Stock Price Prediction Using the ARIMA Model

[1]Ayodele A. Adebiyi., [2]Aderemi O. Adewumi
[1,2]School of Mathematic, Statistics & Computer Science
University of KwaZulu-Natal
Durban, South Africa

email: {adebiyi, adewumia}@ukzn.ac.za

[3]Charles K. Ayo
[3]Department of Computer & Information Sciences
Covenant University
Ota, Nigeria

email: charles.ayo@covenantuniversity.edu.ng

***Abstract*— Stock price prediction is an important topic in finance and economics which has spurred the interest of researchers over the years to develop better predictive models. The autoregressive integrated moving average (ARIMA) models have been explored in literature for time series prediction. This paper presents extensive process of building stock price predictive model using the ARIMA model. Published stock data obtained from New York Stock Exchange (NYSE) and Nigeria Stock Exchange (NSE) are used with stock price predictive model developed. Results obtained revealed that the ARIMA model has a strong potential for short-term prediction and can compete favourably with existing techniques for stock price prediction.**

Keywords- ARIMA model, Stock Price prediction, Stock market, Short-term prediction.

I. INTRODUCTION

Prediction will continue to be an interesting area of research making researchers in the domain field always desiring to improve existing predictive models. The reason is that institutions and individuals are empowered to make investment decisions and ability to plan and develop effective strategy about their daily and future endevours. Stock price prediction is regarded as one of most difficult task to accomplish in financial forecasting due to complex nature of stock market [1, 2, 3]. The desire of many investors is to lay hold of any forecasting method that could guarantee easy profiting and minimize investment risk from the stock market. This remains a motivating factor for researchers to evolve and develop new predictive models [4].

In the past years several models and techniques had been developed to stock price prediction. Among them are artificial neural networks (ANNs) model which are very popular due to its ability to learn patterns from data and infer solution from unknown data. Few related works that engaged ANNs model to stock price prediction are [5, 6, 7]. In recent time, hybrid approaches has also been engaged to improve stock price predictive models by exploiting the unique strength of each of them [2]. ANNs is from artificial intelligence perspectives.

ARIMA models are from statistical models perspectives. Generally, it is reported in literature that prediction can be done from two perspectives: statistical and artificial intelligence techniques [2]. ARIMA models are known to be robust and efficient in financial time series forecasting especially short-term prediction than even the most popular ANNs techniques ([8, 9, 10]. It has been extensively used in field of economics and finance. Other statistics models are regression method, exponential smoothing, generalized autoregressive conditional heteroskedasticity (GARCH). Few related works that has engaged ARIMA model for forecasting includes [11, 12, 13, 14, 15, 16].

In this paper extensive process of building ARIMA models for short-term stock price prediction is presented. The results obtained from real-life data demonstrated the potential strength of ARIMA models to provide investors short-term prediction that could aid investment decision making process.

The rest of the paper is organized as follows. Section 2 presents brief overview of ARIMA model. Section 3 describes the methodology used while section 4 discusses the experimental results obtained. The paper is concluded in section 5.

II. ARIMA MODEL

Box and Jenkins in 1970 introduced the ARIMA model. It also referred to as Box-Jenkins methodology composed of set of activities for identifying, estimating and diagnosing ARIMA models with time series data. The model is most prominent methods in financial forecasting [1, 12, 9]. ARIMA models have shown efficient capability to generate short-term forecasts. It constantly outperformed complex structural models in short-term prediction [17]. In ARIMA model, the future value of a variable is a linear combination of past values and past errors, expressed as follows:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \ldots - \theta_q \varepsilon_{t-q} \quad (1)$$

where,

Y_t is the actual value and ε_t is the random error at *t*, ϕ_i and θ_j are the coefficients, *p* and *q* are integers that are often referred to as autoregressive and moving average, respectively.

978-1-4799-4923-6/14 $31.00 © 2014 IEEE
DOI 10.1109/UKSim.2014.67

The steps in building ARIMA predictive model consist of *model identification, parameter estimation and diagnostic checking* [18].

III. METHODOLOGY

The method used in this study to develop ARIMA model for stock price forecasting is explained in detail in subsections below. The tool used for implementation is Eviews software version 5. Stock data used in this research work are historical daily stock prices obtained from two countries stock exchanged. The data composed of four elements, namely: open price, low price, high price and close price respectively. In this research the closing price is chosen to represent the price of the index to be predicted. Closing price is chosen because it reflects all the activities of the index in a trading day.

To determine the best ARIMA model among several experiments performed, the following criteria are used in this study for each stock index.

- Relatively small of BIC (Bayesian or Schwarz Information Criterion)
- Relatively small standard error of regression (S.E. of regression)
- Relatively high of adjusted R^2
- Q-statistics and correlogram show that there is no significant pattern left in the autocorrelation functions (ACFs) and partial autocorrelation functions (PACFs) of the residuals, it means the residual of the selected model are white noise.

The subsections below described the processes of ARIMA model-development.

A. *ARIMA (p, d, q) Model for Nokia Stock Index*

Nokia stock data used in this study covers the period from 25th April, 1995 to 25th February, 2011 having a total number of 3990 observations. Figure 1 depicts the original pattern of the series to have general overview whether the time series is stationary or not. From the graph below the time series have random walk pattern.

