FORECASTING FOREIGN EXCHANGE RATES IN KENYA USING TIME SERIES: A CASE OF USD/KES EXCHANGE RATES

BY

KANYIRI MARY NJOKI

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN APPLIED STATISTICS

SCHOOL OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE

MASENO UNIVERSITY

Declaration

This project is my own work and has not been presented for a degree award in any other institution.

Kanyiri, Mary Njoki

MSC/MAT/00106/2016

Signature:....................................Date:................................

This project has been submitted for examination with my approval as the university supervisor.

Dr. Edgar Ouko Otumba

Department of Statistics and Actuarial Science

Maseno University

Signature:......................................Date:...............................

Acknowledgement

My deepest gratitude goes to God who provided patience, hope and kept circumstances in favor for me as I was compiling this report.

I would like to express my deep and sincere gratitude to Dr. Edgar Otumba for providing invaluable guidance throughout this project. Without his help, this research would not have been a success. For his support and insight, I say thank you.

I am also extremely grateful to my parents for their love, prayers and support in educating and preparing me for my future. Your encouragement when times got hard is much appreciated.

Dedication

First and foremost this research is dedicated to the Almighty God for giving me the physical strength and mental capability to undertake this project. I also dedicate this work to my parents whose encouragement, moral and financial support has made sure that I have finished what I started.

Abstract

In October 1993, Kenya adopted a floating exchange rate system where the exchange rates are determined by forces of demand and supply for the local currency. Exchange rate forecasts are necessary to evaluate the foreign denominated cash flows involved in international transactions. Therefore exchange rate forecasting is important to evaluate the benefits and risks attached to the international business environment. This study therefore sought to fit a Seasonal Autoregressive Intergrated Moving Average Model(SARIMA)$(p, d, q)(P, D, Q)[12]$ to United States Dollar vs Kenya Shilling exchange rate since it is the most dominant exchange rate in Kenya. The secondary monthly data from January 1993 to March 2019 from Central Bank of Kenya official website was divided into two parts namely the in-sample data and the out-sample data. The in-sample data was used to fit the model while the out-sample was used to validate the model. Seasonal Mann-Kendall test established that there was seasonal trend. A first regular difference was used to stationarize the series since the ADF test established it was not stationary. Autoregressive Intergrated Moving Average (ARIMA) and Seasonal Autoregressive Intergrated Moving Average (SARIMA) models were fitted in the data. ARIMA$(1, 1, 0)$ and SARIMA$(1, 1, 0)(0, 0, 2)[12]$ were found to be the best models on the basis of Bayesian Information Criterion(BIC) and Akaike's Information Criterion(AIC). In the short run i.e 3 months, the Seasonal Autoregressive Intergrated Moving Average had the least Mean Absolute Error(MAE), Mean Absolute Percentage Error(MAPE) and Root Mean Squared Error(RMSE) values of 0.1651, 0.1636 and 0.2037 respectively. This study therefore recommends the integration of the Seasonal Autoregressive Intergrated Moving Average Model in forecasting United States Dollar vs Kenya Shilling exchange rate in Kenya in the short run.

TABLE OF CONTENTS

DEFINITIONS

Exchange rate in this study referred to the value of a foreign currency versus the local currency i.e USD/KES.

Fixed (pegged) exchange rate is where a currency's value is fixed against either the value of another single currency, a basket of other currencies or another measure of value e.g gold.

Floating (flexible) exchange rate is a regime in which a currency's value is allowed to fluctuate in response to foreign exchange market mechanisms.

Foreign exchange (Forex) market is a market in which the participants are able to buy, sell, exchange and speculate on currencies.

Exchange rate volatility is the tendency for foreign currencies to appreciate or depreciate in value thus affecting the profitability of foreign exchange trades.

Time series forecasting is the use of a time series model to predict future values based on previously observed values.

ABBREVIATIONS AND ACRONYMS

AIC Akaike's Information Criterion

AR Autoregressive Model

ACF Autocorrelation Function

ARIMA Autoregressive Intergrated Moving Average Model

ARMA Autoregressive Moving Average Model

BIC Bayesian Information Criterion

CBK Central Bank of Kenya

KES Kenya Shilling

MA Moving Average Model

MAE Mean Absolute Error

MAPE Mean Absolute Percentage Error

PACF Partial Autocorrelation Function

RMSE Root Mean Squared Error

SARIMA Seasonal Autoregressive Intergrated Moving Average Model

USD United States dollar

USD/KES United States Dollar vs Kenya Shilling exchange rate

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

1.1 Background to the study

Exchange rate is the price of one currency in terms of another currency and so they can be analyzed by the tools of demand and supply. Different countries use different mechanisms to keep their currency stable by identifying an exchange rate regime that best suits their economy. Basically there are two types of exchange rate regimes namely fixed exchange rates and floating exchange rates.

Fixed exchange rates or pegged exchange rate system is one in which the exchange rates for a currency is fixed by the government. The local currency's value is fixed against either the value of another single currency, a basket of other currencies or another measure of value e.g gold, silver or other precious metal. To achieve stability, the government buys foreign currency when the exchange rate becomes weaker and sells foreign currency when the exchange rate gets stronger.

Floating exchange rate system is one in which exchange rate is determined by forces of demand and supply of different currencies in the foreign exchange market i.e the value of currency is allowed to fluctuate freely according to changes in demand and supply of foreign exchange [1]. The volatility in exchange rates results increases exchange rate risks and adversely affects the international trade and investment decisions [36].

The Kenya shilling was pegged to the sterling pound from 1966. The peg was changed to the US dollar from 1971 to 1974 but after discrete evaluations the peg was changed to the special drawing right (SDR) in 1975 [23]. In 1990, the dual exchange rate was adopted where the fixed exchange rate was applied on the essential segments such as imports and exports while the other sectors implemented the floating regime. This regime was used until October 1993 where after further devaluations the official exchange rate was merged with the market rate and the shilling was allowed to float[22].

According to data from CBK,in January 1999 the USD/KES currency exchange was going for 61.802. The KES currency has been depreciating over time since in December 2018 this exchange rate was going for 102.2918. The leading factors that cause exchange rates fluctuations include [18]:

- **Demand and supply**

  The demand for the local currency indicates there is demand for local goods and services. If this demand is high then the value of the local currency rises leading to low exchange rates. The demand for the foreign currency appears from the need to buy goods and services from other countries. If this demand is high then the local currency depreciates leading to a rise in exchange rates.

- **Inflation rate**

  This is the persistent rise in prices of commodities over time, resulting in a fall in the purchasing value of money. Therefore a country with a lower inflation rate compared to another exhibits a rising currency value as its purchasing power increases relative to other currencies. This lowers the value of exchange rates.

- **Interest rates**

  This is the amount charged by a lender to a borrower for the use of assets. Higher interest rates in a country offer lenders in an economy a higher return relative to other countries. This attracts foreign capital causing the value of the local currency to rise leading to low exchange rates.

- **Current account deficit**

  This is a measurement of a country's trade where the value of the goods and services it imports exceeds the value of the products it exports. This means that there is a higher demand for foreign currency. This rise in demand for foreign currency leads to the rise in the exchange rates.

- **Public debt**

  This is how much a country owes to its lenders. A country perceived to have a high national debt has high exchange rates.

- **Terms of trade**

  This is defined as the ratio of export prices to import prices i.e the amount of import goods an economy can purchase per unit of export goods. If this ratio is high it means that a country exports more than it imports. Therefore the value of the local currency appreciates leading to low exchange rates.

- **Political stability and Economic performance**

  A country with these positive attributes draws foreign investment funds since foreign investors seek out stable countries with strong economic performance. Therefore the value of the local currency rises leading to low exchange rates.

For developing countries like Kenya, exchange rates play a highly significant role in the ability of the economy to attain optimal productive capacity. Predicting exchange rates is a challenging task to both traders and practitioners in modern financial markets. Statistical and econometric models have been used in the analysis and prediction of foreign exchange rates.

Meese and Rogoff in 1983 [19] compared out-of-sample forecasts from both structural and time series models. They found that although the models fit very well in-sample, none of the models made more accurate point forecasts than a random walk, when the forecast accuracy was compared by computing the root mean squared forecast error. Since then many researchers have tried to refine models and some have found out that the ARIMA models gives comparable accurate forecasts [3].

This study therefore fitted a Seasonal Autoregressive Integrated Moving Average model to USD/KES that was used to forecast the exchange rate at least at a certain level of confidence.

1.2 Statement of the Problem

The structure of a country's exchange rate is one of the factors that affect the survival of the country in the international trade. Exchange rate forecasts play a fundamental role in nearly all aspects of international financial management. Therefore, exchange rate forecasting is very important to evaluate the benefits and risks attached to the international business environment. High exchange rates have various consequences which include high inflation rates, reduction of exports, slower growth in GDP and increased deficit in the balance of payments. A number of studies and literature that have looked at forecasting exchange rates claim that exchange rates are very difficult to forecast. This makes it difficult to estimate the future value assets and liabilities denominated in foreign currency. This creates uncertainty about the magnitude of profits to be realized from international trade.

1.3 Objective of the Study

The main aim of this study is to forecast the USD/KES exchange rates. The specific objectives were:

1. To fit a SARIMA model to USD/KES exchange rate.
2. To forecast USD/KES exchange rates.
3. To determine the forecasting performance of the SARIMA model.

1.4 Significance of the Study

The results of this study are significant in several respects: First, the findings add to the body of literature in the area under study. Second, policy makers may use these results to adequately measure economic performance. Portfolio managers and corporate finance managers whose clients' cash flows are affected by exchange rate movements would benefit from an estimate as this would provide a basis for decision making. Speculators involved in Forex trade who profit from changes in exchange rates may also benefit from these results.

1.5 Basic Concepts

1.5.1 Time series

A time series $\{X_t\}$ is a set of observations collected sequentially over time. If the observations in a time series are recorded at successive equally spaced points in time it is called a discrete-time time series [6].

Notations

- B,backshift operator, means moving the element back one time i.e $BX_t = X_{t-1}$. If we backshift p times then $B^p X_t = X_{t-p}$.
- $\nabla = 1 - B$ is the differencing operator such that $\nabla X_t = (1 - B)X_t = X_t - X_{t-1}$. Differencing d times is written as $\nabla^d X_t = (1 - B)^d X_t$
- $\nabla_s = 1 - B^s$ is the seasonal differencing operator. Takes the difference between two points in a season.

1.5.2 Stationary time series

A stationary time series is a time series whose statistical properties such as the mean, variance etc. are constant over time. A non-stationary series time series is a time series whose statistical properties are not constant i.e. they exhibit trend and seasonality.

1.5.3 ARIMA family time series models

The models below are as per [9].

Autoregressive (AR) models

These are models in which the dependent variable can be written as a weighted average of the past observations of the dependent variable.

Let p be the maximum order lag of the AR model then AR(p) is written as

$$\begin{aligned} X_t &= \delta + \omega_t + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} \\ X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} &= \delta + \omega_t \\ (1 - \phi_1 B - \ldots - \phi_p B^p) X_t &= \delta + \omega_t \\ \Phi(B) X_t &= \delta + \omega_t \end{aligned} \tag{1.1}$$

Where $\Phi(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$ is the general polynomial of an AR model.

Moving Average (MA) models

These are models in which the dependent variable can be written as a weighted average of the current and past observations which are uncorrelated mean-zero random noises. Let q be the maximum order lag of the MA model then MA (q) is written as

$$\begin{aligned} X_t &= \mu + \omega_t + \theta_1 \omega_{t-1} + \ldots + \theta_q \omega_{t-q} \\ X_t - \mu &= \omega_t + \theta_1 \omega_{t-1} + \ldots + \theta_q \omega_{t-q} \\ X_t - \mu &= (1 + \theta_1 B + \ldots + \theta_q B^q) \omega_t \\ X_t - \mu &= \Theta(B) \omega_t \end{aligned} \tag{1.2}$$

Where $\Theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \ldots + \theta_q B^q$ is the general polynomial of an MA model.

