



REGIONAL FLOOD FREQUENCY ANALYSIS UPSTREAM OF AWASH WITH THE CONFLUENCE OF KESEM RIVER

**By
Siraj Mekoya Yirefu**

A Dissertation submitted in partial fulfillment of the requirements for the degree of Master of Science (Engineering) of the Addis Ababa University

**Addis Ababa University
March, 2010**

CERTIFICATION

The undersigned certify that they have read the dissertation entitled: **Regional Flood Frequency Analysis upstream of Awash with the confluence of Kesem river** and here by recommend for acceptance by the Addis Ababa University in partial fulfillment of the requirements for the degree of Master of Science (Engineering).

Dr. Yilma Seleshi

(Advisor)

Date _____

Dr. Yonas Michael

(Internal Examiner)

Date _____

Ato Deksyos Tarekegn

(External Examiner)

Date _____

DECLARATION AND COPYRIGHT

Siraj Mekoya Yirefu. declare that this dissertation is my own original work and that it has not been presented and will not be presented by me to any other University for similar or any other degree award.

Signature: _____

This dissertation is copyright material protected under the Berne convention, the copyright Act 1999 and other international and national enactments, in that behalf, on intellectual property. It may not be reproduced by any means, in full or in part, except for short extracts in fair dealing, for research or private study, critical scholarly review or discourse with an acknowledgement, without written permission of the Directorate of postgraduate studies, on behalf of both the author and University of Addis Ababa.

ACKNOWLEDGEMENT

I would like to express my wholehearted gratitude to my advisors, **Dr. Yilma Seleshi**, for his priceless support in supervising my work and providing me, with his invaluable guidance, constructive comments and his encouragement from early age of my study.

I would like to acknowledge **Addis Ababa University** for the financial support made through the post graduate program of the University, which is available to me during both my course work and research time .My sincere appreciation goes to all staff members of **Ministry of Water Resource**, particularly Hydrology department, GIS department and Library, for their appreciable support in providing me hydrological data, digitized map and other reference materials.

The **National Metrological Service Agency** helped me in providing rainfall and climatic data of the region. Therefore, they are gratefully acknowledged.

Finally ,I wish to extend my gratitude to all staff members of **Addis Ababa University**, and my friends who helped me one way or the other, in carrying out my research through remarkable encouragement, advice, material support and collaboration in every aspect.

Abstract

Regionalization is the most viable way of improving flood quantile estimation in index flood method of frequency analysis. Systematic ways of estimating flood quantiles for the design of hydraulic structures for the cases of both gauged and ungauged catchments are presented. These have been carried out on Awash river upstream of Awash with the confluence of Kesem river.

The Sub- Basin has been regionalized into similar flood producing characteristics based on established flood frequency procedures. The basin was delineated in to five homogeneous regions. Accordingly, region one includes Berga, Holeta, Ginchi, Teji and Awash-Bello river catchments, Region two covers Melka-Hombola, Melka-Kountera, Akaki, Little-Akaki and Mojo catchments, Region three includes Awash-Below Koka dam, Wonji, Kelata and Nura-Hera, Region four includes Methara, Awash at Awash Station, Arba-Abomsa and Melka-Sedi catchments, Region five incorporates Kesem-Baka and Kesem-Awara Melka. All regions have shown satisfactory results for homogeneity tests. For the identified homogeneous regions three different best fitted distribution types and robust parameter estimation methods were selected to fit the standardized flow data for various return period. General extreme value type of distribution is found to be the best fitting distribution for region one. Whereas, Gama distribution for region two and four, Person-III for region three and five are selected for flood quantil estimation.

For all distributions types method of moment parameter estimation is more efficient. Based on this, regional flood frequency curves are developed for all regions using standardized flow data for the purpose of estimating flood quantiles in the ungauged catchments of the basin. This study is believed to benefit regional states, NGOs and the country at large in providing new information on flood quantiles for designing of hydraulic structures, culverts, bridges, dams, weirs, etc.

TABLE OF CONTENTS

Certification.....	i
Declaration and Copyright.....	ii
Acknowledgment.....	iii
Abstract	iv
Table of contents	v
List of table.....	viii
List of Figures.....	x
List of Annexes.....	xi
Acronyms	xii
1.INTRODUCTION	
1.1. Background	1
1.2. statement of the problem	2
1.3. Objective of the research	3
4.1. Description of the study area	4
1.4.1Awash river basin.....	6
1.4.2 Awash u/s of Kesem.....	6
1.4.3 Land use and Soil type	8
1.5 previous studies of the area	9
1.6. Outline of the research.....	10
2. LITREATURE REVIEW	
2.1. Flood frequency models	11
2.2 Regionalization	14
2.3 Homogeneity Test	15

2.4 procedure for selection of flood frequency distribution	16
2.5 Method of Parameter estimation.....	20
2.5.1 Comparison.....	21
2.5.2 Selection.....	22
2.6 Estimation of Index-Flood for Standardization.....	23

3. METHODOLOGY AND PROCEDURE

3.1. collection of important data for study	26
3.2. Consistency test.....	27
3.3. Independency test.....	28
3.4 Homogeneity test.....	29
3.5 Delineation of homogeneous regions	33
3.6 Selection of best fit distribution with proper method of parameter estimation.....	34
3.7. Quantile estimation and derivation of flood frequency curve.....	35

4. DATA COLLECTION AND ANALYSIS OF THE SUB-BASIN

4.1 Source and availability of data	36
4.2. Filling and extension of data	38
4.3 Consistency and Independence test of data	42
4.3.1 Independence and Stationarity test.....	42
4.4 Test of data for homogeneity of stations in the sub basin	44
4.5 Regionalization.....	48

5. RESULTS AND DISCUSSION

5.1 Results from consistency, independency and homogeneity tests.....	50
5.2. Selection of best fitted statistical distribution and method of parameters estimation.....	51
5.3 Selection of the most robust flood estimation procedure.....	53
5.4 Quantile estimation and derivation of flood frequency curve.	55
5.4.1 Determination of quantile estimates	55
5.4.2 Guidelines for determining of quantiles.....	55
5.4.3 Estimation of return period	56
5.4.4 Guidelines for Estimation of return period.....	56
5.4.5 Properties of quantile estimator.....	57
5.4.6 Evaluation of sampling and distribution error.....	57
5.4.7 Simulation test on quantile estimation	58
5.5 Derivation of dimensional and standardized flood Quantil for each region.....	65
5.6 Prediction of mean annual flood (MAF) for ungauged catchments.....	74

6. SUMMARY, CONCLUSION AND RECOMMENDATION

6.1. Summary and conclusion	77
6.2. Recommendation	79
7.REFERENCE	80
8.ANNEXES.....	82

LIST OF TABLES

Table 1.1 Major drainage Basins of Ethiopia.....	5
Table1.2 General Characteristics of stations and catchments.....	7
Table 2.1 Mathematical expression of analyzed statistical distribution for Annual maximum series.....	18
Table 4.1 -Showing the regression equation of stations for filling and extension	40
Table 4.2- Results of Independency test.....	43
Table 4.3 Results of homogeneity test	45
Table 4.4- Showing stations and their regional classification.....	49
Table 5.1 Combinations of Distribution with that of Parameters.....	52
Table 5.3- Standard error for Region -1.....	60
Table 5.4-Standard error estimation for Region -2.....	61
Table 5.5- Standard error estimation for Region-3.....	62
Table 5.6-Standard error estimation for Region -4.....	63
Table 5.7- Standard error estimation for Region -5.....	64
Table 5.8- Dimensional Quantiles for Region-1.....	65
Table 5.9- Dimensional Quantiles for Region-2.....	66
Table 5.10- Dimensional Quantiles for Region-3.....	66
Table 5.11- Dimensional Quantiles for Region-4.....	67

Table 5.12- Dimensional Quantiles for Region-5.....67

Table 5.13-Standardized Quantiles for Region-168

Table 5.14- Standardized Quantiles for Region-2.....68

Table 5.15- Standardized Quantiles for Region-3.....69

Table 5.16- Standardized Quantiles for Region-4.....69

Table 5.17-Standardized Quantiles for Region-5.....70

Table 5.18-Derived regression equation for each regions to predict Mean Annual
Flood (MAF) for ungauged catchment.....74

LIST OF FIGERS

Table 1.1 Major drainage Basins of Ethiopia.....	5
Table1.2 General Characteristics of stations and catchments.....	7
Fig.4.1-Graphs comparing measured and generated flow data by regression for the purpose of filling and extension.....	39
Fig 4.4: Moment ratio distribution of gauged stations of the homogeneous regions.....	47
Fig 5.1- Standardized Growth curve of Region-1 and each station.....	71
Fig 5.2- Standardized Growth curve of Region-2 and each station.....	71
Fig 5.3- Standardized Growth curve of Region-3 and each Station.....	72
Fig 5.4- Standardized Growth curve of Region-4 and each station.....	72
Fig 5.10- Standardized Growth curve of Region-5 and each station.....	73

LIST OF ANNEXES

8.1 - ANNEX - I

Growth curve of each station including the regions

Dimensional flow (Q_T)	83
Standardized flow (Q_T)	90

8.2 -ANNEX-II

Map of the study area

Hydrometric gauging stations in the study area (Source MoWR)	96
Homogeneous Regions division.....	97

8.3-ANNEX-III

Metrological stations used for the thesis	98
---	----

ACRONYMS

AM- Annual Maximum

CV- Coefficient of variation

Cs- Coefficient of skewness

Ck- Coefficient of Kurtosis

EV1- Extreme value type one

FFA- Flood Frequency Analysis

GEV- Generalized Extreme Value

GIS- Geographical Information System

MAF- Mean Annual Flood

MLM- Method of Maximum Likelihood

MOM- Method of Ordinary Moment

MoWR - Ministry of Water Resource

NMSA - National Meteorological Service Agency

P-III- Person three distribution

PD - Partial Duration

POT- Peaks Over a Threshold

PWM- Probability Weighted Moment

RFFA- Regional Flood Frequency Analysis

SEE- Standard Error of Estimate

U/S- Up Stream

WMO -World Meteorological Organization

CHAPTER - 1

INTRODUCTION

1.1 Background

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

The primary objective of frequency analysis is to relate the magnitude of extreme events to their frequency of occurrence through the use of probability distributions (Chow et al, 1988). Data observed over an extended period of time in a river system are analyzed in frequency analysis. The data are assumed to be independent and identically distributed. The flood data are considered to be stochastic and may even be assumed to be space and time independent. Further, it is assumed that the flood have not been affected by natural or man-made changes in the hydrological regime in the system.(Rao, 2000)

The use of regional information to estimate flood magnitudes at sites with little or no observed data has become increasingly important because many projects which require design flood information are located in areas where observed flood data are either missing or inadequate. In the analysis attention must be given to the at-site data since they are the bases for regional information.

1.2 Statement of the Problem

The main feature of a flood, from the water management point of view is its interface with human activities. The interface is measured in terms of actual and potential economic losses and danger to human life the purpose of flood estimation is to assess the magnitude and frequency of this interference. This flood frequency analysis provides vital information for the planning and design of many hydraulic structures such as culvert, bridge, reservoir, spillway, road embankment and dikes and for the risk of most hydraulic assessment in flood plain used. (Admasu,1989)

Regional flood frequency analysis is becoming an important subject because of most structures are constructed in areas where recorded flood data are either missing or inadequate.

Moreover, most of Ethiopian river basins have sparse network of observation sites with short record length of observed flow that makes the use of single site analysis to estimate design parameters at many potential project sites unreliable. There are also river basins, which are totally ungauged. In such instant, the design of any structure within the region is difficult unless data transposing from gauged site is done which requires a lot of effort. The work described here will attempt to solve these problems using the regional approach whereby statistically homogeneous regions are identified and the parameters of choosing distribution are estimated from the regional averages so that the flow quantile for the ungauged catchment within that region can easily be computed from the regional curve developed. So that the necessity of this study arises from the weight to improve site – specific estimates based on limited data and to make inference about ungauged catchments.

1.3 Objective of the Research

The objective of this paper is to find an appropriate procedure for analysis of flood frequency in the basin. The development of water resources project in the basin needs proper guidance on how to estimate flood quantiles. This study indicates the way towards proper solution.

The specific objective of the thesis are :

- ❖ To identify and delineate hydrologically homogeneous region.
- ❖ To identify the best-fit statistical distributions to the data of each region.
- ❖ To choose suitable parameter estimation method and procedure
- ❖ To establish regional frequency curves for the delineated homogeneous region
- ❖ To derive a relationship between mean annual flood and catchment characteristics that can be used for the ungauged catchments within the region.
- ❖ To develop method of estimating quantiles for the ungauged catchments from regional frequency curve for the design of hydraulic structures within the region.

1.4 Description of the Study Area

Ethiopia is located in the eastern part of Africa between 3° 30' and 18° 12' N latitude and 32° 42' and 48° 12' E longitude. The country has great geographical, topographical and climatological diversity: From high rugged mountains to deep gorges; from lowest altitude at about 120m below sea level to highest altitude of 4600m above sea level; from 2000mm high annual rainfall to 200mm of low annual rainfall. Besides, the Great Rift Valley divides the country in two parts forming the eastern and western high lands. (Master Plan of Awash –By Halcrow)

The country is endowed with large amount of water resource potential. It has 12 major drainage basins as listed in the table below

Table 1.1 Major drainage Basins of Ethiopia

No	Basins	Locations	Area (Km ²)	Coverage in (%) of the total area
1	Mereb Basin	Extreme North	6065	0.5
2	Tekeze Basin	North	81034	7.2
3	Denakel Basin	North	66489	5.9
4	Blue Nile Basin	North - West	192953	17.1
5	Awash Basin	Central east	113604	10.1
6	Aysha Basin	Extreme East	4717	0.4
7	Ogaden Basin	East	82157	7.3
8	Wabi Sheble Basin	South East	207497	18.4
9	Genal Dawa Basin	South	172681	15.3
10	Rift Valley Lakes	Central West	51664	0.6
11	Omo – Gibe Basin	South West	74912	6.6
12	Baro – Akobo Basin	West	73958	6.5
TOTAL			1127730	100

This research will concentrate only on the Awash Basin especially on the upstream of Awash with the confluence of Kesem river.

1.4.1 Awash river basin

The Awash river basin rises at an elevation of 3000m above sea level over the central high land of Ethiopia about 150km west of Addis Ababa. The river flows generally north east wards along the Rift valley and terminates in the lake Abe at an elevation of 250m above sea level near Djibouti. This high variation of elevation difference exposes the basin for high flood to occur due to altitude variation, topographic and geographical diversity.

The Awash basin covers a total area of 110000 km² of which 64000km² most of which comprises the so-called eastern catchments, drain in to a desert area and do not contribute to the main river course (Awash master plan by Hal crow)

The Awash basin has about four major drainage sub- basins the research focus in the upper and part the middle sub-basin (upstream of Kesem) and the Awash basin has 67 gauging stations out of this 35 gauging stations found in the study area. Only parts of them get a good record of data and can be preferable for the analysis.

1.4.2 Awash u/s of Kesem

The catchments of Awash river U/S of Kesem are found in the highlands of the central Ethiopia. The Koka reservoir is located in the upper reaches of the Awash basin. The area comes under the influence of the ITCZ and the seasonal rainfall distribution results from the annual migration of the ITCZ. The mean annual rainfall is 1200mm and reaches 1500mm at eastern high lands of Addis Ababa. The catchments receive its maximum rainfall during June to September and amounts to 70% to 75% of the annual rainfall. The second rainy period covers the period from February to May. This high variation in elevation and rainfall leads for the occurrence of high flood in the area (Awash master plan by Hal crow)

The drainage area of Awash U/S of Kesem is 24413km². The major rivers that flow in the sub basin are Awash, Mojo and Kesem rivers. The other major tributaries are Akaki,

Holeta, Berga, Kelata and Arba rivers. In the sub catchment there are 35 gauging stations that record the flow and lake depth in the sub basin out of this gauging stations only 20 are used in this research.

Table1.2 General Characteristics of stations and catchments

Stat No	Station Name	Catchment Area(Km2)	UTM-North(m)	UTM-East(m)
31001	Berga near Addis -Alem	249	997479	428884
31002	Holeta	119	1003757	446240
31003	Teji near Asgora	663	974683	429482
31004	Akaki	884	981176	476362
31012	Awash at Melka -Kuntura	4456	962272	456660
31013	Awash at Melka-Hombola	7656	926171	456660
31014	Mojo	1264	951552	522480
31015	Kelata	747	915053	543822
31016	Awash at Wonji	11690	938001	525361
31017	Awash below koka	111219	936637	518127
31019	Kesem - Beke	50	1013372	507656
31020	Awash Near Belo	2569	978155	435153
31021	Little Akaki at Akaki	131	999964	465908
31033	Awash at Ginchi	76	997173	404697
32002	Arba - Abomsa	140	939481	591329
32003	Methara	16417	978626	601325
32004	Awash Station	19111	994848	629815
32005	Kesem Awara- Melka	3113	1012050	604532
32015	Awash Melka - Sedi	21520	1017145	623210
32017	Nura -Hera	14173	942607	563362

1.4.3 Land use and Soil type

The land use condition in Awash catchments upstream of Kesem includes mainly of cultivated agricultural land, grassland, forest land, rural and urban settlements. It is estimated that 67% is intensively cultivated, 25.5% is moderately cultivated, 4.5% is bush land or shrub land or wooded grassland, and 3% is urban area and alpine vegetation. Strictly speaking, even the land use within the sub catchment is diverse. In the upper most part where there is high rainfall, land use is complete in May with barley and teff. Steeper slopes are heavily wooded with natural acacia and eucalyptus. On the lower most part, however, rainfall is too unreliable and the sparse dry acacia scrub gives way to wide stretches of bare ground with clumps of coarse grass and occasional thickets of acacia. The soil type in the study area is diverse. The most common soil types are Clay, Sand, Clay-Loam, Silt-Clay -Loam, Sand-Clay, Silt-Clay (Paulose, 1989). Land use and soil type have a direct impact on the flood amount, speed and potential to create damage that the study should give attention for land use and land cover of the sub basin.

1.5 Previous study of the area

For the Awash River basin especially upper and middle sub-basins different researchers and organizations have stated different ideas on the protection of flood in the area but the flood frequency analysis with adequate and updated data hasn't fully studied in the sub-basin. Normally the area is susceptible to high flood due to the natural topography which varies from the very mountainous to the immediate very plate land, which greatly enhances the creation of flood even with less rainfall. Looking at this problems different bodies have high lightly revised the area concerning the flood protection but not on the detail flood frequency analysis that is why the problem has stayed without any solution up to now. From many of the papers written by different researchers such as Admasu Gebeyehu in 1989, in his study of regional flood frequency analysis of the whole country, he has tested the flood frequency of the area with limited data and information, Leulseged from Ministry of water resource has written some research paper on the upper Awash basin on flood frequency analysis and the consultant Halcraw has studied the flood damage of the area, all of them didn't analyzed the flood frequency of the area in detail. Now the target of this paper is to develop flood frequency curve in the upper and part in the middle Awash of major rivers and tributaries.

1.6 Outline of the research

The thesis is organized into six chapters from introduction to the conclusion and recommendation. The first chapter presents about the introduction, background of the problem and the objective of the paper. The second chapter discusses about the literature review and state of art related to RFFA, in this part all the high light of the paper will be discussed. The third chapter illustrates the methodology and procedure to be applied in the paper, from data collection to the result of the analysis, the fourth part of the thesis collection of data and analysis of data will takes place. The fifth chapter discusses the result of the above testes on the distribution selection, parameter estimation, quantile estimation and derivation of flood frequency curve, this part of the thesis presents the result of the analysis and evaluates the result with different physical realities in addition it contains estimation of mean flow for unguaged catchments. The last chapter of the paper concentrates on the conclusion and recommendation of the thesis which concludes and recommends on the result of the analysis.

CHAPTER -2

LITERATURE REVIEW

2. 1 Flood Frequency Models

In flood frequency analysis the objective is to determine a Q-T relationship at any required site along a river. At any river site it is usually assumed that nature provides a unique Q-T relationship and that Q is a monotonically increasing function of T. In order to estimate this natural Q-T relation from a good quality continuous hydrometric record of N years duration, it is necessary to resort to a statistical or stochastic model of the continuous hydrograph, which retains information in the hydrograph relevant to the Q-T relation, and discard the rest.

Two such models are:

- ❖ Annual maximum series model,(AM)
- ❖ Partial duration series (or peak over a threshold) model (PD)

In flood frequency modeling the problems related to the following points have to be solved (Cunnane 1989);

- ❖ Choice of model type (AM, or PD)
- ❖ Choice of distribution to be used in the chosen model
- ❖ Choice of method of parameter and quantile estimation

It should be noted that two separate aspects of such choice are important. These are the descriptive and predictive properties of the chosen method. The descriptive property' relates to the requirements that the chosen distribution shape resembles the observed sample distribution of floods and that random samples drawn from the chosen model distribution must be statistically similar to the properties of real flood series, the

predictive properties relates to the requirement that quantile estimates are robust with small bias and standard error (Cunnane, 1989),

2.1.1 Relative advantage of the two models

a) Annual maximum series model (AM)

Cunnane, 1989 has stated that a series of annual maximum flood is assumed to form a random sample from stationary population in which Q is a random variable with distribution $PR(Q < q) = F(q)$. The variate values with exceedence probability $1/T$ is said to have return period T . Denoting this value Q_T , it is such that:

$$1 - F(Q_T) = 1/T$$

In the annual maximum (AM) flow series, only the peak flow in each year of record is considered, that may involve some loss of information.

b) Partial duration series (peak over a threshold) model (PD)

In this model most of the flow hydrograph is disregarded and the hydrograph is viewed as a series of randomly spaced flood peaks of random magnitude. For case of statistical modeling and also for case of identification of the values, which form the series, only the series of peak exceeding an arbitrary threshold q_0 are considered. In particular, each of these showed that if the number of flood peaks exceeding some value q_0 (a threshold value) in some interval of time such as a year has a Poisson distribution with parameter λ , the number of events exceeding a great value \bar{q} is also Poisson distributed with parameter $\bar{\lambda} = \lambda P$ where $P = PR(q > \bar{q} / q > q_0)$. Here P is a conditional probability, being the proportion of all peaks exceeding q_0 which also exceeds \bar{q} (Cunnane, 1989) In partial duration series all peaks above a certain base value are considered. The base is usually selected low enough to include at least one event in each year (Rao and Hammed, 2000)

C) Statistical efficiency of estimates of Q_T by each model

Denoting the estimates of Q_T obtained by AM method as Q_T and that obtained from the same hydrometric record by the PD method as $\overline{Q_T}$. it is usually observed that these two estimates are unequal. Furthermore the sampling variance of Q_T is not equal to that of $\overline{Q_T}$ i.e. $\text{Var}(Q_T) \neq \text{Var}(\overline{Q_T})$. From a statistical point of view that method which has the smallest sampling variance enjoys an advantage. Under certain common assumption Cunnane (1989) examined the relative values of $\text{Var}(Q_T)$ and $\text{Var}(\overline{Q_T})$ and found that $\text{Var}(Q_T) < \text{Var}(\overline{Q_T})$ provided $\lambda < 1.65$ where λ is the mean number of peaks per years included in the PD series. Where $\lambda > 1.65$ the opposite was true. This shows that the AM method is statistically more efficient than the PD method when λ is small but less efficient when λ is large. In many practical situations the assumptions of the PD model may not be valid if λ is increased to too high a level, certain if $\lambda > 1.65$ (Cunnane 1989).

Therefore, to avoid the problem of dependency on data, annual maximum (AM) series model has been selected in this study. In addition to this, (AM) series is widely and commonly used model by different researchers for the purpose of flood frequency analysis (Cunnane, 1989)

2.2 Regionalization

Regionalization, in the context of RFFA, refers to identification of homogeneous regions through homogeneity test and selection of appropriate frequency distribution for the identified region and stations. There is no universally accepted objective method of regionalization. This is due to the complexity of factors that affect the generation of floods. Several attempts have been made by different researchers to identify hydrologically homogeneous regions based on either geographical considerations or flood data characteristics, or a combination of both. (Mkhandi,1985).

