Journal of Resources Development and Management
ISSN 2422-8397 An International Peer-reviewed Journal
Vol.24, 2016

# Regional Flood Frequency Analysis for Abaya – Chamo Sub Basin, Rift Valley Basin, Ethiopia

Behailu Hussen [1] Negash wagesho[2]
1.Water Resources Research Center, Arba Minch University, PO box 21, Arba Minch, Ethiopia
2.Scientific Director Of Arba Minch Institute Of Technology, Arba Minch University, PO box 21, Arba Minch, Ethiopia

## Abstract

The purpose of this study is to develop the regional flood frequency (growth) curve and index flood estimation equation (model) essential for computing flood quantiles for ungauged catchments in Abaya-Chamo sub basin by using an index flood method. Annual series model is used for the analysis. Identification of homogeneous region was performed initially on the basis of catchment characteristics like slope, elevation, soil type, soil texture and mean annual rainfall and flood statics are used for test of homogeneity of the proposed region. 17 rainfall stations were selected for analysis based on length and quality of the data they have. Concerning on the above procedure the Abaya-Chamo sub basin was divided in to three regions. Region one is located in the South-West part of the sub basin which covers the area of 4583sq. km, region two is located at the middle of the sub basin with an area of 5701sq. km and region three located at North-east and South-West part of the sub basin also covers an area of 7259sq. km. For the above regions selection of parent distribution was performed on the bases of the three common methods which are probability plot by using SPSS software, statistical method like Chi- square test, Kolmogorov-Smirnov test and moment diagram (both conventional and L- moment). From the candidate distribution from all methods best fit distribution is selected by using the software called Easy Fit test. For the selected distribution parameter estimation technique was selected by performing a Robustness Assessment. Accordingly, GEV with PWM selected for region one, LN (2P) with MOM is selected for region two, finally Gamma (2P) with PWM is selected for region three and the regional growth curve were developed for each regions by using the selected parameter estimation techniques. Accordingly the Abaya-Chamo sub basin was divided in to two regions and the index flood estimation equation were developed with a coefficient of determination ($R^{2)}$ 0.93 for upper region and 0.99 for lower region.

**Keywords**: Abaya-Chamo sub basin, index flood method, SPSS software, Easy Fit software.

## Introduction

Flood frequency analysis provides vital information for design of an economic appraisal of a variety of engineering and water resource planning and development projects. Frequency analysis of flood is a very active area of investigation in statistical hydrology. Various distributions, method of parameters estimation, quantiles estimation, and problems related to regionalization, at-site and other related topics are the main components in flood frequency analysis. The analysis involves estimation of a flood magnitude corresponding to a required return period or probability of exceedance. (Mengistu D., 2008).

In flood frequency analysis the objective is to estimate a flood magnitude corresponding to any required return period of occurrence. The resulting relationship between magnitude and return period is referred to as the Q-T relationship. Return period, T, may be defined as the time interval for which a particular flood having magnitude $Q_T$(also known as quantiles) is expected to be exceeded (Admasu, 1989). Return period is also referred to as recurrence interval. The magnitude of flood is inversely related to their frequency of occurrence, high floods occurring less frequently than moderate flood. (Admasu, 1989).

In flood frequency analysis the objective is to estimate a flood magnitude corresponding to any required return period of occurrence. The resulting relationship between magnitude and return period is referred to as the Q-T relationship. Return period T, may be defined as the time interval for which a particular flood having magnitude $Q_T$(also known as quantiles) is expected to be exceeded (Admasu, 1989). Return period is also referred to as recurrence interval. The magnitude of flood is inversely related to their frequency of occurrence, high floods occurring less frequently than moderate flood. (Admasu, 1989).

The use of regional information to estimate flood magnitude at site with little or no observed data has become increasingly important because many projects that 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. (Admasu, 1989)

Therefore, the objective of this study was to investigate an appropriate procedure for analysis of flood frequency for use in the basin. The results from the analysis can be used for the proper planning and design of water resource development projects in the sub-basin.

## Study Area and Data Availability

The Abaya - Chamo lakes basin of the rift valley crosses through Ethiopian midway in the north south direction. The basin comprises the two lower lining lakes, Lake Abaya and lake Chamo.

The main rivers draining into Lake Abaya are listed as: Gelana, Bilate, Gidabo, Hare, Baso, and Amessa. In addition a number of small brooks and ephemeral rivers inter into the Abaya Lake. The rivers draining into lake Chamo are listed as: Sile, Argoba, Wezeka, Sego in addition to the overflow from Lake Abaya which confluences with River Kulfo and eventually drains to Lake Chamo. The Abaya and Chamo Lake are hydro logically interconnected. An overflow from Lake Abaya flows in to Kulfo River that in turn ends up into the Chamo Lake (Seleshi, 2000).

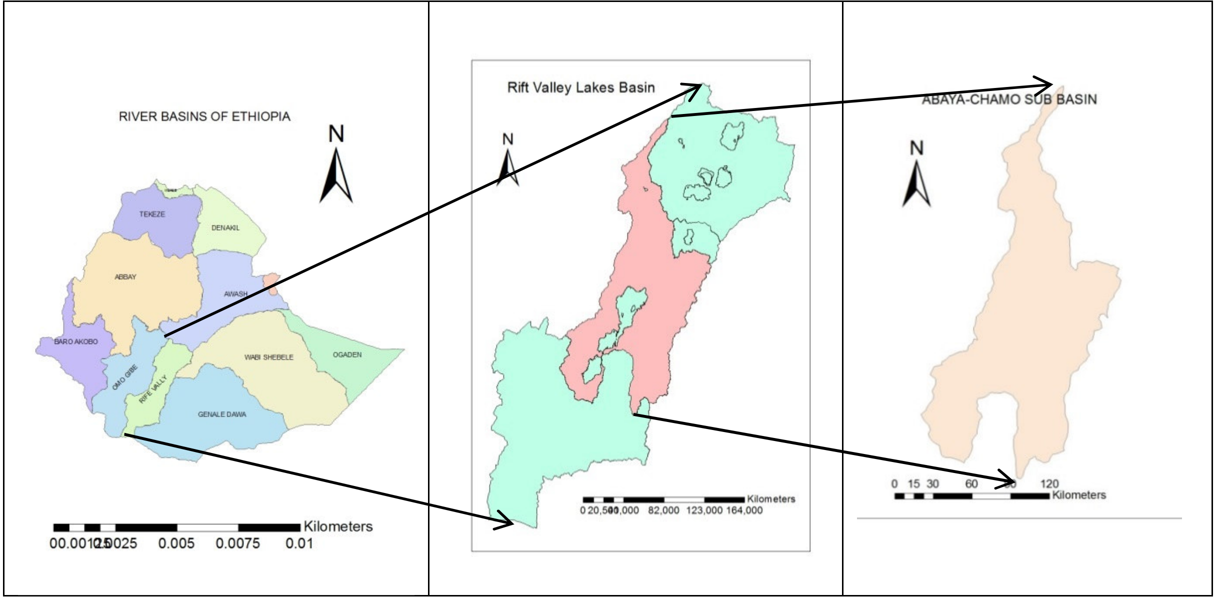

Figure 1: Location of Abaya-Chamo sub basin.

