A Time Series Analysis - Based Forecasting Approach for the Indian Realty Sector

Jaydip Sen 1

1Praxis Business School

October 30, 2023

Abstract

Prediction of stock prices using time series analysis is quite a difficult and challenging task since the stock prices usually depict random patterns of movement. However, the last decade has witnessed rapid development and evolution of sophisticated algorithms for complex statistical analysis. These algorithms are capable of processing a large volume of time series data executing on high-performance hardware and parallel computing architecture. Thus computations which were seemingly impossible to perform a few years back are quite amenable to real-time time processing and effective analysis today. Stock market time series data are large in volume, and quite often need real-time processing and analysis. Thus it is quite natural that research community has focused on designing and developing robust predictive models for accurately forecasting stochastic nature of stock price movements. This work presents a time series decomposition-based approach for understanding the past behavior of the realty sector of India, and forecasting its behavior in future. While the forecasting models are built using the time series data of the realty sector for the period January 2010 till December 2015, the prediction is made for the time series index values for the months of the year 2016. A detailed comparative analysis of the methods are presented with respect to their forecasting accuracy and extensive results are provided to demonstrate the effectiveness of the six proposed forecasting models.
A Time Series Analysis - Based Forecasting Approach for the Indian Realty Sector

Jaydip Sen
Praxis Business School, Bakrahat Road, Off Diamond Harbor Road, Kolkata – 700104
West Bengal, INDIA
email: jaydip.sen@acm.org

ABSTRACT

Prediction of stock prices using time series analysis is quite a difficult and challenging task since the stock prices usually depict random patterns of movement. However, the last decade has witnessed rapid development and evolution of sophisticated algorithms for complex statistical analysis. These algorithms are capable of processing a large volume of time series data executing on high-performance hardware and parallel computing architecture. Thus computations which were seemingly impossible to perform a few years back are quite amenable to real-time time processing and effective analysis today. Stock market time series data are large in volume, and quite often need real-time processing and analysis. Thus it is quite natural that research community has focused on designing and developing robust predictive models for accurately forecasting stochastic nature of stock price movements. This work presents a time series decomposition-based approach for understanding the past behavior of the realty sector of India, and forecasting its behavior in future. While the forecasting models are built using the time series data of the realty sector for the period January 2010 till December 2015, the prediction is made for the time series index values for the months of the year 2016. A detailed comparative analysis of the methods are presented with respect to their forecasting accuracy and extensive results are provided to demonstrate the effectiveness of the six proposed forecasting models.

Key words: Time Series Decomposition, Trend, Seasonal, Random, Holt Winters Forecasting model, Auto Regression (AR), Moving Average (MA), Auto Regressive Integrated Moving Average (ARIMA), Partial Auto Correlation Function (PACF), Auto Correlation Function (ACF).

JEL Classification: G 11, G 14, G 17, C 63
1. Introduction

Developing an accurate and efficient forecasting model for predicting stock prices has been one of the most exciting challenges confronting the research community working in the field of machine learning, applied econometrics and artificial intelligence. Various technical, fundamental and statistical indicators have been proposed in the literature for predicting stock prices. However, each method has its own limitations and none have been fully effective in predicting the stochastic movement of stock prices. (Sen & Datta Chaudhuri, 2016a; Sen & Datta Chaudhuri, 2016b; Sen & Datta Chaudhuri, 2016c) proposed a novel approach towards portfolio diversification and prediction of stock prices. The authors argued that different sectors in an economy do not exhibit identical pattern of variations in their stock prices. Different sectors exhibit different trend patterns, different seasonal characteristics and also differ in the randomness in their time series. While on one side the efficient market hypothesis has focused on the randomness aspect of stock price movements, on the other side, there are propositions to disprove the hypothesis delving into various fundamental characteristics of different stocks. It may be contended that besides the differences in the fundamental characteristics among stocks of different companies, performances of different stocks also have a lot to do with the sectors to which the stocks belong. Since each sector has its own set of factors influencing its behavior, the price movements of stocks belonging to different sectors are guided by these factors. The factors responsible for the phenomenal growth of the information technology (IT) sector in India are different from those which have made the metals sector in the country sluggish, or the realty sector grow at a slow pace. From the point of view of investors in the stock market, it is critical to identify these factors and analyze them effectively for optimal portfolio choice and also for churning of the portfolio.

In this paper, we focus on the time series pattern of the realty sector in India in order to understand its distinguishing characteristics. We use the monthly time series index values of the Indian realty sector during the period January 2010 till December 2016 as per the Bombay Stock Exchange (BSE). We decompose the time series using R programming language. We, then, illustrate how the time series decomposition approach provides us with useful insights into various characteristics and properties of the realty sector time series. It is further demonstrated that a careful and deeper study of the trend, seasonal and random components values of the time series enables one to understand the growth pattern, the seasonal characteristics and the degree of randomness exhibited by the time series index values. We also propose an extensive framework for time series forecasting in which we present six different approaches of prediction of time series index values. We critically analyze the six approaches and also explain the reason why some methods perform better and produce lower values of forecast error in comparison to other methods.

The rest of the paper is organized as follows. Section 2 describes the detailed methodology used in this work. The method of time series decomposition and of the realty sector time series is explained in detail in this section. Section 3 presents the decomposition results the time series index values into its three components - trend, seasonal and random. The behavior exhibited by the time series is analyzed based on the decomposition results. Section 4 presents a detailed forecasting framework consisting of six different models of forecasting that are applied on the realty sector time series. Section 5 presents extensive results on the performance of the six forecasting methods on the realty sector time series data. A comparative analysis of the techniques is also provided on the basis of six different metrics of the forecasting
techniques: maximum error, minimum error, mean error, standard deviation of error, the root mean square error (RMSE), and the ratio of the RMSE value and the mean index value. Section 6 presents a brief discussion on some of the existing work in the literature on time series forecasting with particular focus on the realty sector. Finally, Section 7 concludes the paper.

2. Methodology

The Section presents a brief discussion on the methodology followed in this work. We use the programming language R for data management, data analysis and pictorial presentation of the results. (Ihaka & Gentleman, 1996) provides a detailed description of various capabilities of R programming language. R is an open source language with a very rich set of libraries having in-built functions that makes it one of the most powerful tools in handling data analytics projects. For the current work, we have used the monthly index data from the Bombay Stock Exchange (BSE) of India for the realty sector for the period January 2010 till December 2016. The monthly index values of the realty sector for the 7 years are stored in a plain text (.txt) file. This plain text file contains 84 index values corresponding to the 84 months in the 7 year period under our study. The text file is then read into an R data object using the `scan()` function. The R data object is then converted into a time series object by applying the `ts()` function with a frequency value of 12. The frequency value is chosen to be 12 so that the seasonality characteristics of the time series for each month can be analyzed. The time series data object in R is then decomposed into its three components – trend, seasonal and random – using the `decompose()` function which is defined in the TTR library in the R environment. We plot the graphs of the realty time series data as well as its three components so that further analysis can be made on the behavior of the time series and its three components.

After carrying out a comprehensive analysis of the decomposition results of the time series of the realty sector, we propose six different approaches of forecasting of time series index values. In order to compute the forecast accuracy of each method, we build the forecast models using the realty time series data for the period January 2010 till December 2015, and apply the six models to forecast time series index values for each month of the year 2016. Since the actual values of the time series for all months of 2016 are already available with us, we compute the error in forecasting using each method of forecast that we propose. Comparative analysis of the methods of forecasting is done based on several useful metrics and why a particular method performs better than the other methods for the realty sector time series are critically analyzed in detail.

The work in this paper follows from several previous work. (Sen & Datta Chaudhuri, 2016a; Sen & Datta Chaudhuri 2016b) demonstrated how effectively time series decomposition approach can be utilized in robust analysis and forecasting of the Indian Auto sector. In another different work, (Sen & Datta Chaudhuri, 2016c) analyzed the behavior of two different sectors of Indian economy – the small cap sector and the capital goods sector – the former having a dominant random component while the latter exhibiting a significant seasonal component. Following another approach of time series analysis, (Sen & Datta Chaudhuri, 2016d) studied the behavior of the Indian information technology (IT) sector time series and the Indian capital goods sector time series. In yet another work, using the time series decomposition-based approach, (Sen & Datta Chaudhuri, 2016e) illustrated how time series analysis enables us to check the consistency between the fund style and actual fund composition of a mutual fund. In two different work, (Sen & Datta Chaudhuri, 2017a; Sen & Datta Chaudhuri 2017b) presented detailed analysis of the
behavior of the healthcare sector and the fast moving consumer goods (FMCG) sector of India using time series decomposition approaches.

