Archives of Current Research International

20(7): 49-61, 2020; Article no.ACRI.62569
ISSN: 2454-7077

ARIMA Model for Gross Domestic Product (GDP): Evidence from Nigeria

Atanu, Enebi Yahaya[1*], Ette, Harrison Etuk[2], Nwuju, Kingdom[2] and Nwaoha, William Chimee[3]

[1]Department of Statistics, Federal Polytechnic of Oil and Gas, Bonny Island, Rivers State, Nigeria.
[2]Department of Mathematics/Computer Science, Rivers State University of Science and Technology, Port Harcourt, Nigeria.
[3]School of General Studies, Federal Polytechnic of Oil and Gas, Bonny Island, Rivers State, Nigeria.

Authors' contributions

This work was carried out in collaboration among all authors. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/ACRI/2020/v20i730213
Editor(s):
(1) Dr. Faisal, Abdul Wali Khan University, Pakistan.
Reviewers:
(1) Samuel Asuamah Yeboah, Sunyani Technical University, Ghana.
(2) C. G. Amaefula, Federal University Otuoke, Nigeria.
Complete Peer review History: http://www.sdiarticle4.com/review-history/62569

Original Research Article

Received 10 September 2020
Accepted 14 November 2020
Published 12 December 2020

ABSTRACT

A nation's GDP is an important index reflecting development in economy and incomes. This paper uses the annual data of Nigeria's GDP from 1981 to 2019 as the research data. An Augmented Dick Fuller test was used to test for stationarity of the data and was seen to be stationary at the second differencing. ARIMA (1, 2, 1) was identified as an appropriate model using Eviews 11 software after comparing the AIC values. The Ljung-Box test of the Residual satisfied that the model was adequate and was used to forecast the out of sample data. And with a Theil inequality of 0.022008, the model forecasting ability is deemed be a good.

Keywords: GDP; ARIMA modelling; residual analysis; forecasting.

**Corresponding author: Email: eyatanu@gmail.com;*

1. INTRODUCTION

One of the major objectives of time series model is to forecast future values or activities by studying the behavioral pattern of past data. In government and large organizations, long and short term planning is carried out on the basis of the analysis of past data of various economic variables. In all areas of human endeavor, when we intend to make forecast about the future, our previous experience, if it exist is usually relied on.

Time series analysis deals with the statistical technique of analyzing past data in other to obtain estimates for future values. This is usually done by collecting data on past observations and making forecast about the future.

Box and Jenkins [1], developed a practical procedure for an entire family of models, the autoregressive integrated moving average or ARIMA, applicable to stationary data series where the mean, the variance and the autocorrelation function remains constant through time.

According to Anderson [2] and Pankratz [3], the Box-Jenkins approach is a powerful and flexible method for forecasting because it places more emphasis on the recent past data and where structural shift occur gradually rather than suddenly which makes the ARIMA model valuable when dealing with economic time series data.

The model is generally referred to as an ARIMA (p, d, q) model where p is the autoregressive component, q is the moving average component and d is differencing and are all integers greater than or equal zero. Hence, an ARIMA model describes the non-stationary behavior that can be differenced to obtain a stationary process which is beneficial in modelling Gross Domestic Product (GDP).

Gross domestic product (GDP) refers to the market value of all the final products (goods and services) which is produced or provided by economic society i.e. either a country or a region in a given period. GDP is an important indicator to measure a country's wealth and economic strength. GDP is part of the National income and product accounts which are statistics that enable policy makers to determine whether the economy is contracting or expanding and if either a recession or inflation beckons. GDP is used by economic to determine the level of development of a country.

Gross domestic product (GDP) comprises of consumption(C), investment (I), government (G) purchase of goods and services and net exports (X) produced within the nation during that period.

Hence, GDP = C+I+G+X.

The objective of this paper is to use ARIMA model to model the stochastic mechanism that rise to the GDP series and to forecast future values of the series based on the history of the series. There are many studies that used these models for studying the GDP in different countries, such as Wabomba et al. [4], and Uwimana et al. [5].

2. REVIEW OF RELATED LITERATURE

Box and Jenkins [1] methodology has been used severally by many researchers in highlight the future rates of gross domestic product (GDP). Dritsaki (2015), in his study of Greek GDP utilize ARIMA (1,1,1) in modelling and forecasting 1980-2013 Greece’s GDP rate. Zakai [6] in investigating the forecast of Gross Domestic Product (GDP) for Pakistan using quarterly data from 1953-2012 choose an ARIMA (1,1,0) model and found out that the size of the increase for Pakistan’s GDP for the years 2013- 2025. Bluiyan et al. [7] studied the modelling and forecasting of Gross Domestic Product of manufacturing industries in Bangladesh. The non-stationary data was made stationary by taken second difference of the data. ARIMA (2,2,0) ARIMA (2,2,1), and ARIMA (2,2,2) were selected on the basis of their Akaike information criteria. And finally ARIMA (2,2,1) was selected based on smallest value of standard error and the result shows a GDP of sustainable upward trend and that the estimated value fits the data very well and forecast was made for next thirteen years beginning from 2002/2003. Abiola and Okafor [8], examined the various forecasting models for the Nigerian crude oil prices from 2005Q1 to 2012Q4. The study discovered that ARIMA (1, 1, 4) model is best fitted forecasting model for predicting Nigerian crude oil price benchmark. Iwueze et al. (2013) initially fitted ARMA(1,0,0) to the non-stationary data series. After differencing, ARIMA (2,1,0) was fitted to the Nigeria External Reserves. From the results, ARIMA (2,1,0) provided better estimates than the initial ARMA (1,0,0) which was fitted to the data series.

3. METHODOLOGY

The time series analysis can provide short-run forecast for sufficiently large amount of data on the concerned variables very precisely, see Granger and Newbold [9]. In univariate time series analysis, the ARIMA models are flexible and widely used. The ARIMA model is the combination of three processes: (i) Autoregressive (AR) process, (ii) Differencing process, and (iii) Moving-Average (MA) process.

