
Modelling and Forecasting Inflation Rate in Kenya Using SARIMA and Holt-Winters Triple Exponential Smoothing

Caspah Lidiema

Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

Email address:

caspahlidiema@gmail.com

To cite this article:

Caspah Lidiema. Modelling and Forecasting Inflation Rate in Kenya Using SARIMA and Holt-Winters Triple Exponential Smoothing. *American Journal of Theoretical and Applied Statistics*. Vol. 6, No. 3, 2017, pp. 161-169. doi: 10.11648/j.ajtas.20170603.15

Received: April 5, 2017; **Accepted:** April 13, 2017; **Published:** May 25, 2017

Abstract: In this paper, two models of forecasting are used the Box-Jenkins procedure employing the SARIMA and the Holt-Winters triple exponential smoothing. Published Consumer Price Index Data from Kenya National Bureau of Statistics (KNBS) for the period November 2011 to October 2016 was used. This paper we equate the forecasted values of both the models and we choose the best model based on the least mean Absolute square error (MASE), mean absolute error (MAE) and mean absolute percentage error (MAPE). The three step model building for Box-Jenkins was first employed, followed by the Hold-Winters triple exponential smoothing. The study found the SARIMA Model was a better model than the Holt-winters triple exponential smoothing as per the obtained results using MASE, MAE and MAPE.

Keywords: Inflation, CPI, Holt-Winters, Triple Exponential Smoothing, ARIMA, SARIMA

1. Introduction and Motivation

Inflation is the rate at which the general level of prices for goods and services in an economy rises and consequently affecting the purchasing power of currency to fall. In Kenya Consumer Price Index is used in calculation of Inflation rates on yearly basis even though data is collected and computed monthly basis by the KNBS. KNBS, defines Consumer Price Index (CPI) is defined as a measure of the weighted aggregate change in retail prices paid by consumers for a given basket of goods and services. Price changes are measured by re-pricing the same basket of goods and services at regular intervals, and comparing aggregate costs with the costs of the same basket in a selected base period Price data for constructing the indices are collected by Kenya National Bureau of Statistics through a survey of retail prices for consumption of goods and services. The percentage change of the CPI over a one-year period is what is usually referred to as inflation.

In 2016 the rate has been almost stable with the highest being in January at 7.78% and the lowest being in May at 5.01%. November 2016, inflation was 6.68% which was the highest since February 2016 (7.08%). According to Otu, A. O, et al [1], The effect of inflation is highly considered as a crucial issue for a country. Inflation problems could cause living conditions in a country much harder to lot of people.

People who are living on fixed income suffer most as when prices of commodities rise, since these people cannot buy as much as they could previously. Monetary policy consists of decisions and actions taken by the Central Bank of Kenya (CBK) to ensure that the supply of money in the economy is consistent with growth and price objectives set by the government. The objective of monetary policy is to maintain price stability in the economy. Price stability refers to maintenance of a low and stable inflation.

Inflation is caused by many factors including micro economic factors and even natural factors like rain and drought. According to CBK [16], periods of drought or excessive rain, can cause the prices of food to increase, leading to an increase in the inflation rate. International factors like increases or decreases in oil prices can also lead to changes in inflation reflecting movements in energy and transport costs. Depreciation in the exchange rate against the major currencies can also cause inflation since Kenya is a net importer of goods. Inflation can also be caused by factors that influence the demand for goods and services, like the amount of money ordinary people have available to spend.

The Central Bank's monetary policy decisions are made to maintain a low and stable inflation rate over time, which is an indication of price stability. Inflation is a general increase in price levels over time. It is based on the prices of various

consumer goods and services, which are evaluated and statistically represented in the Consumer Price Index (CPI). The month-on-month (or year-on-year) inflation rate is determined by comparing the CPI for a particular month to the CPI of that same month in the previous year.

Inflation is a key factor in helping people make sound decisions on financial matters. Deflation on the other hand can affect the economy by impacting on profitability of companies and hamper investor confidence. If there is a general decrease of prices over time due to a collapse in demand or increased supply of goods and services, then there is deflation. Hence there is need for the inflationary rates to be stable to enable a more stable economy through price stability that will drive economic growth.

In Kenya inflation, is controlled through the Monetary policy whose objective is to spur economic growth. Monetary policy decisions are made by the Monetary Policy Committee (MPC). The MPC meets at least once every two months and reviews data and analysis from various sources including the Central Bank Departments enabling it to decide on any action to maintain or vary its stance.

1.1. Justification of the Study

Not only have many Forecasting methods been developed in the past but equally many methods of measuring forecasting accuracy have been developed. Hyndman R. J. and Koehler A. B. [2], for example studied and compared all measures accuracy and settled on mean Absolute Scaled Error (MASE) as the best measure of accuracy on forecasting.

Time series researchers have also compared several models of forecasting with a view of determine which is a better model for forecasting depending on the nature of data and industry. For example, Virginia Gathing [3] modelled inflation in Kenya using ARIMA and VAR. Equally Ingabire. J and Mung'atu. J. K. [4] compared ARIMA and VAR models in forecasting inflation rate in Rwanda. In the IMF Working paper series, Tim, C. and Dongkoo, C. [5] studied modelling and forecasting inflation in India using Bivariate VAR. They found out that broad money supply, exchange rate and import prices are relevant indicators that affect inflation especially in the manufacturing sector in India.

