A New Financial Time Series Approach for Volatility Forecasting

R.M. Kapila Tharanga Rathnayaka
Faculty of Applied Sciences,
Sabaragamuwa University of Sri Lanka
Belihuloya, Sri Lanka.
kapiar@apps.sab.ac.lk

D.M.K.N Seneiratna
Faulty of Engineering
University of Ruhuna
Galle, Sri Lanka
seneiratne@is.ruh.ac.lk

H.I. Arumawadu
School of Computer Science
Wuhan University of Technology
Wuhan, P.R. China
Hasitha87@gmail.com

***Abstract*— The investment in capital market is easiest, fastest and securable way for building healthy financial foundation today. Because of the economic outlooks causing directly on these market fluctuations, the making decisions in the equity market has been regarding as one of the biggest challenges in the modern economy. The main purpose of this study is to take an attempt to understand the behavioral patterns and seek to develop a new hybrid forecasting approach under the volatility. The results are successfully implemented on Colombo stock exchange (CSE), Sri Lanka over the three year period from January 2013 to December 2015. The empirical results indicated that the new proposed hybrid approach is more suitable for forecasting price indices than traditional time series forecasting methodologies under the high volatility.**

Keywords; ANN, ARIMA, ARIMA-ANN and Volatility

I. INTRODUCTION

The investment in capital market has becoming more institutionalized during the past few decades and arranging large investment funds to the general public. Day by day investors have been accessing to capital markets and exchanging vast number of financial products.

The successful prediction of a stock price could yield a significant profit for investors. Numerous types of methodologies can be found in the literature to estimate the market predictions. Among them, the time series analysis is an essential methodology which comprises the tools for analyzing the time series data to identify the meaningful characteristics for making future ad-judgements; especially decision makings' should be taken in economic and finance to predict the future patterns under the numerous type of mathematical and economic assumptions or between both with computing assumptions. Based on the behaviors, all these methodologies can be summarized under the two categories as frequently domain and time domain. Furthermore, the methodologies under the time domain can be divided into another two sub-categories namely parametric and non-parametric [1]. Basically, the parametric approaches have been developed under the basis on the stationary stochastic process that can be described by using the very limited number of parameters. According to the literature, most of the previous studies carried out under the parametric assumptions. For an example, the well-balanced statistical assumptions with Box-Jenkins methodology, an autoregressive moving average (ARMA) and it's generalization models of autoregressive integrated moving average (ARIMA) have been widely used for forecasting various type of indices in finance[2][3].

However, these traditional approaches are suitable and appropriated just only for empirical data studies under the normality, linearity and stationary conditions. As a result of these complications, neural network computing model with new hybrid methodology was proposed by McCulloch and Pitts to handle incomplete, noise and uncertain data in the multidisciplinary systems [4]. Because of the flexible nonlinear modelling capability, this novel concept was popular and has been successfully applied to various systems such as financial, economic, military, geological and agricultural systems for signal processing, pattern recognition, classification, time series forecasting and etc.; especially, artificial neural network (ANN) based hybrid methodologies are more suitable for forecasting stock market predictions under the non-linear high volatility [5] [6] and [7].

Overall, the existence of nonlinear relationships with respect to the stock returns and trading volume is widely documented in the literature. However, in the Sri Lankan context, very limited studies can be seen relates to this scenario. The objective of this study is to introduce new forecasting mechanism to predict

the future behaviors under the modern social and economic conditions; especial for the financial data analysis [8].

As a developing market, Colombo stock exchange has been achieving the massive financial development in the fast two decades. Currently, it is one of the most modernized stock exchange in South Asia providing a fully automated trading platform for their locals as well as foreign customers. As a case study, this study is carried out on the basis of secondary data, which were obtained from Colombo Stock Exchange official database during the three year period from January 2013 to December 2015 were extracted and tabulated for calculations [9].

The remainder of this paper is organized as follows. Section 2 briefly explains the theoretical background of the traditional forecasting approaches with ANN methodology. Indeed, the new proposed hybrid methodology explains methodology describe under the subsections. Section 3 gives the illustrations. Section4 is discussion and section 5 end up with conclusions, policy issues and future works with possible extensions.

