

LINEAR STATISTICAL MODELLING OF THE GROSS DOMESTIC PRODUCT GROWTH OF KENYA

Miriam Wamaitha Thuo

Lecturer, Department of Social Sciences, Chuka University

DOI: 10.46609/IJSSER.2021.v06i02.001 URL: <https://doi.org/10.46609/IJSSER.2021.v06i02.001>

ABSTRACT

GDP a basic measure of an economy's economic performance, is the market value of all final goods and services produced within the borders of a nation in a year. Kenya has promoted rapid economic growth through public investment, encouragement of smallholder agricultural production, and incentives for private (often foreign) industrial investment and this has seen the economy experience some periods of growth. The overall objective of the study was to model and forecast the GDP Growth rate for Kenya for the period 2017-2021. The features of annual Kenya GDP Growth rate data over a period of fifty two years are studied. Secondary data spanning from 1961-2016 was used. To model the GDP, a class of ARIMA models was built following the Box-Jenkins approach that involved the stages of identification, estimation, and diagnostic checking. Results indicated that the ARIMA (4,1,2) was the best model to fit the GDP data. Finally, the fitted ARIMA model was used to do an out-of-sample forecasting for GDP Growth rate values for 2017-2021. The forecasting values obtained indicated that the economy would grow initially but later decrease gradually during the forecasted periods.

Keywords: Autoregressive Integrated Moving Average, Economic Growth, Forecasting, Gross Domestic Product, Investment

1.1 Introduction

Gross Domestic Product (GDP) can be defined in simple terms as the value of all final goods and services produced in a country over a period of time, usually one year. Another way in which GDP can be defined is the sum of value added at every stage of production (intermediate stages) by all the industries within a country, plus taxes less subsidies on the products in the period, or also as the sum of the income generated by production in the country in the period; that is compensation of employees, taxes on production and imports less subsidies and profits.

GDP of a country gives a general idea of how well-off a country is. It is also the most common way of describing Economic growth though not a perfect measure. Economic Growth is the

increase in the amount of goods and services produced in an economy over time. It is conventionally measured as a percent rate of increase in real GDP or Gross National Product (GNP). These two measures, which are calculated slightly differently, total the amounts paid for the goods and services that a country produced during the period. As an example of measuring economic growth, if Kenya creates Ksh 8,000,000,000 in goods and services in 2000 and then creates Ksh 8,080,000,000 in 2001, then it has achieved a nominal economic growth rate of 1% for 2001.

1.1 Trends in Economic Growth of Kenya

Kenya is the largest economy in East Africa and is a regional financial and transportation hub. Since independence, the government of Kenya has promoted rapid economic growth through public investment, encouragement of smallholder agricultural production, and incentives for private (often foreign) industrial investment.

The country's Vision 2030 aims at achieving an average economic growth rate of 10 percent per annum (Republic of Kenya, 2007). However, economic growth in Kenya has been low and erratic with the economy growing recording its highest growth in 1971 at 22.17 percent and the lowest growth in 1970 at -4.66. The following graph shows the trends in GDP growth from 1963-2014.

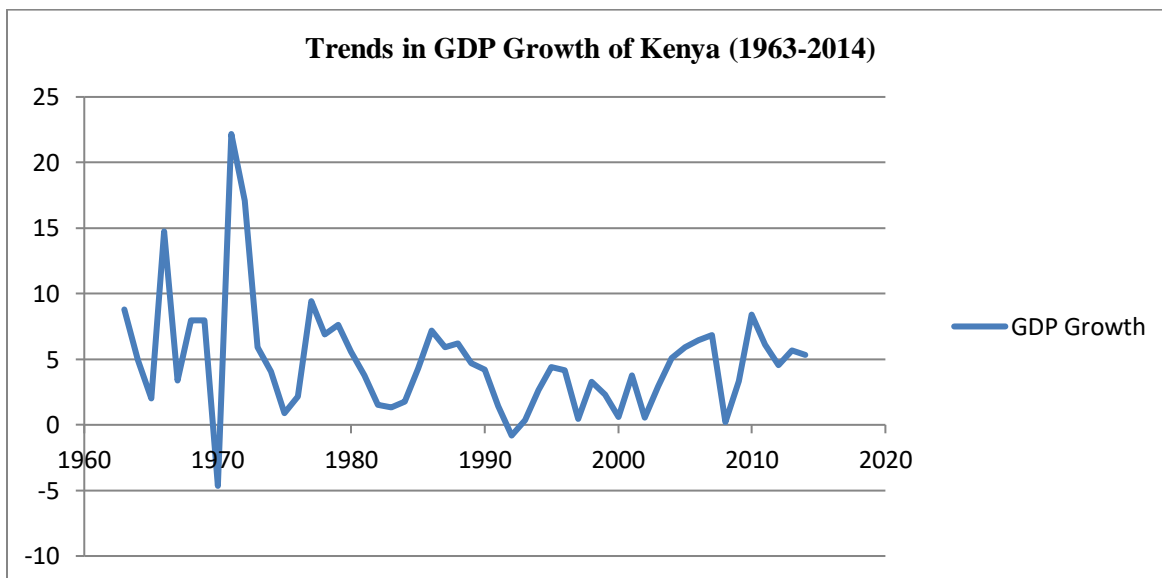


Figure 1: Trends in GDP growth of Kenya (1963-2014).Source: World Bank

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. From 1991 to 1993, Kenya

had its worst economic performance since independence. Growth in GDP stagnated, and agricultural production shrank at an annual rate of 3.9%. Inflation reached a record 100% in August 1993. In the mid-1990s, the government implemented economic reform measures to stabilize the economy and restore sustainable growth. Nevertheless, the economy grew by an annual average of only 1.5% between 1997 and 2002, which was below the population growth estimated at 2.5% per annum, leading to a decline in per capita incomes.

