Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 128(2) (2019) 348-362 EISSN 2392-2192

Time Series ARIMA Model for Predicting Nigeria Net Foreign Direct Investment (FDI)

Akpensuen Shiaondo Henry[1,*], Edeghagba Eghosa Elijah[2], Alhaji Abdullahi Gwani[2], Joel Simon[1]

[1]Department of Mathematical Sciences, Abubakar Tafawa Balewa University, Bauchi, Nigeria

[2]Department of Mathematical Sciences, Bauchi State University Gadau, Nigeria

*E-mail address: henroo85@gmail.com

ABSTRACT

This paper presents an empirical study of modelling and forecasting time series data of Nigeria net foreign direct investment (FDI). The Box-Jenkins ARIMA methodology was used for forecasting the yearly data collected from 1972 to 2018. Result of the analysis revealed that the series became stationary at first difference. The diagnostic checking has shown that ARIMA (1, 1, 2) is appropriate or optimal model based on the Akaike's Information Criterion (AIC), the Bayesian Information Criterion (BIC) and Hannan Quinn criterion (HQ). A twenty (20) year forecast was made from 2019-2039, the result of the forecast showed that the net FDI in Nigeria will continue to grow in the period forecasted. These forecasts will help policy makers in Nigeria to sustain their efforts to expand the tax base, reduce red tape, and strengthen the regulatory framework to investment and also investors friendly policies in order to attract the much needed FDI.

***Keywords*:** ARIMA, ADF, Foreign Direct Investment, Forecasting, Nigeria

1. INTRODUCTION

Time series modelling is a dynamic research area which has attracted attentions of research community over the last few decades. The main aim of time series modelling is to carefully collect and rigorously study the past observations of a time series to develop an

(Received 25 April 2019; Accepted 10 May 2019; Date of Publication 11 May 2019)

appropriate model which describes the inherent structure of the series. This model is then used to generate future values for the series, i.e. to make forecasts. Time series forecasting thus can be termed as the act of predicting the future by understanding the past [1]. Due to the indispensable importance of time series forecasting in numerous practical fields such as business, economics, finance, science, engineering, etc. [2-4], proper care should be taken to fit an adequate model to the underlying time series. It is obvious that a successful time series forecasting depends on an appropriate model fitting. A lot of efforts have been done by researchers over many years for the development of efficient models to improve the forecasting accuracy. As a result, various important time series forecasting models have been evolved in literatures. One of the most popular and frequently used stochastic time series models is the autoregressive integrated moving average (ARIMA), [3-6] 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) [5-16] models.

Nigeria, like other African countries, recognizes the contribution of FDI to economic development and integration into the world economy. Nigeria since pre-independence era till date has being making considerable efforts to improve its investment climate through liberation, deregulation, privatization and enabling laws and incentives. However, the much-expected surge in FDI into Nigeria has not occurred. This is particularly worrisome, as Nigeria possesses almost all the attributes of a good FDI destination. These include size of market, availability of natural resources, low labour cost and high productivity, incentives, high level of human capital development, major markets proximity, etc. Nigeria needs FDI because it is favoured over other forms of private capital flows. Portfolio equity and debt are subject to reversals in financial crises period, while FDI is more resilient. FDI is critical to the country as it is the key source of large pool of capital necessary for the development of the country. However, despite several fiscal incentives by the government, foreign direct investment has remained dismal (The Punch 2002). The cost of not having foreign direct investment is high. A decline in investment reduces the expansion of output, variety and quality, leading to reduced market share and potentially declining non-price competitiveness. Several studies have been done on modelling FDI, [7] forecasted FDI inflows into India using ARIMA models, hinged on the Box – Jenkins technique; over the period 1976 – 2003 and found out that there is an expected increase of FDI volumes over the period 2004 – 2025. Biswas (2015) studied net FDI inflows in India using the Box Jenkins ARIMA model over the period 1992 – 2014 and found out that FDI in India is following an increasing trend over the forecasted period (2015 – 2034). In another Indian study, [8] analysed foreign institutional investment inflows using the ARIMA models (based on the Box – Jenkins methodology) over the period January 2004 to September 2012 and found out that various AR terms and MA terms influence the current inflow or out flow of foreign institutional investment. [9] modelled and forecasted FDI inflows into SAARC using ARIMA models (based on the Box – Jenkins approach) over the period 1970 – 2012 and found out that the total value of FDI expected for the next 25 years (2013 – 2037) is US $1672895.8 million and average FDI expected for the next 25 years is US $66915.81 million for SAARC. Closer to home and more recently, in Zambia, [10] forecasted FDI inflows using ARIMA models (based on the Box – Jenkins technique) over the period 1970 – 2014, and found out that there will be a gradual increase in annual net FDI inflows of about 44.36% by 2024 in Zambia. [11] applied Box-Jenkins ARIMA approach to predicting the net FDI inflow in Zimbabwe over the period (1980-2017), he found that the predicted net FDI inflow over the next two (2) decade show a relatively poor and unimpressive growth trend.

