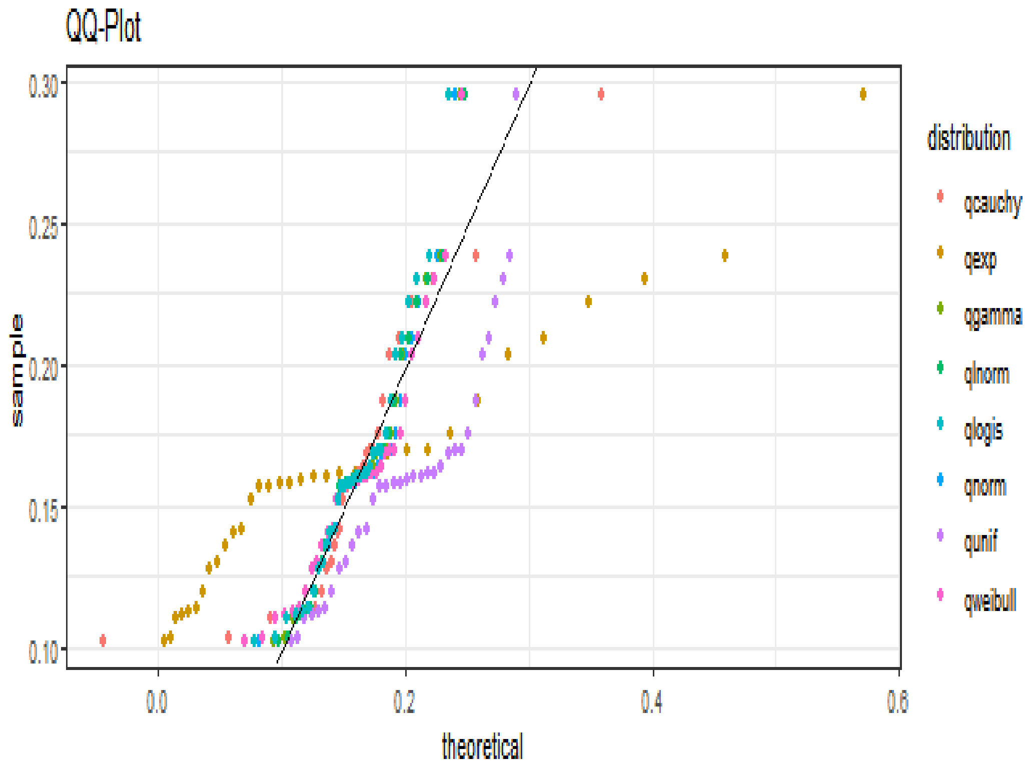
Figure 4. 4: Fitted Cumulative distribution functions (CDF) of the nine selected distribution models:



(a) Region one

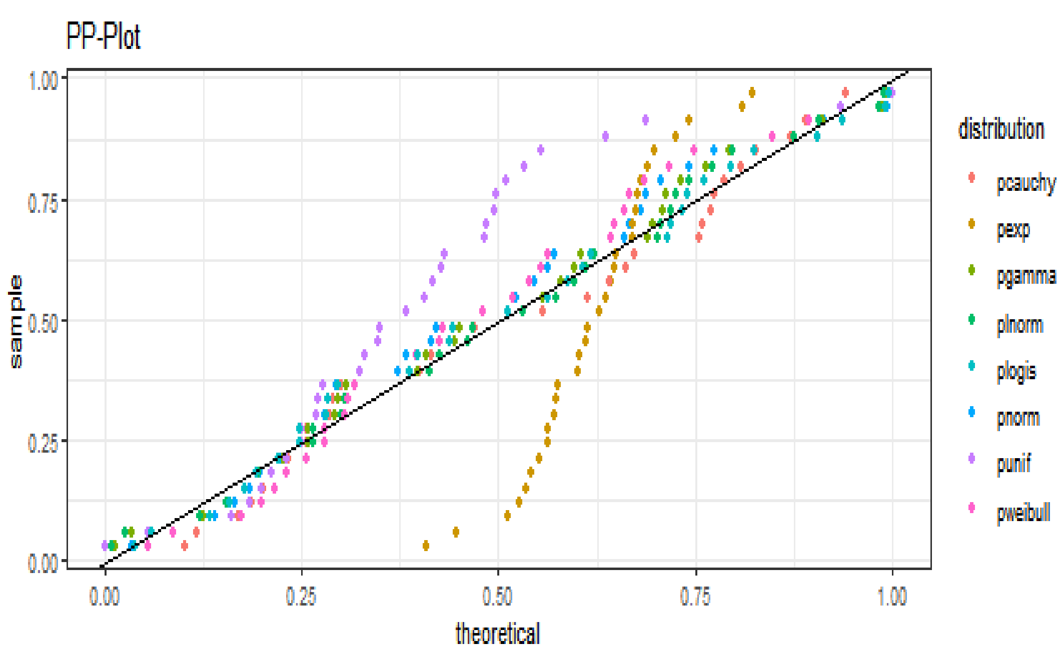


(b) Region two

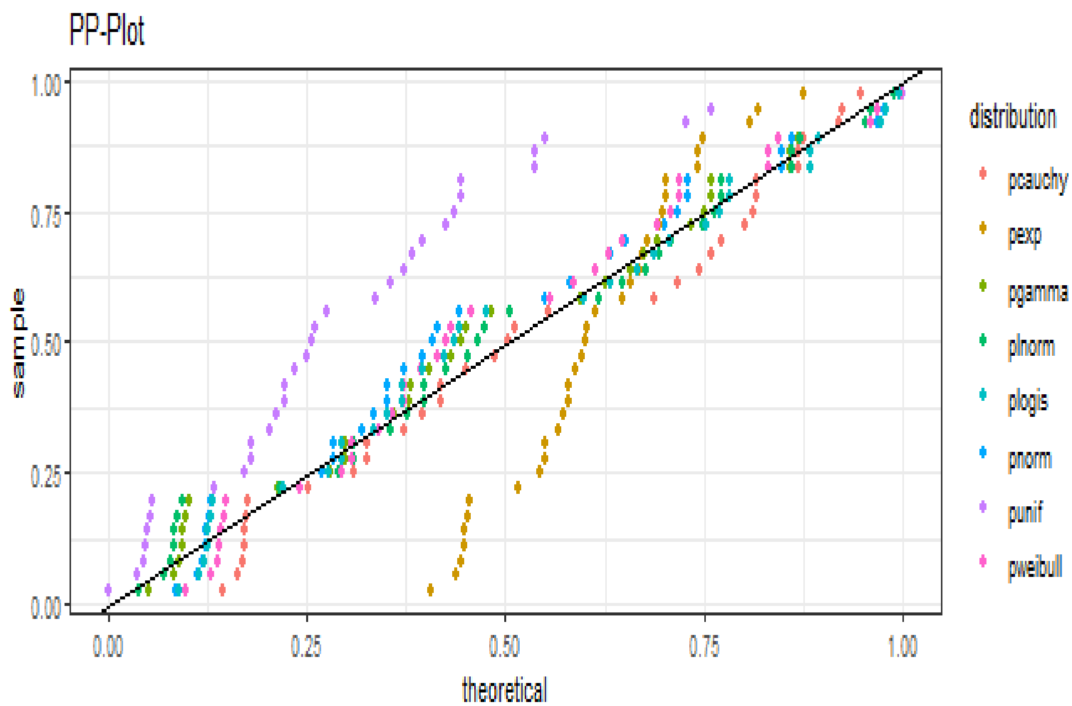


(c) Region three

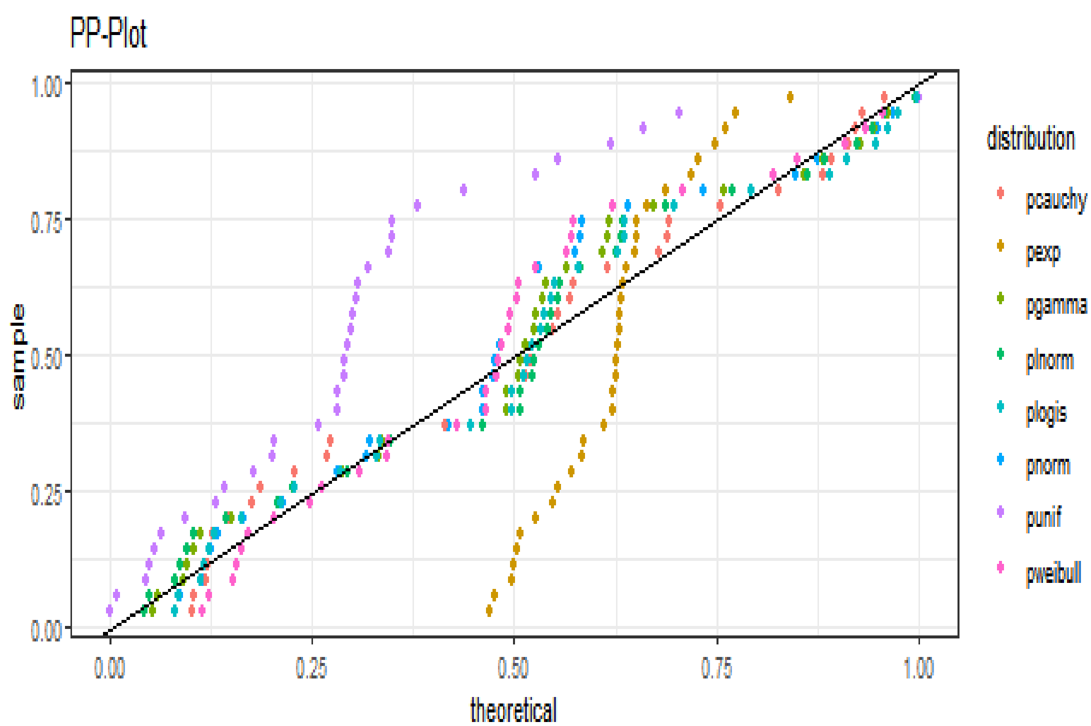
Figure 4. 5: Quantile–quantile (Q–Q) plots for region one, region two, and region three.