The poor economic performance was largely due to inappropriate agricultural, land, and industrial policies compounded by poor international terms of trade and governance weaknesses. Increased government intrusion into the private sector and import substitution policies made the manufacturing sector uncompetitive. The policy environment, along with tight import controls and foreign exchange controls, made the domestic environment for investment unfavorable.

With the new Government of 2003-2007, 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.

The economy started on an upward trend again in 2009 mainly due to the reawakening of the tourism sector and a lot of investment in the building and construction industry. Macroeconomic stability, low inflation, increased credit to the private sector, better weather conditions, increased remittances from abroad after recovery of the global economy and improved prices of major exports improved the economy in 2010 after suppressed growth in 2008 and 2009, (Republic of Kenya, 2011). This favourable economic growth was however constrained in 2011 due to high food and oil prices, unfavorable weather conditions and foreign exchange market instability. From 2012, the economy has been growing again and it remains to be seen for how long this good performance will continue, (Republic of Kenya, 2013).

1.3 Statement of the Problem

Kenya's GDP growth has been inconsistent since its independence reaching a record high of 0.35% in 1992 and 22.17% in 1971. Over the years, little has been done as pertains to forecasting the GDP growth rate of Kenya. For the purpose of spearheading the economic growth of Kenya, a better understanding of the country's economic situation and future expected economic growth is important to initiate proper policies to enhance more domestic and foreign investment as well as to device measures to curb inflation and control population growth which are among factors

that retard growth in developing economies. This study therefore aimed at forecasting the rate for 2017-2021 so that planners, managers and policy makers know what strategies to adopt in case there will be a fall or rise in the Growth rate.

The overall objective of this study was to model and forecast GDP growth of Kenya. The specific objectives were;

- i. To fit the ARIMA model to the GDP time series data of Kenya from 1961-2016
- ii. To evaluate the best model for forecasting GDP growth.
- iii. To provide a forecast of Kenya's GDP growth for the years 2017-2021.
- iv. To reveal the regularity of Kenya's GDP growth using the fitted ARIMA model.

1.4 Significance of the study

The main significance for this study was to find the best tool that would result in the best forecast for Kenyan GDP growth. The forecast will be important for companies, the stock market, investors (both foreign and local), and researchers in business and particularly in economics. Besides this, the government will be able to set up suitable economic development strategies, economic policies and allocation of funds on different priorities for itself as well as individual firms in a particular industry. Hence, a reliable measure of GDP in some period ahead is needed, which is only possible by a suitable time series model.

2.0 Literature Review

The Gross Domestic Product measures the value of economic activity within a country. Strictly defined, GDP is the sum of the market values, or prices, of all officially recognized final goods and services produced in an economy during a period of time. GDP growth is the change in real GDP that is expressed as a percentage of the previous year.

Economic growth in a country has both a positive and negative impact. Growth inevitably brings a range of economic and social benefits but there are also dangers when an economy rides a fast growth path. One of the benefits of economic growth is that growth stimulates higher employment since an increase in real GDP should cause an outward shift in aggregate demand for labour. Economic growth has an indirect potential to alleviate poverty, as a result of a simultaneous increase in employment opportunities and increase labour productivity (Claire, Renate & Ursula, 2011).

Secondly, growth has a positive effect on government finances as it boosts tax revenues and helps to reduce the budget deficit. Economic growth encourages further investment in new capital machinery via the accelerator mechanism also known as income induced investment.

Sustained economic growth has a positive impact on a company's profits and boosts business confidence. Lastly, growth improves living standards as measured by real GDP per capita although real GDP on its own is not an adequate measure of the true standard of living and quality of life. Kuznets (1962) stated that the welfare of a nation can, therefore, scarcely be inferred from a measure of national income.

On the other hand, the costs resulting from economic growth are as follows; rapid economic growth poses an inflation risk to the country as demand races ahead of the ability of the economy to supply goods and services. Producers take advantage by raising the prices for consumers resulting to inflation. Fast growth can create negative externalities (increased pollution and congestion) which damage the overall welfare. Some critics argue that a narrow view of economic growth, combined with globalization, is creating a scenario where we could see a systemic collapse of our planets' natural resources, (Donnella et. al 2004). Fast growth also results to inequality since not all the benefits of growth are evenly distributed to a country's population, (Alesina et. al 1994). It thus results in regional differences in the distribution of rising income and spending.

2.1 Empirical Literature

Over the years, there has been a considerable number of related studies done on forecasting GDP growth. It is a point to note that most of these studies have been done on developed economies but very few on Africa and in particular in Sub-Saharan countries like Kenya. Lu (2009) used statistical models to collect, analyze, explain and present the time series data of China's GDP from 1962-2008. Comparing with other models, the ARIMA (4,1,0) model was selected as the final model. He applied a Chow test, Chow (1960), to test for an existing data break. He provided a method for prediction and forecasting based on data which was useful to China's government and business.

Gruben and Hayes (1991) constructed a forecasting model that helped predict mild overall growth in the Louisiana economy in 1991. The model predicted expansion in seven of nine economic indicators. According to their forecast, only durable and nondurable goods manufacturing would decline. They concluded that the Louisiana economy was likely to experience expansion if not a boom in 1991.

Adrei and Bugudui (2010) modeled the GDP time series data for USA economy. They used several statistical tests in order to transform the GDP data to being stationary. After applying the tests, the time series became stationary and integrated of order 1. Thus they used the Box-Jenkins procedure for the determination of ARMA and estimated the parameters using OLS technique.