In regional flood frequency analysis (RFFA), the established curve of flood variate versus return period can be used for estimating flood quantiles at any site within the region. For ungauged sites, RFFA involves the analysis of flood records of all gauged sites in the region, summarizing each record by few representative statistical values calculated from it and then finding a relation ship among these statistical values and measurable catchment characteristics. This will help to express the ungauged flood variate intermes of the gauged flood variate (Cunnane, 1985a).

Regionalization can be done based on geographic proximity, physiographic & climatic characteristics of the catchments. (Admassu, 1989) Further advancements in the field of RFFA have led to a better approach other than geographic space. Recently, researchers based their division of homogeneity by analyzing the statistical characteristics of flood data of different stations within the basin. (Rao & Hammed, 2000) Finally, the delineated regions have to be checked for their homogeneity using different homogeneity tests.

2.3 Homogeneity Test

Regional flood estimation methods are based on the premises that standardized flood variate, such as $X=Q/E$ (Q) has the same distribution at every site in the chosen region. In particular $C_v(x)$ and $C_s(x)$ are considered to be constant across the region. Serious departures from such assumptions could lead to biased flood estimates at some sites. Those catchments whose C_v and C_s value happen to coincide with the regional mean values would fortuitously not suffer such bias. Nevertheless if the degree of heterogeneity present is not too great its negative effect may be more than compensated for the larger sample of sites contributing to parameter estimates. Thus X_t estimated from M sites, which are slightly heterogeneous, may be more reliable than X_t estimated from a small number, say $M/3$, more homogenous sites, especially if flow records are short.

The importance of homogeneity has been demonstrated by (Hosking,1985). Homogeneity implies that region have similar flood generating mechanism A more specific definition of a homogeneous region is that region which consists of sites having the same standardized frequency distribution form and parameter.

At least five categories of questions arise in this context (Cunnane, 1987)

1. Are flood frequency behaviors of anyone of M sites in a region; with AM records available, inconsistent with that of the remainder of the group?
2. Are geographically defined regions better or worse than regions obtained by partitioning the catchments characteristics data space?
3. How can a large group of catchments be divided in to homogenous sub group of regions?
- 4.. How can un-gauged catchments be allocated to one of a number of pre selected homogenous regions?
5. What degree of departure from regional homogeneity can be tolerated in a flood quantiles estimation procedure?

2.4 Procedures for Selection of Flood Frequency Distribution

Choice of distribution for AM series has received widespread attention. In many countries the selection of AM distribution is actually not made in any objective manner and that the choice of distribution is argued in a general manner as follows:

The choice of distribution would be: (Cunnane, 1989)

- ❖ Widely accepted
- ❖ Simple and convenient to apply
- ❖ Consistent, flexible or robust (low sensibility to outliers)
- ❖ Theoretically well based
- ❖ Documented in the guide

No special method of parameter estimations referred and the graphical method is used as frequently even more used as any other method. The choice of distribution is influenced by many factors such as method of discrimination between distributions, method of parameters estimation, the availability of data etc. the method of parameter estimation goes parallel whit distribution selection.

There are many distributions that have been suggested for AM series models. Some of them are (Cunnane, 1989)

a)Normal and related distributions

- ❖ Normal distribution
- ❖ Log normal two parameter distribution
- ❖ Log normal three parameter distribution

b) The Gamma distribution

- ❖ Exponential distribution
- ❖ Two parameter Gamma distribution
- ❖ Pearson three distribution
- ❖ Log Pearson two distribution

c) Extreme value distribution

- ❖ Generalized extreme value distribution
- ❖ Extreme value type I distribution
- ❖ Extreme value type II distribution
- ❖ Weibull distribution

d) Wake-by distribution

- ❖ Five parameter wake-by distribution
- ❖ Four parameter wake-by distribution
- ❖ Generalized pareto distribution

e) Logistic distribution

- ❖ Log-logistic distribution
- ❖ Generalized logistic distribution

List of tabulated distribution types and their formula will be presented below.

Table 2.1 Mathematical expression of analyzed statistical distribution for annual maximum series

Name of Distribution	Distribution function F(x)	Variate and parameter ranges
Normal Distribution (N)	$F(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{1}{2\sigma^2}(x-u)}$	$-\infty < x < \infty$ μ and σ are parameters
Two Parameter Lognormal distribution(LN2)	$F(X) = \frac{1}{\sqrt{2\pi}ax} \exp\left\{\frac{-1}{2}\left(\frac{\log x - b}{a}\right)^2\right\}$	$0 < x$
Three parameter Lognormal distribution	$F(x) = \frac{1}{\sqrt{2\pi}a(x-m)} \exp\left\{\frac{-1}{2}\left(\frac{\log(x-m)-b}{a}\right)^2\right\}$	$m < x$
Exponential Distribution (EXP)	$F(x) = \frac{1}{a} \exp\left(-\frac{x-m}{a}\right)$	$m < x$ (i.e. P-III with $b = 1$)
Two parameter Gamma distribution (Gam-2)	$F(x) = \frac{\left(\frac{x}{a}\right)^{b-1}}{a/\Gamma(b)} \exp\left(-\frac{x}{a}\right)$	$0 \leq x$ if $a > 0$ $x \leq 0$ if $a < 0$ (i.e. P-III with $m = 0$)
Person-III Distribution (P-III)	$F(x) = \frac{\left(\frac{x-m}{a}\right)^{b-1}}{a/\Gamma(b)} \exp\left\{-\frac{x-m}{a}\right\}$	$m \leq x$ if $a > 0$ $x \leq m$ if $a < 0$
Log Person –III distribution (LP-III)	$F(x) = \frac{\left(\frac{z-c}{a}\right)^{b-1}}{x/a/\Gamma(b)} \left\{-\frac{z-c}{a}\right\}$	$c < z < \infty$ $e^c < x < \infty$ $a > 0$
		$-\infty < z < c$ $0 < x < e^c$ $a < 0$
Generalized Extreme value distribution (GEV)	$F(x) = \exp\left\{-\left[1 - k\left(\frac{x-u}{a}\right)\right]^{1/k}\right\}$	$\alpha > 0$ $u + \frac{\alpha}{k} \leq x \leq \infty$ if $k > 0$ $-\infty < x \leq u + \frac{\alpha}{k}$ if $k < 0$
Extreme value Type-I distribution (EV1)	$F(x) = \exp\left\{-\exp\left(-\frac{x-u}{\alpha}\right)\right\}$	$-\infty < x < \infty$ $\alpha > 0$

Name of Distribution	Distribution function F(x)	Variate and parameter ranges
Five parameters Wake by distribution (WAK-5)	$X = m + \alpha[1 - (1 - F)^b] - [1 - (1 - F)^{-d}]$ Where F=F(x)	Note: It is analytically defined only in inverse form
Four parameters Wake by distribution	$F(x) = \frac{\alpha}{\beta} [1 - (1 - F)^\beta] - \frac{\gamma}{\delta} [1 - (1 - F)^{-\delta}]$	
Log Logistic distribution (LLg)	$F(x) = \left\{ 1 + \left[\frac{x-a}{b} \right]^{-1/c} \right\}^{-1}$	$x > a, c > 0, b > 0$
Generalized logistic distribution (GLg)	$F(x) = \left\{ 1 + \left[1 - \gamma \left(\frac{x-\alpha}{\beta} \right) \right]^{1/\gamma} \right\}^{-1}, \gamma \neq 0$ $F(x) = \left\{ 1 + \exp \left[-\frac{x-\alpha}{\beta} \right] \right\}^{-1}, \gamma = 0$	$\gamma < 0, \alpha + \frac{\beta}{\gamma} \leq x < \infty$ $\gamma > 0, -\infty < x < \alpha + \frac{\beta}{\gamma}$ $-\infty < x < \infty$

2.5 Method of Parameter estimation

In the past only the ordinary methods of moments (MOM) was mentioned for parameter estimation. Now numerous parameter estimation procedures have been proposed and studied from those MOM , PWM and ML are the most efficient method of parameter estimation available due to their smaller error in quantile estimation.

Methods of estimation includes:

- a) Probability weighted moments (PWM)
- b) Method of moment (MOM)
- c) Maximum likely hood (ML)

a) Method of probability weighted moments (PWM)

PWMs are useful in deriving expression for the parameters of distributions whose inverse forms $X=X(F)$ can be explicitly defined. In particular they allow parameter estimates to be obtained for distributions. Methods of parameter estimation are obtained in this method by equating moment of the distribution with the corresponding sample moment of observed data. For a distribution with parameter K , the first K sample moments are set equal to the corresponding population moments. The resulting equation is then solved simultaneously for the unknown parameters. Parameter estimation by PWM, which is relatively new, is as easy to apply as ordinary moments is usually unbiased and is almost as efficient as method of maximum likelihood (ML).

Indeed in small samples PWM may be as efficient as ML. With a suitable choice of distribution PWM estimation also contributes to robustness and is attractive from that point of view (Cunnane, 1989).

b) Method of Moment (MOM)

It is one of the most commonly used methods of estimating parameters of a probability distribution. The estimates of the parameters of a probability distribution function are obtained by equating the moments of the sample with the moments of the probability distribution function. It provides simple calculation, but higher order moment estimates are biased (Wallis, et. al. 1974). Parameter estimation by MOM is known to be biased and inefficient especially with three-parameter distribution but it is more preferable for two parameter distribution types.

c) Method of Maximum Likelihood (ML)

Estimation by the maximum likelihood (ML) method involves the choice of parameter estimates that produce a maximum probability of occurrence of the observations. The parameter estimates that maximize the likelihood function are computed by partial differentiation with respect to each parameters and setting these partial derivatives equal to zero and finally solve the resulting set of equations simultaneously. The equations are usually complex that can only be solved by numerical techniques. As a result of this difficulty, the solution set may not properly found (Cunnane, 1989).

2.5.1 Comparison

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

2.5.2 Selection

The selection of a distribution for flood frequency analysis goes with the selection of the method of parameter estimation. Parameters estimated by any of the above methods are subject to sampling errors. While a method may be efficient for one distribution it is not necessarily efficient for other distributions.

Therefore, to select the most robust flood estimation procedure:

- ❖ **Descriptive ability tests and**
- ❖ **Predictive ability tests have to be applied.**

More detail expression of these two methods will be listed in the next chapter.

2.6 Estimation of Index-Flood for Standardization

The application of regionalization requires a two-step procedure (D.Bacchiola, 1999):

- ❖ Identification of the homogeneous region for which a common probability model of maximum annual flow can be adopted to accommodate normalized flow
- ❖ Searching for the appropriate index flood estimation in the examined river site.

The estimation of the index flood plays a major role in design flood prediction, and it requires merging statistical and physical hydrology concept to reduce the present uncertainty. The basic idea behind the index-flood method is to increase the reliability of the frequency characteristics within a region. If, within a hydro logically homogenous area, a number of hydrometric stations have been operating and recording the effect of the same meteorological factors then a combination of these records will provide, not longer record, but a more reliable record. The index-flood method is based on the hypothesis that flood from different catchments with in a region normalized by their mean annual flood come from a single distribution. An essential prerequisite for this procedure is the standardization of the flood data from sites with different flood magnitudes. The most common practice is to standardize data, i.e., division by an estimate of the at-site means (Admasu, 1989).

$X_T = \frac{Q_T}{Q_T}$ Is the index flood, which is the ratio of normal flood to mean?

The parameters of the distribution of X_T are obtained from the combined set of regional data. If at-site data are not available, the index-flood can be predicted from a regionally derived empirical equation. The form of the relation (empirical equation) for estimating the index flood depends on the amount of physical and climatic data available for developing it.

The other component of the index-flood procedure is the standardized quantile. The accuracy of the standardized quantile estimate depends mainly on:

- ❖ The method of grouping catchments for homogenous regions and testing for homogeneity.
- ❖ The type of probability distribution and the method of parameter estimation and quantile estimation.

Procedure for index-flood determination

- List data of each gauging stations
- Select a common period of record
- Exclude stations less than 5 years of record
- The base period with maximum number of station year is selected
- Missing data may be filled in by inter-station correlation
- Not use filled data directly
- The index-flood method computes the return period, T , for each record events and stations. (G. w.Kite. 1985)

CHAPTER -3

3. Methodologies and procedure

Regional flood frequency analysis is a methodology of using at-site and regional information to predict the flood magnitude and its frequency for water resource planning and management and design of hydraulic structures.

The methodology chosen in this study is index flood method for at site and regional data analysis and it comprises data preparation, testing of data of the stations for homogeneity, selection of frequency distribution, method of parameter estimation and quantile estimation.

Generally the study involves the following procedure:

1. Collection of important data for the study such as hydrological data, metrological data, topographical and digitized map of the sub-basin.
2. Checking of data for quality, continuity, consistency and independency
- 3 Computation of statistical parameters of selected stations within the Sub-basin
4. Carry out homogeneity test for the stations in the region.
5. Delineation of homogeneous regions
6. Selection of frequency distribution for the determination of the quantiles.
7. Selection of parameter estimation method for the selected distribution.
8. Quantile estimation.
9. Derivation of regional and at site flood frequency curve.

3.1. Collection of important data for study

Important data has been collected from different institutions. The data have been collected as a soft copies, hard copies and maps, Hydrological data and digitized map of the sub-basin were collected from the Ministry of Water Resource, from the department of Hydrology and GIS. Meteorological data such as temperature data, humidity, and rainfall data have been obtained from National Meteorological Agency and the top map of the sub-basin will be taken from EMA.

3.2 Consistency Test

Guideline for test of Consistency of data

According to Rao and Hammed, 2000:

A time series plot should be prepared to show:

- ❖ The annual maximal including gaps,
- ❖ A 3- to 5-years moving average of the annual maximal. Missing data and possible jumps or trends should be noted on this plot.

- ❖ A histogram plot should be prepared to show the distribution by magnitude of annual maximal. Anomalies such as bimodal distributions or apparent outliers can sometimes be identified on such plot. However, these features are more commonly identified after initial frequency analysis.

- ❖ Where a substantial degree of error is suspected in one or higher flood peaks, a sensitivity analysis using alternative values may be advisable, In important or dubious cases the original data should be reviewed critically with respect to gauge shift and malfunctions, estimation of missing data, computation of flow, and extrapolation-of rating curve.

The double- mass curve is used to check the consistency of many kinds of hydrologic data by comparing data for a single station with that of a pattern composed of the data from several other stations in the sub-basin. The double mass curve can be used to adjust inconsistency flow data. The graph of the cumulative data of one variable versus the cumulative data of a related variable is a straight line so long as the relation between the variables is a fixed ratio. Breaks in the double mass curve of such variables are caused by change in the relation between the variables. These changes may be due to changes in the method of data collection or to physical changes that affects the relation. Poor correlation between the variables can prevent detection of inconsistencies in a record, but an increase in the length of record tends to offset the effect of poor correlation.

3.3. Independency test

To test the independency of data for the stations, the W-W test has been used: For a given sample size N, the Wald-Wolfowitz (1943) called W-W test is used to test for the independence of a data set and to test for the existence of trend in it .

For a data set $x_1, x_2, x_3, \dots, x_N$ the static R calculated form:

$$R = \sum_{i=1}^{N-1} x_i x_{i+1} + \dots + x_i x_N \quad \dots \dots \dots (3.1)$$

When the elements of the sample are independent, R follows a normal distribution with mean and variance given by;

$$\bar{R} = \left[\frac{(S_1^2 - S_2^2)}{N-1} \right] \quad \dots \dots \dots (3.2)$$

$$var(R) = \frac{(S_2^2 - S_4^2)}{N-1} - (\bar{R})^2 + \left[\frac{(S_1^2 - 4S_1^2 S_2 + 4S_1 S_3 + 2S_2^2 - 2S_4)}{(N-1)(N-2)} \right] \quad \dots \dots \dots (3.3)$$

Where $S_r = \sum_{i=1}^N m_i^r$ and m_i^r is the r^{th} moment of the sample about the origin

$$u = \left[\frac{(R - \bar{R})}{(var(\bar{R}))^{1/2}} \right] \quad \dots \dots \dots (3.4)$$

The statistic 'u' is approximately normally distributed with mean zero variance unity and is used to test the hypothesis of independence at significance level $\alpha=0.05$, by comparing the statistic u with the standard normal variate $u_{\alpha/2}$ corresponding to a probability of exceedence $\alpha/2$. In this paper the value of |u| computed will be compared with $U_{0.025} = 1.96$ which is the standard value. (Rao and Hammed, 2000).

3.4 Homogeneity test

In this research also both conventional and L – moments have been used to calculate CV, LCV and their respective CC, that must be less than 0.3 to be homogeneous. The procedures are described below. value.

i) CV –Based homogeneity test

- a) For each site in a region calculate mean, standard deviation and coefficient of variation (Cv)

$$\bar{Q}_i = \sum_{j=1}^{n_i} \frac{Q_{ij}}{n_i} \dots\dots\dots(3.5)$$

$$\sigma_i = \sqrt{\frac{\sum_{j=1}^{n_i} (Q_{ij} - \bar{Q}_i)^2}{n_i - 1}} \dots\dots\dots(3.6)$$

$$CV_i = \frac{\sigma_i}{\bar{Q}_i} \dots\dots\dots(3.7)$$

Where :

Q_{ij} = the flow rate of station " j" in region " i"

\bar{Q}_i = The mean flow rate for region " i"

σ_i = Standard deviation for region " i"

CV_i = Coefficient of variation of region " i"

b) For each region ,using the statistic calculated CV above ,the regional mean ,CV and finally the corresponding CC value using the following relation

$$\overline{CV} = \sum_{i=1}^N \frac{CV_i}{N} \dots\dots\dots (3.8)$$

$$\sigma_{cv} = \sqrt{\frac{\sum_{i=1}^N (CV_i - \overline{CV})^2}{N-1}} \dots\dots\dots (3.9)$$

$$CC = \frac{\sigma_{cv}}{\overline{CV}} \dots\dots\dots (3.10)$$

Where :

N=Number of site in a region

\overline{CV} =Mean coefficient of at site CV values

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

The region declared to be homogeneous if $CC < 0.3$

ii) LCV-based homogeneity test

LCV- based homogeneity test is more accurate and effective way of testing the homogeneity of the site (station) when compared with that of the CV-based homogeneity test. The procedural calculation is the same as that of the CV;

The following are advantage of LCV (cunnane. 1989):

- ❖ Compared to CV, LCV can characterize a wide range of distribution.
- ❖ Sample estimates of LCV are so robust that they are not affected by the presence of outliers in the data set.
- ❖ They are less subjected to bias in estimation.
- ❖ LCV yields more accurate estimate of the parameter of a fitted distribution.

Hosking (1986) gave the unbiased estimators of M_{10k} and M_{1j0} as:

$$\hat{M}_{10k} = \frac{1}{N} \sum_{i=1}^N \left[\frac{\binom{N-i}{K}}{\binom{N-1}{K}} \right] x_i \quad (K = 0,1,2 \dots \dots N-1) \quad \dots \dots \dots (3.11)$$

($N = \text{number of year of record}$)

$$\hat{M}_{1j0} = \frac{1}{N} \sum_{i=1}^N \left[\frac{\binom{i-1}{j}}{\binom{N-1}{K}} \right] x_i \quad (j = 0,1,2 \dots \dots N-1) \quad \dots \dots \dots (3.12)$$

Where, $i =$ rank of observed flow data in ascending order.

The first few moments are:

$$L_1 = M_{100}$$

$$L_2 = M_{100} - 2 * M_{101}$$

$$L_3 = M_{100} - 6 * M_{101} + 6 * M_{102}$$

$$L_4 = M_{100} - 12 * M_{101} + 30 * M_{102} - 20 * M_{103} \quad \dots \dots \dots (3.13)$$

Like the conventional moments L-moment can be used to specify and summarize probability distribution. In particular L_1 , the first L-moment, is the mean of the statistical distribution and identical to the first conventional moment, and L_2 is a linear measure of spread or dispersion analogous to standard deviation. L-moment ratio, which are analogous to conventional moment ratio are defined by Hosking (1986);

Accordingly the formula used for the homogeneity test will be formulated as follows;

$$\tau_1 = \frac{L_2}{L_1} = \text{Measure of scale and dispersion } (L_{cv})$$

$$\tau_3 = \frac{L_3}{L_2} = \text{measure of skewness } (L_{cs})$$

$$\tau_4 = \frac{L_4}{L_2} = \text{Measure of kurtosis } (L_{ck})$$

Using the above procedural formula we have :

$$L_{cv} = \frac{L_2}{L_1} \quad \text{and} \quad \bar{L}_{cv} = \sum_{i=1}^N \frac{L_{cv}}{N}$$

$$\sigma_{L_{cv}} = \sqrt{\frac{\sum_{i=1}^N (L_{cv} - \bar{L}_{cv})^2}{N-1}} \quad \dots\dots\dots(3.14)$$

$$CC = \frac{\sigma_{L_{cv}}}{\bar{L}_{cv}}$$

The region declared to be homogeneous if $CC < 0.3$

3.5 Delineation of homogeneous regions

Depending on the statistical values, the geographical, topographical and altitude of the sub-basin, the stations can be grouped to form region. In the case of upper and middle Awash sub basin, since the basin is found in a great variation of elevation from higher mountainous to the plain area the influence of altitude plays a great role in variation of flood for upper and lower part of the sub-basin. This differs the flood producing characteristics of the stations. Combining the results from the statistical values, the geographical, topographical and altitude of the area, the flood producing characteristics of the stations and the physical phenomena of the area can be grouped to form regions. on the digitized map of the sub-basin, all stations under analysis were identified according to their geographical location (latitude and longitude) and statistical value.

The procedure adopted in delineation of homogeneous region comprises the following three steps:

- 1) Geographic information was used' to identify likely homogenous regions that are geographically continuous and having similar flood producing characteristics.
- 2) Each region that was identified in step '1' was checked for similarity in the statistics of observed flood data. Based on this step, regions obtained in step "1" were modified.
- 3) The proposed test of homogeneity was applied to confirm that the delineated regions are statistically homogenous.

3.6 Selection of best-fit distribution with proper method of Parameter estimation

To select the best fit distribution and the best method of parameter estimation different soft ware's have been used to fit the data with distribution. Using those soft ware's the data of the stations will fit with the distribution ,. The method of parameter estimation goes parallel with the distribution selection and can be tested for accuracy using standard error estimation, the less in standard error the more the selection of the method of the parameter estimation.

3.7 Quantile estimation and derivation of flood frequency curve

Based on the selected distributions for each station, the quantile will be calculated according to the formula of the selected distributions listed in chapter five. To calculate the quantile for each selected distribution the return period has to be determined first, the way of determining the return period is presented in chapter five. Accordingly the quantile versus return period graph of each region will be drawn as the growth curve for each station and for each region delineated. Depending on the growth curve of each station in the sub-basin and the representative growth curve of the sub-basin, the importance of regionalization of the sub basin can be determined. If the growth curve of the sub-basin can represent the station growth curve with less diversion from curve of each station, the sub basin can be taken as a single region otherwise the sub-basin has to be divided in to the proper homogeneous regions.

CHAPTER 4

4. DATA COLLECTION AND ANALYSIS

4.1 Source and availability of data

Data that are collected and to be analyze can be set under three categories according to their source and availability; Hydrological , Meteorological and Topographic data.

a) Hydrological data

It is a runoff data that is collected from gauging stations in the basin. In A wash basin there are about 76 gauging station, out of these 35 stations are in the study area. All 35 stations are not selected for the analysis because some of the stations are under the influence of the unnatural conditions such as the release of spillway, The very much plate of the area, and the back flow of reservoir; due to these reasons only 20 stations are selected for data analysis and maintained in Table 1.2. The stations have various length of record with fair distribution over the basin. These hydrological data is collected from the MoWR Hydrology department.

b) Meteorological data

Rainfall data are collected from meteorological Agencies for the purpose of comparison corresponding to the hydrological stations. Only 17 metrological stations are obtained for the analysis that can correspond with the hydrological stations and they are enough for comparison purpose. The data obtained are daily rainfall, temperature, humidity, sunshine and evaporation. The role of this data is very less since there is no rainfall-runoff model evolved in this paper. These data are obtained from National Meteorological Agency with the stations maintained in the Annex part .

