ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

MODEL FOR FORECASTING NIGERIAN REAL GROSS DOMESTICS PRODUCT (GDP) USING AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA)

James, Tolulope. O, Babyemi, A.W, Shamsudden A Shayau, and *B. Shehu

Department of Mathematics
Kebbi State University of Science and Technology,
Aliero, Nigeria

**Corresponding Author Email: bsbshehu0@gmail.com*

Abstract

Gross Domestic Product GDP of any given country must also be one of the most difficult measures to predict due to its complexity while modelling. However, this study employed the used of Autoregressive Integrated Moving Average (ARIMA) model for Nigerian GDP 1960 to 2020 data obtained in World Bank database. In the study we found that first difference is insignificant, therefore, we alternatively used a log of the first difference and hence discovered that it is the most suitable model as compare with second difference, where ARIMA (2,1,2) was found with lowest information criteria under parameters estimate. The choosing model was used to forecast the Nigerian GDP using both in sample and out sample prediction method, where 80% of the data was used for training and yield an interesting good performing result with 94.41% accuracy while 20% of the data was presented for testing model (2,1,2) and forecasting.

Keywords: *GDP, ARIMA, forecasting, Akaike Infromation (AIC), Hannan-Quinn (HQ) and Schwarz (SC)*

1.1 Background of the Study

Gross Domestic Product (GDP) is used to represent nations' index that reflects the final results of the production activities in a given state and measures the profits, structure, speed and scale of national economic development. GDP can also be considered as primary index for determining economic development strategic objectives (Zhang *et al.* 2020).

GDP refers to the production of all goods and services of a country or nation within a period of time and is one of the crucial factors of the economy which is to be measured annually. It is the aggregate statistic of all economic activities and captures the broadest coverage of the economy than other macro-economic variables. It is the market value of all final goods and services produced within the borders of a nation in a year. It is often considered the best measure to see how the economy is performing.

GDP can be measured in three different ways. First, the Expenditure approach, which consists of household, business and government purchases of goods and services and net exports. Second the Production approach, it is equal to the sum of the value added at every stage of production (the intermediate stages) by all industries within the country, plus taxes and fewer subsidies on products in the period. Third is Income approach, it is equal to the sum of all factor income generated by production in the country (the sum of remuneration of employees, capital income, and gross operating surplus of enterprises i.e. profit, taxes on production and imports less subsidies) in a period (Yang, 2009).

One of the most popular and frequently used stochastic time series models is Autoregressive Integrated Moving Average (ARIMA) model. The basic assumption made to implement this model is that the considered time series is linear and follows a particular known statistical distribution, such as the normal distribution. ARIMA model has subclasses of other models, such as the Autoregressive (AR), Moving Average (MA) and Autoregressive Moving Average (ARMA) models. For seasonal time series forecasting, Box and Jenkins (Box and Jenkins, 1970) had proposed a quite successful variation of ARIMA model. The popularity of the ARIMA model is mainly due to its flexibility to represent several varieties of time series with simplicity as well as the associated Box-Jenkins methodology for optimal model building process (Balasmeh, *et al.* 2019).

Dynamic nature and time evolvement of data will ever make the forecasting of financial instrument such as real Nigerian Gross Domestic product (GDP) relevant for different researches with different dimension using different statistical techniques. Such was due to the fact that it is critical to precisely estimate the consequences of numerous shocks to the economy by estimating the future dynamics of important aspects of the economy such as GDP, inflation, and exchange rates. Banerjee, *et al.* (2004) investigated forecasting in acceding nations and supported the cautious application of models for forecasting macroeconomic variables. They argued that ARIMA and ANN models have the adequate predictive capacity as seen in (Shahriar *et al.* 2021; Zhang *et al*., 2020; Balasmeh *et al.* 2019). Faisal (2012) argued that the ARIMA model is the best for forecasting GDP in Bangladesh. Salam et al. (2006) investigated the prediction of GDP in underdeveloped countries and recommended that the simple ARIMA (1) model beats the other forecasting models. Olajide, et al. (2012) suggested the ARIMA model (1,1,1) is the best model for projecting the GDP rate as it has the lowest root mean squared error (RMSE).

But today ARIMA (1,1,1) cannot go with the Nigerian GDP due to time involvement and other development carry along in the GDP system, we categorically says that ARIMA (1,1,1) cannot be run or executed (see Table 4.03 for more details). Therefore, Needs arise on reliable estimate of GDP for some period ahead, which is only possible by forecasting GDP as accurately as possible using suitable sophisticated time series modelling. The researchers were motivated to undertake this study dealing with the GDP issues in Nigeria so as to proper a suitable model through the use of ARIMA modelling system. The work will objectively present the forecast of Nigerian GDP using Autoregressive Integrated Moving Average (ARIMA) and performance of the model generated.

COMPUTATIONAL METHODS

2.1 Time Series Analysis using ARIMA Model

ARIMA: the expression ARIMA means Autoregressive Integrated Moving Average. This method was first introduced by Box and Jenkins (1970). It is used to transform a non-stationary time series data by differencing process, known as transformation or removing the trend component represented as order d (1(d)) first difference or order d (2(d)) second difference continuously. An ARIMA(p,d,q) for time series $\{X_t\}$ can be expressed as

$$\boldsymbol{\phi}(\mathbf{L})\Delta^{d} \mathbf{X}_{t} = \boldsymbol{\theta}(\mathbf{L})\boldsymbol{\varepsilon}_{t} \qquad 2.1$$

2.2 Stationarity/Unit Root Test

A time series $\{\boldsymbol{X_t}\}$ is said to be weakly stationary or wide sense stationary or covariance stationary or second order stationary if it satisfies the following three conditions.

