Addis Ababa University
College of Natural and Computational Sciences
Department of Statistics

# Univariate Time Series Analysis of Rainfall in Somali Region, Ethiopia

Abdulahi Abdirashid

A Thesis Submitted to the Department of Statistics in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics (Applied Statistics)

June, 2020
Addis Ababa, Ethiopia

**Addis Ababa University**
**College of Natural and Computational Sciences**
**Department of Statistics**

# Univariate Time Series Analysis of Rainfall in Somali Region, Ethiopia

**Abdulahi Abdirashid**

**Advisor: M.K. Sharma (Prof)**

**A Thesis Submitted to the Department of Statistics in Partial Fulfillment of the Requirements for the Degree of Master of Science in Statistics**

**June, 2020**
**Addis Ababa, Ethiopia**

**Addis Ababa University**

**School of Graduate Studies**

This is to certify that the thesis prepared by Abdulahi Abdirashid, entitled: **Univariate Time Series Analysis of Somali Region, Ethiopia** and submitted in partial fulfillment of the requirements for the Degree of Master of Science in Statistics (Applied Statistics), complies with the University's regulations and meets the accepted standards with respect to originality and quality.

Approved by Examining Board:

Bedilu Alamirie (PhD)

| | | |
|---|---|---|
| Examiner | Signature | Date |

Dejen Tesfaw (PhD)

| | | |
|---|---|---|
| Examiner | Signature | Date |

Merga Belina (PhD)

| | | |
|---|---|---|
| Chairman | Signature | Date |

# Declaration

I, the under signed, declare that this thesis entitled "Univariate Time Series Analysis of Rainfall in Somali Region, Ethiopia" and the information provided in it are correct, valid, and my own. All the sources of material used in it have been duly acknowledged.

Abdulahi Abdirashid

__________________ __________________ ______________

Signature Date

Confirmed by Advisor:
Prof. M.K. Sharma
Name

M.K. Sharma
Signature

07/06/2020
Date

Addis Ababa University

Univariate Time Series Analysis of Rainfall in Somali Region, Ethiopia

Abdulahi Abdirashid

# Abstract

Rainfall is one of the most significant sources of water on earth, supporting the existence of the majority of living things. It is such a critical variable, particularly those whose livelihoods depend on rain-fed agriculture. Statistical approaches such as time series analysis give a way for analyzing and forecasting the patterns of rainfall. In this study, we have attempted to analyze the pattern of rainfall in the Somali region. For this purpose, a monthly rainfall data of 26 meteorological stations covering the region for 34 years of gridded $0.4^o$ *$0.4^o$ resolution have been used. A univariate Box-Jenkins approach was used to build a seasonal ARIMA model to determine the rainfall pattern of the region. Descriptive results show that, during the years in this study, the mean annual and monthly rainfall series in the region are 567mm and 34.9mm respectively, with a higher year-to-year variation (CV $>$ 0.30). It also shows that the region has two rainy periods April-May and August-October, followed by a prolonged dry season. Using Box-Jenkins approach we have obtained a seasonal ARIMA $(1,0,1)(2,1,2)_{12}$ model and then used to establish an 80% and 95% prediction interval to forecast a 5 year monthly rainfall values. The forecasting results revealed a very slight decrease of monthly rainfall over the forecast period from January 2017 to December 2021.

# Acknowledgements

First and foremost, I would like to thank the Almighty Allah for giving me the strength and courage throughout my studies. I wish to express my appreciation to my advisor, Professor M.K Sharma, for his guidance and support. I would also like to extend my heartfelt gratitude to the National Meteorological Agency for its contribution and providing all the relevant data used in this study. I am very much grateful to Professor Eshetu Wencheko for his valuable feedback, encouragement, and friendship throughout this process. My warmest thanks go to my ever-loving family who was always there for giving me the courage to pursue my goals and provide any kind of support. I am whole-heartedly appreciated to my friends for offering me love and support which has been invaluable for this journey. Finally, I wish to thank all the people whose assistance was a milestone in the completion of this thesis.

# Contents

# List of Figures

# List of Tables

# List of Abbreviations

| | |
|---|---|
| **ACF** | Autocorrelation Function |
| **ADF** | Augmented Dickey-Fuller |
| **AIC** | Akaike information criteria |
| **AR** | Autoregressive |
| **ARIMA** | Autoregressive Integrated Moving Average |
| **ARMA** | Autoregressive Moving Average |
| **ASAL** | Arid and Sami-Arid Lowlands |
| **FAO** | Food and Agriculture Organization |
| **GDP** | Growth Domestic Product |
| **MA** | Moving Average |
| **NAPA** | National Adaptation Programme of Action |
| **NMA** | National Meteorological Agency |
| **NMSA** | National Meteorological service Agency |
| **PACF** | Partial Autocorrelation Function |
| **SARIMA** | Seasonal Autoregressive Integrated Moving Average |
| **SBC** | Schwartz's Bayesian Criterion |
| **SRS** | Somali Regional State |

# Chapter 1

# Introduction

## 1.1 Background of the Study

Ethiopia is located in the horn of Africa covering an area of about 1.2 million square kilometers. It is a mountainous country with a rich diversity in climate, biological resources, ethnicity and culture. Its climate varies from the types of hot and arid to cold and humid. The country is also bestowed with rich water resources compared to most African countries (NMA 2007). It is one of the largest countries of Africa which is characterized by a wide variety of landscapes, with an altitudinal range between 210m below sea level and over 4500m above sea level (Fazzini et al., 2015). Agriculture plays a dominant role in the economy of the country, contributing 41% GDP, 80% of the employment and the majority of foreign exchange earnings. However, the agriculture of the country is mostly rain-fed, and it is characterized by extreme dependence on rainfall (Gebreegziabher et al., 2011).

The climate of Ethiopia is highly characterized by rainfall variation. Such climate conditions have caused major constraints to agricultural development (Admassu,2004). Not only in Ethiopia but also in many other regions around the world, the effects of climate change on various environmental variables have been widely observed. Among these variables, rainfall is the most concerned variable affected by climate-change due to its non-homogeneous distributions in time and space (Nirmala and Sundaram,2010). Thus, the variability of rainfall as a result of climate change has been threatening the already fragile rain-fed agriculture and livelihoods of pastoral communities in Ethiopia (Kedir and Tekalign, 2016).

Climate variability and extreme events such as; drought and heavy rains are causing significant damage to life, property, natural resources, and the country's economy, making the most important economic systems and regions of the country highly vulnerable (Kindie et al., 2016). Among the most vulnerable regions in Ethiopia are the arid and semi-arid lowlands (ASAL), which are located in the peripheral border regions of the country below 1,500 m. Large parts of the population in ASAL are mobile pastoralists and with a major cause of rainfall variability, their livelihood is adapted to deal with droughts and floods, which seem to be occurring with increased severity (Rettberg et al., 2017).

Despite the fact that; rainfall is the most crucial and key variable both in the atmospheric and hydrological cycle, its patterns usually have spatial and temporal variability. This variability is assumed to be the main cause of the frequently occurring climate extreme events such as drought and flood (Takele and Gebretsadik,2015). Droughts can vary in intensity, but ASAL regions are stranger to devastating conditions brought on by climate changes and extreme events. More than 18 famine periods were registered in the region's history, in the middle of 1900 and 2011. One of those registered famine periods took place in 1985, which was a highly destructive drought in the area killed nearly 1 million people and in the last decade major droughts have occurred in 2001, 2003, 2005/06, 2008/09, 2011 and 2015/2016 (Griffin and Collins, 2018).

There are many causes of rural poverty including wide fluctuations in agricultural production as a result of drought (Namara et al., 2008). However, examining the spatio-temporal dynamics of meteorological variables in the context of the climate change, particularly in a country like Ethiopia where rain-fed agriculture is predominant, is vital to assess climate-induced changes and suggest feasible adaptation strategies (Asfaw et al., 2018).

## 1.2 Statement of the problem

Rainfall variability has historically been a major cause of droughts, food insecurity and famines in Ethiopia (Bewket, 2009). Thus, the changing rainfall pattern in combination with warming trends could make rain-fed agriculture risky and aggravate food insecurity in the country (Tesfaye et al., 2016). Events such as negative rainfall trends often signify higher probabilities of droughts that have historically affected millions of rural poor farmers and pastoralists and have grave ramifications for both agricultural and environmental issues (Cheung et al., 2008).

In terms of livelihood, smallholder rain-fed subsistence farmers and pastoralists are considered to be the most vulnerable to climate change and its variability, and need interventions to adapt their livelihood systems to the changing climatic conditions (Asfaw et al., 2018). In the case of Somali region, pastoralist and agro-pastoralist livelihood systems are becoming increasingly susceptible to climate change, and irregularity in rainfall patterns affect agricultural production. This has led to extensive food insecurity, drought and malnutrition in the region (SRS, 2016).

Generally, only few studies of rainfall characteristics in the arid and semi-arid regions have ever been conducted. Where as, the Ethiopian arid and semi-arid lowland regions are no exception with almost no study to characterize the rainfall pattern in this area (Tilahun, 2006). Therefore, there is a need for studying rainfall series in detail. For this, we utilize the study from the arid and semi-arid lowlands of Somali Region.

## 1.3 Objective of the Study

**General Objective**

The general objective of this study was to analyze monthly rainfall in Somali Region based on $0.4^o$ x $0.4^o$ resolution gridded data.

**Specific Objectives**

- To characterize monthly rainfall distribution in Somali Region
- To analyze the variability of rainfall in Somali region
- To build a time series model for the monthly rainfall data
- To forecast the rainfall pattern in the study area

## 1.4 Significance of the Study

The findings of this study will be beneficial to the federal government, regional government and regional community. The study will also be of a great assistance to other researchers that intend to carry out similar research on the study topic as it will disclose ways of analyzing the regional data of the last 34 years and also forecasts the following year's expected rainfall in the region from the rainfall data of the previous years. It will also contribute to the body of the existing rainfall literature of Somali region in particular and Ethiopia in general.

# Chapter 2

# Literature Review

## 2.1 Theoretical Literature Review

Climate is often described by the statistical interpretation of rainfall and temperature data recorded over a long period of time for a given region or location (NMA 2007). There is evidence that climate has changed and is continuing to change, and it affects regions across the globe. However, people living in poor countries such as rural Ethiopia are more affected than any other part of the country. Thus, the poor communities in rural Ethiopia are particularly at risk of feeling the effects of climate change, because of their already vulnerable state.

Climate change and the livelihoods of pastoralists are interlinked. Georgis (2010) stated that climate change mostly affects the dry lands and pastoral livelihoods, because of their highly dependence on livestock, which in turn depends on forage and water. Low precipitation and high temperatures affect availability of forage and water supply for livestock (Mahoo et al., 2013).

Climate change has both direct and indirect impacts on agriculture and pastoral livelihoods including reduced or increased rainfall, resulting from change in rainfall patterns (Mahoo et al., 2013). Thus, three of the resulting changes of the climate are changes in rainfall, temperature and extreme events such as droughts and floods.

### Characteristics of Rainfall

Rainfall varies with elevation, latitude, topography, seasons, distance from the sea, and coastal sea-surface temperature. It occurs in different seasons during a year, unlike most of the tropics where two seasons are common (one wet season and one dry season), three seasons are known in Ethiopia, namely Bega (dry season) which increases from October-January, Belg (short rain season) which increases from (February-May), and Kiremt (long rain season) which increases from June-September (NMA 2007).

