Forecasting inflation rate in Kenya using SARIMA model

Susan W. Gikungu[1], Anthony G. Waititu[1], John M. Kihoro[2]

[1]Department of Statistics and Actuarial Science, Jomo Kenyatta University of Agriculture and Technology (JKUAT), Nairobi, Kenya
[2]Computing and E-learning Department, Co-operative University College of Kenya, Nairobi, Kenya

Email address:
wanjirusue@gmail.com (S. W. Gikungu)

Abstract: It is the desire of the policy makers in a country is to have access to reliable forecast of inflation rate. This is achievable if an appropriate model with high predictive accuracy is used. In this paper, Seasonal Autoregressive Integrated Moving Average (SARIMA) model is developed to forecast Kenya's inflation rate using quarterly data for the period 1981 to 2013 obtained from KNBS. SARIMA $(0,1,0)$ $(0,0,1)_4$ was identified as the best model. This was achieved by identifying the model with the least Akaike Information Criterion. The parameters were then estimated through the Maximum Likelihood Estimation method. Diagnostic checks using Jarque-Bera Normality Test indicated that they were normally distributed. ACF and PACF plots for the residuals and squared residuals revealed that they followed a white noise process and were homoskedastic respectively. The predictive ability tests RMSE=0.2871, MAPE=3.9456 and MAE= 0.2369 showed that the model was appropriate for forecasting the inflation rate in Kenya.

Keywords: Seasonal Autoregressive Integrated Moving Average (SARIMA), Kenya National Bureau of Statistics (KNBS),
Autocorrelation function (ACF) and Partial Autocorrelation Function (PACF),
Akaike Information Criterion and Jarque-Bera Test

1. Introduction

Inflation as defined by Webster (2000) is the persistent increase in the level of consumer prices or a persistent decline in the purchasing power of money.

Inflation tends to be a relatively persistent process, which means that current and past values should be helpful in forecasting future inflation (Brent and Mehmet, 2010). Applying that intuition, the two basic models that exploit information embedded in past values of inflation was constructed.

In 1970, Box and Jenkins introduced Autoregressive Integrated Moving Average models, ARIMA. SARIMA model is useful in situations when the time series data exhibit seasonality-periodic fluctuations that recur with about the same intensity periodically, for example, quarterly (Martinez et al., 2011). This property makes the SARIMA model appropriate for studies concerning quarterly inflation rate data.

2. Review of Previous Studies

Various studies have been done on SARIMA model and inflation in different counties. Fannoh et al. (2014) used SARIMA approach to model Liberia's monthly inflation rates which showed that SARIMA model was appropriate for modeling the inflation rates. Otu et al. (2014) used Box-Jenkins methodology to build ARIMA model for Nigeria's monthly inflation. SARIMA (1, 1, 1) (0, 0, 1)12 model was developed and used to forecast monthly inflation for the year 2014.

Saz (2011) studied the efficacy of SARIMA Models in the view of forecasting the inflation rates in Turkish economy. Tests were performed on the seasonality in the Turkish inflation rate. They used the systematic modeling approach in conjunction with the step-wise selection procedure of the HK Algorithm to find the best model of inflation in Turkey.

Kibunja et al. (2014) considered a univariate time series model to forecast precipitation in Mt. Kenya region. SARIMA model was fitted on to the data and the model which exhibited the least AIC and BIC values was selected. The model passed residual normality test and the forecasting evaluation statistics.

In this study, the forecast efficiency of SARIMA model in forecasting inflation rate in Kenya was done.

3. Methodology

SARIMA Model

SARIMA is a Box-Jenkins technique that takes into account time series data and decomposes it into:

-AR (Autoregressive) process- A real valued stochastic process (Y_t) is said to be an AR process of order p, denoted by AR(p) if

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_{1p} y_{t-p} + \varepsilon_t \qquad (1)$$

The value of AR(p) process at time t is therefore regressed on its own past p values plus a random shock.

-MA (Moving Average) process- A real valued stochastic process (Y_t) is said to be an MA process of order q, denoted by MA(q) if there exists $b_1,\ldots,b_q$ and a white noise (ε_t) such that

$$y_t = b_0 \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q} \qquad (2)$$

The value of MA(q) process at time t is therefore regressed on its own past errors.

