Trend Analysis of Indian Stock Market Using Dynamic Mode Decomposition☆

Deepthi Praveenlal Kuttichira*,[1]

Gopalakrishnan E.A[2]

Vijay Krishna Menon[3]

K.P Soman[4]

Center Of Computational Engineering and Networking, Amrita Vishwa Vidyapeetham, Coimbatore

Abstract

Stock price prediction is a challenging problem as the market is quite unpredictable. We propose a method for price prediction using Dynamic Mode Decomposition assuming stock market as a dynamic system. DMD is an equation free, data-driven, spatio-temporal algorithm which decomposes a system to modes that have predetermined temporal behaviour associated with them. These modes help us determine how the system evolves and the future state of the system can be predicted. We have used these modes for the predictive assessment of the stock market.

We worked with the time series data of the companies listed in National Stock Exchange. The granularity of time was minute. We have sampled a few companies across sectors listed in National Stock Exchange and used the minute-wise stock price to predict their price in next few minutes. The obtained price pre-

*Corresponding author

Email addresses: deepthikuttichira93@gmail.com (Deepthi Praveenlal Kuttichira)

[1]Master's Student

[2]Assistant Professor

[3]Assistant Professor

[4]Professor

diction results were compared with actual stock prices. We used Mean Absolute Percentage Error to calculate the deviation of predicted price from actual price for each company. We used three sampling approaches to sample the companies. In the first, we sampled companies belonging to the same sector to predict the future price. In the second, we considered sampled companies from four different sectors and in third, we used all the companies listed in NSE. For all the three approaches we made analysis using DMD as well as adaptive DMD. For DMD, the sampling as well as the prediction window size were fixed. For adaptive DMD, the sampling window was kept fixed, but predictions were made until it crossed a threshold error.
When predicted values and actual values were plotted in the graph, the resulting curves were similar and close. Prediction was found to be more accurate when samples were taken from all the sectors, than from a single sector. When all 1721 companies listed in NSE were considered prediction was more accurate. When sampling window alone was fixed; the predictions could be made for longer period for certain instances of sampling. At other instances, the predictions could me made only lesser than 6 minutes.
Key words: `Dynamic Mode Decomposition`, Proper Orthogonal Decomposition, Mean Absolute Percentage Error

---

1. Introduction

Stock price prediction is a challenging problem as the market is quite unpredictable. Arguments prevail,regarding the possibility of stock price prediction. Efficient market hypothesis states that the current price reflects the current state of the market and nothing further can be inferred from it[1][2]. Fundamental analysts believe that stock price predictions can be made by analyzing the current financial situations and by taking into account the financial statements of a company. Technical analysis involves studying the past data to predict future prices. Though there are a number of studies supporting efficient market hypothesis, recently there are studies that support technical analysis[3][4][5][6].

Technical analysis uses statistical analysis to predict the future price. Technical analysis is gaining wider popularity with the advent of automated trading. This revolution of automated trading was generated by the computer-enabled algorithmic trading schemes. High frequency trading attempts to take advantage of stock price fluctuations that occur on time scales that are as fast as microseconds. These high-frequency trading can be executed only by computer. Technical analysis can be used by day traders to gain profit. Day traders aim to obtain profit by using daily fluctuation of price to make gains within the day. In technical analysis time series analysis and neural networks are most widely used methods. Autoregressive(AR), Moving Average(MA), Autoregressive Moving Average(ARMA), Autoregressive Integrated Moving Average(ARIMA) are some of the models used for time series analysis[7][8].In AR the future values is determined using past values,MA uses past error values to determine future values and ARMA uses both for prediction.Changing dynamics of the financial system adds to the non-stationarity of the data which makes long term predictions inaccurate. Also time series analysis is computationally costly. As neural networks are efficient in pattern learning, these models can be used to learn patterns in time series data. Predictions using neural networks are also widely done in stock market[9].One major drawback of neural network is its requirement of large amount of training data. Each time there is a pattern change in financial market, the model has to be trained for new parameters[10][11]. This pattern change can happen quite often in financial market. Decomposition methods are also used in stock market.Wavelet modes has been used to identify business cycles that China shared with the rest of the world economy[12]. In our proposed method we assume stock market to be a dynamical system. A dynamical system can be decomposed to modes and these modes can be used to identify the evolutionary patterns of the system[13][**?**]. We use Dynamic Mode Decomposition(DMD), a data driven, spatial-temporal coherent algorithm to identify the evolutionary patterns of this system[14][15]. DMD is computationally very efficient as it exploits the low-dimensional structure of the data. Due to this advantageous property of DMD, it finds its application in many fields. A dynamical system can be decomposed to modes and these modes can be used to identify the evolutionary patterns of the system. Dynamic systems are usually described using equation that describe how their trajectories evolve in finite dimensional space. An alternative description of this can be made using an operator that acts on infinite dimensional space of functions.Koopman operator is one such infinite dimensional,linear operator.The evolution of a dynamic system from time t to t+1 can be captured using koopman operator[16][17]. This koopman operator captures the dynamics of the system. DMD decomposes this matrix into modes. These are the dynamic modes of the system. In the context of analyzing stock market, DMD modes can be thought of as coherent structures in the financial activity. DMD has been used in finance to extract cyclic nature in market as well as for price prediction[18]. In our work, we used DMD for short term prediction of price.

