Forecasting Saudi Arabia Daily Stock Market Prices

Mahmoud Al-Zyood

Abstract: The Saudi stock market is regionally and globally important due to the country's economic position as the world's largest oil producer and its inherent socio-political role as a major world economy. Stock market prices are one of the main factors affecting the national economy, indicating economic strength and attracting investment. This paper inspects the best autoregressive integrated moving average (ARIMA) model to forecast daily stock market prices in Saudi Arabia. The results indicate that the optimum model is ARIMA (4, 4, 0, 0), due to the ACF having an exponential deterioration and the PACF having a spike at lag12, which is an indication of its being the best model to forecast Saudi Arabia stock market prices from 2000 to 2018. The least Akaike Information Criterion (AIC) value was used to select the appropriate model from 25 tentative ARMA models. The chosen model is the first one, AIC -5.404104. The selected ARIMA (4, 4) (0, 0) predicts the future values of time series (stock market prices) with 95% prediction intervals for the next year. It is important to focus on the improvement and development of other models to improve the forecasting process and improve the ability of companies to plan. The results expected by the model indicate economic strength in the near future, which stimulates the economic situation of the state and increases confidence in it.

Keywords: Bayesian Information Criterion, Akaike Information Criterion, Saudi Stock Exchange (Tadawul), Stock Prices, Prediction Models, Forecasting.

I. INTRODUCTION

This research applies the Arima model to predict the prices of the Saudi Stock Exchange (Tadawul). Stock prices usually change significantly during short periods of time and are inherently volatile and unpredictable the importance of this research is to help investors and companies make more informed decisions, which represents a significant impact on the Saudi economy and society in general. Prediction models are essential to attract future investment and knowledge of companies' probably future financial situation affects their access to capital. Furthermore, investors are keen to buying and sell shares profitably in order to maximize their value. According to the Annual Statistical Report 2016 of the Saudi Stock Exchange (Tadawul, 2016), the 'Banks & Financial Services' sector directed the market in terms of the value of shares traded during the year, amounting to SAR 219.35 billion, representing 18.96% of the total value traded. This study uses prediction models for stock market forecasting based on the example of previous studies reviewed in the following section and provides consciousness for investors to help them adjust their plans and implement elective., so Investment planning to gaining maximum returns on their investments.

II. LITERATURE REVIEW

ARIMA models are the most general class of models that seek to explain the autocorrelation frequently found in time series data (Hyndman and Athanasopoulos, 2014). Stock markets are described as a compound, evolutionary and nonlinear active systems (Anish and Majhi, 2015), and accurate prediction of their future trends is a prerequisite for successful financial market trading (Bagheri, Peyhani and Akbari, 2014). The statisticians Box and Jenkins developed an approach to forecasting as a direct result of their practice with forecast problems in many areas, such as business, control engineering applications and economics, which developed into modern ARIMA models. ARIMA processes are a class of stochastic processes used to analyze time series (Box, Jenkins and Gwilym, 1970).

Doguwa, S. I., & Alade, S. O. (2013) utilized a framework of time series models to forecast inflation in the Irish economy and they noted the extremely optimized forecast performance of ARIMA, while focusing more on minimizing out-of-sample forecast errors rather than maximizing in-sample 'goodness of fit'. Contreras et al. (2003) used ARIMA to provide a method for spot markets and long-term contracts for mainland Spain and Californian markets to predict next-day electricity prices. The subsequent development of advanced statistical technologies enabled the application of artificial neural networks (ANNs) in stock index forecasting using traditional time series (Dase and Pawar, 2010).

Datta (2011) tested ARIMA models to predict inflation in the Bangladeshi economy, identifying ARIMA (1, 0, 1) model as the best fit for inflation data. Uko and Nkoro (2012) inspected the relative predictive power of ARIMA, VAR and ECM models in forecasting inflation in Nigeria. The results indicated that ARIMA is a good predictor of inflation in Nigeria and serves as a benchmark model in inflation forecasting.

ARIMA has thus been found to be a highly effective tool for the complex and challenging task of stock market forecasting, given that prices are ultimately affected by innumerable factors including government policy, investors' expectations, global economic and political developments and correlations with other markets (Malkiel, B. G. (2003).), thus it is used in this study to find the optimally efficient model for stock market price prediction.