Other studies that concentrated on inflation include Otu et al [1], who discussed the application of SARIMA Models in Modelling and Forecasting Nigeria's Inflation Rates. while, Uwilingiyimana C., Mungatu. J and Harerimana J. [6] conducted a study on forecasting inflation in Kenya using two models, the ARIMA (1, 1, 12) and GARCH (1, 2) and a combination of the two model ARIMA(1, 1, 12)-GARCH(1, 2).

Despite having many researcher compare models and use different models in forecasting most researches have concentrated on other industries like Motor vehicle, telecommunication, medical and biological fields. Few researches that modelled inflation did not compare Box-Jenkins Seasonal ARIMA and Holt-Winters Triple Exponential Smoothing in Kenya.

The objective of this research is to use a 60-month CPI data

from Kenya National Bureau of Statistics to test which between SARIMA and Holt-Winters Triple Exponential Smoothing is best suited to forecast inflation in Kenya in its forecast of variables by use of three measures of Accuracy MASE, MAE and MAPE.

The rest of the study is structured as follows: Section 2 highlights the empirical literature whereas, Section 3 presents the methodology. Section 4 will report the results of the empirical analysis and section 5 will discuss and conclude the study.

1.2. Research Objective

This paper aims to predict the value of CPI in Kenya twelve months ahead and to establish the best forecasting method for forecasting inflation to help in proper economic and financial management.

1.3. Scope of the Study

This research uses secondary data based on facts and figures collected and by the Kenya National Bureau of Statistics (KNBS) on their website to forecast CPI data which is the measure of inflation in Kenya. The data from November 2011 to October 2016 was used.

2. Literature Review

Otu et al [1], discussed the application of SARIMA Models in Modeling and Forecasting Nigeria's Inflation Rates. They employed Box and Jenkins to build the Autoregressive Integrated Moving Average (ARIMA) monthly inflation rates for the period November 2003 to October 2013 with a total of 120 data points. They found that the Seasonal ARIMA (1,1,1)(0,0,1)₁₂ was the best model to forecast Nigeria's inflation rate.

Gathing, V. [3] modelled inflation in Kenya using ARIMA and VAR models using data from January 2005 to June 2013. When the author compared the two models, VAR was a better model than the ARIMA (1, 1, 0) due to the smaller errors for RMSE, MAE and MAPE. However, she concluded that despite ARIMA using univariate historical consumer pricing data to model inflation, ARIMA model (1, 1, 0) resulted as the best model showing strong evidence of substantial inflation inertia even with the exclusion of independent variables

Ingabire. J and Mung'atu. J. K. [4] compared ARIMA and VAR models in forecasting inflation rate in Rwanda. After carrying out all the necessary diagnostic checks, the study indicated that ARIMA (3, 1, 4) model was better than VAR model in predicting inflation in Rwanda. However, thy further stated that ARIMA model may be efficient in forecasting short term periods.

Uwilingiyimana, C et al. [6], conducted a study on forecasting inflation in Kenya using two models, the ARIMA (1, 1, 12) and GARCH (1, 2) and a combination of the two model ARIMA (1, 1, 12)-GARCH (1, 2). The study revealed the combination between ARIMA (1, 1, 12)-GARCH (1, 2) model provided the best and improved results for estimating and

forecasting accuracy compared to the other forecasting models.

Jere, S. and Siyanga, M. [7], studied forecasting of inflation rate in Zambia using Holts exponential smoothing. In this paper, the use of Exponential smoothing uses a linear combination of the previous values of the given data to model and predict future values. Even though both models gave almost similar results ARIMA model (12, 1, 0) was adequate enough than holts exponential smoothing. They further concluded that Holt's exponential smoothing is as good as an ARIMA model due to the smaller deviations in the MAPE and RMSE but the Holt's exponential smoothing model is less complicated since you do not require specialised software to implement it as is the case for ARIMA models.

Hamidreza M. and Leila S.[8] examined a seasonal long memory process, denoted as the SARFIMA to study and predict the Iran's oil supply. They fitted both SARIMA and SARFIMA models, and estimated the parameters using CSS method. The results of their analysis indicated that the best model was SARFIMA (0,1,1)(0,-0.199,0)₁₂ which was used to predict the quantity of oil supply in Iran.

Puthran et al [9], studied and forecasted Indian Motor Cycle industry by comparing SARIMA and Holt Winters models. They found out that even though both models were good, Holt-Winters method was a better model than SARIMA model due to the minimum MSE, MAE, and MAPE values when compared with SARIMA.

Udom. P and Phumchusri. N [10], compared the application of three forecasting methods on the amount of the sales volume for plastic distributor in Thailand for the period between January 2004 and December 2012. They compared, the ARIMA method, Moving average method and Holt's and Winter exponential method. They employed five data sets of raw material from plastic distributor and their results showed that the ARIMA model was a better model when compare with other methods by using MAPE (Mean Absolute Percentage Error).

Hyndman R. J. and Koehler A. B. [2], in their paper titled "Another look at measures of forecast accuracy" analysed and compared measures of accuracy of univariate time series forecasts. They analysed Scale-dependent measures, Measures based on percentage errors, Measures based on relative errors, Relative measures, and Scaled errors. They concluded and proposed that scaled errors become the standard measure for forecast accuracy The Mean Absolute Scaled Error (MASE). They however, stated that depending on the type of data mean absolute error (MAE) and Mean Absolute Percentage Error (MAPE) could be used. In this Paper we shall employ all the three methods for testing the model accuracy. They argued that MAE should be used if the data series are on the same scale on the other hand MAPE should be used when the data is all positive and much greater than zero since its very simple to use and interpret.

