Science Journal of Applied Mathematics and Statistics
2017; 5(5): 174-180
http://www.sciencepublishinggroup.com/j/sjams
doi: 10.11648/j.sjams.20170505.12
ISSN: 2376-9491 (Print); ISSN: 2376-9513 (Online)

Science PG
Science Publishing Group

Methodology Article

Predictive Model with Square-Root Variance Stabilizing Transformation for Nigeria Crude Oil Export to America

Obinna Adubisi[1, *], Titus Terkaa Mom[1], Chidi Emmanuel Adubisi[2], Phillip Luka[1]

[1]Department of Mathematics and Statistics, Faculty of Pure & Applied Sciences, Federal University Wukari, Wukari, Nigeria

[2]Department of Physics, Faculty of Physical Science, University of Ilorin, Ilorin, Nigeria

Email address:

adubisiobinna@fuwukari.edu.ng (O. Adubisi)

[*]Corresponding author

To cite this article:

Obinna Adubisi, Titus Terkaa Mom, Chidi Emmanuel Adubisi, Phillip Luka. Predictive Model with Square-Root Variance Stabilizing Transformation for Nigeria Crude Oil Export to America. *Science Journal of Applied Mathematics and Statistics*. Vol. 5, No. 5, 2017, pp. 174-180. doi: 10.11648/j.sjams.20170505.12

Received: January 20, 2017; **Accepted**: September 19, 2017; **Published**: November 5, 2017

Abstract: In the last few decades, crude oil has claimed the topmost position in Nigerian export list, constituting a very fundamental change in the structure of Nigerian international trade. In this study, secondary data on monthly crude oil export to the United States was obtained from the Energy Information Administration (EIA) database. Using the Box-Jenkins (ARIMA) methodology, the results showed that Seasonal ARIMA (0, 1, 1) $(1, 0, 1)_{12}$ model had the least information criteria after the data was Square-Root transformed and non-seasonally first differenced in order to achieve series stationarity. The diagnostic tests on the selected model residuals revealed the residuals are normally distributed uncorrelated random shocks.

Keywords: Transformation, SARIMA, Unit Root, Crude Oil Export, ARCH-LM

1. Introduction

Crude oil is considered the major source of energy in Nigeria and the world in general. Crude oil being the mainstay of the Nigerian economy plays a vital role in shaping the economic and political destiny of the country. Nigeria exports most of its crude oil to countries like India, United States of America, Brazil, The Netherland, United Kingdom and Spain [1]. As recently as 2010, Nigeria provided about 10 percent of over-all United States oil imports and ranked as the fifth-largest source of oil imports in the United States. However, Nigeria crude oil export to the United States has recently being declining as a result of the boom in shale oil and lifting of the United States 40-years ban on crude oil exports by the U.S. Senate [2]. Therefore, the ultimate aim of this study is to construct a statistical model that could be used to monitor the export pattern of crude oil export from Nigeria to the United States. Using this model, forecast of future values of crude oil export to the United States can be obtained. A lot of studies have been carried out using time series ARIMA modelling approach to identify patterns and appropriate models. Adubisi [3] used ARIMA procedure in modelling the growth pattern of reserve currency in Nigeria, Iwueze et al; [4], modelled the Nigeria external money reserves with Autoregressive Integrated Moving average model, ARIMA modelling approach was used to model yearly exchange rates between USD/KZT, EUR/KZT and SGD/KZT, and the actual data compared with developed forecasts by Daniya [5], Kumar and Anand [6] used ARIMA modelling approach to forecast sugarcane production in Indian. Adubisi and Jolayemi [7] used ARIMA-intervention analysis modelling approach to evaluate and estimate the impact of the financial crisis on Nigeria Crude oil export. Furthermore, Smart [8] explored the feasibility for application of Box-Jenkins Approach (ARIMA) in modelling and forecasting maternal mortality Ratios (MMR) and Adubisi et al; [9] used the seasonal ARIMA to model the Nigeria money in circulation series and also produced a three years forecast values using the fitted model.

2. Main Research

2.1. Material

The data used for this study are secondary data on Nigeria monthly crude oil export to the United States obtained from the Energy Information Administration (EIA) database for twenty-three consecutive years, from January 1993 to December 2015 [10].

