Forecasting of Post-Covid-19 Import Value Index in Nigeria using Box-Jenkins Methodology

OGUNLADE TEMITOPE OLU
Department of Mathematics, Ekiti State University,
Ado-Ekiti, NIGERIA
&
Landmark University SDG 3 Research Group (Good Health and Well-being),
Department of Physical Sciences (Mathematics Programme),
Landmark University, Omu-Aran, Kwara State,
NIGERIA

AKINDUTIRE OPEYEMI ROSELYN
Department of Accounting (Actuarial Science Unit), Ekiti State University,
Ado-Ekiti, NIGERIA

FAWEYA OLANREWAJU
Department of Statistics, Ekiti State University,
Ado-Ekiti, NIGERIA

BALOGUN KAYODE OGUNTUASE
Federal School of Statistics, Oyo State,
NIGERIA

OKORO JOSHUA OTONRITSE
Landmark University SDG 3 Research Group (Good Health and Well-being),
Department of Physical Sciences (Mathematics Programme),
Landmark University, Omu-Aran, Kwara State,
NIGERIA

Abstract: - This study employed the Box-Jenkins Methodology otherwise known as the Autoregressive Integrated Moving Average (ARIMA) modelling to model and forecast the series for the period of 2018 to 2030. The results indicated an upward trend with fluctuations while the series was stationary at first difference, i.e the series was $I(1)$. Based on the Akaike information criterion (AIC) and Bayesian Information Criteria (BIC) choice criteria, it was found that ARIMA (2, 1, 2) model was better suited to the import value index (IVI) series. Diagnosed check of the model reveals that the error was random, normally distributed and there was no serial correlation, in the same vein, thirteen years forecast was made which shows fluctuation pattern in import value index (IVI) series.

Key-Words: -Import value index, COVID-19, Forecast, ARIMA, Box-Jenkins, Autocorrelation.

Received: June 10, 2021. Revised: February 7, 2022. Accepted: February 21, 2022. Published: March 24, 2022.

1 Introduction

The ratio which explains the percentage change in nominal value to base value is said to be import value index (IVI) [1]. The index tells the percentage of import at a given point in time relative to its base point. However, when the number of exported goods in a country is far more than the imported goods, such country is said to be experiencing economic development. Presently, the advent of coronavirus disease 2019 (COVID-19), a global pandemic has created uncertainties in international trade, especially among African countries (for instance, Nigeria, South Africa and Angola) whose markets are driven by the exportation of commodities whose prices have crashed in the international market [2]. In the same vein, it is very essential for a country to involve in foreign trade which thereafter bring about economic growth development [3, 4]. It is generally suggested that most developing countries recorded an indefinite reduction which contributed substantially to the breakdown of oil prices on the market in their foreign exchange revenue from the early 1980s: A forecast on import value index (IVI) using Autoregressive Integrated Moving Average (ARIMA) model.

The transportation of produce, human and resources (financial and nonfinancial) through national disturbances, has been disposed of, particularly in recent time. Academic, trade and technical studies from various economies have emphasized foreign trade on a global scale [5] and external financial flowsas determinants of buoyant growth in countries that are keen on identifying and taking advantage of opportunities. Besides, they argue that agro-based market is a stimulus for expansion, particularly in countries whose main source of incomes (national and foreign) and employment generations are from agricultural produces. They further note that agricultural trade creates varieties of choices for consumers. The country having bountiful products for farming; nevertheless, the country's most agricultural export commodity have not been completely handled for industrial and agricultural business yet. Before the advent of oil discovery and extraction in the country in 1960s, agriculture was the giant wellspring of exportation, but it has been hijacked by the crude oil. Thus, agricultural activities have gradually declined, particularly during the 1970s' oil boom. Ever since, Nigeria is a net importer of grain and agricultural products. Therefore, due to over-dependence on oil and the decrease in agricultural production, Nigeria has started importing goods which can be produced locally, however, loosed its potentials of agriculture [6-10]. Active players in the agricultural sector have claimed that, if assisted by strong, effective and long-lasting government agricultural policies, only the middle belt of the country could supply the rice demand for all of West Africa. Low cereal yield in Nigeria are due to higher production costs, lack of fertilizers, failure to maintain irrigation facilities, and lack of labour. Management methods such as weeding, transplanting and harvesting rely on minimal family labour. Presently, a number of related formal models have been formulated to forecast some selected cereals such as maize, sorghum etc. In this study, we are applying the univariate time series model to justify truly whether past values of Nigeria Cereals Production (CP) series can predict its current and future values using methodology technically known as ARIMA modeling.