Revised Version Manuscript Received on March 16, 2018

Mahmoud Al-Zyood, King Abdulaziz University, Finance College of Business, P.O. Box 344, Rabigh, Jeddah 21911, Saudi Araba. E-mail: amman2121@yahoo.com

III. MATERIALS AND METHODOLOGY

1.1 Methods

This research uses historical data from Tadawul for the period 2000-2016, with an autoregressive technique to forecast stock market prices. Comprehensive data are shown in the Appendix. The analysis uses ARIMA as a statistical method to analyze longitudinal data and identify correlation among neighboring observations (Chatfield, C.,2016), which is particularly useful in dealing with real-world data, in contrast to more academic alternatives such as stationarity and Gaussianity models (Thomson, 1994). According to Heizer and Render (2009), ARIMA basically uses the time series function, which requires a model approach to early identification and assessment of its parameters.

In ARIMA analysis, there are two simple components for representing the behavior of observed time series processes, namely the autoregressive (AR) and moving average (MA) models (Pankratz, 2009). The appropriate model passes through three stages, as shown in Figure 1, which may be reiterated until the best model for forecasting is achieved.

Figure 1: Three Steps of Time Series Modeling

1.2 Identification

The first stage includes checking the stationarity of the series through graphic check and formal statistical tools. To determine the stationarity of the data we used the autocorrelation function (ACF) and partial autocorrelation (PACF). If a data series is stationary then the variance of any major subset of the series will be different from the variance of any other major subset only by chance (Pankratz, 2009). In the identification stage of model building, we determine the possible models based on the data pattern, whereby the first condition is to check whether the series is stationary or not before we can begin searching for the best model for the data (Ofori, 2013).

According to Hamilton (1994), the stationarity condition ensures that the autoregressive parameters are invertible. If this condition is assured, then the estimated model can be prediction. Time series is said to be stationary when these properties remain constant (Cryer and Chan, 2008). To determine the stationarity of the data we used the autocorrelation function (ACF) and partial autocorrelation (PACF). The final model can be selected using penalty function statistics such as the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC) (Akaike, 1974; Sakamoto, Ishiguro and Kitagawa, 1986).

1.3 Estimation

We estimate the model from the results of the analysis to determine if the model is significant.

1.4 Diagnosis

In this stage validation testing is applied using sample data in order to see whether the proposed model fits the data and predicts the future stock prices, comparing the predicted and real stock prices.

