Journal of the Nigerian Association of Mathematical Physics
Volume35,(May, 2016), pp185 – 192
© J. of NAMP

On Modelling Daily Share Price Data in Nigeria using ARIMA Modelling Approach

***Osemwenkhae J. E.*[1] *andEguasa E. B.*[2]**

Department of Mathematics, University of Benin, Benin City.

Abstract

The aim of this work is to develop a time series model for the Nigerian Stock Exchange using the daily stock closing price of Okomu Oil Palm Company Plc from Jan. 2010 to Dec. 2014.
We applied the Box-Jenkins autoregressive integrated moving average (ARIMA) modelling methodology. The time plot and the Augmented Dickey-Fuller (ADF) test showed that the time series data was non-stationary but was made stationary by log transformation and differencing once.
Our results revealed that the ARIMA (0,1,1) was best in modelling the data as indicted by the Akaike Information Criterion (AIC). Hence, the ARIMA (0,1,1) is found as the best model for forecasting the Okomu Oil Palm Company share price data.

Keywords: Autoregressive integrated moving average, Augmented Dickey-Fuller, Akaike Information Criterion, Okomu oil palm

1.0 Introduction

Shares represent part of a company's stock, any of the equal, usually small, parts into which a company's capital stock is divided. A share of a company's stock is a share in the ownership of such a company.
In 1960, the establishment of the Nigerian Stock Exchange (NSE) (formerly known as Lagos Stock Exchange) stimulated private capital investment for growth and development. It is generally believed that investment that promotes economic growth and development involves long term funding far longer than the duration for which regular savers are willing to commit their funds [1].
Dabbling in shares and stocks can be risky but lucrative for the bold, beautiful, and watchful. It can be a gradual ride or a rollercoaster [2]. Financial superpowers have emerged from investing in the stock market.Osamwonyi [2] confirmed that shares transacting is rewarding and indeed particularly so only when the investor remains awake.
However, the seeming risks associated with investing in Nigeria market are high [1]. This in recent times may not be unconnected to the global financial meltdown that has plagued many economies of the World. In the "world" of uncertainties, stock market investors need to keep wide-awake. This entails following up trends, financial reports, market analyses and stock forecasts. Stock market price prediction is viewed as one of the most challenging task of financial time series prediction [3].
In this study, we carry out the technical analysis modelling of stock price data of a company in the Nigerian stock market with emphasis on the agricultural sector. This is to increase investment awareness in the agricultural sector as a way of diversifying the Nigeria economy from the oil sector.
This work seeks to provide a scientific study of the underlying patterns of the Nigerian Stock Exchange using Okomu Oil Palm Company Plc stock price data.The aim is principally to develop a time series model that could be used for forecasting the daily stock closing price of Okomu Oil Palm Company in the short term.

2.0 Literature Review

Stock market prices have been the most analyzed economic data during the past forty years [4]. The basic question is – are (real) price changes forecastable? Wojceiech [5] noted that a great deal of attention has been focused on the idea of predicting stock prices and price information within the last two decades.

Corres. author:Osemwenkhae J. E., E-mail:joseph.osemwenkhae@uniben.edu, Tel.: +2348023713354 & 8053902957(EEB)

Adebiyi et al. [6] presented extensive process of building stock price predictive model using the ARIMA approach. They developed stock price predictive model for the New York Stock Exchange and the Nigeria Stock Exchange using published stock data obtained from both stock markets for Nokia stock index and Zenith Bank stock prices respectively. The experimental results they obtained from the best ARIMA model demonstrated the potential of ARIMA modelsin predicting stock prices satisfactorily on short-term basis. They concluded that the ARIMA models could compete reasonably well with other forecasting techniques in short-term prediction. Their work showed that the ARIMA (2,1,0) model best modeled the Nokia stock index while ARIMA (1,0,1) was selected as the best model for Zenith Bank stock index.
Etuk et al. [7] modeled monthly stock price data using the Box-Jenkins approach. Their work revealed that the ARIMA (2,1,3) model based on the use of automatic model selection criteria: AIC, SIC and R^2 was most appropriate in predicting Nigerian stock prices.
Hadavandi et al.[8] pointed out that stock market prediction is regarded as a challenging task in financial time series forecasting. They attributed this to the uncertainties involved in the movement of the market. They noted that "the central idea to successful stock market prediction is achieving best results using minimum required input data and the least complex stock market model.
Paul et al.[9] examined empirically the best ARIMA model for forecasting average daily share price indices of Square Pharmaceuticals Limited (SPL), Bangladesh. Their work showed that the daily share price indices of SPL data series were non-stationary. They however achieved stationarity by taking the first difference of the logarithmic values of the SPL data. The best ARIMA model was selected using the AIC, AICc, SIC, and RMSE criteria. Paul et al. [9] found ARIMA (2,1,2) to be the best model for forecasting the SPL data series.