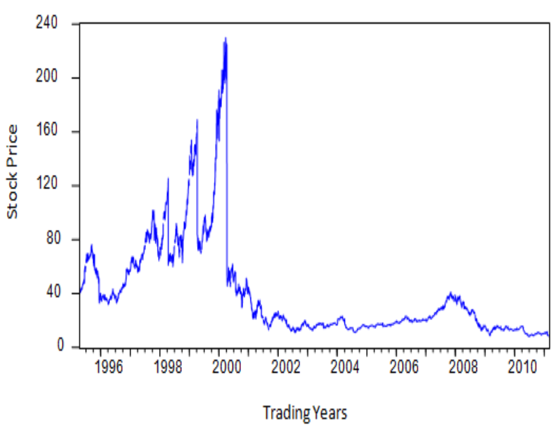
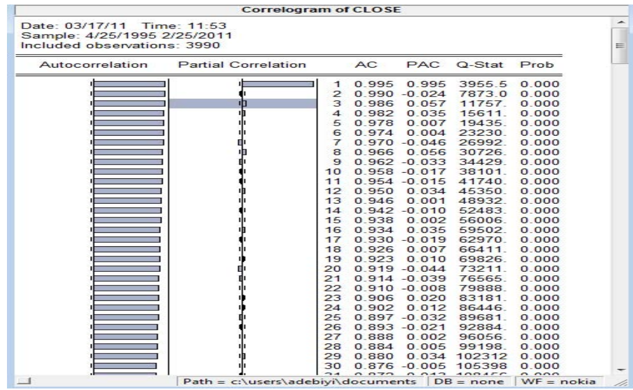


Figure 1: Graphical representation of the Nokia stock closing price index

Figure 2 is the correlogram of Nokia time series. From the graph, the ACF dies down extremely slowly which simply means that the time series is nonstationary. If the series is not stationary, it is converted to a stationary series by differencing. After the first difference, the series "DCLOSE" of Nokia stock index becomes stationary as shown in figure 3 and figure 4 of the line graph and correlogram respectively.



Correlogram of CLOSE

Date: 03/17/11 Time: 11:53
Sample: 4/25/1995 2/25/2011
Included observations: 3990

	AC	PAC	Q-Stat	Prob
1	0.995	0.995	3955.5	0.000
2	0.990	-0.024	7873.0	0.000
3	0.986	0.057	11757.	0.000
4	0.982	0.035	15611.	0.000
5	0.978	0.007	19435.	0.000
6	0.974	0.004	23230.	0.000
7	0.970	-0.046	26992.	0.000
8	0.966	0.056	30726.	0.000
9	0.962	-0.033	34429.	0.000
10	0.958	-0.017	38101.	0.000
11	0.954	-0.015	41740.	0.000
12	0.950	0.034	45350.	0.000
13	0.946	0.001	48932.	0.000
14	0.942	-0.010	52483.	0.000
15	0.938	0.002	56006.	0.000
16	0.934	0.035	59502.	0.000
17	0.930	-0.019	62970.	0.000
18	0.926	0.007	66411.	0.000
19	0.923	0.010	69826.	0.000
20	0.919	-0.044	73211.	0.000
21	0.914	-0.039	76565.	0.000
22	0.910	-0.008	79888.	0.000
23	0.906	0.020	83181.	0.000
24	0.902	0.012	86446.	0.000
25	0.897	-0.032	89681.	0.000
26	0.893	-0.021	92884.	0.000
27	0.888	0.002	96056.	0.000
28	0.884	0.005	99198.	0.000
29	0.880	0.034	102312	0.000
30	0.876	-0.005	105398	0.000

Path = c:\users\adebiyi\documents | DB = none | WF = nokia

Figure 2: The correlogram of Nokia stock price index

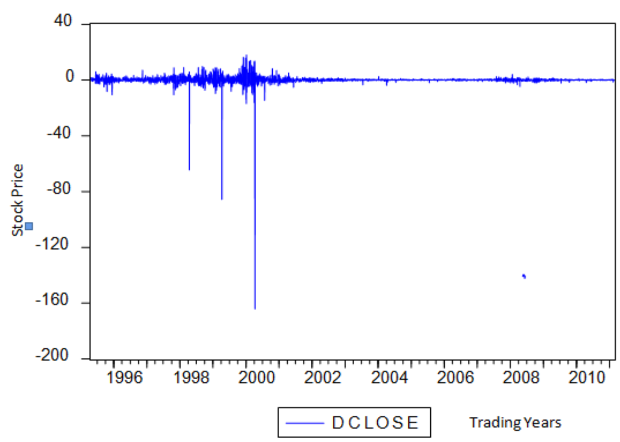
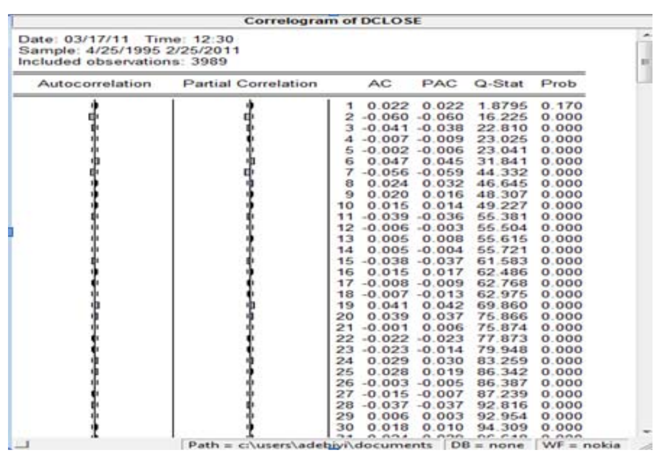


Figure 3: Graphical representation of the Nokia stock price index after differencing .



