



## PREDICTIVE ANALYTICS IN CAPITAL MARKETS

Nitin Singh<sup>1†</sup> --- Prakash Kumar Jha<sup>2</sup><sup>1</sup>Chairperson – Business Analytics IIM Kashipur Advisor KPMG Lodha Excelus NM Joshi Marg Mumbai, India<sup>2</sup>Technical Centre, Ashok Leyland Ltd, Vellivayal Chavadi, Via Manali New Town, Chennai, India

## ABSTRACT

*The paper is a practical application of the random walk model on stock price behaviour. The academic literature has moved beyond this random walk approach and the recent focus is now much more on how to improve the forecasts. Since the performance of the random walk model has been contextual, it is desirable that the model is tested in different contexts. Our model shows good results in the Indian context. This model is also useful for traders and investors looking to predict stock prices in the immediate future as the model accounts for changes in the immediate past.*

© 2017 AEES Publications. All Rights Reserved.

**Keywords:** Capital markets, Predictive analytics, Random walk, Time series, Stochastic process, Smoothing effect.

Received: 12 November 2016/ Revised: 10 December 2016/ Accepted: 28 December 2016/ Published: 7 January 2017

## Contribution/ Originality

The paper's primary contribution is finding the applicability of the random walk model in the context of Indian capital market and improvement in forecast accuracy by introducing a smoothing constant in the model. The modification enables a self-correcting feature that inherently checks for drifts and accordingly 'corrects' the forecast for the next period.

## 1. INTRODUCTION

Many decision-making applications depend on a forecast of some quantity. Broadly, these can be categorised into three groups: (1) the judgemental method, which is non-quantitative method; (2) extrapolation or time series methods that use past data of time series variable; and (3) econometric, causal or regression-based methods that use other explanatory time series variables. When an organisation plans to invest in stocks, bonds or other financial instruments, it typically attempts to forecast movement in prices and interest rates. This paper presents an analytical model to predict stock prices. Typically, we predict based on observations made in the past. We can investigate past behaviour, search for patterns or relationships and make a forecast. However, it is not easy to uncover historical patterns or relationships. It is all the more difficult to separate noise or random behaviour from underlying patterns. Moreover, there is no guarantee that past patterns will continue in the future. Extrapolation is found to be a good way to study past movement of a variable such as stock price, interest rate, GDP, etc. to forecast its future values. Several

time-series methods are available for extrapolation, including moving averages, exponential smoothing, trend-based regression, auto regression, random walk model, etc. The literature also provides evidence where the random walk model is found to perform very well in prediction (Armstrong, 1986; Schnarrs and Bavuso, 1986). In this paper, we attempt to build a forecast by using a random walk model. Furthermore, we investigate the utility of introducing a smoothening constant to stabilise the forecast. It has been found that this helps to shadow the real price movement over the next time period with a reasonable degree of accuracy.

## 2. LITERATURE REVIEW

The present study seeks to review the application of a random walk model in different situations and regions while highlighting some (not all) relevant and seminal work on the topic in order to call attention to pertinent issues. While random walk is considered a standard model of entirely random and irregular behaviour, its application includes methodological differences. Since the methodologies involved differ, the results achieved through the application of such methodologies also vary. It is also noteworthy that the efficacy of the random walk approach depends on the sample of stock or data to which it is applied. In this study, the author seeks to apply the random walk approach to the Indian context and develop a forecasting model. A review of the literature (in chronological order) depicting various contexts in which random walk has been applied is given in the following passages. Studies on the random nature of stock market prices may be traced back as far as the 1960s. Levy (1967) stated that empirical evidence for non-randomness was missing, and sought to disprove the random walk hypothesis to back select tenets of technical analysis while upholding norms of academic evidence. Cochrane (1988) offered a measure of persistence of fluctuations in GNP built on the variance of its long differences. The measure revealed slight long-term persistence in GNP. While prior studies on the subject established significant persistence in GNP and suggested models such as random walk, reconciling the results of Cochrane (1988) study with those of past research showed that standard criteria for time series model building could give deceptive approximations of persistence. Eckbo and Liu (1993) modelled stock prices as a total of a random walk and a general stationary (predictable) component, and suggested an estimable lower bound on the proportion of total stock return variance caused by the predictable component. The lower bound thus proposed fairly estimated the true variance proportion in finite samples also when the temporary component did not adhere to a first-order autoregressive process. The value-weighted market portfolio displayed generally less significant variance proportion estimates in the study.

Au *et al.* (1997) stated that ground-breaking developments in option pricing theory by Black, Scholes, and Merton, and the swift growth of derivative securities in the financial market made it important for finance students to understand the relatively little known stochastic process and geometric Brownian motion. They linked the intuitive discrete time random walks with their corresponding continuous time limits. In their study, stock price movements were illustrated through a logarithmic random walk to assign discerning meaning to the phrase, "geometric Brownian motion". Chen (1999) sought to find an advanced exchange rate forecasting model that could surpass, with respect to the mean square error, the random walk at short-run horizons. The study adopted the Smooth Transition Autoregressive (STAR) and Exponential Generalized Autoregressive Conditional Heteroscedastic in Mean (EGARCH-M) model, which incorporated non-linearity.

Jabbari *et al.* (2001) developed models to characterise time until boundary crossing and associated statistics in cellular wireless networks. They suggested modelling terminal movements in a cell through a discrete two-dimensional random walk process. Furthermore, they determined the time until crossing an exit point from a circular cell by selecting a random direction between starting and exit points. Lai *et al.* (2002) examined predictability of technical trading rules on daily returns of the Kuala Lumpur Stock Exchange Composite Index from January 1977 to December 1999, and found non-randomness of successive price changes.

Using Lo and MacKinlay (1988) variance ratio tests, Rashid (2006) examined the random walk hypothesis for the Pakistani foreign exchange market by considering five pairs of nominal exchange rate series weekly over approximately 10 years. The study found that nominal exchange rates followed random walks. Balsara *et al.* (2007) rejected the random walk null hypothesis for class A and class B stock market indices traded on the Shanghai and Shenzhen stock exchanges with the help of a variance ratio test. The study further found that ARIMA forecasting model forecasted more accurately as opposed to the naïve model founded on the random walk assumption. Jarrett (2008) attempted to examine capital market efficiency in the context of securities traded on organised Hong Kong markets and identified predictable short-term properties in data considered.

Lim and Brooks (2010) used the rolling bivariate test to determine the extent of nonlinear departures from a random walk for aggregate stock price indices of fifty countries pertaining to the years 1995–2005. The results indicated that in countries with lower per capita GDP, stock markets generally witnessed price deviations with greater frequency. The reason behind this appeared to be cross-country variation in the extent of private property rights protection. Lim and Brooks (2010) further opined that inadequate protection deterred informed investors from participating, resulting in sentiment driven noise traders dominating the market. Trading by such investors led stock prices in emerging economies to diverge from random walk standards for prolonged periods of time.

Lakshmi and Roy (2012) examined the Indian equity market for random movements in stock indices by testing random walk hypotheses in daily, weekly and monthly returns of six Indian stock market indices (including Nifty, CNX Nifty Junior, NSE 500, SENSEX, BSE 100 and BSE 500) from January 2000 to October 2009. The study found no random movements in share indices. Furthermore, mixed results were observed when Lo and MacKinlay (1988) applied the variance ratio test with assumptions of homoscedasticity and heteroscedasticity. At times, heteroscedasticity was found to beget non-random behaviour in share indices.