The level differences between the two lakes are 62 meters, Abaya Lake being higher than Chamo Lake. The altitude of the region varies between 4200masl (Mount Guge) and 1108masl (at the out flow from the Chamo lake). The region is located in the range $37^0$-$38^0$ in the eastern longitude and $5^0$-$8^0$ in the northern latitude. The two lakes have been used for transport (Lake Abaya), fishery and to and tourism purpose. The lakes have not been used intensively for irrigation (Ababu T. and Bernd W. 2004).

Two pattern of rainfall are observed in the basin. From Lake Abaya south wards the main rain occurs during the period of March to May. North Lake Abaya, the minimum rain occurs from July to September, often with a secondary peak in March of April. In the northern area, the rainfall pattern is generally typical of the area.

Temperature variability is large due to altitude difference. The mean annual temperature is estimated to vary from $27^0$c on the valley floor (chew bahir) to $10^0$c around 3000m contour.

### *Collection of important data*

Important data are collected from different institutions. Hydrological data and digitized map of the sub-basin are collected from the Ministry of Water irrigation and energy, from the department of Hydrology and GIS. Meteorological data such as temperature data and rainfall data are obtained from National Meteorological Agency of Ethiopia.

### Hydrological data

Stream flow data are collected from 17 stations in the basin. These all stations are not selected for the analysis because some of the stations have no enough data for the analysis. At least 15years of flow data are decided to be taken for flood frequency analysis in this study even though a guideline for FFA allows a minimum 10 years flow data (USWRC, 1976). For those stations whose number of record less than 15years, data filling and extension have been done by filling the missed flow data by using simple arithmetic average procedure.

### Meteorological data

Daily and yearly rainfall data are collected from 21 meteorological stations within the basin from National Meteorological Service Agency. List of meteorological stations within the sub-basin are listed in the table below.

### GIS and DEM data

GIS data of the catchment such as basin boundaries, sub-basins, lakes, river network, land use- land cover, soil type, slope and 90M resolution DEM of the basin are obtained from MoWIE data archive.

### Methodology

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:

- Collection of important data for the study such as hydrological data, methodological data, topographical and digitized map of the sub-basin.
- Checking of data for quality (independence, consistency and outliers test)
- Filling and extension of flow and rainfall data.
- Computation of statistical parameters of selected stations within the sub-basin.
- Carry out homogeneity test for the stations in the region.
- Delineation of homogeneous regions
- Selection of frequency distribution for the determination of the quantiles.
- Selection of parameter estimation method for the selected distribution.
- Derivation of regional and at site flood frequency curve.
- Developing regression model between mean annual flow and catchment characteristics.
- Computing quantiles on the ungauged catchment using the standardized growth curve and the index flood on the regression model.

## Result and Discussion

### Identification and Delineation of Homogeneous Regions

The identification of homogenous regions is usually the most difficult stage and requires greatest amount of subjective judgment. The aim is to form group of sites that approximately satisfy homogeneity condition that the sites frequency distributions are identical. We strongly prefer to base the formation of regions on site characteristics and to use the at site statistics in subsequent testing of homogeneity of a proposed set of regions (Hosking and Wallis, 1997).

### Catchment characteristics as a base for grouping stations

In this study preliminary identification of groups of stations into a certain category is achieved through looking at catchment characteristics. Then those stations having nearly same kind of physiographic and climatic characteristics are clustered and set forth to see their parent distribution on LMRD. And finally they are checked for the regional homogeneity.

For Abaya-Chamo river basin, some of the physiographic characteristics available and are believed to either directly or indirectly affect the catchment flood producing characteristics are land cover, soil type, geology, elevation and terrain slope. On the other hand rainfall pattern of the basin considered from the climatic characteristics side. The physiographic and climatic characteristics considered in this study are shown on the table below.

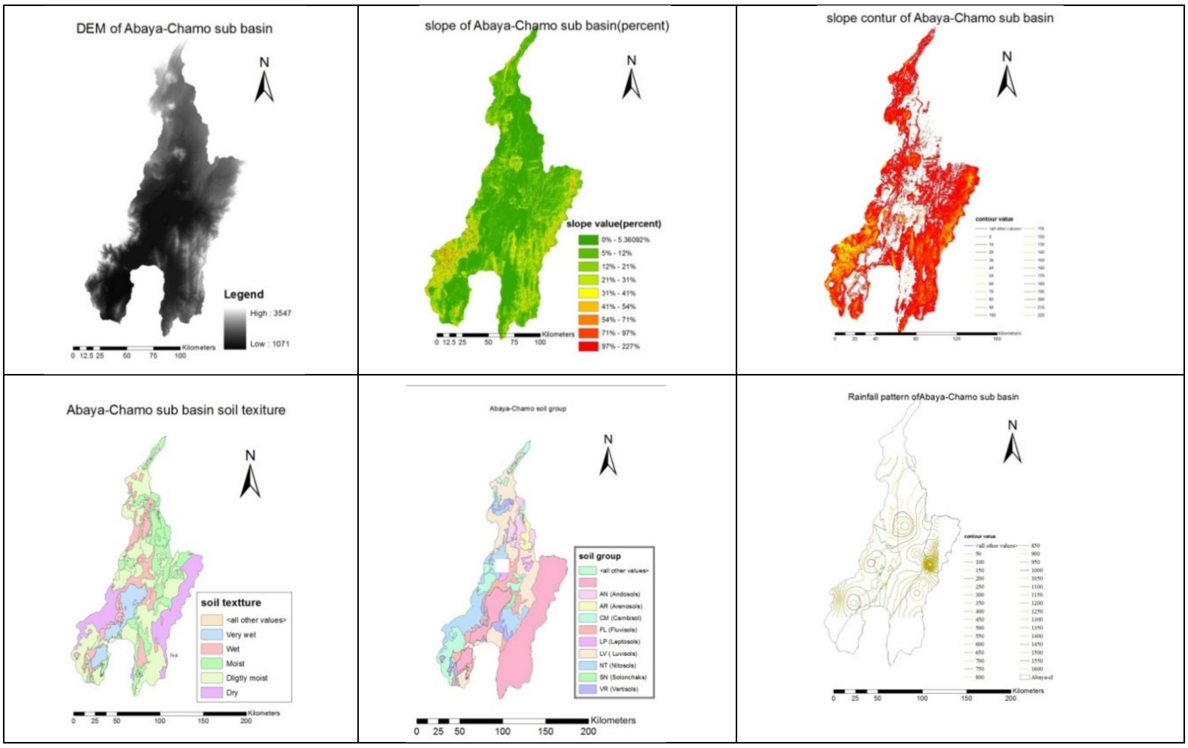

Figure 2: physiographic and climatic characteristics considered in this study for realization

### Using L moment ratio diagram to form a region

After processing a preliminary region based on catchment characteristics the next step will be identifying stations having the same parent distribution. Various approach have been proposed to select stations having the same parent distribution, where as in this paper the L moment ratio diagram is employed to do so. To use the LMRD some of the statistical parameters such as LCs and LCk are first computed and drawn on the LMRD and those stations lie close to a single line or distribution on the diagram are then supposed to come from the same parent distribution and are considered to be in the same region.

Table: 1 provides the summary of LCV, LCs and LCk for all stations on the sub-basin while figure: 3 shows the result of LMRD used to categorize stations in to regions.

