THE ARTIFICIAL NEURAL NETWORK BASED HYBRID STATISTICAL APPROACH FOR VOLATILITY FORECASTING IN COLOMBO STOCK EXCHANGE

R.M Kapila Tharanga Rathnayaka*[1] and D.M.K.N Seneviratna[2]
[1]*Faculty of Applied Sciences, Sabaragamuwa University of Sri Lanka*
[2] *Faculty of Engineering, University of Ruhuna*
*Corresponding author *(email: kapila.tr@gmail.com)*

Introduction

The time series analysis is an essential methodology which comprises the methods for analyzing the time related data to identify the meaningful characteristics for making future ad-judgements. Especially decision making should be done in economic and finance to predict the future patterns under the numerous type of mathematical and economic assumptions [1]. For an example, the well-balanced statistical assumptions with Box-Jenkins methodology, an autoregressive moving average (ARMA) and its generalization models of autoregressive integrated moving average (ARIMA) have been widely used for forecasting finance indices today [1].

However, most of these traditional approaches are suitable and appropriate just only for empirical data studies under the normality, linearity and stationary assumptions [3]. As a result of these complications, "Neural network computing model" with new hybrid methodology was proposed by McCulloch and Pitts in 1943 to handle incomplete, noise and uncertain data in the multidisciplinary systems[4], [5]. Because of the flexible nonlinear modelling capability, this novel concept has been successfully applied in various systems for signal processing, pattern recognition, classification, time series forecasting and etc.[6], [7]. However, in the Asian context, very limited studies can be seen relating to this scenario. Among them, famous Chinese scholars namely Baoan Yang, Shinong Wu, Liang dipped into this area and conducted some remarkable studies [1], [7], [8].

This current study mainly focused and attempted to find out the suitable methodology for forecasting time series data under the high volatility with unstable patterns. So, the main objective of this study was to introduce new forecasting mechanism to predict the future behaviors. The results are to be implemented on Colombo stock exchange (CSE), Sri Lanka.

The paper is organized as follows. Under the materials and methods it briefly explains the theoretical background of the traditional forecasting approaches with ANN methodology. Indeed, the new proposed hybrid methodology is explained under subsections. The research paper ends up with conclusions, policy issues and future works with possible extensions.

Materials and Methods

The methodology of the study can be described under three phases. In the first phase, stock market validations are identified based on traditional time series approaches such as exponential smoothing and autoregressive moving average.

In the second part, new proposed two combined approaches are used for forecasting CSE price indices over the three year period from October 2013 to July 2016. They are; Artificial Neural Network (ANN) with ARIMA (ANN_ARIMA) and ANN with Geometric Brownian Motion (ANN_GBM).

Finally, testing accuracy techniques will be applied to find the suitable model for forecasting time series data in CSE under the high volatility.

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 [6], [7].

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 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 [8], [9].

Geometric Brownian Motion (GBM)

A GBM is a continuous-time stochastic process in which the logarithm of the randomly varying quantity follows a Brownian motion or wiener process with drift. It is an application of stochastic processes satisfying a stochastic differential equation; especially, it is used in mathematical finance to model stock prices in the Black–Scholes model.

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 [1,6 and 10]. So, combined methodologies under the linear autocorrelation structure

and non-linear weighted average component have created high accuracy forecasting than single model approaches.

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

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 behaviours. 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.

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

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.

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

The data were obtained from annual reports of Central Bank of Sri Lanka, the monthly trading reports from CSE, various types of background readings and 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 October 2013 to July 2016 were extracted and tabulated.

Results and Discussion

As an initial stage, 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.