Vanderbei *et al.* (2013) used linear programming duality to resolve optimal stopping problem of a perpetual American option (both call and put) in discrete time while assuming that underlying stock prices followed discrete time and discrete state Markov processes - a geometric random walk. The pricing problem was formed as an infinite dimensional linear programming (LP) problem with the help of excessive-majorant property of the value function. It was discovered that for the call option, such critical values existed in a few cases only, and were dependent on the order of state transition probabilities and the economic discount factor (the interest rate prevailing). However, it was not a concern for the put. Kung and Carverhill (2012) sought to determine, using bootstrap technology, whether the Nikkei 225 evolved with time according to four generally used processes for estimating stock prices: random walk with a drift, AR(1), GARCH(1,1), and GARCH(1,1)-M. It was found that of the four processes, GARCH(1,1)-M gave returns that were most aligned with those estimated from the actual Nikkei series. Bacry *et al.* (2012) proposed a continuous-time stochastic process to satisfy an exact scaling relation, including odd-order moments, thus suggesting a continuous-time model for the price of a financial asset that reflected most major stylised facts observed on real data, including asymmetry and multifractal scaling. Rossi (2013) sought to identify and illustrate, through a literature review and empirical analysis, which variables (if any) forecasted exchange rates, and proposed new methodologies. Predictability was most apparent when one (or more) of the following held: the predictors were Taylor rule or net foreign assets, the model was linear, and a small number of parameters were estimated. The most difficult benchmark was the random walk without drift. Chitenderu *et al.* (2014) examined the Johannesburg Stock Exchange for presence of random walk hypothesis employing monthly time series of All Share Index (ALSI) for the years 2000–2011. It was found that ALSI bore resemblance to a series that followed the random walk hypothesis with significant proof of wide variance between forecasted and actual values, which suggested that the series had weak or no forecasting strength.

Kim and Seo (2015) investigated the influence of transaction costs on market efficiency and price discovery in the EU ETS. They discovered that transaction costs did affect mean reversion behaviour, limit market efficiency and price discovery. Tanrıöver and Çöllü (2015) tested weak form efficiency within the random walk model structure considering price movements of the BIST-100 Index, and assessed forecasting performance of investors in the Turkish stock market. The results revealed that investors in the Turkish capital market could forecast on the basis of historical stock price movements.

The discussion above highlights the various contexts in which the random walk approach has been used. Our study builds a forecast for a time series of stock prices on the basis of stock values pertaining to a period of 242 days. To this end, underlying patterns in a dataset of stock prices were observed and it was found that prices followed random behaviour.

## 2.1. Theory

Studies employing a random walk model to forecast stock prices have produced mixed results. Furthermore, there is evidence in the literature that certain time series models might have better mean forecasts. Studies have also found that the performance of the random walk model is contextual and a function of the underlying nature of data points. This implies that the model warrants testing in various contexts to determine its efficacy according to context. Modifications to the model may help overcome certain shortcomings, which might also enhance the model's predictive ability. Considering the above, we analyse time series stock price data to check on runs, randomness and lags, and develop a modified forecasting model of random walk which is applied on the daily stock prices of different firms. To capture the drift, the model is modified to account for the stabilisation effect. The modification in random walk model enables a self-correcting feature that inherently checks for drifts and accordingly 'corrects' the forecast for the next period. This is a significantly useful feature of the model as daily traders and investors might use it to anticipate changes in stock prices in the immediate future as a result of activity in the immediate past. We have taken the approach from exponential smoothing in time series models and interweave this methodology in a random walk model. We check for mathematical accuracy by deriving equations from the random walk model and validate the model by running several tests. The tests and their results have been described in the results and discussion section. Forecast accuracy has been tested through unit independent measure of forecast accuracy MAPE (Mean Absolute Percentage Error). Predicted values and actual values have been presented in a visual format through spreadsheet applications. Our analysis shows that the modified random walk model applied to stock price data in the Indian context works very well. It also validates the contention that the model must be tested in various contexts to establish its efficacy. This model can also be productised and used by traders and investors in combination with other tools and techniques to predict stock prices. Thus, the present study makes a significant contribution to existing body of literature and carries important implications for practitioners, traders and investors.

## 2.2. Proposition

We address the issue of forecasting time series of stock prices for which we have the information on their past values over the full financial year period. Such information is obtainable from public domain as the stock prices of a company are quoted in various financial dailies and databases. We have selected the growth sectors from Indian equity market that are expected to grow in the near future from the perspective of investment decisions. We have relied on the information given by research agencies such as S&P to identify such sectors. The sectors identified include Cements, ITES, FMCG and Commercial Vehicles. Within each sector, leading companies were taken as investment choices. We have taken a full financial year of longitudinal data of 2014–15 for each of these companies. All of these are listed companies whose share price data were extracted from Bloomberg.

Figure 1 shows time series plot of the closing prices from 1 April 2014 to 31 March 2015 for ACC Ltd, which is traded on the stock exchange in India. Refer to Appendix for the stock price information for the identified companies. We are thus posed with the question of how future value of these prices should be forecast. What are the other underlying patterns and is the time series really random as it seems to be?

A first cut analysis through visual inspection shows that the series itself is not random as a whole since there is a gradual upward trend, especially in the last quarter. Hence, the original series is not random. However, the series does show meandering tendency throughout, i.e., the difference series shows randomness as seen visually from the scatter plot of difference series in Figure 2.

### 3. METHODOLOGY

A way to check for randomness in a time series of prices is to examine autocorrelations. Many forecasting techniques are based on a specific autocorrelation structure of a time series of prices. Such a structure can tell us how prices are related to their own past values in a time series. If this structure is known, it will provide valuable insights for designing a forecast. If successive observations of prices are correlated with one another, then a series cannot be random. For example, if a time series is such that large values of observation follow large values or small values follow small values then this is a kind of positive autocorrelation. However, there can be negative autocorrelation, and even the lags between autocorrelated values can differ. To begin with, we test for randomness of time series using the autocorrelation test. To test for randomness, we first create a lagged series with different lags. Lags are simply past observations, removed by a certain number of time periods from the present time. Autocorrelation of lag  $m$  is essentially the correlation between original series and lagged- $m$  version of the series. We have limited ourselves to 32 lags, which is no more than 25 percent of the number of observations. Given that there is no seasonality, we find 32 lags to be sufficient for capturing the test for autocorrelation. Examination of the autocorrelation coefficient and the standard errors in Figure 3 indicates that autocorrelation is significant up to 25 lags. We can see that autocorrelation coefficient in these cases is more than two standard errors in magnitude, which helps us to conclude that a time series is not random.

At the next stage, we test for autocorrelation in the difference in stock prices over consecutive time periods. In this case, we basically test the hypothesis that changes in prices may or may not be correlated. Evidently, it can provide a glimpse into underlying pattern of stock prices that can help us forecast better.

For this, we take differenced series. Each value in differenced series is obtained as under:

$$C(p)_{t+1} - C(p)_t = c(dif)_{t+1}$$

$C(p)_{t+1}$  = Closing price in time period  $t+1$

$c(dif)_{t+1}$  = Difference in closing price in time period  $t+1$

We create differenced series and apply autocorrelation test on this series. We created lagged series for closing price difference and calculated autocorrelation of lag  $m$ , i.e., the correlation between original differenced series and lagged- $m$  version of the differenced series. Autocorrelation is checked for close price difference up to 32 lags. The number of observations in this case is 242, so 32 lags work out to be sufficient. Examination of the autocorrelation coefficient and the standard error as evident from Figure 4 and Table 1 suggests that autocorrelation is not significant in any of the lags. It is less than two standard errors in magnitude except for lag#5 which is slightly more than two standard error in magnitude; thus, this lag#5 can be ignored. This leads us to conclude that differenced time series is random.

To support the check for randomness we also apply a runs test to the differenced series. First, we determine the number of runs in different time periods. We choose a base value, which is equal to average value of time series.

Then, we define a run as a consecutive series of values that are at one side of the base level. The number of observations above the cut-off is 116 and those below the cut-off is 126 as seen in Table 2.