Performances chosen ARMA (1,1,1) model were verified on the basis of classical statistical tests and forecasting.

Maity and Chatterjee (2012) attempted to forecast the growth rate of India's GDP. Using ARIMA(1,2,2), they fitted the model onto data to estimate the parameters of autoregressive and moving average terms. The absolute values of the forecasted GDP indicated an increasing trend but the respective growth rates revealed an opposite trend in future.

Schulz (2004) studied the application of two different unobserved factor models to a data set from Estonia. There is use of small scale state-space model by Stock and Watson (1991) and large static principal components model by Stock and Watson (2002) to derive common factors. Using the factors, forecasts of real economic growth for Estonia were performed and evaluated against benchmark models for different estimation and forecasting methods.

Kraay *et al.*, (1999) consider two alternative methods of forecasting real GDP per capita at various horizons. The two alternatives are; univariate time series models estimated by country and cross-country growth regressions. They evaluate out-of-sample forecasting performance of both approaches for a large sample of industrial and developing countries. They find only modest differences whereby in most cases difference in median (across countries) forecast performance is small relative to large discrepancies between forecasts and actual performance.

Liu *et al.*, (2011) studied real time forecasts of economic activity. They evaluated 'now casts' and 'forecasts' of real GDP growth using five alternative models for 10 different Latin American countries. Results indicated that the flow of monthly data helps to improve forecast accuracy, and the dynamic factor model consistently produced more accurate forecasts relative to other model specifications.

Laungani, *et al.*, (2011) studied information rigidities in forecasts for real GDP growth in 46 countries over the past two decades. They investigated if rigidities are lower around turning points in the economy such as in times of crisis and if rigidities differ across countries particularly between advanced countries and emerging markets and how quickly forecasters incorporate news about growth in other countries into their growth forecasts, with a focus on how advanced country's' growth forecasts incorporate news about emerging market growth and vice versa.

2.2 The Model

The ARIMA model is composed of an Autoregressive component and a Moving Average component. It is discussed in the following sub-sections in more detail.

2.2.1 The Autoregressive Process (AR)

In general, suppose $\{\varepsilon_t\}$ is a white noise process with mean 0 and Variance σ^2 , then a process $\{X_t\}$ is referred to as an AR process of order p if $X_t = \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \dots + \alpha_p X_{t-p} + \varepsilon_t$

$$\sum_{i=1}^p \alpha_i X_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2) \quad (2.1)$$

Where;

α_i is the model parameter to be estimated

ε_t is the error term

This is abbreviated as AR (p). In an AR process, the current value of a time series depends upon the past values of the time series and the random error (white noise).

2.2.2 The Moving Average Process (MA)

In general, suppose $\{\varepsilon_t\}$ is a white noise process with mean 0 and Variance σ^2 , then a process $\{X_t\}$ is referred to as an MA process of order q if it can be expressed in the form

$$X_t = \beta_0 \varepsilon_t + \beta_1 \varepsilon_{t-1} + \dots + \beta_q \varepsilon_{t-q}$$

$$\sum_{i=0}^q \beta_i \varepsilon_{t-i}, \quad \varepsilon_t \sim WN(0, \sigma^2) \quad (2.2)$$

Where;

β_i is the model parameter to be estimated

ε_t is the error term

2.2.3 ARMA Models

ARMA Models are used in time series analysis to describe stationary time series and predict future values in this series. To build an ARMA model, one essentially uses Box-Jenkins methodology (1976). The model is a combination of autoregressive (AR) model and a moving average (MA) model. The ARMA (p, q) model is as follows, where p is the order of the autoregressive part and q the order of the moving average part;

$$\text{ARMA}(p, q): X_t = C + \sum_{i=1}^p \alpha_i X_{t-i} + \sum_{i=0}^q \beta_i \varepsilon_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2) \quad (2.3)$$

Where;

X_t -GDP Growth in percent

X_{t-i} -GDP Growth in percent at period $t - i$

ε_{t-i} -Random shock at period $t - i$

ε_t -Random error term at period t

$(C, \alpha_i \text{ and } \beta_i)$ - The parameters to be estimated

The ARMA is combined with two parts: Autoregressive Model and Moving Average Model

$$\text{AR}(p): = \sum_{i=1}^p \alpha_i X_{t-i} + \varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2)$$

$$\text{MA}(q): X_t = \sum_{i=0}^q \beta_i \varepsilon_{t-i} + \varepsilon_t \sim WN(0, \sigma^2)$$

The error terms ε_t are generally assumed to be independently and identically distributed random variables.

2.2.4 ARIMA Models

In statistics and econometrics, and in particular in time series analysis, an autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) Model. In theory, the most general class of models for forecasting a time series which can be stationary by transformations such as logging and differencing. ARIMA models form an important part of the Box-Jenkins approach to time series modeling. A non-seasonal ARIMA model is said to be classified as an ARIMA(p, d, q) model where: p is the number of autoregressive terms, d is the number of number of times the series has to be differenced before it becomes stationary, q is the number of moving average lags. $\{X_t\}$ Is said to be ARIMA (p, d, q) if;

$$(1 - L)^d \phi^*(L) y_t = c + \theta(L) \varepsilon_t \quad (2.4)$$

Where:

$\phi^*(L)$ Is defined in $\phi(L) = (1 - L)\phi^*(L), \phi^*(z) \neq 0 \forall |z| \leq 1$

