FORECASTING KENYA'S INFLATION RATE USING A VARMA FOR PRICE OF IMPORTED CRUDE OIL AND KENYA'S PREVIOUS INFLATION RATE TIME SERIES

BY

AMISI, Pascal Ouma
MSC/MAT/00158/2015

A THESIS 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 thesis is my own work and has not been presented for a degree award in any other institution;

AMISI, Pascal Ouma

MSC/MAT/00158/2015

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

This thesis has been submitted for examination with our approval as the university supervisors;

Dr. Edgar Ouko Otumba

Department of Statistics and Actuarial Science,

Maseno University.

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

Dr. Joyce Akinyi Otieno

Department of Statistics and Actuarial Science,

Maseno University.

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

Acknowledgement

Special thanks goes to my supervisors, Dr. Joyce Akinyi Otieno and Dr. Edgar Ouko Otumba, for their time and insightful guidance and assistance at every stage of this work. For your support, patience and generous words and advice, I say thank you. I am sincerely grateful to my family, collegues and friends for their encouragement and support.

Dedication

I dedicate this research to the Almighty God who gave me physical strength and mental capacity to undertake this project. I would also like to give special dedication to my beloved family for their financial and moral support throughout my study.

Abstract

Inflation is the persistent rise in the prices of selected goods and services over time. The rate of inflation measures economic performance of a country and is an important economic indicator to economists of any given government. High rates of inflation lead to slow economic growth and has the effect of lowering the living standards of a population by eroding their purchasing power. In the period November 2016 to June 2017, Kenya experienced an unprecedented rise in the inflation rate to a high of 11.7% causing harsh economic and social repercussions to her population [5]. To cushion its population against such strain, the government should be able to estimate and predict the rate of inflation. Previous research by Bilal Kargi for Turkey's case [6] indicates a relationship between the changes in price levels of imported crude oil and the rate of inflation. The objective of this study was to determine if there is a long-run relationship between Kenya's Inflation rate and the price of imported crude oil, fit a VARMA (p,q) model and use the fitted model to forecast Kenya's inflation rate using the previous rates of inflation and the price of imported crude oil since there was a cointegrated association between the two time series. This will enable the government plan strategically for the mid- and long-term effects of inflation in Kenya. Cross-correlation analysis was used to determine whether there is a significant correlation between the two time series and a test of cointegration was used to determine a significant association. A VARMA model was fitted to the data using the SCM approach. The study showed that there exists a moderate negative correlation between the two time series with a correlation coefficient of -0.21, with a p-value of < 0.05 that implies that the correlation is statistically significant. The study further showed that there is a moderate statistically significant association between the two time series at lags 6 and that there exists cointegration and dependencies between the price of imported crude oil and the Kenya's Inflation rate by a CADF test which a statistically significant Dickey Fuller Statistic of -8.3394, with a $p-value = 0.01$, implying cointegrating association between the two time series. A VARMA $(2, 1)$ model was fitted to the data and used to forecast Kenya's inflation rates to six steps (months) behind for comparison to the actual available data and further a eleven-steps ahead forecast. The forecasts were accurate with a Mean Absolute Error (MAE) of 0.66% which are good forecasts according to [17] for planning purposes. From the study results it shows that there exists a statistically significant association between the price of crude oil and Kenya's previous inflation rates and therefore used in forecasting future Kenya's inflation rates. This study therefore provides better inflation forecasts(Kenya) to be used for strategical planning for the mid- and long-term effects of inflation by the government.

TABLE OF CONTENTS

Abbreviations and Acronyms

ARDL Autoregressive Distributed Lag Model

AR Autoregressive

CPI Consumer Price Index

ECI Employment Cost Index

GDP Gross Domestic Product

IPP International Price Program

KNBS Kenya National Bureau of Statistics

NOCK National Petroleum Corporation of Kenya

PPI Producer Price Index

RFDLM Restricted Finite Distributed Lag Model

VAR Vector Autoregression

CADF Cointegrated Augmented Dickey-Fuller test

SCM Scalar Constant Methodology

MAPE Mean Absolute Percentage Error

MAE Mean Absolute Error

Definition of Terms

Inflation the general increase in prices of goods and services and a corresponding decrease in the purchasing power of money.

Consumer Price Index an economic index that determines the weighted average prices of a fixed basket of consumer goods and services.

Producer Price Index an index that measures the average change in the selling prices received by domestic producers of goods and services over a given time period.

Time series plot a graph of how a specified variable(s)changes over time.

Stationary time series a time-series whose statistical properties such as the mean, variance, autocorrelation are constant over time.

Non-stationary time series a time series whose statistical properties such as the mean and variance are not constant over time. They exhibit trend and or seasonality.

Covariance stationary time series a time series Y_t is said to be covariance stationary if it has a constant mean, constant (and finite) variance and a stable autocovariance function.

Order of integration of a time series the number of differences required to make a time series covariance stationary.

Co-integration of time series two or more non-stationary time series of same order of integration yet a linear combination of these series is stationary.

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

Mean Absolute Percentage Error a statistical measure to determine the accuracy of forecasts.

Mean Absolute Error a statistical measure to determine the accuracy of forecasts.

List of Figures

CHAPTER 1

INTRODUCTION

1.1 Introduction

Inflation is the general rise in prices of a selected basket of goods with reference to a certain base year. This is usually coupled with the falling of the currency's purchasing power, [5]. Inflation is determined using various methods some of which include a change in the GDP, the CPI, the PPI indicator and the ECI. In Kenya, it is mostly determined using the CPI indicator by KNBS. The world's general inflation rate is about 3.1% yet Africa's inflation rate is approximately 6.96% .

High inflation rates in an economy has consequences which include, income redistribution in families due to the increase in food and basic domestic utilities, reduced level of real incomes due to the cut in the workers' wages and increase in the borrowing cost through financial markets who want to protect themselves from the negative effects of high inflation rates [13]. Basically, inflation reduces the people's purchasing power.

Researchers mull over the prospect of involving more factors to forecast inflation rates. The price of imported crude oil would be useful to ensure a better forecast due to their dependencies. A Vector autoregressive (VAR) model has previously been used to try and determine the relationship between crude oil prices, money supply, GDP and the rate of inflation. The main challenge to the researchers was that since the number of time dependent variables were five, it meant estimating parameters using very few data points hence the forecasts determined were not as accurate, [4]. Mihaela [8] showed that when fitting Multivariate Time Series models, models with fewer number of time dependent variables are usually more accurate. This therefore implies that the when fitting these models caution should be taken to ensure that only variables that contribute the most to the association are considered in the model.

In this study therefore, two time series, the price of imported crude oil and the previous rates of inflation, are used to forecast the inflation rates in Kenya. A bivariate time series analysis (VARMA(p, q)) is an appropriate method since both the price of imported crude oil and the rate of inflation are factors of time and there exists a cointegration association between the two time series.

1.2 Statement of the problem

Inflation has a dilapidating effect on any economy. Between the months of November 2016 through to April 2017, the rate of inflation in Kenya hit unprecedented high of 11.7% and lowest of 7%, which led to a strain on the economy of the country [5]. The country experienced an increase in the price of basic commodities, increased unemployment rates and consequently reduced real wage levels due to job cuts. An accurate forecast of this unprecedented inflation rate could have cushioned the Kenyan population from its effects. Although, previous research show that some other factors have an effect on the inflation rate, the price of crude oil has a significant effect on the rate of inflation of any given economy. Further since petroleum, a refined product of crude oil, is an important input in the manufacturing processes in Kenya, its price contributes to the final product prices and hence to the rate of inflation. Previous researches attempted to model the association between Kenya's inflation rate and price of imported crude oil but univariate time series models do not capture the contribution of other variables. On the other hand, they used VAR models that had 5 other variables which could lead to forecasts that are not highly accurate since out of the five variables not all had a cointegrating association hence the need for a better model which includes cointegrating associtions between them.