C) Topographical data

Topographical data are collected from EMA (Ethiopian Map Authority). The map is used to identify the land feature and characteristics of the sub-basin. The digitized contour map of the sub-basin is the most important input to generate different basin outputs and to regionalize the sub-basin; that is collected from Mo WR, GIS department.

4.2. Filling and extension of data

In the study area there are about 35 gauging stations, out of these only 20 stations are selected for flood frequency analysis, the remaining have some defect, they are under the influence of natural and manmade factors. Some of the stations are set at u/s of koka reservoir that are directly subjected to the backup effect of the reservoir; others are set d/s of legedadi and koka dam which have no natural record due to the influence of the release of the water through spillway and the rest are on the very plate and plain area that may be subjected to the over flow and have a wrong record of data, on the other hand three of the stations are newly installed and have short record of data. Due to the above reasons, only 20 stations are selected for the proper analysis.

The selected once by themselves have no fully recorded data; they have less number of years of record that needs to be extended. Some of the stations have missing data that needs to be filled before analysis. Out of those selected 20 gauging stations 11 stations have one or more missed data and 5 stations have less than 15 years of record data;

As a result regression analysis is used to fill the missing monthly data and to extend those short length recorded data with satisfactory correlation coefficients.

To test the accuracy of the existing data as regressed and recorded data has been used; finally we have got a good effect in regressing data for extension or for filling of the missing data.

The graph that shows the relation between the recorded and generated is shown below.

Graphs that show the accuracy and application of the regression equation tested on different stations of the sub-basin, which compares the recorded and the regressed data of the stations that have maximum and minimum coefficient of determination:

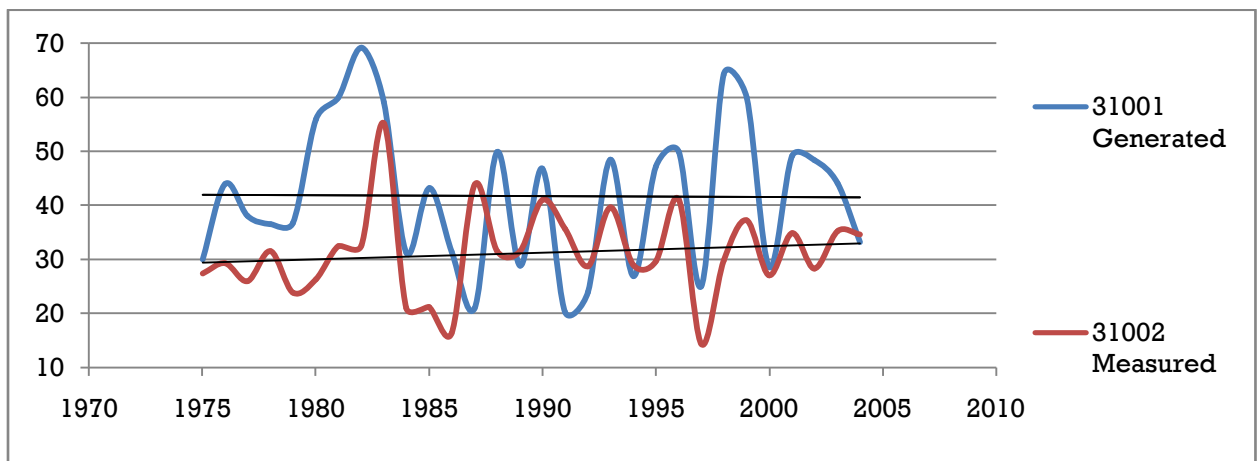
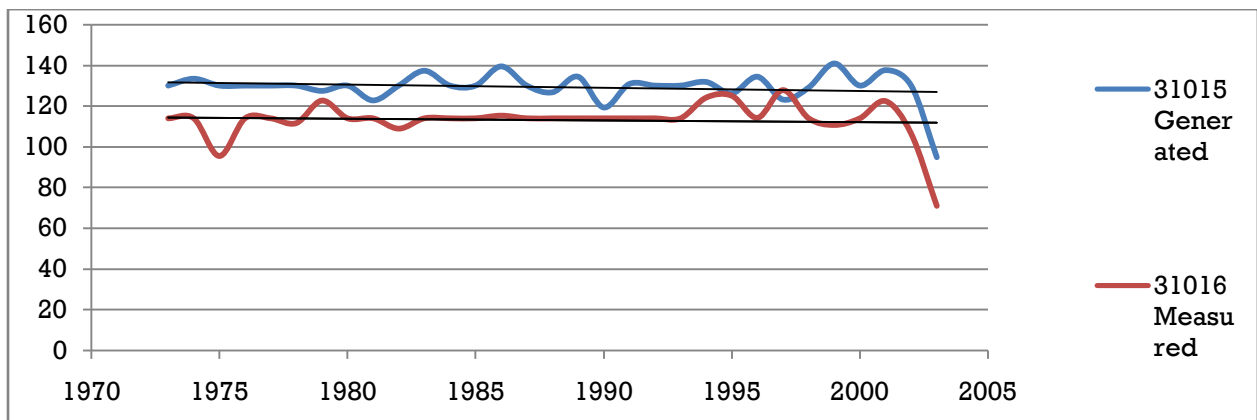
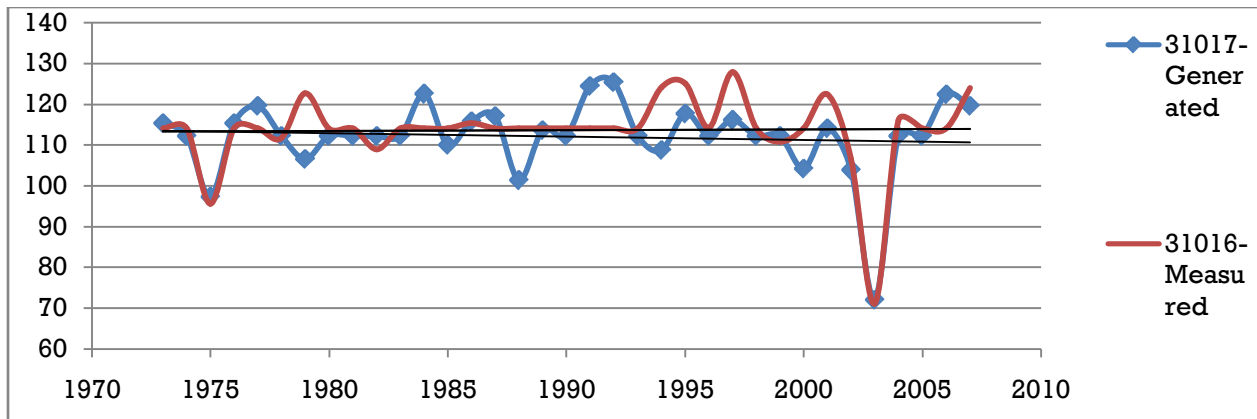


Fig.4.1-Graphs comparing measured and generated flow data by regression for the purpose of filling and extensio

Table 4.1 -Showing the regression equation of stations for filling and extension

Stat. No (Q1)	Station Name	Regretted with(Q)	Equation	R²	Range	N	Extended or Filled
31001	Berga near Addis -Alem	31002	$Q_1 = -0.081Q^2 + 5.952Q - 59.56$	0.782	1975-2004	29	Filled
31002	Holeta	31001	$Q_1 = -0.011Q^2 + 1.608Q - 17.35$	0.661	1970-2004	35	Filled
31003	Teji near Asgora	31012	$Q_1 = -0.031Q^2 + 14.33Q - 1536$	0.665	1980-2004	25	Filled
31004	Akaki	31013	$Q_1 = -0.002Q^2 + 2.153Q - 74.88$	0.781	1981-2003	23	Filled
31012	Awash at Melka -Kuntura		Enough Data		1973-2007	35	None
31013	Awash at Melka-Hombola		Enough Data		1968-2004	37	None
31014	Mojo		Enough Data		1973-2005	33	None
31020	Awash Near Belo	31001	$Q_1 = 0.052Q^2 - 4.733Q + 143.5$	0.652	1980-2004	25	Filled and Extended
31021	Little Akaki at Akaki	31013	$Q_1 = -0.001Q^2 + 0.658Q - 31.71$	0.731	1990-2003	14	Filled and Extended
31033	Awash at Ginchi	31001	$Q_1 = -0.256Q^2 + 22.75Q - 439.9$	0.663	1993-2004	12	Filled and Extended
31015	Kelata	31016	$Q_1 = -0.021Q^2 + 4.748Q - 134$	0.679	1973-2003	30	Filled
31016	Awash at Wonji	31017	$Q_1 = -0.02Q^2 + 4.943Q - 179.4$	0.728	1973-2003	30	Filled

Stat. No (Q1)	Station Name	Regretted with(Q)	Equation	R²	Range	N	Extended or Filled
31017	Awash below koka	31016	$Q_1 = -0.015Q^2 + 3.794Q - 122.2$	0.666	1973-2007	34	Filled
31019	Kesem - Beke	32005	$Q_1 = -0.132Q^2 + 70.57Q - 9345$	0.662	1994-2003	10	Filled
32002	Arba - Abomsa	32004	$Q_1 = -0.015Q^2 + 9.033Q - 1279$	0.777	1994-2003	10	Extended
32005	Kesem Awara Melka		Enough data		1969-2000	32	None
32003	Methara	32004	$Q_1 = -0.081Q^2 + 45.86Q - 63.41$	0.714	1969-2000	32	Filled
32004	Awash Station	32015	$Q_1 = -0.072Q^2 - 39.4Q + 56.57$	0.67	1969-2000	32	Filled
32009	Arba Bordada	32015	$Q_1 = -0.621Q + 185$	0.652	1993-2000	8	Extended
32017	Nura -Hera	32003	$Q_1 = -0.096Q^2 + 26.27Q - 1642$	0.867	1975-1983	9	Filled and Extended
32015	Awash Melka - Sedi	32004	$Q_1 = -0.009Q^2 - 5.5Q + 1041$	0.741	1983-2003	21	Filled

4.3 Consistency and Independence of data

The reliability of estimates of population quintile derived from frequency analysis depends in the first instance on the data series used in the analysis. Before conducting the analysis, the series should be scrutinized for possible errors or inconsistency and for any indication that contravene basic statistical assumption. (Rao and Hammed, 2000).

All the stations in the sub-basin are consistent except Akaki and Awash below Koka dam due to the influence of unnatural flow from Legedadi and Koka dam that does not tide our analysis.

4.3.1 Independence and Stationarity Test

A time series is Stationary if in the long term it is invariant with respect to time. Non Stationary may invalidate the result frequency analysis unless the data are first adjusted. Non-Stationeries in flood series may be of three types: jump, trends, and long-term periodic cycles or segments of the same. (Rao and Hammed, 2000)

It is usually assumed that all the peak magnitudes in the AM series are mutually independent in the statistical sense. Non-independent of the events in the data series may bias the result of frequency analysis. Although comparatively rare in flood data, non-independent may occasionally arise due to year-to year carry over of surface or subsurface storage, or non-independence in input series.

Specific Statistical test for independence are incorporated in various computation in frequency analysis. When applied to short series, however, commonly used tests can be misleading: they may indicate non-independence when the events are actually independent, or failed to indicate it when serial correlation long lags is in fact present. Flood frequency analysis is one of the investigations of extreme values. In any time series data, outliers may or may not exist. These outliers may come due to personal error during recording and inadequacy of measuring device or really due to very extreme condition of natural phenomenon that is important information for flood frequency analysis. Therefore, unless the source of the outliers clearly identified, it is

difficult to remove outliers completely from analysis. Outliers can be excluded from the estimation procedure only if it is certain that a single known distributional form can adequately model AM flood in this paper there is no as such magnified outlier. According to the above expressive idea on independency The independency test has been tested for the station using Wald- Wolfowitz (W-W), The Results of Independency test is tabulated below.

Table 4.2- Results of Independency test

St.No	Statistic	Critical test statistic	Remark
31033	0.524	1.96	Independent
31002	0.64	1.96	Independent
31003	-0.22	1.96	Independent
31004	0.114	1.96	Independent
31013	-0.345	1.96	Independent
31014	0.318	1.96	Independent
31020	0.289	1.96	Independent
31021	0.241	1.96	Independent
31001	0.325	1.96	Independent
31012	0.332	1.96	Independent
31015	-0.005	1.96	Independent
31016	0.266	1.96	Independent
32017	0.255	1.96	Independent
32004	0.183	1.96	Independent
32003	0.105	1.96	Independent
32015	0.254	1.96	Independent
32005	0.183	1.96	Independent
31019	0.183	1.96	Independent
32002	0.514	1.96	Independent
31017	0.354	1.96	Independent

Since all the statistic value are less than the critical test statistics values the data in the stations are independent . it is usually that all the peak magnitude in the AM series are mutually independent in the statistical sense .

4.4 Test of data for homogeneity of stations in the sub basin

Floods are affected by the physical and climatic characteristics of the catchments such as storm duration and intensity size, relief, drainage, morphology, land cover, presence or absence of storage, soil type and land use all these factors vary in space, Therefore, it is unreasonable to expect that a region can be chosen in which the flood frequency distribution at all site are identical what is required is that the region should be sufficiently homogenous that no further divisions of the region in to smaller regions or individual sites would improve the accuracy of flood quintile estimates (Admasu, 1989).

Stations in a region can be tested for homogeneity using statistical value; homogeneity can be taken as a base for many criteria of the basin. There are various homogeneity tests. The tests used in this study are:

- CV- based homogeneity test
- LCV- based homogeneity test

In regionalization, assumptions must be made about the statistical similarity of the sites in a region. To investigate whether those has been met or not many researchers as Cunnane (1989) have used the values of mean coefficient of variation (CV) and the site-to – site coefficient of variation (CC) of both conventional and L – moments of the proposed region. According to the researchers the higher the values of CV and CC the lower the performance of index flood method for the considered region. This is due to the dominance of the flood quantile estimation variance by the variance of the at – site sample mean. Hence for better performance of the index flood method, CC should be kept low ($CC < 0.3$) (Melsew,1996). The results are described below in the table.

Table 4.3 Results of homogeneity test

Region-1

St.No.	Cs	Ck	Cv	Lcv	Lcs	Lck	CC	Remark
31001	-0.719	3.976	0.153	0.029	0.424	0.077	Lcv based	Homogeneous
31002	-1.008	4.087	0.177	0.021	1.256	1.147	0.162	
31003	-0.968	3.523	0.271	0.022	0.376	44.994	Cv based	
31020	1.461	5.413	0.142	0.024	1.486	3.893	0.051	
31033	-1.953	6.433	0.200	0.030	0.571	0.433		

Region-2

St.No.	Cs	Ck	Cv	Lcv	Lcs	Lck	CC	Remark
31004	0.945	3.571	0.111	0.020	3.234	7.359	Lcv based	Homogeneous
31012	-0.291	5.418	0.066	0.018	2.296	3.949	0.142	
31013	-0.489	9.211	0.076	0.021	3.138	6.317	Cv based	
31014	-1.187	4.041	0.059	0.014	0.720	9.253	0.022	
31021	-0.126	3.355	0.098	0.020	0.082	0.127		

Region-3

St.No.	Cs	Ck	Cv	Lcv	Lcs	Lck	CC	Remark
31015	-2.816	16.281	0.056	0.012	4.152	8.535	Lcv based	Homogeneous
31016	-2.829	15.216	0.083	0.011	2.393	6.609	0.042	
31017	-4.718	25.653	0.082	0.011	3.716	8.973	Cv based	
32017	-2.458	12.811	0.126	0.011	0.369	2.013	0.029	

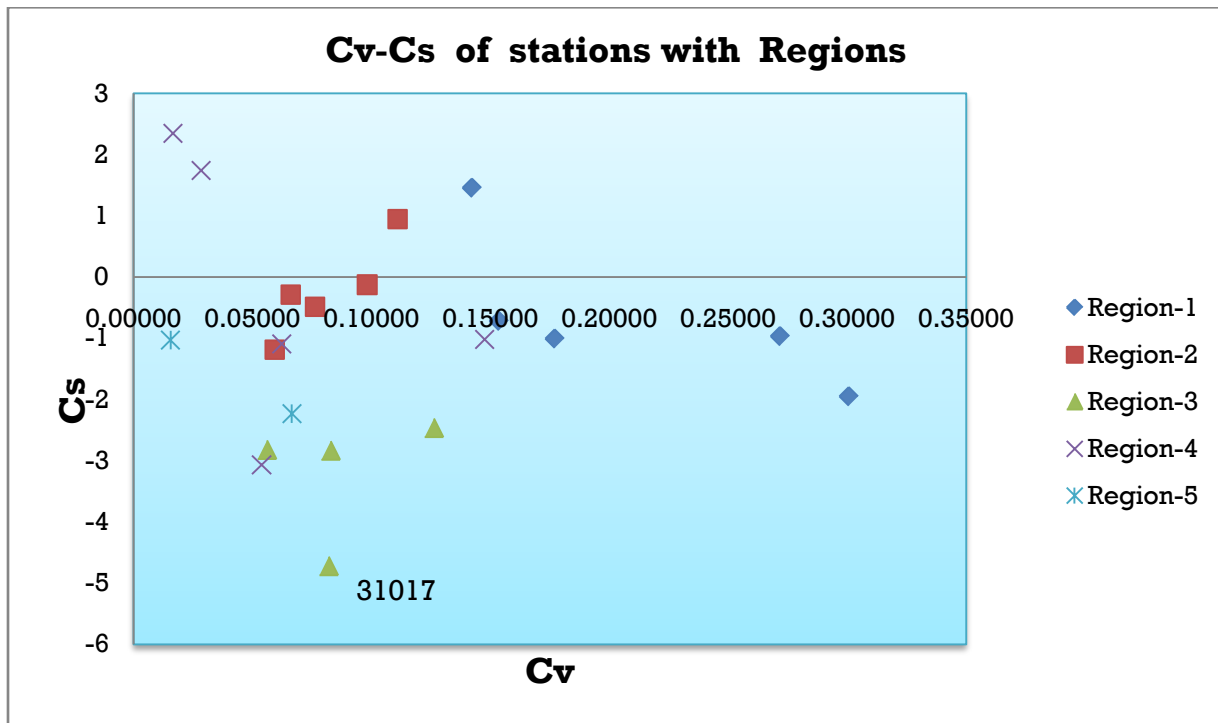
Region-4

St.No.	Cs	Ck	Cv	Lcv	Lcs	Lck	CC	Remark
32002	-1.088	4.087	0.046	0.024	5.432	11.528	Lcv based	Homogeneous
32003	-3.063	16.087	0.053	0.023	2.911	5.961	0.125	
32004	1.739	10.352	0.028	0.029	2.825	5.553	Cv based	
32015	2.345	10.374	0.036	0.024	3.066	6.296	0.011	

Region-5

St.No.	Cs	Ck	Cv	Lcv	Lcs	Lck	CC	Remark
31019	-2.227	9.487	0.066	0.109	3.234	6.556	Lcv based	Homogeneous
							0.154	
32005	-1.028	4.929	0.055	0.135	3.046	6.093	CV based	
							0.008	

From the above statistical values we have got that CC value is less than 0.3, which shows the stations in the region are homogeneous. The sub-basin is delineated under five regions from statistical and physical values all the regions are statistically homogenous. The statistical properties of the station can express the nature of the station; from the stations in the sub-basin, two of the stations are statistically have higher C_s value. The Akaki station is highly influenced by the release from Legedadi reservoir and that of Awash below Koka dam highly influenced by unnatural flow from the release of Koka reservoir the flow is not natural and it is also difficult to naturalize it. The Awash Bello station is found in a very plate area of becho plain area, which is mostly subjected to the over flow . This problem can also be more explained and illustrated on the graph below in the conventional-moment. From the graph the station with different values of C_s , C_v and C_K are differentiated from those remaining station Awash below Koka dam (31017) which has different statistical value as compared to the remaining stations due to the mentioned reasons.



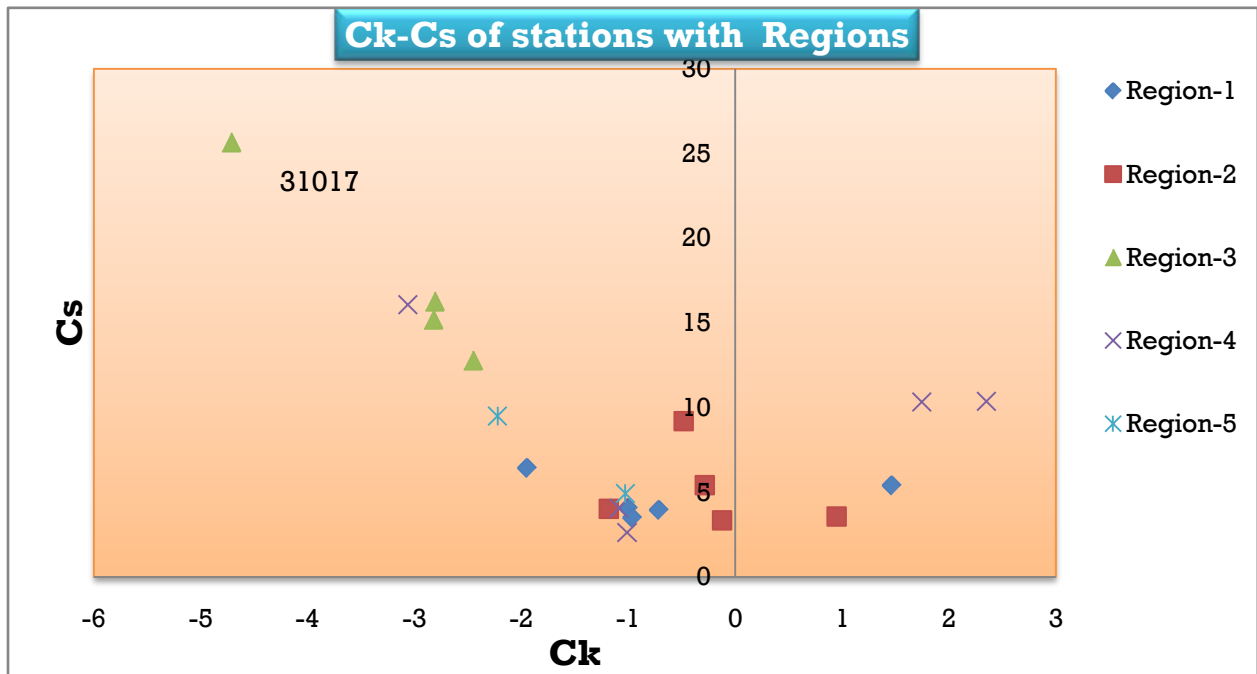
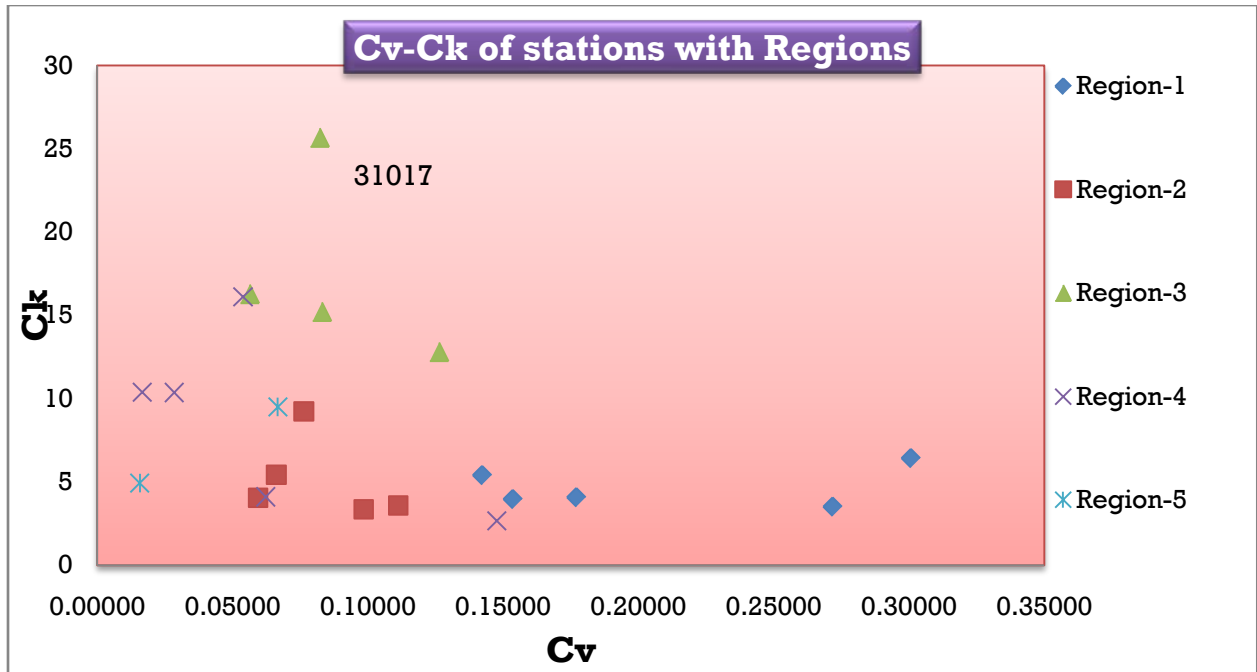


Fig 4.4: Moment ratio distribution of gauged stations of the homogeneous regions

4.5 Regionalization

Regionalization is the identification of homogenous region that contains stations having similar flood producing characteristics. This can be achieved using information that is obtained from geographical proximity, physiographic, climate, and altitude (topography) and drainage characteristics of the catchments. In the upper Awash sub-basin there is a great variation of elevations that produces the variation of flood producing characteristics, accordingly the topographic variation from highly mountainous to that of very plate area around becho plain that mostly attacked with great over land flood. These physical factors latter refined on the basis of statistics of the observed flood.