3. Methodology

3.1. The Box-Jenkins Procedure

The Box-Jenkins procedure is concerned with fitting an

ARIMA model to data. It has four parts: identification, estimation, verification or diagnostic checking and forecasting.

3.1.1. Model Identification

Normally we assume the data series is stationary. The initial model Identification is to estimate the sample autocorrelations function (ACF) and partial autocorrelations function (PACF) and compare the resulting ACF and PACF with expected or theoretical ACF and PACF derived already.

It can also be detected from an autocorrelation plot. Specifically, non-stationarity is often indicated by an autocorrelation plot with very slow decay.

Box and Jenkins recommend the differencing approach to achieve stationarity. In our data the series was not stationary and therefore we had to differentiate once to make it stationary. This was mainly done in order to Straighten out trends and to reduce heteroscedasticity (produce approximately uniform variability in the series over the sample range. If on plotting the data is not stationary on variance, the series has to be transformed through with the objective of making the series stationary on both mean and variance.

3.1.2. Model Estimation

Estimating the parameters for Box-Jenkins models involves numerically approximating the solutions of nonlinear equations. In our Case we employ R software to estimate our model by using Maximum likelihood estimation.

3.1.3. Diagnostic Checks

The model having been identified and the parameters estimated, diagnostic checks are then applied to the fitted model, Box G. E. P et al [11]. In this paper we shall use the residuals methods in our diagnostic checks.

3.1.4. Forecasting

We evaluate forecasts using both subjective and objective means. The subjective examination looks for large errors and/or failures to detect turning points The analyst may be able to explain such problems by unusual unforeseen or unprovided for events Great care should be taken to avoid explaining too many of the errors by strikes etc In an objective evaluation of a forecast we may use various standard measures If x_i is the actual datum for period $\{i\}$ and $\{fi\}$ is the forecast then the error is defined as.

In this study, the selected SARIMA model was used to forecast the mean monthly CPI for the period November-2016 to October-2018 by using the observed data of the period November-2011 to October-2016.

3.1.5. General ARIMA Process

The autoregressive moving average process, ARMA(p, q), is defined by

$$X_t - \sum_{r=1}^p \phi_r X_{t-r} = \sum_{s=0}^q \theta_s \varepsilon_{t-s} \quad (1)$$

where again ε_t is white noise. This process is stationary for appropriate ϕ, θ .

if we consider two models

$$X_t = \phi X_{t-1} + \varepsilon_t \tag{2}$$

$$Y_t = X_t + \eta_t \tag{3}$$

where X_t is unobserved, Y_t is observed and ϕ_t and η_t are independent white noise sequences. Note that X_t is AR(1). We can write

$$\begin{aligned} \varepsilon_t &= Y_t - \phi Y_{t-1} \\ &= (X_t + \eta_t) - \phi(X_{t-1} + \eta_{t-1}) \\ &= (X_t - \phi X_{t-1}) + (\eta_t - \phi \eta_{t-1}) \\ &= \varepsilon_t + \eta_t - \phi \eta_{t-1} \end{aligned}$$

Now ξ_t is stationary and $cov(\xi_t, \xi_{t+k}) = 0, k \geq 2$. As such, ξ_t can be modelled as a MA(1) process and Y_t as ARMA(1, 1).

In our Case, if the CPI data denoted by Y_t is process is not stationary, we will look at the first order difference process

$$X_t = \nabla Y_t = Y_t - Y_{t-1} \tag{4}$$

or the second order differences

$$X_t = \nabla^2 Y_t = \nabla(\nabla Y_t) = Y_t - 2Y_{t-1} + Y_{t-2} \tag{5}$$

This can continue until we find the difference with the expected outcome.

If we ever find that the differenced process is a stationary process we can look for a ARMA model of that. The process Y_t is said to be an autoregressive integrated moving average process, ARIMA(p, d, q), if $X_t = Y_t$ is an ARMA(p, q) process. The general ARIMA(p, d, q) is denoted by

$$(B)(1-B)^d X_t = \theta(B)Z_t \tag{6}$$

3.1.6. The SARIMA Process

Chatfield, C. [12], states that If the series is seasonal, with s time periods per year, then a seasonal ARIMA (abbreviated SARIMA) model may be obtained as a generalization of an ARIMA. Let B^s denote the operator such that

$$B^s X_t = X_{t-s} \tag{7}$$

Then the seasonal differencing will be written as

$$(1-B^s)X_t = (X_t - X_{t-s}) \tag{8}$$

Given that our data is monthly with 12 months per year(s=12) the seasonal difference will be;

$$(1-B^{12})X_t = (X_t - X_{t-12}) \tag{9}$$

The main objective of seasonal differencing will be to remove seasonal trend and season random walks in our data. Chatfield C. [12], further elaborates that A seasonal autoregressive term, for example, is one which depends

linearly on. A SARIMA model with non-seasonal terms of order (p, d, q) and seasonal terms of order (P, D, Q) is abbreviated a SARIMA(p, d, q)(P, D, Q)_s model and may be written

