“Forecasting stock market prices using mixed ARIMA model: a case study of Indian pharmaceutical companies”

AUTHORS	Bharat Kumar Meher https://orcid.org/0000-0002-7264-5882 Iqbal Thonse Hawaldar https://orcid.org/0000-0001-7181-2493 https://publons.com/researcher/1456475/iqbal-thonse-hawaldar/ Cristi Spulbar https://orcid.org/0000-0002-3909-9496 Ramona Birau https://orcid.org/0000-0003-1638-4291

ARTICLE INFO	Bharat Kumar Meher, Iqbal Thonse Hawaldar , Cristi Spulbar and Ramona Birau (2021). Forecasting stock market prices using mixed ARIMA model: a case study of Indian pharmaceutical companies. *Investment Management and Financial Innovations*, *18*(1), 42-54. doi:10.21511/imfi.18(1).2021.04
DOI	http://dx.doi.org/10.21511/imfi.18(1).2021.04
RELEASED ON	Friday, 22 January 2021
RECEIVED ON	Wednesday, 11 November 2020
ACCEPTED ON	Monday, 18 January 2021

JOURNAL	"Investment Management and Financial Innovations"
ISSN PRINT	1810-4967
ISSN ONLINE	1812-9358
PUBLISHER	LLC “Consulting Publishing Company “Business Perspectives”
FOUNDER	LLC “Consulting Publishing Company “Business Perspectives”

NUMBER OF REFERENCES	NUMBER OF FIGURES	NUMBER OF TABLES
18	1	10

BUSINESS PERSPECTIVES

LLC "CPC "Business Perspectives"
Hryhorii Skovoroda lane, 10,
Sumy, 40022, Ukraine

www.businessperspectives.org

Bharat Kumar Meher (India), **Iqbal Thonse Hawaldar** (Bahrain),
Cristi Spulbar (Romania), **Ramona Birau** (Romania)

FORECASTING STOCK MARKET PRICES USING MIXED ARIMA MODEL: A CASE STUDY OF INDIAN PHARMACEUTICAL COMPANIES

Abstract

Many investors in order to predict stock prices use various techniques like fundamental analysis and technical analysis and sometimes rely on the discussions provided by various stock market analysts. ARIMA is a part of time-series analysis under prediction algorithms, and this paper attempts to predict the share prices of selected pharmaceutical companies in India, listed under NIFTY100, using the ARIMA model. A sample size of 782 time-series observations from January 1, 2017 to December 31, 2019 for each selected pharmaceutical firm has been considered to frame the ARIMA model. ADF test is used to verify whether the data are stationary or not. For ARIMA model estimation, significant spikes in the correlogram of ACF and PACF have been observed, and many models have been framed taking different AR and MA terms for each selected company. After that, 5 best models have been selected, and necessary inculcation of various AR and MA terms has been made to adjust the models and choose the best adjusted ARIMA model for each firm based on Volatility, adjusted R-squared, and Akaike Information Criterion. The results could be used to analyze the stock prices and their prediction in-depth in future research efforts.

Keywords Akaike Information Criterion, augmented Dickey-Fuller test, prediction, risk, investment, India

JEL Classification C22, C53, E27

Received on: 11th of November, 2020
Accepted on: 18th of January, 2021
Published on: 22nd of January, 2021

Bharat Kumar Meher, Ph.D., Department of Commerce, D.S. College Under Purnea University, Katihar, Bihar, India.

Iqbal Thonse Hawaldar, Ph.D., Professor, Department of Accounting & Finance, College of Business Administration, Kingdom University, Bahrain. (Corresponding author)

Cristi Spulbar, Ph.D., Faculty of Economics and Business Administration, University of Craiova, Craiova, Romania.

Ramona Birau, Ph.D., Faculty of Education Science, Law and Public Administration, Constantin Brancusi University of Targu Jiu, Romania.

Conflict of interest statement:
Author(s) reported no conflict of interest

INTRODUCTION

The general attitude of the society towards the share market is that it is enormously risky for investment or not suitable for trade. Even though there is the involvement of risk in the stock market, many people are interested in investment. The significant factor for any investor is to maximize the yields on their investments and achieve this element, and the investors always try to predict or forecast the stock prices. Many sound investors use various techniques like fundamental analysis and technical analysis to forecast stock prices. Still, on the other hand, many investors rely on the discussions or suggestions given by various stock market analysts and financial analysts. Some financial analysts and investors use these fundamental or technical analyses and prediction algorithms and functions to predict future share prices and their performance.

The prediction algorithms and functions include time series analysis. The time series analysis is considered an appropriate tool to predict the trend. However, the major limitation of the trend chart is sometimes it might not reflect or predict the variations or steady flow of the market. "The investors are very much interested to know the past trend or flow, seasonal growth, or variations of the stock" (Faisal, 2012; Iqbal & Mallikarjunappa, 2009). "A general view or expectation is that it must give a holistic view of the stock market. As it is essential to identify a

http://dx.doi.org/10.21511/imfi.18(1).2021.04

model to show the trend with adequate information for the investor to decide." Davi et al. (2013) recommend that "ARIMA is an algorithmic approach to transform the series is better than forecasting directly, and it gives more accurate results." Moreover, Wadia et al. (2011) confirmed that "ARIMA model has a fixed structure and is specifically built for time series (sequential) data." Its forecasts are usually more accurate and reliable as it is a univariate model and hence cannot exploit the leading indicators or explanatory variables.

The study is based on applying the ARIMA model to forecast the share prices of pharmaceutical companies in India. Among all the developing countries, the Indian pharmaceutical industry is one of the biggest and the most sophisticated and plays a pivotal role in the economic development of India. "Being a very intense knowledge-based industry, it offers innumerable business opportunities for investors worldwide. Indian pharmaceutical exports account for export to more than 200 countries around the world" (Kumar et al., 2020). Moreover, as per the report of the Indian Brand Equity Foundation, the low cost of production and increasing research and development has led to competitive pharma exports from exports reached US$ 17.15 billion in FY2019. The annual turnover of pharmaceutical products bestows to about US$ 20 billion. The pharmaceutical industry of the country is expected to rise at a CAGR of 22.4% over 2015–2020 to reach US$ 55 billion. The sector was valued at US$ 33 billion in the previous year. The industry is the third-largest in terms of volume in India. It has been a topmost player in the segment of health care in an epoch of aging populations, increasing health care costs, and the perpetual development of novel and extremely beneficial drugs.