1.3 Objectives of the study

1.3.1 Main Objective

The aim of this study was to determine the relationship between the price of imported crude oil and the rate of inflation in Kenya with the intention of forecasting inflation rates.

1.3.2 Specific Objectives

The specific objectives of the study were to:

1. Determine the cross-correlation between the rate of inflation in Kenya and the price of imported crude oil time series.

2. Determine if there is a co-integration between the time series of price of imported crude oil and rate of inflation in Kenya.

3. Fit a multivariate time series model (VARMA(p, q)) to the variables rate of inflation and the price of imported crude oil in Kenya.

4. Forecast inflation rates in Kenya using A fitted VARMA(p, q) model with the previous Kenyas inflation rates and the price of imported crude oil as the time-dependent variables.

1.4 Significance of the study

This study will provide information on the relationship between the rate of inflation and the price of imported crude oil in Kenya. It will define the dependencies between these two factors so that the price of imported crude oil can be used to forecast inflation rates. This will enable the government plan strategically for the mid- and long-term effects of inflation in Kenya. The study will also supplement available literature on the forecasting of inflation rates.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

This chapter presents a literature review of studies on modelling and forecasting of inflation rates. These include studies that indicate the relationship between inflation rates and crude oil prices, studies that model rates of inflation and studies that use the fitted models to forecast inflation rates. It is intended to developing of time series models on the nexus between the price of importation of crude oil and the rate of inflation in Kenya.

2.2 Inflation rate and the price of imported crude oil

It is generally thought that oil prices have an effect on production costs and hence price of commodities eventually. To determine the relationship between crude oil price changes and the inflation rates, [2], was established that most of the countries in the study exhibited a long run asymmetric statistical association of inflation rates to the changes in the price of crude oil. The association exhibited a statistically significant linear association in the shorter periods of the available data and that the cumulative impact of crude oil price increase is greater in comparison to the cumulative impact of a decrease in the crude oil price. The study however investigated data from several countries without regard to whether the country was a crude oil exporter or a crude oil importer and did not account for the difference in the economic ratings of the twenty-nine countries under study. Since Kenya is an importer of crude oil, the model developed by Juan may not accurately model Kenya's inflation rate.

In yet another study, [6] examined the effect of crude oil prices on inflation and economic growth in Turkey over the long-run. It was apparent that imported crude oil price granger causes the GDP of Turkey, while Turkey's inflation rate granger causes the price of imported crude oil. The study also established that an increase in oil price led to inflation since the cost of production increases due to the increase in the cost of energy. Aggregate output is generally affected by energy use and privately by imported crude oil use, therefore influencing economic growth. The research determined that the imported crude oil price has a statistical association with Turkey's inflation rate and further influencing the rate of economic growth which is a consequence of the rate of inflation. Since Turkey imports crude oil as an energy source it gives an indication that Kenya's inflation rate could equally influenced by the imported crude oil price since she also imports crude oil.

The main driver of the rate of inflation on average for the East African countries and Ethiopia is the price of crude oil, [1]. This study also indicated that money growth in Ethiopia and Uganda was the main driver of inflation in the short-run accounting to forty-percent and thirty-two percent of the total inflation of Ethiopia and Uganda respectively. However, Kenya and Tanzania's inflation rate was mostly influenced by imported crude oil price accounting to about 20% and 26% respectively of the countries' total inflation. Money growth also influenced the Kenya's and Tanzania's inflation rate, in the short run, but to a lesser extent in comparison to the contribution of the price of the imported crude oil,14% and 20% respectively. The research further indicated that the world food prices also influence the countries' rate of inflation in the short-run to an average of approximately 12%.

Rising prices is the root cause of inflation, [14]. Price rise has multiple causes which include when the overall demand for goods in an economy increases more rapidly than the economy's production capacity, when there is an increase in money supply, when there is an increase in the prices of production input costs or when wages increase usually to maintain the populations standards of living, [14]. Yet, over the years research links the price of imported crude oil to inflation with changes in the crude oil price accounting for more than fifteen percent of the changes in inflation [1]. It has been shown that there is both a short and mid-term relationship between the price of imported crude oil and inflation rates and that the aggregate output is a function of energy, [6]. Essentially, a rise in the cost of energy increases the cost of production and leads to inflation to some extent.

These changes in the price of imported crude oil and rate of inflation occurred concurrently with the global financial crisis together with the general increase in global crude oil prices [10]. Economists have done substantive research on the relationship between inflation and global crude oil prices using regression analysis [9]. A study conducted by the economist Bilal Kargi determined that the price of imported crude oil has a specific effect on inflation and that inflation and oil importation are long-term integrated variables in the economy of Turkey, [6]. In the 1970‘s the price of imported crude oil in the United States was studied and found to cause the cost-push type of inflation, [10] where inflation is caused by increase in the cost of factors of production.

Research carried out in Ethiopia and Tanzania to determine drivers of inflation indicated that they included crude oil prices, currency value and food prices. It was equally noted that world food prices have a long run impact on the inflation rate of both countries [4]. In a study by African Development Bank on the dynamics of inflation in Ethiopia and the East African Community, oil prices was the main driver of short-run inflation explaining about 25% of the total rate of inflation [1]. These studies indicate that the price of imported crude oil is among the main drivers of the rate of inflation.

In Kenya, it is not clear the extent to which the increase in the price of imported crude oil affects inflation. However, petroleum consumption, a refined product of crude oil, in Kenya constitutes 22% of total energy consumption in the country [12]. Kenya has experienced an increasing trend of inflation since its independence. In the 1990’s there was an increase in the inflation rate over time which continued through to the 2000’s but in 2008 it rose to a high of 18.7% in May, [4]. In the period November 2016 to June 2017, Kenya experienced an unprecedented rise in the levels of rates of inflation which led to reduced economic growth and caused a dent in the growth of Kenya’s economy [12].

Forecasting the rate of inflation is one way to cushion the citizens against the adverse effects of inflation. When forewarned, sustainable measures or polices are developed to ensure that the economy of the country is stable. The government is also able to plan accordingly in response to the expected changes in the rate of inflation. Economists have fronted several ways to forecast the inflation rate of any given country. Some economists use the simple regression models to forecast inflation while others tend to use univariate time series to forecast inflation since there is not sufficient information about factors influencing inflation, [4].

2.3 Modelling of Inflation rates

A number of researchers have modeled inflation rates with a view of making predictions due to its harmful effects on economies. [11], modeled the statistical association between oil price and inflation for selected OPEC and EU countries using monthly data from 2000 to 2014 considering the role of asymmetries and structural breaks. They used both the Linear (Symmetric) ARDL time series models and Asymmetric (Nonlinear) ARDL time series models. They determined that there were asymmetries in the short run for nations that imported crude oil. In other words, positive and negative shocks to crude oil price seem to matter only in the short run for the affected oil importing countries. However, the study did not fully explore the interaction between the price of imported crude oil and rate of inflation and their cointegration which is possible using multivariate time-series analyses. However, the result that crude oil price and the inflation rate exhibit a significant statistical relationship is assumed to hold for the Kenyan situation since Kenya is an importer of crude oil.