In regional flood frequency analysis, the established curve of flood variat versus return period can be used for estimating flood quintile at any site with in the region. In the Awash upstream of kesem the stations are found in evenly distributed way covering the whole area of the sub-basin with proper distribution and area coverage. Therefore, there is no excess place that is ungauged rather the stations with improper record of data have been removed, using those gauged stations the statistical properties and flood producing characteristics of the stations for regionalization purpose have been tested. Statistically we have got five regions; the first region includes the Berga, Holeta, Teji, Ginchi, , and Awash-Bello station and the second region have major and little Akaki, Melka-hombola, Melka-kunture and Mojo stations the third Awash-below koka, Wonji,Kelata and Nura-Hera the fourth Methara,Awash-at Awash-Station,Arba-Abomsa, and Awash Melka-Sedi the fifth Kesem-Baka and Kesem at Awara Melka. The stations topographical and geographical proximity, altitude and other external fixture of the catchments have been visualized. The sub basin can be divided in to five varying elevation and flood producing natures as the first region having higher elevation difference and in the second region which has moderate slope and plate nature the third region is plain area with highly cultivated ; those of Wonji and Nura-Hera farms also this area is affected with unnatural flow from the release of Koka dam the fourth region is also plain area the fifth region have also elevation difference from mountain to plate area. Generally depending on the result obtained from both tests five main

regions are delineated. In the regionalization the statistical results plays a great role in our case the result of both physical and statistical values indicates almost the same result, accordingly the stations in the regions is tabulated below:

Table 4.4- Showing stations and their regional classification

Station Number	Station Name	Region	Remark
31001	Berga near Addis -Alem	Region-1	Homogeneous
31002	Holeta	Region-1	Homogeneous
31003	Teji near Asgora	Region-1	Homogeneous
31020	Awash Near Belo	Region-1	Homogeneous
31033	Awash at Ginchi	Region-1	Homogeneous
31004	Akaki	Region-2	Homogeneous
31012	Awash at Melka- Kuntura	Region-2	Homogeneous
31013	Awash at Melka-Hombola	Region-2	Homogeneous
31014	Mojo	Region-2	Homogeneous
31021	Little Akaki at Akaki	Region-2	Homogeneous
31015	Kelata	Region-3	Homogeneous
31016	Awash at Wonji	Region-3	Homogeneous
31017	Awash below koka	Region-3	Homogeneous
32017	Nura -Hera	Region-3	Homogeneous
32002	Arba Abomsa	Region-4	Homogeneous
32003	Methara	Region-4	Homogeneous
32004	Awash Station	Region-4	Homogeneous
32015	Awash Melka - Sedi	Region-4	Homogeneous
31019	Kesem Beke	Region-5	Homogeneous
32005	Kesem Awara- Melka	Region-5	Homogeneous

CHAPTER 5

5. RESULTS AND DISCUSSION

5.1 Results from consistency, independency and homogeneity tests

The result obtained in the previous chapters shows all the station selected and regions for the analysis are consistent, independent and homogeneous. Due to this the stations and the regions delineated in the sub-basin are homogeneous, that is the base for further analysis such as distribution selection, method of parameter estimation and quintile estimation .

5.2 Selections of best fitted statistical distribution and parameter estimation

In RFFA a single frequency distribution is fitted to data from several sites. In general, the region will be slightly heterogeneous, and there will be single “true” distribution that applies to each site. The aim is therefore not to identify a “true” distribution but to find a distribution that will yield accurate quantile estimate for each site.

The aim of RFFA is not to fit a particular data set but to obtain quantile estimates of the distribution from which future data values will arise. When several distributions fit the data adequately, any of them is a reasonable choice for use in the final analysis, and the best choice among them will be the distribution that is most robust. i.e. most capable of giving good quantile estimate even though future data value may come from a distribution some what different from the fitted distribution. (Cunnane, 1986).

Distribution selection goes parallel with parameter estimation.

The three major parameter estimation used are:

- Method of moment (MOM)
- Method of probability weighted moment (PWM) and
- Method of maximum likely hood (ML)

Combinations of candidate distribution with that of parameter selection for this study are described below :

Table 5.1 Combinations of Distribution with that of Parameters

Se. No	Distribution/ Parametric estimation
1	EV1/MOM
2	EV1/PWM
3	EV1/ML
4	GEV/PWM
5	GEV/MOM
6	LN1/MOM
7	P-III/MOM
8	P-III/PWM
9	LP-III/MOM
10	Gama/MOM
11	Gama/PWM
12	Gama/ML
13	LL/PWM
14	EXP/MOM
15	Wakeby-4/PWM
16	Wakeby-5/PWM

5.3 Selection of the most robust flood estimation procedure

According to cunnane, 1989, several techniques have been used in the past for evaluating the suitability of different distributions for AM series. Two main categories can be identified in the use of these techniques, these are:

a) Test of descriptive ability

One of the recent techniques, which is introduced under descriptive ability test, is goodness of fit measure. For the already identified regions, goodness –of – fit measure helps to test whether a given distribution fits the data or not. For those of candidate distributions, the goodness of fit measure takes place with a significance level of $\alpha=0.05$ which is a confidence level of 95% and it is described in the table below.

b) Test of predictive ability

Predictive ability tests investigate how well a distribution and its associated method of Parameter estimation can estimate the Q – T relationship or the frequency of future events when the population distribution is not identical to that of the proposed distribution. The indicator which is used to test the predictive ability of a distribution is standard-error of estimation. The most efficient the parameter estimation gives the smaller standard- error of estimation.

Table 5.2. Fitting of data with distribution by using goodness of fit test

Station. No	LN		LP3		Gama		GEV		LLG		P3		EXP	
	Xv	Xc	Xv	Xc	Xv	Xc	Xv	Xc	Xv	Xc	Xv	Xc	Xv	Xc
31001	7.81	2.56	2.21	7.81	5.32	7.81	2.56	7.81	5.44	7.81	2.56	7.81		
31002	5.26	9.49	3	9.49			0.1	0.22	2.19	9.49				
31003	4.77	5.99	5.35	5.99	1.89	5.99	1.62	5.99	4.18	5.99	1.86	5.99		
31004	4.97	5.99	3.05	5.99	3.08	5.99	4.94	5.99	0.23	0.89	3.07	5.99	4.46	5.99
31012	7.64	7.81	7.68	7.81	0.11	0.22								
31013	3.83	9.45	4.19	9.45	4.04	7.81	4.38	9.49	3.3	9.89	2.71	9.49		
31014	7.11	7.81	0.25	9.49	0.23	9.49	0.21	9.49	2.32	7.81	0.2	9.49	4.99	9.49
31015	1.6	11.07	2.74	9.49	3.46	9.49	2.94	9.49	0.86	9.49	2.52	9.49		
31016	6.64	7.81	5.66	7.81	3.2	7.81	7.11	7.81	6.19	7.81	7.63	7.81		
31017	3.72	9.49	3.69	9.49	5.96	9.49	3.95	9.49	4.23	9.49	2.22	9.49	1.05	9.49
31019	5.54	5.99	5.47	5.99	5.62	5.99	5.46	5.99	4.57	5.99	5.44	5.99		
31020	1.12	5.99			1.48	5.99	0.99	5.99			5.89	5.99	0.06	3.84
31021	1.34	7.78	1.29	9.49	3.37	9.49	3.6	9.49	1.63	9.49	1.33	9.49	2.12	5.99
31033	4.45	7.81	2.72	7.81	0.32	7.81	2.12	7.81	1.81	5.99	2.22	7.81		
32002					5.11	9.49					2.05	7.81		
32003	7.23	7.81	7.21	7.81	7.27	7.81	7.21	7.81	5.48	7.81	7.3	7.81		
32004	3.67	9.49	4.48	9.49	4.46	9.49	0.28	11.07	2.96	9.49	2.12	9.49		
32015	0.47	3.84	0.3	3.84	0.36	3.84	0.13	3.84	0.18	3.84	0.27	3.84		
32017	1.68	5.99	0.89	5.99	0.89	5.99	3.91	5.99	2.22	5.99	0.89	5.99		

5.4 Quantile estimation and derivation of flood frequency curve.

5.4.1 Determination of quantile estimates

After selection of best-fit distribution, the desired quantile estimates are computed from the statistics of the adopted distribution. In the case of major hydraulic structures, flood estimates are sometimes requested for very long return period depending on their record data, up to 10,000 years or more. It may also be desired to estimate the return period of a deterministically derived probable maximum flood. The reliability of extrapolating of flood frequency curve to such return periods is generally extremely low—a minor change in the data series or in the fitting distribution can make huge differences to the estimates. Where such estimates are required, it is advisable to consider additional studies using methods other than standard frequency analysis (Rao and Hammed, 2000).

5.4.2 Guidelines for determining of quantiles

Quantile estimates should normally be computed from the adopted probability distribution as fitted to the adopted data series. Estimate for return periods of greater than 200 years should normally be proved only when specially registered, and their uncertainty should be indicated. When fitting of a theoretical distribution is not practicable, estimates may be read from a graphical fitting curve. For special purposes such as econometric analysis and determination of flood damage benefits, an 'expected probability' curve may be developed from the basic flood frequency relationship using conversions. (Rao and Hammed, 200)

5.4.3 Estimation of return period

Given a verified or accepted discharge Q for the maximum known flood the plotting position formula used in the frequency analysis provides a rough estimate of its own return period T as a function solely of the record length N . For the formula previously quoted, the plotting-position return periods in years are: $T=1.67N+0.3$ (Cunnane); or $T=N+1$ (Weibull). The Cunnane formula gives a return period substantially longer than the period of record, and generally seems to result in a more compatible plot. A more refined estimate of return period T is given by the position of Q on the adopted probability distribution. For the usual use of single probability distribution, both T and its associated standard error can be calculated from the statistics of the distribution (Roa and Hammed, 2000).

5.4.4 Guidelines for Estimation of return period

- ✓ The return period of the maximum known flood may be determined from the fitted probability distribution and compared with that indicated by the Cunnane plotting position formula. If there are several discrepancies, consideration should be given to the physical circumstances of the maximum event, the reliability of its accepted magnitude and the reliability of the adopted fitting curve.
- ✓ To indicate the error of the estimated return period for the maximum known flood, the range of return periods indicated by the confidence limits or equivalent may be quoted.

Depending on the above guidelines the return period have been selected, based on the recorded number of years, up to 10000 years and calculated the corresponding quantiles. The result quantile estimation is shown below in tabular form and graphical form.

5.4.5 Properties of quantile estimator

- Sources of error of estimation:

- sampling error
- model error

The estimated value, QT , may differ from the true value QT because of:

- a) Inability of model chosen (AM or PD) to reproduce the population Q-T.
- b) In correct choice of distribution to describe the population with in the chosen model.
- c) Bias in the estimating procedure (if this is known to exist, a correction can be made for it)
- d) Sampling error due to the fact that parameters are estimated from a finite sample.
- e) The available record (sample) may not be a truly random sample from the required population. No control can be exercised over this, even though test can be made to test the reasonableness of the assumption. (Cunnane, 1989)

5.4.6 Evaluation of sampling and distribution error

- ❖ It has not been universal practice to report error in quantitative terms.
- ❖ Different researchers use different measures of error. The two most common measures are standard errors and confidence intervals.
- ❖ From the various source of error only sampling error can be evaluated theoretically a consensus seems to be emerging that at least sampling error should normally be reported in quantitative terms. There appears to be no generally accepted basis for determination of distribution error associated with the adoption of a particular theoretical distribution (Rao and Hammed, 2000).

5.4.7 Simulation test on quantile estimation

Estimation method depends on data availability and on the amount of regional pooling of data, which is to be allowed (Cunnane, 1989):

1. Decide on:
 - i) Parent distribution; hence calculate true quantile value Q_T
 - ii) Sample size
 - iii) Estimating distribution
 - iv) Method of parameter estimation
2. Generate a sequence of N values (apply estimating procedure QT)
3. Repeat step two M times
4. Calculate:

$$\ddot{Q} = \frac{\sum \overline{Q}_T}{Q_T}$$

$$S_e = \left[\frac{\sum (\overline{Q} - \ddot{Q})^2}{Q_T} \right]^{0.5}$$

$$rms_e = \left[\frac{\sum (\overline{Q} - \ddot{Q})^2}{Q_T} \right]^{0.5}$$

$$biase(\overline{Q}) = (\ddot{Q}_T - \overline{Q})$$

\overline{Q}_T = Estimated value of standard quantile

Q_T = Population value of standardized flood

5. Apply the respective formula for each test (*biase*, S_e , rms_e) to get their values for the considered distribution procedure, return period (T) and sample size N.

The procedure giving the least estimator values is the most robust flood estimation procedure for a given region and station. Selection of the most efficient method that gives the smallest standard error of estimate. Based on the above steps, guidelines and criteria the standard error for the selected parameter estimation method and return periods have been estimated, depending on the result of SEE we have selected the best fit parameter estimation and distributions as tabulated below :

Table 5.3- Standard error for Region -1

T	1001			1002			1003	
	GEV/MO M	LN/MO M	P3/MO M	GEV/MO M	LN/MO M	P3/MO M	GEV/MO M	LN/MO M
2	0.32	0.31	0.33	0.17	0.17	0.17	0.80	0.62
10	0.54	0.56	0.54	0.30	0.29	0.30	0.73	1.16
20	0.66	0.70	0.66	0.38	0.36	0.38	0.65	1.46
50	0.84	0.90	0.83	0.52	0.48	0.50	0.60	1.91
100	1.00	1.08	0.97	0.63	0.57	0.60	0.67	2.31
200	1.18	1.27	1.12	0.75	0.67	0.70	0.81	2.76
500	1.44	1.54	1.33	0.93	0.81	0.84	1.08	3.43
1000	1.66	1.76	1.50	1.06	0.92	0.96	1.31	4.02
10000	2.49	2.39	1.93	1.55	1.22	1.25	2.14	5.73
Av.SE	1.13	1.17	1.02	0.70	0.61	0.63	0.98	2.60

T	1003	1020	1020	1020	1033	1033	1033
	P3/MO M	GEV/MO M	LN/MO M	LP3/MO M	Gama/MO M	LN/MO M	P3/MO M
2	0.80	0.12	0.12	0.11	0.45	0.43	0.52
10	0.71	0.32	0.29	0.31	0.60	0.78	0.58
20	0.64	0.40	0.36	0.42	0.70	1.01	0.59
50	0.61	0.52	0.44	0.58	0.84	1.37	0.63
100	0.66	0.61	0.51	0.73	0.96	1.69	0.68
200	0.75	0.71	0.57	0.89	1.08	2.05	0.76
500	0.94	0.85	0.65	1.15	1.25	2.57	0.89
1000	1.12	0.97	0.72	1.38	1.38	3.03	1.01
10000	1.64	1.49	0.87	2.12	1.72	4.33	1.35
Av.SE	0.87	0.66	0.50	0.85	1.00	1.92	0.78

Table 5.4-Standard error estimation for Region -2

T	1004		1012			1013	1013
	Gama/MO M	P3/MO M	Gama/MO M	P3/MO M	LN/MO M	GEV/MO M	Gama/MO M
2	0.17	0.19	0.16	0.16	0.14	0.25	0.29
10	0.28	0.28	0.19	0.17	0.17	0.42	0.41
20	0.34	0.32	0.24	0.22	0.20	0.51	0.48
50	0.42	0.37	0.32	0.30	0.28	0.63	0.58
100	0.49	0.42	0.39	0.36	0.35	0.72	0.66
200	0.55	0.47	0.46	0.42	0.45	0.81	0.75
500	0.63	0.55	0.56	0.51	0.60	0.93	0.88
1000	0.70	0.61	0.63	0.58	0.75	1.02	0.99
10000	0.87	0.78	0.82	0.75	1.42	1.25	1.28
Av. SE	0.49	0.45	0.42	0.38	0.48	0.73	0.70

T	1013	1014		1014	1021		1021
	P3/MO M	Gama/MO M	P3/MO M	GEV/MO M	Gama/MO M	P3/MO M	LN/MO M
2	0.24	0.12	0.11	0.11	0.06	0.06	0.05
10	0.47	0.19	0.13	0.13	0.10	0.10	0.12
20	0.60	0.22	0.13	0.14	0.13	0.13	0.17
50	0.79	0.27	0.15	0.15	0.17	0.18	0.25
100	0.95	0.31	0.16	0.16	0.20	0.22	0.33
200	1.13	0.35	0.17	0.18	0.24	0.27	0.42
500	1.37	0.41	0.20	0.21	0.28	0.33	0.56
1000	1.58	0.45	0.22	0.23	0.32	0.38	0.68
10000	2.14	0.57	0.28	0.34	0.42	0.52	1.06
Av. SE	1.03	0.32	0.17	0.18	0.21	0.25	0.41

Table 5.5- Standard error estimation for Region-3

T	31015			31016		
	EXP/MOM	Gama/MOM	P3/MOM	EXP/MOM	Gama/MOM	P3/MOM
2	0.06	0.07	0.08	0.14	0.16	0.19
10	0.11	0.12	0.11	0.33	0.34	0.34
20	0.14	0.14	0.13	0.43	0.42	0.40
50	0.19	0.18	0.16	0.56	0.55	0.50
100	0.22	0.21	0.19	0.66	0.65	0.58
200	0.25	0.24	0.22	0.75	0.76	0.67
500	0.30	0.29	0.26	0.89	0.90	0.80
1000	0.33	0.32	0.29	0.99	1.02	0.91
10000	0.44	0.41	0.38	1.32	1.31	1.20
Av.SE	0.23	0.22	0.20	0.67	0.68	0.62

T	31017			32017		
	EXP/MOM	Gama/MOM	P3/MOM	EXP/MOM	Gama/MOM	P3/MOM
2	0.09	0.08	0.11	0.09	0.10	0.12
10	0.21	0.18	0.19	0.13	0.13	0.12
20	0.27	0.23	0.21	0.15	0.14	0.12
50	0.35	0.29	0.23	0.18	0.17	0.14
100	0.40	0.33	0.25	0.21	0.19	0.16
200	0.46	0.38	0.27	0.24	0.21	0.18
500	0.54	0.44	0.30	0.28	0.24	0.22
1000	0.60	0.50	0.33	0.31	0.27	0.25
10000	0.79	0.62	0.42	0.41	0.33	0.34
Av.SE	0.41	0.34	0.25	0.22	0.20	0.18

Table 5.5- Standard error estimation for Region-4

T	32003			31004		
	Gama/MOM	P3/MOM	EV1/MOM	Gama/MOM	P3/MOM	EV1/MOM
2	0.14	0.15	0.13	0.14	0.16	0.14
10	0.17	0.16	0.16	0.16	0.16	0.16
20	0.19	0.17	0.19	0.18	0.19	0.18
50	0.23	0.20	0.22	0.22	0.24	0.23
100	0.26	0.23	0.25	0.25	0.29	0.26
200	0.30	0.26	0.28	0.29	0.35	0.30
500	0.36	0.32	0.32	0.33	0.43	0.35
1000	0.41	0.36	0.36	0.37	0.49	0.39
10000	0.54	0.48	0.46	0.46	0.65	0.52
Av.SE	0.29	0.26	0.27	0.27	0.33	0.28

T	32004			32002		
	Gama/MOM	P3/MOM	EV1/MOM	Gama/MOM	P3/MOM	EV1/MOM
2	0.14	0.16	0.14	0.05	0.05	0.04
10	0.18	0.17	0.18	0.06	0.06	0.05
20	0.22	0.19	0.22	0.07	0.07	0.07
50	0.26	0.22	0.27	0.09	0.10	0.08
100	0.30	0.25	0.31	0.10	0.12	0.10
200	0.34	0.28	0.35	0.12	0.14	0.11
500	0.39	0.32	0.40	0.14	0.17	0.13
1000	0.43	0.36	0.45	0.16	0.20	0.14
10000	0.54	0.47	0.60	0.20	0.25	0.19
Av.SE	0.31	0.27	0.32	0.11	0.13	0.10

Table 5.7- Standard error estimation for Region -5

T	32015	32015	32015	32019	32019	32019
	Gama/MOM	P3/MOM	EV1/MOM	Gama/MOM	P3/MOM	EV1/MOM
2	0.14	0.13	0.14	0.04	0.05	0.04
10	0.20	0.21	0.20	0.07	0.06	0.06
20	0.25	0.26	0.25	0.08	0.07	0.08
50	0.31	0.32	0.32	0.09	0.09	0.09
100	0.36	0.37	0.37	0.11	0.11	0.11
200	0.41	0.42	0.43	0.12	0.12	0.12
500	0.48	0.47	0.51	0.14	0.15	0.14
1000	0.53	0.52	0.57	0.15	0.17	0.15
10000	0.66	0.63	0.76	0.19	0.22	0.20
Av.SE=	0.37	0.37	0.39	0.11	0.12	0.11

5.5 Derivation of dimensional and standardized flood Quantil for each region

The regional flood frequency has important implication for hydrological processes the slop of a frequency curve graphically represent the standard deviation of the flood frequency distribution and the higher the slop, the greater the standard deviation in flood discharge (pitlick, 1994) the index flood method standardizing flow pays a great role in derivation of flood frequency curve both for the dimensional and standard flow.