Hence, the investment in securities and shares of pharmaceutical companies appears to be cost-effective. When it comes to the infusion of money in publicly traded pharmaceutical companies, investors should closely examine these companies when they reach clinical trials. Clinical trials are always a make-or-break chance for firms, and their products – successful outcomes could lead to significant profits in the market. Moreover, "in the recent past, the mutual fund houses aligned their products to invest in the pharma sector. The sector, which has seen a huge decline in the last couple of years, is now a darling for the mutual fund houses. Recently, Mirae Assets and ICICI Prudential announced new thematic schemes keenly focused on maximizing profits from the pharma sector stocks" (Desai, 2018). Hence, this study attempts to frame models to forecast the share prices of selected pharmaceutical companies, which are under NIFTY 100, using a mixed Auto-Regressive Integrated Moving Average (ARIMA).

1. 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). A study of 10 selected pharmaceutical companies of India, listed in BSE and NSE, was performed with the help of fundamental analysis using the ratio analysis technique (Iqbal & Mallikarjunappa, 2009, 2010; Panigrahi, Sharma, & Dhande, 2018). A study considered New York Stock Exchange and NASDAQ Stock Exchange. The research proposed the ARIMA model for ascertaining the value of future share prices. ARIMA revealed better forecasting results as it can handle the time series data very well, which is suitable for forecasting the future share index (Iqbal & Mallikarjunappa, 2011; Chi & Subramanian, 2019). Again a study focused on forecasting the gold price in Malaysia with the help of ARIMA with symmetric GARCH-type models if there exists heteroscedasticity (Yaziz et al., 2019). The researches done earlier were not enough to provide an appropriate model to predict the stock prices of pharmaceutical companies of India through ARIMA; hence, this study is an attempt to fill this research gap.

2. METHODS

The study is analytical. The data used in this study are from secondary sources. The secondary data involves the daily closing prices of shares of pharmaceutical companies listed in NIFTY 100 of India. The secondary data related to daily closing prices of stocks ranging from January 1, 2017 to December 31, 2019 have been downloaded from Yahoo Finance. Wherever required, an attempt has been made to make the unbalanced data into balanced data, i.e., five days a week. There are nine pharmaceutical companies in India, listed under NIFTY 100, namely Sun Pharmaceutical Industries Ltd., Divi's Laboratories Ltd., Dr. Reddy's Laboratories Ltd., Aurobindo Pharma Ltd., Biocon Ltd., Cadila Healthcare Ltd., Cipla Ltd., Lupin Ltd., and Piramal Enterprises Ltd. The top three pharmaceutical companies based on market capitalization have been selected for modeling and analysis, i.e., Sun Pharmaceutical Industries Ltd., Lupin Ltd., and Dr. Reddy's Laboratories. The total sample size is 2,346, i.e., three companies of 782 observations each. A required number of differencing has been done to make the data stationary, and an augmented Dickey-Fuller test has been utilized to check the stationarity of the data. Autoregressive Integrated Moving Average (ARIMA) has been used to formulate the model for each company for the forecasting share price. A correlogram of ACF and PACF would be plotted to determine the different AR and MA terms. After formulating the models, the models have been utilized to predict the share prices from October 1, 2019 to December 31, 2019. The awareness regarding the growth of the pharma industry to the investors is needed. The output of the study could give a potential model for each pharmaceutical company of India selected for this study. Through this study, the share price of selected pharmaceutical companies could be predicted, which could help the scholars and researchers go through proper research to develop best fitted predicted models in the future as well. Moreover, this could also assist investors with a basic knowledge of algorithms to run the developed models to predict the price of selected pharmaceutical companies.

3. RESULTS AND DISCUSSION

Auto-Regressive Integrated Moving Average (ARIMA) model is a generalization of an Autoregressive Moving Average (ARMA) model. An ARMA model expresses the conditional mean of Y_t as a function of both past observations Y_{t-1}, Y_{t-2}, Y_{t-p} and past innovations, ε_{t-1}, ε_{t-q}. The number of past observations that Y_t depends on, p, is the AR degree. The number of past innovations that Y_t depends on, q, is the MA degree.

In general, these models are denoted by ARMA (p, q). The form of the ARMA (p, q) model is

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \ldots + \beta_p Y_{t-p} + \\ + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \phi_2 \varepsilon_{t-2} + \ldots + \phi_q \varepsilon_{t-q}, \quad (1)$$

http://dx.doi.org/10.21511/imfi.18(1).2021.04

Table 1. The results of the best 5 ARIMA models out of 100 models for Sun Pharmaceutical Industries Ltd.

ARIMA (*p*, *d*, *q*)	AIC (Akaike Information Criterion)	Schwarz criterion	Volatility (SIGMASQ)	*R*-squared	Adjusted *R*-squared
ARIMA (24, 1, 47)	7.466761	7.49063	101.1789	0.031146	0.027405
ARIMA (47, 1, 24)	7.467272	7.491142	101.2417	0.030544	0.026801
ARIMA (35, 1, 47)	7.46908	7.49295	101.4247	0.028792	0.025042
ARIMA (47, 1, 131)	7.467994	7.491864	101.1849	0.031088	0.027347
ARIMA (131, 1, 47)	7.467841	7.491711	101.1739	0.031193	0.027453

where α – constant term, $\beta_1 \ldots \beta_p$ – AR – non-seasonal autoregressive (AR) coefficients, ϕ – MA – Nonseasonal Moving Average (MA) coefficients, $Y_{t-1} \ldots Y_{t-p}$ – non-seasonal AR lags corresponding to non-zero, $\varepsilon_{t-1} \ldots \varepsilon_{t-q}$ – MA lags – lags corresponding to non-zero, non-seasonal MA coefficients, D – degree of non-seasonal differencing, D (if D has value 0 meaning no non-seasonal integration).