Table 1: Summary of LCv, LCS and LCk for all stations of Abaya-Chamo sub basin

| | **Station Name** | **N** | **Mean Flow($m^3$/sec)** | **CV** | **CS** | **Ck** | **LCV** | **LCS** | **LCK** |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Sala | 16 | 20.1 | 1.15 | 0.08 | 0.19 | 0.28 | 0.123 | 0.19 |
| 2 | Hamessa at Humbo | 29 | 9.157 | 0.63 | 0.9 | 3.06 | 0.35 | 0.25 | 0.11 |
| 3 | Hamessa at Wajifo | 34 | 20.235 | 0.47 | -0.4 | 2.45 | 0.27 | -0.08 | 0.06 |
| 4 | Gidabo near Aposto | 37 | 34.957 | 0.56 | 1.03 | 4.39 | 0.31 | 0.2 | 0.13 |
| 5 | Bedessa at Dilla | 32 | 17.054 | 0.39 | 2.02 | 8.19 | 0.26 | 0.35 | 0.24 |
| 6 | Kolla | 34 | 17.165 | 0.47 | 0.81 | 4.23 | 0.28 | 0.21 | 0.19 |
| 7 | Werija | 21 | 111.948 | 0.5 | 0.29 | 2.55 | 0.29 | 0.09 | 0.07 |
| 8 | Gato | 39 | 3.65 | 1.23 | 1.94 | 2 | 0.72 | 0.02 | 0.02 |
| 9 | Bilate at Tena | 37 | 108.131 | 0.48 | 1.26 | 4.42 | 0.25 | 0.28 | 0.23 |
| 10 | Bilate at Alaba | 43 | 61.577 | 0.49 | 0.88 | 3.03 | 0.28 | 0.24 | 0.07 |
| 11 | Harie | 34 | 6.726 | 0.58 | 1.67 | 5.86 | 0.29 | 0.38 | 0.19 |
| 12 | Kulfo | 38 | 41.499 | 0.52 | 2.13 | 10.43 | 0.36 | 0.26 | 0.16 |
| 13 | upper Gelana | 33 | 34.317 | 0.64 | 0.96 | 3.41 | 0.35 | 0.24 | 0.36 |
| | | | | | | | Min.LCs | -0.25 | |
| | | | | | | | Maxim. LCs | 0.378 | |

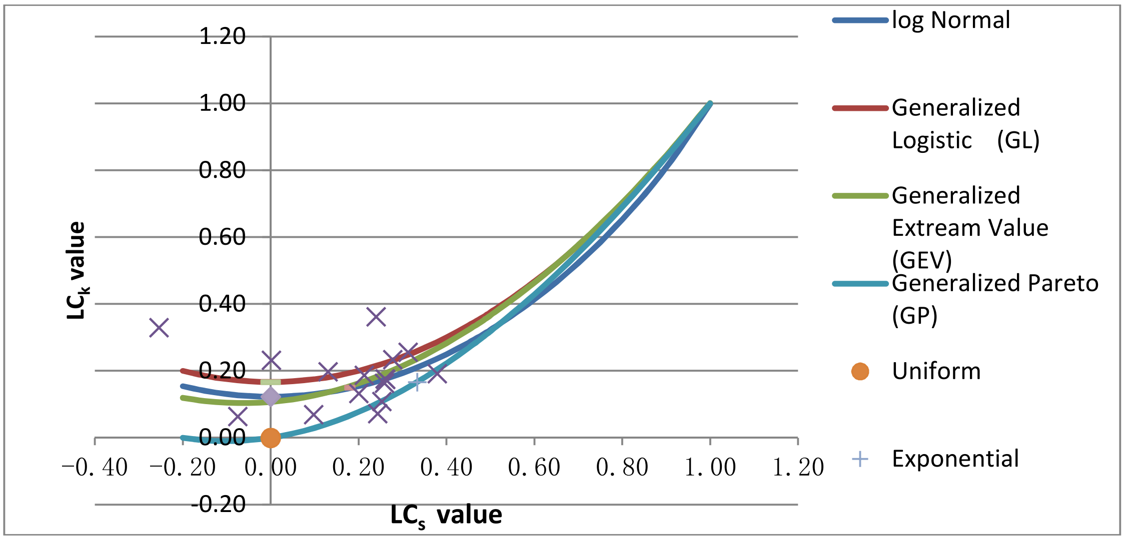

Figure 3: L moment ratio diagram for stations in Abaya-Chamo sub basin

The result of LMRD is almost similar to that of using the site characteristics. The combined result from the site characteristics and LMRD is tabulated as follows.

Table 2: The combined result of catchment characteristics and LMRD for selection of homogenous region

| **region** | **Station name** | **Possible distribution from LMRD** |
|---|---|---|
| Region one | Upper Gelana | Generalized logistic(GL)/ Generalized Extreme(GEX) |
| | Sala | Generalized logistic(GL)/ Generalized Extreme(GEX) |
| | Kolla | Generalized Extreme(GEX)/Log Norma(LN) |
| | Gidabo near Aposto | Generalized Extreme(GEX)/Gamma (2P) |
| Region two | Gato | Generalized logistic(GL)/ Log Norma(LN) |
| | Bilate at Alaba | Generalized Parento(GP)/ Log Norma(LN) |
| | Bilate at Tena | Generalized Extreme(GEX)/Log Norma(LN) |
| | Bedessa at Dilla | Log Norma(LN)/ Generalized Extreme(GEX) |
| Region three | Kulfo | Gamma (2P)/ Generalized Parento(GP) |
| | Hamessa at Wajifo | Generalized Extreme(GEX)/Gamma (2P) |
| | Hamessa Humbo | Generalized Parento(GP)/ Gamma (2P) |
| | Haier | Gamma (2P)/ Generalized Parento(GP) |
| | Werija | Generalized Parento(GP)/ Gamma (2P) |
| | Battena | Log Norma(LN)/ Gamma (2P) |

**N: B the one which is highlighted is the best selected for the region since stations in the same region lie close to the same distribution**.

## Regional homogeneity test

Once a set of physically plausible regions has been defined, it is desirable to assess whether the region are meaningful. This involves statistical testing whether a proposed region may accepted as being homogeneous, and when two or more homogeneous regions are sufficiently similar, they should be combined in to a single region. However, (Hosking, 1985) shows that small departure from perfect homogeneity does not appreciably reduce the beneficial aspects of RFFA, even with moderate heterogeneity (CC between 0.375-0.625). Some of the most commonly used statistical homogeneity tests are discordance measure test and CC – based homogeneity test. Both these homogeneity test are employed to check regional homogeneity of the proposed stations in the Abaya-Chamo sub basin

## Discordance measure test

This test measures sites that are grossly discordant from the group as a whole. The discordancy measure D estimates how far a given site is from the center of the group.

A discordant station will have $D_i$ value greater than or equals to 3.The result of the discordance measure test are shown in the table 3, 4 and 5 for region one, region two and region three respectively.