In the northwestern, southwestern and eastern part of Ethiopia where elevation is up to 1750m above sea level, climate is dominated by tropical rain with mean annual temperature greater than $18^o$c and mean annual rainfall ranging from 680 mm to 1200 mm. The northeastern and southeastern parts of Ethiopia are dominated by dry climate with mean annual temperature ranging from $27^o$c to $30^o$c and the mean annual rainfall less than 450 mm (NMA, 1996). Usually these parts of the country are characterized by strong winds, high temperature and low relative humidity.

## The Effects of Rainfall

The source of livelihood to the overwhelming majority of Ethiopia's population is agriculture, which is characterized by extreme dependence on rainfall (Tadesse, 2002). Rainfall is generally the single most important limiting factor in both agriculture and pastoral activities in the country. According to Von Braun (1991), for example, a seasonal rainfall decrease of 10% from the long-term average generally translates into a 4.4% decrease in the country's food production. Rainfall in much of the country is often erratic and unreliable; rainfall variability and associated droughts have historically been major causes of food shortages and famines (Wood, 1977; Pankhurst and Johnson, 1988). Even though rainfall variability and drought are not new a phenomenon in Ethiopia, its frequency of occurrence has reportedly increased during the past a few decades (Ketema, 1999).

The amount of rainfall determines the quantity, quality and the spatio-temporal distribution of natural pastures. Though the annual rainfall in the country does not show a clear trend, there are various studies which have indicated a rising trend in the frequency of extreme events over the country. This is expected to have a more negative impact over the sub-humid and semi-arid areas of the country (Jember and Tadege, 2010).

## 2.2 Empirical Literature Review

Numerous studies have investigated rainfall in various regions across the world. Different time series methods are employed to analyze rainfall data in various literatures. Focusing mainly on the pattern and distribution of rainfall with different objectives. A review of the available literature relevant to the scope of the study has been presented with a view to survey the various methodologies employed by the researchers along with their findings.

Moloy et al., (2018) considered monthly rainfall data for the period 1960-2016 to examine rainfall in Bangladesh.They suggested that seasonal ARIMA $(0,0,0)(2,1,1)_{12}$ model is the best to predict monthly rainfall data. And they used this model to forecast 120 months (January 2017 - December 2022) seasonal rainfall. Graham and Mishra (2017) described the Box-Jenkins time series seasonal ARIMA approach for prediction of rainfall on monthly scales. Seasonal ARIMA model $(0,0,0)(0,1,0)_{12}$ for rainfall was identified as the best model to forecast rainfall for subsequent 5 years with confidence level of 95% by analyzing last 31 year's data (1985-20015).

Papalaskaris et al., (2016) employed Seasonal ARIMA model, for the period 2006-2014 and concluded that seasonal ARIMA $(0,0,0)(0,1,1)_{12}$ model best fits the total recorded monthly rainfall data to perform short term forecasts of monthly rainfall in Kavala city, Greece. Etuk and Mohamed (2014) analyzed of monthly rainfall data for the Gadaref rainfall station, Sudan. They concluded that the monthly rainfall in Gadaref, Sudan follows a seasonal ARIMA $(0,0,0)(0,0,1)_{12}$ model and it may be used as the basis for forecasting, planning and management of the rainfall in this region.

Momani and Nail (2009) used Box-Jenkins methodology to build ARIMA model for monthly rainfall data taken from Amman airport, Jordan, station for the period from 1922-1999 with a total of 936 readings. They developed ARIMA $(1,0,0)(0,1,1)_{12}$ model to forecast monthly rainfall of the upcoming 10 years. They have also revealed that this approach possesses many appealing features for researchers.

Filder et al., (2019) investigated the monthly time series rainfall data of Embu, County in Kenya. They found Seasonal ARIMA $(1,1,1)(0,1,2)_{12}$ model to be suitable for fitting rainfall data in Embu, County. Furthermore, they showed that the model can be used as a potential alternative for the prediction of annual rainfall values.

Berhane et al., (2018) applied the Box-Jenkins approach, to forecast monthly rainfall of Ethiopia for the period of twelve months ahead using monthly rainfall data from 1901 to 2015. They identified seasonal ARIMA $(1,0,0)(0,1,1)_{12}$ as an appropriate model for forecasting the amount of monthly rainfall of Ethiopia, they also used the model to forecast twelve months ahead. All forecasted values were found to lie within the 95% confidence interval which shows that the model is neither over forecasting nor under forecasting.

Abebe (2018) used a model for annual rainfall series (over the period 1954-2015) of Debre Markos town, Ethiopia. Among three statistically competent ARIMA models, ARMA $(2,1)$ model was selected as the better model. Hence, he concluded that this model can be used to aid the understanding of Debre Markos rainfall process at annual scale and can be used as a potential alternative for prediction of annual rainfall values.

Gebretsadikan and Sharma (2011) attempted to build a seasonal model of monthly rainfall data of Mekele station of Tigray region, Ethiopia, using univariate Box-Jenkins's methodology. The method of estimation and diagnostic analysis results revealed that the forecasting ability of the Seasonal ARIMA $(0,0,1)(1,1,4)_{12}$ model is better for the purpose future monthly rainfall data.

Takele and Gebretsadik (2015) used spectral analysis, cross-spectral analysis and seasonal ARIMA with the objective to predict the pattern and its extreme event frequency of rainfall in Dire Dawa Region, Ethiopia. They established and adequately used Seasonal ARIMA $(5,0,0)(0,1,1)_{12}$ model to forecast 2 years monthly rainfall values. They have also revealed that there is a tendency of relatively increase in the monthly and annual rainfall pattern over the forecast period from January 2014 to January 2016. Takele (2012) also carried out a statistical analysis of rainfall pattern in Dire Dawa, Ethiopia. He used descriptive analysis, spectral analysis and univariate Box-Jenkins method to establish a time series model that he used to forecast two years monthly rainfall. Results showed an extreme event of rainfall occurs every 2.5 years.

NMA (2007) reported that the mean annual rainfall is likely to increase, as a small increase in annual rainfall is also expected in Ethiopia. However, other study made by NMSA (1996) identified that the mean annual rainfall distribution over the country is characterized by large spatial variation which ranges from about 2000 mm over some areas in the Southwest of Ethiopia to less than 250 mm over the Afar and Ogaden lowlands.

Seleshi and Zanke (2004) undertook a study based on 11 key stations in different zones of climatic located in Ethiopia over the common period 1965-2002. They observed changes in annual, June-September and March-May rainfall and rainy days here in (defined as a day with rainfall greater than 1 mm). The annual and the June-September total rainfalls for the eastern (Jijiga, 137 mm per decade), southern (Negele,119 mm per decade) and southwestern (Gore, 257 mm per decade) stations showed significant declines since about 1982. In comparison, the annual and the Kiremt total rainfall in eastern and southern Ethiopia showed significant declines since 1982. The Kiremt rainy days in eastern Ethiopia showed significant declines (at the 5% level) since about 1982 compared with the period 1965-1981.

# Chapter 3

# Data and Methodology

## 3.1 Data Source and Study Area

### Data Source

The data used in this study were collected form National Meteorological Agency of Ethiopia. It is based on a gridded monthly rainfall data at $0.4^o$ x $0.4^o$ resolution for 34 years (1983-2016) covering the entire Somali region.

#### Study Area

Somali Region is the second largest among the nine regions in Ethiopia, which is located in the eastern part of the country. The region lies between 200 meters above sea level in the southern and central parts and 1,800 meters above sea level in Jijiga Zone ($6.6612^o$N and $43.80^o$E). About 86% of the population of the region lives in rural areas, and are mainly pastoralists and, to a lesser extent, agro-pastoral. The vast area, around 80% of the region, is considered lowland, characterized by arid and semi-arid land. Pastoralism is the dominant mode of production as well as way of life for the majority of people, this livestock herding dependent way of life is highly tied to natural resources of pastures and water exploitation to thrive. This makes pastoralism inherently to be sensitive to climate change and extremely exposed to its impacts.

## 3.2 Methodology

Time series is a sequence of dependent observations ordered in time. Mostly these observations are collected at equally spaced, discrete time intervals. When there is only one variable up on which observation are made then we call this a single time series or more specifically a univariate time series. A basic assumption in any time series analysis/modelling is that some aspects of the past pattern will continue to remain in the future.

### 3.2.1 Stationarity

Stationarity plays a crucial role in the analysis of time series data, a series may be stationary or non-stationary. Stationary series are characterized by a kind of statistical equilibrium around a constant mean level as well as a constant dispersion around that mean level (Box and Jenkins, 1976). There are several kinds of stationarity, a series is said to be stationary in the wide sense, weak sense, or second order[1] if it has a fixed mean and a constant variance. A series is said to be strictly stationary if it has, in addition to a fixed mean and constant variance, a constant auto-covariance structure. When a series possesses this covariance stationarity, the covariance structure is stable over time (Diebold, 2001).

According to Box et al. (2015) if a series is non stationary, then it necessitates a differencing to be carried out to transform it to a stationary series in order to continue with the ARIMA modelling. The unit root's presence in the time series data was carried out. The standard tests for unit root is the augmented Dickey-Fuller (ADF) test. This test is based on estimates from an augmented auto-regression. The selection of lag length k is one of the key issues in the ADF test. The presence of a unit root is an indication of non-stationarity of the series. If the series is non-stationary differencing is required to convert it to a stationary data. However, a non-stationary process arises when at least one of the conditions for stationarity does not hold.

#### 3.2.2.1 Unit Root Test

The first step for an appropriate analysis is to determine whether the series is stationary or not. Thus, a test of stationarity (or non-stationarity) that has become widely popular over the past several years is the unit root test. The test first poses the null hypothesis that the given time series has a unit root, which means that the time series is non-stationary, and tests if the null hypothesis is to be statistically rejected in favor of the alternative hypothesis that the given time series is stationary. The widely used unit-root tests are Augmented Dickey Fuller test (ADF) and Phillips Perron (PP) tests. However, in this study we used ADF test to examine the unit-root test of the monthly rainfall series.

[1]It follows that second-order stationarity implies strict stationarity for normal processes

**The Augmented Dickey Fuller (ADF) Test**

Similar to the original Dickey-Fuller test, the augmented Dickey-Fuller test is one that tests for a unit root in a time series sample. It is comparable with the simple DF test, but it is augmented by adding lagged values of the first difference of the dependent variable as additional regressors which are required to account for possible occurrence of autocorrelation. Consider the AR(p) model:

$$x_t = \Phi_1 x_{t-1} + ..... + \Phi_p x_{t-p} w_t \tag{3.1}$$

from equation (3.1) we have:

$$\nabla_{x_t} = \mu + \rho x_{t-1} + \sum_{i=2}^{p} \psi_i \Delta x_{t-p} + w_t \tag{3.2}$$

The test statistics is $t_{DF} = \frac{\widehat{\rho}}{se\widehat{\rho}}$ where $\widehat{\rho}$ is the OLS estimate of $\rho = (\phi - 1)$, $t_{DF}$ has to be compared against the 95% critical value of the appropriate DF distribution, which depends on the inclusion of the linear trend and the lag structure. Then we use the t statistic on the $\rho$ coefficient to test whether we need to difference the data to make it stationary or we need to put a time trend in the regression model to correct for the variables deterministic trend. The null hypothesis for the test is given as:

$$H_0 : \rho = 0 \quad \textit{There exists a unit root problem}$$
$$H_1 : \rho < 0$$

## 3.3 ARIMA Models

Often interesting features of a time series cannot be explained by classical regression. The presence of correlation among observations requires application of stochastic methods which are not based on the assumption of observation independence. Autoregressive (AR), moving average (MA) or mixed autoregressive moving average (ARMA) models are variants of such a stochastic method. The common thing in these three models is that they all acknowledge the presence of correlation and are applicable on a stationary time series. If non-stationarity is an issue, as in most practically encountered time series, another stochastic model called autoregressive integrated moving average (ARIMA) can be used.