If the P-value is sufficiently small ( $\leq 0.05$  for 95% confidence level) then we can reject the null hypothesis of randomness and conclude that the series does not alternate enough (too few runs) or alternates too much (too many runs). In this case, the number of runs (126) is very close to the expected number of runs  $E(R) = 121.7934$ . The p value (2-tailed) = 0.5872 is much greater than 0.05, which implies that we cannot reject the null hypothesis of randomness, i.e., the number of runs is equal to the expected number of runs  $E(R)$ ; therefore, we conclude that the series is random. In the plot of differences in Figure 3, we can see that differences do not vary around a mean of zero; rather, these are actually moving across the mean of 0.84, and there is an upward drift. Evidently, observation in time series plots follows similar behaviour. The series itself is not random, but the changes from one period to next are random. It is interesting to test this phenomenon by the autocorrelations and runs test we have completed so far. However, if we were to forecast stock prices over the next few days, we cannot really use the average of past values as a forecast given that it may be either too low or too high. Stock prices may follow a trend and, in this case, the forecast will either undershoot or overshoot the actual value of the stock price. In such situations, we may be more prudent if we were to base our forecast on the most recently observed values. The closing price difference series is of  $N$  periods of length  $t$ . Therefore, we can define an additive process  $y$  by:

$$y(t_k) = y(t_{k-1}) + \xi(t_{k-1})\sqrt{\Delta t} \quad t_k = t_{k-1} + \Delta \quad (1)$$

for  $k = 1, 2, 3, \dots, N$ . This process is termed as a random walk.

$\mathcal{E}(t_{k-1})$  is a normal random variable with mean 0 and variance 1 - a standardised normal random variable. The process is started by setting  $y(t_1) = 0$  after which a path emerges that meanders around depending upon the chance of random variables.

Differenced random variables can be written as:  $y(t_k) - y(t_{k-1})$ . Such a difference is related to the standardised normal variable  $\xi(t_i)$ :

$$y(t_k) - y(t_{k-1}) = \sum_i^{k-1} \xi(t_i)\sqrt{\Delta t} \quad (2)$$

$$E[y(t_k) - y(t_{k-1})] = 0 \quad (3)$$

Difference random variables are found to be normally distributed and have a mean of 0 as equation (3) shows. Most importantly, if difference variables made up of different  $\xi(t_i)$  are independent, we find that:

$$\Psi[y(t_k) - y(t_{k-1})] = E[\sum_i^{k-1} \xi(t_i)\sqrt{\Delta t}]^2 \quad (4)$$

where  $\Psi[y(t_k) - y(t_{k-1})]$  = variance of  $[y(t_k) - y(t_{k-1})]$

$$= E[\sum_i^{k-1} \xi(t_i)^2 \Delta t] \quad (5)$$

$$= t_k \rightarrow t_{k-1}$$

It is clear that variance of  $y(t_k) - y(t_{k-1})$  is exactly equal to 1 if the time difference between two non-overlapping intervals is also 1.

The functional form of random walk model can be given by equation (6).

$$Y_t = Y_{t-1} + \mu + \xi(t_i) \quad (6)$$

where

$\mu$  = mean of differences and represents the expected value of differenced variable

$\xi(t_i)$  = random series (noise) with mean zero and standard deviation  $\sigma$  that remains constant with time.

$Y_t$  = observation in time  $t$ .

If  $Y_t$  is the change in series from time  $(t)$  to time  $(t-1)$  at  $t$ , then equation (6) can be written as:

$$Y_t = \mu + \xi(t_i) \quad (7)$$

It is apparent that the series tends to have an upward trend if  $\mu > 0$  and will have a downward trend if  $\mu < 0$ . Basically, we add the estimated trend to the current value in order to forecast the next value.

$$F_{t+1} = Y_t + Y_D^- \quad (8)$$

To forecast future closing prices, we use the following equation:

$$F_t = Y_{t-1} + \mu + \xi(t_i) \quad (9)$$

where

$F_t$  = forecast in time period  $t$ .

We measure forecast accuracy by taking standard error of forecasting  $k$  periods. We compute the standard error using following equation:

$$SE_k = \sigma \sqrt{k}$$

$SE_k$  = Standard error of forecasting  $k$  periods ahead

$\sigma$  = standard deviation of differences

We suggest improving the forecast by factoring the difference in actual price and predicted price in preceding period. We use a stabilisation coefficient to factor in the difference itself and, as a result, we get a revised forecast model as given by the following equation:

$$F_t = Y_{t-1} + \mu + \xi(t_i) + \alpha (A_{t-1} - F_{t-1}) \quad (10)$$

where

$\alpha$  = stabilisation coefficient

$A_{t-1}$  = Actual price in time period  $t-1$

$F_{t-1}$  = Forecast in time period  $t-1$

To validate the modified random walk model modified by smoothening or stabilising coefficient, we apply this modified model on 12 leading stocks of Indian equity stocks from across the industry as described in the next section.

#### 4. DATA ANALYSIS

Companies in the cement sector are ACC, Ambuja, Dalmia, India Cements and Ultra Cements. In the ITES sector, the firms that we have picked up are TCS, Infosys and Wipro. In FMCG and Pharma, we picked one company each as Hindustan Unilever and Sun Pharma, respectively. For auto, we selected two companies - Ashok Leyland Ltd. and Tata Motors Ltd. We previously observed in case of ACC stocks that closing prices follow an upward drift. If forecast is overshooting the actual price, then forecast in the next time period will take into account the difference and modify the forecast value appropriately. In these cases, the forecast for the next time period would tend to reduce its value by the stabilisation coefficient multiplied by the difference. In a similar vein, the forecast is modified if the actual price overshoots the forecast. In both cases, we must grapple with the task of choosing a suitable stabilisation coefficient. This can be any number between 0 and 1. We would be giving a very high credence to the price-forecast difference if we were to choose one as the stabilisation coefficient. The best way forward is to choose ' $\alpha$ ' such that we give an adequate amount of weight to the difference in order to stabilise the forecast. We experiment with different values of ' $\alpha$ ' and check forecast accuracy for each of the experiments over a time period. We present the random walk and modified version on a stock (we chose ACC to illustrate this model) in Figures 5 and 6, respectively. We show the actual values, the predicted values using the random walk model as presented in the paragraphs above in Figure 5. Additionally, we attempted to simulate the movement of stock prices at different time periods and observed their deviation vis-à-vis predicted values. We chose 44 days as the time horizon. The replication was done 25 times and the expected value for each day (period) was computed and presented in the figure. We find that the expected value approaches closer to the predicted value of a random walk as we move closer to 25 replications. However, the simulated values that we obtain in any given run fall within a 95% confidence interval.

The same is also depicted in the figure. We also found that predicted values could be brought closer to the actual values if we could probe further into the model to minimise the error. For the sake of better visibility in this approach, we have zoomed in on the prediction horizon of 44 days, showing the actual and predicted values and presented them in a separate figure (Figure 6). Additionally, the modified random walk model prediction is also shown here.

Next, we wanted to see the performance of predicted values against different values of stabilisation coefficient ( $\alpha$ ). We start the experimentation by keeping  $\alpha = 0.2$  and then increased it systematically to 0.5, 0.8 and 1, respectively, in each experiment. That way we have a total of 5 experiments including one in which  $\alpha = 0$ . Figure 7 depicts different line plots for different values of  $\alpha$ .