In this work, we demonstrate how time series decomposition-based approach enables one in analyzing and understanding the behavior and different properties of the realty time series of the Indian economy based on time series data for the period January 2010 till December 2016. We also investigate what forecasting approach is most effective for the realty time series. For this purpose, we compare several approaches of forecasting and identify the one that produces the minimum value of forecasting error. We critically analyze all the proposed forecasting approaches, and explain why a particular approach has worked most effectively while some others have not done so for the realty time series data.

3. Time Series Decomposition Results

We present the decomposition results for the time series of the realty sector index values as per the records of the BSE for the period January 2010 till December 2016. First, we create a plain text (.txt) file containing the monthly index values of the realty sector for the period January 2010 till December 2016. This file contains 84 records corresponding to the 84 months in the 7 years under our study. We use the `scan()` function in R language to read the text file and store it in an R data object. Then, we convert this R data object into a time series object using the R function `ts()`. We used the value of the frequency parameter in the `ts()` function as 12 so that the decomposition of the time series is carried out on monthly basis. After creating the time series data object, we used the `plot()` function in R to draw the graph of the realty sector time series for the period January 2010 till December 2016. Figure 1 depicts the pictorial representation of the realty sector time series.

![Figure 1: Time series of realty Sector Index in India (Period: Jan 2010 – Dec 2016)](image)

To obtain further insights into the characteristics of the time series, we decomposed the time series object into its three components – trend, seasonal and random. The decomposition of the time series object is done using the `decompose()` function defined in the TTR library in R programming environment. The
The `decompose()` function is executed with the realty time series object as its parameter and the three components of the time series are obtained. Figure 2 presents the graphs of realty sector time series and its three components. Figure 2 consists of four boxes arranged in a stack. The boxes display the overall time series, the trend, the seasonal and the random component respectively arranged from top to bottom in that order.

From Figure 1, it may be seen that the time series of the realty sector has consistently fallen during the period January 2010 till December 2016 with occasional minor upward swings. The index value for the realty sector in the month of January 2010 was 3500, while in the month of December 2016, the index was found to be 1264. Except for two short periods - August 2012 to January 2013 and January 2014 to June 2014 - in which the realty sector index exhibited a modest increase in its value, in all other months during the period of our study, the index showed a downward movement. The three components of the time series are shown separately so that their relative behavior can be visualized.

![Figure 2: Decomposition results of the realty sector index time series into its three components (Period: Jan 2010 – Dec 2016)](image)

Table 1 presents the numerical values of the time series data and its three components. The trend and the random components are not available for the period January 2010 – June 2010 and also for the period July 2016 – December 2016. This is due to the fact that trend computation needs long term data. Coughlan (2015) illustrated that the `decompose()` function in R uses a 12-month moving average method to compute the trend values in a time series. Hence, in order to compute trend values for the period January 2010 – June 2010, the `decompose()` function needs time series data for the period July 2009 – December 2009. Since, the data for the period July 2009 – December 2009 are not available in the dataset under our study, the trend values for the period January 2010 – June 2010 could not be computed. For similar reason, the trend values for the period July 2016 - December 2016, could not be computed due to non-availability of the time series record for the period January 2017 – June 2017 in our dataset. It may be
noted from Table 1 that the seasonal value for a given month remains constant throughout the entire period of study. For example, the seasonal component has a constant value of -36 for the month of January in every year from 2010 till 2016. It is interesting to note that due to the non-availability of the trend values for the periods January 2010 – June 2010 and July 2016 – December 2016, the random components for these periods could not also be computed by the \textit{decompose()} function. In other words, since the aggregate time series values are given by the sum of the corresponding trend, seasonal and random component values, and because of the fact that the seasonal value for a given month remains the same throughout, non-availability of trend values for a period makes the random components values also unavailable for the same period.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
\textbf{Year} & \textbf{Month} & \textbf{Aggregate Index} & \textbf{Trend} & \textbf{Seasonal} & \textbf{Random} \\
\hline
\textbf{2010} & January & 3500 & -36 &  \\
& February & 3237 & -75 &  \\
& March & 3274 & -9 &  \\
& April & 3491 & -5 &  \\
& May & 3098 & 54 &  \\
& June & 3197 & 64 &  \\
& July & 3373 & 3251 & 8 & 114 \\
& August & 3332 & 3146 & -112 & 298 \\
& September & 3727 & 3054 & 41 & 632 \\
& October & 3635 & 2961 & 89 & 586 \\
& November & 2925 & 2868 & -5 & 62 \\
& December & 2856 & 2780 & -15 & 91 \\
\hline
\textbf{2011} & January & 2229 & 2676 & -36 & -411 \\
& February & 1982 & 2554 & -75 & -497 \\
& March & 2337 & 2406 & -9 & -60 \\
& April & 2180 & 2252 & -5 & -68 \\
& May & 2178 & 2124 & 54 & 0 \\
& June & 2020 & 2006 & 64 & 50 \\
& July & 2041 & 1923 & 8 & 110 \\
& August & 1740 & 1900 & -112 & -49 \\
& September & 1763 & 1876 & 41 & -154 \\
& October & 1920 & 1832 & 89 & -1 \\
& November & 1571 & 1787 & -5 & -211 \\
& December & 1376 & 1747 & -15 & -356 \\
\hline
\textbf{2012} & January & 1708 & 1716 & -36 & 28 \\
& February & 1956 & 1690 & -75 & 341 \\
& March & 1777 & 1683 & -9 & 102 \\
& April & 1693 & 1681 & -5 & 17 \\
& May & 1579 & 1692 & 54 & -167 \\
& June & 1668 & 1741 & 64 & -137 \\
& July & 1638 & 1794 & 8 & -164 \\
& August & 1511 & 1818 & -112 & -195 \\
& September & 1847 & 1820 & 41 & -15 \\
& October & 1772 & 1829 & 89 & -146 \\
& November & 1998 & 1842 & -5 & 160 \\
& December & 2111 & 1840 & -15 & 286 \\
\hline
\end{tabular}
\caption{Time series index values of the Indian realty Sector and its components (Period: Jan 2010 – Dec 2016)}
\end{table}
From Table 1, some important observations are made. First, we can see that seasonal component has the maximum value of 89 in the month of October, while the lowest value of seasonality -112 is observed in the month of August. The seasonal component is found to have positive values during the months of May, June, September and October, while negative values of seasonality are observed during the months of
January, February, August and December. Other months have exhibited negligible seasonality. In order to analyze the impact of seasonality on the realty time series, we computed some statistics of the seasonal component values. We computed the percentage contribution of the seasonal component on the aggregate time series values and found the following. The maximum, the minimum, the mean of the absolute values, and the standard deviation of the percentage of the seasonal components with respect to the aggregate time series values were found to be 6.63, -9.54, 2.53 and 3.44 respectively. The maximum percentage of seasonal component was found in the month of October 2012, while the minimum was observed in the month of August 2012. While the very low value of the mean percentage indicated that the time series was not seasonal, the high value of standard deviation in comparison to the mean value implied that the seasonal percentages exhibited high level of dispersion among themselves across the mean.

Second, the trend component of the time series has exhibited a consistent fall with occasional modest increase. During the period July 2010 till May 2012, there was a consistently decreasing trend. However, the trend started showing a modest increase in its value during June 2012 till December 2012. The trend started falling in its value again in January 2013 and continued to fall till November 2013. The period December 2013 till November 2014 witnessed a modest upward trend, before it started decreasing again in December 2014. The fall in trend continued till December 2015. However, the trend started increasing again in January 2016, which continued till June 2016. The maximum, the minimum, the mean, and the standard deviation of the percentage of the trend component with respect to the aggregate time series were found to be 128.92, 76.99, 101.85 and 11.78 respectively indicating that the trend was the single most dominant component in the time series. The maximum percentage of trend component was found in the month of February 2016, while the minimum was found in the month of June 2014.