-ARMA - Moving averages MA(q) and autoregressive AR(p) processes are special cases of so called autoregressive moving average processes (ARMA). A real valued stochastic process (Y_t) is said to be an ARMA process of order p, q, denoted by ARMA (p, q) if it satisfies the equation

$$y_t = \varepsilon_t + (a_1 Y_{t-1} + a_2 Y_{t-2} + \cdots + a_p Y_{t-p}) + (b_0 \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q} \qquad (3)$$

This can be re-written as $\phi(z) Y_t = \Theta(z) \varepsilon_t$

Where $\phi(z) = 1 + a_1 z + \cdots + a_p z^p$ and $\Theta(z) = 1 + b_1 z + \cdots + b_q z^q$ are the characteristic polynomials of the AR part and of the MA part of an ARMA (p, q) process (Y_t). z is the back-shift (lag) operator.

-ARIMA process- The Autoregressive Integrated Moving Average (ARIMA) model is a broadening of the class of ARMA models to include differencing. A process Y_t is said to be an ARIMA (p, d, q) if $(1 - z)^d Y_t$ is a causal ARMA (p,q). The corresponding ARIMA equation is

$$\phi(z)(1 - z)^d X_t = \Theta_q(z) \varepsilon_t \qquad (4)$$

-SARIMA process- For a non- stationary time series possibly containing seasonality, that is, seasonal periodic component repeats itself after every s observations, Box-Jenkins (1976) have defined a general multiplicative Seasonal ARIMA model (SARIMA) as

$$\phi_p(z) \Phi_P(z^s)(1 - z)^d (1 - z^s)^D Y_t = \theta_q(z) \Theta_Q(z^s) \epsilon_t \qquad (5)$$

Where $\phi_p(z)$, $\Phi_P(z^s)$, $\theta_q(z)$ and $\Theta_Q(z^s)$ are characteristic polynomials of orders p, P, q and Q respectively. d and D are the orders of non-seasonal and seasonal differencing respectively. Box-Jenkins(1976) bases the model selection on three stages ie. Model identification, estimation of coefficients and diagnostic checking.

Stage1: Model Identification

The objective of this step is to determine the possible SARIMA model that best fit the time series data under consideration. SARIMA model is appropriate for stationary time series therefore stationarity condition must be satisfied. Augmented Dickey Fuller (ADF) test is used to see whether the seasonal differenced series is stationary. The values of p, q, P and Q are then determined at this step by looking at the patterns of the Autocorrelation function (ACF) and the Partial Autocorrelation Function (PACF).

Stage2: Estimation of coefficients

The parameters are estimated by the maximum likelihood estimation method. In this study, the model with the minimum value of AIC is judged as the best model. It is given by

$$AIC = -2 \ln(L) + 2k \qquad (6)$$

where k= p+q+1 and

L is the maximized likelihood value.

Stage3: Diagnostic checking

This step involves checking whether the residuals follow a white noise process and ensure that the estimated parameters are statistically significant. This is done by inspecting the i) ACF plots of the residuals ii) the probability plots of the residuals iii) the residual q-q plots. Jarque Bera test can also be used to verify normality among the residuals. If the model fails these diagnostic checks then return to the identification stage to find a better model. The best model is then used to forecast.

4. Empirical Results

The data used in this study is quarterly inflation rate of Kenya from 1981 to 2013, which is made up of 132 observations. It was obtained from Kenya National Bureau of Statistics. Figure 1 shows the plot of Kenya's quarterly inflation.
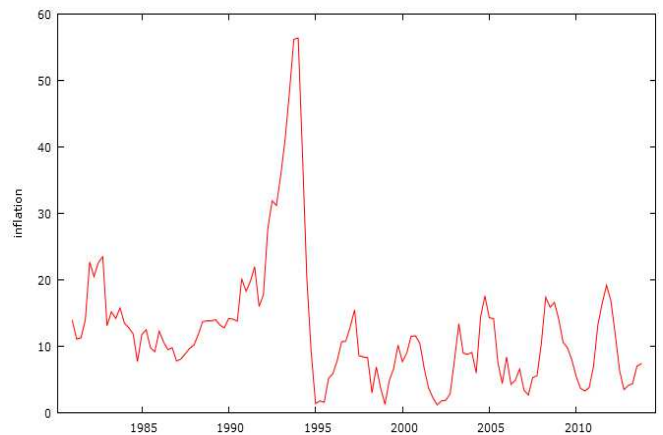