Table 3: The result of discordancy measure for region one

| St. Name | t | $t_3$ | $t_4$ | Discordancy measure | Remark |
|---|---|---|---|---|---|
| Upper Gelana | 0.354 | 0.239 | 0.362 | 0.35076 | Homogenous |
| Sala | 0.279 | 0.130 | 0.197 | -0.3148 | Homogenous |
| Kolla | 0.278 | 0.213 | 0.187 | 0.5300 | Homogenous |
| Gidabo near Aposto | 0.310 | 0.200 | 0.133 | -16.4345 | Homogenous |

Table 4: The result of discordancy measure for region two

| St. Name | t | $t_3$ | $t_4$ | Discordancy measure | Remark |
|---|---|---|---|---|---|
| Gato | 0.355 | 0.312 | 0.256 | 0.9459 | Homogenous |
| Bilate at Alaba | 0.278 | 0.243 | 0.072 | 0.8788 | Homogenous |
| Bilate at Tena | 0.254 | 0.277 | 0.234 | 0.9656 | Homogenous |
| Bedessa at Dilla | 0.291 | 0.260 | 0.177 | 0.3836 | Homogenous |

Table 5: The result of discordancy measure for region three

| St. Name | t | $t_3$ | $t_4$ | Discordancy measure | Remark |
|---|---|---|---|---|---|
| Kulfo | 0.360 | 0.258 | 0.164 | 0.2244 | Homogenous |
| Hamessa at Wajifo | 0.272 | -0.075 | 0.064 | 1.4263 | Homogenous |
| Hamessa at Humbo | 0.347 | 0.253 | 0.109 | 0.9782 | Homogenous |
| Haier | 0.295 | 0.379 | 0.192 | 1.5683 | Homogenous |
| Werija | 0.292 | 0.097 | 0.096 | 0.6199 | Homogenous |
| Battena | 0.412 | 0.262 | 0.174 | 1.1374 | Homogenous |

As shown on the result of the discordancy measure test all values are below 3, which implies that all the regions are homogeneous.

## CC –Based homogeneity test

In this test the site-to-site coefficient of variation of the coefficient of variation (CC) of both conventional and L-moments of the proposed region are used.

Table 6: CC values for both conventional moment and L- moment cases.

| | Station name | N | Mean flow | Cv in con. moment | Cv in L moment | conclusion |
|---|---|---|---|---|---|---|
| **Region one** | **Upper Gelana** | **33** | **38.9** | **0.64** | **0.35** | **Homogenous** |
| | **Sala** | **16** | **19.22** | **1.15** | **0.28** | |
| | **Kolla** | **34** | **18.07** | **0.47** | **0.31** | |
| | **Gidabo near Aposto** | **37** | **33.65** | **0.56** | **0.28** | |
| | | | **Mean** | **0.71** | **0.31** | |
| | | | **SDEV** | **0.19** | **0.04** | |
| | | | **CC** | **0.28** | **0.12** | |
| **Region Two** | **Gato** | **39** | **3.86** | **0.71** | **0.36** | **Homogenous** |
| | **Bilate at Alaba** | **43** | **70.46** | **0.49** | **0.28** | |
| | **Bilate at Tena** | **37** | **117.53** | **0.48** | **0.25** | |
| | **Bedessa at Dilla** | **31** | **18.34** | **0.31** | **0.21** | |
| | | | **Mean** | **0.50** | **0.26** | |
| | | | **SDEV** | **0.14** | **0.06** | |
| | | | **CC** | **0.28** | **0.22** | |
| **Region Three** | **Kulfo** | **38** | **42.16** | **0.58** | **0.36** | **Homogenous** |
| | **Hamessa at Wajifo** | **34** | **19.93** | **0.47** | **0.27** | |
| | **Hamessa at Humbo** | **29** | **10.43** | **0.63** | **0.35** | |
| | **Haier** | **34** | **7.72** | **0.58** | **0.29** | |
| | **Werija** | **21** | **111.72** | **0.50** | **0.29** | |
| | **Battena** | **23** | **8.45** | **0.92** | **0.41** | |
| | | | **Mean** | **0.61** | **0.33** | |
| | | | **SDEV** | **0.16** | **0.05** | |
| | | | **CC** | **0.26** | **0.16** | |

From the result it is concluded that all the regions are homogeneous since all CC values of the regions are below 0.3.

## Delineation of the homogeneous region

After organizing and assembling the data set, some of the important statistical parameters have been computed and interpolation of these statistical values ($LC_s$, $LC_k$) in collaboration with catchment characteristics are then used to come up to the following result of delineation of Abaya-Chamo sub basin on the Arc GIS interface.

Accordingly, the Abaya-Chamo sub basin is divided in to three regions; the first region includes the Upper Gelana, Sala, Kolla and Gidabo near Aposto stations. The second region has Gato, Bilate at Alaba, Bilate at Tena and Bedessa at Dilla stations. The third region has Kulfo, Hamessa at Wajifo, Hamessa at Humbo, Haier, Werija and Battena stations. Finally delineated individual regions of the Abaya –Chamo sub basin are shown in the figure 4.

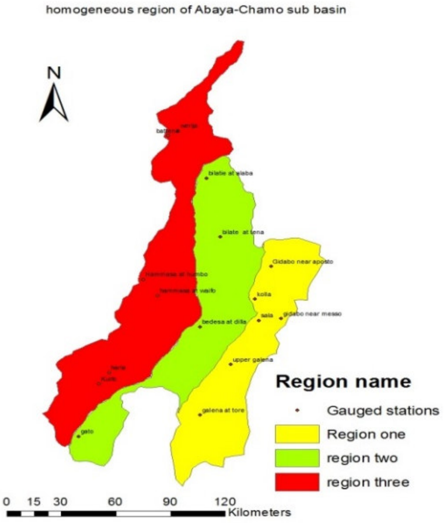

Figure 4: The final delineated individual regions of the Abaya –Chamo sub basin.

## Selection of regional parent distribution

Method of probability plot and L-moment ratio diagram were used for preliminary selection of distribution for the stations and regions of Abaya- Chamo sub basin.

## Method of probability plot

Observed data are plotted against the vales estimated from the fitted distribution by using software SPSS (statistical package for the social science). If the fitted distribution is exact parent distribution, the relationship should appear as a straight line through the origin with $45^0$slope.

**Method of L moment ratio diagram**

This popular and widely accepted method is used for preliminary selection of distribution for the stations and regions of Abaya- Chamo sub basin. Here regional average L-moment statistical value of stations ($LC_s$ & $LC_k$) are used and plotted on LMRD for initial parent distribution LMRD and corresponding results obtained from regional data are shown below.

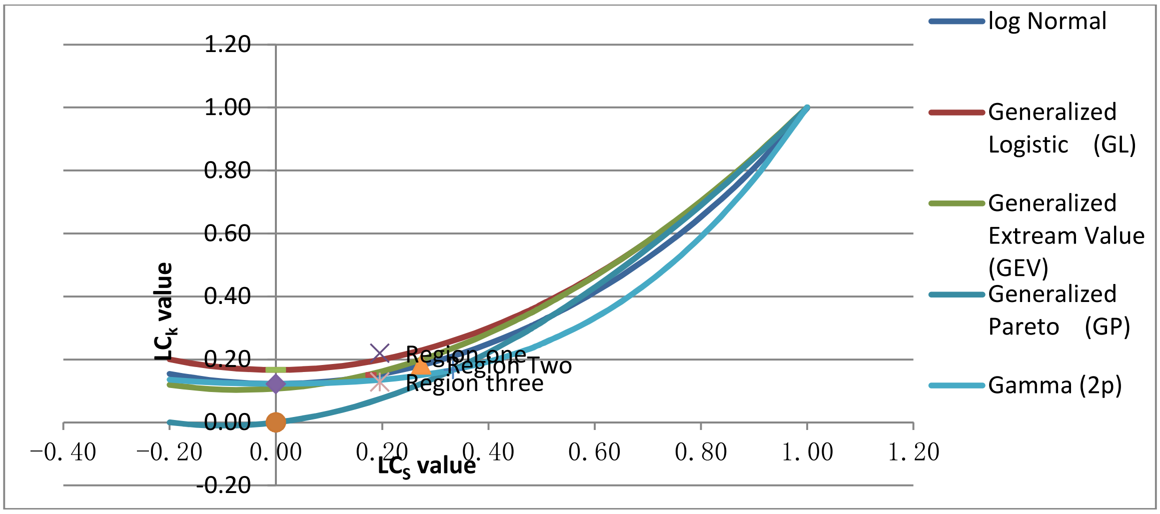

Figure5: Moment diagram for the average values of each region

Summary of preliminary selection results from the LMRD shown above is tabulated as follows. Screening for final best fit candidate distribution and corresponding parameter estimation for each region will be made by using the software called Easy Fit.