We measure the efficacy of the forecast model by comparing one-period ahead forecast  $F_{t-1}$  from the model and compare it to the known or actual values,  $Y_t$ , for each  $t$  in the future time period. We report Mean Absolute Percentage Error for different forecast models (with different values of  $\alpha$ ), which is given by

$$MAPE = \sum_t^T \frac{|\varepsilon_t|/A_t}{N} \quad (11)$$

where  $N$  = number of future time periods in which forecast and actual values are compared. In each case, we computed MAPE and compared it with predicted values (with different  $\alpha$ ) and evaluated as to which forecast was working better. We present this analysis in Table 3, which shows the forecast values against actual prices for different values of the stabilisation coefficient.

It is obvious from the chart that the forecast with  $\alpha = 1$  is performing better as most forecast values are meandering very close to actual prices.

We select this stabilisation coefficient value to prepare our forecasts. We also note that the forecast profile in this case closely matches the actual values. This leads us to interpret that the forecast is largely able to capture the random behaviour of these stock prices.

#### 4.1. Discussion

We find that successive price movements of stocks follow a random walk movement; we have used this assumption in applying the random walk model and its modified form to test its validity for another 11 equity stocks for the 44 days of the horizon period based on these companies' full-year stock price data. The abridged analysis is as below:

- All the stocks' predictions show that modified random walk predictions closely follow the actual price movement. In fact, they shadow the actual price movement unlike the linear trend line of random walk model for the predicted mean price.
- Whereas accuracy of the random walk model in terms of MAPE for the various stocks ranges from 3.6% to 14.5%, the same value for modified random walk model is 1.2% to 3.74%; the latter consistently outperforms the former in each case as seen from Figure 8.
- Stocks predictions of ACC, Ambuja, Ashok Leyland, Sun Pharma FY 14-15, TCS and Wipro show a similar nature with both the models converging at the end as seen from the representative case of Ashok Leyland in Figure 9. Random walk predictions for Ambuja, Dalmia, India Cements, Ultra Cements, Hindustan Unilever Ltd, TML, Sun Pharma FY 2015-16 and Infosys Company show that the random walk model for mean value are drifting away from the actual price, especially for the last month of the data as seen in the representative figure of HUL as below in Figures 10(a) and 10(b).

This drift in random walk model is also reflected by higher forecast error -MAPE of 7% to 11%. This is because the random walk model was based on the year-long trend, which may be different from temporal trends and especially to that of prediction horizon. On the other hand, it can be seen that the modified model is still able to



closely follow actual price movement in the entire prediction horizon for all the stocks with the MAPE in the range of 1.5 to 3.75% only (Figure 8). Two datasets of Sun Pharma for FY 1014-15 and FY 2015-16 were taken; these displayed the opposite year-long trend in these two years. FY 2014-15 has an increasing price trend whereas FY 2015-16 has a decreasing trend. In both cases, a modified random walk model is able to shadow the actual price movement (Figure 11(a) and Figure 11(b)). Whereas the forecast error by random walk is observed with higher forecast error (MAPE-9.2%) – especially in FY 2015-16 – the modified random walk gave a forecast error of only MAPE of 1.8- 1.9% (Figure 8). Furthermore, ITES stocks of Infosys, TCS and Wipro were studied. We can see that even though there was stock split on 2 December 2014 for Infosys, a modified random walk model was still able to self-correct and follow the actual price movement. It should be noted that we have used the entire year-long data without treating the stock price differently before and after the split (Figure 12).

#### 4.2. Managerial Implications

We adopt a random walk approach to make a predictive model that captures the random nature of the dataset. The forecast is stabilised using a proper choice of coefficients in a revised forecasting model and checked for forecast accuracy. Forecast accuracy is evaluated by comparing the forecast with actual values and is found to be reasonably good in terms of mean absolute percentage error. It could be observed from the discussion of results that the predicted value for successive time periods tends to closely follow the actual values over the predicted window period. This is unlike the standard random walk model which predicts the next period's value being the same as current period's value plus the trend value (if any) and the error term added to that. With the revised model, we have been able to shadow the actual price in all cases with improved prediction accuracy. A predictive analytics approach, as presented in this paper, is wide-ranging in nature and can also be applied to predict uncertainty in stock prices of all equity markets, interest rates and even macro-economic indicators. Thus, the suggested approach could be of good utility for business traders for predicting daily stock prices even without the knowledge of complex extrapolation models, econometric models or alternatives such as technical analysis.

#### 5. CONCLUSION

We have examined a decision problem related to predicting the stock price for next-day trading. We have shown the modified random walk model gives an average accuracy of 1.5% for different stocks of firms that we selected for analysis. In the modified random walk model, introduction of stabilisation coefficient essentially helps to simulate the errors terms of the random walk model thus improving the prediction accuracy. This simple model and its simplicity of application with reasonably good accuracy can provide good utility to ordinary stock traders for the purposes of daily trading. Although this concept has been applied only to equity stocks, this model can be extended to any stochastic time series data such as exchange rates, futures, and forecasting for supply chain decisions. We observed that the strength of this model lies in its flexibility and practical approach. The simple skill in using this model over other complex models could be worth applying in various stochastic data series of any normal day-to-day business.

**Funding:** This study received no specific financial support.

**Competing Interests:** The authors declare that they have no competing interests.

**Contributors/Acknowledgement:** All authors contributed equally to the conception and design of the study.

#### REFERENCES

Armstrong, S., 1986. Research on forecasting: A quarter-century review. *Interfaces*, 16(1): 89-103.

- Au, K.T., M. Raj and D.C. Thurston, 1997. An intuitive explanation of brownian motion as a limit of a random walk. *Journal of Financial Education*: 91-94. [View at Google Scholar](#)
- Bacry, E., L. Duvernet and J.F. Muzy, 2012. Continuous-time skewed multifractal processes as a model for financial returns. *Journal of Applied Probability*, 49(2): 482–502. [View at Google Scholar](#) | [View at Publisher](#)
- Balsara, N.J., G. Chen and L. Zheng, 2007. The Chinese stock market: An examination of the random walk model and technical trading rules. *Quarterly Journal of Business & Economics*: 43-63. [View at Google Scholar](#)
- Chen, Y., 1999. Can the star or the EGARCH-M model outperform the random walk model for short run exchange rate forecasts? The case of Taiwan and Japan. Dissertation, The University of Hong Kong.
- Chitenderu, T.T., A. Maredza and K. Sibanda, 2014. The random walk theory and stock prices: Evidence from Johannesburg stock exchange. *International Business & Economics Research Journal*, 13(6): 1241. [View at Google Scholar](#) | [View at Publisher](#)
- Cochrane, J.H., 1988. How big is the random walk in GNP? *Journal of Political Economy*, 96(5): 893-920. [View at Google Scholar](#) | [View at Publisher](#)
- Eckbo, B.E. and J. Liu, 1993. Temporary components of stock prices: New univariate results. *Journal of Financial & Quantitative Analysis*, 28(2): 161–176. [View at Google Scholar](#) | [View at Publisher](#)
- Jabbari, B., Y. Zhou and F.S. Hillier, 2001. A decomposable random walk model for mobility in wireless communications. *Telecommunication Systems*, 16(3-4): 523–537. [View at Google Scholar](#)
- Jarrett, J.E., 2008. Random walk, capital market efficiency and predicting stock returns for Hong Kong exchanges and clearing limited. *Research News*, 31(2): 142–148. [View at Google Scholar](#) | [View at Publisher](#)
- Kim, J. and B. Seo, 2015. Transaction costs and nonlinear mean reversion in the EU emission trading scheme. *Hitotsubashi Journal of Economics*: 281–296. [View at Google Scholar](#)
- Kung, J.J. and A.P. Carverhill, 2012. A bootstrap analysis of the Nikkei 225. *Journal of Economic Integration*, 27(3): 487–504. [View at Google Scholar](#) | [View at Publisher](#)
- Lai, M.M., K.G. Balachandher and F.M. Nor, 2002. An examination of the random walk model and technical trading rules in the Malaysian stock market. *Quarterly Journal of Business & Economics*: 81–104. [View at Google Scholar](#)
- Lakshmi, V.D. and B. Roy, 2012. Testing the random walk model in Indian stock markets. *IUP Journal of Applied Finance*, 18(2): 63. [View at Google Scholar](#)
- Levy, R.A., 1967. Random walks: Reality or myth. *Financial Analysts Journal*, 23(6): 69–77. [View at Google Scholar](#) | [View at Publisher](#)
- Lim, K.P. and R.D. Brooks, 2010. Why do emerging stock markets experience more persistent price deviations from a random walk over time? A country-level analysis. *Macroeconomic Dynamics*, 14(S1): 3–41. [View at Google Scholar](#) | [View at Publisher](#)
- Lo, A.W. and A.C. MacKinlay, 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. *Review of Financial Studies*, 1(1): 41–66. [View at Google Scholar](#) | [View at Publisher](#)
- Rashid, A., 2006. Do exchange rates follow random walks? An application of variance-ratio test. *Pakistan Economic & Social Review*: 57–79. [View at Google Scholar](#)
- Rossi, B., 2013. Exchange rate predictability. *Journal of Economic Literature*, 51(4): 1063–1119. [View at Google Scholar](#)
- Schnarrs, S. and J. Bavuso, 1986. Extrapolation models on very short-term forecasts. *Journal of Business Research*, 14(1): 27-36. [View at Google Scholar](#) | [View at Publisher](#)
- Tanrıöver, B. and D.A. Çöllü, 2015. Analysis of forecasting performance of investors in Turkey within framework of the random walk model (Türkiye’de Yatırımcıların Öngörü Performanslarının Rassal Yürüyüş Modeli Çerçevesinde Analizi). *Business & Economics Research Journal*, 6(2): 127–139. [View at Google Scholar](#)
- Vanderbei, R.J., M.Ç. Pınar and E.B. Bozkaya, 2013. Discrete-time pricing and optimal exercise of American perpetual warrants in the geometric random walk model. *Applied Mathematics & Optimization*, 67(1): 97–122. [View at Google Scholar](#) | [View at Publisher](#)