4. AUTOREGRESSIVE (AR) PROCESS

Autoregressive models are based on the idea that current value of the series, X_t, can be explained as a linear combination of ρ past values, $X_{t-1}, X_{t-2}, \dots, X_{t-\rho}$, together with a random error in the same series. An autoregressive model of order ρ, abbreviated $AR(\rho)$, is of the form:

$$X_t = \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \dots + \emptyset_\rho X_{t-\rho} + w_t = \sum_{i=1}^{\rho} \emptyset_i X_{t-i} + w_t \quad (3.1)$$

where X_t is stationary, $w_t \sim wn(0, \sigma_w^2)$, and $\emptyset_1, \emptyset_2, \dots, \emptyset_\rho$ are model parameters. The hyper parameter p represents the length of the series.

5. MOVING AVERAGE (MA) PROCESS

In AR models above, current observation X_t is regressed using the previous observations $X_{t-1}, X_{t-2}, X_{t-3}, \dots, X_{t-p}$, plus an error term w_t at current time point. One problem of AR model is the ignorance of correlated noise structures (which is unobservable) in the time series. { In other words, the imperfectly predictable terms in current time, w_t, and previous steps, $w_{t-1}, w_{t-2}, w_{t-3}, \dots, w_{t-q}$, are also informative for predicting observations.

A moving average model of order q, or MA(q), is defined to be

$$X_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \theta_3 w_{t-3} + \cdots + \theta_q w_{t-q} = w_t + \sum_{j=1}^{q} \theta_j w_{t-j} \quad (3.2)$$

Where $w_{t \sim wn(0,\sigma^2)}$, and $\theta_1, \theta_2, \theta_3, \dots, \theta_q$ $(\theta_q \neq 0)$ are parameters.

5.1 ARMA Model

In the statistical analysis of time series, the class of autoregressive-moving-average (ARMA) models is mostly utilized for the prediction of second-order stationary stochastic process. The ARMA model is a tool for understanding and analyzing the causal structure, or to obtain the predictions of the future values in this series. The model consists of two parts, one for autoregressive (AR) and the second for moving average (MA). The model is usually referred to as the ARMA (p, q) process where p is the order of the autoregressive part and q is the order of the moving average part.

A second-order stationary process (X_t) is called an ARMA (p, q)

process, if there exist real coefficients $c, \emptyset_1, \emptyset_2, \emptyset_3, \dots, \emptyset_p, \theta_1, \theta_2, \theta_3, \dots, \theta_q$, where p and q are integers, so

$$X_t - \sum_{i=1}^{p} \emptyset_i X_{t-i} = c + w_t + \sum_{j=1}^{q} \theta_j \epsilon_{t-j}, \ \ \forall_t \in \mathbb{Z} \quad (3.3)$$

where $\{\epsilon_t\}$ *is the white noise* $(0, \sigma^2)$.

Let's denote *B* as the back-shift operator such that $B^k X_t = X_{t-k}$. Using *B*, rewrite the ARMA (p, q) equation above as $\emptyset(B) X_t = \theta(B) w_t$. (3.4)

5.2 ARIMA Models

The ARMA models can further be extended to non-stationary series by allowing the differencing of the data series resulting to ARIMA models. The general non-seasonal model is known as ARIMA (p, d, q): where with three parameters; p is the order of autoregressive, d is the degree of differencing, and q is the order of moving-average. For example, if X_t is non-stationary series, we will take a first-difference of X_t so that ΔX_t becomes stationary, then the ARIMA (p, d, q) model is:

$$\hat{X}_t = \mu + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \emptyset_3 X_{t-3} + \cdots + \emptyset_p X_{t-p} - \theta_1 w_{t-1} - \theta_2 w_{t-2} - \theta_3 w_{t-3} - \cdots - \theta_q w_{t-q} \quad (3.5)$$

After considering the differencing for an ARMA model in order to be able to extended the model to a non-stationary series, we have:

$$\nabla^d X_t = \emptyset_1 \nabla^d X_{t-1} + \emptyset_2 \nabla^d X_{t-2} + \emptyset_3 \nabla^d X_{t-3} + \cdots + \emptyset_p \nabla^d X_{t-p} + w_t + \theta_1 \nabla^d w_{t-1} + \theta_2 \nabla^d w_{t-2} + \theta_3 \nabla^d w_{t-3} + \cdots + \theta_p \nabla^d w_{t-p}.$$

where $\{w_t\}$ is the error term in the equation; a white noise process, a sequence of independently and identically distributed (iid) random variables with $E\ (w_t)$ = 0 and $var\ (w_t)$ = σ^2 and $\emptyset$'s and θ's are the model parameters.

The autoregressive (AR) order may be determined by the lap at which the partial autocorrelation function (PACF) cuts off. The moving average (MA) order may be estimated as the lap at which the autocorrelation function (ACF) cuts oft. Estimation of α's and β's may be done by the method of lest squares.

6. RESULTS AND DISCUSION

Autoregressive integrated moving average was used to determine an appropriate model for estimating Nigeria's annual Gross Domestic Product (GDP). The data used in this paper is the yearly Nigeria's GDP data from 1981 to 2019. The data was obtained from the Central Bank of Nigeria (CBN. These data are transformed into logged data in other to stabilize the variance [10].

From the logged time plot of the real gross domestic product above, it can be observed that the data shows a certain trend. Hence, we can check the data's stationarity, correlogram and randomness in other to identify a suitable model for the series.

Testing for stationarity: Since ARIMA model can only be applied to non-stationary time series data only when the data is stationary. Then, before we perform the analysis of the time series data it is expected that we determine the stationarity of the data. The stationarity test of the GDP is performed by Augmented Dick Fuller (ADF) Test. The test result obtained is shown in the Tables below.

Table 1 shows the ADF result for the actual real gross domestic plot data, showing its non-stationarity behavior.