i. $\mathbf{E}(\boldsymbol{X_t}) = \boldsymbol{\mu} < \infty$ (i.e constant mean)
ii. $\mathbf{Var}\ (\boldsymbol{X_t}) = \boldsymbol{\sigma} < \infty$(i.e constant variance)
iii. $\mathbf{Cov}(\boldsymbol{X_t}, \boldsymbol{X_{t-j}}) = \boldsymbol{\gamma}_{|j|} < \infty$(i.e covariance is independent of t) and
$|\boldsymbol{j}| = \pm\mathbf{1}, \pm\mathbf{2}, \pm\mathbf{3} \dots\dots, \infty$

2.3 Augmented Dickey-Fuller (ADF) Test

This test was first introduced by Dickey and Fuller in 1979 to test for the presence of unit root and the ADF equation is given as

$$\Delta\boldsymbol{y_t} = \boldsymbol{\alpha y_{t-1}} + \boldsymbol{x_t\delta} + \boldsymbol{\beta_1}\Delta\boldsymbol{y_{t-1}} + \boldsymbol{\beta_2}\Delta\boldsymbol{y_{t-2}} + - - - + \boldsymbol{\beta_p}\Delta\boldsymbol{y_{t-p}} \qquad 2.2$$

The hypothesis testing is

$\boldsymbol{H_0}: \boldsymbol{\alpha} = \mathbf{0}$ (The series contains unit root)

$\boldsymbol{H_1}: \boldsymbol{\alpha} < 0$ (The series is stationary)

The test statistic: $\boldsymbol{t_\alpha} = \widehat{\boldsymbol{\alpha}}/\boldsymbol{se}(\widehat{\boldsymbol{\alpha}})$

Decision rule: Reject $H_0\ if\ t_\alpha$ is less than the asymptotic critical value (tabulated)

Test statistics: $\boldsymbol{KPSS} = \frac{1}{T^2}\sum_{t=1}^{T}\frac{S_t^2}{\hat{\sigma}_\infty^2}$ 2.3

Decision rule: Reject the null hypothesis if the test statistic is greater than the asymptotic critical values (tabulated).

2.4 ARIMA MODEL BUILDING STRATEGIES

The time series ARIMA modeling is a selection of the appropriate model for the data in achieving an iterative procedure based on the four (4) fundamental steps of Box – Jenkins methodology (Box and Jenkins, 1976).

1. Model identification
2. Model estimation
3. Model checking
4. Forecasting

3.4.1 ARIMA (p,d,q) MODEL IDENTIFICATION

The autoregressive integrated moving average [ARIMA (*p, d, q)]* model is given as

$$\boldsymbol{\phi(L)(1-L)^{d}X_{t}=\theta(L)\varepsilon_{t}} \qquad 2.4$$

Where L is the lag operator $\phi(\boldsymbol{L})$ and $\boldsymbol{\theta(L)}$ are the polynomial of orders p and q representing with $\phi_{\boldsymbol{0}} = \boldsymbol{-1}$ and $\boldsymbol{\theta_0 = 1}$.

Our aim is to find the order of AR and MA processes. The order of differencing has been decided already.

The mathematical formulation of AIC is defined as

$$\boldsymbol{AIC(M) = \ln \hat{\sigma}^2_{\varepsilon_t} + 2M} \qquad 2.5$$

3.4.2 METHOD OF ESTIMATING MODEL

These methods make it possible to estimate simultaneously all the parameters of the process, the order of integration coefficient and the parameter of an ARIMA structure. The estimator of the exact maximum likelihood proposed by Sowell (1992a) is the vector

$\boldsymbol{\hat{\beta}} = (\,\boldsymbol{\hat{d}}\,, \hat{\phi}'\boldsymbol{\hat{\theta}'})'$ This maximizes the log-likelihood function $\boldsymbol{L(\beta)}$

$$\mathbf{L}\,(\boldsymbol{\beta}) = -\left(\frac{n}{2}\right)\ln(2\pi) - \left(\frac{1}{2}\right)\ln(\boldsymbol{R}) - \left(\frac{1}{2}\right)\boldsymbol{x'R^{-1}x} \qquad 2.6$$

Where R is the variance, covariance matrix of the process, the matrix R is a complicated algebraic expression and it is difficult to calculate.

DATA ANALYSIS

Gretl is one of the important statistical package used in Time Series analysis. Such a package is used in the respective ARIMA modelling and forecasting in this study.

4.01 ARIMA MODELING PRESENTATION

Table 4.01: Summary Statistics, using the observations 1960 – 2020 for the variable NIGERIANGDP (61 valid observations)

Mean	Median	Minimum	Maximum
1.32855e+011	5.44578e+010	4.19609e+009	5.47000e+011
Std. Dev.	C.V.	Skewness	Ex. kurtosis
1.59398e+011	1.19979	1.27950	0.184527

Sources of Data: World Bank

The lowest Nigerian GDP in the history from 1960 to 2020 is 4.19609e+009 and highest level it attained was 5.47000e+011. The statistic also shows that both the Skewness and Ex. Kurtosis distribution of the data were not good for time series analysis, since the Kurtosis is less than 3 and skewness is high indicating a high deviation from the normality. Hence a need arise to extract a time plot and see how mean of 1.32855e+011 and square of the standard deviation (variance) of the data exist across the timely dat of study. In order word to check if the mean and variance are constant or stationary.