Table7: Possible regional candidate distribution obtained from LMRD diagram.

| | Average statistical values | | Possible parent distribution on LMRD |
|---|---|---|---|
| Regions | LCs [$t_3$] | LCk [$t_4$] | |
| Region One | 0.1956 | 0.2195 | GEV,LN(2P), GL |
| Region two | 0.2733 | 0.1847 | Gamma(2P), GEV, LN |
| Region three | 0.1956 | 0.1286 | Gamma(2P), GP, LN(2P) |

**Goodness of fit test**

In this study selection of best fit distribution from candidate distribution obtained from the above mentioned analysis is done by using software which is called Easy Fit. Within this software all goodness of fit test such as Chi-Square, Kolmogorov Smirinove and Anderson Darling tests are done and the best fit distribution from the candidate is displayed automatically.

Table8: Goodness of Fit summary for region one

Goodness of Fit - Summary

| # | Distribution | Kolmogorov Smirnov | | Anderson Darling | | Chi-Squared | |
|---|---|---|---|---|---|---|---|
| | | Statistic | Rank | Statistic | Rank | Statistic | Rank |
| 1 | Gen. Extreme Value | 0.09327 | 1 | 0.28798 | 1 | 1.3612 | 1 |

Table9: Goodness of Fit summary for region two

**Goodness of Fit - Summary**

| # | Distribution | Kolmogorov Smirnov | | Anderson Darling | | Chi-Squared | |
|---|---|---|---|---|---|---|---|
| | | Statistic | Rank | Statistic | Rank | Statistic | Rank |
| 1 | Lognormal | 0.1543 | 1 | 0.63832 | 1 | 3.3329 | 1 |

Table10: Goodness of Fit summary for region three

**Goodness of Fit - Summary**

| # | Distribution | Kolmogorov Smirnov | | Anderson Darling | | Chi-Squared | |
|---|---|---|---|---|---|---|---|
| | | Statistic | Rank | Statistic | Rank | Statistic | Rank |
| 1 | Gamma | 0.18647 | 1 | 2.3871 | 1 | 1.056 | 1 |

Best regional distribution which is selected from the candidate distribution by using Easy Fit software is tabulated below.

Table 11: Best regional distribution selected for each regions using goodness of fit test (Easy Fit software)

| Region | Candidate distribution | Finally selected best fit distribution |
|---|---|---|
| One | GEV,LN(2P), GL | GEV |
| Two | Gamma(2P), GEV, LN | LN |
| Three | Gamma(2P), GP, LN(2P) | Gamma(2P) |

***Parameter and quantile estimation***

The methods used for parameter estimation in this study are method of moment (MOM) and probability weighted moments (PWM) which are also the most commonly used methods by many researchers.

The summary of parameter estimation of each region using both method of moment (MOM) and probability weighted moment are shown in Table: 12.

Table 12: Parameter of each region summary

| | Selected distribution | Parameter of MOM | | Parameter of PWM | |
|---|---|---|---|---|---|
| | | parameter | value | Parameter | Value |
| Region One | Generalized Extreme Value (GEV) | $\hat{\kappa}$ | 0.081 | $\hat{\kappa}$ | -0.98 |
| | | $\hat{\alpha}$ | 0.004 | $\hat{\alpha}$ | 0.004 |
| | | $\hat{\mu}$ | 1.044 | $\hat{\mu}$ | 0.716 |
| Region Two | Two-Parameter Lognormal (LN(2P)) | $\delta_y$ | 0.218 | $\delta_y$ | 0.531 |
| | | $\widehat{\mu_y}$ | 0.231 | $\widehat{\mu_y}$ | 1.654 |
| Region Three | Two-Parameter Gamma (G(2)) | $\hat{\alpha}$ | 0.469 | $\hat{\alpha}$ | 0.406 |
| | | $\hat{\beta}$ | 2.297 | $\hat{\beta}$ | 2.659 |

***Quantile estimation***

***Quantile estimation for region one***

**Quantile estimate by MOM parameters:**

Magnitude of T-year flood estimated by using an Equation:

$$Q_T = \hat{\mu} + \frac{\hat{\alpha}}{\hat{\kappa}}\left[1 - \left\{-log\left(1 - \frac{1}{T}\right)\right\}^{\hat{\kappa}}\right] \qquad (1)$$

Where, $\hat{\mu}$, $\hat{\alpha}$ and $\hat{\boldsymbol{\mu}}$ are parameter values of obtained by MOM

**Quantile estimate by PWM parameters:**

Here Equation: 1 also applied for estimation of T-year flood magnitude but the values of $\hat{\mu}$, $\hat{\alpha}$ and $\hat{\mu}$ are parameter values obtained by PWM

**Quantile estimation for region two**

**Quantile estimate by MOM parameters:**

Magnitude of T-year flood estimated by using an Equation:

$$\log Q_T = \widehat{\mu_y} + \widehat{\delta_y} u \quad \text{or} \quad Q_T = \exp^{\widehat{\mu_y} + \widehat{\delta_y} u} \qquad (2)$$

Where, $\widehat{\mu_y}$ and $\widehat{\delta_y}$ are parameter values obtained by MOM

**Quantile estimate by PWM parameters:**
Here Equation: 2 also applied for estimation of T-year flood magnitude but the values of $\widehat{\mu_y}$ and $\widehat{\delta_y}$ are parameter values obtained by PWM.

**Quantile estimation for region three**

**Quantile estimate by MOM parameters:**
Magnitude of T-year flood estimated by using an Equation:

$$Q_T = \hat{\alpha}\hat{\beta} + K_T\sqrt{\hat{\alpha}^2\hat{\beta}} \qquad (3)$$

Where, $K_T = \frac{2}{C_s}\left[\left\{\frac{C_s}{6}\left(u - \frac{C_s}{6}\right) + +\right\}^3 - 1\right]$ (4)

**Quantile estimate by MOM parameters:**
Here Equation: 5.21 also applied for estimation of T-year flood magnitude but the values of $\hat{\alpha}$ and $\hat{\beta}$are parameter values obtained by PWM

Table 13: Summary of quantile estimated for each region

| **Region** | **Estimated Quantile** | | | |
|---|---|---|---|---|
| | **Method of moment (MOM)** | | **Probability weighted moment (PWM)** | |
| | **T** | $Q_T$ | **T** | $Q_T$ |
| Region one | 2 | 1.0487 | 2 | 0.073 |
| | 5 | 1.0529 | 5 | 0.756 |
| | 10 | 1.0555 | 10 | 0.809 |
| | 50 | 1.061 | 50 | 1.208 |
| | 100 | 1.062 | 100 | 1.70 |
| Region two | 2 | 1.259 | 2 | 5.22 |
| | 5 | 1.513 | 5 | 8.17 |
| | 10 | 1.666 | 10 | 10.33 |
| | 50 | 1.971 | 50 | 15.56 |
| | 100 | 2.0919 | 100 | 18.01 |
| Region three | 2 | 0.975 | 2 | 0.983 |
| | 5 | 1.623 | 5 | 1.586 |
| | 10 | 2.024 | 10 | 1.96 |
| | 50 | 2.839 | 50 | 2.718 |
| | 100 | 3.167 | 100 | 3.024 |

**Robustness Assessment**

In order to have a reliable estimate of the regional quantile values, a robustness assessment has to done among the various available parameter estimation procedures. The combination of all the available parameter estimation provides us a lot of procedures, where as it will be difficult and cumbersome to fit all these procedures to a certain regional data set. Therefore, the robustness of only those distributions selected from the descriptive analysis so far from (from L-moment and Goodness of fit) and those commonly used parameter estimate are used in this paper.