**BIBLIOGRAPHY**

Ming, L.M., F. Nor and B. Guru, 2000. Technical analysis in the Malaysian stock market: An empirical evidence. West Malaysia: Multimedia University.

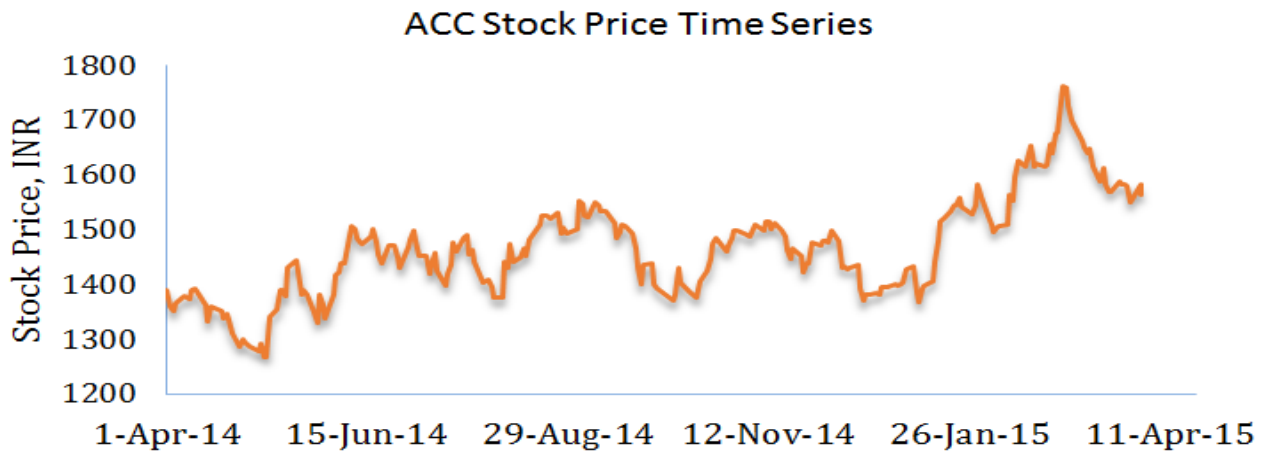


Figure-1. Scatter plot of ACC Stock Price of FY 2014-15

Source: Bloomberg

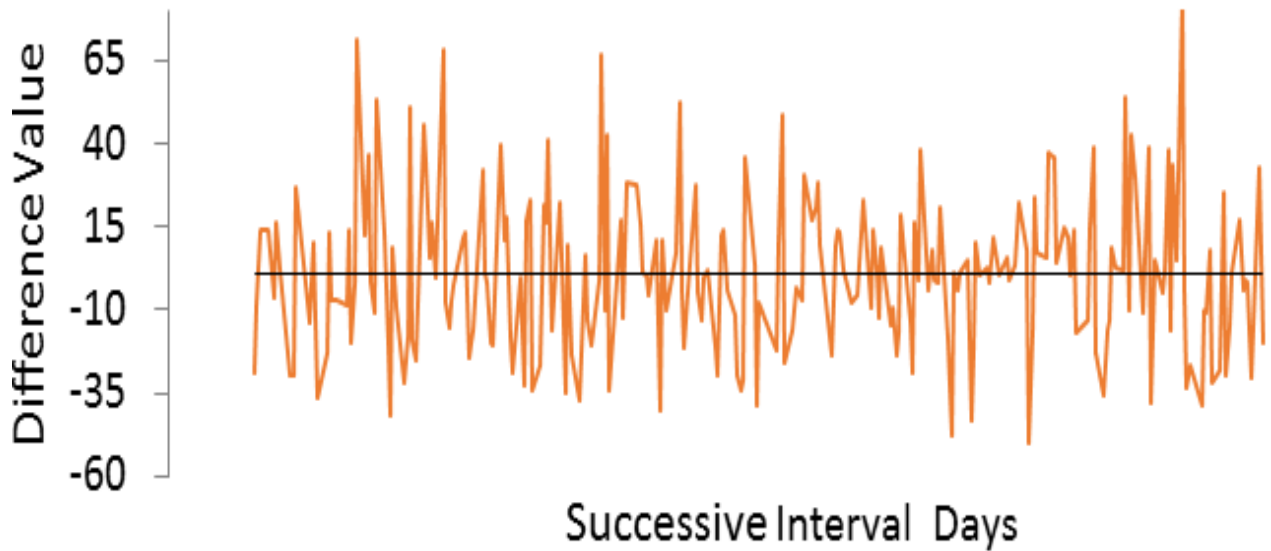


Figure-2. Scatterplot of Difference Series from Time series from figure 1.