4.02 GRAPHICAL REPRESENTATIONS

Time Series plot of Nigerian GDP 1960 to 2020 before differencing

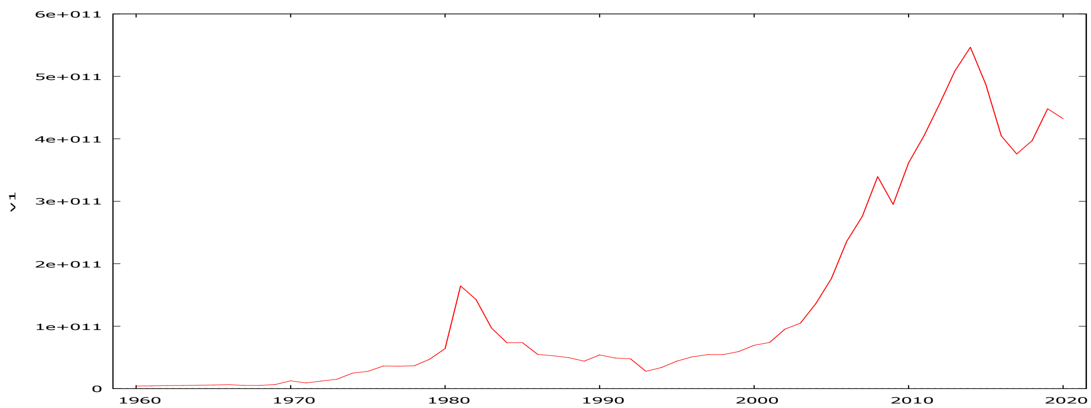

Figure 4.1: Time series plot

Figure 4.1 of time plot shows that from 1960 to the late 70s Nigerian GDP is said to be almost stable, possible shifts in both the mean and the dispersion over time for the series started in the early 70s till early 80s where it rise and slowly decline in 90s. In the early 2000 Nigerian GDP also rise upward and fluctuate in between 2015 to 2020. The mean may be edging upwards, and the variability is increasing more especially in the early 2000. This plot indicates that the data is not stationary and such indicates that a proper transformation is needed for generating a good model.

4.2.1 Correlogram of Nigerian GDP 1960 to 2020 before differencing

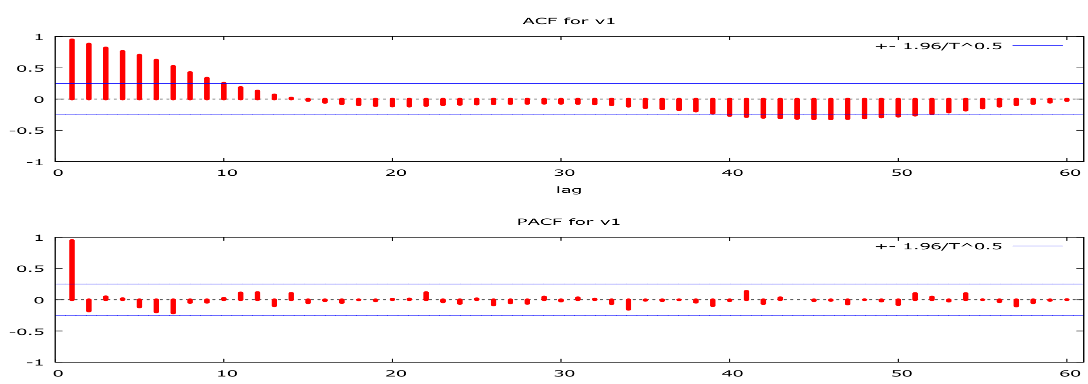

Figure 4.2: Correlogram before difference

ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)
JMSCM, Vol.4, No.2, January, 2023 (Special issue)

Fig 4.2 of Correlogram before taking difference of Nigerian GDP have shown that PACF has a single spike at the first lag and cut, but the ACF shows a decaying pattern. The plot shows more of Moving Average process MA(9) than the Autoregressive process of AR(1) as indicated. Therefore, the data need to be difference in other to obtained a reasonable order of MA(9) process that could be used in model selection.

4.2.2 Time Series plot of Nigerian GDP 1960 to 2020 after first differencing

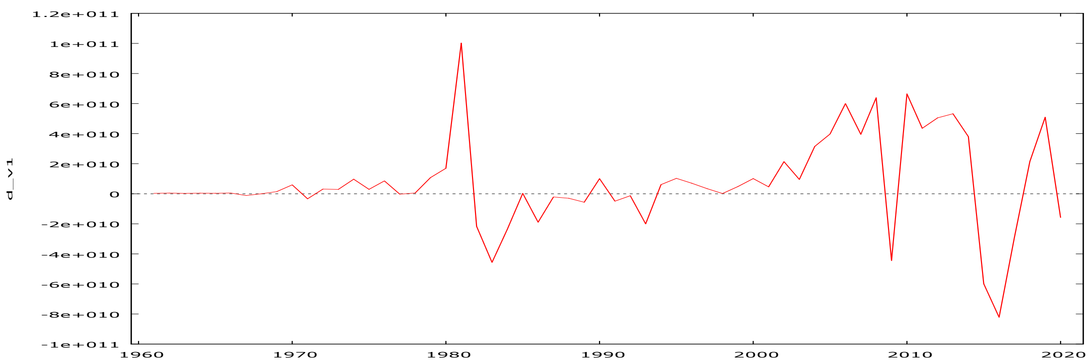

Figure 4.3: Time plot after first difference

The series appeared stationary with respect to central tendency. However, the mean and variance were constant and stable over the periods indicated. Therefore need arise to check the order of MA and AR using correlogram.

