KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

KUMASI

COLLEGE OF SCIENCE

A TIMES SERIES ANALYSIS INTO THE RAINFALL PATTERNS IN FOUR SELECTED REGIONS OF GHANA

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS IN PARTIAL FULFILLMENT OF THE REQUIRMENTS FOR THE AWARD OF MASTER OF PHILOSOPHY IN MATHEMATICS

BY

KINGSLEY KORMLA MORTEY

SEPTEMBER, 2011

DECLARATION

I hereby declare that this submission is my own work towards the award of the Mphil. (Mathematics) degree and that to the best of my knowledge, it contains no material previously published by another person or material which had been accepted for the award of any other degree of the university, except where due acknowledgement had been made in the text.

…………………………… ……………………………

KINGSLEY KORMLA MORTEY Date

…………………………… ……………………………

NANA KENA FREMPONG Date

(SUPERVISOR)

…………………………… ……………………………

Mr. F. K. DARKWAH Date

(HEAD OF DEPARTMENT)

DEDICATION

I dedicate this work to the Most High God who has shown me mercy in many ways to have brought me this far in life. To my biological mother Margaret Yawa Agbatey, Others who have shown me Motherly love Mrs Elizabeth Agbodra, Mama Ernestina Agbesinyale, Mama Grace Adulai, Mama Victoria Fiakponu, Gwendolyn Amamoo, Naa Ahinee Amamoo, Lucy Adzo Dzoagbe, Hannah Maame Acquaye and also to my family, friends and loved ones.

ACKNOWLEDGEMENT

My sincere appreciations first and foremost go to Almighty God who has seen me through this thesis. To my supervisor Nana Kena Frempong for his valuable time, guidance and productive criticisms, which enabled me to produce this thesis. Dr. S. K. Amponsah, Dr. Adebanji, Mr Adakwah, the late Mr Agyemang Dr. Dapaah, Dr Prempeh and the lecturers and staff of the Department of Mathematics, for their knowledge shared, encouragement and advice.

To Ghana Meteorological Agency for providing the relevant data for the Analysis of this thesis. To Mathematics-KNUST Class of 2009 for creating an enabling environment for studies making it a success, to them I say thank you. This thesis would not be complete if not for you great support through day and night reading and supporting me when the analysis got tough, I say Naa Ahinee Amamoo, thank you and God bless you for your effort. To my Boss Professor Nene Jonas Akpanglo-Nartey

Finally to my friends and families especially Lucy Adzo Dzoagbe for standing on my neck all the time and all others who have contributed in one way or the other to make the writing of this thesis possible, thanks you all.

ABSTRACT

In recent times there have been calls on the climatic changes that are happening and the need to put measures in place to forestall drastic changes in the climate.

In this study therefore, we attempted to study the trend of rainfall patterns in Ghana for four regions: Northern, Western, Eastern, and the Greater Accra Region, using four models namely: Linear Trend with Seasonal Terms, Seasonal Exponential Smoothing, ARIMA, Simple Exponential Smoothing, and Linear (Holt) Exponential Smoothing. All the ARIMA models are one model family, and the Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing which is one model family, since the Simple Exponential Smoothing is just basics of the Linear (Holt) Exponential Smoothing.

We have been able to show that the best model for the prediction of rainfall in the regions under study was Linear Trend with Seasonal Terms; in choosing this model we used the Means Square Error (MSE) and the R-Square as a criterion for the selection of the best model for prediction. We also further were able to come to the conclusion that the levels of rain in this four regions are not going to fall but rather will rise at least for the year 2011.

Table of Contents

TABLE OF FIGURES

TABLE OF TABLES

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

Rain is liquid precipitation, as opposed to non-liquid kinds of precipitation such as snow, hail and sleet. Rain requires the presence of a thick layer of the atmosphere to have temperatures above the melting point of water near and above the Earth's surface. On Earth, it is the condensation of atmospheric water vapor into drops of water heavy enough to fall, often making it to the surface. Two processes, possibly acting together, can lead to air becoming saturated leading to rainfall: cooling the air or adding water vapor to the air. Virga is precipitation that begins falling to the earth but evaporates before reaching the surface; it is one of the ways air can become saturated. Precipitation forms via collision with other rain drops or ice crystals within a cloud. Rain drops range in size from oblate, pancake-like shapes for larger drops, to small spheres for smaller drops.

Moisture moving along three-dimensional zones of temperature and moisture contrasts known as weather fronts is the major method of rain production. If enough moisture and upward motion is present, precipitation falls from convective clouds (those with strong upward vertical motion) such as cumulonimbus (thunderstorms) which can organize into narrow rain bands. In mountainous areas, heavy precipitation is possible where upslope flow is maximized within windward sides of the terrain at elevation which forces moist air to condense and fall out as rainfall along the sides of mountains. On the leeward side of mountains, desert climates can exist due to the dry air caused by down slope flow which causes heating and drying of the air mass. The movement of the monsoon trough, or inter tropical convergence zone, brings rainy seasons to savannah climes. Rain is the primary source of freshwater for most areas of the world, providing suitable conditions for diverse ecosystems, as well as water for hydroelectric power plants and crop irrigation. Rainfall is measured through the use of rain gauges. Rainfall amounts are estimated actively by weather radar and passively by weather satellites. (http://en.wikipedia.org/wiki/Rain retrieved on November, 2010)

The standard way of measuring rainfall or snowfall is the standard rain gauge, which can be found in 100-mm (4-in) plastic and 200-mm (8-in) metal varieties. The inner cylinder is filled by 25 mm (0.98 in) of rain, with overflow flowing into the outer cylinder. Plastic gauges have markings on the inner cylinder down to 0.25 mm (0.0098 in) resolution, while metal gauges require use of a stick designed with the appropriate 0.25 mm (0.0098 in) markings. After the inner cylinder is filled, the amount inside it is discarded, then filled with the remaining rainfall in the outer cylinder until all the fluid in the outer cylinder is gone, adding to the overall total until the outer cylinder is empty. Other types of gauges include the popular wedge gauge (the cheapest rain gauge and most fragile), the tipping bucket rain gauge, and the weighing rain gauge. For those looking to measure rainfall the most inexpensively, a can that is cylindrical with straight sides will act as a rain gauge if left out in the open, but its accuracy will depend on what ruler you use to measure the rain with. Any of the above rain gauges can be made at home, with enough know-how. When a precipitation measurement is made, various networks exist across the United States and elsewhere where rainfall measurements can be submitted through the

Internet, such as COCORAHS or GLOBE. If a network is not available in the area where one lives, the nearest local weather or met office will likely be interested in the measurement. One millimeter of rainfall is the equivalent of one liter of water per square meter. This makes computing the water requirements of crops simple.

1.2 OBJECTIVE

The objective is to study the rainfall patterns in Northern, Western, Eastern and Greater Accra Regions taking the years 1995 – 2009 and getting an appropriate times series model for forecasting rainfall in four regions in Ghana.

1.3 PROBLEM STATEMENT

As a result of climate change more work is now being done on climate indices such as rainfall, sunshine, temperature and so on. But much has not been done in this direction in Ghana as a country especially looking at a number of models at the same time to conclude on the best model; this research studies rainfall patterns in four regions of Ghana, in the selection of the four regions purposive sampling (**Judgmental sampling** or **Purposive sampling** - The researcher chooses the sample based on who or where they think would be appropriate for the study) is used.

1.4 SIGNIFICANCE

The significance of this research work will be enormous since the world's population is increasing and with the events of global warming.

1. It will provide empirical evidence to stakeholders on rainfall trends to formulate policies that can benefit the regions concern and the nation at large.

2. It will help the regions under consideration to know the trends and rainfall patterns.

3. It will help the farmers to know how to plan their farming activities ahead of time.

1.5 LIMITATIONS

Just like any other academic thesis, this very work will also have some limitations:-

I. Time: the time needed to complete a conduct this thesis in a very elaborate manner is not available.

II. Finance: being an academic thesis, it need some amount of financial commitment to be able to get more logistics for the thesis but clearly that is not available

III. Literature: getting data needed for any purpose is not easy in Ghana, and that was a major limitation for this work.

1.6 SCOPE

This thesis, "The Application of Times Series on rainfall Patterns in Ghana" will cover four selected regions in Ghana, namely Eastern Region, Western Region, Northern Region and the Greater Accra Region and that is selected to represent the regions in the extreme ends of the country. The analysis will be based on a secondary data collected by the Ghana Meteorological Agency.

1.7 PURPOSE OF THE STUDY

The purpose of this study will be to:-

- ✓ To advance the frontiers of knowledge
- ✓ To add to the existing academic literature on Rainfall pattern using Times series
- ✓ To serve as a guide to further studies in this field of studies
- ✓ To fulfill the requirements for graduation.

1.8 DATA COLLECTION

The main data that will be used for the research is from secondary source. The rainfall data in millimeters (mm) will be collected from the Ghana Meteorological Agency (GMET) Accra who receives data from their 22 synoptic stations all over Ghana. The data is about monthly average of rainfall from January 1995 and December 2009. (Unit: millimeter)

1.9 OUTLINE OF DISSERTATION

This section deals with the outline of the project:-

Chapter 1, will talk about the background and overview of Times series and Rainfall.

Chapter 2, talks about the literature that was reviewed to help the writer on this thesis.

Chapter 3, talks about the theory behind the methods that will be used in analysis of the data collected.

Chapter 4, will contain analysis from the four models which will be made using SAS.

Chapter 5 will be on Conclusion and recommendations.

CHAPTER 2

LITERATURE REVIEW

2. 1 INTRODUCTION

The Quantitative Precipitation Forecast (abbreviated QPF) is the expected amount of liquid precipitation accumulated over a specified time period over a specified area. A QPF will be specified when a measurable precipitation type reaching a minimum threshold is forecast for any hour during a QPF valid period. Precipitation forecasts tend to be bound by synoptic hours such as 0000, 0600, 1200 and 1800 GMT. Terrain is considered in QPFs by use of topography or based upon climatological precipitation patterns from observations with fine detail. Starting in the mid to late 1990s, QPFs were used within hydrologic forecast models to simulate impact to rivers throughout the United States. Forecast models show significant sensitivity to humidity levels within the planetary boundary layer, or in the lowest levels of the atmosphere, which decreases with height. QPF can be generated on a quantitative, forecasting amounts, or a qualitative, forecasting the probability of a specific amount, basis. Radar imagery forecasting techniques show higher skill than model forecasts within 6 to 7 hours of the time of the radar image. The forecasts can be verified through use of rain gauge measurements, weather radar estimates, or a combination of both. Various skill scores can be determined to measure the value of the rainfall forecast.

The urban heat island effect leads to increased rainfall, both in amounts and intensity, downwind of cities. Global warming is also causing changes in the precipitation pattern globally, including wetter conditions across eastern North America and drier conditions in the tropics. Precipitation is a major component of the water cycle, and is responsible for depositing most of the fresh water on the planet. The globally-averaged annual precipitation is 990 millimeters (39 in). Climate classification systems such as the Köppen climate classification system use average annual rainfall to help differentiate between differing climate regimes. Antarctica is the Earth's driest continent. Rain is also known or suspected on other worlds, composed of methane, iron, neon, and sulfuric acid rather than water.

Although drought is seen as an extreme event, long periods of low rainfall are common in Australia. The most recent drought story begins in 2001-02, when drought began in areas of south-western Queensland in 2002-03. Extreme drought occurred across much of eastern Australia, further exacerbating drought conditions in those areas (McKeon and Hall, 2000; McKeon *et al*., 2004; Sivakumar and Ndegwa, 2007). Following average conditions in 2003-04, severe drought returned in many regions in 2004-05. For many regions of Australia, the overall five-year period from April 2000 to March 2005 represents extremely low rainfall compared to the historical records commencing in 1890.

A similar but more cautionary story emerges from an analysis of the last 40 years of rainfall records. In much of Australia, this recent drought started after a sequence of above-average years of rainfall from the second half of 1998 to the first half of 2001. Central coastal Queensland and south-west Western Australia had already experienced drier conditions for at least 15 years. For example, eastern Australia received significantly less rainfall during the three years from 2002-2005 than during 1961-90 (the current international standard reference period). Coastal areas experienced the greatest rainfall deficits (the difference between actual rainfall in a year and the long-term average). In contrast, during the same 2002-05 period, rainfall in the north of the Northern Territory and in parts of north-western Western Australia was significantly greater than that experienced from 1961-90 (Bureau of Meteorology, 2005). The recent drought may be unusual in that it has been warmer than previous droughts in the last 50 years (the length of temperature records).

The enhanced greenhouse scenario suggests that temperatures in Australia may rise by 1-0-2°C, summer rainfall may increase and the frequency of high rainfall and flooding events may also increase (Whetton *et al*., 1993; 1994). Hence, there is growing awareness of the possible consequences of global climate change. High rainfall may be a blessing to Australia but this has not been noted in recent times. Indeed, such a scenario is difficult to grasp for those living in Queensland and parts of eastern Australia that have been subject to a 'severe and persistent' drought since 1991 (Bureau of Meteorology, 1995).

Observational studies, however, lend some support to these projected changes. Over the period 1910 to 1988, summer rainfall appears to have increased over much of eastern Australia (Nicholls and Lavery, 1992). This increase occurred rather abruptly around 1950, confirming earlier findings (Pittock, 1975). There is also evidence that annual rainfall intensity and the frequency of heavy rainfall events have increased over tropical Australia over the period 1910-1989 (Suppiah and Hennessy, 1996). Observational studies also provide evidence of temperature increases over Australia (Jones *et al*., 1990; Plummer, 1991).

Temperatures have been equally variable. The average temperature across Australia has risen by 0.82°C between 1910 and 2004 with much of the warming occurring in the second half of the twentieth century. The warmest year on record is 2005. Until 2004, the warmest year had been 1998. These temperature changes have been greater for minimum than maximum temperatures with a consequent decline in the Daily Temperature Range (DTR) in recent decades (Plummer *et al*., 1995), matching trends in DTR found in other parts of the world (Karl *et al*., 1993). Some of these trends in Australian climate over recent decades have also been identified in climate regions of the south-west Pacific (Salinger *et al*., 1995).

While there has been research on rainfall and temperature interactions the majority of them was conducted in earlier times and therefore notes able to benefit from the large changes that may have occurred in recent times when climate change issue has come to the fore. Not only that, the literature review showed little work has been done in coastal Queensland on this area given that climate change has been a major driving force for research. As noted earlier, this study uses an updated dataset and time series techniques to further understand the relationship between rainfall and temperature in coastal Queensland (Singhtaun and Charnsethikul, 2007).

Case study: According to Jordanian Ministry of Water and Irrigation Jordan is located 80 kilometers east of the eastern coast of the Mediterranean Sea. Its location between 29°11'N and 33°22'N and between 34°19'E and 39°18'E with an area of 89329 km2. In Jordan, more than 80% of the country is classified as arid areas with an average of rainfall ranges from 600 mm years-1 in the north to less than 50 mm year-1 in the south. The precipitation pattern is both latitude and altitude dependent. In addition, water resources in Jordan are limited and with deteriorating quality due to urban development. Therefore, it is important to know the future water resources budget in order to help decision makers improve their decisions with taking consideration the available and future water resources (Naill, and Momani, 2009).

Additionally, using modeling and forecasting for future water resources becomes possible with advances in forecasting methodologies such as time series analysis. The rainy season is between October and May where 80% of the annual rainfall occurs through December to March. Jordan witnessed rainy seasons above average for the years 1970/1971 and 1991/1992 where the last one considered the highest in the last 75 years.

The climate in Jordan is predominantly of the Mediterranean type. hot and dry summer and cool wet winter with two short transitional periods in autumn and spring: Four climatic regions are distinguishable in Jordan. They are:

- ✓ The Jordan rift valley (Al-Ghor): The climate of Al-Ghor is classified as tropical. It is very hot in summer and warm in winter with an annual rainfall of 150-250 mm. The elevation of the Ghor is below Mean Sea Level ranging from 200-400 m. Its width ranges from 15 km in the North to 30 km in the South
- ✓ The mountainous (hilly) region: The climate of these Regions is rather mild in summer and cold in winter. The amount of rainfall ranges from 300-600 mm year-1. Snowfall occurs over the mountains. This region lies to the east of the Jordan rift valley extending from North to South. Its elevation varies from 750-1200 m with some tops exceeding 1700 m.

- ✓ The badia region: A flat terrain that lies to the east of the high lands with an elevation varies from 600-700m. It is characterized by dry hot summer and relatively cold dry winters with rainfall varying from 40-100 mm year-1.
- ✓ The Gulf of Aqaba: This area is considered very hot in summer and warm in winter with an amount of rainfall <50 mm year-1 Amman is the capital city of Jordan where more than one million people live. Since rainfall data for this station is around the average among these stations it is used as a case study in our analysis.

In an agricultural country like India, the success or failure of crops and water scarcity in any year is a matter of greatest concern and these problems are highly associated with the behavior of the summer monsoon rainfall. Mean monsoon rainfall over India, as a whole during June–September, is 88 cm with a coefficient of variation of 10% (Rajeevan, 2004). Accurate, long lead prediction of monsoon rainfall can improve planning to mitigate the adverse impacts of monsoon variability and to take advantage of favorable conditions (Reddy and Salvekar, 2003). The summer monsoon precipitation over India is dominated by the semi-permanent monsoon trough, which extends from West Pakistan to the North Bay of Bengal across Northwest India and the Westward moving synoptic disturbances developing over the North Bay of Bengal (Mohanty and Mohapatra, 2007). In the paper by Rajeevan (2001), the status and future prospects of long-range forecasts of Indian summer monsoon have been reviewed. In another work, Krishnamurthy and Shukla (2000) studied the inter seasonal and inter annual variability of summer monsoon rainfall over India. In the following paragraphs, we present a survey of the studies on rainfall time series over different parts of the world. Subsequently, we have mentioned the newness in our study with respect to the studies available on monsoon rainfall forecasting over India.

The study of rainfall time series is a topic of great interest in the field of climatology and hydrology. Some significant examples in such areas include Delleur and Kavvas (1978), Shin et al. (1990), Singh (1998). Both univariate (e.g. Soltani et al., 2007) and multivariate (Grimaldi et al., 2005) approaches have been attempted to model the rainfall time series. Impact of other atmospheric variables on rainfall has been discussed in various literatures (e.g. Cracknell and Varotsos, 2007; Varotsos, 2005; Chattopadhyay, 2007b). The association between rainfall and agro meteorological processes is well discussed (e.g. Jhajharia et al., 2009; Chattopadhyay et al., 2009).

Several stochastic models were attempted to forecast the occurrence of rainfall, to investigate its seasonal variability and to forecast monthly/yearly rainfall over some given geographical area. Study of the rainfall is interesting because of the associated problems, such as forecasting, corrosion effects and climate variability and various literatures have discussed these issues (e.g. Kondratyev et al., 1995; Kondratyev and Varotsos, 2002; Ferm et al., 2005, 2006; Tzanis and Varotsos, 2008). In a study by Chin (1977), where daily precipitation records for 25 years at more than 100 stations in the conterminous United States were analyzed, it was proved that the proper Markov order describing the daily precipitation process has to be determined and cannot be assumed a priori. Gregory et al. (1993) applied a Markov chain model to investigate inter annual variability of area averaged total precipitation. Wilks (1998) applied mixed exponential distribution to simulate precipitation amount at multiple sites exhibiting realistic spatial correlation. Chaotic features associated with the atmospheric phenomena have attracted the attention of modern scientists (e.g., Varotsos et al., 2007, Varotsos and Krik-Davidoff, 2006, Khan et al., 2005; Bandyopadhyay and Chattopadhyay, 2007). In recent times, the study of the possible presence of chaotic behavior in rainfall time series has been of much interest (Sivakumar, 2001). Mathematical tools based on the theoretical concepts underlying the methodologies for detection and modeling of dynamical and chaotic components within a hydrological time series have been studied extensively by various scientists like Islam and Sivakumar (2002); Khan et al. (2005) and Jayawardena and Lai (1994). The existence of deterministic chaos within rainfall time series is well documented in the literature (e.g. Rodriguez-Iturbe et al., 1989; Sharifi et al., 1990).

Phase space reconstruction and artificial neural networks (ANN) are non-linear predictive tools that have been proposed in the modern literature as effective mathematical methodologies to be useful to hydrological time series characterized by chaotic features (Chattopadhyay and Chattopadhyay, 2008a; Elsner and Tsonis, 1992; Khan et al., 2005). The suitability of ANN over conventional statistical approaches has been demonstrated in many research papers dealing with hydrological processes (e.g. Chattopadhyay, 2007a ; Chattopadhyay and Chattopadhyay, 2008a). Applicability of ANN to rainfall time series is well documented in the literature. In recent times, the competence of ANN(Rojas,1996) in forecasting chaotic time series has been established by several authors (e.g. Principe et al., 1992; Oliveira et al., 2000; Silverman and Dracup, 2000). Prediction of atmospheric events, especially rainfall, has benefited significantly by voluminous developments in the application field of ANN and rainfall events and quantities have been predicted statistically (e.g. DelSole and Shukla, 2002; Mohanty and Mohapatra, 2007). The advantages of ANN over traditional statistical and numerical weather prediction approaches have been discussed by McCann (1992), Kuligowski and Barros (1998) and Silverman and Dracup (2000). Several research papers are available where the suitability of the ANN approach has been established quantitatively over conventional statistical rainfall prediction procedures (e.g. Chattopadhyay, 2007a; Hastenrath, 1995; Toth et al., 2000; Ramirez et al., 2005; Chattopadhyay, 2007b; Chattopadhyay and Chattopadhyay, 2008b). Guhathakurata (2008) generated an ANN based model that captured the input-output non-linear relationship and predicted the monsoon rainfall in India quite accurately. The purpose of the present article is to investigate the stationarity within the average monsoon rainfall time series in India and subsequently to model this time series through autoregressive approach.

The Asian monsoon circulation influences most of the tropics and subtropics of the Eastern Hemisphere and a major portion of the Earth's population (Chattopadhyay, 2007b). The southwest (summer) and the northeast (winter) monsoons influence weather and climate between

30N and 30S over the African, Indian and Asian land masses (Reddy and Salvekar, 2003, Chattopadhyay, 2007b). The variability in the monsoon rainfall depends heavily upon the sea surface temperature anomaly over the Indian Ocean (Clark et al., 2000). As the extra tropical circulation anomalies display energy dispersion away from the region of anomalous tropical convection, they have been taken to mean a Rossby wave response to the latent heat release associated with the tropical convection (Ferranti et al., 1990). In regions of anomalous tropical heating, there is a dynamical response with anomalous large-scale ascent and upper tropospheric divergence, which acts as a Rossby wave source (Sardeshmukh and Hoskins, 1988) for extra tropical waves. The summer monsoon (June–August) is the most productive period in India with respect to its agricultural practices. Moreover, the conserved rainwater for this period is used for future irrigation purposes in this country. Therefore, forecasting of average summer monsoon rainfall is necessary for future agricultural and irrigation modeling over this country. A plethora of literature is available where the summer-monsoon rainfall over this country has been predicted through multivariate approach (e.g. Chattopadhyay, 2007a, 2007b; Chattopadhyay and Chattopadhyay, 2008a, 2008b and references therein). However, the multivariate approach requires various other parameters which themselves are characterized by chaotic properties.

The autoregressive approach, which is a univariate approach, depends solely upon the concerned variable and therefore is free from the effect of other variables.

Dahale and Singh (1993) adopted autoregressive approach to monsoon rainfall time series over India and identified third order autoregressive model as the best fit. In the present paper, we have viewed the autocorrelation structure of the monsoon rainfall time series and consequently adopted autoregressive integrated moving average (ARIMA) approach instead of tradition autoregressive (AR) approach. Finally, we have implemented ANN in autoregressive manner and compared its performance statistically with the ARIMA-based model. The methodology and implementation procedure are described in the subsequent sections.

A time series analysis is a valuable tool to get information about analyzed data structures and their components, which provides a good basis for successful future predictions. The formation of predictive linear models for time series of discharges in Slovakia have been dealt with by, e.g. Pekárová et al. (2004), Pekárová et al. (2005) as Komorníková and Szökeová (2004a), who tested mean monthly and annual discharges on the Kysuca River. Jones and Smart (2005) as well as Worrall and Burt (1999) dealt with autoregressive modeling too. Tesfaye, et al. (2006) used seasonal ARMA models for the identification of mean annual discharges. The seasonal models were analyzed by Jones and Brelsford (1967), Pogano (1978), Troutman (1979), Anderson and Vecchia (1993), Ula (1990 a, b), Ula and Smaldi (1997), Shao and Lund (2004) and Tesfaye, et al. (2006) too.

The testing of mean monthly discharge time series with Long Memory models was done by Komorníková and Szökeová (2004b). Állóová (2006) tested time series of mean monthly precipitation totals data in the regions of the Belianske Lúky, Abrod and Kláštorské Lúky wetlands. Amendola (2003) dealt with the prediction of mean rainfall data using regime switching models. These models are suitable for time series analysis of extreme events.

CHAPTER 3

METHODOLOGY

3.0 INTRODUCTION

In this chapter we discuss the theory of time series in terms of its definition, types of time series models as well as the theory of time series.

3.1 DEFINITION

Time series is a time dependent sequence denoted $Y_1, Y_2, \dots, Y_t$ or Y_t where $t \in N$ where 1, 2,…n denote time steps.

3.1.1 Deterministic Time Series

If from past knowledge, the future of a time series can be exactly predicted, it is a deterministic series and requires no further investigation. It can be expressed as a known function. That is $Y_t = f(t)$.

3.1.2 Stochastic Time Series

If a time series can be expressed as $Y_t = X(t)$, where X is a random variable, then Y_t is a stochastic time series.

3.2 OBJECTIVES OF TIME SERIES ANALYSIS

The main objectives of analyzing a time series are classified as description, explanation, prediction and control.

3.2.1 Description

When presented with time series data, the first step in the analysis is usually to plot the data to obtain simple descriptive measure of the main properties of the series as seasonal effect, trend etc.

Apart from trend and seasonal variations, the outliers to look for in the graph of the time series are the possible presence of turning points, where for example, an upward trend has suddenly changed to a downward trend.

3.2.2 Explanation

When observations are taken on two or more variables, it may be possible to use the variable in one time series variable to explain the variation in the other time series variable. This may give a deeper understanding of the mechanism which generated a given time series. For example, sales are affected by price and economic condition.

3.2.3 Prediction

Given an observed time series one may want to predict the future values of the series. This is an important task in sales forecasting and in the analysis of economic and industrial time series. Prediction is closely related to control problems in many situations. For example if we can predict that manufacturing process is going to move off target, then appropriate corrective action can be taken.

3.2.4 Control

When a time series is generated which measures the quality of a manufacturing process, the aim of the analysis may be to control the process. In statistical quality control, the observations are plotted on control charts and the controller takes action as a result of studying the charts. Box and Jenkins have described a more sophisticated control strategy which is based on fitting a stochastic model to the series from which future values of the series are predicted. The values of the process variables predicted by the model are taken as target values and the variables conform to the target values.

3.3 COMPONENTS OF TIME SERIES

Traditional methods of time series analysis are mainly concerned with decomposing the variation in series into the various components of trend, periodic and stochastic.

3.3.1 Periodic Component

If $Y_t = Y_{t+T} + e_t$, for all $t \in N$, then the time series has a periodic component of period T.

3.3.2 Trend Component

If $Y_t = y + \beta_t + e_t$, then there exist a linear trend with the slope being β.

3.4 STATIONARY TIME SERIES

A time series is said to be stationary if the joint distribution of $X_{t_1}, \ldots, X_{t_n}$ is the same as the joint distribution of $X_{t_1+T}, \ldots, X_{t_n+T}$ for all $t_{1+T}, \ldots, t_{n+T}$. In other words shifting the time origin by an amount *T* has no effect on the joint distribution which must therefore depend only on the intervals between $t_1, \ldots, t_n$.

3.4.1 Autocorrelation Function (ACF)

The autocorrelation function measures the degree of correlation between neighboring observations in a time series. The autocorrelation function at lag k is defined as

$$\rho_k = \frac{E[(Y_t + \mu_y)(Y_{t+k} - \mu_y)]}{[E(Y_t + \mu_y)^2 E(Y_{t+k} - \mu_y)^2]} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 3.1$$

$$\rho_k = \frac{co(Y_t, Y_{t+k})}{\sigma_{t+k}} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 3.2$$

The autocorrelation coefficient is estimated from sample observation using the formula;

$$r_k = \frac{\sum_{t=2}^{n}(Y_t - \mu_y)(Y_{t+k} - \mu_y)}{\sum_{t=1}^{n}(Y_t - \mu_t)^2} \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots 3.3$$

(Hamilton J.D., 1994)

3.4.2 Partial Autocorrelation Coefficient

Partial autocorrelation function measures the degree of association between Y_t and Y_{t+k} when the effects of other time lags on Y are held constant. The partial autocorrelation function PACF denoted by $\{\emptyset_{kk}: k = 1,2, \ldots \ldots\}$ the set of partial autocorrelation at various lags k are defined by

$\emptyset_{kk} = \frac{|P_k^*|}{|P_k|}$ where P_k is the KxK autocorrelation matrix and P_k^* is P_k with the last column replaced by $[P_1, P_2, \dots, P_k]^T$ and an example is $\emptyset_{11} = \emptyset_1 = P_1$

and $$\emptyset_{22} = \frac{\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{bmatrix}}{\begin{bmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{bmatrix}} = \frac{\rho_2 - \rho_1^2}{1-\rho_1^2} \dots\dots\dots\dots\dots\dots\dots\dots 3.4$$

And estimates of $\emptyset_{kk}$ can be obtained by replacing the ρ_i by r_i.

3.4.3 An Autoregressive Model of Order p [AR(p)]

An autoregressive model of order p denoted by AR(p) is a special kind of regressive in which the explanatory variables are past values of the process. An autoregressive model of order p is given by

$$Y_t = \sum_{k=1}^{p} \alpha_k Y_{t-k} + \mu + e_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 3.5$$

where μ is the mean of the time series data and e_t is the white noise.