To determine the best model, [15] modeled the Swedish Inflation rates using economic data from the Swedish economic market. The used ARIMA distributed lag model (DLM), finite distributed lag model (FDLM) and the restricted finite distributed lag model (RFDLM). The RFDLM was found to be the best approach to model inflation with 20% RMSE compared to 32% of the naive forecast. The study also established that inflation correlates with the market data and that RFDLM together with ARIMA is the best model in comparison to DLM and ARIMA. However, the research did not fully explore the interaction between the price of imported crude oil and rate of inflation and their cointegration that is only possible through multivariate time-series analyses. Further the research findings would not generally apply since the market forces of Sweden are very different to that of Kenya hence the need for this study.

Gathingi [4], modeled the statistical association on the influence of money supply, imported crude oil prices (Murban) oil prices and exchange rate on inflation using time series by comparing two statistical time series models. The researcher compared between the univariate statistical models for each of the variables against a multivariate time series models. The researcher compared ARIMA models to VAR models. The study by [4] used ARIMA model to fit historical CPI time series resulting to the model ARIMA $(1, 1, 0)$. The researcher fitted a VAR $(1, 1)$ model to the data as well. The comparison of both models indicated that the VAR model has the least error of 0.23% mean absolute percentage hence the better model in forecasting inflation while the ARIMA model had a mean absolute percentage error of 0.57%. The study by [4] therefore concluded that the VAR models were the better models. However, the VAR model did not account for the short-term effects, which the MA terms in any model capture hence a VARMA model will suffice, and not VAR in the modeling of Kenya's inflation rate and the contribution of the price of imported crude oil. This study therefore used the VARMA model to capture the effects of past Kenya's inflation rates(1966 to 2018) and the price of imported crude oil (1966 to 2018) on current Kenya's inflation rate.

2.4 Forecasting inflation rates

Fannoh [3] forecasted Liberia's inflation using the seasonal autoregressive integrated moving average (SARIMA) approach. The SARIMA approach tends to capture the seasonality component of Liberia's inflation rate. A SARIMA $(0, 1, 0)(2, 0, 0)_{12}$ model was fitted to Liberia's rate of inflation. The forecasts accuracy was determined and the SARIMA $(0, 1, 0)(2, 0, 0)_{12}$ had the lowest values for the RMSE, MAE, MPE, MAPE and MSE of 1.626%, 1.413%, -18.05%, 20.117% and 2.644% respectively as the percentage forecast errors. [3], conducted both in sample and out of sample forecasts for the next twelve periods, twelve months, that indicated the model was able to explain the behavior of the actual observations although the values were not exactly the same. As much as it covered the seasonality component of inflation, it failed to capture the dynamism of the rate of inflation and other market forces by determining the interaction in the same system of models motivating multivariate time series analysis.

Lidiema [7], forecasted inflation rate in Kenya Using SARIMA and Holt-Winters Triple Exponential Smoothing time series models incorporating the Box-Jenkins procedure. The study evaluated the forecasts both at the objective and subjective levels. The forecasts of both models were compared and the best model was selected on the basis of the least mean absolute square error (MASE), mean absolute error (MAE) and mean absolute percentage error (MAPE). The researcher modeled the Kenyan inflation by comparing two time series models of GARCH $(1,2)$ with ARIMA $(1,1,12)$ eventually combining the two models to determine the accuracy of the forecast. From the analysis, the researchers determined that both models gave similar output but upon comparing MAE, MASE and MAPE for the models, SARIMA model was more accurate as it had the highest accuracy [7]. The Holt-Winters forecast errors were 0.643% for MASE, 0.595% for MAE and 0.400 for MAPE while the SARIMA forecasts had 0.059% for MASE, 0.004% for MAE and 0.073 for MAPE. The results of the forecast errors indicated that SARIMA model had the lowest forecast errors between the two models. However, this research did not cover the interaction in the same system of equations (multivariate time series models) hence the need for this research.

Mihaela [8] used VARMA models to forecast specified macroeconomic indicators such as GDP, Inflation etc. while comparing the accuracy of these forecasts to those forecasts determined by the VAR models. The researcher decided to use generalized forecast error second moment as measure of accuracy (GFESM). The research by [8] determined that VARMA forecasts were more accurate than those of VAR or AR models. For instance, since the researcher used a ratio of the other models' forecast to the VARMA's forecast, the ratio for VAR(AIC) model was 1.01, VAR (SIC) model was 1.06, AR model's ratio was 1.03 and the naive forecast's ratio was 1.35. Therefore this research concludes that the VARMA's forecast for inflation rate are the more accurate forecasts in relation to other models, [8].

In summary the researches that employed the use of univariate timeseries models [7] to forecast inflation rates missed out in including one of the other factor that has a statistical association(price of crude oil) hence the produced forecasts were not as accurate. The researches that employed the use of multivariate timeseries models to forecast inflation rates [4] missed out by including many time dependent variables hence lowering the accuracy of the fitted model [8]. This therefore created have the need to use a better multivariate time series model which includes other time dependent variable(s) that have proven statistically statisical assocation to the rate of inflation.
From the literature above it indicates that there is a statistical relationship between the price of imported crude oil and inflation rates [10]. This study explored the interaction between the two variables and their cointegration using multivariate time series analysis(VARMA) as discussed in the preceding chapters.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Introduction

The methodology used in this chapter will determine whether there is a statistical association between the two time series. A multivariate time series model, (VARMA(p, q) model), is fitted to the data. A suitable model was selected based on RMSE statistics and diagnostics of the residuals. The residuals from the fitted model are estimated and analyzed to determine the model's accuracy. The fitted model is used to forecast Kenya's inflation rates. Using Kenya's inflation rate data from Kenya National Bureau of Statistics (KNBS − 1966 to 2018) in percent and the National Oil Corporation Kenya (NOCK - 1966 to 2018) the prices per barrel are in US Dollars ($). Cross-correlation between the time series was performed to find out if the two time series are correlated.
Residuals from the fitted model were analysed to determine the model's accuracy and further used the selected model to forecast Kenya's inflation rates.

3.2 Product Moment Correlation

Correlation analysis was performed to determine whether there exists a linear relationship between the price of imported crude oil time series and Kenya's rate of inflation. The correlation coeffecient takes values between $-1 \leq r \leq 1$, the bounds indicating maximum correlation and 0 indicating no correlation. A high negative correlation implies a strong inverse linear association while a high positive correlation implies a direct linear association (an increase in one quantity leads to an incresease in the other quantity as well). The formula used is as shown below in (3.1):

$$r = \frac{\sum_i [(x_i - m_x) * (y_i - m_y)]}{\sqrt{\sum_i (x_i - m_x)^2} \sqrt{\sum_i (y_i - m_y)^2}} \tag{3.1}$$

3.3 Cross - Correlation Function

To achieve the objective one on determination of the cross-correlation between the price of imported crude oil time series and Kenya's rate of inflation, to develop the cross correlation matrices after determining the respective delay d. The procedure is as follows:
The cross correlation at delay d was calculated as in (3.2).

$$r_d = \frac{\sum_i[(x_i - m_x) * (y_{i-d} - m_y)]}{\sqrt{\sum_i(x_i - m_x)^2}\sqrt{\sum_i(y_{i-d} - m_y)^2}} \tag{3.2}$$

where m_x and m_y are the means of the corresponding series.
If computed for the all the delays $d = 0, 1, 2, ..., N - 1$ then a cross correlation matrix r_d is formed which has twice the length of the original size as the series. The cross correlation coefficients in the matrix assumes values such that $-1 \leq r_d \leq 1$, the bounds indicating maximum correlation and 0 indicating no correlation. A high negative correlation indicates a high correlation but of the inverse of one of the series. A CCF plot is a graphical data analysis technique for determining if correlation exists between lags for two time series. It plots the lag (an integer between 1 and $n/4$ where n is the number of observations) on the horizontal axis against the correlation coefficient (a value between -1 and 1) on the vertical axis. In addition, vertical lines are drawn at zero and at levels indicating statistically significant correlation. On the CCF plot we consider the dominant correlations at negative values of d.
Canonical correlation is defined as; for two scalar variables, the (single) canonical correlation between them is the absolute value of their ordinary correlation coefficient.