And $\theta(L)$ is defined in $\theta(z) \neq 0 \forall |z| \leq 1$

When the process $\{X_t\}$ is stationary iff $d = 0$, in which case it reduces to ARMA (p, q) process:

$$\phi(L)x_t = c + \theta(L)\varepsilon_t, \quad \varepsilon_t \sim WN(0, \sigma^2) \quad (2.5)$$

2.3 Overview of the Box- Jenkins Methodology

The Box-Jenkins methodology is an econometric framework, named after statisticians Box (1933-1982) and Jenkins (1919) which applies autoregressive moving average (ARMA) Models or Autoregressive integrated moving average (ARIMA) Models to time series data to facilitate forecasting. The original model uses an iterative four-stage modeling approach consisting of; model identification, model parameter estimation, diagnostic checking and forecasting. Each stage is analyzed in the following subsections,

2.3.1 Model identification

The foremost step in the process of modeling is to check for the stationarity of the series, as the estimation procedures are for stationary series. Challis and Kitney (1991) define stationarity as a quality of a process in which the statistical parameters (mean and standard deviation) of the process do not change with time. A time series is stationary if $E(X_t) = \mu, Var(X_t) = \sigma^2$

Thus if $x_1, x_2, x_3, \dots, x_n$, values of the time series fluctuate around a constant mean and variance then the time series is stationary. If a time series is non-stationary, first order difference of the time series is calculated. If it exhibits a linear trend and $Var(x_t)$ is constant, first difference will yield a stationary time series. If the $Var(x_t)$ is non-constant, a log transform may yield a time series with constant variance.

A look at the graph of the data and structure of autocorrelation coefficients may provide clues for the presence of stationarity. A sample autocorrelation function at lag k is defined as;

$$r_k = \text{Corr}(z_b, z_{b+k})$$

$$\frac{\sum_{t=b}^{n-k} (z_t - \bar{z})(z_{t+k} - \bar{z})}{\sum_{t=b}^n (z_t - \bar{z})^2} \quad (2.6)$$

Where $\bar{z} = \frac{\sum_{t=b}^n z_t}{n-b+1}$

According to Bowerman and O'Connell (1979) and Box (1976), if a time series is non-stationary, the sample autocorrelation function will neither cut off nor die quickly, but rather it will die down extremely slowly.

Another way of checking for stationarity is to fit a first order autoregressive model for the raw data and test whether the coefficient ‘ ϕ_1 ’ is less than one. A KPSS test is used to verify whether or not the differenced series is stationary. The Augmented Dickey Fuller (ADF) test proposed by Dickey and Fuller (1979) is used to establish whether there exists a unit root. If the model is found to be non-stationary, stationarity could be achieved by differencing the series. For the purpose of analysis, the stationarity property is a good thing to have in one’s data since it leads to many simplifying assumptions.

2.3.2 Model parameter estimation

At the identification stage, one or more models are tentatively chosen that seem to provide statistically adequate representations of the available data. The next step is to obtain precise estimates of parameters of the model by Ordinary Least Square (OLS) as advocated by Box and Jenkins. Standard computer packages like SAS, SPSS etc. are available for finding the estimates of relevant parameters using iterative procedures.

2.3.3 Diagnostics

Different models can be obtained for various combinations of AR and MA individually and collectively. The best model is obtained with the following diagnostics.

The criterion introduced by Akaike in 1969 is used for choosing between competing statistical models and is presented as follows;

$AIC = (-2\text{Log}L + 2m)$, Where $m = p + q + P + Q$ and L is the likelihood function.

Since $-2\text{Log}L$ is approximately equal to $\{n(1 + \text{Log} 2\pi) + n \text{Log} \sigma^2\}$ where σ^2 is the model MSE, AIC can be written as $AIC = \{n(1 + \text{Log} 2\pi) + n \text{Log} \sigma^2 + 2m\}$ and because the first term in this equation is a constant, it is usually omitted while comparing between models.

As an alternative to AIC, the Schwartz Bayesian Information Criteria (SBC) is also used which is given by $SBC = \text{Log} \sigma^2 + \frac{(m \log n)}{n}$

The mean square error of an estimator is one of the ways to quantify the difference between values implied by an estimator and the true values of the quantity being estimated. The MSE of an estimator $\hat{\theta}$ with respect to the estimated parameter θ is defined as;

$$MSE(\hat{\theta}) = E [(\hat{\theta} - \theta)^2] \tag{2.7}$$

Once the appropriate ARIMA model has been fitted, one can examine the goodness of fit by means of plotting the ACF of residuals of the fitted model. If most of the sample autocorrelation coefficients of the residuals are within the limits $\pm 1.96/\sqrt{N}$ where N is the number of observations upon which the model is based then the residuals are white noise indicating that the model is a good fit.

2.3.4 Forecasting

In forecasting, we distinguish between in-sample and off-sample forecasting. In sample forecasting essentially tells us how the chosen model fits a given sample while off-sample forecasting is concerned with determining how a fitted model forecasts future values of the regressors. In forecasting, more consideration is put on the degree of accuracy required, the availability of data and information and the time horizon the forecast is intended to cover.

3.0 Methodology

The research design adopted was exploratory where we aimed at forecasting the Economic Growth rate of Kenya in percent basis for 2017-2021. Yearly GDP percent growth values over the period 1961-2016 were used as a sufficient measure of economic growth. The research fitted the ARIMA model to enable us model and forecast the GDP growth value for the periods.

Secondary data was used in the study. This data was obtained from the IMF, World Economic Outlook and the Kenya National Bureau of Statistics. Data obtained from these sources is considered authentic and can therefore be relied upon to derive conclusions. Such data is also considered credible and free from any error or any bias.