Our empirical literature review indicates that no similar study has been done in Nigeria so far. Hence the aim of this work is to apply ARIMA approach to predict net FDI inflows in Nigeria.

2. MATERIALS AND METHODS

2. 1. An Autoregressive Process AR (p)

The notation AR (p) refers to autoregressive model of order p. The AR(p) model specify as

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_{t.} \tag{1}$$

where $\varphi_1,\ldots,\varphi_p$ are AR(p) parameters to be estimated, c is a constant and the random variable $\varepsilon_{t.}$ is error term.

2. 2. Moving Average Model

The notation MA (q) is the moving average model of order q

$$X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}, \text{ t} = 1,2 \ldots \tag{2}$$

where $\theta_1,\ldots,\theta_q$ are MA(q) parameters to be estimated, μ is the mean of X_t and ε_t is the error term

2. 3. Autoregressive moving average

The autoregressive moving average model is denoted by ARMA (p, q) with p autoregressive term and q moving average term.

$$X_t = c + \varepsilon_{t.} \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{3}$$

2. 4. Autoregressive Integrated Moving Average.

Autoregressive integrated moving average (ARIMA) models are specific subset of univariate modelling, in which a time series is expressed in terms of past values of itself (the autoregressive component) plus current and lagged values of a "white noise" error term (the moving average component). ARIMA models are univariate models that consist of an autoregressive polynomial, an order of integration (d), and a moving average polynomial.

A process (X_t) is said to be an autoregressive integrated moving average process, denoted by ARIMA (p, d, q) if it can be written as:

$$\varphi(B)\nabla^d X_t = \theta(B)\varepsilon_t. \tag{4}$$

where $\nabla^d = (1 - B)^d$ (1–B) with $\nabla^d X_t$ and d^{th} consecutive differencing.

If E($\nabla^d X_t$) = μ , we write the model as

$$\varphi(B)\nabla^d X_t = \alpha + \theta(B)\varepsilon_t. \tag{5}$$

where, α is a parameter related to the mean of the process (X_t), α = μ, ($\varphi_1 \ldots \varphi_p$) and this process is called a white noise process, that is, a sequence of uncorrelated random variables from a fixed distribution (often Gaussian) with constant mean E(X_t) = μ ,usually assumed to be "zero" and constant variance. If d = 0, it is called ARMA (p, q) model while when d = 0 and q = 0, it is referred to as autoregressive of order p model and denoted by AR (p). When p = 0 and d = 0, it is called Moving Average of order q model, and is denoted by MA (q).

2. 5. Box-Jenkins Modelling Approach

The Box-Jenkins model uses iterative three-stage modelling approach which is:

1) Model identification
2) Model estimation
3) Model checking

2. 5. 1. Model identification

The first step in developing a Box–Jenkins model is to determine if the time series is stationary. Time series Stationarity can be access from time plot of the series plot. It can also be detected from an autocorrelation plot. Specifically, non-stationary is often indicated by an autocorrelation plot with very slow decay. Finally, unit root tests provide a more formal approach to determining the degree of differencing such as Augmented Dickey Fuller test (ADF), Kwiatkowski-Phillips-Schmidt-Shin (KPSS) and Phillips-Perron Unit Root Tests. But for the purpose of the work we shall consider the ADF test. Once state of being stationary have been addressed, the next step is to identify the order (i.e. the p and q) of the autoregressive and moving average terms. These are determined by examining the values of the autocorrelations and the partial autocorrelations with their corresponding plots (12).

2. 5. 1. 1. Dickey Fuller Test

The test was introduced by Dickey and Fuller (1979) to test for the presence of unit root(s). The regression model for the test is given as

$$\Delta FDI_t = \rho FDI_{t-1} + \beta X_{t-1} + \delta_1 \Delta FDI_{t-1} + \delta_2 FDI_{t-2} + \cdots + \delta_p FDI_{t-p} + e_t. \tag{6}$$

The hypothesis testing

$H_{0:} = \rho = 0$(the series contain unit root).
$H_1: \rho < 0$ (the series is stationary).

where:

ΔFDI_t is the difference series.
FDI_{t-1} is the immediate previous observation.
δ_1,…,δ_p is the coefficient of the lagged differenced term up to p.

X_t is the optimal exogenous regresses which may be constant or constant trend.
ρ and β parameters to be estimated.