3.4 Dickey-Fuller (DF) Test for Unit Root

If a time series has a systematic unpredictable pattern then it has a unit root. The DF Test tests the null hypothesis that the AR time series is non-stationary. Consider a simple

AR(1)−process

$$\begin{aligned}X_t &= \alpha + \rho X_{t-1} + \epsilon_t \\ X_t - X_{t-1} &= \alpha + (\rho - 1)X_{t-1} + \epsilon_t \\ \Delta X_t &= \alpha + (\rho - 1)X_{t-1} + \epsilon_t \\ \Delta X_t &= \alpha + \delta X_{t-1} + \epsilon_t\end{aligned}$$

The hypothesis tests whether;

$$\begin{aligned}H_0 &: \delta = 0 \text{ non-stationary time series implying there is a unit root} \\ H_1 &: \delta \neq 0 \text{ time series is stationary}\end{aligned}$$

where Δ is the First difference operator, $E(X_t) = 0$ and $V(X_t) = \delta^2$. X_t is said to be AR(1) stationary if $|\delta| < 1$. It calculates the t-statistic on $\hat{\delta}$, the estimator of $\delta = (1 - \rho)$. However under H_0, X_t is non-stationary so the Central Limit Theorem does not apply and we cannot compare the calculated t-statistic with the critical value of the t-distribution. Fortunately, Dickey-Fuller tabulated the asymptotic distribution of the least squares estimator of δ under H_0. Hence we compare our ordinary t-statistic, t, with values of the DF distribution and if $t < DF_{critical}$, we reject H_0.

3.5 Cointegration

Testing the hypothesis that there is a statistically significant connection between two series is done by testing for the existence of a cointegrated combination of the two series. While the individual series might be non−stationary, if there is a linear combination of the residuals that is stationary, then the two time series are cointegrated. If we can find a particular parameter β such that the process $Y_t - \beta X_t$ is $I(0)$, a stationary process, we say that X_t and Y_t are cointegrated. This implies that there exists a true statistical relationship between X_t and Y_t which holds across time.

A visual test for cointegration between time series involves looking at the time series plot of Y_t and βX_t if β is known. The distance between them should seem constant i.e. $Y_t - \beta X_t = \varepsilon_t$, is a stationary process which is independent. Most of the time, β is unknown and have to be estimated using the residuals from the fitted regression model using the Least Square Estimation procedure as:

$$Y_t = \hat{\alpha_1} + \hat{\beta} X_t + \hat{\mu}_t \tag{3.3}$$

where $\hat{\mu}_t$ are the estimated values of the residuals. From Equation (3.3),

$$\hat{\mu}_t = Y_t - \hat{\alpha_1} + \hat{\beta} X_t \tag{3.4}$$

If X_t and Y_t are cointegrated then $\hat{\mu}_t$ should be a $I(0)$ process. Hence, to test for cointegration between X_t and Y_t, we carry out a Dickey-Fuller (DF) Test on our residuals.

Applying a standard DF Test on our residuals involves running a regression of the change in our residual, $\Delta\hat{\mu}_t$, on $\delta_0 + \delta_1 \hat{\mu}_{t-1}$ and in principle, we might want to add future lags to correct for any zero correlation. Hence

$$\Delta\hat{\mu}_t = \delta_0 + \delta_1 \hat{\mu}_{t-1} + \cdots + \nu_t$$

where

ν_t is a stationary cointegrating error component. Intuitively, can be thought of as short-term deviations from the long-run equilibrium.

The t-statistic, t, for $\hat{\delta}_1$ is calculated and compared with the critical values for the standard DF_1 t-distribution values. If $t < DF_1$ then the null hypothesis H_0 is rejected and we conclude that the error is an $I(0)$ process. This implies that X_t and Y_t are cointegrated. The hypothesis being tested is $H_0 : \beta = 0$ versus $H_1 : \beta \neq 0$ and if $\beta = 0$, there is no way that X_t and Y_t can be cointegrated.

The standard DF Test assumes that the value of β is known (this is usually not the case). More often than not, β is estimated. The DF distribution is amended to take into account the fact that β is estimated, hence the augmented DF (ADF) Test. The procedure remains the same however we adjust the DF distribution to DF_2. Its critical values are more negative than the original ones (DF_1) making it harder to accept H_0, $t < DF_2 < DF_1$. The more negative the critical value is, the stronger the rejection of the hypothesis that there is a unit root at some level of confidence.

3.6 Vector Autoregression Moving Average (VARMA) Model

A VARMA (p, q) model is a multivariate time series model with both autoregressive and moving average terms. The assumption for the e_t component is and e_t:t $\in \mathbb{Z} \sim$ WN(0, δ^2). $\Theta_i \neq 0$. A VARMA (p, q) takes the form (3.5)

$$y_t = \Phi_1 y_{t-1} + ... + \Phi_p y_{t-p} + x_t + \Theta_1 x_{t-1} + ... + \Theta_q x_{t-q} + e_t \tag{3.5}$$

where

$y_t = \begin{bmatrix} \text{inflation rate at time t} \\ \text{price of imported crude oil at time t} \end{bmatrix}$, $\Phi_i = \begin{bmatrix} \beta_{1i} & \alpha_{1i} \\ \beta_{2i} & \alpha_{2i} \end{bmatrix}$, the matrix of constants to be estimated for AR terms, $\Theta_i = \begin{bmatrix} \gamma_{1i} & \delta_{1i} \\ \gamma_{2i} & \delta_{2i} \end{bmatrix}$, the matrix of constants to be estimated for MA terms, $Y_{t-i} = \begin{bmatrix} y_{1t-i} \\ y_{2t-i} \end{bmatrix}$ and $e_t = \begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix}$.

3.7 Fitting a VARMA(p, q) model

From the determined cross correlation matrix, statistically significant lags are selected after checking their corresponding p-values. All the statistically significant lags are picked for model testing and fitting and the best model is picked to be fitted to the data.
Next, the number of parameters to be estimated in the model are determined. Further the statistical significance of the determined parameters are assessed by considering the p-value of the estimated parameters. Then the covariance of the residuals of the parameters are considered to ensure no serial correlation between the residuals. However the estimated model can be refined by reducing the number of parameters by setting the statistically insignificant parameters of the model to 0 to achieve a model with better and more accurate forecasts.