Analysis was facilitated by the Time Series Model Software Version 4.25. The foremost step in ARIMA modeling is to check for stationarity. To achieve this, tests like the unit root Dickey Fuller test, the KPSS and use of the autocorrelation function and correlogram were used. Inspection of the time plots for the GDP growth data also helped to see whether the data is stationary at a glance. Since GDP time series data is non-stationary because it's characterized by changes from one period to another, stationarity was achieved by taking the logarithm of the first order difference of the dependent variable (GDP Growth).

A comparison of various combinations of different ARIMA Models was then carried out to establish the best model to fit. Plotting graphs for sample autocorrelation functions and sample partial autocorrelation functions was useful in determining the p autoregressive terms and q lagged error terms. Applying the Box-Jenkins approach, the values of p and q were chosen from the ACF and PACF plots. For diagnostic checking, the Akaike Information Criteria, Mean Square Error, Hannan-Quinn and Schwartz Bayesian Information Criteria were used where the

best model to fit was established. The final model was estimated by the Gaussian Likelihood estimation method, including the estimation of coefficients $(\varphi_1, \varphi_2, \varphi_3, \varphi_4, \varphi_5, \dots)$. Using the model obtained, at 95% confidence interval, we forecasted Kenyan GDP growth values for the next 5-year period i.e. 2017-2021.

4.0 Results and Discussion

4.1 Preliminary Data Analysis

The GDP Growth data for the whole period portrayed to have a mean percentage rate of 5.121%, standard deviation of 4.126, minimum and maximum value of 0.35% and 22.17% respectively. In order to set the time series, the researcher used the GDP growth data from 1961-2016. Time series values were observed at the same frequency, namely annual. A series plot for GDP Growth data is portrayed in Figure 1.

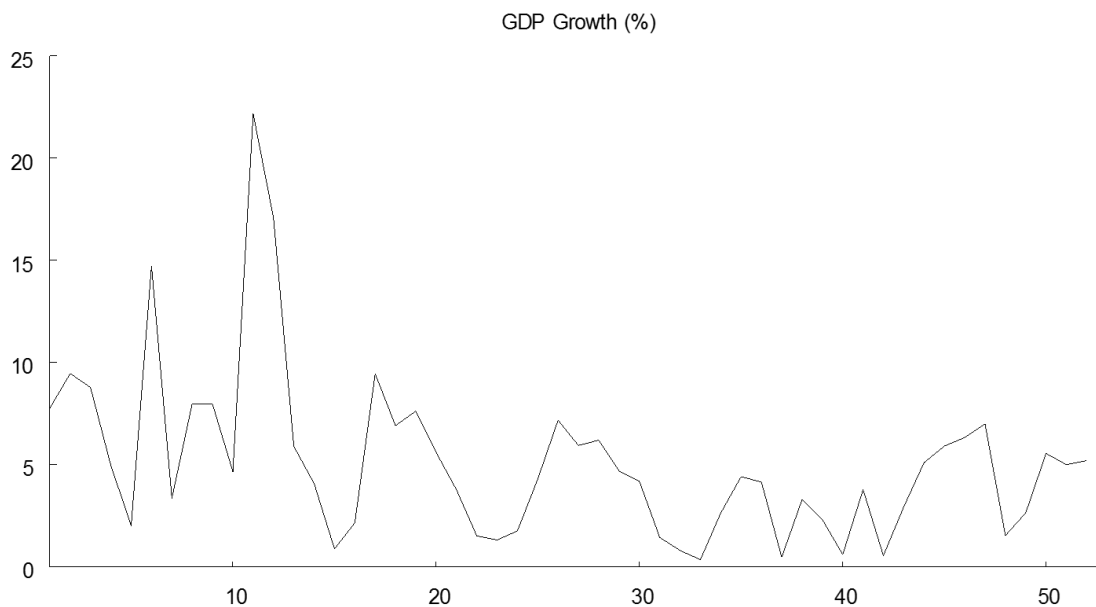


Figure 1: Time series plot for Kenya's GDP Growth

4.2 Testing for Stationarity

A stationary series has a constant mean, constant variance and a constant autocorrelation structure. By inspection of Figure 1, the data seems not to be stationary as the mean and variance are changing with different observations.

A plot of the correlogram of GDP growth is as below.

