ADDIS ABABA UNIVERSITY

INSTITUTE OF TECHNOLOGY SCHOOL OF GRADUATE STUDIES

DEPARTMENT OF CIVIL AND ENVIRONMNTAL ENGINEERING

HYDROLOGICAL TREND, VARIABILITY AND TIME SERIES MODELING OF WEYIB CATCHMENT

BY

ABEL LEGESSE

ADVISOR: DR.Ing.DEREJE HAILU

August, 2017

HYDROLOGICAL TREND, VARIABILITY AND TIME SERIES MODELING OF WEYIB CATCHMENT

BY

ABEL LEGESSE

A thesis submitted to the School of Graduate Studies of Addis Ababa University Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering (Major in Hydraulic Engineering).

Addis Ababa University Institute of Technology

August, 2017

Declaration and Copyright

I, Abel Legesse Habtegebriel, do here by declare to the Senate of Addis Ababa University that thesis entitled **Hydrological trend, variability and time series modeling of weyib catchment** is entirely original work and all other materials are duly acknowledged. This work has not been submitted for any academic degree award at any University.

Abel Legesse Habtegebriel

Obelo.fs@gmail.com

Signature: ----------------------

Addis Ababa Ethiopia, July 2017

ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

SCHOOL OF GRADUATE STUDY

HYDROLOGICAL TREND, VARIABILITY, AND TIME SERIES MODELING OF WEYIB CATCHMENT

BY ABEL LEGESSE

APPROVED BY BOARD EXAMINERS

Dereje Hailu, PhD	…………………	……..…………..
(Advisor)	Signature	Date
Agizew Nigussie, PhD	…………………	……..…………..
(Internal Examiner)	Signature	Date
Belete Birhanu, PhD	…………………	……..…………..
(External Examiner)	Signature	Date
____________________	…………………	……..…………..
(Chairman)	Signature	Date

Dedicated to my parents

Legesse Habtegebriel and Yeshi Debela

ABSTRACT

The main goal of this study is to assess trend, spatio- temporal variability of rainfall and runoff, and time series modeling of major tributaries of Weyib River catchment, which is located in Genale Dawa basin between $6^{0}22'$ to $7^{0}06'$N and $39^{0}37'$ to $41^{0}15'$E. Stream flow records from six flow-gauging stations in four major river catchment of the weyib between (1981- 2008) and six precipitation stations all over the weyib catchment between (1984-2014) were studied at monthly, seasonal, and annual scales. In the study, the data were firstly subjected to quality checks through the cumulative deviations test and the standard normal homogeneity test. To detect correlation Lag-1 autocorrelation coefficient was used. ArcGIS 10 environment is used to investigate the spatial variability. The Coefficient of Variation (CV) was used for variability analysis. Mann-Kendall trend analysis and Sen's slope estimator in MAKESENSE version 1.0 Excel template are used in determining the changes in the rainfall and runoff. ARIMA model in Minitab 17 Statistical software was used for time series modeling. By using statistical measure (R^2, AIC and BIC) it was determined that, ARIMA time series model is better appropriate to stream flow forecasting. The interesting observation is that all sub catchment of the Weyib catchment for both rainfall and runoff exhibit different behavior as far as trend and variability analysis is concerned. Trend analyses with Mann-Kendall test show that, there is increase in significant trend in annual mean runoff at Togona and Tebel station while the other four stations (Shaya, Agarfa, Alemkerem and Denbel) shows no significant trend. From precipitation stations, only Ginir station shows significant increasing trend in annual basis. The direction of rainfall and runoff trend was, in general, upward and statistically significant (upward) trends (95% to 99.9%) significant levels were observed. Global climate change influence on the changes in rainfall and runoff characteristics was not conclusive. The key conclusion is that, the results of this study are expected to assist water resources managers and policy makers in making better planning decisions in the Weyib River catchment.

Keywords: Weyib River catchment ; Minitab 17 ;Trend analysis ;Precipitation ; Mann-Kendall Test ; Sen's Slope Estimator; probability distribution

ACKNOWLEDGEMENT

First, I would like to thank ‘Almighty God ‘for granting me his limitless care, love, and blessings all along the way. I would like to express my wholehearted gratitude to my advisor Dr.Ing. Dereje Hailu, for his unreserved, constructive and timely comments at all stages of my work .My genuine acknowledgment also goes to Ethiopian Road Authority (ERA) for funding this MSc scholarship made through the postgraduate program of AAiT. I would like to extend my acknowledgement to Ministry of Water, Irrigation and Electricity, Ministry of Agriculture and Natural resource, National Meteorological Service Agency and Oromia water works design and supervision enterprise (OWWDSE) for the provision of all relevant data, documents and information essential for the study. My friends are all acknowledged. Finally, I would like to thank my family, who are always behind my success.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

ACF: Autocorrelation Function

AIC: Akaike's Information Criterion

ANOVA: Analysis of Variance

AR: Autoregressive

ARIMA: Autoregressive Integrated Moving Average

BIC: Bayesian Information Criterion

CDF: Cumulative Distribution Function

FAO: Food and Agriculture Organization

ITCZ: Inter-tropical Convergence Zone

MA: Moving Average

Mk: Mann- Kendall

PACF: Partial Autocorrelation Function

SAR: Seasonal Autoregressive

ARIMA: Autoregressive Integrated Moving Average

Sqkm: Square kilometer

SMA: Seasonal Moving average

CHAPTER ONE

INTRODUCTION

1.1. Background

Trend can be defined as a general direction in which something is developing or changing. Therefore, Hydrological trend is a general direction in which a hydrological component such as stream flow and precipitation in the hydrological time series is developing or changing (Haji, 2015). The knowledge of trend analysis is important for characterizing the change of behavior of input parameter for a sub catchment with respect to time. Analysis of trend in time series data is also required to determine what will be the condition of the values with respect to a certain specified range of time.

Most water resource systems have been designed and operated based on the assumption of stationary hydrology, if this assumption is incorrect then existing procedures for designing levees, dams, reservoirs, etc. will have to be revised, without revision, there is a danger that systems are over or under designed and either does not serve their purpose adequately(Robson et al., 2000). Therefore, detection of trends in long time series of hydrological data has paramount scientific and practical significance.

Watson et al. (2003) stated that the study of past trends in stream flow and rainfall is very essential for the better utilization of water resources and appropriate decision-making. Studies of change are also important because of our need to understand the impact that man is having on the natural world. Urbanization, deforestation, emissions of greenhouse gases, changes in agricultural practice and dam construction are just a few examples of anthropogenic activities that may be altering important aspects of the hydrological cycle.

The principal water related problems always related to having too much water (floods) or too little water (low flows or droughts) this means that studying changes in characteristics of hydrological extremes is of major importance (Robson et al., 2000).

Mathew (2004) stated that precipitation is an important weather element whose future change will have a large impact on the society, a decrease in mean precipitation would lead to increased risk of drought, while an increase in mean precipitation would lead to increased risk of flooding. This implies that analysis of trend in the recorded data of precipitation is necessary to reduce the catastrophic effect of drought and flood.

The timing, variability, and quantity of seasonal and annual rainfall are important factors in the relationship between climate and cultivation. In Ethiopia, several studies have been done on hydrological trends. Wing and Cheung (2008) studied the trends and spatial distribution of annual and seasonal rainfall in Ethiopia. He stated that the rainfall trends are one of the more important factors in explaining various socio-economic problems such as food insecurity for the countries, which their economy is heavily dependent on low-productivity rain-fed agriculture, therefore investigation of the temporal dynamics of rainfall and its spatial distribution within Ethiopia help policymakers and developers to make more decisions that are informed. In addition, (Bekele, 2016) stated the importance of the analysis of rainfall trend and variability for agricultural water management.

Therefore, Hydrological studies that determine the hydrological characteristics are obviously the fundamental basis to quantify the available water resources effectively. Adequate and long-range hydrologic data are essential in the planning of development schemes, the design of hydraulic structures and the optimum development and management of water resources. These data should be collected over a sufficient time-span to define the time variability of the data and must be collected on a nationwide basis to define the areal variability.

The study of Hydrological trend, Variability and modeling of time series provide a basis for future scenario analysis of water resource management of Weyib River catchment. Berryman et al. (1988) proved that even though there are various types of statistical trend assessments, most often Mann-Kendall rank correlation is the toughest method because this method is suitable for data that are not normally distributed, tolerates missing values, and is relatively genuine by extreme values or skewed data.

The trend slope and its statistical significance express the output of the Mann-Kendall analysis. Taking into account this, the trend analysis of areal rainfall, and river discharge at six-sub catchment of Weyib catchment was analyzed using the Mann-Kendall statistical test.

On the other hand, Autoregressive Integrated Moving Average (ARIMA) time series modeling were used for the modeling of the stream flow time series. Furthermore statistical software called Minitab17 was used for trend analysis and modeling of the time series of precipitation and stream flow. This study also considered comprehensive analysis of spatio-temporal variability of Hydrological Parameters like stream flow and Rainfall in monthly, seasonal and annual basis. This will help for better understanding of the change in discharge caused by implementing past practices and use this information in developing strategies for better utilization of water resources in the future and to make appropriate decision(Watson et al., 2003).

1.2. Problem Statement

The changes in the trends of extreme rainfall events have predominant effect on hydrological extreme events such as floods and droughts. The hydrological extreme events (flood and drought) have significant impact on human, ecology and agricultural sector particularly on crop production. This droughts and flood event have been recurring phenomena in Weyib catchment. The drought prone south-east part of the catchment, including Raitu, Dawe Qachen and Ginir sub catchment, which contributes areas of about 1871sqkm, 1751 sqkm and 1187 sqkm to Weyib catchment respectively, have suffered greatly from recurrent drought often followed by devastating famine.

According to (ACT Alliance, 2017) in the year 2016/17 Rayitu and Dawe Kachen sub catchment was suffered from serious drought, which led to loss of cattle lives and paralyze socio economic activities of the area. The communities along Weyib river course are also highly and seasonally threatened by flooding disasters since long time, particularly in downstream plain.

Due to absence of long-term concurrent rainfall and stream flow observations in weyib river catchment, the use of available data is warranted to conduct as much analysis as possible. Similarly, complete data sets are required on many variables such as rainfall, stream flow, evaporation, temperature etc. Unfortunately, hydro meteorological station network available in

Weyib catchment is too scant and often have a breaks. The scarcity of this vital information poses a great difficulty in proper water resources evaluation and development, such as irrigation, water supply, hydropower and flood control etc.

Again, large amount of investment committed in the catchment recently, like Weyib irrigation project must have to consider hydrological variability, the future water availability, and the driving force of the change in water resources in detail. Otherwise, the systems would be either under designed or overdesigned, i.e. either missing the target or an economical.

Therefore, this thesis is intended to study the Weyib catchment hydrological trend and variability to minimize the disasters on the water resource projects, Agricultural development, ecology and human being and study time series modeling, to combat the major challenges in data scarce area within weyib catchment.

1.3. Thesis Objectives

1.3.1. General objective

The general objective of this study is to analyze and identify whether there is a significant trend of precipitation and stream flow using statistical indicators, Analyze variability of rainfall and stream flow, to select suitable probability distribution for hydrological extreme event and describing the type of time series modeling equation for each stream flow in the Weyib River catchment.

1.3.2. Specific objectives

The specific objectives of this study will be:

- To analyze the stream flow and precipitation trend of each sub catchment.
- To model and forecast the future stream flow using hydrological time series model called ARIMA.
- To analyze spatio-temporal variability of stream flow and precipitation in the catchment.
- To identify best-fit annual and monthly rainfall probability distribution and analyze dry and wet return period.

1.4. Significance of the study

The significance of this study will be to update and expand information on stream flow, Precipitation characteristics and provide updated stream flow and precipitation statistics to water resource managing and planning sector. These help them to have good information about how to prepare for and reduce a devastating drought or flooding if there will be any in the future. In addition to this, the outcome of this study will give the updated information for the downstream users about the characteristics of the trend in precipitation and stream flow in the upper Weyib river Catchment so that they will have good information on the future water availability in the catchment.

1.5. Limitation of the study

In this study, the trend of stream flow and the trend of Precipitation was assessed using Mann-kendall non-parametric trend test. However, in real world a number of variables directly affect stream flow and probably will affect the results of trend analysis. Some of these variables are precipitation, land use/land cover change, water diversion and wastewater discharge. Similarly different factors such as, temperature, evapotranspiration and evaporation can affect the trend of precipitation. To determine the effect of each factor, the sensitivity analysis should have to be done which is not concerned in this thesis work due to time constraint.

1.6. Overview of the thesis

This thesis report consists of five chapters. The contents of each chapter are organized as follows: In the first chapter the background information, problem statement, general and specific objectives, Significance of the study and Scope of the study are discussed. In the second chapter, literature review about the subject matter is presented and it gives a scientific review this study is mainly based on. In the third chapter methodologies followed for determination of hydrological trend, variability and time series modeling are presented step-by-step and the models used for this exercise are described. Description of the study area, Data used in the study, their sources and the methods used for data quality control are mentioned. The fourth chapter presents the outcome of model application. It gives a detailed account of the model set up, model parameters, calibration and validation. The fifth chapter summarizes the contribution of this research and suggests related future research issues.

CHAPTER TWO

LITERATURE REVIEW

Essential information has been gathered from the available grey literatures published and from the different available sources, such as studies conducted on the Genale Dawa Basin and in the other different part of the world. Based on these, the useful methodology to achieve the objectives of this study has been developed.

2.1. Hydrological trends

Changes in hydrological series can take place in many different ways. A change may occur abruptly (step change) or gradually (trend) or may take more complex forms. Changes can be seen in mean values, in variability (variance, extremes, persistence) or within-year distribution. Abrupt changes can be expected because of a sudden alteration within the catchment. They can also inadvertently arise from changes to gauging structures, or to rating curves (stage-to-flow relationships), or to observation methods. Gradual hydrological changes typically accompany gradual causative changes such as urbanization, deforestation, climate variability, and other change. Although climate change is often thought of in terms of progressive trend, it is also possible for it to result in a step-like change because of complex dependencies on non-linear dynamic processes that feature cumulative effects and thresholds.

There is a huge variety of hydrological data that it is possible to analyses for trend and step change. These may be collected at a range of temporal intervals: continuous, hourly, daily, monthly, annually, or sampled irregularly. Data records contain either instantaneous values or totals for a time interval. Data may also pertain to different spatial scales, from point or experimental plot to large areas (Robson et al., 2000).

Studies of hydrological change are typically complicated by factors such as missing values, seasonal and other short-term fluctuations (climate variability) and by lack of homogeneity. In some cases, there are further problems because of censored data and data series that are not sufficiently long. The hydrological trends can be secular, periodic, and cyclical trend. All or some of these trends may occur in both runoff and precipitation time series, which is the main concern of this study.

2.2. Parametric and non-parametric trend test methods

Many tests for trend detection have been used in studies of long time series of hydrological data. Yet, every test requires a number of assumptions to be satisfied. When underlying test assumptions are not fulfilled, acceptance and rejection regions of the test statistic cannot be rigorously determined. Therefore, such tests should be treated as methods of exploratory data analysis rather than as rigorous testing techniques. Many approaches can be used to detect trends and other forms of non-stationery data in hydrology. In deciding which approach to take it is necessary to be aware of which test procedures are valid (the data meets the required test assumptions) and which procedures are most useful (likely to correctly find change when it is present).

Parametric test is a test that involves estimation of parameters and it is not rank based. Parametric testing procedures are widely used in classical statistics. In parametric testing, it is necessary to assume an underlying distribution for the data (often the normal distribution), and to make assumptions that data observations are independent of one another. For many hydrological series, these assumptions are not appropriate. Firstly, hydrological series rarely have a normal distribution. Secondly, there is often temporal dependence in hydrological series particularly if the time series interval is short. If parametric techniques are to be used, it may be necessary to (a) transform data so that its distribution is nearly normal and (b) restrict analyses to annual series, for which independence assumptions are acceptable, rather than using the more detailed monthly, daily or hourly flow series.

As stated by McCuen (1998), a parametric test is based on theory or concepts that require specific conditions about the underlying population and/or its parameters from which sample information will be obtained. Non-parametric test is a test that does not involve estimation of parameters and it is rank-based tests. In non-parametric and distribution-free methods, fewer assumptions about the data need to be made. With such methods, it is not necessary to assume a distribution. However, many of these methods still rely on assumptions of independence. More advanced approaches must therefore be used for daily or hourly series. A very useful class of non-parametric tests is permutation tests. They are based on changing the order (shuffling) of data points, calculating statistics, and comparing these with the observed test statistics.

McCuen (1998) stated that, a nonparametric test is based on theory or concepts that have not required the sample data to be drawn from a certain population or have conditions placed on the parameters of the population. Even within the basic categories above it is necessary to choose tests that are appropriate for the situation. Some tests are very good at detecting a very specific type of change; other tests may be good at picking up any one of a broad range of possible changes. Since one does not know the pattern of variability beforehand, using a number of tests is sensible.

The primary difference between the assumptions made for the two classes of tests is that those made for nonparametric tests are not as restrictive as those made for parametric tests, such as complete specification of the underlying population (Huth.et al., 2004). The assumption of normality, needed in the case of parametric tests, may be an unacceptably simplifying one in the context of strongly positively skewed hydrological data. In the case of non-parametric, robust tests, one does not need to assume a distribution.

Hirsch et al (1991) found that non-parametric procedures offer large advantages when the data are strongly non-normal, and suffer only small disadvantages (in terms of efficiency of power) for normally distributed data. Even though no distribution needs be assumed, non-parametric tests still make assumptions. Usually, an assumption of temporal independence must be made. When analyzing a time series of river flows, this assumption may be adequate for annual flow records. However, for shorter time intervals, such as months or days, it is not likely to hold.

There are many parametric and non-parametric tests for change detection. These are, Mann's test (non-parametric), Normal scores linear regression (non-parametric), Spearman's rank correlation (non-parametric), Linear regression (parametric), Jump fitting to normal scores (non-parametric), Jump fit to ranks (non-parametric) and Jump fit (parametric). Some parametric tests can be applied in a non-parametric way by testing either the ranks or the so-called "normal scores", i.e. the series transformed in such a way that the marginal distribution becomes normal, while the relative ranks of the values are preserved(Radziejewski et al., 2000).

Two common types of non-parametric tests used for detecting monotonic trend in a time series are Mann-Kendall (MK) and Spearman's rho (SR) test. Both MK and SR methods are rank based non-parametric tests. However, the MK test has been popularly to applied assess the significance of trends in hydrometer-logical time series (Yue et al., 2002(2)). The simulation experiments they made have demonstrated that SR test provided results almost identical to those obtained by the MK test. However, the SR test is seldom used in hydro meteorological trend analysis.

2.3. Non parametric Mann-Kendall trend test

Mann-Kendall's test is a non-parametric method, which is less sensitive to outliers and test for a trend in a time series without specifying whether the trend is linear or non-linear (Yenigun et al., 2008). The non-parametric mann-kendall test on the sample variable was first suggested by (Mann, 1945) using the test for significance of Kendall's tau. The Mann-Kendall test is based on the null hypothesis that a sample of data is independent and identically distributed, which means that there is no serial correlation or trend among the data points.

According to McCuen (1998), this test is designed to detect a monotonically increasing or decreasing trend in the data rather than an episodic or abrupt event. The Kendall test may detect watershed changes due to either gradual trends or abrupt events. However, it is more sensitive to changes that result from gradually changing trends. This test is non-parametric test, and has been widely used to test for randomness against trend in hydrology and climatology (Zhang et al., 2000).

The problem in using Mann-Kendall test is that the result would be affected by serial correlation of the time series. If there is a positive serial correlation (persistence) in the time series, the test will suggest a significant trend in a time series, which is actually random more often than specified by the significance level (Kulkarni and Von Storch, 1995). To remove the effects of serial correlation (Von Storch, 1995) suggest that the series be "pre-whitened" before applying the Mann-Kendall test. Here assumption of normality is not needed, but there must be no serial correlation for the resulting p-values to be correct.

The Mann-Kendall test possesses the useful property of other nonparametric tests in that it is invariant to (monotonic) power transformations such as those of the ladder of powers. Since only the data or any power transformation of the data need be distributed similarly over time except for their central location in order to use the Mann-Kendall test, it is applicable in many situations (Hirsch and Helsel, 2002).

2.4. Monotonic change versus step change

Monotonic change is a change that is consistently in one direction (either always upwards or always downwards) and step change is abrupt change in a time series. Two primary types of long-term trends can be considered in hypothesis testing and trend estimation, one is the monotonic trend and the other is the step trend (Hirsch et al., 1991).

Step trend test will be applied for those stations with naturally broken in to two distinct periods with relatively long time gap between them and the other is when there is human influence or diversion structure which likely result in a change in stream flow. Monotonic trend test will be applied for those stations with no human influence and diversion (Helsel and Hirsch, 1992).

2.5. Purpose of trend testing

The issue of change detection in hydrological data is of much practical importance. Most water resources systems are designed and operated under the assumption of stationary. This implies that the essential characteristics of variability of hydrological processes do not change with time. If this assumption were abandoned, existing codes of design of water resources systems, dams, levees and other water engineering works would have to be revised. Otherwise, the systems would be either under designed or overdesigned, i.e. either missing the target or an economical.

According to Jain and Kumar (2012), to understand rainfall variability and trend is crucial and necessary in order to appreciate the impacts of climate change and also as a basic and important requirement for the planning and management of water resources.

According to Hirsch et al (2002), the purpose of trend testing is that, if we have a series of random variable that have been collected over some period of time, we would like to determine whether their values are decreasing or increasing (getting better or worse). In statistical term, this can be determined as whether the probability distributions from which they arise have been changed over time. We would also like to describe the amount or rate of that change, in terms of changes in some central value of the distribution such as a mean or median.

2.6. Type of changes

Change in a series can occur in numerous ways: steadily (a trend) and abruptly (a step- change) or in a more complex form. It may affect the mean, median, variance, autocorrelation or almost any other aspect of the data. The most widely used tests for change are look for, trend in the mean or median of a series and step-change in the mean or median of a series.

Trend and step change are special cases of a change in distribution. Tests for a change in distribution are generally not particularly powerful. If trend were present, it would be best detected by a test for trend. However, such tests may be useful as a general check for evidence of change. Test for change that is more complex and for measures other than the mean/median generally requires use of advanced techniques such as Maximum Likelihood (Robson et al., 2000). Typically, Maximum Likelihood techniques are beyond the scope of this study.

2.7. Hypotheses testing

Hypothesis is a supposition or proposed explanation made based on limited evidence as starting point for further investigation. Hypothesis testing represents a class of statistical techniques that are designed to extrapolate information from samples of data to make inferences about populations. In problems of hydrologic analysis, it is frequently desirable to make a statistical test of some hypothesis. Specifically, we may want to use sample data to draw inferences about the underlying population. Decisions should be made with the population parameters, not the sample statistics where the population parameter is computed from sample information; a statistical hypothesis test may be made to provide verification of some hypothesis (Maidment, 1986).

The starting point for a statistical test is to define the null and alternative hypotheses; statements that describe what the test is investigating. The null and alternative hypotheses are usually framed in terms of the types of change. For example, to test for trend in the mean of a series the null hypothesis (Ho) would be that there is no change in the mean of a series, and the alternative hypothesis (H_A) would be that the mean is either increasing or decreasing over time. To test for step-change in the mean of a series, the null hypothesis would again be that there is no change in the mean of the series, but the alternative hypothesis would be that the mean of the series has suddenly changed. Assume that the null hypothesis is true, and then to check whether the observed data are consistent with this hypothesis. The null hypothesis is rejected if the data are not consistent. As discussed, the procedure for trend test depends on the hypotheses test and regression. Therefore, discussion about hypotheses test and procedure of testing hypotheses is very essential.

2.8. Hydrological System

Complex systems can be decomposed into subsystems, each having an input output linkage as a component. Hydrologic systems can be considered as a subsystem of water resources representing the physical functioning of that system in a region. A hydrologic system is defined as a structure or volume in space, surrounded by a boundary, that accepts water and other inputs, operates on them internally, and produces them as outputs (Chow et al., 1988).

2.8.1. Hydrologic Cycle

The hydrologic cycle is defined as "the pathway of water as it moves in its various phases through the atmosphere to the Earth, over and through the land, to the ocean, and back to the atmosphere" (National Reaserch, 1991). The hydrologic cycle can be considered as a closed system for Earth, because the total amount of water in the cycle is constant. The hydrologic cycle is generally described in terms of six major components: and some subcomponents Figure 2.1

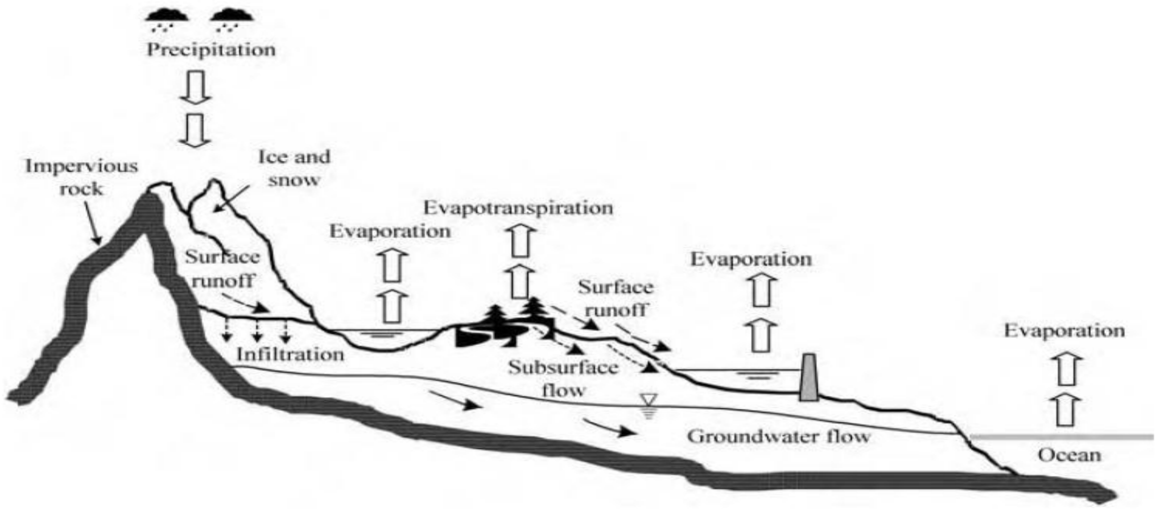


Figure2. 1: Schematic diagram of the hydrologic cycle after Mohammad et al. (2003)

Any change and/or fluctuations in the hydrological cycle directly affect the availability and quality of fresh water, which is a major environmental issue of the 21st century (Kim et al., 2008).

2.8.2. Hydrologic variables and parameters

All natural physical processes are subject to variability. For example, rainfall intensity, flood magnitude, or low flows in droughts have wide variations. To study these variations and incorporate them in the planning and operation of water resources, engineers gather and investigate samples of data.

The hydrologic cycle is composed of various phenomena, such as precipitation, runoff, infiltration, evaporation, evapotranspiration, and abstraction. Different characteristic variables, which can simply be called hydrologic variables, have been defined to describe each of these phenomena. Depth or intensity of rainfall in different time steps of a rainstorm, monthly inflow discharge to a reservoir, or daily evaporation are some examples of hydrologic variables (Shahim et al., 1993).

A dataset consists of a number of measurements of a phenomenon, and the quantities measured are variables. A continuous variable can have any value on a continuous domain; examples include volume of water flowing in a river or the amount of daily evaporation measured in a climatic station. A discrete variable represents an interval or the number of occurrences within each interval of time and space; the number of rainy days in a certain period of time (e.g., a year) is an example of discrete hydrologic variables.

The hydrologic variables can also be classified as qualitative or quantitative. A qualitative variable can be expressed as a real number in a sensible way; type of soil is an example of a qualitative hydrologic variable. A quantitative variable can be measured as a real number; the numbers of rainy days in a year and rainfall intensity in a day are examples of discrete and continuous quantitative hydrologic variables, respectively.

Hydrologic variables usually vary in time and space. A time series is a sequence of values arranged in order of their occurrence in time. Conflicts can arise when trying to analyze the impact of various hydrologic variables on each other and model them during the decision-making process (Mohammad et al, 2003).

2.9. Factors affecting hydrological trend

Drivers that could affect natural flow regimes and probably will affect the results of trend analysis are mainly climate variability and human activities such as construction of water retention structures (Beavis et al., 1997), deforestation and clearing of land cover, expansion of agricultural land (Masih et al., 2011), urbanization and catchment. Change and increased abstraction of water for irrigation and industries, impoundment of water (Alemayehu et al., 2007) and modification of the morphology of the riverine system. To determine the effect of each factor on the runoff trend, the sensitivity analysis should have to be done which is not concerned in this thesis work.

2.9.1. Precipitation effect

Rainfall is known as the main contributor to the generation of surface runoff trend. Therefore, there is a significant and unique relationship between rainfall and surface runoff. By basic principle of hydrologic cycle, when rain falls, the first drops of water are intercepted by the leaves and stems of the vegetation. This is usually referred to as interception storage. Once they reach the ground surface, the water will infiltrate through the soil until it reaches a stage where the rate of rainfall intensity exceeds the infiltration capacity of the soil.

Apart from rainfall characteristics such as intensity, duration and distribution, there are other specific factors, which have a direct bearing on the occurrence and volume of runoff. The most common factor is the soil type.

An upward trend in precipitation likely results in an upward trend in high and low flows. The reverse results with a downward trend in precipitation. The season of the year along with an upward or downward precipitation trend may not affect the low or high stream flows. If the trend in precipitation is upward in winter and spring and downward in summer and fall, the result could be an upward trend for high flows but a downward trend for low flows (K. Schopp, 1897).

2.9.2. Land use/Land cover effect

Another factor that can affect the runoff trend is vegetation. Meta-analyses of paired catchments studies have found that afforestation typically results in decreased stream flow while deforestation typically leads to increased stream flow (Brown et al., 2005).

However, the hydrologic response to deforestation is in general more consistent than the response to afforestation. This difference may be due to higher variability in land cover following afforestation compared to deforestation, and the effects of different transitional species and/or changes in forest physiology (Andreassian, 2004). An area, which is densely covered with vegetation, produces less runoff than bare ground while the amount of rain lost to interception storage on the foliage depends on the kind of vegetation and its growth stage. Vegetation has a significant effect on the infiltration capacity of the soil. A dense vegetation cover shields the soil from the intense raindrop impact which eventually will cause a breakdown of the soil aggregate as well as soil dispersion with the consequence of driving fine soil particles into the upper soil pores. This results in clogging of the pores, formation of a thin but dense and compacted layer at the surface, which highly reduces the infiltration capacity.

This particular effect is often referred as to capping, crusting or sealing. In addition, the root system as well as organic matter in the soil increases the soil porosity thus allowing more water to infiltrate. Vegetation also retards the surface flow particularly on gentle slopes, giving the water more time to infiltrate and to evaporate. (Baharudin, 2007).

2.10. Time series analysis and modeling

A time series is defined as an ordered chronological sequence of observations on particular variable. A time series can be composed of a quantity, observed at discrete times, averaged over a time interval, or recorded continuously with time. An ensemble of time series is a set of several time series measuring the same variable. A single time series is called a realization. Thus, an ensemble is made up of several realizations. A time series may be composed of only deterministic events, only stochastic events or a combination of the two. Most generally, a hydrological time series will be composed of a stochastic component superimposed on a deterministic component. For example, average daily temperature and monthly stream flow data, at some point would contain seasonal variation, the deterministic component, plus random deviation from the seasonal values.

The deterministic components may be classified as a periodic component, a trend, a jumper a combination of these. Trend in hydrological time series can result from gradual natural or man-induced changes in the hydrologic environment producing the time series. Jumps in time series may result from catastrophic natural events such as earthquakes or large forest fires that may quickly and significantly alter the hydrologic regime of an area and man made changes such as constructing a new dam and pumping of ground water may also cause jumps in certain hydrologic time series. The properties of a time series can be obtained based on a single realization over a time interval or based on several realizations at a particular time.

McCuen (1993) found that, time series are analyzed for a number of reasons. One might be to detect a trend due to another random variable. Second, time series may be analyzed to formulate and calibrate a model that would describe the time-dependent characteristics of a hydrologic variable. Third, time series models can be used to predict future values of a time-dependent variable.

The properties based on several realizations at a given time are known as the ensemble properties. If the time average properties and the ensemble properties are the same, the time series is said to be argotic. In actual practice, random data representing stationary physical phenomena are generally argotic. In hydrology, time series models are often used for determining the design capacity of reservoirs. From a shorter record of flow data, it is difficult to determine the required reservoir capacity in a reliable manner.

However, if a time series model is developed, we can synthetically generate many realizations of length equal to the design life of reservoir and having similar statistical properties to those of the available data. Time series have also found application in the area of short-term river flow and an intermediate flow forecast.

2.10.1. Stationary and Non stationary time series

If the statistics of the sample (mean, variance, covariance, etc.) are not functions of the timing or the length of the sample, then the time series is said to be stationary to the second order moment, weekly stationary, or stationary in the broad sense. In hydrology, moments of the third and higher orders are rarely considered because of the unreliability of their estimates. Second order stationarity, also called covariance stationarity, is usually sufficient in hydrology. A process is strictly stationary when the distribution of variables does not depend on time and when all simultaneous distributions of the random variables of the process are only dependent on their mutual time lag. In another words, a process is said to be strictly stationary if its n^{th} (n for any integers) order moments do not depend on time and are dependent only on their time lag.

If the values of the statistics of the sample (mean, variance, covariance, etc.) are dependent on the timing or the length of the sample, then the time series is said to be non-stationary i.e. if a definite trend and periodicity are visible in the series, then it is a non-stationary (Chong-yuXu, 2002).

2.10.2. Types of time series model

A mathematical model representing a time series or stochastic process is called a time series model. Since 1970, researchers in hydrology have used methodologies suggested by (Box and Jankin, 1970) especially for improving the estimate of the parameters of the model, for verifying the conditions to be met by those parameters, for verifying or checking the assumptions of the model and for selecting among competing models.

2.10.2.1. Autoregressive modeling

Autoregressive (AR) models have been used widely in hydrologic time-series modeling. These models incorporate the correlation between time sequences of variables. These models are the simplest models and their development goes back to the application of Markov lag-1 models. They can be classified into the following subsets:

- AR models with constant parameters, which are typically used for modeling of annual series.
- AR models with timely variable parameters, which are typically used for modeling of seasonal (periodic) series.