3.8 Model diagnostics

Model diagnostics is necessary to check the validity of the model fitted to data. It involves checking whether the assumptions of the model are met, assessing the model structure and studying subgroups of observations such as those that are poorly represented or have high influence on the model. It involves inspecting the residual plots for any collinearity or serial correlation between the residuals.

To check if the residuals generated by the fitted model are correlated or not to check suitability of the model the Multivariate Ljung-Box statistic is determined and interpreted. The Multivariate Ljung-Box statistic tests the hypothesis:

$$\begin{aligned} H_0 &: Q_k(1) = ... = Q_k(m) = 0 \\ H_1 &: Q_k(i) \neq 0 \text{ for } i \in \{1, ..., m\} \end{aligned}$$

The calculated Multivariate Ljung-Box statistics determines if there is autocorrelation and cross-correlations in the residuals matrix of the fitted Multivariate time series model r_l.

The test statistic is given by Equation (3.6).

$$Q_k(m) = T^2 \sum_{l=1}^{m} \frac{1}{T-1} tr(\hat{\Gamma}_l' \hat{\Gamma}_0^{-1} \hat{\Gamma}_l \hat{\Gamma}_0^{-1}) \tag{3.6}$$

where T is the sample size, k is the dimension of r_l, $tr(A)$ is the trace of matrix A, $\hat{\Gamma}_0$ is the cross covariance matrix at lag 0. and $\hat{\Gamma}_l$ is the cross covariance matrix at lag l.

the $Q_k(m)$ statistic follows a Chi$-$ distribution with mk^2 degrees of freedom. Statistical significance of $Q_k(m)$ were determinedusing the p-values, if p$-$value is less than the level of significance indicates serial correlation between the residuals, or otherwise.

However, the Multivariate Ljung-Box statistic can be represented in terms of cross correlation matrices $\hat{\rho}_l$ using the kronecker product $\otimes$ and matrix vectorization. The test statistic is then given by Equation (3.7).

$$Q_k(m) = T^2 \sum_{l=1}^{m} \frac{1}{T-1} \left(\hat{\rho}_0^{-1} \otimes \hat{\rho}_0^{-1}\right) b_l \tag{3.7}$$

where $b_l = vec(\hat{\rho}_l)$.

3.9 Forecasting using a VARMA(p, q) model

To forecast Kenya's inflation rate using the selected VARMA(p, q) model, say Z_t, the minimum mean-squared error criteria is used [16]. Forecasts can be determined up to a given h periods, using the estimated parameters obtained in the VARMA model estimation stage. Assume that the forecast origin is h, the series of fitted residuals denoted a_h and let F_h denotes the information available at h. For one-step ahead prediction, we have, from the model is as shown by Equation (3.8).

$$\begin{aligned} Z_h(1) &= E\{Z_{h+1}|F_h\} \\ &= \phi_0 + \sum_{i=1}^{p} \phi_i Z_{h+1-i} - \sum_{j=1}^{q} \theta_j a_{h+1+j} \end{aligned} \tag{3.8}$$

where p and q are non negative integers, ϕ_0 is a $k-$dimensional constant vector, ϕ_i and θ_j are $k * k$ constant matrices and the a_i, j is a sequence of independent and identically distributed random vectors with mean 0 and positive$-$ definite covariance matrix $\sum_a$. The associated forecast error is given by the Equation (3.9).

$$e_h(1) = Z_{h+1} - Z_h(1) = a_{h+1} \tag{3.9}$$

where $e_h(1)$ is the associated forecast error obtained from the difference of the observed value and the forecasted value.

Therefore to proceed with additional testing to determine forecast accuracy the MAPE and MAE methods were considered. The Table for interpretation of accuracy using MAPE forecasts(economic) as determined by Romadhon and Aprianto [17] is shown in the Table 4.5.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

To forecast Kenya's inflation rate, secondary data on past inflation rates for the years 1966 to 2018 was obtained from Kenya National Bureau of Statistics. To make the forecasts more accurate, annual data on crude oil importation price thought to have an association with annual inflation rates in Kenya was collected from the National Oil Corporation, Kenya (NOCK − 1966 to 2018).

4.2 Limitation of the study

The study was limited in that it relied on the accuracy of the secondary data collected by KNBS and NOCK. The model formulated was only useful to forecast the rate of inflation for a limited number of future time points to ensure accuracy as the p-values of the Ljung Box statistics in Table 4.4 indicate. It therefore may not be useful in giving accurate forecasts of more than 12 months ahead.

4.3 Product Moment Correlation

The linear correlation between the two time series of the crude oil prices and the Kenyan inflation rate were determined and the results are summarised in Table 4.1; Correlation results from the Table 4.1 it shows that there exists a negative statistically significant linear correlation between the two time series of the crude oil prices and the Kenyan inflation rate.

Table 4.1: Linear correlation results

	value
correlation coeffecient	$-\ 0.2143973$
t-statistic	-4.385
degrees of freedom (df)	399
p−value	0.00001
95% CI	$-0.306 to - 0.119$

4.4 Cross Correlation Analysis

The stationarity of both the crude oil prices and the Kenyan inflation rate time series were tested using the Dickey-Fuller test. A time series is said to be stationary if its statistical properties such as mean, variance, autocorrelation, etc. are all constant over time and not stationary if its statistical properties such as mean, variance, autocorrelation, etc. are not constant over time. The Kenyan Inflation rate time series, was not stationary until after applying a first difference to the series. The Dickey Fuller test results are Dickey-Fuller statistic $= -5.5902$, with the lag order $= 7$, p-value of the Dickey-Fuller statistic $= 0.01$ and the level of significance for the test is 0.05. Since Kenya's inflation rate time series was differenced first to make it stationary it is $I(1)$ stationary. The output of the differenced time series is as shown in Figure 4.1.

Similarly the price of imported crude oil time series is stationary after a first difference. Testing for stationarity using a DF Test gave the results: DF statistic $= -7.9771$, with the lag order $= 7$, p-value of the Dickey-Fuller statistic $= 0.01$ and the level of significance for the test is 0.05. It is also $I(1)$ stationary. The output of the differenced time series is as shown in Figure 4.2.