2 Materials and Methods

An annual time series data on import value index (IVI) in Nigeria was used for this research work. In this study, an annual time series data on food production index (FPI) in Nigeria ranges from 1980 to 2017 was originated from the record of World Bank through [11]. In this study, Box-Jenkins methodology which is also known as ARIMA modeling propounded by Box and Jenkins [12] was used in analysing the import value index in Nigeria. According to Pankratz [13], the autoregressive integrated moving average (ARIMA) model reveals the relationship between the time series data and its former valence. There are four major steps in creating a good model which are; identification, estimation, diagnostic checking and forecasting [12]. The methodology is however focusing on making non-stationary time series data stationary by differencing. The general equation of the ARIMA (p, d, q) model is as follows;

$$X_t = \mu + \emptyset_1 X_{t-1} + \emptyset_2 X_{t-2} + \cdots + \emptyset_p X_{t-p} + \theta_1 e_{t-1} + \theta_2 e_{t-2} + \cdots + \theta_q e_{t-q} + \varepsilon_t \quad (1)$$

Also, studies [14,15] showed that the mathematical equation above can be expressed using lag polynomial as shown in (2).

$$\phi(L)(1-L)^d X_t = \theta(L)\varepsilon_t \quad (2)$$

More so, (2) express polynomial factorization property with p=p'-d which is given written as;

$$\left(1-\sum_{i=1}^{p}\phi_i L^i\right)(1-L)^d X_t = \left(1+\sum_{i=1}^{p}\theta_j L^j\right)\varepsilon_t \quad (3)$$

Where L is the lag operator, ϕ_i are the parameters of the autoregressive part of the model, the θ_j are the parameters of the moving average part, ε_t are the error terms which are generally assumed to be independent, identically distributed (i.i.d). Also, p means the number of preceding ("lagged"), X values to be added or subtracted from X in the model, in order to make better predictions based on local growth time or decrease in data capturing of the ARIMA's auto-regressive existence. More so, d is the number of times that variations need to be made between the data to generate a stationary signal that i.e a signal that has constant mean over time covering the integrated (I) essence of ARIMA. And q represents the number of lagged values for the error word, which captures the moving average (MA) portion of ARIMA.

2.1 Model Modification

Box-Jenkins methodology is one of the popular method used in fitting an appropriate models to a given time series. Thus, AR and MA of orders p and q must be critically examine the Sample Autocorrelation Function (SACF) as well as the Sample Partial Autocorrelation Function (SPACF) for X_t prior to the estimation of ARMA (p, q) for a given time series X_t.The SAFC is used for detecting the order ***q*** of the MA term while the SPACF is used for detecting the order ***p*** of the AR term from the sample correlogram. By sample correlogram, we mean a plot of SACF $\hat{\rho}_k$ against the lags k [16]. Mathematically, the sample Autocorrelation Function (SAFC) $\hat{\rho}_k$at lag k is defined as:

$$\hat{\boldsymbol{\rho}}_{\boldsymbol{k}} = \frac{\hat{X}_k}{\hat{X}_0} = \frac{\text{Covariance at lag K}}{\text{Variance}} = \frac{\sum(X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum(X_t - \bar{X})^2}, \; k = 0, 1, 2, \dots \qquad (4)$$

Where: $\hat{\boldsymbol{\gamma}}_{\boldsymbol{k}}$ is the sample covariance at lag k, $\hat{\boldsymbol{\gamma}}_{\boldsymbol{0}}$ is the sample variance and $\bar{\boldsymbol{X}}$ is the sample mean.
The most popular model selection criteria with specific utility are Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) which is otherwise known as Schwarz Information Criterion (SIC) [17]. The mathematical expression for AIC and BIC are as follow:

$$\text{AIC} = 2\text{k} - 2\ln(\hat{L}) \qquad (5)$$

$$\text{BIC} = \text{k}\ln(\text{n}) - 2\ln(\hat{L}) \qquad (6)$$

Where $\hat{\boldsymbol{L}}$is the maximized value of the likelihood function of the model M, i.e$\hat{\boldsymbol{L}}$ = **p(x / $\hat{\boldsymbol{\theta}}$, M),** also $\hat{\boldsymbol{\theta}}$are the parameter values that maximize the likelihood function, x is the observed data, n is the number of effective data point in x and k is the number of estimated parameters in the model. Konishi and Kitagawa [18] derived the BIC to approximate the distribution of the data, integrating out the parameters using Laplace's method which is given below;

$$\text{P}(\text{x} / \text{M}) = \int p(x / \theta, \text{M})\pi(\theta / \text{M})d\theta \qquad (7)$$