2.10.2.2. Moving average process

Moving average (MV) models account for the possibility of relationship between a variable and the residuals from previous periods. The autoregressive models can be used as an effective tool for modeling hydrologic time series such as low flow season stream flow that is mainly supplied from groundwater and has low variations. However, the previous studies have shown that the stream flows in high flow season can be better formulated by adding a moving average component to the autoregressive component (Salas et al., 1988).

2.10.2.3. Autoregressive moving average (ARMA) modeling

The ARMA model of order (p, q) can be defined by combining an autoregressive model of order p and a moving average model of order q. The experience of applying AR (p) models to the real-world data might result in significant simulation errors. The reason for this is related to the fact that the occurrence of Z_t (the data value at the time t) is considered as a function of its previous values. It is a fact that the effects of other predictors are neglected in the conventional AR (p). Suppose that we are modeling the stream flow of a river by an AR (p) model. Obviously, the stream flow is a function of the base flow, snow budget, soil moisture, etc., which are not considered in the AR model at all. A solution to mitigate this shortcoming of the AR models is to consider the error term, ε_t, as a representative of all other predictors except the persistency of the time series data.

The order of an ARIMA model is usually denoted by the notation ARIMA (p, d, q), where, p is the order of the autoregressive part, d is the order of the differencing, q is the order of the moving average. If no differencing is done (d = 0), the models are usually referred to as ARMA (p, q) models.

2.10.2.4. Autoregressive integrated moving average (ARIMA) modeling

ARIMA stands for Autoregressive Integrated Moving Average with each term representing steps taken in the model construction until only random noise remains. An ARIMA model predicts a value in a response time series as a linear combination of its own past values, past errors (also called shocks or innovations), and current and past values of other time series. Box and Jenkins first popularized the ARIMA approach, and ARIMA models are often referred to as Box-Jenkins models. Use ARIMA to model time series behavior and to generate forecasts. ARIMA modeling differs from the other time series methods discussed in this chapter in the fact that ARIMA modeling uses correlation techniques.

The ARIMA models are suitable for the data that have the following two basic characteristics:-

- No apparent deviation from stationarity
- Rapidly decreasing autocorrelation function

If these conditions are not met by a time series, a proper transformation should be performed to generate time series satisfying the above two conditions. This has usually been achieved by differencing, satisfying the essence of ARIMA models. This class of models is also very powerful for describing stationary and no stationary time series.

CHAPTER THREE

METHODS AND MATERIALS

3.1. Description of the study area

3.1.1. Location

Weyib River catchment is located in the Genale Dawa river basin, which is one of the twelve drainage basins in the country between $6^{0}22'$ to $7^{0}06'$N and $39^{0}37'$ to $41^{0}15'$E. The Weyib River originates from the northern flanks of the Bale Mountains and first flows generally north-eastwards before changing direction to east and south- eastwards for the remainder of its course. The catchment of Weyib River covers the area around 7609.6sqkm. Its major tributaries are Shaya, Tegona, and Tebel Rivers.

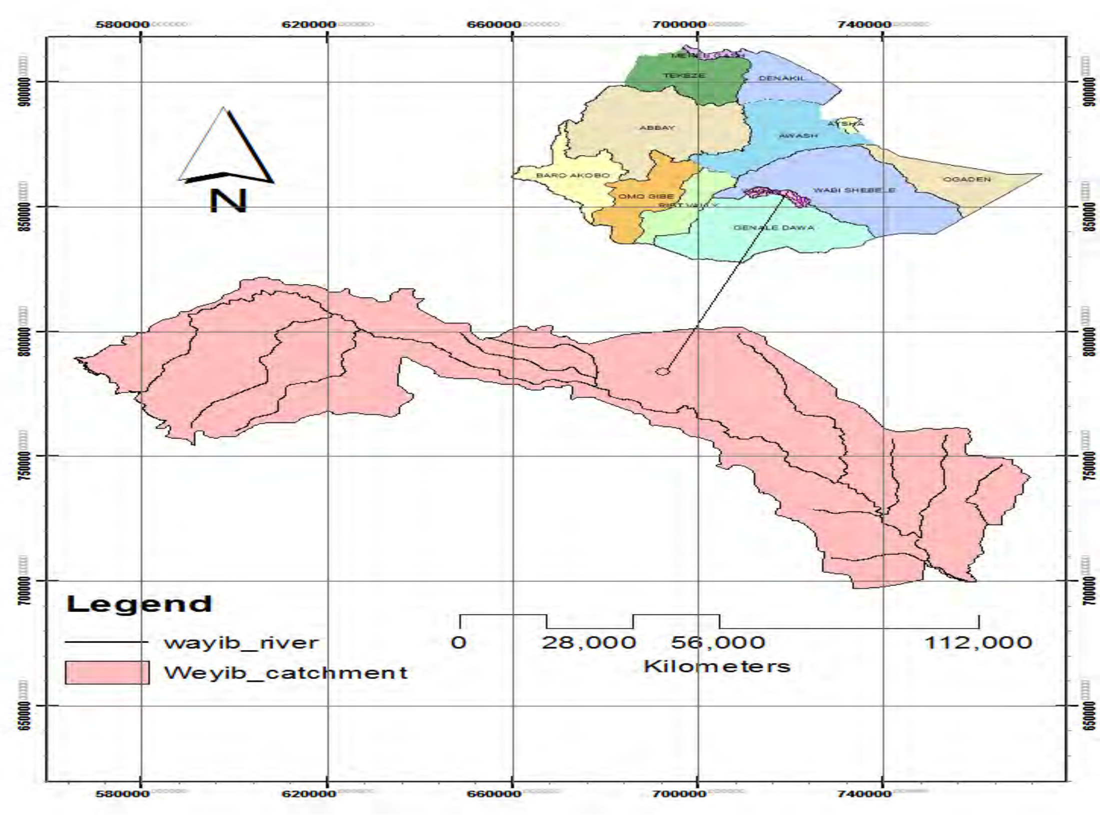


Figure3. 1: Map showing Weyib River catchment locations

3.1.2. Climate

The main wet season, locally known as Kiremt and a minor rainy season, locally known as Belg are the two distinct seasonal weather patterns that characterize the climate of the Weyib catchment .The wet season runs from July to mid-October. The dry season spans from November to February. However, in some part of the catchment, there is a third season with moderate rainfall (Belg) occurring from mid-March to mid-Jun.

3.1.3. Rainfall

The variation in the seasonal distribution of rainfall in Ethiopia can be attributed by the reference to the position of the Inter-Tropical Convergence Zone (ITCZ), the relationship between upper and lower air circulation, the effects of topography and the role of local convection currents and the amount of rainfall (Daniel, 1977).

The seasonal rainfall distribution within the study area results from the annual migration of the ITCZ. The rainfall pattern of Weyib Catchment follows symmetric bimodal profile with peaks in September and April Months.

In the analysis, for each month in a year, spatial rainfall statistics was computed using all gauges with available data for the period.

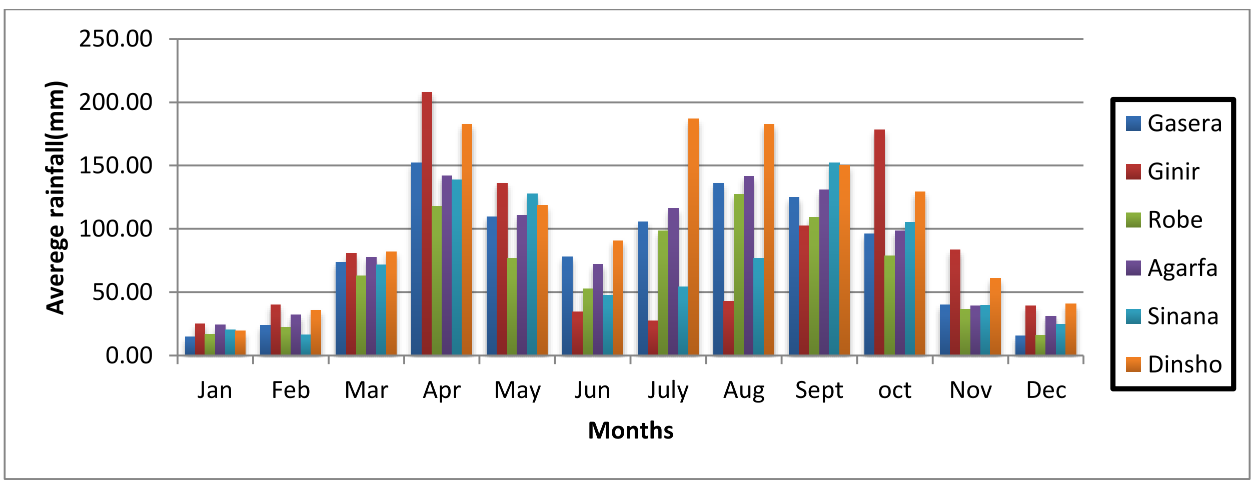


Figure3. 2: Mean monthly rainfall distribution

3.1.4. Temperature

Air temperature of the Weyib River catchment was analyzed using monthly minimum and maximum data from four stations. As depicted in (Figure 3.3), the monthly distributions of the Weyib River catchment temperature suggest that the maximum occurs in the month of March

(15.559 °c) and the minimum in the month of November (13.369 °c). On the average in the Weyib River catchment, there is a drop in average air temperature of 1 °c for every 161 m increase in elevation.

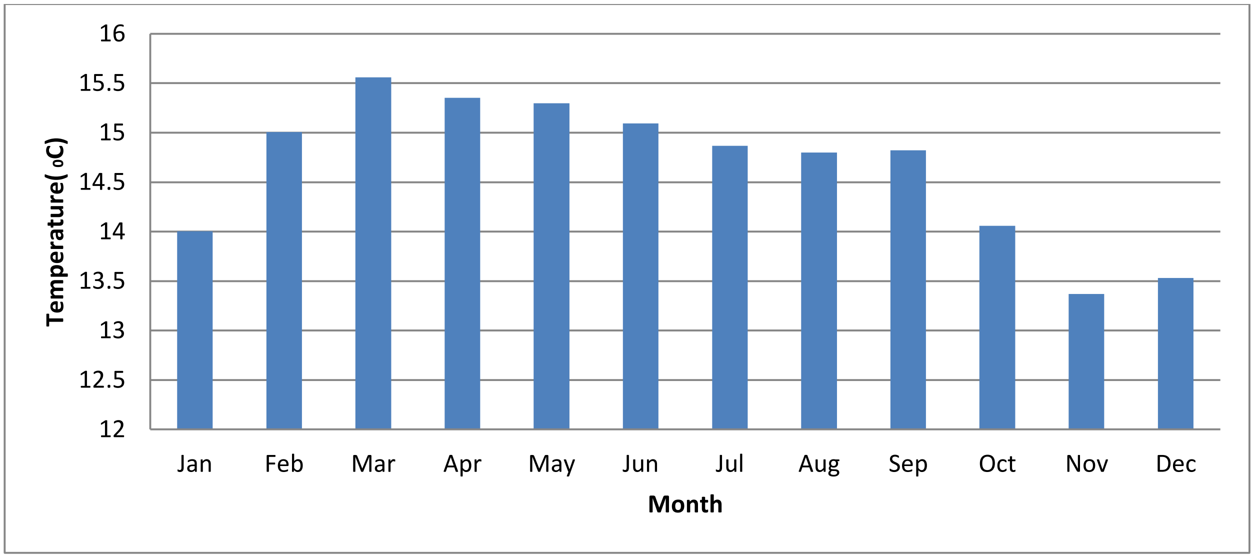


Figure3. 3: Monthly distribution of mean air temperature in the Weyib River catchment

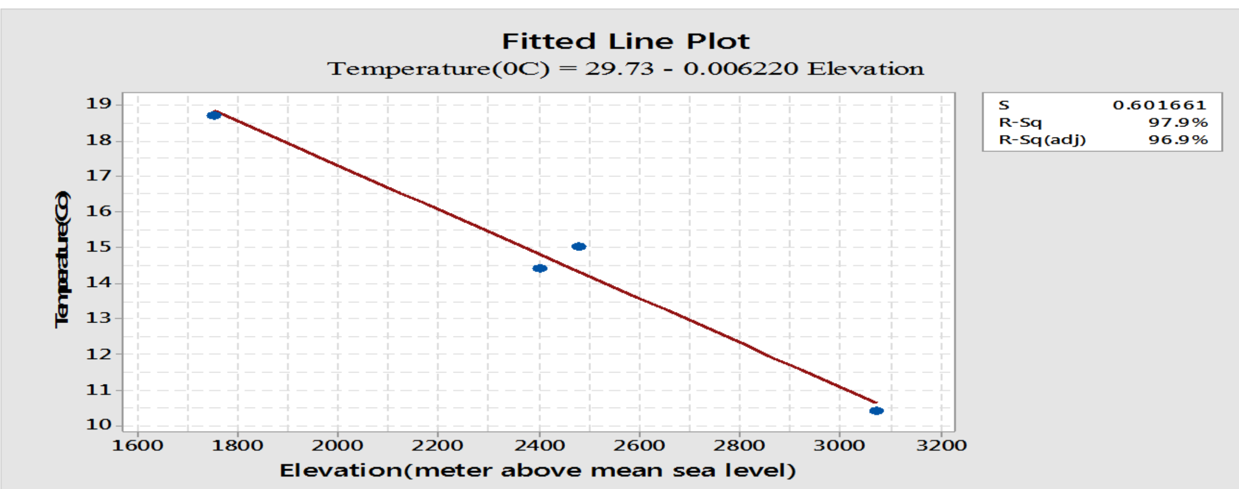


Figure3. 4: Air temperature and elevation relationships in the Weyib River catchment

3.1.5. Topography

The upper reaches of the River is fairly forest land with rugged slope of Batu mountainous ridge, while the lower part of the drainage area is narrowed gorge and finally flows to Genale River which is located in Genale Dawa River Basin.

3.1.6. Soil

According to (FAO, 2002), the soils of the Weyib River catchment can be divided into four broad groups for agro-ecological purposes.

a) The shallow soils, with limited or zero agricultural potential are generally grouped as Leptosols and Regosols and are commonly associated with steep slopes. These cover large areas of the catchment, and their depth is determining in recognition of the associated agro-ecological units.
b) The moderately deep soils, classed as Cambisols, some Luvisols include most of the lighter textured soils of the catchment, and they are therefore suitable for cultivation. However, the limited depth demands careful management to avoid erosion leading to further depth constraints.
c) The heavy textured vertisols are deep soils with poor drainage and properties of swelling and cracking. Under careful management, and assuming that drainage is not too problematic, they can be highly productive. Due to the clay texture, they are difficult to manage with traditional agricultural management and even pose problems for mechanized production.
d) Solonetz are slowly permeable to water and subsequent ponding of water on top is a common problem normally associated with flat lands in a climate with hot and dry summer. Contains a high proportion of sodium ions, which affect arable cropping either directly or indirectly. Solonetz in semi-arid regions are mostly used as rangeland or lie idle. Arenosols have a course texture accountable for the generally high permeability and low water and nutrient storage capacity and used for little more than extensive grazing but they could be used for arable cropping if irrigated.

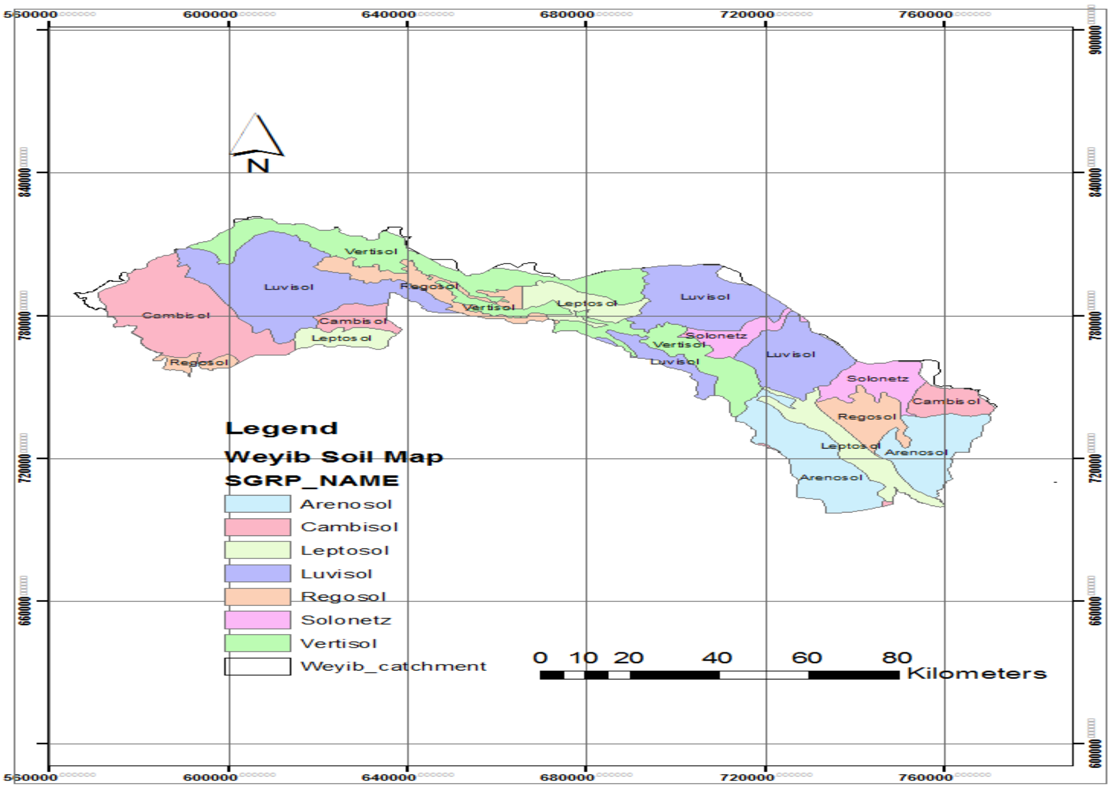


Figure3. 5: Soil map of weyib catchment

3.1.7. Land cover and Land use

Land use map of the Weyib River catchment was obtained from Oromia water works design and supervision enterprise in soft copy. Due to the variation in the topography, rainfall distribution and landscape, the catchment has a mosaic of land cover including 37.2% (Cultivated land), 36.2% (Shrub land), 15.9% (Bush land), 5.6% (Afro alpine), 2.4% (Settlement), 2% (Forest), 0.4% (Grass land), 0.1% (Plantation) and 0.1% (Exposed surface). Population, remoteness, and traditional factors attribute to the type of use and the natural vegetation as they are presently expressed in the catchment. In the arid and semiarid areas grasses, shrubs, and sparsely cultivated lands are common. On the extreme where the climate is arid, exposed rock or sand surface is predominant land cover. The wide occurrence of shrubs is associated with the less population pressure and cultural orientation of the people, which is of pastoral farming system Figure 3.6.

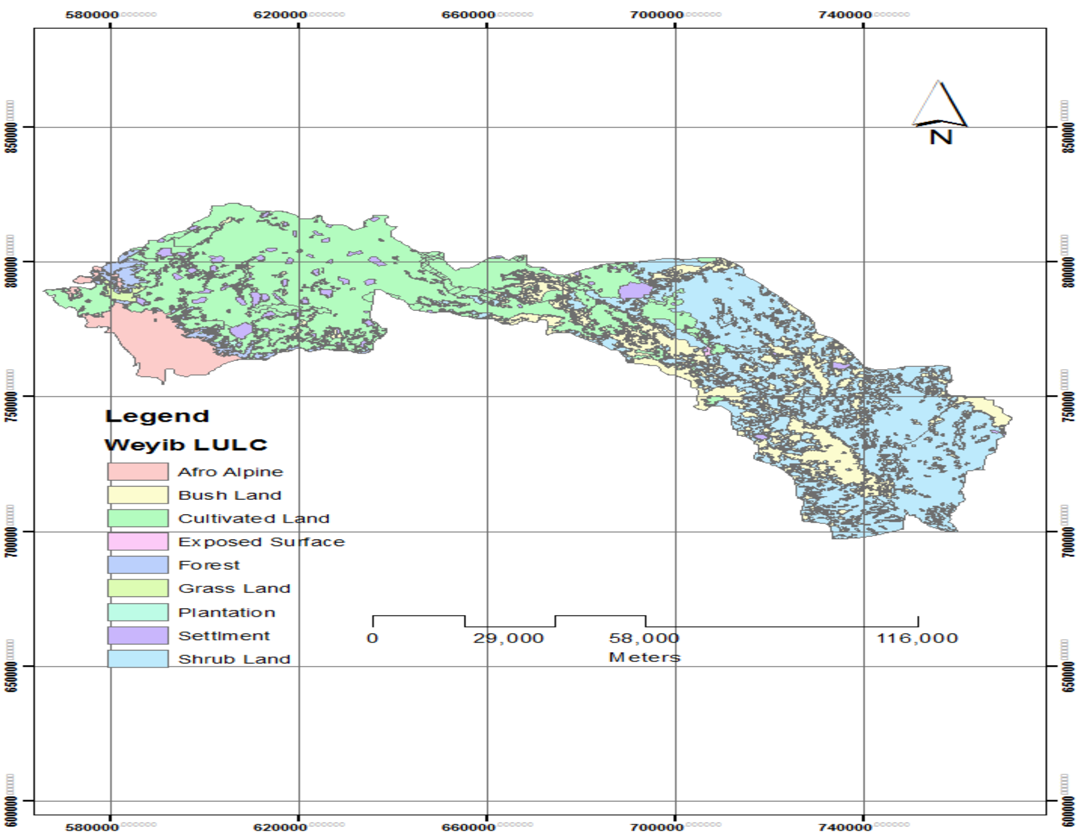


Figure3. 6: Land cover and Land use of Weyib River catchment.

3.2. Data source and Analysis

To be able to access the long-term trends in stream flow and precipitation in the Weyib River catchment the whole weyib catchment is divided into nine-sub catchment.

Dinsho sub catchment, Sinana sub catchment, Agarfa sub catchment, Gassera sub catchment, Ginir sub catchment, Goro sub catchment, Goba sub catchment, Dawe Kachen sub catchment, and Raitu sub catchment. Each of this sub catchment can be divided further in sub catchments according to the long-term data availability and completeness. The first seven sub catchments had a gauging station located within their catchment while the last two sub catchments do not have gauging station but these sub catchments had still contribution to the catchment as runoff and precipitation.

This division of catchments in to smaller woreda sub catchments is essential for better understanding of hydrological trend and variability at woreda scale.

Furthermore, since the data of the research is collected only for seven woreda sub catchments to represent the total catchment, it is necessary to identify which catchment's data were used for this study.

Those of sub catchments in which runoff-gauging stations present were described in Figure 3.8 briefly. These sub catchments include, Agarfa sub catchment, Ginir sub catchment, Gasera sub catchment, Goba sub catchment and Sinana sub catchment.

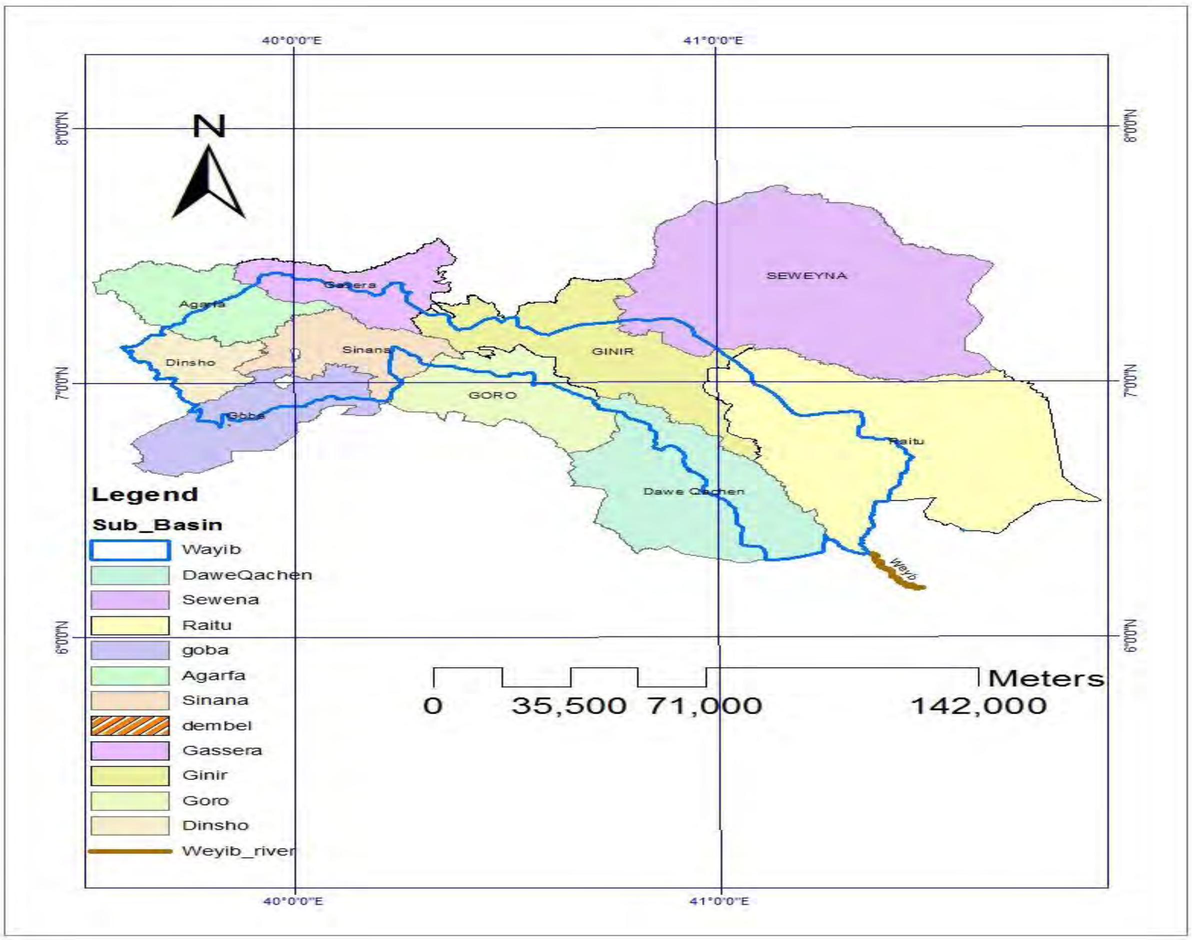


Figure3. 7: Map showing the catchment and the sub catchment of the study area.

3.2.1. Agarfa sub catchment

This sub catchment is located to Left hand side of the study area Figure 3.7. This woreda catchment covers the area of 1221.8sqkm.The area which contributes to the catchment is 533.7sqkm while the reset is outside the catchment. Agarfa stream flow stations and Agarfa precipitation station is located in this sub catchment.

Agarfa stream flow station

The discharge of Agarfa gauging station was recorded since 1981. This gauging station is located on the Weyib River at the top most of the Agarfa sub catchment and after Denka intermittent stream joining it. The geographical location of this gauging station is at latitude of 7:22:0 N and 39:48:0E longitude. The high flow period of this gauging station, run from July to September while the low flow period extend from November to January.

3.2.2. Ginir sub catchment

Ginir sub catchment is found at the central part of the study area covering an area around 2364.745sqkm; out of which 1751.42sqkm is contributing to Weyib River catchment. This sub catchment includes Tebel flow station and Ginir precipitation station. Although the Sofumer flow, station is located on Weyib River in between boundary of Ginir and Dawe Qachen sub catchment it was not used for this thesis work as the data have very poor quality and very short recording year.

Tebel gauging station

The record of this gauging station starts from 1983 E.C. The river of this gauging station starts from Melka oda district which is located at the top most of Ginir sub catchment and then which latter on join the major river. The geographic location of this gauging station is in between 7:11:0N and 40:44:0E.

Alemkerem gauging station

This gauging station is found at 6:59:0N and 40:58:0E. The drainage area of this gauging station is 3576.9 sqkm. The record period of Alemkerem gauging station starts from 1984.This gauging station is located on the Weyib River at the boundary of Goro, Sinana and Ginir sub catchment and after Asmedaba intermittent stream joining it.

3.2.3. Gasera sub catchment

Gasera sub catchment has an area of 1128.9sqkm out of which 503.9sqkm is contributing to the Weyib catchment and the remaining cannot contribute to this catchment. Denbel flow station and Gasera precipitation station was located in this sub catchment. The flow station is located on the

Weyib River at the top most of the Gasera sub catchment and after Wegerge intermittent stream joining it.

3.2.4. Sinana sub catchment

Sinana sub catchments have an area of 1078.96sqkm; out of which 991.08sqkm is, contributing to the Weyib catchment and the remaining cannot contribute to this catchment. Robe Precipitation station, Sinana Precipitation station and Shaya flow station was located in this sub catchment.

Shaya gauging station

This gauging station was located on the Shaya River, which is one of the rivers, which contribute to weyib water shade. The Shaya River starts from the Bale Mountain in the Goba sub catchment. The topography of the river composed of highlands, hills, valleys and occasional rock peaks. Act as boundary between Agarfa sub catchment and sinana sub catchment. Most of the streams feeding the Shaya River are perennial and includes Bamo, Lola, Anole, Chorino and Kokete and this River finally meet Weyib River at Gasera sub catchment.

3.2.5. Goba Sub catchment

Goba sub catchment has an area of 1490.3sqkm out of which 593.7sqkm is contributing to the Weyib catchment and the remaining cannot contribute to this catchment. Togona flow station was located on the Togona River, which is one of the rivers, which contributes to weyib water shade.

Togona gauging station

This gauging station is located at 7:0:0N and 39:59:0E.The drainage area of this gauging station is 83.1sqkm. The record period of Togona gauging station starts from 1983 E.C. The River of this gauging station starts from Goba sub catchment and then joins the major river at the down in Gasera sub catchment after it joins two rivers Micha and Magida intermittent river.

3.3. Data collection and preprocessing

3.3.1. Data collection

It is very important to collect adequate and quality data to do successful thesis work on any field of study. For this thesis work, a number of data were collected from their respective sources. These data includes metrological data, stream flow data and shape file data. The material and the data which are collected from different agencies are given in Table 3.1 with their respective sources, types of the data and the amount (number of the data item).

Table 3. 1: Source and types of data

Types of data	**Agencies (organization)**	**Type and number of data**
Meteorological data	National Meteorological Agency	Ten precipitation station
Stream flow data	Ministry of Water, Irrigation and Electricity	Seven stream flow station
Soil data	Ministry of Agriculture and Natural Resource	One shape files
Land use/cover data	Oromia Water Works Design and Supervision enterprise	Two shape file

In general, seven stream flow station, three shape files, and ten precipitation station data were collected from their respective data source.

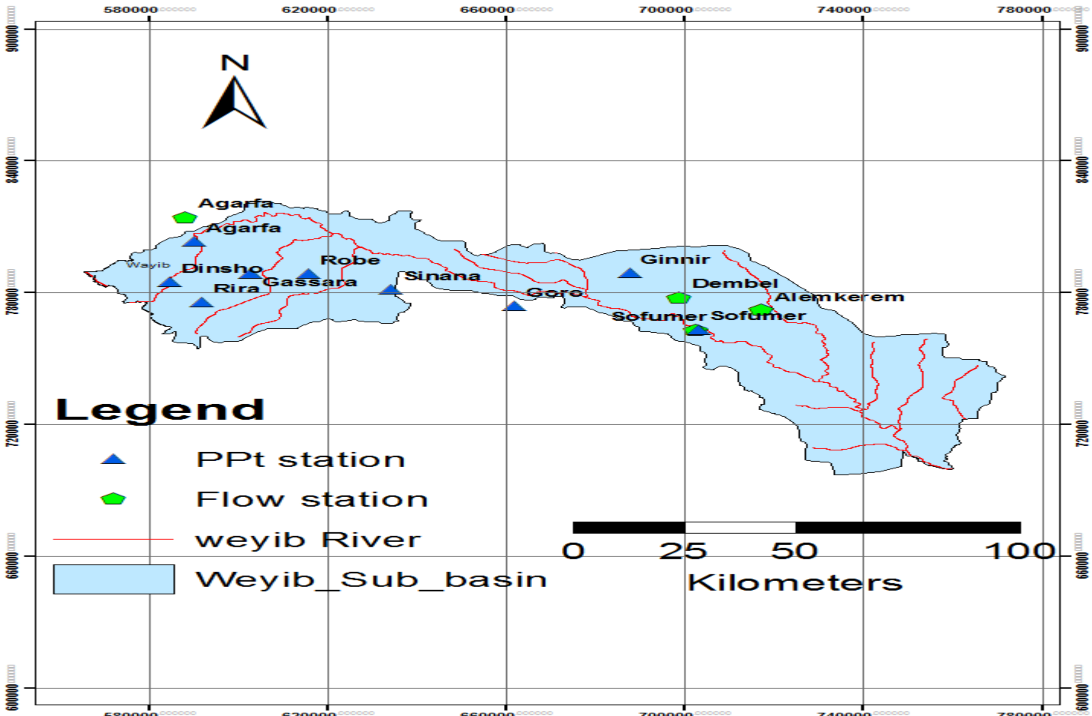


Figure3. 8: Flow and Precipitation gauging station network in the Weyib catchment

Table 3. 2: List of runoff data availability

No	Flow station	River	Station No	Catchment area(sqkm)	Record year
1	Agarfa	Weyib	073003	771.9	1981-2008
2	Shaya	Shaya	073006	433.8	1981-2007
3	Togona	Togona	073007	83.1	1983-2007
4	Tebel	Tebel	073011	79.0	1983-2007
5	Denbel	Denbel	073004	1215.0	1984-2004
6	Alemkerem	Alemkerem	073005	3576.9	1986-2004
7	Sofumer	Weyib	073009	3792.7	1992-2005

The Sofumer gauging station was not used for analysis because of the data of this station was not recorded for long period and have poor quality. This gauging stations was left unconsidered again due the fact that the selection criteria of the gauging stations were depending on selection criteria adopted by (D.H.Burn and A. Mohammed, 2002) which says station can be selected from the available stations with a minimum record length of 25 years to insure the statistical validity of the trend results.