Third, the maximum and the minimum values of the random component of the time series were found to be 632 and -497 respectively. These values are quite significant in comparison to the aggregate time series values. In order to understand the contribution of the random component on the overall time series, we computed the maximum, the minimum, the mean of the absolute values and the standard deviation of the percentage of random component values with respect to the aggregate time series values. These values were found to be 20.32, -25.87, 8.98 and 11.21 respectively. Since the maximum and the minimum values exhibited quite a significant randomness in the time series, the random component is found to be a dominant component. The large value of the standard deviation in comparison to the mean value also supported the fact that the randomness in the time series was quite significant. The random component contributed its maximum percentage to the aggregate time series in the month of January 2013, while the lowest percentage was found in the month of December 2011.

The overall conclusion is that the realty time series is primarily dominated by its trend component, while the random component is also significant in its contribution the aggregate time series. While the seasonal component is not a dominant component, the seasonal component values exhibited significant variations across their mean.

4. Proposed Forecasting Methods

This Section presents a set of six forecasting approaches that we propose for predicting the time series index of the realty sector. We discuss the details of the six different methods of forecasting and present the performance of these approaches in predicting the realty sector time series index. For the purpose of
comparative analysis of different approaches of forecasting, we use five different metrics and identify the method that yields the lowest value of forecasting error. We also critically analyze the approaches and argue why one method performs better than the others on the given dataset of realty sector time series index for the period January 2010 – December 2016. In the following, we first describe the six approaches, and then provide the detailed results as these forecasting methods are applied on the realty sector dataset.

Method 1: In this method, we use the realty sector time series data for the period January 2010 till December 2015 for the purpose of forecasting the monthly index values for each month of the year 2016. The `HoltWinters()` function in R library `forecast` has been used for this purpose. In order to build a robust forecasting framework, the `HoltWinters` model is used with a changing trend and an additive seasonal component that best fits the realty time series index data. The forecast horizon in the `HoltWinters` model is chosen to be 12 so that the forecasted values for all months of 2016 can be obtained by using the method at the end of the year 2015. Forecast error is computed for each month of 2016 and an overall RMSE value is also derived for this method.

Method II: In this approach, the realty sector index value for each month of the year 2016 is forecasted using the `HoltWinters()` method with a forecast horizon of 1 month. For example, for the purpose of forecasting the index for the month of March 2016, the index values of the realty sector from January 2010 till February 2016 are used to develop the forecasting model. As in Method I, the `HoltWinters` model is used with a changing trend and an additive seasonal component. Since the forecast horizon is short, the model is likely to produce higher accuracy in forecasting compared to the approach followed in Method I that used a forecast horizon of 12 months. The forecast error corresponding to each month of 2016 and an overall RMSE value for the model is computed.

Method III: In this forecasting approach, we first use the time series data for the realty sector index for the period January 2010 till December 2015 and derive the trend and the seasonal component values. This method yields the values of the trend component of the time series for the period from July 2010 till June 2015. Using the series of trend values for the period July 2010 till June 2015, forecast for the trend values for the period July 2015 till June 2016 are made using the `HoltWinters()` function in R with a changing trend component level but without any seasonality component. In other words, for forecasting the trend values, in the `HoltWinters()` function in R, we set the parameter ‘beta’ = TRUE and the parameter ‘gamma’ = FALSE, in the `HoltWinters()` function in R. The forecasted trend values are added to the seasonal component values of the corresponding months (based on the time series data for the period January 2010 till December 2015) to arrive at the forecasted aggregate of the trend and seasonal components. Now, we consider the time series of the realty sector index values for the entire period, i.e., from January 2010 till December 2016, and decompose it into its trend, seasonal and random components. Based on this time series, we compute the aggregate of the actual trend and the actual seasonal component values for the period July 2015 till June 2016. We derive the forecasting accuracy of this method by calculating the percentage of deviation of the aggregate of the actual trend and the actual seasonal component values with respect to the corresponding aggregate values of the forecasted trend and past seasonal components for each month during July 2015 to June 2016. An overall RMSE value for this method is also computed.

Method IV: The approach followed in this method is exactly similar to that used in Method III. However, unlike Method III that used `HoltWinters()` function with changing trend component and a zero seasonal
component, this method uses a linear regression model for the purpose of forecasting the trend component values for the period from July 2015 till June 2016. The `lm()` function in R is used for building a bivariate linear regression model with trend component as the response variable and time as the predictor variable. The regression model is built using the trend values for the period July 2010 till June 2015. The aggregate of the predicted trend values and the past seasonal values are compared with the aggregate of the actual trend values and the actual seasonal values for the period from July 2015 till June 106. The error in forecasting and an overall RMSE value is computed as in Method III.

**Method V:** We use *Auto Regressive Integrated Moving Average* (ARIMA) based approach of forecasting in this method. For building the ARIMA model, we use the realty sector time series data for the period January 2010 till December 2015. Using this training data set and executing the `auto.arima()` function defined in the forecast package in R, we compute values of the three parameters of the *Auto Regressive Moving Average* (ARMA) model, i.e. the *auto regression* parameter (*p*), the *difference* parameter (*d*), and the *moving average* parameter (*q*). Next, we build the ARIMA model of forecasting using the `arima()` function in R with the two parameters: (i) the realty time series R object (based on data for the period from January 2010 till December 2015), (ii) the order of the ARMA i.e., the values of the three parameters (*p*, *d*, *q*). Using the resultant ARIMA model, we call the function `forecast.Arima()` with parameters: (i) ARIMA model object, and (ii) forecast horizon = 12 months (in this approach). Since a forecast horizon of 12 months is used, we compute the forecasted values of each month of the year 2016 at the end of the year 2015. The error in forecasting and the RMSE value are also computed.

**Method VI:** As in Method V, we use the ARIMA model of forecasting in this method. However, unlike Method V that used forecast horizon of 12 months, this method uses a short forecast horizon of 1 month. For the purpose of forecasting, the ARIMA model is built using time series data for the period January 2010 till the month previous to the month for which forecasting is being made. For example, for the purpose of prediction of the time series index for the month of May 2016, the time series data from January 2010 till April 2016 is used for building the ARIMA model. Since the training data set for building the ARIMA model constantly changes in this approach, we evaluate the ARIMA parameters (i.e., *p*, *d*, and *q*) before every round of forecasting. In other words, for each month of the year 2016, before we make the forecast for the next month, we compute the values of the three parameters of the ARIMA model.

**Table 2: Computation Results using Method I**

| Month | Actual Index (A) | Forecasted Index (B) | % Error (C-B)/B *100 | RMSE |
|-------|------------------|----------------------|----------------------|------|
| Jan   | 1209             | 953                  | 21.17                |      |
| Feb   | 1051             | 610                  | 41.96                |      |
| Mar   | 1228             | 604                  | 50.81                |      |
| Apr   | 1357             | 661                  | 51.29                |      |
| May   | 1421             | 789                  | 44.48                |      |
| Jun   | 1533             | 767                  | 49.97                |      |
| Jul   | 1607             | 730                  | 54.57                |      |
| Aug   | 1542             | 683                  | 55.71                |      |
| Sep   | 1512             | 957                  | 36.71                |      |
| Oct   | 1556             | 1141                 | 26.67                |      |
| Nov   | 1282             | 1005                 | 21.61                |      |
| Dec   | 1264             | 765                  | 39.48                | 608  |
5. Forecasting Results

In this Section, we provide results on the performance of the six forecasting approaches.

Method I: The results obtained using these methods are presented in Table 2. In Figure 3, we have plotted the actual values of the realty sector index and the corresponding predicted values for all the twelve months of the year 2016.

Observations on Method I: We observe from Table 2 that the forecasted values are very much deviant from the actual values of the realty sector index for all months in the year 2016. It is evidently clear that the Method I is not at all accurate in forecasting the realty sector index values. The highest value of the error percentage (55.71) has been too high, which corresponds to the month of August 2016. The forecast for the month of January 2016 exhibited the lowest percentage of error (21.17), which is also unacceptably high. The reason for the large error of forecasting in this method is due to the fact that the method consistently underestimated the index. While the actual index increased in its value for the period February 2016 till July 2016, the forecast consistently exhibited downward trend following the past behavior of the time series till December 2015. Hence, the two reasons for failure of this method can be identified as: (i) the long forecast horizon of 12 months used in the HoltWinters method could not really understand the actual behavior of the time series in the year 2016, (ii) the actual time series of the realty index showed a reversal in its trend and increased in its value during February 2016 till July 2016, which the forecasting method could not predict in December 2015 when the time series was exhibiting a consistent fall in its value. The RMSE value for this method is found out to be 608 that is 44.06 percent of the mean value of 1380 of the actual index of the realty sector index during 2016. The unacceptably high percentage of RMSE value with respect to the mean index value clearly indicates that Method I is largely ineffective in forecasting the realty sector time series index.