2.2. Power Transformation

The parametric family of transformations from the original series to a transformed series was originally proposed by Box and Cox [11]. The power transformation is a continuously varying function with respect to the power parameter lambda λ, in a piecewise function form that makes it continuous at the point of singularity $(\lambda = 0)$. Suppose we observe a $(n\times 1)$ vector of observations $(x_1,...,x_n)$ in which each $x_i > 0$, the power transform is expressed as

$$x_t^{(\lambda)} = \begin{cases} \dfrac{(x_t^{\lambda}-1)}{\lambda (GM(x))^{\lambda-1}} & ,\lambda \neq 0 \\ GM(x)\log x_t & ,\lambda = 0 \end{cases} \tag{1}$$

$GM(x) = (x_1,...,x_n)^{\frac{1}{n}}$ is the Geometric mean of the data. It generalizes both the square root and the log transformation and admits a likelihood ratio test to select the best fitting parameter. The one-parameter Box-Cox transformation is expressed as

$$x_t^{(\lambda)} = \begin{cases} \dfrac{(x_t^{\lambda}-1)}{\lambda} & ,\lambda \neq 0 \\ \log x_t & ,\lambda = 0 \end{cases} \tag{2}$$

For more details on Box-Cox variance stabilization transformation procedures, see Box and Cox [11], Yan [12], Carroll and Ruppert [13], Nishili [14], Sakia [15] and Bickel et al; [16]. The various transformation parameter lambda (λ) values and the appropriate transformation attached to each are summarized in the Table 1.

Table 1. *Transformation for Some Values of Parameter Lambda* (λ).

S/No.	1	2	3	4	5	6	7
Lambda (λ)	0	0.5	-0.5	-1	1	-2	2
Transformation	$\log x_t$	$\sqrt{x_t}$	$1/\sqrt{x_t}$	$1/x_t$	No transformation	$1/x_t^2$	x_t^2

Source: Minitab (2010).

2.3. Box-Jenkins Methodology

The Autoregressive integrated moving average (ARIMA) model procedure popularized by Box and Jenkins [17] and Box *et al*. [18]. The ARIMA (p,d,q) model which contain both the non-seasonal and the seasonal parameters is written as

$$\varphi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D Y_t = \theta_q(B)\Theta_Q(B^S)\varepsilon_t \tag{3}$$

The observed values is Y_t, (B) represent the Backshift operator, t is the time, (φ_p,Φ_P) represent the non-seasonal and seasonal autoregressive coefficients parameters with the roots within the unit circle

$$\varphi_p(B) = 1-\varphi_1 B - \varphi_2 B^2 - ... - \varphi_p B^p \tag{4}$$

$$\Phi_P(B^S) = 1-\Phi_1 B^S - \Phi_2 B^{2S} - ... - \Phi_P B^{PS} \tag{5}$$

(θ_q,Θ_Q), represent the non-seasonal and seasonal moving-average coefficients parameters with the roots within the unit circle

$$\theta_q(B) = 1-\theta_1 B - \theta_2 B^2 - ... - \theta_q B^q \tag{6}$$

$$\Theta_Q(B^S) = 1-\Theta_1 B^S - \Theta_2 B^{2S} - ... - \Theta_Q B^{QS} \tag{7}$$

While $(1-B)^d$ is the regular differencing which is applied to remove the stochastic trend in the series, $(1-B^S)^D$ is the seasonal differencing applied to remove the series seasonal effects and ε_t is the white noise error i.e. $\varepsilon_t \sim WN(0,\sigma^2)$. More details on seasonal ARIMA can be found in Box and Jenkins [17], Box *et al*; [18], Pankratz [19].

2.4. Unit Root Tests

The Augmented Dickey-Fuller (ADF) test and the KPSS tests are performed to determine if the series contains a unit root. The tests are based on the assumption that the time series data Y_t follows a random walk. The Augmented Dickey-Fuller (ADF) test, corresponding to modelling a random walk pattern with drift around a stochastic trend

$$Y_t = \alpha + \rho Y_{t-1} + \sum_{i=1}^{p-1} \partial_i \nabla y_{i-i} + \beta t + \varepsilon_t \tag{8}$$