Table 3. 3: List of rainfall stations

N<u>o</u>	**Station name**	**Latitude(° N)**	**Longitude(° E)**	**Elevation(m)**	**Record year**
1.	Dinsho	7.1	39.76667	3072	1984-2015
2.	Robe	7.133333	40.05	2480	1985-2015
3.	Sinana	7.066667	40.21667	2400	1984-2015
4.	Agarfa	7.26667	39.81667	2550	1984-2015
5.	Gasera	7.13333	39.93333	1680	1984-2015
6.	Goro	7	40.4667	2334	1999-2015
7.	Ginir	7.133333	40.7	1750	1989-2014
8.	Sofumer	6.9	40.84	1680	1987-2013
9.	Rira	7.01667	39.83333	2700	2007-2012
10.	Megalo sabro	6.8459	40.7999	1573	2007 -2012

From ten precipitation stations that are collected from National Meteorological Agency, only six of them are analyzed as representative stations of the study area. Rira, Magalo sabro and Sofumer is not analyzed because of the first two station data's were recorded for a short period of time while that of Sofumer gauging station is left an analyzed because the data collected have a lot of missing data and filling all the missing is led to an error by any method filling missing data. In addition, although the precipitation data of Agarfa, Robe, Dinsho, Sinana and Gasera stations are collected since 1984- 2015, there trend is analyzed up to 2014. This is because of the data after year 2014 were of poor quality and having a lot of missing value.

3.3.2. Data Quality Analysis

The raw data is found to be highly distorted and non-sequential. Therefore, before any analysis of data, it is very necessary to go through the collected raw data deeply to avoid any misbehavior of the data such as, missing data, outliers and consistency. In particular, the following steps were taken in this study to improve the quality of the raw data collected.

- Checking for consistency
- Filling the missing data
- Test for outliers

3.3.2.1. Consistency checking

A hydrologic time series data may be inconsistent with time due to change in physical condition of the catchment, method of data collection and negligence of the observer (James, 1960). In order to check the consistency of the data, a mass curve technique is used. Cumulative of mean annual precipitation data of the stations is used as a pattern for testing the individual station records. The mass curve for individual station is plotted against the cumulative of the pattern. If a break in slope is observed then the data of the station is adjusted by multiplying it with the ratio of the two slopes (Equation 3.1).

$$P_a = \left(\frac{b_a}{b_o}\right) P_o \tag{3.1}$$

Where;

P_a = adjusted precipitation

P_o = observed precipitation

b_a= slope of graph to which records are adjusted

b_o= slope graph at time P_o was observed

3.3.2.2. Test for outliers

An outlier is an observation that deviates significantly from the bulk of the data, which may be due to errors in data collection, recording, due to natural causes etc. Outliers should have to be investigated because they can provide useful information about your data or process. Unless the outliers are detected and corrected, they may result in unreliable result in both trend test and time series modeling case. Several explanations for the occurrences of outliers are:-

- Data entry error: Correct the error and reanalyze the data.
- Process issue: Investigate the process to determine the cause of the outliers.
- Missing factor: Determine whether you failed to account for a factor that influences the process.
- Random chance: Investigate the process and the outliers to determine if it occurred by chance, perform the analysis with and without the outliers to see its impact on the results.

Often, it is easier to identify outliers graphically. In model fitting procedures such as regression and ANOVA, outliers are points that are not explained well by the fitted model. These points are outlaying in the Y direction relative to the fitted regression line and have extreme residual values. The presence of outliers in the data causes difficulties when fitting a distribution to the data. Low and high outliers are both possible and have different effect on the analysis. Using the suspect value for the station considered, the suspicion may be dropped or accepted by confirm, using the comparison plot of the neighboring stations.

3.3.2.3. Filling missing observations

Hydrological data records may be missed because of different reasons, including extreme natural phenomena and human induced phenomena such as mishandling of the observed data by field personnel, wars etc. Therefore, in any hydrological data analysis, filling the missed observation is the foremost work. Filling the missed observation can be done through numerous methods. These methods includes, a classical method of filling the hydrological data, such as, Arithmetic method, normal ratio method, weighted distance interpolation method, time series analysis method and regression method (Hastie et al., 2001). For the case of this study missing data was filled in with regression equation between the neighboring stations.

3.4. Method of trend analysis

In order to be able to access the long term trends in stream flow and precipitation in the Weyib catchment as discussed earlier , the whole catchment were divided into six sub-catchments for which data were available, as, Agarfa sub catchment, Sinana sub catchment, Goba sub catchment, Dinsho sub catchment, Gasera sub catchment and Ginir sub catchment. So trend is analyzed for each sub catchment using the procedures, which would be discussed.

When attempting to detect trends in natural series, one must be aware of the inherent variability of hydrological time series (Burn D , 1994b). Askew (1987) indicate that there is a difficulty associated with differentiating natural variability and trend. This argues for the development of rigorous procedure for detecting trends. For each sub catchment, the steps followed are applied:-

- Selection of variables to be studied (precipitation and stream flow variable are used)
- Selection of stations that have sufficient long record of stream discharge and precipitation
- Data analysis and interpretation, which include checking for the presence of trend and determine the significance of the detected trends.

a) Selection of variables

Hydrologic variables are important indicators of climatic change. These variables tend to reflect climatic change and can help in understanding the relation between hydrology and climate. (Burn et al., 1992) suggested the study of a large number of hydrologic variables, since climate change is expected to affect various variables in different ways. In case of this study, stream flow and precipitation were taken as the hydrological variables of study for the trend analysis and different variables were selected for both run off and precipitation. These variables include the annual, monthly and seasonal mean flows and precipitations. Stream flow variables were given a great emphasis herein as they tend to reflect the integrated response of the catchment area as whole.

b) Selection of stations

Selection of station is one of the more important steps in climate change research. The work of selection follows the selection criteria adopted by (D.H.Burn and A. Mohammed, 2002) which says, stations were selected from the available stations with a minimum record length of 25 years to insure the statistical validity of the trend results.

c) Trend detection test

The time series of hydrologic variables were analyzed using the Mann-Kendall non-parametric test for trend. This test was found to be excellent tool for trend detection by other researchers in similar applications (Hirsch et al., 1982).The Mann-Kendall has two parameters that are important to trend detection. These parameters are the significance level that indicates the trends strength and the slope magnitude estimates that indicates the direction as well as the magnitude of the trend.

d) Significance of trend results

The results of trend test can be used to determine whether those observed collection of time series for hydrologic variable exhibits a number of trends that is greater than the number that is expected to occur by chance. However, to do this, it is necessary to consider the correlation structure of the data.

e) Serial correlation

The presence of serial correlation can complicate the identification of trend, in that a positive serial correlation can increase the expected number of false positive outcome for the Mann-Kendall test (Von Storch, 1995). Several approaches have been suggested to remove the serial correlation from the data set prior to applying the trend test. The two most common approaches are to pre-whiten the series and to prune the data set to form a sub set of observation that are sufficiently separated temporally to reduce the serial correlation. Method of pre-whitening the data was used in this study.

3.4.1. Models for trend analysis

Different statistical methods, both nonparametric tests and parametric tests, for identifying trend in time series are available in the literature. Two commonly used steps for identifying the trend are discussed in Figure 3.9.

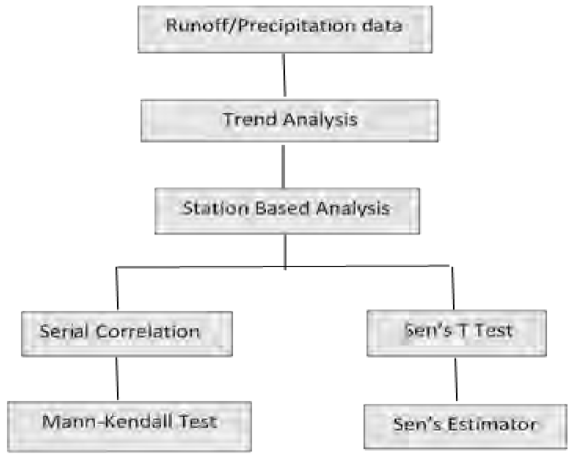


Figure3. 9: Trend Analysis Procedure used in this study

3.4.2. Mann-Kendall, Sen's method and Seasonal Kendall test

First, the presence of a monotonic increasing or decreasing trend is tested with the nonparametric Mann-Kendall test and secondly the slope of a linear trend is estimated with the nonparametric Sen's method.

Gilbert (1987) stated that, Mann-Kendall test is suitable for cases where the trend may be assumed monotonic and thus no seasonal or other cycle is present in the data. The Sen's method uses a linear model to estimate the slope of the trend and the variance of the residuals should be constant in time. These methods offer many advantages that have made them useful in analyzing hydrological data. Missing values are allowed and the data need not conform to any particular distribution. Besides, the Sen's method is not greatly affected by single data errors or outliers.

3.4.2.1. Mann-Kendall test

The Mann-Kendall test is applicable in cases when the data values x_i of a time series can be assumed to obey the model of the form:

$$X_i = f(t) + \varepsilon_i \qquad (3.2)$$

Where, f(t) is a continuous monotonic increasing or decreasing function of time and the residuals ε_i can be assumed to be from the same distribution with zero mean. It is therefore assumed that the variance of the distribution is constant in time. To test the null hypothesis (H_o) of no trend the observations x_i are randomly ordered in time, against the alternative hypothesis (H_A) where there is an increasing or decreasing monotonic trend.

The Mann-Kendall test statistic S is calculated using the formula

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}\left(x_j - x_k\right) \tag{3.3}$$

Where, x_j and x_k are the annual values in years j and k, j >k, respectively, and

$$\operatorname{sgn}\left(x_j - x_k\right) = \begin{cases} 1 & \text{if } x_j - x_k > 0 \\ 0 & \text{if } x_j - x_k = 0 \\ -1 & \text{if } x_j - x_k < 0 \end{cases} \tag{3.4}$$

At certain probability level H_0 is rejected in favor of H_A if the absolute value of S equals or exceeds a specified value $S\alpha/2$, where $S\alpha/2$ is the smallest S which has the probability less than $\alpha/2$ to appear in case of no trend. A positive (negative) value of S indicates an upward (downward) trend.

However, if there are several tied values (equal values) in the time series, it may reduce the validity of the normal approximation. First, the variance of S is computed by the following equation, which takes into account that ties may be present.

$$\mathrm{VAR}\,(S) = \frac{1}{8}\left[n(n-1)(2n+5) - \sum_{p=1}^{q} t_p(t_p - 1)(2t_p + 5)\right] \tag{3.5}$$

Here q is the number of tied groups and t_p is the number of data values in the p^{th} group.

The values of S and VAR(S) are used to compute the test statistic Z as follows

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{\mathrm{VAR}(S)}} & \text{if } S > 0 \\ 0 & \text{if } S = 0 \\ \dfrac{S+1}{\sqrt{\mathrm{VAR}(S)}} & \text{if } S < 0 \end{cases} \tag{3.6}$$

The presence of a statistically significant trend is evaluated using the Z value. A positive (negative) value of Z indicates an upward (downward) trend. The statistic Z has a normal distribution. To test for either an upward or a downward monotonic trend (a two-tailed test) at α level of significance, H_0 is rejected if the absolute value of Z is greater than Z_{1}-α/2, where, Z_{1}-α/2 is obtained from the standard normal cumulative distribution tables.

3.4.2.2. Sen's method

To estimate the true slope of an existing trend (as change per year) the Sen's non-parametric method is used. The Sen's method can be used in cases where the trend can be assumed linear. This means that f (t) is calculated as in equation (3.7).

$$f(t) = Qt + B \tag{3.7}$$

Where Q is the slope and B is a constant.

To get the slope estimate Q in equation (3.7) first calculate the slopes of all data value Pairs

$$Q_i = \frac{x_j - x_k}{j-k} \tag{3.8}$$

Where j>k

If there are n values x_j in the time series we get as many as N = n (n-1)/2 slope estimates Q_i. The Sen's estimator of slope is the median of these N values of Q_i. The N values of Q_i are ranked from the smallest to the largest and the Sen's estimator is

$$Q = Q_{[(N+1)/2]}, if\ N\ is\ odd$$

Or (3.9)

$$Q = \frac{1}{2}\left(Q_{\left[\frac{N}{2}\right]} + Q_{\left[\frac{N+1}{2}\right]}\right) if\ N\ is\ even$$

A 100(1-α) % two sided confidence interval about the slope estimate is obtained by the non-parametric technique based on the normal distribution. The method is valid for n as small as 10 unless there are many ties. MAKESENSE (2002) Excel Template Version 1.00 computes the confidence interval at two different confidence levels; α = 0.01 and α = 0.05, resulting in two different confidence intervals the procedure is as follow.

At first compute

$$C_{\alpha=Z_{1-\alpha/2}}\sqrt{\mathrm{VAR(s)}} \tag{3.10}$$

Where, VAR(S) has been defined in equation (3.6) and Z_1-α/2 is obtained from the standard normal distribution.

Next M_1 = (N - Cα)/2and M_2 = (N + Cα)/2 are computed. The lower and upper limits of the confidence interval, Q_{min} and Q_{max}, are the M_1 the largest and the (M_2 +1) the largest of the N ordered slope estimates Q_i. If M_1 is not a whole number, the lower limit is interpolated. Correspondingly, if M_2 is not a whole number the upper limit is interpolated. To obtain an estimate of B in equation (3.7) the n values of differences x_i – Qt_i are calculated. The median of these values gives an estimate of B (Sirois, 1998). The estimates for the constant B of lines of the 99% and 95% confidence intervals are calculated by a similar procedure.

3.4.2.3. Seasonal Kendall test

The trend test for annual stream flow series gives us an overall view of the change in stream flow volumes. To examine the possible changes occur in smaller time scale, we need to investigate the monthly or seasonal flow series. Monthly stream flow usually exhibit strong seasonality. Trend test techniques for dealing with seasonality of univariate time series fall into three major categories (Helsel and Hirsh, 1992).

(1) Fully nonparametric method (seasonal Kendall test)

(2) Mixed procedure (regression of deseasonalized series on time)

(3) Parametric method (regression of original series on time and seasonal terms)

The first approach, namely, seasonal Kendall test is used in this research considering the benefit that the seasonal Kendall test considers the effect of fluctuation in season in both runoff and precipitation. Hirsch et al (1982) introduced a modification of the MK test, referred to as the seasonal Kendall test that allows for seasonality in observations collected over time by computing the Mann-Kendall test on each of m seasons separately, and then combining the results.

Compute the following overall statistic S'.

$$s' = \sum_{j-1}^{m} S_j \tag{3.11}$$

Where S_j is simply the S-statistic in the MK test for season j (j = 1, 2... m).When no serial dependence exhibit in the time series, the variance of S is defined as

$$\mathrm{Var}(s') = \sum_{j-1}^{m} \mathrm{Var}(S'_j) \tag{3.12}$$

Then the quantity z' defined in the following equation is approximately standard normally distributed:

$$Z' = \begin{cases} \frac{S' - 1}{\sqrt{\mathrm{VAR}(S')}} & \text{if } S' > 0 \\ 0 & \text{if } S' = 0 \\ \frac{S' + 1}{\sqrt{\mathrm{VAR}(S')}} & \text{if } S' < 0 \end{cases} \tag{3.13}$$

The positive values of S indicate upward trends whereas negative S value indicate downward trend.

3.5. Method of Time series analysis

3.5.1. Autoregressive modeling

Autoregressive (AR) models are models in which the value of variables in one period is related to its values in previous periods.

The basic form of the AR model of order p with constant parameters

$$Z_t = \sum_{i-1}^{p} \phi_i Z_{t-i} + \varepsilon_i \tag{3.14}$$

Where, z_t is the time-dependent normal and standardized series {N (0, 1)}. ϕ_i are the autoregressive coefficients. ε_t is the time independent variable (white noise). p is the order of the autoregressive model. The above formulation is based on the standardized series, which can be obtained as follows:

$$Z_t = \frac{x_{t-\mu}}{\sigma} \tag{3.15}$$

Where, μ and σ are the mean and standard deviation of the series x_t. The parameter set of the model is

$$\{\mu\ \sigma\ \phi_1\ \ldots.\phi_p \sigma^2(\varepsilon)\}$$

Where σ^2 (ε) is the variance of the time independent series. In order to forecast or generate annual AR models, the following relation can be used:

$$Z_{t=}\ \phi_1\ Z_{t-1} + \cdots + \phi_{pZ_{t-p}} + \sigma(\varepsilon)\xi_t \tag{3.16}$$

Where, ζ_t is the standardized normal variable.

3.5.2. Moving average process

Moving average (MA) models account for the possibility of relationship between a variable and the residuals from previous periods.

The moving average model of order q (MA (q)) can be formulated as:

$$Z_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \theta_2\varepsilon_{t-2} - \theta_3\varepsilon_{t-3} - \cdots - \theta_q\varepsilon_{t-q} \tag{3.17}$$

It can also be written as:

$$Z_{t=} \sum_{j-0}^{q} \theta_j\varepsilon_{t-j} \qquad (\theta_0 = -1) \tag{3.18}$$

Where, $\theta_1 \ldots \theta_q$ are q orders of MA (q) model parameters. The parameter set of the model can be summarized as:

$$\{\mu\ \ \theta_1 \ldots.\theta_q\ \sigma^2(\varepsilon)\}$$

3.5.3. Autoregressive moving average (ARMA) modeling

The ARMA model of order (p, q) can be defined by combining an autoregressive model of order p and a moving average model of order q as follows:

$$Z_t - \phi_1 Z_{t-1} - \cdots - \phi_{pZ_{t-p}} = \varepsilon_t + \theta_1\varepsilon_{t-1} + \cdots + \theta_q\varepsilon_{t-q} \tag{3.19}$$

The ARMA (p, q) model can also be shown in the following compact form:

Where ϕ (z) and Θ (z) are the p^{th} and q^{th} degree polynomials:

$$\phi(Z) = 1 - \phi_1 Z - \cdots - \phi_p Z^p \quad (3.20)$$

$$\theta(Z) = 1 + \theta_1 Z + \cdots + \phi_q Z^q \quad (3.21)$$

The parameter set of the ARMA (p, q) model can be summarized as:

$$\{\mu \;\; \theta_1 \ldots, \phi_{1,} \ldots, \phi_p, \theta_q \, \sigma^2(\varepsilon)\}$$

The parameters of the ARMA (p, q) model should satisfy the conditions of stationarity. Once the ARMA model is fitted to a time series, the following procedures can be used for generating or forecasting the values of that time series. The synthetic generated values conserve the statistical properties of the historical data. In addition, ARMA (p, q) can be used to forecast Z values for lead-time L. If Z_t (L) denotes the value of Z at lead-time L, the following equations could be used for forecasting:

$$Z_t(L) = \phi_1 Z_{t+L-1} + \phi_2 Z_{t+L-2} + \cdots + \phi_p Z_{t+L-p} - \theta_1 \varepsilon_{t+L-1} - \cdots - \theta_q \varepsilon_{t+L-q} \text{ for } L \leq q \quad (3.22)$$

$$Z_t(L) = \phi_1 Z_{t+L-1} + \phi_2 Z_{t+L-2} + \cdots + \phi_p Z_{t+L-p} \text{ for } L > q \quad (3.23)$$

3.5.4. Autoregressive integrated moving average (ARIMA) modeling

The non-seasonal form of ARIMA models of order (p, d, q) can be formulated as:

$$\phi(B)(1 - B)^d Z_t = \theta(B)\varepsilon_t \quad (3.24)$$

Where (B) and ϕ (B) are polynomials of degree p and q, respectively:

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p \quad (3.25)$$

$$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q \quad (3.26)$$

The model has p + q + 1 parameters to be estimated

$$\{\phi_{1,} \ldots, \phi_{p,} \theta_1, \ldots, \theta_q, \theta_{q,} \;\; \sigma^2(\varepsilon)\}$$

In order to model the seasonal hydrologic time series, the seasonal form of ARIMA models of non-seasonal order (p, d, q) and of seasonal order (P, D, Q) w with seasonality w has been developed. The general multiplicative ARIMA (p, d, q) (P, D, Q) w can be formulated as:

$$(1-\Phi_1 B^w-\Phi_2 B^{2w}-\cdots-\Phi_p B^{pw})(1-\phi_1 B-\phi_2 B^2-\cdots-\phi_p B^p)(1-B^w)^D(1-B)^d Z_t \quad (3.27)$$

$$\left(1-\Theta_1 B^w-\Theta_2 B^{2w}-\cdots-\Theta_Q B^{QW}\right)\left(1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q\right)\varepsilon_t \quad (3.28)$$

Where, ε_t is an independently distributed random variable, B is the backward operator, B (z_t)=z_{t-1}, (1–Bw)D is the D[th] seasonal difference of season w, (1–B) d is the d[th] non-seasonal difference, p is the order of the non-seasonal autoregressive component, q is the order of the non-seasonal moving average component, P is the order of the seasonal autoregressive component, Q is the order of the seasonal moving average component, Φ is the seasonal autoregressive parameter, Θ is the seasonal moving average parameter, φ is the non-seasonal autoregressive parameter and θ is the non-seasonal moving average parameter.

The ARIMA models are non-stationary and cannot be used for synthetic generation of stationary time series but they are useful for forecasting. The previous ARIMA model equations can be used for data forecasting. Forecasting equations for ARIMA (1, 0, 1) (1, 0, 1)12 could be represented:

$$Z_t(L)=\phi_1 Z_{t-1+L}+\Phi_1 Z_{t-12+L}-\phi_1\Phi_1 Z_{t-13+L}-\theta_1\varepsilon_{t-1+L}-\Theta_1\varepsilon_{t-12+L}-\theta_1\Theta_1\varepsilon_{t-13+L}$$

$$for\ L\leq 12 \quad (3.29)$$

$$Z_t(L)=\phi_1 Z_{t-1+L}+\Phi_1 Z_{t-12+L}-\phi_1\Phi_1 Z_{t-13+L}\ for\ L>L \quad (3.30)$$

3.6. ARIMA time series modeling and Hydrological Justification of ARMA model

3.6.1. Hydrologic justification of ARMA model

A physical justification of ARMA models for annual stream flow simulation is as follows. Consider a watershed with annual precipitation X_t, infiltration aX_t and evapotranspiration bX_t. The surface runoff is (1-a-b) X_t= dX_t (Figure 3.10)

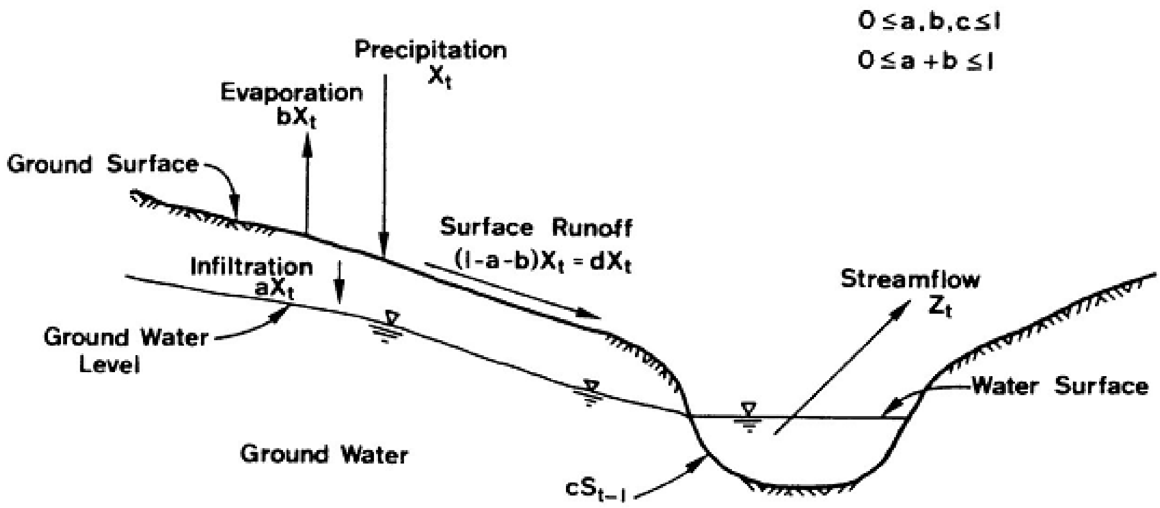


Figure3. 10: Conceptual representation of the precipitation-stream flow process after Salas and Smith (1980)

$Z_t = cS_{t-1} + dX_t$ 3.31)

The conservation of mass for the groundwater storage is:

$S_t = S_{t-1} + aX_t - cS_{t-1}$ (3.32)

Thus, $Zt = (1 - c) Z_{t-1} + dX_t - [d (1 - c) - ac] X_{t-1}$ (3.33)

Which has the form of an ARMA (1, 1) model when the precipitation, X_t is an independent series and when (1-c) = β_1, d = 1, and [d (1-c)-ac)] = θ_1

3.6.2. ARIMA time series modeling

Box and Jenkins (1970) provided hydrologists with an alternative model type the autoregressive integrated moving average (ARIMA) model and other forms of time series stochastic models. In order to understand how each section works within the modeling process, it is important to understand the conceptual framework of each step.

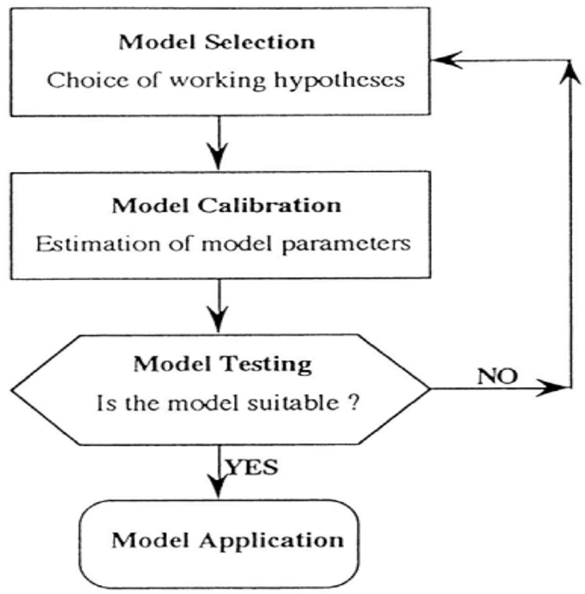


Figure3. 11: Phases of model analysis after Chong-yuXu (2002)

3.6.3. Model selection

ARIMA modeling differs from the other time series methods in fact that ARIMA modeling uses correlation techniques. This time series model is selected for this research work due to the fact that; this model is more versatile than AR model, Greater flexibility in fitting time series models and is achieved by including both autoregressive and moving average terms in the model. The combination of AR and MA models makes it possible to simulate many hydrologic processes by using a small number of parameters. The model identification step generally requires judgment from the analyst.

1. Decide if the data are stationary. That is, do the data possess constant mean and variance?

Examine a time series plot to see if a transformation is required to give constant mean and variance. Examine the ACF to see if large autocorrelations do not die out, indicating that differencing may be required to give a constant mean. A seasonal pattern that repeats every k^{th} time interval suggests taking the k^{th} difference to remove a portion of the pattern. Most series should not require more than two difference operations or orders. If spikes in the ACF die out rapidly, there is no need for further differencing.

2. Examine the ACF and PACF of stationary data in order to identify what autoregressive or moving average models, terms are suggested.

Table 3. 4: ACF and PACF plot properties and model type

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tail off	Cut off after lag q	Tail off
PACF	Cut off after lag p	Tail off	Tail off

An ACF with large spikes at initial lags that decay to zero or a PACF with a large spike at the first and possibly at the second lag indicates an autoregressive process. An ACF with a large spike at the first possibly at the second lag and a PACF with large spikes at initial lags that decay to zero indicates a moving average process. The ACF and the PACF both exhibiting large spikes that gradually die out indicates that both autoregressive and moving averages processes are present.

3. Use the ARIMA procedure

Fit the likely models, examine the significance of parameters, and select one model that gives the best fit. Check that the ACF and PACF of residuals indicate a random process, signified when there are no large spikes. If large spikes remain, consider changing the model. Perform several iterations in finding the best model. When satisfied with the fit, go ahead and make forecasts.

3.7. Model calibration

Whatever the model form is chosen, there are some unknown constants used to represent the physical process. These so-called parameters of the model must be assigned fixed numerical values before the model may be used to predict the runoff, in other words one need to estimate these parameters such that the best agreement between modeled and observed runoff can be obtained. The process by which the parameters are selected is called model calibration. The emphasis here is directed towards the calibration of conceptual hydrologic model of stream flow.

3.7.1. Manual calibration

In manual calibration, we use a trial and error process of parameter adjustment; after each parameter adjustment is made, the simulated and observed hydrographs are visually compared to see if the match is improved. The main weakness of manual calibration is that the absence of generally accepted objective measures of comparison makes it difficult to know when the process should be terminated. Different persons may obtain very different parameter values for the same watershed. For this study, manual calibration was used and the regression coefficient (R^2), were checked in accordance to (Green, 2004) recommendation (R^2 >0.5).

3.7.2. Automatic calibration

Automatic optimization procedures are mathematical search algorithms that seek to minimize differences between selected features of modeled and observed stream flows by systematic trial alterations in the values of the model parameters. These trial alterations are called iterations. Successful iterations are those, which cause a reduction in the value of the objective function (for direct search method).

3.7.3. Model Calibration data

It is generally agreed that proper choice of the calibration data can do much to reduce the difficulties encountered during calibration of a hydrologic model. From the viewpoint of model calibration, the quality of the data is dependent on the information (about the parameters) contained in the data and the noise (errors) in the data. Clearly, the information content to be as large as possible and the noise to be as small as possible. The best choice seems to be a data set that contains a lot of hydrologic variability and it is desirable that the data be carefully examined for various errors.

3.8. Model validation

Validation is comparison of the model outputs with an independent data set without making further adjustments. The process continues until simulation of validation period stream flows confirm that the model performs satisfactorily. In the validation process, data for a period of five years was used in all sub catchment of Weyib catchment to evaluate the model accuracy.

3.9. Model Evaluation

Another technique that can be used to select variables within a model or to choose among various types of models is known as model selection criterion. This work focus on two of the most representative and widely applied model selection criteria. AIC (Akaike Information Criteria) and BIC (Bayesian Information Criteria) are derived from distinct perspectives: AIC intends to minimize the Kull back-Leibler divergence between the true distribution and the estimate from a candidate model and BIC tries to select a model that maximizes the posterior model probability. Due to the rather different motivations, it is not surprising that they have different properties .Minitab does not contain a default means for calculating this statistic, so information can be taken from Minitab and plugged manually into a formula.

$$\mathrm{AIC} = \mathrm{N} \ln \hat{\delta}^2 + 2\mathrm{k} \tag{3.34}$$

$$\mathrm{BIC} = \mathrm{N} \ln \hat{\delta}^2 + \mathrm{k} \ln \mathrm{N} \tag{3.35}$$

$$\hat{\delta}^2 = \frac{\mathrm{SSE}}{\mathrm{N}} \tag{3.36}$$

Where N=sample size, k=number of variables in model and SSE=sum of squares error. When comparing multiple AIC and BIC values to determine which model statement is best, the lowest AIC and BIC value is selected as the best-fit model.

CHAPTE FOUR

RESULT AND DISCUSION

4.1. Filling missing observations

For this study missing data was filled in with regression equation between the neighboring stations and using time series modeling.

Regression result

There are a number of data missing in the given catchment gauging stations. Filling these missing data using method of regression between the neighboring stations was used to fill in the missed data. The neighboring stations of the Agarfa gauging station used in this regression analysis were Alemkerem gauging station and Denbel gauging station. When one of the this two neighboring gauging station and the station for which the missing data was to be filled were missing at same period, the regression with the second neighboring gauging station were used to fill in missing data.

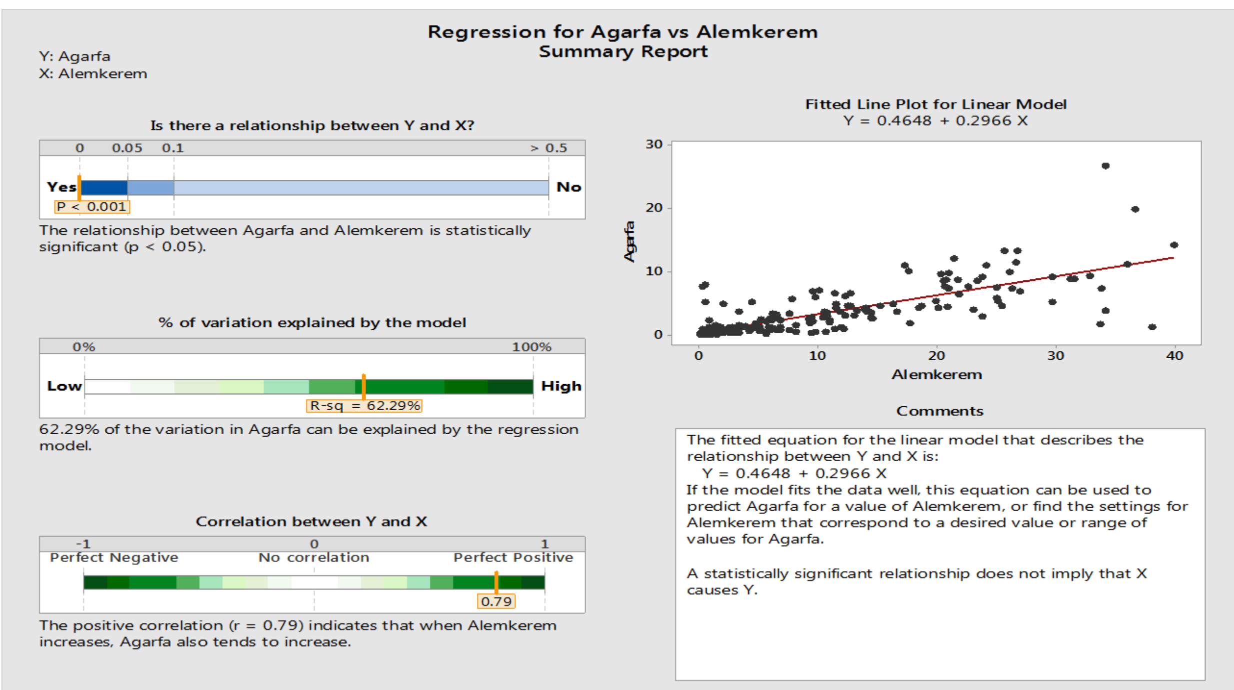


Figure 4. 1: Summery report plot of regression between Agarfa and Alemkerem.

The relationship between Agarfa and Alemkerem is statistically significant with ($p< 0.05$) Figure 4.1. This regression model can explain around 62.29% of variation in Agarfa. The positive correlation (r =0.79) indicates that when value of Almkerem increase, the value of Agarfa increases too.

Some of the missing data were filled by the regression equation shown Figure 4.1 while other was filled by the equation Figure 4.2.