(a) region one



(b) region two



(c) region three

Figure 4. 6: Probability–probability (P–P) plots for region one, region two, and region three

4.4.3 Best-Fit Distribution Model

In the selection of models for comparing the multiple probability distributions, AIC employs the same response variables for all of the candidate distributions while keeping all of the components of each likelihood and does not mix null hypothesis testing with information criterion models. The best model is chosen based on the determined minimal AIC value. The BIC, like the AIC, is a model selection criterion based on the likelihood function. Lower BIC values indicate that there are fewer explanatory variables and that the model fits better. The best-fit distribution for extreme datasets is the one that best matches the goodness-of-fit statistics and information criteria-based model selection criteria. Table 12 and Table 14 shows the findings of the best-fit distribution model for region-A based on LL, AIC, and BIC inside the extreme type of distribution. Because they had the higher LL values and lowest AIC and BIC values, the lognormal, gamma, and logistic distribution models were chosen as the best-fitting functions for the region-A streamflow, as shown in Table 15, Table 17, Table 18, and Table 20, because both the AIC and BIC returned low values and higher LL values, the lognormal, gamma, and beta functions were the best-fit functions for the both Region-B and Region-C stream flows at the location located in the upper Omo river basin.

4.4.4 Goodness of fit Test-Based Analysis

The Kolmogorov–Smirnov test, chi-square test, and Anderson–Darling tests were used to examine the nine probability distributions at a 95% significance level ($\alpha = 0.05$). All except exponential and uniform distribution, the classical goodness-of-fit test statistics for the Kolmogorov–Smirnov, chi-square test, and Anderson–Darling tests were acceptable at this stage for the estimation of Region-A streamflow based on the test hypothesis at a 95% significance level. The computed values of test statistics are lower than the critical values for the seven probability distributions at the chosen significance level. Each distribution was assigned a rank between one and nine as indicated on Table 20, with one indicating the best-fitting distribution, and nine indicating the worst fitting distribution. The goodness-of-fit tests indicated the lognormal model as the best quality of fit for the region-A dataset for the upper Omo River basin, followed by the gamma and logistic distribution models in the ranking order based upon the goodness-of-fit results at the site. The beta, Weibull and Cauchy distribution models respectively had the least quality of performance for streamflow at this site. As for the Region-B and Region-C streamflow datasets, lognormal, gamma, beta and logistic showed the best goodness-of-fit statistics for all the test statistics. The purpose of the goodness-of-fit test is to determine the best fitting frequency distribution by computing the difference of the L-kurtosis between the sample data and using R studio software.

In this study, the goodness of fit tests was performed for all distributions using Kolmogorov Smirnov, Anderson-Darling, and Chi-Squared methods for the data of gauging stations. They were applied to determine whether the distribution to be fitted to the data or not. The best-fit result of each station was taken as the nine different distributions with the lowest sum of the rank orders from each of the three test statistics. The GOFs at a5% level of significance was used to define the best-fit ranking using R studio statistical software. The probability distribution having the accepted along with their test statistic was presented in Table 4.21 for Region-A, form the nine distribution the exponential and uniform distribution were not accepted distribution because their values of GOFs at 5% level of significance were greater than the critical value. The same is true for Region-B based on Table 22. Whereas the Weibull, exponential and uniform distributions were not accepted distribution due to their higher values of the test result for Region-C as presented in Table 23.

Table 4. 21 Region-A Goodness-of-fit information Kolmogorov–Smirnov, Chi-squared and Anderson–Darling criteria

	Kolmogorov–Smirnov (Critical Value at 0.05 = 0.20517)		Chi-squared test (Critical Value at 0.05 = 14.067)		Anderson–Darling (Critical Value at 0.05 = 2.5018)	
Distribution	Statistic	Remark	Statistic	Remark	Statistic	Remark
Normal	0.07223	Accepted	0.23988	accepted	1.62899	Accepted
Lognormal	0.10945	Accepted	0.2116	accepted	0.46318	Accepted
Gamma	0.09057	Accepted	0.65489	accepted	0.81632	Accepted
Beta	0.0761	Accepted	1.6239	accepted	0.26539	Accepted
Cauchy	0.11714	Accepted	2.3704	accepted	0.66346	Accepted
Weibull	0.12003	Accepted	0.27273	accepted	2.455	Accepted
L2ogistic	0.10358	Accepted	0.60733	accepted	0.96873	Accepted
Exponential	0.45545	Unaccepted	60.631	unaccepted	8.7586	Unaccepted
Uniform	0.15549	Accepted	N/A	unaccepted	8.521	Unaccepted

Table 4. 22 Region-B Goodness-of-fit information Kolmogorov–Smirnov, Chi-squared and Anderson–Darling criteria

	Kolmogorov–Smirnov (Critical Value at 0.05 = 0.20517)		Chi-squared test (Critical Value at 0.05 = 14.067)		Anderson–Darling (Critical Value at 0.05 = 2.5018)	
Distribution	Statistic	Remark	Statistic	Remark	Statistic	Remark
Normal	0.13832	Accepted	1.6634	accepted	0.69392	Accepted
Lognormal	0.1084	Accepted	0.72721	accepted	0.33539	Accepted
Gamma	0.08385	Accepted	1.6293	accepted	0.34973	Accepted
Beta	0.10903	Accepted	0.37143	accepted	2.21213	Accepted
Cauchy	0.14301	Accepted	1.2269	accepted	0.92233	Accepted
0.76709	0.10074	Accepted	0.87016	accepted	2.455	Accepted
Logistic	0.13952	Accepted	3.2856	accepted	0.64159	Accepted
Exponential	0.40991	Unaccepted	32.036	unaccepted	7.3859	Unaccepted
Uniform	0.15533	Accepted	N/A	unaccepted	11.655	Unaccepted

Table 4. 23 Region-C Goodness-of-fit information Kolmogorov–Smirnov, Chi-squared and Anderson–Darling criteria

	Kolmogorov–Smirnov (Critical Value at 0.05 = 0.20517)		Chi squared test (Critical Value at 0.05 = 14.067)		Anderson–Darling (Critical Value at 0.05 = 2.5018)	
Distribution	Statistic	Remark	Statistic	Remark	Statistic	Remark
Normal	0.20297	accepted	13.201	accepted	2.1767	Accepted
Lognormal	0.10547	accepted	0.473	accepted	0.44467	Accepted
Gamma	0.17761	accepted	2.9342	accepted	1.2396	Accepted
Beta	0.19647	accepted	4.8571	accepted	2.0871	Accepted
Cauchy	0.20403	accepted	3.4518	accepted	2.4763	Accepted
Logistic	0.25452	unaccepted	8.7703	accepted	0.96873	Accepted
Weibull	0.1978	accepted	1.8303	accepted	13.185	Unaccepted
Exponential	0.34709	unaccepted	15.725	unaccepted	4.5967	Unaccepted
Uniform	0.21677	accepted	N/A	unaccepted	12.607	Unaccepted

4.4.5 The Cullen and Frey graph

Figure 4.3 illustrates an initial skewness-kurtosis graph of the unbiased distribution of the extreme data to aid in visualization and model selection. For the skewness and kurtosis, uniform, exponential, normal, and logistic models have only one potential distribution value, whereas the possible lognormal, Weibull, and gamma areas are depicted bylines, and the possible Beta areas are represented by larger areas. In the Cullen and Frey graph, the kurtosis and squared skewness of extreme datasets is represented as a blue point denoting "observation". Lognormal, gamma, normal, and Weibull, being common right-skewed distributions, are indicated as possible model distribution candidates because of the positive skewness, and a kurtosis value that is close to three, while the GEV, Gumbel, and GP distributions should not be considered. As a result, this justified that the lognormal, gamma and normal distributions would be acceptable and the dominant probability distributions in the upper Omo-River Basin for estimation of regional flood frequency.

4.5 Estimation of Regional Flood Frequency Curves

Following the acceptance of areas as homogeneous, appropriate distributions for the regions were found. For each region, flood frequency curves were created based on an appropriate distribution to calculate the variances in the standardized flow of various return periods.

4.5.1. Parameter and Quantile Estimations

According to the Likelihood value (LL), Akaike information criterion (AIC), and Bayesian information criterion (BIC), the maximum likelihood technique provides a better parameter estimation procedure since it provides greater LL values and lower AIC and BIC values in all regions, Region-A, Region-B, and Region-C. The results obtained from maximum streamflow data series analysis by MLE for the Upper Omo-Gibe river basin and its vicinity indicate that, based on most credible functions of probability distribution by LL, AIC, and BIC, lognormal, and gamma, which are both two-parameter distributions, are the most suited models for maximum flood prediction studies. The MLE approach has been used in this regional flood frequency distribution analysis and recommendations of distribution functions made for extreme streamflow data series. The MLE estimates are consistent, and the maximum likelihood estimators are asymptotically unbiased and efficient; thus, the MLE is generally preferred over the method-of-moments (MoM) approach, the Quantile matching approach, and the maximizing goodness of fit approach. Therefore, the recommended distributions for the frequency regimes in this study may be used to estimate the regional flood frequency distribution (return periods for various flood extremes). These findings largely agree with other recommendations elsewhere in Ethiopia RFFA conducted on the upper awash river basin, Kenya, and the world. The Gamma distributions have been extensively used in the study of floods in the United States and Australia (Beard, 1992) since their adoption and recommendation in Bulletin 17B. Recent work recommended lognormal annual maximum flood analysis in the Tel Basin of the Mahanadi River System, India using MLE (Guru and Jah, 2015).