Correlogram of DCLOSE

Date: 03/17/11 Time: 12:30
Sample: 4/25/1995 2/25/2011
Included observations: 3989

	AC	PAC	Q-Stat	Prob
1	0.022	0.022	1.8795	0.170
2	-0.060	-0.060	16.225	0.000
3	-0.041	-0.038	22.810	0.000
4	-0.007	-0.009	23.025	0.000
5	-0.002	-0.006	23.041	0.000
6	0.047	0.045	31.841	0.000
7	-0.056	-0.059	44.332	0.000
8	0.024	0.032	46.645	0.000
9	0.020	0.016	48.307	0.000
10	0.015	0.014	49.227	0.000
11	-0.039	-0.036	55.381	0.000
12	-0.006	-0.003	55.504	0.000
13	0.005	0.008	55.615	0.000
14	0.005	-0.004	55.721	0.000
15	-0.038	-0.037	61.583	0.000
16	0.015	0.017	62.486	0.000
17	-0.008	-0.009	62.768	0.000
18	-0.007	-0.013	62.975	0.000
19	0.041	0.042	69.860	0.000
20	0.039	0.037	75.866	0.000
21	-0.001	0.006	75.874	0.000
22	-0.022	-0.023	77.873	0.000
23	-0.023	-0.014	79.948	0.000
24	0.029	0.030	83.259	0.000
25	0.028	0.019	86.342	0.000
26	-0.003	-0.005	86.387	0.000
27	-0.015	-0.007	87.239	0.000
28	-0.037	-0.037	92.816	0.000
29	0.006	0.003	92.954	0.000
30	0.018	0.010	94.309	0.000

Path = c:\users\adebiyi\documents | DB = none | WF = nokia

Figure 4: The correlogram of Nokia stock price index after first differencing

In figure 5 the model checking was done with Augmented Dickey Fuller (ADF) unit root test on "DCLOSE" of Nokia stock index. The result confirms that the series becomes stationary after the first-difference of the series.

Augmented Dickey-Fuller Unit Root Test on DCLOSE

Null Hypothesis: DCLOSE has a unit root
Exogenous: Constant
Lag Length: 1 (Automatic based on SIC, MAXLAG=30)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-46.89879	0.0001
Test critical values:	1% level	-3.431805	
	5% level	-2.862068	
	10% level	-2.567094	

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DCLOSE)
Method: Least Squares
Date: 03/17/11 Time: 12:37
Sample (adjusted): 4/28/1995 2/25/2011
Included observations: 3987 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DCLOSE(-1)	-1.037429	0.022121	-46.89879	0.0000
D(DCLOSE(-1))	0.060437	0.015814	3.821732	0.0001
C	-0.008294	0.056703	-0.146267	0.8837

R-squared	0.491018	Mean dependent var	0.000100
Adjusted R-squared	0.490762	S.D. dependent var	5.017265
S.E. of regression	3.580364	Akaike info criterion	5.389558
Sum squared resid	51070.94	Schwarz criterion	5.394292
Log likelihood	-10741.08	F-statistic	1921.694
Durbin-Watson stat	2.004589	Prob(F-statistic)	0.000000

Path = c:\users\adebiyi\documents | DB = none | WF = nokia

Figure 5: ADF unit root test for DCLOSE of Nokia stock index.

Table 1 shows the different parameters of autoregressive (p) and moving average (q) among the several ARIMA model experimented upon . ARIMA (2, 1, 0) is considered the best for Nokia stock index as shown in figure 6. The model returned the smallest Bayesian or Schwarz information criterion of 5.3927 and relatively smallest standard error of regression of 3.5808 as shown in figure 6.

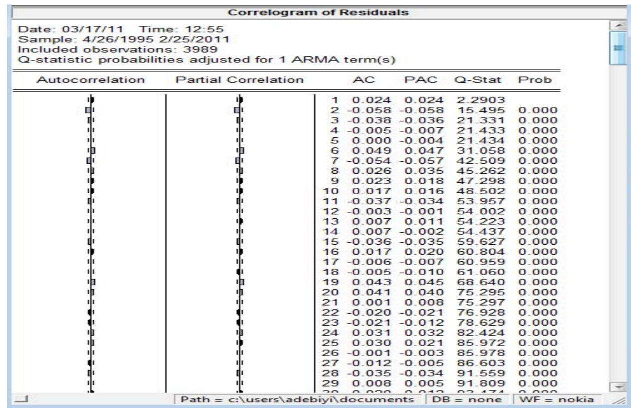
Dependent Variable: DCLOSE
Method: Least Squares
Date: 03/19/11 Time: 16:08
Sample (adjusted): 4/28/1995 2/25/2011
Included observations: 3987 after adjustments
Convergence achieved after 3 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.007997	0.053504	-0.149458	0.8812
AR(2)	-0.059938	0.015813	-3.790524	0.0002

R-squared	0.003593	Mean dependent var	-0.007988
Adjusted R-squared	0.003343	S.D. dependent var	3.586866
S.E. of regression	3.580866	Akaike info criterion	5.389588
Sum squared resid	51098.07	Schwarz criterion	5.392744
Log likelihood	-10742.14	F-statistic	14.36807
Durbin-Watson stat	1.958555	Prob(F-statistic)	0.000153

Figure 6: ARIMA (2, 1, 0) estimation output with DCLOSE of Nokia index.

Figure 7 is the residual of the series. If the model is good, the residuals (difference between actual and predicted values) of the model are series of random errors. Since there are no significant spikes of ACFs and PACFs, it means that the residual of the selected ARIMA model are white noise, no other significant patterns left in the time series. Therefore, there is no need to consider any AR(p) and MA(q) further.


