Article Number: 0227914

A Paper presented at the 1st International Conference Nigeria Statistical Society, University of Ibadan, Oyo State, Southwest Nigeria.
3rd – 6th April 2017

ACADEMIC JOURNALS
expand your knowledge

Conference Proceedings

Full Length Research Paper

Fitting of seasonal autoregressive integrated moving average to the Nigerian stock exchange trading activities

Nureni Olawale Adeboye* and Idowu Fagoyinbo

Department of Mathematics and Statistics, Federal Polytechnic, Ilaro P. M. B 50, Ilaro, Nigeria.

This research work was set to examine the activities of the Nigerian Stock Exchange using the All-Share Index monthly data published between the year 2000 and 2015. Based on the plotted ACF graph of the original series, it was observed that the series was non-stationary and also exhibited some elements of seasonality which necessitated the series to be differenced to attain stationarity as well as reducing the seasonal effect. This deseasonalised stationary series data was modeled in order to determine the stability of the parameter estimation. The plots of the ordinary and seasonal differenced series autocorrelation and partial autocorrelation functions suggested some models for selection but the Akaike and Bayesian Information Criterion was used to select the model that really provided the best fit for the series. From the family of the seasonal models generated using R-Console, seasonal ARIMA $(2, 1, 1)\times(0, 1, 1)_{12}$ model was found to be the most adequate model that really captured the dependence in the series and that also tracked the seasonal effect. The adequacy of the chosen model was subsequently checked using both the Shapiro-Wilk and Ljung-Box test approaches. The Shapiro-Wilk test for normality of residuals while Ljung-Box test for dependence in residuals of the fitted model. Method of maximum likelihood was used to determine the estimates of the parameters of the identified models and each parameter was statistically tested for significance. The model was used for a short term forecast (2016-2018).

Key words: Deseasonalised, autocorrelation function, partial autocorrelation function, stationarity, All-Share Index.

INTRODUCTION

The Autoregressive (AR), Moving Average (MA) and the mixed autoregressive moving average (ARMA) models are often very useful in modeling most time series data. However, they have the assumption of homoscedasticity (or equal variance) for the errors. A time series that exhibits some elements of seasonality can only be modeled using seasonal models such as Seasonal AR, Seasonal MA, Seasonal ARMA and Seasonal ARIMA.

*Corresponding author. E-mail: nureni.adeboye@federalpolyilaro.edu.ng.

(Madsen, 2008).

A Seasonal ARIMA model contains seasonal part and non-seasonal part in which both parts have the same structure. It may have an AR factor, an MA factor, and/or an order of differencing. In the seasonal part of the model, all of these factors operate across multiples of lags (the number of periods in a season). If the series has a long and consistent pattern, then one should consider using an order of seasonal differencing. It is statistically good not to use more than one order of seasonal differencing or more than two orders of total differencing (seasonal and non-seasonal). If the autocorrelation at the seasonal period is positive, one should consider a seasonal AR term to the model but if the autocorrelation at the seasonal period is negative, consider adding a Seasonal MA term to the model. Seasonal AR and Seasonal MA terms cannot be mixed in the same model and one should always avoid using more than one of either kind. Usually, the most commonly used seasonal model is the seasonal ARIMA (0, 1, 1) x (0, 1, 1)12 model, that is, an MA (1) x SMA (1) model with both a seasonal and non-seasonal difference. In a rare case, a seasonal ARIMA (0, 1, 2) x $(0, 1, 1)_{12}$ might be more adequate depending on the nature of the data. When a seasonal ARIMA model is fitted to log data, it is capable of tracking a multiplicative seasonal pattern (Ediger and Akar, 2007). Box et al. (1994) introduced a SARIMA model as an adaptation of an autoregressive integrated moving average (ARIMA) model, which was earlier proposed to specifically explain the variation of seasonal time series. The best forecasts in Box et al. (1994) as judged by the root mean-square error (RMSE) and other criteria were obtained with the family of periodic autoregressive models. It was found that a periodic auto-regression which was determined by choosing ρ_m as small as possible to achieve an adequate fit gave the best model forecasts. This was accomplished by initially determining ρ_m based on a plot of the periodic partial autocorrelation function and then checking the adequacy of the fitted model. This approach is thus a natural extension of that of Box and Jenkins (1976). On the other hand, a subset periodic auto regression approach was found to produce comparatively very poor forecasts. In this approach, for each period all possible auto regressions with some parameters constrained to zero and with ρ_m=12 were examined (212 possibilities) and the best model was selected (Akaike, 1974) as well as Anderson (1999). In the work of Hamilton (1994), the model fitted in the case of Domestic Inflation Rate and Exchange Rate were ARIMA (1, 1, 0) and ARIMA (1, 1, 1) respectively which were used to make inflation and exchange rates forecasts in and the validity of the model was tested. In the work of Dielman (1991), on Application of SARIMA and Exponential Smoothing in Urban Freeway Traffic Flow Prediction, the application of time series models to the single interval traffic flow forecasting problem of urban freeway was addressed. Seasonal Time Series approaches have not been used in previous forecasting research. However, time series of traffic flow data are characterized by definite periodic cycles. Seasonal ARIMA and Winters Exponential Smoothing models were developed and tested on the data sets belonging to two sites. A direct comparison with Smith (1980) findings revealed that ARIMA (2, 0, 1) (0, 1, 1) at a lag of 96 (daily period) and SARIMA (1, 0, 1)x$(0, 1, 1)_{12}$ at a lag of 96 (daily period) were the best fit models for the Telegraph Road and Wilson Bridge Sites. The single step forecasting results indicate that SARIMA out-perform neural network and historical average models as reported by Smith (1980). Noakes et al. (1985) fitted a seasonal ARIMA model using data that consisted of thirty mean monthly river flows for periods between 37 and 64 years. Various models and model selection procedures were used to calibrate a model to each data set omitting the last three years of data. The one step forecasts ahead were then compared for the last three years (36 values) of the data. Modeling of Nigerian Naira foreign exchange rates with other currencies has also engaged the attention of many researchers, a few of whom are Olowe (2009), Etuk (2012, 2013) etc. Many economic and financial time series are known to exhibit some level of seasonality in their behavior. Foreign exchange rates are among such series, their observed volatility notwithstanding. For instance, Etuk (2012) has shown that monthly Nigerian Naira-US Dollar exchange rates are seasonal with period 12 months. He fitted an (0, 1, 1) x $(1, 1, 1)_{12}$ SARIMA model to it and on this basis forecasted the 2012 values. Etuk (2013) also fitted another (0, 1, 1) x (1, 1, 1)12 SARIMA model to the monthly Naira-Euro exchange rates. A few other authors who have written extensively on the theoretical properties as well as on the practical applications of SARIMA models, highlighting their relative benefits are Priestly (1981), Farrah (2009), Philips (1994), Bollerslev (1986), Surhatono (2011), Oduro-Gyimah et al. (2012), Sami et al. (2012) and Bigovic (2012).