$$\phi(B)\Phi(1-B)^d(1-B^s)^D X_t = \theta(B)\Theta(B^s)Z_t \tag{10}$$

where ϕ and Θ denote polynomials in B^s of order P, Q respectively. One model, which is particularly useful for seasonal data, is the SARIMA model of order (0,1,1)(0,1,1)₁₂. Shumway R. H and Stoffer D. S [13], confirm that The multiplicative seasonal autoregressive integrated moving average model, or SARIMA model, of Box and Jenkins (1970) is given by;

$$\phi_p(B^s)\phi(B)\nabla_s^D \nabla^d x_t = \alpha + \Theta_Q(B^s)\theta(B)Z_t \tag{11}$$

where Z_t is the usual Gaussian white noise process. The general model is denoted as ARIMA(p, d, q(P, D, Q)_s). The ordinary autoregressive and moving average components are represented by polynomials $\phi(B)$ and $\theta(B)$ of orders p and q, respectively and the seasonal autoregressive and moving average components by $\Phi_p(B^s)$ and $\Theta_Q(B^s)$ of orders P and Q and ordinary and seasonal difference components by $\nabla^d = (1-B)^D$ and $\nabla_s^D = (1-B^s)^D$

Since our data for CPI is monthly data, s = 12, hence our equation will be written as.

$$(1-B)(1-B^{12})X_t = (1+\theta B)\Theta(B^{12})Z_t \tag{12}$$

Under this method we shall employ monthly CPI data with 12 seasons per year (s = 12), where the first order AR(1) model will use X_{t-12} to predict X_t , while the seasonal first order MA(1) will use Z_{t-12} as its predictor.

3.2. Holt-Winters Tripple Exponential Mmoothing

Triple exponential smoothing takes into account seasonal changes as well as trends. This was an extension of Holt's method to capture seasonality. The Holt-Winters seasonal method comprises the forecast equation and three smoothing equations - one for the level l_t , one for trend b_t , and one for the seasonal component denoted by s_t , with smoothing parameters α , β^* and γ . Holt-Winter's exponential smoothing model is used for data that exhibit both trend and seasonality. Seasonality is the tendency of time-series data to exhibit behavior that repeats itself every t periods. The term season is used to represent the period of time before behavior begins to repeat itself. The Holt-Winters method has two versions, additive and multiplicative, the use of which depends on the characteristics of the particular time series.

3.2.1. Holt-Winters Additive Method

$$l_t = \alpha(y_t - s_{t-m} + (1-\alpha)(l_{t-1} + b_{t-1})) \tag{13}$$

Level

$$b_t = \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1} \tag{14}$$

Growth

$$s_t = \gamma(y_t - l_{t-1} - b_{t-1}) + (1 - \gamma)s_{t-m} \tag{15}$$

Seasonal

$$\hat{y}_{t+h|t} = l_t + hb_t + s_{t-m+h_m^+} \tag{16}$$

Forecast

where m is the length of seasonality in months, l_t represents the level of the series b_t denotes growth, s_t is the seasonal component, $\hat{y}_{t+h|t}$ is the forecast for h periods ahead, and $h_m^+ = [(h-1) \text{mod} m] + 1$. The parameters $(\alpha, \beta^*$ and $\gamma)$ are usually restricted to lie between 0 and 1.

$$s_t = \gamma^*(y_t - l_t) + (1 - \gamma^*)s_{t-m} \tag{17}$$

if we substitute for l_t from the seasonal smoothing equation (15) for the equation for the level of the component form we get

$$s_t = \gamma^*(1 - \alpha)(y_t - l_{t-1} - b_{t-1}) + [1 - \gamma^*(1 - \alpha)]s_{t-m} \tag{18}$$

which is identical to the smoothing equation for the seasonal component $\gamma = \gamma^*(1 - \alpha)$.

The Error correction term of the smoothing equation is given by

$$\begin{aligned} l_t &= l_{t-1} + b_{t-1} + \alpha e_t \\ b_t &= b_{t-1} + \alpha \beta e_t \\ s_t &= s_{t-m} + \alpha e_t \end{aligned} \tag{19}$$

where $e_t = y_t - (l_{t-1} + b_{t-1} + s_{t-m}) = y_t - \hat{y}_{t|t-1}$ are one-step training forecast errors.

3.2.2. Holt-Winters Multiplicative Method

$$l_t = \alpha \frac{y_t}{s_{t-m}} + (1 - \alpha)(l_{t-1} + b_{t-1})t - 1 \tag{20}$$

Level

$$b_t = \beta^*(l_t - l_{t-1}) + (1 - \beta^*)b_{t-1} \tag{21}$$

Growth

$$s_t = \gamma \frac{y_t}{l_{t-1} + b_{t-1}} + (1 - \gamma)s_{t-m} \tag{22}$$

Seasonal

$$\hat{y}_{t+h|t} = (l_t + hb_t)s_{t-m} + h_m^+ \tag{23}$$

Forecast

See (23). and the error correction component is given by;

$$l_t = l_{t-1} + b_{t-1} + \alpha \frac{e_t}{s_{t-m}} \tag{24}$$

$$b_t = \beta_{t-1} + \alpha \beta^* \frac{e_t}{s_{t-m}} \tag{25}$$

$$s_t = s_t + \gamma \frac{e_t}{l_{t-1} + b_{t-1}} \tag{26}$$

where $e_t = y_t - (l_{t-1} + b_{t-1})s_{t-m}$ and α , β , and γ are constants that must be estimated in such a way that the MASE of the error is minimized.