4.2.3 Correlogram after First Difference of Nigerian GDP 1960 to 2020

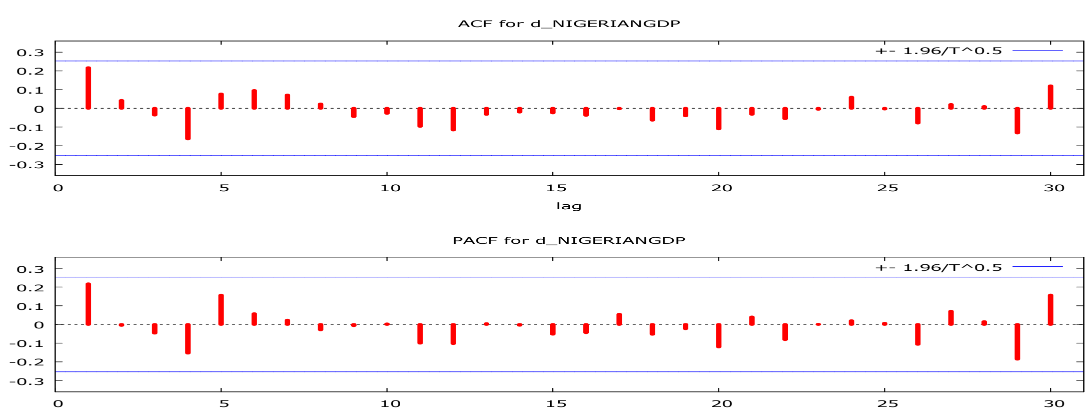

Figure 4.4: **Correlogram** after first difference

The Correlogram in Fig. 4.4 of Nigerian GDP indicates the absence of both AR and MA (order) from PACF and ACF respectively, in other word both AR and MA were not significant. Hence this will affect the proper selection of the best model. Therefore, we alternatively choose to go for log difference and avoid ordinary differencing.

4.2.4 Time Series plot of Nigerian GDP 1960 to 2020 after first Log Difference

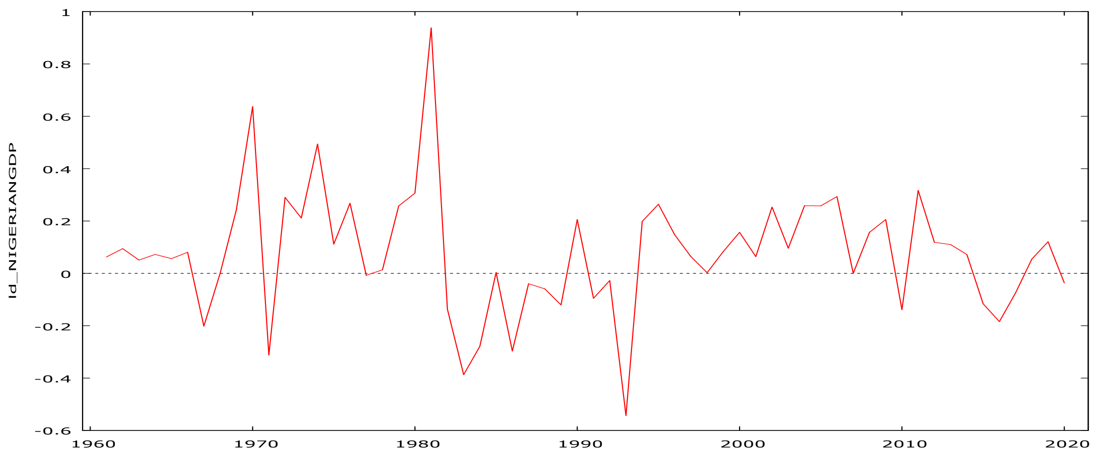

Figure 4.5: Time plot after second difference

The series now appears stationary with respect to central tendency, so further differencing does not appear to be necessary. If correlogram can give us a choosing order of MA and AR

4.2.5: Correlogram after first Log Difference of Nigerian GDP 1960 to 2020

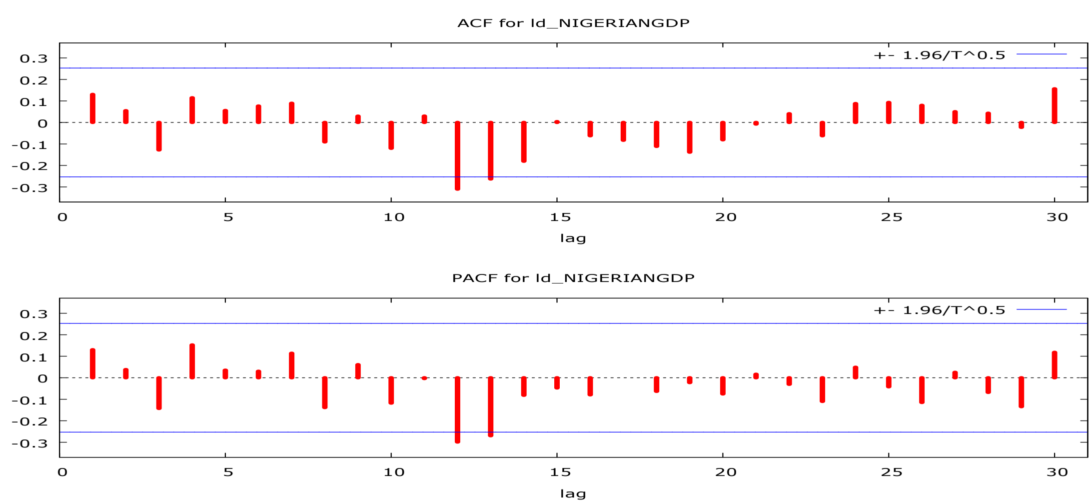

Figure 4.6: **Correlogram** after first Log Difference

Fig. 4.6 of Correlogram depicted that both ACF and PACF were significant, precisely at order (2) each. Fig 4.6 signified MA(2) and AR(2) respectively. This indicates the high possibility of arriving at ARMA (2,2) as our best model might emerge.

4.03 UNIT ROOT TEST AFTER FIRST LOG DIFFERENCE

TABLE 4.02: Result of ADF and KPSS test after first log difference

TEST	LAG Order	TEST STATISTIC	CRITICAL, VALUE/P-VALUE
ADF test with constant	10	-6.60516	3.858e-007
	5	-6.26223	3.858e-007
KPSS, without trend	5	0.0938657	0.351*, 0.469** 0.727***
KPSS, without trend	3	0.0994731	0.351* 0.469** 0.727***

Sources of Data: World Bank **Key:** * at 10% , **at 5% , ***at 1%

The unit root test after first log difference of Nigerian GDP revealed that the data is stationary. The result shows that ADF test statistics is less than critical value, hence we failed to accept the null hypothesis. Therefore, we conclude that the data is stationary. We also found that they converged at lag 4, meaning from lag 4, say 5, 6 etc., yield the same result. KPSS test statistic of both lag 3 and 5 are less than the respective critical values at 10% , 5% and 1% respectively. We therefore, conclude that Nigerian GDP is stationary after taking log of first difference of the original data.