2. Methodology

A dynamic system is defined using a set of governing equations.We consider financial system as a non-linear dynamic system whose governing equations are not known to us.The snapshots we take corresponds to a state of the system. Snapshot of a system consists of observed measurements of that system at time t. In our case the observed measurements are the stock price of each company at the time t. A dynamic system is described using a governing set of differential equations.

$$\frac{dx}{dt} = F(x,t) \tag{1}$$

At each state we can make different kinds of measurements of the observables.The measurement function can be denoted as

$$G(x, t_k) = 0 \tag{2}$$

where $k = 1, 2, \ldots, M$ where M is measurement time.The initial condition is stated as

$$x(0) = x_0 \tag{3}$$

The non-linear function F that defines the set of governing equations is unknown. All we have is the initial conditions and the measurements taken.In our case the measurement is the stock price. In DMD procedure, we construct an approximate linear evolution of the system.

$$\frac{d\tilde{x}}{dt} = A\tilde{x} \tag{4}$$

The well-known solution for equation(4) is,

$$\tilde{x}(t) = \sum_{k=1}^{K} b_k \psi_k \exp(\omega_k t) \tag{5}$$

Here ψ_k and ω_k are the eigenvectors and eigenvalues of matrix A. If real part of eigenvalues are positive and greater than one, then it means a growing mode and growing money. If eigenvalues are negative, then it means decaying modes and losing money. The data we used was the minute-wise data of transactions from 1 July 2014 to 30 June 2015.The data was structured as date,transaction id,time,company name,price of the stock and volume of the stock. For our analysis,we selected 57 companies across sectors IT,financial services,pharma,automobiles. These 4 sectors combined hold 63.94 %market share. Our sampling interval is one minute. Not all companies have transactions in every minute. In such cases we substituted the previous stock price in the current minute. The reasoning behind this is that, even though no transaction was done in that minute,the stock price at that time is the previously transacted price.The data matrix given for DMD algorithm is snapshots of the system taken at equispaced time intervals.Each snapshot consists of the stock price of all the companies considered at that particular time instant.DMD decomposes data matrix to modes that spans spatially and has temporal frequency associated with them[14].The data matrix will be an $n \times m$ matrix where n=number of companies considered m=number of snapshots taken The matrix will be of the form,

$$X = [x_1, x_2, x_3, \dots x_m] \tag{6}$$

From this matrix we need to capture the underlying dynamics of the system.For this we splits the data matrix into,

$$X_1 = [x_1, x_2, x_3, \dots x_{m-1}] \tag{7}$$

$$X_2 = [x_2, x_3, x_4, \dots x_m] \tag{8}$$

X_2 is one time slot shifted from X_1.Let A be a linear operator that maps X_i to x_{i+1}.Then X_2 can be expressed as

$$X_2 = [Ax_1, Ax_2, Ax_3, \dots Ax_{m-1}] \tag{9}$$

$$X_2 = AX_1 \tag{10}$$

DMD expects X_1 to have low-rank structure. If not DMD fails immediately. Since$N \gg M$,the matrix X_1 is sure to have a low rank structure.At some point addition of a new snapshot to matrix X_1 will not add on to the vector space spanned by X_1. We can express x_m as combination of previous columns of X_1.