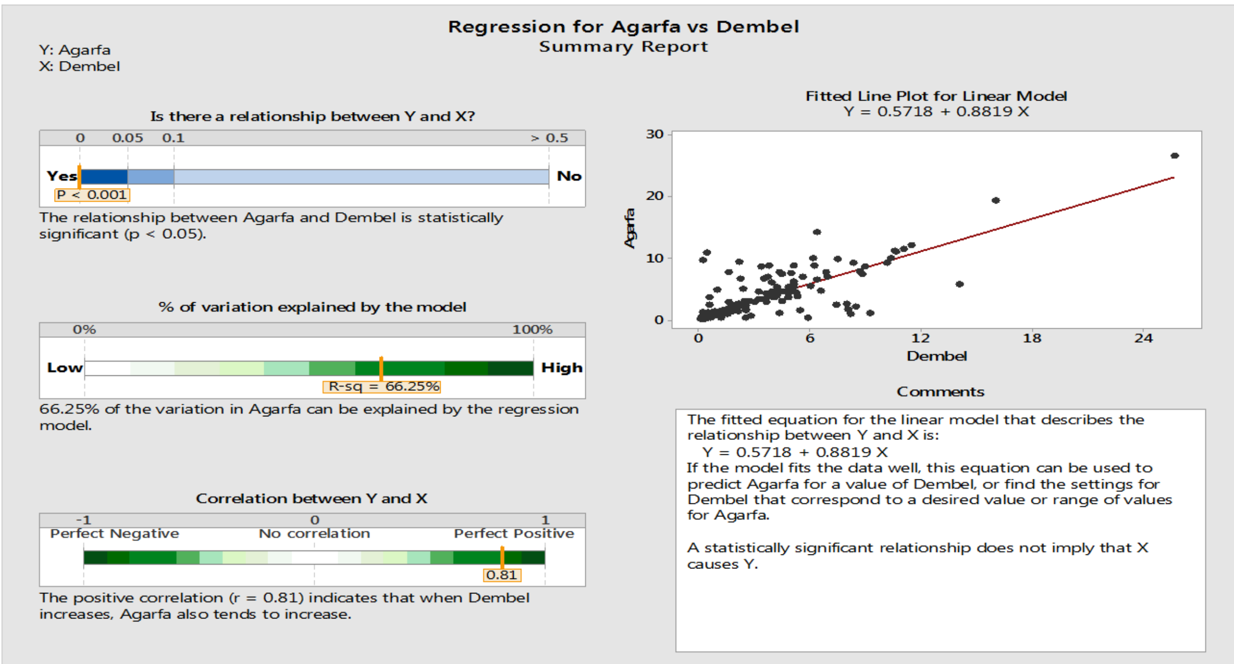


Figure 4. 2: Summery report plot of regression between Agarfa and Denbel

Similar procedures were followed for all other gauging stations in the catchment except for Ginir Precipitation station. For Ginir case, missing value was not filled by regression between its neighboring stations due to lack of good regression between all of the other neighboring stations. So, time series modeling was used for filling missing data in this gauging station.

The result of the regression equations with their model selection used for filling missing data of other gauging stations were given in the Appendix I. Generally, around 20 data value of Alemkerem, 31 data value of Agarfa, 23 data value of Denbel, 14 data value of Shaya, 25 data value of Tebel and 9 data value of Togona, were filled in using the method of regression.

4.2. Result for Consistency of data

The most common method of checking for inconsistency of record is Double Mass Curve analysis. The curve is a plot on arithmetic graph paper, of cumulative rainfall collected at a gauge where measurement condition may have changed significantly against the average of the cumulative rainfall for the same period of record collected at several gauges in the same region.

Table 4. 1: Summery for robe station double-mass analysis.

Year	Avg of (Sinana, Dinsho & Agarfa)station	Robe station	Cum of (Sinana, Dinsho & Agarfa) station	cum of Robe station	Robe/Avg of other station	Cum (Avg)-Bavg* cum(Robe)	Adjusted Robe station	Cum of Adjusted Robe station
1	1127	667	1127	667	0.6	527	667	667
2	1262	670	2389	1337	0.5	1186	670	1337
3	931	861	3320	2197	0.9	1342	861	2197
4	1242	875	4562	3073	0.7	1796	875	3073
5	1361	787	5922	3859	0.6	2449	787	3859
6	1144	910	7066	4770	0.8	2773	910	4770
7	1120	596	8186	5366	0.5	3357	596	5366
8	1424	935	9610	6301	0.7	3940	935	6301
9	1260	816	10871	7116	0.6	4466	816	7116
10	1243	882	12114	7999	0.7	4915	882	7999
11	1093	793	13207	8792	0.7	5294	793	8792
12	1021	901	14228	9693	0.9	5504	901	9693
13	1193	967	15421	10661	0.8	5826	967	10661
14	1430	869	16851	11530	0.6	6474	869	11530
15	1163	884	18014	12414	0.8	6842	884	12414
16	1036	785	19050	13199	0.8	7171	785	13199
17	1113	827	20163	14025	0.7	7540	827	14025
18	949	627	21112	14652	0.7	7925	401	14426
19	869	950	21981	15602	1.1	7939	608	15034
20	941	844	22922	16447	0.9	8120	540	15574
21	844	717	23767	17164	0.8	8319	459	16033
22	866	1123	24632	18287	1.3	8174	718	16752
23	1055	1140	25687	19427	1.1	8203	730	17482
24	543	842	26230	20269	1.6	7988	538	18020
25	838	1070	27067	21339	1.3	7863	685	18705
26	744	696	27811	22035	0.9	7980	445	19150
27	635	823	28446	22858	1.3	7874	527	19677
28	911	911	29357	23769	1	7965	583	20260
29	776	932	30133	24701	1.2	7902	596	20856
30	643	432	30776	25133	0.7	8156	277	21133
				B avg=	0.86			

B avg = average slope, cum = cumulative

Plot of cumulative differences, also called a residual-mass plot, the maximum and minimum values correspond to break points in the original double-mass line, making interpretation easier.

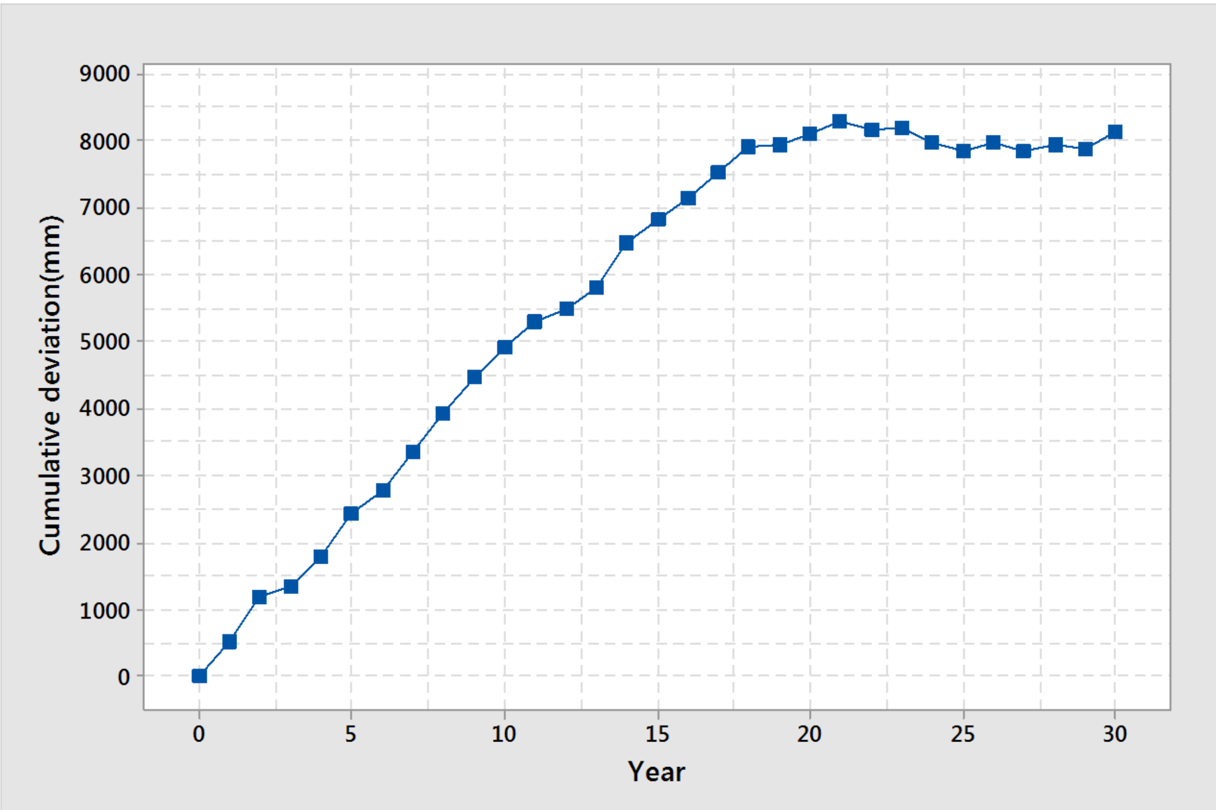


Figure 4. 3: Cumulative differences (residual-mass plot)

In Figure 4.3, possibly significant break point is at year 18. The average ratio or slope of the data in Table 4.1, from year 1 to year 18 and from year 19 to year 30 are S_1=0.69 and S_2=1.08 respectively. So the ratio of S_1/S_2 (r=0.64) is used to adjust the broken line through the origin.

The data series, which is inconsistent, adjusted to consistent values by proportionality. Double mass curve plot was made for all stations. Figure 4.4 shows only for Robe observatory. From the double mass curve plot the stations are consistent each other after correction.

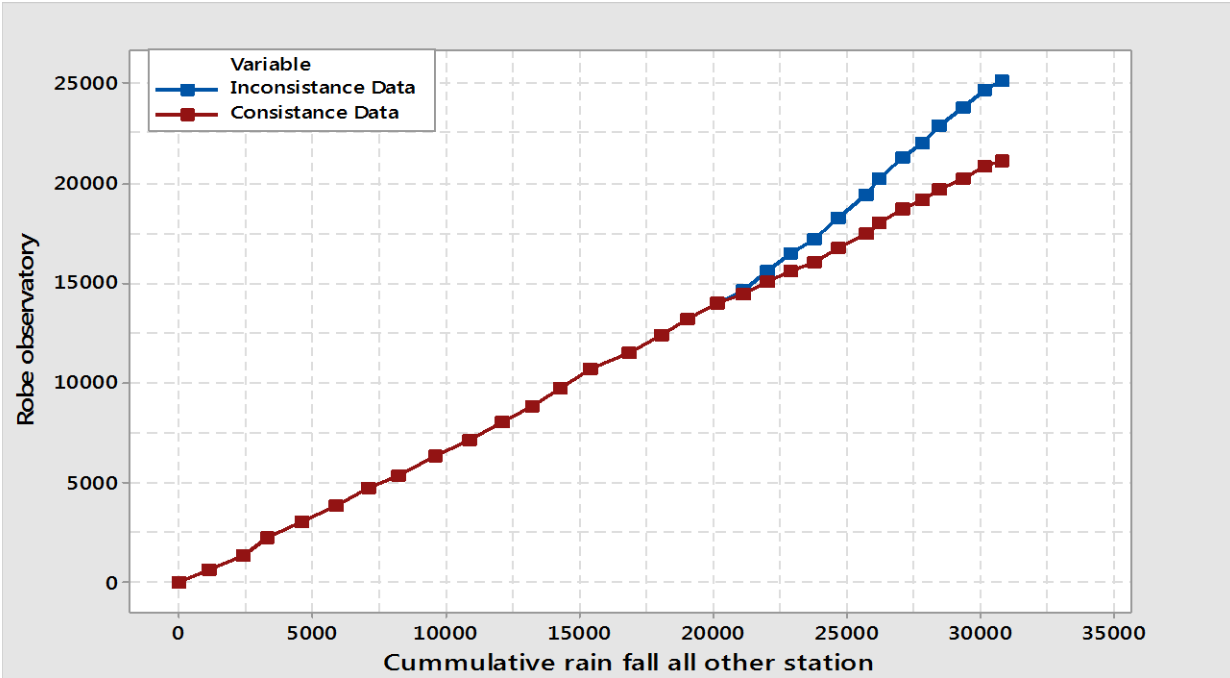


Figure 4. 4: Double mass curve for Robe observatory Meteorological stations

4.3. Test for outliers

An outlier is an observation that deviates significantly from the bulk of the data, which may be due to errors in data collection, or recording, or due to natural causes. Outliers should have to be investigated because they can provide useful information about your data or process. Unless the outliers are detected and corrected, they may result in unreliable result in both trend test and time series modeling case.

4.3.1. Outliers test result

First visual screening showed a few mistyped decimal points in the gauging stations of these sub catchments. These were corrected and flagged in the data files. Then data were subjected to a few validations for upper and lower boundary limit to identify outliers. Outliers were detected using box plot by using Minitab 17 statistical software for all sub catchments. Herein, only the result of Alemkerem gauging station in the Ginir sub catchment were given in Figure 4.5 while result for the other gauging stations were given in the Appendix II.

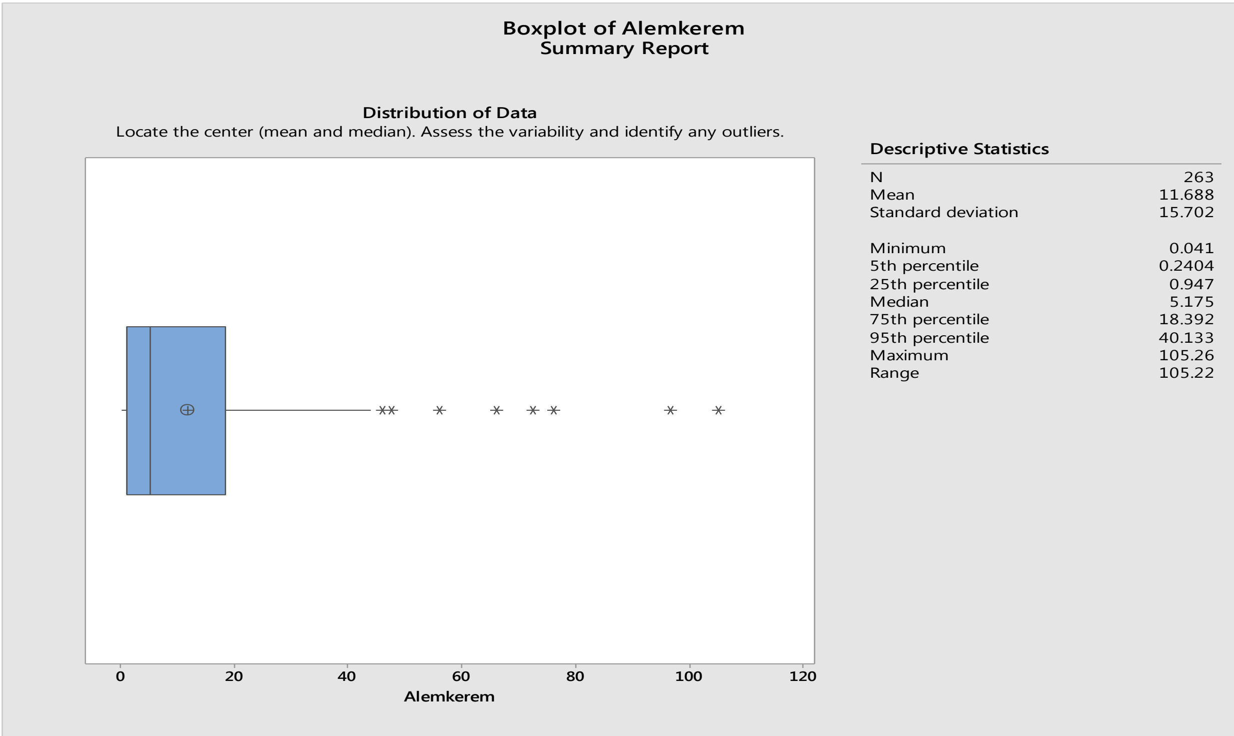


Figure 4. 5: Summery report of box plot of Alemkerem gauging Station

As indicated in Figure 4.5 the data at Alemkerem gauging station was right skewed and the symbol in the form of star shows the outliers. It was found that around eight data were detected as outliers.

Table 4. 2: Suspected outliers and time of occurrence for Alemkerem gauging station

Year	**1988**	**1988**	**1989**	**1994**	**1997**	**1998**	**1998**	**2001**
Month	August	October	August	August	September	July	October	Jun
Value (m^3/sec)	56.03	47.56	66.03	105.26	72.53	45.95	96.67	76.11

Sometimes, errors occur from mistyping one number and/or misplacing the decimal point and were checked if there is any. However, the data values given in the Table 4.2 were violated the data limit and they were flagged and waited to be confirmed by the next methods, which is comparison of plot between adjoining stations. Those methods are multi-station validation and relation curve (regression analysis).

4.3.2. Multi-station validation

The value, which is set as suspicious in one of the gauging station at some period should not have to be regard as correct, rather there should have to be a confirmation, or validation for that suspected value comparing to the others neighboring value at the same time of record. Multi station validation should have to be done for stations, which their hydrological catchment characteristics seem to have similar characteristics in the neighboring station.

Comparison of series may also permit the acceptance of the value flagged as suspect in screening of data. When two or more station show the same behavior there is strong evidence to suggest that the value are correct. However, if the behavior of the neighboring station appears to be different, this implies that, the suspicion of the value is correct and further consideration is necessary to deal with it.

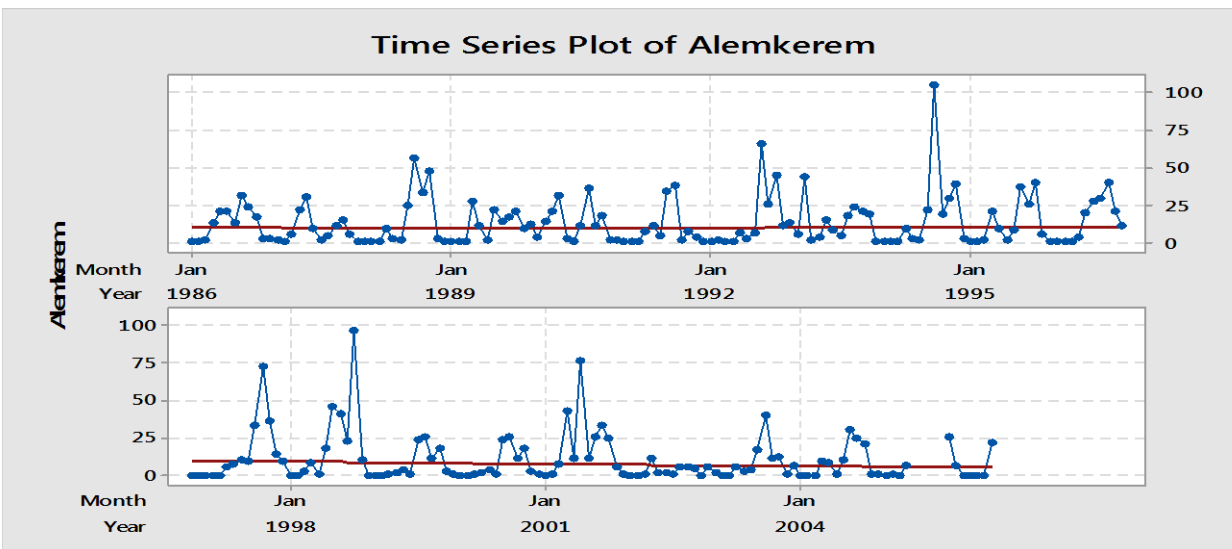


Figure 4. 6: Outliers test time series plot of Alemkerem

After this comparison was made between each other and with the others gauging stations in the catchment, it is confirmed that some of data were not outliers rather they appeared so due to the fact that months in which these data occurred are the high flow months. However, the data of September 1997, which is 72.53m^3/sec, and data of June 2001, which is 76.11m3/s was clearly outliers. This value can be confirmed from time series plot of the gauging station (Figure 4.5). Similarly, the outliers' test of the other gauging stations were done for all stations and corrected as needed.

4.3.3. Relation curve (Regression analysis)

The relation curve between neighboring stations, which developed also, strengthens the flagged value by looking the relation curve of two sequential station data values. Then the suspect values previously identified should be removed before deriving the relationship, which may then be applied to compute corrected values to replace the suspect ones. The validation was displayed river by river for all sub catchments. Each sub catchment station was compared with other neighboring stations in the sub catchment.

4.4. Result for Trend analysis

The first step in time series analysis is visually inspecting the data. Significant changes in level or slope usually are obvious. From the visual inspection, it seems that the annual flow series and monthly mean flow series of all the sub catchment showed no significant change over the period under consideration.

Significance of serial correlation

Prior to conducting the Mann-Kendall test, the Trend Free Pre-Whitening procedure was applied to the series. Clearly, using an I-MR chart would be helpful to know if the data are auto correlated. Fortunately, the Assistant in Minitab Statistical Software will check this for us without even needing to use sophisticated options deep down in software. Then the lag-1 serial correlation test was applied to all the time series data in which there is a significant serial correlation.

All most all monthly, annual and seasonal time series data in weyib river catchment shows no significant serial correlation coefficient at 5% significant level. This indicates that the data series are not violating the assumption of independence. The Ginir sub catchment stations (Tebel) showed a significant serial correlation in annual series. So in order to avoid the correlation lag-1 serial correlation test was applied.

4.4.1. Runoff trend results

4.4.1.1. Annual Runoff trend results

The annual stream flow trends for different stations are represented in Table 4.3. The positive vale of Z and Q Statistics showed that there is rising trend with upward slope of stream flow on annual basis. Two of the six stations (Tebel and Togona) in annual series showed positive trends at >99.9% statistical significance with test statistics (Z=5.18, and Z=4.32) and respective sen's slope of (S=0.007 and S=0.096) respectively. The Sen's estimator summarizes the results of change per unit time of the trend detected.

Table 4. 3: Annual runoff trend analysis

Station	Time series			Mann-Kendall trend			Sen's slope
	First year	Last Year	N	Test Z	P-value	Significant level	Q
Agarfa	1981	2008	28	0.97	0.166	0.05	0.037
Alemkerem	1984	2006	23	-0.34	0.365	0.05	-0.029
Denbel	1986	2006	21	-0.63	0.262	0.05	-0.035
Shaya	1981	2007	27	-0.10	0.458	0.05	-0.005
Tebel	1983	2007	25	5.18	< 0.001	0.001	0.007
Togona	1983	2007	25	4.32	< 0.001	0.001	0.096

4.4.1.2. High flow trend results

The positive vale of Z and Q Statistics showed that there is rising trend with upward slope of stream flow on high flow season basis. Two of the six stations (Tebel and Togona) in high flow season series showed positive trends at >99.9% statistical significance with test statistics (Z=4.13, and Z=3.71) and respective sen's slope of (S=0.006 and S=0.142) respectively. From those stations, no station showed significant negative trend.

Table 4. 4: Runoff trend analysis for high flow season

Station	Time series			Mann-Kendall trend			Sen's slope
	First year	Last Year	N	Test Z	P-value	Significant level	Q
Agarfa	1981	2008	28	1.60	0.054	0.05	0.194
Alemkerem	1984	2006	23	0.77	0.221	0.05	0.165
Denbel	1986	2006	21	0.00	0.5	0.05	0.002
Shaya	1981	2007	27	1.02	0.153	0.05	0.070
Tebel	1983	2007	25	4.13	< 0.001	0.001	0.006
Togona	1983	2007	25	3.71	< 0.001	0.001	0.142

4.4.1.3. Low flow season trend results

The Mann-Kendall trend and its statistical significance along with magnitude of Sen's slope for the all station are shown in Table 4.5. Accordingly, Tebel and Togona stations showed positive trend in low flow season at >99.9% statistical significance with test statistics (Z=4.18 for Tebel and Z=3.22 for Togona) with respective sen's slope of (S=0.005 and 0.03) respectively. Which means that there is 0.1% probability that we make a mistake when rejecting Ho. Shaya station data having (Z=1.73 & S=0.053) showed increasing significant trend with 95 % significance interval. Agarfa station with test statistic (Z=2.75) and trend magnitude(S=0.049) showed positive trend at 99% statistical significance.

Table 4. 5: Runoff trend analysis for low flow season

Station	Time series			Mann-Kendall trend			Sen's slope
	First year	Last Year	N	Test Z	P-value	Significant level	Q
Agarfa	1981	2008	28	2.75	0.003	0.01	0.049
Alemkerem	1984	2006	23	0.45	0.326	0.05	0.017
Denbel	1986	2006	21	0.69	0.243	0.05	0.031
Shaya	1981	2007	27	1.73	0.041	0.05	0.053
Tebel	1983	2007	25	4.18	<0.001	0.001	0.005
Togona	1983	2007	25	3.22	< 0.001	0.001	0.030

4.4.1.4. Monthly runoff trend analysis

The Mann-Kendall (MK) Test for monotonic analysis of trend together with nonparametric Sen's Slope Estimator was used to estimate the magnitude of trend for time series data. Accordingly, month January and February for Shaya station showed significant increase in trend with z statistics (Z=2.315 and 2.189) and sen's slope of (S=0.019 and 0.011) respectively. Meanwhile, all of the monthly records of the Tebel gauging station showed significant increasing trend. The increasing trend are shown in month of January (Z=2.648), July (Z=2.391) and November (Z=2.49) with respective slope of (S=0.019, 0.229 and 0.062) respectively for Agarfa station. Similarly Denbel station showed significant increase in trend in months of January and December (Z=2.989 and Z=2.144) with their respective slope of (S=0.023 and 0.033) respectively. Almost all station of Togona station showed increase in trend except in the months of April, May and Jun. Alemkerem station is the only station showed no significant trend in all moths of the years. This indicates that the flow remain the same in this sub catchment gauging stations (no additional flow added to the catchment). A positive (negative) value of Z indicates an upward (downward) trend. Generally, the trend values for all months of this sub catchment were shown on the Appendix IV.

4.4.2. Precipitation trend results

4.4.2.1. Annual Precipitation trend results

The trend is said to be decreasing if Z is negative and the computed probability is greater than the level of significance, and increasing if Z is positive and the computed probability is greater than the level of significance. If the computed probability is less than the level of significance, there is no trend**.** During the 1989–2014 periods, annual precipitation showed significant increasing trends for Ginir stations in Ginir sub catchment at >99% statistical significance with test statistic (Z=2.64) and respective sen's of (S=17.889)respectively. Dinsho station data having (Z= -1.67 & S= -9.85) showed significant decreasing trend at 0.05 significance level having 95% significance interval of Q.

Table 4. 6: Annual Precipitation trend analysis

Station	Time series			Mann-Kendall trend			Sen's slope
	First year	Last Year	N	Test Z	P-value	Significant level	Q
Agarfa	1984	2014	31	-1.16	0.124	0.05	-6.711
Robe	1985	2014	30	0.79	0.216	0.05	2.030
Sinana	1984	2014	31	-0.85	0.198	0.05	-1.631
Dinsho	1984	2014	29	-1.67	0.048	0.05	-9.851
Ginir	1989	2014	26	2.64	0.004	0.01	17.889
Gasera	1984	2014	30	-1.03	0.15	0.05	-3.678

4.4.2.2. Wet season Precipitation trend results

Table 4.7 is an evident that, most stations in the weyib catchment in wet season show no significant trend. However, Ginir station shows a very pronounced positive trend at >99% statistical significance with z statistics (Z=2.8) and respective sen's slope of (S=1.769) respectively.

Table 4. 7: Precipitation trend analysis for wet season

Station	Time series			Mann-Kendall trend			Sen's slope
	First year	Last Year	N	Test Z	P-value	Significant level	Q
Agarfa	1984	2014	31	-0.95	0.171	0.05	-0.750
Robe	1985	2014	30	1.30	0.096	0.05	0.775
Sinana	1984	2014	31	-0.65	0.259	0.05	-0.344
Dinsho	1984	2014	29	-0.44	0.329	0.05	-0.544
Ginir	1989	2014	26	2.80	0.003	0.01	1.769
Gasera	1984	2014	30	-1.07	0.142	0.05	-0.617

4.4.2.3. Dry season Precipitation trend results

Dry season consist of November to February months and this seasons showed no significant trend for all stations at (>95% to 99.9%) statistical significance. Which means that precipitation remain the same for dry season of the catchment.

Table 4. 8: Precipitation trend analysis for dry season

Station	Time series			Mann-Kendall trend			Sen's slope
	First year	Last Year	N	Test Z	P-value	Significant level	Q
Agarfa	1984	2014	31	0.24	0.406	0.05	0.094
Robe	1985	2014	30	0.64	0.26	0.05	0.147
Sinana	1984	2014	31	0.03	0.486	0.05	0.012
Dinsho	1984	2014	29	0.00	0.5	0.05	-0.011
Ginir	1989	2014	26	0.44	0.33	0.05	0.323
Gasera	1984	2014	30	0.39	0.347	0.05	0.151

4.4.2.4. Monthly Precipitation Trend Analysis

The variation in Precipitation data (trend) on monthly basis is calculated individually for each station using Mann-Kendall statistical method and magnitude of slope is calculated with Sen's slope estimator. It was analyzed that there is significant changes in monthly rainfall data. Some of the months showed increasing (upward) trend and some showed decreasing trends. The monthly Mann- Kendall trend value of Sinana station (September), Dinsho station (March) and Ginir station (February) shows significant decreasing monotonic trend with (Z=-2.006, -1.972 and -2.074) and(S=-2.055, -2.32 and -0.543) respectively. Similarly Dinsho station (November), Robe station (July) and Two month in Ginir station (May and September) shows significant increasing monotonic trend with (Z=1.989, 1.998, 2.315 and 2.074) with respective slope of (S=1.422, 1.997, 4.181 and 2.185) respectively.

As represented in Appendix V those months which have single star (*) sign shows significance level 0.05 having 95 % significance interval of Q and there is 5 % probability that we make a mistake when rejecting Ho. The month of Jun (Dinsho station), month of April (Robe station), month of February (Sinana station), months February and April(Agarfa station) and month of

September (Gasser and Ginir station) data having plus (+) sign showed 0.1 significance level having 90 % significance interval of Q and there is 10 % probability that we make a mistake when rejecting Ho. While the other months with blank sign showed significance level greater than 0.1.

Generally those time series data give quite significant trends whose both Mann-Kendall trend (Z statistics) and Sen's Slope magnitude (Q statistics) having either increasing (positive)or decreasing (negative) values.

4.5. Hydrological Variability

4.5.1. Rainfall

4.5.1.1. Monthly Rain fall statistics

In the analysis, for each month in a year, spatial rainfall statistics was computed using all gauges with available data for the period. As a result, spatially averaged monthly rainfall over the Weyib River catchment ranges from a low of 20.20mm in January to a high of 155.6 mm in April. Spati

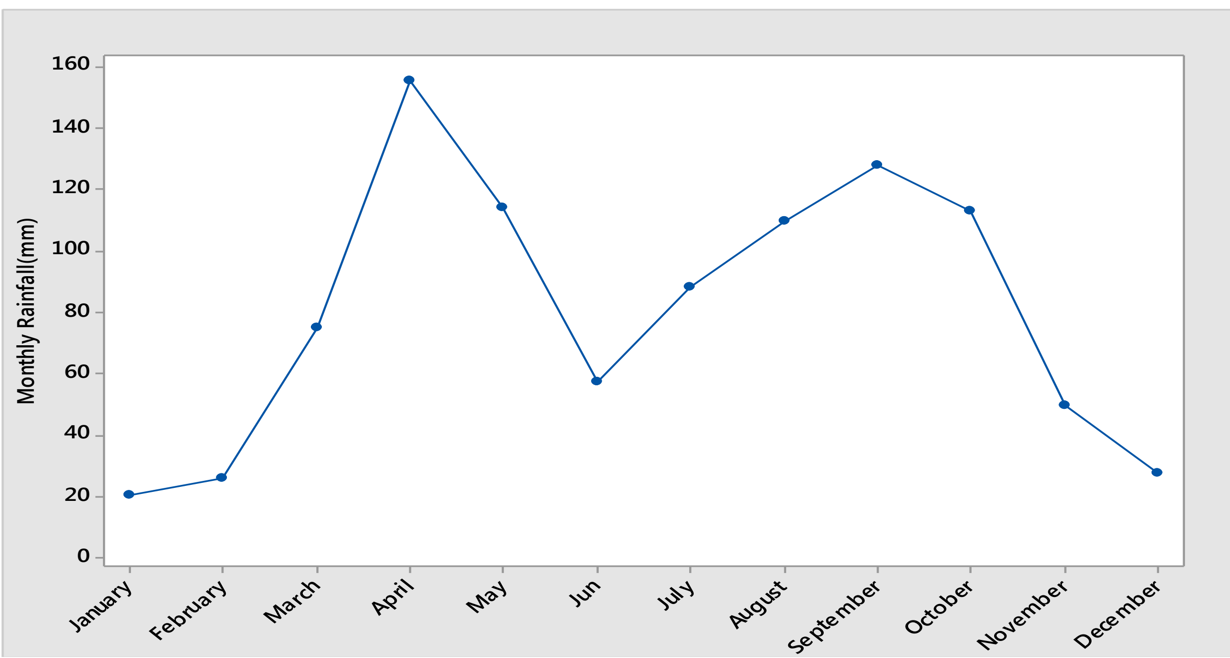


Figure 4. 7: Monthly spatial average rainfall, with temporal mean

The high wet season months are July, August and September, small wet season months are March, April, and May and dry season months are November, December, and January. June is the transition month between two wet season and October and February are the transition months between the wet season and the dry season.

The standard deviation (σ) is relatively high for the Wet season and transition months. The coefficient of variation, c.v and (σ/μ) of the wet season except for July and August months is small indicating that these months have low variation Table 4.9. April, July, October, and November have relatively higher positive skewness. The dry season months have higher coefficient of variation indicating that the year-to-year variation for these months is high.

Table 4. 9: Weyib River catchment monthly rainfall statistics

Variable	Mean	StDev	CoefVar	Skewness	Kurtosis
Jan	20.20	3.85	19.05	-0.03	-0.77
Feb	26.05	10.47	40.20	-0.04	-1.31
Mar	74.94	6.48	8.64	-1.05	1.31
Apr	155.6	30.1	19.38	0.87	0.40
May	114.36	18.93	16.55	-1.36	2.72
Jun	57.12	24.28	42.51	0.02	-1.31
July	88.0	57.1	64.88	0.64	0.12
Aug	109.6	50.6	46.23	-0.02	-1.34
Sept	127.98	18.70	14.61	0.06	-1.09
oct	112.8	32.5	28.85	1.62	2.97
Nov	49.74	17.02	34.23	1.61	2.12
Dec	27.37	10.20	37.28	0.28	-1.58

4.5.1.2. Monthly Rainfall Probability Distribution

Frequency of rainfall depths and duration can be determined from point, single gauge measurements or from spatially averaged rainfall data. In this case, the distribution of monthly rainfall time series averaged over the basin is analyzed. The density function characterizes a continuous probability of monthly rainfall of a given value and less than the given value, such that the area under the curve bounded by the horizontal axis is equal to 1.Theoretical probability distribution fittings test resulted in January, February, March, Jun, and August catchment rainfall fit the normal probability distribution. April, September, and December fit lognormal distribution. May, July, October and November fit Weibul distribution.