2. 5. 2. Model Estimation

After an optimal model has been identified, the model estimation method makes it possible to estimate simultaneously all the parameters of the process, the order of integration coefficient and parameters of an ARMA structure. The parameters are estimated using the method of maximum likelihood

2. 5. 3. Model Verification

The last step in Box-Jenkins methodology is model verification or model diagnosis. The conformity of white noise residual of the model fit will be judged by plotting the ACF and PACF of the residual to see whether it does not have any pattern or we perform Ljung-Box Test on the residual. The test hypothesis:

Ho: There is no serial correlation
H_1: There is serial correlation

The test statistics of the Ljung-Box

$$LB = n(n+2)\sum_{k}^{m}\frac{\rho_k^2}{n-k}. \tag{7}$$

where n is the sample size, $m = lag\ length$ and p is the sample autocorrelation coefficient.

The decision: if LB is less than critical value of x^2, then we do not reject the null hypothesis. This means that a small value of Ljung-Box statistic will be in support of no serial correlation i.e the errors are normally distributed.

2. 6. Information Criteria

The three information criteria used are as follows:

$$AIC(n) = \log(\sigma^2) + \frac{2n}{T}. \tag{8}$$

$$BIC(n) = \log(\sigma^2) + \frac{nlog(T)}{T}. \tag{9}$$

$$HQ(n) = \log(\sigma^2) + \frac{2nlog(T)}{T}. \tag{10}$$

where n is the dimensionality of the model σ^2 is the maximum likelihood estimate of the white noise variance, and T is the sample size

3. DATA ANALYSIS

In order to achieve the goal of the study, we explore the data on yearly net foreign direct investment in Nigeria from 1972 to 2018 (denoted by FDI). The FDI was measured in US Dollars per annum. The software packages used in this study are Gretl and E-views 8

3. 1. Stationarity Test

The time plot of FDI and the ADF test were used to test the stationarity of the data as shown in Figure 1 and Tables 1, 2, 3, 4, 5 and 6

Table 1. ADF test at level with constant.

Null Hypothesis: FDI has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag = 9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.336295	0.6046
Test critical values:	1% level	-3.584743	
	5% level	-2.928142	
	10% level	-2.602225	

*MacKinnon (1996) one-sided p-values.

Table 2. ADF test at level with constant and linear trend

Null Hypothesis: FDI has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 4 (Automatic - based on SIC, maxlag = 9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-3.552131	0.0469
Test critical values:	1% level	-4.198503	
	5% level	-3.523623	
	10% level	-3.192902	

*MacKinnon (1996) one-sided p-values.

Table 3. ADF test at level with none.

Null Hypothesis: FDI has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag=9)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.614210	0.4457

Test critical values:	1% level	-2.617364
	5% level	-1.948313
	10% level	-1.612229

*MacKinnon (1996) one-sided p-values.

From Figure 1 in the appendix, it is observed that the FDI series display a non-stationary pattern with up and down movement. The ADF test from Tables 1, 2 and 3 indicate that the FDI is not stationary at levels. Hence the need to check for stationarity at first difference as shown in Tables 4, 5 and 6.

