P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

Time-series Models - Forecasting Performance inthe Stock Market

Prof. Tinni Chaudhuri*, Dr. Abhijit Pandit#,
*Assistant Professor (Grade I), Department of Statistics, Amity Institute of Applied Sciences, Amity University, Kolkata, Email: tinnistat1991@gmail.com
#Assistant Professor (Grade II), Amity University, Kolkata,
Email: abhijitpandit1978@gmail.com

Abstract
Contradicting evidence on time-series and financial analysts' forecasting performance calls for further research in financial markets. Motivation to use time-series models rather than analysts' forecasts stems from recent research that reports time-series predictions to be superior to analysts' forecasts in predicting earnings for longer periods and for small firms that are hardly followed by financial analysts. The paper aims to explore performance of time series models in forecasting earnings for six firms considering historical data of 11 years from January, 2010 to December, 2020. Monthly average stock data of last 11 years for five firms namely HCL, TCS, Infosys, Reliance, Tech Mahindra and Wipro was considered from NSE site. Every company had 132 values whose graphical plotting and stationarity check was performed. Data series for each of the five companies was found to be non-stationary. After differencing each of them, the series became stationary and graphical plotting was again done. Then best suited ARIMA Model for each stationary time series was determined upon comparison of goodness of fit statistics. After choosing the best suited ARIMA model, residuals were extracted and were found to be random with no external influence whatsoever. Hence forecasting was done based on chosen model for the monthly average stock price of these top six companies of India in 2020. The paper finds that premier ARIMA family models outperform naive time-series models in terms of mean percentage errors, AIC and average ranks. The findings suggest that investors use the selected ARIMA model to form their expectations.

Keywords : ARIMA, time-series, forecasting, stock, financial market

INTRODUCTION

Financial Time Series Data (TSD) mining provides useful information for the investors, banks and insurance companies to channel their funds properly for better returns. This decision making is the primary motivation for prediction of financial TSD. Accurate multi-step ahead prediction of financial TSD becomes more difficult as the financial TSD is highly volatile. If prediction is one-step ahead preserving data trend is irrelevant. However, as the forecast horizon increases, preserving data trend becomes significant. In either case, prediction accuracy should remain high. Hence, the two basic requirements for multi-step ahead prediction model are maintaining high prediction accuracy and preserving the data trend across the prediction horizon.

Traditional models cannot meet both the requirements simultaneously. However, a hybrid model may provide scope for preserving data trend across the forecast horizon while maintaining good prediction accuracy, which motivates the research work of this paper. Most of the traditional models such as ARIMA, GARCH and ANN are applied for one-step ahead forecasting in many works of the

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

literature, where prediction accuracy is of major concern. However, in the present paper, we target multi-step ahead prediction, which requires preserving data trend in addition to high prediction accuracy. Such a model should account for the nature of TSD at every stage in the model.

The paper is organized in the following sections: First we present a literature survey of different prediction models existing in TSD. Next, we discuss the research methodology and the proposed ARIMA model is detailed in subsequent section. A quantitative analysis of the proposed model is discussed henceforth. The proposed and the traditional models are applied on selected NSE India data and the performance is compared. The paper ends with conclusion.

LITERATURE REVIEW

There are many kinds of research works in the area of forecasting using time series analysis. Some of the important tasks are mentioned here. A study deals with the implication of support vector machines (SVMs) regression, a novel neural network technique, in predicting the share price to examine the feasibility of SVM regression in predicting stock price. A data set related to Shanghai Stock Exchange in China has been used to test the validity of SVMs regression. The experiment depicts SVMs regression as a valuable method in forecasting the stock price (Bao et al., 2004; Pinto et al., 2020; Kumar et al., 2020). Again, a study focused on forecasting the price of Infosys Technologies, taking into consideration the previous open, close, high, and low price using different neural classifier functions like Least Mean Square, Multilayer Perceptron, Pace Regression, Linear Regression, Gaussian Processes, Simple Linear Regression, Isotonic Regression, and SMO Regression (Sureshkumar & Elango, 2011; Meher et al., 2020). Besides, a study examines the relative predictive power of ARIMA, VAR, and ECM models in predicting inflation in Nigeria. In doing this, a domestic Consumer Price Index (CPI) was lumped into the headline (all-item). Annual data from 1970 to 2010 were used.