Table 1. Augmented Dickey- Fuller Test

Null Hypothesis: LGDP has a unit root				
Trend Specification: Intercept only				
Break Specification: Intercept only				
Break Type: Innovational outlier				
Break Date: 1991				
Break Selection: Minimize Dickey-Fuller t-statistic				
Lag Length: 0 (Automatic - based on Schwarz information criterion, maxlag=9)				
			t-Statistic	**Prob.***
Augmented Dickey-Fuller test statistic			-4.666607	0.0271
Test critical values:	1% level		-4.949133	
	5% level		-4.443649	
	10% level		-4.193627	
*Vogelsang (1993) asymptotic one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: LGDP				
Method: Least Squares				
Date: 09/19/20 Time: 18:24				
Sample (adjusted): 1982 2019				
Included observations: 38 after adjustments				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
LGDP(-1)	0.953616	0.009940	95.94142	0.0000
C	0.385739	0.060058	6.422781	0.0000
INCPTBREAK	0.257472	0.053380	4.823344	0.0000
BREAKDUM	-0.178643	0.091799	-1.946029	0.0600
R-squared	0.998803	Mean dependent var		8.716068
Adjusted R-squared	0.998697	S.D. dependent var		2.320507
S.E. of regression	0.083762	Akaike info criterion		-2.022365
Sum squared resid	0.238548	Schwarz criterion		-1.849988
Log likelihood	42.42494	Hannan-Quinn criter.		-1.961035
F-statistic	9454.274	Durbin-Watson stat		1.205090
Prob(F-statistic)	0.000000			