Correlogram of Residuals

Date: 03/17/11 Time: 12:55
Sample: 4/26/1995 2/25/2011
Included observations: 3989
Q-statistic probabilities adjusted for 1 ARMA term(s)

	AC	PAC	Q-Stat	Prob
1	0.024	0.024	2.2903	
2	-0.058	-0.058	15.495	0.000
3	-0.038	-0.036	21.331	0.000
4	-0.005	-0.007	21.433	0.000
5	0.000	-0.004	21.434	0.000
6	0.049	0.047	31.058	0.000
7	-0.054	-0.057	42.509	0.000
8	0.026	0.035	45.262	0.000
9	0.023	0.018	47.298	0.000
10	0.017	0.016	48.502	0.000
11	-0.037	-0.034	53.957	0.000
12	-0.003	-0.001	54.002	0.000
13	0.007	0.011	54.223	0.000
14	0.007	-0.002	54.437	0.000
15	-0.036	-0.035	59.627	0.000
16	0.017	0.020	60.804	0.000
17	-0.006	-0.007	60.959	0.000
18	-0.005	-0.010	61.060	0.000
19	0.043	0.045	68.640	0.000
20	0.041	0.040	75.295	0.000
21	0.001	0.008	75.297	0.000
22	-0.020	-0.021	76.928	0.000
23	-0.021	-0.012	78.629	0.000
24	0.031	0.032	82.424	0.000
25	0.030	0.021	85.972	0.000
26	-0.001	-0.003	85.978	0.000
27	-0.012	-0.005	86.603	0.000
28	-0.035	-0.034	91.559	0.000
29	0.008	0.005	91.809	0.000

Path = c:\users\adebiyi\documents | DB = none | WF = nokia

Figure 7: Correlogram of residuals of the Nokia stock index.

In forecasting form, the best model selected can be expressed as follows:

$$Y_t = \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \varepsilon_t \qquad (2)$$

where, $\varepsilon_t = Y_t - \hat{Y}_t$ (i.e., the difference between the actual value of the series and the forecast value)

TABLE I: STATISTICAL RESULTS OF DIFFERENT ARIMA PARAMETERS FOR NOKIA STOCK INDEX

ARIMA	BIC	Adjusted R^2	S.E. of Regression
(1, 0, 0)	5.3936	0.9907	3.5824
(1, 0, 1)	5.3950	0.9907	3.5817
(2, 0, 0)	6.1061	0.9811	5.1157
(0, 0, 1)	8.8324	0.7126	19.9942
(0, 0, 2)	8.8871	0.6964	20.5490
(1, 1, 0)	5.3956	0.0002	3.5860
(0, 1, 0)	5.3937	0.0000	3.5859
(0, 1, 1)	5.3953	0.0002	3.5854
(1, 1,2)	5.3941	0.0035	3.5800
(2, 1, 0)	**5.3927**	**0.0033**	**3.5808**
(2, 1, 2)	5.3947	0.0031	3.5812

The bold row represent the best ARIMA model among the several experiments.

B. ARIMA (p, d, q) Model for Zenith Bank Index

The stock data of Zenith bank used in this study covered the period from 3rd January, 2006 to 25th February, 2011 with total of 1296 observations. Figure 8 is the original pattern of the series. From the graph there was upward movement of the index from 2006 and downward movement is observed from 2008 possibly because of world financial crisis experienced at that time.

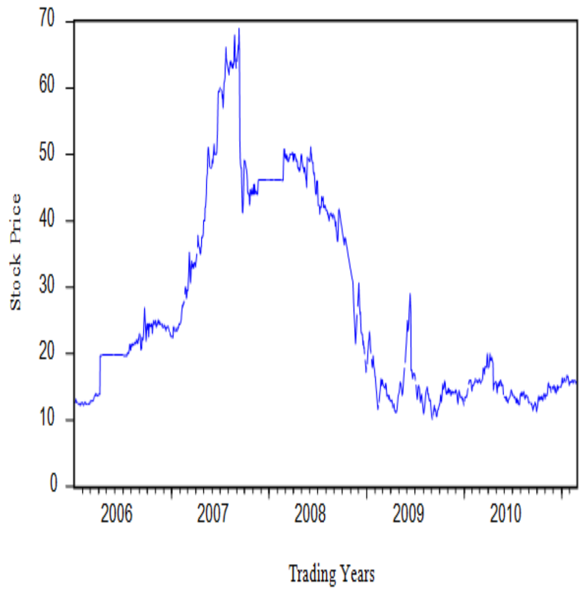
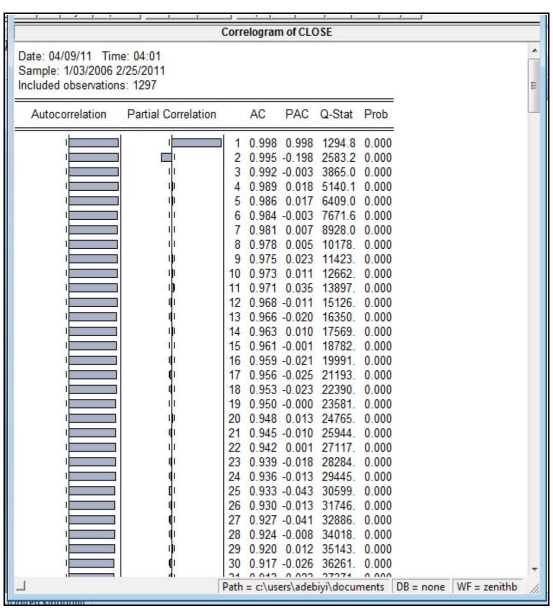


Figure 8: Graphical representation of the Zenith Bank stock index closing price.

Figure 9 is the correlogram of the time series of Zenith bank stock index. From the graph of the correlogram, the ACF dies down extremely slowly which simply means that the time series is nonstationary. If the series is not stationary, there is need to convert to stationary series by differencing. After the first difference, the series "DCLOSE" of Zenith bank stock index becomes stationary as shown in figure 10 and figure 11 of the line graph and correlogram of the series after first differencing.