The study examines the performance of the forecasting ability of the models. It was observed that different models performed well in different periods. While ARIMA is useful as a benchmark model, VAR for short-term forecasting and ECM are suitable for long-run forecasting (Uko & Nkoro, 2012; Bolar et al., 2017). Furthermore, in a study, the authors reviewed some of the approaches, which could be used for a stock market forecast like Hidden Markov Model, Non- linear Regression Analysis, Naive Bayes Classifier, Artificial Neural Networks, Decision Trees Classifier, Support Vector Machines, Random Forest Method, PCA (Principal Component Analysis), WB-CNN (Word Embeddings Input and Convolutional Neural Network prediction model) and CNN (Convolutional Neural Network) and finally concluded that neural network showed better results compared to other methods (Sharma & Kaushik, 2018).

RESEARCH METHODOLOGY

Any time series analysis has the following major steps in methodology in Box Jenkins approach:

a) Stationary check

b) Performing Augmented Dickey Fuller (ADF) Test / Unit Root test. If p-value of the test is less than level of significance(α), we reject null hypothesis of presence of unit root which means series is stationary.

c) Plotting the ACF and PACF of stationary series in order to determine the parameters of ARIMA model.

d) Fitting all possible combinations of ARIMA model by changing the values of parameter p and q in ARIMA.

e) Choosing the best ARIMA model with minimum value of AIC, since AIC is a measure of goodness of model fitting. Smaller the value of AIC, better is the fit.

f) Carrying out "Residual Analysis" of the selected model. Ljung -Box test is performed in this regard, where if p- value greater than level of significance(α) we conclude that the model fits the TSD well.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

g) Forecasting is done based on the chosen ARIMA model for the subsequent years.
There are three important parameters in the ARIMA(p,d,q) model, which are the auto-regressive order p, the difference order d and the moving average q. In the front, after one difference to determine the smooth of the sequence, we can get d=1 Next, it is necessary to determine the value of parameter p and parameter q in order to determine the final model ARIMA (p, d, q). By analyzing the auto-correlation and partial correlation graphs of time series $\{Y_t\}$ the value of p and q can be determined.

RESULTS AND DISCUSSION

The original time series data of six companies (2010-2020) are shown below:

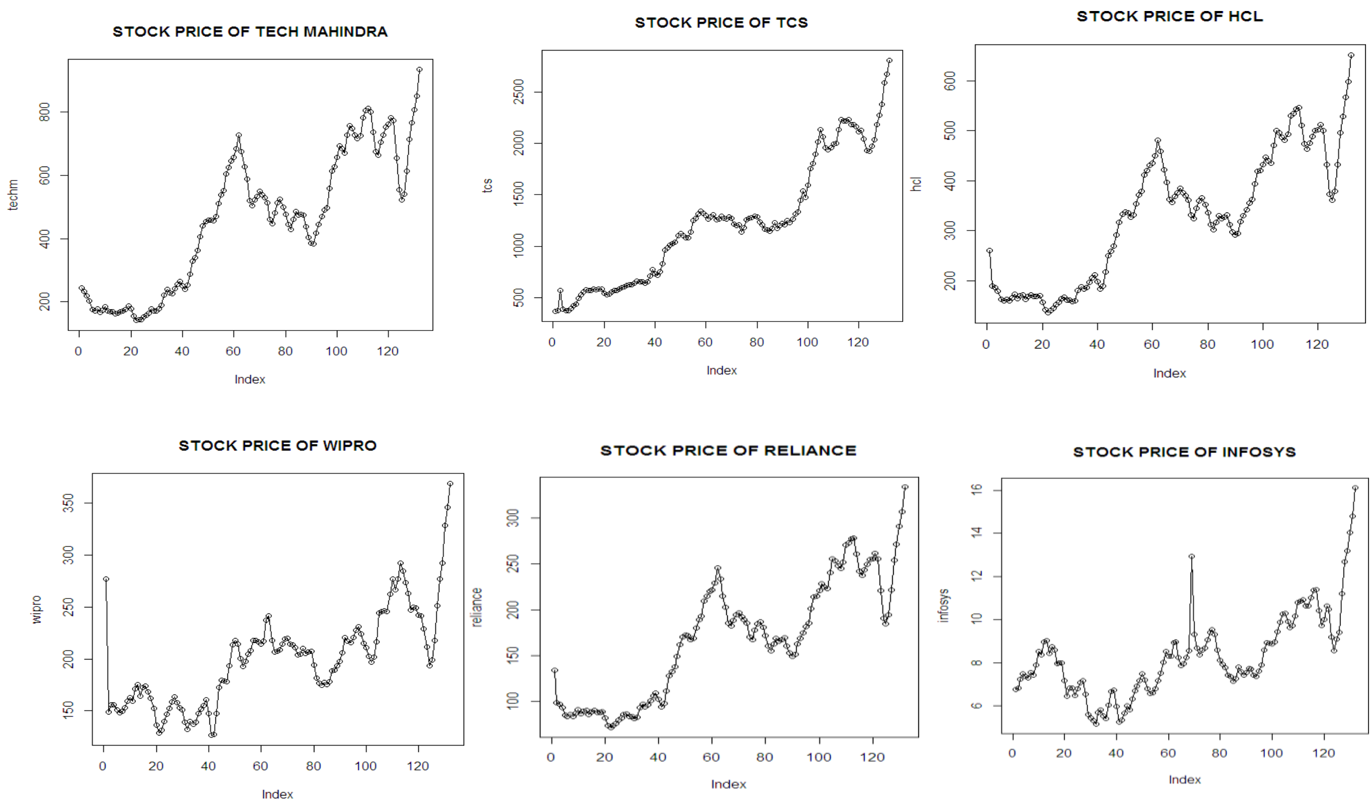


Fig. 1 Graphs of stock price of six companies (2010-2020)

It is clear from the above graphs that all the series are *non-stationary*. We difference each of the series appropriate number of times till stationarity is achieved. We carry out the Augmented Dickey Fuller / Unit root test on the differenced series to be certain that they are converted to stationary series.

Unit root test/ADF Test

We want to test whether each of the series is stationary under this test. The null hypothesis is presence of unit root i.e. non- stationary.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

COMPANY	**TEST STATISTIC VALUE**	**P-VALUE**	**DECISION**
TECHMAHINDRA	-3.733	0.02447	STATIONARY
TCS	4.0532	0.01	STATIONARY
HCL	-3.4255	0.04341	STATIONARY
WIPRO	-3.3564	0.01	STATIONARY
RELIANCE	-7.6845	0.01	STATIONARY
INFOSYS	-9.058	0.01	STATIONARY

TABLE 1. Unit Root Test

As it can be seen from Table 1, the p-value corresponding to each series is less than level of significance (0.05) making accept the alternative hypothesis implying that all the differenced series are stationary.

We have also check the stationarity by ACF and PACF plots of each series.

Model Fitting

The cut offs (spikes) from the ACF and PACF function of the stationary series we can take various combinations of p and q to fit ARIMA (p,d,q). Here d is 1, as we difference each of the series once to make it stationary (for six companies HCL, TCS, Infosys, Reliance, Tech Mahindra and Wipro). We choose the best model looking at the value of Akaike's Information Criteria (AIC). Lesser the value of AIC, better is the fit.The following Table 2 gives the best fitted ARIMA model for each state along with the estimated parameters and some diagnostic measures.