![Figure 3. Actual and predicted values of realty sector index using Method I of forecasting](Period: Jan 2016 – Dec 2016)

Method II: The results of forecasting using Method II are presented in Table 3. In Figure 4, the actual index values and their corresponding predicted values are plotted.
Table 3: Computation Results using Method II

| Month | Actual Index | Forecasted Index | % Error | RMSE |
|-------|--------------|------------------|---------|------|
| Jan   | 1209         | 953              | 21.17   |      |
| Feb   | 1051         | 819              | 22.07   |      |
| Mar   | 1228         | 1015             | 17.35   |      |
| Apr   | 1357         | 1291             | 4.86    |      |
| May   | 1421         | 1495             | 5.21    |      |
| Jun   | 1533         | 1426             | 6.98    |      |
| Jul   | 1607         | 1508             | 6.16    |      |
| Aug   | 1542         | 1581             | 2.53    |      |
| Sep   | 1512         | 1868             | 23.54   |      |
| Oct   | 1556         | 1771             | 13.82   |      |
| Nov   | 1282         | 1425             | 11.15   |      |
| Dec   | 1264         | 1052             | 16.77   |      |

Observations on Method II: We observe from Table 3 and Figure 4 that the forecasted values match reasonably well with the actual values of the time series index. The lowest value of error percentage is found to be 2.53 which occurred in the month of August 2016, while the highest error percentage value of 23.54 was found in the month of September 2016. The RMSE value for this method is 191 while the mean value of the realty sector index in 2016 is 1380. Hence, the RMSE value contributes 13.84 percent of the mean value of the actual index. This clearly demonstrates that HoltWinters additive model with a prediction horizon of 1 month is quite effective in accurately forecasting the realty sector index values.

Figure 4. Actual and predicted values of realty sector index using Method II of forecasting (Period: Jan 2016 – Dec 2016)
Method III: The results of forecasting using this method are presented in Table 4. Figure 5 depicts the actual index values and their corresponding predicted values for all months of the year 2016 using Method III of forecasting. The actual trend and seasonal component values for the period from July 2015 till June 2016 (computed based on the time series data for the period from January 2010 till December 2016) and their aggregated monthly values are noted in Columns B, C and D respectively in Table 4. The forecasted trend values (using *HoltWinters* method with changing trend component and nil seasonal component and with a forecast horizon of 12 months) and the past seasonal component values (based on time series data for the period from January 2010 till December 2015) and their corresponding aggregate values are noted in columns E, F and G respectively in Table 4. The error value for each month and an overall RMSE value for this method are also computed.

**Table 4: Computation Results using Method III**

| Month | Actual Trend | Actual Seasonal | Actual (Trend + Seasonal) | Forecasted Trend | Past Seasonal | Forecasted (Trend + Seasonal) | % Error | RMSE |
|-------|--------------|-----------------|---------------------------|------------------|--------------|-------------------------------|---------|------|
| A     | 1469         | 8               | 1477                      | 1480             | 17           | 1497                          | 1.35    | 85   |
| Aug   | 1412         | -112            | 1300                      | 1457             | -113         | 1344                          | 3.38    |      |
| Sep   | 1361         | 41              | 1402                      | 1434             | 33           | 1467                          | 4.64    |      |
| Oct   | 1334         | 89              | 1423                      | 1411             | 90           | 1501                          | 5.48    |      |
| Nov   | 1320         | -5              | 1315                      | 1388             | -19          | 1369                          | 4.11    |      |
| Dec   | 1320         | -15             | 1305                      | 1365             | -32          | 1333                          | 2.15    |      |
| Jan   | 1334         | -36             | 1298                      | 1342             | -27          | 1315                          | 1.31    |      |
| Feb   | 1355         | -75             | 1280                      | 1319             | -38          | 1281                          | 0.08    |      |
| Mar   | 1372         | -9              | 1363                      | 1296             | 9            | 1305                          | 4.26    |      |
| Apr   | 1384         | -5              | 1379                      | 1273             | -9           | 1264                          | 8.34    |      |
| May   | 1389         | 54              | 1443                      | 1250             | 49           | 1299                          | 9.98    |      |
| Jun   | 1383         | 64              | 1447                      | 1227             | 38           | 1265                          | 12.58   |      |

Observations on Method III: It may be observed from Table 4 and Figure 5 that the errors in forecasting have been quite small for all months during the period July 2015 till June 2016. The lowest value of the
error percentage had been 0.08 that occurred in the month of February 2016, while the highest value of error percentage of 12.58 was observed in the month of June 2016. The computed RMSE value is 85, while the mean value of the sum of the actual trend and the actual seasonal components during the period July 2015 till June 2016 components is 1369. Hence, the RMSE represents only 6.21 percent of the mean value of the sum of the actual trend and the actual seasonal components. The low value of RMSE indicates that this method is very effective in forecasting the reality index time series index values. It is evident from Table 4 that the actual trend values of the reality sector exhibited a fall in its value, albeit at a very slow rate, during the period July 2015 till December 2016, while the trend values increased at a very sluggish rate during January 2016 till July 2016. As the magnitudes of the actual values of the seasonal component in the time series are quite small compared to the trend values, the sum of the actual trend and the actual seasonal values also exhibited almost the same pattern of change as the trend - decreasing till December 2015 and then exhibiting an upward movement till June 2016. In case of three months, there were some exceptions observed. These months were, September and October in the year 2015, and the month of January in the year 2016. For the months of September and October 2015, the sum of the trend and the seasonal components increased with respect to its values in the corresponding previous months because of the presence of moderately large positive values of the seasonal components during those months and a large negative value of the seasonal component during the month of August 2015. For the month of January 2016, the sum of the actual trend and actual seasonal component value decreased with respect to the corresponding value in the month of December 2015 due to a reasonably large negative value of the seasonal component in month of January 2016. As the trend values are forecasted using HoltWinters method with a forecast horizon of 12 months, the forecasted trend values are smoothened out over the period, resulting in a series of forecasted trend values exhibiting a very slow rate of decrease. In fact, the forecasted trend started with a value of 1480 in the month of July 2015, and attained a value of 1227 in the month of June 2016. Since the past seasonal values (i.e., the seasonal values based on the reality sector time series for the period January 2010 till December 2015), were also very low compared to the corresponding trend values, the sum of the forecasted trend values and their respective past seasonal values exhibited similar as the forecasted trend values. The similar behavior of both the actual series and the forecasted series made the two series match closely with each other, thereby making this method of forecast extremely accurate for the reality sector index time series.

| Month | Actual Trend | Actual Seasonal | Actual (Trend + Seasonal) | Forecasted Trend | Past Seasonal | Forecasted (Trend + Seasonal) | % Error | RMSE |
|-------|--------------|----------------|--------------------------|-----------------|--------------|-----------------------------|---------|------|
| A     | B            | C              | D                        | E               | F            | G                           | (G-D)/D |      |
| Jul   | 1469         | 8              | 1477                     | 1243            | 17           | 1260                        | 14.69   | 250  |
| Aug   | 1412         | -112           | 1300                     | 1223            | -113         | 1110                        | 14.62   |      |
| Sep   | 1361         | 41             | 1402                     | 1203            | 33           | 1236                        | 11.84   |      |
| Oct   | 1334         | 89             | 1423                     | 1182            | 90           | 1272                        | 10.61   |      |
| Nov   | 1320         | -5             | 1315                     | 1162            | -19          | 1143                        | 13.08   |      |
| Dec   | 1320         | -15            | 1305                     | 1142            | -32          | 1110                        | 14.94   |      |
| Jan   | 1334         | -36            | 1298                     | 1121            | -27          | 1094                        | 15.72   |      |
| Feb   | 1355         | -75            | 1280                     | 1101            | -38          | 1063                        | 16.95   |      |
| Mar   | 1372         | -9             | 1363                     | 1081            | 9            | 1090                        | 20.03   |      |
| Apr   | 1384         | -5             | 1379                     | 1060            | -9           | 1051                        | 23.79   |      |
| May   | 1389         | 54             | 1443                     | 1040            | 49           | 1089                        | 24.53   |      |
| Jun   | 1383         | 64             | 1447                     | 1019            | 38           | 1057                        | 16.95   |      |

**Method IV:** Table 5 presents the results of forecasting for Method IV. Figure 6 shows the actual index values and their corresponding predicted values using Method IV of forecasting for all months during the
period July 2015 till June 2016. The values of the actual trend component and the actual seasonal component for the period July 2015 till June 2016 (computed based on the time series data for the period January 2010 till December 2016) and their aggregated monthly values are listed in Columns B, C and D respectively in Table 4. In Method III, the trend values are forecasted using $HoltWinters()$ function with a forecast horizon of 12 months. However, in Method IV, the forecasting of the trend values are done using a linear regression approach with the realty index values as the response variable and the month as the predictor variable. The forecasted trend values and the past seasonal component values (based on time series data for the period from January 2010 till December 2015) and their corresponding aggregate values listed noted in columns E, F and G respectively in Table 5. The percentage of error in forecasting for each month during the period July 2015 till June 2016, and an overall RMSE value are also listed.