4. Data Analysis

Our analysis will use two methods Box-Jenkins which is sometimes referred to as ARIMA model and the Holt-Winters Triple Exponential smoothing. Firstly, the plot of the monthly CPI data (from November 2011 to October 2016) done and observed the presence of trend and check for stationarity. The data collected is shown in table 1 below.

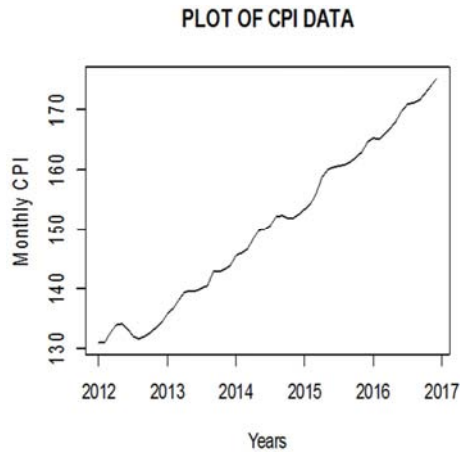
Table 1. CPI data, Source KNBS.

| | Jan | Feb | Mar | Apr | May | Jun | Jul | Aug | Sep | Oct | Nov | Dec |
|------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 2012 | 130.82 | 130.76 | 132.51 | 133.74 | 134.09 | 133.06 | 131.92 | 131.51 | 131.89 | 132.46 | 133.33 | 134.25 |
| 2013 | 135.62 | 136.57 | 137.96 | 139.25 | 139.52 | 139.59 | 139.87 | 140.29 | 142.82 | 142.75 | 143.14 | 143.85 |
| 2014 | 145.40 | 145.95 | 146.61 | 148.20 | 149.70 | 149.91 | 150.60 | 152.02 | 152.24 | 151.92 | 151.85 | 152.51 |
| 2015 | 153.43 | 154.14 | 155.86 | 158.70 | 159.98 | 160.46 | 160.57 | 160.90 | 161.33 | 162.13 | 162.97 | 164.72 |
| 2016 | 165.37 | 165.06 | 165.92 | 167.07 | 167.99 | 169.76 | 170.84 | 170.97 | 171.56 | 172.62 | 173.85 | 175.18 |

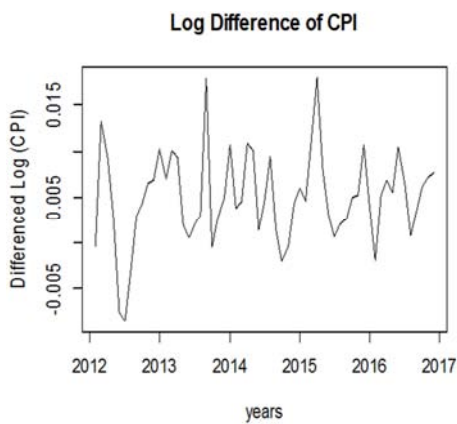
4.1. Box-Jenkins/SARIMA Model

4.1.1. Model Identification

From Fig. 1 (a), we can clearly see that the data has the presence of trend since its increasingly moving upward.



(a) Plot of Monthly CPI in Kenya



(b) Plot of Time Series CPI Data

Figure 1. (a)Non-stationary and (b)Stationary plots.

Our data exhibited trend but seasonality and hence no seasonal difference but transformed the data by taking log and first difference and re-evaluated the trend and Fig. 1(b) clearly indicates no presence of trend. We confirmed stationarity through Augmented Dickey Fuller (ADF) test and the results indicated stationarity as per the results below.

The hypothesis of the Augmented Dickey-Fuller t-test is:

H_0 : The data needs to be differenced to make it stationary

H_1 : The data is stationary and doesn't need to be differenced

Test regression trend

Call: lm(formula = z. diff z. lag. 1 + 1 + tt)

Table 2. ADF results.

| | Estimate | Std. Error | t value | Pr(> t) |
|-------------|------------|------------|---------|-------------|
| (Intercept) | 2.827e-03 | 1.369e-03 | 2.064 | 0.0437* |
| z. lag. 1 | -6.609e-01 | 1.260e-01 | -5.24 | 2.57e-06*** |
| tt | 1.874e-05 | 3.803e-05 | 0.493 | 0.6242 |

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 3. Critical values for test statistics.

| | 1pct | 5pct | 10pct |
|-----------|-------|-------|-------|
| tau3 | -4.04 | -3.45 | -3.15 |
| z. lag. 1 | 6.50 | 4.88 | 4.16 |
| tt | 8.73 | 6.49 | 5.47 |

4.1.2. Model Estimation and Selection

From Fig. 2. we can see that ACF and PACF The seasonal spikes at ACF and PACF after 1 lag (11, 12, 16) is observed which indicates that it is not taking seasonal difference of the series. This also indicates the seasonal model of SAR (1) and SMA(0). Therefore, model (1, 0, 0) will form the (P, D, Q) part. The non-seasonal part the PACF shows that there is a spike at lag 1 and no spike till lag 11, 12, 15 16 and then a discontinuation. This indicates AR (1) and since we differentiated our data once the model (p, d, q) will be (1, 1, 0). From this estimation we see that $D=0$ implying that the data does not have seasonality and hence there is no need for seasonal differencing. This means that the appropriate SARIMA model for forecasting Monthly CPI and hence inflation will be $(1,1,0)(1,0,0)_{12}$. Plot ACF and PACF to identify potential AR and MA model is shown below.