### 3.3.1 Autoregressive Models

The central idea of autoregressive models is that the current value of a time series is affected by one or more past values of the same series. The extent to which past values affect the current value can be examined using sample autocorrelation function (ACF). The general equation of an autoregressive model of order p, AR(p), is given as:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + .... + \phi_1 x_{t-1} + w_t \tag{3.3}$$

Where $x_t$ is current value of the series, $\phi_1, \phi_2, ..., \phi_p$ are model coefficients with $\phi_p \neq 0$, $x_{(t-1)}, x_{(t-2)}, ...., x_{(t-p)}$are past values of the series and $w_t$ is white noise term which is a collection of uncorrelated random variables with mean of 0 and finite variance, $\sigma_w^2$. From the equation, one can see that current values are assumed to be combined linearly with past values. In a more concise form the AR(p) equation can be written using a backshift operator B defined as $Bx_t = x_{t-1}$ as follows

$$\phi(B)x_t = w_t \tag{3.4}$$

Where the autoregressive operator $\phi(B)$ is defined as

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - .... - \phi_p B^p \tag{3.5}$$

It can be shown that equation (3.4) has a unique causal stationary solution provided that the roots of $\Phi(B) = 0$ lie outside the unit circle. If this is so, the solution can be expressed in the form

$$x_t = \sum_{j=o}^{\infty} \phi_j w_{t-j} \tag{3.6}$$

For some constants $\phi_j$ such that $\sum_{j=0}^{\infty} \mid \phi_j \mid < \infty$ typically an AR process is stationary provided that the roots of $\phi(B) = 0$ lie outside the unit circle (Chatfield, 2000). A useful property of an AR(p) process is that the partial ACF is zero at all lags greater than p.
Particular equations of the various order p can be derived from the general equation (3.3). For example; AR (1) model can be given as:

$$x_t = \phi x_{t-1} + w_t \tag{3.7}$$

### 3.3.2 Moving Average Models

The moving average model is an alternative to the autoregressive model, it assumes the current value as a linear combination of white noises $w_t$. The general equation of moving average model of order q, MA(q), is given as:

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + .... + \theta_q w_{t-q} \tag{3.8}$$

Where $\theta_1, \theta_2, ..., \theta_q$ are model coefficients with $\theta_q \neq 0$. We may write MA(q) model equation as

$$x_t = \theta(B) w_t \tag{3.9}$$

Where the moving average operator $\theta(B)$ is defined as:

$$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + .... + \theta_q B^q \tag{3.10}$$

If all roots of $\theta(B) = 0$ lie outside the unit circle, the MA process has an autoregressive representation of generally infinite order $w_t = \sum_{j=0}^{\infty} \phi_j x_{t-j}$ with $\sum_{j=0}^{\infty} \mid \phi_j \mid < \infty$
The order of such a model can be determined by analysis of the autocorrelation function, ACF, which cuts off after q lags and partial ACF that decays exponentially fast. Unlike the AR process, the MA process is stationary for any values of the parameters $\theta_1, \theta_2, ..., \theta_q$.
Particular equations of the various order q can be derived from the general equation (3.8). For example; MA (1) model can be given as:

$$x_t = w_t + \theta w_{t-1} \tag{3.11}$$

### 3.3.3 Autoregressive Moving Average Models

For a stationary process, a mixed type of model i.e. autoregressive moving average (ARMA) can be used. This model is a linear combination of AR(p) and MA(q) models and its general equation for orders p and q, ARMA (p,q), is given as:

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + ... + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + ... + \theta_q w_{t-q} \tag{3.12}$$

With $\phi_p \neq 0$ and $\theta_q \neq 0$. The concise form of ARMA (p,q) model is given as:

$$\phi(B) x_t = \theta(B) w_t \tag{3.13}$$

To have ARMA (p,q) model, both ACF and PACF should show a pattern of decaying to zero. The autocorrelation of an ARMA (p, q) process is determined at greater lags by the AR (p) part of the process as the effect of the MA part dies out. Thus, eventually the ACF consists of mixed damped exponentials and sine terms. Similarly, the partial autocorrelation of an ARMA (p, q) process is determined at greater lags by the MA (q) part of the process. Thus, eventually the partial autocorrelation function will also consist of a mixture of damped exponentials and sine waves.

From equation (3.12) above one can notice that when $p = 0$, the model becomes a moving average model of order q, MA(q), and when $q = 0$, the model becomes autoregressive model of order p, AR(p).

### 3.3.4 Autoregressive Integrated Moving Average Models

Yet another model to be used especially when non-stationarity is apparent in the time series is autoregressive integrated moving average (ARIMA) model. In this model the time series is made stationary by applying a certain order of differencing, d. The general equation of this model for orders p, d and q, ARIMA (p,d, q) is given as:

$$\nabla^d x_t = (1 - B)^d x_t \tag{3.14}$$

Where $\nabla^d$ is a d order difference operator is on $x_t$ which is defined as $\nabla x_t = (1-B)x_t$. The model can also be written as:

$$\phi(B)(1 - B)^d x_t = \theta(B)w_t \tag{3.15}$$

From Equation (3.15) we can notice that if $d = 0$, the resulting model is pure ARMA (p,q). Hence ARIMA (p,d,q) can be considered as a general case of ARMA (p,q), MA (q) and AR (p) models.

## 3.4 Seasonal ARIMA (SARIMA) Model

Seasonality usually causes the series to be non-stationary because the average values at some particular times within the seasonal span (months, for example) may be different than the average values at other times. Box et al (1994) have generalized the ARIMA model to deal with seasonality, and define a general multiplicative seasonal ARIMA model, which are commonly known as SARIMA models. In short notation the SARIMA model described as ARIMA (p, d, q)(P, D, Q)s, which is mentioned below:

$$\phi(B)\Phi_P(B^s)\nabla^d\nabla_s^D(x_t) - \mu = \theta(B)\Theta_Q(B^s)w_t \tag{3.16}$$

Where $x_t$ is an observation at time t,t is a discrete time, S is seasonal length, $\mu$ is mean level of the process which is usually taken as the average of the $w_t$ series (often, if D+d>0 then $\mu \equiv 0$), and $w_t \sim NID(0, \sigma_w^2)$, p is the order of non-seasonal AR, d the number of regular differencing, q the order of non-seasonal MA, P the order of seasonal AR, D the number of seasonal differencing, Q the order of seasonal MA, s is the length of season, $x_t$ is an observation at a time t, $\Phi_P$ and $\Theta_Q$ are the seasonal polynomials of order P and Q, respectively.

Hence, in equation (3.16) the ordinary and seasonal difference components can be written as:

$$\nabla^d = (1-B)^d \;\; and \;\; \nabla_s^D = (1-B^s)^D \tag{3.17}$$

Thus, equation (3.16) can be written as:

$$\phi(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D x_t - \mu = \theta(B)\Theta_Q(B^s)w_t \tag{3.18}$$

The operators of the seasonal autoregressive operator and the seasonal moving average operator of orders P and Q, respectively, with seasonal period s are;

$$\Phi_P(B^s) = 1 - \Phi_{1s}B^s - \Phi_{2s}B^{2s} - .... - \Phi_{Ps}B^{Ps} \tag{3.19}$$

and

$$\Theta_Q(B^s) = 1 + \Theta_{1s}B^s + \Theta_{2s}B^{2s} + .... + \Theta_{Qs}B^{Qs} \tag{3.20}$$

Thus, The resulting pure seasonal autoregressive moving average model, say, ARMA(P, Q)$_s$, then takes the form

$$\Phi_P(B^s)x_t = \Theta_Q(B^s)w_t \tag{3.21}$$

## 3.5 Building ARIMA Models

### 3.5.1 Model Identification

Once the series process is made stationary, the next step is to find the appropriate ARIMA form to model the stationary series (i.e. to identify appropriate values of the model; that is p, d, and q.). There are two main approaches to identification of ARIMA models in the literature. These are penalty function criteria and traditional Box-Jenkins procedure, in which traditional Box-Jenkins an iterative process of model identification, model estimation and model evaluation is followed. The Box-Jenkins procedure is a quasi-formal approach with model identification relying on subjective assessment of plots of correlogram and partial auto correlogram of the series.

The Box-Jenkins methodology is not only about model identification but is, in fact, an iterative approach incorporating model estimation and diagnostic checking in addition to model identification. Theoretically the Box-Jenkins model identification is relatively easy if one has a pure AR and pure MA process. However, in the case of mixed ARMA models (especially of higher order) it can be difficult to interpret sample ACFs and PACFs, and Box-Jenkins identification become a highly subjective exercise depending on the skill and experience of the analyst.

**ACF/PACF Plot**

Many researchers noted that; The ACF plot is a spike graph, which is a special type of bar chart, of the ACF as a function over lags. They gave the following description of the basic ideas: The ACF is the correlation between any two values in a time series with a specific time shift, called lag. The PACF is the correlation between any two points with a specific lag, where the linear effects of the points in between is removed. This PACF plot combined with the ACF plot, where this linear dependence is included, is called ACF/PACF plot and enables us to choose the number of parameters for the model.

Using the definitions of the autoregressive models, moving average models, ACF, and PACF it is possible to identify the basic behavior of the ACF and the PACF for AR, MA, and ARMA models. Likewise, it is possible to describe the behavior for the seasonal component of the model in a similar way. The behaviors of the ACF and the PACF indicate which class of model and what number of parameters could be adequate for the non-seasonal and seasonal part of the model. The behavior is shown in Table 3.1 (Shumway and Stoffar, 2017).

Table 3.1: ACF and PACF behavior for ARMA and Seasonal ARMA models

| | AR(p) | MA(q) | ARMA(p,q) |
|---|---|---|---|
| ACF | Tails off | Cuts off after lag q | Tails off |
| PACF | Cuts off after lag p | Tails off | Tails off |
| | $AR(p)_s$ | $MA(q)_s$ | $ARMA(p,q)_s$ |
| ACF* | Tails off at lags $k \cdot s$ | Cuts off after lag $Q_s$ | Tails off at lags $k \cdot s$ |
| PACF* | Cuts off after lag $P_s$ | Tails off at lags $k \cdot s$ | Tails off at lags $k \cdot s$ |

$*Where\ k \cdot s\ \ are\ multiples\ of\ s,\ for\ k = 1, 2, ....$

**Sample Autocorrelation Function (ACF) $\gamma_k$**

$$\gamma_k = \frac{\sum_{i=1}^{n-k}(x_i - \overline{x})(x_{i+k} - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \tag{3.22}$$

*Where $\overline{x}$ is the average of the $n$ observations.*

**Sample Partial Autocorrelation Function (PACF) $\widehat{\phi}_{kk}$**

$$\phi_{kk} = \frac{(\gamma_k - \sum_{j=1}^{k-1}\phi_{k-1,j}\gamma_{k-j})}{(1 - \sum_{j=1}^{k-1}\phi_{k-1,j}\gamma_j)}, k = 3, ... \tag{3.23}$$

**Criteria for Model Selection**

Because economic theory does not provide any guidance to the appropriate choice of model, some additional criteria can be used to choose from alternative models that are acceptable from a statistical point of view. As a more general model will always provide a better fit (within the sample) than a restricted version of it, all such criteria provide a trade-off between goodness-of-fit and the number of parameters used to obtain that fit.

$$AIC = log\widehat{\sigma}^2 + 2\left[\frac{P+q+1}{T}\right] \tag{3.24}$$

Where p is number of autoregressive term, q is number of moving average term and $\widehat{\sigma}^2$ is the estimated variance of $w_t$. An alternative is Schwarz's Bayesian Information Criterion (SC, BIC or SBC), proposed by Schwarz (1978), which is given as:

$$BIC = log\widehat{\sigma}^2 + \left[\frac{P+q+1}{T}\right] logT \tag{3.25}$$

Both criteria are likelihood-based and represent a different trade-off between 'fit', as measured by the log likelihood value, and 'parsimony', as measured by the number of free parameters, $p + q + 1$ (assuming that a constant is included in the model). Usually, the model with the smallest AIC or BIC value is preferred, although one can choose to deviate from this if the differences in criterion values are small for a subset of the models.