![Figure 6. Actual and predicted values of the sum of trend and seasonal components of realty sector index using Method IV of forecasting (Period: Jan 2016 – Dec 2016)](image)

**Observation on Method IV:** It is evident from Table 5 and Figure 6 that Method IV has produced higher values of error as compared to those produced by Method III. For Method IV, the lowest percentage of error (10.61) was found in the month of October 2015, while the highest error percentage value (24.53) was observed in the month of May 2016. It may also be noted from Figure 5 that the error, in general, increased with time. The RMSE value for this method has been found to be 250 while the mean value of the actual sum of trend and seasonal components during the period July 2015 till June 2016 is 1369. Hence, the RMSE value is 18.26 percent of the mean value of the index, indicating that the method of forecasting is moderately accurate in its forecasting results. It is not difficult to understand the reason for why Method IV has produced higher values of error produced as compared to most of the other methods used by us. This method used linear regression method for predicting the trend values for the period July 2015 till June 2016. Based on the previous trend values for the period July 2010 till June 2015, the linear regression computed the regression coefficients and used the values of the coefficients to compute the trend values for the period July 2015 till June 2106. Since, the trend for the period July 2010 till June 2015 had a negative slope (as the trend consistently decreased in its value during this period), the
regression coefficient was also negative. The negative regression coefficient produced forecasted values of trend which consistently decreased in its value - starting from a value of 1243 in July 2015 to 1019 in June 2016. Since seasonal component values were small compared to trend values, the sum of the forecasted trend and past seasonal values also exhibited the same pattern as the forecasted trend values – the sum of the forecasted trend values and the past seasonal values decreased consistently from 1260 in July 2015 to 1057 in June 2016. Only for three months, the sum of the forecasted trend and the past seasonal component increased with respect to the previous months. These months are September and October in the year 2015, and May in the year 2016. For these three months, the presence of reasonably large positive values of the seasonal component more than compensated the decrease in the trend values for those months. While the sum of the predicted values of the trend and the past seasonal component consistently decreased, the sum of the actual trend and the actual seasonal component values increased in its value for most of the months. Moreover, the linear regression underestimated the trend values for majority of the months for the period under study. Since the forecasted trend values decreased while the actual trend increase for most part of the period, the method yielded increasingly higher values of error with time.

![First-order difference of the realty sector time series](image)

**Figure 7. First-order difference of the realty sector time series (Period: Jan 2016 – Dec 2015)**

**Method V:** This method of forecasting is based on ARIMA technique with a forecast horizon of 12 months. Applying `auto.arima()` function defined in the library `forecast` in R, on the realty sector time series index values for the period January 2010 till December 2015, we obtain the parameter values for the time series as: $p = 0$, $d = 1$, and $q = 0$. We cross verify the values of $p$, $d$, and $q$ by plotting the partial auto correlation function (PACF), the first-order difference of the time series, and the auto correlation function (ACF) respectively. The first order difference of the realty time series is presented in Figure 7. It is clear that the first-order difference time series is a stationary one, as the mean and the variance of the first-order difference time series are approximately constant. Hence the value of $d = 1$ is cross-verified. Figure 8 depicts the PACF of the realty sector time series for the period January 2010 till December 2015. It is clear that except for $lag = 0$, the partial correlation values at all other lags are insignificant. Hence the
value of $p = 0$ is also verified. Figure 9 shows that minimum integral value of lag beyond which all autocorrelation values are insignificant is 1. Therefore, $q = 1$ is also verified. Hence, we have verified that the realty sector time series for the period January 2010 till December 2015 is an ARMA $(0, 1, 1)$ model. Using the \textit{arima()} function in R with its two parameters: (i) the realty sector time series R object and (ii) the order $(0, 1, 0)$ of ARMA, we build the ARIMA model. Finally, we use the function \textit{forecast.Arima()} with two parameters: (i) the ARIMA model and (ii) the time horizon of forecast = 12 months, for forecasting the index values of the time series for all the twelve months of the year 2016. Table 6 presents the results of forecasting using this method, and Figure 10 depicts the actual index values with their corresponding predicted values for all months in the year 2016.

![Figure 8. Plot of the partial auto correlation function (PACF) of the realty sector time series with max lag of 2 years (Period: Jan 2016 – Dec 2015)](image)

![Figure 9. Plot of the auto correlation function (ACF) of the realty sector time series with max lag of 2 years (Period: Jan 2016 – Dec 2015)](image)
Table 6: Computation Results using Method V

| Month | Actual Index (A) | Forecasted Index (C) | % Error (C-B)/B *100 | RMSE |
|-------|------------------|----------------------|------------------------|------|
| Jan   | 1209             | 1344                 | 11.17                  |      |
| Feb   | 1051             | 1344                 | 27.88                  |      |
| Mar   | 1228             | 1344                 | 9.45                   |      |
| Apr   | 1357             | 1344                 | 0.96                   |      |
| May   | 1421             | 1344                 | 5.42                   |      |
| Jun   | 1533             | 1344                 | 12.33                  |      |
| Jul   | 1607             | 1344                 | 16.37                  |      |
| Aug   | 1542             | 1344                 | 12.84                  |      |
| Sep   | 1512             | 1344                 | 11.11                  |      |
| Oct   | 1556             | 1344                 | 13.62                  |      |
| Nov   | 1282             | 1344                 | 4.84                   |      |
| Dec   | 1264             | 1344                 | 6.33                   |      |

Observations on Method V: It is evident from Table 6 that Method V is quite effective in forecasting the realty sector time series. The lowest value of error percentage had been 0.96 that occurred in the month of April 2016, while the highest value of error percentage was 27.88, observed in the month of February 2016. The RMSE value for this method is found to be 171 which is 12.39 per cent of the mean value of the realty sector index during the period January 2016 till December 2016. The mean value of the realty sector index has been 1380. Considering the fact that this method uses a long forecast horizon of 12 months, the error percentage values are quite moderate. This is attributed to the fact that the realty sector index experienced a very small dispersions in its values during the year 2016. The index started with a value of 1209 in January 2016, experienced its lowest value of 1051 in February, then attaining its peak value of 1607 in the month of July and then again decreasing in its value finally reaching a value of 1264 in the month of December 2016. Method V with a forecast horizon of 12 forecasted a constant average value of 1344 for the series with ARIMA parameters (0, 1, 1), so that the average error for all the forecasted values is minimized. Figure 10 presents a graphical depiction of the actual index values and their corresponding predicted values using Method V.

Figure 10. Actual and predicted values of realty sector index using Method V of forecasting (Period: Jan 2016 – Dec 2016)
**Method VI:** In this approach, we build an ARIMA model with a forecast horizon of one month. The methodology used for building the ARIMA model, however, is exactly identical to that used in Method V. The difference in Method V and Method VI lies in different values of forecast horizon used in these methods. While Method V used a forecast horizon of 12 months, we use a forecast horizon of 1 month in Method VI. Since, in Method VI, forecast is made only one month in advance, the training data set used for building the ARIMA model constantly increases in size, and hence, we re-evaluate the parameters of the ARIMA model every time we use it in forecasting. In other words, for every month of 2016, before we make the forecast for the next month, we compute the values of the parameters of the ARIMA model. Computation of the values of ARIMA parameters $p$, $d$, and $q$ showed that for the period January 2016 till July 2016, the ARIMA model was (1, 1, 1), while the for the remaining period it was (0, 1, 0). Table 7 presents the forecasting results for Method VI. Figure 11 depicts the actual index values of the realty sector and their corresponding predicted values for this method of forecasting.