Company	Best fitted ARIMA model	Estimated parameters along with its standard error (s.e)	AIC	Normalized BIC
Techmahindra	(1,1,1)	ar1 ma1 0.2031 0.788 s.e.0.1100 0.0836	1197.85	1206.47
TCS	(1,1,0)	ar1 drift 0.3000 18.8968 s.*e. 0.0843 6.8406*	1427.65	1436.27

Journal of Contemporary Issues in Business and Government Vol. 27, No. 2,2021
https://cibg.org.au/

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

HCL	(1,1,1)	ar1 ma1 drift 0.2201 0.8065 2.7505 s.e.0.1127 0.0736 2.9955	1079.38	1090.88
WIPRO	(0,1,1)	ma1 0.6877 s.e. 0.1500	1075.83	1081.58
RELIANCE	(1,2,3)	ar1 ma1 ma2 ma3 -0.4911 -1.1138 -0.7511 0.8787 s.e. 0.1057 0.0615 0.1019 0.0592	931.76	1030.61
INFOSYS	(1,2,3)	ar1 ma1 ma2 ma3 -0.0617 -2.8402 2.6887 -0.8484 s.e. 0.0946 Na Na Na	310.39	870.23

Table 2: Best fitted ARIMA model with parameter estimation and diagnostic values

Diagnostic Check

1) *Ljung Box test for residuals:* Here given at 5% level of significance, if p-value is less than 0.05 we reject Ho implying that the SSE is large (or fit is not good) , otherwise we accept Ho i.e. SSE is small or fit is good. In the following Table 4 we have given the value of Ljung-Box test statistic along with its corresponding p-value for each series for the chosen ARIMA model.

Company	Ljung Box test statistic	P-value
TECHMAHINDRA	37.736	0.899
TCS	28.881	0.9928
HCL	40.931	0.8161
WIPRO	46.767	0.6039
RELIANCE	48.448	0.5358
INFOSYS	28.707	0.9933

Table 3. Ljung-Box Test

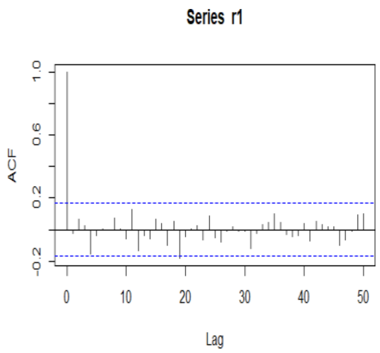
Journal of Contemporary Issues in Business and Government Vol. 27, No. 2,2021
https://cibg.org.au/

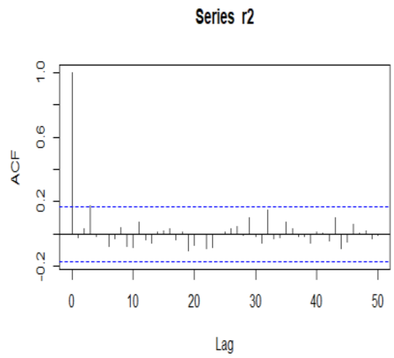
P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

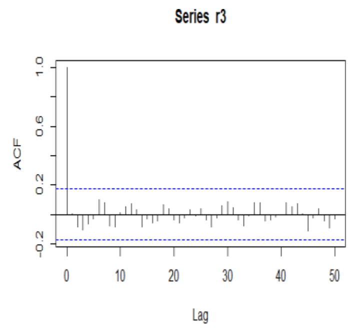
We can see from Table.3, Ho is accepted in all the cases which mean that the models we have chosen for each company stock is appropriate.

2) *Residual Acf:* As we know a primary measure of goodness of fit is that errors of the fitted models should be uncorrelated. This is a desired property that ensures that the model fitted is appropriate. If the spikes of residual ACF are on both sides of the horizontal axis in a random pattern and gradually decreases to zero, we may call it an ideal situation. Hence, eventually as the lag increases the spikes get within the control band. We can see from the residual ACFs plotted in Fig.2 that all the ARIMA models fitted are proper.