IV. RESULTS AND DISCUSSION

1.1. Model Selection

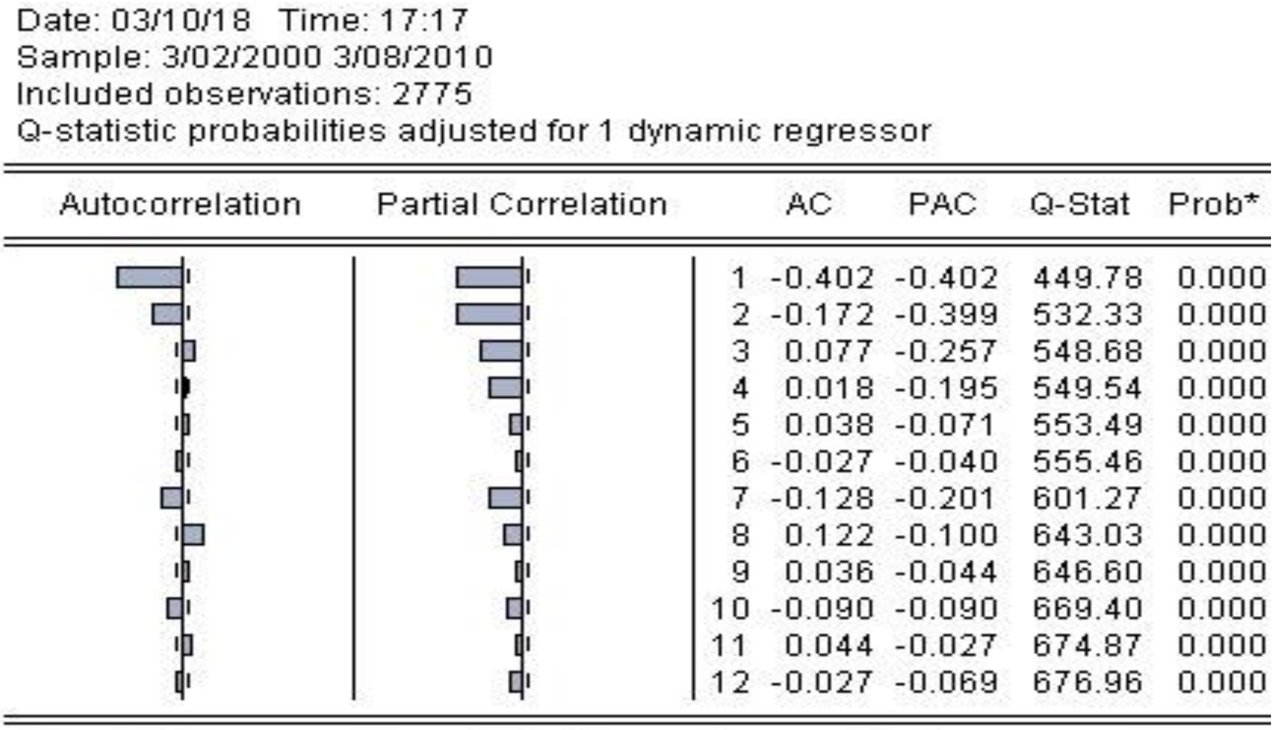
The Box-Jenkins methodology for predicting needs the series to be stationary (Hamilton, 1994). The stationary condition certifies that the autoregressive parameters are invertible. If this condition is guaranteed, the estimated model can be forecast (Table 1). If a data series is stationary then the variance of any major subset of the series will differ from the variance of any other major subset only by chance (Pankratz, 2009). Figures 2 and 3 plot the time series and ACF and PACF of data, providing a good indication of a non-stationary series.

Table 1: Least Squares Method

Dependent Variable: OPEN				
Method: Least Squares				
Date: 03/10/18 Time: 17:14				
Sample (adjusted): 3/04/2000 3/08/2010				
Included observations: 2775 after adjustments				
Variable	**Coefficient**	**Std. Error**	**t-Statistic**	**Prob.**
C	-0.259221	3.683192	-0.070379	0.9439
CLOSE	0.501497	0.008669	57.84951	0
OPEN (-1)	0.49854	0.008666	57.52781	0
R-squared	0.999362	Mean dependent var		6440.115
Adjusted R-squared	0.999361	S.D. dependent var		4172.103
S.E. of regression	105.4439	Akaike info criterion		12.15532
Sum squared resid	30820230	Schwarz criterion		12.16172

Log likelihood	-16862.5	Hannan-Quinn criter.	12.15763
F-statistic	2170034	Durbin-Watson stat	2.807778
Prob(F-statistic)	0.000000		

Figure 2: Open Stock Market Prices Plot



Date: 03/10/18 Time: 17:17
Sample: 3/02/2000 3/08/2010
Included observations: 2775
Q-statistic probabilities adjusted for 1 dynamic regressor

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	-0.402	-0.402	449.78	0.000
		2	-0.172	-0.399	532.33	0.000
		3	0.077	-0.257	548.68	0.000
		4	0.018	-0.195	549.54	0.000
		5	0.038	-0.071	553.49	0.000
		6	-0.027	-0.040	555.46	0.000
		7	-0.128	-0.201	601.27	0.000
		8	0.122	-0.100	643.03	0.000
		9	0.036	-0.044	646.60	0.000
		10	-0.090	-0.090	669.40	0.000
		11	0.044	-0.027	674.87	0.000
		12	-0.027	-0.069	676.96	0.000

Figure 3: Autocorrelations (ACF) and Partial Autocorrelations (PACF) of First Differenced Series by lag

1.2. Model Estimation

The model with the smallest AIC value was selected from 25 tentative ARMA models, which was the first one: AIC (-5.404104), as explained in Tables 2 and 3 and Figures 4 and 5. Table 2 clearly shows that the lowest AIC and BIC values are for the ARIMA (4, 4) (0,0), model with (p=4, d=0 and q=0), identifying it as the best model for forecasting the future values of our time series data. The best model is identified in Figure 4 in red.