The MLE's estimation entails selecting parameter estimates that produce the highest chance of the observations occurring. Table 4.24 shows the best parameter estimates from R studio software for a selection of distribution models. The remark and descriptive statistics of the goodness fit tests presented in Table 4.21, Table 4.22, and Table 4.23 were used to create these results. As a result, those distributions could be accepted as the most appropriate and dominant distributions in the upper Omo-Gibe River Basin for accurate evaluation and estimation of floods.

Table 4. 24 Results of Estimation parameters for fitted distributions in the region

Name of Regions	The Best Parameter Estimation Model	Best-fitted distribution	Value of parameters	
			shape	Scale
Region-A	Maximum likelihood Estimation	1- Lognormal	0.833	0.255
		2- Gamma	15.654	34.843
		3- Logistic	0.44	0.064
Region-B	Maximum likelihood Estimation	1- Lognormal	2.216	0.483
		2- Gamma	4.251	34.519
		3- Beta	3.685	26.152
Region-C	Maximum likelihood Estimation	1- Lognormal	1.86	0.2464
		2- Gamma	16.1887	100.84
		3- Beta	13.346	69.712

4.5.2 Estimation of Index-Flood for Standardization

The average of the growth curves was used to illustrate the flood frequency curves of regions in this scenario. The standardized quantiles for regions using the given distribution and parameters with their corresponding return periods are shown in Table 4.25. For specified parameter estimation, the amount of flood rises as the return period increases for all stations. This could be related to the varying flood regimes of the meteorological phenomena that cause flood this can significantly help in risk assessment works, water resources management, and engineering decisions and actions in the study area.

Table 4. 25 Estimated standardized flood quantiles of regions

Gumbel reduced variant	RGC-A	RGC-B	RGC-C
0.366512921	-3.25	-0.15254	-0.154
1.499939987	-0.15347	0.866361	0.850368
2.250367327	0.856173	1.540963	1.515345
2.673752092	1.524646	1.921568	1.89052
2.970195249	1.901792	2.188058	2.153208
3.198534261	2.165861	2.393325	2.355547
3.901938658	2.369263	3.025655	2.978856
3.901938658	2.995848	3.025655	2.978856
4.310784111	2.995848	3.39319	3.341147
4.600149227	3.360043	3.653316	3.597562
5.007292664	3.617806	4.019321	3.958345
5.295812143	3.980485	4.278688	4.214012
6.213607264	4.237495	5.103746	5.027299
6.907255071	5.055057	5.727306	5.641963
9.21029037	5.672951	7.797636	7.682756

(RGC: Regional Growth Curve)

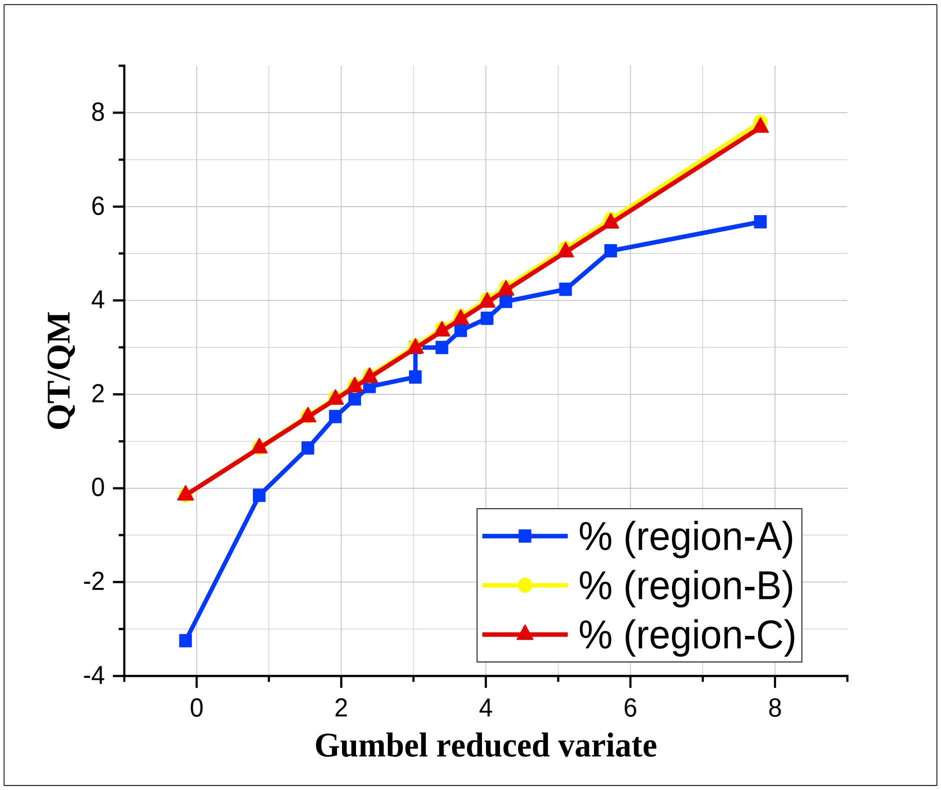


Figure 4. 8: Regional growth curves for delineated homogeneous regions

The three generated regional frequency curves show that each curve has different flood characteristics. This could be because flood statistics vary depending on the location. For the identical return periods, as shown in Figure 4.8, the estimated regional growth curve of Region-B and Region-C yielded greater quantile estimates than Region-A. This large-scale flood in the region could inflict significant damage and disruption to nearby residents. This could be because that their flood regimes and contributing areas are both variables. Higher variances in regional curves may be owing to significant spatial fluctuations in altitudes, as well as the spatially undulating mountainous topography of regional boundaries, which generates flood forecast uncertainty.

4.5.3 Estimation of flood quantiles

Estimation of flood quantiles was applied for 2, 5, 10, 15, 20, 25, 30, 50, 75,100,150, 200, 500, and 1000 years return periods are shown in Table 4.26, and flood frequency curves for regions were developed as shown in Figure 4.9. Flood frequency curves were estimated using lognormal distribution flood estimation equations. This estimation of the flood can be utilized in the designing of vital hydraulic structures in the river reach.

Table 4. 26 Estimated flood quantiles of regions

T(year)	Region-A	Region-B	Region-C
2	0	111.4490342	154.1639675
5	427.6593441	176.3407042	196.1514292
10	547.2488022	219.3046362	223.9507822
15	626.4274293	243.5445235	239.6349423
20	671.0993416	260.5166712	250.6165891
25	702.3774693	273.5896781	259.0753385
50	726.4698345	313.8614232	285.1327409
50	800.6870057	313.8614232	285.1327409
75	800.6870057	337.2688969	300.2782966
100	843.8248551	353.8358081	310.9977395
150	874.3561689	377.145837	326.0802446
200	917.3144364	393.6643332	336.7683611
500	947.7565258	446.2105093	370.7678028
1000	1044.594359	485.9236641	396.4637752