The following table shows the estimated quantile (QT) and standardized (QT/Q-mean) using the above selected distribution and parameter for the stations with their corresponding return period and the region delicate in the sub-basin

Table 5.8- Dimensional Quantiles for Region-1

		31002	31003	1020	31033
T	GEV/MOM	GEV/MOM	GEV/MOM	GEV/MOM	GEV/MOM
2	43.3	31.51	66.66	36.86	51.53
10	68.86	44.88	110.67	46.92	85.66
20	77.91	49.48	123.1	51.17	96.6
50	89.07	55.07	136.43	57.02	109.29
100	97.03	59	144.72	61.7	117.81
200	104.66	62.7	151.77	66.61	125.55
500	114.25	67.29	159.52	73.51	134.72
1000	121.16	70.54	164.39	79.07	140.95
10000	142.19	80.15	176	99.78	157.93

Table 5.9- Dimensional Quantiles for Region-2

T	31004	1012	31013	1014	1021
	Gama/MOM	Gama/MOM	Gama/MOM	Gama/MOM	Gama/MOM
2	188.32	215.65	405.32	116.37	51.57
10	310.48	340.08	628.99	218.3	107.1
20	352.76	382.42	704.65	255.29	128
50	404.26	433.65	795.94	301.24	154.34
100	442.09	471.06	862.48	335.53	174.21
200	478.33	506.75	925.86	368.75	193.63
500	525	552.52	1007.02	412	219.11
1000	560.6	587.32	1068.62	445.33	238.89
10000	645.97	670.34	1215.36	526.27	287.35

Table 5.10- Dimensional Quantiles for Region-3

T	31015	31016	31017	32017
	P3/MOM	P3/MOM	P3/MOM	P3/MOM
2	76.59	98.91	80.7	108.13
10	141.29	198.83	170.26	164.74
20	162.22	232.96	203.94	181.88
50	187.01	274.31	246.38	201.58
100	204.8	304.55	278.4	215.37
200	221.55	33.43	309.67	228.12
500	242.75	370.49	350.71	243.95
1000	258.68	398.69	382.56	255.64
10000	296.11	466.04	460.57	282.48

Table 5.11- Dimensional Quantiles for Region-4

T	32002	32003	32004	32015
	Gama/MOM	Gama/MOM	Gama/MOM	Gama/MOM
2	56.41	111.24	256.07	241.2
10	91.66	219.65	371.92	341.7
20	103.8	259.73	410.12	374.52
50	118.55	309.88	455.73	413.54
100	129.36	347.51	488.67	441.63
200	139.71	384.13	519.85	468.14
500	153.01	432.01	559.51	501.78
1000	163.15	469.04	589.43	527.11
10000	187.42	559.37	660.18	586.81

Table 5.12- Dimensional Quantiles for Region-5

T	32005	31019
	P3/MOM	P3/MOM
2	258.27	59.81
10	437.91	96.08
20	496.04	108.01
50	564.88	122.25
100	614.27	132.53
200	660.79	142.25
500	719.68	154.62
1000	763.91	163.95
10000	867.86	186

Table 5.13-Standardized Quantiles for Region-1

T	31001	31002	31003	1020	31033
	GEV/MOM	GEV/MOM	GEV/MOM	GEV/MOM	GEV/MOM
2	0.95	0.97	0.98	0.96	0.87
10	1.51	1.38	1.63	1.23	1.44
20	1.71	1.52	1.81	1.34	1.62
50	1.95	1.69	2.01	1.49	1.84
100	2.13	1.81	2.13	1.61	1.98
200	2.29	1.92	2.24	1.74	2.11
500	2.50	2.06	2.35	1.92	2.26
1000	2.65	2.16	2.42	2.07	2.37
10000	3.12	2.46	2.59	2.61	2.65

Table 5.14- Standardized Quantiles for Region-2

T	1004	1012	31013	1014	1021
	Gama/MOM	Gama/MOM	Gama/MOM	Gama/MOM	Gama/MOM
2	0.94	0.95	0.96	0.91	0.88
10	1.56	1.50	1.48	1.71	1.82
20	1.77	1.69	1.66	2.00	2.18
50	2.03	1.92	1.88	2.35	2.63
100	2.21	2.08	2.04	2.62	2.97
200	2.40	2.24	2.19	2.88	3.30
500	2.63	2.44	2.38	3.22	3.73
1000	2.81	2.60	2.52	3.48	4.07
10000	3.24	2.96	2.87	4.11	4.89

Table 5.15- Standardized Quantiles for Region-3

T	31015	31016	31017	32017
	P3/MOM	P3/MOM	P3/MOM	P3/MOM
2	0.95	0.92	0.88	0.98
10	1.75	1.85	1.85	1.50
20	2.01	2.17	2.21	1.65
50	2.32	2.55	2.67	1.83
100	2.54	2.83	3.02	1.96
200	2.75	0.31	3.36	2.08
500	3.01	3.44	3.80	2.22
1000	3.21	3.71	4.15	2.33
10000	3.67	4.33	4.99	2.57

Table 5.16- Standardized Quantiles for Region-4

T	32002	32003	32004	32015
	Gama/MOM	Gama/MOM	Gama/MOM	Gama/MOM
2	0.95	0.89	0.97	0.97
10	1.54	1.77	1.41	1.38
20	1.74	2.09	1.55	1.51
50	1.99	2.49	1.72	1.67
100	2.17	2.79	1.85	1.78
200	2.34	3.09	1.97	1.89
500	2.57	3.47	2.12	2.02
1000	2.74	3.77	2.23	2.13
10000	3.15	4.50	2.50	2.37

Table 5.17-Standardized Quantiles for Region-5

T	32005	31019
	P3/MOM	P3/MOM
2	0.96	0.96
10	1.63	1.54
20	1.84	1.73
50	2.10	1.96
100	2.28	2.12
200	2.45	2.28
500	2.67	2.48
1000	2.84	2.63
10000	3.22	2.98

5.6 Derivation of flood Frequency curve for each region

Only standardized flood Quantiles are presented here for comparison purpose the rest of the graphs are found in the annex part.

(Standard – $Q_T = Q_T/Q_{mean}$ and T in years)

Fig 5.1- Standardized Growth curve of Region-1 and each station

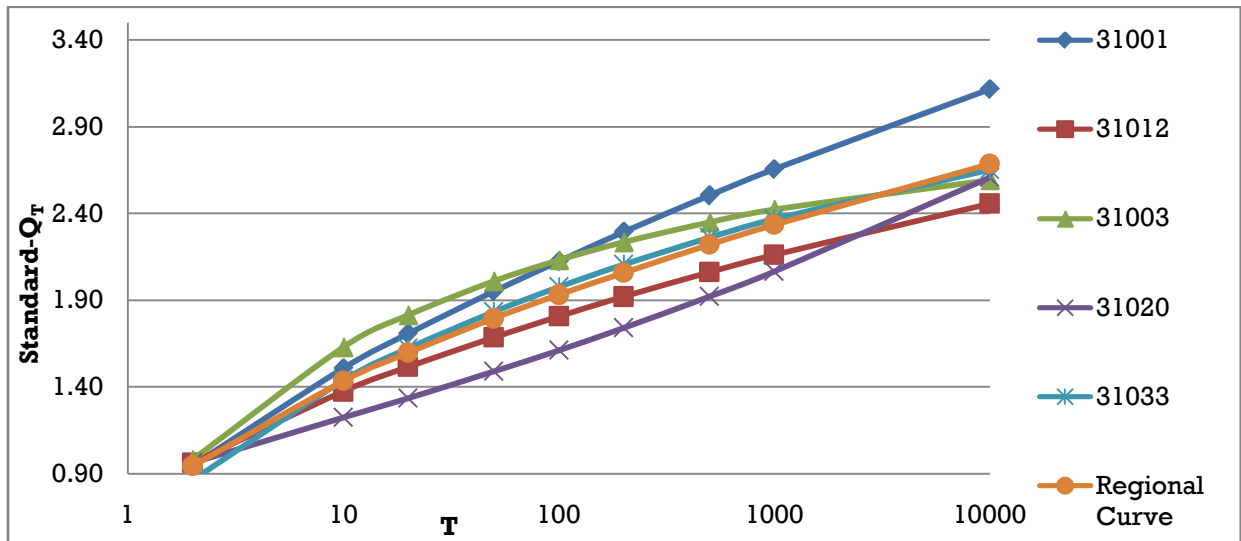


Fig 5.2- Standardized Growth curve of Region-2 and each station

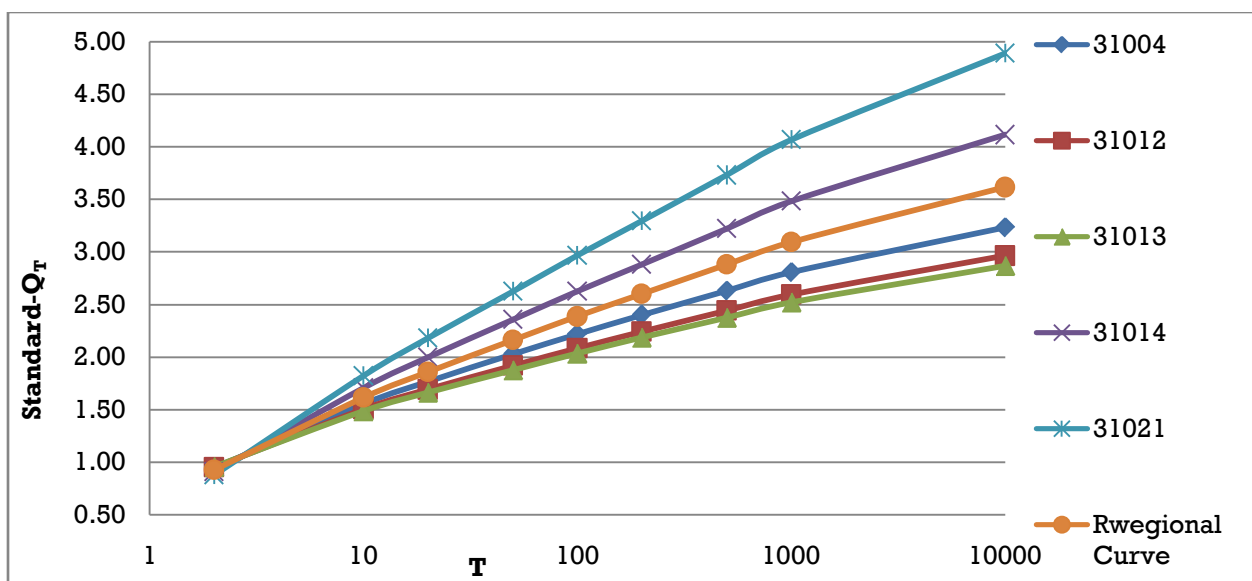


Fig 5.3- Standardized Growth curve of Region-3 and each Station

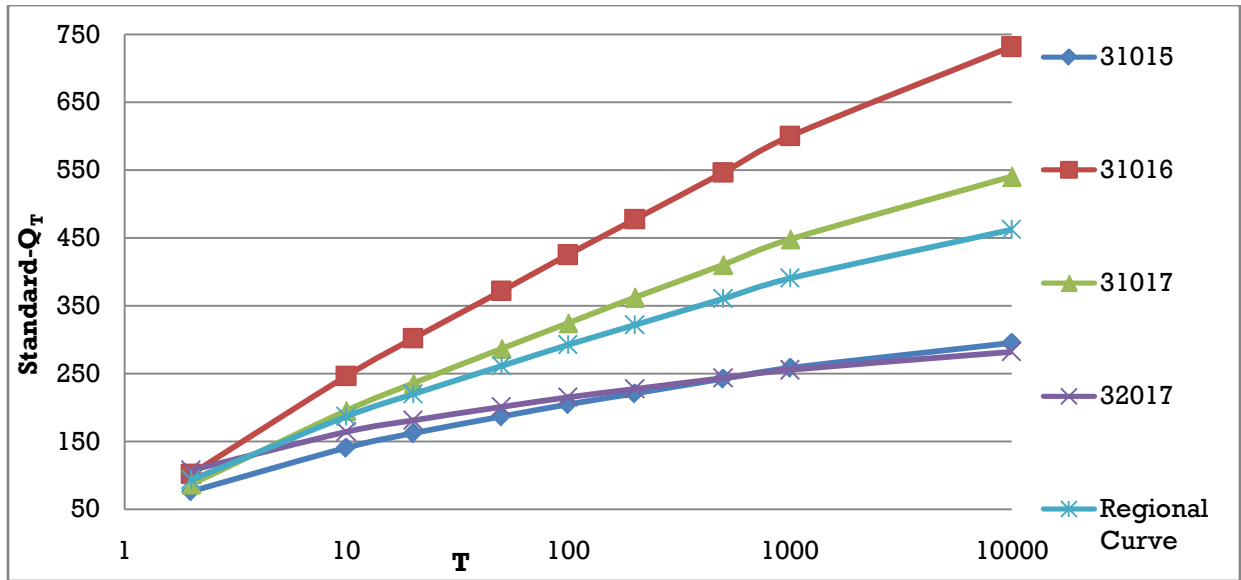


Fig 5.4- Standardized Growth curve of Region-4 and each station

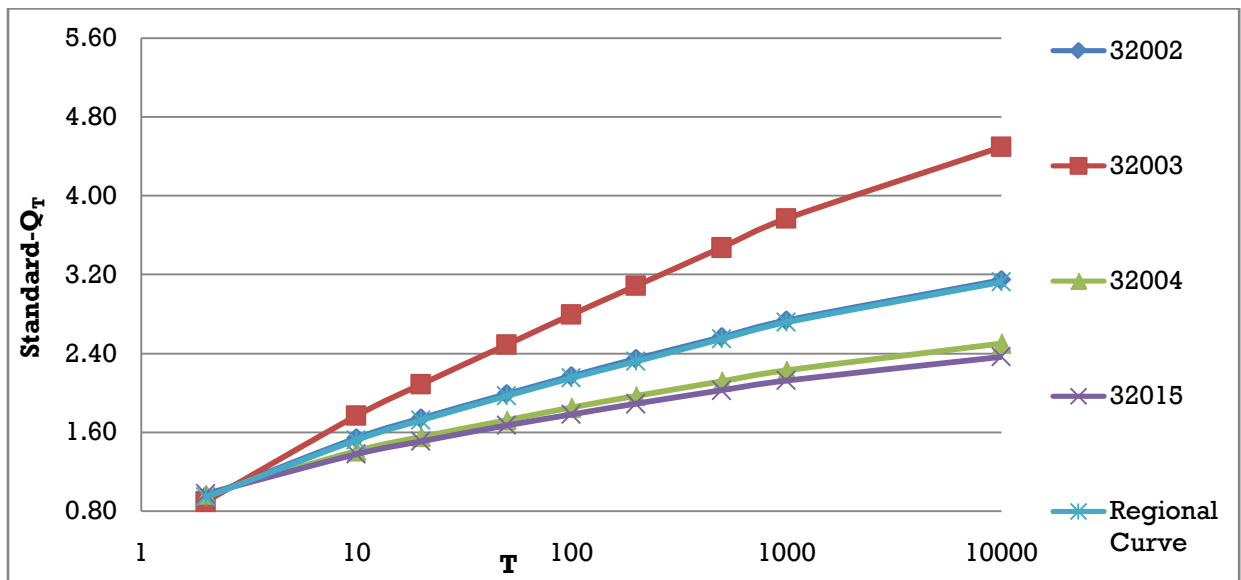
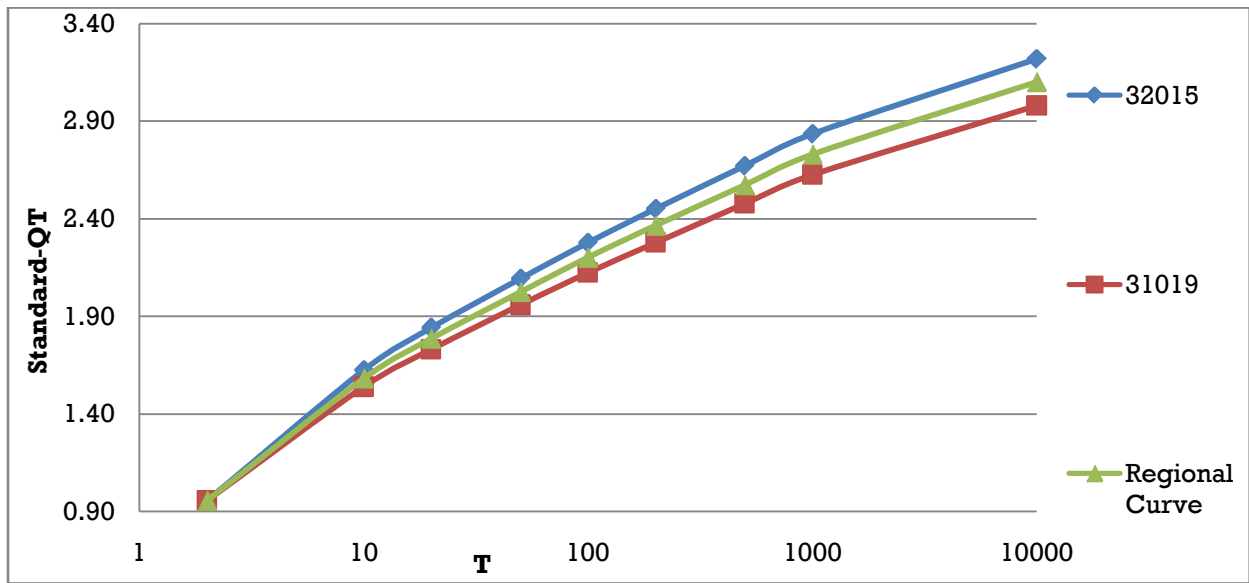


Fig 5.10- Standardized Growth curve of Region-5 and each station



5.6 Prediction of mean annual flood (MAF) for ungauged catchments

In this study regression analysis was applied to predict mean annual flood (MAF) for ungauged catchments in the study area. The first task in the regression analysis was to obtain the correlation matrix, which consists of coefficients of correlation, computed between pairs of catchment characteristics. In the analysis, MAF value was taken as dependable variable where as the catchments characteristics (such as area, slope, channel length, Elevation, shape factor) and average annual rainfall was considered as independent variables. Actually the rainfall and elevation are dependent variables so that during the analysis either of the two was used .The results of the correlation were then used to build the regression equation .The first variable to be added in the regression equation was the one with the highest simple correlation with the dependent variable. Other variables that were observed to explain the largest variation in the dependent variable that remained unexplained by the first variable were then added.

The information on catchments characteristics that were available for consideration in the regression analysis are:

- ❖ Catchments area A , (Km²)
- ❖ Mean rainfall of the catchments P , (mm)
- ❖ Average elevation of the catchments E_l , (m)
- ❖ Average slope of the catchments S , (%)
- ❖ Stream length L , (m)
- ❖ Shape factor of the catchments $F(L^2/A)$

The development of the relationship between the mean annual flood or index flood and the catchment characteristics was a necessary step in predicting flood magnitudes at any point in a region where the frequency curve has been derived. The quantiles at ungauged sites can be estimated by first estimating the magnitude of the index flood and then multiplying it by ratio that can be found from the regional frequency growth curve for various return periods.

The ratio X_T , $X_T = \frac{Q_T}{\bar{Q}}$ can be found from the regional Standardized Growth curve of its respective region for different return period.

Q_T is computed from the regression model of the region

Then the Quantile discharge for return period T can be computed using

$$Q_T = \bar{Q} * X_T$$

The regression models relating to catchments characteristics were developed for each individual region in the study area. The simple linear regression models have been developed for those regions where the establishment of the regression relationship involved only mean annual floods and catchment area. In the cases where two or more independent variables were applied, the selection of the best regression model was based on the computed value of the coefficient of multiple regressions, R^2 which expresses the proportion of the variance of the MAF which is explained by the regression.

The regression analysis is done in the form:

$$\bar{Q} = b_0 A^{b_1} S^{b_2} P^{b_3} F^{b_4}$$

Where: A, S, P, F are independent variables (Catchment characteristics Area, Slope, Rainfall and Shape factor respectively)

To make the analysis simpler, the above equation is transformed to a linear form using logarithmic transformation.

$$\log Q = \log b_0 + A \log b_1 + S \log b_2 + P \log b_3 + F \log b_4$$

Then the estimation was performed on the parameters namely b_1 , b_2 , b_3 , b_4 and $\log b_0$.

Table 5.18- Derived regression equation for each regions to predict Mean Annual Flood (MAF) for ungauged catchment

Region	Regression equation	R^2	Remark
One	$\bar{Q} = 3.27A^{0.486}$	0.77	
Two	$\bar{Q} = 1.082A^{0.631}$	0.72	$A < 5000 \text{Km}^2$
	$\bar{Q} = 1.623A^{0.631}$		$A > 5000 \text{Km}^2$
Three	$\bar{Q} = 41.47A^{0.148}$	0.68	$A < 1000 \text{Km}^2$
	$\bar{Q} = 29.02A^{0.148}$	0.67	$A > 1000 \text{Km}^2$
Four	$\bar{Q} = 63.89A^{0.0054}$	0.75	$A < 500 \text{Km}^2$
	$\bar{Q} = 170.37A^{0.054}$	0.66	$A > 500 \text{Km}^2$
Five	$\bar{Q} = 4.67A^{0.65}$	1.00	$A < 300 \text{Km}^2$
	$\bar{Q} = 1.41A^{0.65}$	1.00	$A > 300 \text{Km}^2$

CHAPTER-6

SUMMARY CONCLUSION AND RECOMMENDATIONS

6.1 Summary and Conclusion

The study has demonstrated regional flood frequency analysis (RFFA) for gauged catchments in the upper and part the middle Awash sub-basin (u/s of Kesem) through the application of index flood method. One of the objectives of the study was to test the homogeneity of the stations and regions in the sub-basin which leads into delineation of the sub-basin into hydrological homogeneous regions, which would form the basic units to develop frequency curves for each region and station.

The collected time series data were checked for consistency using mass curve and filling and extension was done using regression analysis The discharge data was also tested for independence by using W-W test and that showed serial autocorrelation (dependence behavior) were discarded from further analysis.

Despite many attempts made by research hydrologists to delineate homogeneous regions, no general methodology is accepted universally. Here regionalization was made on the statistical values of LCs of each stations based on the concept that stations from the same region, their index flood series come from the same parent distribution. The proposed five regions satisfied the homogeneity test applied in the study .Even if two of the stations are under the influence of unnatural flow conditions and very plate area of flow uncontrolled by the gauge .They are homogeneous which can be categorized in region two and three that of Akaki (31021) and Awash below koka dam(31017) especially Awash below Koka dam is highly influence due to excessive release from Koka reservoir.

Software have been used for the selection best fit distributions; Most of the stations satisfy the goodness of fit test , and those fitted distribution combined with the three parameter estimation methods and tested with standard error of estimation to identify the most robust combination of distribution with that of parameter. The one with smaller standard error adopted for quantile estimation. Finally the quantiles estimated for region one with GEV/MOM, region two Gama/MOM, region three P-III/MOM, region four Gama/MOM and region five P-III/MOM this shows us MOM found to be the vigorous method for estimation of parameter.

With the selected distributions and parametric estimations Quantile versus return period graph is developed for each stations and regions with dimensional and standardized flow .

Regression analyses have been applied to develop regression models to predict mean annual flood from ungauged catchment and to predict quantiles for various return periods with the help of standardized frequency curve.

6.2 Recommendations

Based on the result obtained in this research, the following points are forwarded:

I. Index flood method of regional FFA is powerful means of approaching for estimation of flood quantiles used for the purpose of hydraulic structure including for ungauged catchment.

II. Delineation of homogeneous regions based on statistical parameter of gauged site could be one of an alternative method of regionalization to identify stations of similar flood producing characteristics.

III. Catchments with similar topography have follow similar frequency curve.

VI. It is advisable to extend this approach of RFFA for other parts of Awash and for the other Ethiopian river basins. Specially for that of Omo river basin next to Abay in water resource and very huge new project establishment and its flood nature in the border areas.

7. REFERENCES

Chow, V.T., Maidment, D.R. and Mays. L.W.(1988) Applied Hydrology McGRAW, Singapore

Chow V.T (1964) Handbook of Applied Hydrology, McGraw – Hill company USA.

Cunnane, C. (1989) Statistical Distributions For Flood Frequency Analysis.

World Meteorological Organization Operational Hydrology Report, No.33.

Cunnane, C, (1985) Hydrological frequency and Time series Analysis, section B.

volume-1

Gebeyehu A. (1989) Regional Flood Frequency Analysis PHD, thesis report

Hosking, J.R.M. and Wallis, J.R. (1993). Some Statistics Useful in

Regional Frequency Analysis. Water Resource Research, Volume 29, No.2.

New York.

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

LLC, Florida.

Jayarami R. (1987). Stochastic Hydrology ,Deepak printing service, Delhi.

Sine .A. (2004) regional Flood Frequency model for Blue Nile thesis report .