Where $\pi(\theta / \text{M})$ is the prior for θ under model M. Another model selection criterion which is also common but less used is the Hannan–Quinn information criterion (HQC) [19]. The mathematically expression is as follows;

$$\text{HQC} = -2\text{L}_{\text{max}} + 2\text{k}\ln(\ln(\text{n})) \qquad (8)$$

Where: $\boldsymbol{L_{max}}$ is the log-likelihood, **k** is the number of parameters and ***n*** is the number of observations.
After a tentative model has been identified, the next step is to estimate the parameters in the model.The last steps in creating a good model is forecast. Therefore, forecast is then made for the series

3 Data Analysis

This section focuses on the data analysis and the interpretation of results carried out on Import Value Index (IVI) series using GRETL (Gnu Regression, Econometrics and Time – series Library) version 1.9.8.

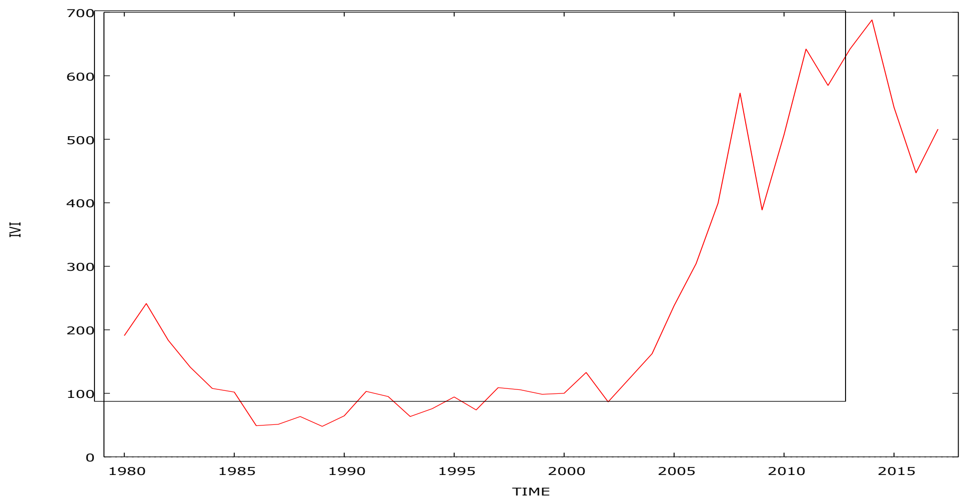

Fig. 1: Time series plot of Import Value Index (IVI)

Fig. 1 shows that the import value index (IVI) is increasing with time (i.e upward trend) although with fluctuations. This means that the series will need to be differenced for it to become stationary. To determine the order of integration(d), this study conducts unit root tests such as Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) tests. Tables 1a and 1b indicate the results of unit root tests conducted on the IVI series.

Table 1a. Augmented Dickey-Fuller (ADF) Test for IVI Series

		ADF		
Test Variable	ADF Statistics	Critical Values	p-value	Integration Order
IVI	-6.439140	-2.945842	<0.0001***	I(1)

Table 1b. Phillips-Perron (PP) Test for IVI Series

		PP		
Test Variable	PP Statistics	Critical Values	p-value	Integration Order
IVI	-6.424313	-2.945842	0.0001***	I(1)

The results from tables 1a and 1b as shown above reveals that the IVI series is differenced stationary series of order one {I(1)}. It can be deduced from their p-values (0.0001) since both less than the chosen level of significance (α=0.05). The ACF and PACF are employed to determine the type and order of the model since the series has become stationary [20,21].