3.0 Methodology

3.1 Stationarity

A time series $\{X_t, t = 0,\pm1,\pm2...\}$ is said to be stationary if it has statistical properties similar to those of the "time-shifted" series $\{X_{t+h}, t = 0,\pm1,\pm2...\}$, for each integer (interval time),*h*. In a univariate time series context, we consider that the pair of values (X_{t-h}, X_t) constitutes the predictor and response respectively. The condition that these values come from a well-defined population as *t* varies is satisfied by the assumption of stationarity [10].
Specifically, stationarity could either be weak stationarity or strict stationarity. The underlying difference is that for weak stationarity the following holds:

$$i.\mathrm{E}(X_t) = \mu_X$$
$$ii.Var(X_t) = \sigma_X^2$$
$$iii.Cov(X_{t-h}, X_t) = \gamma_{X,h} \forall t, h \in Z \quad (1)$$

while for strict stationarity; the joint distribution of $(X_{t1}, X_{t2},...X_{tk})$ is the same as that of $(X_{t1+h}, X_{t2+h},...X_{tk+h})$ for every value of h.

This definition says that the stochastic behaviour of the process does not change through time. If { X_t } is stationary, then
$\mu(t) = \mu(0)$ and

$$\gamma(s,t) = \gamma(s-t,0). \quad (2)$$

3.1.1 Differencing and Logarithmic (log) Transformation

In order to convert non-stationary series to stationary, differencing method can be used in which the series is lagged 1 or more steps and subtracted from original series. If x_t denotes the value of the time series X at period t, then the first difference of X at period t is equal to $x_t - x_{t-1}$.
Time series data can first be transformed by taking the log of the series and then the differencing is performed. This is because some time series especially financial series are exposed to exponential growth, and thus log transformation can smooth out (linearize) the series and differencing will help stabilize the variance of the series.

3.2 Autoregressive Integrated Moving Average (ARIMA) Models

A series { x_t } is integrated of order d, denoted as I(d), if the dth difference of { x_t } is white noise $\{w_t\}$; i.e. $\nabla^d x_t = w_t$ [11].

A time series { X_t } follows an ARIMA (p,d,q) process if the dth difference of the { X_t } series are an ARMA (p,q) process. If we introduce

$$y_t = (1-B)^d x_t , \quad (3)$$

then

$$\theta_p(B)y_t = \phi_q(B)w_t . \quad (4)$$

Substituting for y_t given in equation (3), we obtain a more succinct form for an ARIMA (p,d,q) process as

$$\theta_p(B)(1-B)^d x_t = \phi_q(B)w_t \quad (5)$$

where θ_p and ϕ_q are polynomials of orders p and q, respectively, x_t represent a non-stationary time series at time t, w_t is a white noise (zero mean and constant variance), d is the order of differencing[11].

3.3 ARIMA Modelling Procedure.

Having defined the ARIMA model, the next issue is how to select an appropriate ARIMA model that can produce accurate forecast hinged on a description of historical pattern in the series and how to determine the optimal model orders.
Box and Jenkins [12] developed the procedure for building an ARIMA model, which best fit to a given time series and also satisfy the parsimony principle. The Box-Jenkins methodology does not assume any particular pattern in the historical data of the series for forecasting. Rather, it uses a three step iterative approach of; model identification, parameter estimation and diagnostic checking, to determine the best parsimonious model from a general class of ARIMA models[13]. This three-step process is repeated several times until a satisfactory model is finally selected. Then this model can be used for forecasting future values of the time series.