Figure 1. *Quarterly inflation rates of Kenya, 1981:1 to 2013:4.*

The plots of the autocorrelation and partial autocorrelation functions are presented in figure 2. The ACF plot dies down in a sine wave pattern which implies that there is a seasonal and a non seasonal component of the series.
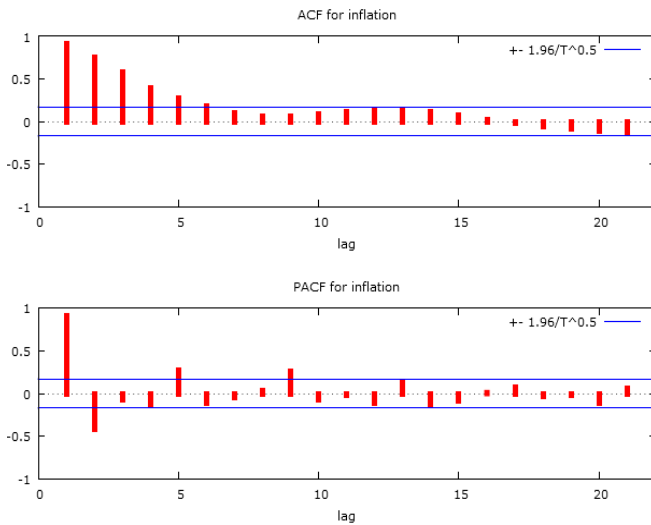
Figure 2. ACF and PACF plots

Figure 3. Residual Q-Q plot.

4.1. SARIMA Model

4.1.1. Model Identification and Estimation of Coefficients

SARIMA modeling requires the stationarity condition to be satisfied. Augmented Dickey Fuller (ADF) test was used to test for it. The p-value obtained from the ADF test without a constant is 0.1216. We therefore failed to reject the null hypothesis at 5% level of significance and concluded that the data is non-stationary. This meant that differencing was necessary to make the data stationary. The differenced data was tested again and the value obtained was 1.215e-010 and we concluded that the data were stationary.

The best model was selected by picking the model with the least AIC value. Model $(0,1,0)(0,0,1)4$ was selected as the most appropriate. Using the Maximum Likelihood estimator, the model parameters ϕ, Θ, Φ and Θ were estimated. The fitted model is given by

$$(1 - z)w_t = w_{t-1} + e_t + \Theta_1 e_{t-s} \tag{7}$$

Replacing the coefficients with the corresponding values, the following is obtained;

$$w_t = w_{t-1} + e_t + 0.216634 e_{t-s} \tag{8}$$

4.1.2. Diagnostic Checking

The residuals were checked to find out if they followed a white noise process. This was achieved by plotting the residual Q-Q and normality test plots as shown in figure 3 and 4 respectively. The Q-Q plot is reasonably straight so normality is okay. The histogram shows a bell shaped distribution with a p-value=0.0639237 >0.05 which is an indicator for normality.