Paulose , (2006) the release of Koka dam on dry season

Leuleseged (2002) At – site and regional flood frequency Analysis MoWR

Gebricle and N. Arenll (1911) water resource research . A hierachical approach to regional flood frequency Analysis

Mkhandi et al, (2000) regionalization In flood frequency analysis

Pitilick (1994) Drought and Flood Frequency Analysis

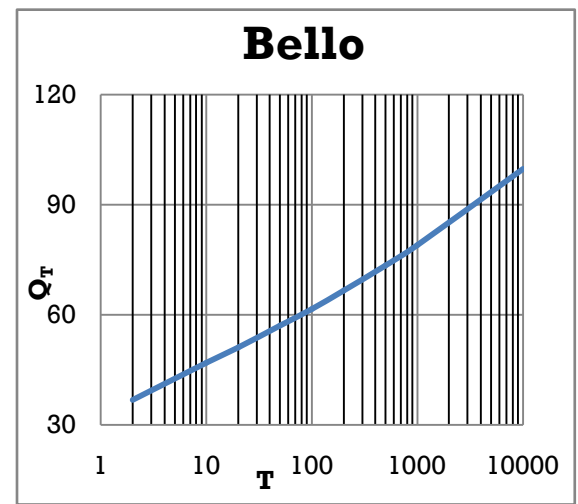
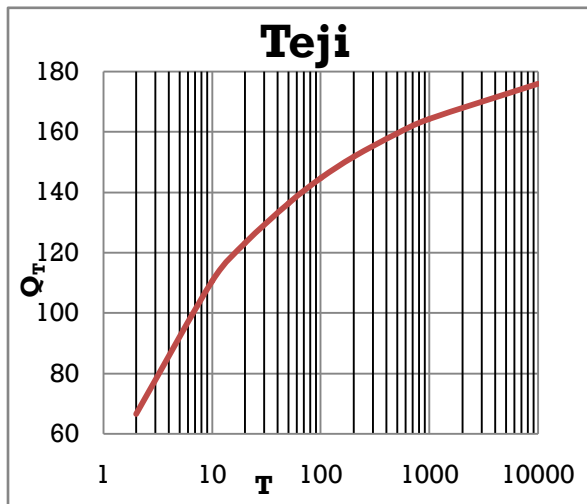
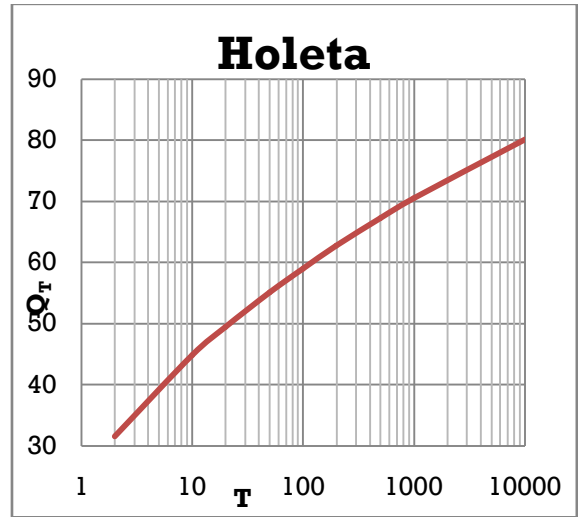
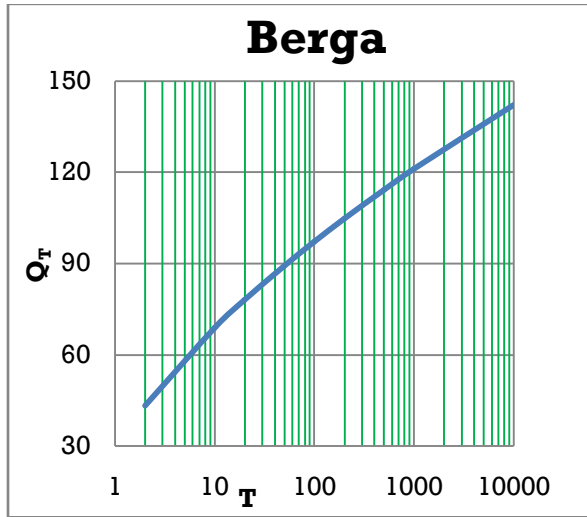
8 - ANNEXES

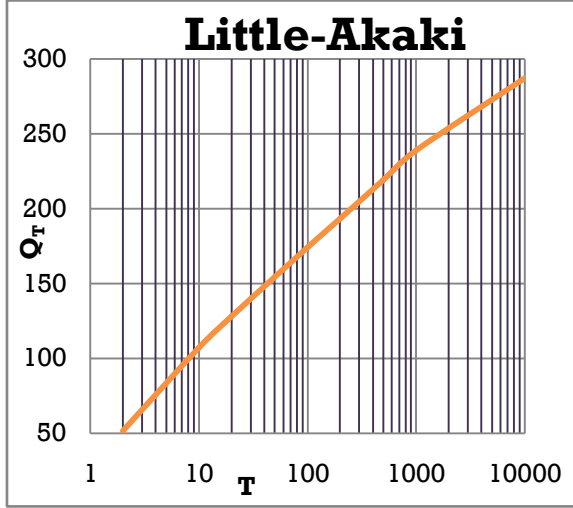
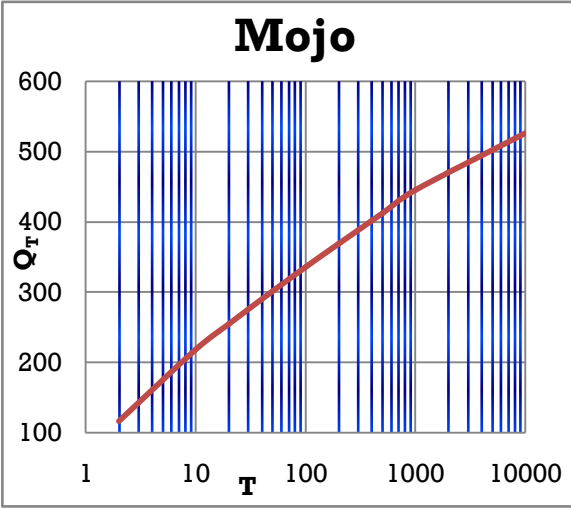
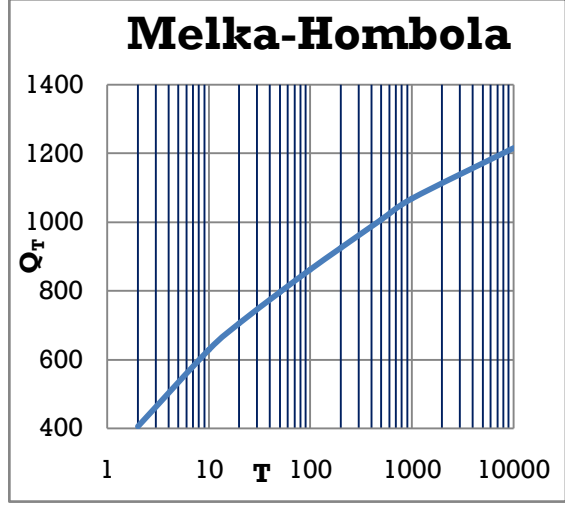
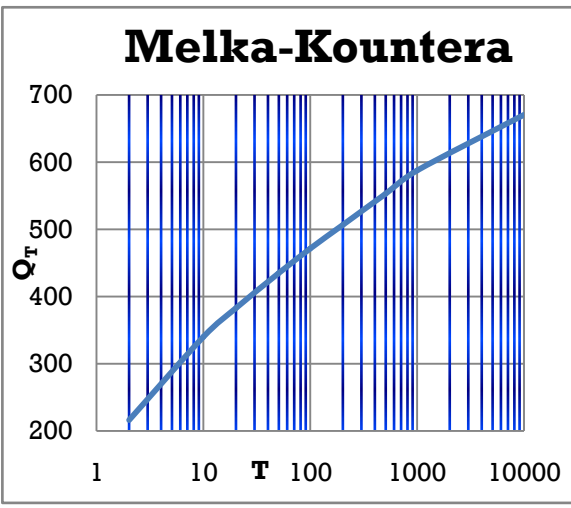
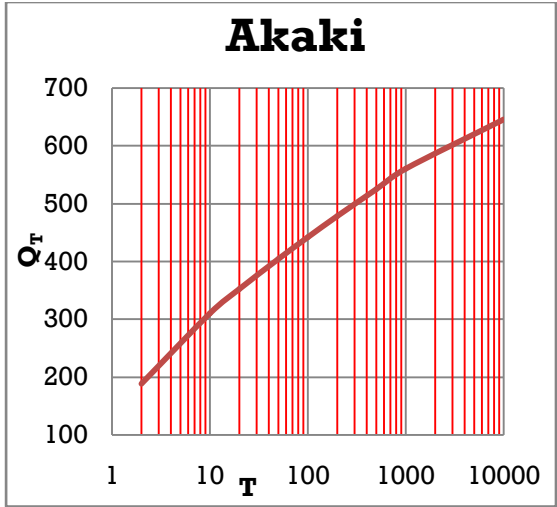
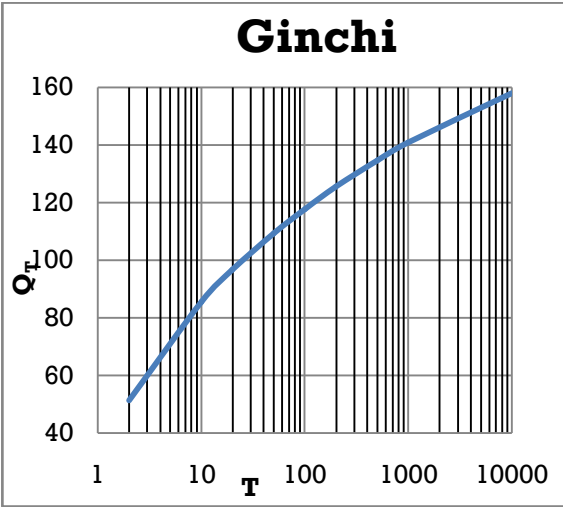
8.1- ANNEX-I

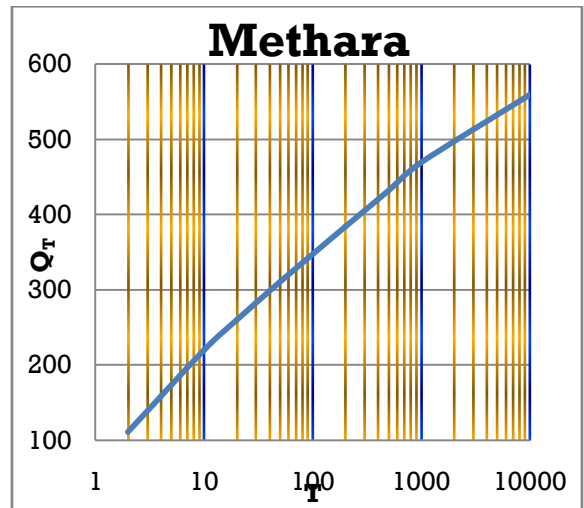
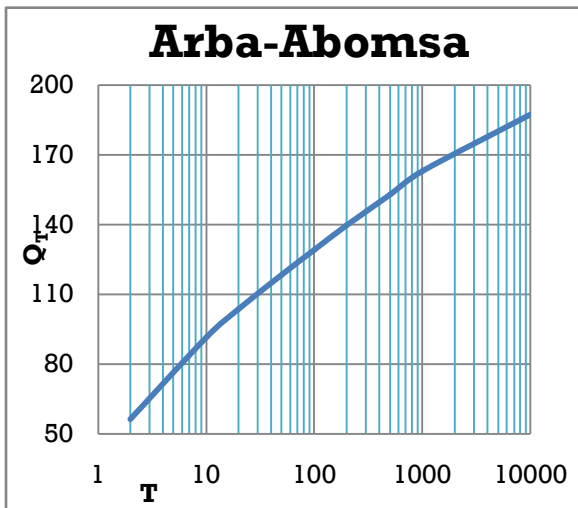
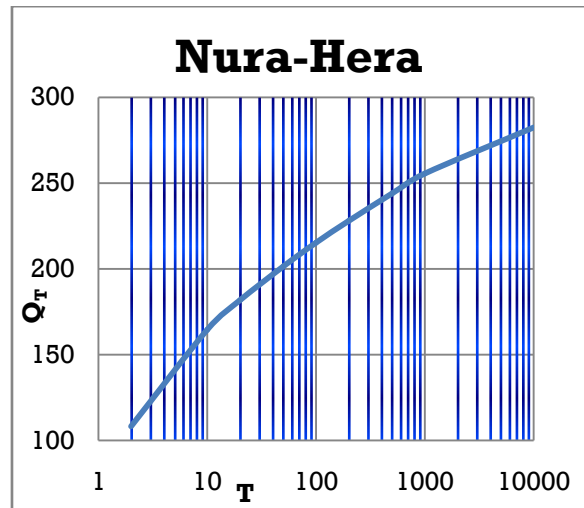
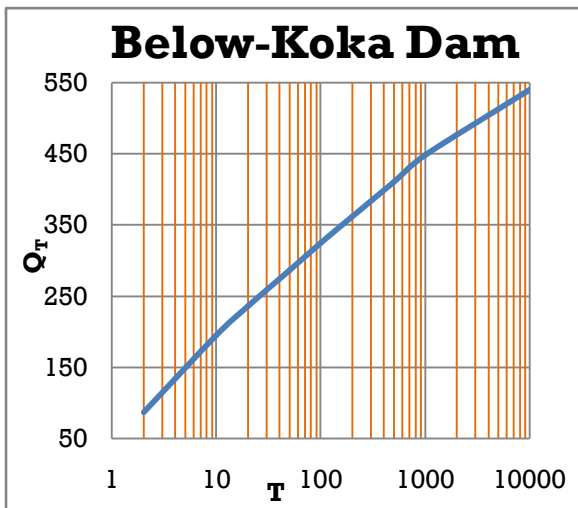
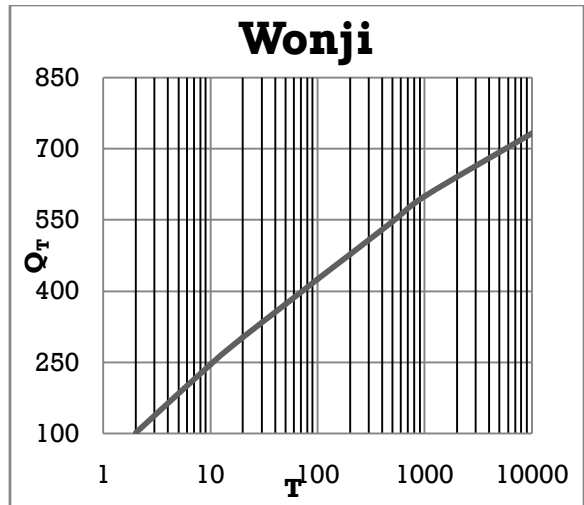
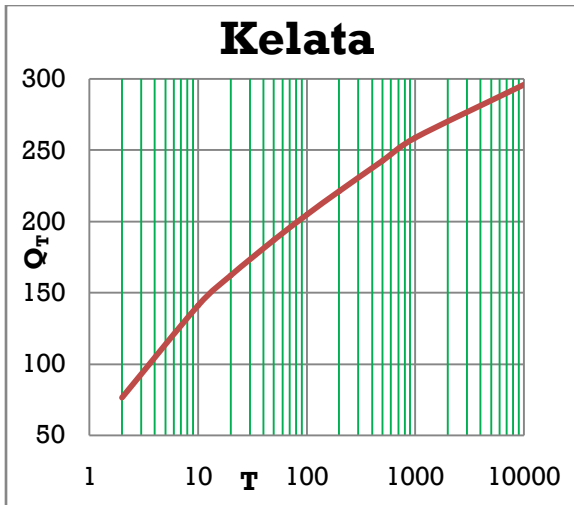
Growth curve of each station including the regions

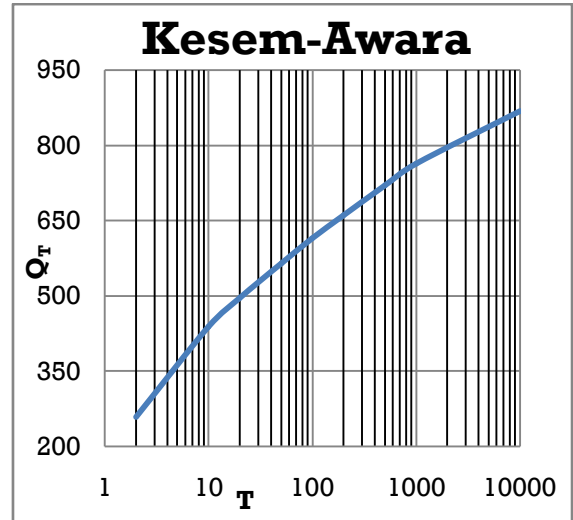
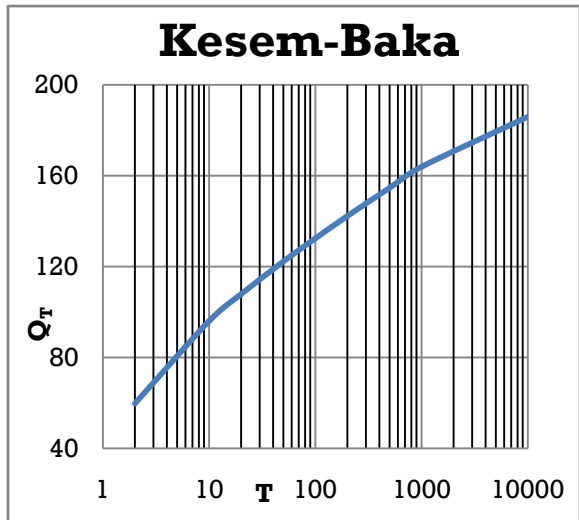
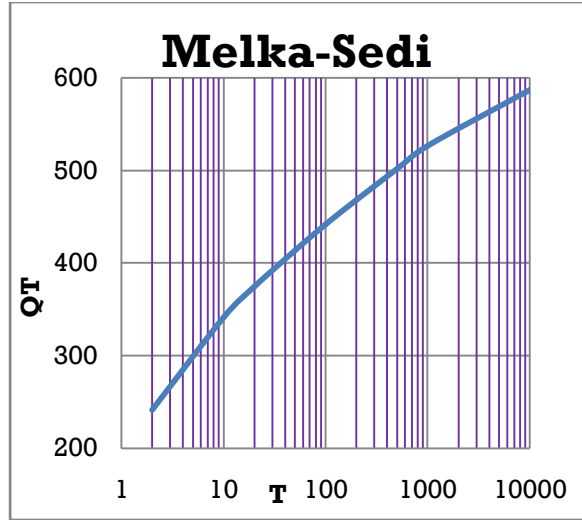
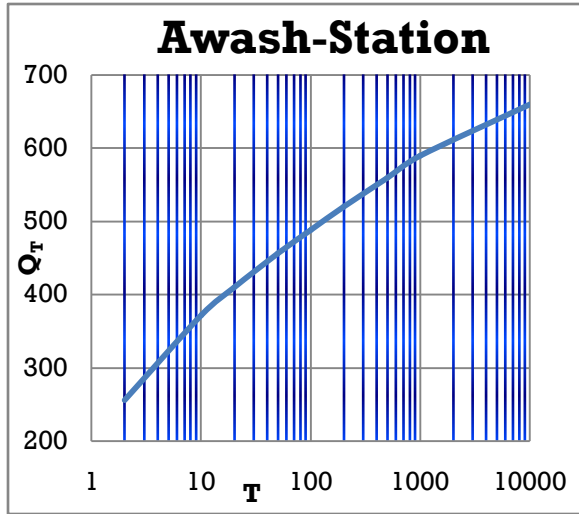
Dimensional flow (Q_T)

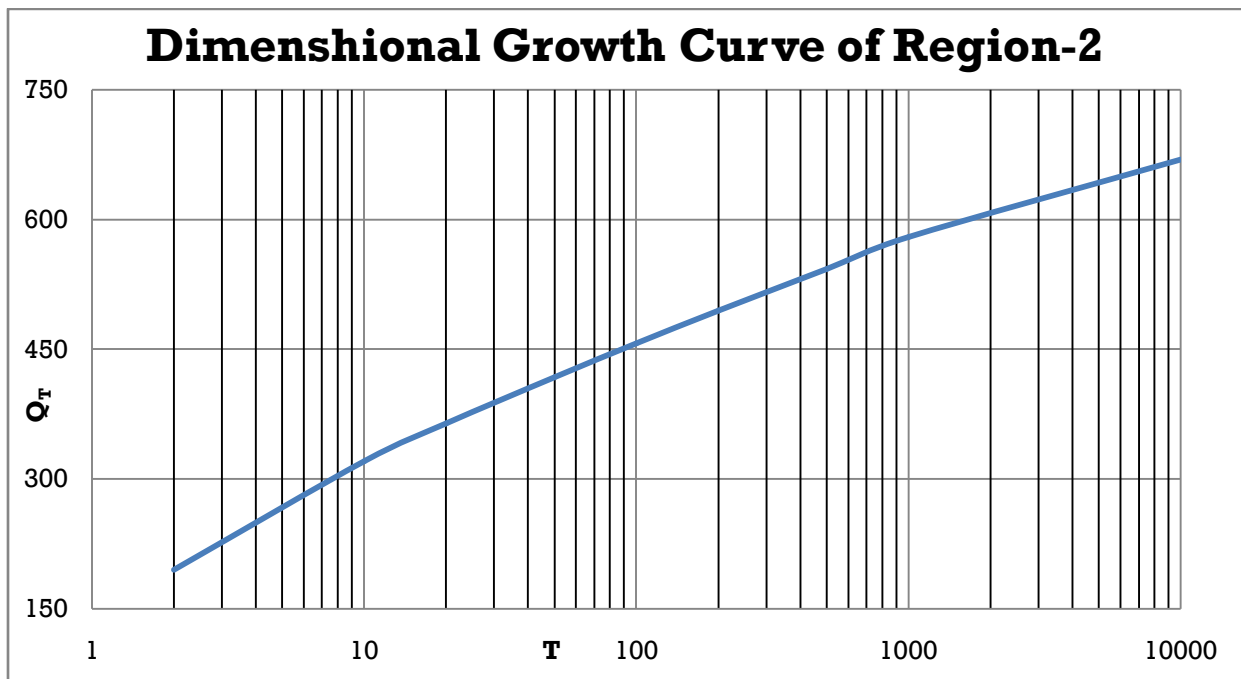
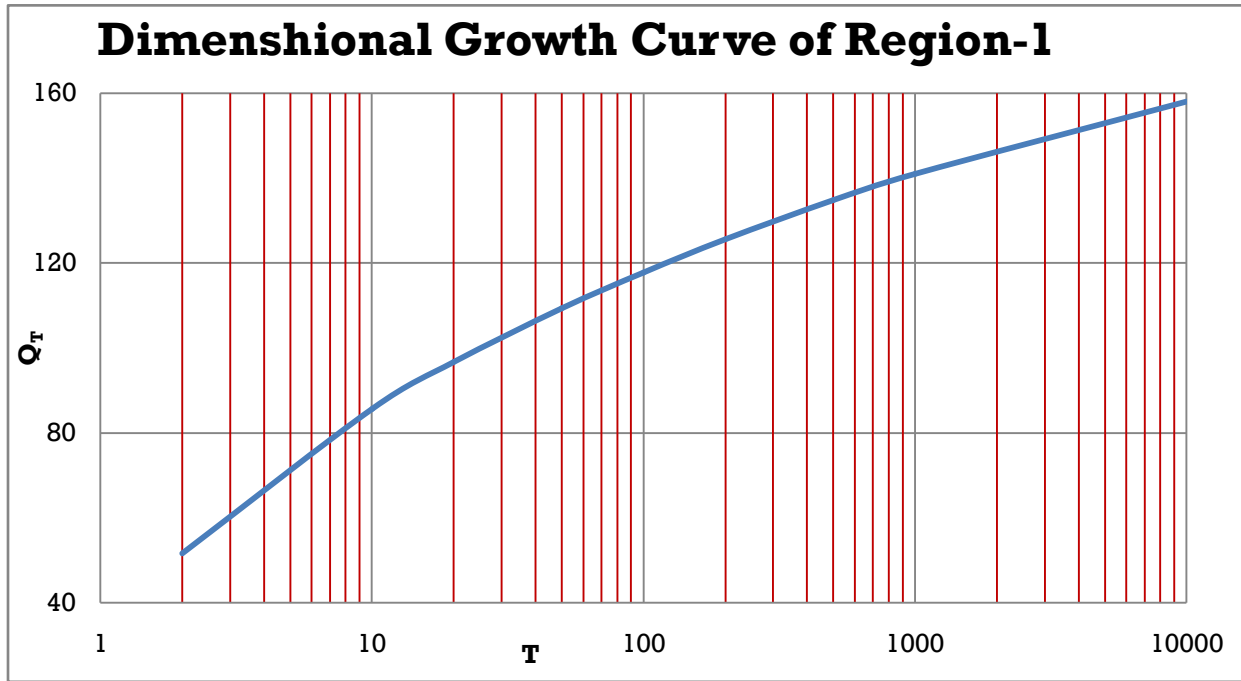
(Q_T in m^3/s and T in Years)