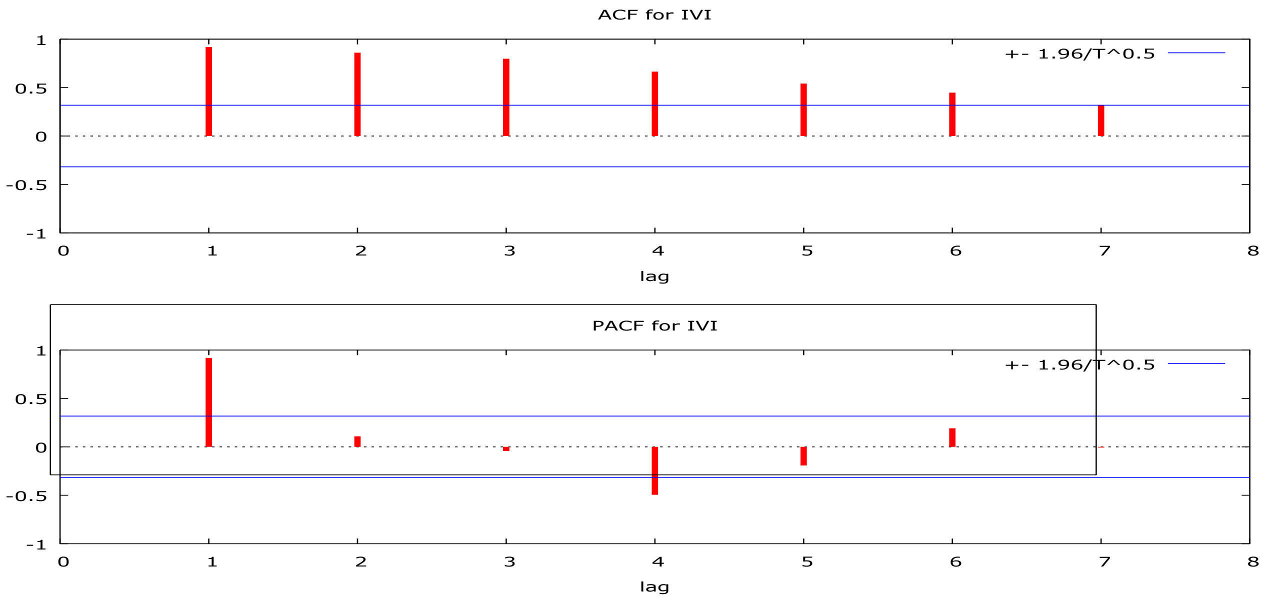

Fig. 2: ACF and PACF for the IVI stationary series after first difference

Figure 2 reveals that the spikes of the ACF and PACF are decaying exponentially while spikes of ACF are significant at lag 1, lag 2, lag 3 and lag 4 respectively and the PACF has only two significant spikes at lag 1 and lag 4. Consequently, ARIMA (p, d, q) model has any possible order from p=1, 2, 3, 4 of AR term and q=1, 2, 3 and 4 of MA term is therefore suggested to be fitted for the data. According to the result revealed by the ACF and PACF, selection criteria such as Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC) and Hannan–Quinn information criterion (HQC) were also employed for further detection of order p and q. This is important so as to select the best model from all the possible fitted models as it is shown in the table 2 below.

Table 2. Possible Fitted ARIMA (p, d, q) Models

S/N	(p, d, q)	AIC	BIC	HQC
1	(1, 1, 1)	426.3203	432.7640	428.5920
2	(1, 1, 2)	424.5162	432.5708	427.3558
3	(1, 1, 3)	423.6620	433.3275	427.0695
4	(1, 1, 4)	421.6775	432.9539	425.6529
5	(2, 1, 0)	425.4543	431.8979	427.7260
6	(2, 1, 1)	424.1405	432.1951	426.9801
7	**(2, 1, 2)**	**421.2783***	**430.9438***	**424.6859***
8	(2, 1, 4)	423.2348	436.1222	427.7782
9	(3, 1, 1)	423.2917	432.9572	426.6992
10	(3, 1, 3)	421.9339	434.8213	426.4773
11	(3, 1, 4)	423.9124	438.4106	429.0237
12	(4, 1, 0)	423.2506	432.9161	426.6582
13	(4, 1, 1)	425.0123	436.2887	428.9877
14	(4, 1, 2)	421.9353	434.8226	426.4787
15	(4, 1, 3)	422.4256	436.9238	427.5369

It can be seen from table 2 that ARIMA (2, 1, 2) model has the smallest values of the AIC, BIC and HQC selection criteria. Thus, it can be confirmed as the best model for the import value index (IVI) series. Since the best model has been known. Therefore, model estimation of ARIMA (2, 1, 2) model can be done which is presented on Table 3.