Each model was fitted to each month rainfall and the best-fit model was selected using Minitab17 .This handy tool allows you to compare how well your data fit 16 different distributions. Before walking through the output, there are three measures need to be known.

Anderson-Darling statistic (AD): Lower AD values indicate a better fit. However, to compare how well different distributions fit the data, you should assess the p-value.

P-value: It is generally valid to compare p-values between distributions and go with the highest. A low p-value (e.g., < 0.05) indicates that the data do not follow that distribution. For some 3-parameter distributions, the p-value is impossible to calculate and is represented by asterisks.

LRT P: For 3-parameter distributions only, a low value indicates that adding the third parameter is a significant improvement over the 2-Parameter version. A higher value suggests that you may want to stick with the 2-Parameter version.

Goodness of Fit Test for Month of January

Goodness of Fit Test

Distribution	AD	P	LRT P
Normal	**0.217**	**0.743**	
Box-Cox Transformation	0.272	0.547	
Lognormal	0.226	0.711	
3-Parameter Lognormal	0.250	*	0.697
Exponential	2.206	0.004	
2-Parameter Exponential	0.634	0.128	0.000
Weibul	0.267	>0.250	
3-Parameter Weibul	0.260	>0.500	0.589
Smallest Extreme Value	0.311	>0.250	
Largest Extreme Value	0.271	>0.250	
Gamma	0.249	>0.250	
3-Parameter Gamma	0.858	*	1.000
Logistic	0.241	>0.250	
Log logistic	0.242	>0.250	
3-Parameter Log logistic	0.241	*	0.764

ML Estimates of Distribution Parameters

Distribution	Location	Shape	Scale	Threshold
Normal*	**20.20191**		**3.84895**	
Box-Cox Transformation*	0.05115		0.01025	
Lognormal*	2.98968		0.19572	
3-Parameter Lognormal	8.92652		0.00047	-7508.79046
Exponential			20.20191	
2-Parameter Exponential			6.54526	13.65660
Weibul		6.46902	21.70235	

3-Parameter Weibul		2.54358	9.34529	11.92059
Smallest Extreme Value	21.97430		3.22232	
Largest Extreme Value	18.42270		3.28116	
Gamma		31.22632	0.64695	
3-Parameter Gamma		1859.33765	0.08267	-135.03287
Logistic	20.19533		2.12822	
Log logistic	2.99694		0.10736	
3-Parameter Log logistic	8.92572		0.00028	-7502.80576

* Scale: Adjusted ML estimate

A good place to start is to skim through the p-values and look for the highest. The highest p-value is for normal distribution. Similar procedures were followed for all Months to select best-fit Probability distribution. Figures 4.8, 4.9, and 4.10 depict density function for all 12 months.

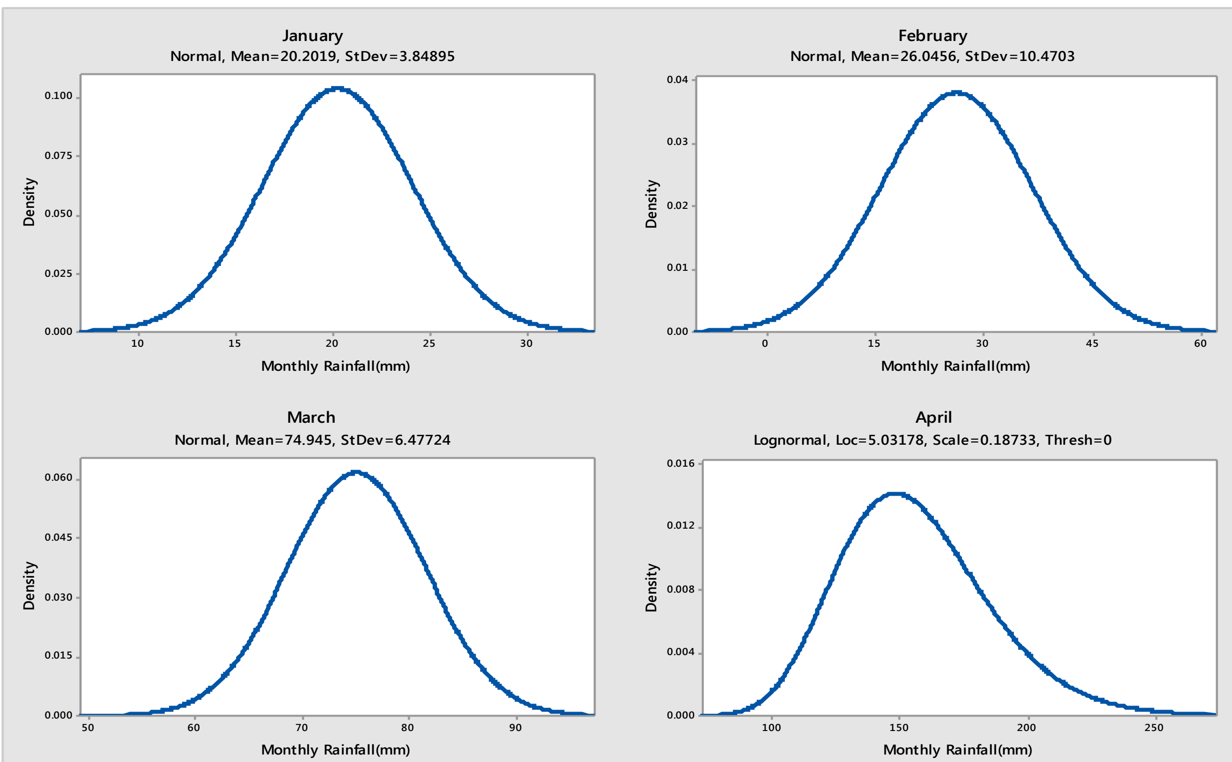


Figure 4. 8: Probability density functions for January, February, March, and April

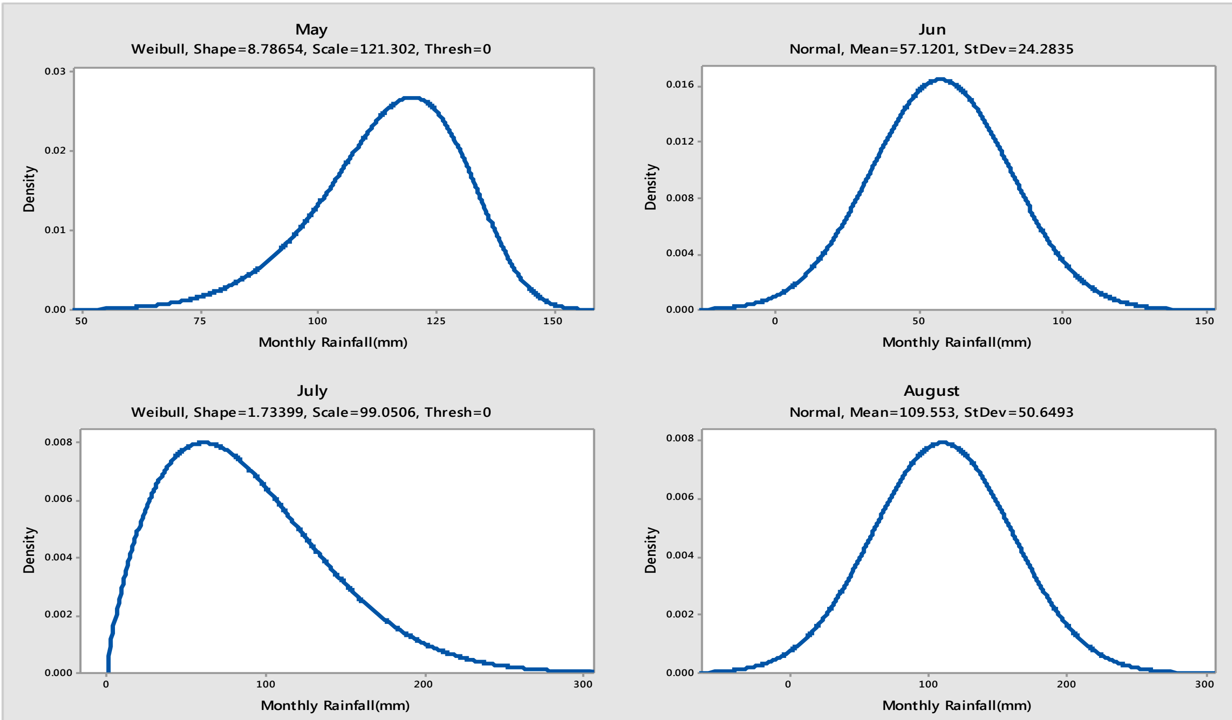


Figure 4. 9: Probability density functions for May, Jun, July, and August

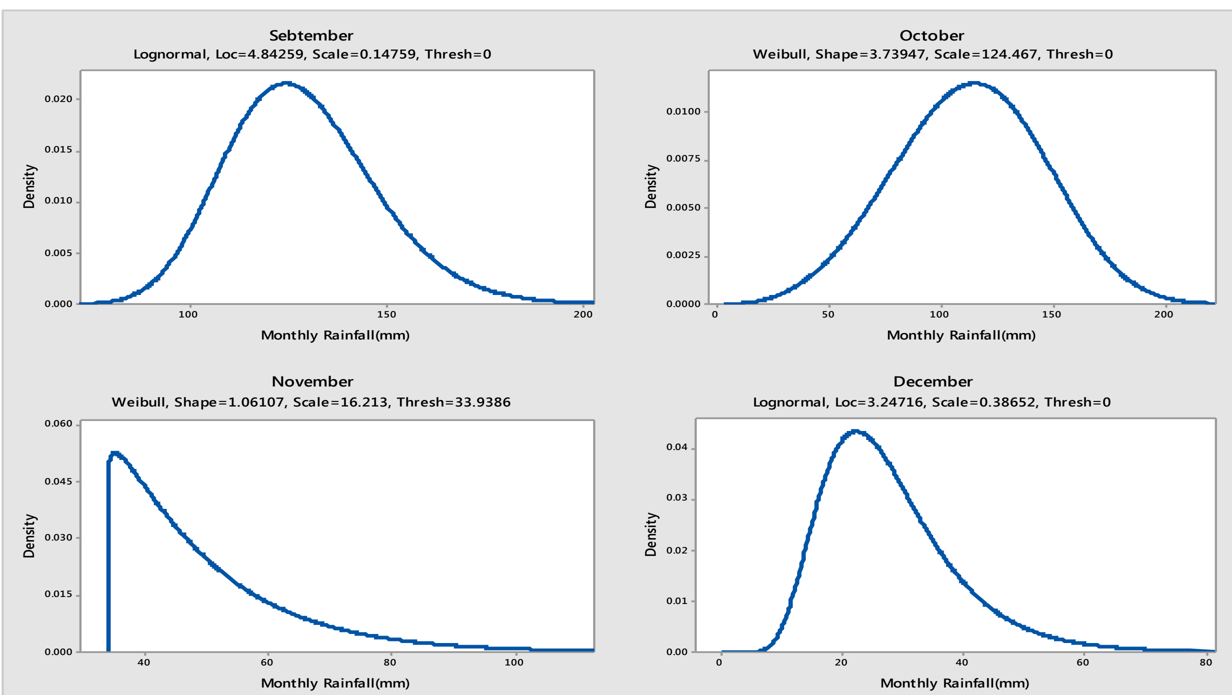


Figure 4. 10: Probability density functions for September, October, November, and December

4.5.1.3. Monthly Rainfall Return Periods

Probabilistic measure of the likelihood of occurrence of a given depth of rainfall is measured in return periods. Return periods are the expected interval of years a given depth of rainfall is expected to occur. Return periods in years were computed from the Cumulative Density Function (CDF) of each month's respective distribution fitting. Return periods for dry (below average) and wet (above average) rainfall patterns were computed for return periods of 2, 5, 10, 25, 50 and 100-year.

Dry Return Period = 1/CDF for CDF $\leq$ 0.5

Wet Return Period = 1/1 – CDF for CDF $\geq$ 0.5

Return period for dry and wet pattern rainfall for each month is depicted in Fig. 4.11. The six months(December to February and Jun to August) rainfall for dry and wet return periods show that for drought return periods of 100 years and higher, all the 6 months experience small rainfall. However, for wet return periods, distinct groups emerge with December, January, and February producing lower rainfall amount than Jun, July, and August.

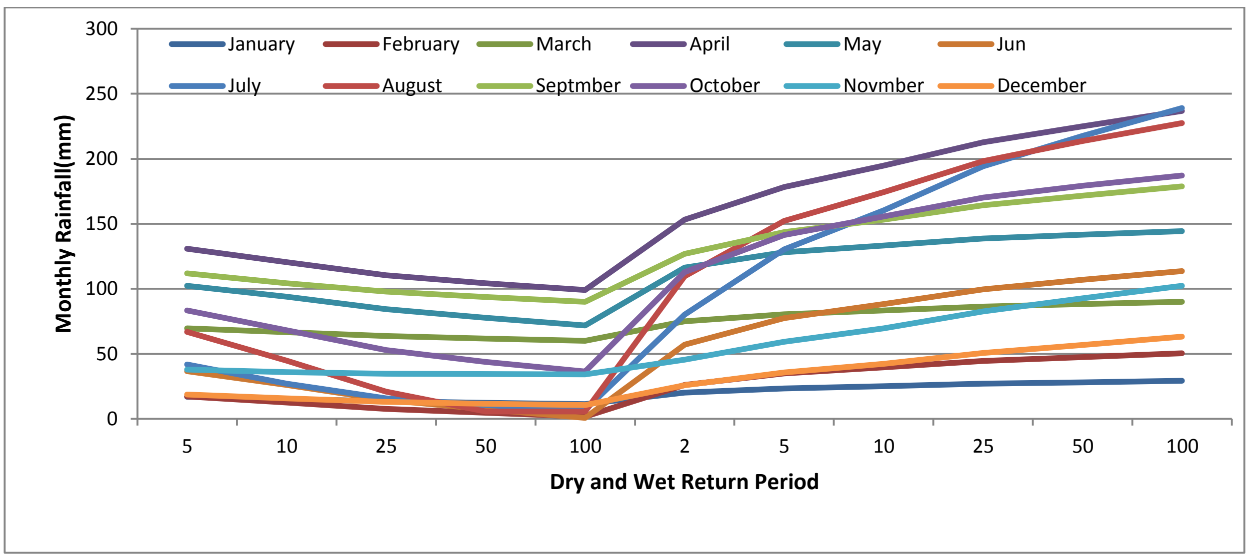


Figure 4. 11: Dry and wet return period rainfall for each month

4.5.1.4. Annual Rainfall Statistics and Probability Distribution

With mean annual rainfall of 1002.07 mm, the Weyib catchment is relatively wet. Figure 4.12 depicts catchment annual rainfall (1985–2014), mean annual rainfall, and ±standard deviation

(108.2 mm) catchment annual rainfall. The Weyib River catchment annual rainfall has a lognormal distribution. The coefficient of the temporal variation of annual rainfall is 0.108, which is relatively small. The rainfall statistics is based on five rainfall stations with the same length of record.

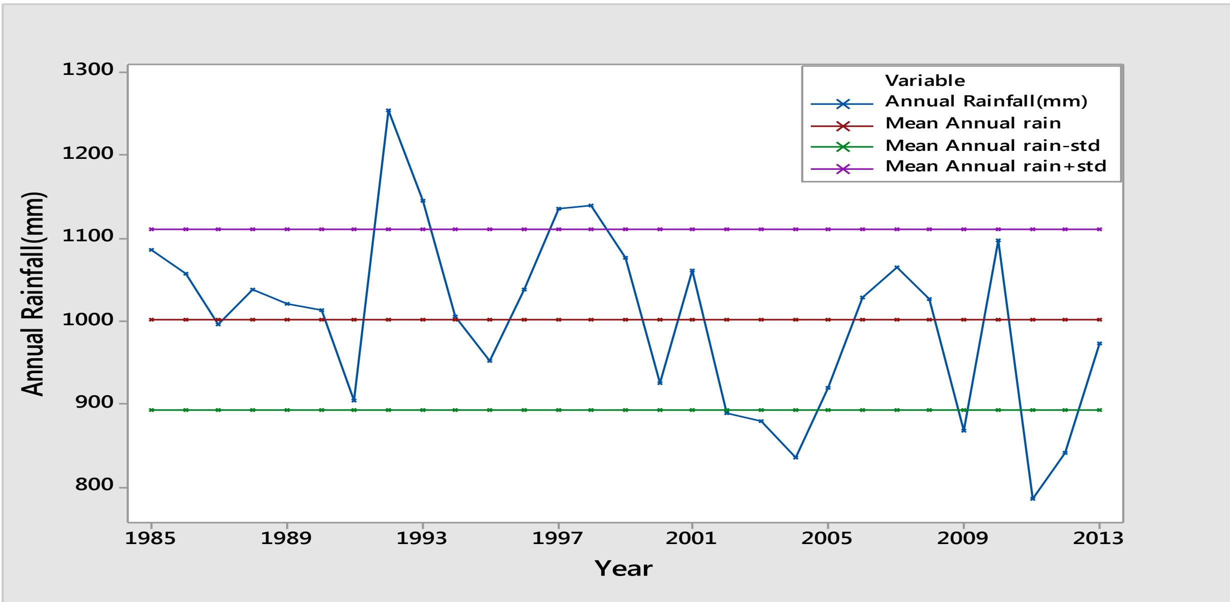
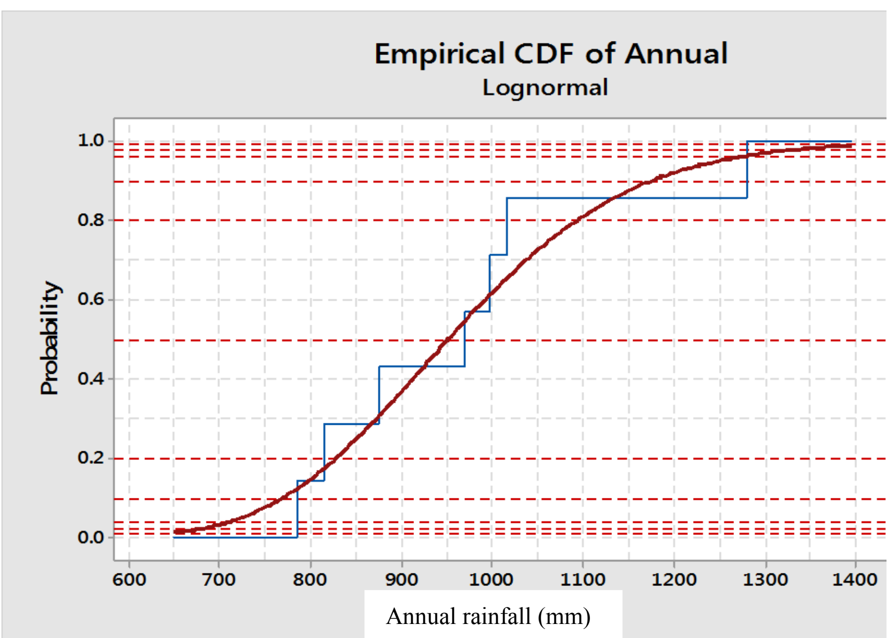


Figure 4. 12: Weyib catchment average annual rainfall (1985–2014)



Probability	Annual
0.01	649.992
0.02	679.738
0.03	699.313
0.04	714.408
0.05	726.928
0.06	737.757
0.07	747.384
0.08	756.110
0.09	764.135
0.1	771.598
0.2	829.380
0.3	873.711
0.4	913.463
0.5	952.252
0.6	992.689
0.7	1037.85
0.8	1093.33
0.9	1175.20
0.91	1186.68
0.92	1199.28
0.93	1213.28
0.94	1229.11
0.95	1247.42
0.96	1269.28
0.97	1296.68
0.98	1334.02
0.99	1395.07

Figure 4. 13: Weyib catchment Average annual rainfall cumulative distribution function

Hydrologic systems are impacted by extreme events, such as severe storms, floods and droughts. The magnitude of an extreme event is inversely related to its frequency of occurrence, very extreme events occurring less frequently than more moderate events.

For each year, spatial mean annual rainfall analysis was computed as a sum of spatial monthly average rainfalls for the year. An arithmetic spatial mean monthly rainfall is computed from all gauges with monthly rainfall record for each month. The sum of the monthly spatial average rainfall produces the spatial average rainfall for the year under consideration. Annual dry and wet return periods were computed from Lognormal probability cumulative density functions as shown in Table 4.9 for dry and wet return periods of 2, 5, 10, 25, 50 and 100-year. The 100-year drought annual catchment rainfall is 650 mm while the 100-year wet annual rainfall is 1,395 mm. That is, 650 and 1395 mm rainfall are expected approximately in every 100 year for drought annual and wet annual return period, respectively. In other words, this rainfall have a 1% probability of occurrence in every year in the two respective events. A catchment wide irregularity of minus 300 mm of rainfall would result in extreme drought and plus 440 mm of rainfall would result in high stream flows.

Table 4. 10: Dry and wet return periods for the Weyib River catchment annual rainfall

Return Period(Years)	**Annual Rainfall(mm)**
100	650
50	680
25	714
10	772
5	829
2	952
5	1093
10	1175
25	1269
50	1334
100	1395

4.5.1.5. Spatial Variation of Rainfall

Spatial variation of rainfall over a catchment is essential knowledge for water resources planning and management at catchment and sub-catchment level. Mapping spatial distribution of rainfall from data set observed at a limited number of sites provides the means to infer estimates of point (local) and areal (regional) values from the map. Spatial variation of annual rainfall over the Weyib River catchment is high with a coefficient of variation of 0.17. Rainfall amount varies generally from the Northwest to the Southeast decreasingly. Isohyetal map was produced using the Kriging interpolation package in ArcGIS 10 with linear variogram Fig.4.14. Since the Kriging method is a best linear unbiased estimator, the spatial distribution of average annual rainfall over the catchment is the best reflection of spatial variation of annual rainfall for the given rain gauge network.

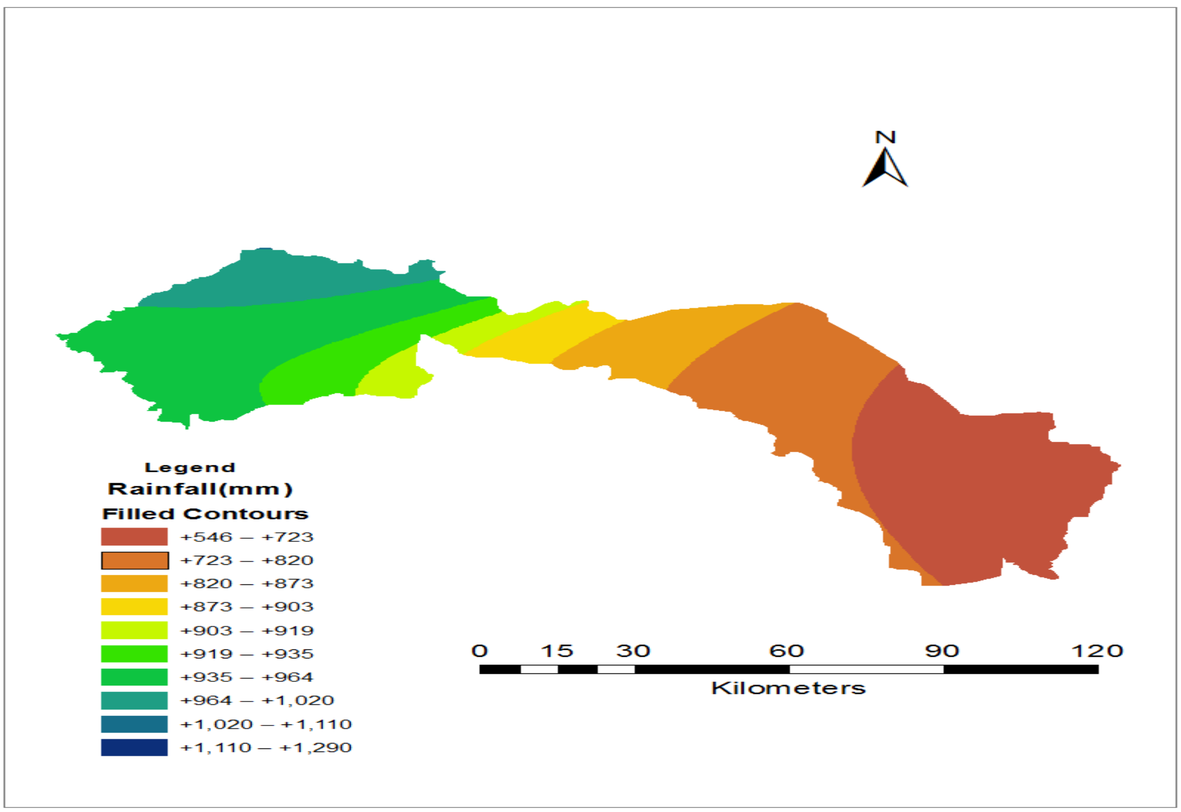


Figure 4. 14: Spatial variation of the Weyib River Catchment average annual rainfall

4.5.2. Result for High and Low flow analysis

In order to understand the behavior of the extreme flows and their recurrence interval, monthly high and low flow were analyzed for selected representative stations. The recurrence interval (in months) for the high and low flows based on the recorded data for selected rivers in the Weyib catchment is shown in Figure 4.15, 4.16, 4.17, 4.18, 4.19 and 4.20. It is shown that, based on the flow size and river, different lengths of recurrence intervals are found. Weyib River at Shaya has large flows compared to the other two rivers. Corresponding high flow with the recurrence interval of 10 years is 148.77m^3/s for Shaya at Sinana, 74.27m^3/s for Togona at Goba and 6.02m^3/s for Tebel at Ginir. Low flow analysis for the records at the three rivers also showed a varied low flow for a recurrence interval of 10 years 0.083 m^3/s for Weyib River at Shaya, 0.007m^3/s for Togona, 0.0002m^3/s for Tebel.

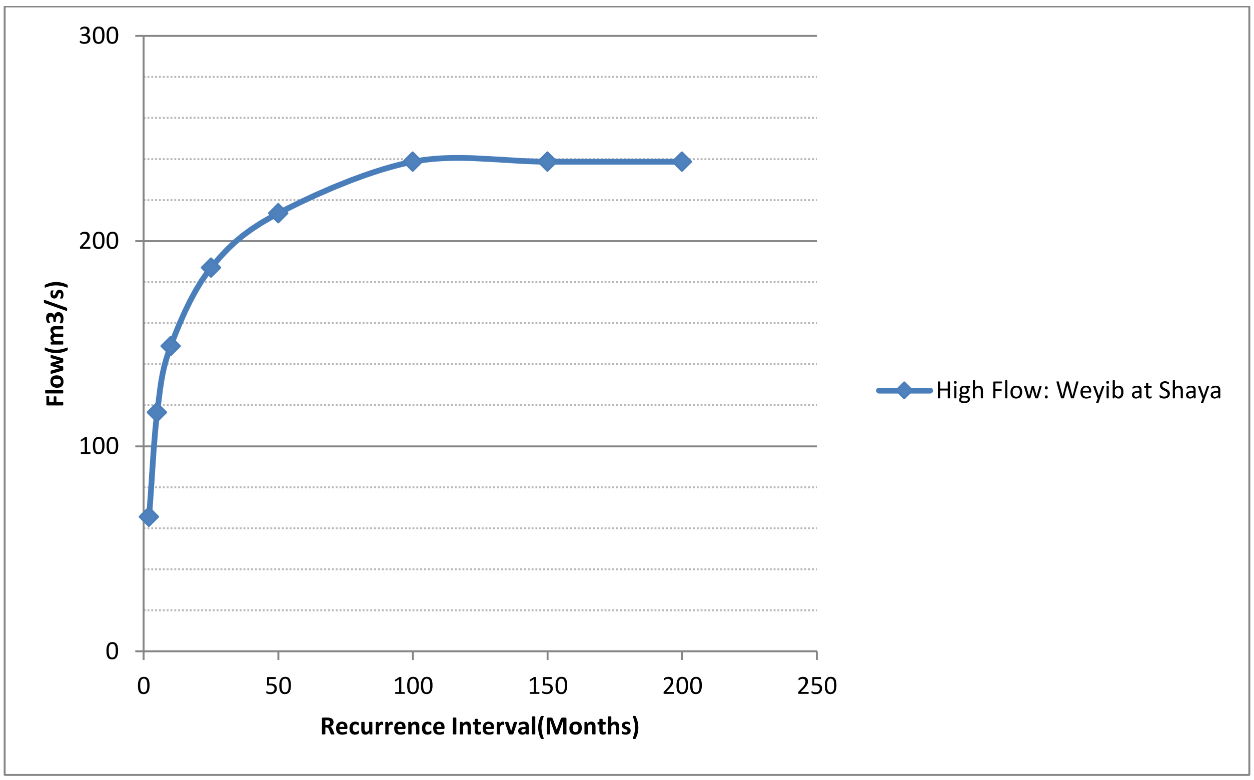


Figure 4. 15: Monthly high flow recurrence interval for Weyib River flow at Shaya