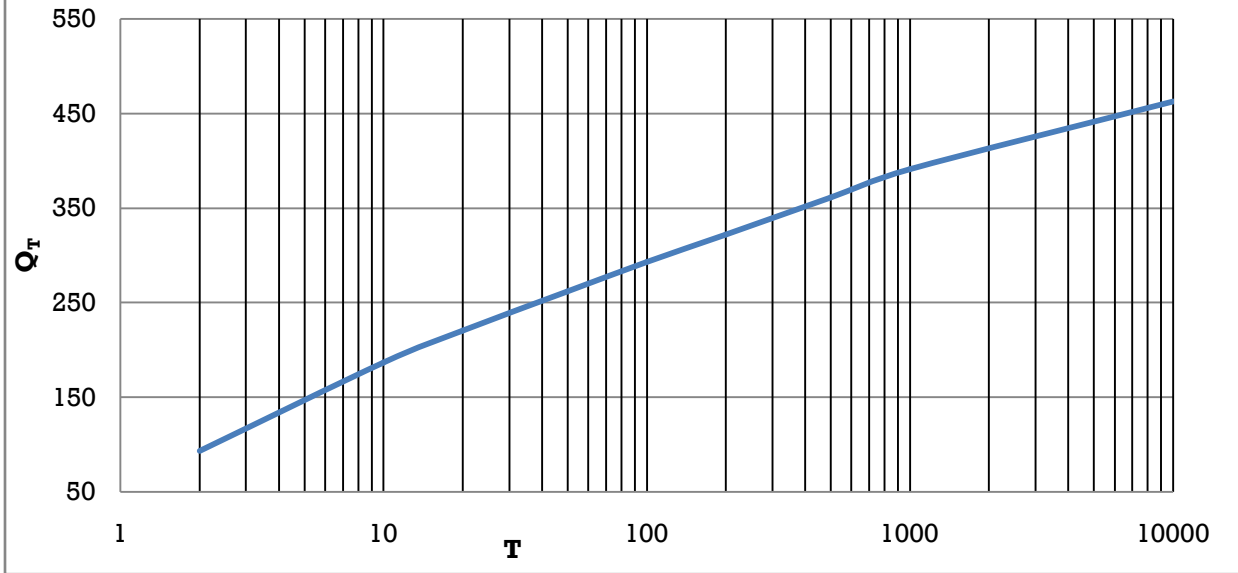




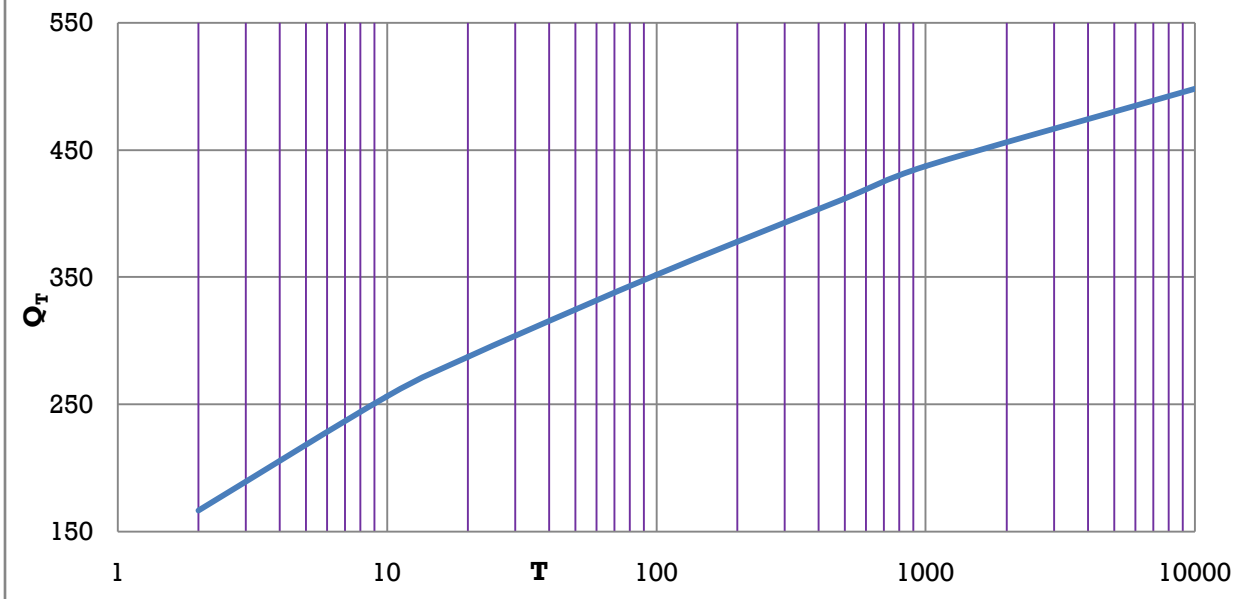


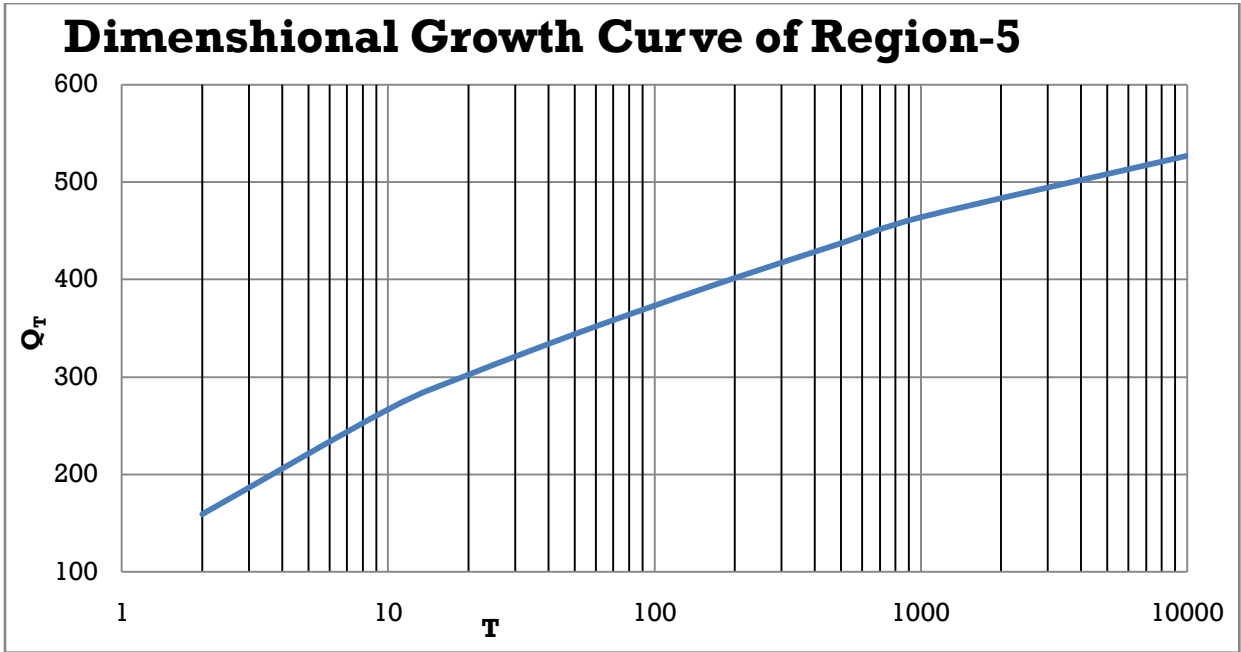


Dimensional Growth Curve of Region-3

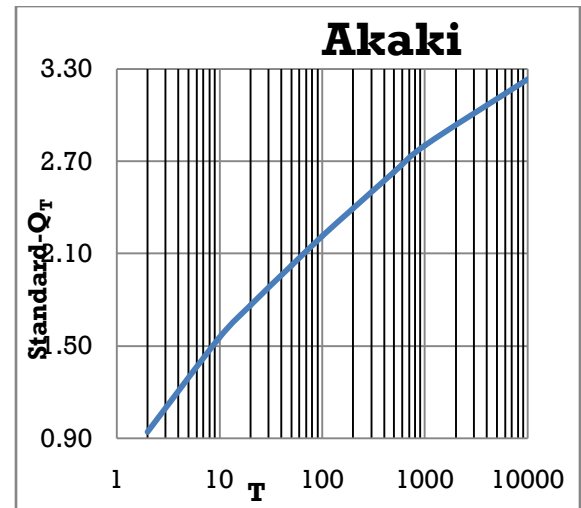
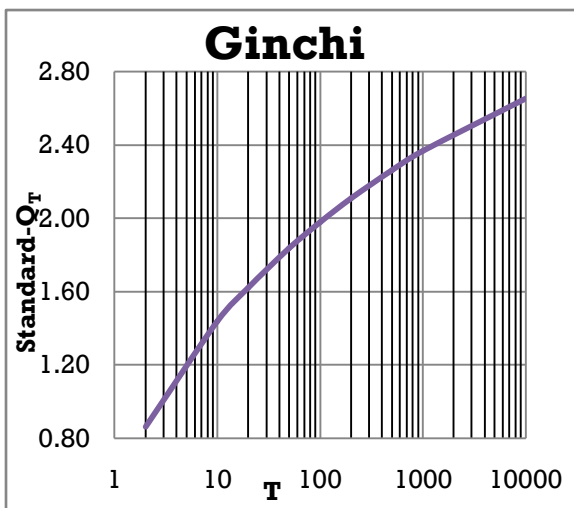
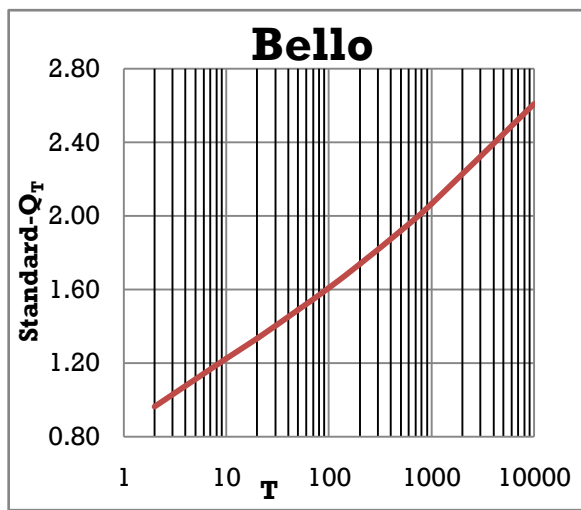
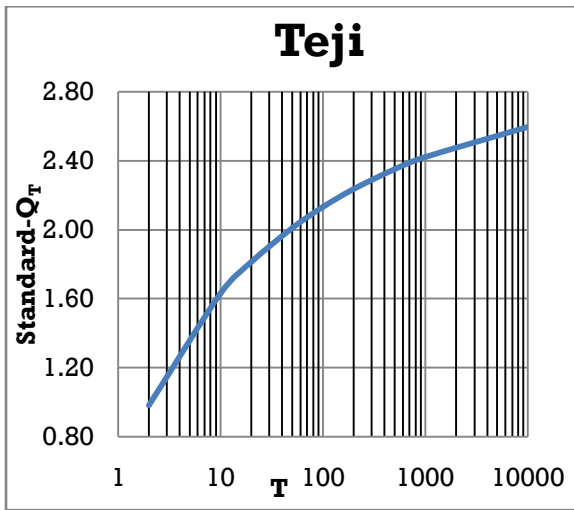
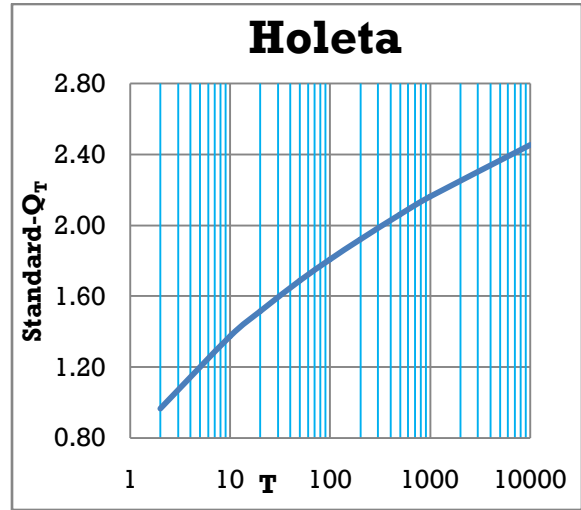
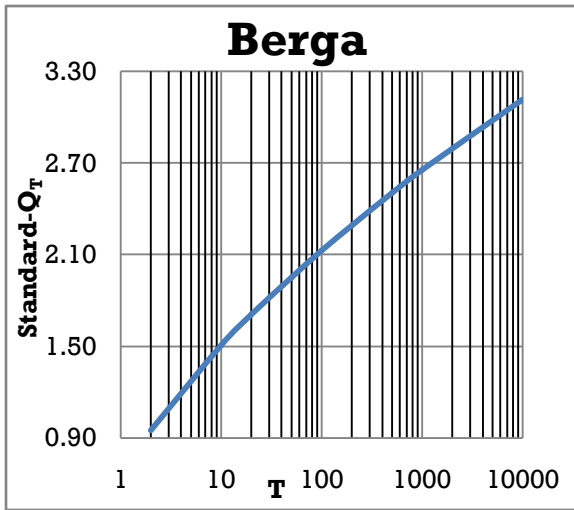


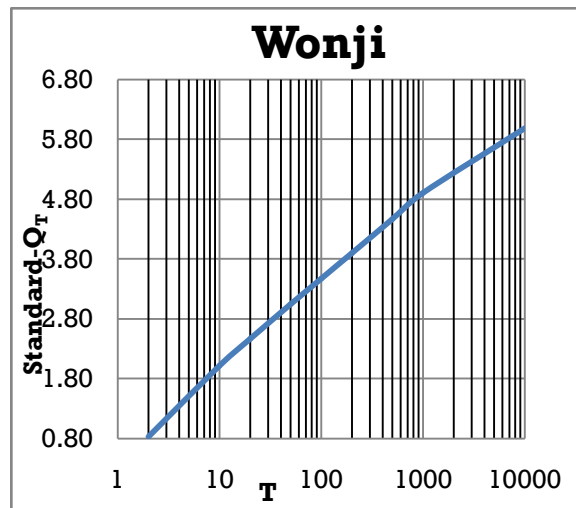
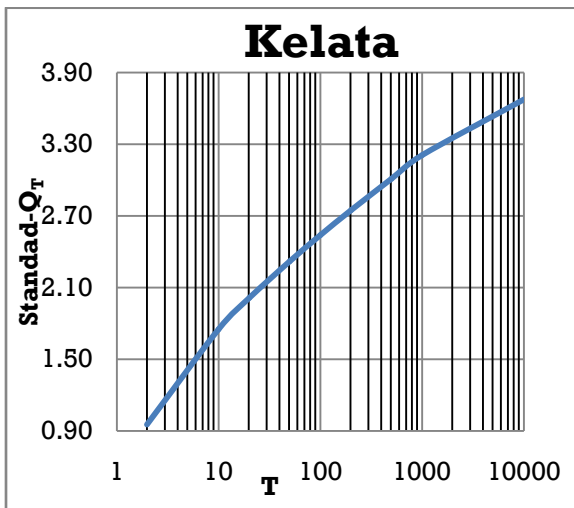
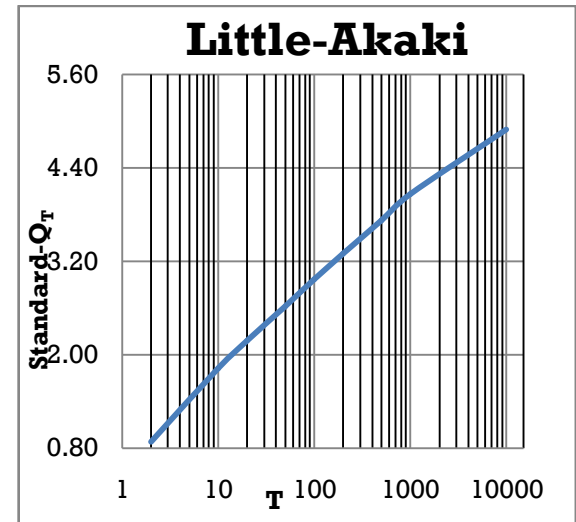
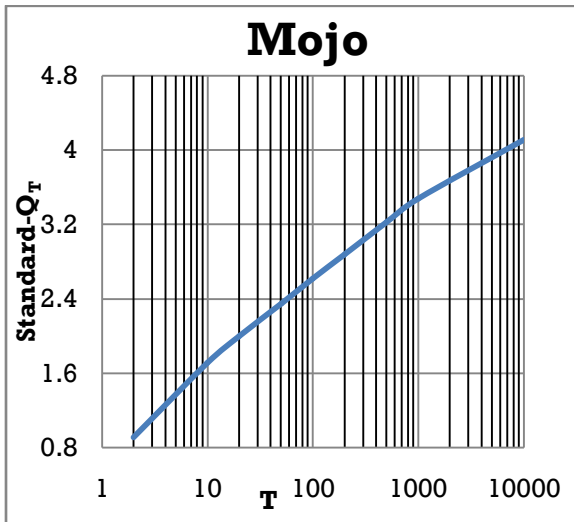
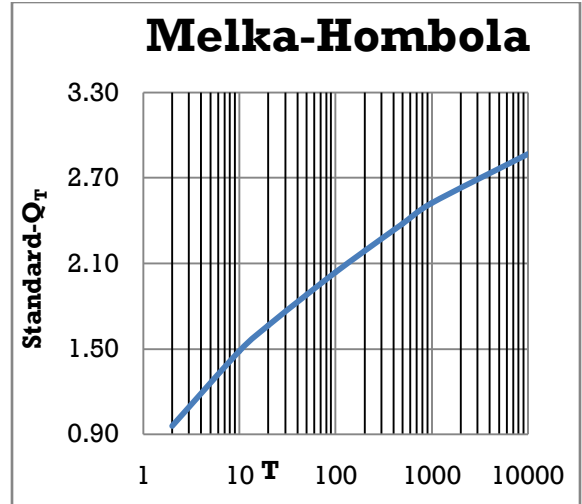
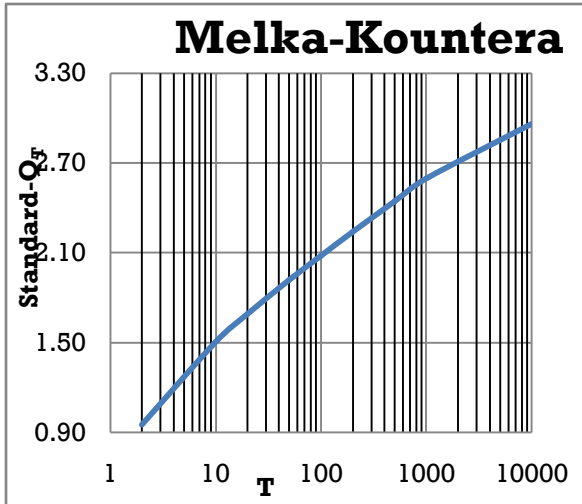
Dimensional Growth Curve of Region -4

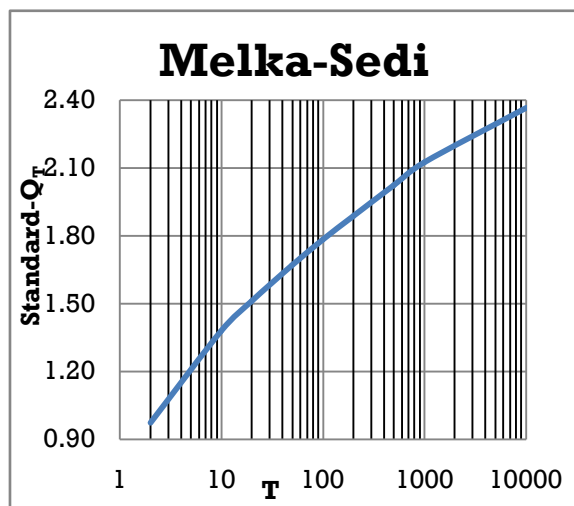
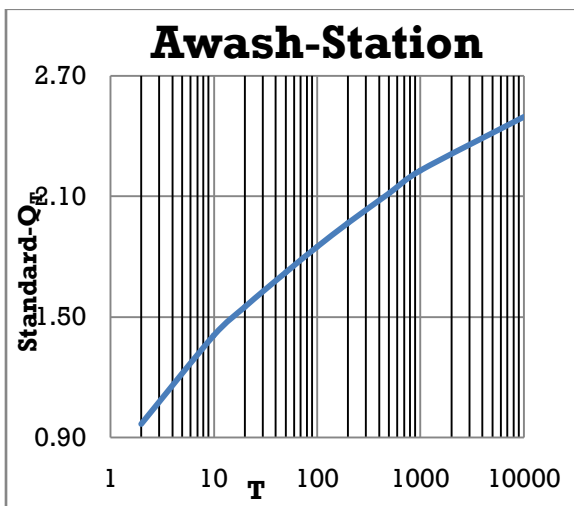
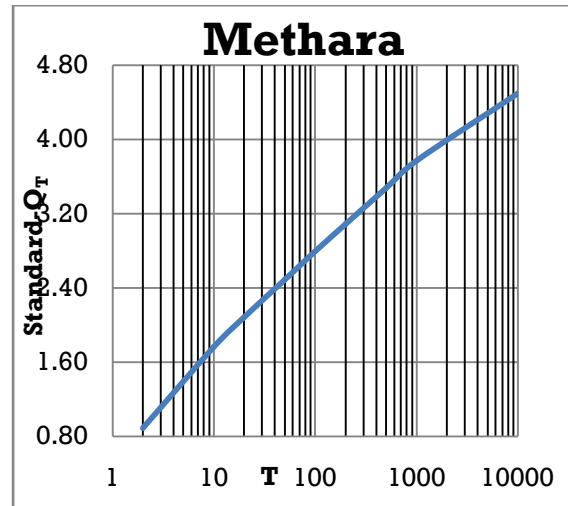
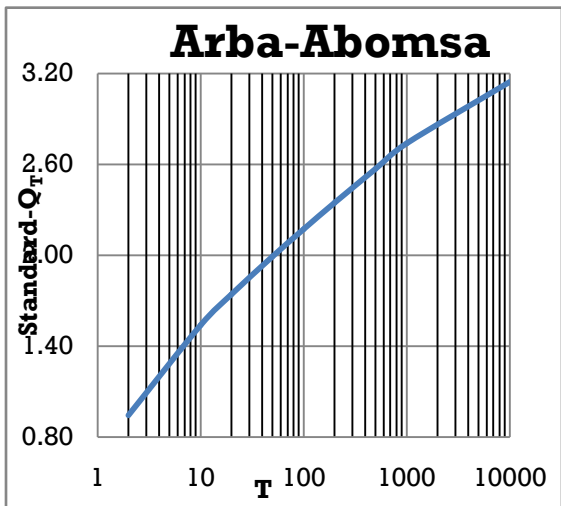
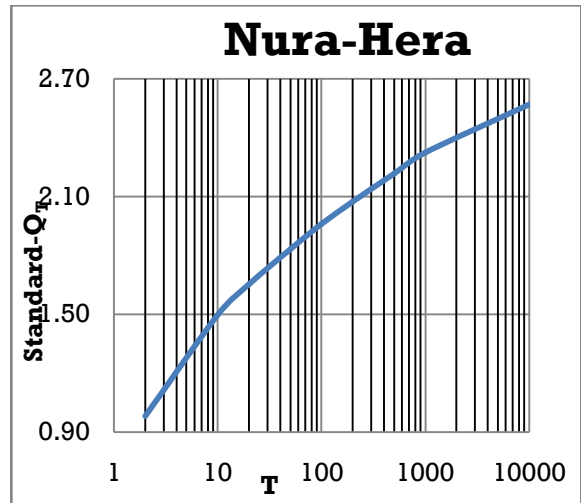
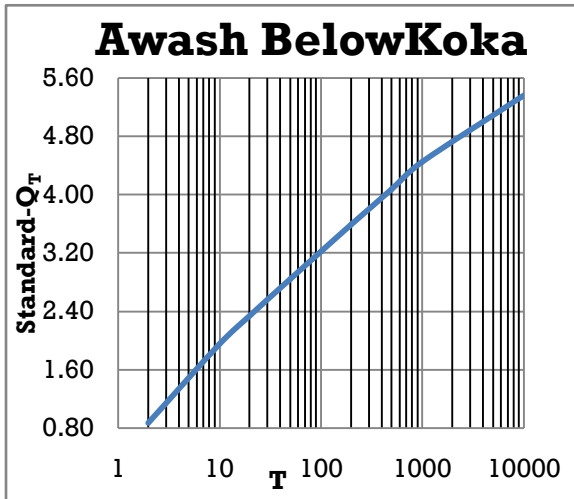