To determine the range of AR and MA degree by examining and comparing the significant spikes in the correlogram of Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) of daily closing prices. To determine the five effective ARIMA (*p*, *d*, *q*) model by comparing the models with the help of AIC, adjusted *R*-squared, and significant coefficients were used. To adjust the selected models by inculcating or eliminating more AR and MA terms to develop a best mixed ARIMA model having high R squared with significant coefficients. The rule of thumb to select better models from the list of developed models with different AR and MA terms is the models with lesser Akaike Information Criterion (AIC), Schwarz criterion, and Volatility, i.e., SIGMASQ, and with higher *R*-squared and adjusted *R*-squared are better. This ARIMA model has been applied in 3 pharmaceutical companies of India, the procedure and results of which are mentioned further.

The closing price of a share of Sun Pharmaceutical Industries Ltd., ranging from January 1, 2017 to December 31, 2019, has been differentiated once to make the data stationary. The closing price data stationarity has been examined with the assistance of a unit root test named augmented Dickey-Fuller test with the inclusion of the test equation as Intercept, Trend, and Intercept and None. The correlogram of ACF and PACF has been framed by taking 1st difference. By analyzing the correlogram, it has been found that the spikes in ACF and PACF are significant at 3, 9, 13, 15, 17, 18, 24, 35, 47, and 131 lags. Considering these lags, 100 models have been framed by taking different AR and MA terms. These 100 models have been compared, and 5 best models have been selected, the results of which are given in Table 1. All the models that have been selected have a significant coefficient as the significance values of AR and MA terms are less than 0.05.

Table 1 represents the best five selected ARIMA models of Sun Pharmaceutical Industries Ltd. It can be observed that in all the selected models, the value of *d* is 1, which means that the closing prices have become stationary only by differencing once. By comparing these five best models, the ARIMA (24, 1, 47) has the least AIC and Schwarz criterion followed by ARIMA (47, 1, 24) followed by ARIMA (131, 1, 47). The ARIMA (131, 1, 47) has the least volatility with the lowest SIGMASQ and has the highest *R*-squared and adjusted *R*-squared. There is also a possibility that some residuals have not been considered in the above models, due to which the *R*-squared and adjusted *R*-squared is still less. Hence, it is necessary to check the residual diagnostic. For this, again, the correlogram of *Q* statistics has been plotted, and by observing the significant spikes, different AR and MA terms have been inculcated and experimented with adjusting the above five models to achieve higher *R*-squared and adjusted *R*-squared. The results of the best five adjusted ARIMA models are mentioned in Table 2.

Table 2 represents the AIC, Schwarz criterion, Volatility with SIGMASQ, *R*-squared, and adjusted *R*-squared of best five adjusted ARIMA models with different AR and MA terms. The models that have been selected above are having a significant coefficient as the significance value is less than

Table 2. Results of the best 4 adjusted ARIMA models for Sun Pharmaceutical Industries Ltd.

ARIMA with different AR and MA terms	AIC (Akaike Information Criterion)	Schwarz criterion	Volatility (SIGMASQ)	*R*-squared	Adjusted *R*-squared
AR (9, 18, 47, 61, 79) MA (13, 24, 131)	7.44407	7.503744	97.06115	0.070576	0.059726
AR (13, 35, 47, 194) MA (24, 131)	7.443417	7.491157	97.2971	0.068316	0.059879
AR (9, 13, 24, 35, 194) MA (47, 131)	7.441396	7.495103	96.86233	0.072479	0.062868
AR (13, 24, 35, 61, 131, 194) MA (9, 47)	7.442778	7.502452	96.81995	0.072885	0.062063

0.05. By comparing the values of AIC, Schwarz criterion, and SIGMASQ, it can be said that the ARIMA with AR 9, 13, 24, 35, 194 and MA with 47, 131 terms, has the least AIC, Schwarz criterion, and SIGMASQ. Moreover, the same ARIMA model has the highest *R*-squared and adjusted *R*-squared. Hence, this model would be appropriate for predicting share price. The result of the selected model is mentioned in Table 3.

The mentioned above is the result of the selected model for predicting the share price of Sun Pharmaceuticals. It can be observed that the values of the coefficient are significant, but the *R* and adjusted *R*-squared are less. With the help of the values of different AR and MA terms and constant term, the following model can be framed:

$$D(Y_t) = -0.245651 + 0.062704Y_{t-9} + \\ +0.088239Y_{t-13} - 0.098491Y_{t-24} - \\ -0.090551Y_{t-35} - 0.097406Y_{t-194} + \\ +0.158197\varepsilon_{t-47} + -0.115884\varepsilon_{131}.$$

With the help of the above model, the stock prices for the last three months, i.e., October 1, 2019 to December 31, 2019, have been predicted. Then the actual price and the predicted price have been plotted in the forecasting section to check whether the model can predict the price properly or not.

Table 3. Result of the selected model for Sun Pharmaceutical Industries Ltd.

Dependent variable: D(CLOSING_PRICE_SUNF)
Method: least squares
Date: 02/25/20 Time: 15:07
Sample: 1/03/2017 – 12/31/2019
Included observations: 781
Convergence achieved after 13 iterations

Coefficient covariance computed using the outer product of gradients				
Variable	**Coefficient**	**Std. Error**	***t*-statistic**	**Prob.**
C	–0.245651	0.350233	–0.701393	0.4833
AR (9)	0.062704	0.031807	1.971361	0.0490
AR (13)	0.088239	0.034887	2.529319	0.0116
AR (24)	–0.098491	0.036820	–2.674966	0.0076
AR (35)	–0.090551	0.036673	–2.469141	0.0138
AR (194)	–0.097406	0.031965	–3.047220	0.0024
MA (47)	0.158197	0.033848	4.673810	0.0000
MA (131)	–0.115884	0.038279	–3.027322	0.0025
SIGMASQ	96.86233	3.116799	31.07750	0.0000
R-squared	0.072479	Mean dependent var		–0.258003
Adjusted *R*-squared	0.062868	S.D. dependent var		10.22572
S.E. of regression	9.899068	Akaike info criterion		7.441396
Sum squared resid	75649.48	Schwarz criterion		7.495103
Log-likelihood	–2896.865	Hannan-Quinn criteria		7.462051
F-statistic	7.540814	Durbin-Watson stat		1.892796
Prob(*F*-statistic)	0.000000			

http://dx.doi.org/10.21511/imfi.18(1).2021.04

Table 4. Results of the best 5 ARIMA models out of 81 models for Lupin Ltd.