Table1. The Model Accuracy for Coming Week

	Model Accuracy	**Forecasting Accuracy (%)**			
		ARIMA	**GBM**	**ARIMA-ANN**	**ARIMA-GBM**
ASPI	**MAD (%)**	16.08	10.91	6.42	2.11
	MAE	297.74	201.49	66.71	37.95
	RMSE	18.36	20.70	11.76	8.59
SL20	**MAD (%)**	15.86	14.95	9.42	5.47
	MAE	202.74	109.89	56.42	33.38

RMSE	16.51	15.05	9.17	8.83

In the same time, geometric Brownian motion (GBM) algorithm was applied to out-of-sample forecasting performance for the horizon of one day ahead (testing sample).

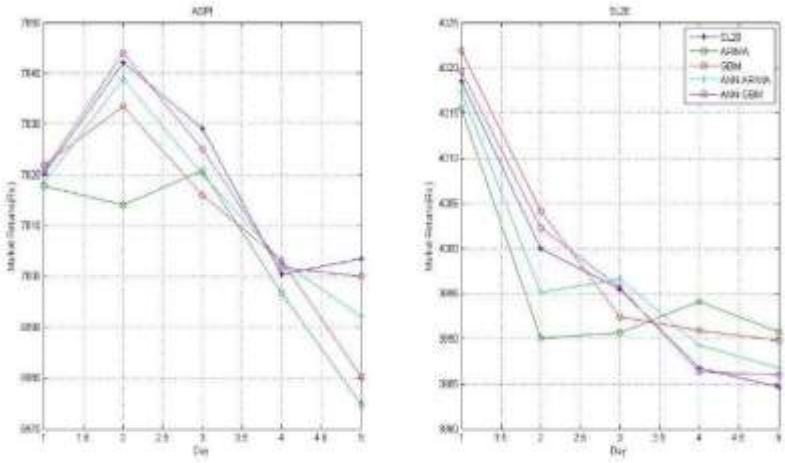

Figure 1. Coming week forecasting from 1st August to 5th August 2016

As a next stage, proposed ANN-ARIMA / ANN-GBM hybrid methodologies applied to forecast non-linear composite in the price indices based on MATLAB training algorithms. The corresponding results are summarized in Table I. According to the error analysis results, new proposed ARIMA-GBM is highly accurate (less than 10%) with lowest RMSE error values. Moreover, MAD accuracy testing results also confirmed that proposed hybrid algorithms (ARIMA-ANN, GBM-ANN) are more significant than traditional ARIMA methods for forecasting financial time series predictions. The point-to-point comparisons between actual and forecasted predictions are given in Figure 1.

Conclusions and Recommendations

In the current study, newly proposed GBM and ARIMA based hybrid approaches were widely used to discuss our result. The model accuracy results of the root mean square error (RMSE) reveal that (RMSE[ARIMA] > RMSE[ANN_ARIMA]), new proposed ANN_GBM model is more significant and gives best solution for predicting short term predictions in high volatility fluctuations than traditional forecasting approaches; especially for the nonlinear models under the stationary and non-stationary frameworks.

It is expected that the current study will make a significant contribution to policy makers as well as the government to open up new direction to develop the CSE investments in Sri Lanka.

References

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

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

[3] S. Abidin and M. Jaffar. A review on Geometric Brownian Motion in Forecasting the share prices in Bursa Malaysia. *World Applied Sciences Journal,* 5(2), pp. 87-93. 2012.

[4] R.M.K.T Ratnayaka, Z. Wang, S. Anamalamudi and S. Cheng, Enhanced greedy optimization algorithm with data ware housing for automated nurse scheduling system. *E-Health Telecommunication Systems and Networks,* 2(1), pp. 43-48. 2012.

[5] R. Rathnayaka, W. Jianguo and D. Seneviratna. *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). 2014.

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

[7] P. Areekkul, T. Senjyu, H. Toyama and Atsushi Yona. *Combination of Artificial Neural network and ARIMA Time Series Models for short term price forecasting in Deregulated Market.* IEEE T&D Asia. 2009

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

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

[10] X. Long, L. Su and A. Ullah. *Estimation and Forecasting of Dynamic Conditional Covariance: A semiparametric Multivariate Model.* Journal of Business & Economic Statistics. 2011