4.03 MODEL IDENTIFICATION

In other to identify the good and better performing model, the following three information criterion, AIC, SC and HQ were tested base on different orders of AR and MA set and we are to determine the criterion with minimum information and consider it as our best model.

TABLE 4.03: Model Identification using Information Criteria

Models	Criteria		
ARIMA Models	Akaike information criterion(AIC)	Hannan-Quinn (HQ)	Schwarz criterion(SC)

ARIMA(2,1,1)	3015.657	3020.523	3028.122
ARIMA(2,1,2)	**3012.413**	**3018.090**	**3026.956**
ARIMA(3,1,2)	3013.717	3020.205	3030.338
ARIMA(3,1,3)	3015.514	3022.813	3034.212
ARIMA(1,1,3)	3016.978	3022.655	3031.521

Source of Data: World Bank

Table 4.03 of Model Identification using Information Criteria revealed that ARIMA (2,1,2) Model has the minimum value of all the criteria. Several models were tested but the presented ones produced the most minimum values while some models failed to be executed e.g (1,1,1), (1,1,2) etc. The present tests have affirmed our correlogram in Fig 4.6 of the first Log Difference. Hence ARIMA (2, 1, 2) emerged as our best model and chose for the forecasting the Nigerian GDP.

4.04 MODEL ESTIMATION

These methods make it possible to estimate simultaneously all the parameters of the process, the order of integration coefficient and the parameter of an ARIMA structure. ARIMA (2,1,2) model was found most significant. We therefore present it parameters in Table 4.04 as an estimated value.

Table 4.04: Model Estimation

Model 2: ARMAX, using observations 1962-2020 (T = 59)
Dependent variable: (1-L) NIGERIANGDP
Standard errors based on Hessian

	Coefficient	*Std. Error*	*z*	*p-value*	
const	6.76093e+09	4.29917e+09	1.5726	0.11581	
phi_1	0.508299	0.0923047	5.5067	<0.00001	***
phi_2	-0.831493	0.0877985	-9.4705	<0.00001	***
theta_1	-0.267336	0.0502647	-5.3186	<0.00001	***
theta_2	1	0.0903753	11.0650	<0.00001	***
ld_NIGERIANGDP	4.36694e+010	8.9619e+09	4.8728	<0.00001	***

Mean dependent var	7.25e+09	S.D. dependent var	3.20e+10
Mean of innovations	-66520350	S.D. of innovations	2.52e+10
Log-likelihood	-1499.207	Akaike criterion	3012.413
Schwarz criterion	3026.956	Hannan-Quinn	3018.090

	Real	*Imaginary*	*Modulus*	*Frequency*

AR				
Root 1	0.3057	-1.0532	1.0967	-0.2050
Root 2	0.3057	1.0532	1.0967	0.2050
MA				
Root 1	0.1337	-0.9910	1.0000	-0.2287
Root 2	0.1337	0.9910	1.0000	0.2287

Source: Gretl

The model parameters were found to be significant by comparing the choosing alpha (α) at 5% level of significant with respective p-values. It can be seen that the p-values are all less than α at 5% level of significant in all the estimates of phi_1, phi_2, theta_1 and theta_2 which represent AR(1), AR(2), MA(1). And MA(2). With this estimated results, we provide more efficient evidence that our model parameters are good for better forecast of Nigerian GDP.

4.05 MODELS CHECKING

After model estimation, the Box – Jenkins model building strategy entails in diagnostic of the adequacy need to be carefully study. More specifically, it is necessary to ascertain in what way the model is adequate and in what way it is inadequate if at all.

The following residuals tests are applied for proper diagnostics model checking.

i. Correlogram of the residuals.
ii. Portmanteau test for residual
iii. ARCH test for residual and
iv. Jarque-Bera test for non normality

4.5.1 RESIDUALS CORRELOGRAM

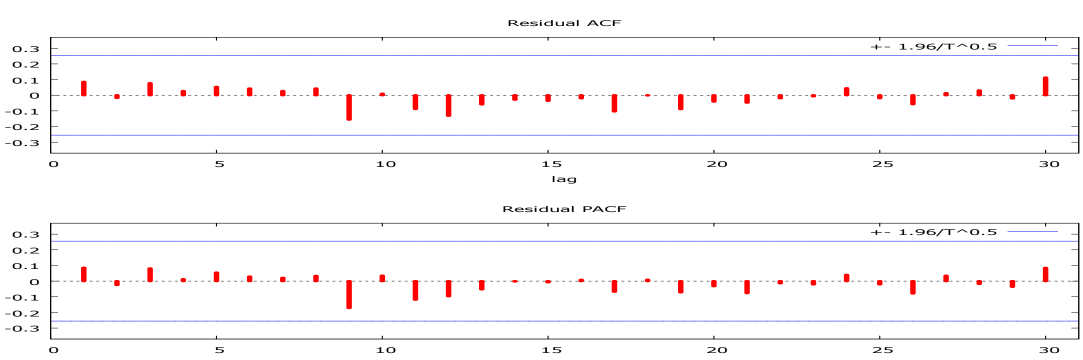

Figure 4.7: Residuals correlogram of the first log difference Nigerian GDP 1960 to 2020

Fig. 4.7 of the residual correlogram shows no significant sign at all lags, this signified that there is no presence autocorrelation in both ACF and PACF and such indicates that the choosing model is good and adequate for Nigerian GDP.