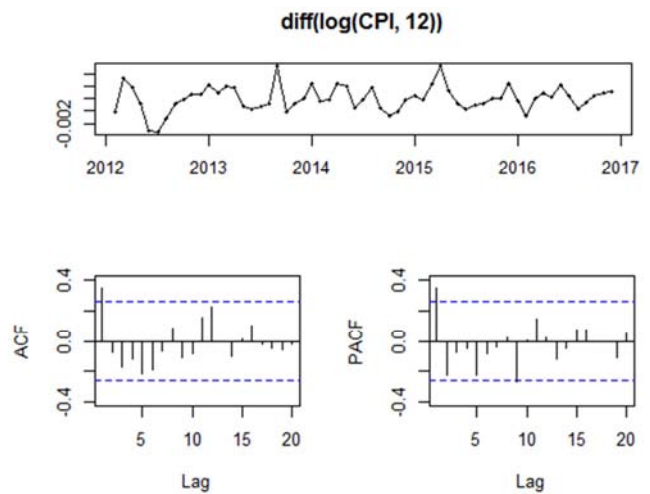


Figure 2. Plot of differenced Transformed Monthly CPI.

4.1.3. Diagnostics Checks

After identification and estimation of the model, diagnostic checks was analysed using the residuals of the model. Figure 4 shows the standardized residuals, using the Normal QQ Plot of Standardd Residuals, ACF plot of the residuals and p-values for the Ljung-Box statistic. From Figure 4, the Standardized Residuals indicate that there's no trend in the residuals and no changing variance with time. The ACF of the residuals shows no significant autocorrelations and hence the estimated model is good and the Normal QQ Plot of Std Residuals portray a normal distribution which are independent and identically distributed sequence with a mean of zero and a constant variance. We confirmed the plot results with The Shapiro-Wilk test for normality where the null hypothesis for this test is that the residual are normally distributed. Our p-value was found to be 0.1019 which greater than 0.05, hence we fail to reject the null hypothesis and conclude that the residuals are normally distributed. From the plot of p-values for the Ljung-Box-Pierce statistics, the presented statistics consider the accumulated residual autocorrelation from lag 1 up to and including the lag on the horizontal axis. The dashed blue line is at 0.05. we can clearly see that All p-values are above it which means that the results are very good. We can also see that the calculated p-value of Ljung-Box-Pierce

statistics at lag 24 is $p\text{-value} = 0.5168$. This confirms non-significant values for this statistic when looking at residuals. The diagnostic results confirm that we cannot reject

the null hypothesis of independence in this residual results. Hence we conclude that our Sarima $(1,1,0)(1,0,0)_{12}$ model is best for our forecasting inflation.

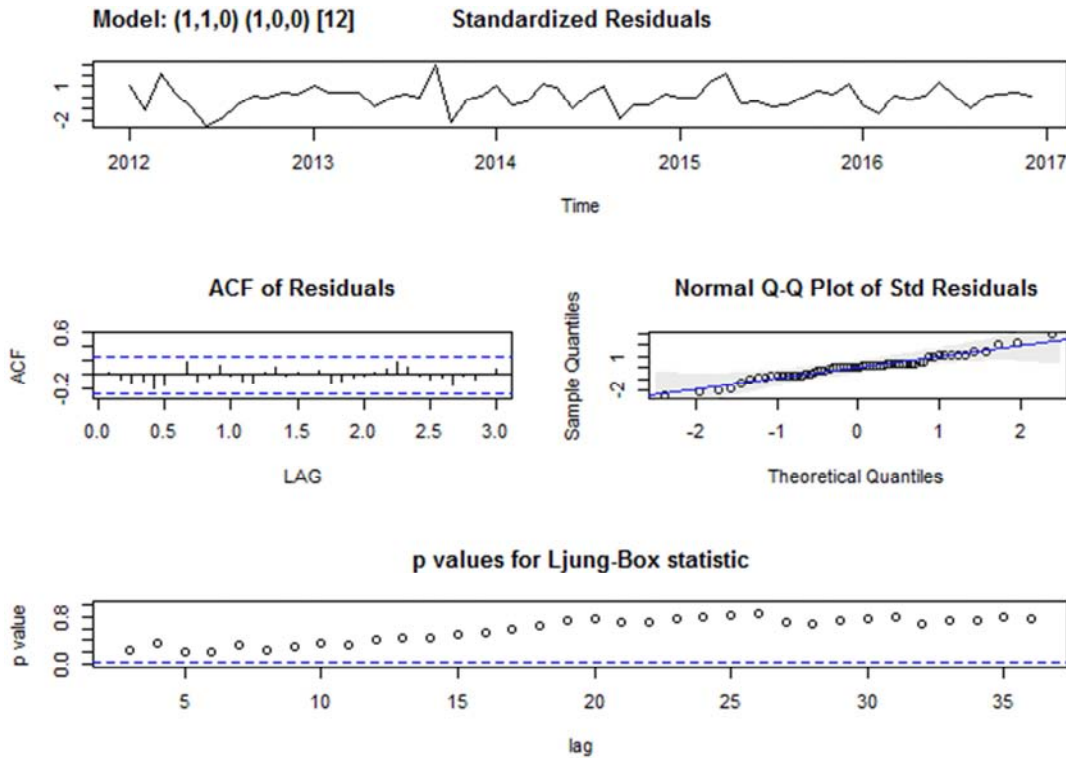


Figure 3. Residual Diagnostics.