Correlogram of CLOSE

Date: 04/09/11 Time: 04:01
Sample: 1/03/2006 2/25/2011
Included observations: 1297

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.998	0.998	1294.8	0.000
		2	0.995	-0.198	2583.2	0.000
		3	0.992	-0.003	3865.0	0.000
		4	0.989	0.018	5140.1	0.000
		5	0.986	0.017	6409.0	0.000
		6	0.984	-0.003	7671.6	0.000
		7	0.981	0.007	8928.0	0.000
		8	0.978	0.005	10178.	0.000
		9	0.975	0.023	11423.	0.000
		10	0.973	0.011	12662.	0.000
		11	0.971	0.035	13897.	0.000
		12	0.968	-0.011	15126.	0.000
		13	0.966	-0.020	16350.	0.000
		14	0.963	0.010	17569.	0.000
		15	0.961	-0.001	18782.	0.000
		16	0.959	-0.021	19991.	0.000
		17	0.956	-0.025	21193.	0.000
		18	0.953	-0.023	22390.	0.000
		19	0.950	-0.000	23581.	0.000
		20	0.948	0.013	24765.	0.000
		21	0.945	-0.010	25944.	0.000
		22	0.942	0.001	27117.	0.000
		23	0.939	-0.018	28284.	0.000
		24	0.936	-0.013	29445.	0.000
		25	0.933	-0.043	30599.	0.000
		26	0.930	-0.013	31746.	0.000
		27	0.927	-0.041	32886.	0.000
		28	0.924	-0.008	34018.	0.000
		29	0.920	0.012	35143.	0.000
		30	0.917	-0.026	36261.	0.000

Path = c:\users\adebiyi\documents DB = none WF = zenithb

Figure 9: The correlogram of Zenith Bank stock price index

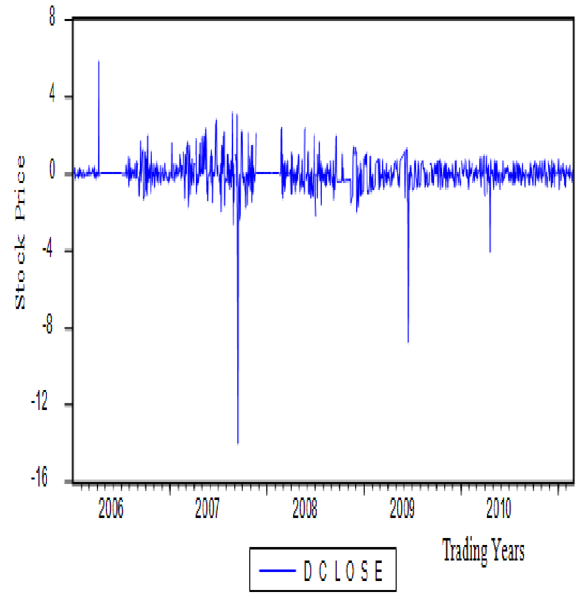
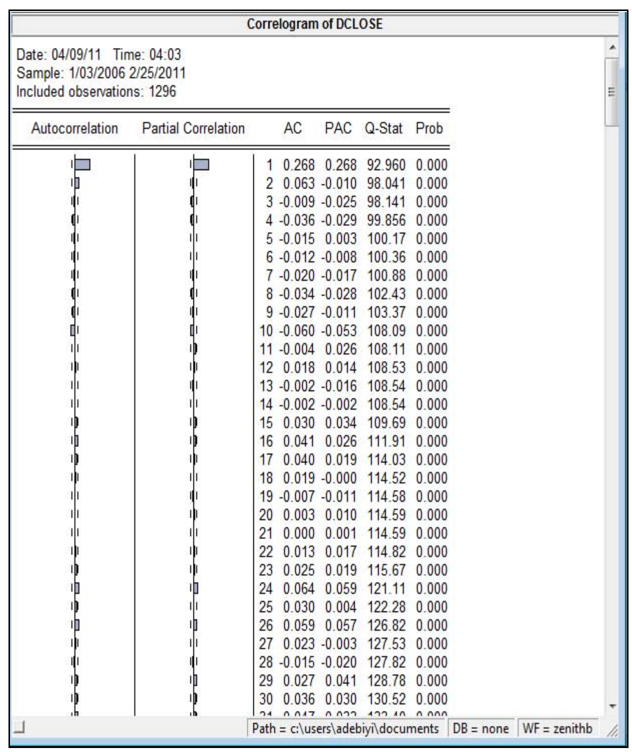


Figure 10: Graphical representation of the Zenith bank stock index first differencing



Correlogram of DCLOSE

Date: 04/09/11 Time: 04:03
Sample: 1/03/2006 2/25/2011
Included observations: 1296

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.268	0.268	92.960	0.000
		2	0.063	-0.010	98.041	0.000
		3	-0.009	-0.025	98.141	0.000
		4	-0.036	-0.029	99.856	0.000
		5	-0.015	0.003	100.17	0.000
		6	-0.012	-0.008	100.36	0.000
		7	-0.020	-0.017	100.88	0.000
		8	-0.034	-0.028	102.43	0.000
		9	-0.027	-0.011	103.37	0.000
		10	-0.060	-0.053	108.09	0.000
		11	-0.004	0.026	108.11	0.000
		12	0.018	0.014	108.53	0.000
		13	-0.002	-0.016	108.54	0.000
		14	-0.002	-0.002	108.54	0.000
		15	0.030	0.034	109.69	0.000
		16	0.041	0.026	111.91	0.000
		17	0.040	0.019	114.03	0.000
		18	0.019	-0.000	114.52	0.000
		19	-0.007	-0.011	114.58	0.000
		20	0.003	0.010	114.59	0.000
		21	0.000	0.001	114.59	0.000
		22	0.013	0.017	114.82	0.000
		23	0.025	0.019	115.67	0.000
		24	0.064	0.059	121.11	0.000
		25	0.030	0.004	122.28	0.000
		26	0.059	0.057	126.82	0.000
		27	0.023	-0.003	127.53	0.000
		28	-0.015	-0.020	127.82	0.000
		29	0.027	0.041	128.78	0.000
		30	0.036	0.030	130.52	0.000

Path = c:\users\adebiyi\documents DB = none WF = zenithb

Figure 11: The correlogram of Zenith bank stock price index after first differencing.