Autoregressive Intergrated Moving Average(ARIMA)models

These are models that relate the present value of a series to the past value and past prediction errors.

Let p be the AR order, q be the MA order and d be the differencing order then an ARIMA(p, d, q) is a discrete time linear equation with noise of the form

$$\Phi(B)(1 - B)^d X_t = \Theta(B) \omega_t \tag{1.3}$$

A random walk is a special ARIMA$(0, 1, 0)$ model [21]. It assumes that in each period the variable takes a random step away from its previous value, and that the steps are independently and identically distributed in size.

It is of the form

$$X_t = X_{t-1} + \omega_t \tag{1.4}$$

1.5.4 Autocorrelation Function (ACF)

The autocorrelation function (ACF) for a series gives correlations between the series X_t and lagged values of the series for lags of $1, 2, \ldots$. It is given by

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

Where ρ_k is the ACF at lag k, $\gamma_k =$Covariance(X_t, X_{t-k}) and $\gamma_0 =$Covariance$(X_t, X_t) =$Var(X_t). The estimate of the autocovariance is given by

$$c_k = \hat{\gamma_k} = \frac{1}{T} \sum_{t=1}^{T-k} (X_{t+k} - \bar{X})(X_t - \bar{X})$$

Where T is the total number of observations and $\bar{X}$ is the mean of the observations. The sample ACF is given by

$$r_k = \hat{\rho_k} = \frac{\hat{\gamma_k}}{\hat{\gamma_0}} = \frac{c_k}{c_0}$$

1.5.5 Partial Autocorrelation Function(PACF)

The partial autocorrelation between X_t and X_{t-k} is the conditional correlation between X_t and X_{t-k}, conditional on $X_{t-k+1}, \ldots, X_{t-1}$ the set of observations that come between the time points t and $t - k$. For example a second order PACF is given by

$$\frac{Cov(X_t, X_{t-2}|X_{t-1})}{\sqrt{Var(X_t|X_{t-1})Var(X_{t-2}|X_{t-1})}}$$

1.5.6 Trend detection

Mann-Kendall test

This is a non-parametric test used to identify trend in a time series. The hypothesis tested is

H_0 : There is no trend in the series

H_1 : There is a (increasing, non-null or decreasing) trend in the series

The computations assume that the observations are independent. The S statistic is computed as

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1} sign(X_j - X_i)$$

Where n is the number of observations, X_i and X_j are monthly data values in months i and j respectively since monthly data was used in this study.

$$sign(X_j - X_i) = \begin{cases} 1 \text{ if} & X_j - X_i > 0 \\ 0 \text{ if} & X_j - X_i = 0 \\ -1 \text{ if} & X_j - X_i < 0 \end{cases}$$

This test analyzes difference in signs between earlier and later data points. If a trend is present, the sign values tend to increase constantly or decrease constantly. Every value is compared to every value preceding it in the time series which gives a total of $\frac{n(n-1)}{2}$ pairs of data.

The variance of S is computed as

$$\text{Var(S)} = \frac{1}{18}\left[n(n-1)(2n+5) - \sum_{p=1}^{g} t_p(t_p-1)(2t_p+5)\right]$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group. The mann-kendall test statistic is computed as follows

$$\begin{aligned} \text{Z} &= \frac{S-1}{\sqrt{\text{Var(S)}}} \text{ if } S > 0 \\ &= 0 \text{ if } S = 0 \\ &= \frac{S+1}{\sqrt{\text{Var(S)}}} \text{ if } S < 0 \end{aligned}$$

The null hypothesis is rejected if $|\text{Z}| \geq \text{Z}_{1-\alpha/2}$. If the p-value is less than 0.05 then we reject the null hypothesis and conclude that the series has a trend.

To measure the ordinal association between data points, the Kendall's τ coefficient is calculated as follows

$$\tau = \frac{S}{D}$$

where

$$D = \left[\frac{1}{2}n(n-1) - \frac{1}{2}\sum_{j=i}^{g} t_p(t_p-1)\right]^{1/2} \left[\frac{1}{2}n(n-1)\right]^{1/2}$$

Where g is the number of tied groups and t_p is the number of observations in the p^{th} group [12].

If the τ value is positive then the series has an increasing trend. If the τ value is negative then the series has an decreasing trend.

Seasonal Mann-Kendall test

This is used to identify seasonal trend in the time series. The hypothesis tested is

$$
\begin{aligned}
H_0 &: \text{There is no seasonal trend in the series} \\
H_1 &: \text{There is a (increasing, non-null or decreasing) seasonal trend in the series}
\end{aligned}
$$

The seasonal statistic is computed by performing a mann kendall calculation for each season, then combining the results for each season. The S statistic for the g^{th} season is calculated as

$$S_g = \sum_{i=1}^{n-1} \sum_{j=i+1} sign(X_{jg} - X_{ig}), g = 1, 2, \ldots, m$$

The variance of S_g is computed as follows:

$$\text{Var}(S_g) = \frac{1}{18}\left[n_g(n_g - 1)(2n_g + 5) - \sum_{p=1}^{y_g} t_{gp}(t_{gp} - 1)(2t_g + 5)\right]$$

where y_g is the number of tied groups for the g^{th} season.

According to Hirsch *et al.*(1982), the seasonal statistic, $\hat{S}$, for the entire series is given by

$$\hat{S} = \sum_{g=1}^{m} S_g$$

with the variance

$$\text{Var}(\hat{S}) = \sum_{g=1}^{m} \text{Var}(\hat{S})$$

where m is the total number of seasons.

The seasonal mann kendall test statistic is calculated as follows:

$$
\begin{aligned}
\text{Z} &= \frac{\hat{S} - 1}{\sqrt{\text{Var}(\hat{S})}} \text{ if } \hat{S} > 0 \\
&= 0 \text{ if } \hat{S} = 0 \\
&= \frac{\hat{S} + 1}{\sqrt{\text{Var}(\hat{S})}} \text{ if } \hat{S} < 0
\end{aligned}
$$

The null hypothesis is rejected if $|\text{Z}| \geq \text{Z}_{1-\alpha/2}$.

1.5.7 Augmented Dickey Fuller (ADF) test

This is a formal test of stationarity of a time series. This test examines the null hypothesis of an ARIMA against stationary and alternatively. The hypothesis tested is

$$H_0 : \text{There is a unit root}$$

$$H_1 : \text{The time series is stationary}$$

Consider the AR(1) model

$$X_t = \phi X_{t-1} + \omega_t$$

where X_t is the value of the time series at time t, ϕ is the AR(1) parameter and ω_t is the residual at time t. If $\phi = 1$, then the AR(1) process is said to have a unit root.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

During the period of Bretton Woods system which was implemented in 1945, exchange rates were pegged to the USD as the key currency. Following the collapse of this system on March 1973, fixed exchange rates among major industrial countries were abandoned and floating exchange rate regime was adopted. Since then there has been a considerable effort in forecasting the exchange rate movements where each model possesses a private strong-point as well as a private weak-point.

2.2 Forecasting of exchange rates in other countries

Newaz(2008) made a comparison on the performance of time series models for forecasting exchange rate for the period of 1985 – 2006. He compared ARIMA model, NAÏVE 1, NAÏVE 2 and exponential smoothing techniques to see which one fits the forecasts of exchange rate and revealed that ARIMA model provides a better forecasting of exchange rate than either of the other techniques.

Shittu *et al* (2008) measured the forecast performance of ARMA and Autoregressive Fractional Integral Moving Average(ARFIMA) model on the application to US/UK pounds foreign exchange. They revealed that ARFIMA model was found to be better than ARMA model. Their result revealed that ARFIMA model is more realistic and closely reflects the current economic reality in the two countries which were indicated by their forecasting evaluation tool.

Adetunde *et al* (2011) used time series analysis to obtain a forecast model for exchange rates of the Ghanaian cedi to the US dollar, which they then used to produce a forecast plot for 2011 and 2012, using the exchange rate data from 1994 to 2010. ARIMA(1,1,1) was found to be the best model and their findings revealed that predicted rates were consistent with the depreciating trend of the observed series.

Onasanya *et al* (2013) used Box Jenkins approach to forecast the naira/dollar exchange rate in Naigeria for the period January 1994 to December 2011 using ARIMA. The result revealed that there is an upward trend and basing on the selection criteria AIC and BIC, the best model that explains the series was found to be ARIMA (1, 2, 1) model. A forecast for period of 12 months terms was made which indicated that the Naira will continue to depreciate within the forecasted time period.

Yao *et al* (2015) conducted a research on forecasting the exchange rate of the Ghanaian cedi to the US dollar using seasonal ARIMA and the Random Walk models. The researchers found modest differences between these two models based on the out-of-sample forecast. However, both models performed similarly based on forecast values. Forecast values showed that the exchange rate of the Ghana cedi to the American dollar would increase continuously in the next three (3) years.

Tran *et al* (2016) developed an ARIMA model to forecast the foreign exchange rate VND/USD using real foreign exchange rate data from 2013 to 2015. The model was used to forecast the exchange rate between VND/USD for the year 2016. The results showed that ARIMA model was suitable for estimating foreign exchange rate in Vietnam in the short-time period.

2.3 Modeling exchange rates volatility in Kenya

Omar *et al* (2013) modeled monthly exchange rate in Kenya US, UK, Euro and Japanese Yen data using Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models for the period January 2001 to December 2010. Their results showed uncertainty of exchange rates between 2001 to 2005, gaining relative stability upto 2008 where the shilling became weaker against the foreign currencies due to post election violence. The results also showed that the Kenyan economy undergoes a cycle of about five years approximately i.e from one election period to the next. GARCH (1,1) was found to be the best model.

Kipkoech (2014) modeled USD/KES,EUR/KES and GBP/KES exchange rate volatility under normal and student-t distributional assumptions. He used Exponential Generalized Autoregressive Conditional Heteroscedasticity (EGARCH) model and noted that the model is more favorable compared to the normal distribution because of evidence of the heavy tailed nature of financial time series. He also claimed that the normal GARCH model could neither explain the entire fat tail nature of the data nor could it explain the asymmetric responses. The GBP/KES and the EUR/KES were both fitted by an EGARCH(1,1) model while USD/KES was fitted by an AR(1)/EGARCH(1,1).

Sylvia (2014) modeled KES/USD exchange rate volatility using GARCH family models where symmetric and asymmetric models were used to capture volatility characteristics of exchange rates on data from the period October 1993 to March 2014. A comparison was made under four different conditional distribution assumptions namely Normal, Skewed Normal, Student-t and Skewed Student-t distributions. The findings showed that the asymmetric EGARCH model with skewed student-t distribution was the best model based on AIC and log-likelihood.

Omari *et al* (2017) modeled USD/KES Exchange Rate Volatility using GARCH Models using daily observations over the period starting January 2003 to December 2015. The performance of the symmetric GARCH (1, 1) and GARCH-M models as well as the asymmetric EGARCH (1, 1), Glosten-Jagannathan-Runkle Generalized Auto Regressive Conditional Heteroscedasticity (GJR-GARCH) (1, 1) and Asymmetric Power Autoregressive Conditional Heteroscedasticity (APARCH) (1, 1) models with different residual distributions were applied to data. The most adequate models were found out to be the asymmetric APARCH model, GJR-GARCH model and EGARCH model with Student's t-distribution.

