Journal of Statistical and Econometric Methods, vol.3, no.4, 2014, 33-46
ISSN: 2241-0384 (print), 2241-0376 (online)
Scienpress Ltd, 2014

Forecasting Gross Domestic Product In Nigeria Using Box-Jenkins Methodology

O.E. Okereke[1] and C.B. Bernard[2]

Abstract

Gross domestic product(GDP) is an important tool for measuring the quality of the overall economic activity in a country within a specified period of time. This study aimed at providing a model that can be used to forecast gross domestic product in Nigeria using the Box-Jenkins approach. Quarterly data on Nigerian GDP from the first quarter of 1990 to the second quarter of 2013 were used for this purpose. The time plot and preliminary analyses of the data called for log transformation and first order differencing of the data to achieve stationarity. Plots of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the transformed and differenced series suggested that SARIMA $(2, 1, 2)x(1, 0, 1)_4$ be fitted to the data. The ACF and PACF of the residuals from the fitted model behaved like those of a white noise process. These confirmed the adequacy of the

[1] Department of Statistics, Michael Okpara University of Agriculture, Umudike.
E-mail: emmastat5000@yahoo.co.uk
[2] Department of Statistics, Michael Okpara University of Agriculture, Umudike.
E-mail:chaleben@yahoo.com

Article Info: *Received* : August 12, 2014. *Revised* : September 30, 2014.
Published online : December 27, 2014.

fitted model. The model was then used to forecast GPD in Nigeria for a period of one year.

Mathematics Subject Classification: 62M10

Keywords: Gross domestic product; log transformation; differencing; stationarity; white noise process

1 Introduction

The gross domestic product (GDP) is an important economic tool for determining how good the economy of a country is. It is generally viewed as the monetary value of all finished goods and services produced within a country at a particular period of time (Abdulrahem, 2011). A country is said to have good economy if its GDP is relatively high.

Several studies involving the gross domestic product have been carried out. A good number of such works dealt the relationships between GDP and other economic variables. For instance, Abdulrahem (2011) examined the sectorial contribution of GDP in Nigeria using a set of time series data and multiple linear regression analysis. An empirical analysis of the contribution of agriculture and petroleum sector to the growth and development of the Nigerian economy for the period 1960-2010, was carried out by Umaru and Zubaru (2012). These authors pointed out that all the variables in their proposed model were stationary and there was no structural break within the period under review. Efiok et al. (2012) determined the extent to which human capital cost influences gross domestic product in Nigeria with the help of ordinary least squares (OLS) regression. His findings remains that human capital cost significantly affects GDP in Nigeria. Nwabueze (2009) investigated the causal relationship between gross domestic product and personal consumption expenditure of Nigeria using regression analysis.

Gross domestic product data comprise of observations made at equally spaced intervals of time. For example, the gross domestic product in Nigeria is usually computed and recorded on quarterly basis. It can therefore be deduced that such data are time series data. Thus, the analysis of GDP as a single variable requires an appropriate method of time series analysis. A well known method of analyzing time series is discussed in Box et al. (1994). The properties of GDP series of many countries have been identified from time series analysis point of view. On this note, Reininger and Fingerlos (2007) established that gross domestic product in Belgium is non stationary in mean and variance. Similar findings were made by Etuk (2012) who emphasized that Nigeria GDP series is seasonal and non stationary in mean level.

It is a common practice to obtain the log of GDP series before it is analyzed by Box and Jenkins method (Gujarati and Porter, 2004). The consequences of ignoring transformation when it should actually be applied cannot be overemphasized. A variance that changes over time affects the validity and efficiency of statistical inference about the parameters that describe the dynamics of the level of a time series (Hamilton, 1994). Moreover, proper transformation of a time series is one way of avoiding model misspecification (Delurgio, 1998). The objective of this study is to build a suitable model for forecasting GDP in Nigeria. For this purpose, the GDP series will be examined for both level nonstationarity and variance nonstationarity prior to its analysis.