II. METHODOLOGY

The methodology of the study can describe under the three phases. In the first phase, stock market validations are identified based on traditional time series approaches such as exponential smoothing and auto regressive moving average. In the second part, new proposed combined approach of Artificial Neural Network with ARIMA (ANN-ARIMA) methodologies are used. Finally, testing accuracy techniques will apply to find the suitable model for forecasting time series data in CSE under the high volatility.

A. The Artificial Neural Network (ANN) Approach for Time Series Modelling

The ANN algorithms are universal and highly flexible approximates that have been widely used to identify the complex relationships between inputs and outputs. Because of the less sensitivity for error term assumptions, high tolerate noises, robustness and heavy tails, ANN algorithms are more suitable for mapping non-linear data patterns than others. The proposed methodology is briefly discussed under eight-steps as follows [10][11].

Step 1: Variable Selection

Step 2: Data collection

Step 3: Data preprocessing

Step 4: Training, testing and validations

Step 5: Define Network paradigms

(Hidden layers, Hidden neurons, Output neurons)

Step 6: Evaluation

Step 7: Training (Number of iterations and learning rate)

Step 8: Validate the network for post-training analysis

The new proposed network architectural model in the current study consists of single hidden layer fully connected feedforward network include single input layer, hidden layer and output layer [12].

B. The Hybrid (ANN-ARIMA) Methodology for forecasting

As a result of high volatility and unstable patterns, the traditional forecasting approaches haven't achieved successes in both linear and non-linear domains. So, combined methodologies under the linear autocorrelation structure and non-linear weighted average component have created high accuracy forecasting than single model approaches [13][14].

$$Y_t = L_t + N_t \quad (1)$$

Where; L_t and N_t denote the linear autocorrelation and non-linear component of the time series pattern Y_t respectively. So, the new proposes hybrid methodology can be described under the two phases based on their linear and non-linear behaviors. As a next stage, residual from the linear models will used to capture the nonlinearity. The residuals of the linear component can be defined as follows.

$$e_t = Y_t - \widehat{L_t} \quad (2)$$

Where, e_t denotes the residual of linear model and $\widehat{L_t}$ presents the forecast value for the estimated time series models at time$t$. If we can see any non-linear significant pattern in residuals, as a next step, ANN modeling approach can be applying to discover the non-linear relationships.

$$e_t = f(e_{t-1}, e_{t-2}, e_{t-3} \cdots, e_{t-n}) + \varepsilon_t \quad (3)$$

Where n represent the input nodes and f is the non-linear function which determined based on ANN approach. However, if the non-linear model is not an appropriate, it means that, the error term ε_t is not necessarily random.

$$\widehat{y_t} = \widehat{L_t} + \widehat{N_t} \quad (4)$$

Based on these assumptions, neural network can be equivalent as a nonlinear autoregressive model [16].

III. Empirical Results

A. Case Study: Colombo Stock Exchange

The study was carried out on the basis of secondary data, which were obtained from Colombo Stock Exchange official database, Central Bank of Sri Lanka Financial reports, different types of background readings, other relevant sources and etc. Two principal price indices namely ASPI (All Share Price Index) and SL 20 (S&P Sri Lanka 20 Index) daily trading 802 data observations from January 2013 to December 2015 were extracted and tabulated for calculations. To assets the forecasting accuracies, the composition for the data set is divided in to two sections as training (882) and testing (176).

B. Stationary/Non Stationary Model Checking

As an initial step, stationary and non-stationary conditions were measured using two unit root approaches namely Augmented Dickey-Fuller test (ADF) and Phillips-Perron test statistic (PP). The Table 1 results suggested that first difference level data significantly stationary under the 0.05 level of significance. As a result, ARIMA is more significant and applicable for predicting future results.