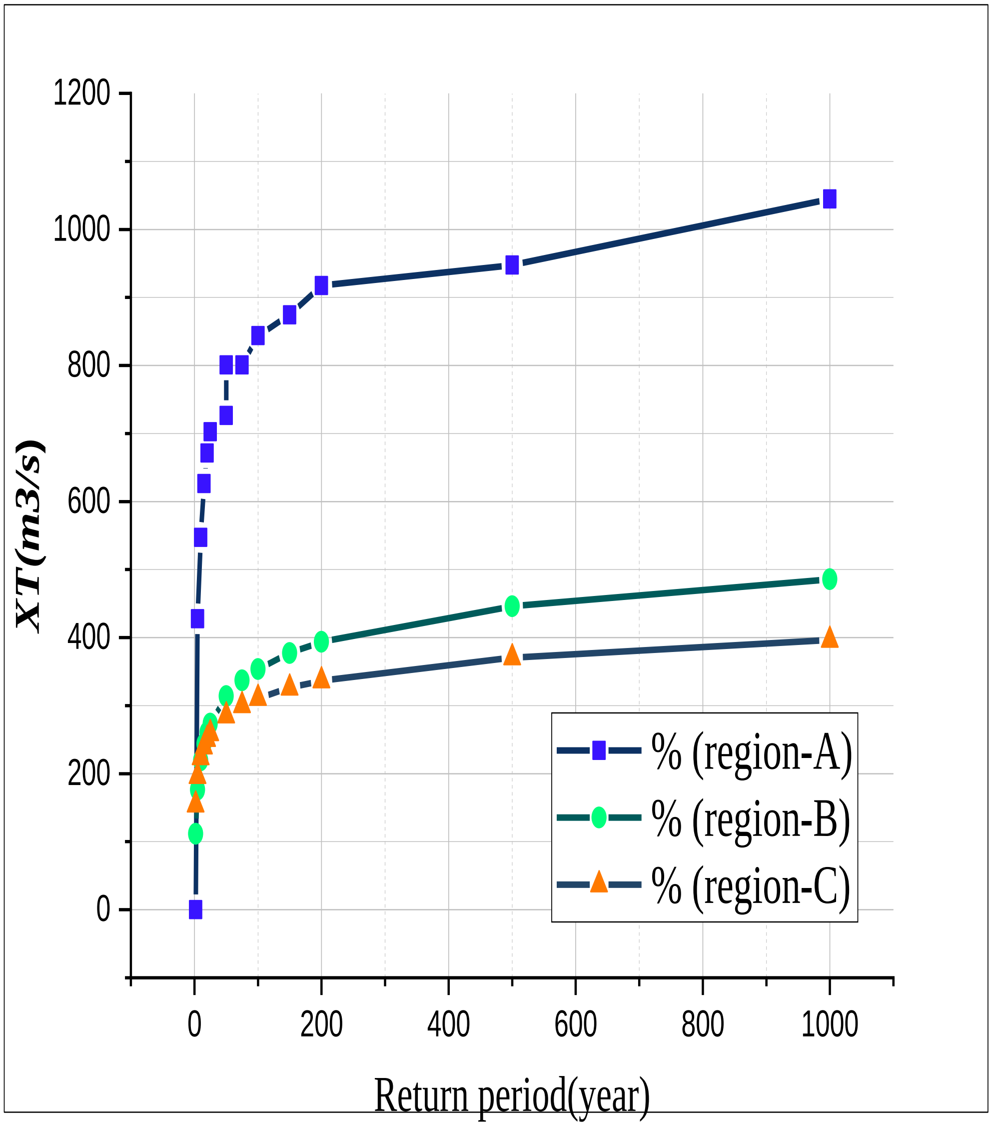


Figure 4. 9: Flood frequency curves of regions

5. CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusions

In this study, regional flood frequency analysis was performed using the data of eleven stream gauging stations to ensure reliable estimation of flood in the Upper Omo-Gibe River Basin. The basin has been defined and delineated into three hydrologically homogeneous regions using the AMF frequency model. The regions were named Region-A, Region-B, and Region-C comprising five, three, and three gauging sites respectively. The delineation of the regions was done using ArcGIS10.4.1. L-moment ratio diagram and R programming software wse used to check whether all stations in the same region are found to lie on the same type of distribution. Further, a discordance measure using Matlab2018a and CC test was conducted to check their homogeneity. A case study is also presented where long-term flood discharge magnitudes and frequencies were extracted from streamflow data for the study area to give an annual maximum time series for this hydrological frequency analysis. Nine probability distribution models were assessed using the MLE approach, GoF tests-based analysis, and information criteria-based selection procedures to identify the most suitable distribution model for RFFA for the study. Lognormal and Gamma distribution models were selected as the best-fit functions for all the three regions stream flows respectively. The GoF tests-based analysis and procedures are useful in the selection of suitable distribution model functions for the site. Regional average values of LCs and LCk were used to select the best fit statistical distribution of each region and the goodness of fit test by using the R programming software was used to approve the best fit distribution. Different distribution functions may be suitable for the flood frequency estimations at the same site; therefore, the choice of a suitable model for flood frequency analysis with the same climatic, catchment, and hydrological characteristics depends on the frequency regime of the data series. Flood data are stochastic in nature and often assumed to be spatially and temporally independent. Practically, a true probability distribution of the data at a given site or region remains uncertain. However, to date, distributions are often used to characterize the relationship between flood magnitudes, and their frequencies are evaluated to assess their performance and selected by using statistical tests. For all the three regions in the study area, the lognormal and gamma distribution functions were selected as the best-fit distribution function since they provide greater LL values and lower AIC and BIC values in all regions, Region-A, Region-B, and Region-C. In addition, these distributions with the method of parameter estimations are finally used to develop a regional growth curve of each homogeneous region. The regional growth curve can be used to safely and feasibly design hydrologic projects under-prediction in both gauged and ungauged catchments. The derived results can be useful as a reference in any hydrological considerations like flood risk management, proper planning, and designing of pivotal hydraulic structures such as dams, spillways, bridges, culverts, and urban drainage systems in the study area.

5.2. Recommendations

The study's findings are used to show the directions in which extra effort should be made. The following suggestions are offered for further research in this area. According to the findings, defining hydrologically homogeneous regions based on statistical parameters of gauged sites is a suitable way of regional analysis. R-Studio Statistical Software can be used for other relevant investigations due to the adequacy of best-fit distributions and the acceptability of results.

Further analyses should incorporate the influence of climatic variables such as precipitation on the variability of L-moments of AMFs in the research area, given the evidence of climate change scenarios. Estimated floods should be used as an input to create hydraulic models such as flood danger, risk, and inundation mapping of delineated homogeneous regions separately for proper land and watershed management. To compare the results and acquire a more appropriate flood estimation for ungauged, flood frequency curves should be produced utilizing several types of catchment variables such as elevation, slope, area, precipitation, soil type, land use land cover, and form factor. More hydrologic stations should be established in the basin to acquire an accurate estimate of regional flood quantile. This study's methodological framework could be used to develop comparable research in other river basins.

REFERENCES

Abdullah, M.N., 2011. Catchment area delineation using GIS technique for Bekhma dam. FIG Working Week 2011Bridging the Gap between Cultures, Marrakech, Morocco

Ahmad, I., Fawad, M., Akbar, M., Abbas, A., and Zafar, H., 2016. Regional Frequency Analysis of Annual Peak Flows in Pakistan Using Linear Combination of Order Statistics. Polish Journal of Environmental Studies, 25(6).

Ahmad, U.N., Shabri, A. and Zakaria, Z.A., 2011. Flood frequency analysis of annual maximum stream flows using L-moments and TL-moments approach. Applied Mathematical Sciences, 5(5), pp.243-253. Akirso, B., 2017. Flood Frequency under Changing Climate in the Upper Omo-Gibe River.

Alam, J., Muzzammil, M. and Khan, M.K., 2016. Regional flood frequency analysis: comparison of L-moment and conventional approaches for an Indian catchment. ISH Journal of Hydraulic Engineering, 22(3), pp.247-253.

Jan, N.A.M., Shabri, A. and Badyalina, B., 2016, June. Selecting probability distribution for regions of Peninsular Malaysia streamflow. In AIP Conference Proceedings (Vol. 1750, No. 1, p. 060014). AIP Publishing LLC.

Amirataee, B., Montaseri, M. and Rezaei, H., 2014. Assessment of goodness of fit methods in determining the best regional probability distribution of rainfall data.

Anusha, M. & Surendra, H. J., 2017. Regional Flood Frequency Analysis Using Computer Simulations. International Journal of Advanced Engineering Research and Science, Volume Vol. 4 (1).

Alam, M.A., Emura, K., Farnham, C. and Yuan, J., 2018. Best-fit probability distributions and return periods for maximum monthly rainfall in Bangladesh. Climate, 6(1).

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

Badreldin, G. & Fengo, P., 2012. . Regional Rainfall Frequency Analysis for the Luanhe.

Beard, L.R. Book review: The gamma family and derived distributions applied in hydrology. By bernardbobee and fahim ashkar, water resources publications, littleton, co 80161-2841, USA, 1991, 203 pp., soft cover, $38.00, isbn 0-918334-68-3. J. Hydrol. 1992, 132, 383. [CrossRef]

Bhagat, N., 2017. Flood frequency analysis using Gumbel's distribution method: a case study of Lower Mahi Basin, India. Journal of Water Resources and Ocean Science, 6(4), pp.51-54.

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

Burn, D.H., 1988. Delineation of groups for regional flood frequency analysis. Journal of hydrology, 104(1-4), pp.345-361.

Castellarin, A., Kohnová, S., Gaál, L., Fleig, A., Salinas, J.L., Toumazis, A., Kjeldsen, T.R. and Macdonald, N., 2012. Review of applied statistical methods for flood frequency analysis in Europe.

Chavoshi, S. B. & Azmin, W. S., 2009. Development of L-Moment Based Models for Extreme Flood Events.. Malaysian Journal of Mathematical Sciences., Volume Vol. 3(2), pp. 281- 296.

Cullen, A.C.; Frey, H.C.; Frey, C.H. Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Models and Inputs; Springer Science & Business Media: New York, NY, USA, 1999.

Cunnane, C., 1989. Statistical distributions for flood frequency analysis. Operational hydrology report (WMO).

Dahmen, E.R. and Hall, M.J., 1990. Screening of hydrological data: tests for stationarity and relative consistency (No. 49). ILRI.

Das, S. and Simonovic, S.P., 2012. Assessment of Uncertainty in Flood Flows under Climate Change: The Upper Thames River Basin (Ontario, Canada). Department of Civil and Environmental Engineering, University of Western Ontario, London, Ontario.

Demissie, M., 2008. Regional flood frequency analysis for upper awash sub-basin (upstream of Koka). A dissertation, Addis Ababa University, Ethiopia.

Denboba, M.A., 2005. Forest conversion-soil degradation-farmer's perception nexus: Implications for sustainable land use in the southwest of Ethiopia (Vol. 26). Cuvillier Verlag

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

Dubey, A., 2014. Regional Flood Frequency Analysis Utilizing L-Moments: A Case Study of Narmada Basin. : International Journal of Engineering Research and Applications, Volume 4(2): pp: 155-161. ISSN: 2248-9622.

Engeland, K., Wilson, D., Borsányi, P., Roald, L. and Holmqvist, E., 2018. Use of historical data in flood frequency analysis: a case study for four catchments in Norway. Hydrology Research, 49(2), pp.466-486.

Ergish, N.J., 2010. Flood frequency analysis for regulated watersheds. University of California, Davis. Ganamala, K. and Kumar, P.S., 2017. A case study on flood frequency analysis. International Journal of Civil Engineering and Technology, 8(4), pp.1762-1767.

Flynn, K.M.; Kirby, W.H.; Hummel, P.R. User's Manual for Program Peakfq, Annual Flood-Frequency Analysis Using Bulletin 17b Guidelines; U.S. Geological Survey: Reston, VA, USA, 2006; pp. 2328–7055.

Gebeyehu, A., 1989. Regional Flood Frequency Analysis, Ph.D. Dissertation Submitted to Stockholm University. Hydraulics Laboratory, the Royal Institute of Technology, s.l.: Florida:CRC Press LLC.

Gebeyehu, A., 1989. Regional Flood Frequency Analysis, Ph.D. Dissertation Submitted to Stockholm University, Florida: CRC Press LLC.

Gebereyesus, K and Abegan, F.,2014 Impact assessment of climate change on the hydrology of Gojeb River catchment in western Omo-Gibe River basin, Ethiopia (Doctoral dissertation, Haramaya University).