Figure 12 is the ADF unit root test on "DCLOSE" of the series which also indicates the first difference of the series becomes stationary.

Augmented Dickey-Fuller Unit Root Test on DCLOSE

Null Hypothesis: DCLOSE has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic based on SIC, MAXLAG=22)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-27.33425	0.0000
Test critical values:	1% level	-3.435188	
	5% level	-2.863564	
	10% level	-2.567897	

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

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(DCLOSE)
Method: Least Squares
Date: 04/09/11 Time: 04:04
Sample (adjusted): 1/05/2006 2/25/2011
Included observations: 1295 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DCLOSE(-1)	-0.732475	0.026797	-27.33425	0.0000
C	0.001779	0.021843	0.081457	0.9351

R-squared	0.366226	Mean dependent var	0.000139
Adjusted R-squared	0.365736	S.D. dependent var	0.986995
S.E. of regression	0.786049	Akaike info criterion	2.357948
Sum squared resid	798.9102	Schwarz criterion	2.365927
Log likelihood	-1524.772	F-statistic	747.1610
Durbin-Watson stat	1.994624	Prob(F-statistic)	0.000000

Path = c:\users\adebiyi\documents DB = none WF = zenithb

Figure 12: ADF unit root test for DCLOSE of Zenith bank stock index.

Table 2 shows the different parameters of autoregressive (p) and moveing average (q) of the ARIMA model in order to get the best fitted model. ARIMA (1, 0, 1) is relatively the best model as indicated in figure 13. The model returned the smallest Bayesian or Schwarz information criterion of 2.3736 and relatively smallest standard error of regression of 0.7872 as shown in figure 13.

Dependent Variable: CLOSE
Method: Least Squares
Date: 04/09/11 Time: 04:24
Sample (adjusted): 1/04/2006 2/25/2011
Included observations: 1296 after adjustments
Convergence achieved after 7 iterations
Backcast: 1/03/2006

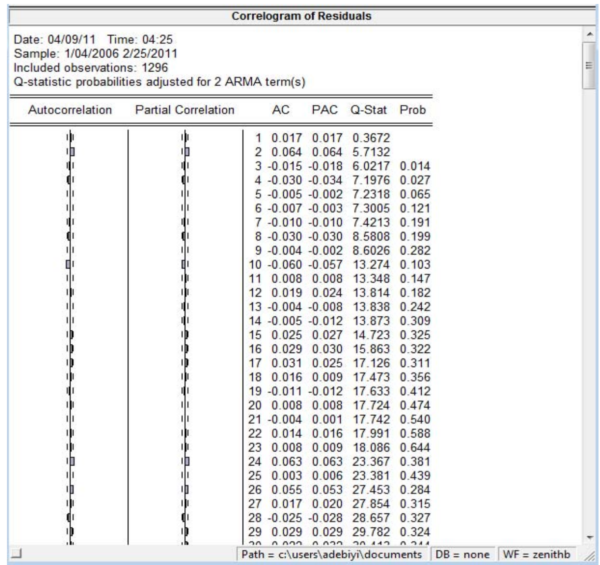
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	27.76649	11.97506	2.318694	0.0206
AR(1)	0.997704	0.001829	545.5149	0.0000
MA(1)	0.254140	0.026934	9.435526	0.0000

R-squared	0.997254	Mean dependent var	26.70431
Adjusted R-squared	0.997250	S.D. dependent var	15.01153
S.E. of regression	0.787215	Akaike info criterion	2.361682
Sum squared resid	801.2819	Schwarz criterion	2.373642
Log likelihood	-1527.370	F-statistic	234805.8
Durbin-Watson stat	1.966274	Prob(F-statistic)	0.000000

Inverted AR Roots	1.00
Inverted MA Roots	-.25

Figure 13: ARIMA (1, 0, 1) estimation output with DCLOSE of Zenith bank index.

Figure 14 is the correlogram of residual of the seies. From the figure it is obvious there is no significant spike of ACFs and PACFs. This means that the residual of this selected ARIMA model are white noise. There is no other significant patterns left in the time series and there is no need for further consideration of another AR(p) and MA(q).


Correlogram of Residuals

Date: 04/09/11 Time: 04:25
Sample: 1/04/2006 2/25/2011
Included observations: 1296
Q-statistic probabilities adjusted for 2 ARMA term(s)

	AC	PAC	Q-Stat	Prob
1	0.017	0.017	0.3672	
2	0.064	0.064	5.7132	
3	-0.015	-0.018	6.0217	0.014
4	-0.030	-0.034	7.1976	0.027
5	-0.005	-0.002	7.2318	0.065
6	-0.007	-0.003	7.3005	0.121
7	-0.010	-0.010	7.4213	0.191
8	-0.030	-0.030	8.5808	0.199
9	-0.004	-0.002	8.6026	0.282
10	-0.060	-0.057	13.274	0.103
11	0.008	0.008	13.348	0.147
12	0.019	0.024	13.814	0.182
13	-0.004	-0.008	13.838	0.242
14	-0.005	-0.012	13.873	0.309
15	0.025	0.027	14.723	0.325
16	0.029	0.030	15.863	0.322
17	0.031	0.025	17.126	0.311
18	0.016	0.009	17.473	0.356
19	-0.011	-0.012	17.633	0.412
20	0.008	0.008	17.724	0.474
21	-0.004	0.001	17.742	0.540
22	0.014	0.016	17.991	0.588
23	0.008	0.009	18.086	0.644
24	0.063	0.063	23.367	0.381
25	0.003	0.006	23.381	0.439
26	0.055	0.053	27.453	0.284
27	0.017	0.020	27.854	0.315
28	-0.025	-0.028	28.657	0.327
29	0.029	0.029	29.782	0.324

Path = c:\users\adebiyi\documents DB = none WF = zenithb

Figure 14: Correlogram of residuals of the Zenith bank stock index.