TABLE I. UNIT ROOT TEST RESULTS

Price Index	Significant Result		Price Index	Significant Result	
	Level Data			Level Data	
	ADF	PP		ADF	PP
ASPI	0.5874	0.5836	ASPI	0.0000	0.0000
SL 20	0.3934	0.3760	SL 20	0.0000	0.0000

Note: All the tests assume asymptotic normality.

As a next step, minimum values of Akaike info criterion (AIC), Schwarz criterion (SC) and Hannan-Quinn criterion (HQC) were used to select the suitable ARIMA approach. The results suggested that, ARIMA (2, 1, 3) (AIC (10.03684), SC (10.06618), and HQC (10.04811)) and ARIMA (2, 1, 1) (AIC (9.373996), SC (9.397023), and HQC (9.382984)) are most suitable for predicting future patterns of ASPI and SL20 respectively.

In the next stage, residual of the selected models (ARIMA) were evaluated with respect to the actual indices. Then, the proposed ANN-ARIMA hybrid methodologies applied to forecast non-linear composite in the price indices based on MATLAB training algorithms. The proposed model is a three layer back forward network which includes one input layer, one hidden layer and one output layer respectively. Furthermore, consecutive last 882 daily days and next coming day (one-step-ahead forecasting) were used as an inputs and output as respectively. To find best accuracy model, single hidden layer with 2, 3, 4, 5, 10 and 15 neural with one output layer used. The out of sample error results suggested that 2 neural with one hidden layer is more accurate than others.

Finally, the selected ARIMA and proposed ARIMA-ANN hybrid method results are summarized in Table II for the horizon of one day ahead (testing sample).

According to the error analysis results, new proposed ARIMA-ANN is highly accurate (less than 10%) with lowest RMSE error values. Moreover, MAD accuracy testing results also confirmed that proposed hybrid ARIMA-ANN algorithm is more significant than traditional ARIMA methods for forecasting financial time series predictions.

TABLE II. THE MODEL ACCURACY FOR ONE DAY FORECASTING

	Model Accuracy	Forecasting Accuracy (%)	
		ARIMA	ARIMA-ANN
ASPI	MAD (%)	3.75	2.29
	MSE	14.06	5.26
	MAPE	0.05	0.03
SL20	MAD (%)	4.71	3.25
	MSE	22.18	10.58
	MAPE	0.11	0.08

***denotes the model with the minimum error values**

The point-to-point comparisons between actual and forecasted predictions for the coming week results are given in Figure 01.

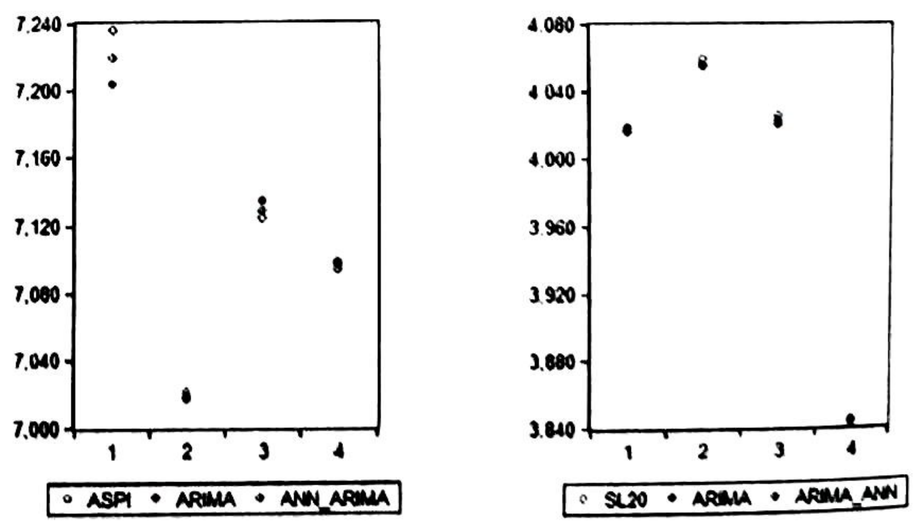

Figure 1. Forecasting Results for Coming week

IV. Conclusion

The Economical Time series models are widely used to develop the economic relationships, especially for the nonlinear models under the stationary and non-stationary

frameworks. Different type of research studies have been carried out to find the best forecasting methods to predict long and short term-predictions in the real-world applications; especially, after 2000 Artificial Intelligence based miscellaneous type of algorithms have been successfully applied for solving real-world problems.