This study shall contribute significantly to the body of knowledge and the development of financial time series study in particular; as a well diagnosed SARIMA model that can be fit into the All-Share Index series of the Nigerian Stock Exchange in order to actualize reasonable forecast.

MATERIALS AND METHODS

The data for this research comprises the All-Share Index of the Nigerian Stock Exchange (NSE) on monthly basis for the periods of sixteen years (2000-2015). The time series SARIMA methodology adopted is subsequently discussed and analyzed.

Seasonal models

Seasonal movement is usually due to the recurring events which takes place annually or quarterly as the case may be. Seasonal models have pronounced regular ACF and PACF patterns with a periodicity equal to the order of seasonality. If the seasonality is annual, the ACF spikes are heightened at seasonal lags over and above the regular non-seasonal variation once per year. If the seasonality is quarterly, there will be prominent ACF spikes four times per year.

Seasonal Autoregressive (SAR) model

Seasonal Autoregressive model contains autoregressive parameters at seasonal lags. The time sequence plot of ACF or PACF can be used as a primary instrument for identifying seasonal autoregressive model.

Seasonal autoregressive models is given as

$$X_t = \Phi X_{t-12} + \varepsilon_t \tag{1}$$

where |Φ| < 1 and ε_t is independent of $X_{t-1}, X_{t-2}, \ldots$, it is obvious that |Φ| < 1 ensures stationarity.

Generally, a seasonal AR (P) model and a seasonal period's s is given as:

$$X_t = \Phi_1 X_{t-s} + \Phi_1 X_{t-2s} + \cdots + \Phi_{\mathrm{p}} X_{t-ps} + \varepsilon_t \tag{2}$$

It is required that ε_t is independent of X_{t-1}, X_{t-2},… and, for stationarity, that the roots of Φ(x) = 0 be greater than 1 in absolute value.

Seasonal Moving Average (SMA) model

A seasonal moving average model of order Q with seasonal period s is given as:

$$X_t = \varepsilon_t + \Theta_1 \varepsilon_{t-s} - \Theta_2 \varepsilon_{t-2s} - \Theta_3 \varepsilon_{t-3s} - \cdots - \Theta_Q \varepsilon_{t-Qs} \tag{3}$$

Seasonal Autoregressive Integrated Moving Average (SARIMA) model

An important tool in modeling non-stationary seasonal processes is the seasonal difference. The seasonal difference of period **s** for the series $[X_t]$ is denoted by $\nabla_s X_t$ and is defined as:

$$\nabla_s X_t = X_t - X_{t-s} \tag{4}$$

For a series of length n, the seasonal difference series will be of length n-s; that is, **s** data values are lost due to seasonal differencing.

In a non-stationary seasonal model, a process X_t is said to be a multiplicative seasonal ARIMA model with non-seasonal (regular) orders p, d and q, seasonal orders P, D and Q and seasonal period **s** if the differenced series:

$$W_t = \nabla^{\mathrm{d}} \nabla_{\mathrm{s}}^{\mathrm{D}} X_t \tag{5}$$

satisfies an ARMA (p,q) x (P Q)$_s$ model with seasonal period **s**. We say that $[X_t]$ is an ARIMA (p, d, q) x (P, D, Q)s model with seasonal period **s**.

