Journal of Statistics: Advances in Theory and Applications
Volume 11, Number 1, 2014, Pages 31-43

ANALYSIS OF THE FLUCTUATION OF GROSS DOMESTIC PRODUCT IN KENYA USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE MODEL

DENNIS K. MURIITHI[1], A. N. NGERETHA[2], R. G. MURIUNGI[3] and E. W. NJOROGE[4]

[1,2,4]Chuka University
P. O. Box 109-60400
Chuka
Kenya
e-mail: kamuriithi@yahoo.com

[3]Meru University of Science and Technology
P. O. Box 972-6200
Meru
Kenya

Abstract

Modelling and forecasting the Kenyan economy is a vital concern. In this paper, the annual gross domestic product (GDP) is forecasted using autoregressive integrated moving average models (ARIMA) so as to determine the most efficient and adequate model for analyzing the Kenyan GDP. The study employed the Box-Jenkins (1976) methodology that involves stages of identification, estimation, diagnostic checking, and forecasting of a univariate time series. An exploratory research design was adopted for a sample of 51 observations. The annual data was obtained from the World Bank national

2010 Mathematics Subject Classification: JEL, C22, E01, REL, 8C, 10G.
Keywords and phrases: Akaike information criteria, gross domestic product, auto-regressive, moving average, economic growth.

Received February 10, 2014

accounts data, and OECD National Accounts data files for the period of 1960 to 2011. Analysis was done using Gretl-package. The results indicate that autoregressive integrated moving average models (ARIMA) (2, 1, 2) is the most adequate and efficient model. This was ascertained by comparing the various model selection criterion and the diagnostic tests for various models. A better understanding of a country's GDP situation and future economic growth will facilitate the government in making appropriate policy measures to maintain high and stable economic growth.

1. Introduction

Kenya is the largest economy in east Africa and is a regional financial and transportation hub. Maintaining a high GDP throughout the years has become a vital concern. The Kenyan GDP has been fluctuating in the past years. After independence, Kenya promoted rapid economic growth through public investment, encouragement of smallholder agricultural production, and incentives for private (often foreign) industrial investment. Gross domestic product (GDP) grew at an annual average of 6.6% from 1963 to 1973. Agricultural production grew by 4.7% annually during the same period, stimulated by redistributing estates, diffusing new crop strains, and opening new areas to cultivation. After experiencing moderately high growth rates during the 1960s and 1970s, Kenya's economic performance during the 1980s and 1990s was far below its potential [6]. During the period 2003-2007, the Government of Kenya began an ambitious economic reform program and resumed its cooperation with the World Bank and the IMF. There was some movement to reduce corruption in 2003, but the government did not sustain that momentum. Economic growth began to recover in this period, with real GDP growth registering 2.8% in 2003, 4.3% in 2004, 5.8% in 2005, 6.1% in 2006, and 7.0% in 2007. However, the economic effects of the violence that broke out after the December 27, 2007 general election, compounded by drought and the global financial crisis, brought growth down to less than 2% in 2008. In 2009, there was modest improvement with 2.6% growth, while the final 2010 growth figure was expected to be about 5%. From 2010 henceforth; the economic growth has remained high. Maintaining it to the highest peak has become a challenge to the government due to increased corruption in the public sectors.

2. Problem Statement

Kenya's economic performance has been hampered by numerous interacting factors: heavy dependence on a few agricultural exports that are vulnerable to world price fluctuations, population growth that has outstripped economic growth, prolonged drought that has necessitated power rationing, deteriorating infrastructure, and extreme disparities of wealth that have limited the opportunities of most to develop their skills and knowledge. Poor governance and corruption also have had a negative impact on growth, making it expensive to do business in Kenya. According to Transparency International, Kenya ranks among the world's half-dozen most corrupt countries. Bribery and fraud cost Kenya as much as US$1 billion a year. Kenyans, most living on less than US$1 per day, pay some 16 bribes a month-two in every three encounters with public officials. Another large drag on Kenya's economy is the burden of human immunodeficiency virus/acquired immune deficiency syndrome (HIV/AIDS). This makes it difficult to predict the Kenyan economy as it keeps on fluctuating due to the factors discussed. This study therefore aims at coming up with an effective and efficient model for forecasting the Kenyan economic growth so as to make it possible for the government to make appropriate decisions on monetary and fiscal policies.