2.4 Summary

The Vector Auotoregressive Models(VAR) were for a while the most dominant models in forecasting exchange rates. However they proved to be complex and had varying predictive performance. Models based on random walk later became popular. However they produced poor short run forecasts and for some countries the hypothesis that exchange rates were random could be rejected. Based on the above, some researchers have found out that the ARIMA models gives comparable accurate forecasts.

According to Sarno and Taylor (2002), models perform differently with different currencies i.e some models produce good out-of sample forecasts but when applied to different currencies the satisfactory results can not be replicated. In Kenya, most studies have concentrated on modeling the volatility of exchange rates. Little effort has been done in forecasting exchange rates. This will fit a SARIMA model to the USD/KES exchange rate and forecast the exchange rate to a certain level of confidence since ARIMA family models give comparable accurate results.

CHAPTER THREE

RESEARCH METHODOLOGY

3.1 Introduction

The main focus will be on the exchange rate USD/KES since it is the most dominant foreign currency in Kenya according to CBK currency rankings. Secondary data of the monthly exchange rate(period average) from Central Bank of Kenya (CBK) will be used.

3.2 Seasonal Autoregressive Intergrated Moving Average(SARIMA) model

This model incorporates both non-seasonal and seasonal factors. Seasonality in a time series is a regular pattern of changes that repeats over S time periods, S defines the number of time periods until the pattern repeats again.
This model is denoted ARIMA$(p, d, q) \times (P, D, Q)S$ where:[4]

- p is the non-seasonal AR order.
- q is the non-seasonal MA order.
- d is the non-seasonal differencing order.
- P is the seasonal AR order.
- Q is the seasonal MA order.
- D is the seasonal differencing order.

This model is of the form[4]

$$\Phi_P(B^S)\phi_p(B)\nabla_S^D\nabla^d X_t = \Theta_Q(B^S)\theta_q(B)\omega_t \tag{3.1}$$

where

- $\Phi_P(B^S) = (1 - \Phi_1 B^S - \ldots - \Phi_P B^{SP})$ is the seasonal AR polynomial operator of order P.

- $\phi_p = (1 - \phi_1 B - \ldots - \phi_p B^p)$ is the regular AR polynomial operator of order p.
- $\nabla_S^D = (1 - B^S)^D$ represents the seasonal differences.
- $\nabla^d = (1 - B)^d$ represents the regular differences.
- $\Theta_Q(B^S) = (1 + \Theta_1 B^S + \ldots + \Theta_Q B^{SQ})$ is the seasonal MA polynomial operator of order Q.
- $\theta_q(B) = (1 + \theta_1 B + \ldots + \theta_q B^q)$ is the regular MA polynomial operator of order q.

3.3 Stationarity Analysis

When an AR(p) is represented as $\omega_t = \Phi(B)$, then $\Phi(B) = 0$ is known as the characteristic equation of the process. For the AR(p) process to be stationary, all the roots of the characteristic equation must fall outside the unit circle. For example from Equation 1.1, the AR(1) process written as $X_t = \delta + \phi_1 X_{t-1} + \omega_t$ is stationary when $|\phi_1| < 1$ with a constant:

- mean calculated as

$$\begin{aligned} E(X_t) &= E(\delta + \phi_1 X_{t-1} + \omega_t) \\ \mu &= \delta + \phi\mu \\ \mu(1-\phi) &= \delta \\ \mu &= \frac{\delta}{1-\phi} \end{aligned}$$

- variance calculated as

$$\begin{aligned} Var(X_t) &= Var(\delta + \phi_1 X_{t-1} + \omega_t) \\ \gamma_0 &= \phi^2\gamma_0 + \sigma^2 \\ \gamma_0(1-\phi^2) &= \sigma^2 \\ \gamma_0 &= \frac{\sigma^2}{1-\phi^2} \end{aligned}$$

Since $Var(X_t)$ is always greater than zero, it follows that $(1 - \phi^2) > 0$ implying that $|\phi_1| < 1$ if the series is stationary.

An MA(q) process is always stationary irrespective of the values of the MA parameters. An ARMA(p, q) is stationary if all the roots of the characteristic equation $\Phi(B) = 0$ lie outside the unit circle.

3.4 Fitting a SARIMA model

The steps used to fit this model are as follows:[32]

The first stage in fitting a SARIMA model is model identification. The objective of this stage is determining the values of p, d, q, P, D and Q in ARIMA$(p, d, q) \times (P, D, Q)S$. The following are the steps involved in this stage.
Step 1:Plot a time series plot of the data and examine it for trend and seasonality.
Step 2: Do any necessary differencing.

- For a data with seasonality and no trend, take a difference of lag S. Seasonality appears in the ACF by tapering slowly at multiples of S.
- For a data with linear trend and no obvious seasonality, take a first difference.
- For a data with both trend and seasonality, apply a seasonal difference to the data and then re-evaluate the trend. If a trend remains, take a first difference.
- For data with neither obvious trend nor seasonality, don't take any differences.

Step 3:Look at the ACF and PACF of the differenced data.

- *Non-seasonal terms* :Look at the early lags to determine non-seasonal terms. Spikes in the ACF at these low lags indicate possible non-seasonal MA terms. Spikes in the PACF indicate non-seasonal AR terms.
- *Seasonal terms*: Look at the patterns at lags that are multiples of S. Spikes in the ACF at lags that are multiples of S indicate seasonal MA terms. Spikes in the PACF lags that are multiples of S indicate possible seasonal AR terms.

The next stage after model identification is parameter estimation. After making a guess or two of a possible model, the next stage is to estimate the coefficients. An optimization criterion like maximum likelihood is used. Given a sample $x_1, x_2, \ldots, x_n$ of n independent and identically distributed set of observations from a distribution $f(x)$ with an unknown parameter θ, then the joint density function is

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1, \theta) f(x_2, \theta) \ldots f(x_n, \theta)$$

The likelihood function is given by

$$L(\theta; x_1, x_2, \ldots, x_n) = f(x_1, x_2, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

In practice, the log-likelihood is more convenient to use:

$$Ln(\theta; x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} Lnf(x_i, \theta)$$

After the coefficients are estimated, look at their significance. For each coefficient calculate

$$Z_{\text{calc}} = \frac{\text{estimated coefficient}}{\text{standard error of the coefficient}}$$

At $\alpha = 0.05$ if $|Z_{\text{calc}}| > 1.96$, the estimated coefficient is statistically significant.

The next step is diagnostic checking. This is done by analyzing residuals to determine the goodness of fit to the data. The following procedures are done in this stage:

- Examine the ACF of the residuals. For a good model, the ACF of the residuals should be non-significant.

- Plot a normal Q-Q plot to check whether the residuals are normally distributed.

- Plot the p-values of the Ljung-Box statistic. This statistic is a function of accumulated sample autocovariances,r_j, upto a specified time lag m. It is determined as

  $$Q(m) = n(n+2) \sum_{j=1}^{m} \frac{r_j^2}{n-j}$$

  where n is the number of usable data points after any differencing. $Q(m)$ has a chi-square distribution with m degrees of freedom. The Ljung-Box tests the null hypothesis that the residuals are random i.e they are independently distributed against the alternative hypothesis that the residuals are not random. If the p-values are greater than than 0.05 then the model is a good fit.

Compare Akaike's Information Criterion (AIC) or Bayesian Information Criterion (BIC) values if you tried several models.

$$\text{AIC} = -2\log(L) + 2(p+q+k+1)$$

and

$$\text{BIC} = \text{AIC} + log(T)(p+q+k-1)$$

where L is the likelihood of the data, T is the total number of observations, $k = 1$ when $p, q = 0$ and $k = 1$ when $p, q \neq 0$. The model with smallest value is the best model. In the two equations, the first part i.e $-2\log(L)$ is a measure of "lack of fit" while the remainder part is for penalizing the number of estimated parameters. It is clear that BIC induces a higher penalty.

The final step is forecasting using the selected model. Forecasting is done to predict future values of a time series based on data collected to the present. A forecast error is the difference between an observed value and its forecast. The forecasting procedure assuming a sample size n is as follows;

- for any ω_j with $1 \leq j \leq n$, use the sample residue for time point j.
- for any ω_j with $j > n$, use 0 as the value of ω_j.
- for any X_j with $1 \leq j \leq n$, use the observed value of X_j.
- for any X_j with $j > n$, use the forecasted value X_j.

The accuracy of forecast is the degree of closeness of the statement of quantity to that quantity's actual value. Mean absolute error (MAE), Mean absolute percentage error (MAPE) and Root mean square error (RMSE) will be used to determine the effectiveness of the model. Accuracy of the forecasts can be measured using the following measures of accuracy.

- *Mean Absolute Error(MAE)*: This is the average vertical or horizontal distance between each point and the identity line. It measures the average absolute deviation of forecasted values from original ones i.e it shows the magnitude of overall error, occurred due to forecasting. For a good forecast, the MAE should be as small as possible. It is calculated as

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|X_t - \hat{X}_t|$$

- *Mean Absolute Percentage Error(MAPE)*: This is the mean or average of the absolute percentage errors of forecasts. The smaller the value, the better the forecast. It is calculated as

$$\text{MAPE} = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{X_t - \hat{X}_t}{X_t}\right|$$

- *Root Mean Square Error (RMSE)*: This is the standard deviation of the forecast errors. the smaller the RMSE the better the model. It is calculated as

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(X_t - \hat{X}_t\right)^2}{n}}$$

where X_t is the actual value at time t, $\hat{X}_t$ is the forecasted value at time t and n is the number of observations.

This chapter has provided an outline and description of the research methodology used in this project. The next task was to fit a SARIMA model using the above procedure and determining the forecasting performance.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

In this chapter the performance of the USD/KES exchange rate is analyzed using monthly data available in the CBK official website from January 1993 to March 2019. The data is divided into two parts, the training set from January 1993 to March 2016 and the test set from April 2016. The training set is used to identify and fit the models while the test set is used to test the forecast accuracy.

Table 4.1: Descriptive Statistics of USD/KES Exchange Rates

Mean	76.888198
SE Mean	0.853164
Median	77.262
Skewness	-0.033745
Kurtosis	-0.378534
Minimum	36.23
Maximum	105.275
Range	69.045
Count	315
Sum	24219.782493
Standard deviation	15.142163

The descriptive statistics presented in Table 4.1 revealed that the average rate of USD/KES exchange rate is 76.888198 with a standard deviation of 15.142163. The largest and smallest exchange rates within the selected period was 105.275 and 36.23 respectively, thus the range was 69.045. The data was negatively and highly skewed since it had a skewness value of -0.033745. The data had a kurtosis value of -0.378534 thus it had light tails.

From the exploratory plots in Figure 4.1 and Figure 4.2, we noted that the USD/KES exchange rate indicated an overall upward trend i.e USD/KES exchange rates increased over time. This was also evidenced by the results of the mann-kendall test in Figure 4.3. The p-value was less than 0.05 while the value of τ was positive indicating an increasing trend. The boxplot in Figure 4.4 showed that the exchange rates in the months of June, July and August were slightly higher than the rest. There was no obvious seasonality since the months seemed to have almost the same mean and variance. A correlogram was plotted to determine stationarity.