4.5.2 PORTMANTEAU TEST with 12 lags

Portmanteau: 4.6788

p-Value (Chi^2): 0.9679

Ljung & Box: 5.7017

p-Value (Chi^2): 0.9304

We reject $\boldsymbol{H_0}$ if the **P-value (**0.9679 or 0.9304) is less than the significant level $\boldsymbol{\alpha}$ $(\mathbf{0.05})$

Hence 0.05 < 0.9679 or **0.9304**, therefore, we reject H_0 in favour of H_1 and conclude that there is at least one lag with non-zero correlations.

4.5.3 JARQUE-BERA TEST:

test statistic: 10.0660

p-Value(Chi^2): 0.0065

skewness: 0.1628

kurtosis: 5.0328

The test statistics is 10.06 and P-value is 0.0065. Now we have enough evidence to reject the null hypothesis and conclude that the data is not normally distributed. We however suggest to present a plot of normality of residual.

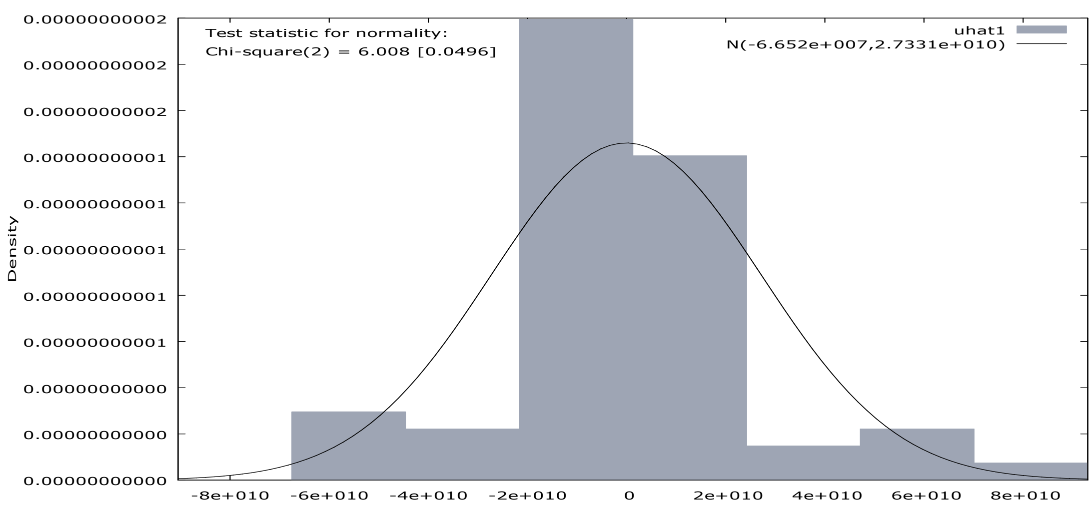

Figure 4.8: Plot of Normality of Residual

The plot in Figure 4.8 of Normality for Residual of the selected model (2,1,2) is a clear indication that the residual is not normal.

ARCH-LM TEST of order 1:

```
             coefficient      std. error   t-ratio   p-value
  -------------------------------------------------------
alpha(0)   4.57274e+020  1.72162e+020    2.656    0.0103  **
alpha(1)   0.339908        0.125711         2.704    0.0091  ***
```

Null hypothesis: no ARCH effect is present

Test statistic: LM = 6.69767

with p-value = P(Chi-square(1) > 6.69767) = 0.00965387

ARCH is test a hypothesis that

$\boldsymbol{H_0}: \boldsymbol{\alpha = 0}$ (The distribution is symmetry)

$\boldsymbol{H_1}: \boldsymbol{\alpha \neq 0}$ (The distribution is asymmetry)

The test statistics has an asymptotic χ^2- distribution. The null hypothesis is rejected if the test statistics is greater than the significant level $\boldsymbol{\alpha}$. Since Test statistic: LM = 6.69767 > 0.05 we reject the null hypothesis and conclude that the distribution is asymmetry

4.06 MODEL FORECAST

ARIMA (2,1,2) MODEL is choosing as the best model for Nigerian GDP 1960 to 2020..

4.6.1 Forecast for Nigerian GDP

Table 4.05: Forecast of Nigerian GDP 2021 to 2030

Obs	NIGERIANGDP	prediction
1960	4.19609e+009	
1961	4.46720e+009	
1962	4.90930e+009	1.26152e+010
1963	5.16549e+009	7.71215e+009
1964	5.55282e+009	1.19381e+010
1965	5.87442e+009	1.12705e+010
1966	6.36679e+009	1.31763e+010
1967	5.20314e+009	-9.44200e+008
1968	5.20090e+009	2.19680e+010
1969	6.63419e+009	2.03001e+010
1970	1.25458e+010	2.46272e+010
1971	9.18177e+009	-2.68962e+010
1972	1.22744e+010	5.34769e+010
1973	1.51629e+010	1.62593e+010
1974	2.48466e+010	2.26873e+010
1975	2.77789e+010	8.95822e+009