4.1.4. Forecasting

Forecasted Values for January, February and March are, 176.0305, 176.5741 and 177.3643 respectively.

Forecasts from ARIMA(0,1,1)(0,0,1)[12] with drift

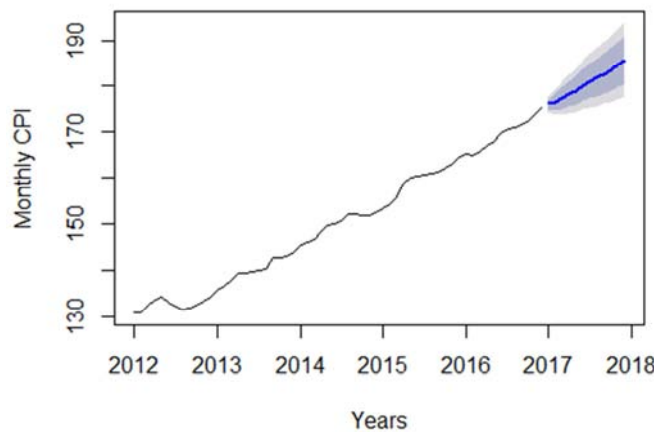


Figure 4. Plot of Actual and forecasted Values

4.2. Triple Exponential Smoothing (Holt-Winters Method)

4.2.1. Model Estimation and Selection

Hyndman et al [14] coined the three components of smoothing triplet (E, T, S) as error, trend and seasonality. In this paper we employed ETS in R for automatic selection of our model in a similar manner as used by Hyndman et al [15] in

Application of the automatic forecasting Strategy to the M-competition data and IJF-M3 competition data. The selected model was then tested for accuracy based on errors.

Figure 5 below indicates that the model space selected is the model ETS(A, A, N). This means that our model has additive errors, additive trend and no seasonality which is an indicator of Holt’s linear method with additive errors. This confirms lack of seasonality just hence confirming our SARIMA model results.

Decomposition by ETS(A,A,N) method

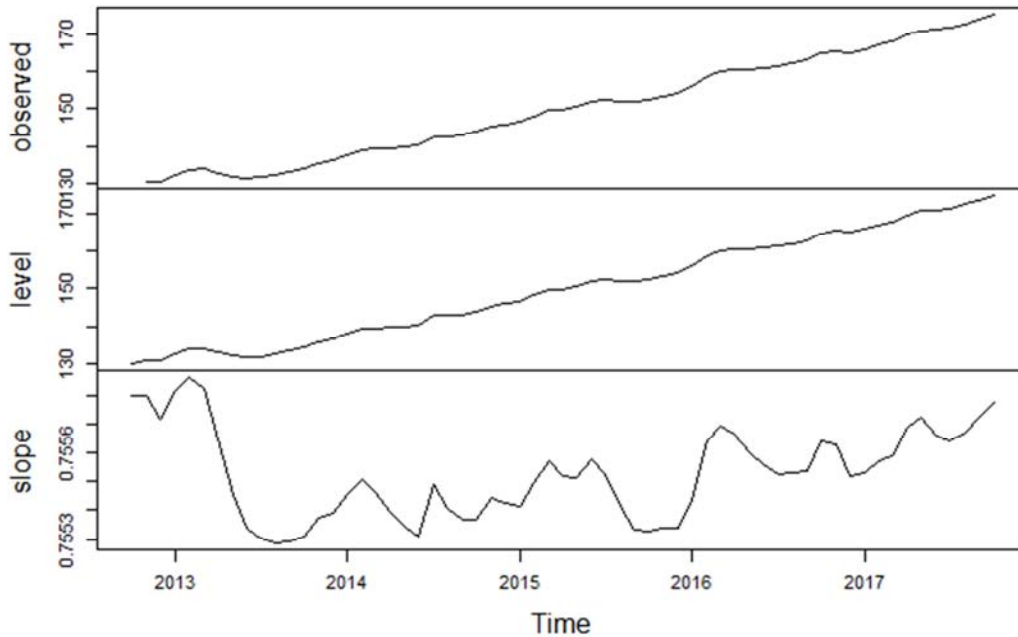


Figure 5. Decomposition by ETS.

4.2.2. Determination of Smooth Parameters

From the results above we see that $\alpha = 0.9999$ which is very close to 1. This indicates that more weight is on the recent value of the independent variable and less on the previous data. Our $\beta = 0.0001$ which is very close to zero. This means less weight on recent data. We do not have the γ which confirms the non existence of seasonality component in our model.

4.2.3. Diagnostic Check

Forecasted Values for January, February and March 2017 are 177.4458, 178.2016 and 178.9574 respectively

4.3. Discussion

The two prediction models were compared using MASE, MAE and MAPE. From Table 4 below, the values for the SARIMA model $(1,1,0)(1,0,0)_{12}$ were 0.059, 0.0036 and 0.073 for MASE, MAE and MAPE respectively, while the Holt-winters Triple exponential Smoothing model gave value of 0.643, 0.595 and 0.400 as values for MASE, MAE and MAPE respectively. This result imply that SARIMA model is better than Holt-Winters method based on the accuracy parameters chosen. This result contradict results by Puthran et al [9] whose results in the study of Indian Motor Cycle industry found that Holt Winters was more precise than SARIMA Model. On model selection, our result agree with Udom. P and Phumchusri. N (2014). who confirmed that ARIMA $(1,0,1)(1,0,1)_{12}$ was a better model compared to Holt-winters based on mean Absolute Percentage Error. Considering that MAPE if one of the accuracy parameters proposed by Hyndman R. J. and Koehler A. B. (2006), this paper is confident to make the comparison.