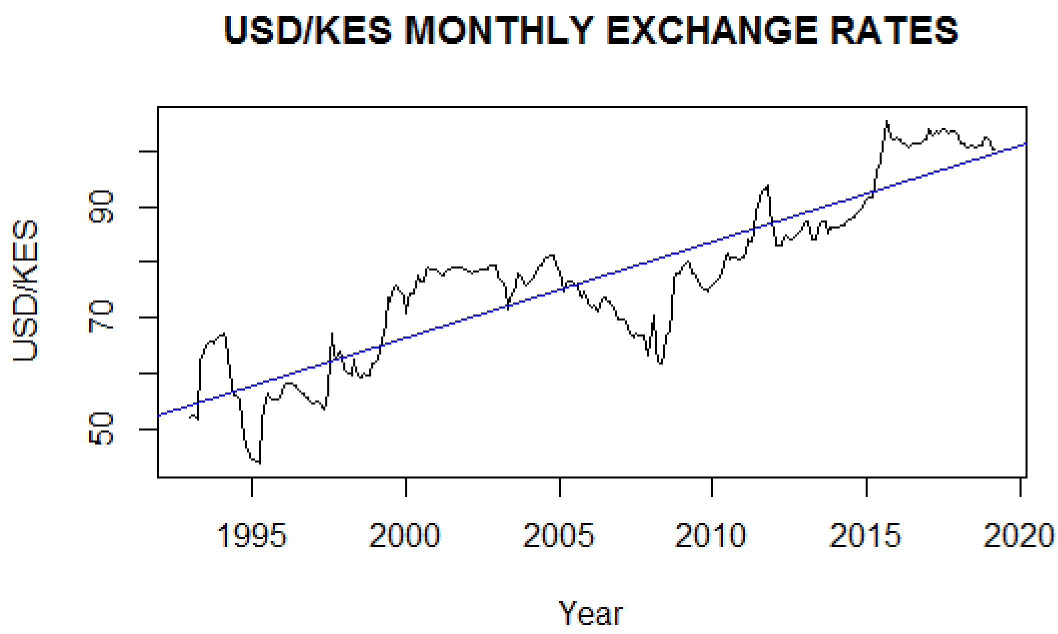

Figure 4.1: Time series plot of the USD/KES exchange rate

ACF plot in Figure 4.5 clearly showed that the autocorrelations were large at many lags indicating lack of stationarity. Using the Augmented Dickey Fuller (ADF) test to test for stationarity the result in Figure 4.6 was obtained. Since the p-value 0.2591 was greater than 0.05 we did not reject the null hypothesis and concluded that the series was not stationary.

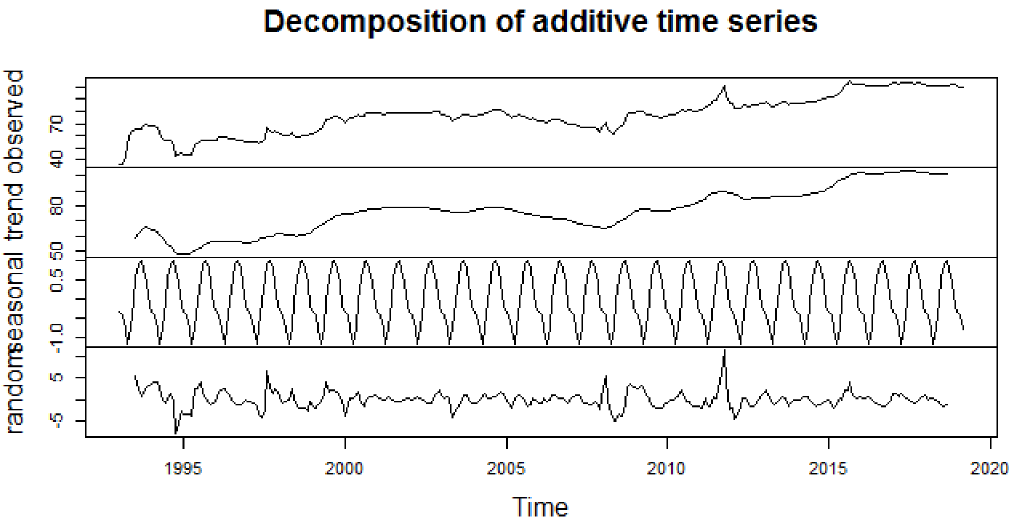

Figure 4.2: Decomposition plot

```
tau = 0.69, 2-sided pvalue =< 2.22e-16
```

Figure 4.3: Mann-Kendall test

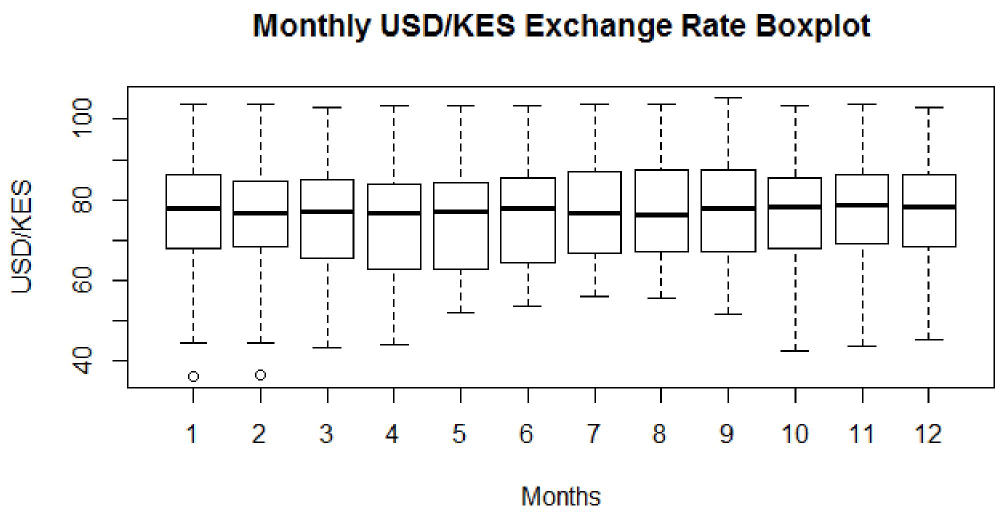

Figure 4.4: Boxplot

To achieve stationarity differencing method was used. In this case regular differencing was used since there was no obvious seasonality in the series. The ADF test of the differenced series yielded the result in Figure 4.7. Since the p-value 0.01 was less than 0.05 we rejected

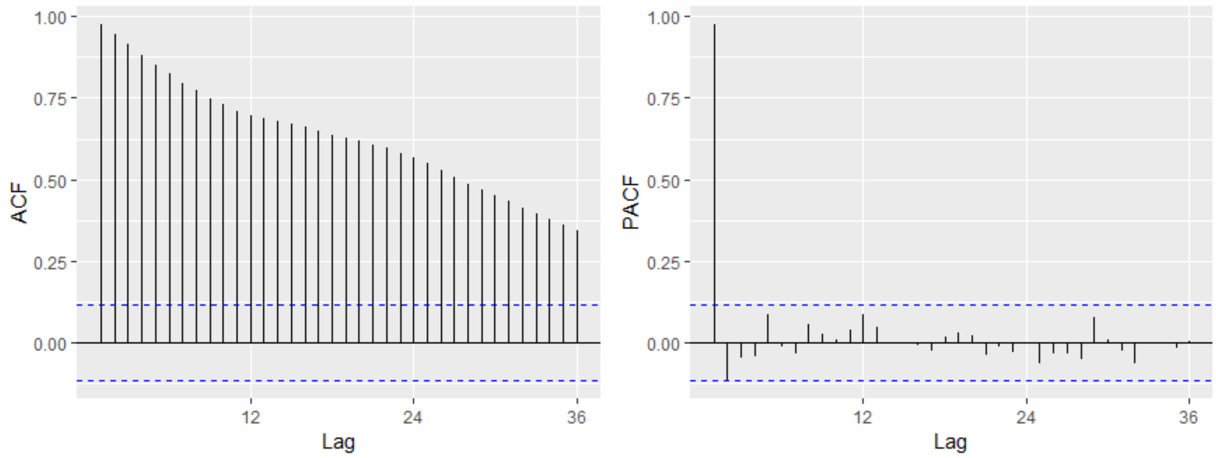

Figure 4.5: Correlogram plots

the null hypothesis and concluded that the differenced series was stationary.

```
	Augmented Dickey-Fuller Test

data:  cl.train
Dickey-Fuller = -2.7524, Lag order = 6, p-value = 0.2591
alternative hypothesis: stationary
```

Figure 4.6: Augmented-Dickey Fuller test of the series

```
	Augmented Dickey-Fuller Test

data:  cl.diff1
Dickey-Fuller = -6.7167, Lag order = 6, p-value = 0.01
alternative hypothesis: stationary
```

Figure 4.7: Augmented-Dickey Fuller test of the differenced series

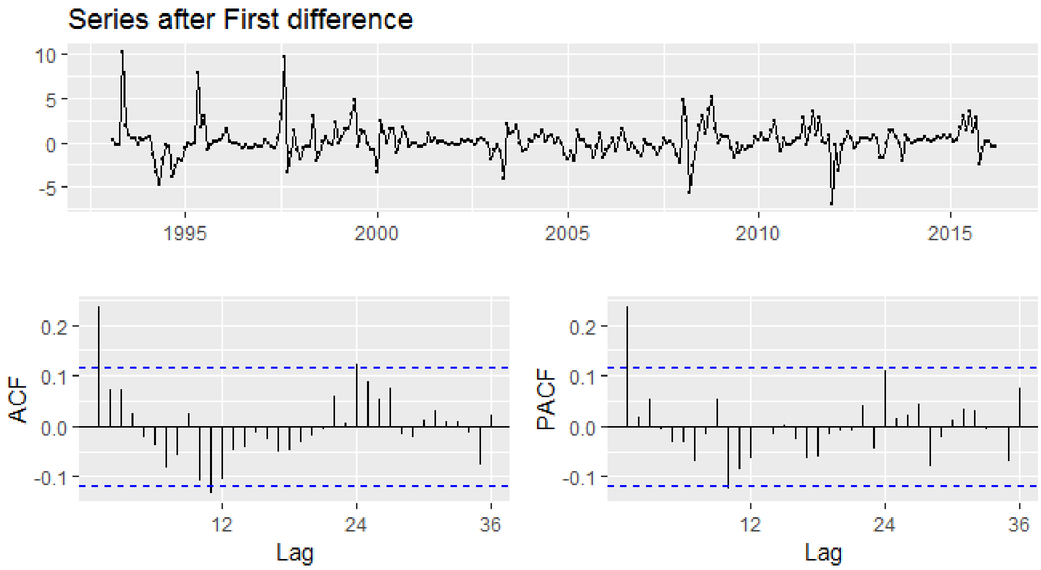

Figure 4.8: Differenced series and Correlogram plots

Table 4.2: Behavior of ACF and PACF

Model	ACF	PACF
AR(p)	Spikes decay toward zero.	Spikes decay toward zero after lag p Coefficients may oscillate.
MA(q)	Spikes decay toward zero after lag q Coefficients may oscillate.	Spikes decay toward zero.
ARMA(p,q)	Spikes decay to zero after lag q	Spikes decay to zero after lag q.