**Table 7: Computation Results using Method VI**

| Month | Actual Index | Forecasted Index | % Error   | RMSE |
|-------|--------------|------------------|-----------|------|
| Jan   | 1209         | 1344             | 11.17     |      |
| Feb   | 1051         | 1209             | 15.03     |      |
| Mar   | 1228         | 1051             | 14.41     |      |
| Apr   | 1357         | 1228             | 9.51      |      |
| May   | 1421         | 1357             | 4.50      |      |
| Jun   | 1533         | 1421             | 7.31      | 4.60 |
| Jul   | 1607         | 1533             | 4.60      |      |
| Aug   | 1542         | 1607             | 4.21      |      |
| Sep   | 1512         | 1542             | 1.98      |      |
| Oct   | 1556         | 1512             | 2.83      |      |
| Nov   | 1282         | 1556             | 21.37     |      |
| Dec   | 1264         | 1282             | 1.42      |      |

![Figure 11. Actual and predicted values of realty sector index using Method VI of forecasting](image)

(Period: Jan 2016 – Dec 2016)
Observations on Method VI: From Table 7, it is evident that the error percentage values for all months of the year 2016 are quite low. The lowest value of error percentage was found to be 1.42 in the month of December 2016, while the highest value of error was 21.37 per cent in the month of November 2016. The RMSE value for Method VI of forecasting is found to be 128. The mean value of the index of the realty sector for the period January 2016 till December 2016 is 1380. Hence, the RMSE value is 9.28 percent of the mean value of the actual index of the realty sector. This indicates that Method VI has been highly accurate in forecasting the realty sector index values. The high level of accuracy of Method VI may be attributed to its short forecast horizon of one month. The short forecast horizon is able to catch the changing pattern of the time series very effectively. This has resulted into a very small error in forecasting. It is clearly evident from Figure 11 that the forecasted time series values exactly followed the pattern of the time series of the actual index values of the realty sector.

Summary of Forecasting Results

In Table 8, we summarize the performance of the six forecasting methods that we have used. For the purpose of comparison between these methods, we have chosen six metrics: (i) minimum (Min) error rate, (ii) maximum (Max) error rate, (iii) mean error rate, (iv) standard deviation (SD) of error rates, and (v) root mean square error (RMSE), and (vi) the ratio of RMSE to the mean of the actual index values in percentage. For Method I, II, IV and V the mean of the index values are the same, being the mean of the actual index values of the realty sector for the twelve months in the year 2016. This mean value is found to be equal to 1380. For Method III and IV, however, the mean of the index values are the mean of the sum of the trend and seasonal values during the period July 2015 till June 2016. This mean value of the sum of the trend and seasonal values is found to be equal to 1369. Since the RMSE values for Methods I, II, IV and V and those for Method III and IV are computed against different set of actual values, hence, instead of the raw RMSE value, the percentage of RMSE to the mean value of the index serves as a better metric for comparing different methods of forecasting. Accordingly, we have ranked the forecasting methods.

Table 8: Comparison of the performance of the forecasting methods

| Methods | Min Error | Max Error | Mean Error | SD of Errors | RMSE | RMSE / Mean Index Value Percentage |
|---------|-----------|-----------|------------|--------------|------|----------------------------------|
| Method I | 21.17     | 55.71     | 41.20      | 12.40        | 608  | 44.06                            |
| Method II | 2.53      | 23.54     | 12.63      | 7.479        | 191  | 13.84                            |
| Method III | 0.08      | 12.58     | 4.81       | 3.77         | 85   | 6.21                             |
| Method IV | 10.61     | 24.53     | 16.48      | 4.35         | 250  | 18.26                            |
| Method V  | 0.96      | 27.88     | 11.03      | 6.87         | 171  | 12.39                            |
| Method VI | 1.42      | 21.37     | 8.20       | 6.22         | 128  | 9.28                             |

Table 8 presents the comparative analysis of the six forecasting methods. It can be seen that Method III that uses the sum of the forecasted trend values using HoltWinters( ) function of horizon 12 months and
the past seasonal values to predict the sum of the future trends values and the new seasonal values, has performed has produced the lowest percentage value of the ratio of RMSE to the mean index value. In fact, Method III has produced lowest values for all other metrics too. Hence, Method III clearly turns out to be the most accurate among all the six methods. On the other hand, the performance of Method I has been the worst since it has produced the highest values for all the five metrics of error percentages. Method II turns out to be a close second, while Method V and Method III follow it in order of their performance. The performance of Method V turns out to be worse than that of Method III, although it has performed much better than Method I.

6. Related Work

Several approaches and techniques are proposed by researchers in the literature for forecasting of daily stock prices. Among these approaches, neural network-based approaches are extremely popular. Mostafa (2010) proposed a neural network-based technique for predicting movement of stock prices in Kuwait. Kimoto et al. (1990) presented a technique using neural network based on historical accounting data and various macroeconomic parameters to forecast variations in stock returns. Leigh et al. (2005) demonstrated methods of predicting stock prices and stock market index movements in the New York Stock Exchange (NYSE) during the period 1981 – 1999 using linear regression and simple neural network models. Hammad et al. (2009) illustrated how the output of an artificial neural network (ANN) model can be made to converge and produce highly accurate forecasting of stock prices. Dutta et al. (2006) used ANN models for forecasting closing index values the BSE during the period January 2002 till December 2003. Ying et al. (2009) used Bayesian Network (BN) – based approach to forecast stock prices of 28 companies listed in DJIA (Dow Jones Industrial Average) during 1988-1998. Tsai and Wang (2009) illustrated how and why the forecasting accuracy of BN-based approaches usually is higher than that obtained using traditional regression and neural network-based methods. Tseng et al. (2012) demonstrated the application of traditional time series decomposition, HoltWinters models, Box-Jenkins method and artificial neural networks in forecasting prices of randomly selected 50 stocks during the period September 1998 till December 2010. The study found that forecasting errors values are low for Box-Jenkins method, HoltWinters model and normalized neural network model, while higher values of error were observed in time series decomposition method and non-normalized neural network model. Moshiri & Cameron (2010) built a back propagation network (BPN) with econometric models for forecasting the level of inflation using the techniques: (i) Box-Jenkins Autoregressive Integrated Moving Average (BJARIMA) model, (ii) Vector Autoregressive (VAR) model and (ii) Bayesian Vector Autoregressive (BVAR) model. Thenmozhi (2001) applied chaos theory for examining the pattern of changes of stock prices in Bombay Stock Exchange (BSE) during the period August 1980 till September 1997, and found that the daily and weekly returns of BSE index exhibited nonlinear trends, while the movement pattern of the time series of BSE index was found to be weakly chaotic. Hutchinson et al. (1994) proposed a novel approach using the principles of learning networks for estimating the price of a derivative.