The proposed procedures selected for robustness assessment are: GEV/MOM, GEV/PWM, LN/ MOM, LN/PWM, Gamma (2P)/MOM, Gamma (2P)/PWM. In this study standard error of estimate was applied for selecting the best parameter estimation method. The standard error of the estimate is a measure of the accuracy of predictions.

Standard error of estimate is calculated by using

$$SEE = \sqrt{\frac{\sum_{I=1}^{n}(E-O)^2}{n-1}} \qquad (5)$$

Where, SEE=standard error of estimate, E= estimated value, O= observed value

The performance of the parameter estimation method under test is then compared through the magnitude of SEE. The parameter estimation procedure giving the smallest values for SEE was selected to be the best parameter estimation procedure.

Table 14: Standard error of estimate value used to select the most robust parameter estimation procedure.

| | Region One | | Region Two | | Region Three | |
|---|---|---|---|---|---|---|
| | GEV/MOM | GEV/PWM | LN/ MOM | LN/PWM | Gamma(2P)/MOM | Gamma(2P)/PWM |
| | 1.396 | 1.248 | 0.305 | 8.130 | 0.218 | 0.167 |

Table 15: Best procedure selected from robustness assessment

| Regions | Candidate distribution | Best parameter estimation method |
|---|---|---|
| Region one | Generalized extreme value(GEV) | Probability weighted moment (PWM) |
| Region Two | Log normal (LN) | Method of moment (MOM) |
| Region Three | Gamma (2P) | Probability weighted moment (PWM) |

***Regional frequency curve***

The most robust parameter estimation procedures discovered in the previous section were used to develop regional frequency curve as follows.

**Region one:**

By using probability weighted moment (PWM) as a parameter estimation method, data from all stations in the region are used to fit a generalized extreme value distribution and the following frequency curve is developed.

Table 16: Parameter and quantile estimate of region one

| Parameters of Generalized Extreme distribution | | Regional quantile value | |
|---|---|---|---|
| $\hat{\kappa} =$ | -0.9856 | Return period[T] | Quantile values[$Q_T$] |
| $\hat{\alpha} =$ | 0.0046 | 2 | 0.7266 |
| $\hat{\mu} =$ | 0.7160 | 5 | 0.756 |
| | | 10 | 0.809 |
| Corresponding quantile estimator equation is:<br>$X_T = \hat{\mu} + \frac{\hat{\alpha}}{\hat{\kappa}}\left[1 - \left\{-log\left(1 - \frac{1}{T}\right)\right\}^{\hat{\kappa}}\right]$<br>$X_T = 0.716 + \frac{0.0046}{(-0.9856)}\left[1 - \left\{-log\left(1 - \frac{1}{T}\right)\right\}^{(-0.9856)}\right]$ | | 50 | 1.208 |
| | | 100 | 1.7 |
| | | 500 | 5.561 |
| | | 1000 | 10.32 |
| | | 5000 | 47.671 |
| | | 10000 | 93.701 |

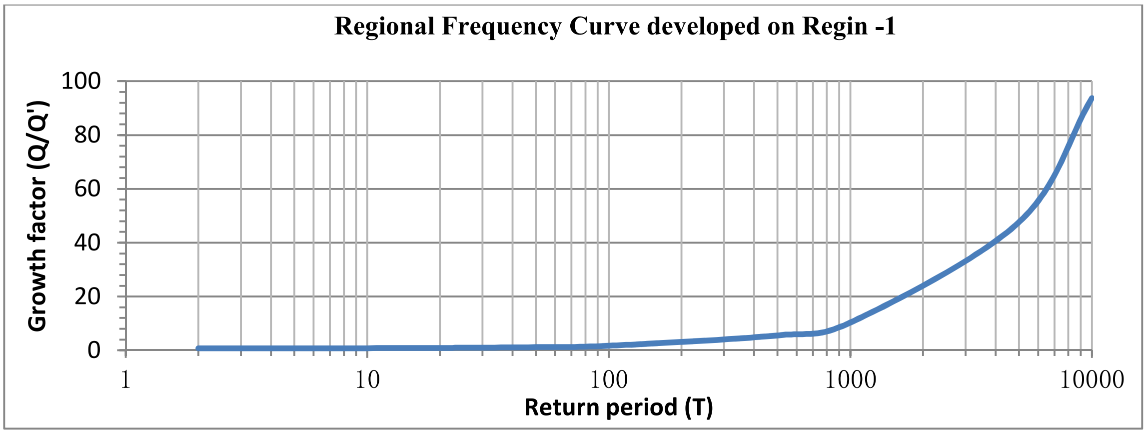

Figure 6: Standardized frequency curve developed for region one

**Region Two**

Log Normal distribution with method of moment as a parameter estimation was used to prepare the regional frequency curve.

Table 17: Parameter and quantile estimate of region two

| Parameters of Log Normal distribution | | Regional quantile value | |
|---|---|---|---|
| $\widehat{\sigma_y}$ | 0.218 | Return period[T] | Quantile values[$Q_T$] |
| $\widehat{\mu_y}$ | 0.231 | 2 | 1.259 |
| Corresponding quantile estimator equation is:<br>$X_T = e^{\widehat{\mu_y} + u\widehat{\sigma_y}}$<br>$u = w - \frac{2.516+0.803w+0.01w^2}{1+1.433w+0.189w^2+0.001w^3}$<br>$w = \left[ln\left(\frac{1}{P^2}\right)\right]^{\frac{1}{2}}$ | | 5 | 1.513 |
| | | 10 | 1.666 |
| | | 30 | 1.879 |
| | | 50 | 1.971 |
| | | 100 | 2.092 |
| | | 500 | 2.362 |
| | | 1000 | 2.473 |
| | | 5000 | 50.04 |
| | | 10000 | 66.74 |

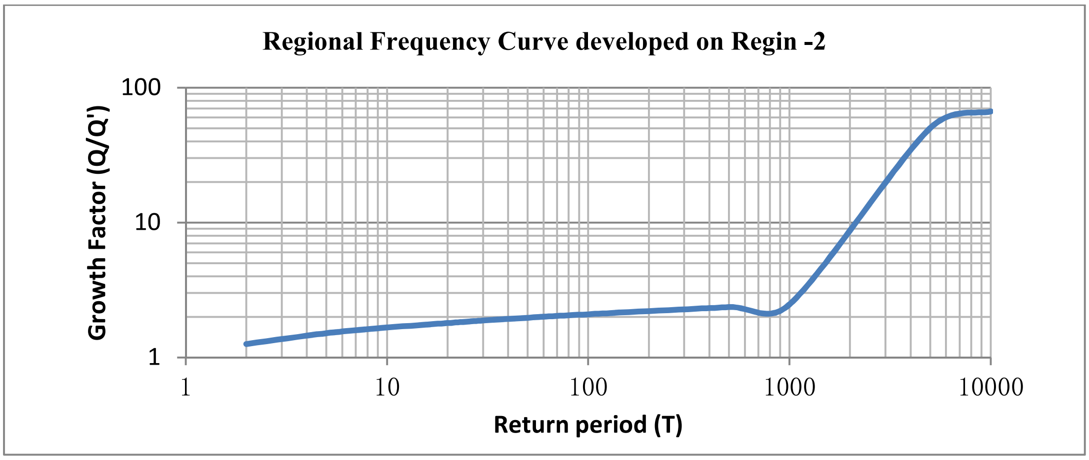

Figure7: Standardized frequency curve developed for region two

**Region Three**

A two parameter Gamma distribution with probability weighted moment as a parameter estimation was used to prepare the regional frequency curve.