The expression $\rho y_{t-1} + \sum_{i=1}^{p-1} \partial_i \nabla y_{i-i}$ is the augmented part, y_{t-1} is the lagged term, ∇y_{t-1} shows the lagged change, t and α represent the deterministic trend and drift components respectively, the ε_t is the error term and ρ, ∂ are coefficients to be estimated. If $\rho = 1$ the model is said to be non-stationary and this implies the presence of a unit root in the series. The null hypothesis is H_0 : ($\rho = 1$ or $\partial = 0$) against the alternative ($\rho < 1$ or $\partial < 0$). When the p-value is greater than the alpha, this would lead to none rejection of the null hypothesis. The KPSS test with a random walk $\alpha_t = \alpha_{t-1} + \varepsilon_t$ allowed is expressed as

$$Y_t = \alpha_t + \beta_t + \varepsilon_t \tag{9}$$

The procedure has a null hypothesis of stationary series and an alternative of non-stationary series. A p-value less than alpha (α) at 5% level of significance from the result of the KPSS test would be enough to reject the null hypothesis. Details on ADF and KPSS see, Dickey and Fuller [20] and Kwiatkowski, et al; [21].

3. Results and Discussion

The series plot in Figure 1 depicts sharp peaks and toughs in the crude oil export series, suggesting some influence and also the variance was observed not to be stable over the periods used in this study.

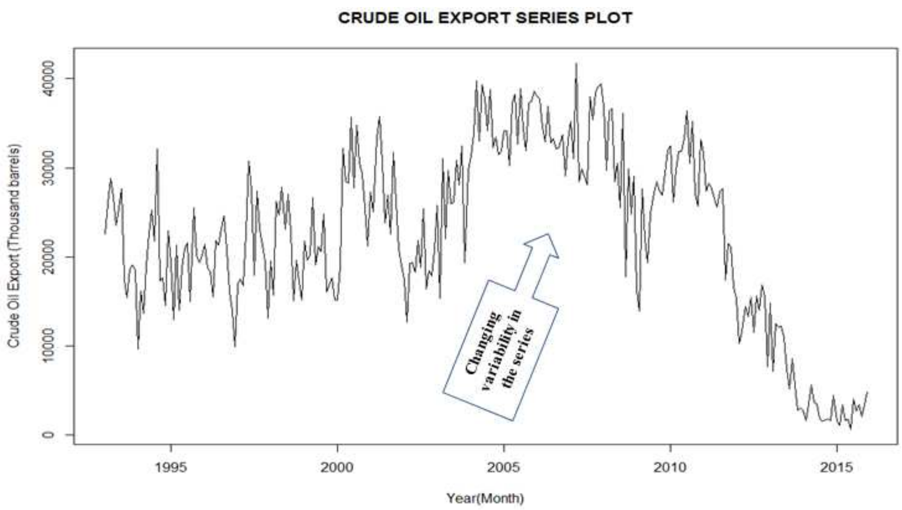

Figure 1. *Crude Oil Export Series Plot.*

The non-stationarity claim from the series plot was affirmed from the slow decay in the autocorrelation function (ACF) of the data while significant spikes at lag 1 and 2 of the partial-autocorrelation function (PACF) are noticed in Figure 2. Therefore, the data requires some form of transformation and differencing to make it stable (stationary).