### 3.5.2 Model Estimation

Once the model is tentatively established and identified the appropriate (p,d,q) for the non-seasonal values and (P,D,Q) for the seasonal values the next stage is to estimate the parameters of the autoregressive and moving average terms included in the model. This calculation can be done by simple least squares or maximum likelihood estimation.

### 3.5.3 Model Diagnostics

This step will be the formal assessment of each of the time series models. This will involve a rigorous assessment of the diagnostic tests for each of the competing models. As different models may perform reasonably similarly, a number of alternative formulations may have to be retained at this stage to be further assessed at the forecasting stage.

There are a number of diagnostic tools available for ensuring a satisfactory model is arrived at. Plotting the residuals of the estimated model is a useful diagnostic check. This should indicate any outliers that may affect parameter estimates and also point towards any possible autocorrelation or heteroskedasticity problems. If the model is correctly specified the residuals should be 'white noise'. Therefore, a plot of the auto-correlogram should immediately die out from one lag on. Any significant autocorrelations may indicate that the model is misspecified and may point to the solution.

### Autocorrelation Test

Before an estimated model is used for statistical inference (e.g. hypotheses tests, forecasting, etc), the residuals must be examined for the presence of serial correlation.

### Ljung-Box Q-statistic

Is the modification of the Box and Pierce (1970) test. It is often used as a test of whether the residual series is white noise. The Q-statistic, to test the null hypothesis of no autocorrelation up to lag k. The test may be defined as:

> $H_0$: The data are independently distributed (i.e. the correlations in the population from which the sample is taken are 0, so that any observed correlations in the data result from randomness of the sampling process).
>
> $H_1$: The data are not independently distributed; they exhibit serial correlation.

The test statistic is:

$$Q_{LB} = T(T+2)\sum_{i=1}^{K}\left(\frac{\gamma_i^2}{T-i}\right) \tag{3.26}$$

Where $\gamma_i$ is the i-th lag autocorrelation and $T$ is the number of observations. Under the null hypothesis, $Q_{LB}$ is asymptotically distributed as chi-square with degrees of freedom equal to $(k-m)$, where m denotes the number of parameters in the model. The critical region for rejection of the hypothesis of randomness is:

$$Q > X^2_{1-\alpha,h} \tag{3.27}$$

(Ljung and Box, 1978).

### Normality

This displays the frequency distribution of your series in a histogram. The histogram divides the series range (the distance between the maximum and minimum values) into a number of equal length intervals or bins and displays a count of the number of observations that fall into each bin. Complements of standard descriptive statistics are displayed along with the histogram and Q-Q plot of a data set give information related to normality. There are several statistical tests used for the diagnostic checking of normality. In this study, The Shapiro-Wilk statistic is used to test the residuals of the fit for normality.

**Shapiro-Wilk test**

The Shapiro-Wilk test is a way to tell if a random sample comes from a normal distribution. The test gives a W value, small values indicate the sample is not normally distributed (reject the null hypothesis). The formula for the W value is:

$$W = \frac{\sum_{i=1}^{n} a_i x_{(i)}}{\sum_{i=1}^{n} (x_i - \bar{x})} \tag{3.28}$$

$H_0$: *The sample belongs to normal distribution*

$H_1$: *The sample does not belong to normal distribution*

(Shapiro and Wilk, 1965).

**Jarque-Bera Test**

Is used for testing whether the series is normally distributed. It measures the difference in the skewness and kurtosis of the series with those from the normal distribution. The test statistic is always nonnegative. If it is far from zero, it signals the data do not have a normal distribution. The statistic is computed as:

$$JB = \frac{n-k}{6}\left(S^2 - \frac{(K-3)^2}{4}\right) \tag{3.29}$$

Where n is the number of observations (or degrees of freedom in general); S is the sample skewness, K is the sample kurtosis. Under the null hypothesis of a normal distribution, the JB statistic asymptotically has a chi-squared distribution with two degrees of freedom. The reported Probability is the probability that a JB statistic exceeds (in absolute value) the observed value under the null hypothesis - a small probability value leads to the rejection of the null hypothesis of a normal distribution.

### 3.5.4 Forecasting

The last step in time series modeling is forecasting. It is the process of estimation in unknown situations, which is commonly used in discussion of time-series data. There are two kinds of forecasts: sample period forecasts and post-sample period forecasts. The former one is used to develop confidence in the model and the latter to generate genuine desired forecasts. In forecasting, the goal is to predict future values of a time series, $x_{(t+m)}, m = 1, 2, ...$ based on the data collected to the present, $x = x_t, x_{(t-1)}, x_{(t-2)}, ...., x_1$. Throughout this section, we will assume $x_t$ is stationary and the model parameters are known.

## Forecasting Accuracy Measures

The accuracy of a forecasting method is determined by analyzing forecast error experiences. The forecasting performance of the estimators is judged on the basis of the differences between predictions and realizations. The smaller the difference between the predictions and the actual values of the dependent variable, the better is the forecasting performance of the estimator. The within sample forecasting performance of the system should be assessed using standard statistical tools such as Root Mean Square Error, Mean Absolute Error, Mean Absolute Percentage Error and Theil?s inequality Coefficient. The first two forecast error statistics depend on the scale of the dependant variables, and the remaining two statistics are scale invariant (i.e. unit free). In most instances unit free measures are preferable (Challen and Hagger, 1983). As a result Theil's Inequality coefficient (TIC) and Mean Absolute Percentage Error (MAPE) are preferable. If the forecast is good, the Mean Absolute Percentage Error and Theil's Inequality coefficient should be as small as possible.

## Mean Squared Errors (MSE)

The mean squared error is an accuracy measure computed by squaring the individual error for each item in a data set and then finding the average or mean value of the sum of those squares.

$$MSE = \frac{1}{n} \sum_{t=1}^{n} e_t^2 \tag{3.30}$$

Where MSE is mean squared error, n is time periods and $e^2$ is forecast error.

## Mean Absolute Percentage Error (MAPE)

The mean absolute percentage error is the mean or average of the sum of all of the percentage errors for a given data set taken without regard to sign so as to avoid the problem of positive and negative values cancelling one another.

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \mid PE_t \mid \tag{3.31}$$

Where PE is percentage error, $Y_t$ is actual observation for time period t, $F_t$ is forecast for the same period, MAPE is mean absolute percentage error and n is time periods.

## Theil's U-statistic

The U-statistic developed by Theil (1966) is an accuracy measure that emphasizes the importance of large errors (as in MSE) as well as providing a relative basis for comparison with forecasting methods. Theil's equation is written as shown below:

$$U = \sqrt{\frac{\frac{\sum_{n=1}^{t-1} F_{t+1} - Y_{t+1}}{Y_t}}{\sum_{i=0}^{n} \left( \frac{F_{t+1} - Y_{t+1}}{Y_t} \right)^2}} \tag{3.32}$$

Where U is Theil's U-statistic, $F_t$ is forecast value and $Y_t$ is actual value Theil's U-statistic. The scaling of U always lies between 0 and 1. If U=0, there is a perfect fit; if U=1 the predictive performance is as bad as it possibly could be, this means the proposed model is as good as the naive model.

## Root Mean Square Error

The RMSE is one of the most widely used measures of forecast accuracy.

$$RMSE = \sqrt{\frac{1}{n}(F_t - Y_t)^2} \tag{3.33}$$

Where RMSE is mean squared error, n is time periods, $F_t$ is the forecasted value and $Y_t$ is the actual value.

# Chapter 4

# Results

## 4.1 Descriptive Results

Monthly rainfall data for 26 meteorological stations almost covering the whole region were obtained for 34 years from the NMA. Thus, to characterize and analyze th monthly rainfall, preliminary statistical analysis including the measure of central tendency (mean and median) and dispersion (standard deviation, coefficient of variation), were performed for the average monthly rainfall series in Somali region. The statistical software package used for the analysis of this study is R-1.2.5042-1.

Figure 4.1 and Tables 4.1, 4.2 shows the monthly rainfall series distribution and the average monthly rainfall series in Somali Region, respectively. From the figure and the tables, it is illustrated that the region has two rainy periods, April-May and August-October, with an average monthly rainfall of 34.9 mm and a maximum monthly rainfall of 206 mm. Thus, the period December-March was considered to be the dry season in the region.

Table 4.1: The Sum for the Monthly Average Rainfall Distribution from 1983-2016 in mm

| Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 88.7 | 200.6 | 1028.6 | 3187.3 | 2559 | 659 | 982.9 | 1371.3 | 1489.4 | 1724.4 | 785 | 149.9 |

Table 4.2: Summary of rainfall amount in mm (1983 - 2016)

| Rainfall | Mean | Min | Median | Max | SD | CV |
|---|---|---|---|---|---|---|
| Annual | 567 | 213.9 | 648.2 | 771 | 184.15 | 0.32 |
| Monthly | 34.9 | 0 | 26.7 | 206 | 35.4 | 1.01 |

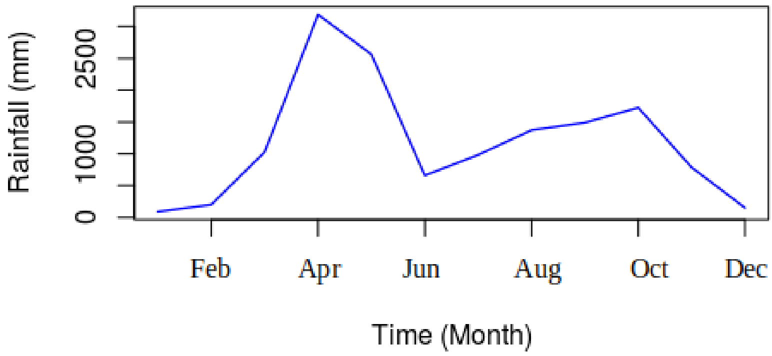


Figure 4.1: Plot for Average Monthly Rainfall Distribution Sum from 1983-2016

NMSA (1996) stated that a series with CV less than 0.20 (20%) can be considered as less variable, while CV between 0.20 (20%) to 0.30 (30%) is moderately variable, and CV greater than 0.30 (30%) is highly variable. Since the CV of a series provides the year-to-year variation in the series, this clearly shows that there is a high monthly rainfall variability in the region.

## 4.2 Tests for Stationarity

**Graphic Inspection**: The graphical inspection of the autocorrelation function plot provides useful information to identify the type of time series. Figure 4.2 and Figure A in the *Appenix* exhibits the time series plot of average monthly rainfall data. The three years' visible plot (2014 - 2016) shows no obvious trend but it rather illustrates a regular pattern of ups and downs, which is an indicator of seasonality, fro this reason, we tried to examine the ACF and PACF of the $12^{th}$ differences (seasonal differencing).