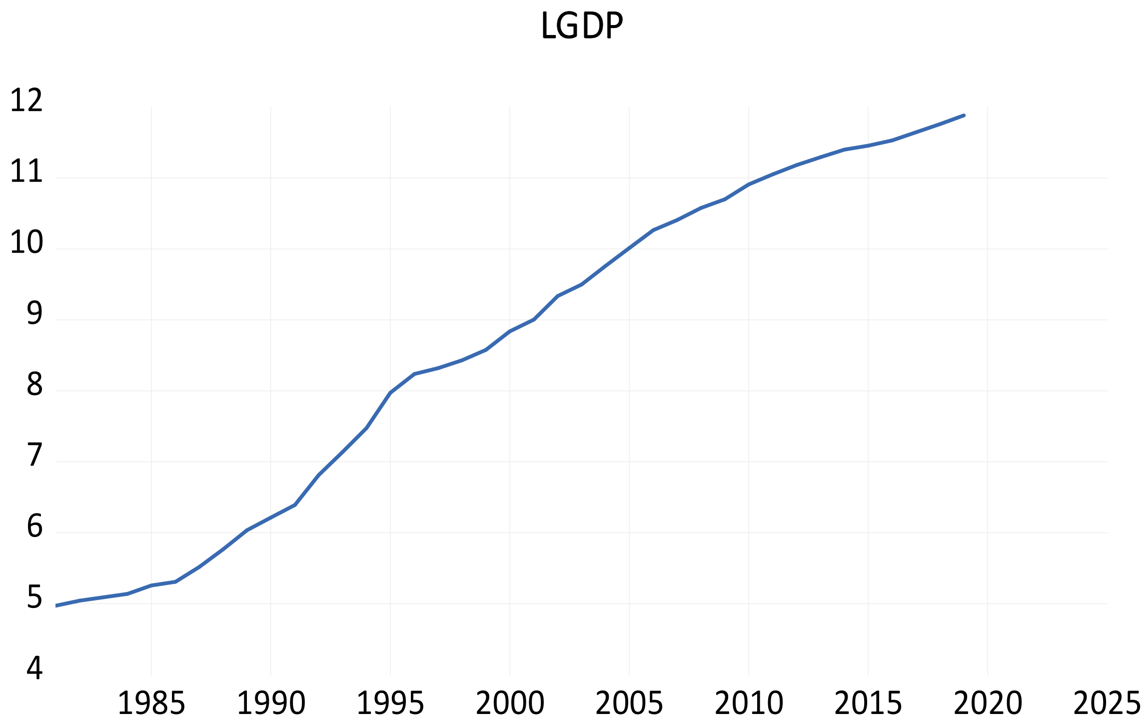


Fig. 1. Time plot of the real GDP

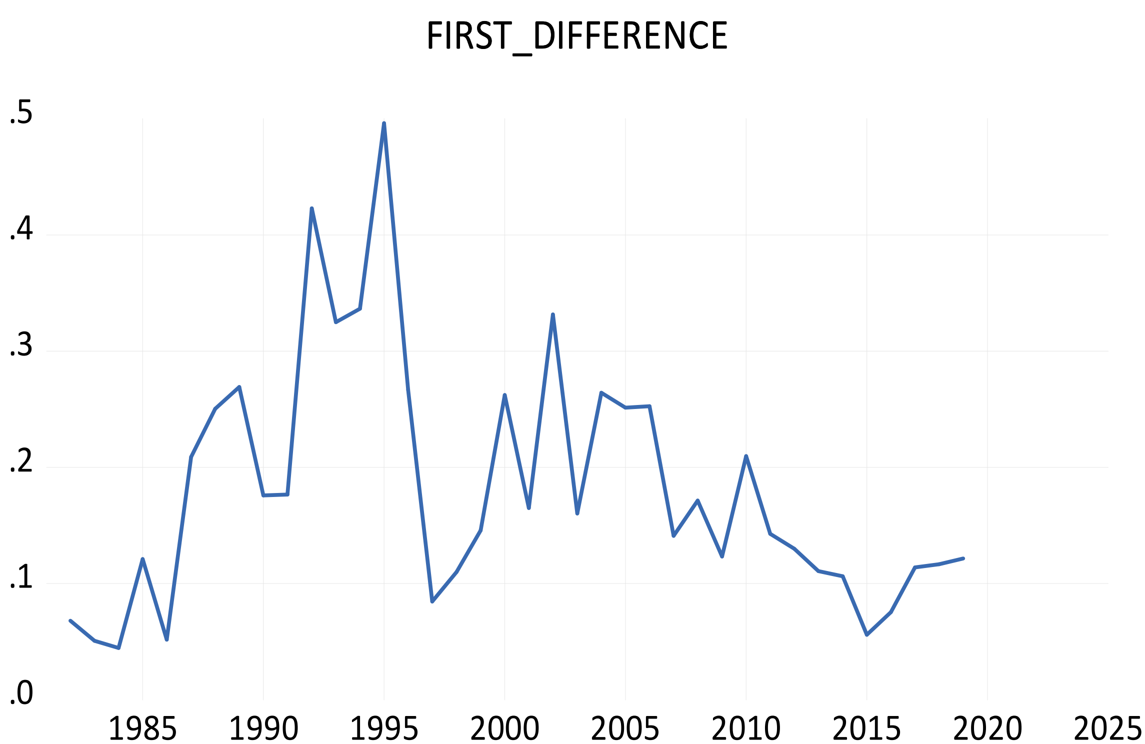


Fig. 2. The plot of the first differenced real GDP

Table 2. Augmented dickey- fuller test of the first differenced data

Null Hypothesis: D(LGDP) has a unit root				
Trend Specification: Intercept only				
Break Specification: Intercept only				
Break Type: Innovational outlier				
Break Date: 1995				
Break Selection: Minimize Dickey-Fuller t-statistic				
Lag Length: 0 (Automatic - based on Schwarz information criterion, maxlag=9)				
			t-Statistic	**Prob.***
Augmented Dickey-Fuller test statistic			-4.221697	0.0925
Test critical values:	1% level		-4.949133	
	5% level		-4.443649	
	10% level		-4.193627	
*Vogelsang (1993) asymptotic one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LGDP)				
Method: Least Squares				
Date: 09/19/20 Time: 18:31				
Sample (adjusted): 1983 2019				
Included observations: 37 after adjustments				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
D(LGDP(-1))	0.464004	0.126962	3.654655	0.0009
C	0.118994	0.032295	3.684567	0.0008
INCPTBREAK	-0.038958	0.027915	-1.395588	0.1722
BREAKDUM	0.259955	0.083043	3.130358	0.0036
R-squared	0.487370	Mean dependent var		0.184750
Adjusted R-squared	0.440768	S.D. dependent var		0.105579
S.E. of regression	0.078954	Akaike info criterion		-2.138107
Sum squared resid	0.205711	Schwarz criterion		-1.963954
Log likelihood	43.55498	Hannan-Quinn criter.		-2.076710
F-statistic	10.45799	Durbin-Watson stat		2.088542
Prob(F-statistic)	0.000055			

From Table 2, the statistic shows that the data is non-stationary and hence, the data is differenced further.

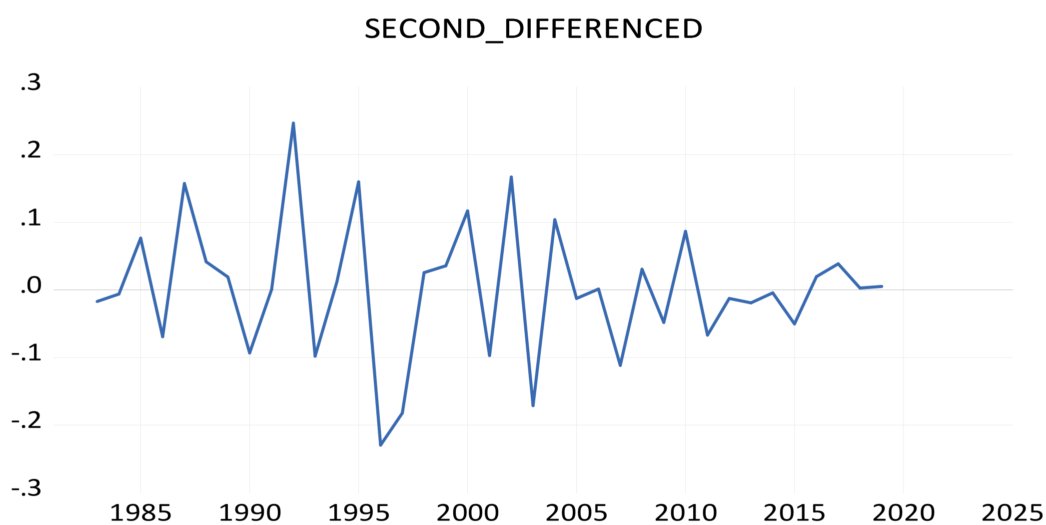


Fig. 3. The plot of the second differenced real GDP

Table 3. Augmented dickey-fuller test of the second differenced data

Null Hypothesis: D(LGDP,2) has a unit root				
Trend Specification: Intercept only				
Break Specification: Intercept only				
Break Type: Innovational outlier				
Break Date: 1997				
Break Selection: Minimize Dickey-Fuller t-statistic				
Lag Length: 0 (Automatic - based on Schwarz information criterion, maxlag=9)				
			t-Statistic	**Prob.***
Augmented Dickey-Fuller test statistic			-10.19480	< 0.01
Test critical values:	1% level		-4.949133	
	5% level		-4.443649	
	10% level		-4.193627	
*Vogelsang (1993) asymptotic one-sided p-values.				
Augmented Dickey-Fuller Test Equation				
Dependent Variable: D(LGDP,2)				
Method: Least Squares				
Date: 09/19/20 Time: 18:36				
Sample (adjusted): 1984 2019				
Included observations: 36 after adjustments				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
D(LGDP(-1),2)	-0.581246	0.155103	-3.747479	0.0007
C	0.035769	0.023576	1.517207	0.1390
INCPTBREAK	-0.038055	0.029677	-1.282292	0.2090
BREAKDUM	-0.313341	0.091618	-3.420072	0.0017
R-squared	0.377067	Mean dependent var		0.001968
Adjusted R-squared	0.318667	S.D. dependent var		0.100531
S.E. of regression	0.082981	Akaike info criterion		-2.035971
Sum squared resid	0.220347	Schwarz criterion		-1.860025
Log likelihood	40.64748	Hannan-Quinn criter.		-1.974561
F-statistic	6.456618	Durbin-Watson stat		2.412365
Prob(F-statistic)	0.001526			

As show in Table 3, the result shows that the ADF test at the second difference is 0.01 and seen to be stationary at the second differencing and hence, we can proceed to determine a suitable model for the series.