$$x_m = \sum_{i=1}^{m-1} a_i x_i + r \tag{11}$$

Here r is the residual vector.DMD algorithm does the minimization of r.

$$X_2 = X_1 S + r e_{m-1}^* \tag{12}$$

Here S is the companion matrix.e_{m-1} is (m-1)th unit vector.It can be seen that $AX_1 = X_2 \approx X_1 S$. Eigen values of A are approximately that of S. To utilize the low rank structure of X_1 we apply SVD to it. Singular Value Decomposition decomposes X_1 to $U\Sigma V^*$ where $U \in C^{n\times r}$, $\Sigma \in C^{r\times r}$ and $V \in C^{(m-1)\times r}$. Here Σ is diagonal matrix,whereas U and V are unitary matrix. Decomposing X_1 using SVD and substituting it,we get

$$X_2 \approx U\Sigma V^* S \tag{13}$$

$$S \approx V\Sigma^{-1} U^* X_2 \tag{14}$$

For robust implementation of the algorithm, instead of finding the eigen values and eigen vectors of S,we find the same for a similar matrix $\tilde{S}$

$$\tilde{S} \approx U^* X_2 V \Sigma^{-1} \tag{15}$$

The eigen values of similar matrices are same.Eigen vector can be found using similar matrix property.DMD eigen values are,

$$\tilde{S}v_j = \lambda_j v_j \tag{16}$$

DMD eigen vectors are,

$$\phi_j = Uv_j \tag{17}$$

For ease of representation DMD modes can be converted to fourier modes as $\omega_j = \ln(\lambda_j)$.So now using these eigen vectors and eigen values we can reconstruct the system as follows

$$X_{DMD}(t) = \sum_{j=1}^{r} b_j \phi_j e^{\omega_j t} = \phi diag(e^{\omega t})b \tag{18}$$

$X_{DMD}(t)$ defines the state of the system at time t.In our case,as we had taken stock prices to represent the state of the system,$X_{DMD}(t)$ gives stock prices at time t.Equ (10)is basically a regression equation.It tries to find a least square fit for the points considered. In our method we considered 6 snapshots of the system to make future predictions. In one approach the prediction window was kept fixed at the size of sampling window. In another approach the sampling window was kept same, but prediction was made until it crossed a threshold value. We sampled our data in three different ways. In the first method we used, the snapshots consisted of the companies sampled from same sector. In the second method the snapshots consisted of the companies sampled across different sectors. The sectors we considered are IT, pharma, financial services, Automobiles. The companies selected for sector-wise sampling were the ones listed in Nifty sector-wise index. The sectors were selected in such a way that, these combined held majority of the market share. In the third approach, the snapshots consisted of all the 1721 companies listed in the National Stock Exchange(NSE). In all these samplings, the size of the sampling window and prediction window was kept same. The sampling window as well as prediction window were both 6 minutes.

For all these approaches we also tried adaptive DMD. In adaptive DMD the sampling window was kept same, but the the prediction was made until it crossed a threshold error. The error was set to 7 percent. It was observed that at certain time instances, the prediction could be made for longer periods of time, say 20 minutes. But at certain instances of time predictions could be made only for a couple of minutes. MAPE was used on the results obtained by all the three approaches.

The DMD algorithm was coded using Scala. For the accuracy of comparison among three different methods of sampling, we selected the prices of one single company for this presentation. MAPE was done for different sets of companies.