Table 3. The estimates of AR and MA terms of ARIMA (2, 1, 2) model

Variables	Coefficient	Std.Error	Z	p-value
Const.	7.23672	8.77671	0.8245	0.4096
AR1	-0.662325	0.0811522	-8.162	3.31e-016 ***
AR2	-0.962503	0.0564810	-17.04	4.06e-065 ***
MA1	0.493345	0.152095	3.244	0.0012***
MA2	0.850897	0.139246	6.111	9.92e-010 ***

$$X_t = 7.23672 - 0.662325Y_{t-1} - 0.962503Y_{t-2} + 0.493345e_{t-1} + 0.850897e_{t-2} + \varepsilon_t \quad (9)$$

Table 3 shows the estimates of AR and MA terms of ARIMA (2, 1, 2) model. It was shown that prior two years values **t-2** (i.e2015 and 2016) of IVI series for the AR term and the random shocks in the MA term are related to the IVI in current time t (i.e 2017). The p-values from their test statistic also proved that it was statistically significant since the chosen significant level (α = 0.05) is greater than their p-values.

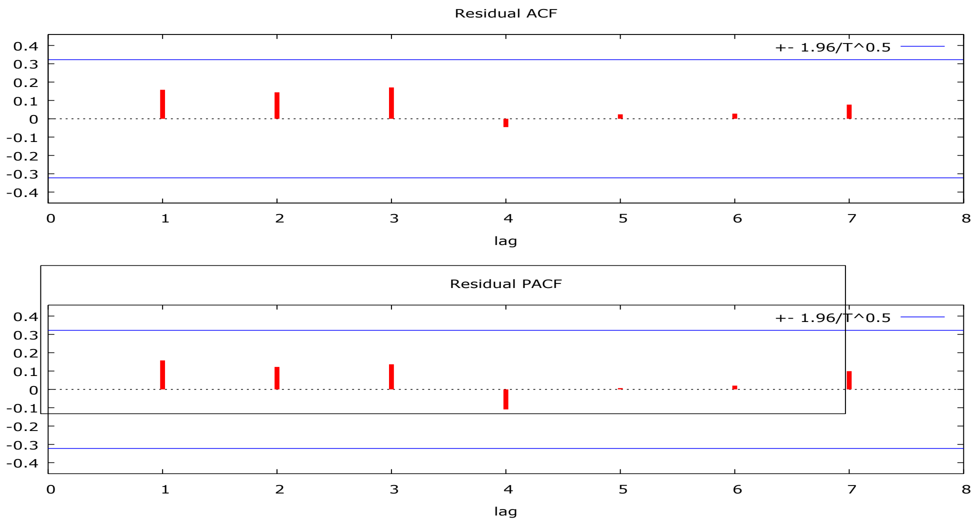

Fig. 3: Correlogram residuals of the fitted model

Figure 3 shows that the spikes for both ACF and PACF of the residual correlogram are not statistically significant. It indicates that ARIMA (2, 1, 2) model is a best fit to the Import Value Index (IVI) series.

Table 4. Forecast of the Import Value Index (IVI) series based on Thirteen Years

Year	Prediction	Std.Error	[95% Conf. Interval]	
2018	483.868299	59.573133	367.107104	-600.629494
2019	441.446518	77.458684	289.630288	-593.262749
2020	519.444463	91.937944	339.249403	-699.639522
2021	527.610735	109.35016	313.288352	-741.933117
2022	466.123909	121.41926	228.146533	-704.101285
2023	517.983243	130.45924	262.287830	-773.678657
2024	561.811925	142.46949	282.576865	-841.046986
2025	501.863493	153.00903	201.971298	-801.755688
2026	518.378736	160.23799	204.318054	-832.439418
2027	584.135975	169.2281	252.454995	-915.816956
2028	543.68251	178.82857	193.184959	-894.180061
2029	526.179453	185.41616	162.770459	-889.588447
2030	595.703891	192.45468	218.499641	-972.908141