In forecasting form, the best model selected can be expressed as follows:

$$Y_t = \phi_1 Y_{t-1} - \theta_1 \varepsilon_{t-1} + \varepsilon_t \qquad (3)$$

where,

$\varepsilon_t = Y_t - \hat{Y}_t$ (i.e., the difference between the actual value of the series and the forecast value)

TABLE II: STATISTICAL RESULTS OF DIFFERENT ARIMA PARAMETERS FOR ZENITH BANK STOCK INDEX

ARIMA	BIC	Adjusted R^2	S.E. of Regression
(1, 0, 0)	2.4385	0.9970	0.8151
(1, 0, 1)	**2.3736**	**0.9972**	**0.7872**
(2, 0, 0)	3.3682	0.9925	1.2974
(0, 0, 1)	6.9285	0.7372	7.6951
(0, 0, 2)	6.9815	0.7228	7.9018
(1, 1, 0)	2.3659	0.0708	0.7860
(0, 1, 0)	2.4338	0.0000	0.8151
(0, 1, 1)	2.3693	0.0669	0.7873
(1, 1,2)	2.3714	0.0701	0.7863
(2, 1, 0)	2.4370	0.0031	0.8144
(2, 1, 2)	2.4412	0.0036	0.8142

The bold row represent the best ARIMA model among the several experiments

IV. RESULTS AND DISCUSSION

The experimental results of each of stock index are discussed in the subsection below.

A. Result of ARIMA Model for Nokia Stock Price Prediction

Table 3 is the result of the predicted values of ARIMA (2, 1, 0) considered the best model for Nokia stock index. Figure 15 gives graphical illustration of the level accuracy of the predicted price against actual stock price to see the performance of the ARIMA model selected. From the graph, is obvious that the performance is satisfactory.

TABLE III: SAMPLE OF EMPIRICAL RESULTS OF ARIMA (2,1,0) OF NOKIA STOCK INDEX.

Sample Period	Actual Values	Predicted Values
1/3/2010	13.28	13.58
2/3/2010	13.51	13.69
3/3/2010	13.86	13.80
4/3/2010	13.78	13.91
5/3/2010	14.13	14.02
8/3/2010	14.17	14.13
9/3/2010	14.12	14.24
10/3/2010	14.56	14.35
11/3/2010	14.49	14.45
12/3/2010	14.84	14.56
15/3/2010	14.81	14.67
16/3/2010	15.14	14.77
17/3/2010	15.42	14.88
18/3/2010	15.28	14.98
19/3/2010	15.07	15.09
22/3/2010	15.11	15.19
23/3/2010	15.26	15.30
24/3/2010	15.07	15.40
25/3/2010	15.20	15.50
26/3/2010	15.46	15.60
29/3/2010	15.42	15.71
30/3/2010	15.41	15.81
31/3/2010	15.54	15.91

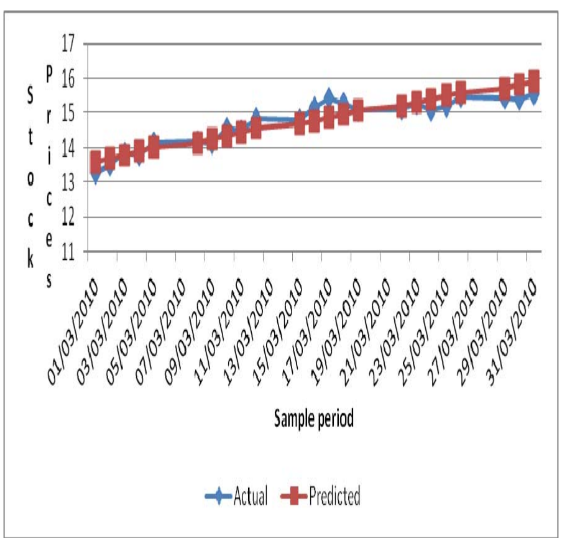


Figure 15: Graph of Actual Stock Price vs Predicted values of Nokia Stock Index

B. Result of ARIMA Model for Zenith Bank Stock Price Prediction

In this case, ARIMA (1, 0, 1) was selected as the best model for Zenith bank stock index after several adjustment of the autoregressive (p) and moving average (q) parameters in Eviews software used. Table 4 contained the predicted values of the model selected and figure 16 is the graph of predicted price against actual stock price to demonstrate the correlation of accuracy. From the graph, the performance of the ARIMA model selected is quite impressive as there are some instances of closely related of actual and predicted values.