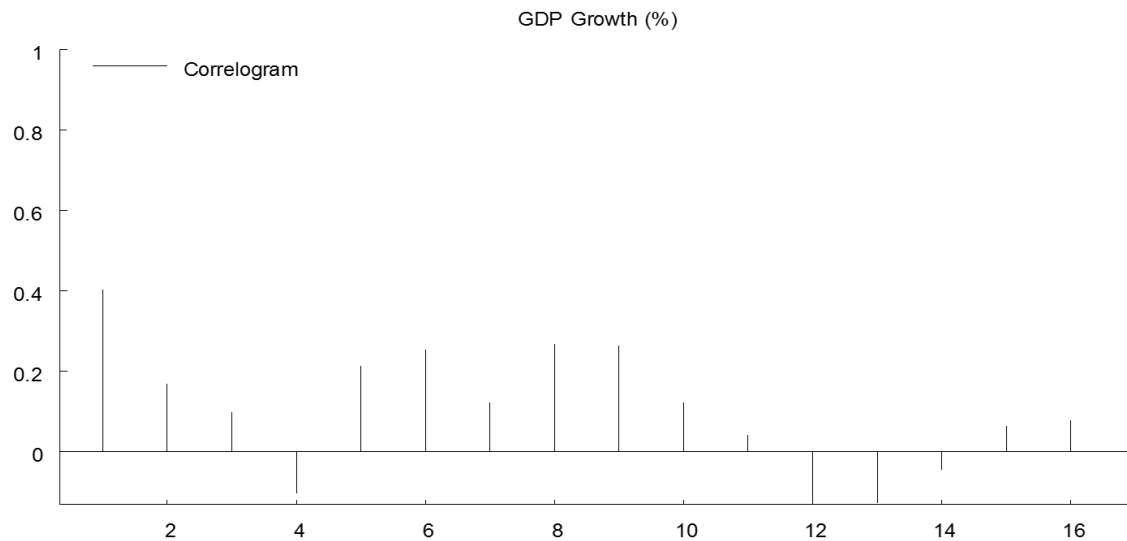


Figure 2: Correlogram of GDP Growth

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 2 indicates the cycles gradually dying away, thus the series is non-stationary. To make the data stationary, we obtain the logarithm of the first order difference of the dependent variable (GDP Growth). The plot for log-differenced GDP growth time series data is shown in Figure 3.

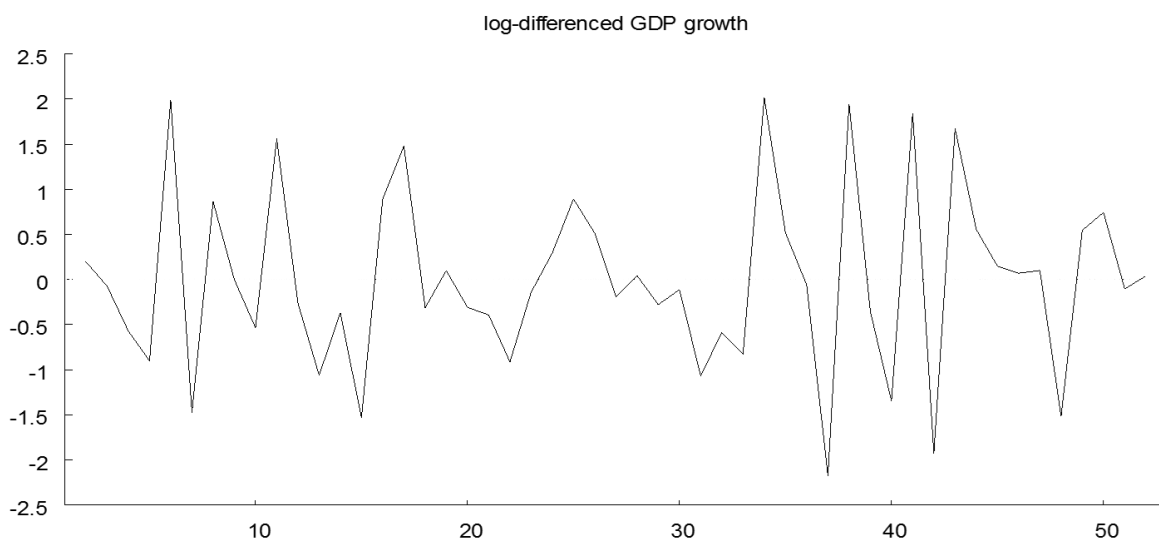


Figure 3: Series plot for Log-Differenced GDP Growth

Summary statistics for the log-differenced GDP Growth data display features of a stationary time series and the results are shown in Table 2.

Table 2 Summary Statistics for Log-Differenced GDP Growth

Statistic	Value
Minimum	-2.178
Maximum	2.020
Mean	-0.008
Standard Deviation	1.004
Skewness	0.203
Kurtosis	2.804
Jarque Bera Statistic	0.433

KPSS test 0.0525 ADF test -6.778

The summary statistics for the Log-differenced GDP Growth series can be interpreted as follows. The mean value is -0.008 a value that tends to zero which is an important feature of stationary time series data. The standard deviation is 1.004. The KPSS and Augmented Dickey Fuller Tests are used to verify whether or not the differenced series is stationary and whether or not there exists a unit root. The p-value (0.0525) of the KPSS is smaller than 0.1 so it rejects the null hypothesis and hence the series is stationary. The p-value (-6.778) of the ADF is smaller than 0.1 so its rejects the null hypothesis that there exists a unit root. From the above results, we find that the differenced series is stationary and there is no unit root.

Further, the series is positively skewed with Skewness equal to 0.203. The Jarque-Bera Statistic by Jarque and Bera (1987) is a test for normality of observations and regression residuals. The test statistic is given by;

$$JB = \frac{n}{6} \left[s^2 + \frac{(k - 3)^2}{4} \right]$$

Where n is the sample size, s is the sample Skewness and k the sample kurtosis.

From the results, the Jarque Bera gives a result of 0.433; this suggests a normal distribution of the series in terms of asymmetry and flattening.

4.3 Estimation results

Box-Jenkins procedure takes into account the value of the d parameter, equal to the initial time series differences order applied to obtain a stationary time series. Since the first differences time series for this data is stationary, the d value is 1. Modeling results that have been estimated by Gaussian Likelihood method of estimation are presented in Table 3. Compared to the Student's t-test, the Gaussian Likelihood method of estimation yielded the lowest SBC, H-Q, AIC Criteria and the lowest Mean Square Error (MSE). Thus it was chosen to be the best method of estimation of the model. The estimation concluded that ARIMA (4,1,2) was the best model to fit.