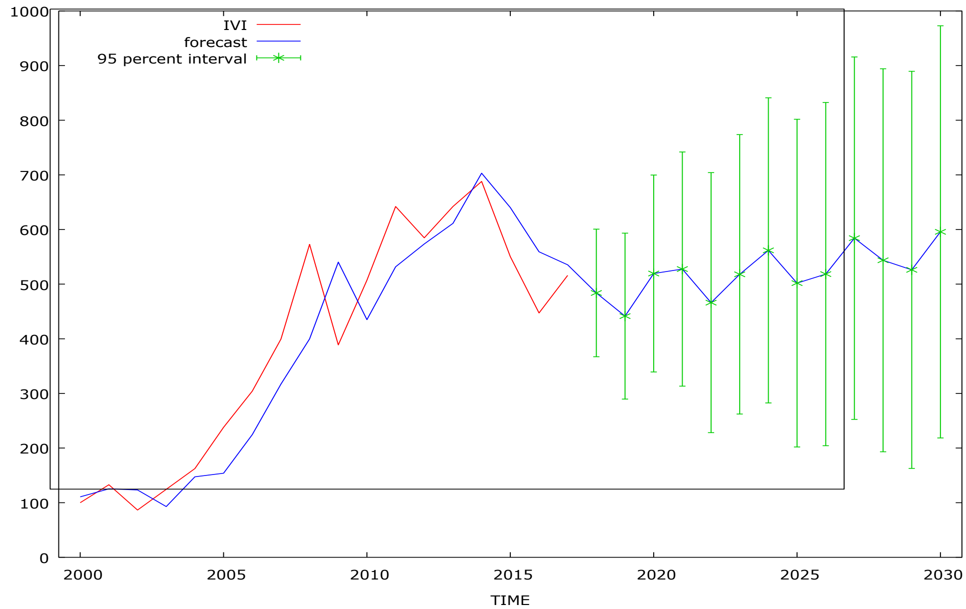

Fig. 4: Forecast model values superimposed on the original values' plot

4 Discussions

This study investigated the import value index (IVI) in Nigeria using ARIMA modelling technique. The stationary condition of import value index (IVI) series was fore determined by using time series plot as shown in Figure 1. The figure indicated enough facts that the series was not stationary at level. The augmented Dickey-Fuller (ADF) and Philips-Perron (PP) tests were used to ascertain the suitability of the order of integration (d) of the variables or the unit roots. The ADF and PP test statistics as shown in table 1a and 1b respectively revealed that all the series were stationary at first difference, that is I(1). However, Box-Jenkins methodology was adopted in this study and all steps laid down for modelling cycle were meticulously applied. It was established that ARIMA (2, 1, 2) model was the best model for the IVI data used as indicated in Figure 2 and Table 2 respectively. Furthermore, Table 3 revealed the estimation of the model where two years past values t-2 (i.e2015 and 2016) of IVI series for the AR term and the random shocks in the MA term are related to the IVI series in the current time t (2017) which was statistically significant as their p-values less than the level of significance. More so, Figure 3 indicates the residual correlogram obtained from the estimated model which showed that spikes of ACF and PACF are not statistically significant thus connote that ARIMA (2, 1,2) model is a perfect fit for the IVI series.

Lastly, forecast of thirteen years was made for IVI series from 2018-2030 as shown in Table 4, the forecast series were within the95% confidence bounds. Thus, it reveals that theforecast values were good. More so, there was fluctuation in the forecast series which shows that the series will continue fluctuating for these forecasted time period.This is an indication of present COVID-19 condition. However, this holds true as there is a significant tendency toward import of goods at the moment. In the same vein, Fig. 4 indicates that there will a continuous fluctuation for the forecast time period of the import value index.

5 Conclusion

Overall, this study models and forecasts import value index (IVI) in Nigeria for the periods 2018 to 2030 employing ARIMA modeling techniques which is also known as Box-Jenkins methodology. The findings show that rate of IVI in recent years has currently been affected by the present COVID-19 pandemic. Model identification in one of the stages of the modeling techniques shows that IVI series was not stationary at level due to the result of ADF test as well as the time plot as it was shown in Table 1a and Figure 1 respectively. The IVI series was found stationary at its first difference and it was the ARIMA (2, 1, 2) model found to be the best fit for the data based on the selection criteria employed i.e AIC and BIC. Furthermore, model estimation was the next stage where the IVI of current is shown to be related to one period lag of its own value and one period lag of error term. More so, it was shown in the correlogram of the residuals (Figure 3) that error emanated from ARIMA (2, 1, 2) model was random, normally distributed and absence of serial correlation. In the same vein, forecast of thirteen years was made which reveals fluctuation in import value index (IVI) series and therefore, indicates that the series will continue fluctuating for the forecast time period. It is therefore recommend that Nigeria should limit the level of importation of goods due to the availability of some commodities in the country as this renders the home-made goods valueless. Furthermore, level of risk exposure will be deteriorating through limiting the level of goods importation to the country.