The order of an AR(p) process is determined by the partial autocorrelation function (PACF). An AR(p) process has its PACF cutting off after lag p and the ACF decays. For example the PACF of an AR(1) process cuts off after lag one (1). *(Hamilton J.D, 1994)*

3.4.4 Autoregressive Process of Order one (1) AR(1)

The AR(1) process is

$$Y_t = \alpha_1 Y_{t-k} + \mu + e_t$$

Putting $\mu = 0$ we have

$$Y_t = \alpha_1 Y_{t-k} + e_t$$

Multiplying through by Y_{t-k} we have

$$Y_{t-k} Y_t = \alpha_1 Y_{t-k} Y_{t-k} + e_t Y_{t-k}$$

$$cov\ (Y_{t-k}, e_t) = \alpha_1 cov(Y_{t-k}, Y_{t-1}) + cov(Y_{t-k}, e_t)$$

But $cov\ (Y_{t-k}, e_t) = 0$ since Y_{t-k} depends only on $e_{t-k}, e_{t-k-1}, \dots \dots$ which are not correlated with e_t for k>0. Hence

$$\gamma_k = \alpha_1 \gamma_{k-1}$$

Dividing through by γ_0 we have $\frac{\gamma_k}{\gamma_0} = \alpha \frac{\gamma_{k-1}}{\gamma_0}$

$$\rho_k = \alpha_1 \rho_{k-1}\ where\ \rho_0 = 1$$

We have $\rho_1 = \alpha_1 \rho_0 = \alpha_0\ since\ (\rho_0 = 1)$

$$\rho_1 = \alpha_1$$

For k = 2

$$\rho_2 = \alpha_1 \rho_1 = \alpha_1(\alpha_1) = \alpha_1^2$$

For k = 3

$$\rho_3 = \alpha_1\rho_2 = \alpha_1(\alpha_1^2) = \alpha_1^3$$

And in general (*Box and Jenkins, 1971)*

$$\rho_k = \alpha_1^k$$

3.4.5 Moving Average of Order q MA(q)

MA models provide predictions of Y_t based on a linear combination of past forecast errors. In particular the MA model of order q is given by *(Hamilton J.D., 1994)*

$$Y_t = \sum_{k=1}^{q} \theta_k e_{t-k} + \mu + e_t \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 3.6$$

3.4.6 Autocorrelation Function (ACF) of MA(q)

$$\gamma(k) = cov(Y_t, Y_{t+k})$$

=

$$cov\left(\theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q}, \theta_1 e_{t-k-1} + \theta_2 e_{t+k-2} + \cdots + \theta_q e_{t+k-q}\right) \dots\dots..3.7$$

$$= \begin{cases} 0 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . k > q \\ \emptyset_e^2 \sum_{t-1}^{q-k} \theta_i \theta_{i+k} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots .\, k = 0,1, \dots q \\ \gamma(-k) \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . k < 0 \end{cases}$$

Since

$$cov(e_s, e_t) = \begin{cases} \sigma_e^2 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots s = t \\ 0 \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots . s \neq t \end{cases}$$

Hence the autocorrelation function (ACF) of MA(q) process is given by

$$\rho(k) = \begin{cases} 1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots k = 0 \\ \dfrac{\sum_{i=1}^{q-k} \theta_i \theta_{i+u}}{\sum_{i=1}^{q} \theta_i^2} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots k = 1,2, \ldots q \\ \rho(-k) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots k < 0 \end{cases} \ldots\ldots\ldots\ldots\ldots\ldots 3.8$$

The order of the MA(q) is given by the autocorrelation function. The ACF cuts after lag q and the partial autocorrelation function decays to zero. Thus an MA(1) process cuts off after lag one. In other words the ACF after lag one will not be significantly different from zero.

3.4.7 Moving Average process of Order one MA(1)

The MA(1) process is given by

$$Y_t = \theta_1 e_{t-1} + \mu + e_t$$

And its autocorrelation is given by

$$\rho(k) = \begin{cases} 1 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots k = 0 \\ \dfrac{\theta_1}{(1+\theta^2)} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots k = +1 \\ 0 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots otherwise \end{cases}$$

Thus

$$\rho_1 = \frac{\theta_1}{1+\theta_1^2}$$

$$\rho_1 + \rho_1 \theta_1^2 - \theta_1 = 0$$

$$= \rho_1 \theta_1^2 - \theta_1 = 0$$

The parameters are thus roots of a quadratic. This means that we can find two MA(1) processes that corresponds to the same ACF. To establish a one-to-one correspondence between the ACF and the model and obtain a converging autoregressive representation, we restrict the moving average parameter such that $|\theta| < 1$. This restriction, known as the invertibility, implies that the process can be written in terms of an autoregressive model. *(Hamilton J.D., 1994)*

3.4.8 The Duality of AR and MA processes

We show that the Random Walk process given by

$$Y_t = Y_{t-1} + e_t$$

Can be rewritten as an infinite moving average. Indeed, consider the following moving average,

$$Y_t = e_t + e_{t-1} + e_{t-2} + \cdots = \sum_{t=0}^{\infty} e_{t-i}$$

$= (1 + B + B^2 + B^3 + \ldots)\, e_t$

$= (\sum_{t=0}^{\infty} B^i)\, e_t$……………………………3.9

We recall that $(\sum_{t=0}^{\infty} y^t) = 1/(1 - y)$ is valid when $|y| < 1$

Hence $Y_t = \left(\frac{1}{1-B}\right) e_t$

So that

$(1 - B)\, Y_t = e_t$

$Yt - BY_t = e_t$

$$Yt - Y_{t-1} = e_t$$

$$Yt = Y_{t-1} + e_t$$

This is the random walk process. This means that a finite autoregressive process. For example, we show that an MA(1) process is an infinite autoregressive process. For such a process,

$$Y_t = e_t - \theta_1 e_{t-1}$$

Using the B operator notation, we have

$$Y_t = (1 - \theta_1 B)\, e_t$$

$$\frac{Yt}{1-\theta 1B} = e_t$$

$$(1 + \theta_1 B + \theta\frac{2}{2}B^2 + \ldots)\ Y_t = e_t$$

$$Y_t + \theta_1 Y_{t-1} + \theta\frac{2}{2}Y_{t-2} + \ldots = e_t$$

This is an infinite autoregressive process.

3.4.9 Autoregressive Moving Average Model (ARMA)

A more general model is a mixture of the AR(p) and MA(q) models and is called an autoregressive moving average model (ARMA) of order (p,q).

The ARMA(p,q) is given by

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{i-1} + \sum_{i=1}^{q} \theta_i e_{i-1} + \mu + e_t \text{ ---3.10}$$

An example of an ARMA(1,1)

$$Y_t = \alpha_i Y_{i-1} + \theta_i e_{i-1} + \mu + e_t \text{---3.11}$$

An important characteristic of ARMA models is that both the ACF and PACF do not cut off as in AR and MA models, *(Box and Jenkins, 1971).*

3.4.10 ARMA or "Mixed" Process

Consider the process given by:

$$Y_t = \alpha_1 Y_{t-1} + \theta_1 e_{t-1} + e_1$$

This can be rewritten as

$Y_t - \alpha_1 Y_{t-1} = e_t + \theta_1 e_{t-1}$ or

$$(1 - \alpha\beta) Y_t = (1 + \theta\beta)\ e_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.12$$

$$AR(\beta)\ Y_t = MA(\beta)\ e_t$$

This is called a mixed or autoregressive moving average (ARMA) process of order (1,1). Since equation (1) is ARMA(1,1) if $|\theta| < 1$, it can be rewritten as

$$(1 - \alpha\beta)\left(\frac{1}{1 + \theta\beta}\right) Y_t = e_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.13$$

$$((1 - \alpha\beta)(1 - \theta\beta + \theta^2\beta^2 + \theta^3\beta^3 +) Y_t = e_t$$

$$[(1 - \alpha + \theta)\beta + (\alpha\theta + \theta^2)\ \beta^2 + \ \ldots]\ Y_t = e_t$$

There is an infinite order AR process. This is true if $|\alpha| < 1$ and $|\theta| < 1$ i.e. if the AR is stationary and MA is invertible. If we have two polynomial in B, MA(B) and AR(B), and an ARMA model.

3.4.11 The Autoregressive Integrated Moving Average Model (ARIMA)

If a non-stationary time series which has variation in the mean is differenced to remove the variation the resulting time series is called an integrated time series. It is called an integrated model because the stationary model which is fitted to the differenced data has to be summed or integrated to provide a model for the non-stationary data. Rotationally, all AR(p) and MA(q) models can be represented as ARIMA(1,0,0) that is no differencing and no MA part.

The general model is ARIMA(p,d,q) where p is the order of the AR part, d is the degree of differencing and q is the order of the MA part.

Writing $W_t = \nabla^d Y_t = (1-B)^d Y_t$ ………………………………………3.14

The general ARIMA process is of the form

$$W_t = \sum_{t=1}^{p} \alpha_i W_{t-i} + \sum_{t=1}^{q} \theta_1 e_{t-i} + \mu + e_t \text{ ……………………………..3.15}$$

3.4.12 ARIMA(1,1,1) Process

An example of ARIMA(p,d,q) is the ARIMA(1,1,1) which has one autoregressive parameter, one level of differencing and one MA parameter is given by

$$W_t = \alpha_i W_{t-i} + \theta_1 e_{t-i} + \mu + e_t$$

$$(1-B)Y_t = \alpha_1 (1-B)Y_{t-1} + \theta_1 e_{t-i} + \mu + e_t \text{ ………………………………………3.16}$$

Which can be simplified further as

$$Y_t - Y_{t-1} = \alpha_1 Y_{t-1} - \alpha_1 Y_{t-2} + \theta_1 e_{t-i} + \mu + e_t$$

$$Y_t - Y_{t-1} = \alpha_1 (Y_{t-1} - Y_{t-2}) + \theta_1 e_{t-i} + \mu + e_t \text{ ………………………………………3.17}$$

3.4.13 Estimating the parameters of an ARIMA Model

In practice most time series are non-stationary and the series is differenced until the series becomes stationary. An AR, MA or ARMA model is fitted to the differenced series and estimation procedures are as described for the AR, MA, and ARMA above.

3.4.14 Stationarity and Invertibility Conditions of Specific Time Series model

In the Table below we display the stationarity and invertibility conditions of specific time series models and the behavior of their theoretical ACF and PACF functions.

3.4.15 The Box-Jenkins Method of Modeling time Series

The Box-Jenkins methodology is a statistical sophisticated way of analyzing and building a forecasting model which best represents a time series. The first stage is the identification of the appropriate ARIMA models through the study of the autocorrelation and partial autocorrelation functions. For example if the partial autocorrelation cuts off after lag one and the autocorrelation function decays then ARIMA(1,0,0) is identified. The next stage is to estimates the parameters of the ARIMA model chosen.

The third stage is the diagnostic checking of the model. The Q-statistic is used for the model adequacy check. If the model is not adequate then the forecaster goes to stage one to identify an alternative model and it is tested for adequacy and if adequacy then the forecaster goes to the final stage of the process.

The fourth stage is where the analysis uses the model chosen to forecast and the process ends.

Below is a schematic representation of the Box-Jenkins process.

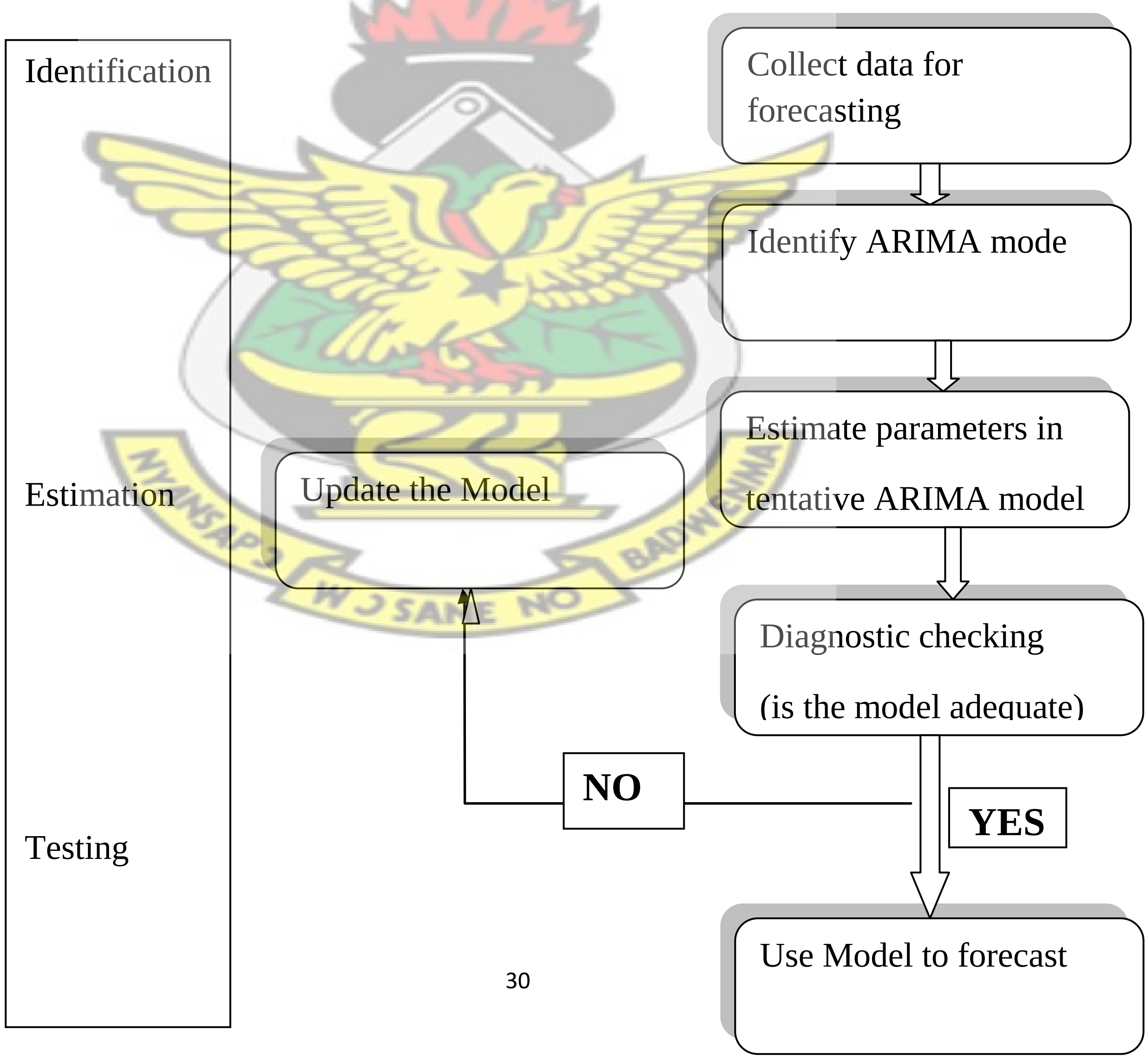

Figure 3.13 Box-Jenkins Process

3.4.16 Identification techniques

Identification methods are rough procedures applied to a set of data indicate the kind of representational model that will be further investigated. The aim here is to obtain some idea of the values p, d and q needed in the general linear ARIMA model and to obtain initial estimates for the parameters.

The task here is to identify an appropriate subclass of models from the general ARIMA family α (B) $\nabla^d Y_t = \theta(\beta)\ e_t$ which may be used to represent a given time series. The approach will be as follows;

(a) To differentiate Y_t as many times as is needed to produce stationary, reducing the process under study to the mixed autoregressive moving average process

$$\alpha\ (B)\ W_t = \theta_0 + \theta(B)\ e_t \quad where\ W_t = (1 - B)\ d\ Y_t = \nabla^d Y_t$$

(b) To identify the resulting ARMA process

The principle tools for putting (a) and (b) into effect is the sample autocorrelation function and the sample partial autocorrelation function. Apart from helping to guess the form of the model, they are used to obtain approximate estimates of the parameters of the model. These approximations are useful at the estimates stage to provide starting values for iterative procedures employed at that stage.

3.4.17 Use of the autocorrelation and Partial Autocorrelation functions in Identification

A stationary mixed autoregressive moving average process of order (p, 0, q),

$\alpha(B)Y_t = \theta(B)e_t$, the autocorrelation function satisfies the different equation

$$\alpha(B) P_k = 0 \qquad k > q$$

Also, if

$$\alpha(B) = \prod_{i=1}^{p} (1 - G_i B)$$

The solution of this difference equation for the kth autocorrelation is, assuming distinct roots, of the form

$$P_k = A_1 G_1^k + A_2 G_2^k + \ldots + A_p G_p^k \qquad k > q - p$$

The stationarity requirement that the zeros of $\alpha(B)$ lie outside the unit circle implies that the roots $G_1, G_2, G_3, \ldots, G_k$, lie inside the unit circle. Inspection of the equation

$$P_k = A_1 G_1^k + A_2 G_2^k + \ldots + A_p G_p^k \qquad k > q - p$$

shows that in the case of a stationary model in which none of the roots lie close to the boundary of the unit circle, the autocorrelation function will quickly "die out" or decay for moderate and large k.

Suppose that a single real root, say G_1 approaches unity, so that $G_1 = 1 - \delta$ where δ is a small positive quantity. Then, since for k large, $P_K = A_1(1 - k\delta)$ the autocorrelation function will not die out quickly and will fall off slowly and very nearly linearly. Similarly if more than one root approaches unity the autocorrelation function will decay slowly. Therefore if the autocorrelation function dies out slowly it implies there is at least a root which approaches unity. As a result failure of the estimated autocorrelation function to die out rapidly might logically suggest that the underlying stochastic process is non-stationary in Y_t but possible stationary in $\nabla^d Y_t$, or in some higher difference.

It is therefore assumed that the degree of differencing d_1 necessary to achieve stationarity has been reached when the autocorrelation function of $W_t = \nabla^d Y_t$ lie out fairly quickly.

3.4.18 Forecasting

The fourth stage of the Box-Jenkins approach is to forecast with model selected. Suppose the model chosen to fit a hypothetical data is

$$Y_t = Y_{t-1} + \alpha_1 (Y_{t-1} - Y_{t-2}) + e_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.19$$

And suppose further that the data is of length 60, $\alpha = 0.2178$

$Y_{60} = 131.2 \qquad\qquad Y_{59} = 134.8$

Then $Y_{61} = Y_{60} + 0.2178\,(Y_{60} - Y_{59})$

$Y_{61} = 131.2 + 0.2178\,(131.2 - 134.8)$

$Y_{61} = 130.097$

Hence, a forecast value for period 61 is 130.097.

3.5 SEASONAL EXPONENTIAL SMOOTHING

It is used when the time series exhibits seasonality but no trend is present. The parameters estimated from this model are the seasonal smoothing weight, the level smoothing weight and the seasonal smoothing factors. A seasonal weight that is close in value to one implies that a non-seasonal model might be more appropriate. Whereas, a seasonal weight near zero implies that deterministic seasonal factors might be present.

The seasonal exponential smoothing model is given by the **model equation**

$$Y_t = \mu_t + S_p(t) + \varepsilon_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.20$$

The **smoothing equations are**

$$L_t = \alpha\left(Y_t - S_{t-p}\right) + (1-\alpha)L_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.21$$

$$S_t = \delta(Y_t - L_t) + (1-\delta)S_{t-p} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.22$$

The k-**step-ahead prediction equation** is

$$\ddot{Y}_t(k) = L_t + S_{t-p+k} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.23$$

That is, you forecast y k-steps ahead by using taking the last available estimated level state and then add the last available smoothed seasonal factor, S_{t-p+k}□, which matches the month of the forecast horizon.

The smoothing weights consist of the following at time *t*:

L_t is a smoothed level of the series that estimates μ_t

$S_{t-j}, j = 0, \ldots p-1$ are seasonal factors that estimate $S_p(t)$

$\boldsymbol{Y_t} =$ the value of the observation at time t

$\alpha =$ a level smoothing weight.

$\delta =$ a seasonal smoothing weight

3.6 EXPONENTIAL SMOOTHING

The formulation of exponential smoothing forecasting methods arose in the 1950's from the original work of Brown (1959, 1962) and Holt (1960) who were working on creating forecasting models for inventory control systems. One of the basic ideas of smoothing models is to construct forecasts of future values as weighted averages of past observations with the more recent observations carrying more weight in determining forecasts than observations in the more distant past. By forming forecasts based on weighted averages we are using a "smoothing" method. The adjective "exponential" derives from the fact that some of the exponential smoothing models not only have weights that diminish with time but they do so in an exponential way, as in $\lambda_j = \lambda^j$ where $-1 < \lambda < 1$ and $j = 1,2, \ldots$ represents the specific period in the past.

At least three major points can be raised about exponential smoothing models:

- As a methodology, exponential smoothing methods suffer from not having an objective statistical identification and diagnostic system for evaluating the "goodness" of competing exponential smoothing models. For example, the smoothing parameters of the smoothing models are determined by fit and are not based on any statistical criteria like tests of hypotheses concerning parameters or tests for white noise in the errors produced by the model. In this sense, exponential smoothing models are ad hoc models, statistically speaking. Of course, if one continues to monitor the forecasting performance of a given exponential smoothing model, and, if the model's forecasts become more and more inaccurate over time, then one has, in a sense, an ex post evaluation method for picking and choosing between competing exponential smoothing models. The only problem is that this approach comes with a cost. Bad forecasting for a certain amount of time while learning can be expensive when, for example, dealing with inventories that run into the millions of dollars. But instead of pursuing this ex post monitoring approach, one can attempt to make a good choice of exponential smoother beforehand by using out-of-sample forecasting experiments. In this approach, the forecaster reserves some of the available data for a "horse race" between competing exponential smoothing methods. To carry these horse races out, one divides the data into two parts: the in-sample data set (say 60% of the first part of the available time series data) and with the remaining latter part of the time series assigned to the out-of-sample data set. Then one "runs" the competing exponential smoothing methods through the out-of-sample data while forecasting h-steps ahead each time (we assume h is the forecast horizon of interest) while updating the "smoothing" parameter(s) as one moves through the out-of-sample data. In the process of generating these h-step-ahead forecasts for the competing methods we can compare the competing forecasts with the actual values that we withheld as we generated our forecasts and then use the standard forecasting accuracy measures like MSE, MAE, RMSE, PMAE, etc. to choose the best (most accurate) exponential smoothing forecasting method, as indicated by the out-of-sample forecasting experiment, for further use (subject to monitoring of course).

- Most exponential smoothing methods, as we will see below, can be shown to be special cases of the class of Box-Jenkins models. For this reason, Box-Jenkins forecasters have been critical of using exponential smoothing models for forecasting. They usually say, "Why be satisfied with a special case of a Box-Jenkins model when we can fit any Box-Jenkins model we want to the data? Moreover, we can use the various diagnostic tests that are available to choose a good Box-Jenkins model without being ad hoc in our model choice." This, of course, is a very persuasive argument and is the reason why many forecasters use standard Box-Jenkins computer software for doing their forecasting.

- A counterargument to the traditional Box-Jenkins criticism of exponential smoothing methods is that, once basic decisions like the presence or absence of trend in the time series and the presence or absence of seasonality is determined, pretty accurate forecasting can be obtained even to the point of being "almost" as accurate as the "fully flexible" and non-ad hoc Box-Jenkins models. The fact that carefully chosen exponential smoothing models do almost as well as Box-Jenkins models has been documented in two large-scale empirical studies by Makridakis and Hibon (1979) and Makridakis et. al. (1982). So if the Box-Jenkins models are more time consuming to build yet only yield marginal gains in forecasting accuracy relative to less time-consuming well informed choices of exponential smoothing models, we have an interesting tradeoff between time (which is money) and accuracy (which, of course, is also money). This trade-off falls in the favor of exponential smoothing models sometimes when, for example, one is working with 1500 product lines to forecast and has only a limited time to build forecasting models for them. In what we will argue below, an informed choice consists of knowing whether the data in hand has a trend in it or not and seasonality in it or not. Once these basic decisions are made (and if they are correct!), then pretty accurate forecasts are likely via the appropriately chosen exponential smoothing method.

3.6.1 SIMPLE EXPONENTIAL SMOOTHING (SES)

It is one of the most commonly used forecasting methods that are very good for forecasting a few periods. This model works best if there is no trend and no seasonality. Similar to moving average forecasting models, SES suppresses short- run fluctuation by smoothing the series and adding greater weights on more recent observations. The smoothing parameter is the weight that is determined for series and has the following characteristics:

1. A value between 0 and 1
2. If α=1 it becomes a naïve model; if α is close to 1, more weights are put on recent values. The model fully utilizes forecast errors.
3. If α is close to 0, distant values are given weights comparable to recent values. Choose α close to 0 when there are big random variations in the data.
4. α is often selected as to minimize the Mean Square Error

Let the observed time series be $Y_1, Y_2, \ldots, Y_n$

The equation for the Simple Exponential Smoothing model is

$$\ddot{Y}_{t+1} = S_t = \alpha Y_{t-1} + (1-\alpha)S_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.24$$

where **α**: the smoothing parameter, $0 \leq \alpha \leq 1$

$\boldsymbol{Y_t}$: the value of the observation at time t

S_t: the value of the smoothed observation at time t

.

3.6.2 THE HOLT LINEAR EXPONENTIAL SMOOTHING MODEL

It is an extension to the simple exponential smoothing model. Holt's method introduces a trend factor to the SES method to allow for forecasting series with trends. This model is used for time series that appear to have a trend but no seasonality. The Holt smoothing method utilizes two equations. Equation one adjusts for the trend of the previous time period by adding it the last smoothed value.

$$L_t = \alpha y_t + (1-\alpha)(L_{t-1} + b_{t-1}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .3.25$$

The second equation updates the trend by differencing the two previous trend values.

$$b_t = \beta(L_t - L_{t-1}) + (1-\beta)b_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .3.26$$

The forecast value is determined as the sum of the trend term and the smoothed level.

$$\ddot{Y}_{t+m} = L_t + mb_t \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.27$$

Where

L_t = Estimate of the level of the series at time t

α = smoothing constant for the data.

y_t = new observation or actual value of series in period t.

β = smoothing constant for trend estimate

b_t = estimate of the slope of the series at time t

m = periods to be forecast into the future.

The initialization of the model requires that an initial trend (b_1)and smoothing value (L_1)must be determine along with identifying α, β.

Generally,

L_1 is set to Y_1

b_1 is set to $Y_2–Y_1$, $(Y_n–Y_1)/(n–1)$, or 0.

The weight α and β are selected subjectively or by minimizing the MSE as similarly done in SES.

3.7 LINEAR TREND MODEL

Most naturally-occurring time series in business and economics are not at all stationary (at least when plotted in their original units). Instead they exhibit various kinds of trends, cycles, and seasonal patterns. For example, here is a time series (Series #2) which exhibits steady, if somewhat irregular, linear growth:

The mean model described above would obviously be inappropriate here. Many persons, upon seeing this time series, would naturally think of fitting a simple **linear trend** model--i.e., a sloping line rather than horizontal line. The forecasting equation for the linear trend model is:

$$Y(t) = \alpha + \beta t$$

where t is the time index. The parameters alpha and beta (the "intercept" and "slope" of the trend line) are usually estimated via a simple regression in which Y is the dependent variable and the time index t is the independent variable. Here is a plot of the forecasts produced by the "Linear Trend"

Although linear trend models have their uses, they are often inappropriate for business and economic data. Most naturally occurring business time series do not behave as though there are straight lines fixed in space that they are trying to follow: real trends change their slopes and/or their intercepts over time. The linear trend model tries to find the slope and intercept that give the best average fit to all the past data, and unfortunately its deviation from the data is often greatest near the end of the time series, where the forecasting action is!

If the model has succeeded in extracting the entire "signal" from the data, there should be no pattern at all in the errors: the error in the next period should not be correlated with any previous errors, and the bars on the autocorrelation plot therefore should all be close to the zero line. The linear trend model obviously fails the autocorrelation test in this case.

When trying to project an assumed linear trend into the future, we would like to know the *current* values of the slope and intercept--i.e., the values that will give the best fit to the *next few periods'* data. We will see that other forecasting models often do a better job of this than the simple linear trend model.

3.8 MEAN SQUARED ERROR

The mean squared error is arguably the most important criterion used to evaluate the performance of a predictor or an estimator. (The subtle distinction between predictors and estimators is that random variables are predicted and constants are estimated.) The mean squared error is also useful to relay the concepts of bias, precision, and accuracy in statistical estimation. In order to examine a mean squared error, you need a target of estimation or prediction, and a predictor or estimator that is a function of the data. Suppose that the target, whether a constant or a random variable, is denoted as *U*. The mean squared error of the estimator or predictor *T(Y)* for *U* is $MSE[T(Y):U] = E[(T(Y) - U^2)]$

The reason for using a squared difference to measure the "loss" between *T(Y)* and *U* is mostly convenience; properties of squared differences involving random variables are more easily examined than, say, absolute differences. The reason for taking an expectation is to remove the randomness of the squared difference by averaging over the distribution of the data.

Consider first the case where the target *U* is a constant—say, the parameter β—and denote the mean of the estimator *T(Y)* as μ_τ. The mean squared error can then be decomposed as

$$MSE[T(Y):\beta] = E[T(Y) - \beta)^2]$$

$$=E[T(Y) - \mu_\tau)^2 - E[(\beta - \mu_\tau)^2]$$

$=Var[T(Y)] + (\beta - \mu_\tau)^2$

The mean squared error thus comprises the variance of the estimator and the squared bias. The two components can be associated with an estimator's precision (small variance) and its accuracy (small bias).

If *T(Y)* is an unbiased estimator of β—that is, if $E[(T(Y)] = \beta$—then the mean squared error is simply the variance of the estimator. By choosing an estimator that has minimum variance, you also choose an estimator that has minimum mean squared error among all unbiased estimators. However, as you can see from the previous expression, bias is also an "average" property; it is defined as an expectation. It is quite possible to find estimators in some statistical modeling problems that have smaller mean squared error than a minimum variance unbiased estimator; these are estimators that permit a certain amount of bias but improve on the variance. For example, in models where regressors are highly collinear, the ordinary least squares estimator continues to be unbiased. However, the presence of collinearity can induce poor precision and lead to an erratic estimator. Ridge regression stabilizes the regression estimates in this situation, and the coefficient estimates are somewhat biased, but the bias is more than offset by the gains in precision.