Table 2: Model Selection Criteria

Dependent Variable: DLOG(OPEN)
Date: 03/10/18 Time: 21:00
Sample: 3/02/2000 3/08/2010
Included observations: 2775

Model	LogL	AIC*	BIC	HQ
(4,4)(0,0)	7508.194715	-5.404104	-5.382741	-5.396389
(3,2)(0,0)	7504.084816	-5.403304	-5.388350	-5.397904
(4,3)(0,0)	7505.781338	-5.403086	-5.383858	-5.396142
(2,4)(0,0)	7504.524234	-5.402900	-5.385809	-5.396728
(3,3)(0,0)	7500.090288	-5.399705	-5.382614	-5.393533
(2,3)(0,0)	7498.766320	-5.399471	-5.384517	-5.394071
(4,2)(0,0)	7499.731129	-5.399446	-5.382355	-5.393274
(3,4)(0,0)	7500.122960	-5.399008	-5.379780	-5.392064
(3,1)(0,0)	7496.134243	-5.398295	-5.385477	-5.393666
(0,4)(0,0)	7496.129978	-5.398292	-5.385474	-5.393663
(4,0)(0,0)	7496.127578	-5.398290	-5.385472	-5.393661
(1,3)(0,0)	7496.094151	-5.398266	-5.385448	-5.393637
(1,4)(0,0)	7496.997706	-5.398197	-5.383242	-5.392796
(2,2)(0,0)	7495.888045	-5.398118	-5.385299	-5.393489
(4,1)(0,0)	7496.886722	-5.398117	-5.383162	-5.392716
(3,0)(0,0)	7493.939237	-5.397434	-5.386752	-5.393576
(0,3)(0,0)	7493.667814	-5.397238	-5.386556	-5.393381
(1,1)(0,0)	7492.558417	-5.397159	-5.388614	-5.394073
(0,2)(0,0)	7491.676181	-5.396523	-5.387978	-5.393437
(2,1)(0,0)	7492.570053	-5.396447	-5.385765	-5.392589
(1,2)(0,0)	7492.569075	-5.396446	-5.385764	-5.392589
(2,0)(0,0)	7491.240061	-5.396209	-5.387664	-5.393123
(0,1)(0,0)	7489.525102	-5.395694	-5.389285	-5.393379
(1,0)(0,0)	7489.017948	-5.395328	-5.388919	-5.393014
(0,0)(0,0)	7483.270246	-5.391906	-5.387634	-5.390363