4.2 Autoregressive (ARIMA) Model

4.2.1 Statistics and Data Analysis

From Figure 4.8 the ACF and PACF had significant values at lags 1. From Table 4.2, the AR terms are depicted by the PACF while the MA terms are depicted by the ACF i.e The ACF cuts-off after lag q, MA terms, while the PACF cuts-off after lag p, AR terms. The possible ARMA models were ARMA$(1,0)$, ARMA$(0,1)$ and ARMA$(1,1)$. Including the differencing term, possible ARIMA models were ARIMA$(1,1,0)$, ARIMA$(0,1,1)$ and

ARIMA$(1, 1, 1)$. Fitting these models yielded the results in Figure 4.9.

```
Series: cl.train
ARIMA(1,1,0)

Coefficients:
         ar1
      0.2446
s.e.  0.0580

sigma^2 estimated as 2.917:  log likelihood=-542.82
AIC=1089.65   AICc=1089.69   BIC=1096.9
Series: cl.train
ARIMA(0,1,1)

Coefficients:
         ma1
      0.2307
s.e.  0.0561

sigma^2 estimated as 2.931:  log likelihood=-543.44
AIC=1090.89   AICc=1090.93   BIC=1098.14
Series: cl.train
ARIMA(1,1,1)

Coefficients:
         ar1      ma1
      0.3966  -0.1632
s.e.  0.2660   0.2892

sigma^2 estimated as 2.925:  log likelihood=-542.69
AIC=1091.38   AICc=1091.46   BIC=1102.26
```

Figure 4.9: Fitted ARIMA models

On the basis of AIC and BIC, ARIMA$(1, 1, 0)$ was the best model since it had less values. From Equation 1.3, $\Phi(B) = 1 - \phi_1 B$ and $\Theta(B) = 1$ since the model consisted of AR and differencing terms only. Therefore the equation was

$$\begin{aligned} (1 - \phi_1 B)(1 - B)X_t &= \omega_t \\ (1 - \phi_1 B)(X_t - X_{t-1}) &= \omega_t \\ X_t - X_{t-1} - \phi_1 X_{t-1} + \phi_1 X_{t-2} &= \omega_t \end{aligned}$$

$$X_t = X_{t-1} + \phi_1 X_{t-1} - \phi_1 X_{t-2} + \omega_t \tag{4.1}$$

From Figure 4.9 $\phi_1 = 0.2446$. Testing the significance of the coefficient we used the test statistic

$$Z_{\text{calc}} = \frac{\text{estimated coefficient}}{\text{standard error of the coefficient}} = \frac{0.2446}{0.0580} = 4.2172$$

At $\alpha = 0.05$, $|Z_{\text{calc}}| > 1.96$ therefore we concluded that the coefficient was significant. Thus Equation 4.1 was

$$X_t = X_{t-1} + 0.2446X_{t-1} - 0.2446X_{t-2} + \omega_t \tag{4.2}$$

4.2.2 Residual Analysis

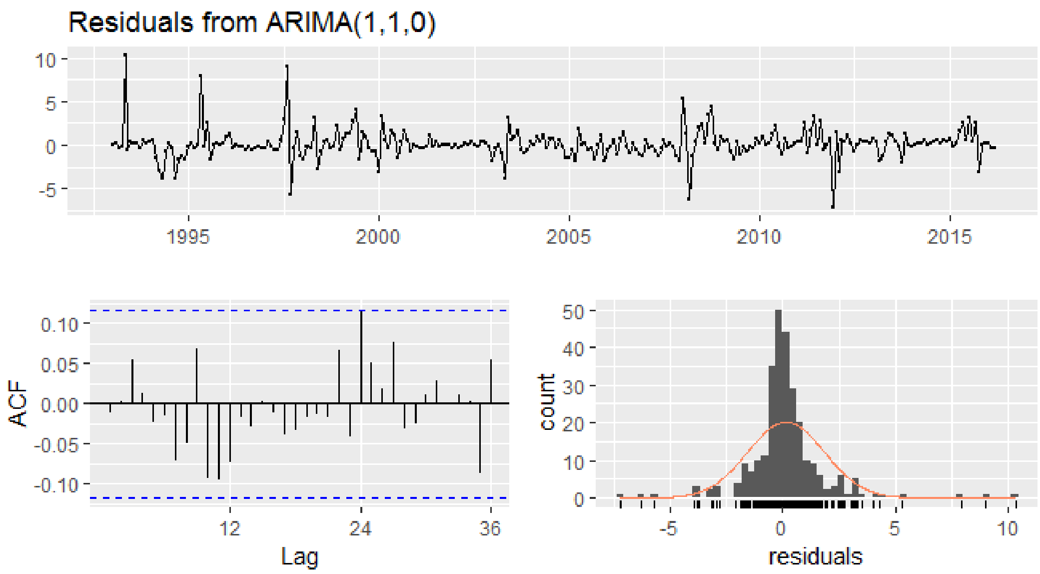

Figure 4.10: Residuals from ARIMA $(1, 1, 0)$

Plotting the ACF of the residuals in Figure 4.10 from ARIMA $(1, 1, 0)$ showed that they were uncorrelated since none of them were significant. The histogram of the residuals showed that they were normally distributed. This was also verified by a normal Q-Q plot in Figure 4.11. From Ljung Box test in Figure 4.12 it was clear that the residuals were independently distributed since all the p-values were greater than 5% significance level.

This showed that ARIMA $(1, 1, 0)$ is a good model.

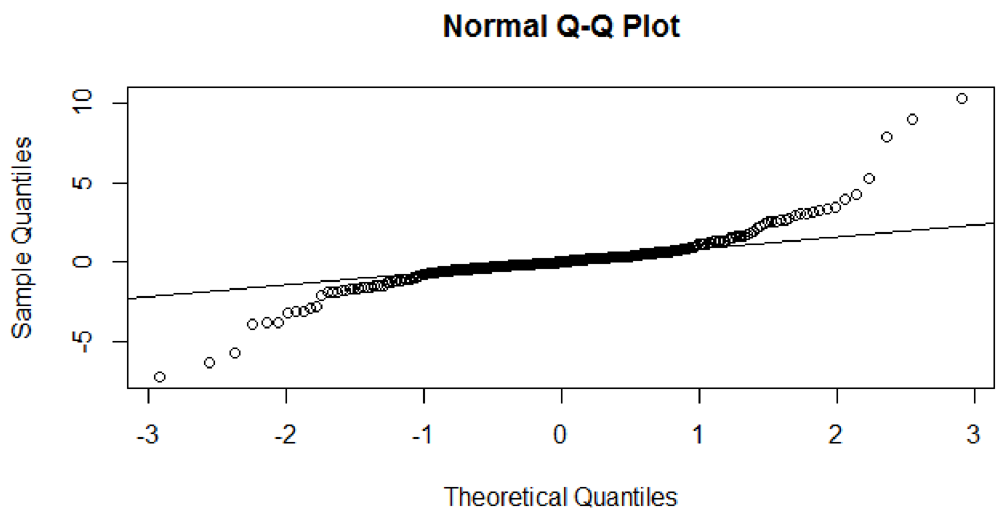

Figure 4.11: Normal Q-Q plot of residuals from ARIMA $(1, 1, 0)$

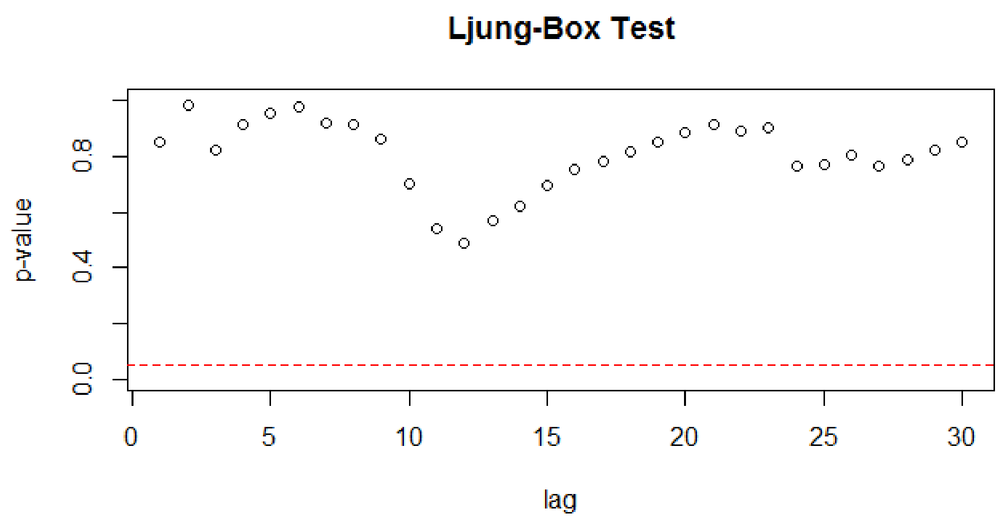

Figure 4.12: Ljung-Box test of residuals from ARIMA $(1, 1, 0)$

4.2.3 Forecasting using ARIMA

Using Equation 4.2

- One-step forecast

$$\begin{aligned}
\hat{X_{t+1}} &= X_t + 0.2446X_t - 0.2446X_{t-1} + \omega_{t+1} \\
\hat{X_{280}} &= X_{279} + 0.2446X_{279} - 0.2446X_{278} + \omega_{280} \\
&= 101.485 + 0.2446(101.485) - 0.2446(101.932) + 0 \\
&= 101.3757
\end{aligned}$$

Where $\hat{X_{280}}$ is the forecasted value of April 2016, $X_{t-1} = X_{278} = X_{278} - \hat{X_{278}}$ and $\omega_{t+1} = \omega_{280} = 0$ since the expected value of future residuals is not known.

- Two-step forecast

$$\begin{aligned}
\hat{X_{t+2}} &= \hat{X_{t+1}} + 0.2446\hat{X_{t+1}} - 0.2446X_t + \omega_{t+2} \\
\hat{X_{281}} &= \hat{X_{280}} + 0.2446\hat{X_{280}} - 0.2446X_{279} + \omega_{281} \\
&= 101.3757 + 0.2446(101.3757) - 0.2446(101.485) + 0 \\
&= 101.3490
\end{aligned}$$

Where $\hat{X_{281}}$ is the forecasted value of May 2016, and $\omega_{t+2} = \omega_{281} = 0$.

The following forecast plot in Figure 4.13 for the next three years was obtained. The model produced more accurate results for the first few months. The long term forecasts eventually went to a straight line. The MAE, MAPE and RMSE for the next three months was as follows:

Table 4.3: ARIMA$(1, 1, 0)$ Short-run Measures of Forecast Accuracy

MAE	MAPE	RMSE
0.32067	0.3178	0.3836

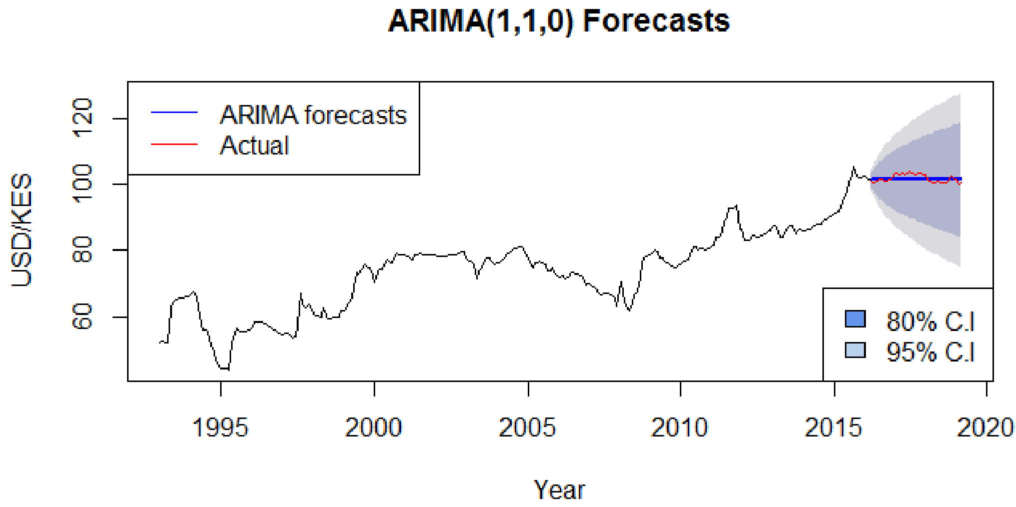

Figure 4.13: Forecasts from ARIMA $(1, 1, 0)$

The MAE, MAPE and RMSE for the next three years was as follows:

Table 4.4: ARIMA$(1, 1, 0)$Long-run Measures of Forecast Accuracy

MAE	MAPE	RMSE
1.0327	1.0051	1.3203

4.3 Seasonal Autoregressive Intergrated Moving Average (SARIMA) model

4.3.1 Statistics and Data Analysis

In this study we considered that the time series is seasonal with a seasonality of 12 months. We used the seasonal Mann-Kendall test to test for a seasonal trend. This tested

$$H_0 \quad : \quad \text{There is no seasonal trend in the monthly exchange rate series}$$
$$H_1 \quad : \quad \text{There is a seasonal trend in the monthly exchange rate series}$$

From Figure 4.14, the p-value is significant since it is less than 0.05. Therefore, we reject the null hypothesis and conclude that the series exhibited a seasonal trend. Since $\tau = 0.714$ we concluded that the series had an increasing seasonal trend.

```
tau = 0.714, 2-sided pvalue =< 2.22e-16
```

Figure 4.14: Seasonal Mann-Kendall test results

From Figure 4.8, seasonality appeared in the ACF by tapering slowly at multiples of 12 months. There was a tapering around lag 12 which showed presence of a possible seasonal MA(1) and a spike at lag 24 which showed presence of a seasonal MA(2). Therefore possible SARIMA models were SARIMA$(0,1,1)(0,0,2)[12]$, SARIMA$(1,1,0)(0,0,2)[12]$ and SARIMA$(1,1,1)(0,0,2)[12]$. Fitting the models yielded the following results in Figure 4.15.

```
Series: cl.train
ARIMA(1,1,0)(0,0,2)[12]

Coefficients:
         ar1     sma1    sma2
      0.2318  -0.0882  0.1412
s.e.  0.0585   0.0625  0.0607

sigma^2 estimated as 2.866:  log likelihood=-539.63
AIC=1087.26   AICc=1087.4   BIC=1101.77
Series: cl.train
ARIMA(0,1,1)(0,0,2)[12]

Coefficients:
         ma1     sma1    sma2
      0.2267  -0.0940  0.1461
s.e.  0.0574   0.0623  0.0604

sigma^2 estimated as 2.871:  log likelihood=-539.9
AIC=1087.8   AICc=1087.95   BIC=1102.31
Series: cl.train
ARIMA(1,1,1)(0,0,2)[12]

Coefficients:
         ar1      ma1     sma1    sma2
      0.2973  -0.0695  -0.0874  0.1399
s.e.  0.3611   0.3828   0.0627  0.0612

sigma^2 estimated as 2.876:  log likelihood=-539.61
AIC=1089.22   AICc=1089.44   BIC=1107.36
```

Figure 4.15: Fitted SARIMA models

On the basis of AIC and BIC, SARIMA$(1,1,0)(0,0,2)[12]$ was the best model since it had minimum values. From Equation 3.1, the SARIMA$(1,1,0)(0,0,2)[12]$ equation was $\phi_1(B)\nabla X_t = \Theta_2(B^{12})\omega_t$ where $\phi_1(B) = 1 - \phi_1 B$, $\nabla = 1 - B$ and $\Theta_2(B^{12}) = 1 + \Theta_1 B^{12} +$

$\Theta_2 B^{24}$. Therefore

$$\begin{aligned}
(1-\phi_1 B)(1-B) &= (1+\Theta_1 B^{12}+\Theta_2 B^{24})\omega_t \\
(1-\phi_1 B)(X_t - X_{t-1}) &= \omega_t + \Theta_1\omega_{t-12} + \Theta_2\omega_{t-24} \\
X_t - X_{t-1} - \phi_1 X_{t-1} + \phi_1 X_{t-2} &= \omega_t + \Theta_1\omega_{t-12} + \Theta_2\omega_{t-24} \\
X_t - (1+\phi_1)X_{t-1} + \phi_1 X_{t-2} &= \omega_t + \Theta_1\omega_{t-12} + \Theta_2\omega_{t-24} \\
X_t &= (1+\phi_1)X_{t-1} - \phi_1 X_{t-2} + \omega_t \\
&\quad + \Theta_1\omega_{t-12} + \Theta_2\omega_{t-24} \qquad (4.3)
\end{aligned}$$

From Figure 4.15 $\phi_1 = 0.2318$, $\Theta_1 = -0.0882$ and $\Theta_2 = 0.1412$. Hence Equation 4.3 was

$$X_t = 1.2318X_{t-1} - 0.2318X_{t-2} + \omega_t - 0.0882\omega_{t-12} + 0.1412\omega_{t-24} \qquad (4.4)$$

4.3.2 Residual Analysis

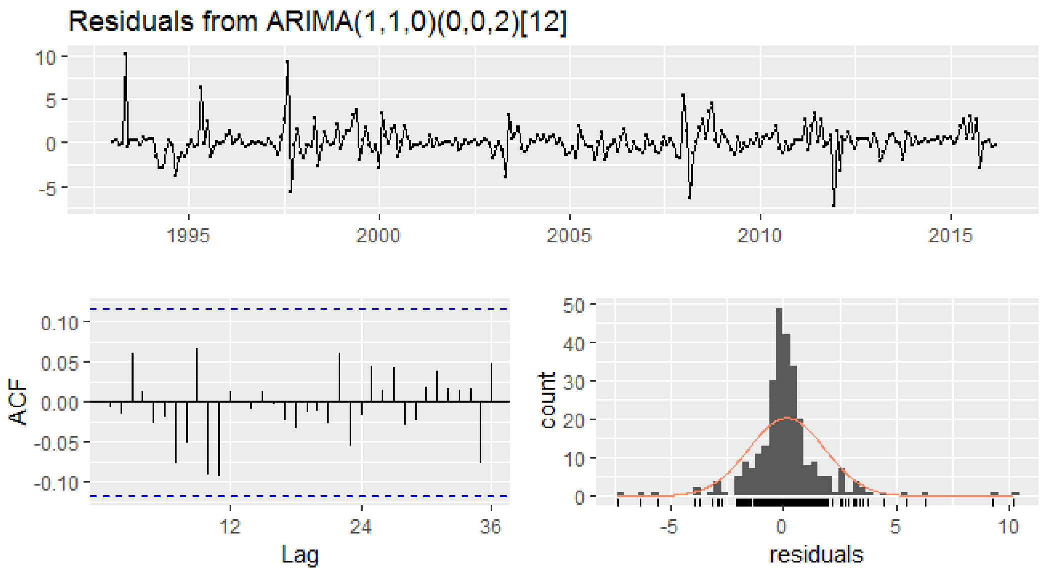

Figure 4.16: Residuals from SARIMA $(1,1,0)(0,0,2)[12]$

Plotting the ACF of the residuals from SARIMA $(1, 1, 0)(0, 0, 2)[12]$, Figure 4.16, showed that they were uncorrelated since none of them were significant. The histogram of the residuals showed that they were normally distributed. Plotting the normal Q-Q plot of the residuals also showed that they were normally distributed. The Ljung-Box Q test plot showed that the residuals were independently distributed since all p-values were greater than 0.05.

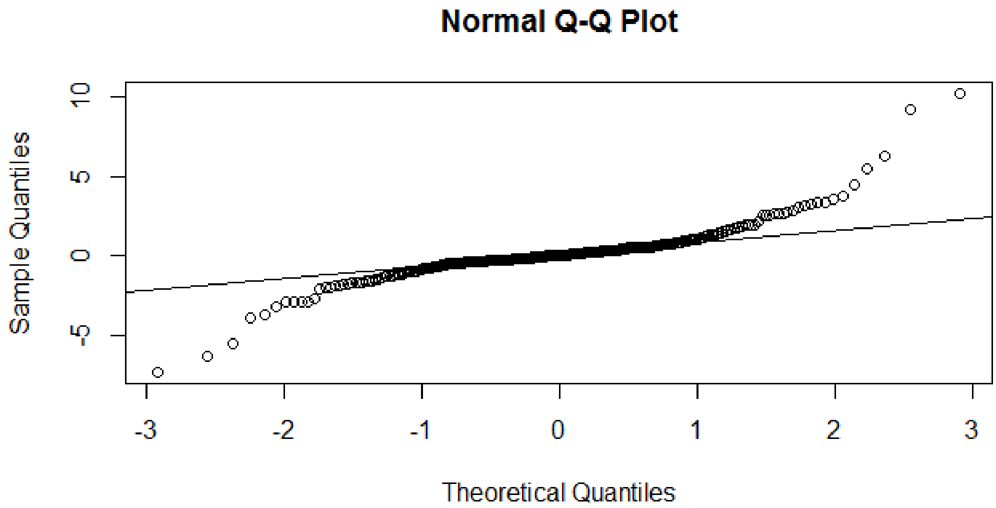

Figure 4.17: Normal Q-Q plot of residuals from SARIMA $(1, 1, 0)(0, 0, 2)[12]$

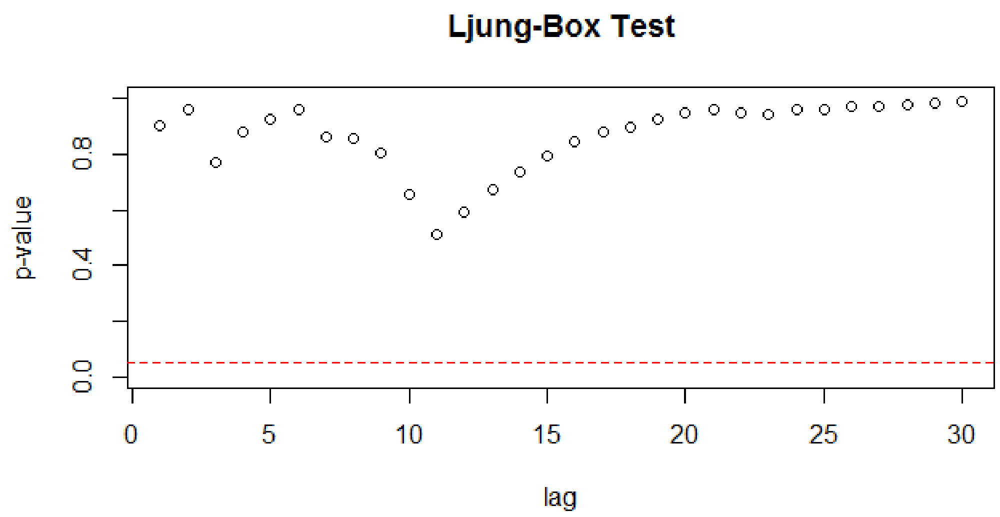

Figure 4.18: Ljung-Box test of residuals from SARIMA $(1, 1, 0)(0, 0, 2)[12]$

This showed that SARIMA $(1, 1, 0)(0, 0, 2)[12]$ is a good model.