2 Transformation of Time Series Data

In time series analysis, a time series that is not originally stationary can be made stationary in variance by dint of an appropriate transformation. Many forms of transformation have been proposed in the literature. These include log transformation, square transformation, square root transformation, inverse transformation among others. Akpanta and Iwueze (2009) proposed a method of determining the type of transformation to apply to time series data. Their method is based on the equation:

$$\log_e \hat{\sigma}_i = \hat{\alpha} + \hat{\beta} \log_e \overline{X}_i \tag{2.1}$$

The model (2.1) simply describes the relationship between annual standard deviations $\left(\hat{\sigma}_i, i = 1, 2, \ldots, k\right)$ and the annual means $\left(\overline{X}_i, i = 1, 2, \ldots, k\right)$, where k is the number of years. Consequently, the following transformation is made based on the value of $\hat{\beta}$ (Akpanta and Iwueze, 2009)

$$Z_t = \begin{cases} \log_e \overline{X}_t, if\ \hat{\beta} = 1 \\ X_t^{1-\hat{\beta}}, if\ \hat{\beta} \neq 1 \end{cases} \tag{2.2}$$

3 Review of Box-Jenkins Methodology

Model building using Box-Jenkins methodology involves three main stages namely identification, estimation and diagnostic checking (Box and Jekins,1994).

3.1 Identification

At this stage of time series modeling the analysts intends to suggest a tentative model to a time series by examining the time plot and the graphical representation of each of the autocorrelation function and partial autocorrelation function. Such plots could reveal certain properties of a time series like nonstationarity and outlier. The sample correlogram and partial correlogram help us to determine the orders of p, d, q in the autoregressive integrated moving average (ARIMA (p, d, q)) model:

$$\phi(B)(1-B)^d X_t = \theta(B)e_t \tag{3.1}$$

where $\phi(B)$ and $\theta(B)$ are polynomials of orders p and q respectively, d is the order of non seasonal differencing and e_t is a white noise process.

Sometimes, the sample autocorrelation function (ACF) and partial autocorrelation function (PACF) of a time series are characterized by spikes at multiples of the seasonal lag 4 for quarterly time series data. To account for seasonal variations in a time series of this kind, there is need to generalize the ARIMA (p, d, q) to the seasonal autoregressive integrated moving average (SARIMA) model. The multiplicative seasonal autoregressive integrated moving average (SARIMA (p, d, q) x (P, Q, D)) $_S$ model is given by Box et al 1994 as

$$\Phi(B^S)\phi(B)(1-B)^d X_t = \Theta(B^S)\theta(B)e_t \tag{3.2}$$

where B is the backshift operator, $\Phi(B^S)$ and $\Theta(B^S)$ are polynomials in B^S of degrees P and Q respectively, $\phi(B)$ and $\theta(B)$ are polynomials in B of degrees p and q respectively and S is the periodicity of the time series. Again, d and D refer to the orders of nonseasonal differencing and seasonal differencing in that order. For the series to be stationary, the zeros of $\Phi(B^S)$ and $\phi(B)$ must lie outside the unit circle while it is invertible whenever the absolute values of zeros of $\Theta(B^S)$ and $\theta(B)$ exceed unity.

3.2 Estimation

Parameters of the tentatively proposed models are estimated using one of the method of moments, maximum likelihood method and non-linear least squares approach (Ion and Adriana, 2008; Montgomery, 2008). These methods of estimation can now be employed with the help of statistical software. In this study, MINITAB is used for estimation of the parameters of the proposed model.

3.3 Diagnostic Checking

Once a tentative model has been fitted to a time series, the adequacy of such a model has to be ascertained. If the model is suitable, its associated residuals must possess characteristics of a white noise process. We therefore expect the residuals to come from a fixed distribution with a constant mean (usually zero) and a constant variance when the fitted model is appropriate (Ion and Adriana, 2008; Wei, 2006). Failure of the residuals to satisfy these assumptions simply suggests that a more appropriate is required.

4 Results and discussion

In this section, the steps discussed in section 3 are used to analyze the Nigeria GDP series from the first quarter of 1990 to second quarter of 2013. This data can be retrieved from the website www.nigeriastat.gov.ng. The time plot of the series is shown in Figure 1.