Predictive models based on ANN are found to be extremely accurate in forecasting stock prices. Shen et al. (2007), Jaruszewicz & Mandziuk (2004), Ning et al. (2009), Pan et al. (2005), Hamid & Iqbal (2004), Chen et al. (2005), Chen et al. (2003), Haniyas et al. (2007) and de Faria et al. (2009) demonstrated the effectiveness of ANN-based models in their forecasting ability of stock price movements. Many applications of hybrid systems in stock market time series data analysis have also been proposed in the
In the literature, researchers have also proposed several forecasting techniques which have particularly focused on various issues in the realty sector. Karakozova (2005) carried out a detailed empirical evaluation of alternative econometric methods for modeling and forecasting rents and returns in property markets. The authors used the Finnish property market as a case study and observed that the choice of econometric model depends on whether we use the model to test the theory, analyze the policy or for making forecasts. The study also revealed that theory-based econometric methods are more effective in evaluating suitability of theoretical frameworks for modeling rents and returns. On the other hand, in forecasting of rents and returns, time series techniques are found to produce better results. an de Meulen et al. (2011) observed that house price fluctuations in Germany are significantly affected by the financial stability and economic development of the country. The authors measured price movements in different real estate markets in Germany and forecasted the short-term price fluctuations in the real estate sector. Using the auto regressive AR(p) model as a base, it was found that vector auto regression (VAR) and auto regressive distributed lag (ARDL) models with additional macroeconomic information improves the forecasting accuracy. Grinis (2015) observed that global outlook in realty sector for the period 2016 - 2018 would be positive, if the macro-economic fundamentals are sound. The author observed that: (i) the asset prices will be positively affected by stable inflationary policy over the long term, (ii) unemployment figures would continue exhibit downward trend in most major markets in the globe which would lead to a boom in the real-estate sector, (iii) commodity prices that affect consumers and the real estate market would not reflect any significant rise in the near future. Dietzel et al. (2014) examined the role of Internet search data in the commercial real estate sector. The authors constructed various forecast models and found that the inclusion of Google search data significantly improves the forecast results of commercial real estate prices for the US market. In other words, the investigation revealed that Google search data are extremely effective in measuring sentiment in the commercial real estate markets. Glaeser et al. (2017) studied the recent boom in the real estate sector in China and analyzed the factors affecting the demand and supply of housing in that country. On the demand side, the authors examined the economic, demographic, cultural, and speculative factors that influences the demand in the real-estate sector. In the supply side, housing prices to the physical costs of construction and assessment of the long-run price of land were considered. The study revealed that a housing crash is not inevitable in China in the near future provided suitable policies are adopted by the Chinese government. Booth & Marcato (2003) presented formulated various methods of constructing real estate indices using the available data of the real estate sector. The authors also demonstrated how the use of publicly available real-estate data can sometimes lead to imperfect models and how the real estate data can be effectively exploited in designing stochastic investment modeling for actuarial purposes. Gaspareniene et al. (2014) proposed models for identifying and analyzing the factors that influences housing price level formation in the economies of the developing countries. The study revealed that the structure of the model of housing price level formation should be an integral multi-stage aggregate of microeconomic, macroeconomic and other elements that effectively describe the price fluctuations in the real estate sector. Gelain & Lansing (2014) analyzed the behavior of the equilibrium price-rent ratio for housing in a standard asset pricing model and compared the model predictions to survey evidence on the returns expectations of real-world housing investors.

In contrast to the work mentioned above, our work in this paper deals with a structural decomposition of the time series of the realty sector index in India during the period January 2010 till December 2016.
Based on the decomposition results of the time series, we identified several important characteristics of the Indian realty sector. We particularly investigated the nature of the trend, seasonality pattern and degree of randomness exhibited by the time series. After analyzing the nature of the realty time series, we proposed six forecasting techniques for predicting the index values of the sector for each month of the year 2016. We computed the accuracies of each of the forecasting techniques, and critically analyzed under what situations a particular technique performs better than the other techniques. Since the forecasting methods proposed in this paper are all generic in nature, these methods can be very effectively applied in forecasting the future trends and behavior of time series index values of other sectors of economy of India or other countries in the world.

7. Conclusion

This paper has presented a time series decomposition-based approach for analyzing the behavior of the time series of the realty sector of the Indian economy during the period January 2010 till December 2016. Algorithms and library-defined functions in the R programming language have been used to decompose the time series index values into three components- trend, seasonal, and random. The decomposition results of the time series provided with several important insights into the behavior exhibited by the realty sector time series during the period under our study. Based on the decomposition results, the degree of seasonality and randomness in the time series have been computed. Particularly, it has been possible to identify the months during which the seasonal component in the realty time series plays a major role. The seasonal component is found to have positive values during the months of May, June, September and October, while negative values of seasonality are observed during the months of January, February, August and December. The mean percentage of the contribution of the seasonal component on the aggregate time series value is found to be low, the high value of standard deviation in comparison to the mean value implied that the seasonal percentages exhibited high level of dispersion among themselves across the mean. The maximum and the minimum values of the percentage of random component with respect to the aggregate time series index exhibited quite a significant randomness in the time series and the random component was found to be a dominant component in the time series. The large value of the standard deviation in comparison to the mean value also supported the fact that the randomness in the time series was quite significant. The trend was found to be the most dominant component in the time series. However, the trend component of the time series exhibited a consistent fall with occasional modest increase. During the period July 2010 till May 2012, the trend experienced a consistent fall in its value, which started rising modestly during June 2012 till December 2012. The period December 2013 till November 2014 witnessed a modest upward trend, before it started decreasing again in December 2014. The fall in trend continued till December 2015. However, the trend started increasing again in January 2016, which continued till June 2016.

After a careful analysis of the decomposition results of the realty sector index time series, we proposed six methods for forecasting the time series index values. The six method of forecasting involved different algorithms and different lengths of forecast horizon. It was observed that Method III that used the sum of the forecasted trend values using `HoltWinters()` function of horizon 12 months and the past seasonal values to predict the sum of the future trends values and the new seasonal values had performed best yielding the lowest percentage value of the ratio of RMSE to the mean index value. However, Method I that predicted the trend values using `HoltWinters()` forecasting approach with a forecast horizon of 12 months was found to produce the highest value of the ratio of the RMSE to the mean index value, thereby
exhibiting the worst performance among the six methods of forecasting. The performance of Method VI that was based on an ARIMA model with a forecast horizon of 1 month, also performed very efficiently with RMSE to mean index ratio value of 9.28 percent. The other three methods, i.e., Method II (i.e., HoltWinters model with a forecast horizon of 1 month), Method IV (i.e., the model that used the sum of the forecasted trend values using a linear regression and the past seasonal values to predict the sum of the actual trend and seasonal values) and Method V (i.e., ARIMA model with a forecast horizon of 12 months) performed moderately well with none of their RMSE to mean index values exceeding the threshold of value of 20 percent.

While the results in this work provide enough valuable insights into the characteristics of the realty index time series in India, and they also serve as guidelines for choosing an appropriate forecasting framework for predicting the future index values of the time series, these results can be extremely useful for constructing an optimized portfolio of stocks. Performing similar exercise on different sectors will enable analysts to understand the individual characteristics of the trend, seasonality and randomness of those sectors. This information can be suitably leveraged by portfolio managers in identifying the timing of buy and sell of stocks from different sectors thereby designing an efficient and optimized portfolio.

References

van de Meulen, P., Micheli, M., & Schmidt, T. (2011). Forecasting House Prices in Germany. RUHR Economic Papers, Document # 294. Available online at: Link (Accessed on June 22, 2017)

Booth, P. M & Marcato, G. (2004). The measurement and modelling of commercial real estate performance. British Actuarial Journal, 10(1), 5-61. Available online at: Link (Accessed on June 22, 2017)

Chen, A.-S., Leung, M. T. & Daouk, H. (2003). Application of neural networks to an emerging financial market: forecasting and trading the Taiwan stock index. Operations Research in Emerging Economies, 30(6), 901–923. DOI: 10.1016/S0305-0548(02)00037-0

Chen, Y., Dong, X. & Zhao, Y. (2005). Stock index modeling using EDA based local linear wavelet neural network. Proceedings of International Conference on Neural Networks and Brain, Beijing, China, pp. 1646–1650. DOI: 10.1109/ICNNB.2005.1614946

Coghlan, A. (2015). A Little Book of R for Time Series, Release 02. Available at: Link (Accessed on: May 10, 2017)

de Faria, E. L., Albuquerque, M. P., Gonzalez, J. L., Cavalcante, J. T.P., & Albuquerque, M. P. (2009). Predicting the Brazilian stock market through neural networks and adaptive exponential smoothing methods. Expert Systems with Applications, 36(10), 12506-12509. DOI: 10.1016/j.eswa.2009.04.032

Dietzel, M.A., Braun, N. & Schaefers, W. (2014). Sentiment-based commercial real estate forecasting with Google search volume data. Journal of Property Investment & Finance, 32(6), 540 - 569. DOI: https://doi.org/10.1080/09765270600500305

Dutta, G. Jha, P. , Laha, A. & Mohan, N. (2006). Artificial neural network models for forecasting stock price index in the Bombay Stock Exchange. Journal of Emerging Market Finance, 5(3), 283-295. DOI: 10.1177/097265270600500305

Gaspareniene, L., Venclauskiene, D. & Remeikiene, R. (2014). Critical review of selected housing market models concerning the factors that make influence on housing price level formation in the countries with transition economy. Procedia - Social and Behavioral Sciences, Vol 110, January 2014, pp. 419-427. DOI: https://doi.org/10.1016/j.sbspro.2013.12.886