Table 18: Parameter and quantile estimate of region three

| Parameters of Gamma (2P) distribution | | Regional quantile value | |
|---|---|---|---|
| $\hat{\alpha} =$ | 0.4056 | Return period[T] | Quantile values[$Q_T$] |
| $\hat{\beta} =$ | 2.6585 | 2 | 0.983 |
| $X_T = \hat{\alpha}\hat{\beta} + K_T\sqrt{\hat{\alpha}^2\hat{\beta}}$<br>$K_T = \frac{2}{C_s}\left[\left\{\frac{C_s}{6}\left(u - \frac{C_s}{6}\right) + 1\right\}^3 - 1\right]$, for $C_s > 3$<br>$K_T = u + \frac{C_s}{2}\left(\frac{u^2-1}{3}\right) + \frac{{C_s}^2}{2^4}\left(\frac{u^3-7u}{9}\right)\left(\frac{{C_s}^3}{2^5}\right)\left(\frac{6u^4+14u^2-32}{405}\right) + \frac{{C_s}^4}{2^7}\left(\frac{9u^5+256u^3-433u}{4860}\right) + \frac{{C_s}^5}{2^8}\left(\frac{12u^6 143u^4-923u^2+1472}{25515}\right) - \frac{{C_s}^6}{2^{10}}\left(\frac{3753u^7 4353u^5-28957u^3-289717u}{9185400}\right)$ ,for $\lvert c_s \rvert \le 2$ | | 5 | 1.586 |
| | | 10 | 1.96 |
| | | 30 | 1.491 |
| | | 50 | 2.718 |
| | | 100 | 3.024 |
| | | 500 | 3.701 |
| | | 1000 | 3.979 |
| | | 5000 | 60.223 |
| | | 10000 | 71.697 |

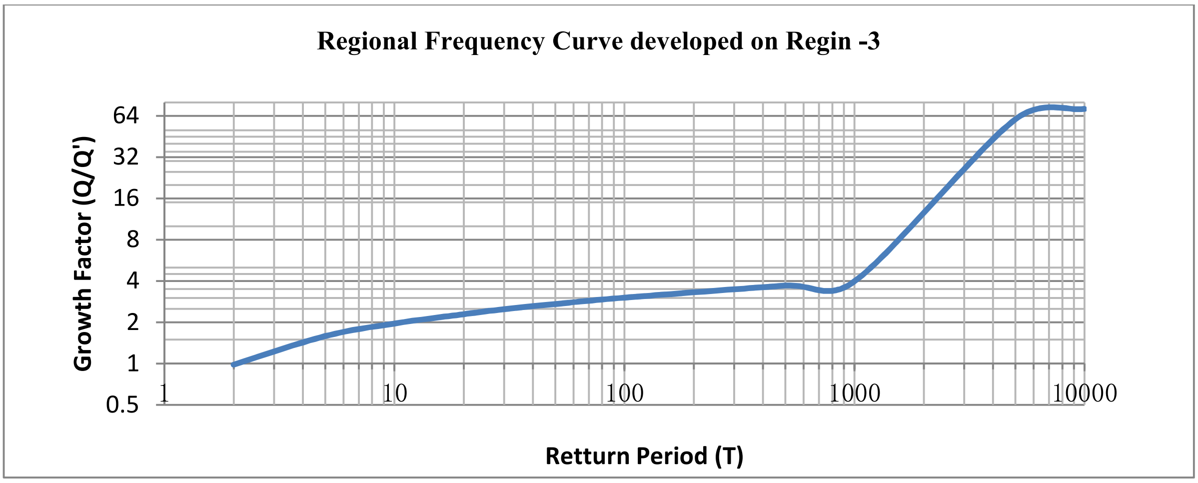

Figure 8: Standardized frequency curve developed for region three.

**Design Flood for the Ungauged Catchment**

If the catchment has no records OF flow, an index flood ($Q_{index}$) must be estimated from a relation between $\overline{Q}$ and measurable catchment characteristics, which are calibrated from gauged catchments in the region. Finally, growth factor will used to get final estimate of quantile on the ungauged places. In this study characteristics like catchment area, slope, soil type, soil texture and mean annual rainfall are used to develop ungauged catchment model (index flood equation). The number of catchment characteristics used on the model depends on the number of station in the region.

**Index-flood equation**

The number of stations within the homogeneous region restricts the number of variables (parameters) that can be used in the index flood equation. It is better to use one single explanatory variable by combining several of those parameters together. Some of the measurable parameters used in this study are catchment area (A), catchment elevation (Z) and mean annual rainfall (P). The possible combinations of parameters to form a single explanatory variable are: Variable-1 = AP, Variable-2 =AZ, Variable-3 =PZ, Variable-4 = AP/Z, Variable-5 =AZ/P and Variable-6 =PZ/A.

The index flood estimation equation (model) is developed by classifying of the whole sub-basin in two broad categories, upper Abaya-Chamo sub basin (Region three) and lower Abaya-Chamo sub-basin (region one plus region two), the necessity of merging the two region is to make the model statistically strong because more parameters will be used if there are more stations in the region. Upper and lower parts of the sub-basin are shown in the figure below.

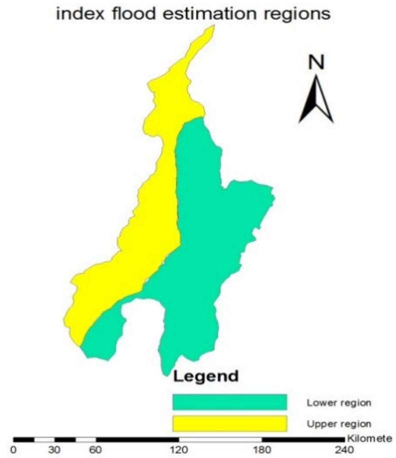

Figure 9: The upper and lower Abaya-Chamo sub basin to which the index flood equation is developed

A scatter plot of each possible variable with the mean annual flow for each station of the region was used to develop the index flood estimation equation. Here, a polynomial relation was the best determine the relation. The one with the coefficient of determination, $R^2$, high value is selected as the best index flood equation (model). In this study a polynomial relation of a variable, (AP/Z) with the mean annual flow shows higher value of coefficient of determination, $R^2$ =0.93for upper region and $R^2$= 0.99for the lower region and an equation which describe this relation were selected as the best index flood estimation equation (model).