1976	3.63089e+010	5.28811e+010
1977	3.60354e+010	3.89501e+010
1978	3.65279e+010	3.61269e+010
1979	4.72599e+010	4.33865e+010
1980	6.42018e+010	5.81127e+010
1981	1.64000e+011	1. 10171e+011
1982	1.43000e+011	1.42260e+011
1983	9.70949e+010	1.43717e+011
1984	7.34844e+010	8.40815e+010
1985	7.37458e+010	6.88046e+010
1986	5.48059e+010	7.58678e+010
1987	5.26760e+010	9.21331e+010
1988	4.96485e+010	4.93187e+010
1989	4.40031e+010	2.87558e+010
1990	5.40358e+010	6.38416e+010
1991	4.91184e+010	6.70884e+010
1992	4.77949e+010	6.41479e+010
1993	2.77522e+010	1.24220e+010
1994	3.38330e+010	5.44564e+010
1995	4.40625e+010	5.01062e+010
1996	5.10758e+010	5.55300e+010
1997	5.44578e+010	5.18288e+010
1998	5.46041e+010	4.93066e+010
1999	5.93726e+010	6.38022e+010
2000	6.94488e+010	7.59950e+010
2001	7.40304e+010	7.43004e+010
2002	9.53858e+010	8.37269e+010
2003	1.05000e+011	9.36101e+010
2004	1.36000e+011	1.26753e+011
2005	1.76000e+011	1.51837e+011
2006	2.36000e+011	1.89533e+011
2007	2.36000e+011	2.39563e+011
2008	2.76000e+011	2.55293e+011
2009	3.39000e+011	2.84427e+011
2010	2.95000e+011	3.41877e+011
2011	4.05000e+011	3.23435e+011
2012	4.56000e+011	4.08558e+011
2013	5.09000e+011	4.86235e+011
2014	5.47000e+011	5.33610e+011
2015	4.87000e+011	5.41980e+011
2016	4.05000e+011	4.61077e+011
2017	3.76000e+011	3.83199e+011
2018	3.97000e+011	3.86731e+011
2019	4.48000e+011	4.35111e+011
2020	4.32000e+011	4.68289e+011

ISSN 2688-8300 (Print) ISSN 2644-3368 (Online)

The data used in the study range from 1960 to 2020 of Nigerian GDP, we forecast no future value here, the essence of this forecast is to generate the general performance of the model selected. The evaluation started from 1962 where the first prediction was found.

4.6.2 Forecast evaluation statistics

Table 4.05: Forecast evaluation statistics

Mean Error	**-6.652e+007**
Mean Squared Error	6.7103e+020
Root Mean Squared Error	2.5904e+010
Mean Absolute Error	1.8319e+010
Mean Percentage Error	-18.277
Mean Absolute Percentage Error	48.15
Theil's U	3.1753
Bias proportion, UM	6.5942e-006
Regression proportion, UR	0.0001398
Disturbance proportion, UD	0.99985

These are the set of errors committed from the original GDP in the prediction process. These are the values used for comparisms and these are the value set to know if the researcher have committed errors above expectation or the choosing model have done well or not.

4.6.3 Forecast Plot of Nigerian GDP

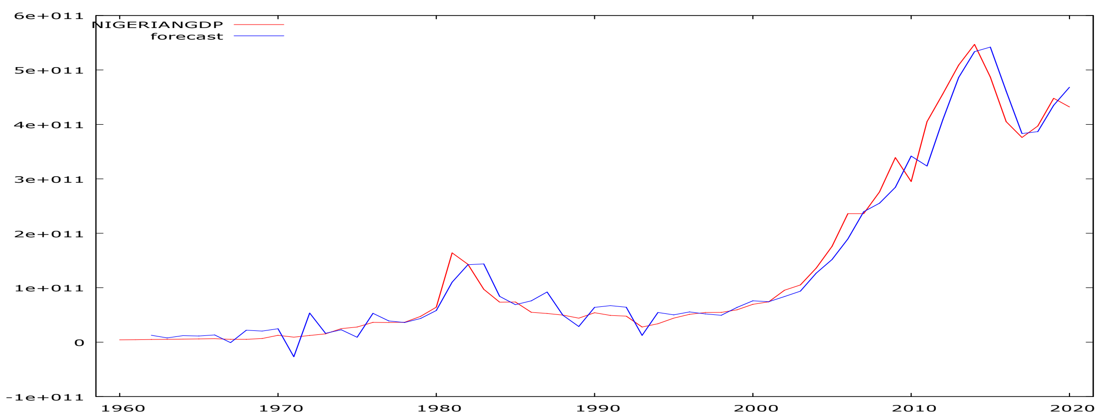

Figure 4.9: Plot of Forecast against Actual

Red line shows the original Nigerian GDP while the blue line is the generate forecast of ARIMA (2,1,2) which shows a close relationship with almost similar pattern. With this plot we can graphically accept our chosing model is good and adequate.

4.6.4 in sample Forecast for Nigerian GDP

Using the same model ARIMA (2,1,2) we trained 80% of the data and used 20% for prediction. This forecast is also another way of checking model performance.

Table 4.06: Forecast with 20% tested data

For 95% confidence intervals, z(0.025) = 1.96

Obs	NIGERIANGDP	Prediction	std. error	95% interval
2018	3.97000e+011	3.75125e+011	2.52937e+010	(3.25550e+011, 4.24700e+011)
2019	4.48000e+011	4.04202e+011	4.00683e+010	(3.25670e+011, 4.82735e+011)
2020	4.32000e+011	4.25526e+011	5.64424e+010	(3.14901e+011, 5.36151e+011)

Table 4.07: Forecast evaluation statistics

Mean Error	2.4049e+010
Mean Squared Error	8.1289e+020
Root Mean Squared Error	2.8511e+010
Mean Absolute Error	2.4049e+010
Mean Percentage Error	5.595
Mean Absolute Percentage Error	5.595
Theil's U	0.83447
Bias proportion, UM	0.71147
Regression proportion, UR	0.031095
Disturbance proportion, UD	0.25744

With this forecast using ARIMA (2,1,2) both Mean Percentage Error and Mean Absolute Percentage Error revealed that we committed an error of 5.595%. Therefore, we are 94.41% confidence with the chosing model when predicting a short term of Nigerian GDP.