In practice, many time series contains a seasonal periodic component which repeats every s observations. Box-Jenkins has generalized the ARIMA model to deal with seasonality and defines a general multiplicative seasonal ARIMA model in the form:

$$\emptyset(B)\Phi(B)(1-B)(1-B^{12})X_t = \theta(B)\Theta(\mathrm{B}^{12})\varepsilon_t \tag{6}$$

where B denotes the backward shift operator, $\emptyset, \Phi, \theta$ and Θ are polynomials for order p, P, q, and Q respectively. X_t is the observed time series and ε_t represent an unobserved white noise series, that is, a sequence of independently (uncorrelated) identically distributed random variables with zero mean and constant variance σ_ε^2.

All the identified parameters shall be estimated using the method of maximum likelihood.

Upon the fitting of the above discussed model, diagnostic checks shall be carried out to ensure normalcy using the following validity checks

i) Residual analysis
ii) Shapiro-Wilk Test of Normality
iii) The Ljung-Box Test - A portmanteau test according to Box and Pierce (1970) proposed the statistic:

$$Q = n(\hat{\rho}_1^2 + \hat{\rho}_2^2 + \ldots\ldots + \hat{\rho}_k^2) = n\sum_{k=1}^{n} \hat{\rho}_k^2 \tag{7}$$

Thus, a general “portmanteau” test would reject the ARMA (p, q) model if the observed value of Q exceeds an appropriate critical value in a Chi-Square distribution with k-p-q degrees of freedom.

The Akaike Information Criterion (AIC)

The AIC is defined as

$$\text{AIC} = -2l/T + 2K/T \tag{8}$$

Where l is the log likelihood computed as:

$$\text{l} = -\frac{T}{2}\{1 + \log(2\pi) + \log(\frac{\hat{e}'\hat{e}}{T})\} \tag{9}$$

The AIC is often used in model selection for non-nested alternative and models with smaller values of AIC are considered best.

RESULTS ANALYSIS

Checking for stationarity and determination of the appropriate SARIMA Order

Tables 1 and 2 and Figures 1 to 7 were derived showing the stationarity and determination of the appropriate SARIMA order.

Parameter estimation of identified model

Model specification for the best fit SARIMA order in Table 2 is given as:

$$(1-\phi_1 B - \phi_2 B^2)(1-B^{12})(1-B)X_t = (1-\theta_1 B)(1-\Theta_1 \mathrm{B}^{12})\varepsilon_t \tag{10}$$

Substituting the values of the parameters, we have

$$(1 - 0.4995B + 0.2191B^2)(1-B^{12})(1-B)X_t = (1-0.7342B)(1+0.7395)\varepsilon_t \tag{11}$$

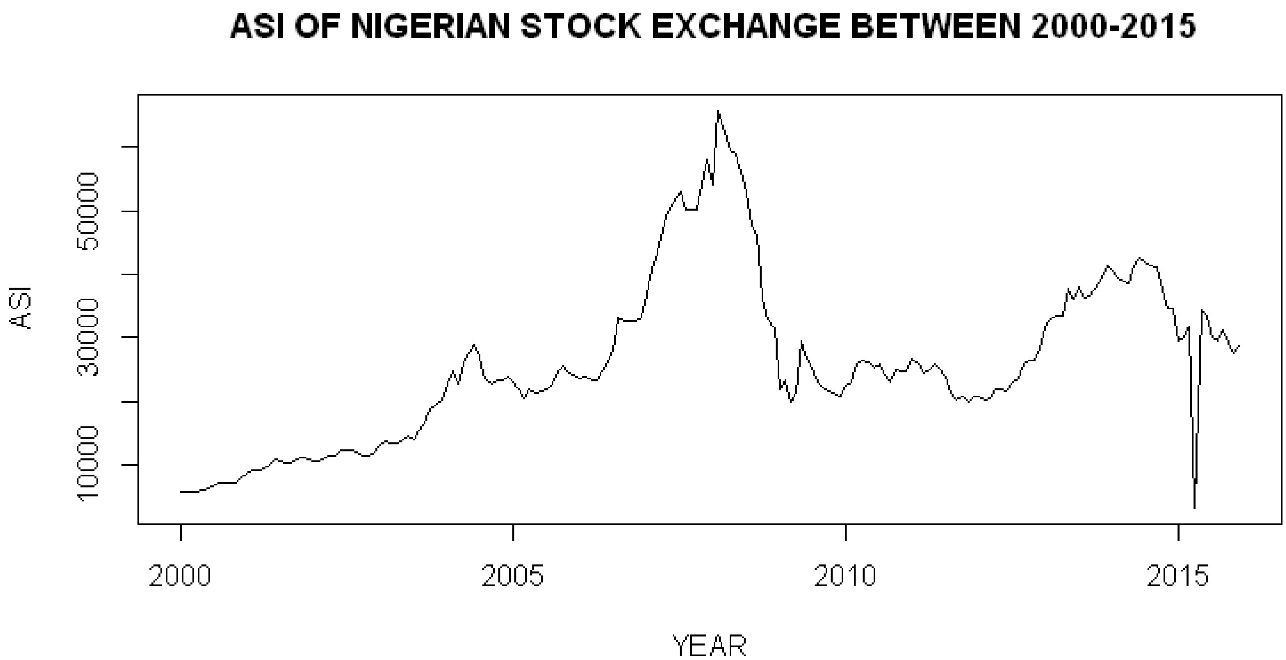

Figure 1. Time Series Plot of NSE ASI Series.