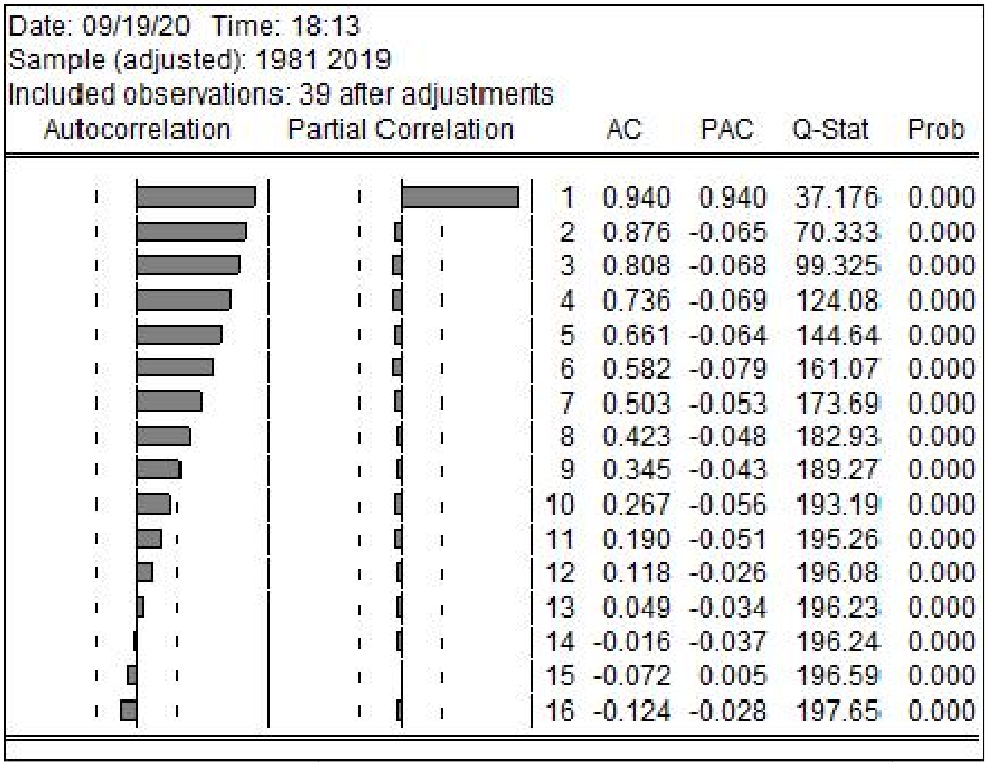
The actual correlogram of the real GDP is presented in Fig. 4, showing significant autocorrelation that are outside the error bound, showing its non-stationarity properties. This lead to taking the first and second difference, where it was found to be stationary.

As seen above, it is reasonable to say that the data plot above is stationary with only the first ACF and PACF statistically significant.

6.1 Estimated Model

The model, ARIMA(1, 2, 1) was selected using the Akaike criterion as seen in Table 4.

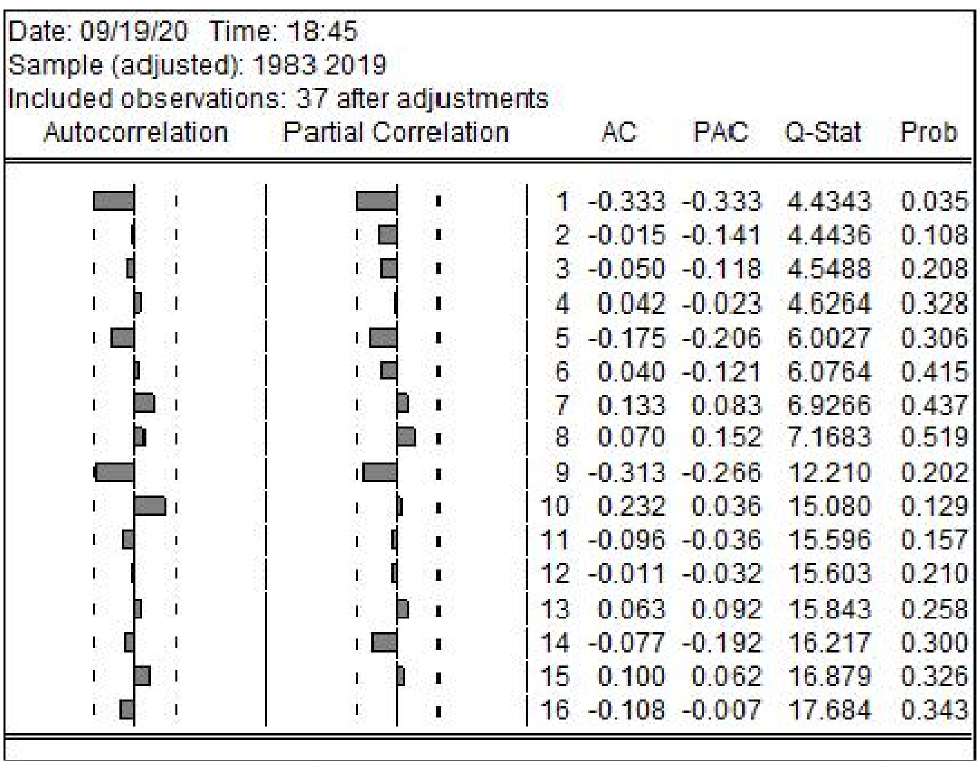
From the results in the Table 4, the best model is ARIMA (1, 2, 1), having the minimum values of AIC and BIC and the model coefficient is given below which are all significant at 5% level of significance.



Date: 09/19/20 Time: 18:13
Sample (adjusted): 1981 2019
Included observations: 39 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.940	0.940	37.176	0.000
		2	0.876	-0.065	70.333	0.000
		3	0.808	-0.068	99.325	0.000
		4	0.736	-0.069	124.08	0.000
		5	0.661	-0.064	144.64	0.000
		6	0.582	-0.079	161.07	0.000
		7	0.503	-0.053	173.69	0.000
		8	0.423	-0.048	182.93	0.000
		9	0.345	-0.043	189.27	0.000
		10	0.267	-0.056	193.19	0.000
		11	0.190	-0.051	195.26	0.000
		12	0.118	-0.026	196.08	0.000
		13	0.049	-0.034	196.23	0.000
		14	-0.016	-0.037	196.24	0.000
		15	-0.072	0.005	196.59	0.000
		16	-0.124	-0.028	197.65	0.000