3. Results

When the predicted values and actual values were plotted in the graph, the obtained curves were similar and close, in three approaches. When only 20 companies were considered, the predicted value showed sharp rises and falls, which were lesser in frequency and intensity when 57 companies were considered(Figure 1). When 1721 companies were considered, these shocks were still lesser(Figure 2). MAPE values for second and third sampling technique was significantly lower than the first sampling technique. MAPE values obtained for the first two sampling approaches for simple and adaptive DMD was comparable. But in the third method MAPE values obtained for simple DMD was evidently lower than adaptive DMD. When adaptive DMD was used, it was observed that sudden spikes in the predicted values could be mitigated. When sampling was done considering only one sector, these spikes were sharper. When sampling considered all the companies, these spikes were scarce. In figure 4, it is observed that three eigen values are on the circle and two inside. This means that the current trend would decay in future as to values correspond to decaying trend. Eigen vectors corresponding to eigen values inside the circle are decaying modes, eigen vectors corresponding to eigen values outside the circle are growing modes and eigen vectors corresponding to eigen values on the circle are stable modes.

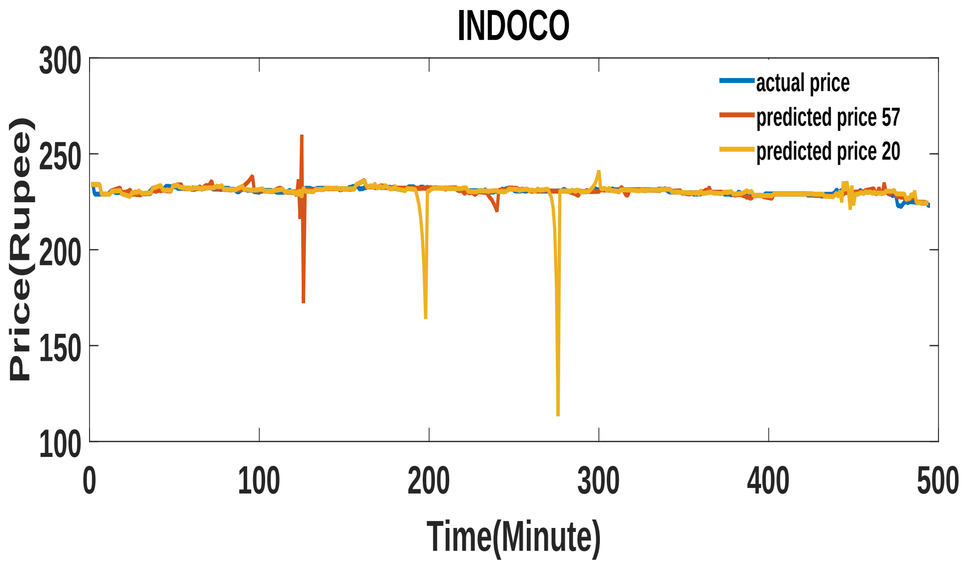


Figure 1: DMD considering Intra-sectoral sampling and Inter-sectoral sampling

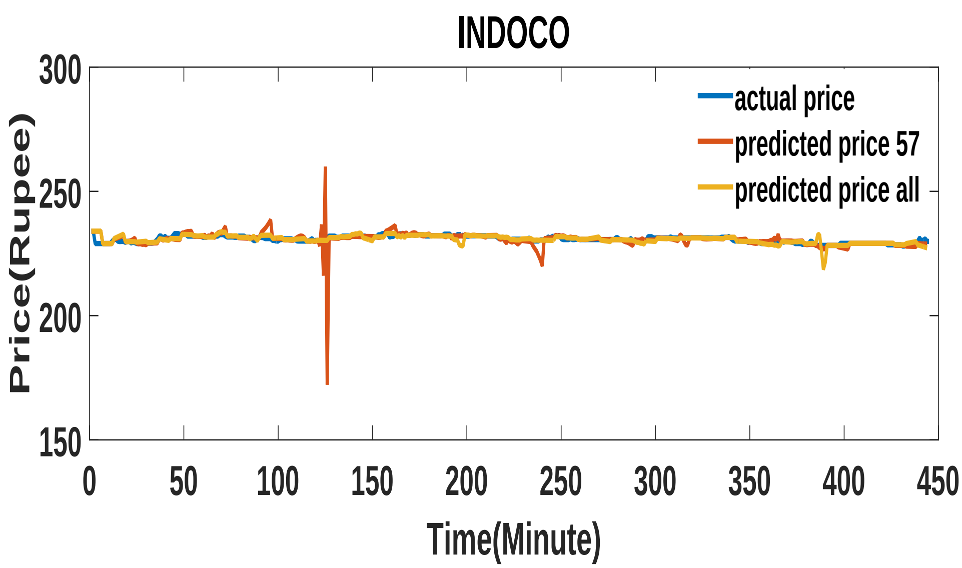


Figure 2: DMD using 57 companies VS all companies

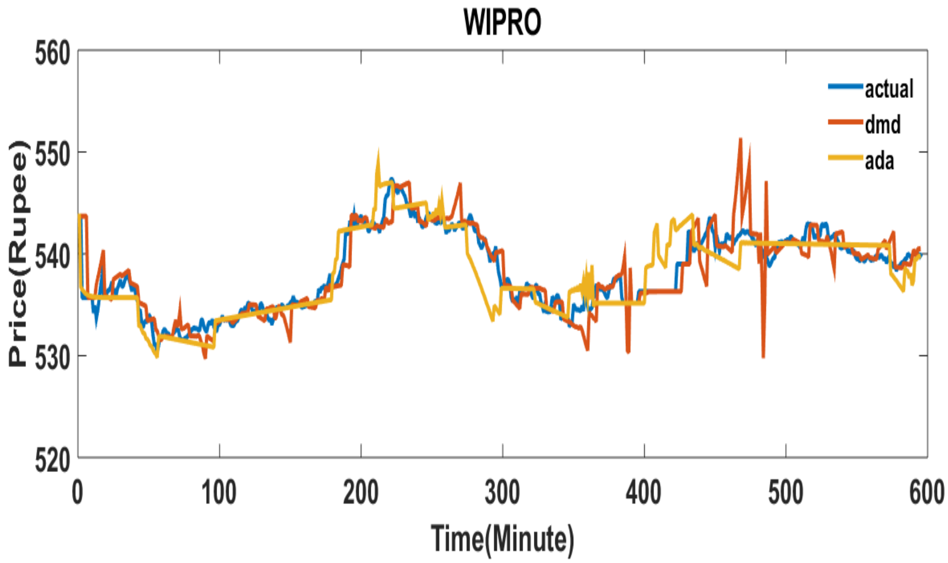


Figure 3: DMD using 57 companies VS all companies

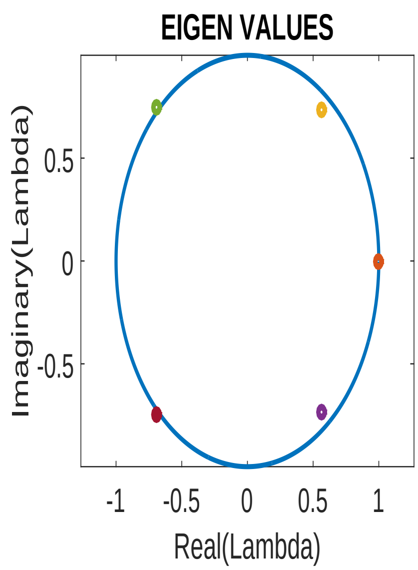
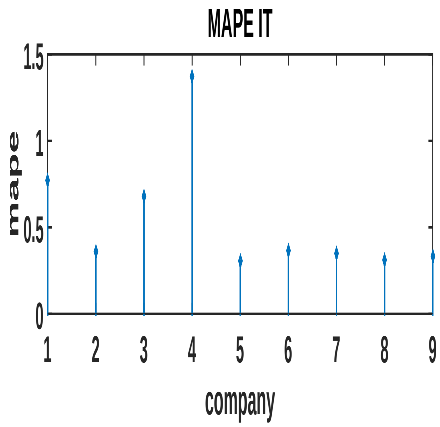
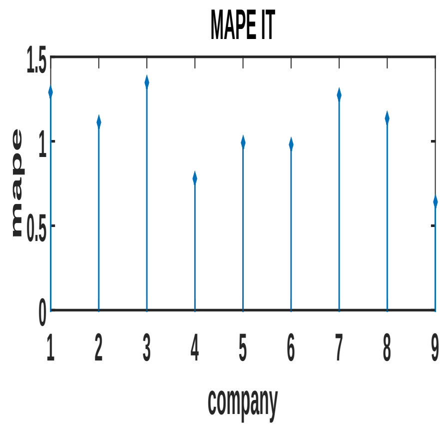


