

# Comparative Study of Time Series and ARIMA Model on Stock Data

Prakatheeswari S R, Pujithaa Vijayanand, Deivarani S, Vani K H

**Abstract**— Stock prediction is an important topic in the field of finance as well as engineering and mathematics. The stock market is attracting more and more people's notice with its high challenging risks and high turnover. A stock exchange market depicts savings and investments that are advantageous to increase the effectiveness of the national economy. The future stock returns have some predictive relationships with the publicly available information of present and historical stock market indices. This paper's main objective is to compare the accuracy of ARIMA Model and Time Series analysis on stock data of 32 companies for about 12 years. From the data, the parameters taken for this study are date, open, high, low, close, volume and the name. This model will automate the process of direction of future stock price and aids with financial specialists to choose the better timing for purchasing and/or selling of stocks.

**Index Terms**— Time series, ARIMA model, forecasting techniques, stock data

## I. INTRODUCTION

Recently, forecasting stock market return is gaining more attention may be because of the fact that, if the direction of the market is successfully predicted, investors may be better guided. The profitability of investing and trading in the stock market to a large extent depends on the predictability. If any system be developed which can consistently predict the trends of the dynamic stock market, would make the owner of the system wealthy. More over the predicted trends of the market will help the regulators of the market in devising corrective strategies. Another motivation for research in this field is that, it possesses many theoretical and experimental challenges.

The systematic approach by which one goes about answering the mathematical and statistical questions posed by these time correlations is commonly referred to as time series analysis. Time series analysis involves developing models that best captures or describes an observed time series in order to understand the underlying causes. Recently, data presented in the form of time series as an analysis and application that have become increasingly important in different domains. Time

**Manuscript received September 22, 2019**

**Prakatheeswari S R**, Student, M.Sc., Data Science, Dept. of Computing, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India

**Pujithaa Vijayanand**, Student, M.Sc., Data Science, Dept. of Computing, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India

**Deivarani S**, Assistant Professor, Dept. of Computing, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India

**Vani K H**, Assistant Professor, Dept. of Computing, Coimbatore Institute of Technology, Coimbatore, Tamil Nadu, India

series is a set of ordered observations on quantitative characteristics of a phenomenon at equally spaced time points. One of the main goals of time series analysis is to forecast future values based on existing series. There are various possible domains and fields of research based on time series like business, engineering, economics, medicine and lot others.

The Auto Regressive Integrated Moving Average (ARIMA) model is a statistical model which is known to be efficient for time series forecasting especially for short-term predictions. Auto Regressive Models are based on the idea that the current value of the series, 'xt', can be explained as a function of 'p' past values, 'xt-1', 'xt-2', . . . , 'xt-p', where 'p' determines the number of steps into the past needed to forecast the current value. ARIMA is superior to exponential smoothing techniques when the data is reasonably long and the correlation between past observations is stable. It constantly out-performed complex structural models in short-term prediction. In ARIMA model, the future value of a variable is a linear combination of past values and past errors, expressed as follows:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q}$$

## II. DATASET – IN THE CONTEXT OF STOCK DATA

The data set consists of daily stock value of 32 companies from 1<sup>st</sup> January 2006 to 1<sup>st</sup> December 2017. This study focuses on the use of time series analysis and to find out the best model to forecast future stock values. Also, the study attempts to compare the predicted value and the actual recorded value, and determine the accuracy of the model. ARIMA Model is also used on the same stock data and forecasting successive daily stock values are brought out. Henceforth, predicted values and the actual recorded values are compared to determine the accuracy of the model.