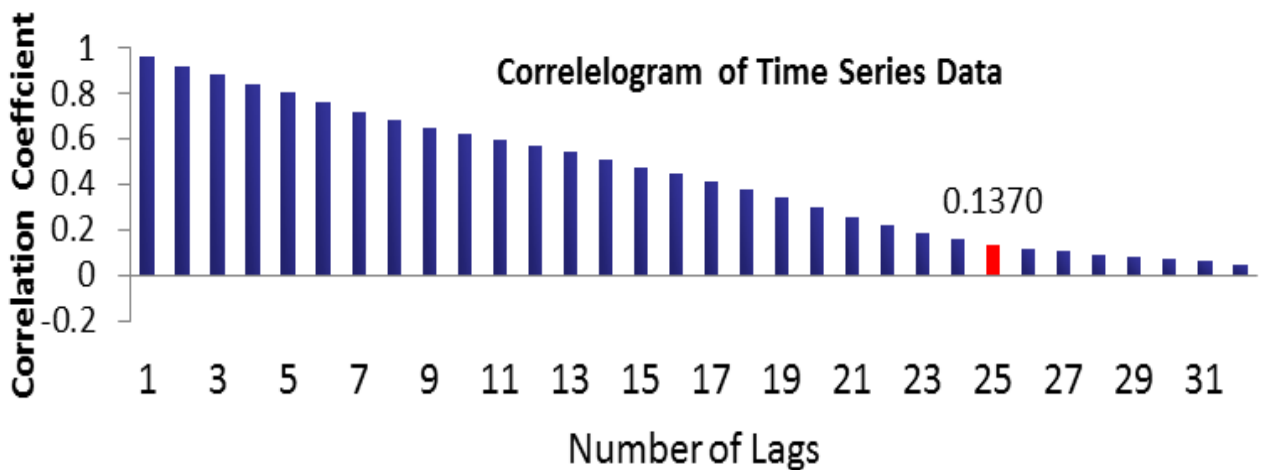


Figure-3. Autocorrelation of time series data of the ACC stock price. Standard Error =0.0642. Correlations is significant up to Lag 25 (0.137) >2 times of Standard Error (0.1284).

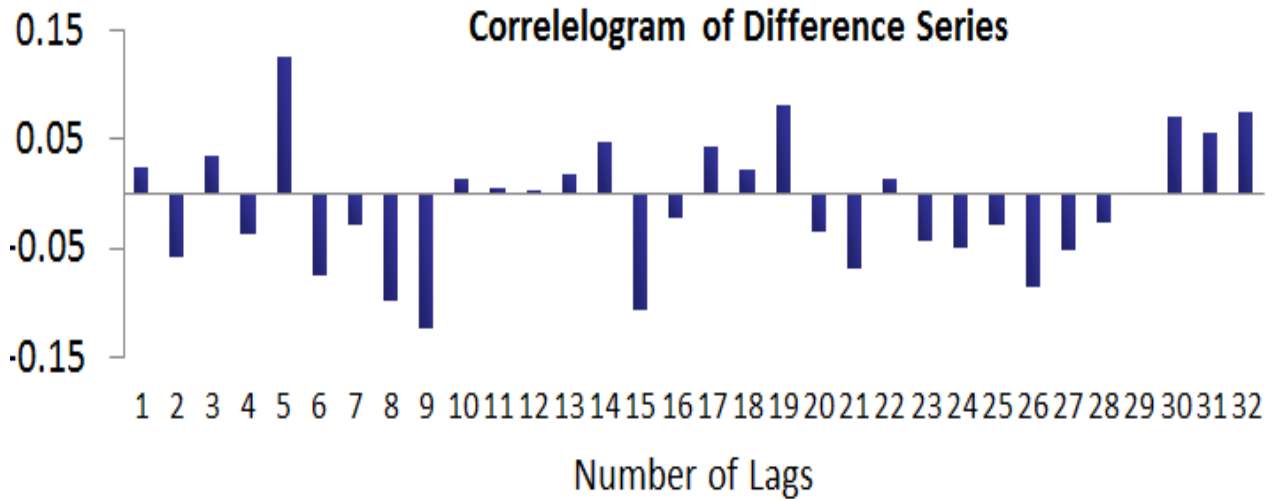


Figure-4. Autocorrelation of Difference Series of the ACC stock price

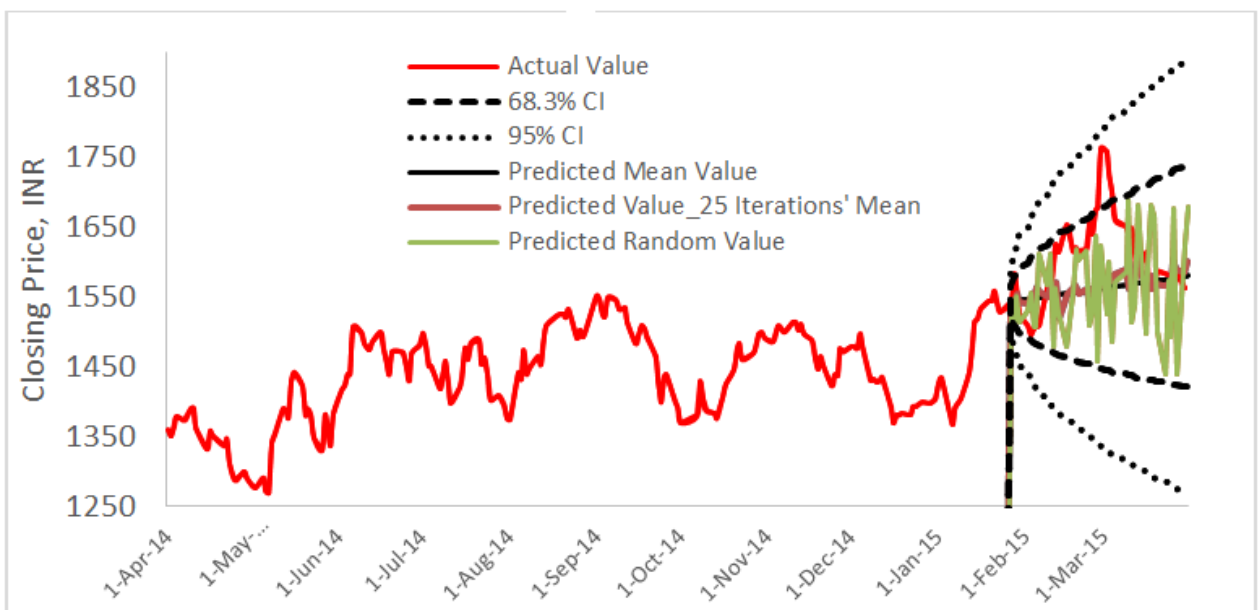


Figure-5. Random Walk Prediction of Last 44 days- Randomised prediction, Mean, Spread Window with confidence intervals

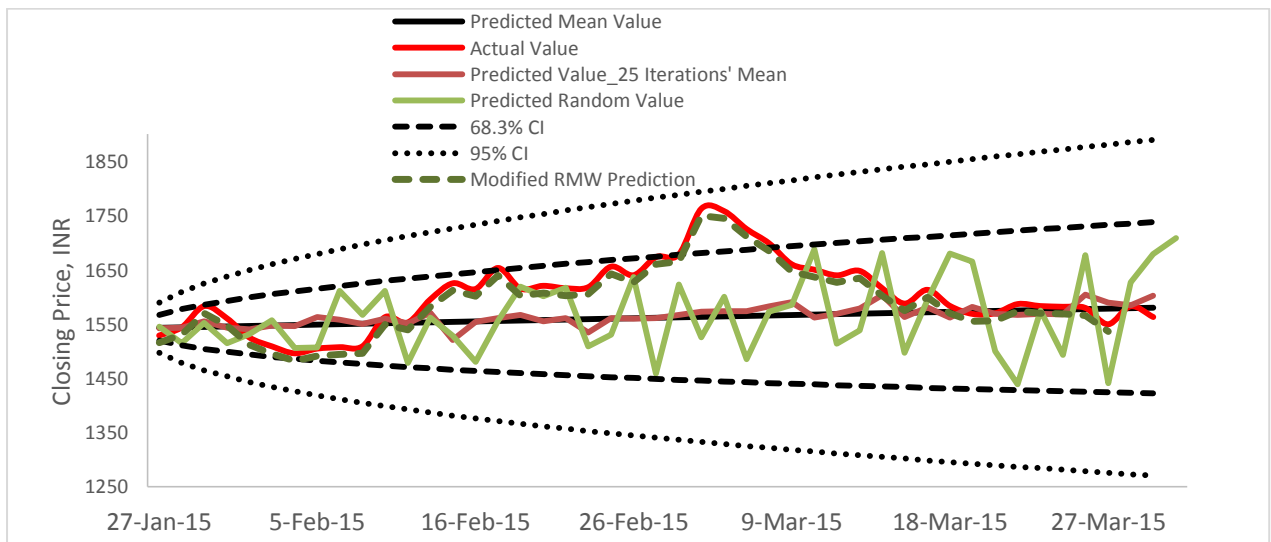


Figure-6. Modified Random Walk Shadowing the Actual Price Movement (zoomed view of prediction window of figure 5 )