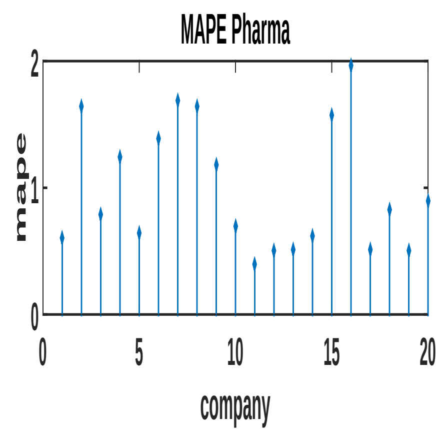
Figure 4: Eigen values on circle indicate stable modes, outside circle indicate growing modes, inside circle indicate decaying modes.



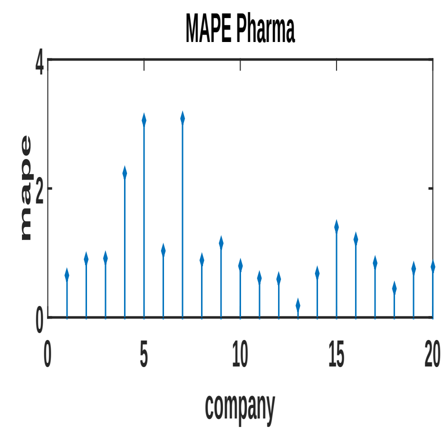
(a) Mape values for DMD



(b) Mape values for Adaptive DMD

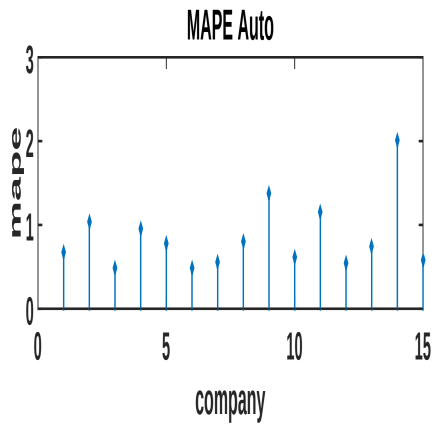


(c) Mape values for DMD

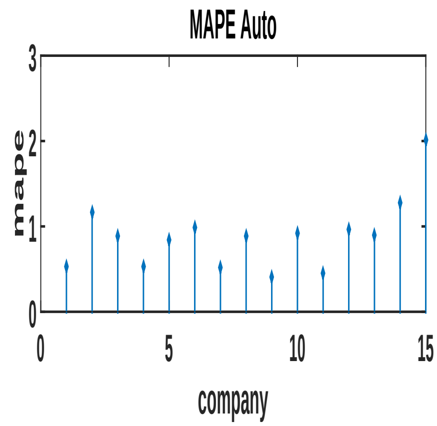


(d) Mape values for Adaptive DMD

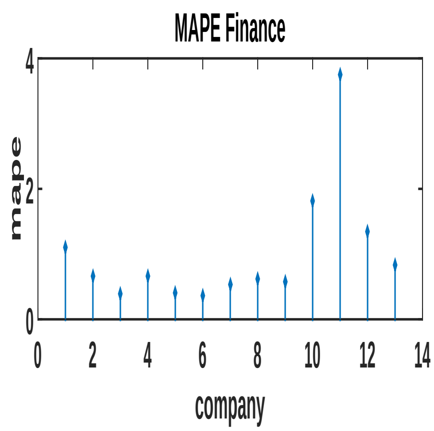
Figure 5: Mape values obtained for DMD and Adaptive DMD