ARIMA (*p*, *d*, *q*)	AIC (Akaike Information Criterion)	Schwarz criterion	Volatility (SIGMASQ)	*R*-squared	Adjusted *R*-squared
ARIMA (22, 1, 98)	8.363473	8.387342	248.1767	0.015311	0.011509
ARIMA (29, 1, 115)	8.359694	8.383564	246.9387	0.020223	0.01644
ARIMA (98, 1, 22)	8.363273	8.387143	248.117	0.015547	0.011746
ARIMA (102, 1, 115)	8.361372	8.385242	247.3055	0.018767	0.014979
ARIMA (115, 1, 102)	8.361278	8.385147	247.2737	0.018893	0.015105

Similarly, the closing price of a share of Lupin Ltd., ranging from January 1, 2017 to December 31, 2019, has been differentiated once to make the data stationary. The stationarity of the data has been checked with the help of a unit root test named augmented Dickey-Fuller test with the inclusion of the test equation as Intercept, Trend, and Intercept and None. The correlogram of ACF and PACF has been plotted by taking 1st difference. By analyzing the correlogram, it has been found that the spikes in ACF and PACF are significant at 1, 6, 22, 24, 29, 98, 102, 115, and 170 lags. By considering these lags, 81 models have been framed by taking different AR and MA terms. These 81 models have been compared, and 5 best models have been selected, the results of which are given in Table 4.

Table 4 represents the best five selected ARIMA models of Lupin Ltd. It can be observed that in all the selected models, the value of *d* is 1, which means that the closing prices have become stationary only by differencing once. By comparing these five best models, the ARIMA (29, 1, 115) has the least AIC and Schwarz criterion, followed by ARIMA (115, 1, 102) followed by ARIMA (102, 1, 115). Again, the model ARIMA (29, 1, 115) has the least volatility with the lowest SIGMASQ and has the highest *R*-squared and adjusted *R*-squared. Hence, ARIMA (29, 1, 115) could be considered the best model. However, there is also a possibility that some residuals have not yet been considered in the above models, due to which the *R*-squared and adjusted *R*-squared are still less. Hence, it is necessary to check the residual diagnostic. For this, again, the correlogram of *Q* statistics has been plotted, and by observing the significant spikes, different AR and MA terms have been inculcated and experimented with adjusting the above five models to achieve higher *R* and adjusted *R*-squared. The results of the best five adjusted ARIMA models are mentioned in Table 5.

Table 5 represents the AIC, Schwarz criterion, Volatility with SIGMASQ, *R*-squared, and adjusted *R*-squared of best five adjusted ARIMA models with different AR and MA terms. The models that have been selected above are having a significant coefficient as the significance value is less than 0.05. By comparing the values of AIC, Schwarz criterion, and SIGMASQ, it can be said that the ARIMA with AR 22, 29, 155, and MA with 98, 115 terms has the least AIC, Schwarz criterion, and SIGMASQ. Moreover, the same ARIMA model has the highest *R*-squared and adjusted *R*-squared. Hence, this model would be appropriate for predicting the share price of Lupin Ltd. The result of the selected model is mentioned in Table 6.

Table 6 mentioned is the result of the selected model for predicting the share price of Sun

Table 5. Results of the best 5 adjusted ARIMA models for Lupin Ltd.

ARIMA with different AR and MA terms	AIC (Akaike Information Criterion)	Schwarz criterion	Volatility (SIGMASQ)	*R*-squared	Adjusted *R*-squared
AR (1, 22, 29) MA (29, 98)	8.354458	8.39623	243.7629	0.032823	0.025326
AR (22, 29) MA (98, 115)	8.350709	8.386514	243.1943	0.035079	0.028854
AR (22, 29, 155) MA (98, 115)	8.34789	8.389662	241.5426	0.041633	0.034204
AR (22, 29, 102) MA (115)	8.352578	8.388383	243.6976	0.033082	0.026844
AR (29, 115) MA (102)	8.357919	8.387757	245.7358	0.024995	0.019969

Table 6. Result of the selected model for Lupin Ltd.

Dependent variable: D(CLOSING_PRICE_LUPI)				
Method: least squares				
Sample: 1/03/2017 – 12/31/2019				
Included observations: 781				
Convergence achieved after 18 iterations				
Coefficient covariance computed using the outer product of gradients				
Variable	**Coefficient**	**Std. Error**	***t*-statistic**	**Prob.**
C	–0.913487	0.496158	–1.841123	0.0660
AR (22)	–0.087667	0.041734	–2.100628	0.0360
AR (29)	–0.084947	0.040117	–2.117448	0.0345
AR (155)	–0.080189	0.034454	–2.327412	0.0202
MA (98)	0.085273	0.043207	1.973582	0.0488
MA (115)	–0.124606	0.043480	–2.865855	0.0043
SIGMASQ	241.5426	6.739731	35.83861	0.0000
R-squared	0.041633	Mean dependent var		–0.950448
Adjusted *R*-squared	0.034204	S.D. dependent var		15.88580
S.E. of regression	15.61176	Akaike info criterion		8.347890
Sum squared resid	188644.7	Schwarz criterion		8.389662
Log-likelihood	–3252.851	Hannan-Quinn criteria		8.363955
F-statistic	5.603931	Durbin-Watson stat		1.847017
Prob(*F*-statistic)	0.000010			

Pharmaceuticals. It can be observed that the values of the coefficient are significant, but the *R* and adjusted *R*-squared are less. With the help of the values of different AR and MA terms and constant term, the following model can be framed:

$$D(Y_t) = -0.913487 - 0.087667Y_{t-22} - \\ -0.084947Y_{t-29} - 0.080189Y_{t-155} + \\ +0.085273\varepsilon_{t-98} - 0.124606\varepsilon_{t-115}.$$

With the help of the above model, the stock prices for the last three months, i.e., from October 1, 2019 to December 31, 2019, have been predicted. Then the actual price and the predicted price have been plotted in the forecasting section to check whether the model can predict the price properly or not.

Similarly, the closing price of a share of Dr. Reddy's Laboratories ranging from January 1, 2017 to December 31, 2019 has been differentiated once to make the data stationary. The stationarity of the data has been checked with the help of a unit root test named augmented Dickey-Fuller test with the inclusion of the test equation as Intercept, Trend, and Intercept and None. The correlogram of ACF and PACF has been plotted by taking 1st difference. By analyzing the correlogram, it has been found that the spikes in ACF and PACF are significant at 6, 7, 8, 24, 35, 39, 53, 99, 102, 144, 172, and 198 lags. Considering these lags, 144 models have been framed by taking different AR and MA terms. These 144 models have been compared, and 5 best models have been selected, the results of which are included in Table 7.