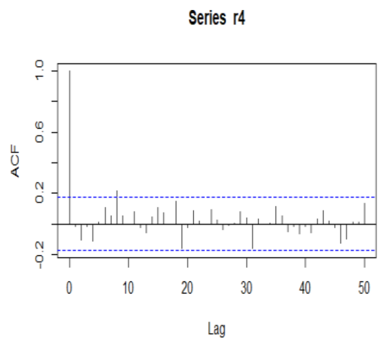
1. Techmahindra 2. TCS 3. Infosys

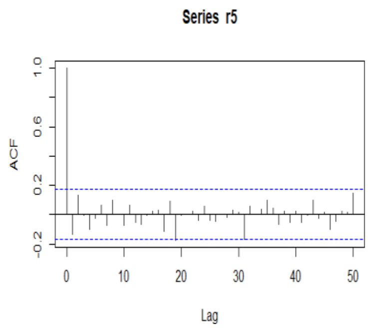






4. Reliance 5. HCL 6. Wipro





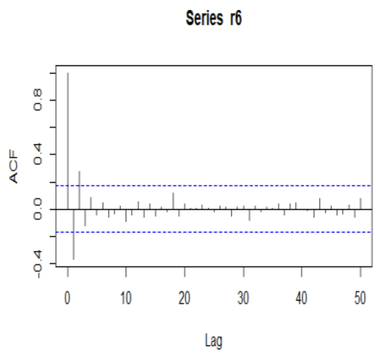


Fig. 2 Residual ACFs.

Forecasting

The forecasted values of monthly average stock data of 6 companies till September 2021 is given in Table 4 for the 6 companies' share price India along with their upper and lower control limits.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

Company		*Jan,2021*	*Feb,2021*	*Mar,2021*	Apl,2021	May,2021	June,2021	July,2021	Aug,2021	Sept,2021
TECHMAHINDRA	U	1070.031	1168.254	1240.597	1293.275	1336.740	1374.604	1408.597	1439.708	1468.566
	F	1006.712	1028.233	1040.146	1040.146	1040.146	1040.146	1040.146	1040.146	1040.146
	L	943.3928	888.2128	839.6950	787.0166	743.5514	705.6879	671.6950	640.5837	611.7258
TCS	U	3013.246	3142.083	3243.376	3329.296	3405.712	3475.693	3540.994	3602.703	3661.543
	F	2857.744	2887.036	2909.052	2928.884	2948.062	2967.043	2985.965	3004.869	3023.768
	L	2702.242	2631.989	2574.728	2528.473	2490.412	2458.393	2430.936	2407.036	2385.994
WIPRO	U	426.4442	465.5598	491.0237	511.4293	528.9557	544.5593	558.7610	571.8822	584.1374
	F	385.7729	385.7729	385.7729	385.7729	385.7729	385.7729	385.7729	385.7729	385.7729
	L	345.10165	305.986	280.5221	260.1165	242.5901	226.9866	212.7848	199.6636	187.4084
HCL	U	743.8218	809.1198	851.7427	884.3130	911.4847	935.2750	956.7053	976.367	994.6377
	F	701.7547	713.6428	716.4185	717.0665	717.2178	717.2531	717.2614	717.2633	717.2639
	L	659.6877	618.1638	581.0943	549.8200	522.9509	499.2313	477.8174	458.1597	439.8898
RELIANCE	U	95.224	82.12	90.32	94.123	99.892	101.143	97.435	98.98	102.34
	F	84.8374	64.7307	75.1777	70.6189	73.4298	72.6212	73.5903	73.6863	74.2111
	L	67.123	52.109	65.45	64.543	67.875	65.233	67.321	66.329	68.432
INFOSYS	U	8.345	8.432	8.123	8.56	8.43	8.234	8.321	8.567	8.543
	F	6.49458	6.844	6.17259	6.0206	6.11433	6.08737	6.11968	6.122877	6.14037
	L	5.567	5.654	5.321	5.432	5.764	5.321	5.234	5.321	5.354