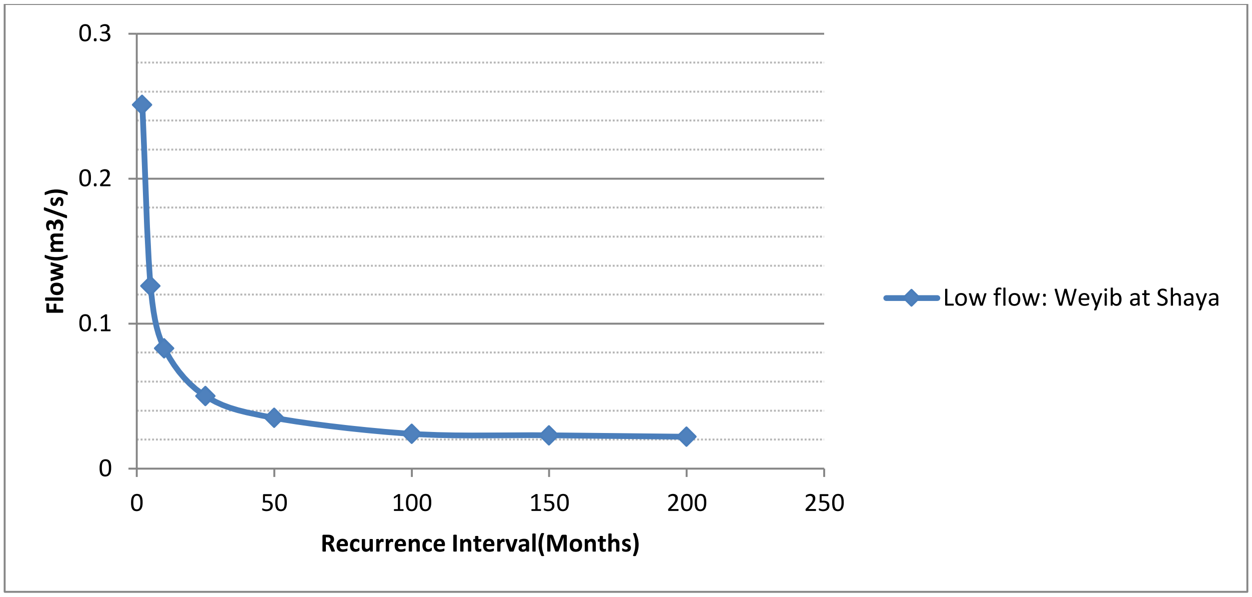


Figure 4. 16: Monthly low flow recurrence interval for Weyib River flow at Shaya

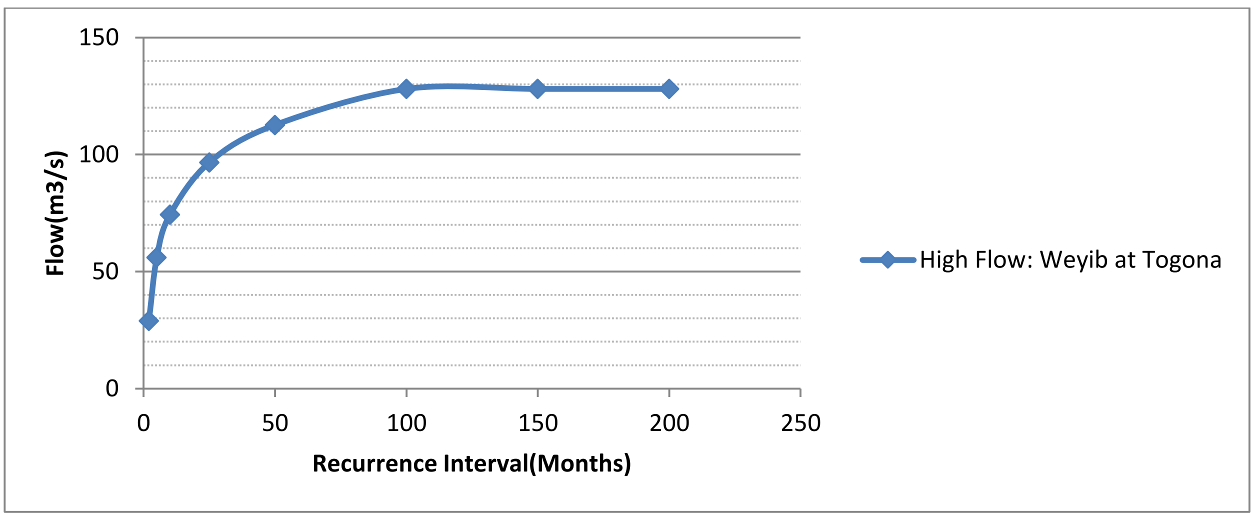


Figure 4. 17: Monthly high flow recurrence interval for Weyib at Togona

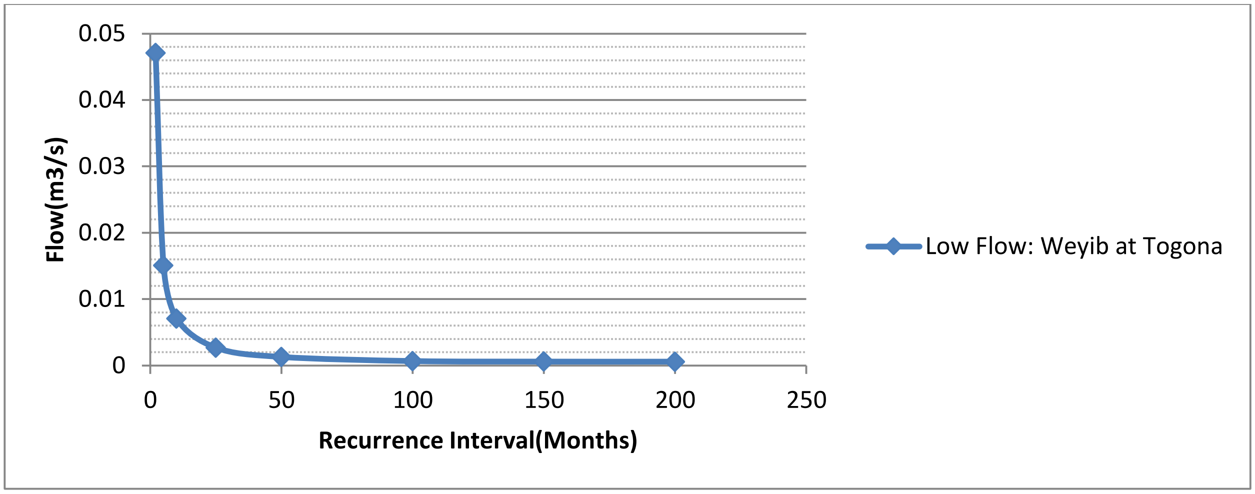


Figure 4. 18: Monthly low flow recurrence interval for Weyib at Togona

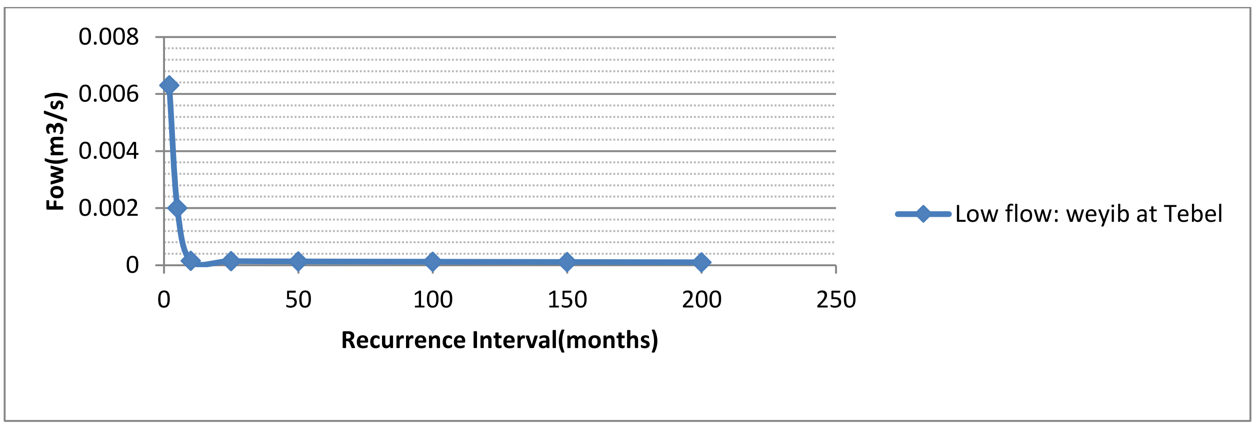


Figure 4. 19: Monthly low flow recurrence interval for Weyib at Tebel

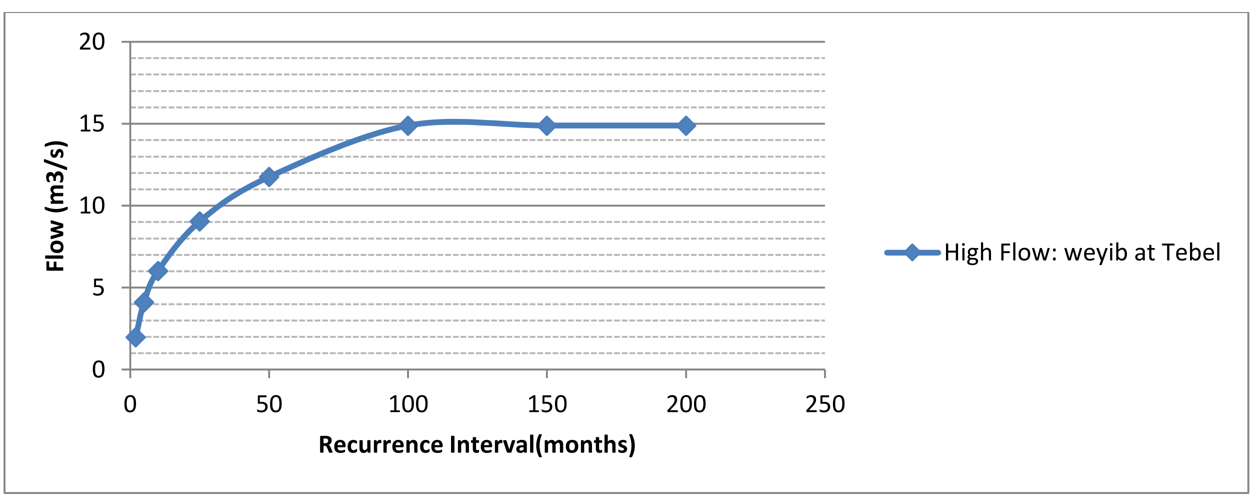


Figure 4. 20: Monthly high flow recurrence interval for Weyib at Tebel

4.6. Time series modeling

Time series data is proving to be very useful these days in a number of different fields of studies. However, fitting a specific model is not always a straightforward process. It requires a good look at the series in question, and possibly trying several different models before identifying the best one. The first step in any time series modeling is to plot the time series graph and visually inspecting that time series plot for those behaviors, which can be seen easily from the plot. This behavior of time series plot includes trend, periodicity, shift, seasonality etc. For this thesis work, the result and the steps followed were presented according to their catchments.

Sinana sub catchment result of time series modeling

Since there is one flow station in sinana sub catchment, the time series modeling has been done for this gauging station and presented systematically in the section below. Since similar steps would be followed for all gauging stations, the time series modeling result of the other five stations of other sub catchment were present in the Appendix III. The first step in any time series modeling is to plot the time series graph and visually inspecting the data. The time series plot of Shaya gauging station showed a seasonal pattern and small trend Figure 4.21 and section (4.4.1.4) indicate the data contain both seasonality and secular trend.

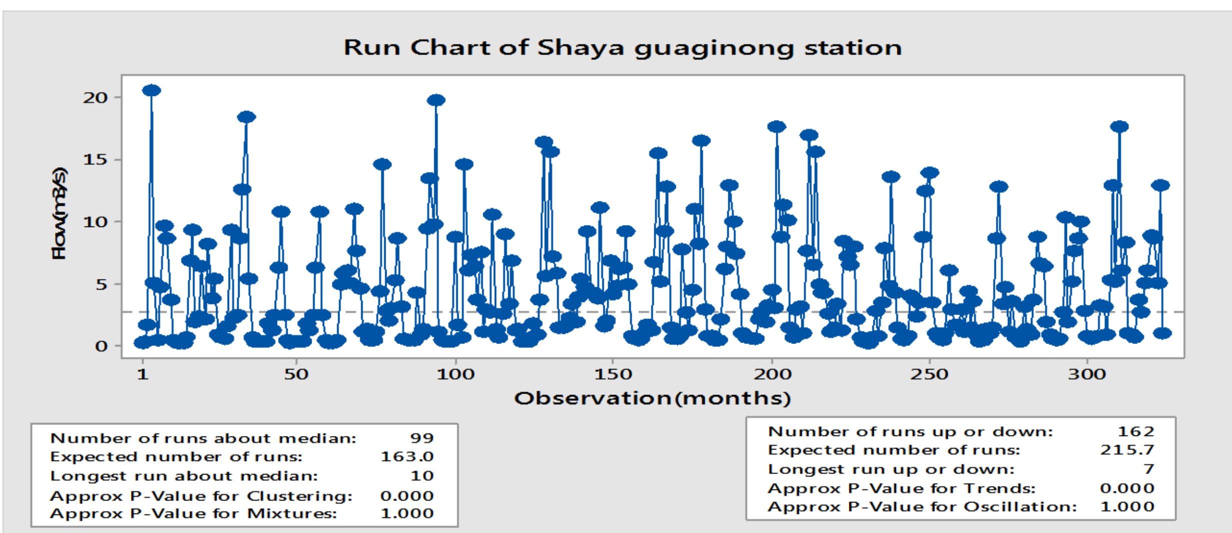


Figure 4. 21: Time series analysis plot of Shaya gauging station

The visible upward trend is confirmed by the very low p value for trends. There is also more clustering as the dash line crosses the median than we would expect if the data were random.

Here are some things to look for. First, a key assumption with these models is that our series has to be stationary. A stationary time series is one whose mean and variance are constant over time. In our case, it is clear that our mean is not constant over time. To resolve this, take a first difference of lag 1of our data, and investigate that. This means that subtract each data point from the one that follows it. When plot this lag 1 difference data, it is now stationary Figure 4.22.

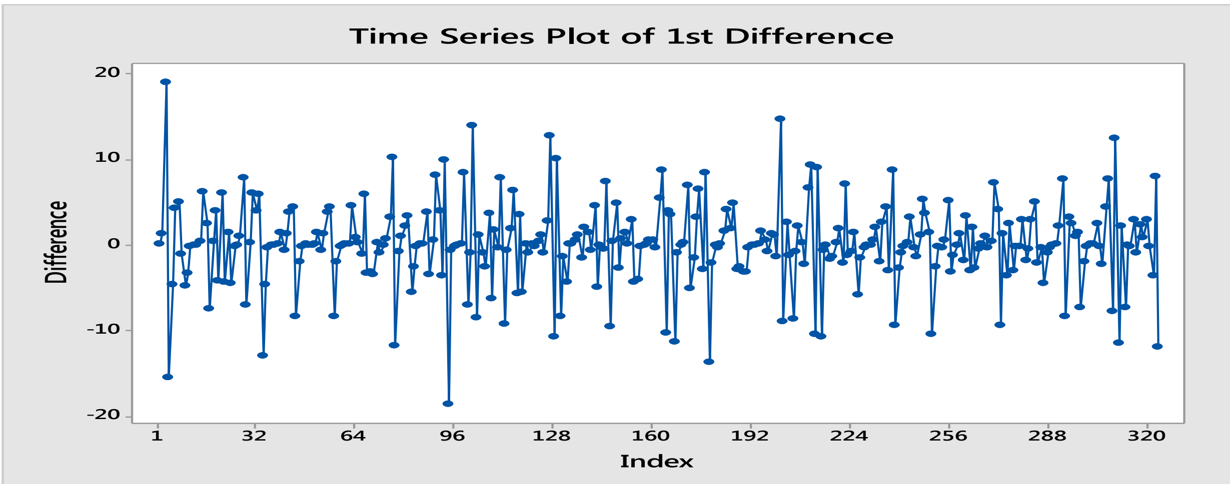


Figure 4. 22: Stationary time series analysis plot of Shaya gauging station.

The data, to which we applied a difference of order 1, seems to have made stationary, with no clear upward or downward trend. That means use this data for determining an ARIMA model. After the secular trend and stationarity were identified, there are two things to look at when trying to fit a time series model. One is past values, which is what to use in AR (autoregressive) models. Essentially, predicting what next point would be based on looking at a certain number of past points. An AR (1) model would forecast future values by looking at one past value. The second thing can look at is past prediction errors. These are called MA (moving average) models and an MA (1) model would be predicting future values using one past prediction error. Both of these concepts make sense individually; they are just different approaches to how we predict future points. An ARIMA model uses both of these ideas and allows us to fit one nice model that looks at both past values and past prediction errors.

ACF and PACF Plots

The ACF stands for Autocorrelation function, and the PACF for Partial Autocorrelation function. Looking at these two plots together can help to form an idea of what models to fit. Autocorrelation function computes and plots the autocorrelations of a time series. Autocorrelation is the correlation between observations of a time series separated by k time units. Similarly, partial autocorrelations measure the strength of relationship with other terms being accounted for, in this case other terms being the intervening lags present in the model. So perform autocorrelation and partial autocorrelation analysis on the differenced data.

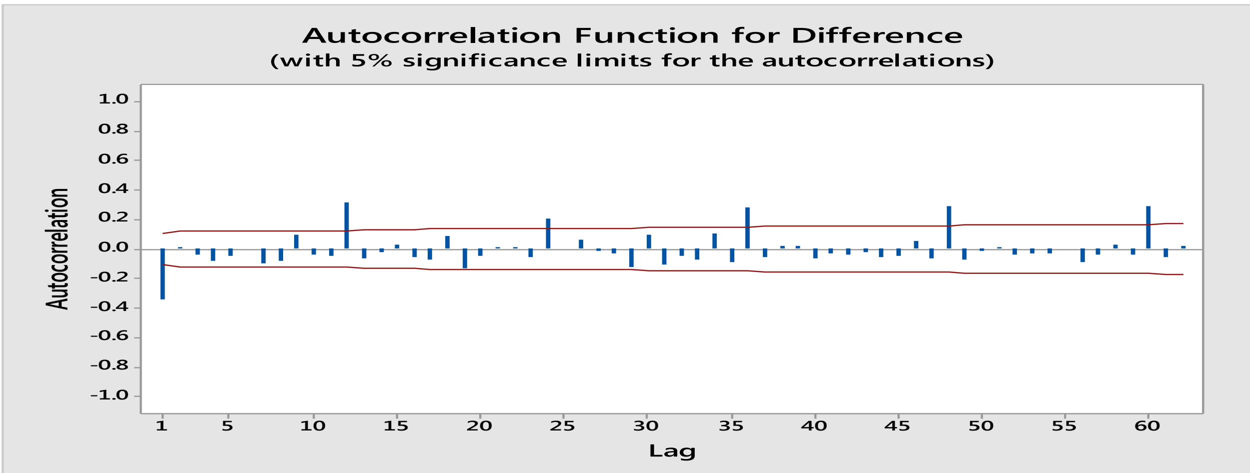


Figure 4. 23: Auto correlation function plot of difference

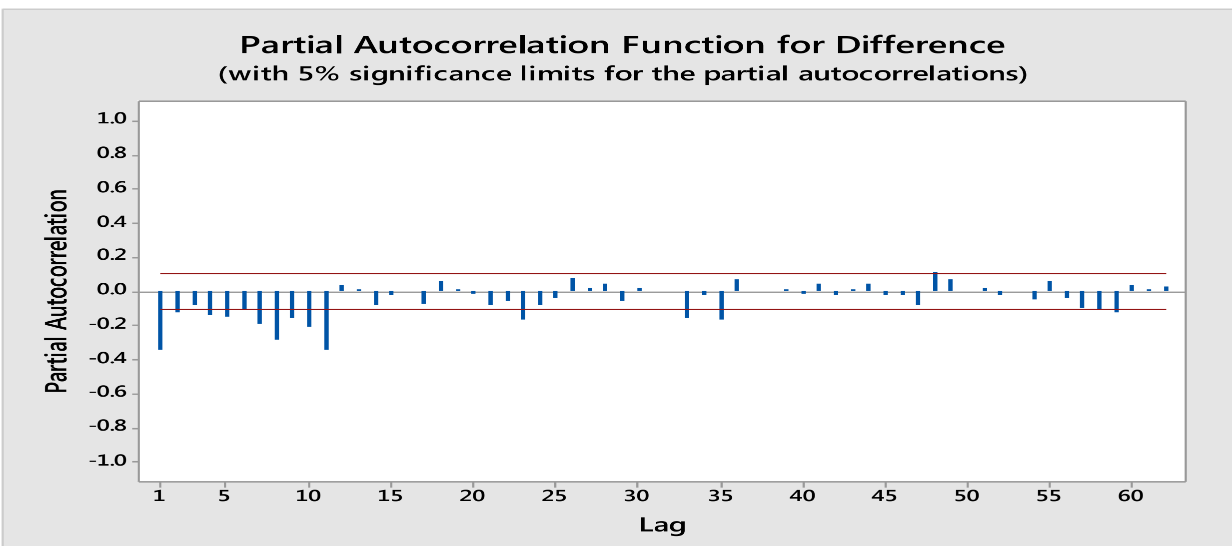


Figure 4. 24: Partial Auto correlation function plot of difference

An autocorrelation function with a sinusoidal (sine-wave-like) pattern and spikes for lags 1, suggests an Autoregressive Model of order 1, or AR(1). Sinusoidal behavior on the partial autocorrelation function and spikes up to lag 2 suggests a moving average model of order 2 or MA (2). Time series modeling can be a bit of an iterative, or even hit-or-miss, process but these graphs suggest the ARIMA (1,1,2) model is a good place to start. Therefore, from the above ACF and PACF it suggests that the use of both the autoregressive term and the moving average term integrated together (ARIMA) model. However, since data is differenced to remove seasonality in the data, ARIMA model is known as SARIMA (seasonal autoregressive integrated moving average) model. Each part of the ARIMA model has a role in the predictions it makes. The autoregressive part of the model predicts the value at time t by considering previous values in the series at time t_{-1}, t_{-2}, etc. The moving average uses past residual values the differences between the actual value and the predicted value based on the model at time t.

Once the likely model is identified, fitting the likely models and examining the significance of parameters and select one model that gives the best fit was the next step. After several iterations were performed, the best-fit model was identified and presented as below. This model also did forecast.

ARIMA Model: Shaya gauging station

```
Estimates at each iteration

Iteration    SSE                   Parameters
0 9499.21   0.100   0.100   0.100 0.100   0.100   0.055
1 7180.55   0.018   0.087  -0.027  0.183  0.250  -0.000
2  6924.19  0.051   0.212  -0.022  0.236  0.400  -0.000
3  6483.01  0.104   0.314  -0.024  0.325  0.550  -0.001
4  5819.07  0.172   0.383  -0.031  0.452  0.700  -0.000
5  5057.60  0.245   0.351  -0.054  0.602  0.778   0.000
6  4341.91  0.306   0.279  -0.083  0.752  0.830   0.001
7  3651.72  0.341   0.168  -0.121  0.902  0.878   0.001
8  3509.96  0.322   0.166  -0.142  0.961  0.933   0.000
9  3264.95  0.243   0.016  -0.125  0.963  0.912   0.001
10  3227.33 0.194   -0.056 -0.175  0.953  0.912   0.001
11  3224.73 0.190   -0.072 -0.217  0.956  0.912  -0.000
12  3223.75 0.189   -0.071 -0.208  0.954  0.914   0.000
13  3223.57 0.189   -0.072 -0.208  0.954  0.914   0.000
14  3223.08 0.189   -0.086 -0.208  0.954  0.907   0.001

Unable to reduce sum of squares any further
```

```
Final Estimates of Parameters

Type           Coef   SE Coef      T      P
AR   1       0.1886    0.0571   3.30  0.001
SAR  12     -0.0861    0.0595  -1.45  0.149
SAR  24     -0.2078    0.0599  -3.47  0.001
MA   1       0.9538    0.0137  69.41  0.000
SMA  12      0.9065    0.0333  27.21  0.000
Constant   0.000730  0.002142   0.34  0.733

Differencing: 1 regular, 1 seasonal of order 12
Number of observations: Original series 324, after differencing
311
Residuals: SS = 2906.58 (back forecasts excluded)
           MS = 9.53  DF = 305
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag             12     24     36     48
Chi-Square    12.8   24.6   34.6   47.3
DF               6     18     30     42
P-Value      0.047  0.138  0.256  0.265
```

The first-order autoregressive coefficient, the seasonal coefficients, and the first-order moving average coefficient are all significant at the 10% alpha level, indicating that this might be an efficient model.

```
Analysis of Variance for Shaya (observed flow vs.simulated flow)

Analysis of Variance

Source          DF   Adj SS   Adj MS   F-Value   P-Value
Regression       1     2610  2610.14    279.54     0.000
Error          309     2885     9.34
Total          310     5495

Model Summary

  S       R-sq       R-sq(adj)   R-sq(pred)
3.05570  47.50%     47.33%      46.70%
```

Then this information can then be plugged into the formula using excel

n	SSE	(Sigma-hat)^2	K	AIC	BIC
311	2885	9.27	1	694.75	698.49

The sum of squares that measures the sum of the squared differences between each original data point and calculated value of AIC and BIC using ARIMA (111) $(211)_{12}$ model are quite small.

What's more, the Ljung-Box Chi-Square statistics show no correlations between points with a difference of 24, 38and 48 lags including the seasonal coefficient has factored these out.

Final equation of ARIMA model

As described earlier in the methodology part the time series that contains both the Autoregressive (AR) and the Moving average term is represented as:

$$z_t = \beta + \phi_1 (z_{t-1} - z_{t-2}) + \varphi_1 (z_{t-12} - z_{t-13}) + \varphi_2 (z_{t-13} - z_{t-14}) - \phi_1\varphi_1 (z_{t-13} - z_{t-14}) - \theta_1 e_{t-1} - \Theta_1 (e_{t-12}-e_{t-13}) - \theta_1\Theta_1 (e_{t-13}-e_{t-14}) \quad (4.1)$$

So, putting the found parameters in to the equation (4.1) will provide the equation (4.2)

$$z_t = 0.000730 + 0.1886(z_{t-1} - z_{t-2}) - 0.0861(z_{t-12} - z_{t-13}) - 0.2078(z_{t-13} - z_{t-14}) + 0.1886*0.0861(z_{t-13} - z_{t-14}) - 0.9538e_{t-1} - 0.9065 (e_{t-12} - e_{t-13}) - 0.9538*0.9065(e_{t-13}-e_{t-14}) \quad (4.2)$$

- z_t is the predicted value at time t based on the model
- e_t is the residual value at time t
- β is constant term
- ϕ_i are the non-seasonal autoregressive coefficients(parameter)
- Φ_i are seasonal autoregressive coefficients(parameter)
- θ_i are non-seasonal moving average coefficients(parameter)
- Θ_i are seasonal moving average coefficient(parameter)

After all the above steps are undertaken, the final steps of the time series modeling are to test the parameter and select the model that best fit the given data. The ARIMA model presented earlier above is converged after 14 iterations until the best fit was found.

Model fitting and forecasting

Assessment of how well the models fits the original values and see what the model predicts for this process in the future.

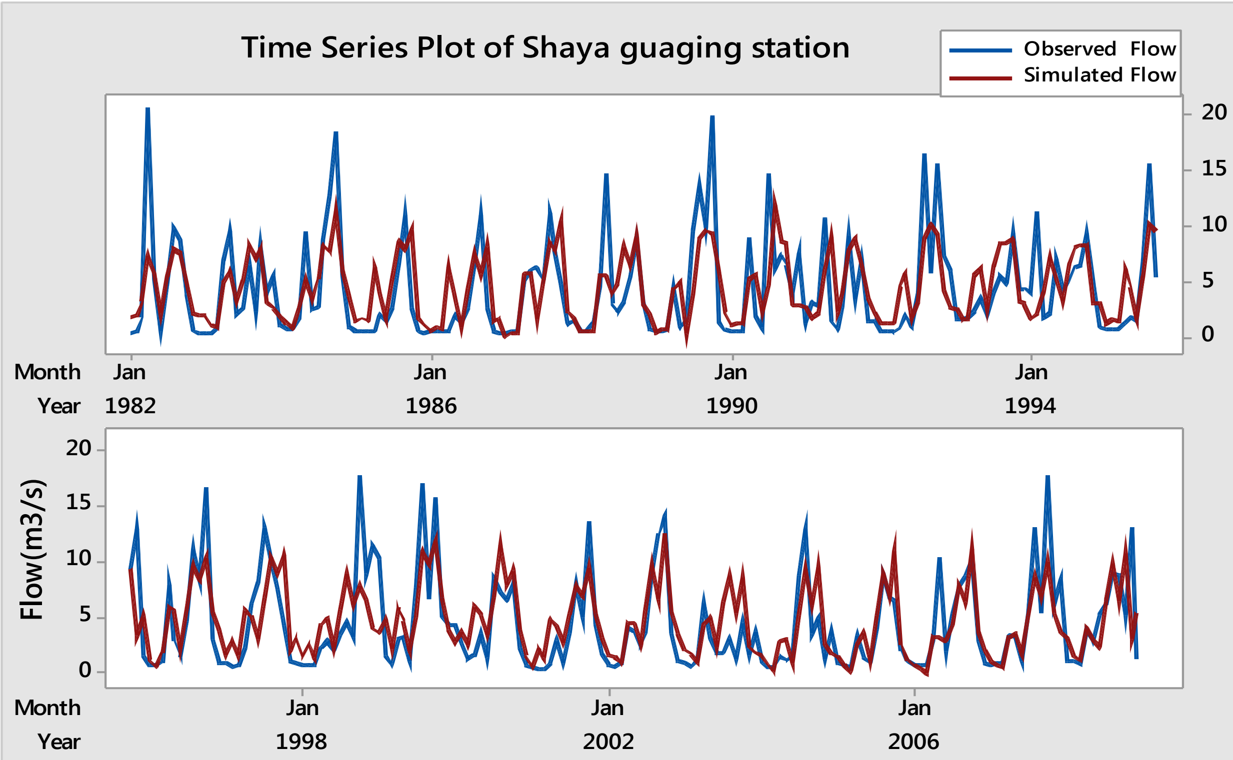


Figure 4. 25: Time series plot between observed data and model output.

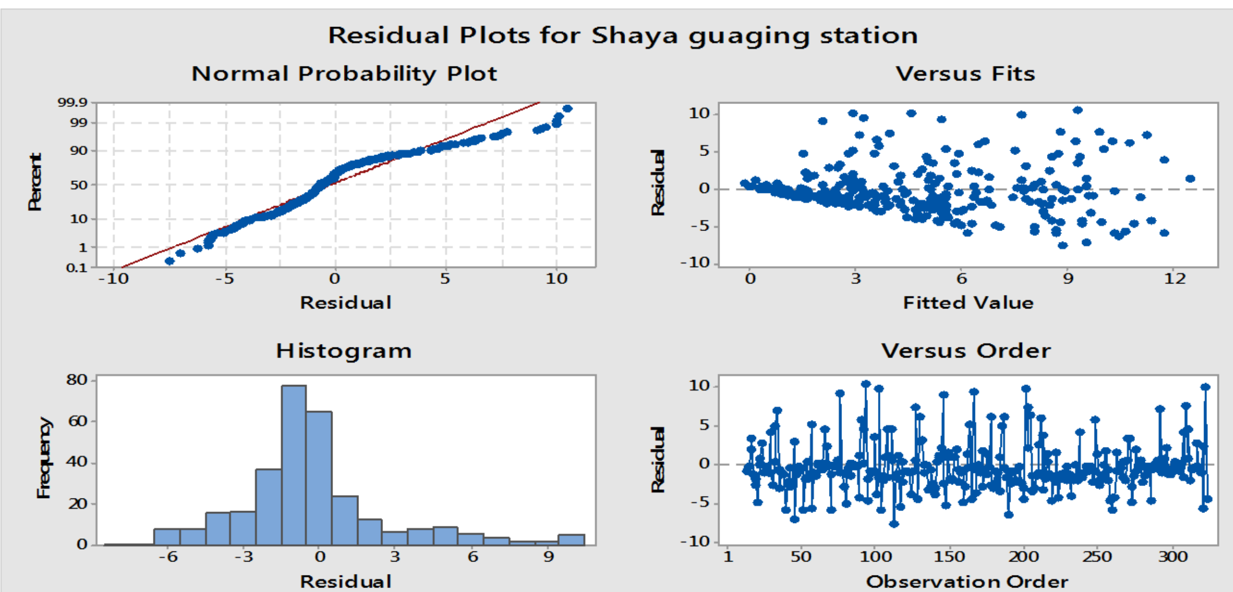


Figure 4. 26: Residual plot of shaya gauging station

As one of the basic test for ARIMA model selection and testing is normality of residuals, once the model parameters are determined, the residuals are found and tested for normality. Therefore, the plot of normality shows that the above model fit the data of a given site very well.

Forecasts from period 324

		95% Limits	
Period	Forecast	Lower	Upper
325	1.7352	-4.3166	7.7870
326	1.7156	-4.5009	7.9321
327	1.4356	-4.8050	7.6762
328	4.2050	-2.0473	10.4574
329	4.8862	-1.3761	11.1484
330	2.7917	-3.4801	9.0635
331	6.1175	-0.1638	12.3987
332	9.1892	2.8985	15.4800
333	7.4168	1.1166	13.7170
334	9.4778	3.1682	15.7875
335	3.9113	-2.4077	10.2304
336	2.5286	-3.7998	8.8571
337	2.1922	-4.1483	8.5326
338	1.8796	-4.4708	8.2300
339	1.5943	-4.7656	7.9543
340	4.2728	-2.0967	10.6423
341	4.9783	-1.4006	11.3573
342	2.3165	-4.0718	8.7049
343	6.1517	-0.2462	12.5495
344	10.2098	3.8025	16.6170
345	7.0066	0.5899	13.4232
346	11.9150	5.4889	18.3410
347	3.4593	-2.9762	9.8947
348	4.1148	-2.3301	10.5596
349	2.2037	-4.2506	8.6579
350	1.9009	-4.5559	8.3577
351	1.6378	-4.8253	8.1009
352	4.3677	-2.1025	10.8379
353	4.7396	-1.7379	11.2170
354	3.0536	-3.4312	9.5384

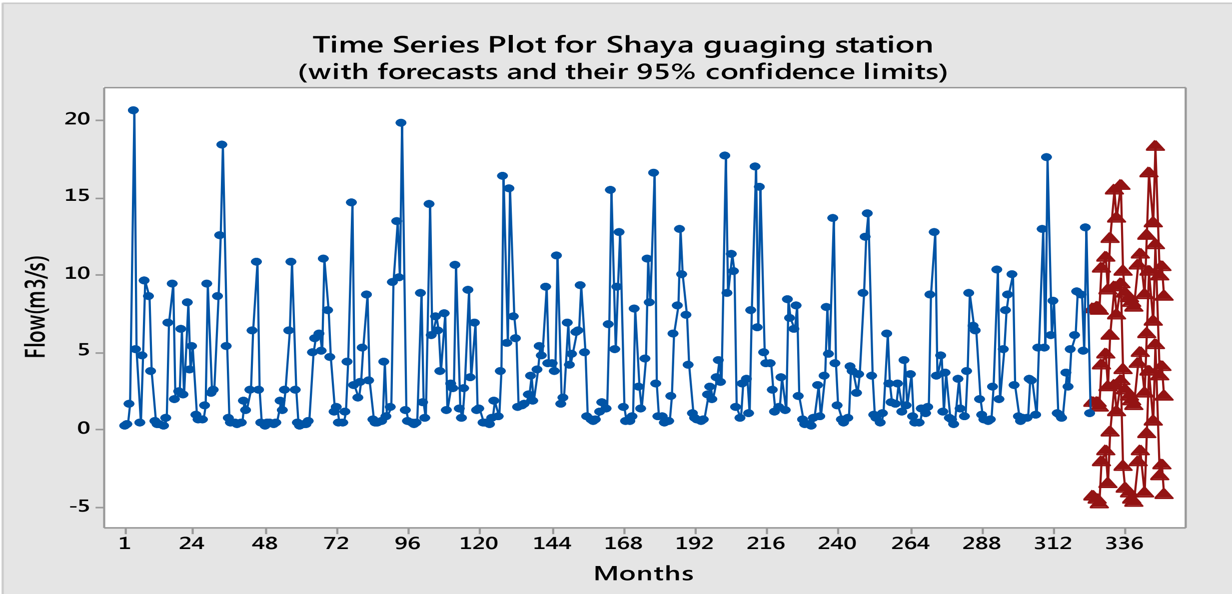


Figure 4. 27: Time series plot of shaya gauging station with forecast

In statistical terminology, k is the lead-time of the forecast. Many stochastic models have a particular advantage for forecasting purposes in that they provide, as a by-product of the procedure for estimating model parameters, confidence limits for forecasts (a pair of values, one less than the forecast and one greater, such that there is a given probability P that these values will bracket the observed value of the variable at time t+k). Confidence limits therefore express the uncertainty in forecasts; the wider apart the confidence limits, the less reliable the forecast. Furthermore, the greater the lead-time k for which forecasts is required, the greater will be the width of the confidence interval, since the distant future in more uncertain than the immediate.

Accordingly, the following is the output of Shaya stream flow station, observed time series data in blue line, forecasted values in red line and the range of expected error (2 times standard deviation) is displayed with red line lines on either side of predicted forecast line Figure 4.27.

Model Calibration and Validation

Flow calibration was performed for a period of five years from January 1996 to December 31, 2000 for all sub catchment and also flow validation was performed from 2001-2005 for monthly stream flow for all sub catchment using the sensitive parameters identified in calibration. The flow was calibrated manually using the observed flow gauged at different sub catchment in Weyib catchment.