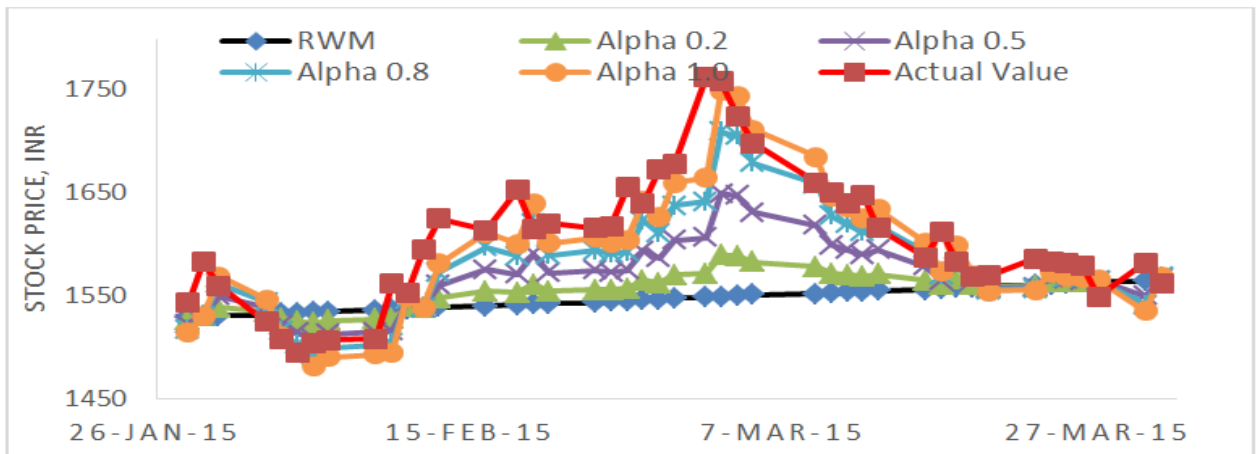


Figure-7. Predicted Values with Different Values of Smoothing/Stabilisation coefficient

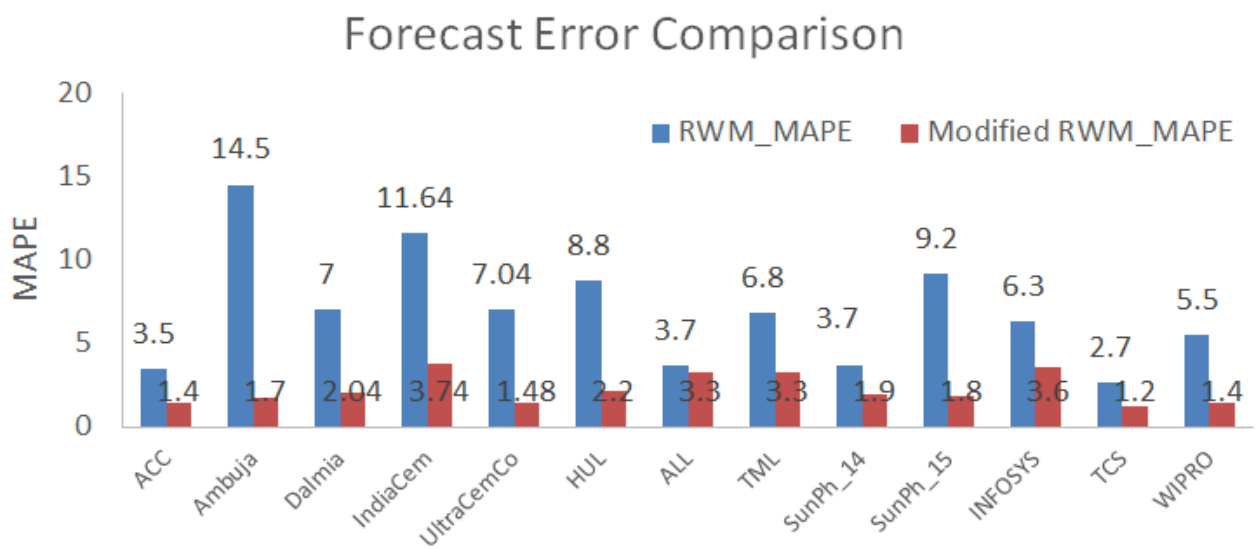


Figure-8. Forecast Error Comparison of Random Walk Model (RMW) vs Modified Random Walk (RMW\_Mod)

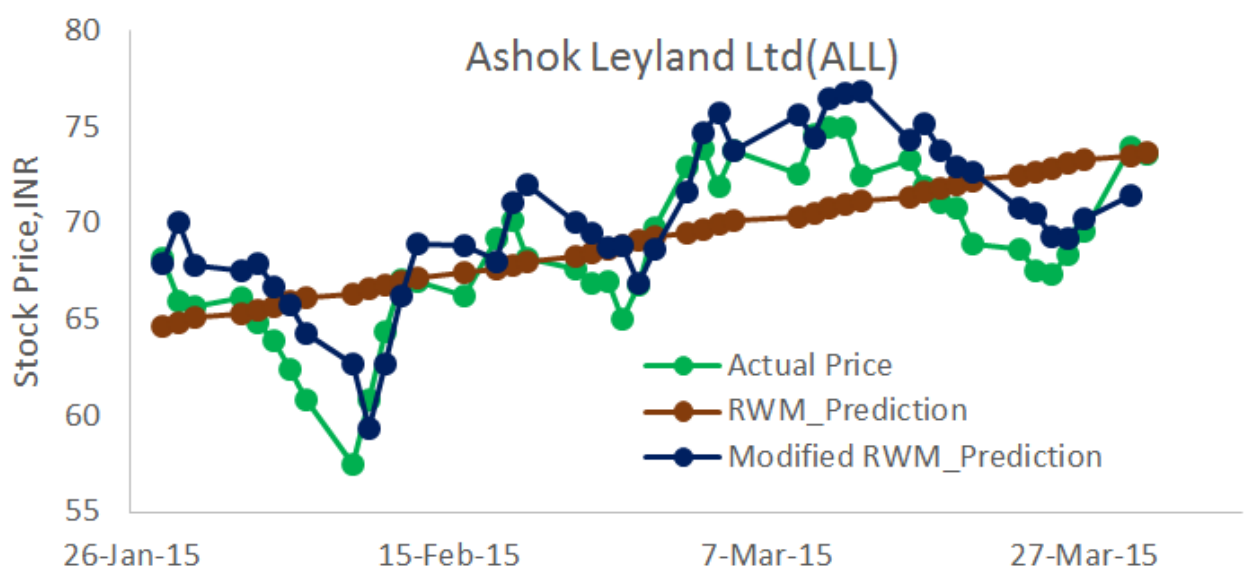


Figure-9. Stock price prediction by both methods show convergence of the values at the end of the prediction window. The modified random walk model, however, shadows actual price movement.