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

Getahun, Y.S. and Gebre, S.L., 2015. Flood hazard assessment and mapping of flood inundation area of the Awash River Basin in Ethiopia using GIS and HEC-GeoRAS/HEC-RAS model. Journal of Civil & Environmental Engineering, 5(4), p.1.

Ghosh, S., Roy, M.K. and Biswas, S.C., 2016. Determination of the best fit probability distribution for monthly rainfall data in Bangladesh. American Journal of Mathematics and Statistics, 6(4), pp.170-174

Grehys, 19966a. a. Inter-comparison of regional flood frequency procedures for Canadian rivers.. J. Hydrol., pp. 186 (1–4), 85–103.

Grubbs, F.E., 1969. Procedures for detecting outlying observations in samples. Technometrics, 11(1), pp.1-21.

Guru, N.; Jha, R. Flood frequency analysis of tel basin of mahanadi river system, india using annual maximum and pot flood data. Aquat. Procedia 2015, 4, 427–434.

Guru, N., 2016. Flood Frequency Analysis of Partial Duration Series Using Soft Computing Techniques for Mahanadi River Basin in India (Doctoral dissertation).

Haberlandt, U. and Radtke, I., 2014. Hydrological model calibration for derived flood frequency analysis using stochastic rainfall and probability distributions of peak flows. Hydrology and Earth System Sciences, 18(1), pp.353-365.

Haddad, K., 2013. Regional flood frequency analysis in the range of small to large floods: development and testing of Bayesian regression-based approaches.

Hailegeorgis, T.T. and Alfredsen, K., 2017. Regional flood frequency analysis and prediction in ungauged basins including estimation of major uncertainties for mid-Norway. Journal of Hydrology: Regional Studies, 9, pp.104-126.

Hosking, J.R.M. and Wallis, J.R., 2005. Regional frequency analysis: an approach based on L-moments. Cambridge university press.

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

Irwin, S., Srivastav, R. and Simonovic, S.P., 2014. Instruction for Watershed Delineation in an Arc-GIS Environment for Regionalization Studies. Department of Civil and Environmental Engineering, Western University, Canada. May. ISBN :(print), pp.978-0.

Islam, R., Kamaruddin, R., Ahmad, S.A., Jan, S.J. and Anuar, A.R., 2016. A Review on Mechanism of Flood Disaster Management in Asiaf. International Review of Management and Marketing, 6(1).

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

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

Kachroo, R.K., Mkhandi, S.H. and Parida, B.P., 2000. Flood frequency analysis of southern Africa: I. Delineation of homogeneous regions. Hydrological sciences journal, 45(3), pp.437-447.

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

Ketsela, H., Tadele, K. and Temam, D., 2017. An Assessment Of predictive Accuracy For Regional flood frequancy distribution Estimation Methods On Awash River Basins (Doctoral dissertation).

Kochito, G., 2014. Impact Assessment of Climate Change on the Hydrology of Gojeb River Catchment in Western Omo-Gibe River Basin, Ethiopia. MSc. Thesis. School of Graduate Studies, Haramaya University, Ethiopia..

Komi, K., Amisigo, B.A., Diekkrüger, B. and Hountondji, F.C., 2016. Regional flood frequency analysis in the Volta River basin, West Africa. Hydrology, 3(1), p.5.

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

Kumar, R., 2019. Development of regional flood frequency relationships using L-moments for gauged and ungauged catchments of India.

Lampros, V. (2009). Hydrological response to metereological drought indices in thessaly. Greece , journal of hydrologic Engineering.

Lilienthal, J., Fried, R. and Schumann, A., 2018. Homogeneity testing for skewed and cross-correlated data in regional flood frequency analysis. Journal of Hydrology, 556, pp.557-571.

Malekinezhad, H., Nachtnebel, H.P. and Klik, A., 2011. Comparing the index-flood and multiple-regression methods using L-moments. Physics and Chemistry of the Earth, Parts A/B/C, 36(1-4), pp.54-60.

Mei, X., Dai, Z., Tang, Z. and van Gelder, P.H.A.J.M., 2018. Impacts of historical records on extreme flood variations over the conterminous United States. Journal of Flood Risk Management, 11, pp.S359-S369.

Mehrannia, H. and Pakgohar, A., 2014. Using easy fit software for goodness-of-fit test and data generation. Intl J Math Arch, 5(1), pp.2229-5046.

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

Mengistu, K.H. and Sivakumar, K., 2018. An Assessment of Predictive Accuracy for Regional Flood Frequency Distribution Estimation Methods on Awash River Basin. International Journal of Latest Trends in Engineering and Technology. Vol. 11 (1), 30-039. ISSN: 2278-621X: DOI: http://dx. doi. org/10.21172/1.111, 6.

Millington, N., Das, S. and Simonovic, S.P., 2011. The comparison of GEV, log-Pearson type 3 and Gumbel distributions in the Upper Thames River watershed under global climate models.

Mishra, B.K., Takara, K., Yamashiki, Y. and Tachikawa, Y., 2009. Selection of Regional Distribution Using Simulated Flood Data. Annual Disaster Prevention Research Institute, Kyoto University, (52B).

Modi, M. and Mitra, A., 2017. Regional flood frequency analysis: a case study of Sabarmati River Basin. Int J Eng Sci Inven (IJESI), 6, pp.55-65.

Murphy, C., Cunnane, C., Das, S. and Mandal, U., 2014. Flood Studies Update Technical Research Report Volume 11: Flood frequency estimation.

Nobert, J., Mugo, M. and Gadain, H., 2014. Estimation of design floods in ungauged catchments using a regional index flood method. A case study of Lake Victoria Basin in Kenya. Physics and Chemistry of the Earth, Parts A/B/C, 67, pp.4-11

Noto, L.V. and La Loggia, G., 2009. Use of L-moments approach for regional flood frequency analysis in Sicily, Italy. Water resources management, 23(11), pp.2207-2229.

Parida, B.P., Kachroo, R.K. and Shrestha, D.B., 1998. Regional flood frequency analysis of Mahi-Sabarmati Basin (Subzone 3-a) using index flood procedure with L-moments. Water Resources Management, 12(1), pp.1-12.

Pathiraja, S., Westra, S. and Sharma, A., 2012. Why continuous simulation? The role of antecedent moisture in design flood estimation. Water Resources Research, 48(6).

Patil, S. and Stieglitz, M., 2012. Controls on hydrologic similarity: role of nearby gauged catchments for prediction at an ungauged catchment. Hydrology and Earth System Sciences, 16(2), pp.551-562.

Pegram, G. and Parak, M., 2004. A review of the regional maximum flood and rational formula using geomorphological information and observed floods. water sa, 30(3), pp.377-392.

Rabba, Z.A., Fatoyinbo, B.S. and Stretch, D.D., 2018. Applications of the PyTOPKAPI model to ungauged catchments. Water SA, 44(2), pp.162-175.

Rahman, A., Haddad, K. and Kuczera, G., 2015. Features of regional flood frequency estimation (RFFE) model in Australian rainfall and runoff. In Partnering with Industry and the Community for Innovation and Impact through Modelling: Proceedings of the 21st International Congress on Modelling and Simulation (MODSIM2015), 29 November-4 December 2015, Gold Coast, Queensland (pp. 2207-2213).

Rao, A.R. and Srinivas, V.V., 2006. Regionalization of watersheds by hybrid-cluster analysis. Journal of Hydrology, 318(1-4), pp.37-56.

Romali, N.S. and Yusop, Z., 2017. Frequency analysis of annual maximum flood for Segamat River. In MATEC Web of Conferences (Vol. 103, p. 04003). EDP Sciences.

Saf, B., 2009. Regional flood frequency analysis using L-moments for the West Mediterranean region of Turkey. Water Resources Management, 23(3), pp.531-551.

Sah, S. and Prasad, J., 2015. Flood frequency analysis of River Kosi, Uttarakhand, India using statistical approach. International Journal of Renewable Energy Technology, 4(8).

Salinas, J.L., Laaha, G., Rogger, M., Parajka, J., Viglione, A., Sivapalan, M. and Blöschl, G., 2013. Comparative assessment of predictions in ungauged basins–Part 2: Flood and low flow studies. Hydrology and Earth System Sciences, 17(7), pp.2637-2652

Schendel, T. and Thongwichian, R., 2017. Considering historical flood events in flood frequency analysis: Is it worth the effort?. Advances in Water Resources, 105, pp.144-153.

Share Bale Eco-Region, 2017. Drivers of Hydrological Dynamics in the Bale Eco-Region, s.l.: Research Report Series No. 7.

Sine, A. and Ayalew, S., 2004. Identification and delineation of hydrological homogeneous regions-The case study of Blue Nile River Basin. In Proceedings, Lake Abaya Research Symposium.

Smith, A., Sampson, C. and Bates, P., 2015. Regional flood frequency analysis at the global scale. Water Resources Research, 51(1), pp.539-553.

Tadesse, A. H., Gottschalk, L. & Tallaksen, L. M., 2011. Regional Flood Frequency Analysis in Southern Africa, MSc.

Tanaka, T., Tachikawa, Y., Ichikawa, Y. and Yorozu, K., 2017. Impact assessment of upstream flooding on extreme flood frequency analysis by incorporating a flood-inundation model for flood risk assessment. Journal of Hydrology, 554, pp.370-382.

Thyer, M.; Renard, B.; Kavetski, D.; Kuczera, G.; Franks, S.W.; Srikanthan, S. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using bayesian total error analysis. Water Resour. Res. 2009, 45.

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

Vivekanandan, N., 2015. Flood frequency analysis using the method of moments and L-moments of probability distributions. Cogent Engineering, 2(1), p.1018704.

Wilson, D., Fleig, A.K., Lawrence, D., Hisdal, H., Pettersson, L.E. and Holmqvist, E., 2011. A review of NVE's flood frequency estimation procedures. NVE report, pp.9-2011.