We used a lag of 12 (S=12) for differences because there is seasonality and the data is in months. The series is, therefore, seasonal due to large rainfall values during the rainy season and a relatively lesser peak due to small values of rainfall in the other months. This indicates that the rainfall data have a seasonal unit root (i.e seasonally non-stationary). However, stationarity of the average monthly rainfall series was further proved by the ADF unit root test at $\alpha$= 0.05 (Table 4.3).

### Unit Root Test

The widely used Augmented Dickey-Fuller test (ADF) Unit Root Test, was applied on the entire monthly rainfall data with the view to investigate and verify whether the rainfall series is stationary or not. The test first poses the null hypothesis that the

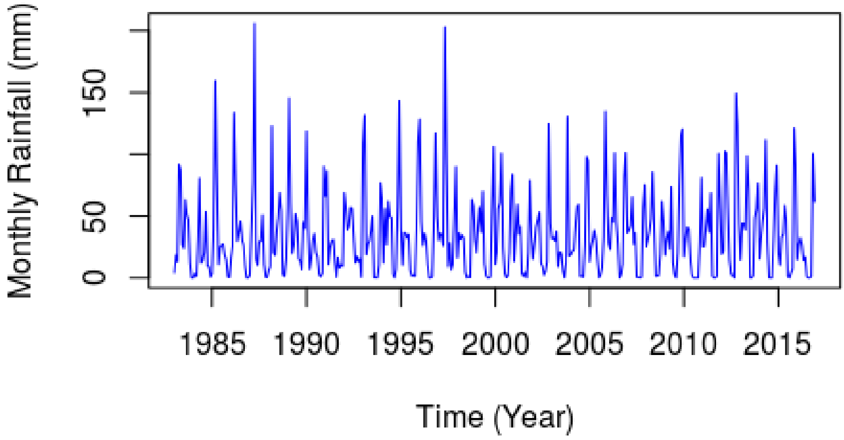


Figure 4.2: Monthly rainfall series plot

rainfall series has a unit root, which means that the rainfall series is non-stationary, and tests if the null hypothesis is to be statistically rejected in favor of the alternative hypothesis that the rainfall series is stationary. The table below depicts the outcomes of the test of the stations: the test statistic value at the original series and first differences are -8.5175 and -13.0749 respectively, which are lower than the critical values, -2.58, -1.95, and -1.62 all at 1%, 5%, and 10% respectively. Evidently, after having the ADF Unit Root test we concluded to reject the null hypothesis ($H_0$), which indicates that the average monthly rainfall series is stationary and does not have a unit root.

Table 4.3: ADF unit-roots test Summary

| Series | ADF test | 1% | 5% | 10% |
|---|---|---|---|---|
| Original Sereis | -8.5175 | -2.58 | -1.95 | -1.62 |
| Seas Defferencing | -13.0749 | | | |

However, the time series plot, the ACF and PACF show that the average monthly rainfall series indicates that the monthly rainfall data have a seasonal unit root. These tests for stationarity seem to agree and suggest that the first seasonal differencing in the series can achieve stationarity around a constant mean, which is approximately zero and its standard deviation is 36.89mm (Figure 4.3). In addition to the seasonally differenced time series plot of the monthly rainfall series, the ACF and PACF (Figures 4.4) also tells that the series is stationary in both mean and variance after the first seasonal difference.

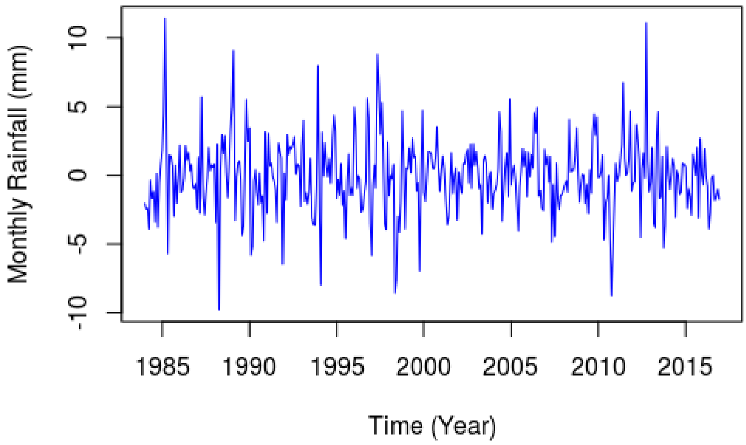


Figure 4.3: First seasonal differenced monthly rainfall series plot

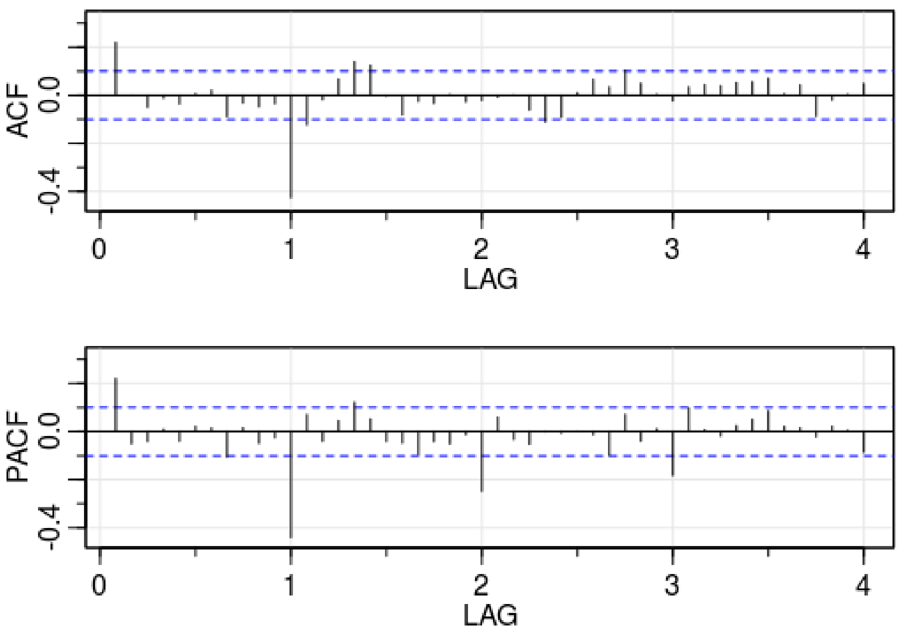


Figure 4.4: First seasonal differenced ACF and PACF monthly rainfall series plot

## 4.3 Model Building for Monthly Rainfall Series

Fitting a model to time series data involves plotting the data, transforming the data when appropriate, identifying the dependence orders of the model, parameter estimation, diagnostics tests, and model choice. In this section, a univariate seasonal ARIMA model is used to analyze the monthly rainfall series in Somali region.

### 4.3.1 Model Identification

The stationarity of the observed average monthly Rainfall series was evidently confirmed in the above section. Hence, there is no trend in the series and it has a seasonal pattern (Figure C), this shows that there is no need for non-seasonal differencing (d=0) but seasonal differencing is required for the seasonal stationarity. Thus, to make the average monthly rainfall series seasonally stationarity, a seasonal differencing (D=1) of the observed rainfall series was done to make series seasonally stationary. For further investigation, the seasonal ARIMA (p, 0, q) $(P, 1, Q)_{12}$ model was suggested. Initial, parameter p, q, P, and Q of these models were determined based on the characteristics of seasonally differenced ACF and PACF plots shown in Figure 4.4.

From the seasonally differenced series, the ACF plot has a positive significant spike at lag 1 and cuts off after that for the non-seasonal component. The plot also illustrates that it has a negative significant spike at lag 1s ($12^{th}$ lag) for the seasonal component. Therefore, one moving average (MA) value (q=1) for non-seasonal and two seasonal moving average (SMA) values (Q = 1 or 2) for the seasonal components were suggested for model identification. Similarly, the PACF of the seasonally differenced series has significant spikes at 1st lag for the non-seasonal and at $12^{th}$, $24^{th}$ and $36^{th}$ lags for the seasonal components. Then, the values of PACF cut off up to lag 48. Hence one autoregressive (AR) parameters (p = 1) for non-seasonal and three seasonal autoregressive (SAR) parameter values (P = 1, 2, or 3) were suggested to be included in the seasonal ARIMA model. Therefore, the following tentative models were identified as an initial model.

$SARIMA(1,0,1)(1,1,1)_{12}$
$SARIMA(1,0,1)(1,1,2)_{12}$
$SARIMA(1,0,1)(2,1,1)_{12}$
$SARIMA(1,0,1)(2,1,2)_{12}$
$SARIMA(1,0,1)(3,1,1)_{12}$
$SARIMA(1,0,1)(3,1,2)_{12}$

### 4.3.2 Methods of Parameter Estimation

Brockwell and Davis (1996) stated that the Non-linear Estimation of the parameters for Box-Jenkins models is quite complicated. Thus, the parameter estimates are usually obtained by the maximum likelihood estimation method. Thus, we got the following estimated values of the parameters of seasonal ARIMA shown in Table 4.4, which also gives the standard errors, the P-value, and the criteria for the model parameters. The seasonal ARIMA $(1,0,1)(2,1,2)_{12}$ compared to the other models shown in Table D in the *Appendix* turns out to be the best model since it has the least values of the information criterions, which means this model is the preferred one as it fits the monthly rainfall series of the region better, and can be further analyzed.

Table 4.4: Summary of Parameter Estimates and selection criteria

| Parameter | Estimate | SE | t-value | p-value | Fit Statistic |
|---|---|---|---|---|---|
| Constant | -0.0002 | 0.0003 | -0.8065 | 0.0020 | AIC = 1790.11 |
| AR1 | 0.8565 | 0.0755 | 11.345 | 0.0000 | BIC = 1824.31 |
| MA1 | 0.7066 | 0.1144 | 6.1743 | 0.0000 | *L = -885.38 |
| SAR1 | -0.0083 | 0.0944 | -0.0883 | 0.0007 | |
| SAR2 | 0.0563 | 0.0755 | 0.7463 | 0.0043 | |
| SMA1 | 1.7527 | 0.1824 | 9.6111 | 0.0000 | |
| SMA2 | 0.7579 | 0.1245 | 6.0860 | 0.0000 | |

* *Likelihood*

### 4.3.3 Diagnostic Checking

The third step is the formal assessment of each of the time series models. In this section, we examined the residuals of the estimated model to see if they are independent, to assess how well the selected model fit the actual monthly rainfall data.

There are several diagnostic tools available for ensuring to come to a satisfactory model. However, plotting the residuals of the estimated models is a useful diagnostic check. This should indicate any outliers that may affect the parameter estimates and also indicate points towards any possible autocorrelation problem. If a model is correctly specified, the residuals should be white noise. Thus, If the model fits well, the standardized residuals should behave as an identically and independently distributed sequence with mean zero and variance one. The time plot should be inspected for any obvious departures from this assumption (Shumay and Stoffar, 2017).