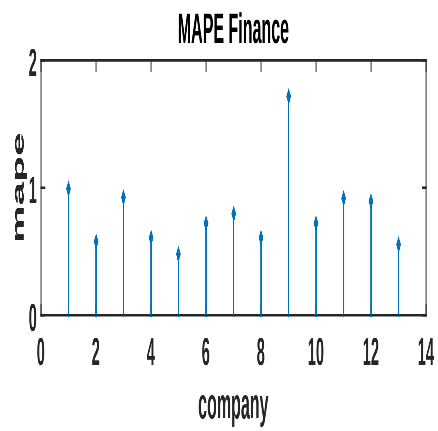
(a) Mape values for DMD



(b) Mape values for Adaptive DMD



(c) Mape values for DMD



(d) Mape values for Adaptive DMD

Figure 6: Mape values obtained for DMD and Adaptive DMD

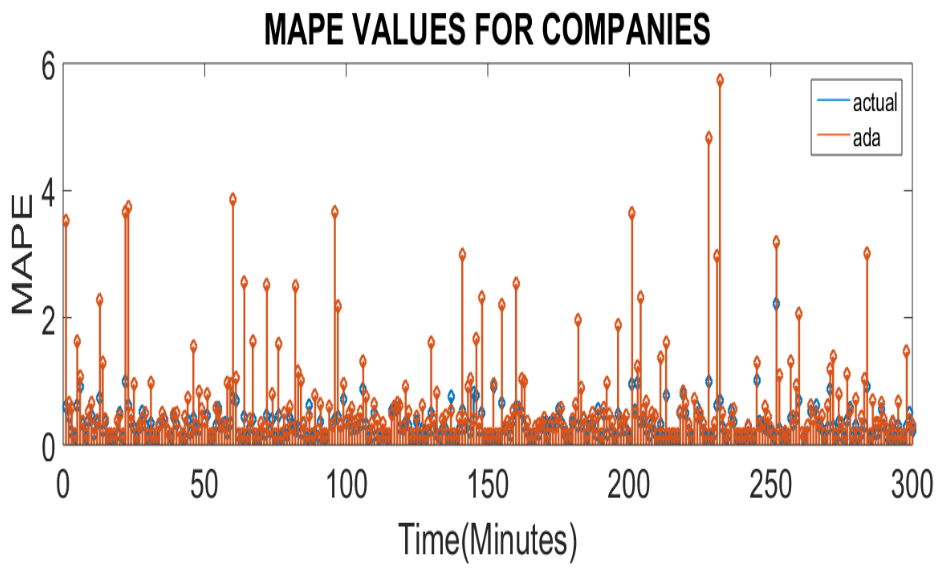


Figure 7: DMD using 57 companies VS all companies

4. Discussion

We observed that better results were obtained when prediction was made by sampling companies across sectors. Some sudden shocks that were not present in the data was predicted by DMD. This is due to some slight spikes in the data got emphasized when projected onto the the lower dimensional space. But this problem was mitigated when all 1721 companies listed in NSE was included in the snapshots. Also adaptive DMD doesnt allow these sudden shocks, as it is conditioned to not to go above a threshold error. Adaptive DMD was found to predict for about half an hour in some instances. Also in other instances adaptive DMD could predict only less than 6 minutes. In the perspective of avoiding sudden shocks adaptive DMD is better. But if the preference is for more accurate prediction, but with tolerance for sudden shocks simple DMD is preferred. Also eigen values give an insight into how the modes propagate in future. The modes can grow, decay oscillate. The eigen values inside the unit circle indicates that the corresponding eigen vectors decay, outside the circle indicates growing vectors, on the circle indicate stable ones. Eigen values with imaginary components indicate oscillating behaviour and these capture the seasonal trends in the data.In our method, we have not taken into account the exogenous variables that affect stock market. So if stock splits are announced or some new product is introduced into the market, our method will not capture those dynamics. Also DMD tends to spread out the shock in one stock to other stocks. So an anomaly that happens in on stock, tends to be reflected in all other stocks. If the number of stocks considered are more, this effect can be mitigated to a certain extend.