Table 7 represents the best five selected ARIMA models of Dr. Reddy's Laboratories. It can be observed that in all the selected models, the value of *d* is 1, which means that the closing prices have become stationary only by differencing once. By comparing these five best models, the ARIMA (24,

Table 7. Results of best 5 ARIMA models out of 144 models for Dr. Reddy's Laboratories

ARIMA (*p*, *d*, *q*)	AIC (Akaike Information Criterion)	Schwarz criterion	Volatility (SIGMASQ)	*R*-squared	Adjusted *R*-squared
ARIMA (24, 1, 144)	10.25701	10.28088	1647.706	0.019745	0.01596
ARIMA (144, 1, 99)	10.25876	10.28263	1648.589	0.01922	0.015433
ARIMA (99, 1, 144)	10.25817	10.28204	1647.362	0.01995	0.016166
ARIMA (144, 1, 24)	10.25732	10.28119	1648.066	0.019531	0.015745
ARIMA (144, 1, 198)	10.26027	10.28414	1649.037	0.018953	0.015166

http://dx.doi.org/10.21511/imfi.18(1).2021.04

Table 8. Results of best 5 adjusted ARIMA models for Dr. Reddy's Laboratories

ARIMA with different AR and MA terms	AIC (Akaike Information Criterion)	Schwarz criterion	Volatility (SIGMASQ)	*R*-squared	Adjusted *R*-squared
AR (24, 35, 151) MA (39, 99, 144, 151)	10.24059	10.29429	1590.011	0.054069	0.044267
AR (24, 35, 151) MA (39, 99, 144, 151, 164)	10.23557	10.29524	1572.647	0.064399	0.053478
AR (24, 35, 144, 151) MA (39, 99, 151, 164)	10.23703	10.2967	1577.854	0.061302	0.050344
AR (24, 35, 99, 144, 164) MA (6, 39, 198)	10.23257	10.29224	1572.68	0.06438	0.053458
AR (24, 35, 99, 144) MA (6, 39, 164, 198)	10.23227	10.29194	1571.879	0.064856	0.05394

1, 144) has the least AIC and Schwarz criterion, followed by ARIMA (144, 1, 24), followed by ARIMA (99, 1, 144). The ARIMA (99, 1, 144) has the least volatility with the lowest SIGMASQ and has the highest *R*-squared and adjusted *R*-squared. Again, there is also a possibility that some residuals have not yet been considered in the above models, due to which the *R*-squared and adjusted *R*-squared are still less. Hence, it is necessary to check the residual diagnostic. For this, again, the correlogram of Q statistics has been plotted, and by observing the significant spikes of ACF and PACF, different AR and MA terms have been inculcated and experimented with adjusting the above five models to improve *R* and adjusted *R*-squared. The results of the best five adjusted ARIMA models are mentioned in Table 8.

Table 8 represents the AIC, Schwarz criterion, Volatility with SIGMASQ, *R*-squared and adjusted *R*-squared of best five adjusted ARIMA models with different AR and MA terms. The models that have been selected above are having a significant coefficient as the significance value is less than 0.05. By comparing the values of AIC, Schwarz criterion, and SIGMASQ, it can be said that the ARIMA with AR 24, 35, 99, and 144 and MA with 6, 39, 164, and 198 terms has the least AIC,

Table 9. Result of the selected model for Dr. Reddy's Laboratories

Dependent variable: D(CLOSINGPRICE_DRED)
Method: least squares
Sample: 1/03/2017 – 12/31/2019
Included observations: 781
Convergence achieved after 25 iterations

Coefficient covariance computed using the outer product of gradients				
Variable	**Coefficient**	**Std. Error**	***t*-statistic**	**Prob.**
C	–0.430753	1.104142	–0.390124	0.6966
AR (24)	–0.090775	0.037453	–2.423679	0.0156
AR (99)	–0.083474	0.037946	–2.199784	0.0281
AR (144)	–0.080736	0.041019	–1.968265	0.0494
AR (35)	–0.083815	0.034932	–2.399393	0.0167
MA (198)	0.110289	0.045107	2.445059	0.0147
MA (6)	–0.079170	0.037050	–2.136841	0.0329
MA (39)	–0.087180	0.032261	–2.702299	0.0070
MA (164)	0.100931	0.046527	2.169284	0.0304
SIGMASQ	1571.879	61.16674	25.69826	0.0000
R-squared	0.064856	Mean dependent var		–0.267413
Adjusted *R*-squared	0.053940	S.D. dependent var		41.02500
S.E. of regression	39.90321	Akaike info criterion		10.23227
Sum squared resid	1227637.	Schwarz criterion		10.29194
Log-likelihood	–3985.700	Hannan-Quinn criteria		10.25522
F-statistic	5.941368	Durbin-Watson stat		1.958696
Prob(*F*-statistic)	0.000000			

Schwarz criterion, and SIGMASQ, moreover, the same ARIMA model has highest *R*-squared and adjusted *R*-squared. Hence, this model would be appropriate for predicting the share price of Dr. Reddy's Laboratories. The result of the selected model is mentioned in Table 9.

It can be observed from Table 9 that the values of the coefficient are significant, but the R and adjusted R-squared are less. With the help of the values of different AR and MA terms and constant term, the following model can be framed:

$$D(Y_t) = -0.430753 - 0.090775Y_{t-24} - \\ -0.083815Y_{t-35} - 0.083474Y_{t-99} - \\ -0.080736Y_{t-144} - 0.079170\varepsilon_{t-6} - \\ -0.087180\varepsilon_{t-39} + 0.100931\varepsilon_{t-164} + \\ +0.110289\varepsilon_{t-198}.$$

With the help of the above model, the stock prices for the last three months, i.e., October 1, 2019 to December 31, 2019, have been predicted. Then the actual price and the predicted price have been plotted in the forecasting section to check whether the model can predict the price properly or not.