In addition, the ACF plot of the residuals in figure 5 shows that for the first 20 lags, all sample autocorrelations fall inside the 95% confidence bounds indicating the residuals appear random. The forecasting evaluation statistics in Table 1 reveals that SARIMA model is appropriate in forecasting inflation in Kenya.
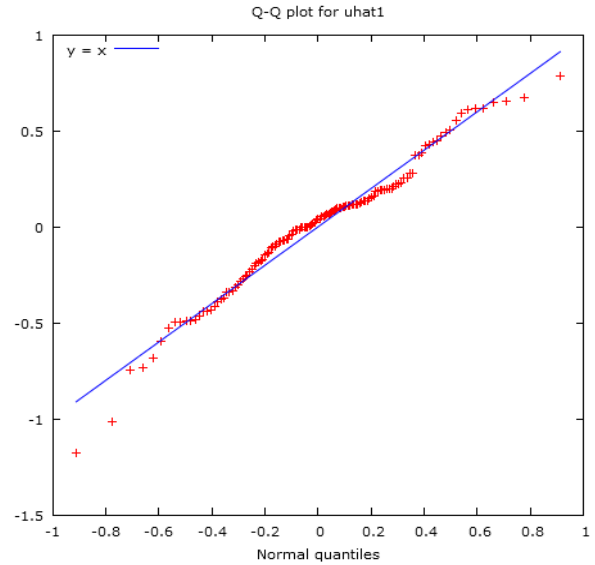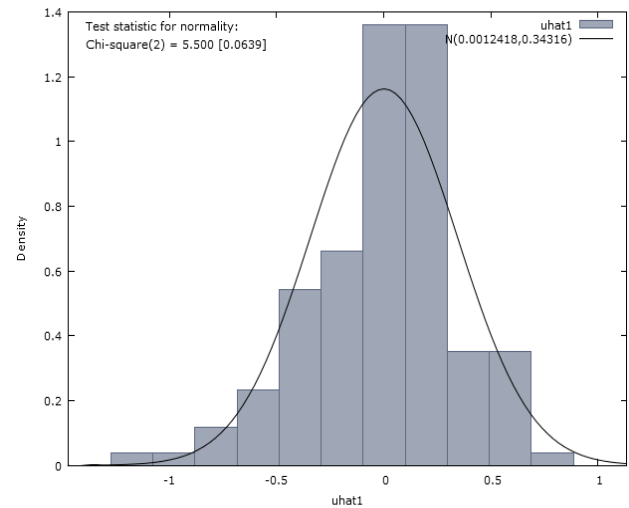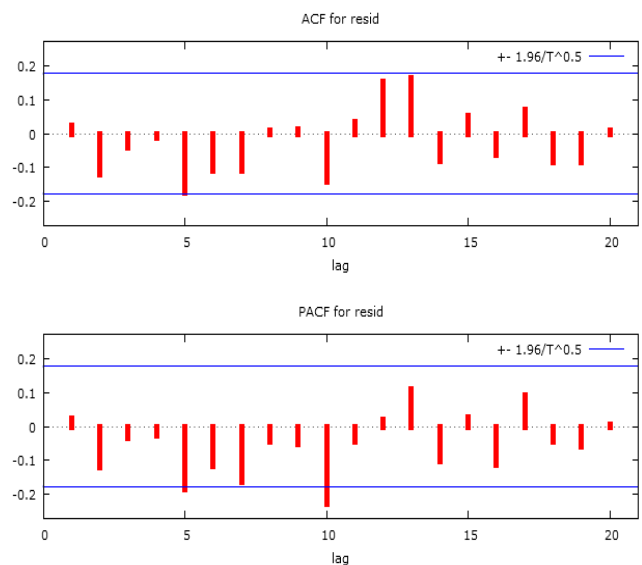

Figure 4. Residual Normality test plot.

Figure 5. Residual ACF-and PACF plots.

Table 1. *Performance Statistics.*

	SARIMA
RMSE	0.2871
MAE	0.2369
MAPE	3.9456

5. Conclusion and Recommendations

In this study, we model the inflation rates of Kenya using Seasonal Autoregressive Integrated Moving Average (SARIMA) model. The model proved to be well-fitted for forecasting inflation rate in Kenya according to the results of the residual diagnostic checks. In addition, RMSE, MAE and MAPE were also employed. Eight quarters out of sample forecast for the year 2014 and 2015 was conducted. It shows a fluctuation in inflation rates and an increasing trend towards the end of year 2015. We recommend that appropriate policies and measures have to be adopted by the government and major stake holders to ensure that the aim of single digit inflation rate value is achieved in Kenya.

References

[1] Box, G.E.P. and Jenkins G.M. (1976), "Time Series Analysis, forecasting and Control, *Holden-Day*, San Francisco.

[2] Brent, M., and Mehmet P. (2010), "Simple ways to forecast inflation: What works best?", *Trade Publication* ,17, 1-9.

[3] Fannoh, R., Orwa G., and Mung'atu J. K.. (2014), "Modeling the Inflation Rates in Liberia SARIMA Approach", *International Journal of Science and Research*, 3, 1360-1367.

[4] Kibunja H., *Kihoro J. and Orwa* G.*(2014),* "Forecasting Precipitation Using SARIMA Model: A Case Study of Mt. Kenya Region", International Institute for Science, Technology and Education, 4(11), 50-58

[5] Martinez E. Z., and Soares E.A. (2011), "Predicting the number of cases of dengue infection in Ribeirão Preto, São Paulo State, Brazil, using a SARIMA model", *Revista da Sociedade Brasileira de Medicina Tropical,* 44(4), 436-440.

[6] OtuA. O., Osuji G. A., Opara J., Ifeyinwa M. H., and Iheagwara A.I. (2014), "Application of SARIMA models in modelling and forecasting Nigeria's inflation rates", *American Journal of Applied Mathematics and Statistics*, 2, 16-28.

[7] Saz G. (2011), "The efficacy of SARIMA models in forecasting inflation rates in developing countries: The case for Turkey", *International Research Journal of Finance and Economics*, 62, 111-142.

[8] Webster, D. (2000), "New Universal Unabridged Dictionary", *Barnes and Noble Books*.