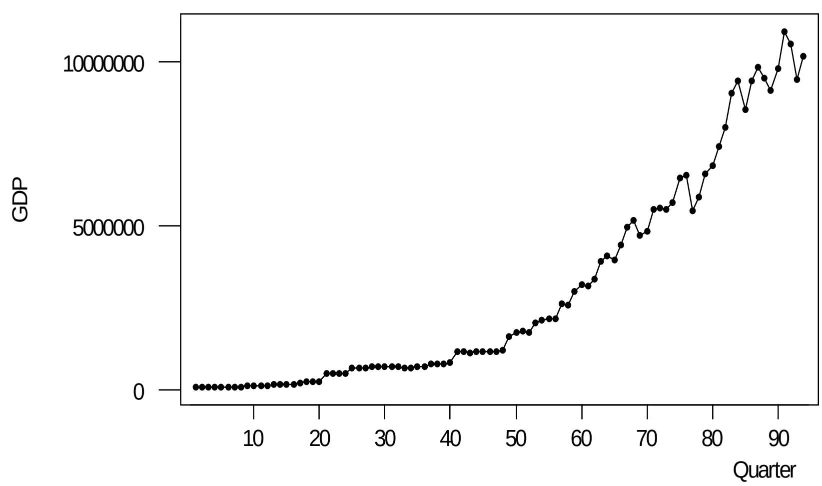


Figure 1: Time plot of Nigerian GDP series

Careful examination of Figure 1 reveals that the GDP series is not stationary in mean and variance. The regression model based on the natural logs annual means and annual standard deviations is

$$\log_e \hat{\sigma}_i = -9.15 + 1.42 \log_e \overline{X}_i \quad (4.1)$$

Since $\hat{\beta}$ =1.42 is approximately equal to 1, we apply the log transformation to the data. Figure 2 contains the time plot of log of the GDP series.

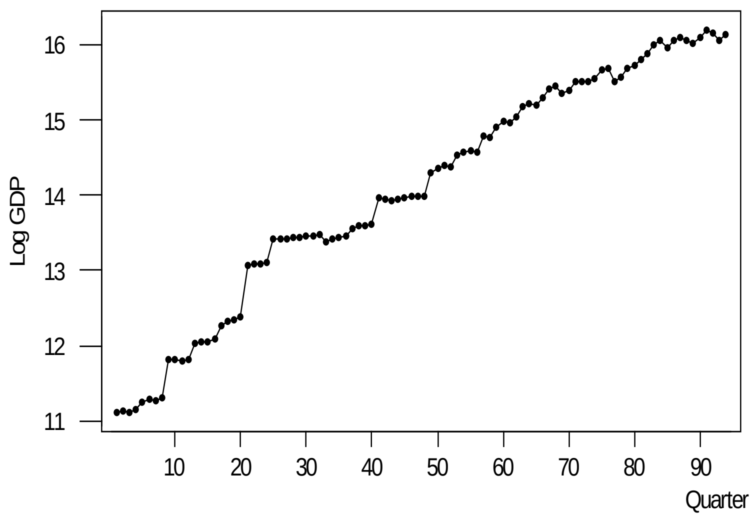
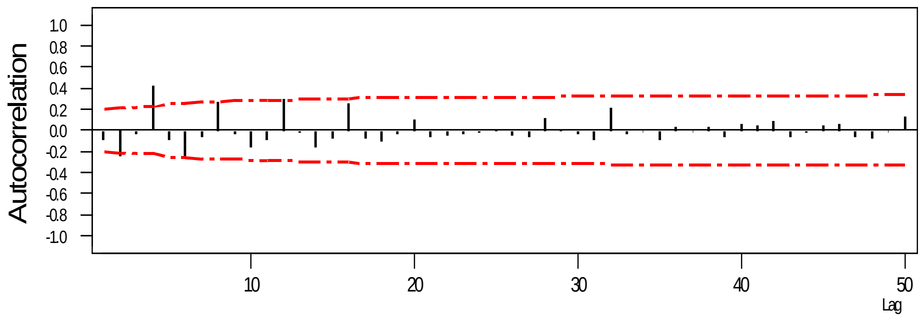