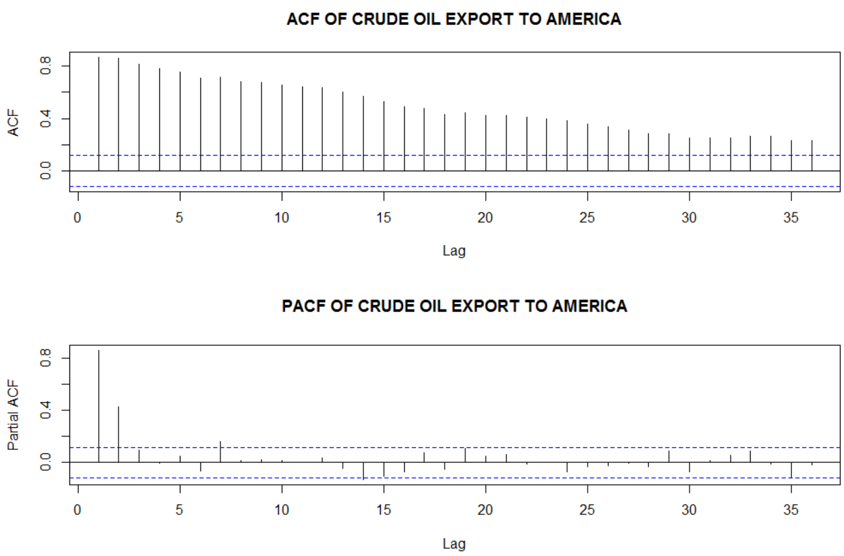

Figure 2. *ACF and PACF of the Actual Data.*

The Square-Root transformation was used to stabilize the series variance based on the result of the computed transformation lambda value ($\lambda = 0.5$), using equation (2) with the help of Minitab. The series was also non-seasonally differenced to achieve stationarity. Figure 3, depicts the non-seasonal first-order differenced crude oil export series.

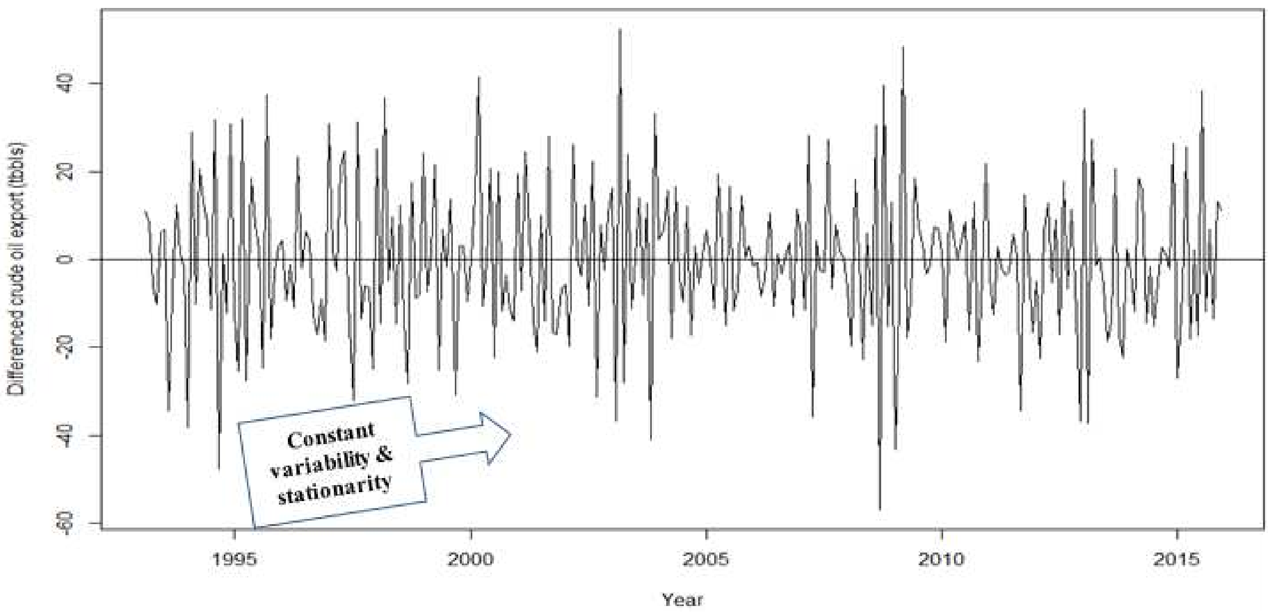

Figure 3. *Non-seasonal Differenced Crude Oil Export Series.*

The results of the ADF and KPSS unit root tests in Table 2, also indicates that the data is stationary after the square root transformation and the non-seasonal first-order difference was applied on the series.

Table 2. *The Unit Root Tests for Differenced Data.*

Summary of test statistics			
Test Type	**Test Statistics**	**Lag Order**	**P-value**
ADF	-8.4569	6	0.01
KPSS	0.0843	3	0.1

The decay in the correlogram of the differenced transformed data in Figure 4, when compared with the 95% confidence limits $\left(\pm\frac{2}{\sqrt{n}}=\pm 0.35\right)$. The PACF decays faster towards zero while the ACF cuts off after the lag 1 with a significant spike at lag 12.