When the target U is a random variable, you need to carefully define what an unbiased prediction means. If the statistic and the target have the same expectation, $E[U] = E[T(Y)]$, then

$$MSE[T(Y):U] = Var[(T(Y)] + Var[U] - 2Cov[T(Y), U]$$

In many instances the target U is a new observation that was not part of the analysis. If the data are uncorrelated, then it is reasonable to assume in that instance that the new observation is also not correlated with the data. The mean squared error then reduces to the sum of the two variances. For example, in a linear regression model where *U* is a new observation Y_0 and *T(Y)* is the regression estimator

$$\hat{Y}_0 = x_0'(X'X)^{-1}X'Y$$

with variance $Var[Y_0] = \sigma^2 x_0'(X'X)^{-1}x_0$, the mean squared prediction error for $Y_0$is

$$MSE[\hat{Y}; Y_0] = \sigma^2(x_0'(X'X)^{-1}x_0 + 1)$$

and the mean squared prediction error for predicting the mean $E[Y_0]$ is

$$MSE\left[\hat{Y}; E[Y_0]\right] = \sigma^2(x_0'(X'X)^{-1}x_0$$

3.9 COEFFICIENT OF DETERMINATION (R^2)

The coefficient of determination in a regression model, also known as the R-square statistic **(R^2)**, measures the proportion of variability in the response that is explained by the regressor variables. In a linear regression model with intercept, **(R^2)**is defined as

$$(\boldsymbol{R^2}) = \mathbf{1} - \frac{\boldsymbol{SSE}}{\boldsymbol{SST}}$$

where SSE is the residual (error) sum of squares and SST is the total sum of squares corrected for the mean. The adjusted $\boldsymbol{R^2}$ statistic is an alternative to $\boldsymbol{R^2}$ that takes into account the number of parameters in the model. This statistic is calculated as

$$ADJRSQ = 1 - \frac{n-i}{n-p}(1 - \boldsymbol{R^2})$$

where n is the number of observations used to fit the model, P is the number of parameters in the model (including the intercept), and i is 1 if the model includes an intercept term, and 0 otherwise.

$\boldsymbol{R^2}$ statistics also play an important indirect role in regression calculations. For example, the proportion of variability explained by regressing all other variables in a model on a particular regressor can provide insights into the interrelationship among the regressors.

Tolerances and variance inflation factors measure the strength of interrelationships among the regressor variables in the model. If all variables are orthogonal to each other, then both the tolerance and variance inflation are 1. If a variable is very closely related to other variables, the tolerance approaches 0 and the variance inflation gets very large. Tolerance (TOL) is 1 minus the $\boldsymbol{R^2}$ that results from the regression of the other variables in the model on that regressor. Variance inflation (VIF) is the diagonal of $(X'WX)^{-1}$, if $(X'WX)$ is scaled to correlation form. The statistics are related as $VIF = \frac{1}{TOL}$

3.10 MAXIMUM-LIKELIHOOD ESTIMATION (MLE):- is a method of estimating the parameters of a statistical model. When applied to a data set and given a statistical model, maximum-likelihood estimation provides estimates for the model's parameters.

CHAPTER 4

ANALYSIS AND DISSCUSSION OF RESULTS

4.1 INTRODUCTION

This chapter displays, discusses and interprets the results obtained from the study. This chapter is generally organized into Preliminary analysis, model fitting and then the forecasting, there will also be a discussion on the accuracy of the forecast. The study has been looking at rainfall figures in fifteen years that is to say from January 1995 to December 2009 within four regions of Ghana i.e. Greater Accra, Northern, Western, and Eastern regions which are representing the north, south east and west of the country.

4.2 RAINFALL PATTERN IN THE NORTHERN REGION

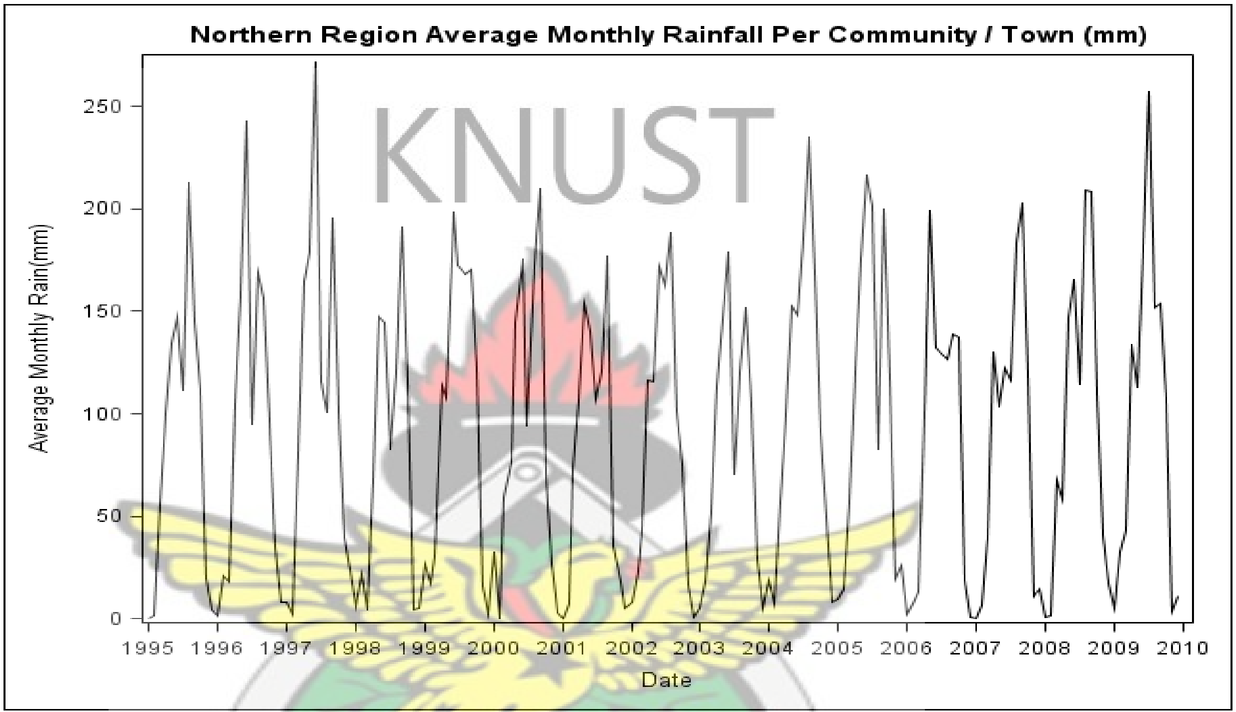

Figure 4.1 A times series plot of the observed data (Northern Region)

From Figure 4.1 which is A times series plot of the observed data in the northern region in Ghana, we can clearly see some form of seasonal trend in the pattern. We observe from the above figure that in the northern region of Ghana over the years under study for all the month, the highest level of rainfall in a particular month was in June 1997 which recorded 271.875mm of rain. Since 1997, the region has not recorded that amount of rainfall. The closest amount of rainfall, in a particular month closest to that of the June 1997 was recorded in July 2009 which was 257.25mm rain. June 1996 which recorded 242.8mm, August 2004 which recorded 235.15mm and June 2005 which recoded 216.6mm just to mention a few followed in that other, interestingly, September also recorded highest figures in the following years 1998, 2000, 2001, and 2007 recording 191.25mm, 209.85mm, 176.825mm, and 203.075mm respectively.

The above argument however strengthens the fact that Ghana experiences its highest amounts of rainfall from June to September each year. From our study, for the years under review June has clearly shown to be the month of highest amount of rainfall recording a total of 2638.08mm of rainfall. The month of January has proven to be the month with the lowest rainfall in the northern region followed by December and February in that order.

A first run of the best possible model came up with some interesting results as presented in the Table below.

The Table 4.2 below has Model Label, Number of Observations, Mean Square Error (MSE), and R-Square. In this research we decided to use (MSE), and R-Square as a criterion for selecting a model. From the Table 4.2 clearly using the summary of the models and the criterion for the selection of the best model, it is clear that with a less MSE (802.04) and a higher R-square (0.8379) the best model that can be used for forecasting the rains of the northern region will be Linear Trend with Seasonal Terms. The second best model so far as these criterions are concerned is the Seasonal Exponential Smoothing and the rest of the models follow as such.

This work is looking at four main models, that is, Linear Trend with Seasonal Terms, Seasonal Exponential Smoothing, ARIMA, and Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing. So all the ARIMA models are one model, and the Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing is also one model since the Simple Exponential Smoothing is just basics of the Linear (Holt) Exponential Smoothing.

Table 4.1 Northern Region summary of models and their MSE's and R-square

Series Name	Model Label	Number of Observations	Mean Square Error	R-Square
RAIN	Linear Trend with Seasonal Terms	180	802.0462	0.8379
RAIN	Seasonal Exponential Smoothing	180	816.6829	0.8349
RAIN	ARIMA(0,1,1)s NOINT	180	1050.6768	0.7890
RAIN	ARIMA(0,2,2)(0,1,1)s NOINT*	180	1274.0108	0.7431
RAIN	ARIMA(2,1,0)(0,1,1)s NOINT*	180	1376.1617	0.7226
RAIN	ARIMA(2,0,0)(1,0,0)s**	180	1734.0558	0.6495
RAIN	Simple Exponential Smoothing	180	3293.7734	0.3342
RAIN	Linear (Holt) Exponential Smoothing	180	3297.0549	0.3335

*** NOINT**

When NOINT is specified the fitting of a constant (or intercept) in the model is omitted. The interpretation of the constant depends on the model that is fit. If there are no autoregressive parameters in the model, then the constant, μ is the mean of the series. The constant represents the intercept if there are autoregressive parameters in the series, and lastly if the series is differenced then the constant represents the mean or the intercept of the differenced series.

**** EXPLANATION FOR ARIMA(P,D,Q)s**

ARIMA models for time series with regular seasonal fluctuations often use differencing operators and autoregressive and moving-average parameters at lags that are multiples of the length of the seasonal cycle. When all the terms in an ARIMA model factor refer to lags that are a multiple of a constant *s*, the constant is factored out and suffixed to the ARIMA(p,d,q)s notation.

ARIMA(p,d,q)×(P,D,Q)s is the general notation for the order of a seasonal ARIMA model with both seasonal and non-seasonal factors. The term (p,d,q) gives the order of the non-seasonal part of the ARIMA model and the term $(P,D,Q)_s$ gives the order of the seasonal part. The number of observations in the seasonal cycle is the value of *s*. For example if it is a monthly series, then the value for s is 12, 4 for quarterly series, 7 for daily series, etc.

We will now look at the various parameters for the various models and then we will state the models and use the models to produce the predicted or forecast values for the various models.

4.2.1 Analysis using the Linear Trend with Seasonal Terms

Below is figure 4.5 which gives us our parameters for the forecast.

The general model for the Linear Trend with Seasonal Terms is given as:-

$$Y = \alpha + LT(A) + SD_1X_1 + SD_2X_2 + SD_3X_3 + SD_4X_4 + SD_5X_5 + SD_6X_6 + SD_7X_7 + SD_8X_8 + SD_9X_9 + SD_{10}X_{10} + SD_{11}X_{11} \ldots\ldots\ldots 4.1$$

$Where\ \alpha = is\ the\ y - intercept$

$LT = is\ the\ linear\ trend$

$(A) = the\ sequence\ on\ the\ table$

$SD_i = Seasonal\ Dummy\ of\ the\ month$

$X_i = the\ month$

i = 1=January, 2=February, 3=March . . . 11=November

Table 4.2 Parameters for the Northern Region With the model "Linear Trend with Seasonal Terms

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Intercept	6.3571	8.6080	0.7385	0.4612
Linear Trend	0.0274	0.0423	0.6480	0.5179
Seasonal Dummy 1	-0.4987	10.7462	-0.0464	0.9630
Seasonal Dummy 2	3.2872	10.7444	0.3059	0.7600
Seasonal Dummy 3	37.4999	10.7429	3.4907	0.0006
Seasonal Dummy 4	98.1391	10.7414	9.1365	0.0000
Seasonal Dummy 5	136.4317	10.7402	12.7029	0.0000
Seasonal Dummy 6	167.0643	10.7391	15.5566	0.0000
Seasonal Dummy 7	125.5903	10.7382	11.6957	0.0000
Seasonal Dummy 8	148.6096	10.7374	13.8403	0.0000
Seasonal Dummy 9	162.6155	10.7369	15.1455	0.0000
Seasonal Dummy 10	92.4548	10.7364	8.6113	0.0000
Seasonal Dummy 11	14.3007	10.7362	1.3320	0.1847

From equation 4.1 above and from the parameters the equation for the forecasting will be given as:-

$$Y = 6.3571 + 0.0274(A) - 0.4987X_1 + 3.2872X_2 + 37.4999X_3 + 98.1391X_4 + 136.4317X_5 + 167.0643X_6 + 125.5903X_7 + 148.6096X_8 + 162.6155X_9 + 92.4548X_{10} + 14.3007X_{11} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \; 4.2$$

Example 1:- So for January 1995 we will have:

$$Y_{Jan1995} = 6.3571 + 0.0274(A) - 0.4987X_1$$

$$Y_{Jan1995} = 6.3571 + 0.0274(1) - 0.4987(1)$$

$$Y_{Jan1995} = 5.8858mm$$

Example 2: for say July 2001 we have

$$Y_{July2001} = 6.3571 + 0.0274(A) + 125.5903X_7$$

$$Y_{July2001} = 6.3571 + 0.0274(79) + 125.5903(1)$$

$$Y_{July2001} = 134.112mm$$

Example 3: for say December 2004 we have

$$Y_{Dec2004} = 6.3571 + 0.0274(A)$$

$$Y_{Dec2004} = 6.3571 + 0.0274(120)$$

$$Y_{Dec2004} = 9.6451mm$$

With the above information we will produce a prediction or forecast table including all the models on the Northern Region for easy comparing, the table will have actual figures and their corresponding predicted figures for the various models. But all these are done with the help of SAS so we will not be going through the calculations for the others and the other regions so as to safe space.

4.2.2 Analysis using Seasonal Exponential Smoothing Model

The general model equation for seasonal exponential smoothing is:-

$$Y_t = \mu_t + S_p(t) + \varepsilon_t \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.3$$

The **smoothing equations are**

$$L_t = \alpha(Y_t - S_{t-p}) + (1-\alpha)L_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.4$$

$$S_t = \delta(Y_t - L_t) + (1-\delta)S_{t-p} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.5$$

The k**-step-ahead prediction equation** is

$$\ddot{Y}_t(k) = L_t + S_{t-p+k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.6$$

That is, you forecast y k-steps ahead by using/taking the last available estimated level state and then add the last available smoothed seasonal factor, S_{t-p+k}, which matches the month of the forecast horizon.

The smoothing weights consist of the following at time *t*:

L_t is a smoothed level of the series that estimates μ_t

$S_{t-j}, j = 0, \ldots p-1$ are seasonal factors that estimate $S_p(t)$

$\boldsymbol{Y_t}$ = the value of the observation at time t

α = a level smoothing weight.

δ = a seasonal smoothing weight

The parameters for the Seasonal Exponential Smoothing Model are given in the Table 4.3 below,

With the above information we will produce a prediction or forecast table including all the models on the Northern Region for easy comparing, the table will have actual figures and their corresponding predicted figures for the various models.

Table 4.3 Parameters for the Northern Region With the model "Seasonal Exponential Smoothing"

Model Parameter	Estimate	Std. Error	T	Prob>|T|
LEVEL Smoothing Weight	0.0188	0.007205718	2.6081	0.0099
SEASONAL Smoothing Weight	0.0010	0.031307409	0.0319	0.9746
Residual Variance (sigma squared)	825.8591			
Smoothed Level	92.8204			
Smoothed Seasonal Factor 1	-82.7832			
Smoothed Seasonal Factor 2	-78.9647			
Smoothed Seasonal Factor 3	-44.7194			
Smoothed Seasonal Factor 4	15.9527			
Smoothed Seasonal Factor 5	54.2782			
Smoothed Seasonal Factor 6	84.9439			
Smoothed Seasonal Factor 7	43.5031			
Smoothed Seasonal Factor 8	66.5556			
Smoothed Seasonal Factor 9	80.5949			
Smoothed Seasonal Factor 10	10.4677			
Smoothed Seasonal Factor 11	-67.6526			

From the equations 4.4 above and from the parameters the equation for the forecasting from the Northern Region will be given as:

$$L_t = 0.019(Y_t - S_{t-p}) + (1 - 0.019)L_{t-1}$$

$$S_t = 0.001(Y_t - L_t) + (1 - 0.001)S_{t-1}$$

$$\hat{Y}_t(k) = L_t + S_{t-p+k}$$

Example 1. To Forecast the rainfall values for May1995 and suppose the following data is

L_{t-1}=87.9131 is a smoothed level of the series that estimates March1995.

S_{t-1}=15.95269,j=0,1,…p-1, are seasonal factor estimate for April 1995

S_{t-1+1}=54.27823 the seasonal factor estimate for May 1995

Y_t=100.2 *is* the value of the observation at time April 1995

α= 0.01879 is a level smoothing weight.

δ= 0.001 is a seasonal smoothing weight.

$$S_t = 0.001(Y_t - L_t) + (1 - 0.001)S_{t-1}$$

$$S_t = 0.001(100.2 - 87.844) + (1 - 0.001) * 15.95269$$

$$= 15.94$$

$$L_t = 0.01879(Y_t - S_{t-p}) + (1 - 0.01879)L_{t-1}$$

$$L_t = 0.01879(100.2 - 15.95269) + (1 - 0.01879)(87.9131)$$

$$= 1.583 + 86.261$$

$$= 87.844$$

Therefore the forecasted value for May 1995 is which is 142.1 is given by:-

$$\hat{Y}_t(k) = L_t + S_{t-p+k}$$

$$= 87.844 + 54.27823$$

$$= 142.1mm$$

4.2.3 Analysis using ARIMA

An autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model, ARIMA models are, in theory, the most general class of models for forecasting a time series which can be stationarized by transformations such as differencing and logging. In fact, the easiest way to think of ARIMA models is as fine-tuned versions of random-walk and random-trend models: the fine-tuning consists of adding *lags of the differenced series* and/or *lags of the forecast errors* to the prediction equation, as needed to remove any last traces of autocorrelation from the forecast errors.

The generalized ARIMA model is of the form **ARIMA (p,d,q)**

Where:-

- **p** is the number of autoregressive terms,
- **d** is the number of non-seasonal differences, and
- **q** is the number of lagged forecast errors in the prediction equation.

In our case the specific ARIMA model we are dealing with in the data of the Northern region is the ARIMA(0,1,1) NOINT.

ARIMA(0,1,1) without constant which is also the same as simple exponential smoothing**:** Another strategy for correcting auto-correlated errors in a random walk model is suggested by the simple exponential smoothing model. Recall that for some non-stationary time series (e.g., one that exhibits noisy fluctuations around a slowly-varying mean), the random walk model does not perform as well as a moving average of past values. In other words, rather than taking the most recent observation as the forecast of the next observation, it is better to use an *average* of the last few observations in order to filter out the noise and more accurately estimate the local mean. The simple exponential smoothing model uses an *exponentially weighted moving average* of past values to achieve this effect. The prediction equation for the simple exponential smoothing model can be written in a number of mathematically equivalent ways, one of which is:

$$\hat{Y}(t) = Y(t-1) - \theta e(t-1) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.4$$

We move on to produce the parameter estimates and the table for the actual values with their corresponding predicted figures will be at the end of the analysis for the northern region.

Table 4.4 Parameters for the Northern Region With the model "ARIMA(0,1,1)"

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Seasonal Moving Average, Lag 12	0.9479	0.1869	5.0724	0.0000
Model Variance (sigma squared)	907.2934			

4.2.4 Analysis Using Simple Exponential Smoothing (SES)

Table 4.5 Parameters for the Northern Region With the model "Simple Exponential Smoothing"

Model Parameter	Estimate	Std. Error	T	Prob>|T|
LEVEL Smoothing Weight (α)	0.9990	0.052852206	19	0.0000

Residual Variance (sigma squared)	3312.1743
Smoothed Level	10.5929

From equation 4.7 below and from the parameters the equation for the forecasting from the Northern Region will be given as:-

$$\hat{Y}_{t+1} = S_t = 0.999Y_t + (1-0.999)S_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots.4.7$$

Example 10. To Forecast the rainfall values for July 1996 and suppose the following data is Y_t(actual rainfall value for June 1996)=242.8 and S_{t-1} (actual smoothed level for May 1996)= 158.046

Then $\hat{Y}_{July1996} = 0.999Y_t + (1-0.999)S_{t-1}$

$= (0.999)*242.8 + (1-0.999)*158.1$

$= 242.7157$

Hence, a forecast value for July 1996 is 242.715.

Example 11. To Forecast the rainfall values for August 1996 and suppose the following data is

Y_t(actual rainfall value for July 1996)= 94.70 and S_{t-1} (actual smoothed level for June 1996)= 242.8

Then $\hat{Y}_{Aug1996} = 0.999Y_t + (1-0.999)S_{t-1}$

*= (0.999)*94.7 +(1-0.999)*242.8*

= 94.848mm

Hence, a forecast value for August 1996 is 94.848mm.

4.2.5 Analysis Using the Holt Linear Exponential Smoothing Model

The Holt Linear Exponential Smoothing Model is an extension to the simple exponential smoothing model. Holt's method introduces a trend factor to the SES method to allow for forecasting series with trends. This model is used for time series that appear to have a trend but no seasonality. The Holt smoothing method utilizes two equations. Equation one adjusts for the trend of the previous time period by adding it the last smoothed value.

$$L_t = \alpha\, y_t + (1-\alpha)(L_{t-1} + b_{t-1}) \qquad \text{.......4.8}$$

The second equation updates the trend by differencing the two previous trend values.

$$b_t = \beta\,(L_t - L_{t-1}) + (1-\beta) b_{t-1} \qquad \text{.......4.9}$$

The forecast value is determined as the sum of the trend term and the smoothed level.

$$\hat{Y}_{t+m} = L_t + m b_t \qquad \text{.......4.10}$$

Where

L_t = Estimate of the level of the series at time t

α = smoothing constant for the data.

y_t = new observation or actual value of series in period t.

β = smoothing constant for trend estimate

b_t = estimate of the slope of the series at time t

m = periods to be forecast into the future.

The initialization of the model requires that an initial trend (b_1)and smoothing value (L_1)must be determine along with identifying α, β .

Generally,

L_1 is set to Y_1

b_1 *is set to* $Y_2 – Y_{1,}$ $(Y_n – Y_1)/(n–1)$, *or* 0.

The weight α and β are selected subjectively or by minimizing the MSE as similarly done in SES.

Table 4.6 Parameters for the Northern Region With the model "Linear (Holt) Exponential Smoothing"

Model Parameter	**Estimate**	**Std. Error**	**T**	**Prob>|T|**
LEVEL Smoothing Weight	0.999	0.05449713	00018.331	0.0000
TREND Smoothing Weight	0.001	0.02565904	00000.039	0.9690
Residual Variance (sigma squared)	3334.100			
Smoothed Level	10.593			
Smoothed Trend	0.043			

From equation 4.8 above and from the parameters the equation for the forecasting from the Northern Region will be given as:

$$L_t = 0.999_{y_t} + (1 - 0.999)(L_{t-1} + b_{t-1})$$

$$b_t = 0.001(L_t - L_{t-1}) + (1 - 0.001)b_{t-1}$$

$$\hat{Y}_{t+m} = L_t + mb_t$$

Example 4. To forecast the rainfall values for June 1996 the following equations must be solved. Assume the following data:

$Y_{May1996}$ =158.1

$L_{May1996}$ (actual rain value for May 1996) =158.047

$L_{April1996}$ (actual rain value for April 1996)= 104.514

$b_{April1996}$ (actual trend level for April 1996) = 0.160

$$b_{May1996} = 0.001(L_{May1996} - L_{April}) + (1 - 0.001)b_{April1996}$$

$$= 0.001(158.047 - 104.514) + (1 - 0.001) * 0.160$$

$$= 0.213mm$$

$$L_{May1996} = 0.999y_{May1996} + (1 - 0.999)(L_{April1996} + b_{April1996})$$

$$= 0.999 * (158.1) + (1 - 0.999)(104.514 + 0.160)$$

$$= 158.047mm$$

Then $\hat{Y}_{June1996} = L_{May96} + b_{May96}$

$= 158.047+0.213=158.260$

Hence, a forecast value for JUNE 1996 is 158.260mm.

With the all the above information on the Northern Region we will produce a prediction or forecast table including all the models on the Northern Region for easy comparing into Table 4.7, the table will have actual figures and their corresponding predicted figures for the various models.

But all these are done with the help of SAS so we will not be going through the calculations for the others and the other regions so as to save space.

Table 4.7 Actual figures of Rainfall with their corresponding predicted figures for all the models for the Northern Region.

Number of Observation	FULL DATE	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
1	Jan1995	0.0	64.28	9.12	65.73	60.44	6.20
2	Feb1995	10.4	63.60	18.53	61.13	59.82	15.53
3	Mar1995	167.0	63.02	62.17	57.68	59.32	59.11
4	Apr1995	93.7	63.97	104.77	68.59	60.42	103.19
5	May1995	104.2	64.19	147.02	74.39	60.75	145.32
6	Jun1995	283.8	64.52	218.19	70.12	61.19	215.92
7	Jul1995	148.8	66.59	63.83	82.52	63.46	62.50
8	Aug1995	17.7	67.33	21.18	85.83	64.32	21.04
9	Sep1995	2.6	66.79	38.36	68.00	63.85	38.18
10	Oct1995	20.1	66.11	69.17	57.93	63.23	68.50
11	Nov1995	50.5	65.61	31.01	55.92	62.79	29.68
12	Dec1995	29.2	65.40	21.81	57.63	62.66	20.77
13	Jan1996	0.0	64.99	8.83	28.77	62.32	8.15
14	Feb1996	56.2	64.30	18.24	32.78	61.69	17.45

Number of Observation	FULL DATE	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
15	Mar1996	72.9	64.16	61.88	118.59	61.64	61.74
16	Apr1996	172.8	64.18	104.48	78.86	61.75	104.38
17	May1996	228.8	65.18	146.74	88.99	62.88	147.55
18	Jun1996	191.8	66.71	217.90	196.81	64.56	219.94
19	Jul1996	99.6	67.87	63.54	120.07	65.85	65.21
20	Aug1996	31.2	68.12	20.90	40.33	66.20	22.96
21	Sep1996	11.7	67.71	38.08	31.00	65.84	40.22
22	Oct1996	3.0	67.10	68.88	39.59	65.29	70.63
23	Nov1996	18.6	66.42	30.72	53.76	64.66	31.61
24	Dec1996	1.9	65.90	21.52	41.22	64.19	22.22
25	Jan1997	0.9	65.22	8.55	25.36	63.56	9.19
26	Feb1997	1.2	64.54	17.96	55.72	62.92	18.54
27	Mar1997	143.8	63.86	61.60	63.62	62.30	62.03
28	Apr1997	177.7	64.58	104.20	123.29	63.12	105.70
29	May1997	171.1	65.62	146.45	160.76	64.29	148.92
30	Jun1997	467.2	66.59	217.62	137.50	65.37	220.40
31	Jul1997	39.5	70.42	63.26	106.77	69.45	69.51
32	Aug1997	9.2	70.07	20.61	67.16	69.15	26.34
33	Sep1997	7.1	69.43	37.79	31.62	68.54	43.22
34	Oct1997	150.4	68.77	68.60	27.95	67.92	73.48
35	Nov1997	31.9	69.51	30.44	46.41	68.75	36.51
36	Dec1997	38.9	69.09	21.24	39.29	68.38	27.23
37	Jan1998	0.0	68.74	8.26	31.23	68.08	14.65
38	Feb1998	7.1	68.02	17.67	29.96	67.39	23.90
39	Mar1998	0.9	67.37	61.31	103.39	66.78	67.50
40	Apr1998	52.0	66.67	103.91	116.49	66.11	109.12
41	May1998	165.8	66.47	146.17	109.63	65.96	150.52
42	Jun1998	57.0	67.38	217.33	276.41	66.98	222.12
43	Jul1998	6.1	67.22	62.97	40.38	66.88	65.30
44	Aug1998	6.1	66.57	20.33	17.83	66.26	21.74
45	Sep1998	8.8	65.93	37.51	29.64	65.65	38.62
46	Oct1998	101.5	65.32	68.31	106.90	65.07	69.08
47	Nov1998	24.3	65.61	30.15	45.46	65.44	31.41
48	Dec1998	15.0	65.15	20.95	49.37	65.02	22.12
49	Jan1999	17.7	64.61	7.98	26.93	64.52	9.26
50	Feb1999	52.6	64.09	17.39	31.33	64.04	18.82
51	Mar1999	16.0	63.92	61.03	31.65	63.92	63.07
52	Apr1999	99.9	63.40	103.63	58.66	63.44	104.96