Fig. 4. Correlogram of the real GDP rate (Level)



Date: 09/19/20 Time: 18:45
Sample (adjusted): 1983 2019
Included observations: 37 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.333	-0.333	4.4343	0.035
		2	-0.015	-0.141	4.4436	0.108
		3	-0.050	-0.118	4.5488	0.208
		4	0.042	-0.023	4.6264	0.328
		5	-0.175	-0.206	6.0027	0.306
		6	0.040	-0.121	6.0764	0.415
		7	0.133	0.083	6.9266	0.437
		8	0.070	0.152	7.1683	0.519
		9	-0.313	-0.266	12.210	0.202
		10	0.232	0.036	15.080	0.129
		11	-0.096	-0.036	15.596	0.157
		12	-0.011	-0.032	15.603	0.210
		13	0.063	0.092	15.843	0.258
		14	-0.077	-0.192	16.217	0.300
		15	0.100	0.062	16.879	0.326
		16	-0.108	-0.007	17.684	0.343

Fig. 5. Correlogram of the real GDP rate (Second Difference)

Table 4. ARIMA model comparison

Model Selection Criteria Table
Dependent Variable: DLOG(LGDP)
Date: 09/19/20 Time: 19:24
Sample: 1981 2025
Included observations: 38

Model	LogL	AIC*	BIC	HQ
(1,1)(0,0)	114.874278	-5.835488	-5.663111	-5.774158
(1,0)(0,0)	114.368893	-5.834521	-5.662238	-5.771523
(2,0)(0,0)	114.807086	-5.831952	-5.659574	-5.770621
(3,0)(0,0)	114.917887	-5.785152	-5.569680	-5.708489
(1,2)(0,0)	114.875009	-5.782895	-5.567423	-5.706232
(2,1)(0,0)	114.874729	-5.782880	-5.567409	-5.706217
(2,3)(0,0)	116.752084	-5.776425	-5.474765	-5.669097
(0,4)(0,0)	115.049984	-5.739473	-5.480907	-5.647477
(0,3)(0,0)	114.029061	-5.738372	-5.522900	-5.661708
(1,3)(0,0)	114.938029	-5.733580	-5.475014	-5.641585
(4,0)(0,0)	114.934153	-5.733376	-5.474810	-5.641381
(3,1)(0,0)	114.923876	-5.732836	-5.474269	-5.640840
(3,3)(0,0)	116.914271	-5.732330	-5.387575	-5.609669
(3,2)(0,0)	115.907242	-5.731960	-5.430300	-5.624632
(2,2)(0,0)	114.886693	-5.730879	-5.472312	-5.638883
(2,4)(0,0)	116.863331	-5.729649	-5.384894	-5.606988
(4,1)(0,0)	115.761056	-5.724266	-5.422606	-5.616938
(0,2)(0,0)	112.723528	-5.722291	-5.549913	-5.660960
(4,2)(0,0)	116.481108	-5.709532	-5.364777	-5.586871
(1,4)(0,0)	115.081267	-5.688488	-5.386827	-5.581159
(4,3)(0,0)	116.912400	-5.679600	-5.291751	-5.541606
(0,1)(0,0)	110.828082	-5.675162	-5.545879	-5.629164
(3,4)(0,0)	116.045578	-5.633978	-5.246128	-5.495984
(4,4)(0,0)	116.534204	-5.607063	-5.176120	-5.453737
(0,0)(0,0)	104.574340	-5.398649	-5.312461	-5.367984

Table 5. The ARIMA model

Dependent Variable: LGDP
Method: ARMA Conditional Least Squares (Marquardt - EViews legacy)
Date: 09/19/20 Time: 19:35
Sample (adjusted): 1982 2019
Included observations: 38 after adjustments
Convergence achieved after 14 iterations
MA Backcast: 1981

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	30.73783	25.51606	1.204647	0.2364
AR(1)	0.991849	0.009323	106.3926	0.0000
MA(1)	0.410431	0.152895	2.684404	0.0110
R-squared	0.998417	Mean dependent var		8.716068
Adjusted R-squared	0.998326	S.D. dependent var		2.320507
S.E. of regression	0.094931	Akaike info criterion		-1.795686
Sum squared resid	0.315413	Schwarz criterion		-1.666403
Log likelihood	37.11804	Hannan-Quinn criter.		-1.749688
F-statistic	11036.66	Durbin-Watson stat		1.722827
Prob(F-statistic)	0.000000			
Inverted AR Roots	.99			
Inverted MA Roots	-.41			

The model coefficient from Table 2, shows that AR(1) and MA(1) coefficient are all statistically significant at 0.05.