5. Conclusion

We found that DMD could capture the trends in stock data. Using these trends price predictions could be effectively made. It was also observed that the predictions were more accurate when all of the companies listed in NSE was considered. This qualitatively proves an inter-sectoral dependency. It affirms that market as a whole has a dynamics and all the stocks are affected by it in one way or other. This observation leads us to suggest that, investment decisions should be made by considering the whole of market dynamics than by focusing on a single stock or a subset of interested stocks. Also DMD doesnt capture sudden shocks in the market. These sudden shocks occur when stock splits are announced or when a new product is introduced in the market. But these are information that can be known prior. Including these prior information in DMD algorithm is expected to improve the results considerably. Also since DMD is computationally fast, it has an upper-hand when capturing the fast changing trend of the market. Most importantly our study emphasizes the existence of trends in stock market. This implies that there is more scope for technical analysis in this area, especially when day trading is concerned.

6. Future Work

Here we have not included the external factors like government policies, political decisions, introduction of new products etc. These can be included by using exogenous DMD[19]. Also if the stocks selected for portfolio are companies listed in Nifty index or not is an analysis that has not been done in this study. Sector-wise stability analysis is another aspect that can be studied. Analysis of eigen values can provide an insight into this. Also causal relationships between stocks has also not been explored in this study. The variational points in the data can also be identified.

References

[1] Malkiel BG. The efficient market hypothesis and its critics. The Journal of Economic Perspectives 2003;17(1):59–82.

[2] Timmermann A, Granger CW. Efficient market hypothesis and forecasting. International Journal of forecasting 2004;20(1):15–27.

[3] Abinaya P, Kumar VS, Balasubramanian P, Menon VK. Measuring stock price and trading volume causality among nifty50 stocks: The toda yamamoto method 2016;:1886–90.

[4] Bessembinder H, Chan K. The profitability of technical trading rules in the asian stock markets. Pacific-Basin Finance Journal 1995;3(2):257–84.

[5] Brock W, Lakonishok J, LeBaron B. Simple technical trading rules and the stochastic properties of stock returns. The Journal of finance 1992;47(5):1731–64.

[6] Matia K, Ashkenazy Y, Stanley HE. Multifractal properties of price fluctuations of stocks and commodities. EPL (Europhysics Letters) 2003;61(3):422.

[7] Ariyo AA, Adewumi AO, Ayo CK. Stock price prediction using the arima model 2014;:106–12.

[8] Gunasekarage A, Power DM. The profitability of moving average trading rules in south asian stock markets. Emerging Markets Review 2001;2(1):17–33.

[9] Mitra SK. Optimal combination of trading rules using neural networks. International Business Research 2009;2(1):86.

[10] Kar A. Stock prediction using artificial neural networks. Dept of Computer Science and Engineering, IIT Kanpur ????;.

[11] Majumder M, Hussian M. Forecasting of indian stock market index using artificial neural network. Available from: www nse-india com/content/research/FinalPaper206 pdf 2007;.

[12] Poměnková J, Fidrmuc J, Korhonen I. China and the world economy: Wavelet spectrum analysis of business cycles 2014;.

[13] Schmid PJ. Dynamic mode decomposition of numerical and experimental data. Journal of fluid mechanics 2010;656:5–28.

[14] Mann J, Kutz JN. Dynamic mode decomposition for financial trading strategies. Quantitative Finance 2016;16(11):1643–55.

[15] Cui Lx, Long W. Trading strategy based on dynamic mode decomposition: Tested in chinese stock market. Physica A: Statistical Mechanics and its Applications 2016;461:498–508.

[16] Mauroy A, Mezić I. Global stability analysis using the eigenfunctions of the koopman operator. IEEE Transactions on Automatic Control 2016;61(11):3356–69.

[17] Williams MO, Kevrekidis IG, Rowley CW. A data–driven approximation of the koopman operator: Extending dynamic mode decomposition. Journal of Nonlinear Science 2015;25(6):1307–46.

[18] Hua JC, Roy S, McCauley JL, Gunaratne GH. Using dynamic mode decomposition to extract cyclic behavior in the stock market. Physica A: Statistical Mechanics and its Applications 2016;448:172–80.

[19] Proctor JL, Brunton SL, Kutz JN. Dynamic mode decomposition with control. SIAM Journal on Applied Dynamical Systems 2016;15(1):142–61.