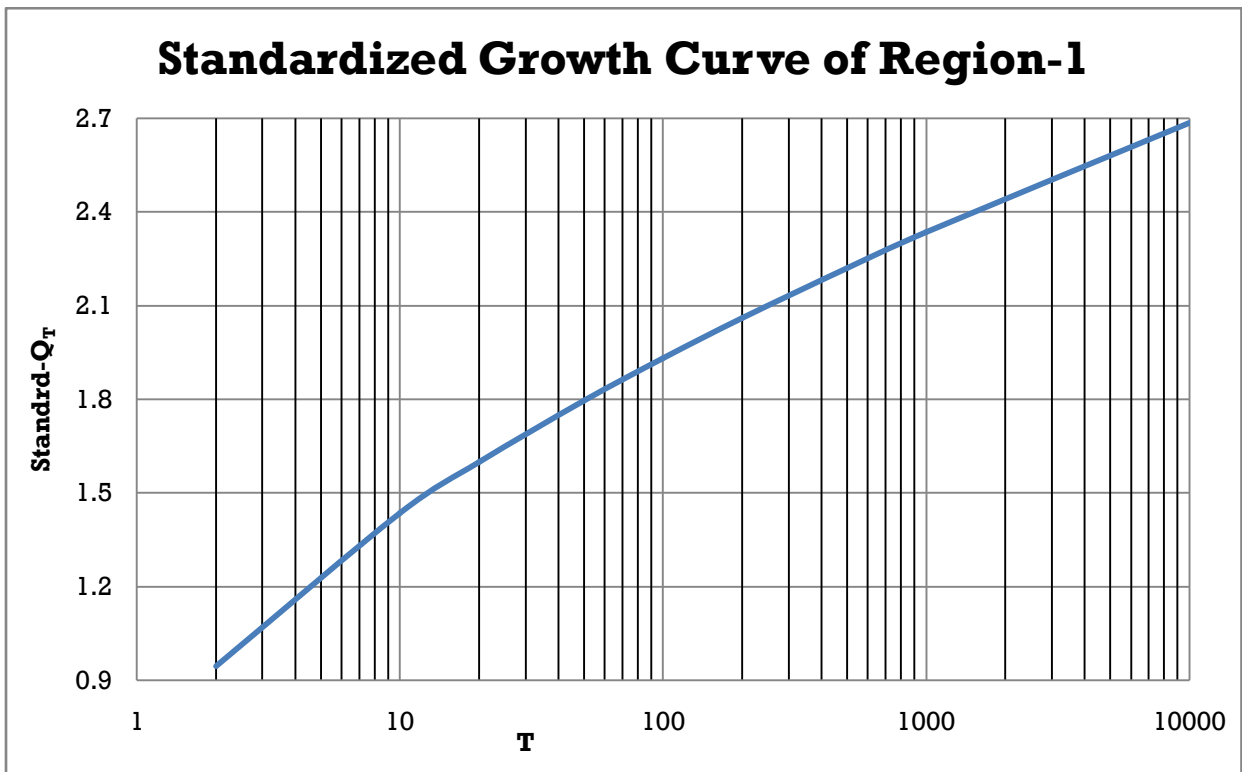
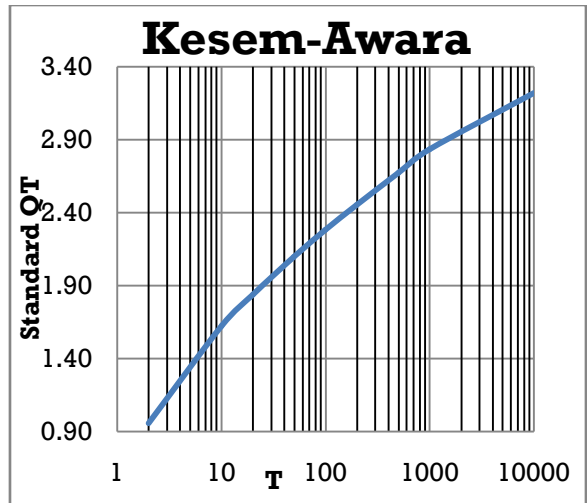
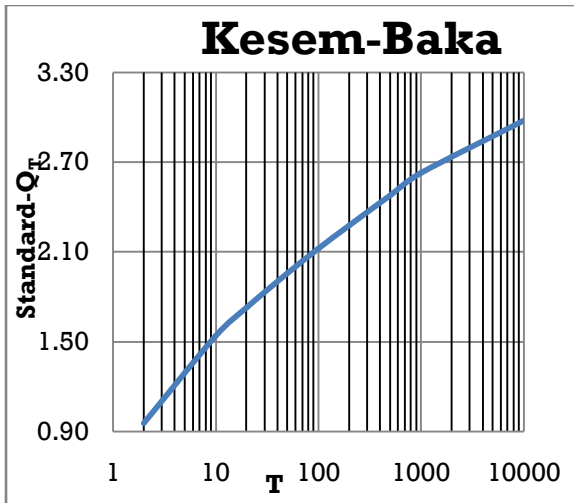


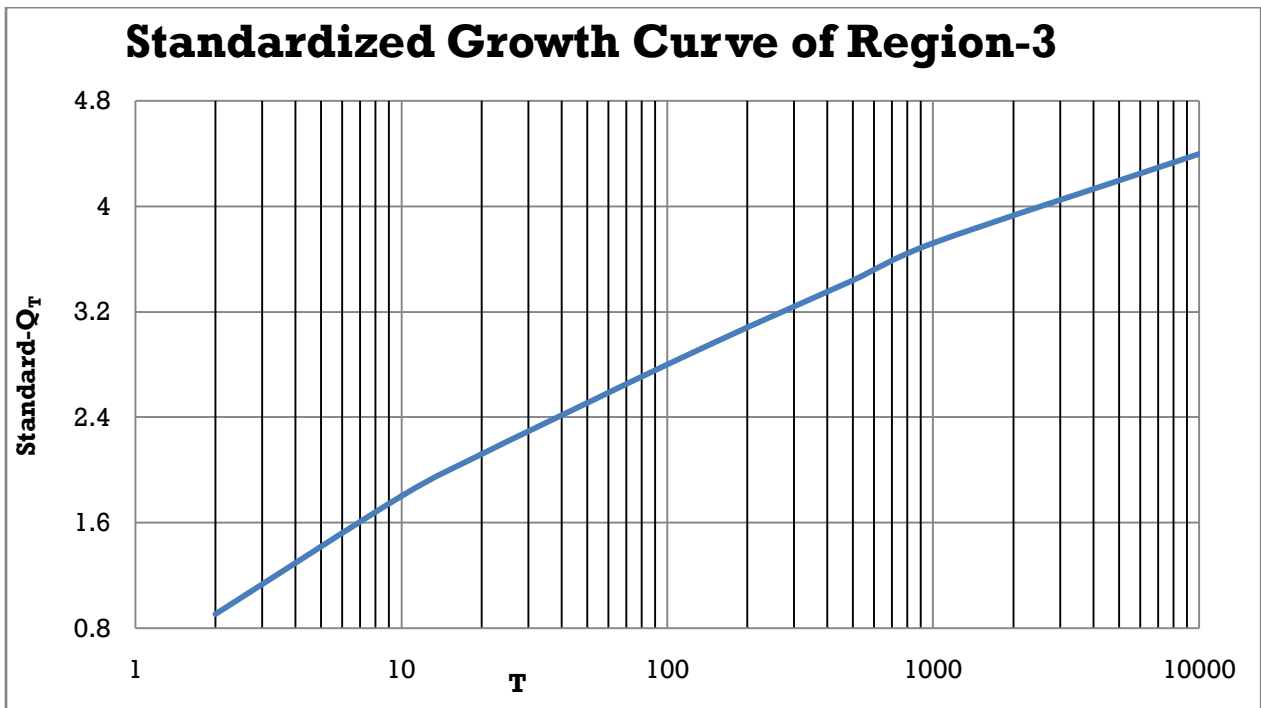
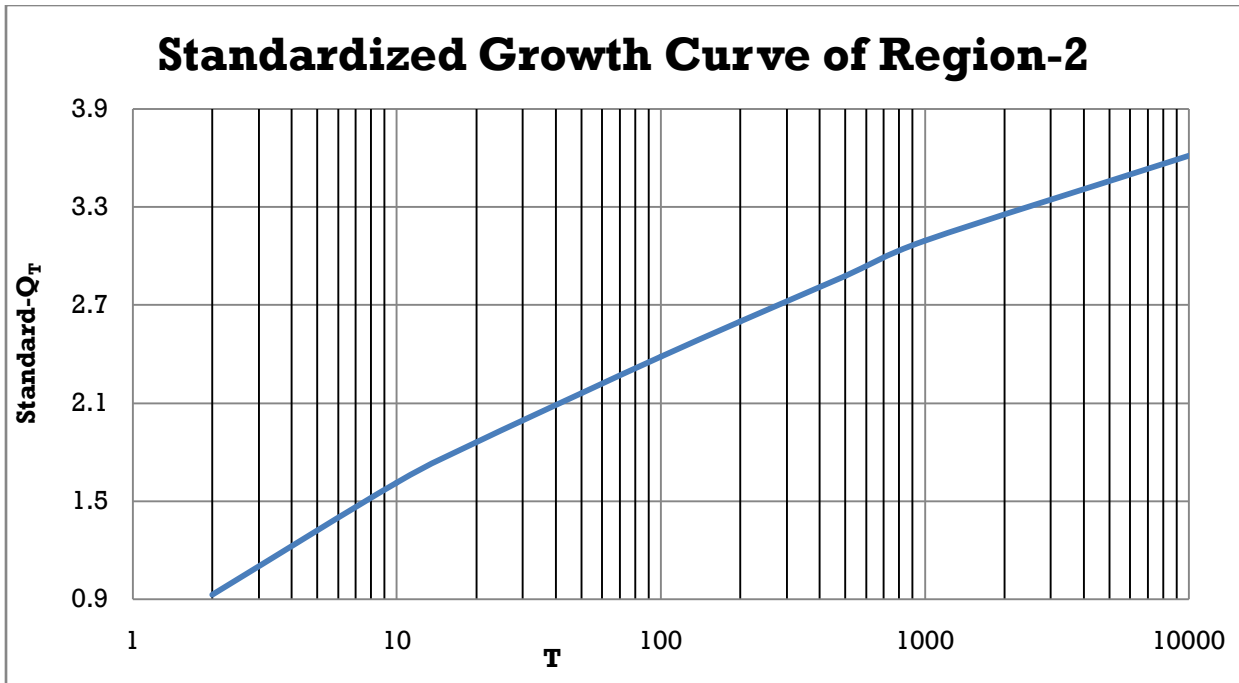
Standardized flow (Q_T) ($Standard - Q_T = Q_T/Q_{mean}$ and T in year s)

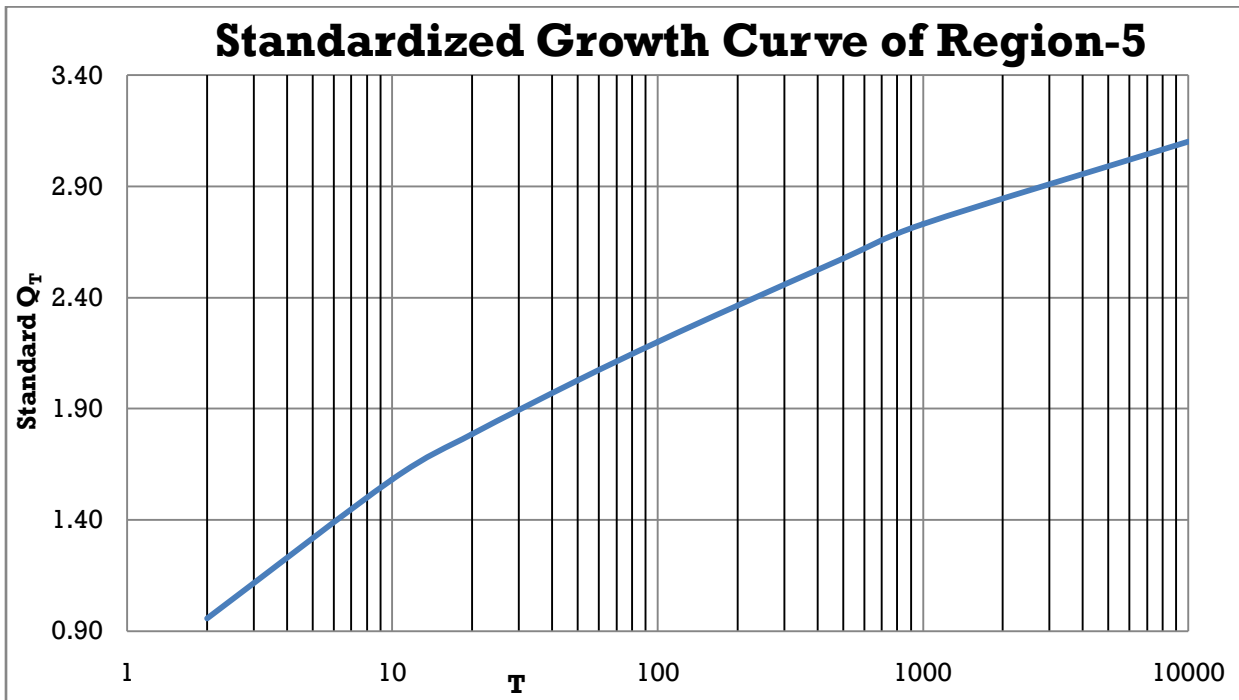
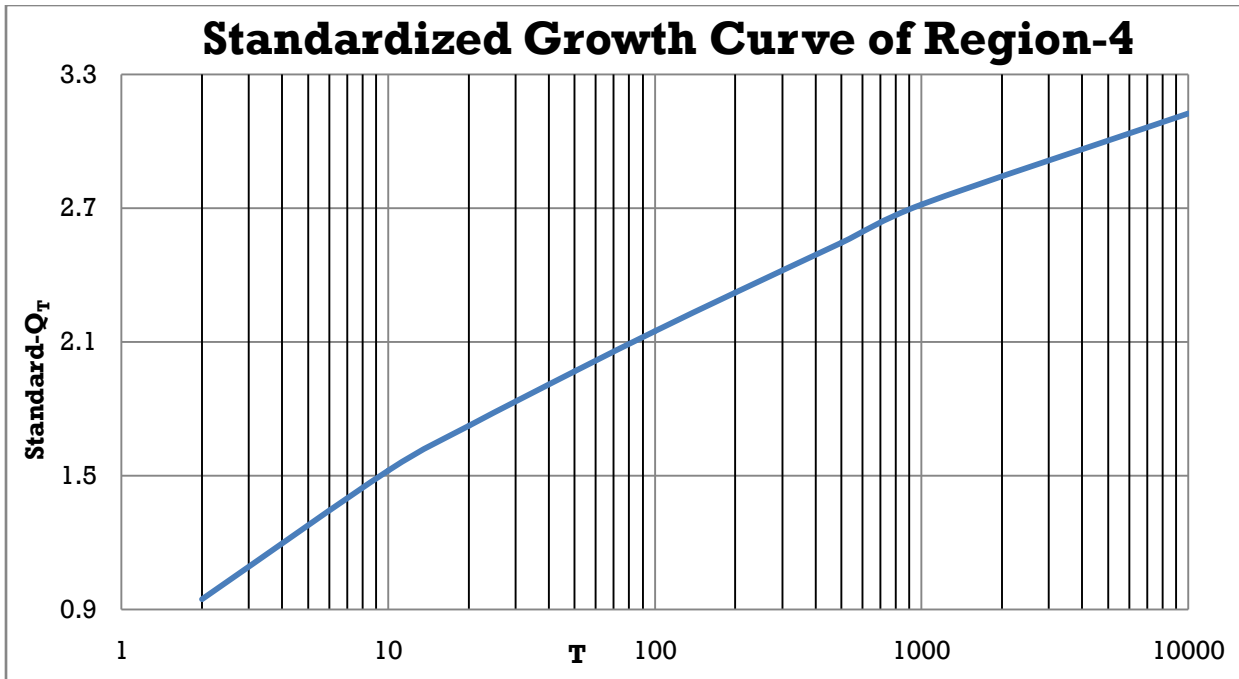












8.2 ANNEX-II

Map of the study area

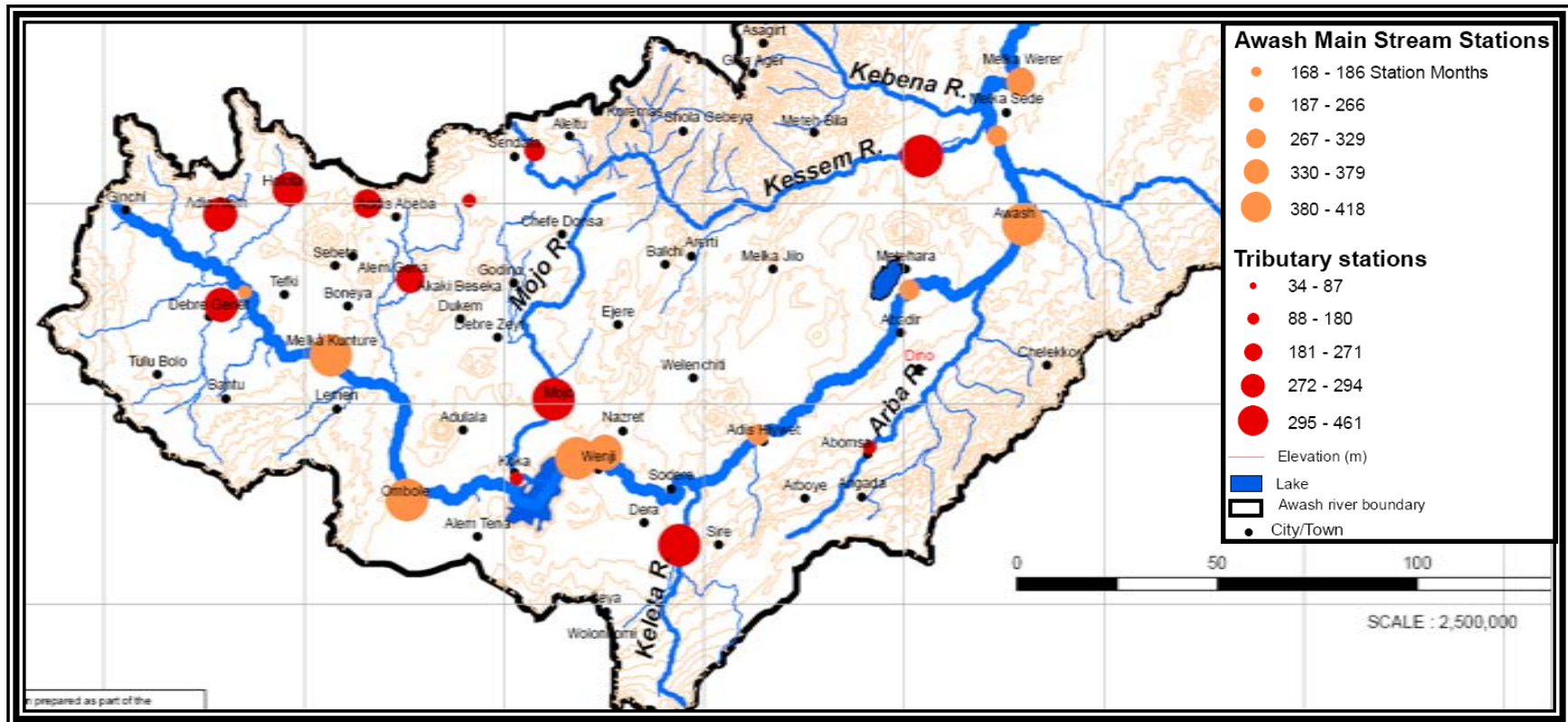


Fig Hydrometric gauging stations in the study area (Source MoWR)

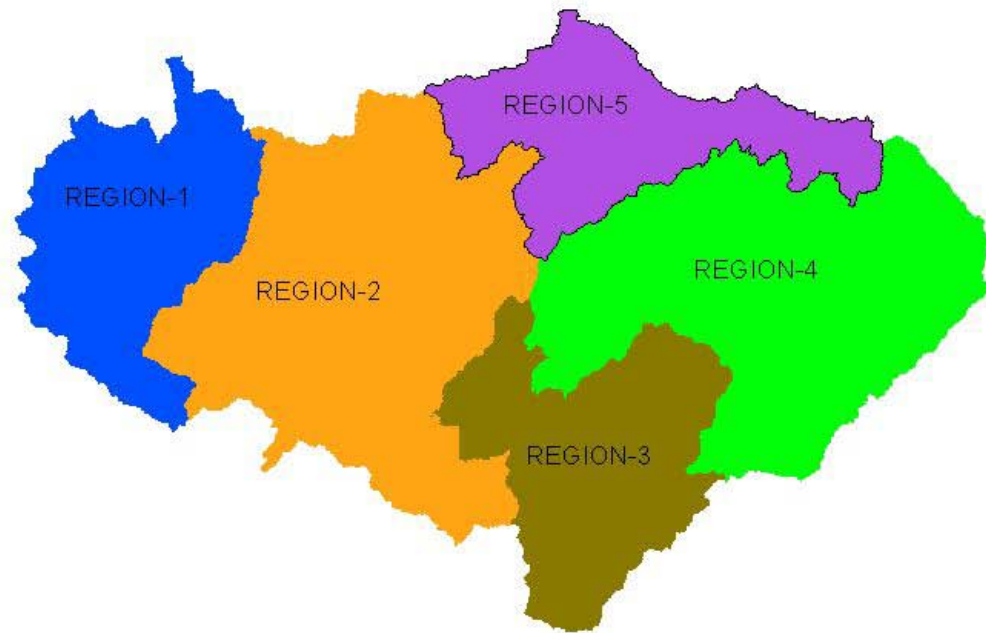


Fig Homogeneous regions division

8.3 – ANNEX -III

The following are the metrological stations used for the thesis:

1. Ginchi
2. Welen Komi
3. Addis Alem
4. Kimoye
5. Asgori
6. Teji
7. Akaki
8. Melka Kuntura
- 9 .Dertu Liben
10. Mojo
11. Hombola
12. Nazreth
- 13.Wolenchita
14. Nura era
15. Shola Gebaya
16. Methara
17. Arba Bordada