Forecasting of stock prices of selected companies

With the help of the above models developed for each selected company under this study, the share prices have been predicted from October 1, 2019 to December 31, 2019. The forecasted prices and the actual prices are represented in the table given in Appendix A.

The values mentioned in that table have been used, and line graphs of Sun Pharmaceutical Industries Ltd., Lupin Ltd., and Dr. Reddy's Laboratories from October 1, 2019 to December 31, 2019 have been plotted, which are mentioned in Figure 1.

The line graphs mentioned in Figure 1 show the actual and predicted price. As the deviations between the actual share price line and predicted share price line are closer, the models can be considered reliable.

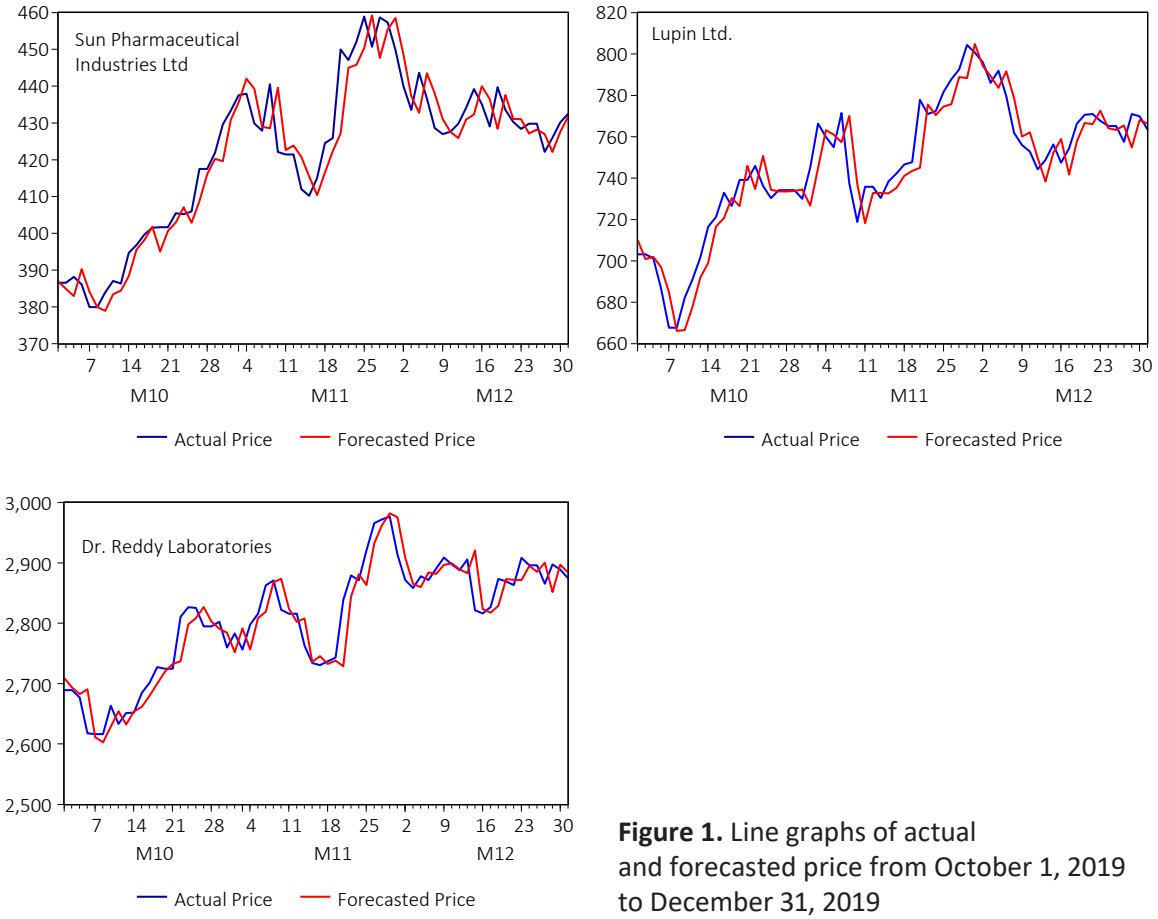

Figure 1. Line graphs of actual and forecasted price from October 1, 2019 to December 31, 2019

http://dx.doi.org/10.21511/imfi.18(1).2021.04

CONCLUSION

It can be observed that different AR and MA terms need to be included in the models to make the model more representative. The prediction of share prices of Sun Pharmaceuticals and Lupin Ltd. seems reliable as the deviations between the actual share price and the forecasted share price of the last three months of the study, i.e., from October 1, 2019 to December 31, 2019, is less. However, so far as the predicted share price of Dr. Reddy Laboratories is concerned, slightly more deviations can be found between the actual and predicted share price, which implies that the model developed for Dr. Reddy Laboratories is less reliable to predict the approximate price of the share on a certain date. In the overall study, the models selected for each firm have significant co-efficient, but so far as *R*-squared values are concerned, the values are less than the desired. There is also a possibility of dropping the constant or drift to increase the value of *R*-squared and adjusted *R*-squared to improve the model. The ACF and PACF have been plotted up to 200 lags, i.e., only one-fourth of the total observations. Many more significant lags might be there if the number of lags increased to one-third. Still, if more significant delays would be considered, then it is necessary to formulate and experiment with more models, which are time-consuming and cumbersome processes. Even after the *R* and adjusted *R*-squared are less in every model, the deviations between the actual values of share and forecasted shares are less, which implies that with the above three mixed ARIMA models, the price can be approximately predicted for the respective companies considered in the study. Moreover, the prices of all pharmaceutical companies are fluctuating. Still, if those prices can be predicted with the help of the models, then investment can be made on pharmaceutical companies of India, and profit can be earned even in the short term. There is also a possibility that the ARIMA models could be formulated by taking fewer periods, which may result in a more reliable model with higher *R* and adjusted *R*-squared. Formulating mixed ARIMA models by taking different periods and comparing them by observing AIC, Schwarz criterion, Volatility (SIGMASQ), *R* and adjusted *R*-squared and selecting a best-fitted model is the area for the future study.