To check whether the residual series is white noise we need to do a Ljung-Box test with the correct degrees of freedom the results are summarized in the table below. From the Table we can observe that the p-value for all lags is greater than 0.05, this implies that the white noise hypothesis is not rejected. Which means, the data are independently distributed

The figure illustrated below (Figure 4.5) displays a plot of the standardized residuals, the ACF of the residuals, their normal Q-Q plot, and the p-values associated

Table 4.5: Residual white noises check with Ljung-Box test for the fitted model

| Total Lags | Chi-Square | DF | P-value |
|---|---|---|---|
| 6 | 3.10 | 4 | 0.5408 |
| 12 | 8.65 | 10 | 0.5651 |
| 18 | 12.75 | 16 | 0.6908 |
| 24 | 17.93 | 22 | 0.7096 |
| 30 | 23.15 | 28 | 0.7253 |
| 36 | 29.32 | 34 | 0.6962 |
| 42 | 36.64 | 40 | 0.6223 |
| 48 | 47.83 | 46 | 0.3981 |
| 54 | 49.35 | 52 | 0.5786 |
| 60 | 65.47 | 58 | 0.2335 |

with the Q-statistic.

Inspection of the time plot of the standardized residuals in the figure shows few outliers in the series, with a few values exceeding 3 standard deviations in magnitude. The ACF of the standardized residuals shows no clear significant autocorrelations, and the Q-statistic is never significant at the lags shown, since the p-values for the Ljung-Box Q test all are well above 0.05, indicating non-significance, which means that the residuals are random, and they are independent and identically distributed. The values of the Q-Q plot of the residuals are normal as they rest on a line and are not all over the place, the Shapiro-Wilk normality test has also a test statistic of W = 0.9724 and a p-value = 0.421 and which means that the residuals of the fitted models are normally distributed for the seasonal ARIMA $(1,0,1)(2,1,2)_{12}$ is adequate for modeling the average monthly rainfall series in Somali region.

From the seasonal ARIMA $(1,0,1)(2,1,2)_{12}$ model, the order of p is 1 means that the current time series ($x_t$) is reliant on its preceding data $x_{t-1}$. The order P is 2 means that $x_t$ is reliant on its preceding years' data of $x_{t-12}$, $x_{t-24}$. The order q is 1 means $x_t$ is reliant on its preceding random shock $w_{t-1}$, and the order Q is 2 means that $x_t$ is reliant on its preceding random shocks of $w_{t-12}$, $w_{t-24}$. As non-seasonal component of this model was stationary, no non-seasonal differencing was used (d = 0). However, seasonal differencing was used only once (D = 1) to remove seasonality in the model.

Hence, all the graphs illustrated in this section are in support of the assumption of the residuals that there is no pattern, we can go ahead and calculate the forecast.

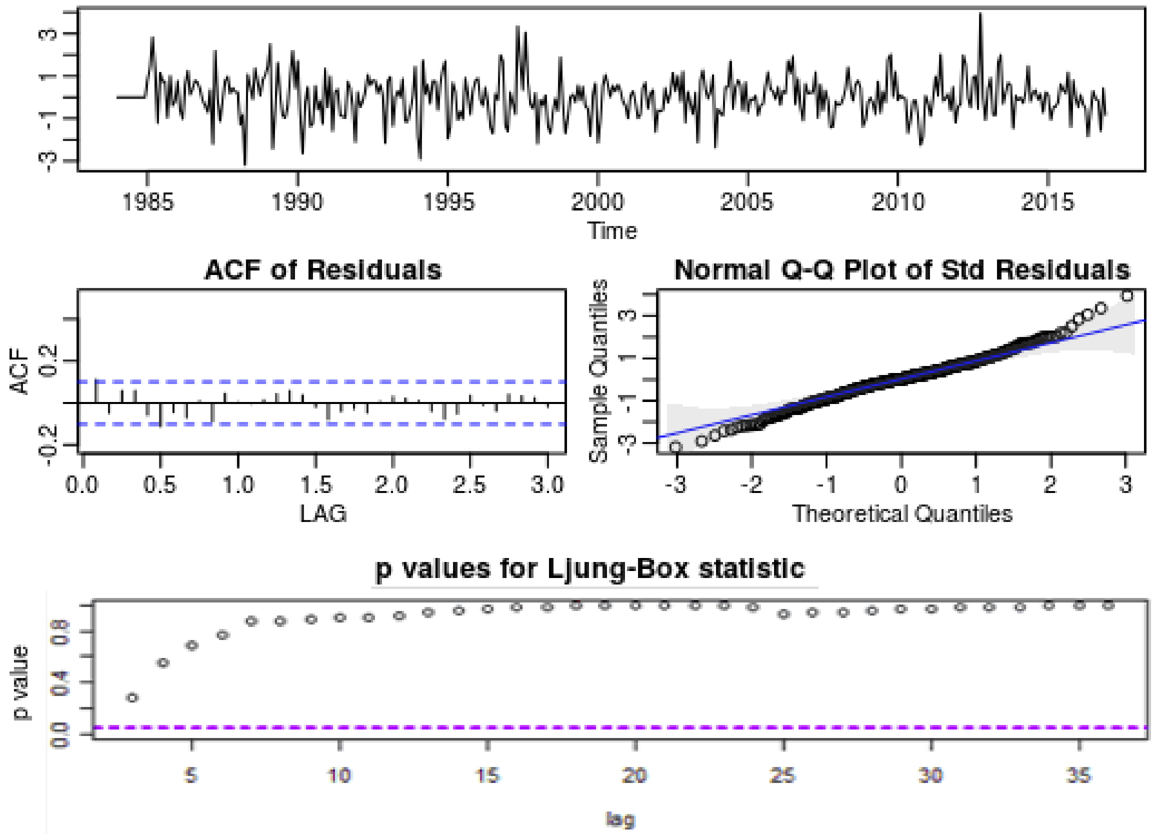


Figure 4.5: Diagnostics of the fitted model on the rainfall series

### 4.3.4 Forecasting

The model diagnostic tests in section 4.4 have shown that all the parameter estimates are significant and the residual series is white noise. We can proceed to forecast the rainfall series with fitted Seasonal ARIMA $(1,0,1)(2,1,2)_{12}$ model. The goal of forecasting is to predict future rainfall values of a known time series, (Shumway and Stoffar, 2017).

In this study, the final model for forecasting the monthly rainfall series of Somali region is given below. The seasonal ARIMA model $(1,0,1)(2,1,2)_{12}$ can be written as:

$$(1-\phi_1 B)(1-\Phi_2 B^{24})(1-B^{12})x_t = (1+\theta_1 B)(1+\Theta_2 B^{24})w_t \qquad (4.1)$$

The forecast in this study is conducted in two-folds, in-sample and out sample forecasts. The rainfall series observed in 2016 is kept out of the main analysis, to make the in sample forecast possible and for this purpose only 1983-2015 data were considered. The in sample forecast provides the rainfall series with a marginal difference compared to the series obtained in 2016, this indicates the reliability of the model estimated. The actual and forecast values for 2016 is displayed in Table 4.7.

**Forecasting Accuracy Measures**

The forecasting accuracy performance is usually measured by using different standardized statistical tools. However, in this study, we have used MSE, MAPE, RMSE, and Theil's Inequality to measure the forecasting performance of the model selected to be the best in our rainfall data. Table 4.6, shown below illustrates the results of the accuracy measures applied to our data. It is seen from the table, that the values of Theil's inequality, Bias proportion, and Variance proportion are relatively close to zero. Thus, the accuracy measures used in this study indicate that the forecasting accuracy (inaccuracy) is high (low).

Table 4.6: Results of Forecast Accuracy Measures

| Accuracy Measures | Monthly rainfall series |
|---|---|
| RMSE | 38.7 |
| MSE | 19.6 |
| MAPE | 12.26 |
| Theil's | 0.031 |
| Bias proportion | 0.023 |
| Variance proportion | 0.018 |

Furthermore, Table 4.7 is summarized the actual and forecasted values of monthly rainfall series from Jan 2016 to Dec 2016, which supports the value of the forecast accuracy measures. Moreover, the month-wise forecast and its interval of monthly rainfall series of Somali region for the following five years (2017-2021) by using the selected seasonal ARIMA $(1,0,1)(2,1,2)_{12}$ model is shown on Table E in the *Appendix*. The table shows that the point forecast results showed that all points lie within the confidence interval. The forecasted figures tend to be very close to the actual points which means the model predicts well.

Table 4.7: Actual and fitted values of the series

| Months | Actual | Forecast | Residual |
|---|---|---|---|
| Jan, 2016 | 0 | 2.1 | -2.1 |
| Feb, 2016 | 0 | 4.2 | -4.2 |
| Mar, 2016 | 0 | 1.9 | -1.9 |
| Apr, 2016 | 116 | 118 | -2 |
| May, 2016 | 108 | 106.4 | 1.6 |
| Jun, 2016 | 45 | 43.2 | 1.8 |
| Jul, 2016 | 50 | 50.9 | -0.9 |
| Aug, 2016 | 85 | 87.3 | -2.3 |
| Sep, 2016 | 66 | 68.2 | -2.2 |
| Oct, 2016 | 22 | 19.8 | 2.2 |
| Nov, 2016 | 17 | 15.9 | 1.1 |
| Dec, 2016 | 0 | 1.7 | -1.7 |

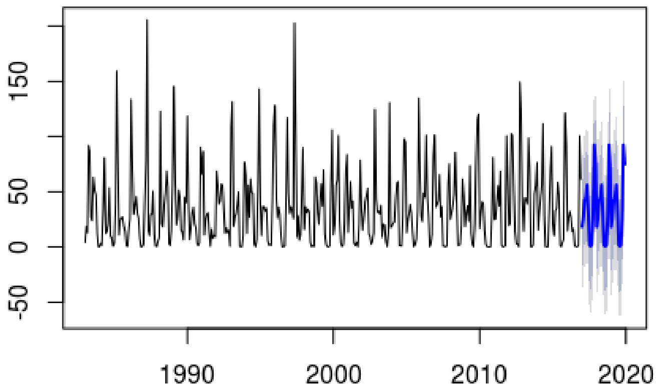


Figure 4.6: Forecast plot for monthly rainfall series in Somlai region

# Chapter 5

# Discussion and Conclusion

## 5.1 Discussion

In this study, a monthly gridded rainfall data at $0.4^o$ * $0.4^o$ resolution gridded rainfall data for 34 years from 26 meteorological stations covering the entire Somali region in Ethiopia, were analyzed. Likewise, many rainfall types of research, a univariate Box-Jenkins methodology were used to examine the modes of variation of the rainfall series.

NMSA (1996) identified that the mean annual rainfall distribution over the country is characterized by a large spatial variation which ranges from about 2000 mm over some areas in the Southwest of Ethiopia to less than 250 mm over the Afar and Ogaden lowlands. Furthermore, results in the current study show that the mean annual rainfall in Somali region is 567 mm, with a higher year-to-year variation ($CV > 0.30$), the results also revealed that the region has two rainy periods April-May and August-October, followed by a prolonged dry season, these findings support Eshetu (2013) who reported that the semi-arid lowlands are characterized by high variability, short and intense rainy season typically occurring in August and no more than 4 months long, followed by a prolonged dry season.

To test whether residual from the fitted model come from normally distributed series, we used QQ-plot of the residual and Shapiro-Wilks test. The QQ-plot of the residual in this study showed normality, as the points rest on the line and are not all over the place. Shapiro-Wilks test also confirmed this fact with a P-value of 0.421, which is greater than 0.05. implying that the residuals from the fitted models come from a normal distribution.

Ljung-Box test with the correct degrees of freedom, was also used to check the residual of the series, in which we observed a P-value greater than 0.05 for all lags, implying that the residual series is white noise. In addition to this, to measure the forecasting performance of the model we have used MSE, MAPE, RMSE, and Theil's Inequality. Hence, the value of Theil's inequality showed to be relatively close to zero, and the value of mean absolute percentage error from this study is 12.26 which means 87.74 accuracies. Thus, from this values, we can say that the model fits the data well.