3. Methodology

3.1. Autoregressive model

The notation AR(p) refers to the autoregressive model of order p. The AR(p) model is written as

$$X_t = C + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t, \tag{1}$$

where $\varphi_i, \ldots, \varphi_p$ are parameters, C is a constant, and the random variable ε_t is white noise.

An autoregressive model is essentially an all-pole infinite impulse response filter with some additional interpretation placed on it. Some constraints are necessary on the values of the parameters of this model in order that the model remains stationary [8].

3.2. Moving average

In time series analysis, the moving-average (MA) model is a common approach for modelling univariate time series models [3]. The notation MA(q) refers to the moving average model of order q:

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q}, \tag{2}$$

where μ is the mean of the series, the $\theta_1, \dots, \theta_q$ are the parameters of the model, and the $\varepsilon_t, \varepsilon_{t-1}, \dots.$ are white noise error terms. The value of q is called the order of the MA model. This can be equivalently written in terms of the backshift operator B as

$$X_t = \mu + (1 + \theta_1 B + \dots + \theta_q B^q)\varepsilon_t. \tag{3}$$

The moving-average model is essentially a finite impulse response filter applied to white noise, with some additional interpretation placed on it. According to [2], the role of the random shocks in the MA model differs from their role in the AR model in two ways. First, they are propagated to future values of the time series directly: For example, ε_{t-1} appears directly on the right side of the equation for X_t. In contrast, in an AR model, ε_{t-1} does not appear on the right side of the equation, but it does appear on the right side of the X_{t-1} equation, and X_{t-1} appears on the right side of the X_t equation giving only an indirect effect of ε_{t-1} on X_t. Second, in the MA model, a shock affects X values only for the current period and q periods into the future; in contrast, in the AR model, a shock affects X values infinitely far into the future, because ε_t affects X_t, which affects X_{t+1}, which affects X_{t+2}, and so on forever. Sometimes the autocorrelation function (ACF) and partial autocorrelation function (PACF) will suggest that an MA model would be a better model choice and sometimes both AR and MA terms should be used in the same model.

3.3. Auto-regressive integrated moving average model: ARIMA (p, d, q)

An autoregressive integrated moving average (ARIMA) model is one of the time series models used in forecasting. This model unlike others does not assume knowledge of any underlying economic model or structural relationships. It is assumed that values of the past series and the previous disturbance terms have information for the purposes of forecasting. The major advantage of ARIMA is that it requires data on the time series in question only. [7], provides that the ARIMA model is a modification of the ARMA model. According to [10], ARMA models represent a combination of the auto regressive (AR) and moving average (MA) models. If we differentiate the time series data d, then the model becomes stationary. By applying ARMA (p, d, q) to it, we say the original time series has been transformed to stationarity and is referred to as an auto-regressive integrated moving average model (p, d, q), where p is the number of autoregressive terms, d is the number of non-seasonal differences, and q is the number of lagged forecast errors in the prediction equation. Stationarity in the model can be detected by the use of the auto-correlation function graph (ACF). If a graph of the time series values either cuts off or dies down fairly quickly, then the times series should be considered stationary. If, however, the graph of the ACF dies down extremely slowly, then the time series values should be considered non-stationary. If the original time series is stationary, $d = 0$ and the ARIMA model reduces to ARMA model.

The application of the ARIMA model to time series analysis is due to [1]. The ARIMA model will be applied in this study to forecast the GDP using annual data from international monetary fund (IMF). ARIMA model predicts future values of a time series by a linear combination of its past values and a series of errors (also known as random shocks or innovations). The ARIMA command performs a maximum likelihood fit of the specified ARIMA model to the time series.