Table 4. 11: Monthly statistical measure using Information criteria (AIC and BIC)

Station	Model Parameter	Calibration	AIC	BIC	Validation	AIC	BIC
Agarfa	$ARIMA(1,1,2)(3,1,2)_{12}$	1996-2000	255.5	257.59	2001-2005	172.08	174.18
Togona	$ARIMA(1,1,2)(3,1,1)_{12}$	1996-2000	70.97	73.07	2001-2005	60.27	62.37
Shaya	$ARIMA(1,1,2)(2,1,3)_{12}$	1996-2000	174.81	176.9	2001-2005	139.32	141.41
Denbel	$ARIMA(1,1,1)(3,1,2)_{12}$	1996-2000	239.72	241.81	2001-2005	114.91	117.00
Alemkrem	$ARIMA(1,1,1)(3,1,2)_{12}$	1996-2000	335.49	337.59	2001-2005	280.35	282.45
Tebel	$ARIMA(2,1,1)(2,1,2)_{12}$	1996-2000	-136.0	-133.9	2001-2005	-136.7	-134.6

Table 4. 12: Monthly statistical measure using coefficient of determination (r^2)

Station	Model Parameter	Calibration	N Months	R^2	Validation	N Months	R^2
Agarfa	$ARIMA(1,1,2)(3,1,2)_{12}$	1996-2000	60	0.97	2001-2005	60	0.92
Togona	$ARIMA(1,1,2)(3,1,1)_{12}$	1996-2000	60	0.96	2001-2005	60	0.96
Shaya	$ARIMA(1,1,2)(2,1,3)_{12}$	1996-2000	60	0.93	2001-2005	60	0.84
Denbel	$ARIMA(1,1,1)(3,1,2)_{12}$	1996-2000	60	0.96	2001-2005	60	0.97
Alemkrem	$ARIMA(1,1,1)(3,1,2)_{12}$	1996-2000	60	0.96	2001-2005	60	0.89
Tebel	$ARIMA(2,1,1)(2,1,2)_{12}$	1996-2000	60	0.92	2001-2005	60	0.69

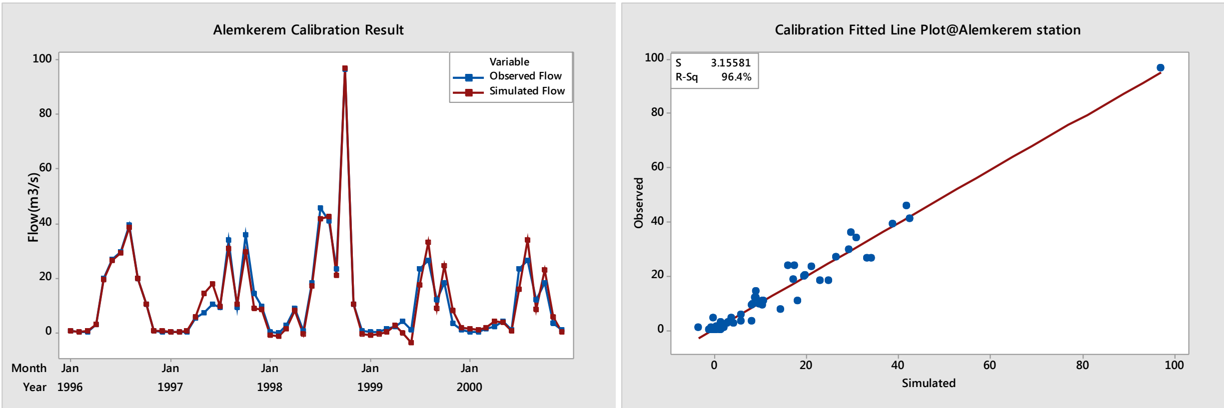


Figure 4. 28: Calibration of observed and simulated monthly flow, for Alemkerem station

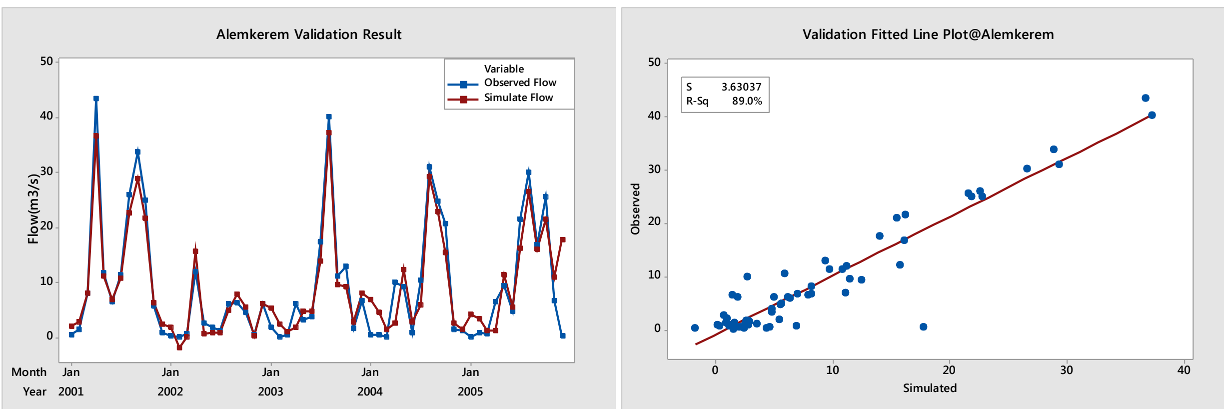


Figure 4. 29: Validation of observed and simulated monthly flow, for Alemkerem station

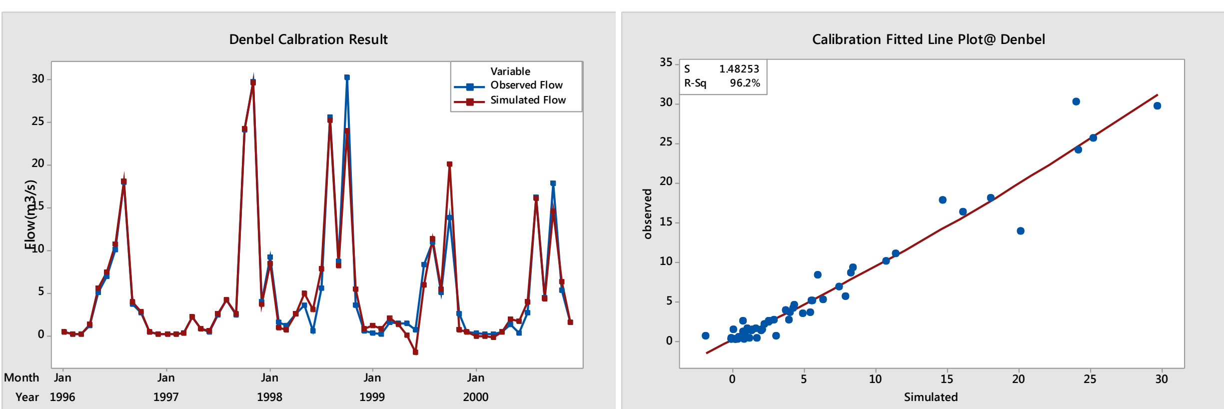


Figure 4. 30: Calibration of observed and simulated monthly flow, for Denbel station

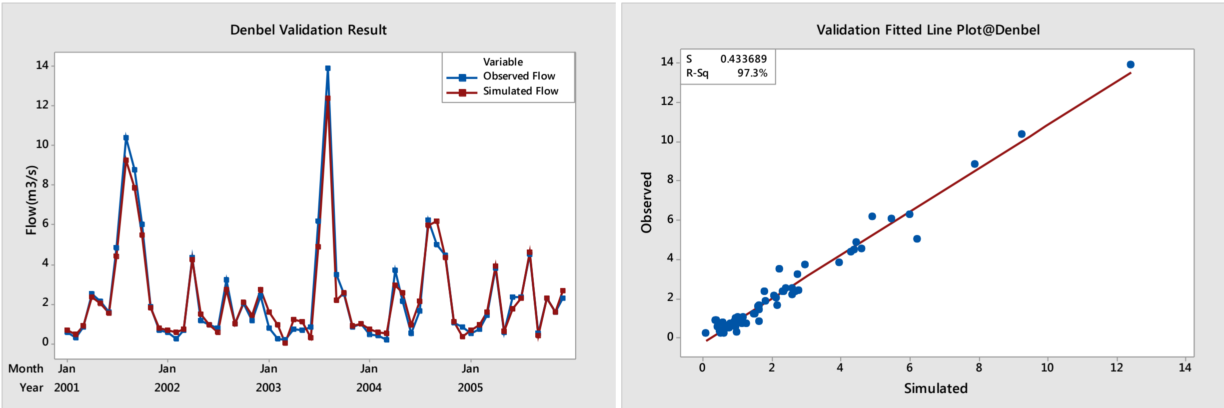


Figure 4. 31: Validation of observed and simulated monthly flow, for Denbel station

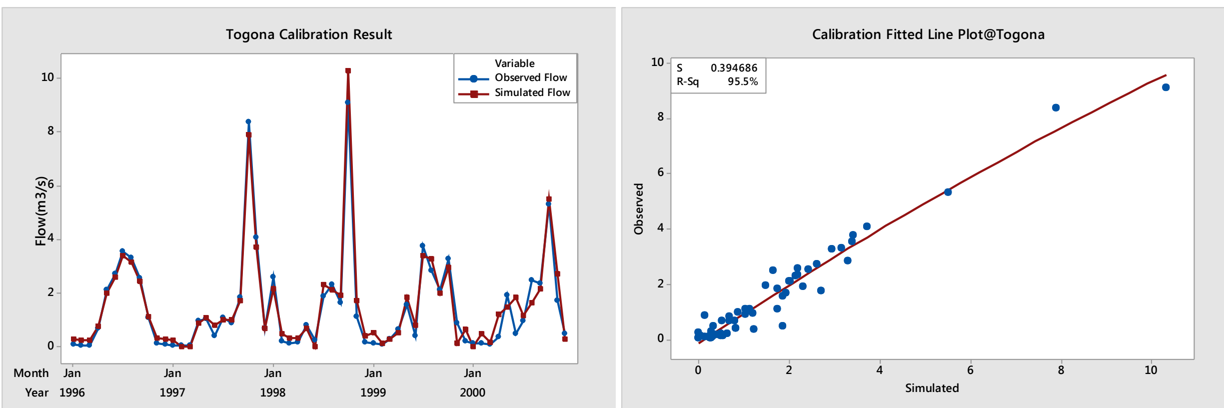


Figure 4. 32: Calibration of observed and simulated monthly flow, for Togona station

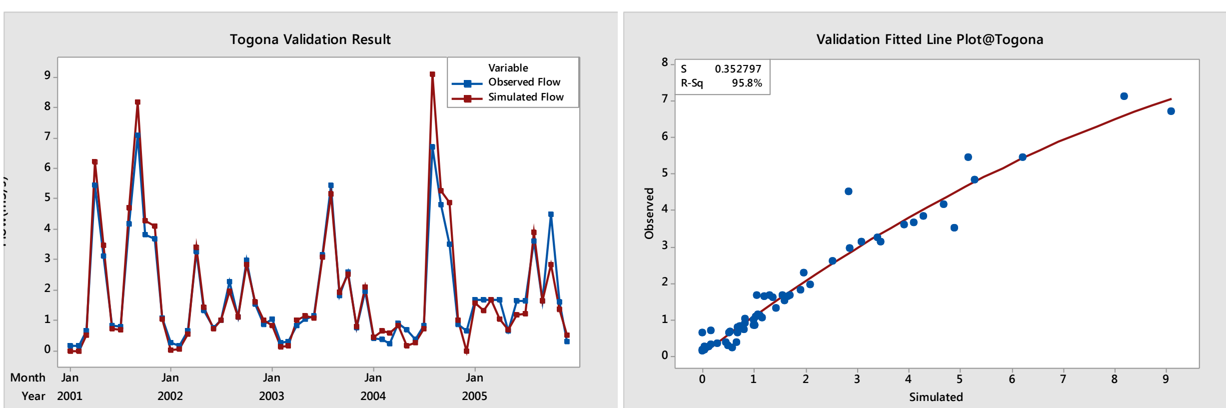


Figure 4. 33: Validation of observed and simulated monthly flow, for Togona station

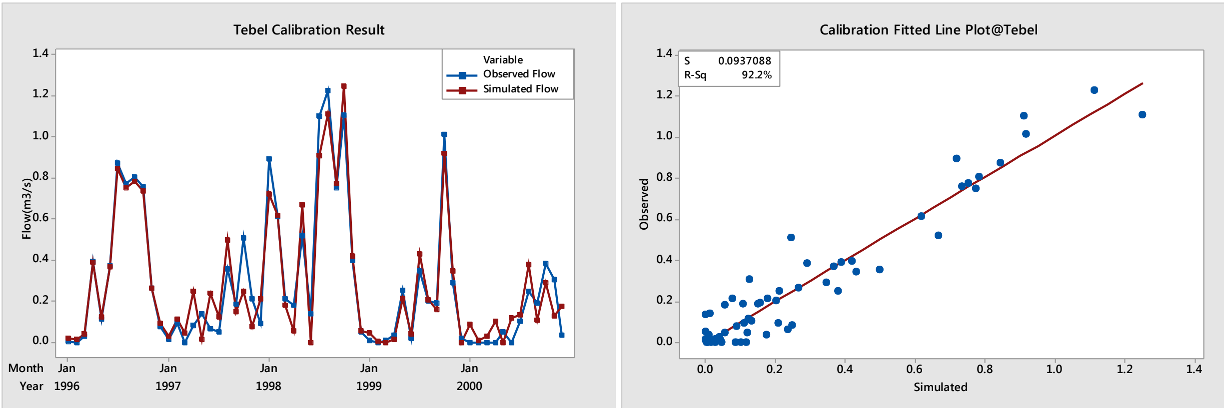


Figure 4. 34: Calibration of observed and simulated monthly flow, for Tebel station

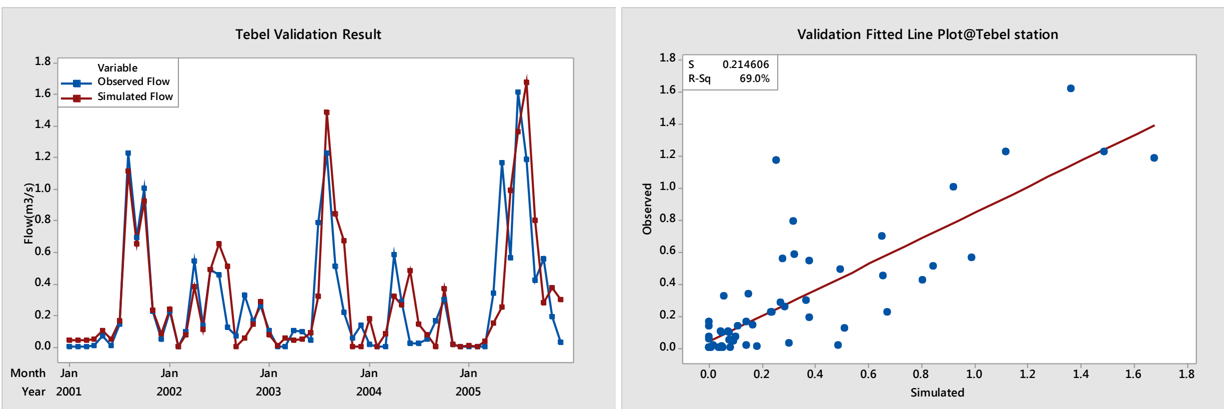


Figure 4. 35: Validation of observed and simulated monthly flow, for Tebel station

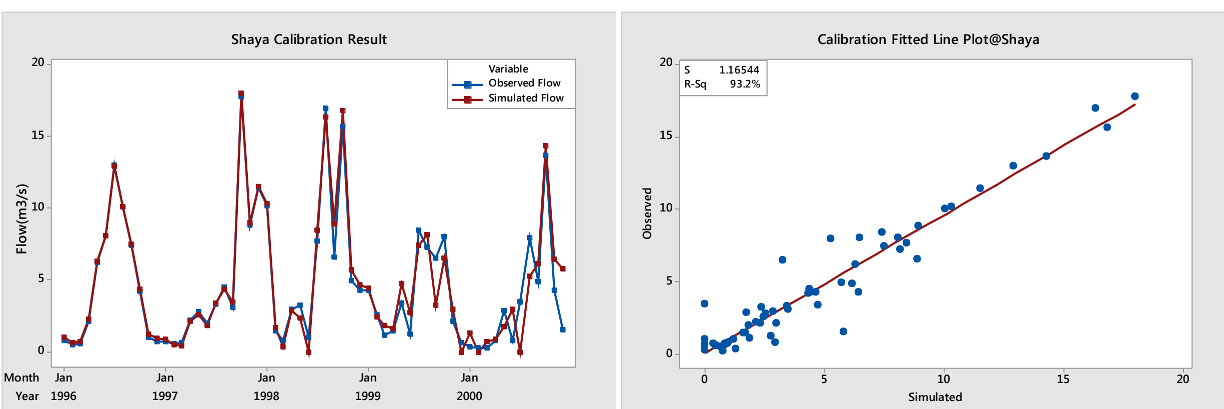


Figure 4. 36: Calibration of observed and simulated monthly flow, for Shaya station

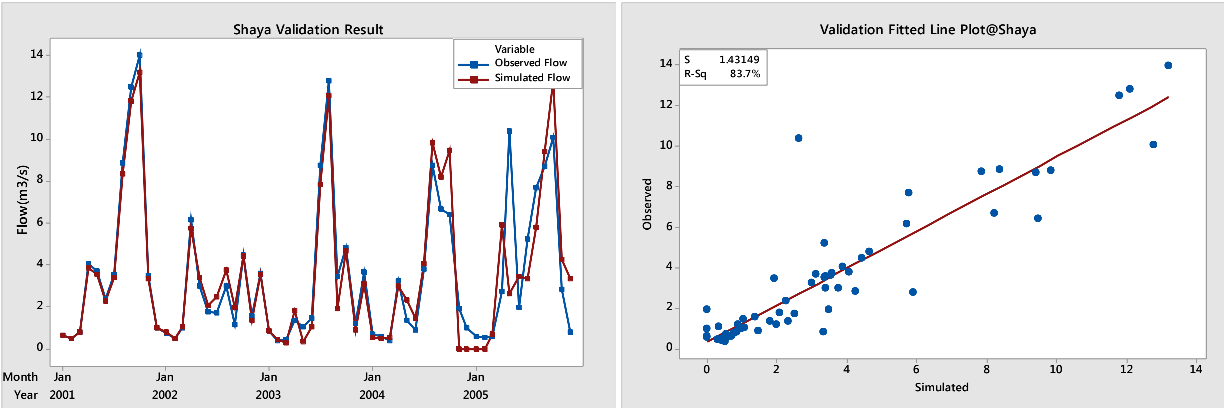


Figure 4. 37: Validation of observed and simulated monthly flow, for Shaya station

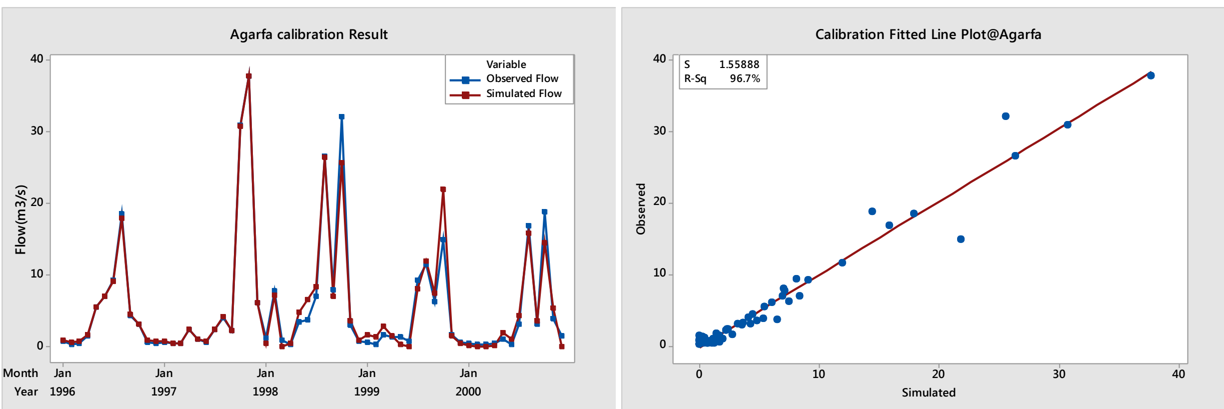


Figure 4. 38: Calibration of observed and simulated monthly flow, for Agarfa station

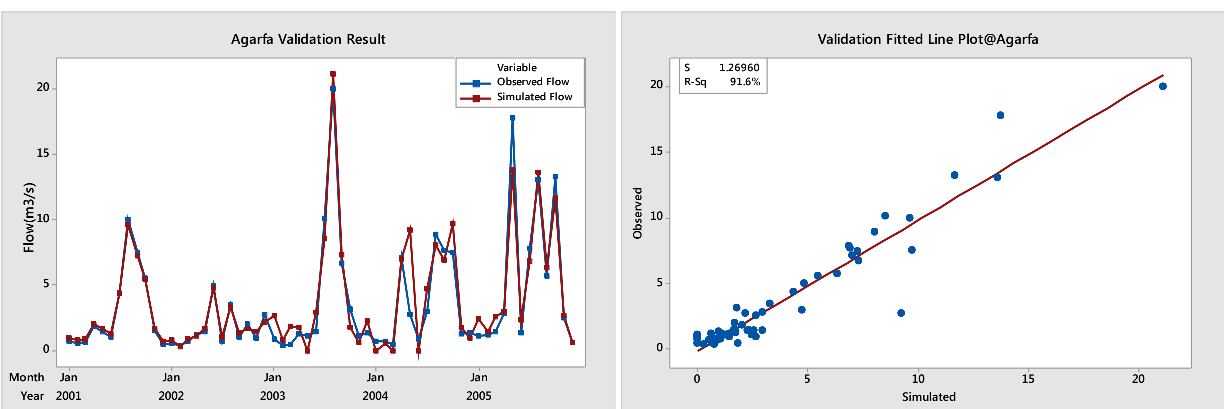


Figure 4. 39: Validation of observed and simulated monthly flow, for Agarfa station

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1. Conclusion

This study has clearly demonstrated that hydrological studies focusing on localized areas are possible and important, since local-scale hydrology has direct relevance to large populations dependent on subsistence agriculture. Based on the results obtained and their analyses, the following conclusions are drawn:

From the Runoff trend analysis, it can be concluded that Weyib catchment had experienced a significant upward trend in Tebel and Togona station annually and seasonally. mean while Mann- Kendall value of all months of Tebel station, January and February for shaya station, January, July, November and low flow season for Agarfa station, January and December for Denbel station and all months except April, may and Jun for Togona station had shown significant increasing trend.

The result of precipitation trend of the catchment showed that of six stations in the weyib catchment, only Ginir stations in Ginir sub catchment had a significant trend for annual and wet season. The monthly Mann- Kendall trend value of Sinana (September), Dinsho (March) and Ginir (February) are station that showed significant decreasing trend. Similarly Dinsho station (November), Robe station (July), Ginir station (May and September) showed significant increase in trend.

Autoregressive integrated moving average (ARIMA) time series model type is selected for all the six sub catchments because all the six sub catchments showed the seasonality behavior in their data records. These model types with different parameters are presented in result part.

Another concluding remark is that, with an annual average rainfall of 1002 mm and standard deviation of 108 mm, the Weyib River basin is relatively wet. The high wet season months are July, August and September the dry season months are, November, December, and January. The small wet months are March, April and May. October and February are the transition months between dry and wet seasons, Jun is the transition month from the small wet season to the high wet season, January has the lowest, and April has the highest monthly rainfall.

Rainfall probability distribution varies from month to month fitting, Normal, Weibul and Lognormal distributions. January, February, March, Jun, and August catchment rainfall fit the normal probability distribution. April, September, and December fit lognormal distribution. May, July, October and November fit Weibul distribution and the annual catchment rainfall fit a lognormal probability distribution.

The year-to-year (temporal) basin rainfall variation is relatively small with a coefficient of temporal variation of 0.108 while the spatial variation is high with coefficient of spatial variation of 0.17. Spatial distribution of annual rainfall over the basin is mapped and shows high variation with the northern tip receiving as high as 1,200 mm and the northeastern tip as low as 546 mm annual average rainfall.

5.2. Recommendations

- Since, Annual, seasonal and monthly rainfall and stream flow totals are studied using only six gauging station. It is essential to increase the spatial coverage by including more number of stations.

- The recent development of irrigation projects almost in all the sub catchments rivers must have to consider the trend of the rivers. Otherwise, the series problems may happen on the irrigation development and the downstream users.

- Factors that affect the hydrological time series like; precipitation, land use/land cover change, water diversion ,wastewater discharge etc must have to be considered and then the sensitivity analysis has to be done to get the degree of effect of each factor.

REFERENCES

Alemayehu, T., Furi, W., and Legesse, D.: Impact of water overexploitation on highland lakes of eastern Ethiopia, Environ. Geol., 52, 147–154, 2007

Andréassian, V.: Waters and forests: from historical controversy to scientific debate, J. Hydrol., 291, 1–27, doi:10.1016/j.jhydrol.2003.12.015, 2004.

Askew, A. (1987). Climate change and water resources. (S. B. Solomon, Ed.) *The Influance of climate change and climate variability on the Hydrologic Regime and Water Resources , 168*, 421-430.

Baharudin, F. B. (2007). Study on Rainfall-Runoff Characteristics of urban catchment of sungai kerayong. 12-16.

Beavis, S. G., Zhang, L., Evans, J. P., Jakeman, A. J., and Smith, D. I.: Impacts and implications of farm dams on catchment hydrology: Methods and application to chafrey catchment, MODSIM 97 IMACSt, University of Tasmania, Hobart, 1997.

Box, G. A. (1970). *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.

Brown, A. E., Zhang, L., McMahon, T. A., Western, A. W., and Vertessy, R. A.: A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation, J. Hydrol., 310, 28–61, doi:10.1016/j.jhydrol.2004.12.010, 2005.

Burn, D. (2002). Detection of hydrological trend and variability. *Journal of hydrology 255* , 107-122.

Burn, D. S. (1994b). Identification of a data collection network for detecting climate change. *Canadian Water Resources Journal 19* , 27-38.

Burn, D. S. (1992). The Use of hydrologic Variables in detecting climatic change. (G. H. Kite, Ed.) *Using Hydrometric Data to Detect and Monitor Climate Change , 8*, 121-130.

Chong-yuXu. (2002). *Hydrologic Model, Uppsala University Department of Earth Sciences.*

Chow, V. M. (1988). *Applied Hydrology.* McGraw-Hill, New York.

Council, N. R. (1991). *Opportunities in the Hydrological Sciences.* Washington, D.C: National Academy Press.

Daniel Bekele, T. (2016). Analysis of rainfall trend and variability for agricultural water managment in Awash River Basin, Ethiopia. *Journal of water and climate change* .

Daniel, G. (1977). Aspects of Climate and Water Budget in Ethiopia. *A Technical Monograph Published , Addis Ababa University Press* .

FAO. (2002). Majore Soils of the World. Land and Water Digital Media Servies(CD-ROM). *Food and Agricultural Orginization of the United Nation, Rome* .

Gilbert, R. (1987). *Statistical mothods for environmental pollution monitoring.* New York: Van Nostrand Reinhold.

Green. (2004). Auto-calibration in Hydrological Modeling using SWAT 2005 in small-scale watersheds. *Science Direct, Environmental Modeling & Software* , 422-434.

Haji, M. (2015). Hydrological Trend Analysis and Time Series Modeling. *Case Study of Wabishebele River Basin, Melkawakana Catchment* , 6-19.

Hastie, T. T. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* New York: Springer.

Helsel, D. (2002). *Statistical Methods in Water Resources techniques of Water Resources Investigation,Book 4, Chapter A3.* U.S.Geological survey, 522 pages.

Helsel, D. H. (1992). Statistical methods in water resources. *Environmental Science, no.49* .

Hirsch, R. A. (1991). Selection of Methods for the Detection and Estimation of Trends in Water Quality. *Water Resource Research, Vol.27, no.5* , 803-813.

Hirsch, R. (1982). Techniques of trend analysis for monthly water quality data. *Water resource Research,vol.18(1)* , 107-121.

Huth. R, P. L. (2004). Parametric versus non-parametric estimates of climate trends. *Theoretical and Applied Climatology 77* , 107-112.

James. (1960). *Double-Mass Curves.* Washington, United Stated Department of the Intrior.

Kara M.Watson, R. G. (2003). *Stream flow characteristic and trend in New Jersy.*

Kendall, M. (1975). *Rank Correlation Method.* London: Charles Griffin.

MAKESENSE. (2002). *Mann-Kendall Test and Sen's Slope Estimates for Trend of Annual Data, Finnish Meteorological Institute.*

Mann, H. (1945). Non-parametric test against trend. *Econometrical, 13* , 245-259.

Masih, I., Maskey, S., Uhlenbrook, S., and Smakhtin, V.: Impact of upstream changes in rainfed agriculture on downstream flow in a semi-arid basin, Agr. Water Manage, 100, 36 45,doi:10.1016/j.agwat.2011.08.013, 2011.

Mathew, R. (2004). *Trend variability in obsrevation of winter precipitation.*

McCuen, R. (1998). Hydrologic analysis and design . *Prentice-Hall, Englewood Cliffs, NJ* .

McCuen, R. (1998). *Hydrologic analysis and design.*

McCuen, R. (1993). Microcomputer Application in Statistical Hydrology. *Prentice-Hall, Ink, Englewood Cliffs, NJ* .

Mohammad. K, F. B. (2003). *Water Resource Systems Anaysis.* United State of America: Lewis.

Radziejewski, M. B. (2000). Detection of change in river flow using phase randomisation. *Hydrol.Sci.J* .

Robson, Z. W. (2000). World Climate Programme-Water. *Detecting trend and other changes in hydrological data, WMO/TD-No.1013* , 46-57.

Salas, J. D. (1988). Applied Modeling of Hydrological Time Series. *Water Resource Puplication, Fort Collins, Co* .

Shahim, M. V. (1993). *Statistical Analysis in Water Resources Engineering.* A.A .Balkema, Rotterdam,the Netherlands.

Sirois, A. (1998). *A Brief and Biased Overview of Time Series Analysis.In WMO Report No.133.* Helsinki, September 1998.

Storch.H.V, V. (1995). Mis uses of statistical analysis in climate research, in analysis of climate variability. (N. H.V storch, Ed.) *Applications of statistical Techniques* .

Wing H.Cheung, G. (2008). Trend and spatial distribution of annual and seasonal rainfall in Ethiopia. *International Journal of climatology,12* .

www.actalliance.org. Drought emergency response in Ethiopia: ETH 171, Nairobi, 3March 2017

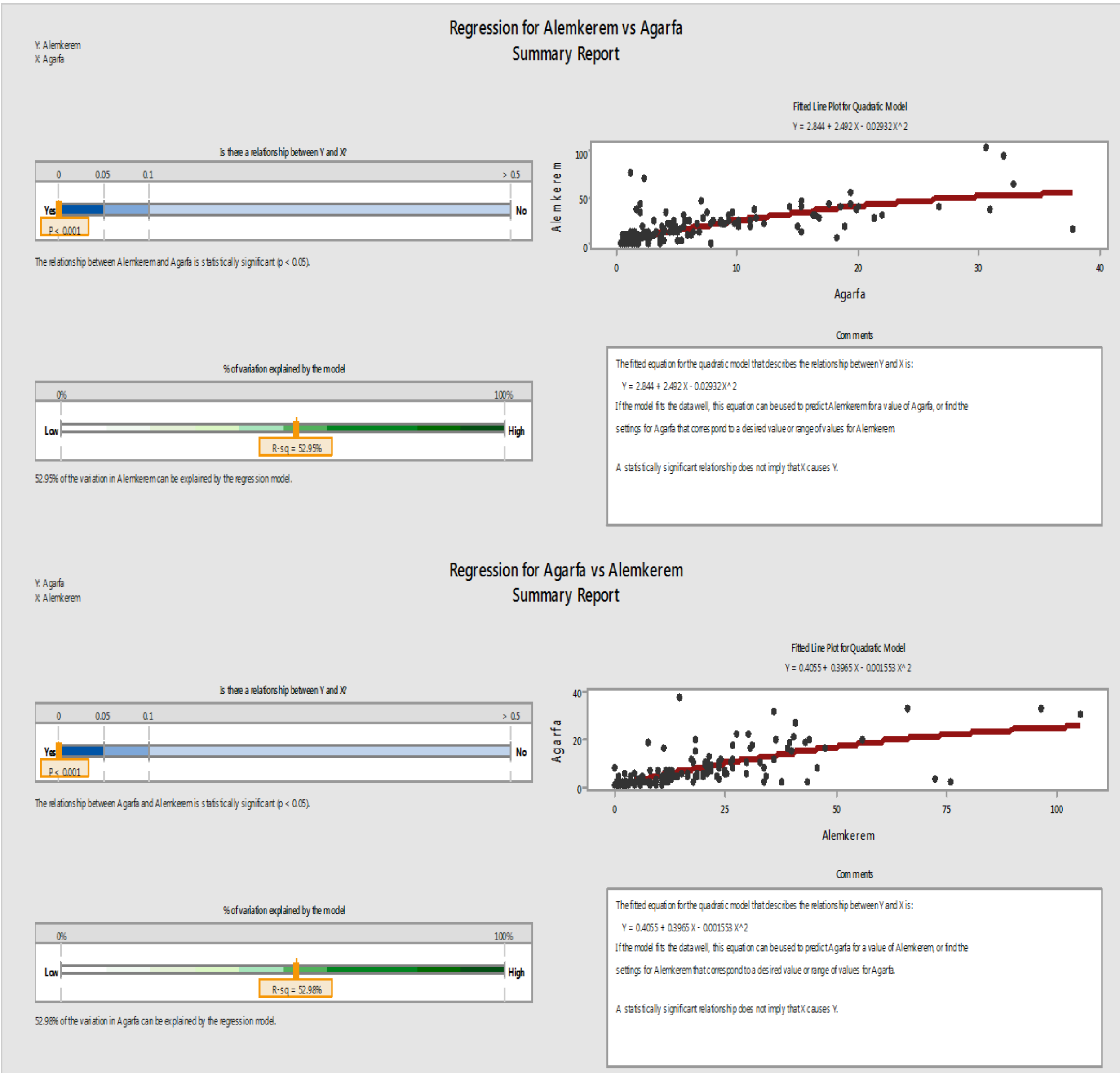
Yenigun K, Gumus V, Bulut H (2008). Trends in stream flow of the Euphrates basin, Turkey. Proc. Inst. Civil Eng. Water Manage. 161: 189–198. doi: 10.1680/wama.2008.161.4.189.