Gelain, P. & Lansing, K. J. (2014). House Prices, Expectations, and Time-Varying Fundamentals. Working Paper, Federal Reserve Bank of San Francisco, 2013 - 03. Available at: Link (Accessed on: June 20, 2017)
Glaeser, E., Huang, W., Ma, Y. & Shleifer, A. (2017). A real estate boom with Chinese characteristics. *Journal of Economic Perspectives, 31*(1), 93-116. DOI: http://doi.org/10.1257/jep.31.1.93

Grinis, M. (2015). *Global Market Outlook 2016 - Trend in Real Estate Private Equity*. EYGM 2015 Global Report. Available online at: Link

Hamid, S. A., Iqbal, Z. (2004). Using neural networks for forecasting volatility of S&P 500 index futures prices. *Journal of Business Research*, 57(10), 1116–1125. DOI: 10.1016/S0148-2963(03)00043-2

Hammad, A. A. A., Ali, S. M. A. & Hall, E. L. (2007). Forecasting the Jordanian stock price using artificial neural network. *Intelligent Engineering Systems through Artificial Neural Networks*, Vol 17, Digital Collection of The American Society of Mechanical Engineers. DOI: 10.1115/1.802655.paper42

Hanias, M., Curtis, P. & Thalassinos, J. (2007). Prediction with neural networks: the Athens stock exchange price indicator. *European Journal of Economics, Finance and Administrative Sciences*, 9, 21–27.

Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance*, 49(3), 851-889. DOI: 10.3386/w4718

Ihaka, R. & Gentleman, R. (1996). A language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3), 299 – 314. DOI: 10.2307/1390807

Jaruszewicz, M. & Mandziuk, J. (2004). One day prediction of NIKKEI index considering information from other stock markets. *Proceedings of the International Conference on Artificial Intelligence and Soft Computing*, 3070, 1130–1135. DOI: 10.1007/978-3-540-24844-6_177

Karakozova, O. (2005). *Modelling and Forecasting Property Rents and Returns*. Doctoral Thesis. Swedish School of Economics and Business Administration, Department of Finance and Statistics, Helsinki, Finland. Available at: Link (Accessed on June 22, 2017)

Kim, K.-J. (2004). Artificial neural networks with feature transformation based on domain knowledge for the prediction of stock index futures. *Intelligent Systems in Accounting, Finance & Management*, 12(3), 167–176. DOI: 10.1002/isaf.252

Kimoto, T., Asakawa, K., Yoda, M. & Takeoka, M. (1990). Stock market prediction system with modular neural networks. *Proceedings of the IEEE International Conference on Neural Networks*, San Diego, pp. 1 – 16, CA, USA,. DOI: 10.1109/IJCNN.1990.137535

Leigh, W., Hightower, R. and Modani, N. (2005). Forecasting the New York Stock Exchange composite index with past price and interest rate on condition of volume spike. *Expert Systems with Applications*, 28(1), 1-8. DOI: 10.1016/j.eswa.2004.08.001

Leung, M. T., Daouk, H. & Chen, A.-S. (2000). Forecasting stock indices: a comparison of classification and level estimation models. *International Journal of Forecasting*, 16(2), 173–190. DOI: 10.1016/S0169-2070(99)00048-5

Moshiri, S. & Cameron, N. (2010). Neural network versus econometric models in forecasting inflation. *Journal of Forecasting*, 19(3), 201-217. DOI: 10.1002/(SICI)1099-131X(200004)19:3<201::AID-FOR753>3.0.CO;2-4

Mostafa, M. (2010). Forecasting stock exchange movements using neural networks: empirical evidence from Kuwait. *Expert Systems with Application*, 37(9), 6302-6309. DOI: 10.1016/j.eswa.2010.02.091

Ning, B., Wu, J., Peng, H. & Zhao, J. (2009). Using chaotic neural network to forecast stock index. *Advances in Neural Networks*, Lecture Notes in Computer Science, Vol 5551, pp. 870–876 Springer-Verlag, Heidelberg, Germany. DOI: 10.1007/978-3-642-01507-6_98

Pan, H., Tilakaratne, C. & Yearwood, J. (2005). Predicting the Australian stock market index using neural networks exploiting dynamical swings and intermarket influences. *Journal of Research and Practice in Information Technology*, 37(1), 43–55. DOI: 10.1007/978-3-540-89378-3_53

Perez-Rodriguez, J. V., Torra, S. & Andrade-Felix, J. (2005). Star and ANN models: forecasting performance on the Spanish IBEX-35 stock index. *Journal of Empirical Finance*, 12(3), 490–509. DOI: 10.1016/j.jempfin.2004.03.001
Sen J. & Datta Chaudhuri, T. (2016a). Decomposition of time series data of stock markets and its implications for prediction – an application for the Indian auto sector. *Proceedings of the 2nd National Conference on Advances in Business Research and Management Practices (ABRMP’16)*, pp. 15 – 28, Kolkata, India, January, 2016. DOI: 10.13140/RG.2.1.3232.0241.

Sen, J. & Datta Chaudhuri, T. (2016b). A framework for predictive analysis of stock market indices – a study of the Indian auto sector. *Calcutta Business School (CBS) Journal of Management Practices*, 2(2), 1-20. DOI: 10.13140/RG.2.1.2178.3448.

Sen, J. & Datta Chaudhuri, T. (2016c). An alternative framework for time series decomposition and forecasting and its relevance for portfolio choice: a comparative study of the Indian consumer durable and small cap sectors. *Journal of Economic Library*, 3(2), 303-326. DOI: 10.1453/jel.v3i2.787.

Sen, J. & Datta Chaudhuri, T. (2016d). An investigation of the structural characteristics of the Indian IT sector and the capital goods sector – an application of the R programming in time series decomposition and forecasting. *Journal of Insurance and Financial Management*, 1(4), 68-132.

Sen, J. & Datta Chaudhuri T. (2016e). Decomposition of time series data to check consistency between fund style and actual fund composition of mutual funds. *Proceedings of the 4th International Conference on Business Analytics and Intelligence (ICBAI 2016)*, Bangalore, India, December 19-21. DOI: 10.13140/RG.2.2.14152.93443.

Sen, J. & Datta Chaudhuri, T. (2017a). A time series analysis-based forecasting framework for the Indian healthcare sector. *Journal of Insurance and Financial Management*, 3(1), 66-94.

Sen, J. & Datta Chaudhuri, T. (2017b). A predictive analysis of the Indian FMCG sector using time series decomposition-based approach. *Journal of Economics Library*, 4(2). (Accepted for publication). Available online at: (Accessed on: June 24, 2017)

Shen, J., Fan, H. & Chang, S. (2007). Stock index prediction based on adaptive training and pruning algorithm. *Advances in Neural Networks*, Lecture Notes in Computer Science, Vol 4492, pp. 457–464, Springer-Verlag, Heidelberg, Germany. DOI: 10.1007/978-3-540-72393-6_55

Thenmozhi, M. (2006). Forecasting stock index numbers using neural networks. *Delhi Business Review*, 7(2), 59-69. Available online at: Link (Accessed on: May 10, 2017)

Tsai, C.-F. & Wang, S.-P. (2009). Stock price forecasting by hybrid machine learning techniques. *Proceedings of International Multi Conference of Engineers and Computer Scientists*, pp. 755 – 765, Hong Kong, March 2009.

Tseng, K.-C., Kwon, O., & Tjung, L.C. (2012). Time series and neural network forecast of daily stock prices. *Investment Management and Financial Innovations*, 9(1), 32-54.

Wang, W. & Nie, S. (2008). The performance of several combining forecasts for stock index. *International Seminar on Future Information Technology and Management Engineering*, pp. 450– 455, Leicestershire, United Kingdom, November 2008. DOI: 10.1109/FITME.2008.42

Wu, Q., Chen, Y. & Liu, Z. (2008). Ensemble model of intelligent paradigms for stock market forecasting. *Proceedings of the 1st International Workshop on Knowledge Discovery and Data Mining*, pp. 205 – 208, Washington, DC, USA, January 2008. DOI: 10.1109/WKDD.2008.54

Zhu, X., Wang, H., Xu, L. & Li, H. (2008). Predicting stock index increments by neural networks: the role of trading volume under different horizons. *Expert Systems with Applications*, 34(4), 3043–3054. DOI: 10.1016/j.eswa.2007.06.023