3.4. Box-Jenkins methodology

Box-Jenkins forecasting models are based on statistical concepts and principles and are able to model a wide spectrum of time series behaviour [1]. It has a large class of models to choose from and a systematic approach for identifying the correct model form. There are both statistical tests for verifying model validity and statistical measures of forecast uncertainty. In contrast, traditional forecasting models offer a limited number of models relative to the complex behaviour of many time series with little in the way of guidelines and statistical tests for verifying the validity of the selected model.

4. Empirical Results

4.1. Model identification

The preliminary analysis was done by use of time plots for the various series presented in Figure 1.

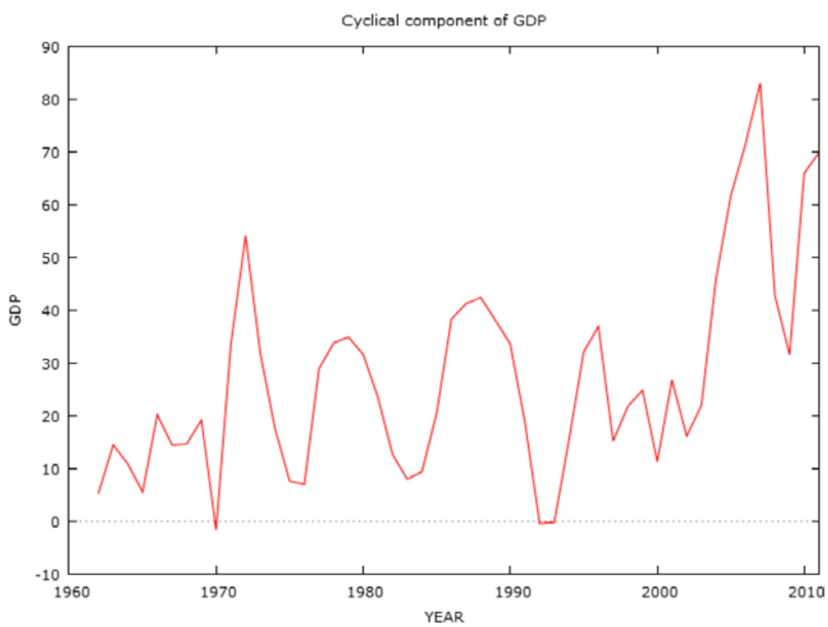

Figure 1. Time series plot.

By inspecting the time plot visually, it clearly shows that the mean and variance are non-constant implying that the data is non-stationary. The non-constant mean and variance provides a suggestion of the utilization of a non-linear model. The series was transformed to attain stationarity by taking the first differences of the natural logarithms of the values in each series. The equation representing the transformation is given by

$$\text{GDP} = In(P_t) - In(P_{t-1}), \tag{4}$$

where P_t represents the annual value for each series. The resulting plots for the returns are presented below:

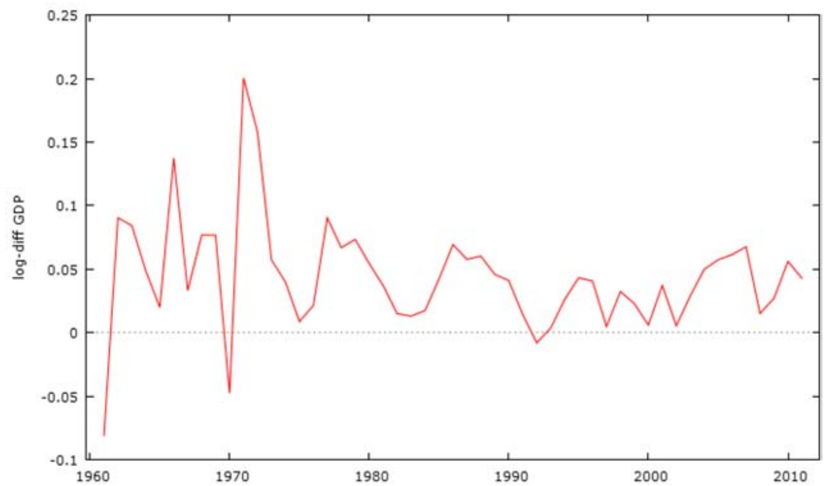

Figure 2. Log-differenced GDP.