Table 4. Mean Errors for the two models.

| Model | MASE | MAE | MAPE |
|--------------|-------|-------|-------|
| SARIMA | 0.059 | 0.004 | 0.073 |
| Holt-Winters | 0.643 | 0.595 | 0.400 |

Finally, we compared the Actual CPI values for the month of January and February 2017 against the forecasted values with SARIMA and Holt-Winters using Analysis of variance with a view of finding out if they are significantly different. we got the results below.

4.4. ANOVA for Actual vs Forecasted Values

Here we test if the forecasted values between the two models are different and also test against the actual values for the months of January and February 2017

H_0 : The values are not significantly different

H_1 : The Values are significantly different

Studying the output of the ANOVA Table 5 above, we see that the F-statistic is 1.38 with a p-value equal to 0.3745. We clearly fail reject the null hypothesis of the data set are not significantly different.

Table 5. Anova results.

| | Df | Sum Sq | Mean Sq | F value | Pr(> F) |
|-----------|----|--------|---------|---------|---------|
| Type | 2 | 4.70 | 2.35 | 1.38 | 0.3752 |
| Residuals | 3 | 5.10 | 1.70 | | |

5. Conclusion

The main objective of this research paper has forecast Inflation in Kenya using SARIMA and Holt-winters (Triple

exponential Smoothing). The study results indicates that both models gives almost similar results but when MASE, MAE, and MAPE was compared SARIMA model seemed more accurate since it had has the minimum MASE, MAE, and MAPE values. This research employed univariate approach to forecast inflation. More research can be done using multivariate analysis to forecast CPI using factors that affect the component used in calculating CPI e.g exchange rates, Oil prices, Weather conditions etc. The results will be helpful for better financial planning and budgeting. The Forecasted CPI for January 2017 is 176.04 meaning that the Forecasted Inflation rate for January and February 2017 is 6.45% and 7.1% respectively while the actual inflation rate for January and February 2017 from KNBS website is 6.99% and 7.9%.

References

- [1] Otu A. O et al (2014) Application of Sarima Models in Modelling and Forecasting Nigeria's Inflation Rates. *American Journal of Applied Mathematics and Statistics*; 2(1): 16-28. doi: 10.12691/ajams-2-1-4.
- [2] Hyndman R. J. and Koehler A. B. (2006), "Another look at measures of forecast accuracy," *International Journal of Forecasting*, vol. 22, no. 4, pg. 679-688.
- [3] Virginia Gathing, 2014. modelling in ation in Kenya using ARIMA and VAR, Project to the School of Mathematics, University of Nairobi.
- [4] Ingabire. J and Mung'atu. J. K. (2016) Measuring the Performance of Autoregressive Integrated Moving Average and Vector Autoregressive Models in Forecasting Inflation Rate in Rwanda, *International Journal of Mathematics and Physical Sciences Research* ISSN 2348-5736 (Online) Vol. 4, Issue 1, pp: (15-25), Month: April 2016 - September 2016, Available at: www.researchpublish.com.
- [5] Tim C and Dongkoo C, (1999), *Modelling and Forecasting inflation in India*, IMF Working Paper, WP/99/119.
- [6] Uwilingiyimana C. et al (2015), Forecasting Inflation in Kenya Using Arima - Garch Models, *International Journal of Management and Commerce Innovations* ISSN 2348-7585 (Online) Vol. 3, Issue 2, pp: (15-27), Month: October 2015 - March 2016, Available at: www.researchpublish.com.
- [7] Jere, S. and Siyanga, M. (2016) Forecasting Inflation Rate of Zambia Using Holt's Exponential.
- [8] Hamidreza M. and Leila S. (2012). Using SARFIMA model to study and predict the Iran's oil supply. *International Journal of Energy Economics and Policy*. Vol. 2, No. 1, 2012, pp. 41-49.
- [9] Puthran D. et al (2014). Comparing SARIMA and Holt-Winters forecasting accuracy with respect to Indian motorcycle industry, *Transactions on Engineering and Sciences*, Vol. 2, Issue 5, May 2014, ISSN: 2347-1964 Online 2347-1875 Print.
- [10] Udom. P and Phumchusri. N (2014). A comparison study between time series model and ARIMA model for sales forecasting of distributor in plastic industry, *IOSR Journal of Engineering (IOSRJEN)*, Vol. 04, Issue 02 (February. 2014), V1, PP 32-38 ISSN (e): 2250-3021, ISSN (p): 2278-8719.
- [11] Box G. E. et al(2008),"Time Series Analysis Forecasting and Control, Fourth Edition", John Wiley & Sons, Inc., Hoboken, New Jersey.
- [12] Chatfield, C. (2000), "Time Series Forecasting", Chapman & Hall.
- [13] Shumway R. H and Stoffer D. S (2006), "Time Series Analysis and Its Applications With R Examples Second Edition" Springer Science+Business Media, LLC Smoothing.
- [14] Hyndman et al (2008) "Forecasting with Exponential Smoothing, The State Space Approach", Springer-Verlag Berlin Heidelberg.
- [15] Hyndman et al (2002) A state space framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*.
- [16] <https://www.centralbank.go.ke/monetary-policy/>.