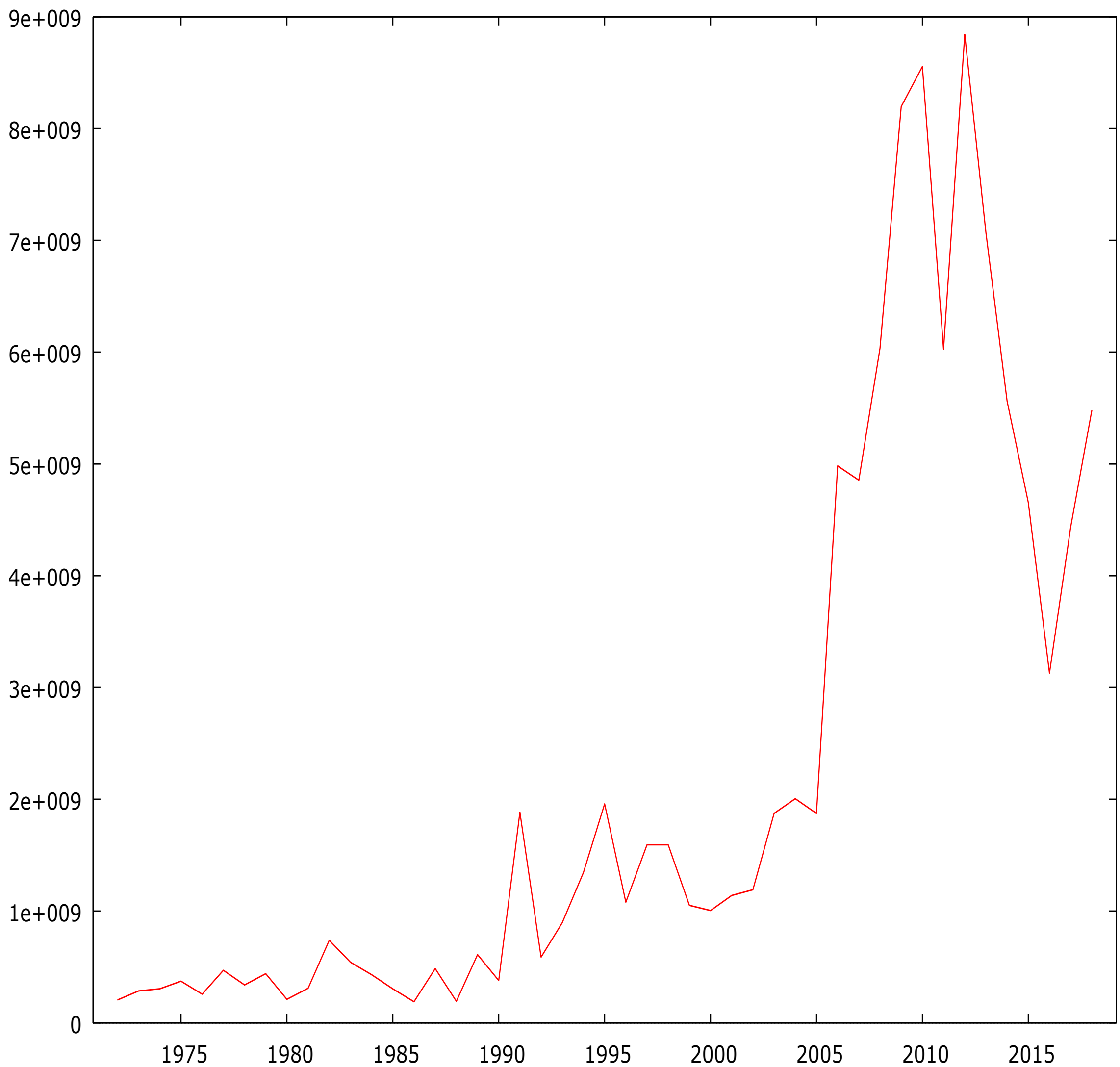

Figure 1. Time plot of FDI

Table 4. ADF test at first difference: constant.

Null Hypothesis: D(FDI) has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag (= 9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.152274	0.0000
Test critical values:	1% level	-3.588509	
	5% level	-2.929734	
	10% level	-2.603064	

*MacKinnon (1996) one-sided p-values.

Table 5. ADF Test at First difference: constant and linear trend

Null Hypothesis: D(FDI) has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 9 (Automatic - based on SIC, maxlag (= 9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-5.277503	0.0007
Test critical values:	1% level	-4.243644	
	5% level	-3.544284	
	10% level	-3.204699	

*MacKinnon (1996) one-sided p-values.

Table 6. ADF test at first difference: none.

Null Hypothesis: D(FDI) has a unit root
Exogenous: None
Lag Length: 0 (Automatic - based on SIC, maxlag (=9)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-8.170947	0.0000
Test critical values:	1% level	-2.618579	
	5% level	-1.948495	
	10% level	-1.612135	

*MacKinnon (1996) one-sided p-values.

Stationarity is reported in Tables 4, 5 and 6.