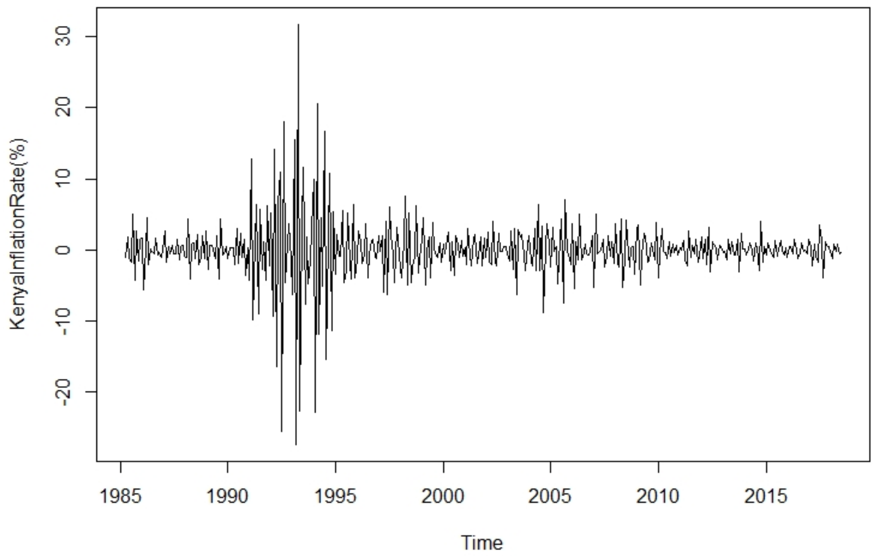

Figure 4.1: Differenced Inflation rates time series

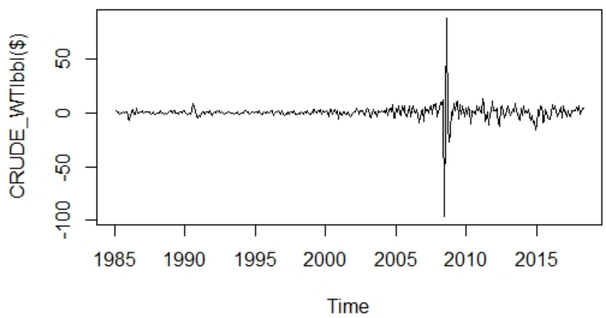

Figure 4.2: Differenced Crude oil import price time series

To determine objective 1, Cross-correlation analysis was used to determine whether there is a significant correlation between Kenya's inflation rate and the price of imported crude oil at different times (lags). The following Table 4.2 is the output of the cross-correlation analysis of the two-time series. The output for 6-lags is shown in Table 4.2. The table has the different p$-$values in the correlation matrix for the $-$MA lags of the time series in the columns and different p$-$values in the correlation matrix AR lags of the time series in the rows. From the Table 4.2 it shows that AR up to lag 2 have statistical significant correlations since the p-values are lower than the level of signficance and MA up to lag 1 have statistical significant correlations since the p-values are lower than the level of significance.

Table 4.2: CCM at lag 6 p-values

	0	1	2	3	4	5	6
0	0	0	0	0	0	0	0
1	0	0	0.0004	0.004	0.001	0.0002	0.018
2	0.017	0.384	0.191	0.001	0.257	0.221	0.795
3	0.589	0.127	0.768	0.746	0.562	0.702	0.705
4	0.953	0.920	0.951	0.853	0.679	0.815	0.859
5	0.989	0.986	0.797	0.801	0.954	0.999	0.991
6	0.999	0.993	0.987	0.999	0.999	0.983	0.288

The results shown in Table 4.1 indicate that there exists a significant correlation between Kenya's inflation rates and the imported crude oil price. This implies that both series have a statistical association which can be modeled and be used for forecasting.

4.5 Test for Cointegration

To test for Cointegration (Objective 2), since both Time series are now Stationary, then to test for cointegration of both series the; the Cointegrated Augmented Dickey Fuller (CADF) test was applied which involves the testing whether the residuals after applying the least squares regression method on the two time series are stationary at a given lag. First a regression model was fit to the Kenya's inflation rate and the imported crude oil price time series. The dependent vector is Kenya's Inflation rate while the independent vector is the imported crude oil price time series. After fitting the linear regression model on the Kenya's inflation rate and the imported crude oil price time series, the results were indicated in Table 4.3;

Table 4.3: LS Regression of the fitted residuals

	Dependent variable: inflation
croil	0.010 (standard error, 0.025)
Constant	−0.017 (standard error, 0.106)
Observations	402
R^2	0.0004
Residual Std. Error	2.127 (df = 400)
F Statistic	0.156 (df = 1; 400)
p-value	0.6933
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4.3 shows that the intercept α_1 for the cointegration equation is -0.01724 while the slope β for the cointegration equation is 0.01001.

An ADF test is then performed on the linear regression residuals in order to determine evidence of stationarity and hence cointegration. The residuals at lag 1 and lag 2 did not yield a statistically significant result. The p$-$value 0.6933 is higher than the value of alpha -0.05 therefore implying that the is no correlation between the residuals of the fitted model.

However, the residuals at lag 3 yielded a statistically significant result implying that a cointegration relationship between the Kenya's Inflation rate and the price of imported crude oil. The Dickey Fuller statistic was -8.3394 with a p-value of 0.01. This implies that the α_1 for the cointegration equation is -0.017 while the β is 0.010. The cointegration equation for the Kenya's Inflation rate and the price of imported crude oil which is a complete form of Equation 3.3 is therefore as shown in Equation 4.1 below

$$y_t = -0.017 + 0.010x_t + \hat{\epsilon}_t \tag{4.1}$$

While the individual series might be non$-$stationary, if there is a linear combination that is stationary, then the two time series are cointegrated. it was therefore necessary to test whether β is statistically different from 0 from the fitted regression model of the residuals. The hypothesis being tested is $H_0 : \beta = 0$ versus $H_1 : \beta \neq 0$ and if $\beta = 0$, then x_t and y_t are not cointegrated. The p-value for the ADF test for the residuals of the fitted model is 0.01. Since the p-value is less than the $\alpha = 0.05$ level of significance, the result is significant hence we reject the null hypothesis that the $\beta = 0$ and assume that β is not equal to 0, at the $\alpha = 0.05$ level is significance. Hence Kenya's rate of inflation is cointegrated to the price of imported crude oil.

The above analysis indicates that there is a statistically significant association between Kenya's Inflation rate and the price of imported crude oil and there exists a stationary linear combination of Kenya's annual inflation rate and the monthly price of imported crude oil. This implies that there exists a long-run relationship between the two time series. We can therefore fit a VARMA model to the data to explain the changes in Kenya's rate of inflation using changes in the price of imported crude oil.

4.6 Fitting a VARMA model

To fit an appropriate VARMA model (Objective 3), the results on the determination of the cross-correlation between Kenya's Inflation rate and the price of imported crude oil and the result showed a statistical association between the two time series as Table 4.2. The other result was for the test for cointegration between Kenya's Inflation rate and the price of imported crude oil which showed that there exists a statistically significant cointegration relationship between the time series of Kenya's Inflation rate and the price of imported crude oil. Therefore, this result implies that a VARMA model can be fit to explain the changes on the behavior of Kenya's rate of inflation to the changes in the price of imported crude oil.

To identify the order for the VARMA (p, q) model the SCM approach was applied. It involved determination of lags with significant canonical correlations from a correlation matrix. Several maximum lags for the model were tested $\{0, 1 \text{ and } 2\}$ for the AR and MA terms and the best model was that with a maximum of $\{2\}$ lags.

From the results, the model with the most significant canonical correlations was the VARMA $(2, 1)$ model. Therefore, from the SCM results the VARMA $(2, 1)$ model is best for the data.

The Cross Correlation Matrix method can equally be applied to determine the order of the VARMA (p, q) model. It considers the p-values of the canonical correlations. Lags with statistically significant correlations are included in the proposed model to fit the data and the model with the lowest Multivariate Ljung Box statistic is chosen. From Table 4.2 and the result from the SCM order identification results above, the VARMA $(2, 1)$ is the best fit for the data. Since the order of the model has been determined the next stage involved the determination of the coefficients of the selected order for the model.

4.7 Fitting the VARMA $(2, 1)$ model

Using R software to fit the VARMA $(2, 1)$ model to the data to give estimates of 14 parameters to be used for forecasting, as shown in Table 1 in the appendix.

The zeroes in the VARMA $(2, 1)$ model imply that the coefficient of the lag value is zero in the determined equation, on the other hand if they are not zeroes then it implies that the coefficient of the lag value is zero in the determined equation. Multiplying out the matrices, the result would be two equations to determine a forecast for either Kenyas inflation rate at time t or the forecasted price of imported crude oil.