Table 4: Forecasts with control limits

U: UCL F: Forecast L: LCL

We plot the forecasted values company-wise from Jan 2021 – Sept 2021

CONCLUSION

Stock and financial markets tend to be unpredictable and even illogical due to the unevenness of market volatility. Because of such characteristics financial or market data should possess a rather turbulent structure which often makes it hard to find reliable patterns. Modeling such structures requires methods and algorithms capable of finding hidden pattern or trend within data and predict how they will affect them in the future. One of the most common and efficient methodology is ARIMA modeling of TSD. Autoregressive Integrated Moving Average (ARIMA) Model converts non-stationary data to stationary data before working and is one of the most popular models to predict nearly linear time series data. ARIMA model has been used extensively in the field of finance and economics as it is known to be robust, efficient and has a strong potential for short-term share market prediction. In this work we have forecasted the share prices of 6 major companies of India in 2020 for the next 9 months and as we have already observed the prediction based on our chosen ARIMA model is fairly consistent. We see a significant rise in the share price of TCS in the near future, which calls for special mention. Another factor to pay attention is we do not go for a very long term prediction of such volatile data as several socio-economic factors are involved which play a major role in the rise and fall of market share prices. Though the companies that we have considered in this study i.e, Reliance, HCL , Tech Mahindra, Infosys, Wipro and TCS do show a steady forecast in the upcoming few years, based on their data till 2020, but investors should not discard the influence of other external factors before making their decision.

References:

[1] G.E.P. Box, G. Jenkins, Time Series Analysis, Forecasting and Control, revised ed. 1976 Edition, Holden- Day, Incorporated, 1990.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

[2] G.L. Griepentrog, J.D. Cummins, Forecasting automobile insurance paid claims using econometric and ARIMA models, Int. J. Forecast. 1 (1985) 203–215.

[3] R. Genesio, S. Pozzi, A. Vicino, U. Di Caprio, Short term load forecasting in electric power systems: a comparison of ARMA models and extended wiener filtering, J. Forecast. 2 (1983) 59–76.

[4] R.E. Spudeck, Scott E. Hein, Forecasting the daily federal funds rate, Int. J. Forecast. 4 (1988)581–591.

[5] C. Babu, B. Reddy, Predictive data mining on average global temperature using variants of ARIMA models, in: 2012 International Conference on Advances in Engineering, Science and Management (ICAESM), 2012, pp. 256–260.

[6] X. Wang, Y. Liu, ARIMA time series application to employment forecasting, in: 4th International Conference on Computer Science Education, 2009, ICCSE '09, 2009, pp. 1124–1127. http://dx.doi.org/10. 1109/ICCSE.2009.5228480.

[7] R.F. Engle, Autoregressive conditional heteroscedasticity with estimates of the variance of the united kingdom inflation, Econometrica 50 (1982) 987–1008.

[8] R.F. Engle, D.B. Nelson, T. Bollerslev, ARCH models, in: R.F. Engle, D.L. McFadden (Eds.), Ch. Handbook of Econometrics, vol. IV, 1994, pp. 2959–3038.

[9] T. Bollerslev, R.Y. Chou, K.F. Kroner, Arch modeling in finance: a review of the theory and empirical evidence, J. Econom. 52 (1–2) (1992) 5–59.

[10]L. Bianchi, J. Jarrett, R.C. Hanumara, Improving forecasting for telemarketing centers by {ARIMA} modeling with intervention, Int. J. Forecast. 14 (4) (1998) 497 –504, http://dx.doi.org/10.1016/S0169-2070(98)00037-5.

[11]H. Ghijsels, P.H. Franses, Additive outliers, GARCH and forecasting volatility, Int. J. Forecast. 15 (1999) 1– 9.

[12]R. Garcia, J. Contreras, M. van Akkeren, J. Garcia, A GARCH forecasting model to predict day- ahead electricity prices, IEEE Trans. Power Syst. 20 (2) (2005) 867 –874, http://dx.doi.org/10.1109/ TPWRS.2005.846044.

[13] E. Scott, Ken Johnston, GARCH models and the stochastic process underlying exchange rate price change, J. Financ. Strateg. Decis. 13 (2000) 13–24.

[14]F. Yusof, I. Kane, Volatility modeling of rainfall time series, Theoret. Appl. Climatol. 113 (1–2) (2013) 247– 258, http://dx.doi.org/10.1007/s00704-012-0778-8.