Number of Observation	FULL DATE	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
53	May1999	63.8	63.69	145.88	120.42	63.81	147.15
54	Jun1999	240.1	63.64	217.05	60.32	63.81	217.22
55	Jul1999	53.4	65.29	62.69	42.09	65.60	63.09
56	Aug1999	12.7	65.12	20.04	47.48	65.47	20.25
57	Sep1999	29.9	64.55	37.22	35.47	64.94	37.23
58	Oct1999	32.1	64.16	68.03	82.87	64.58	68.06
59	Nov1999	8.8	63.79	29.87	39.96	64.25	29.41
60	Dec1999	6.5	63.20	20.67	33.02	63.69	19.93
61	Jan2000	0.7	62.59	7.69	35.81	63.11	7.00
62	Feb2000	0.0	61.93	17.10	54.01	62.47	16.39
63	Mar2000	44.6	61.27	60.74	32.87	61.84	59.86
64	Apr2000	48.6	61.04	103.34	80.21	61.66	102.23
65	May2000	124.7	60.86	145.60	62.76	61.53	143.68
66	Jun2000	81.5	61.42	216.76	159.66	62.17	214.74
67	Jul2000	40.4	61.56	62.40	58.25	62.37	58.43
68	Aug2000	9.1	61.29	19.76	31.23	62.15	15.48
69	Sep2000	6.9	60.72	36.94	43.61	61.61	32.47
70	Oct2000	36.5	60.14	67.74	43.38	61.05	63.03
71	Nov2000	25.8	59.85	29.58	32.13	60.80	24.52
72	Dec2000	31.3	59.46	20.38	32.94	60.45	15.35
73	Jan2001	1.1	59.12	7.41	30.41	60.15	2.83
74	Feb2001	19.2	58.50	16.82	28.73	59.55	12.27
75	Mar2001	53.0	58.05	60.46	51.73	59.14	56.06
76	Apr2001	154.9	57.94	103.06	55.76	59.08	98.57
77	May2001	207.9	58.82	145.31	103.28	60.05	141.56
78	Jun2001	176.6	60.21	216.48	87.95	61.55	213.68
79	Jul2001	14.4	61.28	62.12	66.48	62.72	58.80
80	Aug2001	6.1	60.76	19.47	40.17	62.23	15.50
81	Sep2001	114.6	60.17	36.65	30.06	61.66	32.44
82	Oct2001	17.2	60.64	67.46	53.46	62.20	64.47
83	Nov2001	21.2	60.16	29.30	47.82	61.74	25.71
84	Dec2001	12.4	59.72	20.10	43.73	61.33	16.47
85	Jan2002	53.5	59.20	7.12	27.62	60.83	3.65
86	Feb2002	12.7	59.08	16.53	40.14	60.76	13.81
87	Mar2002	32.3	58.57	60.17	58.62	60.27	57.49
88	Apr2002	146.4	58.25	102.77	109.32	59.99	99.75
89	May2002	116.6	59.05	145.03	141.44	60.86	142.61
90	Jun2002	343.1	59.54	216.19	125.14	61.43	213.36

Number of Observation	FULL DATE	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
91	Jul2002	48.6	62.24	61.83	51.13	64.29	60.77
92	Aug2002	7.4	62.05	19.19	49.55	64.13	17.95
93	Sep2002	27.4	61.46	36.37	90.56	63.56	34.96
94	Oct2002	63.6	61.07	67.17	34.02	63.19	65.63
95	Nov2002	30.4	61.04	29.01	39.33	63.19	27.53
96	Dec2002	2.6	60.68	19.81	38.30	62.86	18.39
97	Jan2003	2.4	60.06	6.83	56.18	62.25	5.47
98	Feb2003	15.8	59.44	16.25	31.57	61.64	14.85
99	Mar2003	31.3	58.95	59.89	42.75	61.17	58.53
100	Apr2003	178.3	58.62	102.49	105.91	60.87	100.83
101	May2003	55.9	59.73	144.74	96.11	62.06	144.05
102	Jun2003	352.6	59.63	215.91	215.02	62.00	214.10
103	Jul2003	33.9	62.42	61.55	63.29	64.95	61.49
104	Aug2003	14.5	62.08	18.90	42.79	64.64	18.46
105	Sep2003	35.9	61.56	36.08	42.56	64.13	35.56
106	Oct2003	114.9	61.26	66.89	62.61	63.84	66.34
107	Nov2003	31.4	61.72	28.73	49.64	64.36	28.95
108	Dec2003	22.2	61.37	19.53	34.66	64.02	19.79
109	Jan2004	10.1	60.93	6.55	30.46	63.60	7.13
110	Feb2004	11.9	60.38	15.96	37.15	63.06	16.60
111	Mar2004	8.0	59.85	59.60	44.21	62.54	60.17
112	Apr2004	26.4	59.28	102.20	120.96	61.98	102.23
113	May2004	120.0	58.91	144.46	51.40	61.62	143.18
114	Jun2004	73.1	59.44	215.62	214.93	62.21	214.34
115	Jul2004	20.5	59.51	61.26	43.45	62.33	57.73
116	Aug2004	36.1	59.07	18.62	26.96	61.90	14.58
117	Sep2004	93.0	58.79	35.80	47.84	61.64	32.04
118	Oct2004	96.8	59.06	66.60	94.42	61.96	63.71
119	Nov2004	32.6	59.37	28.44	50.77	62.31	26.05
120	Dec2004	2.4	59.05	19.24	42.03	62.01	16.94
121	Jan2005	2.2	58.44	6.26	32.62	61.40	4.05
122	Feb2005	3.9	57.84	15.68	32.26	60.80	13.45
123	Mar2005	149.4	57.25	59.32	30.67	60.22	56.91
124	Apr2005	45.1	58.09	101.92	50.09	61.13	100.94
125	May2005	112.7	57.90	144.17	102.22	60.97	142.21
126	Jun2005	157.2	58.37	215.34	70.84	61.49	213.17
127	Jul2005	48.8	59.27	60.98	48.70	62.46	57.83
128	Aug2005	29.9	59.11	18.33	56.40	62.33	15.13

Number of Observation	FULL DATE	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
129	Sep2005	17.6	58.77	35.51	79.34	62.00	32.53
130	Oct2005	91.2	58.31	66.32	76.75	61.55	63.13
131	Nov2005	50.3	58.57	28.16	44.53	61.85	25.38
132	Dec2005	19.0	58.43	18.96	32.92	61.73	16.50
133	Jan2006	10.2	57.99	5.98	31.35	61.30	3.86
134	Feb2006	2.0	57.47	15.39	30.65	60.78	13.36
135	Mar2006	26.9	56.87	59.03	106.57	60.18	56.90
136	Apr2006	39.9	56.52	101.63	47.28	59.84	99.09
137	May2006	216.0	56.29	143.89	84.03	59.64	140.35
138	Jun2006	144.1	57.78	215.05	121.36	61.23	212.73
139	Jul2006	40.5	58.56	60.69	67.18	62.07	57.26
140	Aug2006	7.9	58.33	18.05	47.70	61.85	14.48
141	Sep2006	81.8	57.78	35.23	36.59	61.30	31.56
142	Oct2006	107.2	57.96	66.03	79.31	61.51	63.09
143	Nov2006	3.8	58.37	27.87	62.18	61.97	25.56
144	Dec2006	3.0	57.79	18.67	39.35	61.38	16.02
145	Jan2007	0.0	57.19	5.69	30.20	60.79	3.17
146	Feb2007	6.3	56.58	15.11	27.03	60.17	12.52
147	Mar2007	59.1	56.03	58.75	41.03	59.63	56.11
148	Apr2007	76.7	56.00	101.35	51.22	59.62	98.73
149	May2007	123.3	56.14	143.60	147.91	59.79	140.63
150	Jun2007	219.4	56.73	214.77	107.48	60.44	211.59
151	Jul2007	163.1	58.25	60.41	58.18	62.05	57.22
152	Aug2007	57.9	59.21	17.76	49.88	63.08	16.13
153	Sep2007	57.7	59.14	34.94	83.21	63.03	33.93
154	Oct2007	116.2	59.07	65.75	88.17	62.97	65.10
155	Nov2007	32.4	59.57	27.59	33.65	63.51	27.60
156	Dec2007	24.0	59.25	18.39	34.38	63.20	18.43
157	Jan2008	3.8	58.85	5.41	30.31	62.80	5.84
158	Feb2008	0.0	58.26	14.82	31.90	62.20	15.21
159	Mar2008	60.6	57.63	58.46	58.15	61.57	58.68
160	Apr2008	101.3	57.61	101.06	69.12	61.56	101.26
161	May2008	249.4	57.97	143.32	98.07	61.96	143.47
162	Jun2008	131.1	59.77	214.48	159.15	63.87	216.15
163	Jul2008	100.5	60.41	60.12	126.62	64.55	60.60
164	Aug2008	25.6	60.75	17.48	59.40	64.91	18.55
165	Sep2008	42.4	60.35	34.66	57.99	64.51	35.85
166	Oct2008	47.9	60.12	65.46	88.80	64.29	66.81

Number of Observation	FULL DATE	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
167	Nov2008	73.1	59.95	27.30	43.87	64.12	28.30
168	Dec2008	62.2	60.02	18.10	42.24	64.21	19.68
169	Jan2009	4.2	59.98	5.12	35.81	64.19	7.59
170	Feb2009	48.7	59.39	14.54	30.28	63.58	16.93
171	Mar2009	36.8	59.23	58.18	62.23	63.43	61.06
172	Apr2009	127.9	58.95	100.78	84.07	63.16	103.28
173	May2009	115.2	59.57	143.03	164.11	63.82	145.93
174	Jun2009	324.3	60.05	214.20	100.17	64.34	216.54
175	Jul2009	69.4	62.56	59.84	95.52	66.98	63.77
176	Aug2009	16.4	62.58	17.19	58.63	67.01	21.21
177	Sep2009	8.1	62.08	34.37	50.54	66.49	38.35
178	Oct2009	9.0	61.50	65.18	51.43	65.90	68.78
179	Nov2009	0.1	60.94	27.02	63.49	65.32	29.77
180	Dec2009	26.6	60.29	17.82	56.13	64.66	20.13
181	Jan2010		59.91	4.84	26.00	64.27	7.50
182	Feb2010		59.85	14.25	53.72	64.27	16.91
183	Mar2010		59.80	57.89	49.72	64.27	60.56
184	Apr2010		59.74	100.49	98.43	64.27	103.16
185	May2010		59.69	142.74	91.88	64.27	145.42
186	Jun2010		59.63	213.91	202.72	64.27	216.58
187	Jul2010		59.58	59.55	67.67	64.27	62.23
188	Aug2010		59.52	16.90	39.59	64.27	19.58
189	Sep2010		59.47	34.08	35.19	64.27	36.77
190	Oct2010		59.41	64.89	35.67	64.27	67.58
191	Nov2010		59.36	26.73	30.95	64.27	29.42
192	Dec2010		59.30	17.53	44.99	64.27	20.22
193	Jan2011		59.25	4.55	44.67	64.27	7.50
194	Feb2011		59.19	13.97	59.37	64.27	16.91
195	Mar2011		59.14	57.61	57.25	64.27	60.56
196	Apr2011		59.09	100.21	83.05	64.27	103.16
197	May2011		59.03	142.46	79.59	64.27	145.42
198	Jun2011		58.98	213.63	138.32	64.27	216.58
199	Jul2011		58.92	59.27	66.76	64.27	62.23
200	Aug2011		58.87	16.62	51.88	64.27	19.58
201	Sep2011		58.81	33.80	49.55	64.27	36.77
202	Oct2011		58.76	64.61	49.80	64.27	67.58
203	Nov2011		58.70	26.45	47.30	64.27	29.42
204	Dec2011		58.65	17.25	54.74	64.27	20.22

From Table 4.7 above, clearly there is a lot similarities between the predicted figures for the Holt Exponential Model and the Simple Exponential models that can pretty much be explained from the fact that both models have the same underlying principles and the only difference is the fact that the Holt is an extension of the Simple Exponential Model. Secondly, the figures of the Linear Trend with Seasonal Terms model and the Seasonal Exponential Models are close to each other and again this can also be seen from the fact that both models take into consideration the Seasonality of the data set. The ARIMA model seems to be very independent of the other models. Since we have decided on the fact that the best model for this data set was that of the Northern Region, is the Linear Trend with Seasonal Terms, we will look at the forecast looking at the prediction for 2011, we see that there will be averagely high level of rainfall in the Northern Region, and already there are signals from the Ghana Metrological Service and other happenings.

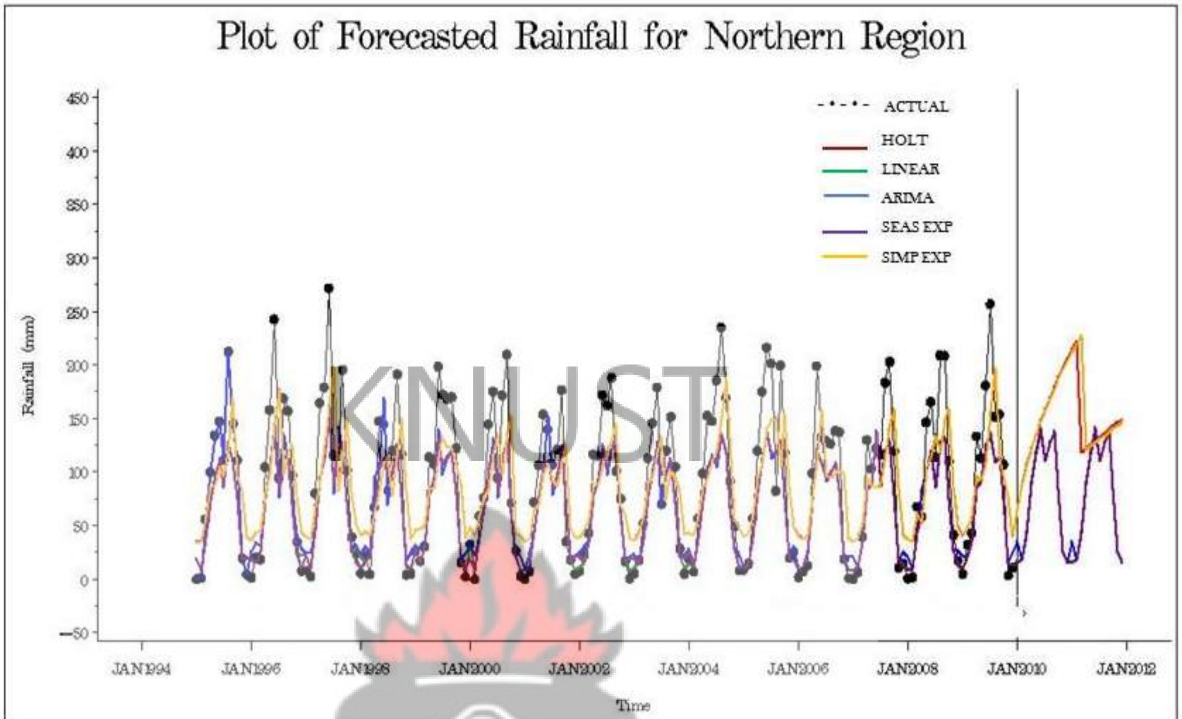

Figure 4.2 Plots actual rainfall and forecast for all the models (Northern Region)

4.3 RAINFALL PATTERN IN THE WESTERN REGION

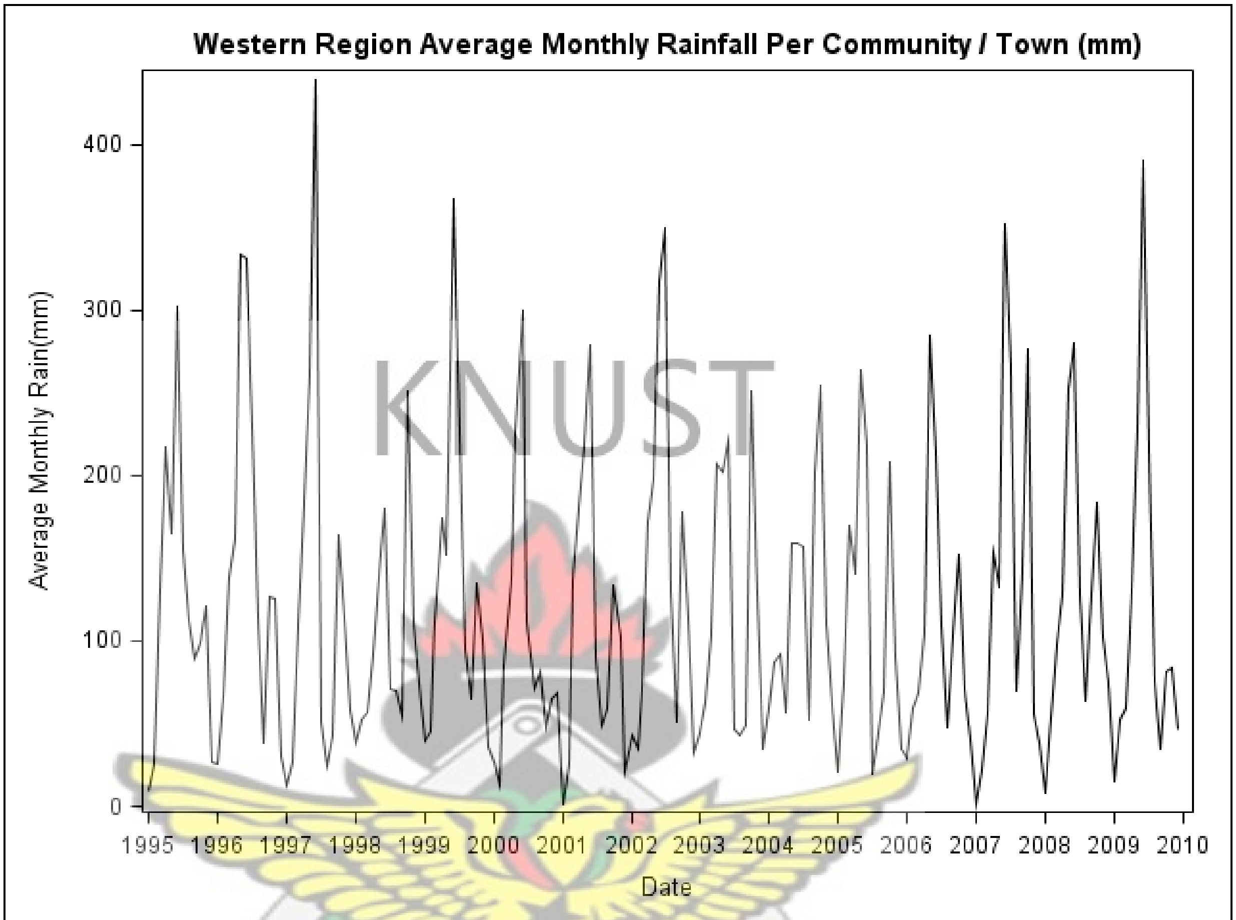

Figure 4.3 A times series plot of the observed data (western region).

From Figure 4.3 which is A times series plot of the observed data in the western region in Ghana, we can clearly see some form of seasonal pattern with some irregularities trend in the pattern. We observe from the above figure that in the western region of Ghana over the years under study for all the month, the highest level of rainfall in a particular month was in June 1997 which recorded 439mm of rain. Again in June 2009, 389.925mm was recorded, 367.05mm was recorded in June 1999, and 352.075mm was recorded in June 2007 just to mention a few. Interestingly, the Western region has recorded higher levels of rainfall than in the Northern region. Unlike in the Northern region where some rainfall levels are too low to measure, rainfall in the western region has been relatively high all through. The month of January has proven to be the month with the lowest rainfall in the Western region followed by December and February in that order. The highest amount of rainfall in total comes from the western region which is a total of 4695.15mm, which can go to explain why the Western region has been tagged as the food basket of the Nation producing the highest amount of Cocoa and other agricultural products.

The Table 4.8 below has Model Label, Number of Observations, Mean Square Error (MSE), and R-Square. In this research we decided to use (MSE), and R-Square as a criterion for selecting a model. From the Table 4.8 clearly using the summary of the models and the criterion for the selection of the best model, it is clear that with a less MSE (2440.6715) and a higher R-square (0.70) the best model that can be used for forecasting the rains of the northern region will be Linear Trend with Seasonal Terms. The second best model so far as these criterions are concerned is the Seasonal Exponential Smoothing and the rest of the models follow as such.

This work is looking at four main models, that is, Linear Trend with Seasonal Terms, Seasonal Exponential Smoothing, ARIMA, and Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing. So all the ARIMA models are one model, and the Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing is also one model since the Simple Exponential Smoothing is just basics of the Linear (Holt) Exponential Smoothing.

We continue to however present the table on the various models and then we will only pick the best model to look at.

Table 4.8 WESTERN REGION MODEL SUMMARY

Series Name	Model Label	Number of Observations	Mean Square Error	R-Square
RAIN	Linear Trend with Seasonal Terms	180	2440.67	0.70
RAIN	Seasonal Exponential Smoothing	180	2480.94	0.69
RAIN	Log ARIMA(2,0,0)(1,0,0)s	180	4174.02	0.48
RAIN	ARIMA(2,0,0)(1,0,0)s	180	4237.47	0.47
RAIN	Simple Exponential Smoothing	180	8118.35	-0.01
RAIN	Linear (Holt) Exponential Smoothing	180	8123.32	-0.01

From Table 4.8 above, clearly the Linear Trend with Seasonal Terms model is the best possible model using our criterions, we considered a model that has the smaller MSE and the highest R-square so we will look at the parameters of this model and conclude.

Table 4.9 Parameters for the Western Region With the model "Linear Trend with Seasonal Terms"

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Intercept	47.8524	15.02	3.1867	0.002
Linear Trend	-0.0262	0.07	-0.3549	0.723
Seasonal Dummy 1	-20.7412	18.75	-1.1064	0.270
Seasonal Dummy 2	1.2583	18.74	0.0671	0.947
Seasonal Dummy 3	53.4445	18.74	2.8519	0.005
Seasonal Dummy 4	103.7440	18.74	5.5366	0.000

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Seasonal Dummy 5	168.7301	18.74	9.0059	0.000
Seasonal Dummy 6	244.4563	18.73	13.0490	0.000
Seasonal Dummy 7	105.9958	18.73	5.6585	0.000
Seasonal Dummy 8	25.5353	18.73	1.3633	0.175
Seasonal Dummy 9	35.1415	18.73	1.8762	0.062
Seasonal Dummy 10	123.8343	18.73	6.6119	0.000
Seasonal Dummy 11	54.5472	18.73	2.9125	0.004

From Table 4.9 that has the Parameters for the Western Region With the model "Linear Trend with Seasonal Terms", the equation for the forecasting will be given as:-

$$\begin{aligned} Y = 47.8524 - 0.0262(A) - 20.7412X_1 + 1.2583X_2 + 53.4445X_3 + 103.7440X_4 \\ + 168.7301X_5 + 244.4563X_6 + 105.9958X_7 + 25.5353X_8 \\ + 35.1415X_9 + 123.8343X_{10} \\ + 54.5472X_{11} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \mathbf{4.11} \end{aligned}$$

And this will be the best model for the forecasting of Western Region of Ghana.

Below is Table 4.10, which is shortened to two years to save space.

Table 4.10 Actual figures of Rainfall with their corresponding predicted figures for all the models for the Western Region (2009-2011).

Number of Observation	Full Date	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
169	Jan2009	15.1	114.40	22.69	55.05	119.35	24.88
170	Feb2009	52.1	113.15	44.66	67.03	118.10	46.70
171	Mar2009	59.3	112.36	96.82	99.58	117.31	98.99
172	Apr2009	130	111.66	147.09	109.58	116.62	148.65
173	May2009	227.2	111.80	212.05	201.08	116.78	213.31

Number of Observation	Full Date	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
174	Jun2009	389.9	113.09	287.75	222.55	118.10	289.14
175	Jul2009	217.7	116.29	149.27	187.12	121.35	152.20
176	Aug2009	75.5	117.42	68.78	106.63	122.50	72.75
177	Sep2009	34.7	116.86	78.36	111.48	121.94	82.43
178	Oct2009	81	115.81	167.03	123.87	120.90	170.43
179	Nov2009	84	115.33	97.71	90.37	120.42	99.72
180	Dec2009	46.2	114.89	43.14	91.24	119.98	44.90
181	Jan2010		114.01	22.37	44.83	119.10	24.40
182	Feb2010		113.94	44.35	79.85	119.10	46.38
183	Mar2010		113.86	96.51	86.17	119.10	98.55
184	Apr2010		113.79	146.78	126.01	119.10	148.83
185	May2010		113.72	211.74	181.02	119.10	213.79
186	Jun2010		113.65	287.44	273.59	119.10	289.50
187	Jul2010		113.58	148.95	175.51	119.10	151.01
188	Aug2010		113.51	68.47	94.53	119.10	70.53
189	Sep2010		113.43	78.05	71.29	119.10	80.12
190	Oct2010		113.36	166.71	97.67	119.10	168.79
191	Nov2010		113.29	97.40	99.37	119.10	99.48
192	Dec2010		113.22	42.83	77.84	119.10	44.92
193	Jan2011		113.15	22.06	77.07	119.10	24.40
194	Feb2011		113.08	44.03	97.01	119.10	46.38
195	Mar2011		113.00	96.19	100.61	119.10	98.55
196	Apr2011		112.93	146.47	123.30	119.10	148.83
197	May2011		112.86	211.43	154.63	119.10	213.79
198	Jun2011		112.79	287.13	207.36	119.10	289.50
199	Jul2011		112.72	148.64	151.50	119.10	151.01
200	Aug2011		112.65	68.15	105.37	119.10	70.53
201	Sep2011		112.57	77.73	92.14	119.10	80.12
202	Oct2011		112.50	166.40	107.16	119.10	168.79
203	Nov2011		112.43	97.09	108.13	119.10	99.48
204	Dec2011		112.36	42.51	95.87	119.10	44.92

From the table above and with our concentration on our choice of model we see that into 2011 we will have higher amount of rainfall. Due to the fact that the Holt and the Simple Exponential Models are not long term predictors they have not given us variations of figures to rely on but clearly we see that even those models are saying we will have higher amount of rain.

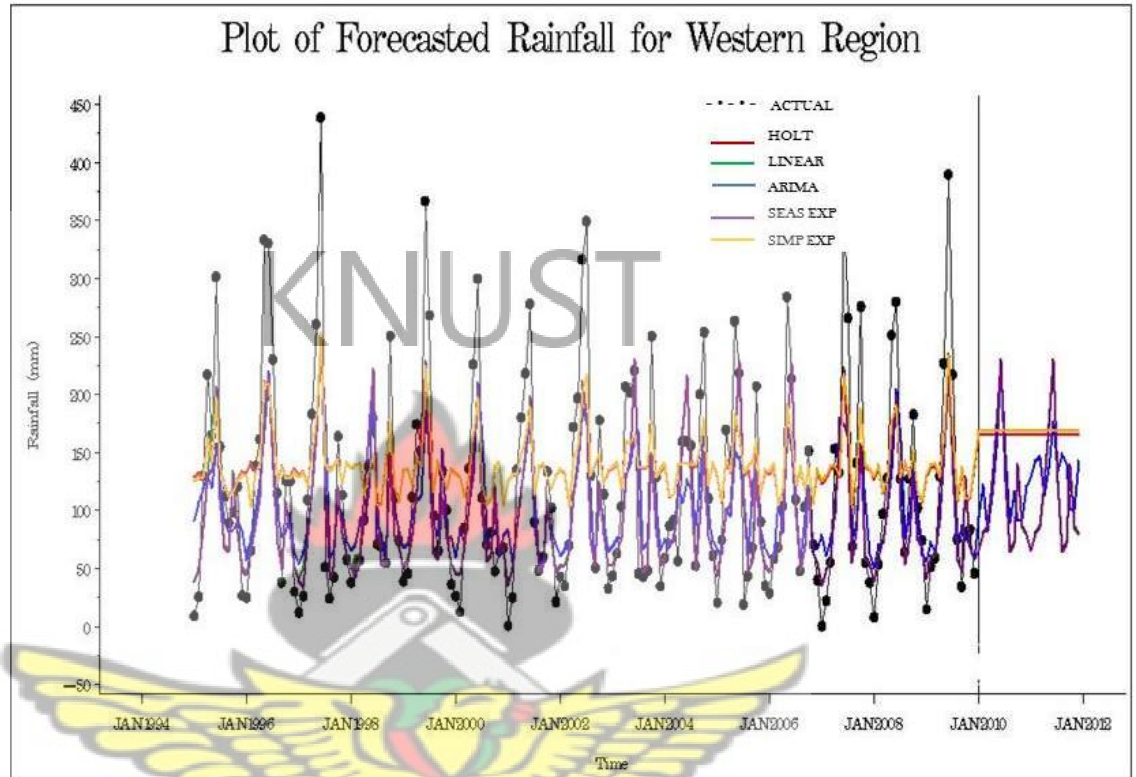

Figure 4.4 Plots actual rainfall and the forecast for all the models (Western Region)

4.4 RAINFALL PATTERN IN THE EASTERN REGION

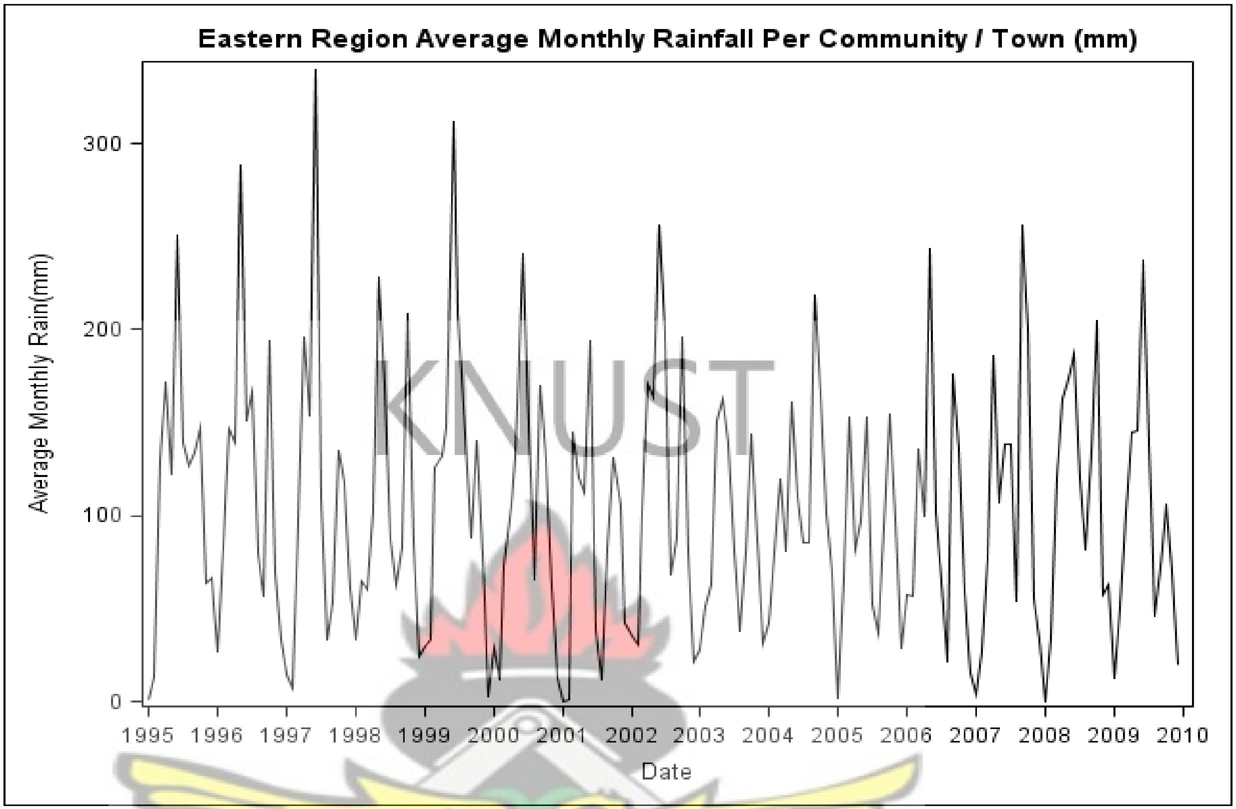

Figure 4.5 A times series plot of the observed data (Eastern region).