4.3.3 Forecasting using SARIMA

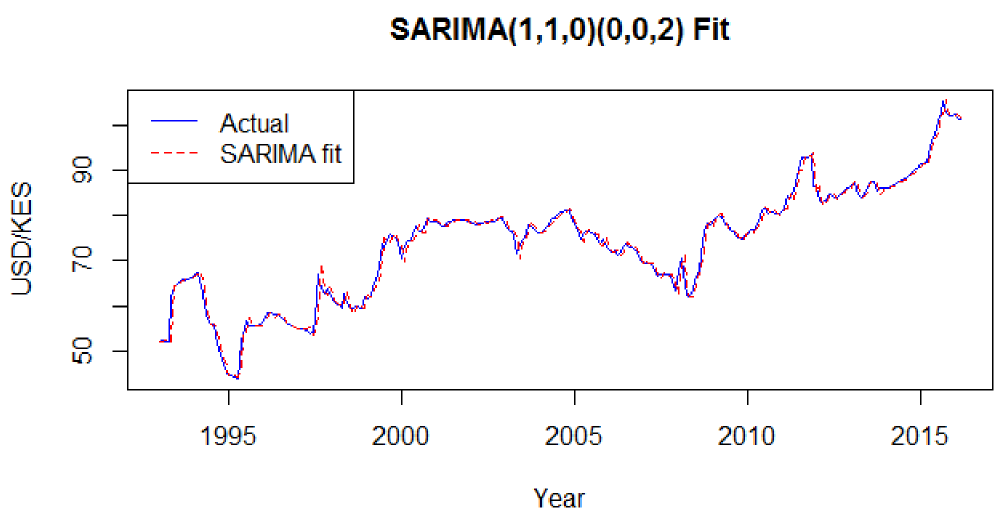

Figure 4.19: SARIMA $(1, 1, 0)(0, 0, 2)[12]$ in-sample forecasts

The model obtains good in-sample fit as evidenced by Figure 4.19. Using Equation 4.4

- One-step out-sample forecast

$$\begin{aligned}
\hat{X_{t+1}} &= 1.2318X_t - 0.2318X_{t-1} + \omega_{t+1} - 0.0882\omega_{t-11} + 0.1412\omega_{t-23} \\
\hat{X_{280}} &= 1.2318X_{279} - 0.2318X_{278} + \omega_{280} - 0.0882\omega_{268} + 0.1412\omega_{256} \\
&= 1.2318(101.485) - 0.2318(101.932) - 0 + 0.0882(1.8177) + 0.1412(0.04676) \\
&= 101.2277
\end{aligned}$$

Where $\hat{X_{280}}$ is the forecasted value of April 2016, $\omega_{t+1} = \omega_{280} = 0$ since the expected value of future residuals is not known, $\omega_{t-11} = \omega_{268} = X_{268} - \hat{X_{268}}$ and $\omega_{t-23} = \omega_{256} = X_{256} - \hat{X_{256}}$.

- Two-step out-sample forecast

$$\begin{aligned}
\hat{X_{t+2}} &= 1.2318\hat{X_{t+1}} - 0.2318X_t + \omega_{t+2} - 0.0882\omega_{t-10} + 0.1412\omega_{t-22} \\
\hat{X_{281}} &= 1.2318\hat{X_{280}} - 0.2318X_{279} + \omega_{281} - 0.0882\omega_{269} + 0.1412\omega_{257} \\
&= 1.2318(101.2277) - 0.2318(101.485) - 0 + 0.0882(2.5749) + 0.1412(0.4963) \\
&= 101.011
\end{aligned}$$

Where $\hat{X_{281}}$ is the forecasted value of May 2016, $\omega_{t+2} = \omega_{281} = 0$ since the expected value of future residuals is not known, $\omega_{t-10} = \omega_{269} = X_{269} - \hat{X_{269}}$ and $\omega_{t-22} = \omega_{257} = X_{257} - \hat{X_{257}}$.

Appendix A provides the forecasted values for the next three years.

The following forecast plot in Figure 4.20 for the next three years was obtained. From Figure 4.20, all the forecasts were well within the confidence limits. Similar to the ARIMA model, this model produced more accurate results for the first few months. The long term forecasts eventually went to a straight line. The MAE, MAPE and the RMSE for the next three months was as follows:

Table 4.5: SARIMA$(1, 1, 0)(0, 0, 2)[12]$ Short-run Measures of Forecast Accuracy

MAE	MAPE	RMSE
0.1651	0.1636	0.2037

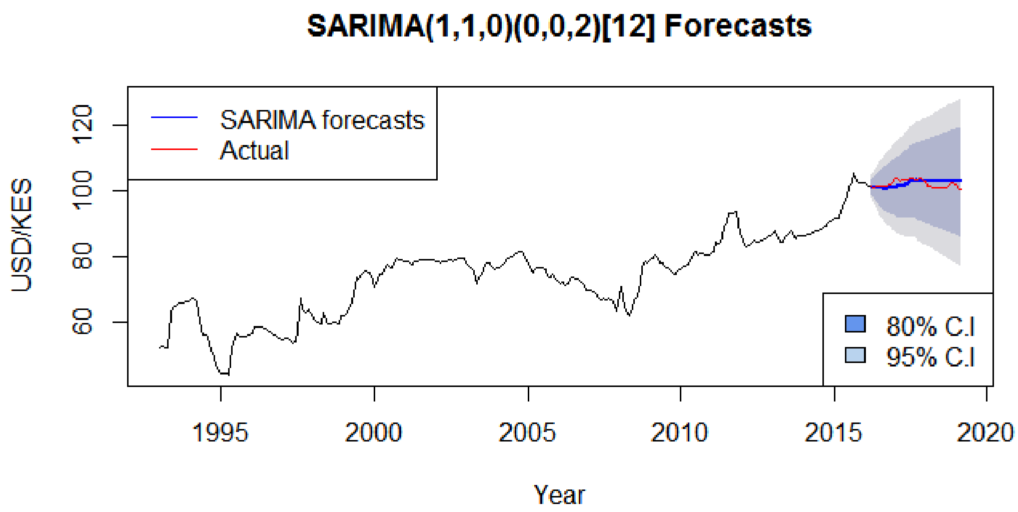

Figure 4.20: Forecasts from SARIMA $(1, 1, 0)(0, 0, 2)[12]$

The MAE, MAPE and the RMSE for the next three years was as follows:

Table 4.6: SARIMA$(1, 1, 0)(0, 0, 2)[12]$ Long-run Measures of Forecast Accuracy

MAE	MAPE	RMSE
1.1933	1.1714	1.4080

4.4 Performance Criterion

Table 4.7: Performance criterion

Model	MAE	MAPE	RMSE
3-months			
ARIMA	0.3207	0.3178	0.3836
SARIMA	0.1651	0.1636	0.2037
3-Years			
ARIMA	1.0327	1.0051	1.3203
SARIMA	1.1933	1.1714	1.4080

It was clear that SARIMA performed better than ARIMA in the short run since the values of the measures of forecast accuracy were less than those from the ARIMA model. In the long run, there was no much difference among the two models as evidenced by the values of MAE, MAPE and RMSE in Figure 4.7.

CHAPTER FIVE

SUMMARY , CONCLUSIONS AND RECOMMENDATIONS

This chapter provides the summary of findings in this study. It also provides conclusions of the study, suggestions for further research and recommendations.

5.1 Summary

The Mann-Kendall test established that the USD/KES monthly exchange rate series had an overall increasing linear and seasonal trend. The ADF test established that the USD/KES exchange rate was not stationary. Stationarity was obtained by taking a first regular difference of the series.

The first objective of this study was to fit a SARIMA model to the USD/KES exchange rate. Three possible models were found i.e SARIMA$(0, 1, 1)(0, 0, 2)[12]$, SARIMA$(1, 1, 0)$ $(0, 0, 2)[12]$ and SARIMA$(1, 1, 1)(0, 0, 2)[12]$. SARIMA$(1, 1, 0)(0, 0, 2)[12]$ was found to be the best model on the basis of AIC and BIC. An ACF plot of the residuals of this model showed that they were uncorrelated since none of them were significant. The residuals were also found to be normally distributed after plotting a histogram and a normal Q-Q plot. The Ljung-Box Q test plot showed that the residuals were independently distributed since all p-values were greater than 0.05. The form of this model was $X_t = 1.2318X_{t-1} - 0.2318X_{t-2} + \omega_t - 0.0882\omega_{t-12} + 0.1412\omega_{t-24}$.

The second objective was to forecast the USD/KES exchange rates. Using the above equation, a forecast plot for the next three years was obtained. It was observed that all the forecasted values were well within the confidence limits. It was also clear from the forecast plot that the model produced more accurate results in the short run.

The third objective was to determine the forecasting performance of the SARIMA model. In the short run i.e 3 months this model had the least MAE, MAPE and RMSE values of 0.1651, 0.1636 and 0.2037 respectively hence had higher forecasting performance than ARIMA. In the long run i.e 3 years there was no much difference between the forecasting performance of the two models.

5.2 Conclusions

The first objective of this study was to fit a SARIMA model to the USD/KES exchange rate. The study concluded that SARIMA$(1,1,0)(0,0,2)[12]$ was the best fitted model on the basis of BIC and AIC. This was also evidenced by the residual analysis.

The second and third objectives were to forecast the USD/KES exchange rates and to determine the forecasting performance of the SARIMA model respectively. It was observed that time horizon plays an important role in forecasting. ARIMA family models are essentially "backward looking" i.e the long term forecasts eventually goes to a straight line. Thus it was concluded that ARIMA family models provide poor exchange rate forecasts as the time horizon increases.

5.3 Recommendations

In light of the above, this study recommends the integration of the SARIMA model in forecasting USD/KES exchange rate in Kenya in the short run.

Further studies are recommended to research on other exchange rate currencies in Kenya such as the Sterling Pound against the Kenyan shilling and the European Pound against the Kenyan shilling i.e STG/KES and EU/KES respectively. In Kenya the GARCH family models have been used to model the volatility of exchange rates but not to forecast the values of the exchange rates. Further studies are recommended to research on this.

REFERENCES

[1] Abdalla, S. Z. (2012). *Modelling exchange rate volatility using garch models: Empirical evidence from arab countries.* International Journal of Economics and Finance Vol. 4, No. 3.

[2] Adam, C. (1992). *On the dynamic specification of money demand in Kenya.* Journal of African Economies, vol. 1, no. 2.

[3] Akincilar, A., Temiz, I. and Erol, Ş. (2011). *An application of exchange rate forecasting in Turkey*, Gazi University Journal of Science 24(4), 817-828.

[4] Alonso A.M. and C. Garcia-Martos (2012), *Time Series Analysis*, Lecture Presentation, University of Carlos III, Madrid, Spain. http://www.etsii.upm.es/ingor/estadistica/Carol/TSAtema4petten.pdf.

[5] Appiah, S. T. and Adetunde, I. A., (2011). *Forecasting Exchange Rate Between the Ghana Cedi and the US Dollar using Time Series Analysis.* Current Research Journal of Economic Theory, 3(2), 76-83.

[6] Brockwell, P. J. and Davis, Richard J. (2002). Second edition.Introduction to time series and forecasting. New York: Springer-Verlag New York Inc.

[7] Box G.E.P.and Jenkins G. (1970),*Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, CA.

[8] Cannetti, E. and J. Greene (1991). Monetary growth and exchange rate as causes of inflation in African countries: An empirical analysis. IMF Working Paper, 91/676, Washington, D.C.

[9] Cryer, Jonathan D. and Chan, Kung-Sik (2008). Time series analysis with applications in R. Second edition. New York: Springer Science+Business Media, LCC.

[10] Ette, H. Etuk (2012). *Forecasting Nigeria naira – US dollar exchange rate by a Seasonal Arima model.* American Journal of scientific research, Issue 59 , pp. 71-78.

[11] Ette Harrison Etuk (2014). *Modeling of Daily Nigerian Naira-British Pound Exchange Rates Using SARIMA Methods*. British Journal of Applied Science and Technology, 4(1):222-234.

[12] Karmeshu, Neha (2012). *Trend detection in temperature and precipitation using mann-kendall test. A case study to assess climate change on select states in the Northeastern United States*. MSc. Thesis, University of Pennsylvania.

[13] Killick, T. and Mwega, F.M.. (1989). Monetary policy in Kenya, 1967-88. ODI Working Paper No. 39.

[14] Kipkoech, R. T. (2014). *Modeling volatility under normal and student-t distributional assumptions (a case study of the Kenyan exchange rates)*. American Journal of Applied Mathematics and Statistics Vol.2, No. 4, pp 179-184.

[15] Kyule, Nicholas N.(2016). *Factors Affecting Kenya Shillings Volatility Against the United States Dollar*. MSc Thesis,University of Nairobi.