The time plot of the series now depicts stationarity. The mean and variance show the property of being constant. At lag 1970, there is a shock caused by a sharp fall in the GDP.

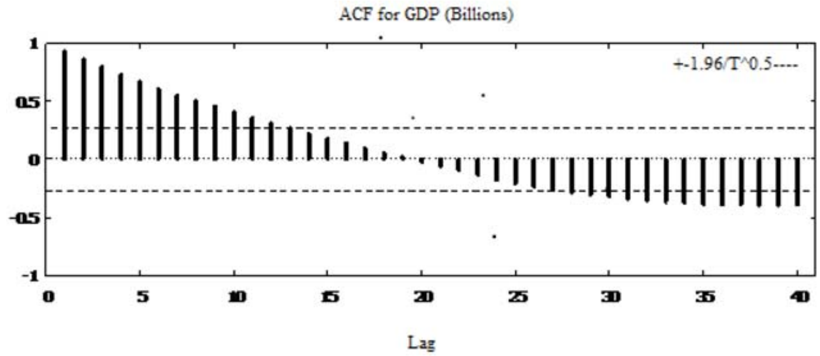

Figure 3. Auto correlation function for GDP.

The autocorrelation function also provides very useful information that is typical of a non-stationary process, where the autocorrelation declines slowly as the number of lags increases. The correlogram in Figure 3 indicates the cycles gradually dying away, thus the series is non-stationary. To make the data stationary, we obtained the logarithm of the first order difference of the dependent variable (GDP growth). The ACF for log differenced GDP is shown in Figure 4.

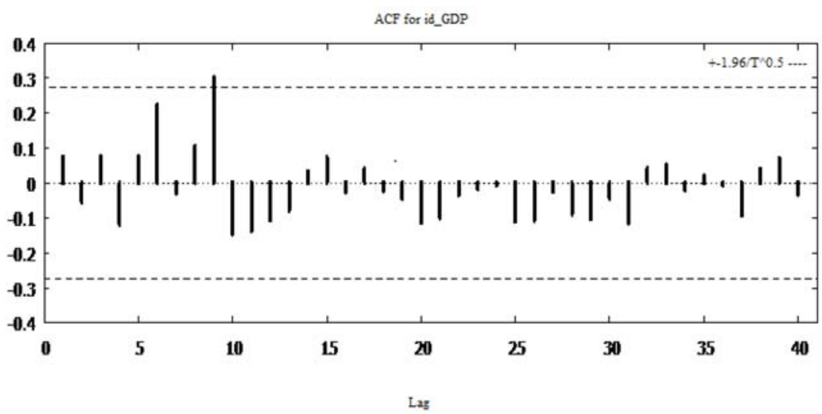

Figure 4. Autocorrelation function for log-diff GDP.

It shows high and decreasing correlations at regular intervals. Figure 4 shows, the correlogram plotted at 95% confidence interval limits, which are approximated to $\pm\frac{2}{\sqrt{n}}$, where $n = 51$ observations. Any value of correlation falling outside $\pm\frac{2}{\sqrt{n}}$ is taken to be significantly different from zero.

4.2. Model estimation and testing

The result in Table 1 shows the model estimates and their significant (if any) in the model at 5% significance level.