Figure 2: Time plot of log of Nigerian GDP series

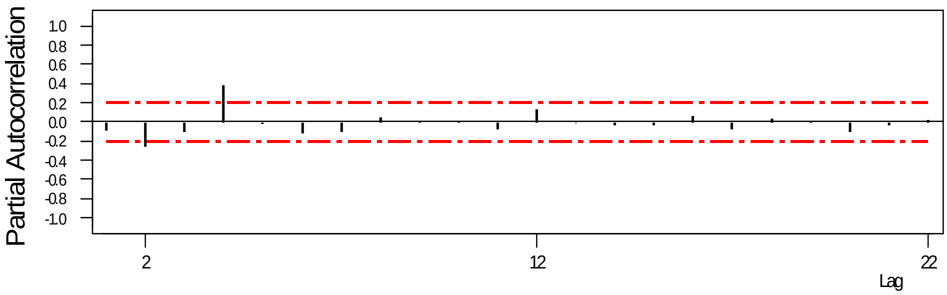
As shown in Figure 2, log transformation appears to make the variance of the series more stable. However, there is need for differencing to attain stationarity. In Figure 3, we give the graphical representation of the autocorrelation function of the first differences of the log GDP series.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.10	-0.92	0.87	13	-0.03	-0.18	52.41	25	0.00	0.01	68.13	37	-0.01	-0.08	80.35	49	-0.01	-0.08	88.23
2	-0.25	-2.34	6.69	14	-0.16	-1.08	55.30	26	-0.05	-0.31	68.45	38	0.04	0.24	80.60	50	0.14	0.82	92.14
3	-0.04	-0.33	6.81	15	-0.07	-0.49	55.92	27	-0.06	-0.38	68.95	39	-0.06	-0.39	81.29				
4	0.43	3.82	24.68	16	0.26	1.72	63.68	28	0.12	0.74	70.86	40	0.06	0.39	81.98				
5	-0.08	-0.66	25.39	17	-0.08	-0.50	64.39	29	0.00	0.02	70.86	41	0.06	0.33	82.50				
6	-0.25	-1.90	31.39	18	-0.11	-0.69	65.76	30	-0.03	-0.18	70.98	42	0.10	0.58	84.11				
7	-0.06	-0.47	31.79	19	-0.04	-0.26	65.96	31	-0.09	-0.56	72.11	43	-0.06	-0.38	84.80				
8	0.27	2.01	39.25	20	0.10	0.64	67.19	32	0.22	1.33	78.77	44	-0.02	-0.10	84.85				
9	-0.04	-0.29	39.42	21	-0.06	-0.39	67.65	33	-0.03	-0.21	78.94	45	0.05	0.28	85.27				
10	-0.17	-1.18	42.28	22	-0.05	-0.30	67.94	34	-0.01	-0.04	78.95	46	0.07	0.41	86.16				
11	-0.08	-0.59	43.05	23	-0.04	-0.23	68.10	35	-0.09	-0.53	80.12	47	-0.07	-0.41	87.08				
12	0.29	2.05	52.34	24	-0.02	-0.09	68.13	36	0.04	0.22	80.32	48	-0.08	-0.45	88.19				

Figure 3: Correlogram of first differences of logs of Nigerian GDP

There seem to be spikes at lag 2 and lag 4 of first differences of the series as shown in Figure 3. This simply indicates that the GDP series is seasonal. For proper model identification, we examine the partial autocorrelogram in Figure 4.



Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	-0.10	-0.92	8	0.05	0.45	15	-0.04	-0.39	22	0.02	0.17
2	-0.26	-2.47	9	0.00	0.02	16	0.06	0.60			
3	-0.10	-0.94	10	0.00	0.01	17	-0.08	-0.76			
4	0.38	3.61	11	-0.08	-0.75	18	0.04	0.35			
5	-0.02	-0.22	12	0.14	1.33	19	0.00	0.01			
6	-0.12	-1.12	13	-0.01	-0.07	20	-0.11	-1.06			
7	-0.11	-1.03	14	-0.04	-0.33	21	-0.04	-0.39			

Figure 4: Partial correlogram of first differences of logs of Nigerian GDP

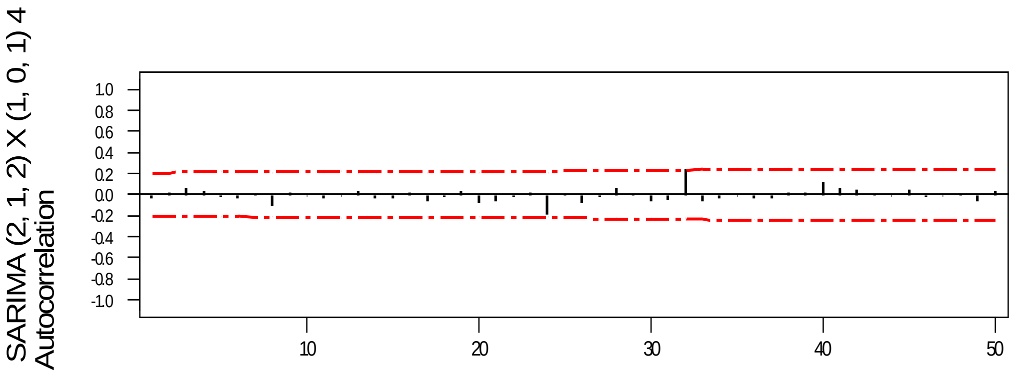
It be can observed that the partial correlogram is characterized by spikes at lag 2 and lag 4. From the foregoing, SARIMA (2, 1, 2) x $(1, 0, 1)_4$ is tentatively fitted to the data.

The estimates of the parameters of the models are shown in Table 1.

Table 1: Estimates of Parameters of SARIMA (2, 1, 2) x $(1, 0, 1)_4$ Model

Type	Coefficients
AR 1	-0.0042
AR 2	0.7590
SAR 4	0.9866
MA1	0.0183
MA2	0.8938
SMA 4	0.6713
Constant	0.0001560

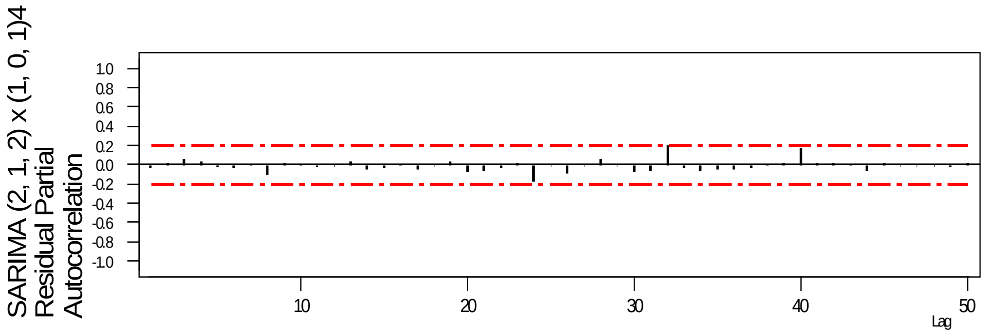
Figure 5 shows the correlogram of residuals from the fitted model.



Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ	Lag	Corr	T	LBQ
1	-0.04	-0.35	0.13	13	0.03	0.32	2.11	25	0.01	0.06	8.89	37	-0.04	-0.31	20.60	49	-0.06	-0.50	25.71
2	0.02	0.21	0.18	14	-0.04	-0.39	2.31	26	-0.07	-0.65	9.59	38	0.02	0.17	20.67	50	0.04	0.33	26.05
3	0.06	0.54	0.49	15	-0.03	-0.28	2.41	27	-0.02	-0.14	9.62	39	0.02	0.20	20.76				
4	0.04	0.35	0.62	16	0.02	0.17	2.44	28	0.07	0.61	10.27	40	0.12	1.02	23.27				
5	-0.02	-0.19	0.66	17	-0.06	-0.55	2.85	29	0.01	0.05	10.27	41	0.06	0.50	23.91				
6	-0.03	-0.30	0.76	18	-0.01	-0.14	2.87	30	-0.06	-0.50	10.71	42	0.05	0.41	24.34				
7	0.01	0.09	0.77	19	0.04	0.35	3.04	31	-0.04	-0.38	10.98	43	0.01	0.11	24.37				
8	-0.10	-0.95	1.80	20	-0.08	-0.76	3.85	32	0.24	2.13	19.41	44	-0.00	-0.01	24.37				
9	0.02	0.19	1.85	21	-0.07	-0.63	4.42	33	-0.06	-0.53	19.99	45	0.05	0.42	24.86				
10	-0.00	-0.00	1.85	22	-0.02	-0.17	4.46	34	-0.04	-0.31	20.20	46	-0.02	-0.14	24.91				
11	-0.04	-0.33	1.98	23	0.02	0.20	4.52	35	-0.01	-0.09	20.22	47	-0.01	-0.09	24.94				
12	-0.00	-0.03	1.98	24	-0.19	-1.70	8.88	36	-0.03	-0.27	20.38	48	0.00	0.03	24.94				

Figure 5: SARIMA (2, 1, 2) X (1, 0, 1) $_4$ residual ACF plot

Obviously, Figure 5 represents the autocorrelation function of a white noise process. Next, we critically examine the residual partial correlogram in Figure 6.



Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T	Lag	PAC	T
1	-0.04	-0.35	13	0.03	0.30	25	-0.01	-0.13	37	-0.03	-0.32	49	-0.03	-0.26
2	0.02	0.20	14	-0.04	-0.41	26	-0.08	-0.80	38	0.01	0.10	50	0.02	0.21
3	0.06	0.56	15	-0.03	-0.29	27	-0.00	-0.03	39	0.02	0.17			
4	0.04	0.39	16	0.00	0.04	28	0.07	0.63	40	0.17	1.61			
5	-0.02	-0.19	17	-0.05	-0.52	29	-0.00	-0.00	41	0.03	0.24			
6	-0.04	-0.37	18	-0.01	-0.11	30	-0.07	-0.69	42	0.03	0.24			
7	0.00	0.03	19	0.04	0.35	31	-0.07	-0.63	43	0.01	0.09			
8	-0.10	-0.94	20	-0.08	-0.76	32	0.20	1.91	44	-0.06	-0.56			
9	0.02	0.18	21	-0.07	-0.64	33	-0.04	-0.34	45	0.01	0.13			
10	0.01	0.07	22	-0.03	-0.33	34	-0.06	-0.59	46	-0.01	-0.06			
11	-0.03	-0.25	23	0.02	0.19	35	-0.06	-0.53	47	-0.01	-0.10			
12	-0.00	-0.01	24	-0.17	-1.63	36	-0.05	-0.43	48	-0.01	-0.08			

Figure 6: SARIMA (2, 1, 2) X (1, 0, 1) $_4$ residual PACF plot

Figure 6 indicates that there is certainly no pattern left in residual partial autocorrelation function. We then conclude that the residuals form a white noise process. This confirms the adequacy of the fitted model. Plots of the actual data and the predicted values of GDP are contained in Figure 7. The original series is graphed using the circles while each plus sign in the plot represent a predicted value of the series.

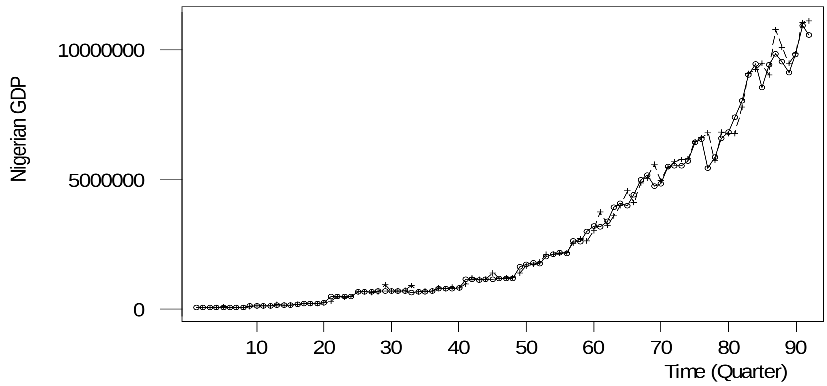


Figure 7: Plots of quarterly Nigerian GDP series and its predicted values

Visual inspection of Figure 7 simply suggest that the SARIMA (2, 1, 2) X (1, 0, 1) $_4$ model fit the data well.