Figure 4.5 is the plot of observed rainfall data for the Eastern region of Ghana. June 1997 was again an interesting month for Eastern Region too, recording a very high figure of 339.6667mm rainfall. This was followed by June 1999 which recorded 311.8667mm of rainfall in the region. May 1996, June 2002, Sept 2007 and June 1995 followed in that order recording 288.7mm, 256.3333mm, 256.333mm and 250.5333mm respectively. The months of January, December and February in that order have been the lowest contributors of rain per month in the region.

TABLE 4.11 EASTERN REGION MODEL SUMMARY

Series Name	Model Label	Number of Observations	Mean Square Error	R-Square
RAIN	Linear Trend with Seasonal Terms	180	1613.29	0.65
RAIN	Seasonal Exponential Smoothing	180	1649.87	0.64
RAIN	ARIMA(0,1,1)s NOINT	180	2176.34	0.52
RAIN	ARIMA(2,1,0)(0,1,1)s NOINT	180	2759.33	0.39
RAIN	ARIMA(2,0,0)(1,0,0)s	180	2874.44	0.37
RAIN	Linear (Holt) Exponential Smoothing	180	4602.21	-0.01
RAIN	Simple Exponential Smoothing	180	4606.42	-0.01

From Table 4.11 above, clearly the Linear Trend with Seasonal Terms model is the best possible model, considering a model that has the smaller MSE and the highest R-square so we will look at the parameters of this model and conclude.

Table 4.12 Parameters for the Eastern Region With the model "Linear Trend with Seasonal Terms"

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Intercept	43.4371	12.2084	3.5580	0.0005
Linear Trend	-0.0871	0.0599	-1.4521	0.1483
Seasonal Dummy 1	-14.9443	15.2409	-0.9805	0.3282
Seasonal Dummy 2	4.6495	15.2384	0.3051	0.7607
Seasonal Dummy 3	73.9299	15.2362	4.8523	0.0000
Seasonal Dummy 4	100.4569	15.2342	6.5942	0.0000
Seasonal Dummy 5	126.8373	15.2324	8.3268	0.0000
Seasonal Dummy 6	163.2643	15.2308	10.7193	0.0000
Seasonal Dummy 7	85.3581	15.2296	5.6048	0.0000

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Seasonal Dummy 8	28.3918	15.2285	1.8644	0.0640
Seasonal Dummy 9	83.6922	15.2277	5.4961	0.0000
Seasonal Dummy 10	124.3459	15.2271	8.1661	0.0000
Seasonal Dummy 11	44.5663	15.2267	2.9268	0.0039

From Table 4.12 that has the Parameters for the Eastern Region With the model "Linear Trend with Seasonal Terms", the equation for the forecasting will be given as:-

$$Y = 43.4371 - 0.0871(A) - 14.9443X_1 + 4.6495X_2 + 73.9299X_3 + 100.4569X_4 + 126.8373X_5 + 163.2643X_6 + 85.3581X_7 + 28.3918X_8 + 83.6922X_9 + 124.3459X_{10} + 44.5663X_{11} \quad \ldots\ldots\ldots\ldots 4.12$$

And this will be the best model for the forecasting of Eastern Region of Ghana.

Table 4.13 Actual figures of Rainfall with their corresponding predicted figures for all the models for the Eastern Region (2009-2011).

Number of Observation	FullDate	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
169	Jan2009	12.5	94.41	13.78	20.28	103.05	20.39
170	Feb2009	50.7	93.16	33.29	41.78	101.77	39.74
171	Mar2009	101.8	92.44	102.48	111.39	101.04	109.16
172	Apr2009	144.5	92.43	128.92	135.66	101.05	135.45
173	May2009	145.3	93.00	155.21	160.05	101.67	161.94
174	Jun2009	237.1	93.57	191.55	176.22	102.29	197.93
175	Jul2009	140.1	95.38	113.56	111.79	104.21	120.68
176	Aug2009	45.7	95.85	56.51	60.46	104.72	64.03
177	Sep2009	71.8	95.04	111.72	140.92	103.88	118.94
178	Oct2009	106.4	94.59	152.29	168.51	103.42	158.66
179	Nov2009	70	94.61	72.42	76.32	103.46	77.80
180	Dec2009	19.8	94.14	27.77	37.55	102.99	33.01
181	Jan2010		93.00	12.74	19.07	101.81	18.69
182	Feb2010		92.86	32.24	43.16	101.81	38.21

Number of Observation	FullDate	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
183	Mar2010		92.72	101.44	109.90	101.81	107.41
184	Apr2010		92.58	127.88	137.03	101.81	133.85
185	May2010		92.45	154.17	157.76	101.81	160.15
186	Jun2010		92.31	190.51	185.67	101.81	196.50
187	Jul2010		92.17	112.52	116.19	101.81	118.51
188	Aug2010		92.04	55.46	58.17	101.81	61.47
189	Sep2010		91.90	110.68	130.19	101.81	116.69
190	Oct2010		91.76	151.24	158.87	101.81	157.26
191	Nov2010		91.62	71.38	75.34	101.81	77.40
192	Dec2010		91.49	26.72	34.80	101.81	32.76
193	Jan2011		91.35	11.69	19.07	101.81	18.69
194	Feb2011		91.21	31.20	43.16	101.81	38.21
195	Mar2011		91.08	100.39	109.90	101.81	107.41
196	Apr2011		90.94	126.83	137.03	101.81	133.85
197	May2011		90.80	153.12	157.76	101.81	160.15
198	Jun2011		90.67	189.46	185.67	101.81	196.50
199	Jul2011		90.53	111.47	116.19	101.81	118.51
200	Aug2011		90.39	54.42	58.17	101.81	61.47
201	Sep2011		90.25	109.63	130.19	101.81	116.69
202	Oct2011		90.12	150.20	158.87	101.81	157.26
203	Nov2011		89.98	70.33	75.34	101.81	77.40
204	Dec2011		89.84	25.68	34.80	101.81	32.76

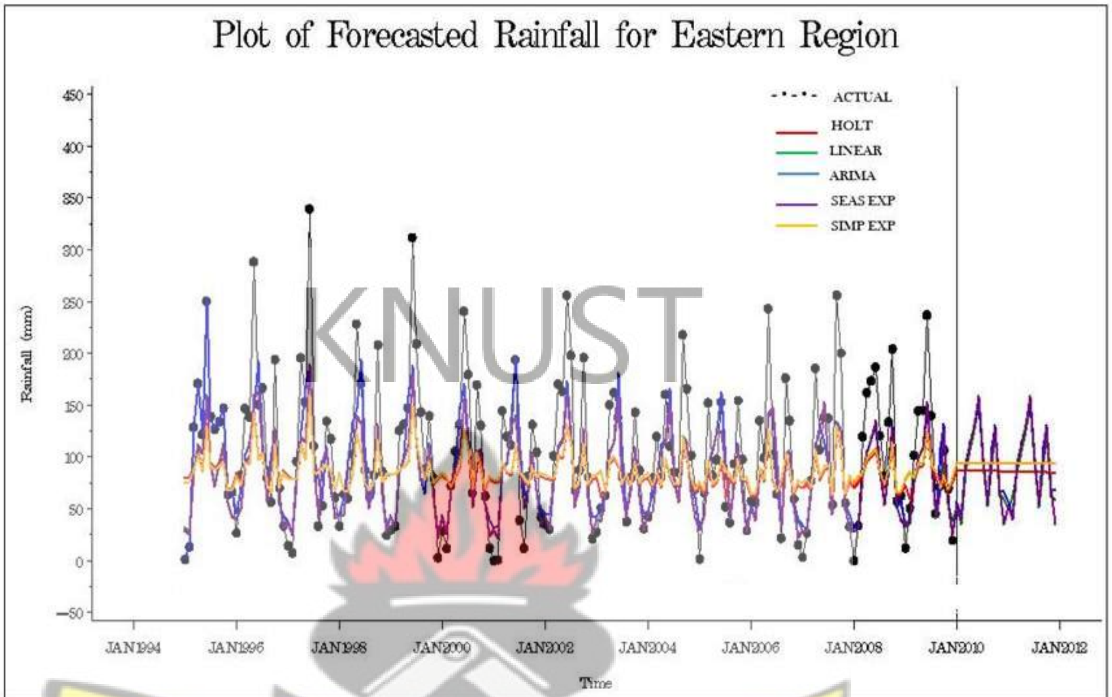

Figure 4.6 Plots actual rainfall and the forecast for all the models (Eastern Region)

4.4 RAINFALL PATTERN IN GREATER ACCRA REGION

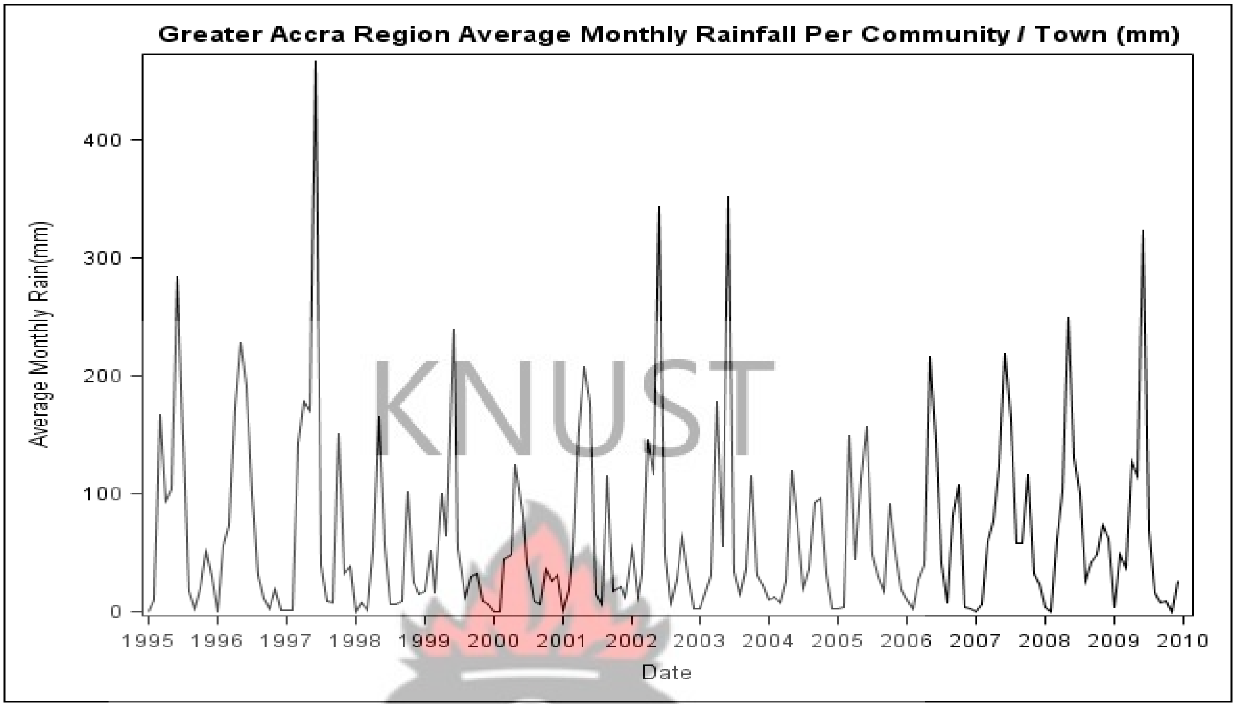

Figure 4.7 A times series plot of the observed data (Greater Accra region).

Figure 4.7 is the plot of observed rainfall data for the Greater Accra region of Ghana. June 1997 was again an interesting month for Greater Accra too, recording a very high figure of 467.2333mm rainfall. This was followed by June 2003 which recorded 352.6mm of rainfall in the region. June 2002, June 2009 May 2008 and June 1999 followed in that order recording 343.1333mm, 324.3mm, 249.4333mm and 240.0667mm respectively. The months of January, February and December in that order have been the lowest contributors of rain per month in the region.

Table 4.14 and figure 4.7 are the raw rainfall data and initial plots for the Greater Accra region of Ghana respectively. We continue to present the table on the various models and then we will only pick the best model to look at.

TABLE 4.14 GREATER ACCRA REGION MODEL SUMMARY

Series Name	Model Label	Number of Observations	Mean Square Error	R-Square
RAIN	Linear Trend with Seasonal Terms	180	2346.027058	0.602669709
RAIN	Seasonal Exponential Smoothing	180	2373.053138	0.59809249
RAIN	ARIMA(2,0,0)(1,0,0)s	180	4247.797256	0.280580113
RAIN	Linear (Holt) Exponential Smoothing	180	5966.106923	-0.010438048
RAIN	Simple Exponential Smoothing	180	5967.483985	-0.010671271

From Table 4.14 above, clearly the Linear Trend with Seasonal Terms model is the best possible model, considering a model that has the smaller MSE and the highest R-square so we will look at the parameters of this model and conclude.

Table 4.15 Parameters for the Greater Accra Region With the model "Linear Trend with Seasonal Terms"

Model Parameter	Estimate	Std. Error	T	Prob>|T|
Intercept	22.0952	14.72214635	1.5008	0.1353
Linear Trend	-0.0238	0.072292566	-0.3288	0.7427
Seasonal Dummy 1	-12.9548	18.3789543	-0.7049	0.4819
Seasonal Dummy 2	-3.5177	18.37596829	-0.1914	0.8484
Seasonal Dummy 3	40.1461	18.37326625	2.1850	0.0303
Seasonal Dummy 4	82.7698	18.37084829	4.5055	0.0000
Seasonal Dummy 5	125.0469	18.36871454	6.8076	0.0000
Seasonal Dummy 6	196.2374	18.36686508	10.6843	0.0000
Seasonal Dummy 7	41.9012	18.36530001	2.2815	0.0238
Seasonal Dummy 8	-0.7217	18.3640194	-0.0393	0.9687
Seasonal Dummy 9	16.4820	18.36302331	0.8976	0.3707
Seasonal Dummy 10	47.3125	18.36231178	2.5766	0.0108
Seasonal Dummy 11	9.1762	18.36188485	0.4997	0.6179

From Table 4.15 that has the Parameters for the Greater Accra Region With the model "Linear Trend with Seasonal Terms", the equation for the forecasting will be given as:-

$$Y = 22.0952 - 0.0238(A) - 12.9548X_1 - 3.5177X_2 + 40.1461X_3 + 82.7698X_4 + 125.0469X_5 + 196.2374X_6 + 41.9012X_7 - 0.7217X_8 + 16.4820X_9 + 47.3125X_{10} + 9.1762X_{11} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4.13$$

And this will be the best model for the forecasting of Greater Accra Region of Ghana.

Table 4.16 Actual figures of Rainfall with their corresponding predicted figures for all the models for the Greater Accra Region (2009-2011).

Number of Observation	FULL DATE	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
169	Jan2009	4.2	59.98	5.12	35.81	64.19	7.59
170	Feb2009	48.7	59.39	14.54	30.28	63.58	16.93
171	Mar2009	36.8	59.23	58.18	62.23	63.43	61.06
172	Apr2009	127.9	58.95	100.78	84.07	63.16	103.28
173	May2009	115.2	59.57	143.03	164.11	63.82	145.93
174	Jun2009	324.3	60.05	214.20	100.17	64.34	216.54
175	Jul2009	69.4	62.56	59.84	95.52	66.98	63.77
176	Aug2009	16.4	62.58	17.19	58.63	67.01	21.21
177	Sep2009	8.1	62.08	34.37	50.54	66.49	38.35
178	Oct2009	9.0	61.50	65.18	51.43	65.90	68.78
179	Nov2009	0.1	60.94	27.02	63.49	65.32	29.77
180	Dec2009	26.6	60.29	17.82	56.13	64.66	20.13
181	Jan2010		59.91	4.84	26.00	64.27	7.50
182	Feb2010		59.85	14.25	53.72	64.27	16.91
183	Mar2010		59.80	57.89	49.72	64.27	60.56
184	Apr2010		59.74	100.49	98.43	64.27	103.16
185	May2010		59.69	142.74	91.88	64.27	145.42
186	Jun2010		59.63	213.91	202.72	64.27	216.58
187	Jul2010		59.58	59.55	67.67	64.27	62.23
188	Aug2010		59.52	16.90	39.59	64.27	19.58
189	Sep2010		59.47	34.08	35.19	64.27	36.77
190	Oct2010		59.41	64.89	35.67	64.27	67.58
191	Nov2010		59.36	26.73	30.95	64.27	29.42
192	Dec2010		59.30	17.53	44.99	64.27	20.22
193	Jan2011		59.25	4.55	44.67	64.27	7.50
194	Feb2011		59.19	13.97	59.37	64.27	16.91
195	Mar2011		59.14	57.61	57.25	64.27	60.56
196	Apr2011		59.09	100.21	83.05	64.27	103.16
197	May2011		59.03	142.46	79.59	64.27	145.42
198	Jun2011		58.98	213.63	138.32	64.27	216.58
199	Jul2011		58.92	59.27	66.76	64.27	62.23
200	Aug2011		58.87	16.62	51.88	64.27	19.58
201	Sep2011		58.81	33.80	49.55	64.27	36.77
202	Oct2011		58.76	64.61	49.80	64.27	67.58

Number of Observation	FULL DATE	Actual Rain	Holt	Linear	ARIMA	Simple Exponential	Seasonal Exponential
203	Nov2011		58.70	26.45	47.30	64.27	29.42
204	Dec2011		58.65	17.25	54.74	64.27	20.22

From Table 4.16 which is the Actual figures of Rainfall with their corresponding predicted figures for all the models for the Greater Accra Region (2009-2011). We again realized that from the chosen model, we will have enough rains coming down in the whole of 2011.

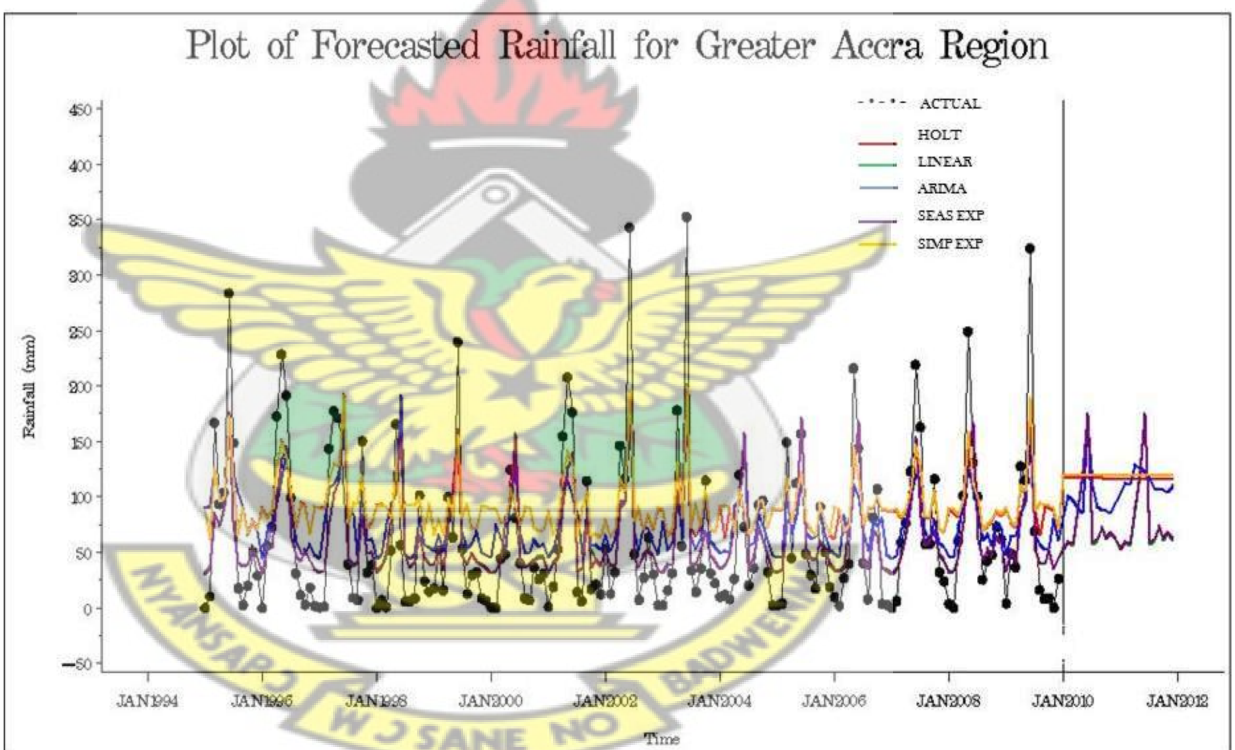

Figure 4.8 Plots actual rainfall and the forecast for all the models (Greater Accra Region)

CHAPTER 5
DISCUSSION, CONCLUSION AND RECOMMENDATIONS

5.0 DISCUSSION

The study has been looking at rainfall figures in fifteen years that is to say from January 1995 to December 2009 within four regions of Ghana i.e. Greater Accra, Northern, Western, and Eastern regions which are representing the north, south east and west of the country.

This work looked at four main models, that is, Linear Trend with Seasonal Terms, Seasonal Exponential Smoothing, ARIMA, and Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing. So all the ARIMA models are one model, and the Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing is also one model since the Simple Exponential Smoothing is just basics of the Linear (Holt) Exponential Smoothing.

We observe that in the northern region of Ghana over the years under study for all the month, the highest level of rainfall in a particular month was in June 1997 which recorded 271.875mm of rain. Since 1997, the region has not recorded that amount of rainfall. The closest amount of rainfall, in a particular month closest to that of the June 1997 was recorded in July 2009 which was 257.25mm rain. June 1996 which recorded 242.8mm, August 2004 which recorded 235.15mm and June 2005 which recoded 216.6mm just to mention a few followed in that other, interestingly, September also recorded highest figures in the following years 1998, 2000, 2001, and 2007 recording 191.25mm, 209.85mm, 176.825mm, and 203.075mm respectively. The month of January has proven to be the month with the lowest rainfall in the northern region followed by December and February in that order.

It became clear that the best model that can be used for forecasting the rains of the northern region will be Linear Trend with Seasonal Terms taking into consideration our decision to use (MSE), and R-Square as a criterion for selecting the best model.

We also observed that there is a lot similarities between the predicted figures for the Holt Exponential Model and the Simple Exponential models that can pretty much be explained from the fact that both models have the same underlying principles and the only difference is the fact that the Holt is an extension of the Simple Exponential Model. Secondly, the figures of the Linear Trend with Seasonal Terms model and the Seasonal Exponential Models are close to each other and again this can also be seen from the fact that both models take into consideration the Seasonality of the data set. The ARIMA model seems to be very independent of the other models. Since we have decided on the fact that the best model for this data set was that of the Northern Region, is the Linear Trend with Seasonal Terms, when we looked at the forecast looking at the prediction for 2011, we see that there will be averagely high level of rainfall in the Northern Region, and already there are signals from the Ghana Metrological Service and other happenings.

A times series plot of the observed data in the western region in Ghana, clearly showed some form of seasonal pattern with some irregularities trend in the pattern. We observe from the previous chapter that in the western region of Ghana over the years under study for all the month, the highest level of rainfall in a particular month was in June 1997 which recorded 339.6667mm of rain. Again in June 1999 311.8667mm was recorded, 288.7mm was recorded in may 1996, and 256.3333mm which recorded in September 2007. Interestingly, the Western region has recorded higher levels of rainfall than in the Northern region. Unlike in the Northern region where some rainfall levels are too low to measure, rainfall in the western region has been relatively high all through. The month of January has proven to be the month with the lowest rainfall in the Western region followed by December and February in that order. The highest amount of rainfall in total comes from the western region which is a total of 4695.15mm, which can go to explain why the Western region has been tagged as the food basket of the Nation producing the highest amount of Cocoa and other agricultural products.

5.1 CONCLUSION

To achieve the set objectives, a theoretical basis was presented in chapter 3 which has to do with our four main types of models which are Linear Trend with Seasonal Terms, Seasonal Exponential Smoothing, ARIMA, Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing. So that as discussed already, all the ARIMA models are one model, and the Simple Exponential Smoothing and Linear (Holt) Exponential Smoothing is also one model family since the Simple Exponential Smoothing is just basics of the Linear (Holt) Exponential Smoothing. Generally in arriving at a model, three basic steps were followed that is; Preliminary analysis, model fitting and then the forecasting.

There have been some amount of academic work in the form of thesis that has been done individually on rainfall and sunshine (weather generally) using some of the above mentioned models separately, and in many of these works there is a mention that there are other models that can be looked at on the subject of forecasting rainfall in the regions of Ghana.

In this work, we have gone on to look at four major models. We have realized that the linear trend with seasonal terms is the best model to use in the modeling of rainfall so far as these four regions are concerned.

In our analysis we have realized that the best model to use in the forecasting of rainfall in these four regions under study should be the Linear Trend with Seasonal Terms. We also think that fifteen years was more than enough for the study and the analysis can be considered as accurate. The exponential models are not the best model for forecasting the rainfall pattern in these four regions in Ghana.

From the Parameters table for the northern region, it’s clear that the month of June has been the month with the highest average significant rainfall.

Also, in the western region, the month of June has been the month with the highest average rainfall.

5.2 RECOMMENDATIONS

We recommend that:-

- At least for the four regions under study, the best model should be a Linear Trend with Seasonal Terms.
- There should be ways of handling the higher levels of rain we expect.
- There should be some more study in the rainfall pattern for the rest of the six regions in Ghana.
- There should be another study of the entire country to determine if the Linear Trend with Seasonal Terms will be suitable for the Country.
- After the end of 2011 there should be some evaluation on the accuracy of if the Linear Trend with Seasonal Terms on the rainfall patterns.

Clearly we believe that there is still a lot of work that can be done in this very interesting topic.

REFERENCES

1. Abraham, Bovas and Ledolter, Johannes, (1983), Statistical Methods for Forecasting,

2. Állóová, M. (2006): Analýza homogenity a trendov zrážok na lokalitách mokradí Abrod, Belianske lúky a Kláštorské lúky. diplomová práca, Slovak Technical University, Bratislava, Slovakia, 80 pp.

3. Amendola, A. (2003): Forecasting performance of regime-switching models in hydrological time series. Giornata di Studio: Metodi Statistici e Matematici per le analisi Idrologiche-Roma, Am. Soc. Eng. 116, pp. 770–779.

4. Anderson, K. and Vecchia, A., (1993). Testing for periodic autocorrelations in seasonal time series data. Biometrika 78, 53-56

5. Bandyopadhyay, G., Chattopadhyay, S., (2007). Single hidden layer artificial neural network models versus multiple linear regression model in forecasting the time series of total ozone. Int. J. Environ. Sci. Tech., 4 (1), 141-149.

6. Bloomfield, Peter, (1976), Fourier analysis of Time Series: An Introduction. 1976. John

7. Box G. E. P. and Jenkins, G. M. (1976), Times Series Analysis Forecasting and control, Oakland, CA: Holden Day

8. Brown, R. G. (1962), *Smoothing, Forecasting and Prediction*. Englewood Cliffs, N.J.: Prentice Hall.

9. Brown, R. G. (1959), “Less-risk in inventory estimates,” *Harvard Business Review*, July-August, 104-116.

10. Caldwell J.G., (2006), The Box-Jenkins Forecasting Technique, Chapman and Hall. London

11. Chatfield, C, (1988), The Analysis of Time Series. An Introduction, Fourth Edition,

12. Chattopadhyay, A., S. Gupta, V.K. Sharma and P. Kumari, (2009), Propagation of Shear waves in visco- elastic medium at irregular boundaries. Acta Geophysica, 58: 195-214

13. Chattopadhyay, G., (2007), Univariate modelling of summer-monsoon rainfall time series: Comparison between ARIMA and ARNN Comptes Rendus Geoscience, Volume 342, Issue 2, Pages 100-107,

14. Chattopadhyay, S. and Chattopadhyay, G. (2008), Comparative study among different neural net learning algorithms applied to rainfall time series. Meteorological Applications, 15: 273–280. doi: 10.1002/met.71

15. Chin, E.H. (1977) Modelling daily precipitation occurrence process with Markov chain. Wat. Resour. Res. 13 (6), 949-956.

16. Clark, C. O., J. E. Cole, and P. J. Webster, 2000: Indian Ocean SST and Indian summer rainfall: Predictive relationships and their decadal variability. *J. Climate*, 13**,** 2503–2519.

17. Cracknell, A. P., and Varotsos, C., (2007), Global climatology and ecodynamics: anthropogenic changes to planet Earth, Springer Berlin Heidelberg, New York

18. Dahale, S. D. and Singh, S. V., (1993) Modeling of Indian monsoon rainfall series by univariate box—Jenkins type of models Advances in Atmospheric Sciences 10: 211-220, May 01, 1993

19. Del Sole, T., and J. Shukla, 2002: Linear prediction of Indian monsoon rainfall. Journal of Climate, vol. **15,** 3645-3658.

20. Delleur, J, W., and Kavvas, (1978) M. L., Stochastic Models for Monthly Rainfall Forecasting and Synthetic Generation, Journal of Applied Meteorology, vol. 17, Issue 10, pp.1528-1536

21. Elsner, J. B. and Tsosis, A. A., (1992), nonlinear prediction, chaos and noise, bull, Amer Meteor. Soc, 73, 49-60

22. Ferm, M., De Santis, F., and Varotsos, C. (2005): Nitric acid measurements in connection with corrosion studies, Atmos. Environ., 39, 6664–6672.