3.3.1 Model Identification

The foremost step in the process of ARIMA modelling is to check for the stationarity of the series, since estimation procedures are available only for stationary series. There are two types of stationarity; stationarity in 'mean' and stationarity in 'variance'. Examining the graph of the data and structure of autocorrelation and partial autocorrelation coefficients may provide clues for the presence of stationarity. The first step is to fit a first order autoregressive model for the raw data and test whether the coefficient "θ_1" is less than one.
If the model is found to be non-stationary, stationarity could be achieved by differencing the series. Stationarity in variance could be achieved by some modes of transformation, say, log-transformation. Augmented Dickey-Fuller (ADF) test is used to determine whether the time series after differencing is stationary or not.
The next step in the identification process is to find the initial values for the order of the model, p and q. they could be obtained by searching for significant autocorrelation and partial autocorrelation coefficients. For example, if second order autocorrelation coefficient is significant, then an AR (2), or MA (2) or ARMA (2,2) model could be tried to start with. It is noteworthy however that up to order 2 for p, d, or q are sufficient for developing a good model in practice.

3.3.2 Diagnostic Checking

Various models can be obtained for different combinations of AR and MA individually and collectively. The best model is obtained with the following diagnostics.

a. Minimum Akaike Information Criterion (AIC), Akaike Information Criterion Corrected (AICc) and Schwarz Information Criterion (SIC)

b. Plot of residual ACF

Once the appropriate ARIMA models have been fitted, one can examine the goodness of fit by means of plotting the ACF of the residuals of the fitted model. If most of the sample autocorrelation coefficients of the residuals are within the limits $\pm 1.96/\sqrt{N}$ where N is the number of observations upon which the model is based, then the residuals are white noise indicating that the model is a good fit.

4.0 Modelling and Analysis of Data

In this section, identification and fitting of tentative ARIMA models to the data based on the autocorrelation and partial autocorrelation plots of the stationary series are considered. Appropriate models for our data were selected following the various diagnostic checking tests in literature.

Furthermore, we carry out some tests of non-linearity on the residual of the best ARIMA model to our data. These tests were conducted in order to reveal (if present) characteristics of non-linearity in our data for possible modelling.

4.1 Data Presentation

The closing stock prices of Okomu Oil Palm Plc of the Nigerian Stock Exchange spanning 2010 to 2014 were retrieved from the website of [14], a registered stock broking firm in Nigeria.

4.2 Analysis of Okomu Oil Stock Series

4.2.1 Time Plot of Okomu Stock Series

The first step in modelling any time series data is to plot and examine the behaviour of the series such as stationarity. The stock price data of Okomu Oil is plotted fig. 1.

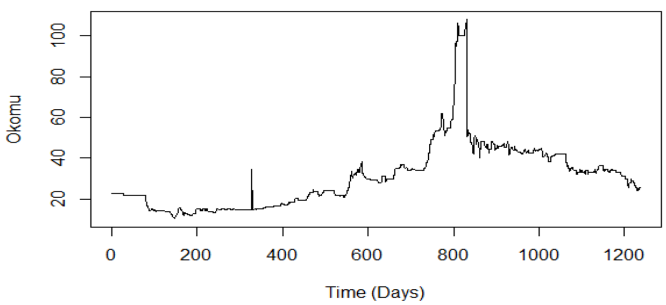

Figure 1: Time series plot of Okomu Oil Plc (2010-2014)

The plot in Figure 1 shows that Okomu share series is non-stationary, however, further test will be conducted to ascertain this. The graph also exhibits exponential increase at different points (levels) in the series

4.2.2 Test of Stationarity: Augmented Dickey Fuller Unit Root Test

Dickey-Fuller = -2.7403, Lag order = 10, p-value = 0.2649

The test shows that the null hypothesis of the presence of unit root in the series is not rejected since p-value > 0.05 (level of significance). Hence, the series is non-stationary.