Table 3 Box-Jenkins Modelling Results of GDP Growth in Kenya

Variable	Coefficient	Std. Error	t-Ratio	P-value
AR(1)	-0.308	1.089	-0.283	0.0001
AR(2)	0.007	0.317	0.023	0.0023
AR(3)	-0.164	0.165	-0.996	0.0025
AR(4)	-0.345	0.258	-1.337	0.0005
MA(1)	0.279	1.908	0.146	0.0000
MA(2)	0.340	1.004	0.338	0.0000

Box-Pierce Q (residuals) = 4.84 {p-value-0.000}

Box-Pierce Q^2 (squared residuals) = 5.5568 {p-value- 0.000}

Table 3 shows the coefficient estimates of various Autoregressive and Moving average schemes of GDP Growth in Kenya. All coefficients are statistically significant at 5 and 10 percent. The Akaike Information Criterion (1974) indicates the usefulness of model selection for non-nested alternative. Smaller values for AIC are preferred. As per this analysis, we have chosen the length of the lag distribution by choosing the lowest value of the AIC which is 65.37. The Schwartz Criterion (1978) is 71.85 and is an alternative to the AIC and penalizes more greatly for additional coefficients. The Box Pierce Q Statistics for the residuals is 4.84 with p-value of 0.963 and that for the squared residuals is 5.557 with p-value of 0.937, the results indicate that the model is adequate since the p-values exceed 5 percent. The Durbin Watson Statistic is the most universally accepted testing criterion for measuring the magnitude of autocorrelation in the residuals from a model. A statistic of 1.88 would lead one to conclude that there is absence of both positive and negative autocorrelation.

4.4 Comparison with other 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 and the Durbin-Watson Statistic, it was considered to be the best model to fit the GDP Data.

Table 4 Evaluation of various ARIMA Models

ARIMA	(0,0,0)	(1,0,1)	(2,0,1)	(3,0,1)	(4,0,1)	(1,0,2)	(2,0,2)	(3,0,2)	(4,1,2)
Log likelihood	72.05	64.03	63.71	66.86	62.51	63.99	61.4	62.41	58.37
SBC	74.01	69.93	71.57	76.69	74.31	71.86	71.14	74.02	71.85
H-Q	73.42	68.14	69.18	73.71	70.72	69.47	68.20	70.53	67.81
AIC	73.05	67.03	67.71	71.86	68.51	68.00	66.40	68.41	65.37
MSE	0.97	0.71	0.70	0.79	0.67	0.71	0.68	0.73	0.63
Resid.Skewnes	0.18	-0.92	-0.93	-0.31	-0.31	-0.92	-0.55	-0.48	-0.34
Resid.Kurtosis	2.80	2.68	2.76	2.20	3.12	2.69	2.53	2.25	2.81
Jarque-Bera	0.361	7.48	7.45	2.15	0.86	7.46	2.91	3.00	1.00
Box -Pierce Q	15.7	11.77	11.22	17.65	4.90	11.78	7.98	13.86	4.84
Box-Pierce Q ²	9.94	7.16	7.46	6.95	7.28	7.69	12.75	6.83	5.55
Durbin Watson	2.69	1.97	2.01	2.09	2.02	1.99	2.09	2.04	1.88

Comparing the different ARIMA models shown in Table 4, the diagnostic checks were performed using the Log likelihood, Schwartz Criterion, Hannan-Quinn Criterion, and the Akaike Information Criteria. The model's MSE is 0.63 which was the smallest as compared to other models fitted. The model with the lowest value of the various criteria was the ARIMA (4,1,2) and was thus chosen as the best model to fit. Residuals of the models fitted were seen to display negative Skewness except for the ARIMA(0,0,0) whose residuals were positively skewed with a Skewness of 0.18. The Jarque Bera Statistic values for all the models fitted showed that they were all normally distributed. The Durbin Watson statistics values for all the

fitted models gave values near 2 and thus all the models indicate absence of both negative and positive autocorrelation.

4.5 The Best Model

Based on the estimation results presented in Table 2, and evaluation of various ARIMA models as presented in Table 3, the model can be re-written as follows.

$$Y = 0.0006 - 0.3076Y_{t-1} + 0.0073Y_{t-2} - 0.1636Y_{t-3} - 0.3447Y_{t-4} + 0.2794\varepsilon_{t-1} + 0.3389\varepsilon_{t-2} + \varepsilon_t \varepsilon_t \sim WN(0, \sigma^2) \quad (5.1)$$

Where;

Y- Represents the GDP Growth in percent

Y_{t-i} , $i = 0,1,2, \dots$ -Represents the GDP Growth in percent at period $t - i$

ε_{t-i} , $i = 0,1,2, \dots$ -Represents the random shock at period $t - i$

t - Represents the time

The model has an intercept value of 0.0006. All the coefficients are statistically significant at 5 percent. The coefficients of Y_{t-2} , ε_{t-1} and ε_{t-2} imply that they are positively correlated to their immediate formers, Y_{t-1} , Y_{t-4} , and ε_{t-1} respectively. The coefficients of Y_{t-3} and Y_{t-4} indicate a negative correlation with the immediate formers which are Y_{t-2} and Y_{t-3} respectively since they have negative values.

4.6 Out-of-Sample Forecasts

Apart from within sample forecasts mentioned in the appendix, the study also estimates five years out-of-sample forecasts of the model in order to measure its forecasting ability. From the Appendix, the model is seen to have a high forecasting power as indicated by small difference between actual and fitted values. A graph showing the forecasted GDP Growth for the next 5-year period is as shown.