Hence, the ARIMA equation is given as

$$\hat{X}_t = 30.73783 + 0.991849X_{t-1} + 0.410431w_{t-1} + w_t.$$

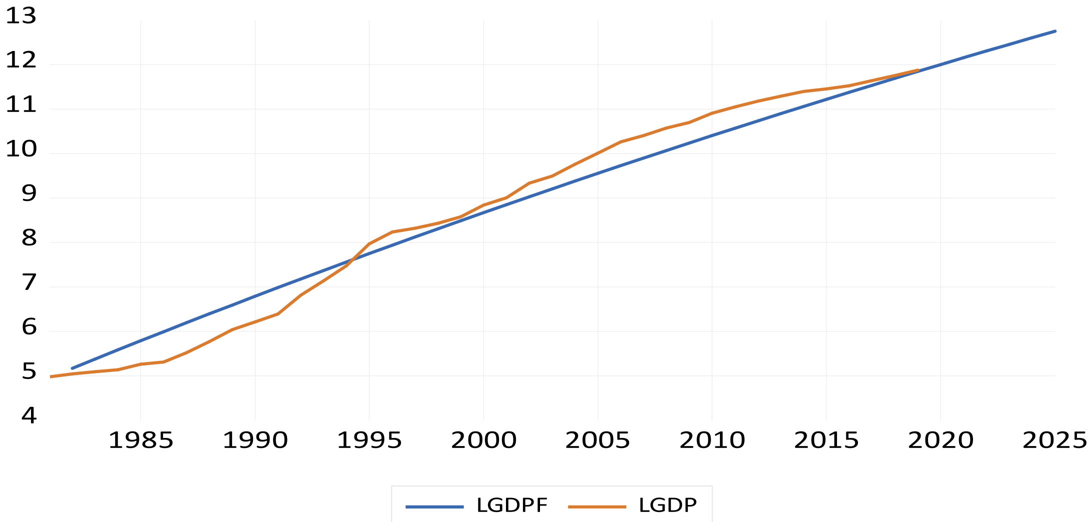
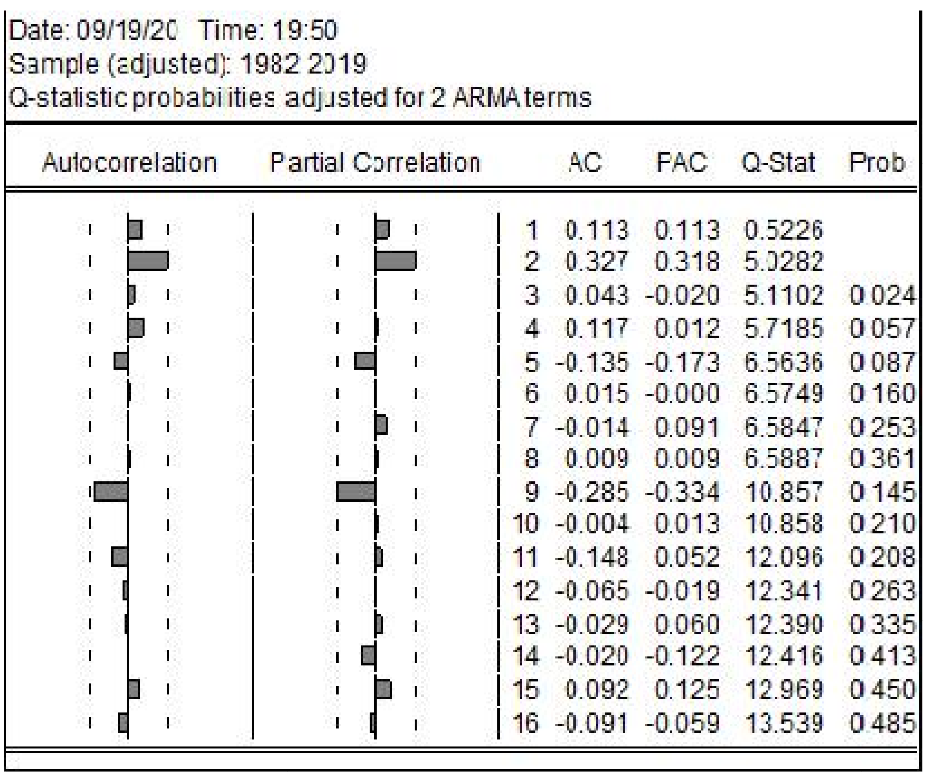


Fig. 6. Fitting the model



Date: 09/19/20 Time: 19:50
Sample (adjusted): 1982 2019
Q-statistic probabilities adjusted for 2 ARMA terms

	AC	PAC	Q-Stat	Prob
1	0.113	0.113	0.5226	
2	0.327	0.318	5.0282	
3	0.043	-0.020	5.1102	0.024
4	0.117	0.012	5.7185	0.057
5	-0.135	-0.173	6.5636	0.087
6	0.015	-0.000	6.5749	0.160
7	-0.014	0.091	6.5847	0.253
8	0.009	0.009	6.5887	0.361
9	-0.285	-0.334	10.857	0.145
10	-0.004	0.013	10.858	0.210
11	-0.148	0.052	12.096	0.208
12	-0.065	-0.019	12.341	0.263
13	-0.029	0.060	12.390	0.335
14	-0.020	-0.122	12.416	0.413
15	0.092	0.125	12.969	0.450
16	-0.091	-0.059	13.539	0.485

Fig. 7. Ljung-Box Q test of the residual

Diagnostic checking of the model will be performed to help us check the acceptability and statistical significance of the estimated model i.e. if the model residuals are not autocorrelated. Q statistic will be used to test for autocorrelation for the model ARIMA(1, 2, 1).

From the Fig. 7 indicates a 24 lag Q statistic of Ljung-Box hav values greater than 0.05 helping us to state that the null hypothesis cannot be rejected. Hence, there is no autocorrelation for the examined residuals of the series.

6.2 Forecast

We use the ARIMA(1, 2, 1) model to forecast the GDP from 2020 to 2025 comparing it with the actual data. As seen below, the Theil inequality of 0.022008, making the model a reasonable model for forecasting future data, meaning that our model may have a very good forecasting ability.