Further comparison based on the forecasting accuracy of the model is performed with the hold-out of some rainfall values. The point forecast results showed a very closer match with the pattern of the actual data and better forecasting accuracy in the validation period. In addition to this, the forecasted values in comparison to the actual values from January, 2016 to December, 2016 showed a relatively very slight increase in the periods December-April and July-September, and on the other hand a relatively very slight decrease in the other months of the year (May, June, October and November).

## 5.2 Conclusion

A univariate time series model for monthly rainfall series of Somali region was adjusted, processed, checked diagnostically, and finally a seasonal ARIMA model was established. Thus, the parameter estimation and diagnostic analysis of the results revealed that models' are adequately fitted to the rainfall data. In particular, the residual analysis, which is important for diagnostic checking confirmed that there is no violation of assumptions concerning the model adequacy. Comparing with other models, a seasonal ARIMA $(1,0,1)(2,1,2)_{12}$ model has been selected as the final and best model for the rainfall series and then used to establish an 80% and 95% forecasting interval to forecast a 5 years monthly rainfall values from January 2017 to December 2021. This forecasting results revealed a relatively very slight decrease of monthly rainfall over the forecast period.

# Chapter 6

# Limitation of the Study

Because of the unavailability of local rain gauge data collected from the rainfall stations in Somali region, the data used in this study is based on $0.4^o$ Latitude * $0.4^o$ Longitude gridded rainfall data.

# References


Abebe, S. (2018). Application of Time Series Analysis to Annual Rainfall Values in Debre Markos Town, Ethiopia. Computational Water, Energy, and Environmental Engineering, 7, 81-94.

Admassu, S. (2004). Rainfall variation and its effect on crop production in Ethiopia . (Masters Dissertation, Addis Ababa University).

Alamerew, B. and Eshetu, W. (2009). Assessment of Local Climate in Addis Ababa, . Journal of Ethiopian Statistical Association, Vol. 18, 55-57.

Asfaw, A., Simane, B., Hassen, A. and Bantider, A. (2018). Variability and time series trend analysis of rainfall and temperature in north-central Ethiopia: A case study in Woleka sub-basin. . Weather and Climate Extremes, 19, 29-41.

Berhane, T., Shibabaw, N., Awgichew, G. and Kebede, T., . (2018). Modeling and Forecasting Rainfall in Ethiopia. . International Journal of Computing Science and Applied Mathematics, 4(2), 42-46.

Bewket, W. (2009). Rainfall variability and crop production in Ethiopia: Case study in the Amhara Region. Proceedings of the 16th International Conference of Ethiopian Studies.(ed), 2:827-836.

Box, G. E. P. & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control, 2nd ed. San Francisco: Holden-Day.

Box, G.E., Jenkins, G.M., Reinsel, G.C. and Ljung, G.M. (2015). Time series analysis: forecasting and control. . John Wiley & Sons.

Box, G.E.P. and Pierce,D.A. (1970). Distributions of residual autocorrelations in autoregressive integrated moving average models . J. American Stat. Assoc, Vol.72, 397-402.

Challen, J. and Hagger, S. (1983). Economics; Mathematical models. . New York: St. Martin's Press.

Cheung, W.H., Senay, G.B. and Singh, A. . (2008). Trends and spatial distribution of annual and seasonal rainfall in Ethiopia. . International Journal of Climatology: A Journal of the Royal Meteorological Society, 28(13), 1723-1734.

Desta S, Tezera S, Gebru G, and Kristjanson P. (2011). Summary of baseline household survey results: Borana, Ethiopia. CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS). Copenhagen, Denmark.

Diebold, F. (2011). Elements of forecasting, 2nd ed. South-Western: Cincinnati, Ohio.

Doti, T., 2010. Climate variability, pastoralists' vulnerability and options. The case of the Borana of Northern Kenya. Climate change and natural resources conflicts in Africa. Monograph, 170, pp. 189-204.

Elias, E. (2009). Threats of climate change on pastoral production under restricted herd mobility: A case study in Borena. Work done for SOS Sahel Ethiopia.

Enders, W. (1995). Applied Econometric Time Series. New York: John Wiley and Sons, Inc.

Eshetu, Z. (2013) Forestry and agroforestry land use for climate change adaptation and mitigation. Climate Change Vulnerability and Risk Assessment of Agriculture and Food Security in Ethiopia

Etuk, E. H. and Mohamed, M. (2014). Time Series Analysis of Monthly Rainfall data for the Gadaref rainfall station, Sudan, by Sarima Methods. . International Journal of Scientific Research in Knowledge, 2(7): 320 – 327.

FAO. (2011). Framework programme on climate change adaptation. Rome: Food and Agriculture Organization of the United Nation.

Fazzini, M., Bisci, C. and Billi, P. (2015). The climate of Ethiopia. In Landscapes and landforms of Ethiopia ( 65-87). Springer, Dordrecht.

Filder, T.N., Muraya, M.M. and Mutwiri, R.M. (2019). Application of Seasonal Autoregressive Moving Average Models to Analysis and Forecasting of Time Series Monthly Rainfall Patterns in Embu County, Kenya. . Asian Journal of Probability and Statistics, 4(4), 1-15.

Gebreegziabher Z, Stage J, Mekonnen A, Alemu A . (2011). Climate change and the Ethiopian economy: A computable general equilibrium analysis. Environment for Development Discussion Paper Series,.

Georgis, K. (2010). Agricultural based livelihood systems in drylands in the context of climate change: Inventory of adaptation practices and technologies of Ethiopia. Environment and Natural Resources Management Working Paper 38. Food and Agriculture Organi.

Gebretsadikan, A. and Sharma, M.K. (2011). Modeling and forecasting of rainfall data of mekele for Tigray region (Ethiopia). Statistics and Applications, 9(1-2), 31-53.

Graham, A. and Mishra, E.P. (2017). Time series analysis model to forecast rainfall for Allahabad region. . Journal of Pharmacognosy and Phytochemistry, 6(5), 1418-1421.

Griffin, T. and Collins, G. (2018). Economics of Resilience oo Drought: Ethiopia Analysis: USAID.

Jember, G. and Tadege A. (2010). Climate Change, Variability, Trends and Potential Impacts and Risks in Major Agro-Ecological Zones in Oromia region of Ethiopia. In proceedings of strengthening, capacity for climate change adaptation in the Agricultural sector.

Kedir, H. and Tekalign, S. (2016). Climate variability and livelihood strategies pursued by the pastoral community of the karrayu people, Oromia region, Central Ethiopia. . East African Journal of Sciences, 10(1), 61-70.

Ketema, T. (1999). Test of homogeneity, frequency analysis of rainfall data and estimate of drought probabilities in Dire Dawa, Eastern Ethiopia. Ethiopian Journal of Natural Resources, Vol. 1: 125-136.

Kindie, T., Jemal, S., Mezegebu, G. and Girma, M. (2016). Agriculture under a changing climate in Ethiopia: challenges and opportunities for research. Ethiopia Institute of Agricultural Research (EIAR) 50th Year Jubilee Anniversary Special Issue, 67-86.

Ljung, G.M. and Box, G.E.P. (1978). "On a Measure of Lack of Fit in Time Series Models", . Biometrika,, Vol.65,297-303.

Mahoo, H., Radeny, M., Kinyangi, J. and Cramer, L. (2013). Climate change vulnerability and risk assessment of agriculture and food security in Ethiopia: which way forward? CCAFS Working Paper No. 59. Copenhagen, Denmark: CGIAR Research Program on Climate Change.

Mersha, E. (2002). Determination of Rainfall Cyclicity over Selected Location in Ethiopia. Ethiopian J. Water Science and Technology, Vol.5.

Moloy, M.D., Chowdhury, M., Binyamin, M. and Mondal, S.K. (2018). Using SARIMA Approach to Modeling and Forecasting Monthly Rainfall in Bangladesh. IJSAR, 5(5), 06-14.

Momani, N.M., and Naill, P.E. (2009). Time series analysis model for rainfall data in Jordan: Case study for using time series analysis. American Journal of Environmental Sciences, 5(5), 599-604.

Mulat, D., Fantu, G. and Tadele, F. . (2004). Agricultural development in Ethiopia: are there alternatives to food aid? . Addis Ababa: Unpublished research report.

Namara, R., Makombe, G., Hagos, F. and Awulachew S.B. . (2008). Rrural poverty and inequality in Ethiopia: Does access to small-scale irrigation make a difference? In: Awlachew SB, Loulseged, M and Yilma AD (eds.) Impact of Irrigation on Poverty and Environment.

Nirmala, M. and Sundaram, S.M. (2010). A seasonal ARIMA model for forecasting monthly rainfall in Tamilnadu . National Journal on Advances in Building Sciences and Mechanics, 1(2),43-47.

NMA. (1996). Climate & agroclimate resources of Ethiopia. Addis Ababa: NMSA Meteorological Research Report Series, Vol.1.

NMA. (2007). National Adaptation Programme of Action of Ethiopia (NAPA). Addis Ababa: Final draft report. National Meteorological Agency.

NMSA. (1996). Climate & agroclimate resources of Ethiopia. . Addis Ababa: NMSA Meteorological Research Report Series, Vol.1, No.1.

Pankhurst, R. and Johnson, D.H. (1988). The great drought and famine of 1888-92 in northeast Africa. In The ecology of survival: case studies from northeast African history. London: Lester Crook Academic Publishing.

Papalaskaris, T., Panagiotidis, T. and Pantrakis, A. (2016). Stochastic Monthly Rainfall Time Series Analysis, Modeling and Forecasting in Kavala City, Greece, North-Eastern Mediterranean Basin. . Procedia Eng, 162, 254–263.

Phillips, P.C.B and Perron, P. . (1988). Testing for a Unit Root in Time Series Regression. Biometrica, Vol. 75, 335–346.

Rettberg, S., Beckmann, G., Minah, M. and Schelchen, A. (2017). Ethiopia‘s Arid and Semi-Arid Lowlands: Towards Inclusive and Sustainable Rural Transformation.

Seleshi, Y. and Zanke, U. (2004). Recent changes in rainfall and rainy days in Ethiopia. International Journal of Climatology: A Journal of the Royal Meteorological Society, 24(8), 973-983.

Shumway, R.H. and Stoffer, D.S. . (2017). Time Series Analysis and Its Applications: With R Examples. Fourth Edition, Springer Texts in Statistics, . Springer, Cham.

Tadege, A. (2007). Climate Change National Adaptation Program of Action (NAPA) of Ethiopia. . Addis Ababa, Ethiopia: NMS (National Meteorological Agency: Federal Democratic Republic of Ethiopia, Ministry of Water Resources.

Tadesse, K. (2002). Trends in agricultural production, technology dissemination and price movements of outputs and inputs. Addis Ababa: Ethiopian Development Research Institute.

Takele, R. (2012). Statistical Analysis of Rainfall Pattern in Dire Dawa Eastern Ethiopia (Masters Dissertation, Addis Ababa University).

Takele, R. and Gebretsidik, S. (2015). Prediction of Long-term Pattern and its Extreme Event Frequency of Rainfall in Dire Dawa Region, Eastern Ethiopia. J Climatol Weather Forecasting , 3: 130, 1-15.

Theil's, H. (1996). Economic Forecasts and Policy . Florida: University of Florida.

Tilahun, K. (2006). The characterization of rainfall in the arid and semi-arid regions of Ethiopia. Water SA, 32(3), 429-436.