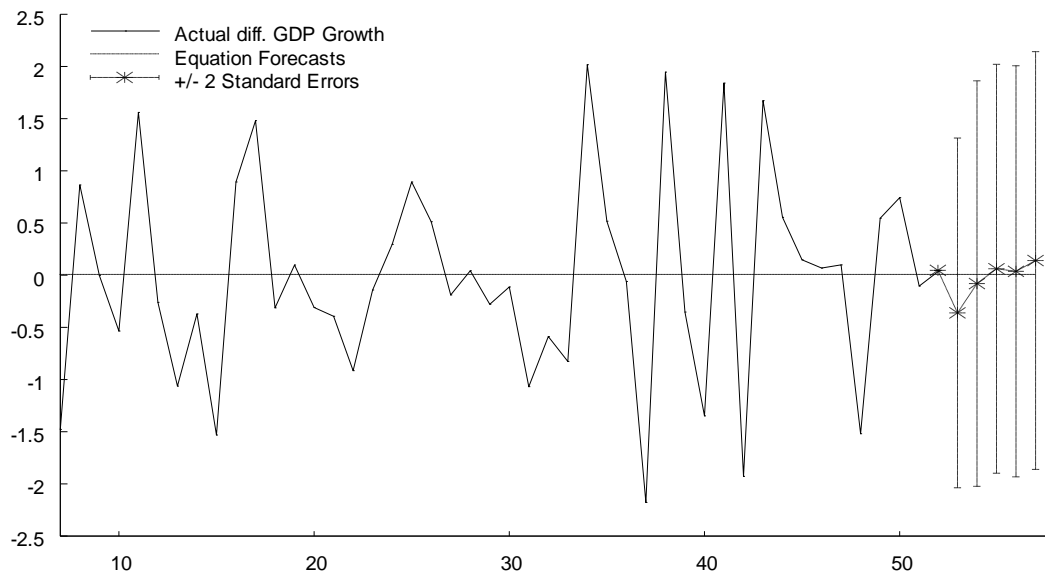


Figure 4: Series plot showing GDP Growth forecasts

The results are shown in Table 5.

Table 5 Five Years- Ahead GDP Growth Forecasts

Year	Point Forecast	Std. Error
2017	5.722	3.616
2018	4.586	3.899
2019	4.922	3.987
2020	5.113	4.051
2021	4.713	4.112

Results indicate that Kenya’s GDP will continue to rise at an annual rate of 5.72% in 2017 but will start falling in 2018 where it will hit a rate of 4.92% in 2019. It is expected to start rising in 2020 followed by a slight drop to 4.713% in 2021.

5.3 Conclusions

In this study, we used statistical methods to collect, analyze, explain and present the time series data of Kenya’s GDP growth from 1961-2016. Comparing with other models, ARIMA (4, 1, 2) Model was selected as the final model for forecasting Kenya’s GDP Growth for 2017-2021. In conclusion, the study found out that Kenya’s GDP will continue to rise but for a short while then start to drop again. 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 and most importantly the political environment as the country moves closer to the 2022 elections.

From the findings, it calls for policy interventions to counter the expected risks following the forecast. From the findings of the study, it is recommended that policy interventions for both domestic and internationally such as the food scale up policy for more incentive for agriculture, maintenance of the recent reduction in taxes on fuel prices, continued infrastructural developments and the sensitization of peace as the country nears the general elections be made to safeguard Kenya's economic growth for specifically 2022.

Acknowledgements

I wish to acknowledge my supervisor and mentor Dr. A. Wagala for his insights in writing this paper.

References

- Allan, S.(2010). *The Environment; from Surplus to Scarcity*. New York, Oxford University Press.
- Ard, H.J, den Reijer. (2010). *Macroeconomic Forecasting using Business cycle leading indicators*, Stolckholm: US-AB
- Bipasha, M. and Bani, C. (2012). Forecasting GDP Growth Rates of India; an Empirical Study. *International Journal of Economics and Management Sciences*, Vol. 1, No. 9, pp. 52-58
- Bowerman, B. L., and O'Connell, R. T. (1979). *Time Series and Forecasting*, Duxbury Press, North Scituate, Massachusetts. USA
- Box, G.E., Jenkins, G. M., and Reinsel, G.C. (1994). *TimeSeries Analysis: Forecasting and Control*. New Jersey: Engelwood Cliffs. USA
- Challis, R. E., and Kitney, R. I. (1991). Biomedical signal processing (in four parts). Part 1 Time-domain methods. *Medical & Biological Engineering & Computing*, vol 28, pp 509-524.
- Claire, M., Renate, H., and Ursula, G. (2011). *Jobs, Growth and Poverty; What do we know, what don't we know, what should we know?* London, Overseas Development Institute.
- Donella, H.M., Jorgen, R., and Dennis, L.M.(2004). *Limits to Growth; The 30-Year Update*. White River Junction, Vermont.

- Elena-Adriana, A. and Elena, B. (2011). *Econometric Modeling of GDP Time Series*. Theoretical and Applied Economics, University of Bucharest, Vol. XVIII, No. 10(563), pp. 91-98
- Jarque, C. M., and Bera, A. K. (1987). *A test for normality of observations and regression residuals*. International Statistical Review, No. 55(2), pp. 163–172
- Ning, W., Kuan –Jiang, B. and Zhi-fa, Y. (2010). *Analysis and forecast of Shaanxi GDP based on ARIMA Model*. Asian Agricultural Research, Vol. 2 No. 1, pp 34-41.
- Makridakis, S., Wheelwright, S.C., and Hyndman, R.J (1998). *Forecasting Methods and Applications*, 3rd Ed, John Wiley, New York.
- Muchiri, B., Audi, P (2004). *Statistical Measure of Growth and Their Changes over Time*, pp 22.
- Kuznets, S(1962). *How to judge Quality*. The New Republic.
- Yang, Lu. (2009). *Modeling and Forecasting China's GDP data with time series model*. D-level essay in Statistics, Department of Economics and Society, Hogskolan Dalarna, Sweden.