TABLE IV: SAMPLE OF EMPIRICAL RESULTS OF ARIMA (1,0,1) OF ZENITH BANK INDEX

Sample Period	Actual Values	Predicted Values
1/3/2010	16.19	15.83
2/3/2010	15.98	15.86
3/3/2010	15.71	15.89
4/3/2010	15.50	15.91
5/3/2010	15.70	15.94
8/3/2010	15.75	15.97
9/3/2010	15.86	15.99
10/3/2010	16.00	16.02
11/3/2010	16.19	16.05
12/3/2010	16.99	16.08
15/3/2010	17.83	16.10
16/3/2010	17.71	16.13
17/3/2010	17.50	16.16
18/3/2010	16.85	16.18
19/3/2010	17.69	16.21
22/3/2010	18.00	16.24
23/3/2010	18.00	16.26
24/3/2010	17.85	16.29
25/3/2010	17.89	16.31
26/3/2010	18.11	16.34
29/3/2010	19.01	16.37
30/3/2010	19.96	16.39
31/3/2010	18.97	16.42

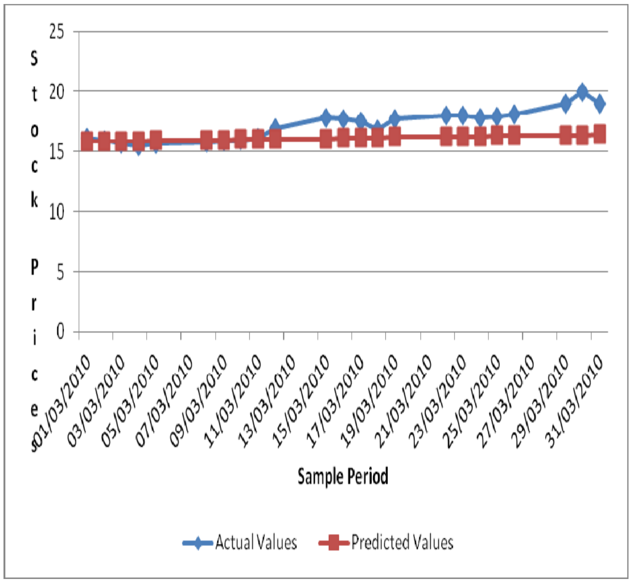


Figure 16: Graph of Actual Stock Price vs Predicted values of Zenith Bank Stock Index

V. CONCLUSION

This paper presents extensive process of building ARIMA model for stock price prediction. The experimental results obtained with best ARIMA model demonstrated the potential of ARIMA models to predict stock prices satisfactory on short-term basis. This could guide investors in stock market to make profitable investment decisions. With the results obtained ARIMA models can compete reasonably well with emerging forecasting techniques in short-term prediction.

REFERENCES

[1] P. Pai and C. Lin, "A hybrid ARIMA and support vector machines model in stock price prediction", Omega vol.33 pp. 497-505, 2005

[2] J.J. Wang, J.Z. Wang, Z.G. Zhang and S.P. Guo, "Stock index forecasting based on a hybrid model", Omega vol.40 pp.758-766, 2012.

[3] L.Y. Wei, "A hybrid model based on ANFIS and adaptive expectation genetic algorithm to forecast TAIEX", Economic Modelling vol. 33 pp. 893-899, 2013.

[4] G.S. Atsalakis, E.M Dimitrakakis. and C.D. Zopounidis, "Elliot Wave Theory and neuro-fuzzy systems, stock market prediction: The WASP system", Expert Systems with Applications, vol. 38, pp.9196–9206, 2011.

[5] S. K.Mitra, "Optimal Combination of Trading Rules Using Neural Networks", International Business Research, vol. 2, no. 1, pp. 86-99, 2009.

[6] G.S. Atsalakis and P.V. Kimon, "Forecasting stock market short-term trends using a neuro-fuzzy methodology", Expert Systems with Applications, vol. 36, no. 7, pp.10696–10707, 2009.

[7] M.M. Mohamed, "Forecasting stock exchange movements using neural networks: empirical evidence from Kuwait", Expert Systems with Applications, vol. 27, no. 9, pp.6302–6309, 2010.

[8] L.C. Kyungjoo, Y. Sehwan and J. John, "Neural Network Model vs. SARIMA Model In Forecasting Korean Stock Price Index (KOSPI), Issues in Information System, vol. 8 no. 2, pp. 372-378, 2007.

[9] N. Merh, V.P. Saxena, and K.R. Pardasani, "A Comparison Between Hybrid Approaches of ANN and ARIMA For Indian Stock Trend Forecasting", Journal of Business Intelligence, vol. 3, no.2, pp. 23-43, 2010.

[10] J. Sterba and Hilovska, "The Implementation of Hybrid ARIMA Neural Network Prediction Model for Aggregate Water Consumption Prediction", Aplimat- Journal of Applied Mathematics, vol.3, no.3, pp.123-131, 2010.

[11] C. Javier, E. Rosario, J.N. Francisco and J.C. Antonio, "ARIMA Models to Predict Next Electricity Price", IEEE Transactions on Power Systems vol. 18 no.3, pp.1014-1020, 2003.

[12] N. Rangan and N. Titida, "ARIMA Model for Forecasting Oil Palm Price", Proceedings of the 2nd IMT-GT Regional Conference on Mathematics, Statistics and Applications, Universiti Sains Malaysia, 2006.

[13] M. Khasel, M. Bijari, and G.A.R Ardali, , "Improvement of Auto-Regressive Integrated Moving Average models using Fuzzy logic and pp. 956-967, 2009.

[14] C. Lee, C. Ho, "Short-term load forecasting using lifting scheme and ARIMA model", Expert System with Applications, vol.38, no.5, pp.5902-5911, 2011.

[15] M. Khashei, M. Bijari, G. A. R. Ardal, "Hybridization of autoregressive integrated moving average (ARIMA) with probabilistic neural networks", Computers and Industrial Engineering, vol. 63, no.1, pp.37-45, 2012

[16] C. Wang, "A comparison study of between fuzzy time series model and ARIMA model for forecasting Taiwan Export", Expert System with Applications, vol.38, no.8, pp.9296-9304, 2011.

[17] A. Meyler, G. Kenny and T. Quinn, "Forecasting Irish Inflation using ARIMA Models", Central Bank of Ireland Research Department, Technical Paper, 3/RT/1998.

[18] B.G. Tabachnick and L.S. Fidell, "Using multivariate statistics", 4th ed., Person Education Company, USA 2001.