Diagnostic check on SARIMA (2,1,1) × $(0,1,1)_{12}$ model

The Shapiro-Wilk test of normality gives a test statistic of W = 0.9235, with a P-value of 0.0000 which indicates that the residuals are normally distributed at 1, 5 and 10% significance levels.

The Ljung-Box test statistic examines the null hypothesis of independence in the residuals of the All-Share Index series with a Chi-squared value of 0.01927 with a P-value of 0.8896 which lead to the acceptance of null hypothesis that all the autocorrelation functions are zero.

DISCUSSION

From Figure 1, it was observed that the pattern of the graph indicates that the series is non-stationary. There were both upward and downward trend as well as seasonal variation which also show that the series is stochastic in nature. The autocorrelation plot in Figure 2 indicates significant spikes from lag 1 to 24 as well as seasonal variation, a downward trend for subsequent lags and off to zero at lag 40 which also indicates an element of non-stationary. In view of this, stationarity was therefore achieved by applying differencing as evidenced in Figures 4, 5 and 6 for the series acceptability ascertained by performing Augmented Dickey Fuller Test.

Since the Dickey-Fuller test statistic is -4.7637 and the P-value is 0.01 as obtained in Table 1, we therefore fail to accept H_0 and hence conclude that the alternative hypothesis is true, that is, the series is stationary in its mean and variance. This test brings to reality the fitting of a suitable SARIMA model.

Having made the series stationary, decision was made on reasonable values of the orders of the Autoregressive (AR (ϕ)), Seasonal Autoregressive (SAR (Φ)), ordinary differencing, Moving Average (MA(θ)) and Seasonal Moving Average (SMA(Θ)). With a few iterations on this model-building strategy, we arrive at the parameters estimation presented in Table 2, hence the SARIMA model of order (2, 1, 1) x $(0, 1, 1)_{12}$ presented as Equation 10. The model parameters have been parsimoniously fitted, the standard errors and log-likelihood have improved while the model has a smaller AIC and variance which confirms that it captures the dependence in the series more than any other iterative models suggested by the sample ACF and PACF of first order and seasonal differencing.

Forecasting

Based on the set objectives of this research, forecasting was done using the fitted SARIMA model and the forecast values exhibited downward trend for All-Share Index in the Nigeria Stock Exchange in the month of January, March, June to August and October every year and a pick up on the month of February, April, May and December on yearly basis. However, the ASI has been confirmed to have seasonal effect over time since there is an upward trend in the stock market every month of December for the forecasted years of 2016-2018 respectively (Figure 8).

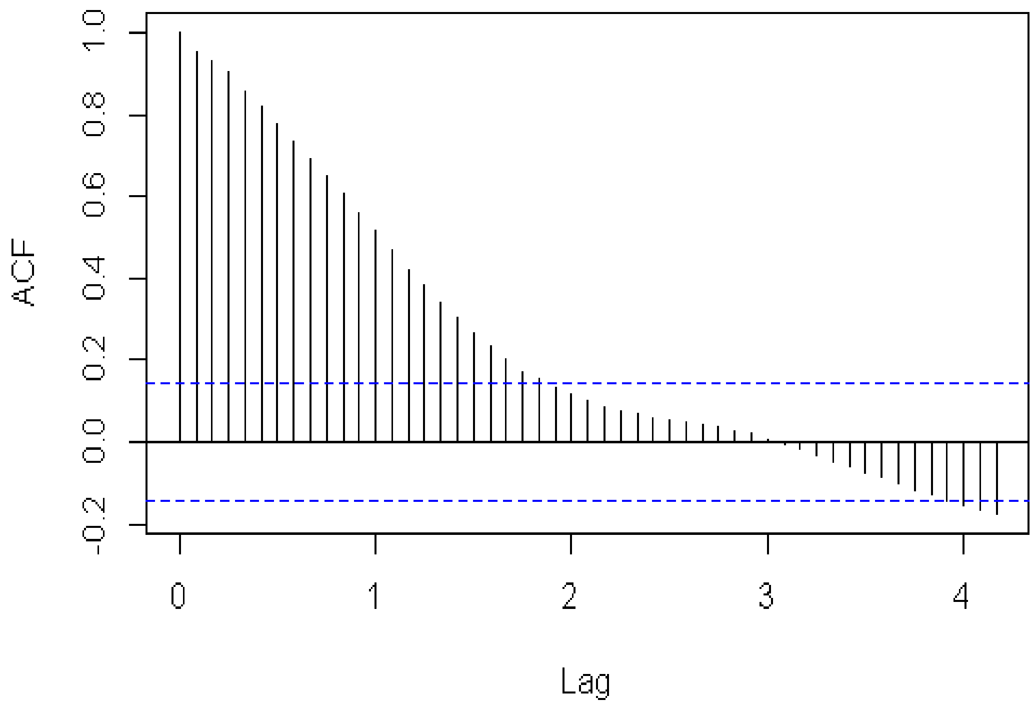

Figure 2. Sample ACF of of NSE ASI Series.