Since the VARMA model when multiplied out should give two equations one for the prediction of or Kenya's inflation rate and the other for the prediction of the price of imported

oil. If the diagonal matrix is zero then it would imply that there is statistical association between the two series.
However, these included some non-zero statistically insignificant parameters. Since, these parameters do not contribute to the accuracy of the fitted model, they are set to zero and the model is refined without loss of generality. After refining the model by dropping statistically insignificant coeffecients, the number of parameters reduces to 7 as shown in equation 4.7.

The estimated model takes the form:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} 1.38 & 0.00 \\ 0 & 1.82 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} -0.396 & 0.00 \\ 0.00 & -0.831 \end{bmatrix} \begin{bmatrix} y_{t-2} \\ x_{t-2} \end{bmatrix} + \begin{bmatrix} 0.00 & 0.00 \\ 0.00 & 0.68 \end{bmatrix} \begin{bmatrix} \epsilon_t \\ \epsilon_t \end{bmatrix}$$

A model diagnostic check was performed to determine the validity of the fitted VARMA $(2,1)$ model. The calculated Multivariate Ljung Box statistic $Q_k(i)$ in Table 4.4 are calculated using the equation 3.6. The tested hypothesis for the model diagnostic check is;

$$\begin{aligned} H_0 &: Q_k(1) = ... = Q_k(m) = 0 \\ H_1 &: Q_k(i) \neq 0 \text{ for } i \in \{1, ..., m\} \end{aligned}$$

Consider the Table 4.4 of the Multivariate Ljung box statistics for the fitted model at different lags. The results show that the Multivariate Ljung box at lags less than 12 are not statistically different from 0 implying that the fitted model will work well while forecasting for the first 11 lags after the last observed observation.

The determined Multivariate Ljung Box statistics M_q in Table 4.4 implies that we do not reject the null hypothesis implying there is no autocorrelation and cross-correlations in the residuals matrix of the fitted Multivariate time series model. The fitted VARMA $(2,1)$ is adequate to forecast inflation rate for Kenya.

Table 4.4: Ljung−Box Statistics

	m	$Q(m)$	df	$p-value$
[1,]	1.00	1.47	4.00	0.83
[2,]	2.00	10.13	8.00	0.26
[3,]	3.00	12.27	12.00	0.42
[4,]	4.00	17.56	16.00	0.35
[5,]	5.00	20.07	20.00	0.45
[6,]	6.00	29.55	24	0.20
[7,]	7.00	31.29	28	0.30
[8,]	8.00	38.27	32.00	0.21
[9,]	9.00	39.68	36.00	0.31
[10,]	10.00	42.09	40.00	0.38
[11,]	11.00	49.47	44.00	0.26
[12,]	12.00	126.15	48.00	0.00
[13,]	13.00	128.81	52.00	0.00
[14,]	14.00	140.30	56.00	0.00
[15,]	15.00	146.30	60.00	0.00
[16,]	16.00	147.66	64.00	0.00
[17,]	17.00	152.09	68.00	0.00
[18,]	18.00	153.61	72.00	0.00
[19,]	19.00	156.02	76.00	0.00
[20,]	20.00	156.58	80.00	0.00
[21,]	21.00	162.04	84.00	0.00
[22,]	22.00	165.58	88.00	0.00
[23,]	23.00	166.16	92.00	0.00
[24,]	24.00	175.82	96.00	0.00

4.8 Forecasting Kenya's Inflation Rates

To forecast Kenya's inflation rate (Objective 4), the fitted VARMA $(2, 1)$ model was used see the refined equation 4.7, a forecast of Kenya's Inflation rates for 12 ahead forecasts with 0 steps backwards so as to determine the accuracy of the forecasts. The forecast errors were determined as shown in Table 4.6, implying that the model was adequate for the prediction of Kenya's inflation rate up to 11 steps ahead. This is because Table 4.4 has the Multivariate Ljung box statistics at steps less than 12 are not statistically different from 0. Consider equation 4.7, it implies that the cointegration as was determined is uni−directional indicating that the price of imported crude oil explains some behavior changes of Kenya's Inflation rate (statistically significant association) but Kenya's Inflation rate does not necessarily explain the price of imported crude oil. This is because the coeffecients of the matrices associated with the inflation prices while determining the forecast of the price of imported crude oil are not statistically different from 0.
Table 2 shows the predicted imported crude oil price for June 2018 upto May 2019 together with the predicted Kenya's inflation rate for the same period. The Table 2 also shows the standard error for both the imported crude oil price and Kenya's rate of inflation. Considering the Multivariate Ljung Box statistics determined 4.4 the fitted model is sufficient to forecast Kenya's Inflation rate.
In VARMA analysis do not require goodness of fit tests since it is covered when testing for the model diagistics in determination of the suitability of the fitted model. Therefore to proceed with additional testing to determine forecast accuracy the MAPE and MAE methods were considered. The Table for interpretation of accuracy using MAPE forecasts(economic) as determined by Romadhon and Aprianto [17] is shown in the Table 4.5;

Table 4.5: MAPE value for prediction evaluation

MAPE Value	**Prediction Accuracy**
MAPE $\leq$ 10%	High
10% < MAPE $\leq$ 20%	Good
20% < MAPE $\leq$ 50%	Reasonable
MAPE > 50%	Low

Table 4.6 shows a comparison between the forecast values between August 2018 and the observed values for the same period. The last column has the absolute error for each of the forecasts.

Table 4.6 shows the Mean Absolute Percentage Error (MAPE) which is a measure of the forecasts' accuracy is 13.48% which is a good accuracy forecast implying that the forecasts of the fitted model are accurate. However, since MAPE is asymmetric i.e. it is bounded by 100% of the observed value then forecasts that are higher than the observed values are highly penalized giving a higher MAPE value. A better test for the fitted model will be the interpretation of the Mean Absolute Error (MAE) which is also a measure of the forecasts accuracy which gives a mean value of 0.66 % for Kenya's forecasted inflation values. This result implies that the fitted VARMA $(2,1)$ model can be used to predict future Kenya's rate of inflation accurately.

In Summary, the results from this section indicate that there exists a crosscorrelation

Table 4.6: observed and forecasted values comparison

Period	Observed inflation	Predicted inflation	MAPE	MAE
Aug 2018	4.04	4.564	15.20%	0.61
Sep 2018	5.70	4.997	12.33%	0.70
Oct 2018	5.53	5.368	2.93%	0.16
Nov 2018	5.58	5.757	3.17%	0.18
Dec 2018	5.71	6.155	7.79%	0.45
Jan 2019	4.70	6.554	39.45%	1.85
Average			13.48%	0.66%

between Kenya's inflation rate and the price of imported crude oil see Table 4.2. The results also showed that tthere exists a cointegration association between the two time series. Since there existed a cointegrated association between the two time series a VARMA $(2, 1)$ was fitted to model the association between the two time series see equation 4.7. Using the fitted VARMA $(2, 1)$ model see the refined equation 4.7, a forecast of Kenya's Inflation rates for 12 ahead forecasts with 0 steps backwards were determined. The accuracy of the forecasts were determined using the MAPE approach and the conclusion was that they were good forecasts see Table 4.5 hence can be used for planning purposes.