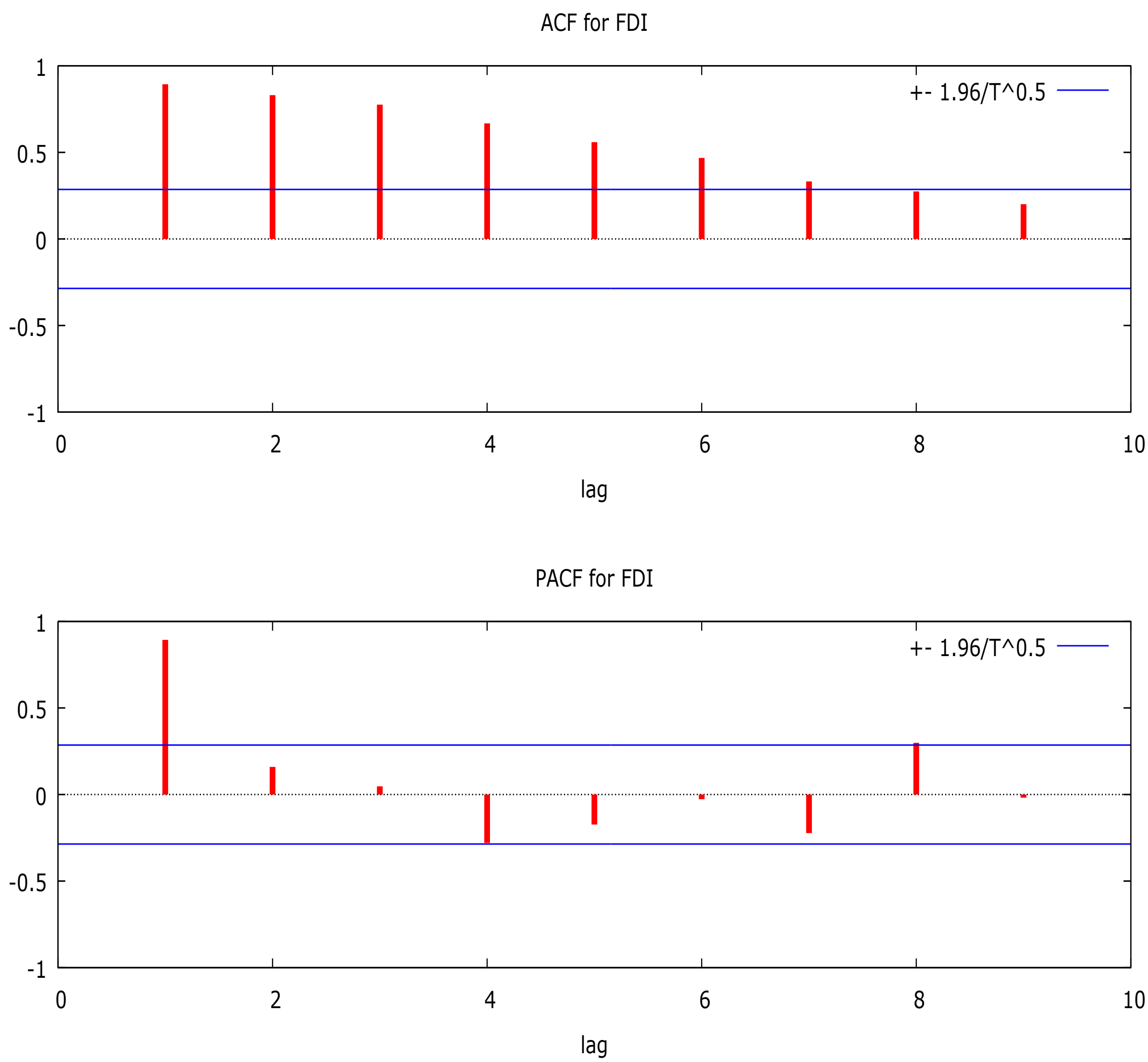

Figure 2. Plot of ACF and PACF of FDI at Levels

3. 2. ARIMA Model Identification

Table 7. Tentative ARIMA Models.

Model	AIC	BIC	HQ
ARIMA(1, 1, 0)	2042.101	2045.759	2043.473
ARIMA(0, 1, 1)	2042.275	2045.932	2043.648

ARIMA(1, 1, 1)	2044.101	2049.587	2046.156
ARIMA(2, 1, 1)	2045.817	2053.132	2048.556
ARIMA(1, 1, 2)	**2041.557**	**2042.871**	**2042.279**
ARIMA(2, 1, 0)	2044.100	2049.586	2046.1555
ARIMA(1, 1, 3)	2041.427	2050.570	2044.852

From Table 7, the optimal model is ARIMA (1, 1, 2) with minimum information criteria

3. 3. Model Estimation

Table 8. Estimated Model.

Model 1: Model 3: ARIMA, using observations 1973-2018 (T = 46)
Dependent variable: (1-L) FDI
Standard errors based on Hessian

	Coefficient	*Std. Error*	*z*	*p-value*	
phi_1	0.480324	0.267813	1.794	<0.0001	***
theta_1	−0.792193	0.338215	−2.342	<0.0001	***
theta_2	0.472374	0.278797	1.694	<0.0001	***

Mean dependent var	1.15e+08	S.D. dependent var	1.04e+09
Mean of innovations	1.04e+08	S.D. of innovations	9.78e+08
Log-likelihood	−1017.778	Akaike criterion	2041.557
Schwarz criterion	2042.871	Hannan-Quinn	2042.297

From Table 8, Parameters of ARIMA (1, 1, 2) was estimated using the maximum likelihood method. The parameters were found to be significant.

3. 4. Model Checking

We plot the ACF and PACF of the residual of ARIMA (1, 1, 2) and performed the Ljung-Box test on the residual of ARIMA (1, 1, 2) model to check if there exists serial correlation in the residual. From figure 3 un appendix, the ACF and PACF plot of residual from ARIMA (1, 1, 2) model shows that all correlations are within the threshold limits indicating that the residuals are behaving like white noise. A Ljung-Box test result in table 9 returns a large p-value (0.8524), also suggesting that the model is adequate.
Our optimal ARIMA (1, 1, 2) model is specify as

$$FDI_t = 0.480324FDI_t - 0.792193\boldsymbol{\varepsilon_{t-1}} + 0.472374\varepsilon_{-2}. \qquad (11)$$

Table 9. Ljung-Box test on Residuals of ARIMA (1, 1, 2).

Model	Test Statistic	p-value
ARIMA (1,1,2)	6.39473	0.8458

Having satisfied that the model is adequate, the next step is to use the model to forecast future values.