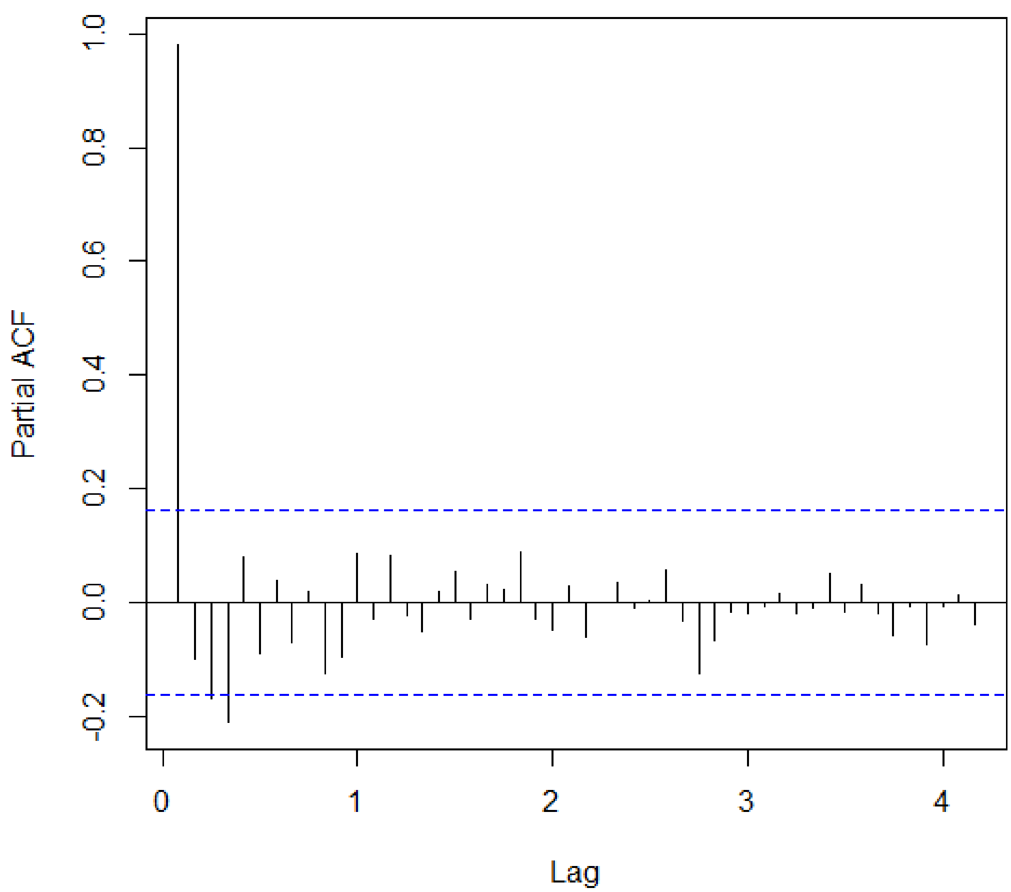

Figure 3. Sample PACF of NSE ASI Series.