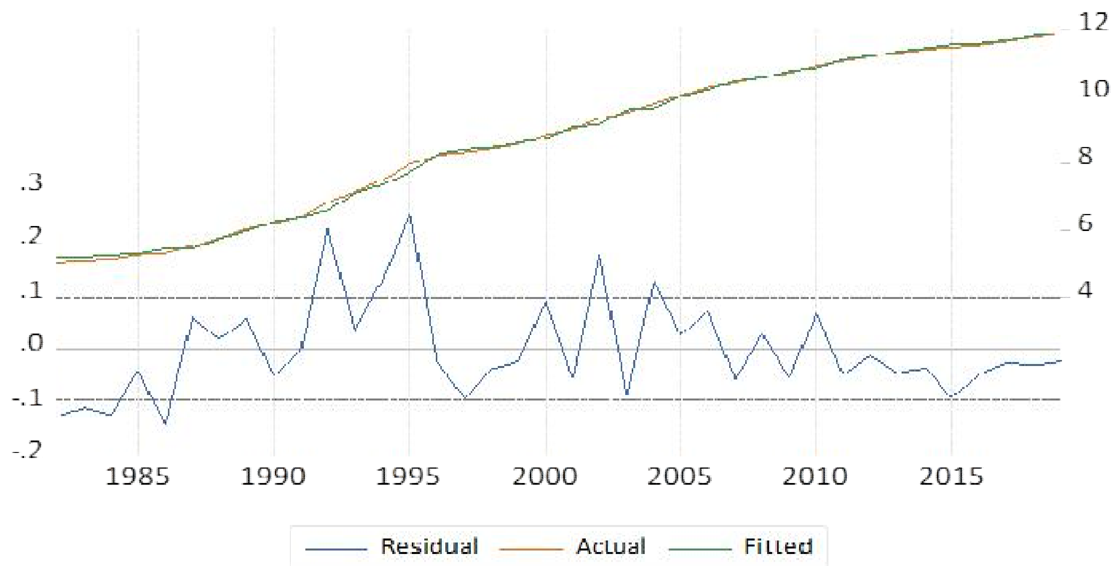


Fig. 8. The plot of actual, fitted and residual of the real GDP

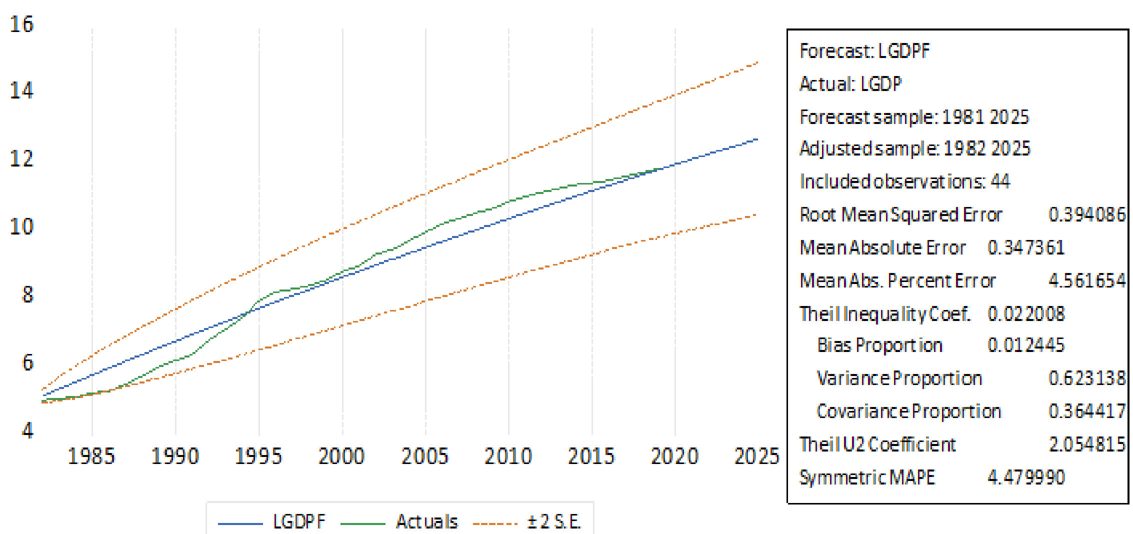


Fig. 9. Model forecast

Table 6. Forecast

Year	Total (GDPF)	Year	Total (GDPF)	Year	Total (GDPF)
1981	144.83	1998	4,588.99	2015	94,144.96
1982	154.98	1999	5,307.36	2016	101,489.49
1983	163	2000	6,897.48	2017	113,711.63
1984	170.38	2001	8,134.14	2018	127,736.83
1985	192.27	2002	11,332.25	2019	144,210.49
1986	202.44	2003	13,301.56	2020	161050.8903
1987	249.44	2004	17,321.30	2021	178404.9527
1988	320.33	2005	22,269.98	2022	196141.7039
1989	419.2	2006	28,662.47	2023	214377.9
1990	499.68	2007	32,995.38	2024	233009.4588
1991	596.04	2008	39,157.88	2025	252129.1643
1992	909.8	2009	44,285.56		
1993	1,259.07	2010	54,612.26		
1994	1,762.81	2011	62,980.40		
1995	2,895.20	2012	71,713.94		
1996	3,779.13	2013	80,092.56		
1997	4,111.64	2014	89,043.62		

7. CONCLUSION

In this study, we use ARIMA model in trying to model the real GDP rate of Nigeria. After stationarity was checked using Augmented Dick Fuller test, correlogram was used in identifying the most suitable model with minimum value of Akaike Informtaion Criterion and this result was used in forecasting with Theil inequality of 0.022008, making the model a reasonable model for forecasting future data, after the residual of the model was checked using Ljung-Box that shows no sign of autocorrelation in the residual. The ARIMA (1, 2, 1) was considered the most appropriate model for the data since the model diagnostic tests showed significant parameter estimates and randomness in the plot of the residuals. Out of sample forecast was generated for 2020 through 2025 using Eviews version 11.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Box George EP, Gwilym M. Jenkins. Time series analysis: Forecasting and control, revised edition, Oakland, Ca: Holden-Day; 1976.
2. Anderson OD. Analyzing time series: Amsterdam. Oxford University Press, Holland; 1980.
3. Pankratz A. Forecasting with Univariate Box–Jenkins Models: Concepts and Cases. New York: John Wiley & Sons; 1983.
4. Wabomba MS, Mutwiri MP, Fredrick M. Modeling and forecasting Kenyan GDP using autoregressive integrated moving average (ARIMA) models. Science Journal of Applied Mathematics and Statistics. 2016;4:64-73.
   Available:https://doi.org/10.11648/j.sjams.20160402.18.
5. Uwimana A, Xiuchun B, Shuguang Z. Modeling and forecasting Africa's GDP with time series models. International Journal of Scientific and Research Publications. 2018;8:41-46.
   Available:https://doi.org/10.29322/ijsrp.8.4.2018.p7608
6. Zakai M. A time series modelling on GDP of Pakistan. Journal of Conporary Issues in Business Research. 2014;3(4): 200-210.
7. Bhuiyan MNA, Kazi SA, Roushan J. Modelling and forecasting of the GDP of

manufacturing industries in Bangladesh. Jahangirnagar University Press, Dhaka, Bangladesh; 2007.

8. Abiola AG, Okafor HO. Searching for appropriate crude oil price benchmarking method in the Nigerian Budgeting Process. Developing Country Studies. 2013;3(12).

9. Granger C, Newbold P. Forecasting in business and economics. Academic Press, New York; 1989.

10. Lütkepohl H, Xu F. The role of the log transformation in forecasting economic variables. Empirical Economics. 2012;42(3):619–638.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/62569