Table 3: AIC and BIC Values of Fitted ARIMA Model

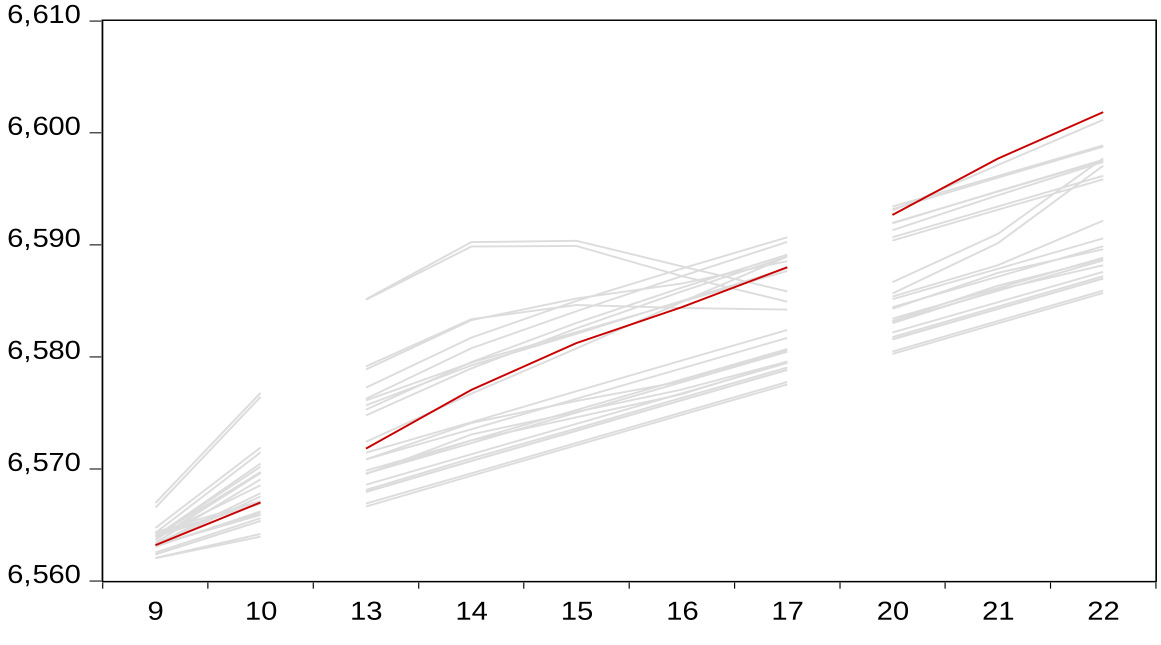


Figure 4: Forecast Comparison Graph

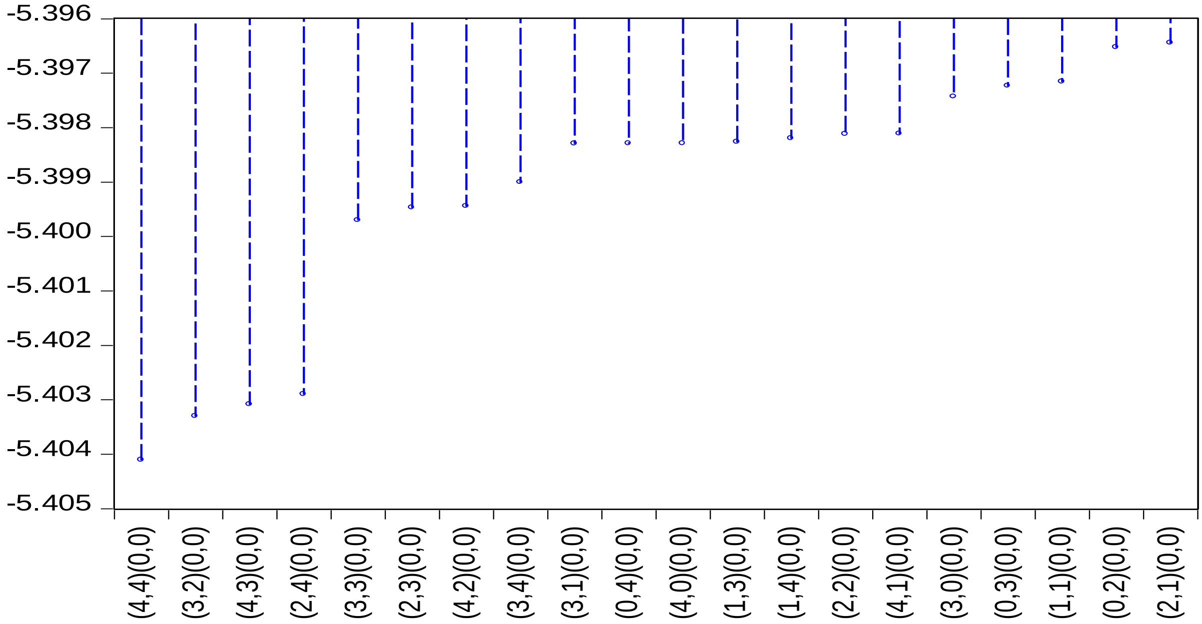


Figure 5: Akaike Information Criteria (Top 20 Models)

1.3. Forecasting using Selected ARIMA Model

The selected model ARIMA (4, 4) (0, 0) is the optimal model for our time series data, to forecast the future values of our time series (stock market prices). Table 4 shows the forecast for the next day open prices for one year with 95% prediction intervals. Figure 6 plots the next day open prices for one-year forecasting of the stock market prices by fitting ARIMA (4, 4, 0, 0).