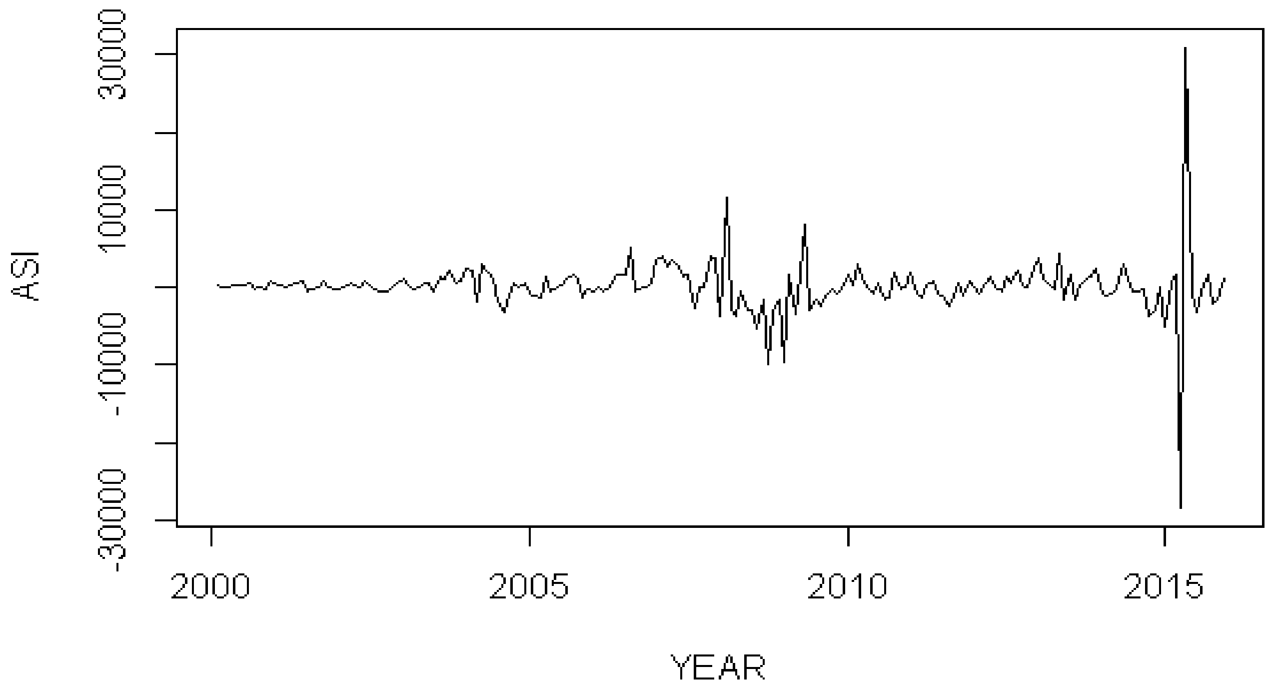

Figure 4. Plot of the first and seasonal differences of NSE ASI Series.

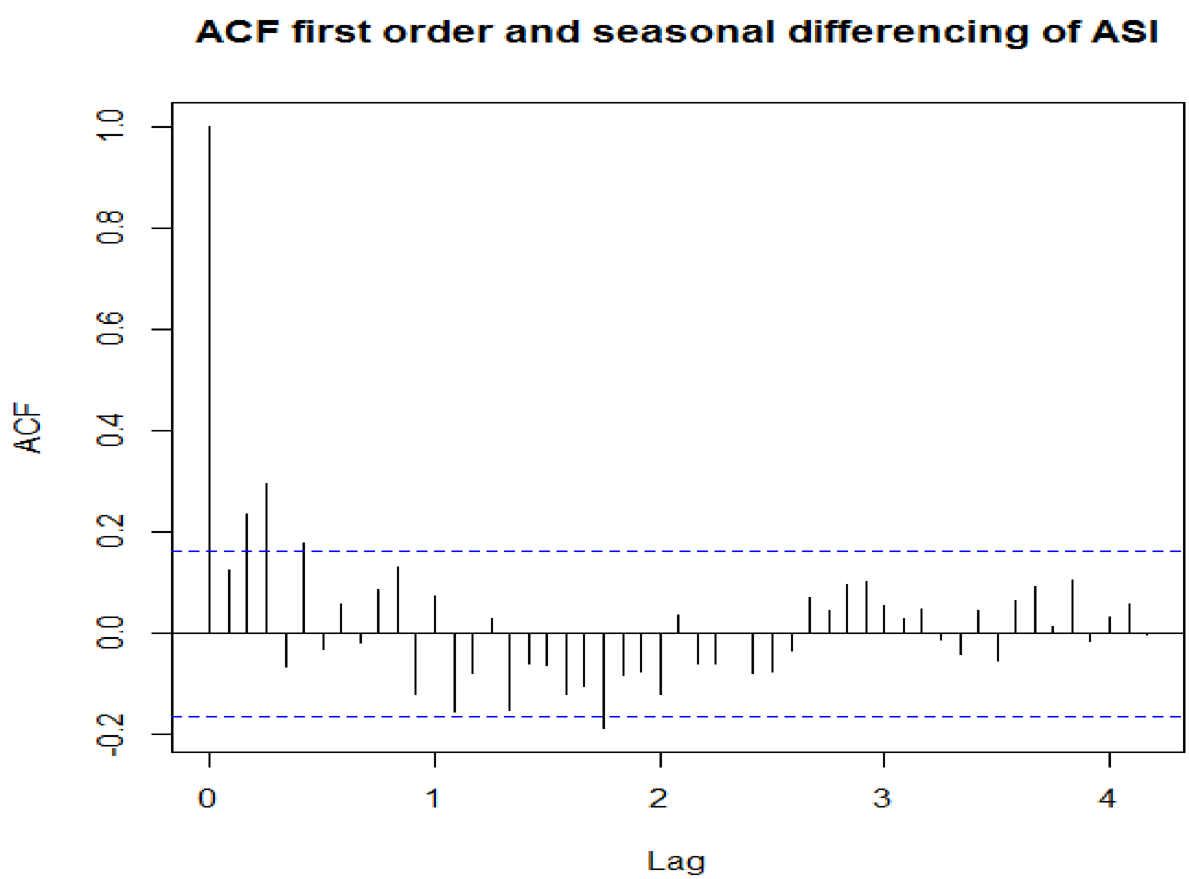

Figure 5. Sample ACF of First Order and Seasonal Differencing of ASI.