Troeger, S. (2011). "Everything that is happening now is beyond our capacity" Nyangatom livelihoods under threat. In PFE proceeding. Unpublished report.

UNDP. (2008). Human Development Report 2007/2008. Fighting Climate Change: Human solidarity in a divided world. New York: United Nations Development Programme.

Von Braun, J. (1991). A policy agenda for famine prevention in Africa. Washington DC: Food Policy Statement No.13. IFPRI.

Wood, A. (1977). A preliminary chronology of Ethiopian droughts. In Drought in Africa. International African Institute, London, Vol. 2, 68-73.

Zakaria, S., Al-Ansari, N., Knutsson, S. and Al-Badrany, T. (2012). ARIMA Models for weekly rainfall in the semi-arid Sinjar District at Iraq. Journal of Earth Sciences and Geotechnical Engineering, 2(3). 25-55.

# Appendix

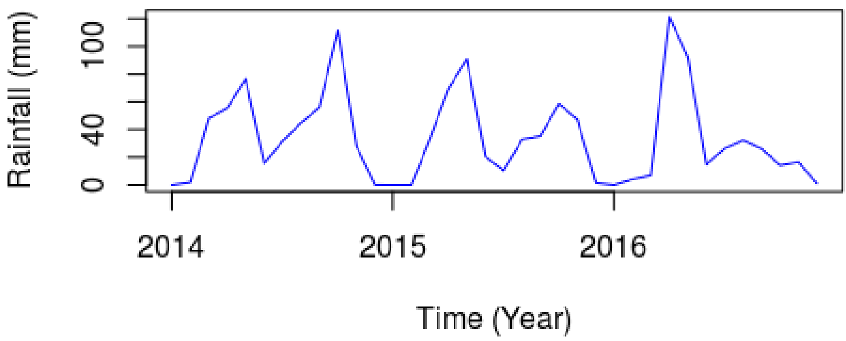


Figure A: Rainfall series plot for visibility

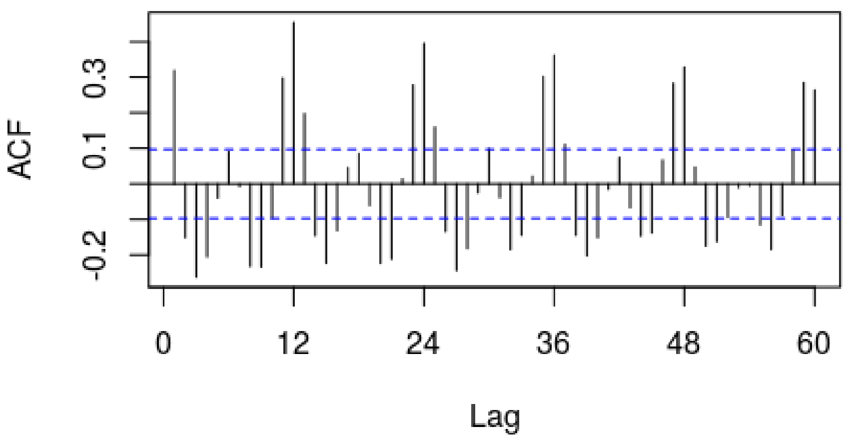


Figure B: ACF plot for monthly rainfall series

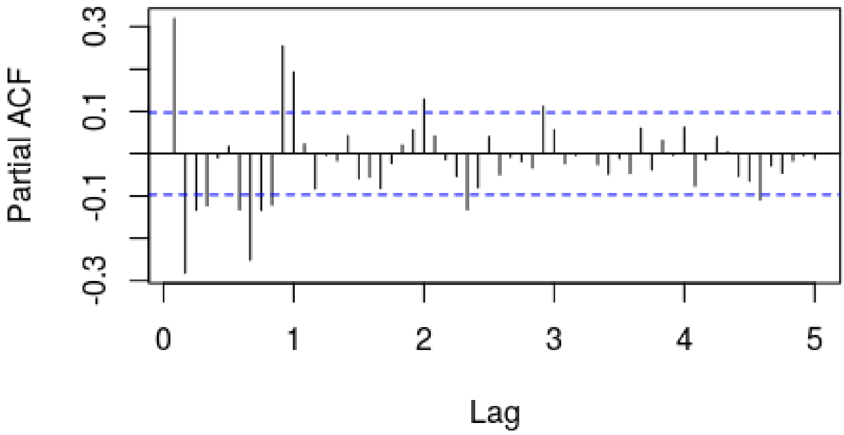


Figure C: PACF plot for monthly rainfall series

Table D: Seasonal ARIMA models

| Model | AIC | BIC |
|---|---|---|
| ARIMA $(1,0,1)(1,1,1)_{12}$ | 1804.09 | 1837.21 |
| ARIMA $(1,0,1)(1,1,2)_{12}$ | 1812.10 | 1859.13 |
| ARIMA $(1,0,1)(2,1,1)_{12}$ | 1794.23 | 1826.04 |
| ARIMA $(1,0,1)(2,1,2)_{12}$ | 1790.11 | 1824.31 |
| ARIMA $(1,0,1)(3,1,1)_{12}$ | 1797.02 | 1829.10 |
| ARIMA $(1,0,1)(3,1,2)_{12}$ | 1802.31 | 1834.31 |

Table E: Forecast of the Rainfall series from the period Jan 2017 - Dec 2021

| Months | Forecast | Lo 80% | Hi 80% | Lo 95% | Hi 95% |
|---|---|---|---|---|---|
| Jan 2017 | 0.00 | -20.89 | 57.39 | -41.617 | 78.12 |
| Feb 2017 | 0.00 | -12.90 | 65.39 | -33.62 | 86.11 |
| Mar 2017 | 2.31 | -13.54 | 81.83 | -17.17 | 12.56 |
| Apr 2017 | 92.23 | 4.60 | 82.90 | -16.11 | 103.62 |
| May 2017 | 96.01 | 17.02 | 95.31 | -3.69 | 116.04 |
| Jun 2017 | 35.03 | -5.54 | 72.74 | -26.27 | 93.46 |
| Jul 2017 | 28.00 | -37.48 | 40.81 | -18.20 | 61.53 |
| Aug 2017 | 82.19 | -38.74 | 39.55 | -59.46 | 60.27 |
| Sep 2017 | 61.07 | -37.22 | 41.06 | -57.94 | 61.79 |
| Oct 2017 | 19.81 | -11.63 | 66.65 | -32.36 | 87.37 |
| Nov 2017 | 9.11 | -13.23 | 31.52 | -30.50 | 51.42 |
| Dec 2017 | 0.00 | -34.56 | 15.23 | -13.21 | 36.58 |
| Jan 2018 | 0.00 | -22.08 | 18.58 | -43.43 | 29.93 |
| Feb 2018 | 0.00 | -14.08 | 16.57 | -35.43 | 27.92 |
| Mar 2018 | 2.69 | -12.35 | 24.02 | -18.99 | 14.37 |
| Apr 2018 | 108.75 | 3.42 | 124.08 | -17.92 | 135.43 |
| May 2018 | 80.71 | 15.83 | 106.50 | -5.51 | 117.85 |
| Jun 2018 | 26.49 | -6.73 | 73.92 | -28.08 | 65.28 |
| Jul 2018 | 23.76 | -38.66 | 41.99 | -60.01 | 63.34 |
| Aug 2018 | 84.40 | -19.92 | 90.73 | -11.27 | 102.08 |
| Sep 2018 | 54.92 | -18.41 | 81.25 | -19.76 | 73.60 |
| Oct 2018 | 21.50 | -12.82 | 47.84 | -24.17 | 49.19 |
| Nov 2018 | 14.37 | -17.04 | 32.71 | -10.69 | 44.06 |
| Dec 2018 | 2.90 | -13.41 | 16.38 | -11.45 | 18.34 |
| Jan 2019 | 0.00 | -23.23 | 19.73 | -15.19 | 21.69 |
| Feb 2019 | 0.00 | -15.24 | 17.72 | -17.20 | 19.69 |
| Mar 2019 | 0.00 | -1.20 | 14.17 | -20.75 | 16.13 |
| Apr 2019 | 98.75 | 22.27 | 118.24 | 19.69 | 127.20 |
| May 2019 | 99.56 | 14.68 | 107.65 | 17.27 | 119.61 |
| Jun 2019 | 20.59 | -7.88 | 45.08 | -9.84 | 52.04 |
| Jul 2019 | 74.90 | 32.29 | 117.50 | 9.73 | 140.06 |
| Aug 2019 | 81.66 | 9.82 | 93.14 | 13.78 | 105.11 |
| Sep 2019 | 60.40 | 9.08 | 81.88 | 13.04 | 93.85 |
| Oct 2019 | 21.92 | -9.56 | 43.40 | -12.52 | 65.36 |
| Nov 2019 | 6.50 | -13.97 | 28.99 | -15.93 | 30.54 |
| Dec 2019 | 2.37 | -15.89 | 13.86 | -28.93 | 15.82 |

| Months | Forecast | Lo 80% | Hi 80% | Lo 95% | Hi 95% |
|---|---|---|---|---|---|
| Jan 2020 | 0.00 | -14.51 | 16.32 | -16.90 | 13.41 |
| Feb 2020 | 0.00 | -16.36 | 18.85 | -18.91 | 19.47 |
| Mar 2020 | 3.21 | -12.08 | 18.29 | -22.46 | 17.85 |
| Apr 2020 | 112.75 | 31.14 | 126.36 | 31.40 | 128.91 |
| May 2020 | 87.17 | 13.56 | 99.77 | 18.98 | 101.33 |
| Jun 2020 | 22.59 | -9.00 | 36.20 | -1.56 | 38.75 |
| Jul 2020 | 21.66 | -4.94 | 44.27 | -3.49 | 46.82 |
| Aug 2020 | 79.40 | 22.20 | 93.01 | 14.75 | 95.56 |
| Sep 2020 | 59.92 | 12.68 | 74.52 | 23.23 | 87.08 |
| Oct 2020 | 20.50 | -15.09 | 50.11 | -17.65 | 52.66 |
| Nov 2020 | 14.37 | -14.77 | 34.98 | -11.21 | 37.53 |
| Dec 2020 | 0.00 | -11.20 | 11.59 | -8.06 | 14.73 |
| Jan 2021 | 0.00 | -15.44 | 16.95 | -14.57 | 15.08 |
| Feb 2021 | 0.00 | -17.45 | 19.94 | -14.58 | 13.07 |
| Mar 2021 | 0.00 | -11.00 | 16.39 | -12.13 | 19.52 |
| Apr 2021 | 107.75 | 25.05 | 127.45 | 23.07 | 122.58 |
| May 2021 | 86.17 | 22.47 | 102.87 | 19.65 | 103.00 |
| Jun 2021 | 35.59 | 10.10 | 57.29 | 13.23 | 60.42 |
| Jul 2021 | 33.66 | 12.03 | 45.36 | 15.16 | 58.49 |
| Aug 2021 | 83.40 | 23.29 | 98.10 | 26.42 | 107.23 |
| Sep 2021 | 62.92 | 21.77 | 85.62 | 14.91 | 89.75 |
| Oct 2021 | 22.50 | -16.19 | 41.20 | -19.32 | 44.33 |
| Nov 2021 | 14.37 | -18.67 | 26.07 | -15.54 | 29.20 |
| Dec 2021 | 2.02 | -12.13 | 16.66 | -6.43 | 14.36 |

** The predicted rainfall values for the period January 2017 - December 2021 given in the above table is calculated from the first seasonal differenced actual data of the rainfall series (1984 - 2016).*