Wolka, K., Sterk, G., Biazin, B. & Negash, M., 2018. Benefits, limitations, and sustainability of soil and water conservation structures in Omo-Gibe basin, Southwest Ethiopia. Land Use Policy, 73, 1–10.

Wu, Y., Zhong, P. & Lall, U., 2018. Local and Regional Flood Frequency Analysis Based on Hierarchical Bayesian Model: Application to Annual Maximum Streamflow for the Huaihe River Basin. Journal of Hydrology and Earth System Science Discussions. https://doi.org/10.5194/hess-2018-22.

Wu, Y., Zhong, P., Lall, U. & Lima, H., 2018. Local and Regional Flood Frequency Analysis Based on Hierarchical Bayesian Model: Application to Annual Maximum Streamflow for the Huaihe River Basin.. Journal of Hydrology and Earth System Science Discussions. https://doi.org/10.5194/hess-2018-22.

Yang, L., 2016. Regional flood frequency analysis for Newfoundland and Labrador using the L-Moments index-flood method (Doctoral dissertation, Memorial University of Newfoundland).

Yucel, I.S.M.A.I.L. and Keskin, F., 2011. Assessment of flash flood events using remote sensing and atmospheric model-derived precipitation in a hydrological model. *Int. Assoc. Hydrol. Sci*, *344*, pp.245-251.

Zaman, M., Rahman, A. & Haddad, K., 2012. Regional Flood Frequency Analysis in Arid Regions: A Case Study for Australia. Journal of Hydrology, pp. 475,74–83.

APPENDIX

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

Code	Gauging station (Y)	Nearby station (X)	Regression equation	R^2	Remark
111111	Tollay	Abelti	Y=12905X+146792	0.66542	WC
61015	Abelti	Tollay	Y=65311X-29831	0.97993	WC
91004	Wolkite	Wolkite	Y=11673X-8448.2	0.98291	WC
91010	Wolkite	Endeber	Y=4133.9X-3459.8	0.96229	WC
91007	Endeber	Wolkite	Y=239.63X+25048	0.71648	WC
91008	Asendabo	Serbo	Y=14966X-12104	0.99225	WC
91032	Serbo	Jimma	Y=1854.1X+14818	0.99225	WC
91023	Jimma	Asendabo	Y=125.3X+2310.8	0.75612	WC
92002	Bonga	Shebe	Y=2010.3X+12793	0.87247	WC
91012	Shebe	Andaracha	Y=21439X-13236	0.99649	WC
92004	Andaracha	Bonga	Y=121.82X+625.77	0.88277	WC

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

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

(Source: Grubbs,1969)

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

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

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

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

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

Station name	Subset-1	Subset-2	V1, v2	Ft2.5 %	Ft	Ft97.5%	v	Tt	Tt2.5%	Tt97.5 %
Tollay	2000-2010	2010-2019	10,10	0.269	0.14	3.72	20	-0.725	-2.1	2.1
Abelti	1985-2001	2002-2017	17,16	0.358	0.253	2.735	33	-0.473	-2.036	2.036
Wolkite	1975-1991	1992-2007	17,16	0.358	2.665	2.735	33	-1.934	-2.036	2.036
Wolkite	1975-1990	1991-2005	16,15	0.348	0.140	2.87	31	0.21	-2.038	2.038
Endeber	1990-2002	2003-2014	13,12	0.308	1.070	3.165	25	0.577	-2.056	2.056
Asendabo	1982-1999	2000-2016	18.17	0.382	0.279	2.655	35	0.528	-2.03	2.03
Serbo	2000-2008	2009-2017	9,9	0.248	4.225	2.056	18	2.055	-2.1	2.1
Jimma	1990-2001	2002-2013	12.12	0.305	1.387	3.28	24	1.178	-2.06	2.06
Bonga	1991-2006	2006-2019	15,14	0.335	44.5	2.95	29	7.842	-2.043	2.043
Shebe	1974-1991	1992-2008	18,17	0.382	1.477	2.655	35	1.215	-2.03	2.03
Andracha	1990-2002	2003-2015	13.13	0.315	2.684	3.245	26	1.638	-2.053	2.053

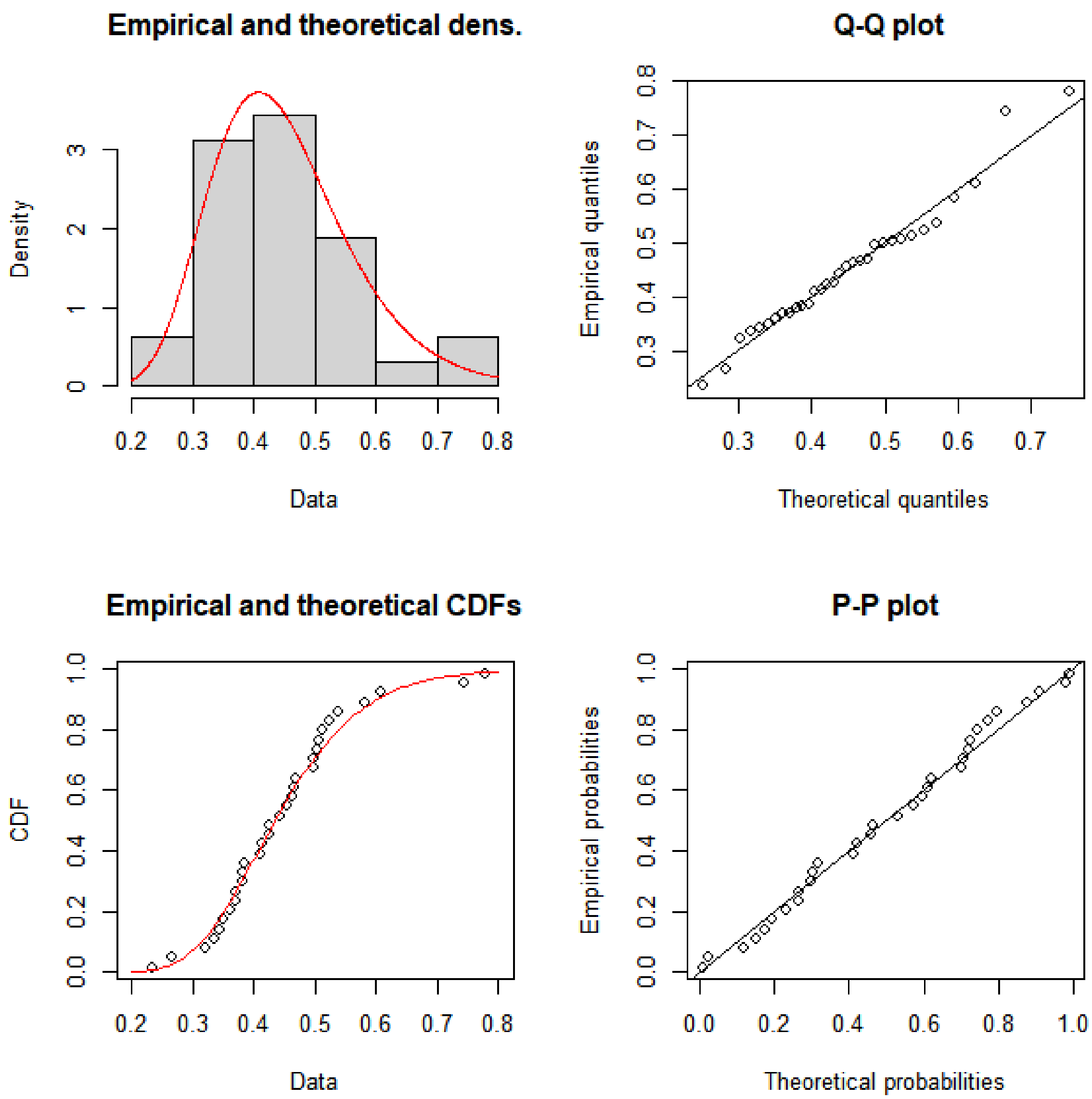
Appendix-F: (Translated Matlab code for Discordancy Measure as provided by Hosking and Wallis, 1997)

```
U=xls. read ('c:\users\name of groups\desktop\U.xls'); % File
% ratios (τ2 i, τ3 i, τ4 i,) of the gauging sites in the region
U= number of gauging sites in the region (Enter the matrix of test statistics);
n=; % input ('enter the number of gauging sites in the group:');
Ubar= [0;0;0];
for i=1: n
Ubar=Ubar+1/n*(U(i,1:3)');
end
S=zeros (3);
for i=1: n
S=S+(U(i,1:3)'-Ubar) *(U(i,1:3)'-Ubar)';
End
for i=1: n
Di(i)=1/3*(U(i,1:3)'-Ubar)'*inv(S)*(U(i,1:3)'-Ubar);
End
disp ('The Di of U Statistics');
disp ('Di, Di+1, Dn');
```

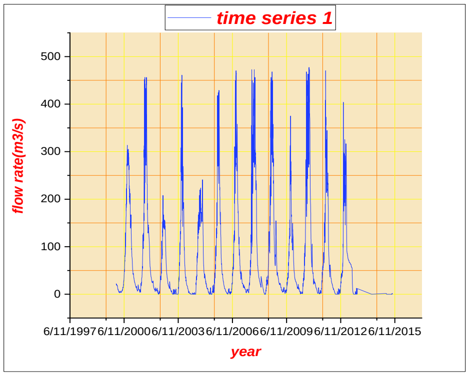
Appendix-G: Distribution sample estimates, standard errors, and correlation matrix of parameter, shape and scale values estimated using the Maximum Likelihood Estimation (MLE) Method.