[16] Latifa, S. Omar, J. K. K. and Mutiso,J. M. (2013). *Garch modelling in monthly foreign exchange in Kenya*. America journal of Mathematical Science and Applications Volume 1, Number 1, PP. 13–39.

[17] Levich, Richard M.(1983). Empirical Studies of Exchange Rates:Price Behavior, Rate Determination and Market Efficiency, National Bureau of Economic Research:NBER paper: 1112.

[18] Mayaka, Cliff M. (2013), Measuring exchange rate risk of Kenya's main trading currencies, Postgraduate Diploma in Actuarial Science, Univerty of Nairobi.

[19] Meese, Richard A. and Rogoff, Kenneth (1983). *Empirical exchange rate models of the seventies - Do they fit out of sample?*, Journal of International Economics 14,3-24.

[20] Mwega, F.M. (1990). An econometric study of selected monetary policy issues in Kenya. ODI Working Paper no. 42, August 1990.

[21] Nau, R.(2014). Notes on the random walk model. Fuqua School of Business.Duke University.

[22] Ndung'u, N. (2000). The Exchange Rate and Monetary Policy in Kenya. African Development Bank.

[23] Ndung'u, N.S (1997). Price and Exchange Rate Dynamics in Kenya: An Empirical Investigation (1970-1993). Nairobi: African Economic Research Consortium.

[24] Ndung'u, N.S. (1993). Dynamics of the Inflationary Process in Kenya, University of Gotenborg, Sweden.

[25] Newaz M.K . (2008) *Comparing the performance of time series models for forecasting exchange rate*. BRAC university journal Vol no 2, 2008 pp. 55-65.

[26] Olanrewaju, I. Shittu. and Olaoluwa, S.Yaya (2008). *Measuring forecast performance of ARMA and ARFIMA model: An application to US dollar/ UK pounds foreign exchange*. European Journal of Scientific Research Vol 32 No.2. Pp. 167-176.

[27] Omari, Cyprian. O , Mwita, Peter N. and Waititu, Antony G. (2017). *Modeling USD/KES Exchange Rate Volatility using GARCH Models* IOSR Journal of Economics and Finance,Volume 8,(Jan-Feb. 2017), PP 15-26.

[28] Onasanya, K. Olanrewaju and Adeniji, E. Oyebimpe (2013). *Forecasting the exchange rate between Naira and US dollar using time domain model*, International Journal of Development and Economic Sustainability, Vol. 1, pp. 45-55.

[29] Rossi, B. (2005), *Are Exchange Rates Really Random Walks?Some Evidence Robust to Parameter Instability*, Macroeconomic Dynamics10(1), 20-38.

[30] Saeid Hosseinipoor (2016).Application of ARIMA and GARCH Models in Forecasting the Natural Gas Prices.

[31] sarno) Sarno, L. and M.P. Taylor (2002),*The economics of exchange rates*. Cambridge University Press.

[32] STAT 510, Department of Statistics Online Programs, *Applied Time Series Analysis,Lesson 4*(The Pennsylvania State University).Retrieved from https://newonlinecourses.science.psu.edu/stat510/node/86/.

[33] Sylvia, A. Omolo (2014). Modeling exchange rate volatility of KES/USD using GARCH faimily models, MSc. Actuarial Science, University of Nairobi.

[34] Tran M. and Uyen Ngan (2016), *Forecasting Foreign Exchange Rate by using ARIMA Model: A Case of VND/USD Exchange Rate*, Research Journal of Finance and Accounting, Vol.7, No.12.

[35] Yao, E. Ayekple, Emmanuel H. , Nana, K. Frempong, and Amevialor,J (2015) *Time Series Analysis of the Exchange Rate of the Ghanaian Cedi to the American Dollar*, Journal of Mathematics Research, Vol. 7.

[36] Yasir Kamal, Hammad-Ul-Haq, U. G. and M. M. Khan (2012). *Modelling exchange rate volatility using generalized autoregressive conditional heteroscedastic (garch) type models.* African Journal of Business Management Vol.6(8), pp. 2830–2838.

APPENDIX A: SARIMA$(1,1,0)(0,0,2)$ OUT-SAMPLE FORECASTS

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Apr 2016	101.2276	99.05796	103.3973	97.90941	104.5458
May 2016	101.0108	97.56847	104.4532	95.74620	106.2755
Jun 2016	100.9289	96.49842	105.3593	94.15309	107.7047
Jul 2016	100.6802	95.43053	105.9299	92.65151	108.7089
Aug 2016	100.6153	94.65500	106.5756	91.49982	109.7308
Sep 2016	100.4495	93.85409	107.0448	90.36271	110.5362
Oct 2016	100.6610	93.48638	107.8356	89.68838	111.6336
Nov 2016	100.8189	93.10843	108.5293	89.02676	112.6110
Dec 2016	100.9150	92.70353	109.1264	88.35667	113.4732
Jan 2017	101.0109	92.32740	109.6945	87.73062	114.2913
Feb 2017	101.0712	91.93988	110.2024	87.10607	115.0362
Mar 2017	101.1821	91.62403	110.7402	86.56430	115.7999
Apr 2017	101.4644	91.55039	111.3785	86.30221	116.6267
May 2017	101.8934	91.64702	112.1398	86.22291	117.5639
Jun 2017	102.0655	91.49974	112.6313	85.90656	118.2244
Jul 2017	102.5391	91.66392	113.4143	85.90694	119.1713
Aug 2017	102.7120	91.53608	113.8880	85.61989	119.8042
Sep 2017	103.1228	91.65398	114.5915	85.58277	120.6627
Oct 2017	102.8124	91.05807	114.5667	84.83572	120.7890
Nov 2017	102.7201	90.68704	114.7532	84.31711	121.1231
Dec 2017	102.7081	90.40256	115.0136	83.88841	121.5277
Jan 2018	102.7343	90.16225	115.3064	83.50701	121.9616
Feb 2018	102.6717	89.83860	115.5047	83.04519	122.2981
Mar 2018	102.6140	89.52517	115.7029	82.59635	122.6317
Apr 2018	102.6007	89.19842	116.0029	82.10369	123.0977
May 2018	102.5976	88.87399	116.3212	81.60915	123.5860
Jun 2018	102.5969	88.55582	116.6379	81.12294	124.0708
Jul 2018	102.5967	88.24443	116.9490	80.64680	124.5466
Aug 2018	102.5967	87.93960	117.2537	80.18062	125.0127
Sep 2018	102.5967	87.64097	117.5523	79.72391	125.4694
Oct 2018	102.5966	87.34818	117.8451	79.27613	125.9172
Nov 2018	102.5966	87.06091	118.1324	78.83679	126.3565
Dec 2018	102.5966	86.77886	118.4144	78.40543	126.7879
Jan 2019	102.5966	86.50175	118.6915	77.98162	127.2117
Feb 2019	102.5966	86.22932	118.9640	77.56499	127.6283
Mar 2019	102.5966	85.96136	119.2319	77.15518	128.0381

Figure 0.1: SARIMA$(1,1,0)(0,0,2)$ Forecasts

APPENDIX B: SARIMA(1, 1, 0)(0, 0, 2) RESIDUALS

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1993	0.052110360	0.233177375	-0.314501321	-0.149007410	10.190846811	-0.385825833	0.336499413	0.275304976	0.263571565	-0.369740883	0.574020788	0.224011840
1994	0.352700270	0.523140624	-1.513929052	-2.946327335	-2.934338895	-0.806615123	0.259998557	-0.358957046	-3.710585930	-1.734711736	-1.169053071	-1.564773038
1995	-0.209212456	0.168552918	-0.420192676	-0.307127109	6.261529551	-0.112936628	2.543462121	-1.650104400	-0.416002365	0.047002835	-0.178755119	0.094139510
1996	0.783989901	1.309753060	-0.073573704	0.319693189	0.837109902	-0.116361046	-0.426785247	-0.259719821	-0.007440658	-0.194336055	0.048317699	-0.191928899
1997	-0.204807304	0.379367088	-0.031692122	-0.432192607	-1.318871935	0.637224138	2.615761632	9.245519337	-5.558020690	-0.377381810	1.587602280	-1.210580081
1998	-1.811725884	-0.361055891	-0.234450379	-0.505408777	2.875346158	-2.704584124	-0.407903846	1.163483685	0.136011597	-0.289815463	-0.075970360	2.162119990
1999	-0.651628770	0.611833559	1.337933620	1.305278727	3.227783840	3.723128346	-2.021567693	0.230898141	1.758655941	-0.375955228	-0.987343085	-0.303068760
2000	-2.867632966	3.399017030	0.774901649	-0.162388581	1.501622797	1.912880177	-1.624835425	0.162106702	1.875238921	0.662364816	-0.722077304	-0.363279781
2001	-0.259289320	-0.113030024	-0.535004093	-0.337415281	0.776666434	-0.518131185	0.521496565	-0.214542312	-0.026721393	0.125104725	0.062813879	-0.260412389
2002	0.356251485	-0.816221199	-0.269180730	0.254887285	-0.152766736	0.022717580	0.328751499	-0.295874758	0.017580762	0.380521250	0.228656919	-0.058547795
2003	-1.740774248	-0.512153853	-0.002950007	-0.797074760	-3.957272091	3.128621912	0.490172677	0.979610707	1.668177652	-0.573662502	-0.983475430	-0.449365214
2004	0.238754417	0.101074174	0.887724010	0.339574659	0.855204826	-0.009111849	0.711579115	0.796096500	-0.153824153	0.400997038	-0.228544220	-1.461847599
2005	-1.245727545	-0.483377603	-1.826331617	1.980339201	0.573873151	-0.216681711	-0.519241429	-0.389453927	-1.856582044	0.117784726	1.239002596	-1.935043861
2006	-0.658595548	-0.259943374	0.285552226	-0.960763510	0.616343950	1.516549781	-0.274623703	-0.992165117	0.036309875	-0.622293231	-0.886674533	-1.195028765
2007	0.723427817	-0.283491112	0.022387028	-1.005495124	-1.246683910	-0.130351814	0.684850410	-0.268826860	0.371594540	-0.268615825	-1.566671278	-1.705193991
2008	5.441710414	1.447252082	-6.327749439	-1.299946824	0.064364944	1.741135128	2.583530731	0.414332559	3.531678405	4.447624325	0.289577312	-0.469813161
2009	1.319538508	0.539804689	0.031382232	-0.776478557	-1.436134090	0.571120094	-0.966411401	-0.049532753	-0.419999666	0.247134232	-0.174593358	1.008334320
2010	-0.457225591	0.705024496	0.894328209	0.371719085	1.080039483	1.983280300	-0.616129650	-1.143431600	0.164868416	-0.913512505	-0.264396123	0.322170801
2011	0.209330996	0.323148509	2.704571522	-0.807025321	1.914293964	3.352731369	0.092966880	2.597322930	-0.443943847	-0.176269346	0.771316826	-7.304021182
2012	1.388477691	-3.163854059	0.567416066	0.231978188	1.144980121	0.143622085	-0.647681218	0.476029843	0.490617756	0.487721003	0.506728058	-0.444785812
2013	0.914360411	0.011223570	-2.086318167	-1.117260984	0.165331565	0.891297438	0.989681068	-0.008456668	-0.120980771	-2.016537731	1.216244260	1.014142681
2014	-0.258093798	0.533684850	-0.068028958	0.046762205	0.496325291	0.097050044	0.293409495	0.227731334	0.572877467	-0.024984276	0.681150352	0.462690168
2015	0.650651828	-0.035339373	0.496182804	1.817698904	2.574880405	0.514753607	3.072145813	0.447152513	2.625382937	-2.872683231	-0.144092211	0.066264127
2016	0.205590371	-0.486814020	-0.305307538									

Figure 0.2: SARIMA(1, 1, 0)(0, 0, 2) Residuals