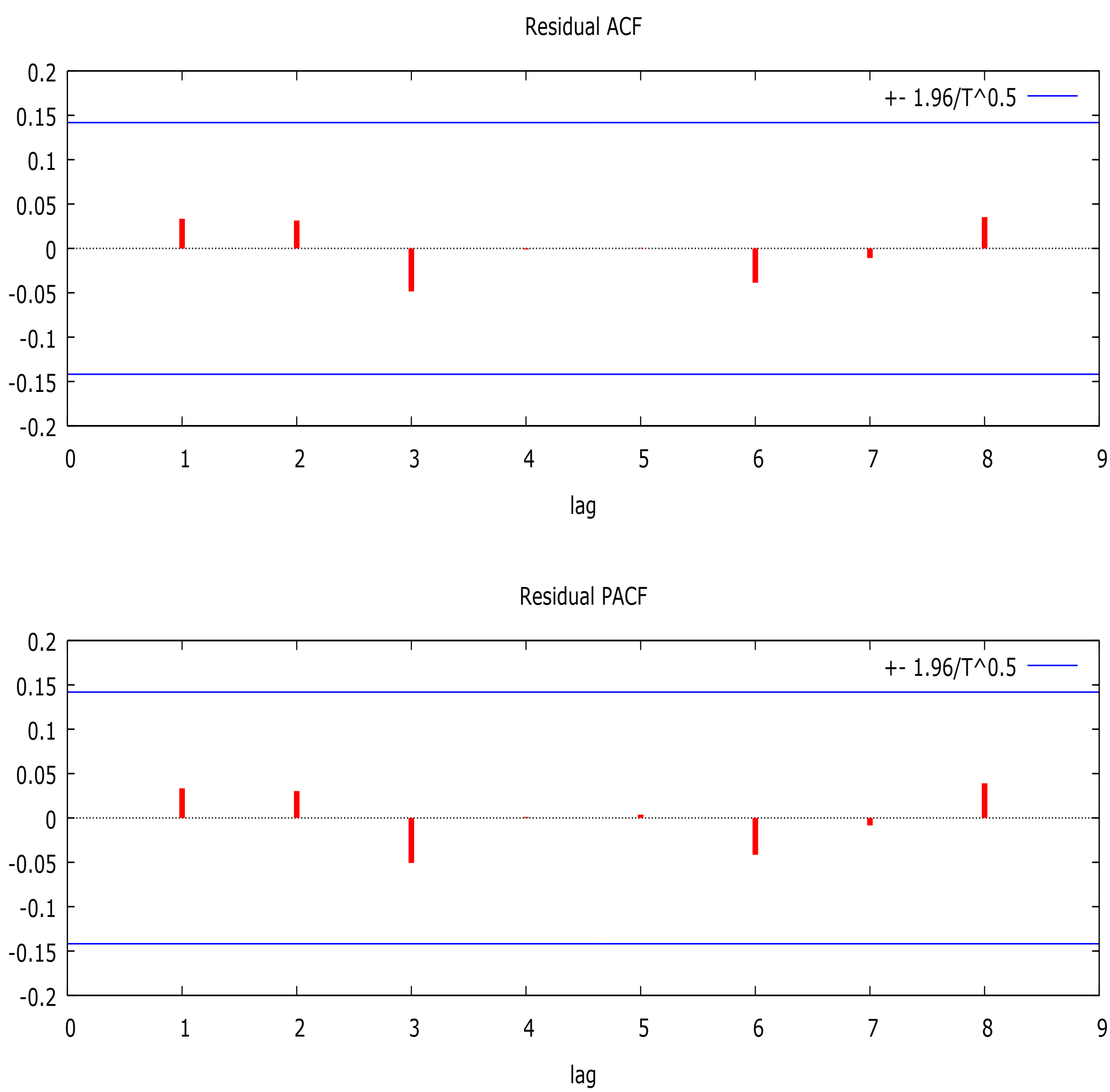

Figure 3. Plot of ACF and PACF of Residuals from ARIMA (1, 1, 2)

Table 10. 20 years forecast FDI from ARIMA (1, 1, 2)

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

OBS	Prediction	std. error	95% interval
2019	4.91494e+009	9.78479e+008	(2.99716e+009, 6.83272e+009)
2020	5.73740e+009	1.18776e+009	(3.40942e+009, 8.06537e+009)
2021	6.13244e+009	1.54558e+009	(3.10316e+009, 9.16173e+009)
2022	6.32219e+009	1.92086e+009	(2.55738e+009, 1.00870e+010)
2023	6.41334e+009	2.27201e+009	(1.96029e+009, 1.08664e+010)
2024	6.45711e+009	2.59238e+009	(1.37615e+009, 1.15381e+010)
2025	6.47814e+009	2.88462e+009	(8.24383e+008, 1.21319e+010)
2026	6.48824e+009	3.15312e+009	(3.08238e+008, 1.26682e+010)
2027	6.49309e+009	3.40194e+009	(-1.74582e+008, 1.31608e+010)
2028	6.49542e+009	3.63441e+009	(-6.27891e+008, 1.36187e+010)
2029	6.49654e+009	3.85318e+009	(-1.05555e+009, 1.40486e+010)
2030	6.49708e+009	4.06031e+009	(-1.46099e+009, 1.44551e+010)
2031	6.49734e+009	4.25744e+009	(-1.84710e+009, 1.48418e+010)
2032	6.49746e+009	4.44587e+009	(-2.21628e+009, 1.52112e+010)
2033	6.49752e+009	4.62664e+009	(-2.57053e+009, 1.55656e+010)
2034	6.49755e+009	4.80062e+009	(-2.91149e+009, 1.59066e+010)
2035	6.49756e+009	4.96851e+009	(-3.24053e+009, 1.62357e+010)