AUTHOR CONTRIBUTIONS

Conceptualization: Bharat Kumar Meher, Iqbal Thonse Hawaldar, Ramona Birau.
Data curation: Bharat Kumar Meher, Cristi Spulbar.
Formal analysis: Iqbal Thonse Hawaldar, Ramona Birau.
Investigation: Bharat Kumar Meher, Ramona Birau.
Methodology: Bharat Kumar Meher, Iqbal Thonse Hawaldar, Cristi Spulbar.
Project administration: Iqbal Thonse Hawaldar.
Validation: Cristi Spulbar, Ramona Birau.
Writing – original draft: Bharat Kumar Meher, Ramona Birau.
Writing – review & editing: Iqbal Thonse Hawaldar,Cristi Spulbar.

REFERENCES

1. 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
2. Bolar, S., Pinto, P., & Hawaldar, I. T. (2017). Semi-monthly effect in stock returns: new evidence from BombayStock Exchange. *Investment Management and Financial Innovations, 14*(3), 160-172. http://dx.doi.org/10.21511/imfi.14(3-1).2017.01
3. Chi, J. L., & Subramanian, P. (2019). Forecasting of Stock Price Using Autoregressive Integrated Moving Average Model. *Journal of Computational and Theoretical Nanoscience, 16*(8), 3519-3524. Retrieved from https://www.researchgate.net/publication/336365885_Forecasting_of_Stock_Price_Using_Autoregressive_Integrated_Moving_Average_Model
4. Desai, N. (2018). *Pharma sector is the next darling of the market.* Retrieved December 5, 2019, from https://www.dsij.in/DSIJArticleDetail/ArtMID/10163/ArticleID/2458/Pharma-sector-is-the-next-darling-of-market

5. Devi, B. U., Sundar, D., & Alli, P. (2013). An Effective Time Series Analysis for Stock Trend Prediction Using ARIMA Model for Nifty Midcap-50. *International Journal of Data Mining & Knowledge Management Process, 3*(1), 65-78. https://doi.org/10.5121/ijdkp.2013.3106

6. Faisal, F. (2012). Forecasting Bangladesh's Inflation UsingTime Series ARIMA Models. *World Review of Business Research, 2*(3), 100-117. Retrieved from https://studylib.net/doc/12375451/forecasting-bangladesh-s-inflation-using-time-series-arim

7. Iqbal, T. H., & Mallikarjunappa, T. (2009). Indian Stock Market Reaction to The Quarterly Earnings Information. *Indian Journal of Finance, 3*(7), 43-50.

8. Iqbal, T. H., & Mallikarjunappa, T. (2010). A Study of Efficiency of the Indian Stock Market. *Indian Journal of Finance, 4*(5), 32-38. Retrieved from https://www.researchgate.net/publication/316239727_A_Study_of_Efficiency_of_the_Indian_Stock_Market

9. Iqbal, T. H., & Mallikarjunappa, T. (2011). *Efficiency of Stock Market: A Study of Stock Price Responses to Earnings Announcements.* Germany: LAP Lambert Academic Publishing Company. Retrieved from https://www.researchgate.net/publication/313599316_Efficiency_of_Stock_Market_A_Study_of_Stock_Price_Responses_to_Earnings_Announcements

10. Kumar, A., Soni, R., Hawaldar, I. T., Vyas, M., & Yadav, V. (2020). The Testing of Efficient Market Hypotheses: A Study of Indian Pharmaceutical Industry. *International Journal of Economics and Financial Issues, 10*(3), 208-216. https://doi.org/10.32479/ijefi.9764

11. Meher, B. K., Hawaldar, I. T., Mohapatra, L., & Sarea, A. M. (2020). The Impact of COVID-19 on Price Volatility of Crude Oil and Natural Gas Listed on Multi Commodity Exchange of India, *International Journal of Energy Economics and Policy, 10*(5), 422-431. https://doi.org/10.32479/ijeep.10047

12. Panigrahi, A. K., Sharma, R., & Dhande, H. (2018). Fundamental Analysis of Indian Pharmaceutical Companies. *Global Management Horizon, 7*(1), 2-25. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3202129

13. Pinto, P., Hawaldar, I. T., Guruprasad, K., Rohit, B., Spulbar, C., Birau, R., & Stanciu, C. V. (2020). The Impact of Risk Anomalies on the Pharmaceutical Sector of the Indian Stock Market: A Comparative Analysis between Pharmaceutical, FMCG and IT Companies. *Revista de Chimie Journal, 71*(2), 58-63. Retrieved from https://www.researchgate.net/publication/339687395_The_Impact_of_Risk_Anomalies_on_the_Pharmaceutical_Sector_of_the_Indian_Stock_Market_A_comparative_analysis_between_pharmaceutical_FMCG_and_IT_companies

14. 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

15. 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-Elango/5fe8b5c6126aa79fb58637b0d375fd4d4d4c6a6e

16. Uko, A. K., & Nkoro, E. (2012). Inflation forecasts with ARIMA, vector autoregressive and error correction 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_correction_models_in_Nigeria

17. Wadia, S. A., Ismail, M. T., Alkhahazaleh, M. H., & Karim, S. A. (2011). Selecting Wavelet Transforms Model in Forecasting Financial Time Series Data Based on ARIMA Model. *Applied Mathematical Sciences, 5*(7), 315-326. Retrieved December 3, 2019, from http://www.m-hikari.com/ams/ams-2011/ams-5-8-2011/alwadiAMS5-8-2011.pdf

18. Yaziz, S. R., Zakaria, R., & Suhartono. (2019). ARIMA and Symmetric GARCH-type Models in Forecasting Malaysia Gold Price. *Journal of Physics: Conference Series, 1.* https://doi.org/10.1088/1742-6596/1366/1/012126

http://dx.doi.org/10.21511/imfi.18(1).2021.04

APPENDIX A

Table A1. Actual and forecasted prices of Sun Pharmaceutical Industries Ltd., Lupin Ltd. and Dr. Reddy's Laboratories from October 1, 2019 to December 31, 2019