23. Ferm, M., Watt, J., O'Hanlon, S., De Santis, F., and Varotsos, C. (2006): Deposition measurement of particulate matter in connection with corrosion studies, Anal. Bioanal. Chem., 384, 1320–1330.

24. Ferranti, L., Oldow, J.S., D'Argenio, B., Catalano, R., Lewis, D., Marsella, E., Maschio, L., Pappone, G., Pepe, F., Sulli, A., (1990). Active deformation in Southern Italy, Sicily and southern Sardinia from GPS velocities of the Peri-Tyrrhenian Geodetic Array (PTGA). Boll. Soc. Geol. It., (Ital. J. Geosci.), 127/2, 299-316.

25. Gottman, M John, (1981), Time Series Analysis, A Comprehensive Introduction for Social Scientists, Cambridge University Press.

26. Gregory, J.M., Wigley, T.M.L. & Jones, P.D. (1993) Application of Markov models to area average daily precipitation series and inter annual variability in seasonal totals. Climate Dynamics 8: 299-310.

27. Grimaldi, S., Serinaldi, F., and Tallerini, C., (2005), Multivariate linear parametric models applied to daily rainfall time series, Institute of Research for Hydrogeological Protection, CNR-IRPI, Perugia Italy

28. Guhathakurta, P (2008), Long-range monsoon rainfall prediction of 2005 for the districts and sub-division Kerala with artificial neural network. Current Science **90**:773-779.

29. Hamilton J.D., (1994), Times Series Analysis, Princeton University Press, Princeton-USA

30. Harvey Andrew C., (1990), The Econometric Analysis of Time Series, 2nd ed.(Oxford: Phillip Allan)

31. Hastenrath, S., (1995), Recession of the Yanamarey glacier in Cordillera Blanca, Peru, during the 20th century. In: J. Glaciol. 41, 191–196.

32. Holt, C.C. (1957) Forecasting seasonal and trends by exponentially weighted moving averages, Office of Naval Research, Research Memorandum No. 52.

33. http://en.wikipedia.org/wiki/Rain retrieved on November, 2010

34. http://northeastcricket.forumotion.com/t316-gauges-standard-rain-gauge retrieved on November, 2010

35. Islam, M.N., and B. Sivakumar, (2002), Characterization and prediction of runoff dynamics: A nonlinear dynamical view, Advances in Water Resources, 25(2), 179-190

36. Jayawardena, A. W., and F. Lai, (1994) Analysis and prediction of chaos in rainfall and stream flow time series, J. Hydrol., 153, 23–52.

37. Jhajharia D, Shrivastava SK, Sarkar D, Sarkar S (2009), temporal characteristics of pan evaporation trends under the humid conditions of northeast India. Agric For Meteorol 149:763–770, John Wiley and Sons Press, N.Y.

38. Jones, A.L. and Smart, P.L., (2005), Spatial and temporal changes in the structure of groundwater nitrate concentration time series (1935-1999) as demonstrated by autoregressive modeling Journal of Hydrology, Volume 310, Issue 1-4, August 2005, Pages 201-215

39. Jones, P.D., Groisman, P.Ya., Coughlan, M., Plummer, N., Wang, W-C. and Karl, T.R., (1990) Assessment of urbanization effects in time series of surface air temperature over land. Nature 347, 169-172.

40. Jones, R. and W. Brelsford (1967). Time series with periodic structure. Biometrika *54*, 403-408

41. Khan S., A. R. Ganguly, and Saigal, S. (2005), Detection and predictive modeling of chaos in finite hydrological time series. Nonlinear Processes in Geophysics 12, 41–53.

42. Komorníková, J. K. and Szökeová, D., (2004), Compa-ring of forecasting methods in hydrologic time series. Proceedings of 3th International Conference APLIMAT 2004, February 2004, Bratislava, 131–236.

43. Komorníková, M., Szökeová, D. (2004b): Modelling in hydrologic time series and testing for long memory. In: Proceedings MAGIA 2004, Kočovce, 68; 77 pp.

44. Kondratyev K Ya, Varotsos C (2000): Atmospheric Ozone Variability: Implications for Climate Change, Human Health and Ecosystems, Springer-Praxis, Chichester, UK,

45. Kondratyev, A., Cortadella, J, Kishinevsky, M, Pastor, E, Roig, O, Alex Yakovlev, (1995) Checking Signal Transition Graph Implementability by Symbolic BDD Traversal

46. Krishnamurthy, V. And Shukla, J., (2000), Intraseasonal and Interannual Variability of Rainfall over India, Center for Ocean–Land–Atmosphere Studies, Institute of Global Environment and Society, Inc., Calverton, Maryland

47. Kuligowski, R.J. and Barros, A.P., (1998), Experiment in short-term precipitation forecasting using artificial neural networks. Monthly Weather Rev., 126: 470-482.

48. McCann, D.W., (1992), "A Neural Network Short-Term Forecast of Significant Thunderstorms", Weather and Forecasting, **7**, 525-534, doi: 10.1175/1520-0434

49. McKeon, G.M. and Hall, W.B. (2000), The use of climate information in sheep and beef cattle industries of Queensland, Greg/Papers/Use of Climate Information vol 23,

50. McKeon, G.M., et al., (2004) Historical degradation episodes in Australia: Global climate and economic forces and their interaction with rangeland grazing systems. In: Pasture degradation and recovery in Australia's rangelands. Natural Resource Sciences, Queensland Government, 256 pp.

51. Mohanty, J. and Mohapatra, P., (2007) Univariate modelling of summer-monsoon rainfall time series: Comparison between ARIMA and ARNN, Department of Computer Application, Pailan College of Management and Technology, 700 104 Kolkata, India

52. Mohanty, U.C. and Mohapatra, M., (2007) Prediction of occurrence and quantity of daily summer monsoon precipitation over Orissa (India), Meteorological Applications, 14: 95-103

53. Naill P. E. and Momani M. (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.

54. Nicholls, N. and B. Lavery, (1992), Australian rainfall trends during the twentieth century. Int. J. Climatol., 12: 153-163.

55. Pagano, M. (1978). On periodic and multiple autoregressions. Annals of Statistics 6, 1310-1317.

56. Pekarova, P. & Pekar, J. (2004), Teleconnections of AO, NAO, SO and QBO with interannual streamflow fluctuation in the Hron basin. J. Hydrol. Hydromech. **52**(4), 279–290.

57. Pekárová, P., Koníček, A., Miklánek, P., Stančík, Š. (2004): Lifespan estimation of the considered sediment trapping ponds in the upper Torysa basin using AGNPS model (Part I. Rainfall scenario creation). Acta Hydrologica Slovaca, 5, 2, p. 286-292.

58. Pekárová, P., Miklánek, P., Pekár, J., Rončák, P., (2005): Long-term prediction of water quality determinants in the Danube River using linear autoregressive models., Slovak Technical University, Bratislava, Slovakia

59. Pittock, A.B., (1975), Climatic change and the patterns of variation in Australian rainfall. Search, 6: 498-504.

60. Plummer, N., (1991) Annual mean temperature anomalies over eastern Australia. Bull. Aust. Meteorol. Oceanogr. Soc., 4: 42-44.

61. Principe J. C., Kuo, J.M. and S. Celebi, (1992), An Analysis of Short Term Memory Structures in Dynamic Neural Networks", accepted in the special issue of recurrent networks of IEEE Trans. on Neural Networks.

62. Rajeevan, M, (2004), Tropical Pacific upper ocean heat content variations and Indian summer monsoon rainfall India Meteorological Department, Pune, India GEOPHYSICAL RESEARCH LETTERS, VOL. 31,

63. Ramirez BG, Blazquez C, Gomez del Pulgar T, Guzman M, de Ceballos ML. (2005) Prevention of Alzheimer's disease pathology by cannabinoids: neuroprotection mediated by blockade of microglial activation. J Neurosci. 2005 Feb 23;25(8):1904-13.

64. Salinger M. J, (1995) Climate variability and wheat baking quality, New Zealand Journal of Crop and Horticultural Science, , Vol. 23: 289-298

65. Salinger, M. J.; McGann, R. P.; Coutts, L.; Collen, B.; Fouhy, E. (1992): Rainfall trends in New Zealand and outlying islands. Wellington, New Zealand Meteorological Service, Wellington. 33 p.

66. Sardeshmukh, P.D. and Hoskins, B.J., (1988), On the generation of global rotational flow by steady idealized tropical divergence. J.Atmos.Sci., **45**, 1228-1251.

67. SAS manual SAS/ETS Software: (1995) Time Series Forecasting System, Version 6, First Edition, Cary, NC. SAS Institute Inc., , 225-235

68. Shao, Q., Lund, R.B. (2004):. Computation and characterization of autocorrelations and partial autocorrelations in periodic ARMA models, J. Time Ser. Anal. 25(3), p. 359-372.

69. Shin S, Tsujita M, Tomizawa H, Sakihama N, Kanei K, Oshino R (1990) Proteolytic degradation of ferredoxin-NADP reductase during purification from spinach. Arch Biochem Biophys 279

70. Silverman, D. and Dracup, J. A., (2000), Artificial Neural Networks and Long Range Precipitation Prediction in California," Journal of Applied Meteorology, Vol. 39, No. 1, January, pp. 57-66

71. Singhtaun, C. and Charnsethikul P, (2007) An Efficient Algorithm for Capacitated Multifacility Location Problems, Journal of Computer Science,. 3(8), 283-291

72. Singhtaun, C. and Charnsethikul, P., (2010) Comparison of Exact Algorithms for Rectilinear Distance Single-Source Capacitated Multi-Facility Weber Problems, Journal of Computer Science 6 (2): 112-116, 2010, ISSN 1549-3636, Operations Research and Management Science Units, Department of Industrial Engineering, Kasetsart University, Bangkok 10903, Thailand

73. Sivakumar, B. (2001) Rainfall dynamics at different temporal scales: a chaotic perspective. Hvdrol. Earth System Sci. 5(4), 645-651.

74. Sivakumar, B. (2002). Review on "Hydrology and Floodplain Analysis" by Bedient, P. B., Huber, W. C., Third Edition, Prentice Hall, Inc., New Jersey. Journal of the American Water Resources Association, 38(4), 1169-1171.

75. Sivakumar, M.V.K. and Ndegwa, N. (2007), Climate And Land Degradation, Springer: Heidelberg, Germany, Social Scientists. Cambridge University Press.

76. Soltani, A., M. Gholipoor and K. Ghassemi-Golezani, (2007). Analysis of temperature and CO_2 effects on radiation use efficiency in chickpea (Cicer arietinum L.). J. Plant Sci., 2: 89-95

77. Suppiah, R. and K.J. Hennessy, (1996). Trends in the intensity and frequency of heavy rainfall in tropical Australia and links with the Southern Oscillation. Aust. Meteorol. Mag., 45: 1-17.

78. Tesfaye, Y.G., Meerschaert, MM., Anderson, PL. (2006): Identification of periodic autoregressive moving average models and their application to the modeling of river flows, Water Resource Research, vol. 42, W01419.

79. Toth, E., Brath, A. and Montanari, A (2000), Comparison of short-term rainfall prediction models for real-time flood forecasting. Journal of Hydrology 239, 132-147.

80. Troutman, B. (1979). Some results in periodic autoregression. Biometrika 66, 219-228

81. Tularam, G.A. and Ilahee, (2010) M Time Series Analysis of Rainfall and Temperature Interactions in Coastal Catchments Journal of Mathematics and Statistics 6 (3): 372-380, 2010 ISSN 1549-3644, Science Publications Department of Mathematics and Statistics

82. Tzanis, C., and C. A. Varotsos, (2008): Tropospheric aerosol forcing of climate: A case study for the greater area of Greece. Int. J. Remote Sens., 29, 2507–2517.

83. Ula, T.A., Smaldi, A.A. (1997): Periodic stationarity conditions of multivariate periodic autoregressive moving average processes as eigen-value problems. Water Resource Research 33(8), pp. 1929-1934.

84. Varotsos, C. & Kirk-Davidoff, D. (2006), Long-memory processes in ozone and temperature variations at the region 60 degrees S-60 degrees N. Atmos. Chem. Phys. 6, 4093–4100.

85. Varotsos, C., (2005), The southern hemisphere ozone hole split in 2002, Environ. Sci. Pollut. R., 9, 375–376, 2002.Wiley and Sons.

86. Wilks, D.S. (1998) Multisite generalization of a daily stochastic precipitation generation model. Journal of Hydrology 210: 178-191.

87. Yaffee R and McGee, M (2000), Introduction to Times series Analysis and Forecasting with Applications of SAS and SPSS, Academic Press, Califonia-USA.

APPENDIX A

Table A-1 Northern Region Average Monthly Rainfall Per Community / Town (mm)

Year	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
1995	0.0	1.2	56.1	100.2	134.5	147.5	111.4	212.7	145.4	111.5	19.6	4.4
1996	1.6	20.7	18.4	104.6	158.1	242.8	94.7	168.9	157.1	96.9	35.1	7.9
1997	8.0	2.4	80.1	164.8	179.3	271.9	115.7	100.9	195.7	101.8	39.4	24.0
1998	5.6	21.3	4.7	67.3	147.7	144.5	82.7	120.4	191.3	116.7	4.2	5.1
1999	26.5	17.2	30.6	114.0	107.7	198.7	172.5	168.6	170.2	122.4	15.7	2.3
2000	32.7	0.0	59.1	76.3	144.6	175.5	94.5	171.9	209.9	71.4	27.0	3.0
2001	0.0	7.3	71.9	106.5	153.9	140.1	107.5	121.2	176.8	35.2	18.4	4.9
2002	7.8	22.8	43.2	116.8	115.7	172.1	162.3	188.1	101.1	75.3	16.9	0.6
2003	5.4	17.9	52.0	113.1	146.2	179.3	70.2	120.4	151.5	105.1	28.6	5.1
2004	18.5	7.0	56.9	99.1	152.7	147.9	185.7	235.2	169.4	91.6	49.7	8.1
2005	9.1	14.5	57.0	119.8	175.4	216.6	201.6	82.6	199.6	118.1	19.8	25.8
2006	1.9	7.2	12.7	99.1	199.3	132.4	129.3	126.5	138.7	137.5	18.9	1.0
2007	0.0	6.4	39.9	130.1	103.3	122.2	116.7	183.7	203.1	119.6	10.8	14.6
2008	0.7	1.7	67.9	58.3	146.7	165.7	114.5	209.3	208.8	110.3	41.4	17.4
2009	5.0	32.4	43.1	133.6	113.3	181.1	257.3	151.9	154.2	107.4	3.4	10.6

Table A-2 Western Region Average Monthly Rainfall Per Community / Town (mm)

Year	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
1995	9.3	25.9	130.3	217.5	164.9	301.7	155.4	112.5	89.5	98.8	120.9	27.0
1996	25.3	65.9	138.5	161.7	333.7	330.7	230.5	115.2	38.4	126.2	125.6	30.4
1997	12.5	26.5	109.8	183.5	261.2	439.0	51.5	24.5	42.9	164.3	113.7	57.9
1998	38.2	52.4	57.3	92.1	138.7	180.2	71.4	69.5	55.1	250.8	111.9	74.3
1999	39.3	45.9	112.0	174.7	151.8	367.1	268.5	95.9	65.6	134.4	100.9	36.6
2000	26.5	12.8	85.3	136.6	226.4	300.2	111.5	70.9	81.1	47.9	65.6	68.1
2001	1.0	25.4	135.5	180.5	218.4	278.3	90.7	48.6	59.9	133.8	102.7	21.5
2002	43.2	35.2	70.5	172.3	197.2	316.7	349.6	130.3	51.0	178.2	114.3	33.1
2003	44.2	63.3	103.7	207.2	202.0	221.0	46.0	43.2	48.6	250.6	129.1	35.1
2004	59.6	87.1	91.6	56.6	159.7	159.7	156.3	52.6	200.6	254.1	110.8	61.4
2005	20.9	74.9	169.9	140.8	263.6	218.6	19.3	43.9	68.1	207.4	90.6	35.3
2006	29.1	58.9	68.6	104.3	284.6	213.9	109.7	48.3	103.2	151.8	70.9	40.3
2007	0.8	22.6	55.6	153.6	132.9	352.1	266.2	69.4	141.6	276.1	55.3	38.2
2008	8.3	54.0	97.4	128.0	251.5	280.2	127.7	64.4	128.1	183.0	102.4	74.7
2009	15.1	52.1	59.3	130.0	227.2	389.9	217.7	75.5	34.7	81.0	84.0	46.2

Table A-3 Eastern Region Average Monthly Rainfall Per Community / Town (mm)

Year	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
1995	1.4	13.5	129.1	171.3	122.2	250.5	139.4	127.1	134.2	147.4	64.0	66.3
1996	27.0	82.3	146.8	139.1	288.7	151.0	166.7	80.3	56.7	194.2	70.2	33.7
1997	14.7	7.6	95.9	196.0	153.4	339.7	110.9	33.6	53.4	134.8	117.7	61.3
1998	33.4	64.4	60.3	102.9	228.5	170.9	87.5	62.7	82.8	208.5	86.3	24.6
1999	30.1	33.4	125.8	131.9	147.4	311.9	209.5	143.5	88.5	139.8	79.7	3.1
2000	28.8	11.9	72.6	105.8	131.6	240.8	179.7	65.7	169.8	130.6	62.3	12.5
2001	0.3	1.0	144.6	120.2	112.1	193.9	39.1	12.1	81.6	131.4	104.9	42.3
2002	35.0	30.3	101.5	170.6	163.7	256.3	198.3	68.3	88.0	196.3	80.8	21.6
2003	27.9	51.2	63.2	150.7	162.4	138.2	80.8	37.8	80.2	143.5	87.3	31.1
2004	42.2	79.7	120.0	80.8	160.6	111.0	85.3	85.3	218.2	166.1	101.8	69.5
2005	1.8	65.7	152.5	82.2	97.5	153.2	52.2	37.0	93.7	154.7	97.9	29.2
2006	57.7	56.5	135.6	99.4	243.6	102.8	64.8	21.9	176.4	135.3	59.7	15.3
2007	3.4	26.8	77.4	185.8	107.2	138.6	138.0	54.1	256.3	200.5	55.7	32.9
2008	0.2	34.0	119.8	162.3	173.7	187.1	120.8	82.2	133.9	204.5	57.7	63.0
2009	12.5	50.7	101.8	144.5	145.3	237.1	140.1	45.7	71.8	106.4	70.0	19.8

Table A-4 Greater Accra Region Average Monthly Rainfall Per Community / Town (mm)

Year	Jan	Feb	March	April	May	June	July	Aug	Sept	Oct	Nov	Dec
1995	0.0	10.4	167.0	93.7	104.2	283.8	148.8	17.7	2.6	20.1	50.5	29.2
1996	0.0	56.2	72.9	172.8	228.8	191.8	99.6	31.2	11.7	3.0	18.6	1.9
1997	0.9	1.2	143.8	177.7	171.1	467.2	39.5	9.2	7.1	150.4	31.9	38.9
1998	0.0	7.1	0.9	52.0	165.8	57.0	6.1	6.1	8.8	101.5	24.3	15.0
1999	17.7	52.6	16.0	99.9	63.8	240.1	53.4	12.7	29.9	32.1	8.8	6.5
2000	0.7	0.0	44.6	48.6	124.7	81.5	40.4	9.1	6.9	36.5	25.8	31.3
2001	1.1	19.2	53.0	154.9	207.9	176.6	14.4	6.1	114.6	17.2	21.2	12.4
2002	53.5	12.7	32.3	146.4	116.6	343.1	48.6	7.4	27.4	63.6	30.4	2.6
2003	2.4	15.8	31.3	178.3	55.9	352.6	33.9	14.5	35.9	114.9	31.4	22.2
2004	10.1	11.9	8.0	26.4	120.0	73.1	20.5	36.1	93.0	96.8	32.6	2.4
2005	2.2	3.9	149.4	45.1	112.7	157.2	48.8	29.9	17.6	91.2	50.3	19.0
2006	10.2	2.0	26.9	39.9	216.0	144.1	40.5	7.9	81.8	107.2	3.8	3.0
2007	0.0	6.3	59.1	76.7	123.3	219.4	163.1	57.9	57.7	116.2	32.4	24.0
2008	3.8	0.0	60.6	101.3	249.4	131.1	100.5	25.6	42.4	47.9	73.1	62.2

2009	4.2	48.7	36.8	127.9	115.2	324.3	69.4	16.4	8.1	9.0	0.1	26.6

APPENDIX B

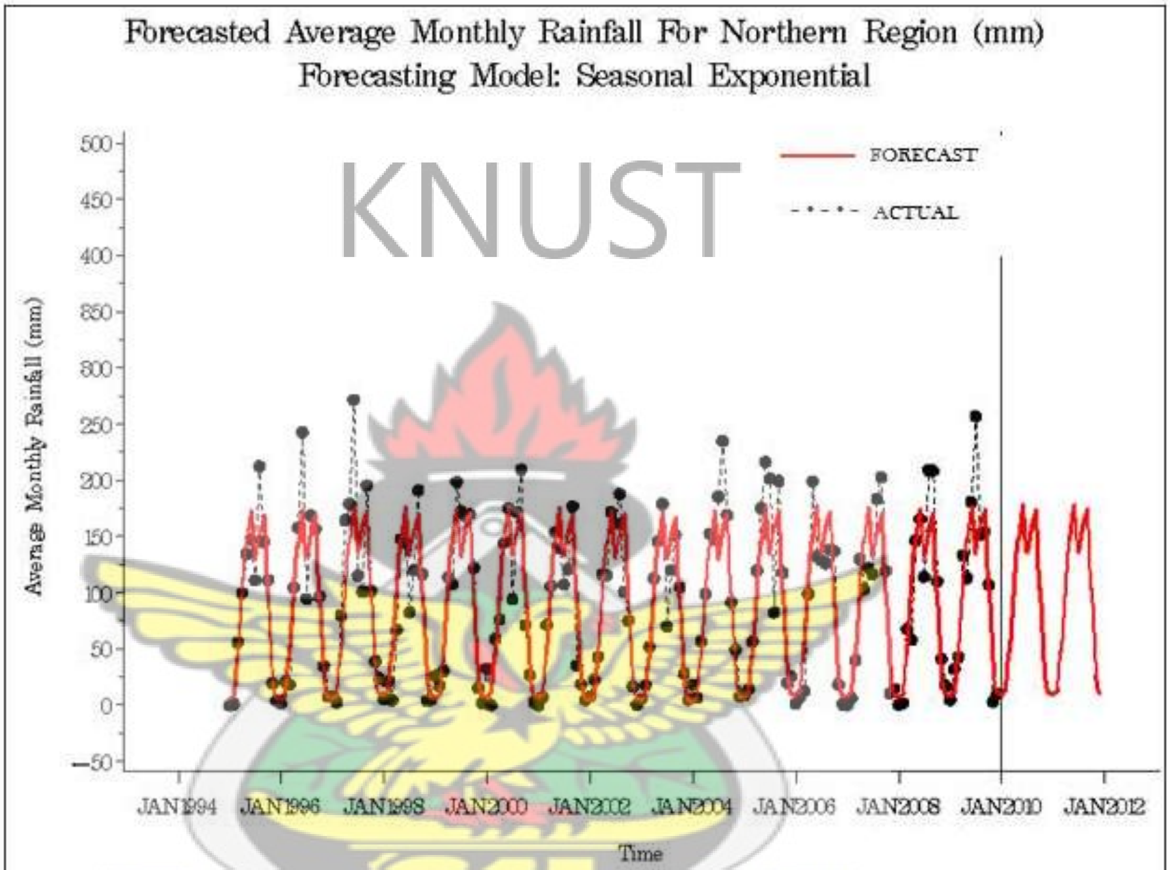

Figure B-1 Forecasting Model: Seasonal Exponential (Northern Region)

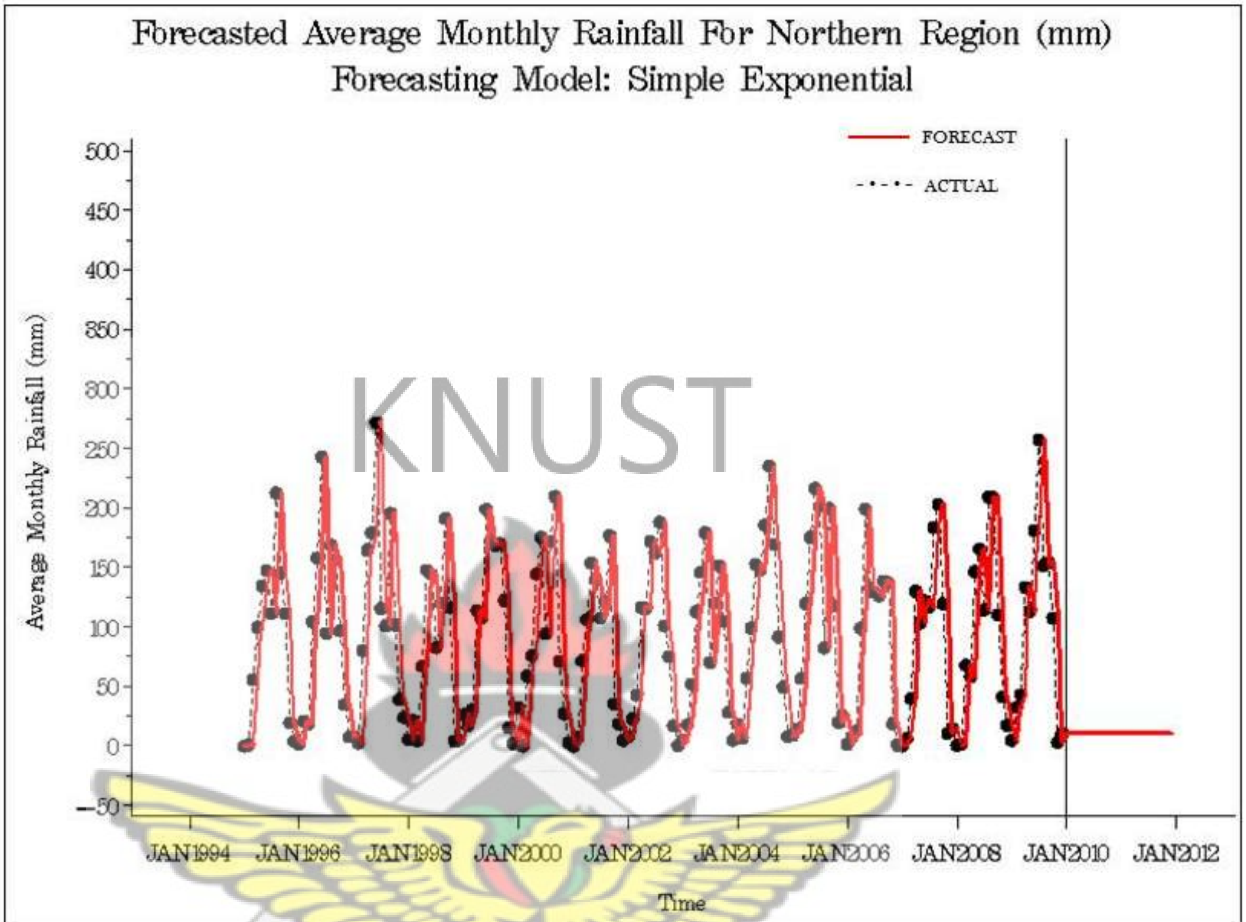

Figure B-2 Forecasting Model Simple Exponential (Northern Region)

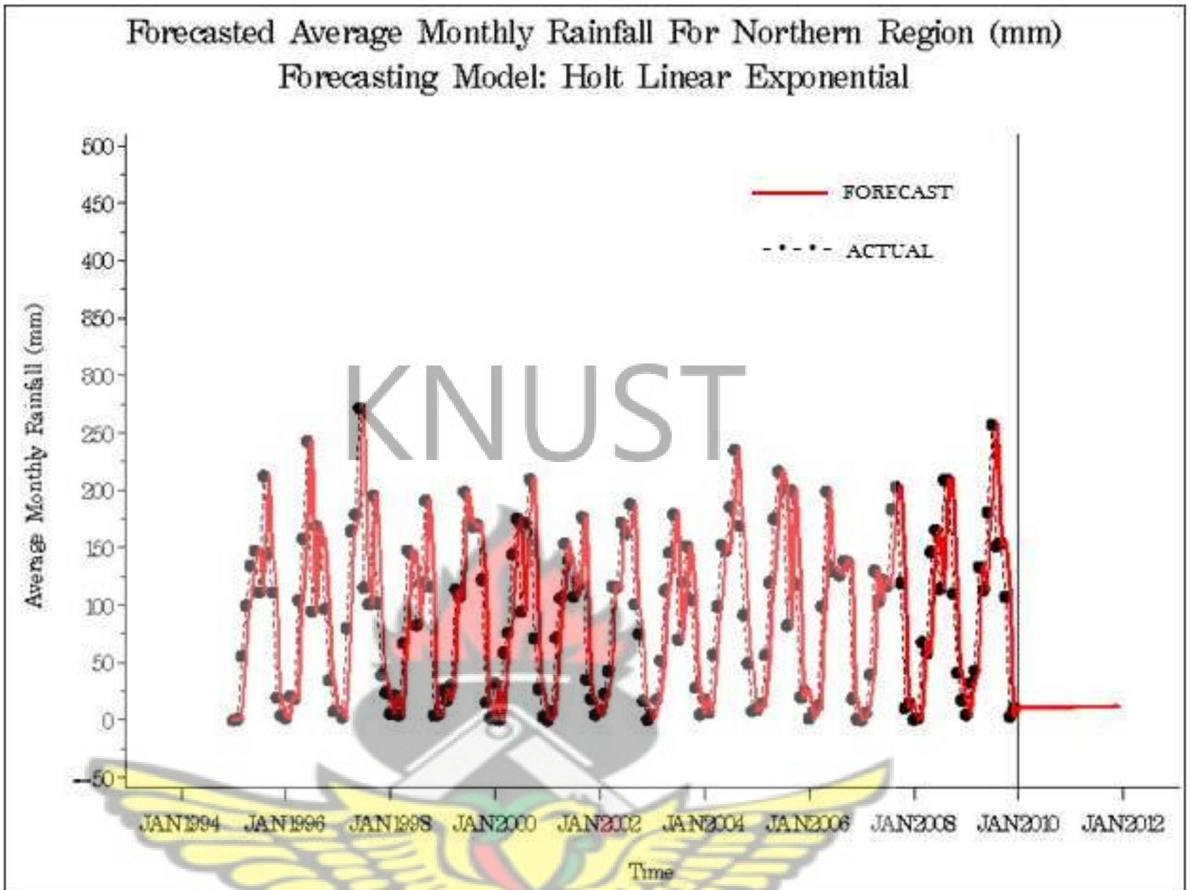

Figure B-3 Forecasting Model: Linear Holt Exponential (Northern Region)