References:

[1] Devarajan, S. (1997). Real exchange ratemisalignment in the CFA zone. *Journal of African Economies*, *6*(1), 35-53.

[2] Addison, T., Sen, K., & Tarp, F. (2020). *COVID-19-macroeconomic dimensions in the developing world* (No. wp-2020-74). World Institute for Development Economic Research (UNU-WIDER).

[3] Mckinnon, R.I. (1964) Foreign Exchange Constraints in Economic Development and Efficient Aid Allocation, *Economic Journal*, vol. 74: 388–409.

[4] Akanni, S. B., Garba, M. K., Banjoko, A. W., & Afolayan, R. B. (2020). Econometric analysis of the effects of land size on cereals production in Nigeria. *Islamic University Multidisciplinary Journal*, *7*(1), 252-258.

[5] Krugman P. R., Obstfeld, M. and Melitz, M. J. (2010). International economics: theory and policy. 9th Edition. Pearson Addison-Wesley.

[6] Ismaila U, Gana AS, Tswanya NM & Dogara D (2010). Cereals production in Nigeria: Problems, constraints and opportunities for betterment. *Am. J. Agric. Res.,* 5(12): 1341-1350

[7] Ozo-Eson, P.I. (1984) "Determinants of imports demand in Nigeria: A monetary approach".The Nigerian Journal of Economic and Social Studies, vol. 26, no. 1: 73–83.

[8] Egwaikhide, F. O. (1999), "Determinants of Imports in Nigeria: A Dynamic Specification," *Department of Economics, University of Ibadan, Nigeria. AERC Research Paper 91*

[9] Afuberoh, D. and Okoye, E. (2014) The impact of taxation on revenue generation in Nigeria: A study of Federal Capital Territory and selected states, *International Journal of Public Administration and Management Research, Vol. 2, No. 2*

[10] Okafor, R. G. (2012) Tax revenue generation and Nigerian economic development, *European Journal of Business and Management,* Vol. 4(19)

[11] World Bank Depository (2018). *Data on Food Production Index.* Retrieved from http://data.worldbank.org

[12] Box, G.E.P. Jenkins, G.M. (1976). Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco.

[13] Pankratz, A (1983). Forecasting with Univariate Box-Jenkins: Concepts and Cases, John Wiley & Sons, Inc. New York

[14] HipelK.W and McLeod A. (1994). Time Series Modelling of Water Resources and Environmental System Elsevier BV

[15] Stiglitz, J. E. and Charlton, E. (2007). Fair trade for all: How trade can promote development. (Revised Edition). Oxford: Oxford University Press.

[16] Gujarati, D.N. (2009) Basic Econometrics. Tata McGraw-Hill Education, New Delhi.

[17] Burnham K. P. and Anderson D. R (2002): Model Selection and Multi-model inference: A practical information-Theoretical Approach. Springer-Verlag, New York USA

[18] Konishi S and Kitagawa G. (2008): information criteria and statistical modelling. Springer Science and Business Media

[19] Burnham K. P. and Anderson D. R (2004): Multiinference: Understanding AIC and BIC in Model Selection. Sociological Methods and Research. 33(261) 1-45

[20] Yaffee, R.A. and McGee, M. (2000) Introduction to time series analysis and forecasting: With applications in SAS and SPSS. Academic Press,USA.

[21] Jönsson, P., & Eklundh, L. (2004). TIMESAT—a program for analyzing time-series of satellite sensor data. *Computers & geosciences*, *30*(8), 833-845

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

All Authors carried out the conceptualization of the research and data collection. Akindutire Opeyemi carried out the Statistical analysis and prepared the original Draft. Ogunlade Temitope, carried out the paper reviewing. Faweya Olanrewaju, Balogun Kayode and Okoro Joshua O had the general project administration

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding to declare