Table 4: Sample of Stock Market Open Prices Forecasting

For 95% Confidence Intervals		
Date	**Close Prices**	**Forecasting Open Prices**
3/2/2000	2012.14	Na
3/4/2000	2015.77	2015.77
3/5/2000	2018.78	2018.78
3/6/2000	2011.84	2011.84
3/7/2000	2003.24	2003.24
3/8/2000	2010.71	2010.71
3/9/2000	2007.82	2007.82
3/11/2000	2004.29	2004.29
3/12/2000	2012.72	2012.72
3/13/2000	2022.13	2022.13
3/14/2000	2022.13	2022.13
3/21/2000	2006.28	2006.28
3/22/2000	2006.28	2006.28
3/23/2000	2006.35	2006.35
3/25/2000	2007.99	2007.99
3/26/2000	2014.02	2014.02
3/27/2000	2009.7	2009.7
3/28/2000	2006.87	2006.87
3/29/2000	2005.13	2005.13
3/30/2000	1988.14	2006.87
4/1/2000	1987.57	1987.57

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

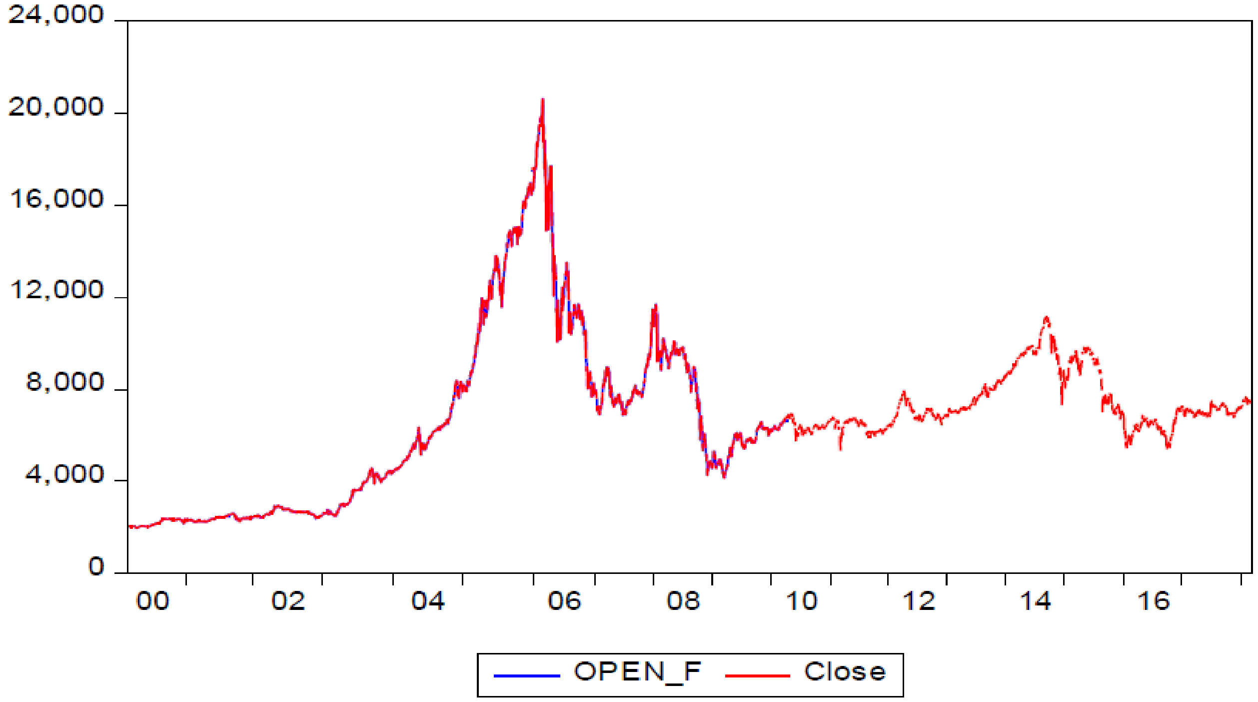


Figure 6: Plot of Actual and Forecast Data

V. CONCLUSION

In this study, the ARIMA model (4, 4, 0, 0) was the best model selected for predictions of next day open prices for one year. It is important to focus on the improvement and development of other models to improve the forecasting process and improve the ability of companies to plan. The results expected by the model indicate economic strength in the near future, which stimulates the economic situation of the state and increases confidence in its market, socioeconomic development and political stability.