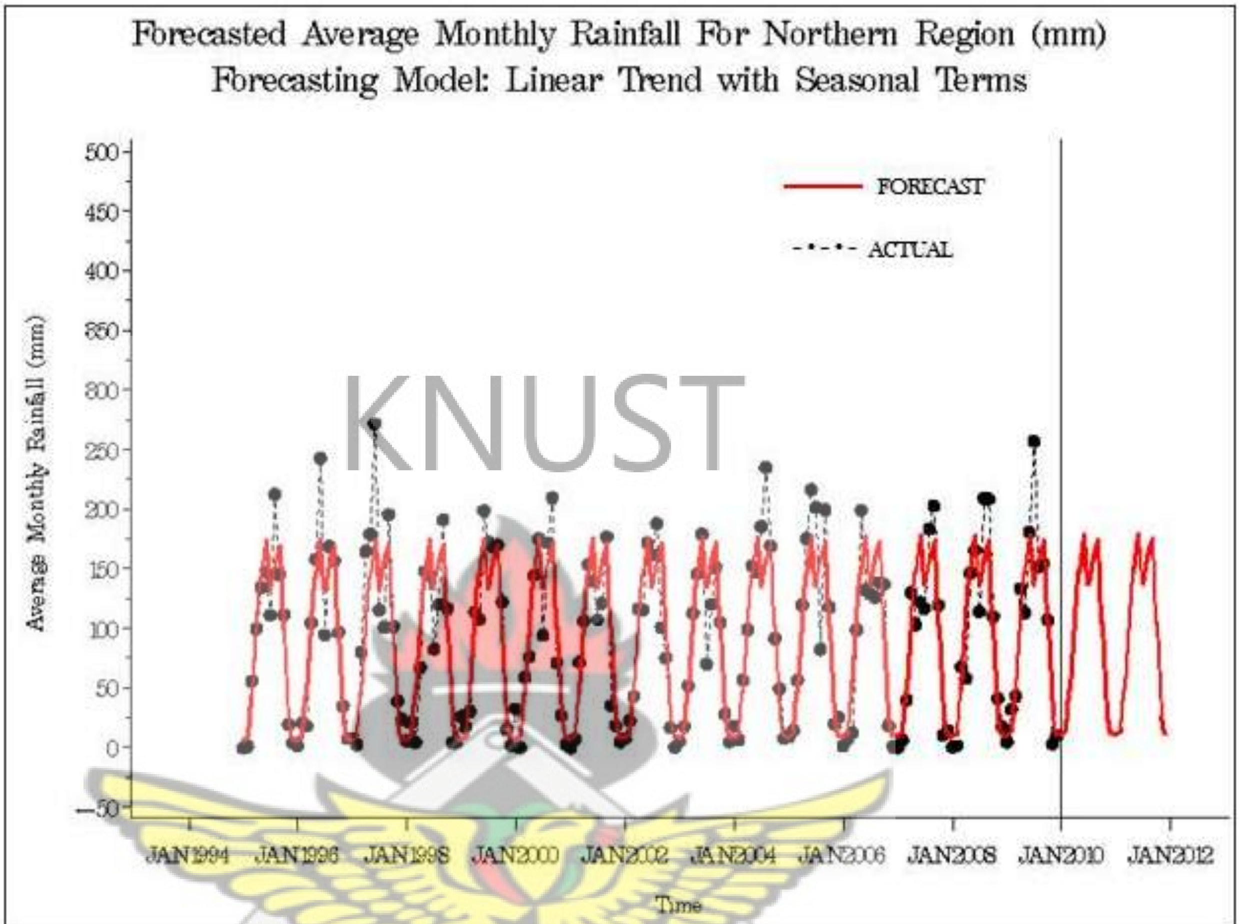

Figure B-4 Forecasting Model- Linear Model with Seasonal Trends (Northern Region)

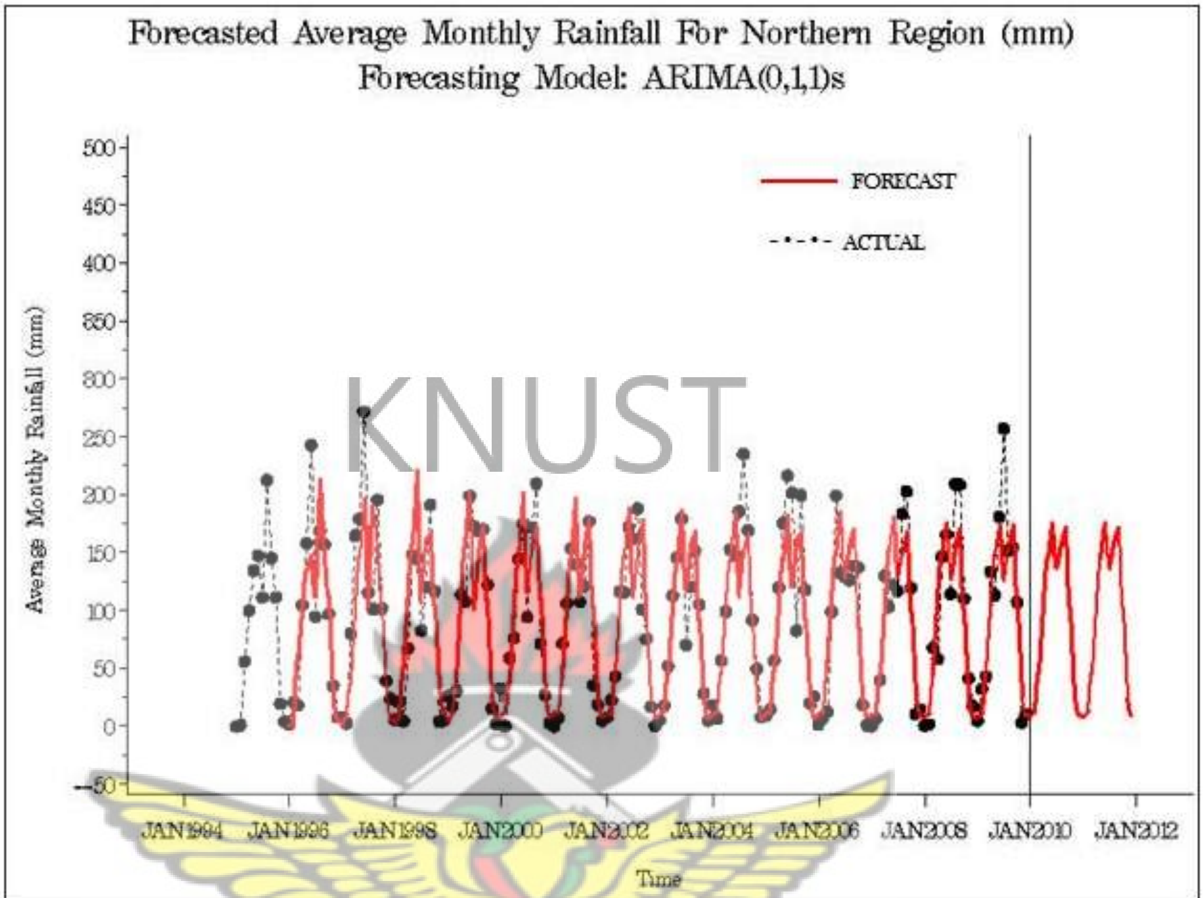

Figure B-5- Forecasting Model: ARIMA (0,1,1)s (Northern Region)

APPENDIX C

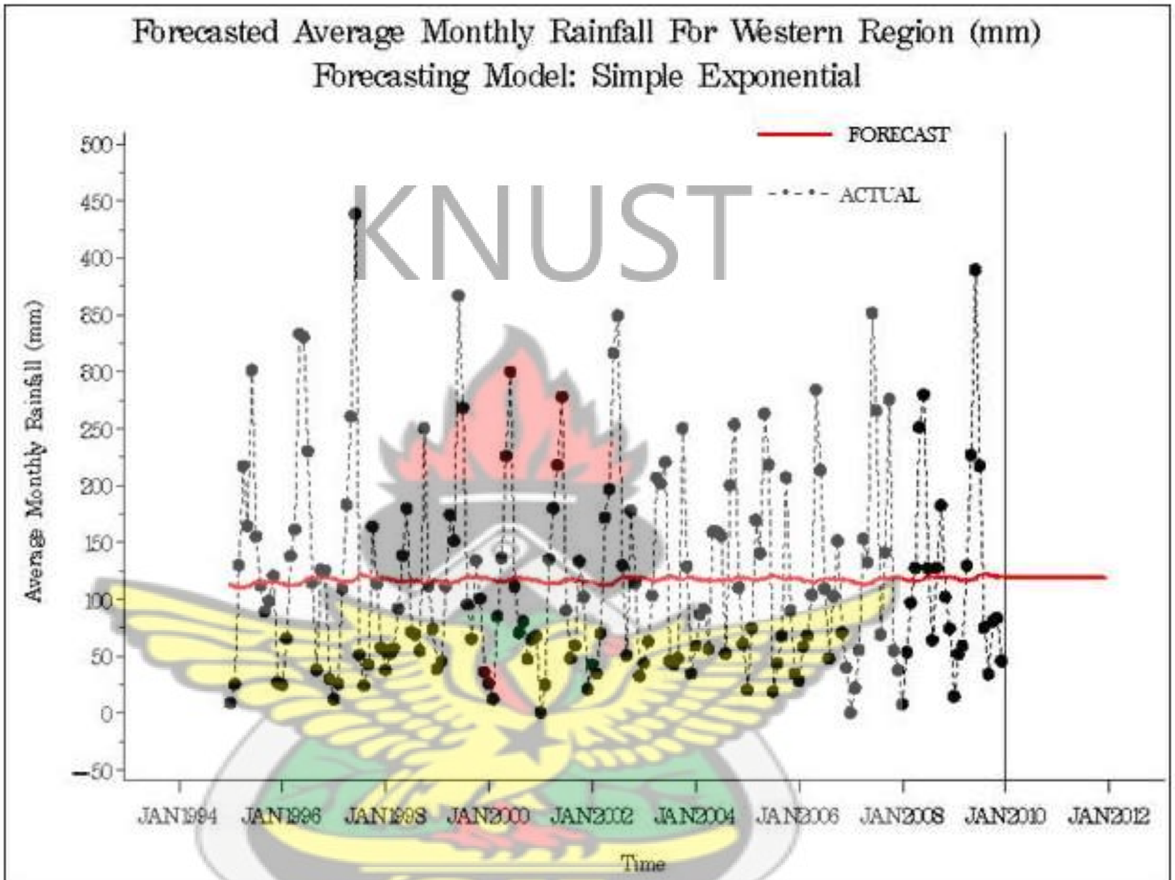

Figure C-1 Forecasting Model-Simple Exponential (Western Region)

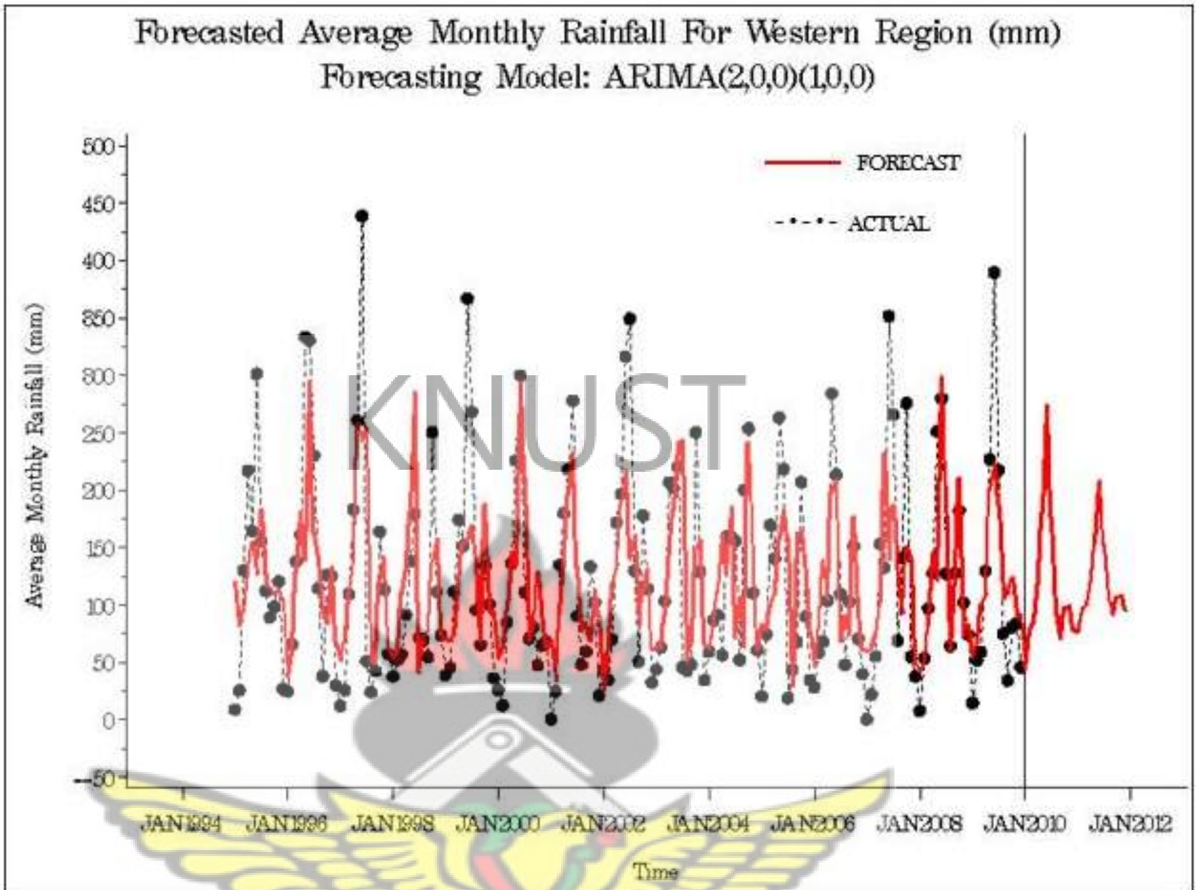

Figure C-2 Forecasting Model: ARIMA(2,0,0)(1,0,0) (Western Region)

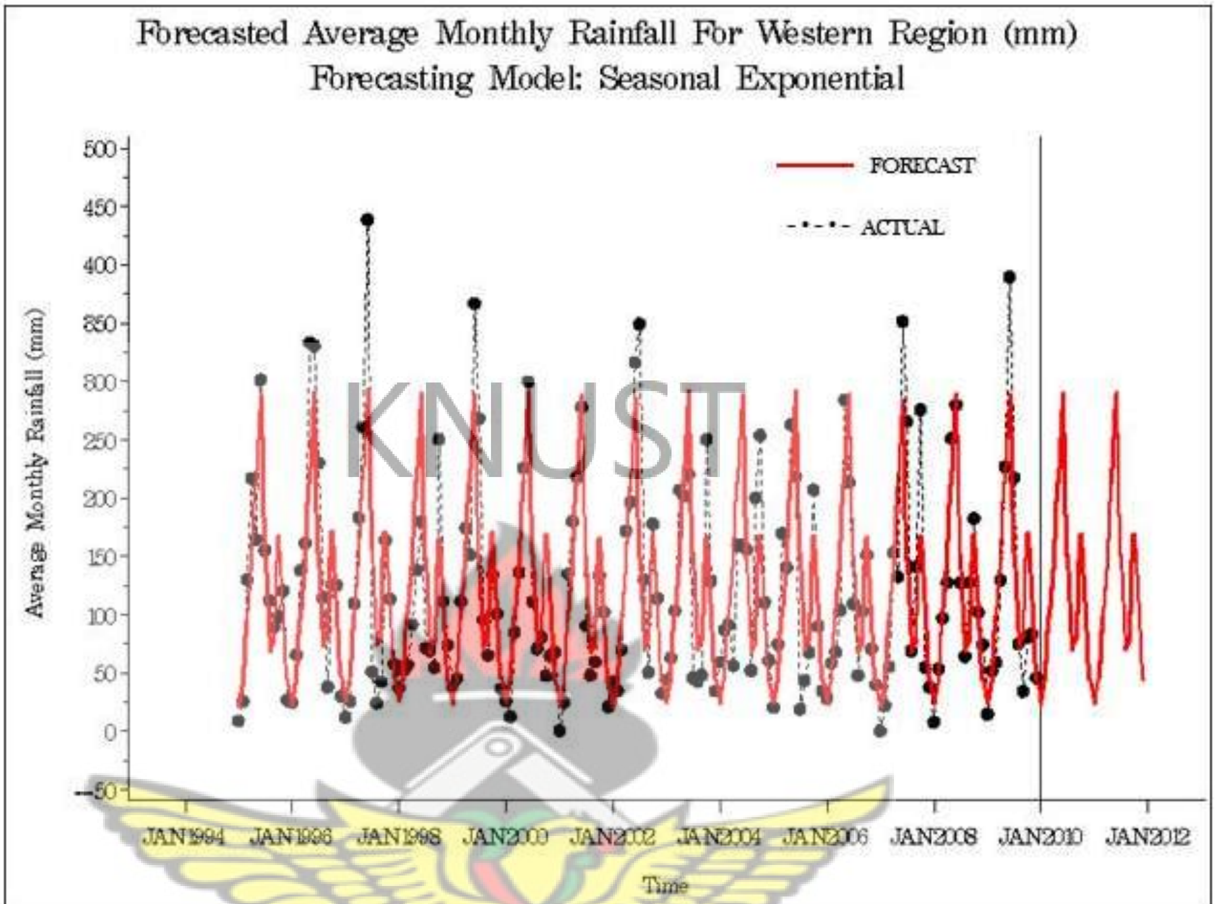

Figure C-3 Forecating Model: Seasonal Exponential (Western Region)

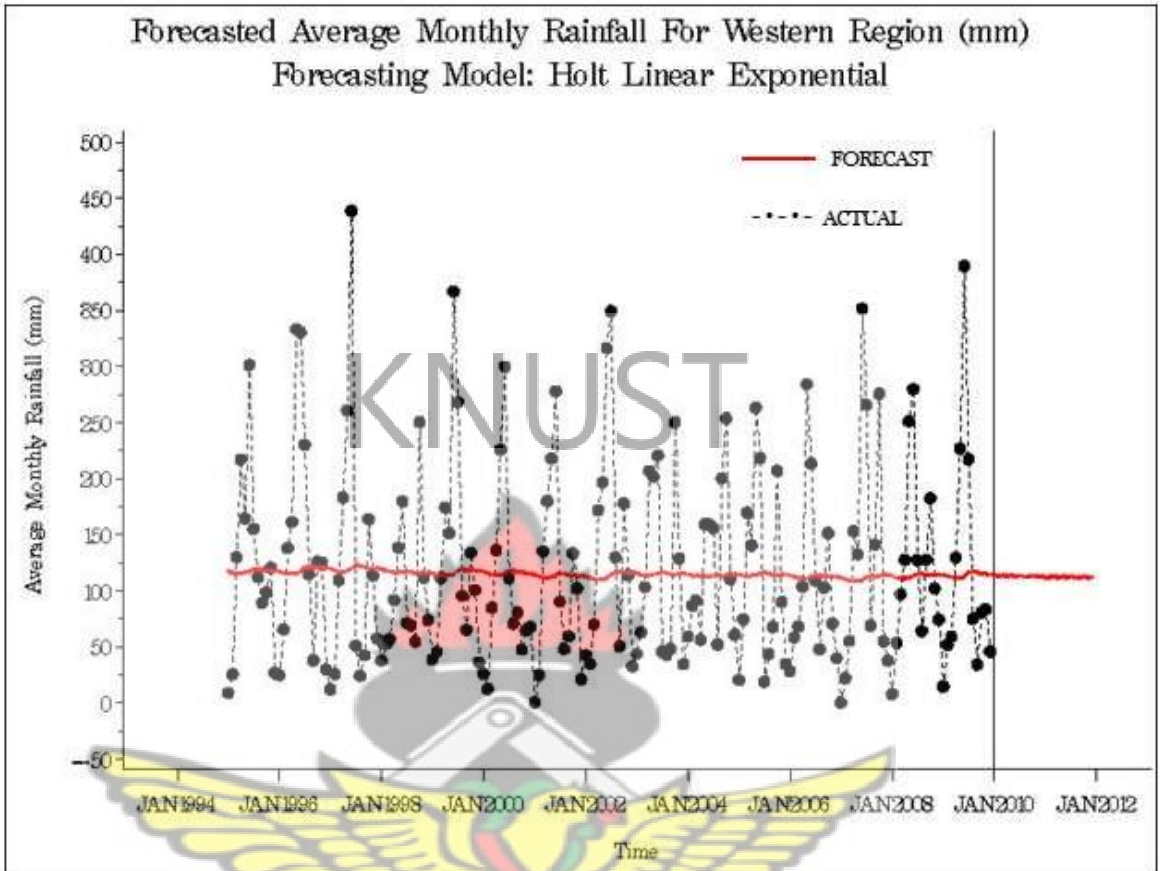

Figure C-4 Forecasting Model- Holt Linear Exponential (Western Region)

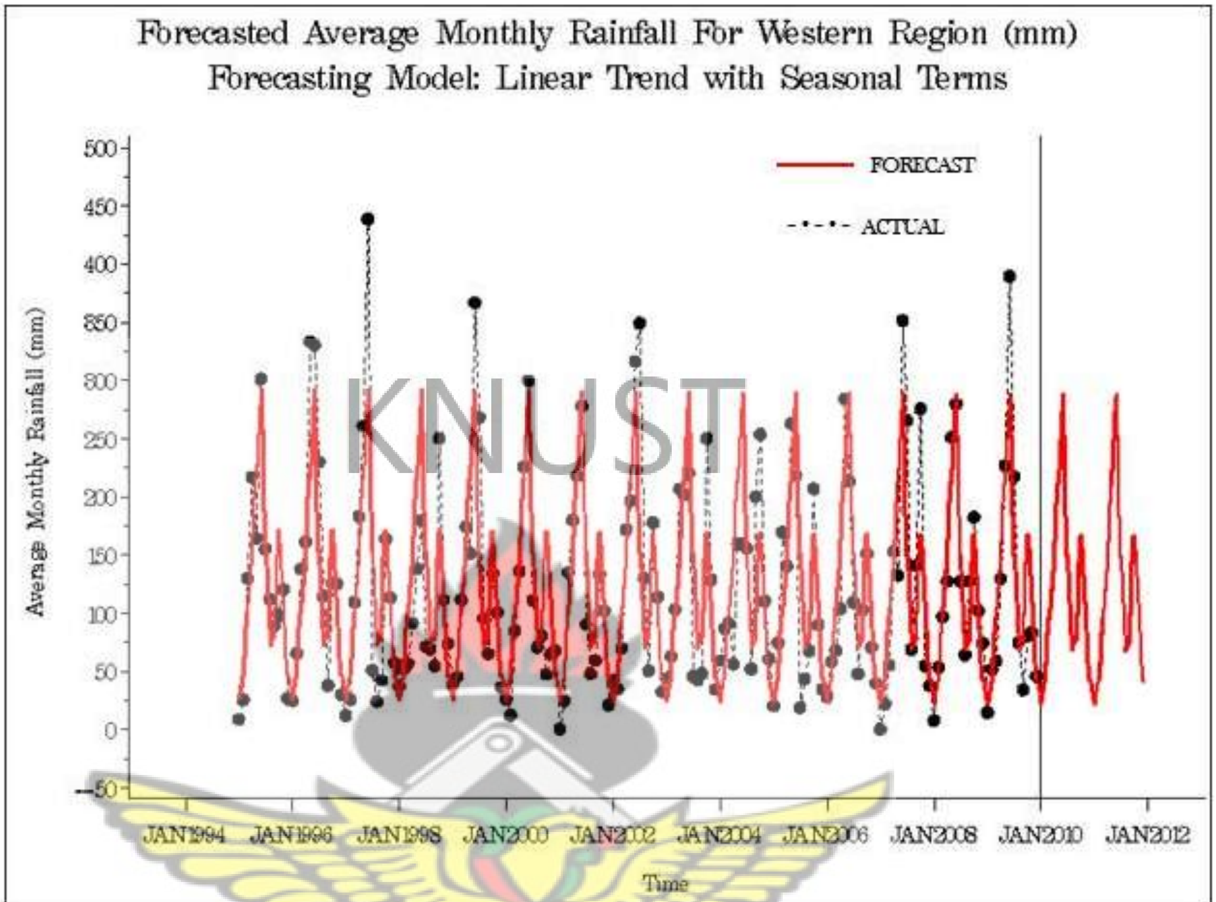

Figure C-5 Forecasting Model: Linear Trend with Seasonal Terms (Western region)

APPENDIX D

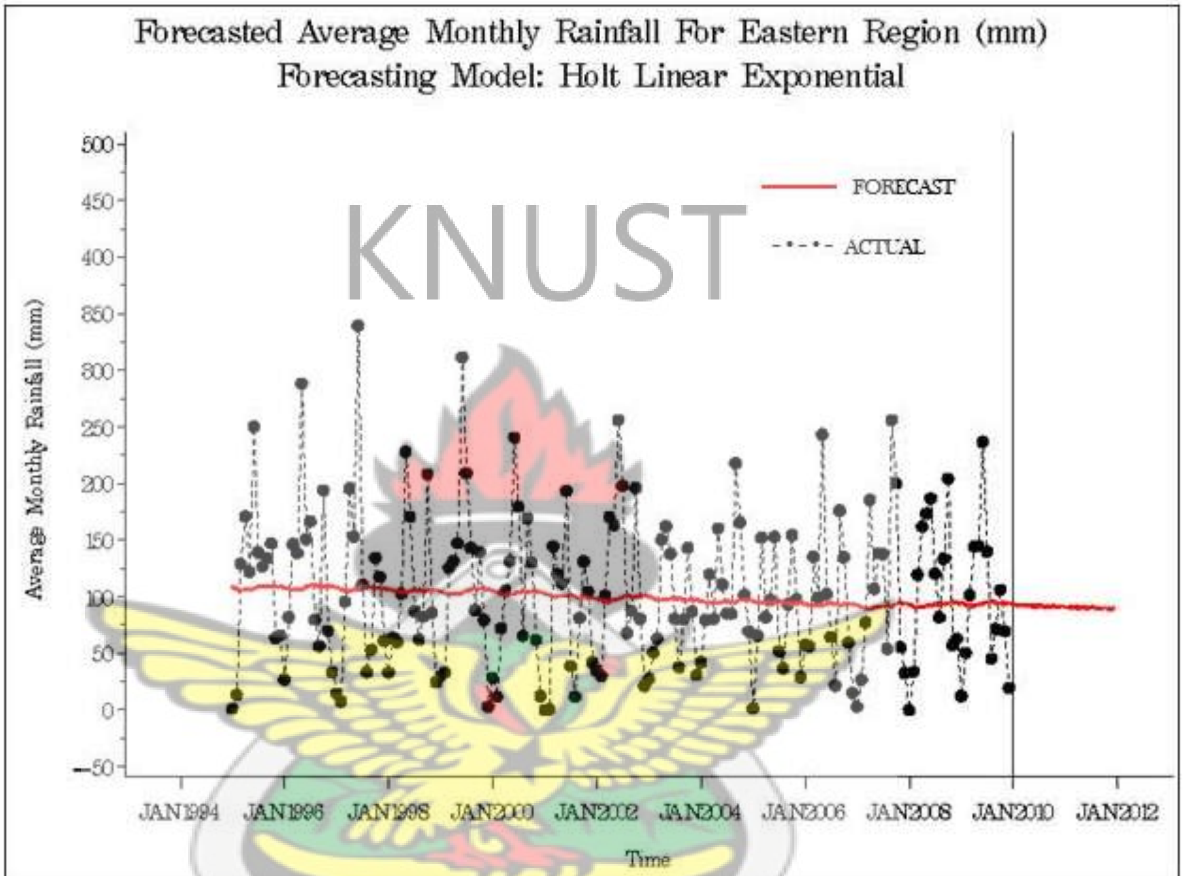

Figure D-1 Forecasting Model: Holt Linear Exponential (Eastern Region)