MM/DD/YEAR	Sun Pharmaceuticals		Lupin Ltd.		Dr. Reddy Laboratories	
	Actual price	Predicted price	Actual price	Predicted price	Actual price	Predicted price
10/01/2019	386.549988	386.9631	703.099976	710.0867	2689.350098	2709.163
10/02/2019	386.549988	384.9125	703.099976	700.8380	2689.350098	2693.781
10/03/2019	388.149994	382.9368	701.200012	701.9341	2677.149902	2682.643
10/04/2019	386.049988	390.2636	686.5	696.9438	2617.699951	2690.666
10/07/2019	379.899994	384.0914	667.599976	684.9523	2616.449951	2611.236
10/08/2019	379.899994	379.9513	667.599976	666.2121	2616.449951	2602.630
10/09/2019	383.950012	378.9204	682.099976	666.5035	2663.399902	2628.464
10/10/2019	387.049988	383.4512	690.950012	677.8391	2633.5	2654.156
10/11/2019	386.299988	384.4038	701.650024	691.9342	2651.399902	2632.143
10/14/2019	394.649994	388.3382	716.400024	698.9567	2651.899902	2653.852
10/15/2019	396.799988	395.5917	721.150024	716.5983	2684.649902	2661.509
10/16/2019	399.700012	398.2319	732.849976	720.6402	2701.449951	2679.654
10/17/2019	401.5	401.8214	726.5	730.3553	2727.399902	2699.898
10/18/2019	401.649994	395.0517	739.099976	726.5316	2724.850098	2719.980
10/21/2019	401.649994	400.6499	739.099976	745.9349	2724.850098	2732.732
10/22/2019	405.450012	402.9271	745.849976	734.7444	2810.75	2737.172
10/23/2019	405.200012	407.1119	736.150024	750.7402	2826.699951	2798.125
10/24/2019	405.950012	402.8990	730.25	734.3838	2825.300049	2808.609
10/25/2019	417.5	408.7538	734.099976	733.5543	2794.949951	2826.936
10/28/2019	417.5	416.0691	734.099976	733.5585	2794.949951	2802.685
10/29/2019	421.950012	420.2553	734.150024	733.8529	2802.5	2791.070
10/30/2019	429.649994	419.6296	729.950012	734.5381	2759.899902	2784.344
10/31/2019	433.399994	430.8921	745.099976	726.7670	2783.199951	2751.992
11/01/2019	437.549988	435.6230	766.299988	744.8425	2756.600098	2791.533
11/04/2019	437.950012	442.0251	759.950012	763.2276	2797.800049	2756.837
11/05/2019	429.899994	439.3374	754.950012	761.0847	2815.699951	2808.621
11/06/2019	427.899994	428.9682	771.450012	757.3803	2862.699951	2819.061
11/07/2019	440.600006	428.5809	737.450012	770.0223	2871.050049	2867.891
11/08/2019	422.100006	439.6006	718.75	737.0883	2822.100098	2873.540
11/11/2019	421.450012	422.6392	735.75	718.2671	2815.75	2824.140
11/12/2019	421.450012	423.8821	735.75	732.7605	2815.75	2802.243
11/13/2019	412	420.7912	730.349976	732.8442	2763.149902	2807.999
11/14/2019	410.200012	415.4825	738.349976	732.6323	2734.050049	2735.976
11/15/2019	415.049988	410.4001	742.25	735.3108	2730.800049	2745.453
11/18/2019	424.549988	416.5293	746.549988	741.1748	2737	2732.606
11/19/2019	425.899994	422.3090	747.700012	743.4958	2743	2738.079
11/20/2019	450	427.1397	777.849976	744.9974	2838.649902	2728.745
11/21/2019	447.149994	445.0476	771	775.5906	2879.25	2844.168
11/22/2019	452.100006	445.8039	772.400024	770.5203	2871.25	2881.208
11/25/2019	458.950012	450.3957	781.700012	774.5878	2920.649902	2863.251
11/26/2019	450.75	459.2639	787.849976	775.7867	2965.850098	2932.737
11/27/2019	458.75	447.7018	792.549988	788.8435	2972	2962.651
11/28/2019	457.350006	455.4809	804.349976	788.3817	2976.75	2982.143
11/29/2019	449.850006	458.5036	800.650024	804.8473	2913.850098	2975.587
12/02/2019	440	448.5920	796.049988	794.2559	2871.699951	2908.763
12/03/2019	433.600006	437.3284	786.049988	789.3396	2858.350098	2864.918
12/04/2019	443.700012	432.7711	791.900024	783.6221	2877.850098	2859.903
12/05/2019	436.700012	443.5603	779.650024	791.6104	2871.600098	2884.277
12/06/2019	428.600006	438.0687	761.849976	778.7119	2890.800049	2881.666

Table A1 (cont.). Actual and forecasted prices of Sun Pharmaceutical Industries Ltd., Lupin Ltd. and Dr. Reddy's Laboratories from October 1, 2019 to December 31, 2019

MM/DD/YEAR	Sun Pharmaceuticals		Lupin Ltd.		Dr. Reddy Laboratories	
	Actual price	Predicted price	Actual price	Predicted price	Actual price	Predicted price
12/09/2019	427	431.0635	755.950012	760.1565	2908.75	2896.728
12/10/2019	427.649994	427.5426	752.799988	762.1870	2897.5	2899.351
12/11/2019	429.799988	425.9348	744.200012	749.7892	2887.800049	2889.428
12/12/2019	434.149994	430.9636	748.75	738.4036	2905.649902	2883.060
12/13/2019	439.25	432.3350	756.25	752.2692	2821.5	2920.722
12/16/2019	435.149994	440.0159	747.450012	758.9412	2816.350098	2823.547
12/17/2019	429.049988	436.4935	754.349976	741.7211	2826.899902	2817.589
12/18/2019	439.75	428.4282	766.349976	757.8267	2873.449951	2828.747
12/19/2019	433.600006	437.6187	770.599976	766.6612	2869.550049	2873.110
12/20/2019	430.299988	431.1171	770.950012	766.1206	2863.350098	2871.974
12/23/2019	428.399994	431.0725	767.700012	772.5785	2908.399902	2871.312
12/24/2019	429.799988	427.1755	765.150024	764.1446	2895.949951	2895.024
12/25/2019	429.799988	428.2074	765.150024	763.3488	2895.949951	2885.271
12/26/2019	422.149994	426.9992	757.5	765.4803	2865.600098	2900.288
12/27/2019	426	422.1544	770.950012	754.8204	2897.600098	2851.648
12/30/2019	430.149994	427.6038	769.900024	768.0761	2888.850098	2897.113
12/31/2019	432.549988	431.8068	763.450012	766.4556	2874.550049	2883.886

http://dx.doi.org/10.21511/imfi.18(1).2021.04