[15]C.-L. Hor, S. Watson, S. Majithia, Daily load forecasting and maximum demand estimation using ARIMA and GARCH, in: International Conference on Probabilistic Methods Applied to Power Systems, 2006, PMAPS 2006, 2006, pp. 1–6.

[16] B. Zhou, D. He, Z. Sun, Traffic modeling and prediction using ARIMA/GARCH model, in: A. Nejat Ince, E. Topuz (Eds.), Modeling and Simulation Tools for Emerging Telecommunication Networks, Springer, US, 2006, pp. 101–121.

[17]C.N. Babu, B.E. Reddy, A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data, Appl. Soft Comput. 23 (0) (2014) 27–38, http://dx.doi.org/10.1016/j.asoc.2014.05.028.

[18]M. O'Mahony, Tigran T. Tchrakian, Biswajit Basu, Real-time traffic flow forecasting using spectral analysis, IEEE Trans. Intell. Transport. Syst. 13 (2012) 519–526.

[19]M. Huang, Y. He, H. Cen, Predictive analysis on electric-power supply and demand in China, Renew. Energy 32 (7) (2007) 1165–1174, http://dx.doi.org/10.1016/j.renene.2006.04.005.

[20]P. Fryzlewicz, Lecture Notes: Financial Time Series, ARCH and GARCH Models, University of Bristol, 2007.

[21]Robert F. Engle III, Risk and Volatility: Econometric Models and Financial Practise, New York University, 2003.

[22]M.S. Heracleous, Volatility Modeling Using the Student-t Distribution, Ph.D. Thesis, Virginia Polytechnic Institute and State University, 2003.

[23]J.-J. Tseng, S.-P. Li, Quantifying volatility clustering in financial time series, Int. Rev. Financ. Anal. 23 (0) (2012) 11–19, http://dx.doi.org/10.1016/j.irfa.2011.06.017 (Complexity and Non-Linearities in Financial Markets: Perspectives from Econophysics).

[24]J.S. Armstrong, Principles of Forecasting: A Handbook for Researchers and Practitioners, Kluwer Academic Publishers, MA, 2001.

P-ISSN: 2204-1990; E-ISSN: 1323-6903
DOI: 10.47750/cibg.2021.27.02.387

[25] B.P. Lathi, Signal Processing and Linear Systems, Oxford University Press, 2000.

[26] T. Bollerslev, R.F. Engle, Common persistence in conditional variances, Econometrica 61 (1) (1993) 167–186.

[27] Bao, Y., Lu, Y., & Zhang, J. (2004). Forecasting Stock Price by SVMs Regression Yukun. In Lecture Notes in Artificial Intelligence (Subseries of Lecture Notes in Computer Science) (pp. 295 -303). https://doi.org/10.1007/978-3- 540-30106-6_30

[28] Sureshkumar, K. K., & Elango, N. M. (2011). An Efficient Approach to Forecast Indian Stock Market Price and their Performance Analysis. International Journal of Computer Applications, 34(5), 44-49. Retrieved from https://www. semanticscholar.org/paper/An- Efficient-Approach-to-Forecast-Indian-Stock-and-Sureshkumar-El ango/5fe8b5c6126aa79fb58637b0d 375fd4d4d4c6a6e

[29] Uko, A. K., & Nkoro, E. (2012). Inflation forecasts with ARIMA, vector autoregressive and errorcorrection models in Nigeria. European Journal of Economics, Finance and Administrative Sciences, 50, 71.Retrieved from https://www.researchgate.net/publication/288453968_Inflation_forecasts_with_ARIMA_vector_autoregressive_and_error_correc- tion_models_in_Nigeria

[30] Sharma, S., & Kaushik, B. (2018). Review Article Quantitative Analysis of Stock Market Prediction for Accurate Investment Decisions in Future. Journal of Artificial Intelligence, 11(1), 48 -54. https://doi. org/10.3923/jai.2018.48.54

[31]