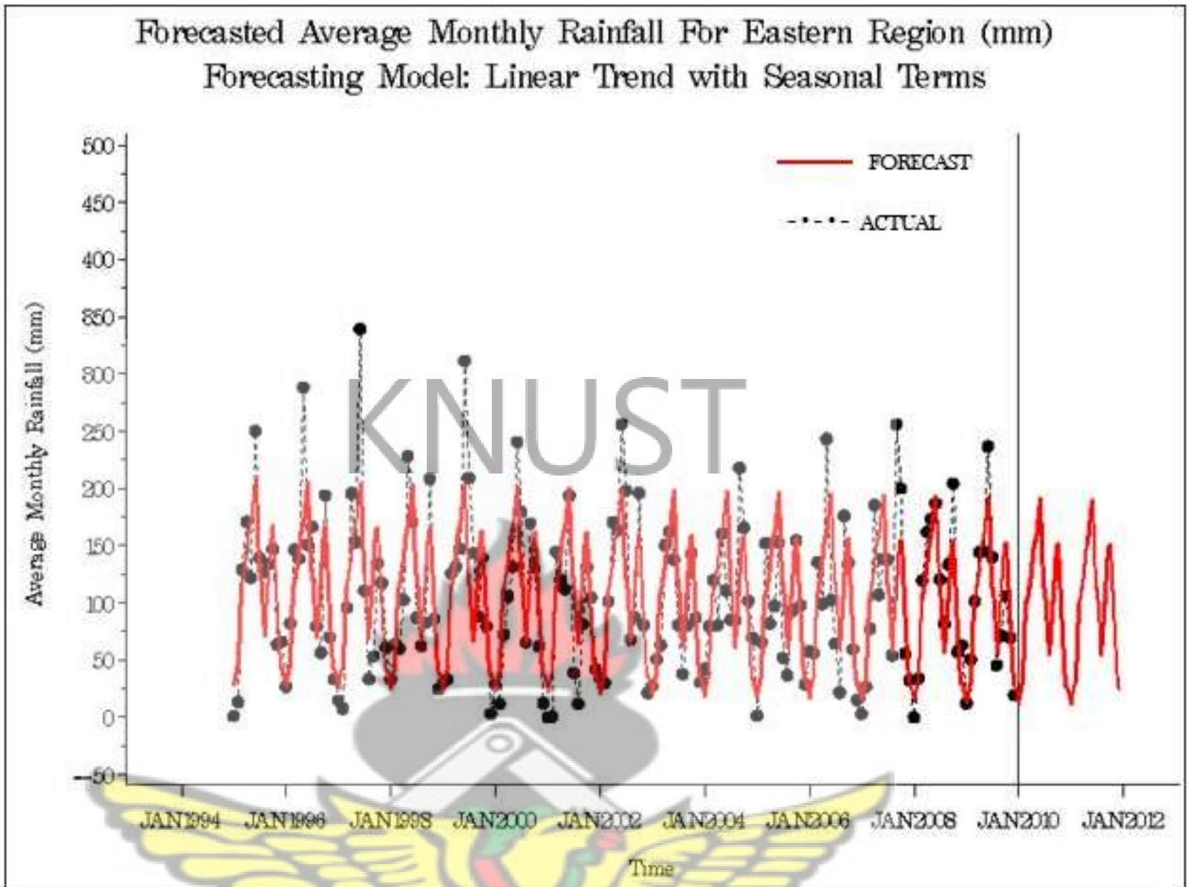

Figure D-2 Forecasting Model: Linear Trend with Seasonal Terms (Eastern Region)

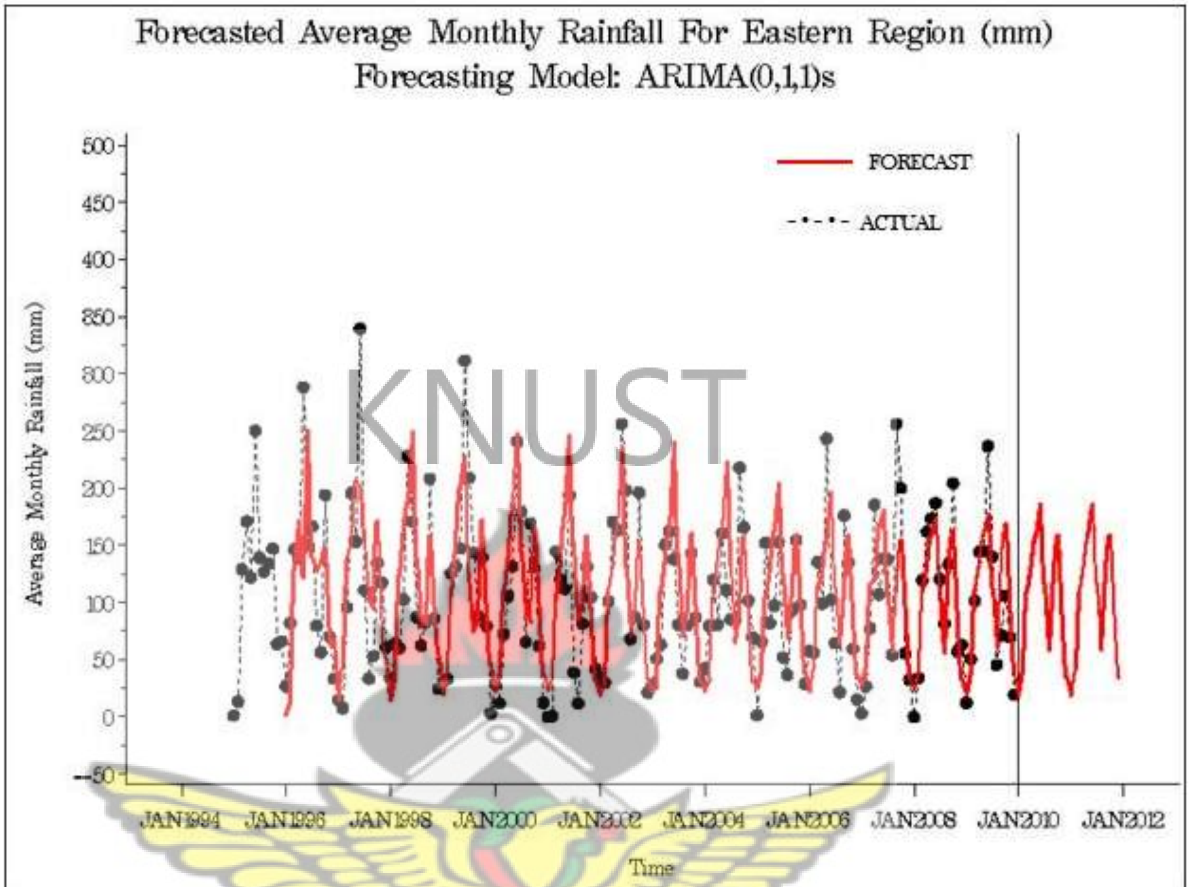

Figure D-3 Forecasting Model: ARIMA (0,1,1)s (Eastern region)

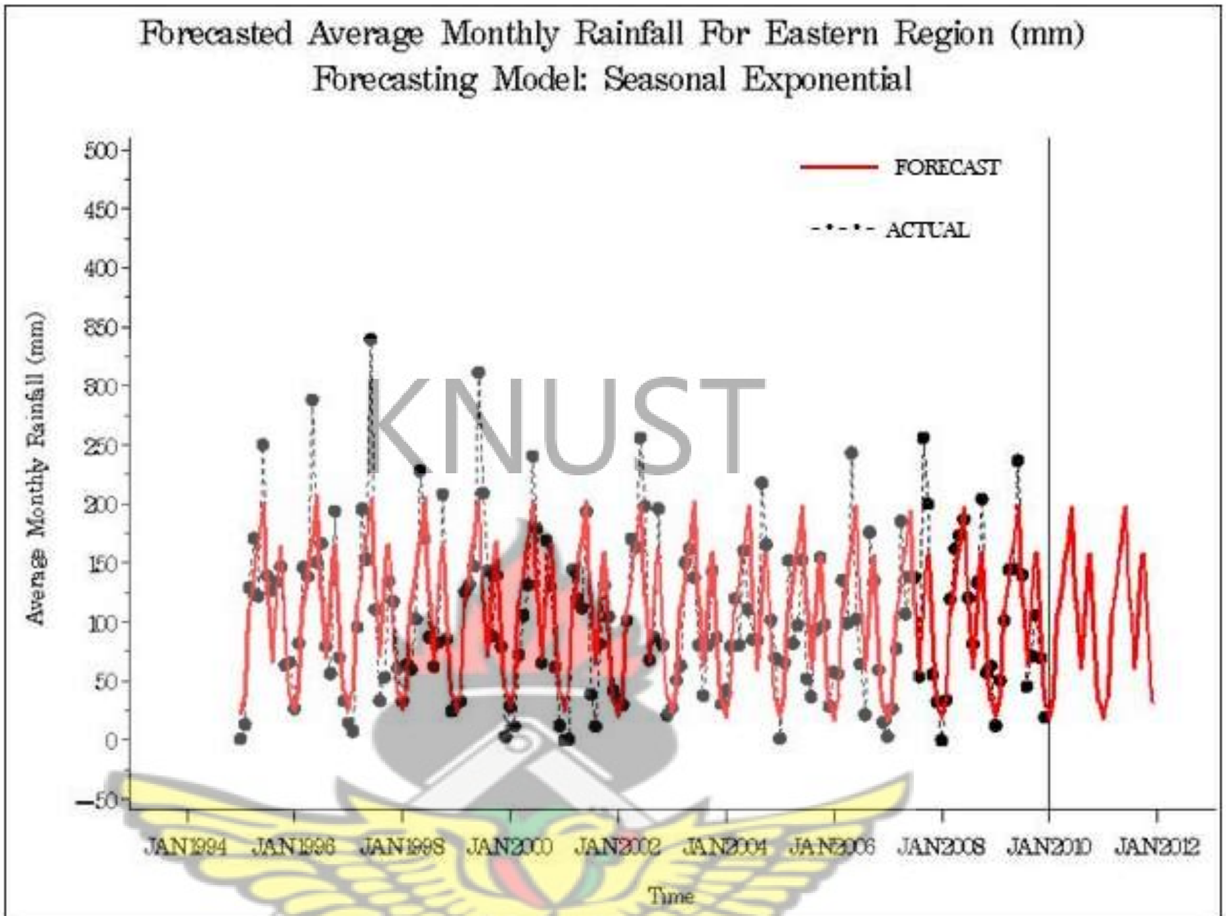

Figure D-4 Forecasting Model: Seasonal Exponential (Eastern Region)

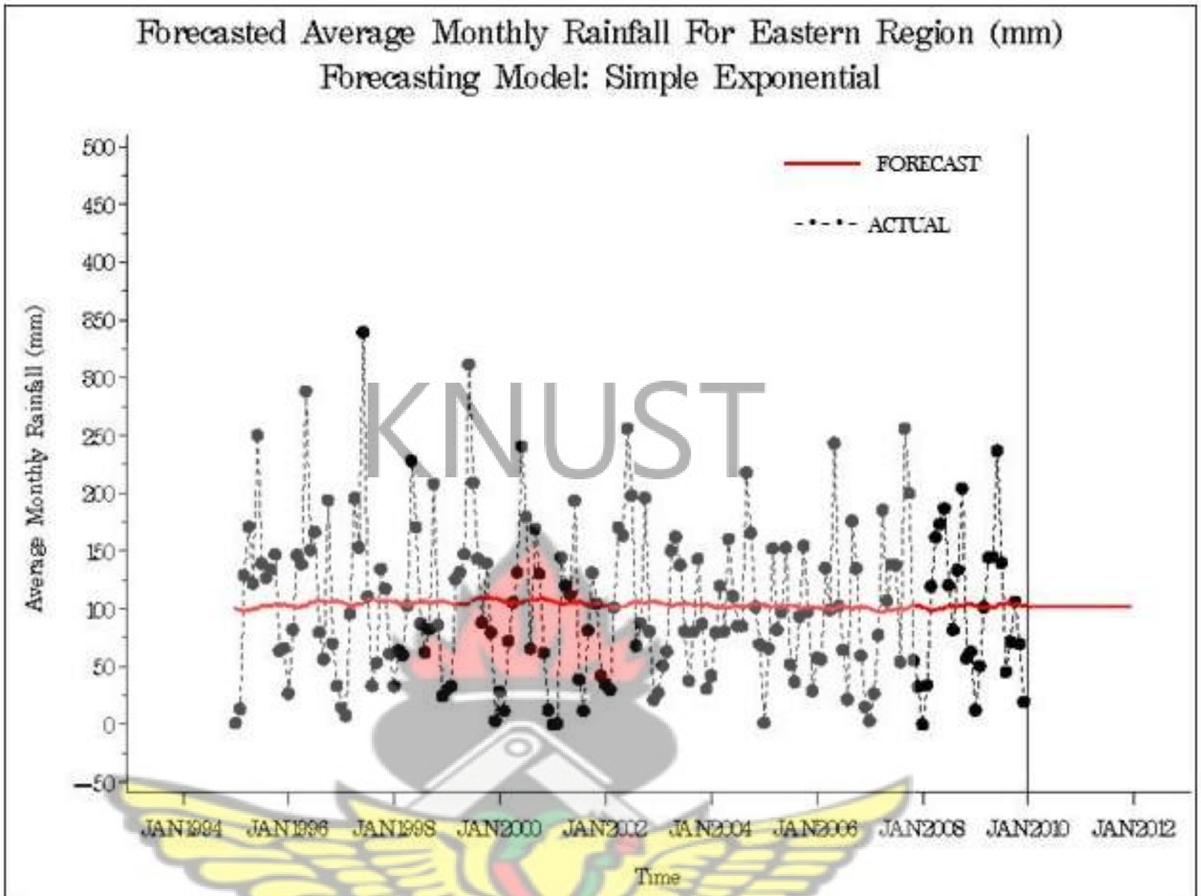

Figure D-5 Forecasting Model: Simple Exponential (Eastern Region)

APPENDIX E

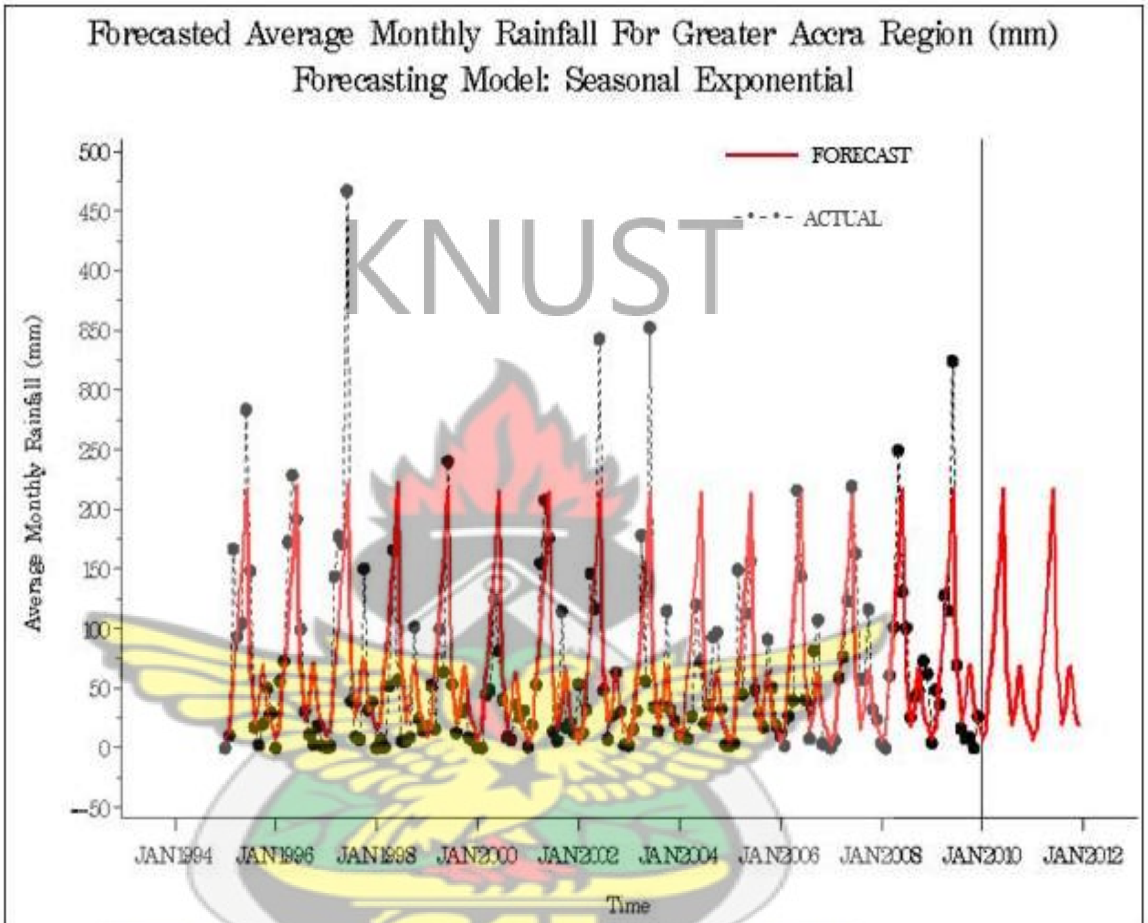

Figure E-1 Forecasting Model: Seasonal Exponential (Greater Accra Region)

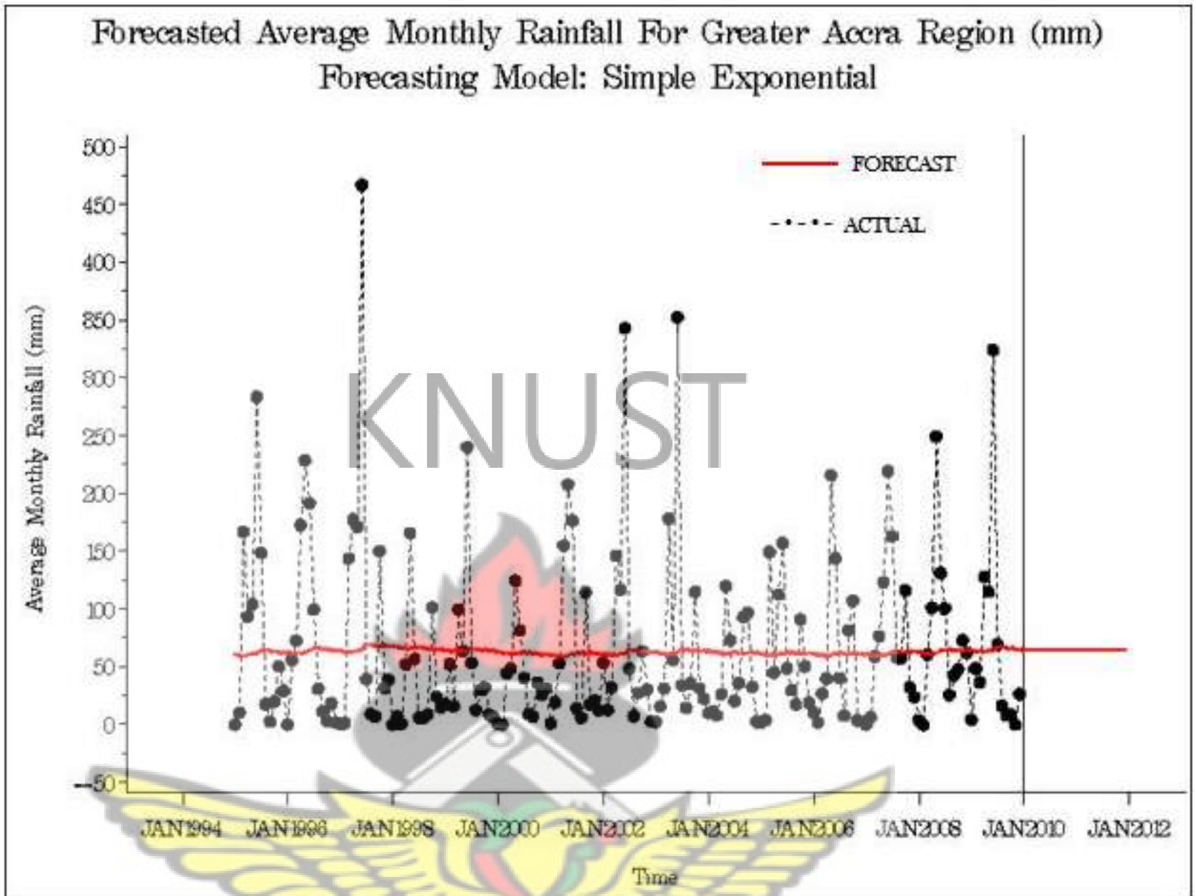

Figure E-2 Forecasting Model: Simple Exponential (Greater Accra Region)

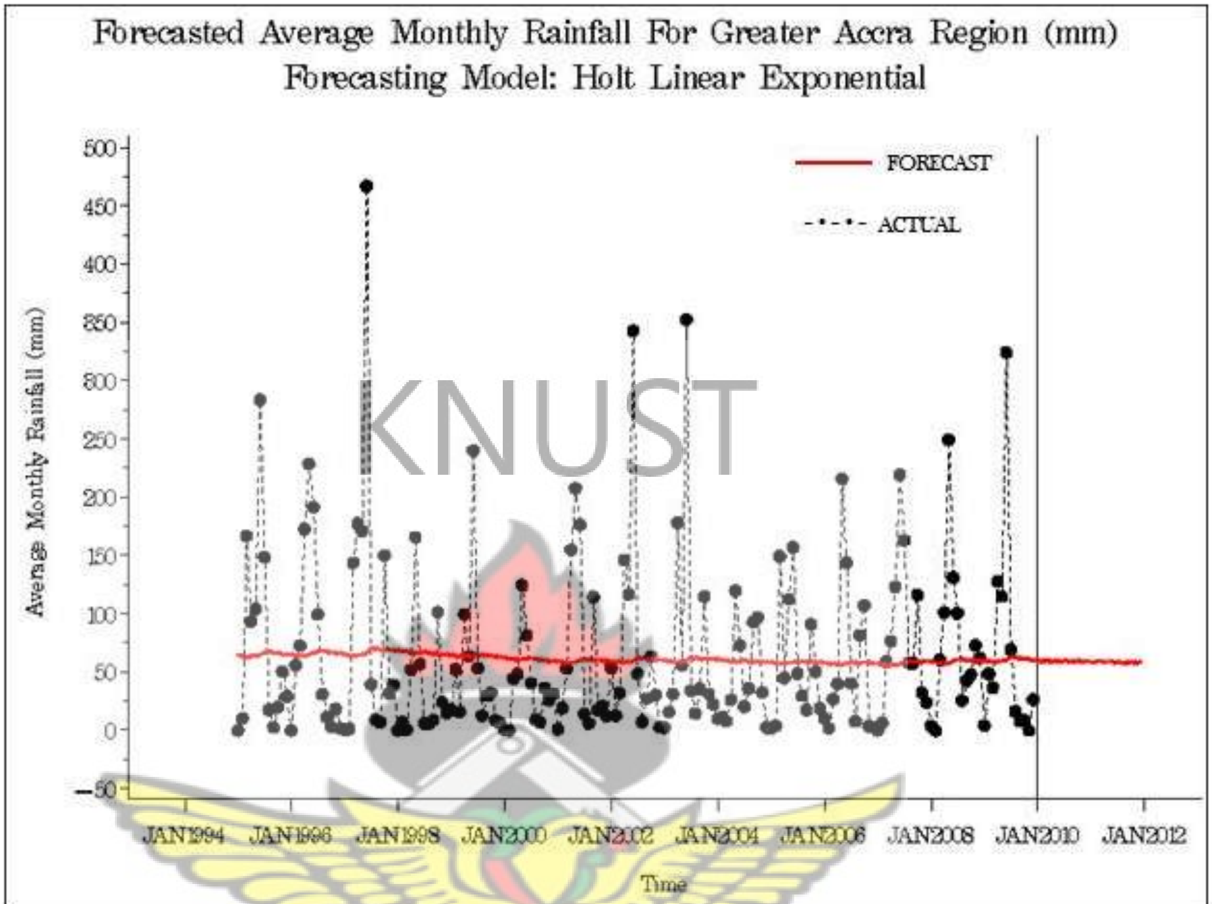

Figure E-3 Forecasting Model: Holt Linear Exponential (Greater Accra Region)

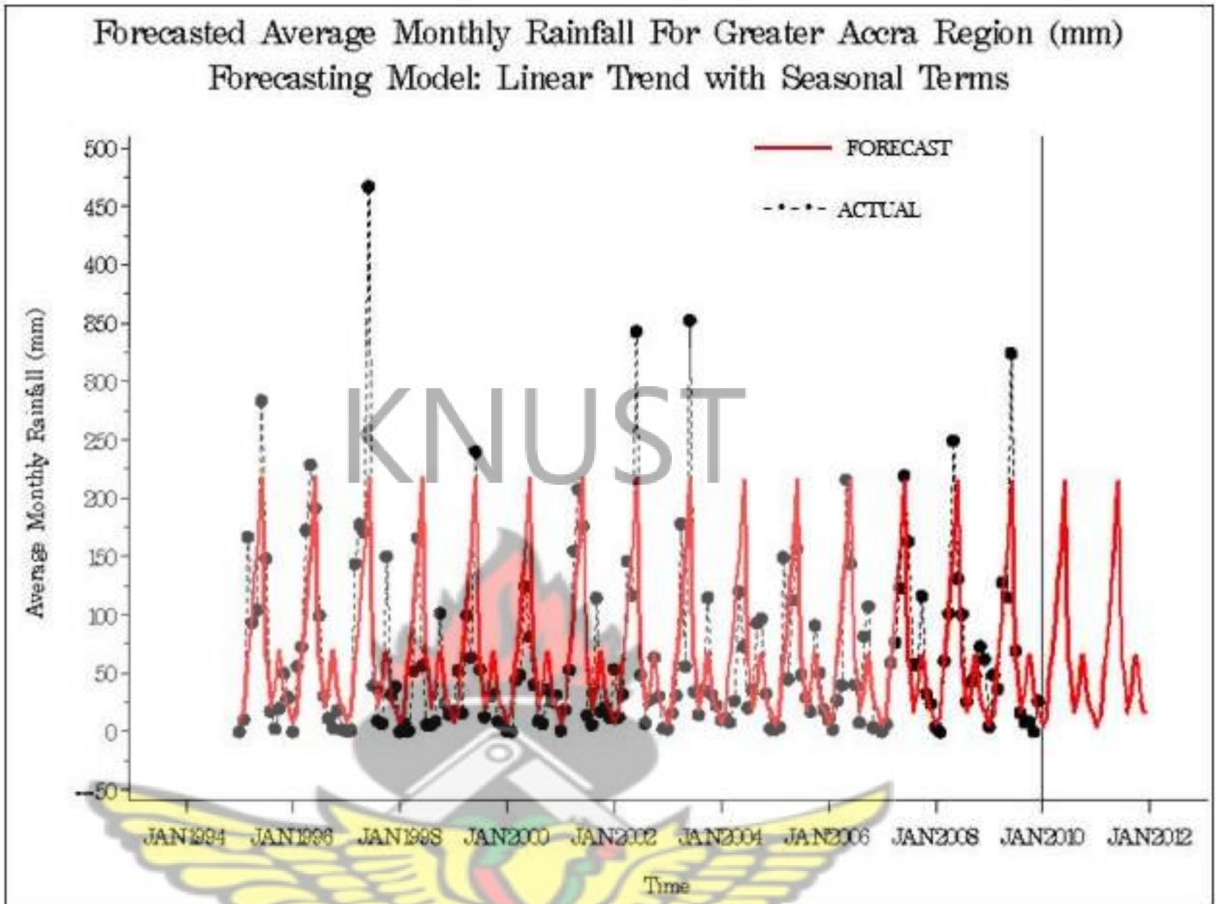

Figure E-4 Forecasting Model: Linear Trend with Seasonal Terms (Greater Accra Region)

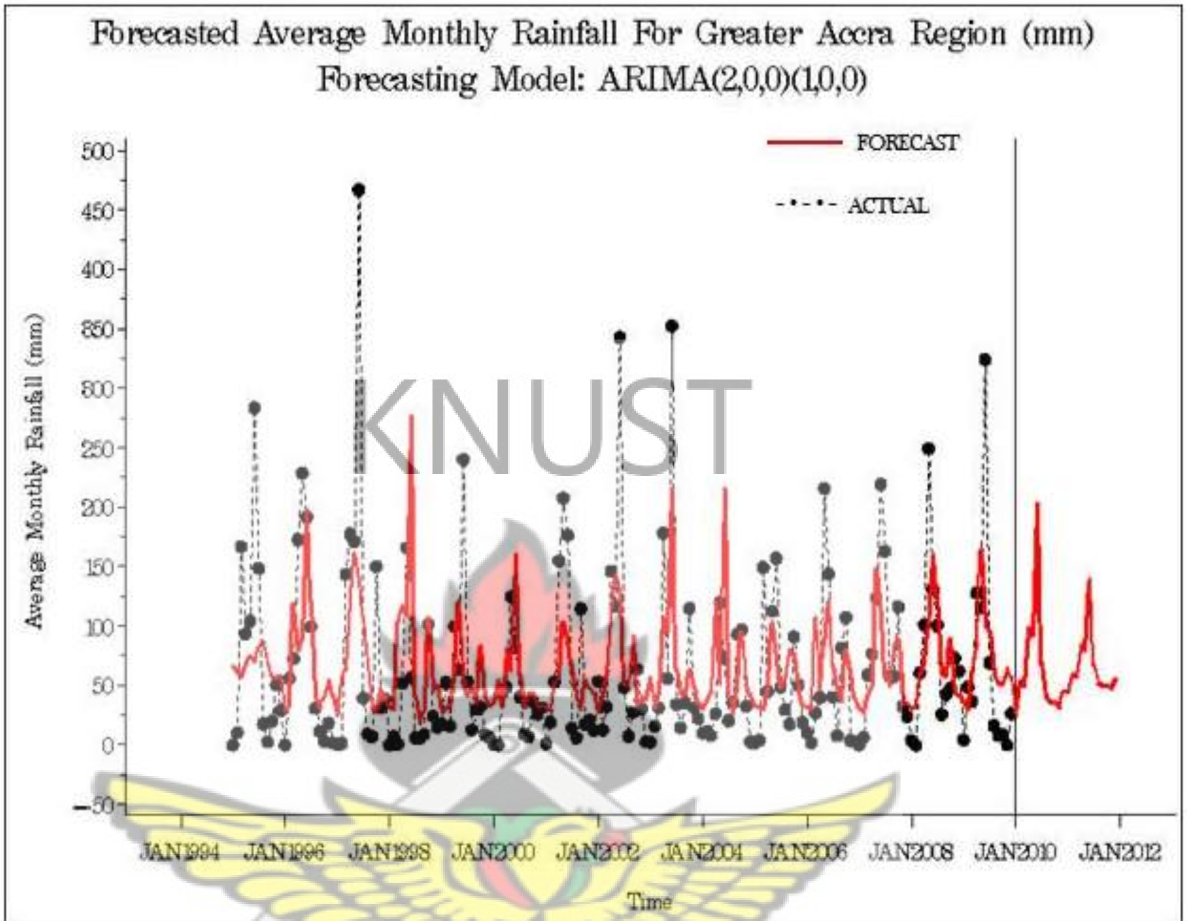

Figure E-5 Figure E-1 Forecasting Model: ARIMA (2,0,0)(1,0,0) (Greater Accra Region)