Table 4.08: Out Sample Forecasted of Nigerian GDP 2021 to 2030

For 95% confidence intervals, z(0.025) = 1.96

Obs	Prediction	std. error	95% interval
2021	4.27264e+011	2.91888e+010	(3.70055e+011, 4.84473e+011)
2022	4.21839e+011	4.38160e+010	(3.35961e+011, 5.07717e+011)
2023	4.33815e+011	5.76201e+010	(3.20882e+011, 5.46748e+011)
2024	4.53521e+011	6.80835e+010	(3.20080e+011, 5.86962e+011)
2025	4.60964e+011	7.50559e+010	(3.13858e+011, 6.08071e+011)
2026	4.56538e+011	8.10922e+010	(2.97601e+011, 6.15476e+011)
2027	4.58127e+011	8.81186e+010	(2.85418e+011, 6.30836e+011)
2028	4.72698e+011	9.55006e+010	(2.85520e+011, 6.59876e+011)
2029	4.87248e+011	1.01515e+011	(2.88283e+011, 6.86213e+011)
2030	4.90282e+011	1.06252e+011	(2.82033e+011, 6.98531e+011)

Using ARIMA (2,1,2) by 2030 Nigeria is expected to bring about 4.90282e+011 or 490282000000 of its annual GDP.

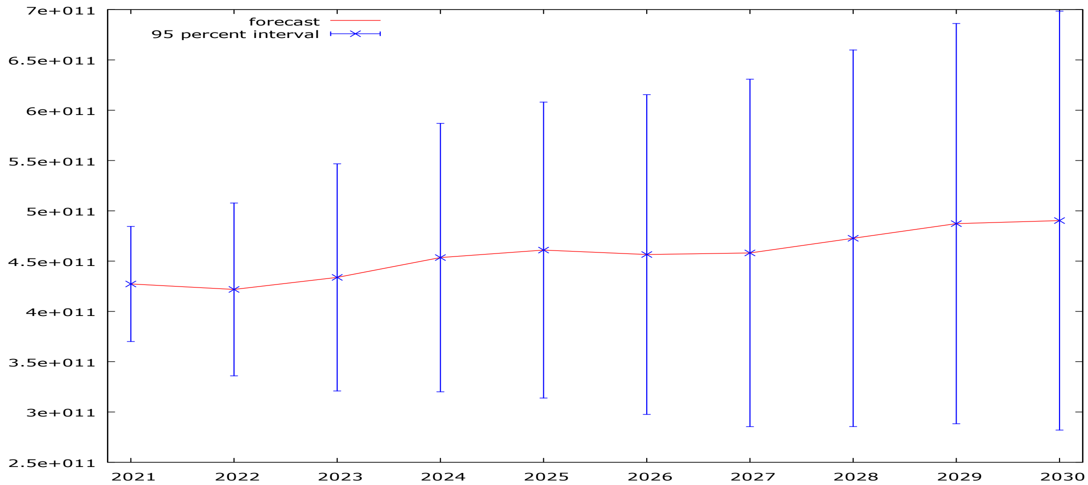

Figure 4.10: Graph representing the Forecast values 2021 to 2030

Fig. 4.10 is a graphical representation of the forecasted Nigerian GDP ranging from 2021 to 2030, which laid an expectation of little increase from the present time. Therefore Nigerian needs to do well to meet the expectation or go above it.

Conclusions

GDP refers to the sum of the annual output of the country's farming, production and expenditure among others. In this study we model Nigerian GDP 1960 to 2020 and able to generate a model that forecast the expected output of Nigerian GDP ahead of 2021 to 2030. When dealing with Model Identification, the Information Criteria revealed that ARIMA (2,1,2) Model has the minimum values of AIC= 3012.413, HQ= 3018.090 and SC= 3026.956. The model parameters were found to be significant by comparing the choosing alpha (α) at 5% level of significant with their respective p-values. Where p-values are all less than α at 5% level of significant in all the estimates of phi_1= 0.00001, phi_2= 0.00001, theta_1= 0.00001 and theta_2= 0.00001 which represent AR(1), AR(2), MA(1). And MA(2) respectively. We finally used both in sample and out sample forecasting method and generated the expected forecast in ten years ahead. We found and conclude that the first difference cannot go with the present Nigerian GDP data due to the fact that it loses order of both AR and AM as seen in Fig 4.4, hence the present study used first order (Log Difference).

References

1. Balasmeh, A. O., Babbar, R., & Karmaker, T. (2019). Trend analysis and ARIMA modeling

for prediction precipitation pattern in Wadi Shueib catchment area in Jordan. Arabian Journal of Geosciences, 12:27.

2. Banerjee, A., Marcellino, M., & Masten, I. (2004). Forecasting Macroeconomic Variables for

the Acceding Countries. IGIER WP, 260. Choudhury, J. P., Sarkar, B., & Mukherjee, S. K. (2002). Forecasting of engineering manpower through fuzzy associative memory neural network with ARIMA: a comparative study. Neurocomputing, 47(1), 241-257.

3. Balasmeh, A. O., Babbar, R., & Karmaker, T. (2019). Trend analysis and ARIMA modeling

for prediction precipitation pattern in Wadi Shueib catchment area in Jordan. Arabian Journal of Geosciences, 12:27.

4. Olajide, J. T., Ayansola, O. A., Odusina, M. T., and Oyenuga, I. F. (2012). Forecasting the

Inflation Rate in Nigeria: Box Jenkins Approach. IOSR Journal of Mathematics. 3(5), 15-19.

5. Shahriar, S. A., Kayes, I., Hasan, K., Hasan, M., Islam, R., Awang, N. R., Hamzah, Z., Rak,

E., & Salam, M. A. (2021). Potential of ARIMA-ANN, ARIMA-SVM, DT and CatBoost for Atmospheric PM2.5 Forecasting in Bangladesh. Atmosphere, 12, 100.

6. Yang, L. (2009) Modeling and Forecasting China's GDP Data with Time Series Models. D-

level Essay in Statistics. Department of Economics and Society, Hogskolan Dalarna, Falun.

7. Zhang, Y., Yang, H., Cui, H., & Chen, Q. (2020) Comparison of the ability of ARIMA,

WNN and SVM models for drought prediction in the Sanjiang Plain, China. Natural Resources Research, 29(6), https://doi.org/10.1007/s11053-019-09512-6