Table 1. Estimation results

Variable	Estimate	Std. error	Z-ratio	p-value
Coefficient	0.0438	0.45662	0.4539	0.76841
AR(1)	0.3893	0.6456	0.6029	0.5466
AR(2)	0.3720	0.1607	2.3146	0.0206
MA(1)	1.1493	0.7548	0.1978	0.8432
MA(2)	0.5059	0.2681	– 1.8868	0.05918

Table 1, shows the coefficient estimates of various autoregressive and moving averages of the Kenyan GDP. The coefficients of AR(1), MA(1), and MA(2) are not statistically significant, although MA(2) is marginally significant at 5% level of significance. Only the coefficient of AR(2) is statistically significant at 5% significance level. Based on estimation of these results, the model can be written as follows:

$$Y_t = 0.04375 + 0.372X_{t-2} + 0.506e_{t-2} + e_t, \tag{5}$$

where Y_t is GDP growth, Y_{t-i} represents the GDP growth at period $t - i$, e_{t-i} represents the random shock at period $t - i$, and t is the time period. For $i = 1, 2,$ all the estimates in the equation are positive indicating that the lags are positively related to previous variables in the previous period.

Evaluation of various ARIMA models

The chosen model was compared with other alternative models and by using the log-likelihood, Schwartz criterion, Hannan-Quinn criterion, mean square error, Akaike information criterion, mean of innovations and S.D. of innovations was considered to be the best model to fit the GDP data.

Table 2. Evaluation of various ARIMA models

ARIMA	(1, 1, 0)	(0, 1, 1)	(1, 1, 1)	(2, 1, 1)	(2, 1, 2)
Log likelihood	83.6874	83.7197	84.0696	85.7894	75.1990
Schwartz criterion	– 155.639	– 151.791	– 148.579	– 148.107	– 138.662
Akaike criterion	– 161.375	– 159.439	– 159.139	– 159.579	– 144.398
Hannan Quinn	– 159.191	– 156.527	– 154.497	– 155.210	– 142.214
Mean square error	0.1488	0.1195	0.1187	0.1174	0.1133
Mean of innovation	0.0014	0.0074	0.0073	0.0076	0.0076
SD. innovations	0.0537	0.0436	0.0436	0.0432	0.0410

The best model is the one with lowest information criteria. The ARIMA (2, 1, 2) model (optimal model) displays lesser results compared to the others and is hence the best model to forecast the Kenyan GDP.

4.3. Forecasting

Apart from within sample forecasts mentioned, the study also estimates eleven years out-of-sample forecasts of the model in order to measure its forecasting ability [1]. From the results below, the model is seen to have a high forecasting power as shown by the graph Figure 5.

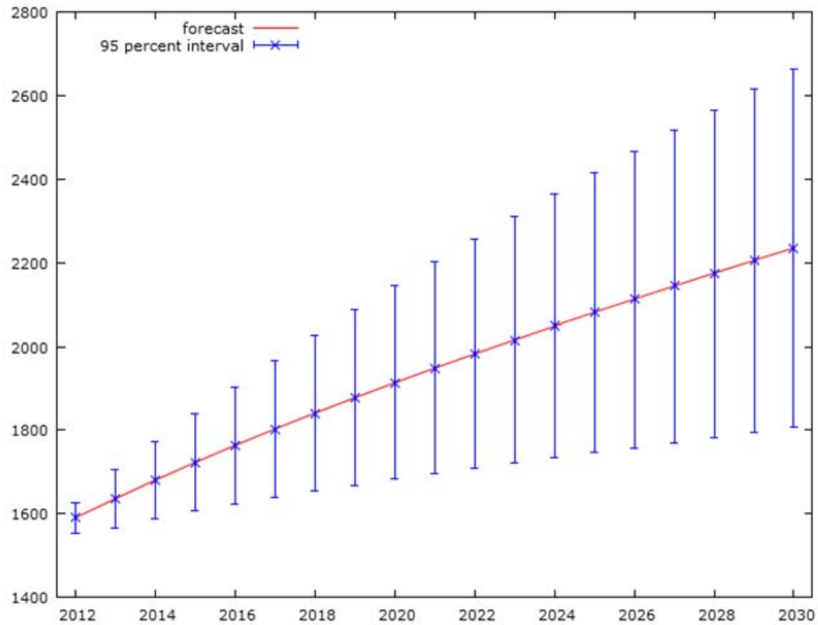

Figure 5. Forecasted growth domestic product (Kshs. in Billions).