2036	6.49757e+009	5.13091e+009	(-3.55883e+009, 1.65540e+010)
2037	6.49757e+009	5.28833e+009	(-3.86735e+009, 1.68625e+010)
2038	6.49757e+009	5.44119e+009	(-4.16696e+009, 1.71621e+010)
2039	6.49758e+009	5.58988e+009	(-4.45838e+009, 1.74535e+010)

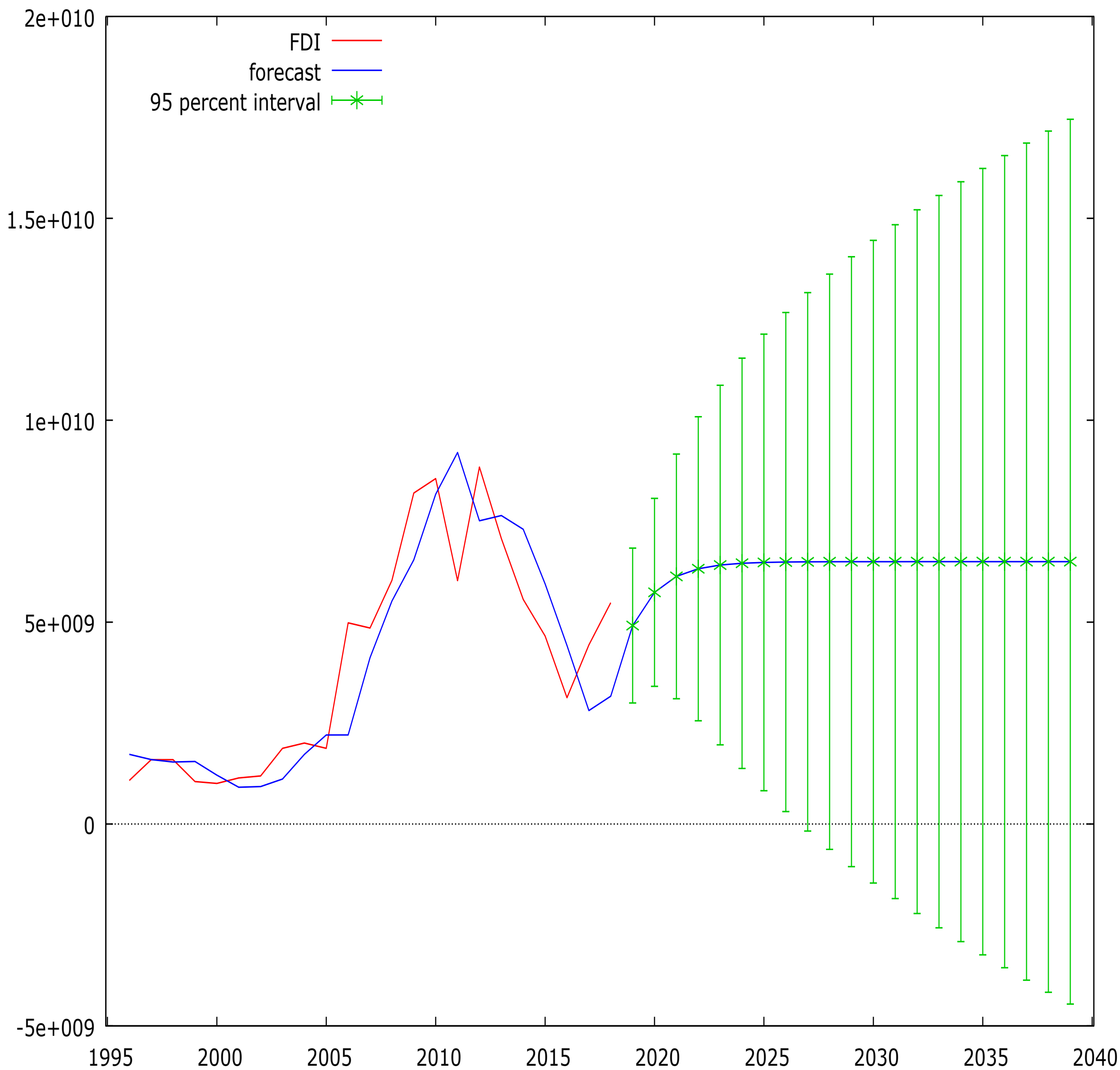

Figure 4. Forecast Graph of FDI

The net FDI investment in Nigeria has been increasing steadily over the years and the forecast of the net FDI indicates that the trend will continues for the period forecasted (2018-2039). This is attributed to the fact that the Nigerian government is working correct anomalies that affect inflow of FDI.