4.2.3 Log Transformation and Differencing of the Series

In financial time series analysis, the series is often transformed by taking its log and then differencing of the series is performed. This is because financial series are usually exposed to exponential growth.

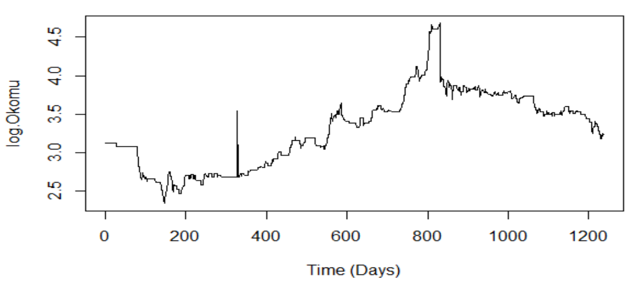

Figure 2: Graph of log transformed Okomu share price series.

Having transformed the data, we carry-out the differencing of the data. If after the first differences, the series is still non-stationary, we carry-out the second differencing of the series.

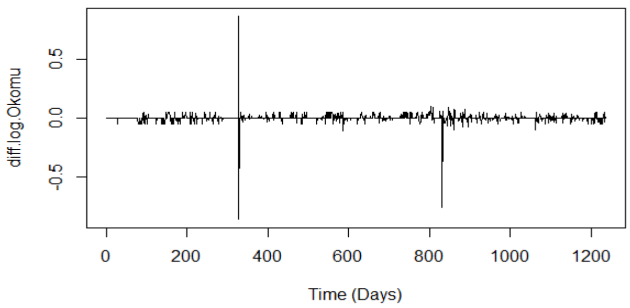

Figure 3: Plot of first differences of the log transformed Okomu series

Figure 3 indicates that the first differences of the series are stationary. The augmented Dickey-Fuller unit root test is performed on the series to ascertain the stationarity of the series. The result is presented below.

Augmented Dickey-Fuller Test Statistic

Dickey-Fuller = -10.4302, Lag order = 10, p-value = 0.01

The augmented Dickey-Fuller test confirms that stationarity is achieved (p-value <0.05) after the first differencing of the log transformed series is performed.

4.2.4 Graphs of the Autocorrelation Function (ACF) And Partial Autocorrelation Function (PACF)

After achieving stationarity, we examine the autocorrelation function and partial autocorrelation function of the series to determine the appropriate orders of the autoregressive and moving average terms.

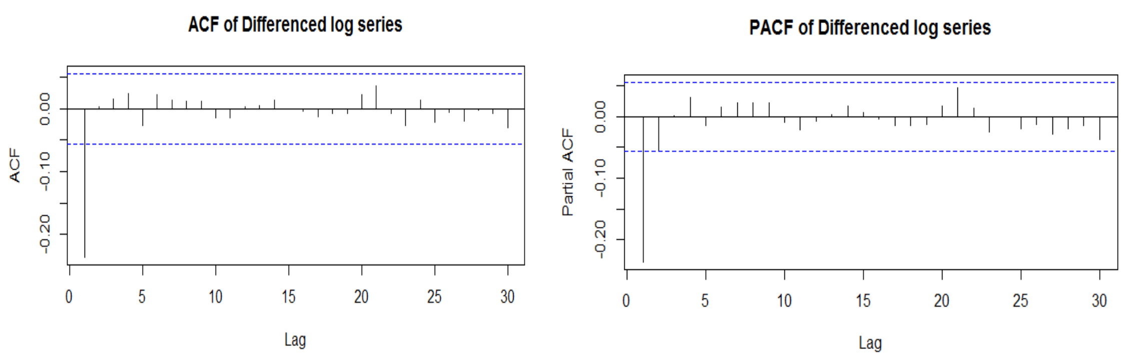

Figure 4: ACF and PACF plots of the first differenced log series of Okomu Oil.