Results indicate that Kenya's GDP will continue to rise from 1,680.36 billion in 2014. It will continue to rise up to 2,234.93 billion according to the data used by the year 2030. This suggests that the economic recovery measures employed by the government should be embraced in order Kenya to remain regional financial and transportation hub.

5. Conclusion

In this study, we used statistical methods to analyze, explain, and present the time series data of Kenya's GDP growth from 1960-2011 (Source: World Bank national accounts data, and OECD National Accounts data files). Comparing with other models, ARIMA(2, 1, 2) model was selected as the final model for forecasting Kenya's GDP growth for 2012-2030. All the model assumptions were satisfied (randomness, independent and normality of the residuals). In conclusion, the study found out that Kenya's GDP will continue to rise if the factors remain constant. Risks likely to shape economic growth would include the high international oil prices, which would result in a drop in economic growth, fluctuations in the exchange rate, inadequate rainfall, which has been insufficient, rising global food prices, unexpected terrorists attack, e.g., the Westgate mall attack and most importantly the political environment as the country moves closer to the 2017 elections.

6. Recommendation

From the findings, it calls for government interventions to maintain the GDP growth following the forecast. Kenya must begin to address the negative politics of ethnicity and historical grievances that many Kenyan ethnic groups may harbor. Kenya needs to institutionalize the fundamentals of democracy and governance with a view to providing and promoting an enabling environment, in which business enterprises flourish. The country's political elite should, in the near future, promote policies that would integrate its populations across ethnic lines than dividing them akin to the South Africa truth and reconciliation mechanism. Kenya, as a developing country, must involve the human capital in order to maintain the economic growth.

References

[1] G. E. Box, G. M. Jenkins and G. C. Reinsel, Time Series Analysis: Forecasting and Control, Engelwood Cliffs, New Jersey, USA, 1994.

[2] R. L. Cooper, The Predictive Performance of Quarterly Econometric Models of the United States, Econometric Models of Cyclical Behaviour, New York, 1972.

[3] A. Elena-Adriana and B. Elena, Econometric Modeling of GDP time series, theoretical and applied economics, University of Bucharest XVIII 10(563) (2011), 91-98.

[4] Kenya economy, Macroeconomic Forecasting using Business Cycle Leading Indicators, Stockholm: US-AB, 2001.

[5] C. C. Kotut, K. I. Menjo and J. Jepkwony, An econometric impact of petroleum oil price fluctuation in Kenya, Journal of Economics and Sustainable Development 3(7) (2012).

[6] A. Legovini, Kenya Macroeconomic Evolution since Independence, UNDP Unpublished Discussion Paper, 2002.

[7] D. K. Muriithi, G. G. Njoroge, E. W. Njoroge and O. Mark, Classification of higher education loans using multinomial logistic regression model, Journal of Mathematical Sciences: Advances and Applications 22 (2013), 1-17.

[8] M. Prerna, Forecasting natural gas price-time series and nonparametric approach, Proceedings of the World Congress on Engineering 1 (2012), 5.

[9] Lu. Yang, Economic Recovery Strategy for Wealth and Employment Creation, (2008). D-level essay in Statistics, Department of Economics and Society, Hogskolan Dalarna, Sweden, 2009.

[10] P. G. Allen, Economic forecasting in agriculture, International Journal of Forecasting 4(10) (1994), 81-135.

[11] S. Makridakis and M. Hibon, ARMA models and the Box-Jenkins methodology, Journal of Forecasting 1(2) (1995), 111-153.

■