One of the primary objectives of time series analysis is to forecast future values of a time series. Using the fitted model, the following forecast values of the GDP series in Table 2 are obtained.

Table 2: GDP forecast for the last two quarters of 2013 and first two quarters of 2014

Period	Forecast
93	10457186.46
94	11342871.01
95	12699840.6
96	12758394.44

Furthermore, there exist some similarities between the results obtained in this study and the one given by Etuk (2012). Our reason is that both studies stipulates that the GDP series is seasonal. However, the fitted models in both cases are different. This could be as a result of the additional characteristic (variance nonstationarity) of the series considered in this study.

5 Conclusion

In this paper, we used Box-Jenkins procedure to model Nigerian GDP series from 1999 to the second quarter of 2013. The time series was found to be non stationary in mean and variance. Log transformation and first order differencing satisfactorily rendered the series stationary. SARIMA (2, 1, 2) X $(1, 0, 1)_4$ was tentatively fitted to the GDP series following careful inspection of its accompanying plots. Analysis of residuals from the fitted model confirmed the adequacy of the model.

A comparison between the model and one previously proposed in the literature showed some differences. We therefore, conclude that proper transformation and differencing of time series can lead to change in model structure. Thus, Care should then be taken to apply these transformations to time series data when necessary.

References

[1] Abdulraheem, A, Sectorial contributions to gross domestic product in Nigeria, 1971-2005, Nigeria Forum, (2011), 1-12.

[2] Akpanta, A.C. and Iwueze, I.S., On applying the Bartlett transformation method to time series data, *Journal of Mathematical Sciences*, **20**(3), (2009), 227-243.

[3] Box, G.E.P., Jenkins, G.M and Reinsel, G.C., *Time series analysis, forecasting and control*, 3rd ed., Prentice and Hall, USA, 1994.

[4] Chatfield, C., *The analysis of time series*, 5th ed., Chapmann and Hall, London, 1995.

[5] Delurgio, S.A., *Forecasting principles and applications,* McGraw Hill, Boston, 1998.

[6] Efiok, S.O., Tapang, A.T. and Eton, O.E., The impact of human capital cost on gross domestic product in Nigeria, *International Journal of Financial Research*, **3**(4), (2012), 116-126.

[7] Gujarati, D.N and Porter, D.C., *Basic econometrics*, 4th ed., McGraw Hill, 2004.

[8] Hamilton, J. D., *Time series analysis*, Princeton University Press, New Jersey, 2008.

[9] Ion, D and Adriana, A., Modelling unemployment rate using Box-Jenkins Procedure, *Journal of Applied Quantitative methods*, **3**(2), (2008), 156-166.

[10] Montgomery, D.C., Jennings, C.L and Kulahci, M., *Introduction to time series analysis and forecasting*, John Wiley and Sons, New Jersey, 2008.

[11] Nwabueze, J.C., Causal relationship between gross domestic product and personal consumption expenditure of Nigeria, *African Journal of Mathematics and Computer Science Research*, **2**(8), (2009), 179-183.

[12] Reininger, M.T. and Fingerlos, M.U., Modelling seasonality of gross domestic product in Belgium, *Working Paper*, (2007),1-22.

[13] Umaru, A. and Zubaru, A.A., An empirical analysis of the contribution of agriculture and petroleum sector to the growth and development of the Nigerian economy, 1960-2010, *International Journal of Social Sciences*, **2**(4), (2012), 758-769.

[14] Wei, W.W.S., *Time series analysis: Univariate and multivariate methods*, 2nd ed., Pearson Addison Wesley, Boston, 2006.