REFERENCES

1. Akaike, H. (1974). A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19(6), 716-723.
2. Ales, J. D., Delleur, J. W., Yevjevich, V. and Lane, W. L. (1980) Applied modeling of hydrologic time series. Littleton, CO: Water Resources Publications.
3. Anish, C. M. and Majhi, B. (2015). An ensemble model for net asset value prediction. 2015 IEEE Power, Communication and Information Technology Conference (PCITC) (pp. 392-396). New York, NY: IEEE.
4. Bagheri, A., Peyhani, H. M. and Akbari, M. (2014). Financial forecasting using ANFIS networks with quantum-behaved particle swarm optimization. Expert Systems with Applications, 41(14), 6235-6250.
5. Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. (1994). Time series analysis: Forecasting and control. New York, NY: Wiley.
6. Box, G. E. P. and Jenkins, G. M. (1970). Time series analysis: Forecasting and control. San Francisco, CA: Holden-Day.
7. Contreras, J., Espinola, R., Nogales, F. J. and Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. IEEE Transactions on Power Systems, 18(3), 1014-1020.
8. Cryer, J. D. and Chan, K. S. (2008). Time series analysis: With applications in R. Berlin, Germany: Springer.
9. Dase, R. K. and Pawar, D. D. (2010). Application of Artificial Neural Network for stock market predictions: A review of literature. International Journal of Machine Intelligence, 2(2), 14-17.
10. Datta, K. (2011). ARIMA forecasting of inflation in the Bangladesh economy. IUP Journal of Bank Management, 10(4), 7.
11. Dutta, A., Bandopadhyay, G. and Sengupta, S. (2015). Prediction of stock performance in the Indian stock market using logistic regression. International Journal of Business and Information, 7(1), 105-132.
12. Hamilton, J. D. (1994). State-space models. Handbook of econometrics, 4, 3039-3080. New York, NY: Elsevier.
13. Hamilton, J. D. (1994). Time series analysis, 2. Princeton, NJ: Princeton University Press.
14. Hyndman, R. J. and Athanasopoulos, G. (2014). Forecasting: Principles and practice. Melbourne, Australia: OTexts.
15. Kumar, M. and Anand, M. (2014). An application of time series ARIMA forecasting model for predicting sugarcane production in India. Studies in Business and Economics, 9(1), 81-94.
16. Nason, G. P. (2006). Stationary and non-stationary time series. In H. Mader & S. C. Coles (Eds.), Statistics in volcanology (pp. 129-142). London, UK: The Geological Society.
17. Ofori, T., Ackah, B. and Ephraim, L. (2013). Statistical models for forecasting road accident injuries in Ghana. International Journal of Modern Mathematical Sciences, 5(2), 99-115.
18. Pankratz, A. (2009). Forecasting with univariate Box-Jenkins models: Concepts and cases. New York, NY: Wiley.
19. Sakamoto, Y., Ishiguro, M. and Kitagawa, G. (1986). Akaike information criterion statistics. Dordrecht, The Netherlands: D. Reidel.
20. Tadawul (2016). Annual statistical report, 2016. Available at: https://www.tadawul.com.sa/wps/wcm/connect/bd537307-975c-462d-8700-a94f3ee05f18/Saudi+Stock+Exchange+%28Tadawul%29%2CStatistical+Report+%E2%80%93+English+2016.pdf?MOD=AJPERES&CONVERT_TO=url&CACHEID=bd537307-975c-462d-8700-a94f3ee05f18 [last accessed 10 March 2018].
21. Uko, A. K. and Nkoro, E. (2012). Inflation forecasts with ARIMA, vector autoregressive and error correction models in Nigeria. European Journal of Economics, Finance & Administrative Science, 50, 71-87.
22. Meyler, A., Kenny, G., & Quinn, T. (1998). Forecasting Irish inflation using ARIMA models.
23. Nason, G. P. (2006). Stationary and non-stationary time series. Statistics in Volcanology. Special Publications of IAVCEI, 1, 000-000.
24. Malkiel, B. G. (2003). The efficient market hypothesis and its critics. Journal of economic perspectives, 17(1), 59-82
25. Chatfield, C. (2016). The analysis of time series: an introduction. CRC press.
26. Doguwa, S. I., & Alade, S. O. (2013). Short-term inflation forecasting models for Nigeria. CBN Journal of Applied Statistics, 4(3), 1-29.