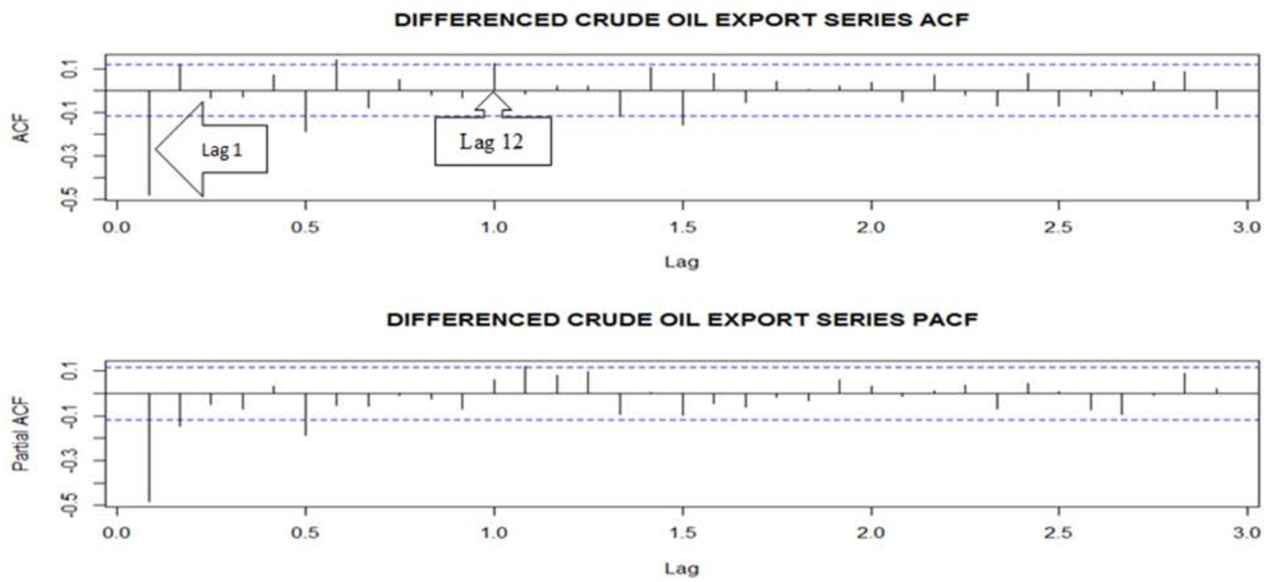

Figure 4. *ACF and PACF of the Transformed Differenced Series.*

The various tentative Seasonal ARIMA model structures extracted from the correlogram plots in Figure 4 are presented in Table 3.

Table 3. *Tentative Seasonal ARIMA models.*

Tentative Models			
MODELS	**AIC**	**AICc**	**BIC**
ARIMA (2, 1, 1) (0, 0, 1)$_{12}$	2279.15	2279.37	2297.24
ARIMA (2, 1, 1) (1, 0, 1)$_{12}$	2281.20	2281.51	2302.90
ARIMA (0, 1, 1) (0, 0, 1)$_{12}$	2279.58	2279.67	2290.43
ARIMA (1, 1, 1) (1, 0, 1)$_{12}$	2273.64	2273.86	2291.72
ARIMA (1, 1, 2) (1, 0, 1)$_{12}$	2275.24	2275.55	2296.94
*ARIMA (0, 1, 1) (1, 0, 1)$_{12}$	2273.19	2273.34	2287.66

"*" Means the best fitted model based on the selection criteria.

The Seasonal ARIMA (0, 1, 1) (1, 0, 1)$_{12}$ was found to fit the series based on the AIC, AICc and BIC selection criteria values. The model parameter estimates in Table 4 are statistically significant with the t-values greater than ± 2. The parameters also satisfy the stationarity and invertibility conditions of the ARIMA methodology.