Efforts to expand the tax base, reduce red tape, and strengthen the regulatory framework to investment are being pursued, albeit with varying degrees of success. This should have enhanced the lure of Nigeria's business environment, which in turn attracts FDI.

The wider business operating environment has improved, and indeed Nigeria jumped 24 places to 145 out of 190 countries surveyed in the 2017 World Bank Doing Business index. Considerable progress has also been made on the drive to reduce Nigeria's dependency on oil. A more diversified economy has made FDI more attractive, and results in a more stratified economy.

4. CONCLUSIONS

In this work, we model Nigeria net FDI investment using the Box Jenkins methodology for the period 1972 to 2018. The modelling was in three stages, the first was model identification stage, where the series became stationary after the first difference base on result provided the ADF test. Based on AIC, BIC and H-Q information criteria, ARIMA (1, 1, 2) was selected as appropriate model with minimum information criteria.

The second stage was parameter estimation, where the parameters conform to stationary condition. Finally, the third stage was model diagnosis where the errors from the residual of the model were normally distributed and no presence of serial correlation. A forecast for the period of 30 years show that the net FDI investment in Nigeria will continue to grow for the period forecasted. It is important to note that predicting economic variables such as FDI is not an easy process because the results of such studies are easily affected by structural breaks in the economy. The result of the estimated model remains relevant as long as there are no structural breaks [11].

The policy implication of this study is that policy makers in Nigeria should sustained their efforts to expand the tax base, reduce red tape, and strengthen the regulatory framework to investment and investors friendly policies in order to attract the much needed FDI.

The study can be extended to non-linear model models to take care of volatility which may occur in the FDI series.

References

[1] Hanzacebi .C (2008). Improving Artificial Neural Network Performance in Seasonal Time Series Forecasting. *Information Science,* 178, 4550-4559

[2] Zhang .G.P (2007). Neural Network Ensemble Method with Jilted Training Data for Time Series Forecasting. *Information Science,* 177, 5329-5346

[3] Zhang. G.P (2003). Time Series Forecasting Using Hybrid ARIMA and Neural Network Model. *Neural Computing,* 50, 159-175

[4] Tong H. Threshold Model in Non-Linear Time Series Analysis, Spring-Verlag, New York, 1983

[5] Box G.E.P and Jenkins G. Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco, 1970.

[6] Hipel K.W and McLeod A.I. Time Series Modelling of Water Resources and Environmental Systems. Amsterdam, Elsevier, 1994.

[7] Adewumi. S (2006). The Impact of FDI on Growth in Developing Countries: An African Experience. Jonkoping International Business School, Jonkoping University, Sweden.

[8] Dhingra, V.S., Bulsara, H.P & Ghandhi, S (2015). Forecasting Institutional Flow Towards India Using ARIMA Modeling, *Management* 75, 13-26

[9] Prasanna, P.W.L (2015). Modelling and Forecasting Foreign Investment (FDI) into SAARC for the Period of 2013-2037 with ARIMA. *International Journal of Business and Social Science* 6(2), 264-272

[10] Jere, S., Kasense, B. and Chilyabanyama, O. (2017) Forecasting Foreign Direct Investment to Zambia: A Time Series Analysis. *Open Journal of Statistics*, 7, 122-131. https://doi.org/10.4236/ojs.2017.71010

[11] Thabani Nyoni (2018). Box Jenkins Approach to Predicting Net FDI Inflows in Zimbabwe. *MPRA Paper* NO 87737

[12] Moffat, I. U. & Akpan, E. A. (2014). Time Series Forecasting: A Tool for Out-Sample Model and evaluation. *American Journal of Scientific and Industrial Research* 5(6), 185-194

[13] Alvaro Santos Pereira, João Tovar Jalles, Martin A. Andresen. Structural change and foreign direct investment: globalization and regional economic integration. *Portuguese Economic Journal* April 2012, Volume 11, Issue 1, pp 35-82. https://doi.org/10.1007/s10258-011-0077-9

[14] Ng S, Perron P (1995) Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *J Am Stat Assoc* 90: 268-281

[15] Coakley J, Fuertes A-M, Smith RP (2006). Unobserved heterogeneity in panel time series models. *Comput Stat Data Anal* 50(9): 2361-2380

[16] Carr DL, Markusen JR, Maskus KE (2001). Estimating the knowledge-capital model of the multinational enterprise. *Am Econ Rev* 91: 693-708