Comparing the autocorrelations with their respective limits, the only significant autocorrelation is at lag 1, indicating an MA(1) (moving average component with lag = 1) behaviour. Similarly, only lags 1 and 2 partial autocorrelations are significant, indicating an AR (1) (autoregressive component with lag = 1) or AR(2) (autoregressive component with lag =2) behaviour. The following models are thus suggested;

i. ARIMA (1,1,0)
ii. ARIMA (1,1,1)
iii. ARIMA (0,1,1)
iv. ARIMA (2,1,0)
v. ARIMA (2,1,1)

To select the best ARIMA model, each model is assessed based on the value of the Akaike information criterion (AIC). The model with the smallest AIC is the most appropriate model.

Table 1: AIC Values for ARIMA Models

MODEL	AIC
ARIMA (1,1,0)	-4153.55
ARIMA (1,1,1)	-4155.12
ARIMA (0,1,1)	-4156.92
ARIMA (2,1,0)	-4155.53
ARIMA (2,1,1)	-4153.53

From Table 1, although the values of the AIC appear close to each other, the most adequate ARIMA model for the OKOMU series is the ARIMA (0,1,1) since this model has the least AIC value. Thus the model for stock price data of Okomu Oil PLC in terms of the differenced series X_t with the moving average lag term equal 1, is expressed as:

$$x_t = \alpha + w_t - \theta_1 w_{t-1} \quad (6)$$

In terms of the observed series, the model in (6) becomes,

$$y_t - y_{t-1} = \alpha + w_t - \theta_1 w_{t-1} \quad (7)$$

which implies

$$y_t = \alpha + y_{t-1} + w_t - \theta_1 w_{t-1} \quad (8)$$

However, equation (8) is re-written as

$$y_t = \alpha + y_{t-1} + w_t - \theta_1 w_{t-1} \quad (9)$$

for a differenced series since the mean is zero after differencing.
The point estimate of each parameter of ARIMA (0,1,1) are as follows:
arima(x = okomulog, order = c(0, 1, 1))
Coefficients:
ma1 intercept
-0.2456 1e-04
s.e. 0.0270 1e-03
sigma^2 estimated as 0.002021: log likelihood = 2080.46, AIC = -4156.92

that is , $\hat{\theta}_1$ =-0.2456 corresponding to the ma1.
The fitted ARIMA (0,1,1) model for the Okomu Stock Price data from 2010-2014 is given by;

$$\hat{y}_t = y_{t-1} + w_t + 0.2456 w_{t-1} \quad (10)$$

Having derived an ARIMA model for our data, we plot the ACF and PACF. These plots would reveal if there exist any significant lags in the model.

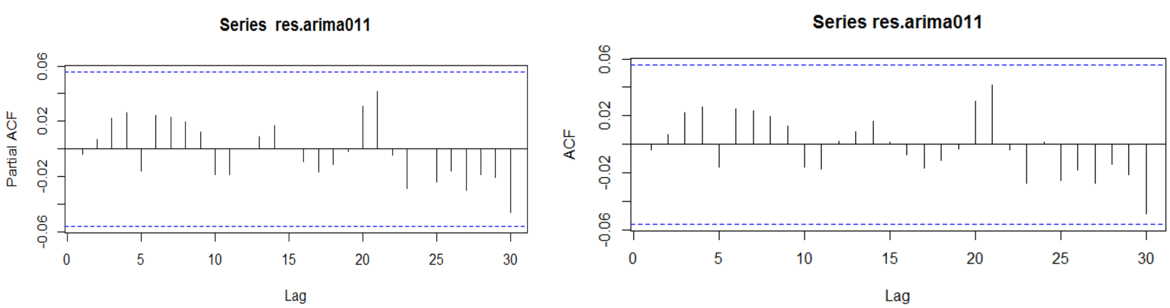

Figure 5: ACF and PACF of ARIMA (0,1,1) Residuals.

The ACF and PACF do not have any significant lag since none of the spikes extends beyond the limits, indicating that ARIMA (0,1,1) model is a good model to represent the series.