Table 4. *Estimates of ARIMA (0, 1, 1) (1, 0, 1)$_{12}$ Model.*

Model Fit Statistics				
Parameter	**Coefficients**	**Standard Error**	**t**	**p-value**
MA1	-0.5507	0.0504	-10.926	0.001
SAR1	0.9518	0.0569	16.727	0.001
SMA1	-0.8690	0.0953	- 9.118	0.001

The seasonal ARIMA (0, 1, 1) (1, 0, 1)$_{12}$ model expressed in backshift format using the parameters in Table 4, is presented as

$$(1-\Phi_1 B^{12})(1-B)y_t = (1+\theta_1 B)(1+\Theta_1 B^{12})\varepsilon_t \qquad (10)$$

$$(1-0.9518B^{12})\hat{y}_t = (1-0.5507B)(1-0.8690B^{12})\varepsilon_t$$

The fitted seasonal ARIMA model from equation (10) is mathematically expressed as

$$\hat{y}_t = 0.9518y_{t-12} - 0.5507\varepsilon_{t-1} - 0.8690\varepsilon_{t-12} + 0.4785\varepsilon_{t-13} + \varepsilon_t \qquad (11)$$

The model adequacy check performed using the correlogram plots of the model residuals coupled with other objective tests like the Box-Ljung test [22], Shapiro-Wilk Normality test [23] and the ARCH-LM test [24] are presented in Table 5. The tests results failed to reject the null hypothesis at the 5% level of significance confirming that the residuals are normally distributed with no autocorrelation and no conditional heteroscedasticity (ARCH) effects. It implies that the residuals from the fitted seasonal ARIMA model are Gaussian white noise.

Table 5. *Sarima (0, 1, 1) (1, 0, 1)$_{12}$ Residuals Diagnostic Tests.*

Ljung-Box,	ARCH-LM	Shapiro-Wilk test statistics
Test type	**Test statistics**	**P-value**
Ljung-Box	14.407	0.2113
ARCH-LM	6.9373	0.7313
Shapiro-Wilk	0.99443	0.4104

The Gaussian white noise residuals is clearly portrayed in Figure 5 by the randomness of the residuals, non-significant spikes in the ACF residuals plot and the probability values falling above the 0.05 limit in the probability plot.

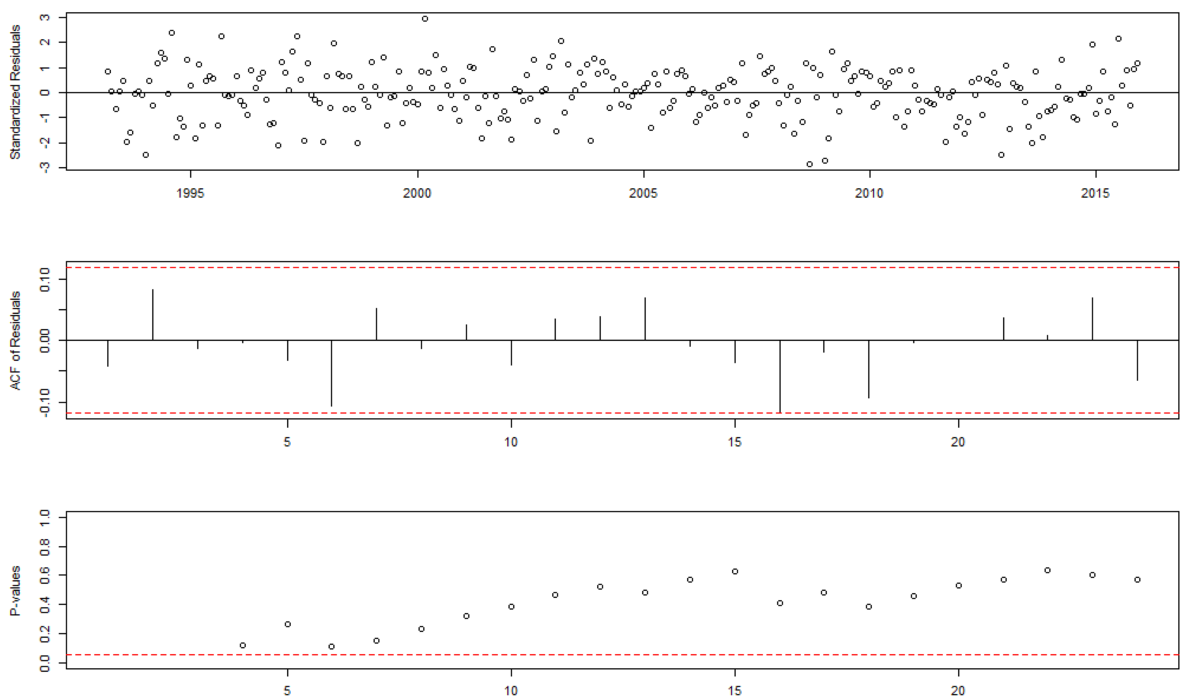

Figure 5. *Fitted Model Residuals Diagnostic Plots.*

Figure 6, depicts the plot of the fitted values against the actual crude oil export series and it clearly shows that the selected model best describe very well the variations in the study data series.