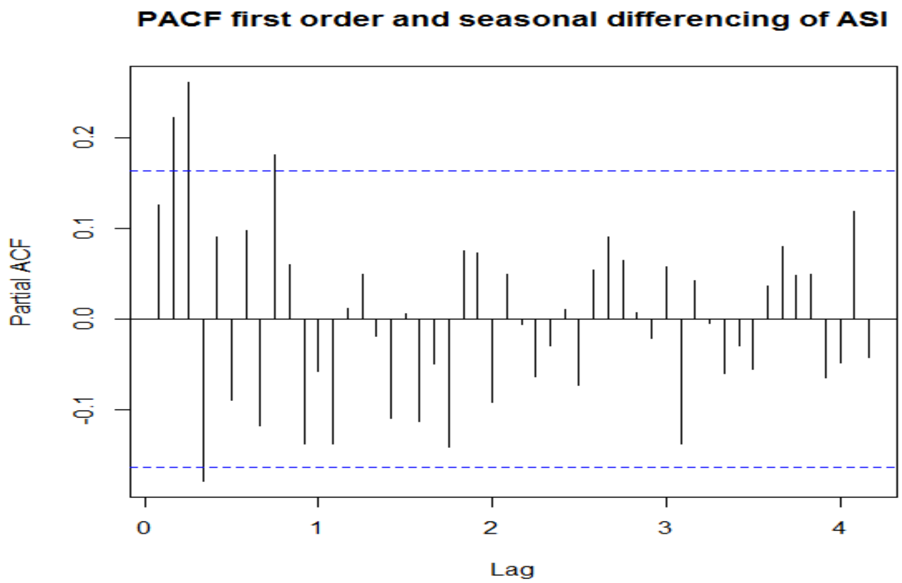

Figure 6. Sample PACF of First Order and Seasonal Differencing of ASI.

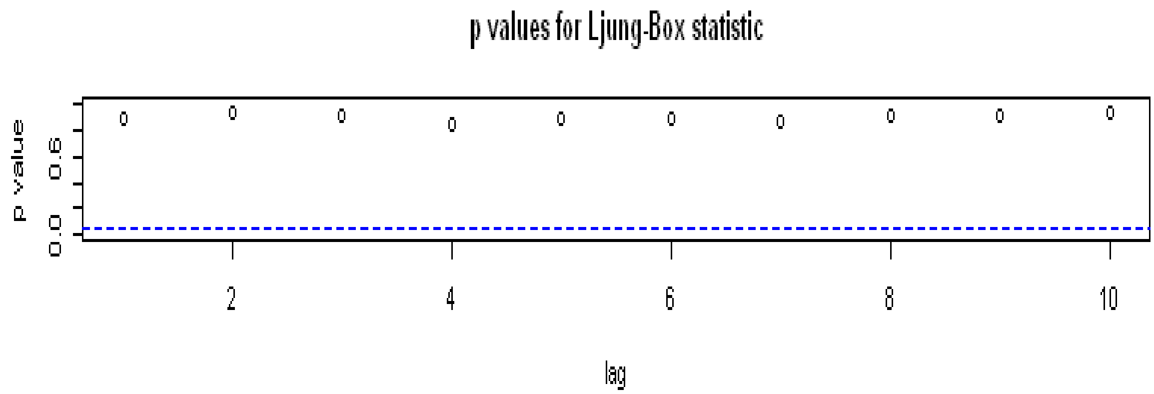

Figure 7. Adequacy Check For Sarima (2,1,1) × $(0,1,1)_{12}$ using ACF of Residuals.

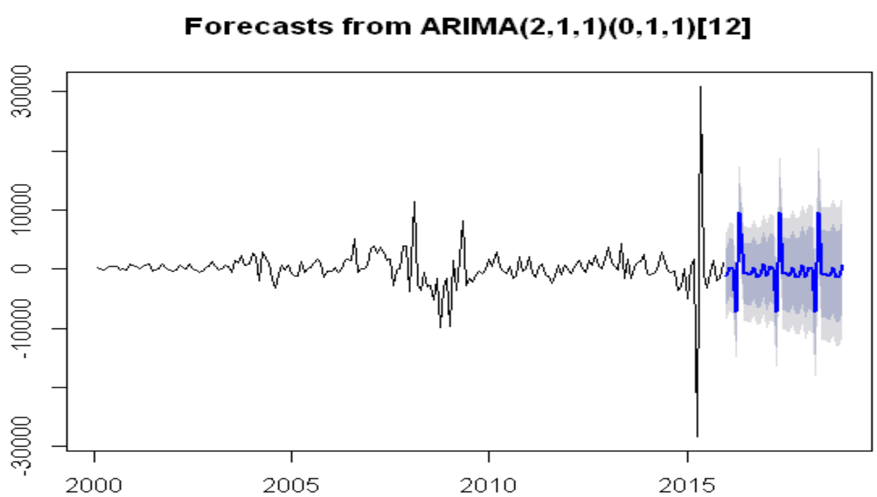

Figure 8. Forecasts from SARIMA (2,1,1) x $(0,1,1)_{12}$.

Table 1. Augmented Dickey-Fuller Test for Series stationarity.

Dickey-Fuller test statistic	Lag order	P-value
-4.7637	5	0.01

Source: R-studio output.

Table 2. Possible SARIMA models for ASI Series.