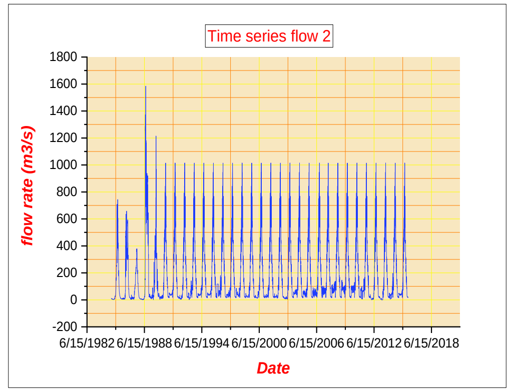
Sample estimates		Region-A				Region-B				Region-C			
				Correlation Matrix		Standard Error		Correlation Matrix		Standard Error		Correlation Matrix	
Distribution		Sample Est	Sd Error	mean	sd	Sample Est	Sd Error	Shape	Scale	Sample Est	Sd Error	Shape	Scale
Normal	mean	0.449	0.021	1	-2.7	0.123	0.011	1	0	0.16	0.01	0	1
	Sd	0.116	0.015	-2.65	1	0.064	0.007	0	1	0.041	0.01	0	1
Lognormal	mean	0.833	0.045	1	0	2.22	0.083	1	8.64	1.86	0.42	1	0
	sg log	0.255	0.032	0	1	0.483	0.058	8.64	1	0.246	0.03	0	1
Gamma	Shape	15.654	3.872	1	1	4.251	0.993	1	0.94	16.19	3.89	1	0.98
	Rate	34.843	8.759	0.99	1	34.52	8.569	0.94	1	100.8	24.6	0.984	1
Beta	Shape	7.64	1.877	1	0.9	3.685	0.856	1	0.92	13.35	3.2	1	0.97
	Scale	9.284	2.295	0.94	1	26.15	6.449	0.92	1	69.71	17	0.972	1
Cauchy	Shape	0.432	0.019	1	0	0.086	0.006	1	0.26	0.157	0	1	-0.23
	Scale	0.064	0.014	0.04	1	0.024	0.006	0.26	1	0.018	0	0.234	1
Weibull	Shape	1	0.336	1	0.3	2.089	0.27	1	0.34	3.825	0.18	1	0.34
	Scale	0.333	1	0.34	1	0.14	0.012	0.34	1	0.176	0.01	0.338	1
Exponential	Scale	2.226	0.393			8.199	1.392			6.229	1.07		
Uniform	Min	0.235				NA				0.101	NA		
	Max	0.779				NA				0.295	NA		
Logistic	Shape	0.44	0.019	1	0	0.114	0.011	1	0.15	0.157	0.01		
	Scale	0.064	0.009	0.04	1	0.036	0.005	0.15	1	0.022	0		

Appendix-H: lognormal distribution using MLE approach



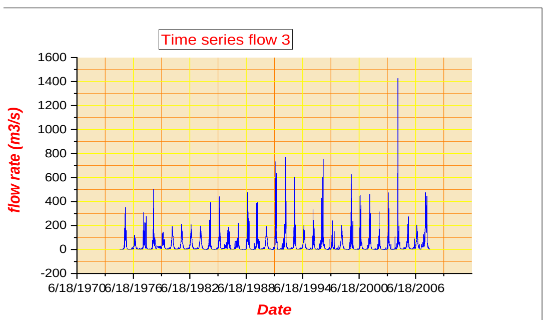
Appendix-I: time series flow of stations