4.3 Discussion of Results

In this work, first we study if the share prices data of Okomu Oil Palm Plc was stationary by plotting the series against time and carrying out the augmented Dickey Fuller test. The results as presented in Figure 1 and Section 4.2.2 respectively indicate that the share prices data was non-stationary. In Section 4.2.3, log transformation and differencing of the data in

order to achieve stationarity were done. Stationarity of the differenced series was confirmed by the augmented Dickey Fuller test (p-value < 0.05) presented in Section 4.2.3.
In Section 4.2.4, the graphs of the autocorrelation function (ACF) and partial autocorrelation function (PACF) were carried out. Significant spikes were observed at lag 1 in the ACF and lags 1 and 2 in the PACF. This led to fitting of five (5) ARIMA models in section 4.3. Table 1 showed the different AIC values for the different ARIMA models with ARIMA (0,1,1) having the smallest AIC value, thus the most appropriate ARIMA model.

4.4 Summary and Conclusion

In this study, various ARIMA models for the stock price data of Okomu Oil Palm Company Plc. were examined. A discussion of the ARIMA modelling procedure was presented and this procedure was practically applied to the stock data.
After fitting various ARIMA models, the Akaike information criterion revealed that the ARIMA (0,1,1) was most adequate in modeling the Okomu share price series. Thus, it is suitable for short-term forecasting for investment decision by stock market investors and policy implication for stock market regulators.
The examination of the ACF and PACF of the residuals of the ARIMA (0,1,1) model were done. These plots (Figure 5) showed no significant lags signifying that the ARIMA model is adequate.
In this work, the nature of the variance of the residuals from the fitted ARIMA (0,1,1) model is not examined. This puts a limit on the scope of the study. Further research is suggested in this area.

5.0 References

[1] Osisanwo, B. G., &Atanda, A. A. (2012) *Determinants of Stock Market Returns in Nigeria: A Time Series Analysis* African Journal of Scientific Research vol. 9,No. 1, pp. 478-496

[2] Osamwonyi, I. O. (2013) *The Nigerian Financial Market and a Tale of Two Foolish Acts* Inaugural Lecture Series 133 University of Benin, Benin City.

[3] Kazem, A., Sharifi, E., Hussain, F. K., Saberi, M., &Hussain, O. K. (2013) *Support vector Regression with Chaos Firefly Algorithm for Stock Market Price Forecasting* Applied Soft Computing 13, pp. 947-958

[4] Granger, C. W. J. (1992) *Forecasting Stock Market Prices: Lessons for forecasters* International Journal of Forecasting 8, pp. 3-13.

[5] Wojciech, G. (2007). "Neural Network Predictions of Stock Price Fluctuations".

[6] Adebiyi, A. A., Adewumi, A. O. & Ayo, C. K. (2014) *Stock Price Prediction using the ARIMA Model*Proceedings of the 2014 UKSim-AMSS 16th InternationalConference on Computer Modelling and Simulation, pp. 1-7

[7] Etuk, E. H., Uchendu, B. &Udo, E. O. (2012) Box-Jenkins Modelling of Nigerian Stock Prices Data Greener Journal of Science Engineering and Technological Research vol. 2 (2), pp. 32-38.

[8] Hadavandi, E., Shavandi, H., &Ghanbari, A. (2010) In*tegration of Genetic Fuzzy Systems and Artificial Neural Networks for Stock Price Forecasting*,Knowledge-Based Systemsvol. 37, no. 9, pp. 6302–6309.

[9] Paul, J. C., Hoque S. M., & Rahman, M. M. (2013) *Selection of Best ARIMA Model for Forecasting Average Daily Share Price Index of Pharmaceutical Companies in Bangladesh: A Case Study on Square Pharmaceutical Ltd* Global Journal of Management and Business Research Finance vol. 13 Issue 3, pp. 15-26.

[10] Pena, D., Tiao, G. C. &Tsay, S. R. (2001) *A Course in Time Series Analysis* John Wiley & Sons, New York

[11] Cowpertwait, P. S. P. & Metcalfe, A. V. (2009) *Introductory Time Series with R* Springer Science, London.

[12] Box, G.E.P. & G.M. Jenkins (1970) *Time series analysis: forecasting and control.* Holden-Day, San Francisco, CA

[13] Adhikari, R. and R. K. Agrawal, (2013) *An Introductory Study on Time Series Modeling and Forecasting*LAP Lambert Academic Publishing, Germany.

[14] http://www.cashcraft.com/pricemovement.asp