*The data of each company in the data set consists of the following:*

- Date - in format: yy-mm-dd
- Open - price of the stock at the open of the market
- Close – price of the stock at the close of the market
- High - Highest price reached in the day
- Low - Lowest price reached in the day
- Volume - Number of shares traded
- Name - the stock's ticker name

## III. DATA ANALYSIS

Once time series and ARIMA is performed, the results are plotted as follows:

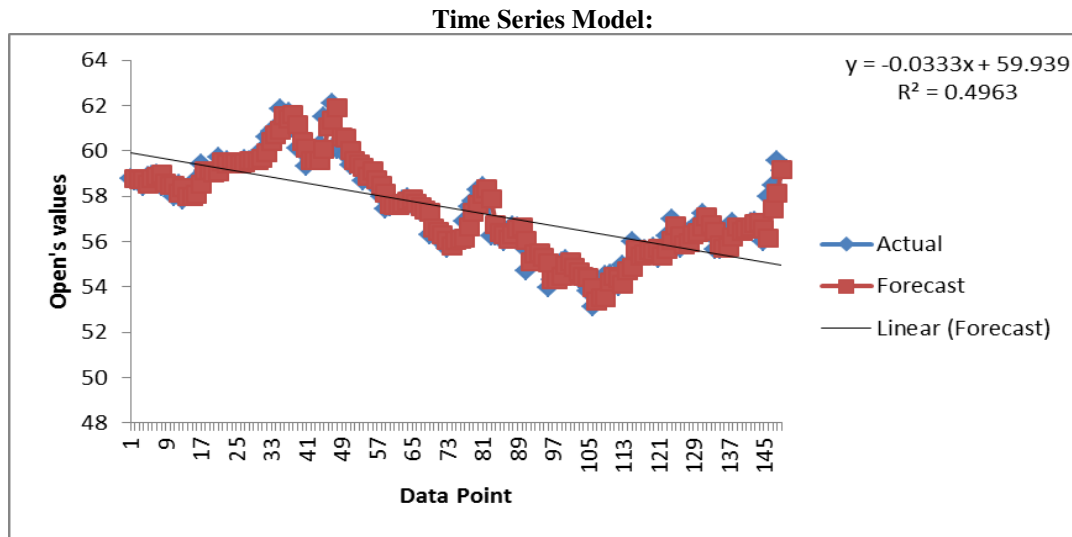


Figure 1: Time series plot for data point and open's values

Figure 1 represents the time series plot for a given set of data. It is inferred that the R2 value depicts the the closeness of the data to the fitted regression equation.

**ARIMA Model Result:**

Regression Statistics	Computed Values
Multiple R	0.543302228
R Square	0.295177311
Adjusted R Square	0.275598903
Standard Error	3033480.282
Observations	149

**ANOVA:**

	Degree of freedom	Sum of squares	Mean sum of squares	F value	Significance of F-value
Regression	4	5.55E+14	1.39E+14	15.07667582	2.57E-10
Residual	144	1.33E+10	9.20E+12		
Total	148	188E+15			

From the above tables, it is inferred that the R<sup>2</sup> values is **0.29**, from which it could be also be concluded that ARIMA model is a better way to forecast stock data compared to time series technique.

**CONCLUSION**

In this paper, the researchers have compared time series forecasting technique with ARIMA model. ARIMA model is a trendy method to analyse stationary univariate time series data. The experimental results obtained with ARIMA model demonstrates the potential of ARIMA model to predict stock prices satisfactorily for short-term forecasting. This could guide investors to make profitable investment decisions in stock market. With the results obtained from ARIMA models, it is safe to say that, it can compete reasonably well with emerging forecasting techniques for short-term prediction. Thus, the survey provides an insight to various time series predictions and forecasting models. Also, a lot of real time applications conducted by various people were studied, and it is clear that ARIMA is an appropriate model for time series prediction, forecasting and analysis with accuracy.

**REFERENCES**

- [1]Levine D M, Krehbiel TC, Berensen M L, 2003, Business Statistics: A First Course, Pearson Education, New Delhi, India.
- [2]Gupta S C, Fundamentals of Statistics, 2014, 7e, Himalaya Publishing, New Delhi.

**ACKNOWLEDGMENTS**

The authors whole heartedly acknowledge and thank all the authors and publishers whose content has been referred and cited in this research paper.