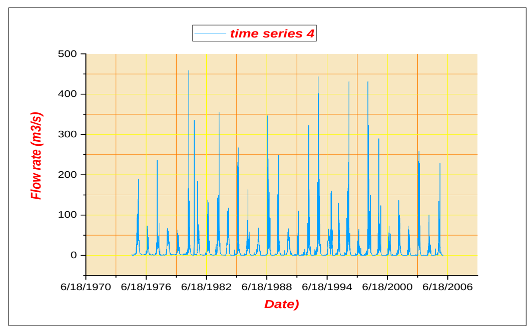




(a) time series flow of station one

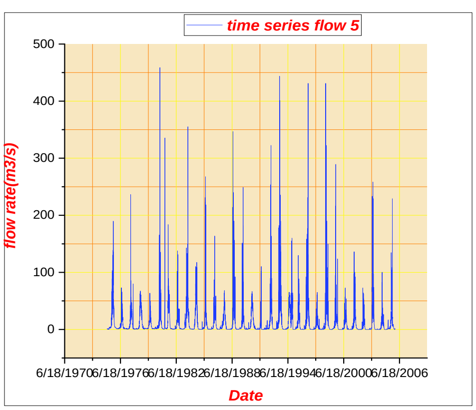
(b) time series flow of station two

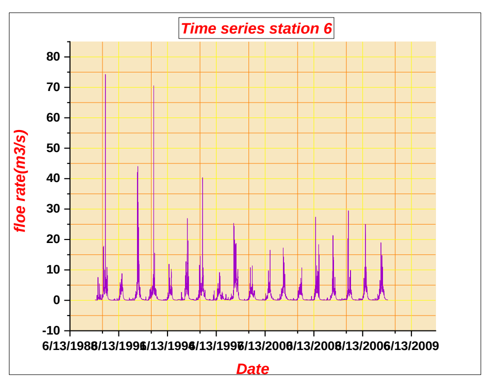




(c) Time series flow of station three

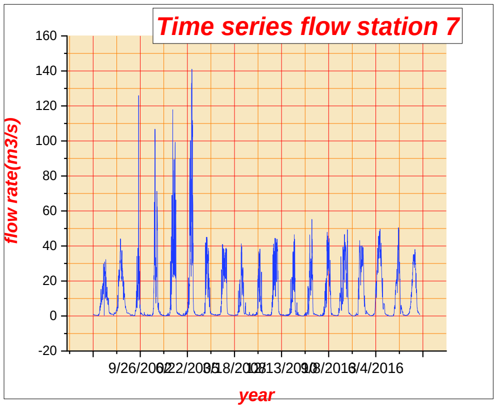
(d) time series flow of station four



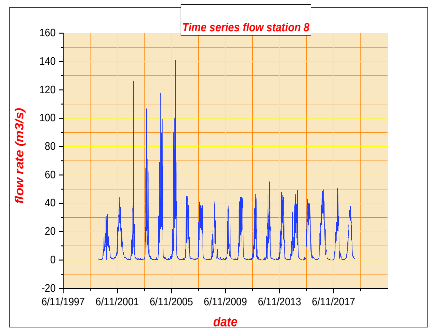


(e) Time series flow of station five

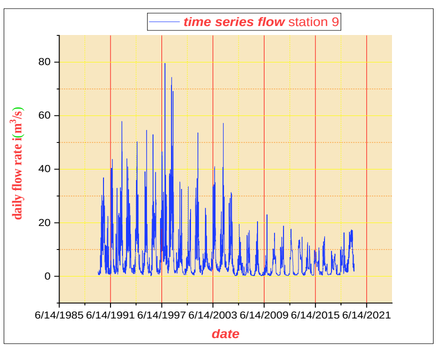
(f) time series flow of station six



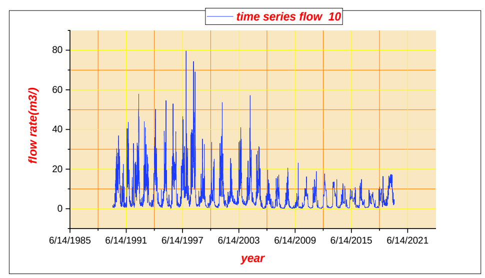
(g) Time series flow of station seven



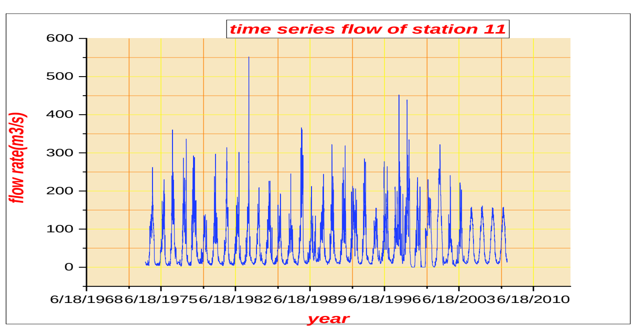
(h) time series flow of station eight



(i) Time series flow of station nine

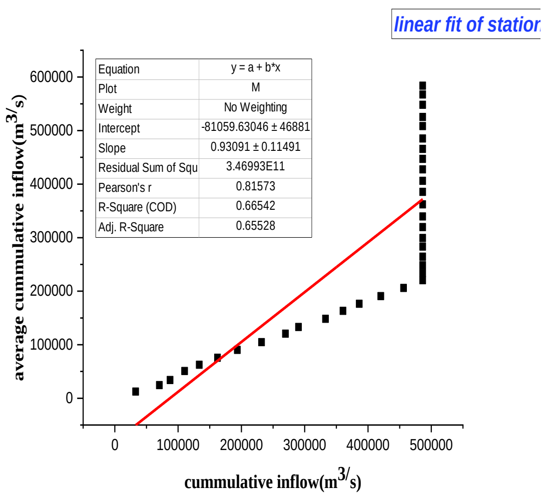


(j) time series flow of station ten

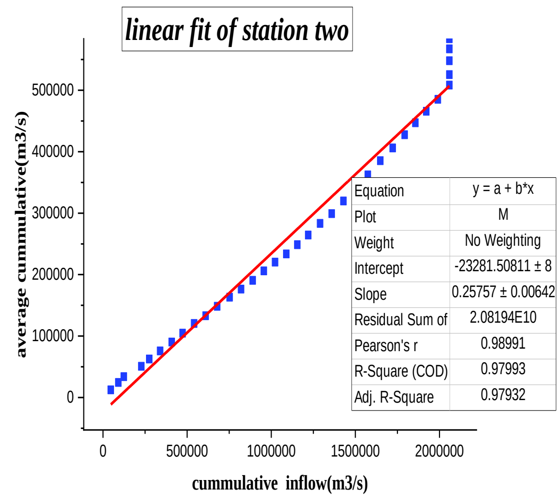


(k) Time series flow of station eleven

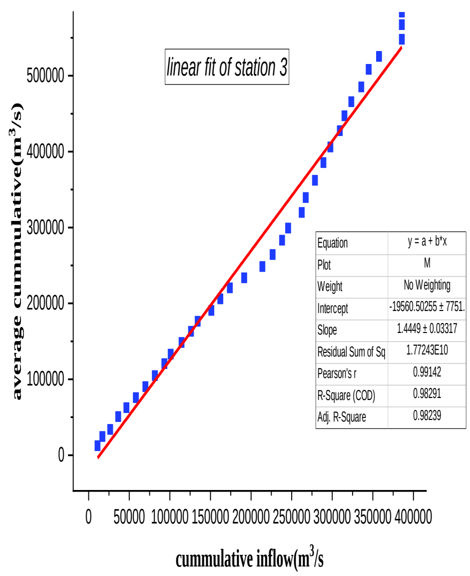
Appendix-J: Linear fit of stations



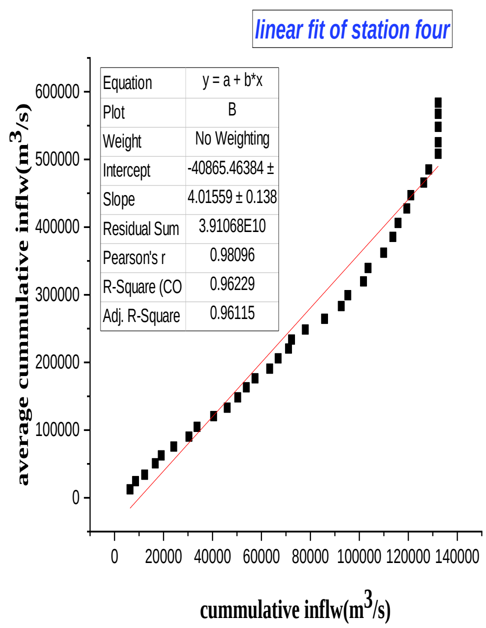
(a) station one



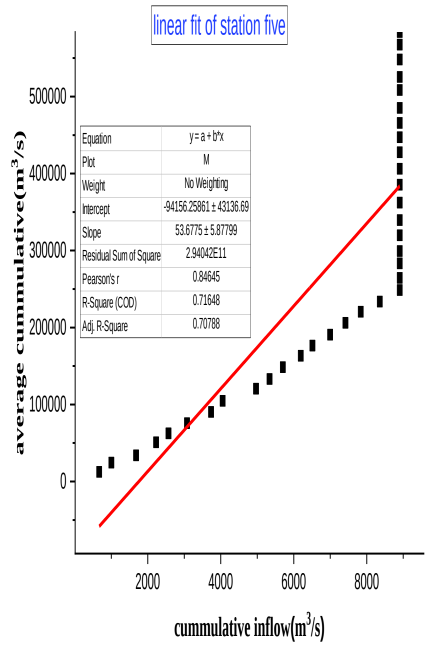
(b) station two



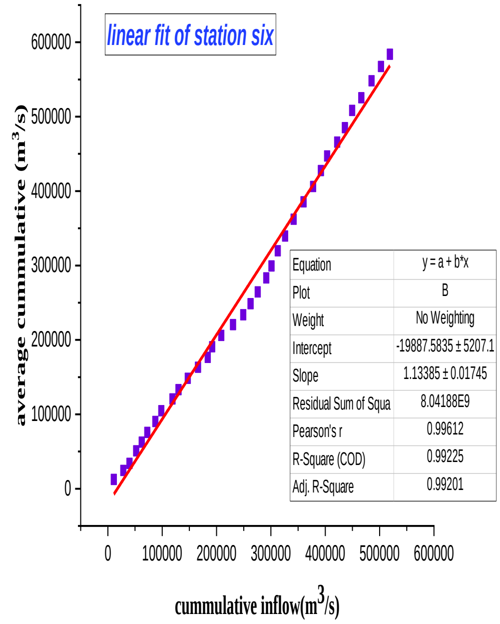
(C) station three



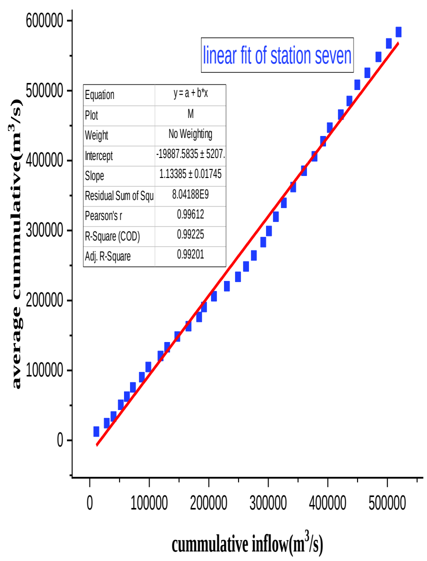
(d) station four



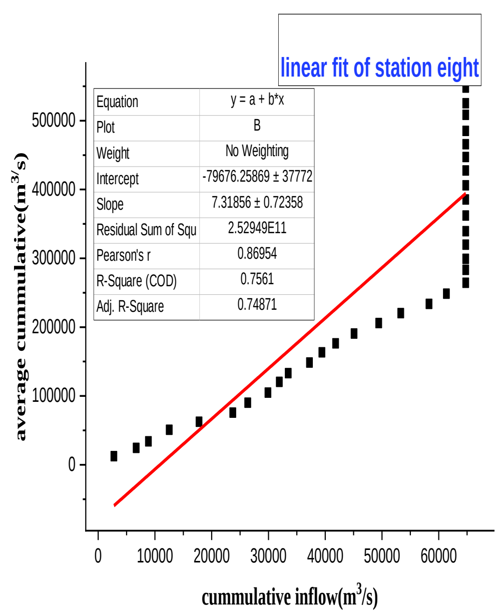
(e) Station five



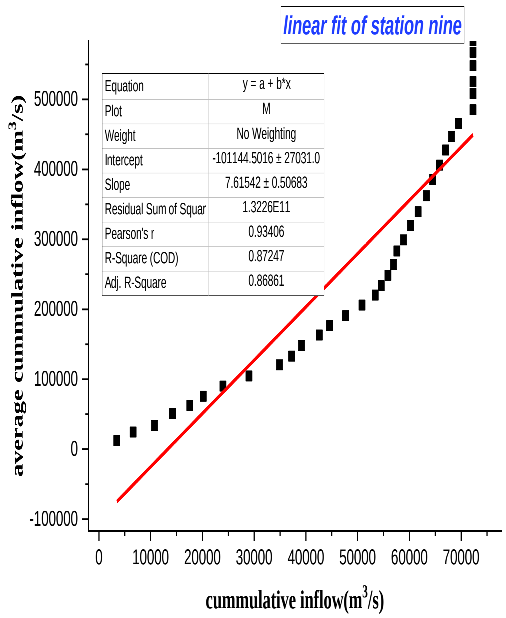
(f) Station six

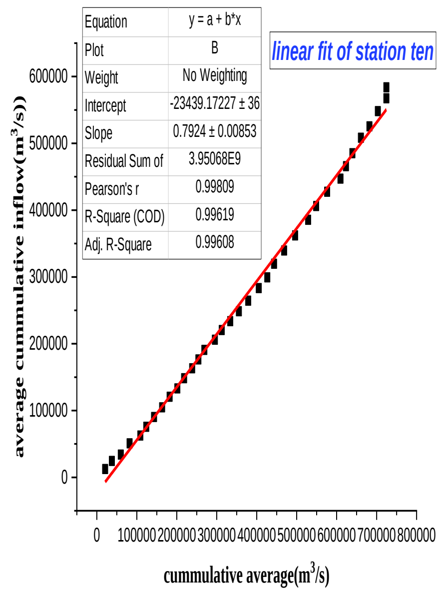


(g) Station seven

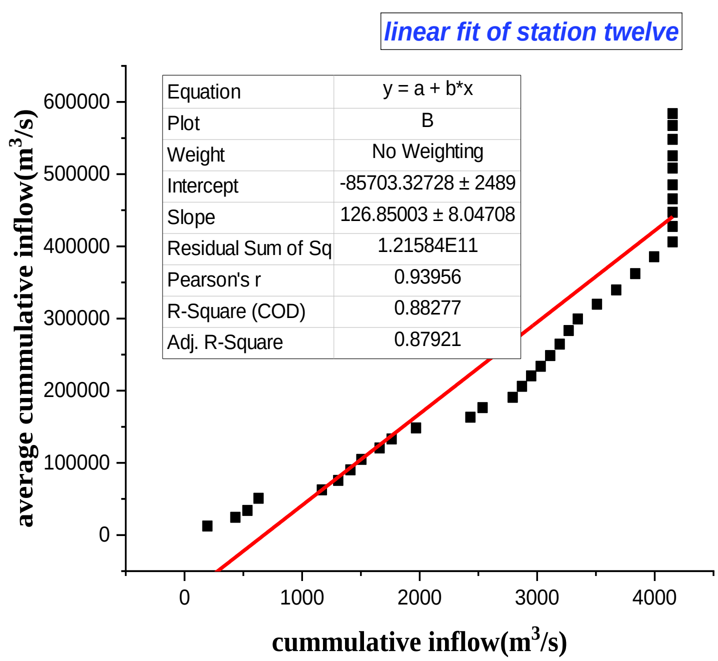


(h) Station eight





(i) Station nine (j) station ten



(k) Station eleven