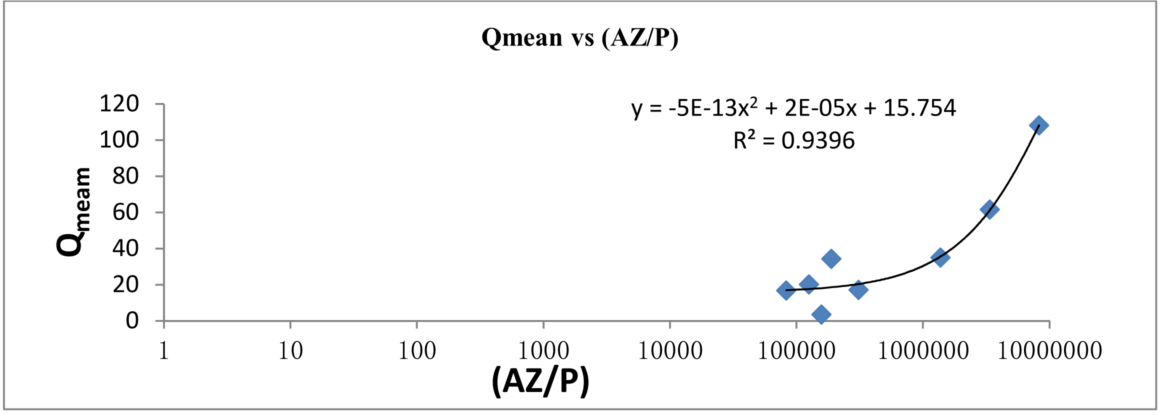

Figure 10: Best selected index flood estimation equations for the upper region:

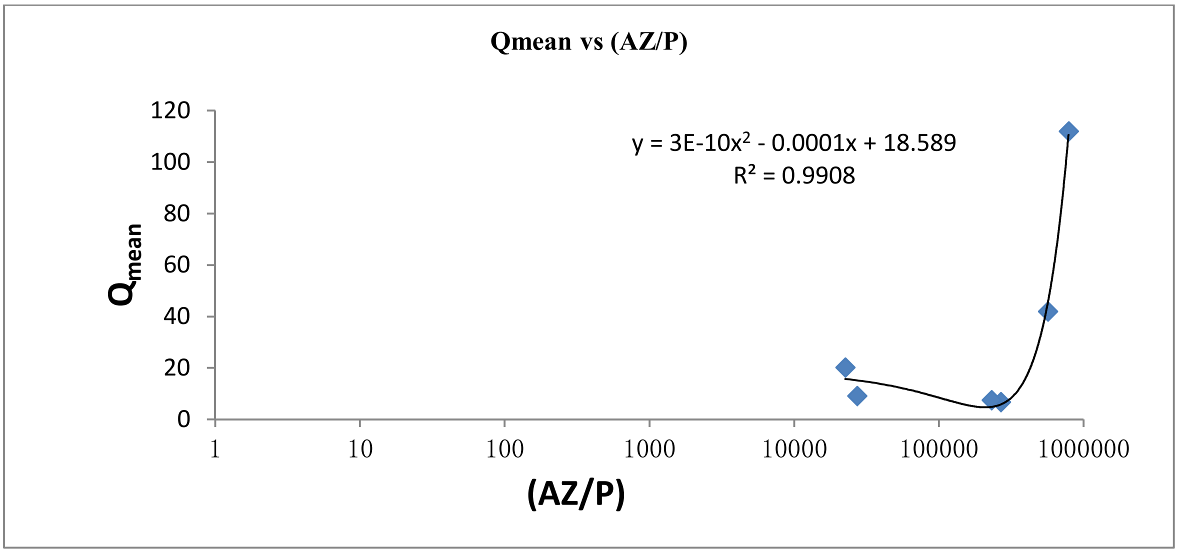

Figure11: Best selected index flood estimation equations for the upper region

The two most important index flood estimation equation (model) of the Abaya-Chamo sub basin are shown in the table below.

Table 19: Index flood estimation equation developed for upper and lower regions of Abaya-Chamo sub basin

| Name of the region | Index flood estimation equation | $R^2$ |
|---|---|---|
| Upper Abaya-Chamo sub basin | $Q_{index} = -5 * 10^{-13} \left(\frac{AZ}{P}\right)^2 + 2 * 10^{-5} \left(\frac{AZ}{P}\right) + 15.75$ | 0.9329 |
| Lower Abaya-Chamo sub basin | $Q_{index} = 3 * 10^{-10} \left(\frac{AZ}{P}\right)^2 - 0.0001 \left(\frac{AZ}{P}\right) + 18.59$ | 0.9908 |

## Conclusion and Recommendation

The regionalization of Abaya Chamo sub basin in to the same flood producing characteristics were initially performed based on site characteristics (elevation, soil type, soil texture, slope, land use land cover and mean annual rainfall), site statics are used for testing of homogeneity of the proposed region. The entire river basin was first classified on the basis of their catchment characteristics and then LMRD is exploited to check weather all stations in the same region are found to lie on the same type of distribution or not. Further, a discordance measure and CC test are conducted to check their homogeneity.

Regional average values of $LC_s$ and $LC_k$ were used to select the best fit statistical distribution of each region and goodness of fit test by using the software Easy Fit is used to approve the best fit distribution. Robustness assessment is conducted to select a superlative method of parameter estimation for the selected distribution. For region one Generalized Extreme Value (GEV) with PWM was selected, for region two Log Normal distributions with MOM was selected and finally Gamma 2Pwith PWM was selected for region three. And these distributions with method of parameter estimations are finally used to develop a regional growth curve of each homogeneous region.

For the sake of index flood estimation equation development, the whole sub basin is classified in two main categories, the upper Abaya-Chamo and lower Abaya-Chamo. Three parameters namely catchment area, catchment elevation and mean annual rainfall are used to describe the index flood. Moreover, an index flood estimation equation (model) with an interesting determination of coefficient($R^2$) of about 0.93 were developed for the upper Abaya-Chamo and 0.99for that of the Abaya-Chamo sub basin and they are considered as a best model to describe the regional index flood value of the region.

## Recommendation

- The regional growth curve and the index flood value estimation equation (model) obtained so far from the study result can be applicable for design of hydraulic structures on the ungauged catchment of Abaya-Chamo sub basin.
- In order to get reliable estimate of regional quantile more hydrometric stations should be installed in the basin to infer something for ungagged sites.
- On regionalization, it is more to lie up on catchment characteristics than be exclusively dependent on statistical values of stations.
- More parameters which are believed to affect catchment characteristics have to be assessed and incorporate to modify the regression equation.
- On this study only the source of flood frequency analysis uncertainty are described but consideration of those uncertainties must be accounted in risk management.

## Referene

1. Ababu T. and Bernd W. (2004). Water Quality Monitoring within The Abaya-Chamo Drainage Basin.
2. Admasu, (1989). Regional flood frequency analysis, PHD thesis report submitted to Stockholm University.
3. Guidelines for FFA USWRC, (1976). Guidelines for determining flood flow frequency.
4. Hosking, J.R.M. and Wallis, J.R. (1997) Regional frequency analysis, Cambridge University press, UK
5. Mengistu D, (2008). Regional flood frequency analysis for upper awash sub- basin (Upstream of koka), MSc thesis submitted to Addis Abeba University.
6. Seleshi Bekele Awelachew and Prof. Dr.-ing. Habil H. B, (2000). Development and application of 2-parameters monthly water balance model in limited data situation, the case of Abaya-Chamo sub basin, *journal of EAEA*.