CHAPTER 5

SUMMARY, CONCLUSION AND RECOMMENDATIONS

5.1 Summary

The research objectives were all met. The annual price of imported crude oil and Kenya's annual Inflation rate are non-stationary time series. However, the time series become stationary after taking the first difference as shown in Table 4.1. The price of imported crude oil and Kenya's inflation rate are correlated as the cross correlation matrix shows in Table 4.2. This implies that there exists a statistical relationship between Kenya's rate of inflation and price of imported crude oil and possible statistical model to explain the association between the two time series.

The research set out to determine if there is cointegration between the price of imported crude oil and Kenya's Inflation rate. Table 4.3 showed that there exists a cointegrating association between Kenya's Inflation rate and the price of imported crude oil. This indicates that both series explain each other at some point and suggests a statistical association that can be modelled.

Since there existed a cointegrated association between the two time series a VARMA $(2, 1)$ was fitted to model the association between the two time series see equation 4.7. Using the fitted VARMA $(2, 1)$ model see the refined equation 4.7, a forecast of Kenya's Inflation rates for 12 ahead forecasts with 0 steps backwards were determined.

5.2 Conclusion

The research ascertained that there exists a moderate statistical association between the price of imported crude oil series and Kenya's rate of inflation which can be modelled using the multivariate time series model VARMA $(2, 1)$ model. The fitted model has the least forecast error among the suggested models. The fitted model gives accurate forecasts since the Mean Absolute Error of 0.66% see table Table 4.6 models the association between the price of imported crude oil series and Kenya's rate of inflation as both series are correlated and cointegrated.

5.3 Recommendation

The research recommends that this model is appropriate for forecasts of less than a year for Kenya's rate of inflation as shown in Table 2 since the forecasts are only accurate up to the 11^{th} step forecast. There is also room to add more variables to the equation but this will lower the accuracy of the model (many parameters to be estimated with few data points available to estimate them).

The government can use the model to forecast inflation for the next year(12 months) and develop policies that can prevent or cushion the citizens from experiencing the negative effects of inflation.

The forecast errors were determined as shown in Table 4.6, implying that the model was adequate for the prediction of Kenya's inflation rate up to 11 steps ahead. This is because Table 4.4 has the Multivariate Ljung box statistics at steps less than 12 are not statistically different from 0.

REFERENCES

[1] African Development Bank, A. D. (2011). Inflation Dynamics in selected East African countries: Ethiopia, Kenya, Tanzania and Uganda, African Development Bank, Tunisia (Report)

[2] Bi-Juan, L. (2009). The Changing Effects of Oil Price Changes on Inflation. (Report)

[3] Fannoh, R. (2014). Modeling Inflation Rates in Liberia; SARIMA Approach. Pan African University Institute for Basic Sciences Technology and Innovation. (Thesis)

[4] Gathingi, V.W., 2014. Modeling Inflation in Kenya using ARIMA and VAR Models, MSc Thesis, University of Nairobi. (Thesis)

[5] Gil-Alana, Luis A. and Mudida, Robert, 2017. (Book) *CPI and inflation in Kenya. Structural breaks, non-linearities and dependence*, International Economics, Elsevier, vol. 150(C), pages 72-79. (Book)

[6] Kargi, B., (2014) *The Effects of Oil Prices On Inflation and Growth: Time Series Analysis In Turkish Economy For 1988:01 - 2013:04 Period*. International Journal of Economics and Research , Vol. 2, No. 5 (March 2014): pp. 29-36. (Journal)

[7] Lidiema, C. (2017). *Modelling and Forecasting Inflation Rate in Kenya Using SARIMA and Holt–Winters Triple Exponential Smoothing*. American Journal of Theoretical and Applied Statistics, 161-169. (Journal)

[8] Mihaela, S. (2013). The use of VARMA models in forecasting macroeconomics indicators. Economics and Sociology, 94-102. (Journal)

[9] Neely, C. J. (2015). How much do oil prices affect inflation. Economic SYNOPSES. `https://www.researchgate.net/publication/324514846\_How\_Much\_Do\_Oil\_Prices\_Affect\_Inflation` (Online Journal)

[10] Neely, J. C., Rapach, E. D. (2011). *International comovements in inflation rates and country characteristics.* Journal of International Money and Finance, 1471 - 1490.(Journal)

[11] Olofin, S. O., Salisub, A. A. (2016). Modeling oil price-inflation nexus: The role of asymmetries and structural breaks. Ibadan: Center for Econometric and Allied Research (CEAR), University of Ibadan. (Journal)

[12] Owuor, R. M., Kageni, E. (2018). Energy. Nairobi, Kenya: Global Legal Insights. (Journal)

[13] Pettinger, T. (2018). What are the effects of a rise in the inflation rate? Retrieved from Economics:
`https://www.economicshelp.org/blog/140824/economics/`
`what-are-the-effects-of-a-rise-in-the-inflation-rate/` (Journal)

[14] Veena Yesikar, Rajendra Kumar Mahore, Sanjay Dixit, Geeta Shivram, Sachin Parmar,Chakresh Jain, (2015) *A Study to Evaluate Inflation and Price Rise: Effect on Common Man.* Journal of Evolution of Medical and Dental Sciences, Vol.4, Issue 30, April 13; Page: 5172 - 5178(Journal)

[15] Zhou, Yang (2017). Modelling Swedish Inflation Using Market Data. Report by the KTH, School of Engineering Sciences (SCI) , Mathematics Department.(Report)

[16] Tsay, S Ruey (2014).Multivariate Time Series Analysis With R and Financial Applications. John Wiley & Sons, Inc.(Book)

[17] Gustriansyah, Rendra (2018). *Single Exponential Smoothing Method to Predict Sales Multiple Products.* Journal of International Series on Integrated Science and Technology. Vol.3, Issue 2, October 20; Page: 176 - 178 (Journal)

APPENDIX

Figures

Table 1: Parameter estimates

```
Estimates in matrix form:
Constant term:
Estimates:  0.9647567 0.1940185
AR coefficient matrix
AR( 1 )-matrix
        [,1]    [,2]
[1,] 1.4759 -0.0547
[2,] 0.0132  1.8293
AR( 2 )-matrix
         [,1]    [,2]
[1,] -0.4905  0.0293
[2,] -0.0141 -0.8436
MA coefficient matrix
MA( 1 )-matrix
         [,1]    [,2]
[1,] 0.11549 -0.0623
[2,] 0.00431  0.7019

Residuals cov-matrix:
             [,1]        [,2]
[1,] 14.61585482 -0.04617365
[2,] -0.04617365  4.04143527
```

Table 2: Standard errors of the forecasts

```
Predictions at origin  403
      CRUDE_WTIbbl KenyaInflationRate
 [1,]        69.69              4.654
 [2,]        69.26              4.997
 [3,]        68.78              5.368
 [4,]        68.30              5.757
 [5,]        67.82              6.155
 [6,]        67.35              6.554
 [7,]        66.89              6.949
 [8,]        66.44              7.335
 [9,]        66.00              7.707
[10,]        65.56              8.062
[11,]        65.14              8.398
[12,]        64.73              8.714
Standard errors of predictions
        [,1]  [,2]
 [1,]  3.838 2.014
 [2,]  6.548 3.049
 [3,]  8.757 3.933
 [4,] 10.578 4.720
 [5,] 12.112 5.427
 [6,] 13.431 6.059
 [7,] 14.585 6.620
 [8,] 15.608 7.115
 [9,] 16.527 7.548
[10,] 17.358 7.923
[11,] 18.115 8.246
[12,] 18.809 8.520
```