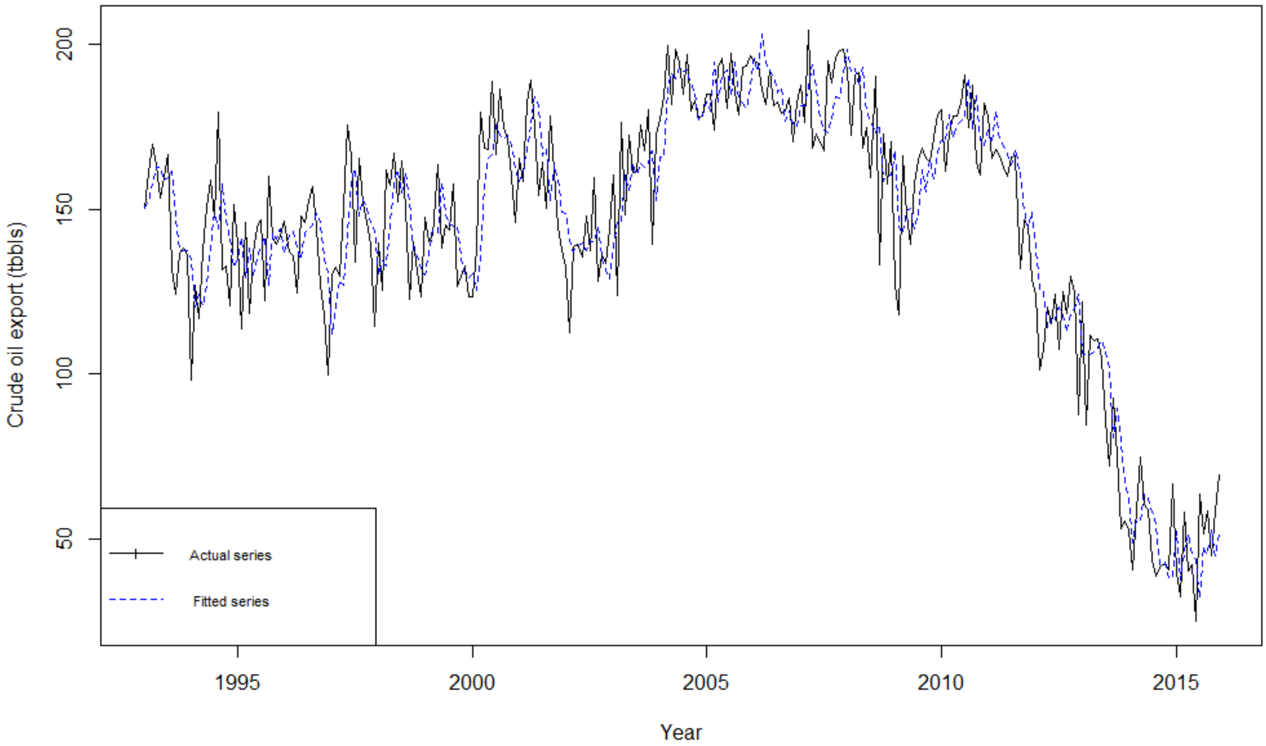

Figure 6. *Actual Values Versus Model Fitted Values.*

4. Conclusion

In this study, the crude oil export to the United States was modelled using the Box-Jenkins (ARIMA) modelling procedures. The data required square-root transformation and differencing to achieve series stationarity. The transformed series was subjected to the Box and Jenkins iterative procedure for ARIMA model building. The results from analysis showed that the appropriate model for the series is the Seasonal ARIMA (0, 1, 1) $(1, 0, 1)_{12}$ model. The model adequacy tests confirmed that the model residuals are normally distributed uncorrelated random shocks. This model is therefore recommended for use in the forecast of Nigeria crude oil export to the United States, until proving otherwise.

References

[1] Central Intelligence Agency (2013). United States central Intelligence Agency (CIA) World Fact-book, www.cia.gov.