Model	Coefficient	Estimate	Standard Error	|Z|-value
SARIMA (3, 1, 3) × (1, 1, 0)$_{12}$	ϕ_1	0.0444	0.3118	0.1424
	ϕ_2	0.2176	0.2337	0.9311
	ϕ_3	0.2975	0.1061	2.8040*
	θ_1	-1.2634	0.3265	3.8879*
	θ_2	0.2095	0.5786	0.3621
	θ_3	0.0539	0.2886	0.1868
	Φ_1	-0.4574	0.0863	5.3001*
	Log-likelihood = -1722		**AIC =** 3459.99	σ^2 **=** 14137733
SARIMA (1, 1, 1) × (1, 1, 0)$_{12}$	ϕ_1	-0.3252	0.0827	3.9323*
	θ_1	-0.8081	0.0617	13.0972*
	Φ_1	0.4241	0.0875	4.8469*
	Log-likelihood = -1730.16		**AIC =** 3468.32	σ^2 **=** 15858073
SARIMA (1, 1, 2) × (0, 1, 1)$_{12}$	ϕ_1	-0.1694	0.1605	1.0555
	θ_1	-1.0572	0.1488	7.1048*
	θ_2	0.2131	0.1497	2.1460*
	Θ_1	-0.7509	0.0993	7.5619*
	Log-likelihood = -1723.1		**AIC =** 3456.21	σ^2 **=** 14018819
SARIMA (2, 1, 1) × (0, 1, 1)$_{12}$ BEST FIT	ϕ_1	0.4995	0.1076	4.6422*
	ϕ_2	-0.2191	0.1009	2.1715*
	θ_1	0.7342	0.0985	7.4538*
	Θ_1	-0.7395	0.0999	7.4024*
	Log-likelihood = -1721.78		**AIC =** 3453.57	σ^2 **=** 13853805

Source: R-studio output.

Conclusion

Having used necessary and suitable methods in line with the set goals of this research, there is no doubt that the main purpose has been fully realized.

Therefore, based on the results obtained by the empirical analysis of the data collected, the following conclusions are therefore arrived at:

i) SARIMA model (2, 1, 1) x (0, 1, 1)$_{12}$ is the most appropriate fit for the All-Shares index of Nigerian Stock Exchange.
ii) The stock market enjoyed the most economic boost in the month of February, April, May and December respectively based on the forecast.
iii) The period of January, March, June, August and October exhibited a downward trend for the stock market.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

Akaike H (1974). A New Look at the Statistical Model Identification", IEEE Transactions on Automatic Control. 19:716-723.

Anderson DD (1999).The Statistical Analysis of time Series. New York: John Wiley& Sons.

Bigovic M (2012). Demand Forecasting within Montenegrin Tourism using Box-Jenkins methodology for Seasonal ARIMA Models," Tourism and Hospitality Management 18(1):1-18.

Bollerslev T (1986). Generalized autoregressive conditional

Heteroskedasticity. Journal of Econometrics 31:307-327.
Box GE, Jenkins GM (1976). Time Series Analysis, revised edition. San Francisco.
Box GEP, Jenkins GM, Reinsel GC (1994). Time Series Analysis, Forecasting and Control, 2nd ed.. New York Prentice-Hall.
Box GEP, Pierce DA (1970). Distribution of the Autocorrelations in Autoregressive Moving Average Time Series Models. Journal of American Statistical Association 65:1509-1526.
Dielman TE (1991). Applied Regression Analysis for Business and Economics, PWSKent, Boston.
Ediger VS, Akar S (2007). ARIMA Forecasting of Primary Energy Demand by Fuel in Turkey, Energy Policy 35(3).
Etuk EH (2012). Forecasting Nigerian naira-us dollar exchange rates by a seasonal Arima model. American Journal of Scientific Research 59:71-78.
Etuk EH (2013). Forecasting Nigerian naira-euro exchange rates by a seasonal Arima model. American Journal of Scientific Research 60:60-67.
Farrah K (2009). Times Series Modelling on Age Specific Breast Cancer Mortality Rate. Journal of Econometric 105:123-127.
Hamilton JD. (1994). Time Series Analysis, Princeton University Press, Princeton, N. Jersey.
Madsen H (2008).Time Series Analysis, Chapman & Hall, London.
Noakes DJ, McLeod AI, Hipel .KW (1985). Time Series Analysis: Forecasting and Control, Washington D.C.
Oduro-Gyimah FK Harris E, Darkwah KF (2012). Sarima Time Series Model Application to Microwave Transimission of Yegi-Salaga (Ghana) Line-of-Sight Link", International Journal of Applied Science and Technology 2(9):40-51.
Olowe RA (2009). Modelling Naira/Dollar Exchange Rate Volatility: Application of Garch and Asymmetric Models", International Review of Business Research Papers 5(3):377-398.
Philips F (1994). Time Series Analysis, Princeton University Press, Princeton, New Jersey.
Priestly M. (1981). Spectral Analysis and Time Series. New York, Academic Press.
Sami M, Waseem A, Jafri YZ, Shah SH, Akbar, S, Siddiqui MA, Murtaza G (2012). Prediction of the rate of dust fall in Quetta city, Pakistan using seasonal ARIMA (SARIMA) modeling", International Journal of Physical Sciences 7(10):1713-1725.
Smith J (1980). Seasonal ARIMA Modelling of Telegraph Road and Wilson Bridges Site", Biometrica 8:101-108.
Surhatono K (2011). Time Series Forecasting by using Autoregressive Integrated Moving Average: Subset, Multiplicative or Additive Model. Journal of Mathematics and Statistic 7(1):20-27.