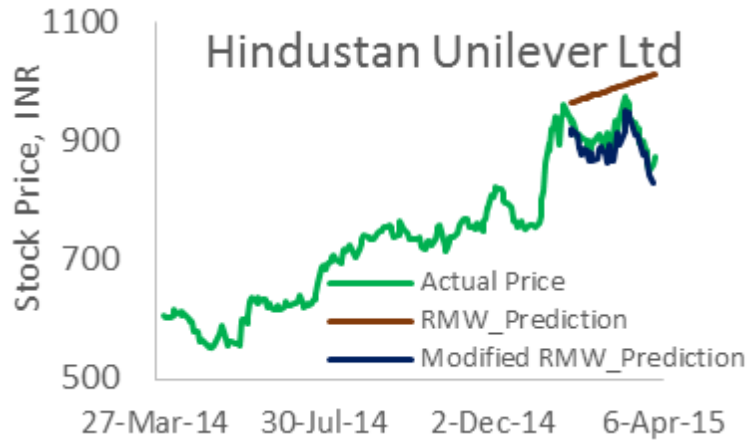


Figure-10(a). Stock Price Prediction of HUL

Source: Bloomberg

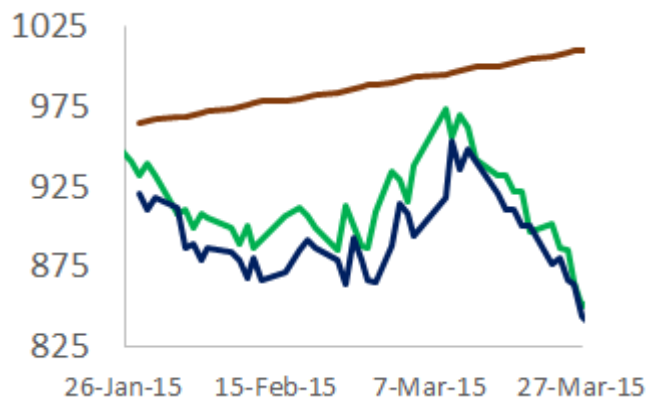


Figure-10(b). Zoomed View of prediction window of HUL stock.

Source: Bloomberg

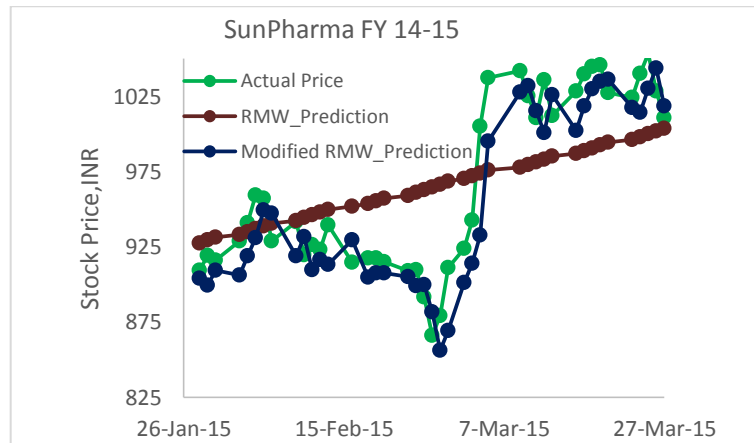
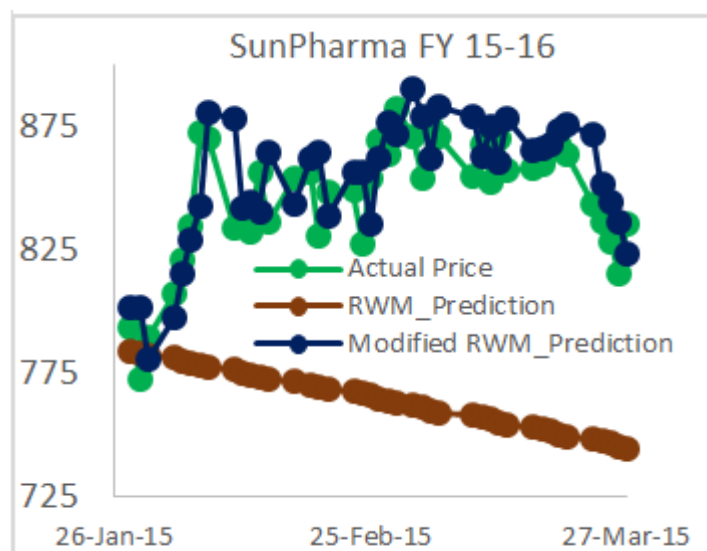
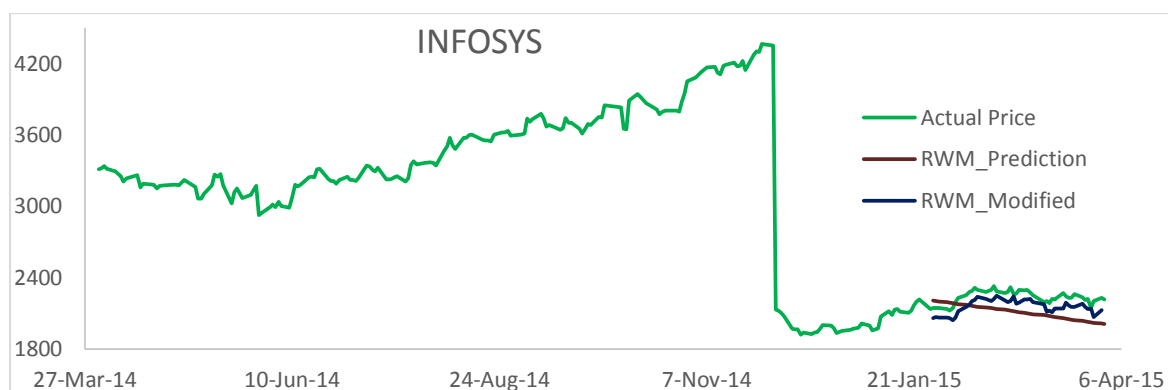


Figure 11(a) Increasing Stock Price.

Source: Bloomberg



**Figure-11(b)** Decreasing Price Trend In FY 2014-15 in FY 2015-16  
 Source: Bloomberg



**Figure-12.** Stock Price Prediction of Infosys with Stock Split Scenario

Source: Bloomberg

**Table-1.** Autocorrelations of Difference Series

Autocorrelation Table	Data Set #1		
Number of Values	242		
Standard Error	0.0643		
Lag #1	0.0230	Lag #17	0.0407
Lag #2	-0.0556	Lag #18	0.0193
Lag #3	0.0385	Lag #19	0.0854
Lag #4	-0.0391	Lag #20	-0.0407
Lag #5	<b>0.1292</b>	Lag #21	-0.0690
Lag #6	-0.0756	Lag #22	0.0295
Lag #7	-0.0354	Lag #23	-0.0401
Lag #8	-0.1068	Lag #24	-0.0422
Lag #9	-0.1182	Lag #25	-0.0288
Lag #10	0.0109	Lag #26	-0.0881
Lag #11	0.0023	Lag #27	-0.0400
Lag #12	0.0024	Lag #28	-0.0234
Lag #13	0.0089	Lag #29	-0.0054
Lag #14	0.0418	Lag #30	0.0616
Lag #15	-0.1052	Lag #31	0.0574
Lag #16	-0.0237	Lag #32	0.0734

**Table-2.** Run Test of Randomness in Difference Series

<i>Runs Test for Randomness</i>	Data Set #1
Observations	242
Below Mean	126
Above Mean	116
Number of Runs	126
Mean	0.84
E(R)	121.7934
StdDev(R)	7.7487
Z-Value	0.5429
P-Value (two-tailed)	0.5872

**Table-3.** MAPEs for Different Values of  $\alpha$

	With Alpha=0	With Alpha=0.2	With Alpha=0.5	With Alpha=0.8	With Alpha=1
MAPE	3.9%	3.3%	2.4%	1.6%	1.4%

*Views and opinions expressed in this article are the views and opinions of the author(s), Asian Economic and Financial Review shall not be responsible or answerable for any loss, damage or liability etc. caused in relation to/arising out of the use of the content.*