[2] Vanguard Nigeria (2016): Importing Nigeria crude oil. http://www.vanguardngr.com/2016/us-importing-nigeria-crude-oil.

[3] Adubisi (2016). Modelling the growth pattern of reserve currency in Nigeria. *FUW Trends in Sciences & Technology Journal.* 1(1): 130 – 133.

[4] Iheanyichukwu S. Iwueze, Eleazar C. Nwogu and Valentine U. Nlebedim (2013). *Time Series Modelling of Nigeria External Reserves*. CBN Journal of Applied Statistics 4 (2).

[5] Daniya Tiegenova (2014). Forecasting Exchange Rates Using Time Series Analysis. The sample of the currency of Kazakhstan.

[6] Kumar Manoj and Anand Manhu (2012). An application of time Series ARIMA Forecasting model for predicting sugarcane production in India. *Studies in Business and Economics*. 81-94.

[7] Adubisi and E. T. Jolayemi (2015). Estimating the impact on the Nigeria Crude oil Export from 2002 to 2013. (An ARIMA-Intervention Analysis*). International Journal of Scientific and Engineering Research.* 6(10): 878-886.

[8] Smart (2013). Modelling and Forecasting Maternal Mortality; An Application of ARIMA Models, *International Journal of Applied Science and Technology.* 3(1): 19 – 28.

[9] O. D. Adubisi, C. C. Eleke, T. T. Mom and C. E. Adubisi (2017). Application of SARIMA to modelling and forecasting money circulation in Nigeria. *Asian Research Journal of Mathematics*. 6(1): 1–10. DOI: 10.9734/ARJOM/2017/35555.

[10] EIA (2009). Energy Information Administration (EIA); *World oil statistics*. OPEC, Zurich, Switzerland. Obtained from www.Eia.org on 24/8/2016.

[11] Box G. E. P and Cox D. R (1964): An analysis of transformations. *Journal of the Royal statistical society*, series B 26(2): 211 – 252.

[12] Yan L (2015): Variance stabilizing properties of Box-Cox transformation for dependent observations. *Advances in science, Technology and Environmentology*. B12 (3): 63 – 70.

[13] Carroll R. J and Ruppert D (1960): On prediction and the power transformation family. *Biometrika* 68: 609 – 615.

[14] Nishii R (2001): Box-Cox transformation Encyclopaedia of Mathematics, Springer, ISBN: 978-1-55-608-010-4.

[15] Sakia R. M (1992): The Box-Cox transformation technique: A review. *The statistician*. 41: 169 – 178.

[16] Bickel P. J and Doksum K. A (1981): An analysis of transformation revisited. *Journal of the American statistical Association*. 76(374): 296 – 311.

[17] Box G. E. P & Jenkins G. M (1976). *Time series Analysis Forecasting and Control*, Revised edn. Holden-Day, San Francisco-USA. ISBN: 0816211043.

[18] Box G. E. P, Jenkins G. M & Reinsel G. C (1994). *Time series Analysis Forecasting and Control,* 3rd edn. Prentice Hall, Englewood Cliffs, New Jersey-USA.

[19] Pankratz, A. (1983). *Forecasting with Univariate Box-Jenkins model: concepts and case,* John Wiley and Sons, ISBN: 0471090239, New-York-USA.

[20] Dickey D. A & Fuller W. A (1979). Distribution of the Estimators for Autoregressive time series with a unit root. *Journal of the American statistics Association.* 7(366): 427 431. JSTOR: 2286348.

[21] Kwiatkowski D., Phillips, P. C. B, Schmidt P., and Shin Y. (1992). *Testing the Null Hypothesis of Stationarity against the Alternative of a unit Root*. Journal of Econometrics 54: 159-178.

[22] Ljung G. M and Box G. E. P (1978). On a measure of a lack of fit in Time Series Models. *Biometrika.* 65 (2): 297 – 303. Doi: 10.1093/biomet/65.2.297

[23] Shapiro, S. S., and Wilk, M. B. (1965): "An analysis of variance test for normality (complete samples). *Biometrika* 52(3-4): 591-611 doi 10.1093/biomet/52.3-4.591. JSTOR 2333709. MR205384. p. 593.

[24] Engle R: (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica, (50): 987-1008.