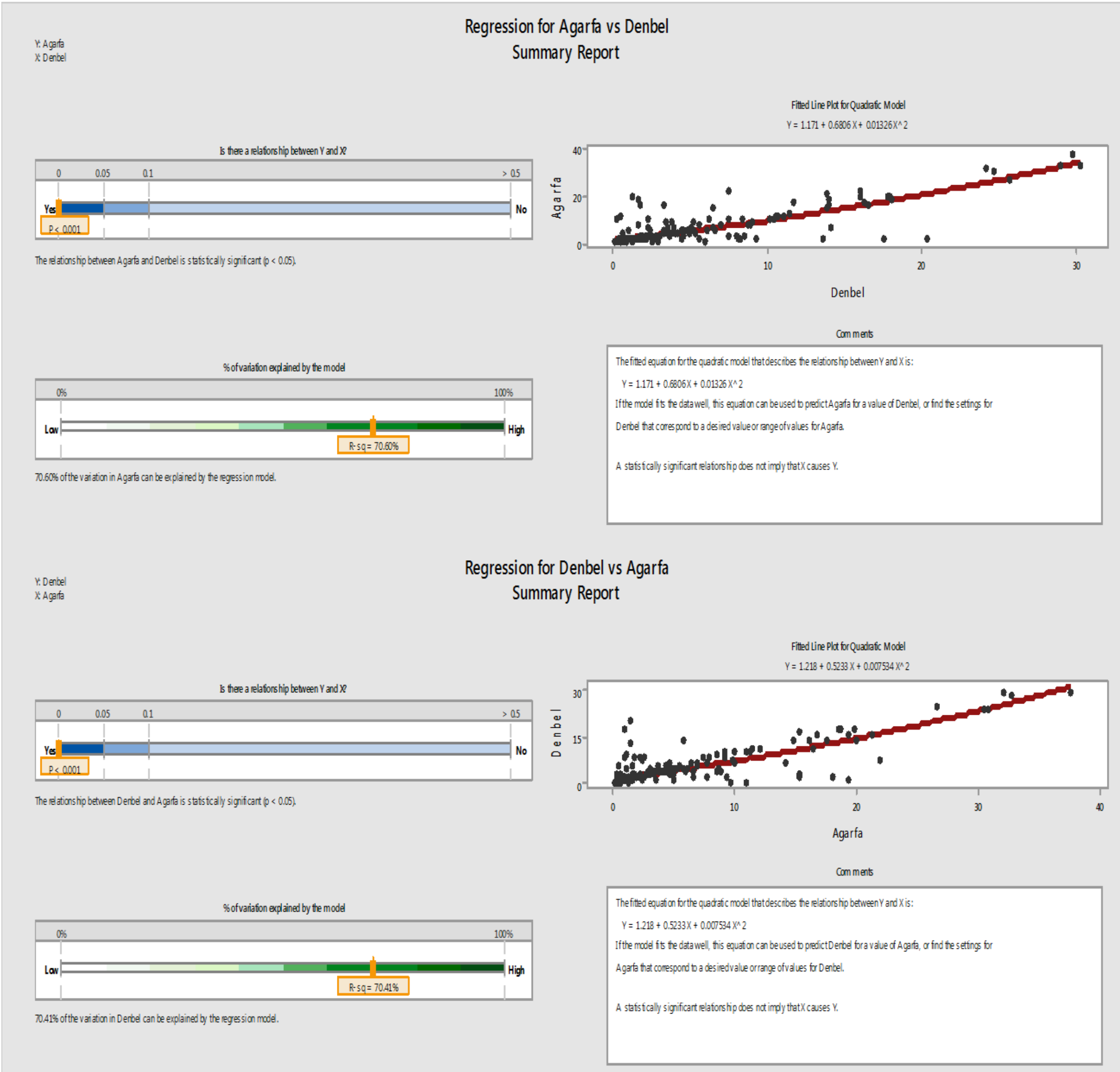
Yue, S., Pilon, P. & Cavadias, G. (2002(2)). Power of the Mann-Kendall test and the Spearman's rho test for detecting monotonic trends in hydrological time series. *J. Hydro/.* 259: 254-271.

Zhang.X, V. H. (2000). *Temperature and Precipitation trends in Canada during 20th century.* Atmos.Ocean, 38(3).

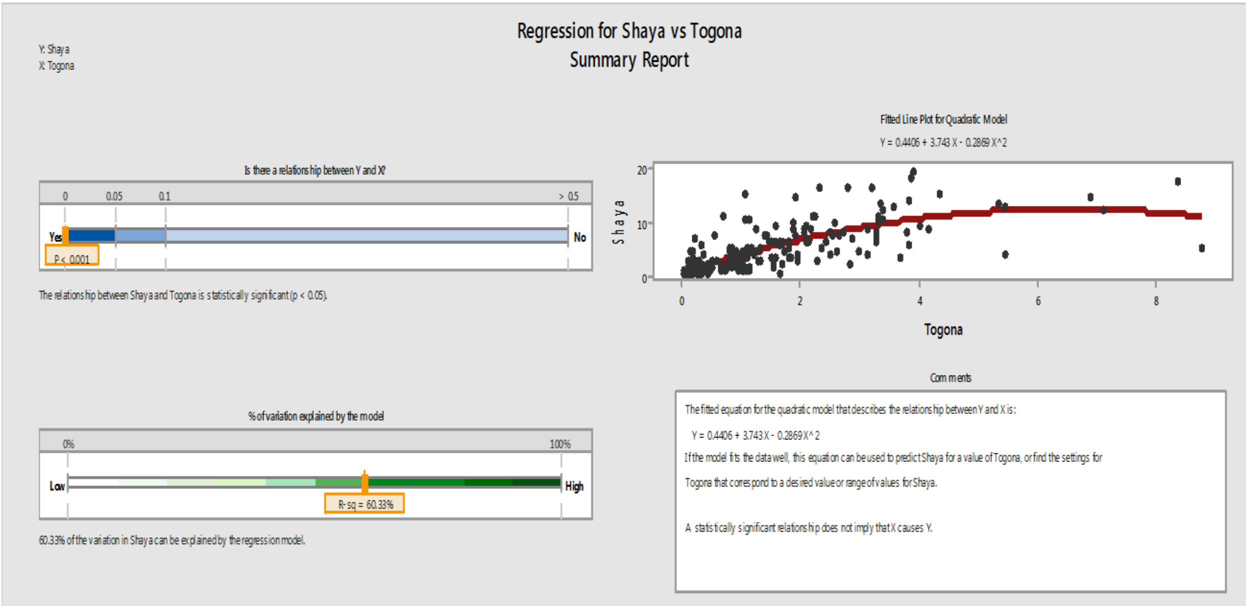
APPENDIX

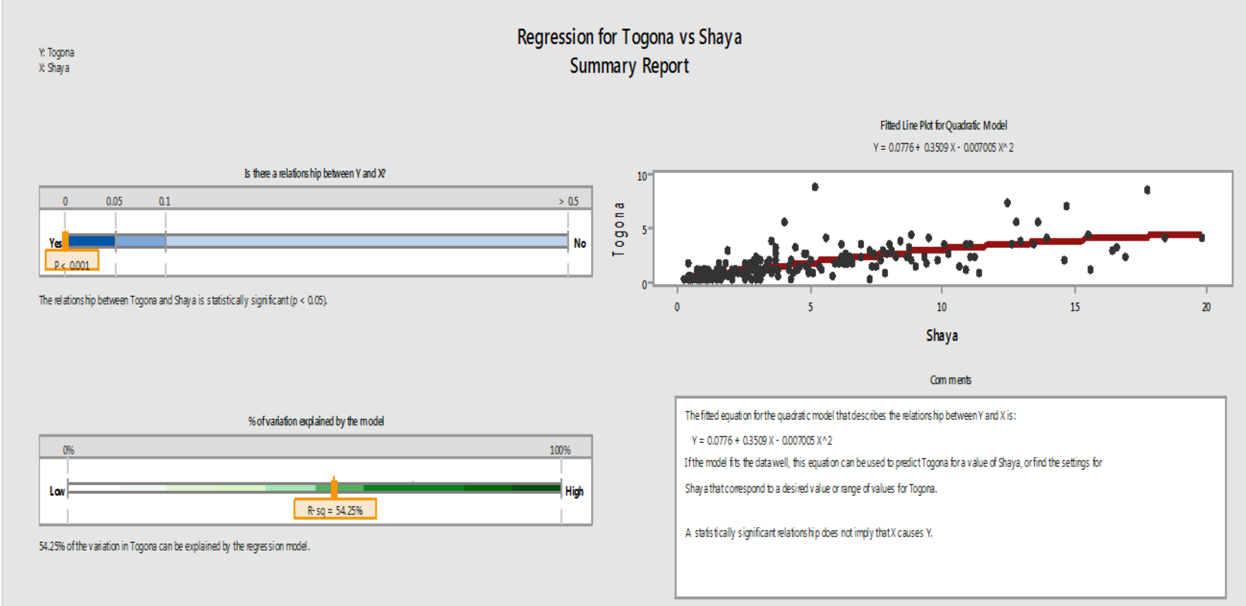
Appendix-I: Regression result of gauging stations



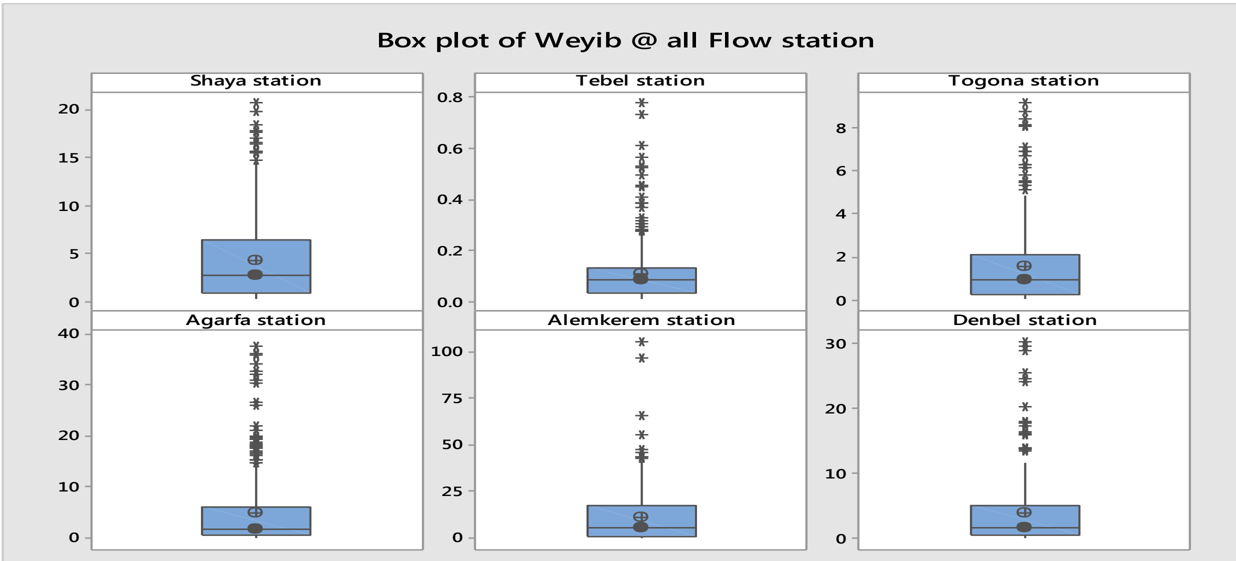


Regression for Agarfa vs Denbel
Summary Report
Y: Agarfa
X: Denbel
Is there a relationship between Y and X?
0
0.05
0.1
> 0.5
Yes
No
P < 0.001
The relationship between Agarfa and Denbel is statistically significant (p < 0.05).
% of variation explained by the model
0%
100%
Low
High
R-sq = 70.60%
70.60% of the variation in Agarfa can be explained by the regression model.
Fitted Line Plot for Quadratic Model
Y = 1.171 + 0.6806 X + 0.01326 X^2
Agarfa
Denbel
Comments
The fitted equation for the quadratic model that describes the relationship between Y and X is:
Y = 1.171 + 0.6806 X + 0.01326 X^2
If the model fits the data well, this equation can be used to predict Agarfa for a value of Denbel, or find the settings for Denbel that correspond to a desired value or range of values for Agarfa.
A statistically significant relationship does not imply that X causes Y.
Regression for Denbel vs Agarfa
Summary Report
Y: Denbel
X: Agarfa
Is there a relationship between Y and X?
0
0.05
0.1
> 0.5
Yes
No
P < 0.001
The relationship between Denbel and Agarfa is statistically significant (p < 0.05).
% of variation explained by the model
0%
100%
Low
High
R-sq = 70.41%
70.41% of the variation in Denbel can be explained by the regression model.
Fitted Line Plot for Quadratic Model
Y = 1.218 + 0.5233 X + 0.007534 X^2
Denbel
Agarfa
Comments
The fitted equation for the quadratic model that describes the relationship between Y and X is:
Y = 1.218 + 0.5233 X + 0.007534 X^2
If the model fits the data well, this equation can be used to predict Denbel for a value of Agarfa, or find the settings for Agarfa that correspond to a desired value or range of values for Denbel.
A statistically significant relationship does not imply that X causes Y.


Regression for Shaya vs Togona
Summary Report
Y: Shaya
X: Togona
Is there a relationship between Y and X?
0
0.05
0.1
> 0.5
Yes
No
P < 0.001
The relationship between Shaya and Togona is statistically significant (p < 0.05).
% of variation explained by the model
0%
100%
Low
High
R-sq = 60.33%
60.33% of the variation in Shaya can be explained by the regression model.
Fitted Line Plot for Quadratic Model
Y = 0.4406 + 3.743 X - 0.2869 X^2
Shaya
Togona
Comments
The fitted equation for the quadratic model that describes the relationship between Y and X is:
Y = 0.4406 + 3.743 X - 0.2869 X^2
If the model fits the data well, this equation can be used to predict Shaya for a value of Togona, or find the settings for Togona that correspond to a desired value or range of values for Shaya.
A statistically significant relationship does not imply that X causes Y.


Regression for Togona vs Shaya
Summary Report
Y: Togona
X: Shaya
Is there a relationship between Y and X?
0
0.05
0.1
> 0.5
Yes
No
P < 0.001
The relationship between Togona and Shaya is statistically significant (p < 0.05).
% of variation explained by the model
0%
100%
Low
High
R-sq = 54.25%
54.25% of the variation in Togona can be explained by the regression model.
Fitted Line Plot for Quadratic Model
Y = 0.0776 + 0.3509 X - 0.007005 X^2
Togona
Shaya
Comments
The fitted equation for the quadratic model that describes the relationship between Y and X is:
Y = 0.0776 + 0.3509 X - 0.007005 X^2
If the model fits the data well, this equation can be used to predict Togona for a value of Shaya, or find the settings for Shaya that correspond to a desired value or range of values for Togona.
A statistically significant relationship does not imply that X causes Y.

Appendix-II: Out layer Test result of gauging stations



Appendix-III: Time series modeling result

The result of the modeling process was presented in bold.

ARIMA Model: Togona

Estimates at each iteration

Iteration	SSE	Parameters			
0	1080.34	0.100	0.100	0.100	0.082
1	882.18	0.030	0.170	0.250	0.021
2	845.25	0.152	0.320	0.283	0.014
3	803.62	0.269	0.470	0.318	0.009
4	754.88	0.377	0.620	0.359	0.005
5	691.42	0.467	0.770	0.417	0.002
6	601.19	0.517	0.920	0.514	0.000
7	546.16	0.484	0.970	0.589	0.001
8	523.85	0.459	0.984	0.626	0.001
9	474.02	0.325	0.984	0.776	0.001
10	468.73	0.296	0.984	0.820	0.001
11	468.23	0.281	0.984	0.830	0.001
12	468.13	0.274	0.983	0.832	0.001
13	468.04	0.271	0.983	0.833	0.001
14	467.93	0.268	0.983	0.834	0.001
15	467.82	0.267	0.983	0.835	0.001
16	467.70	0.267	0.983	0.835	0.001
17	467.58	0.266	0.983	0.836	0.001
18	467.47	0.266	0.982	0.837	0.001
19	467.36	0.265	0.982	0.838	0.001
20	467.24	0.265	0.982	0.839	0.001
21	467.13	0.264	0.982	0.839	0.001
22	467.01	0.264	0.982	0.840	0.001
23	466.89	0.264	0.981	0.841	0.001
24	466.78	0.263	0.981	0.842	0.001
25	466.67	0.263	0.981	0.843	0.001

** Convergence criterion not met after 25 iterations **

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	1	0.2627	0.0584	4.50	0.000
MA	1	0.9809	0.0001	7429.06	0.000
SMA	12	0.8432	0.0441	19.12	0.000
Constant		0.0009315	0.0004295	2.17	0.031

Differencing: 1 regular, 1 seasonal of order 12
Number of observations: Original series 300, after differencing 287
Residuals: SS = 463.484(back forecasts excluded)
MS = 1.638 DF = 283

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	7.0	14.5	26.6	41.5
DF	8	20	32	44
P-Value	0.538	0.803	0.739	0.581

Forecasts from period 300

		95% Limits	
Period	Forecast	Lower	Upper
301	1.17438	-1.33443	3.68319
302	1.35689	-1.24964	3.96343
303	1.50914	-1.10785	4.12614
304	3.59944	0.98016	6.21872
305	3.50373	0.88333	6.12413
306	2.30547	-0.31581	4.92675
307	3.87942	1.25731	6.50153
308	6.42764	3.80472	9.05056
309	5.07873	2.45500	7.70245
310	6.31825	3.69372	8.94278
311	2.84279	0.21745	5.46813
312	2.28256	-0.34358	4.90870
313	1.81302	-0.85284	4.47888
314	1.73113	-0.94052	4.40279
315	1.81485	-0.85874	4.48844
316	3.88807	1.21322	6.56292
317	3.78881	1.11285	6.46477
318	2.59055	-0.08648	5.26758
319	4.16543	1.48734	6.84351
320	6.71482	4.03568	9.39397
321	5.36715	2.68695	8.04735
322	6.60793	3.92668	9.28919
323	3.13374	0.45143	5.81605
324	2.57477	-0.10860	5.25813
325	2.10649	-0.61749	4.83047

ARIMA Model: Tebel

Estimates at each iteration

Iteration	SSE	Parameters			
0	7.21498	0.100	0.100	0.100	0.100
1	4.76068	0.040	-0.050	0.159	0.250
2	4.47557	0.165	-0.030	0.309	0.313
3	4.16697	0.279	-0.078	0.459	0.311
4	3.76599	0.381	-0.064	0.609	0.395
5	3.25113	0.450	-0.070	0.759	0.494
6	2.63024	0.462	-0.124	0.909	0.602
7	2.28568	0.393	-0.189	0.976	0.664
8	2.13568	0.243	-0.264	0.975	0.721
9	2.11829	0.193	-0.285	0.973	0.751
10	2.11665	0.188	-0.289	0.974	0.762

```
Unable to reduce sum of squares any further
```

```
Final Estimates of Parameters
Type        Coef  SECoef       T      P
AR   1    0.1880   0.0588    3.20  0.002
SAR  12  -0.2885   0.0738   -3.91  0.000
MA   1    0.9738   0.0021  470.23  0.000
SMA  12   0.7623   0.0571   13.36  0.000
Differencing: 1 regular, 1 seasonal of order 12
Number of observations: Original series 300, after differencing 287
Residuals: SS = 2.09794 (back forecasts excluded)
           MS = 0.00741 DF = 283
Modified Box-Pierce (Ljung-Box) Chi-Square statistic
Lag            12     24     36     48
Chi-Square    6.0   17.0   24.9   38.5
DF              8     20     32     44
P-Value      0.647  0.655  0.809  0.707
Forecasts from period 300
                      95% Limits
Period Forecast  Lower     Upper
301  0.095906  -0.072884  0.264696
302  0.110876  -0.061744  0.283497
303  0.134298  -0.038687  0.307283
304  0.435042   0.261933  0.608150
305  0.338814   0.165613  0.512015
306  0.131540  -0.041747  0.304828
307  0.118880  -0.054493  0.292254
308  0.114686  -0.058773  0.288145
309  0.121284  -0.052260  0.294829
310  0.342182   0.168552  0.515812
311  0.165720  -0.007995  0.339436
312  0.244577   0.070775  0.418378
313  0.123583  -0.050246  0.297412
314  0.114588  -0.059279  0.288455
315  0.132288  -0.041648  0.306223
316  0.401017   0.227007  0.575028
317  0.312292   0.138205  0.486379
318  0.142596  -0.031568  0.316760
319  0.120602  -0.053638  0.294843
320  0.107519  -0.066799  0.281836
321  0.119715  -0.054679  0.294109
322  0.381327   0.206856  0.555798
323  0.183646   0.009099  0.358194
324  0.189542   0.014918  0.364167
325  0.122766  -0.058275  0.303807
```

ARIMA Model: Denbel

```
Estimates at each iteration

        Iteration      SSE           Parameters
        0  12403.0   0.100   0.100  0.100  0.100
        1   8229.5  -0.026  -0.050  0.227  0.250
        2   7588.4   0.034   0.028  0.351  0.400
        3   6891.3   0.106   0.092  0.501  0.549
        4   6307.8   0.185   0.121  0.651  0.653
        5   5779.2   0.263   0.133  0.801  0.730
        6   5189.3   0.321   0.137  0.951  0.811
        7   5027.3   0.300   0.141  0.980  0.847
        8   4999.7   0.234   0.157  0.978  0.923
        9   4942.1   0.209   0.122  0.972  0.926
       10   4910.7   0.204   0.122  0.964  0.931
```

```
     11   4902.1   0.200   0.122  0.952  0.936
     12   4902.1   0.201   0.123  0.952  0.936
     13   4902.1   0.201   0.123  0.952  0.936
     14   4902.1   0.201   0.123  0.952  0.936
     15   4902.1   0.202   0.124  0.951  0.935
     16   4902.1   0.202   0.124  0.951  0.935
     17   4902.1   0.203   0.124  0.951  0.935
     18   4902.1   0.204   0.125  0.951  0.935
     19   4902.1   0.204   0.125  0.952  0.935
```

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type		Coef	SECoef	T	P
AR	1	0.2042	0.0659	3.10	0.002
SAR	12	0.1249	0.0740	1.69	0.093
MA	1	0.9516	0.0105	90.55	0.000
SMA	12	0.9347	0.0430	21.75	0.000

Differencing: 1 regular, 1 seasonal of order 12
Number of observations: Original series 252, after differencing 239
Residuals: SS = 4839.03 (back forecasts excluded)
MS = 20.59 DF = 235
Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	21.6	31.2	44.0	52.3
DF	8	20	32	44
P-Value	0.006	0.053	0.077	0.184

Forecasts from period 252

		95% Limits	
Period	Forecast	Lower	Upper
253	1.7162	-7.1797	10.6120
254	0.9066	-8.2687	10.0820
255	0.5864	-8.6319	9.8048
256	3.3977	-5.8410	12.6364
257	3.2318	-6.0236	12.4872
258	1.1045	-8.1670	10.3759
259	4.6464	-4.6409	13.9336
260	10.6606	1.3575	19.9636
261	5.2162	-4.1026	14.5350
262	7.4315	-1.9030	16.7660
263	4.0543	-5.2959	13.4045
264	2.0518	-7.3140	11.4177
265	0.9221	-8.7064	10.5505
266	0.8849	-8.7921	10.5620
267	0.6968	-9.0063	10.3998
268	3.3978	-6.3276	13.1232
269	3.2358	-6.5111	12.9827
270	1.2408	-8.5274	11.0090
271	4.7357	-5.0537	14.5252
272	11.7112	1.9006	21.5219
273	5.0119	-4.8198	14.8437
274	8.0789	-1.7739	17.9318
275	4.1248	-5.7492	13.9987
276	1.6154	-8.2795	11.5104
277	0.9503	-9.0483	10.9490

ARIMA Model: Agarfa

Estimates at each iteration

Iteration	SSE	Parameters							
0	29225.5	0.100	0.100	0.100	0.100	0.100	0.100	0.100	0.100
1	26418.6	0.195	0.074	0.092	0.250	0.131	0.105	0.126	0.108
2	24178.5	0.277	0.102	0.229	0.384	0.146	0.103	0.197	0.258
3	21208.7	0.350	0.029	0.303	0.534	0.155	0.096	0.190	0.357
4	18789.9	0.368	0.113	0.287	0.626	0.153	0.087	0.340	0.347
5	17303.6	0.370	0.213	0.234	0.678	0.147	0.080	0.490	0.278
6	14349.5	0.314	0.138	0.234	0.743	0.144	0.079	0.536	0.293
7	12205.6	0.176	-0.012	0.235	0.727	0.158	0.090	0.535	0.329
8	11184.7	0.078	-0.002	0.160	0.722	0.164	0.084	0.685	0.217
9	10757.6	0.040	-0.152	0.150	0.723	0.168	0.078	0.619	0.271
10	10630.9	0.022	-0.302	0.152	0.720	0.173	0.077	0.501	0.380
11	10608.6	-0.019	-0.295	0.159	0.714	0.175	0.078	0.550	0.359
12	10479.7	-0.027	-0.344	0.117	0.712	0.179	0.071	0.562	0.339
13	10444.0	-0.061	-0.494	0.109	0.677	0.213	0.076	0.425	0.475
14	10412.1	-0.087	-0.644	0.096	0.655	0.234	0.080	0.282	0.611
15	10397.2	-0.135	-0.696	0.089	0.614	0.270	0.088	0.237	0.661
16	10379.5	-0.179	-0.725	0.085	0.577	0.305	0.093	0.213	0.691
17	10372.5	-0.198	-0.698	0.088	0.559	0.315	0.097	0.246	0.673
18	10312.8	-0.279	-0.750	0.083	0.482	0.384	0.114	0.193	0.723
19	10309.7	-0.286	-0.751	0.083	0.481	0.384	0.114	0.196	0.719
20	10300.9	-0.296	-0.752	0.085	0.476	0.388	0.116	0.195	0.727
21	10267.2	-0.325	-0.755	0.085	0.440	0.418	0.128	0.192	0.741
22	10223.3	-0.362	-0.775	0.078	0.410	0.446	0.138	0.172	0.758
23	10222.7	-0.357	-0.772	0.074	0.414	0.452	0.143	0.171	0.759
24	10222.1	-0.357	-0.772	0.074	0.414	0.452	0.143	0.171	0.759

Relative change in each estimate less than 0.0010

Final Estimates of Parameters

Type		Coef	SECoef	T	P
AR	**1**	**-0.3572**	**0.0774**	**-4.61**	**0.000**
SAR	**12**	**-0.7719**	**0.2086**	**-3.70**	**0.000**

SAR	**24**	**0.0737**	**0.0685**	**1.08**	**0.283**
MA	**1**	**0.4138**	**0.0589**	**7.02**	**0.000**
MA	**2**	**0.4520**	**0.0440**	**10.28**	**0.000**
MA	**3**	**0.1429**	**0.0303**	**4.71**	**0.000**
SMA	**12**	**0.1714**	**0.2009**	**0.85**	**0.394**
SMA	**24**	**0.7594**	**0.1865**	**4.07**	**0.000**

Differencing: 1 regular, 1 seasonal of order 12
Number of observations: Original series 336, after differencing 323
Residuals: SS = 9187.69 (back forecasts excluded)
MS = 29.17 DF = 315

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	**12**	**24**	**36**	**48**
Chi-Square	**16.1**	**22.5**	**30.8**	**40.7**
DF	**4**	**16**	**28**	**40**
P-Value	**0.003**	**0.129**	**0.326**	**0.439**

Forecasts from period 336

		95% Limits	
Period	Forecast	Lower	Upper
337	0.0297	-10.5578	10.6172
338	0.0319	-10.8933	10.8295
339	0.3065	-10.5691	11.1821
340	4.2798	-6.5997	15.1592
341	4.0831	-6.7963	14.9626
342	1.0848	-9.7951	11.9647
343	6.0199	-4.8601	16.8999
344	17.1438	6.2635	28.0240
345	5.5237	-5.3567	16.4042
346	8.8930	-1.9877	19.7736
347	4.5162	-6.3647	15.3970
348	0.7159	-10.1652	11.5970
349	-0.3328	-11.2269	10.5613
350	0.3332	-10.5611	11.2275
351	0.0465	-10.8479	10.9409
352	4.7417	-6.1530	15.6364
353	3.3349	-7.5600	14.2299
354	0.9780	-9.9172	11.8732
355	5.5709	-5.3245	16.4663
356	14.7818	3.8861	25.6774
357	5.6697	-5.2261	16.5656
358	10.1990	-0.6972	21.0951
359	4.2307	-6.6656	15.1271
360	1.4811	-9.4155	12.3777

ARIMA Model: Alemkerem

Estimates at each iteration

Iteration	SSE	Parameters			
0	97901.8	0.100	0.100	0.100	0.100
1	66000.1	0.012	-0.050	0.188	0.250
2	62080.0	0.129	-0.016	0.338	0.329
3	57526.6	0.234	0.010	0.488	0.410
4	52017.5	0.320	0.024	0.638	0.499
5	45245.4	0.368	0.016	0.788	0.604
6	37450.1	0.341	-0.043	0.938	0.735
7	35490.4	0.307	-0.061	0.978	0.778
8	34286.2	0.210	-0.081	0.975	0.854
9	34076.2	0.193	-0.062	0.968	0.923

```
       10  33840.2  0.172  -0.090  0.954  0.932
       11  33823.4  0.166  -0.096  0.943  0.932
       12  33811.0  0.175  -0.094  0.948  0.932
       13  33811.0  0.175  -0.094  0.947  0.932
```

Unable to reduce sum of squares any further

Final Estimates of Parameters

Type		Coef	SE Coef	T	P
AR	**1**	**0.1749**	**0.0631**	**2.77**	**0.006**
SAR	**12**	**-0.0936**	**0.0678**	**-1.38**	**0.169**
MA	**1**	**0.9475**	**0.0112**	**84.44**	**0.000**
SMA	**12**	**0.9321**	**0.0388**	**24.02**	**0.000**

Differencing: 1 regular, 1 seasonal of order 12
Number of observations: Original series 276, after differencing 263
Residuals: SS = 33002.4 (back forecasts excluded)
MS = 127.4 DF = 259

Modified Box-Pierce (Ljung-Box) Chi-Square statistic

Lag	12	24	36	48
Chi-Square	**13.7**	**20.6**	**23.5**	**31.3**
DF	**8**	**20**	**32**	**44**
P-Value	**0.091**	**0.420**	**0.862**	**0.924**

Forecasts from period 276

		95% Limits	
Period	Forecast	Lower	Upper
277	0.0864	-22.0429	22.2156
278	1.3927	-21.3017	24.0871
279	0.8298	-21.9563	23.6159
280	9.6805	-13.1562	32.5173
281	7.6876	-15.1938	30.5690
282	4.4369	-18.4881	27.3618
283	15.2328	-7.7355	38.2010
284	32.5208	9.5095	55.5322
285	16.3625	-6.6919	39.4170
286	24.0313	0.9339	47.1288
287	5.4156	-17.7248	28.5560
288	1.7373	-21.4459	24.9206
289	0.1209	-23.3193	23.0776
290	1.5006	-21.7331	24.7343
291	1.0814	-22.1919	24.3547
292	11.2066	-12.1069	34.5202
293	8.8052	-14.5487	32.1591
294	4.6143	-18.7799	28.0085
295	15.7642	-7.6702	39.1987
296	30.4708	6.9962	53.9454
297	17.5373	-5.9773	41.0520
298	22.7761	-0.7787	46.3308
299	5.6245	-17.9702	29.2192
300	2.3580	-21.2766	25.9926
301	0.2247	-23.5907	24.0400

Appendix-IV: Monthly runoff trend result of Weyib catchment

Z=Value of Mann-Kendall test   *** = 0.001 level of significance; ** = 0.01 level of significance;

Q= Value of Sen's test   * = 0.05 level of significance   + = 0.1 level of significance

Station	Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	Z	2.315	2.189	-0.083	-0.396	-0.563	0.271	1.689	0.605	-0.98	1.022	1.647	1.564
Shaya	Q	0.019	0.011	-0.0003	-0.035	-0.026	0.004	0.078	0.062	-0.093	0.099	0.091	0.028
	signific	*	*					+				+	
	Z	3.2	3.106	2.594	0.771	0.911	1.705	2.172	3.2	1.752	2.779	2.71	3.06
	Q	0.011	0.01	0.01	0.035	0.032	0.049	0.068	0.071	0.068	0.173	0.049	0.029
Togona	signific	**	**	**			+	*	**	+	**	**	**
	Z	3.256	3.555	3.88	2.383	2.103	2.523	4.043	3.577	3.996	3.317	3.366	2.876
	Q	0.004	0.004	0.004	0.008	0.008	0.004	0.004	0.092	0.004	0.014	0.006	0.004
Tebel	signific	**	***	***	*	*	*	***	***	***	***	***	**
	Z	2.648	1.699	-0.099	-1.561	-1.126	1.363	2.391	0.691	1.047	1.245	2.49	2.509
	Q	0.019	0.042	-0.003	-0.04	-0.079	0.019	0.229	0.132	0.101	0.032	0.062	0.028
Agarfa	signific	**	+					*				*	*
	Z	2.989	1.721	0.091	-0.453	-1.419	0.876	0.151	-0.936	0.755	0.815	0.876	2.144
	Q	0.023	0.007	0.001	-0.023	-0.171	0.048	0.018	-0.217	0.078	0.04	0.045	0.033
Denbel	signific	**	+										*
	Z	-0.555	-0.978	-0.555	-0.026	-0.608	-0.185	0.872	-0.132	-0.132	0.343	0.079	0.766
	Q	-0.008	-0.015	-0.011	-0.009	-0.172	-0.006	0.308	-0.039	-0.035	0.15	0.003	0.018
Alemkrem	signific												

Appendix-V: Monthly Precipitation trend result of Weyib catchment

Z= value of mann-kendall test ***** = 0.001 level of significance ** = 0.01 level of significance**

Q= value of sen's test *** = 0.05 level of significance + = 0.1 level of significance**

Station	Parameter	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
	Z	0.531	-1.009	-1.972	-1.156	-0.408	-1.802	-0.374	-0.986	-0.34	-0.612	1.989	-1.429
Dinsho	Q	0	-0.304	-2.32	-1.302	-0.567	-1.633	-0.27	-1.209	-0.414	-0.733	1.422	-0.831
	signific			*			+					*	
	Z	-0.88	-1.262	-0.428	-1.784	0.963	0.856	1.998	1.463	-0.749	-0.214	0.803	0
	Q	-0.143	-0.42	-0.445	-1.835	0.7	0.567	1.997	2.091	-0.554	-0.137	0.671	0
Robe	signific				+			*					
	Z	-0.294	-1.898	-0.187	0.85	-1.224	-0.119	-0.901	-0.544	-2.006	1.02	0.765	0.119
	Q	0	-0.517	-0.222	0.942	-1.195	-0.07	-0.785	-0.633	-2.055	1.16	0.627	0
Sinana	signific		+							*			
	Z	1.018	-1.688	-0.051	-1.7	-1.462	-1.394	-0.612	-0.544	-1.122	-0.068	1.191	-0.275
	Q	0	-0.852	-0.04	-1.45	-2.105	-1.112	-0.475	-0.738	-1.145	-0.238	0.686	0
Agarfa	signific		+		+								
	Z	0.765	-1.231	-0.999	-0.571	-0.214	-1.142	1.035	0.321	-1.677	-0.678	1.467	0.402
	Q	0	0	-1.173	-0.769	-0.231	-1.191	1.237	0.407	-0.92	-0.743	0.841	0
Gasera	signific									+			
	Z	0.045	-2.074	-0.927	0.707	2.315	0.177	0.353	1.745	2.074	0.75	1.61	0.111
	Q	0	-0.543	-1.66	1.143	4.181	0.089	0.18	1.274	2.185	0.956	1.745	0
Ginir	signific		*			*			+	*			

Appendix-VI: Annual and Seasonal significant trend Weyib at different station