In the current study, new proposed ARIMA based hybrid approaches were widely used to discuss our results. The mean absolute percentage error (MAPE) results reveal that (MAPE[ARIMA] > MAPE[ANN_ARIMA]), new proposed ANN_ARIMA model is more significant and gives best solution for predicting short term predictions in high volatility fluctuations than traditional forecasting approaches.

Finally, we strongly believed that, current study makes significant contribution to policy makers as well as government to open up new direction to develop the CSE investments in Sri Lanka.

REFERENCES

[1] Areekkul, P. &. T. H., 2009. Combination of Artificial Neural network and ARIMA Time Series Models for short term price forecasting in Deregulated Market. s.l., IEEE T&D Asia.

[2] Breckling, J., 1989. The Analysis of Directional Time Series: Applications to Wind Speed and Direction, s.l.: ser. Lecture Notes in Statistics. Berlin, Germany.

[3] C.S Ong, J. H. G. T., 2005. Model identification of ARIMA family using genetic algorithms. Applied Mathematics and Computation , 164(1), p. 885–912.

[4] Chen, C., 1994. Neural Networks for Financial Market Prediction. IEEE, 6(2), pp. 1199-1202.

[5] D. Dickey, W. F., 1979. Distribution of the Estimators for Autoregressive Time Series With a Unit Root.. Journal of the American Statistical Association , 74(1), pp. 427-431.

[6] Ho, S. & M. XIE, 1996. The Use of ARIMA Models For Reliability Forecasting and Analysis.. 23rd International Conference on Computers and Industrial Engineering. Great Britain: Elsevier Science Ltd, pp. 213-216.

[7] J.-J Wang, J.-Z. W. Z.-G. Z. &. S.-P. G., 2012. Stock index forecasting based on a hybrid model.. Omega, 40(1), pp. 758-766.

[8] Jayathileke, P. M. B. & R.M. K.T Ratnayaka, 2013. Testing the Link between Inflation and Economic Growth: Evidence from Asia. Communications and Network, 4(1), pp. 87-92.

[9] L. Xiangdong, S. L., 2009. Estimation and Forecasting of Dynamic Conditional Covariance: A semiparametric Multivariate Model. s.l.:s.n.

[10] Omar, A. & Jaffar, . M., 2011. Comparative Analysis of Geometric Brownian Mortion Model in forcasting FBMHS and FBMKLCI Index in Bursa Malaysia. Langkawi, Malaysia, IEEE symposium on Business , Engineering and industrial Applications.

[11] R. M. K.T Ratnayaka, Z. W. S. A. a. S. C., 2012. Enhanced greedy optimization algorithm with data ware housing for automated nurse scheduling system. E-Health Telecommunication Systems and Networks, 2(1), pp. 43-48.

[12] Rathnayaka, R., Jianguo, W. & Seneviratna, D., 2014. Geometric Brownian Motion with Ito's Lemma Approach to Evaluate Market Fluctuations: A case study on Colombo Stock Exchange. Shanghai, Behavior, Economic and Social Computing (BESC-IEEE).

[13] Rathnayaka, R. K. T., Seneviratne , D., Jianguo , W. & Indika, H., 2014. A Hybrid Statistical Approach for Stock Market Forecasting Based on Artificial Neural Network and ARIMA Time Series Models. Nanging, The 2ndInternational Conference on Behavioral, Economic and Socio-Cultural Computing.

[14] Seneviratne, D. & W. Jianguo, 2013. Analyzing the causal relationship between stock prices and selected macroeconomic variables: evidence from Sri Lanka. ZIJBEMR, 3(9), pp. 1-10.

[15] Zhang, G. P., 2003. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing, 50(1), p. 159 – 175.