

## Design of Volatility Model in Nifty 50 Index Using Thin Plate Spline Regression

Poornima<sup>1</sup>, Vijayalakshmi<sup>2</sup>, Somasundaram<sup>3</sup>

<sup>1</sup>Department of Mathematics, Arunai Engineering College, Thiruvannamalai

<sup>2</sup>SAS, Mathematics Division, MIT University, Chennai

<sup>3</sup>Department of Statistics, Manonmaniam Sundaranar University, Tirunelveli – 627 012, India

\*Corresponding author, e-mail: subramanivit@gmail.com

### Abstract

*The analysis of volatility in stock markets has important consequences for investors and traders. The presence of volatility increases market risks and therefore discourages investment in the stock market. The proper study and understanding of volatility is needed for prudent risk management. In this paper, the market volatility in the National Stock Exchange in India as measured by the India Volatility Index is analyzed. The daily volatility in NIFTY 50 index is regressed on the price to earnings ratio and the volatility of previous day. The market volatility within a period of time is highly correlated and the highly volatile periods coincide with large impact negative events on a national and global scale. The Price to Earnings ratio represent the fundamentals of the market and it also strongly influences the price movements. The nonlinear regression problem is formulated and solved using thin plate spline regression technique. This effectively captures the nonlinear aspect of the problem. Results indicate that volatility has high upward correlation during middle range of P/E ratios than in the upper and lower ranges. Therefore risk management techniques using option derivatives are more important during the middle range of values of P/E ratio.*

**Keywords:** Volatility, Thin Plate Regression Splines, Nonlinear Regression, P/E Ratio, NIFTY 50, INDIAX, National Stock Exchange

**Copyright © 2017 Institute of Advanced Engineering and Science. All rights reserved.**

### 1. Introduction

The Stock Exchange is a centralized market where stocks, bonds and other securities are traded. Across the globe hundreds of stock exchanges function in different time zones providing capital for businesses. The market participants called brokers place buy or sell orders on behalf of their clients ranging from retail investors to large investment firms. The prices of securities are in constant fluctuation as the participants digest information that affect the businesses and make decisions [1]. The National Stock Exchange (NSE) in India was established in 1992 and is the leading stock exchange in the country. The NIFTY 50 index [2] which is a weighted average of 50 significant companies provides a snapshot into the business investment scenario in India much like a heart monitor measures the health level of a patient. In this paper the volatility in stock markets has been studied for the benefit of investors and traders. The presence of volatility increases market risks and it leads to lesser investment in the stock market. Hence the proper study and understanding of volatility is needed in volatility situations.

In financial theory, the Efficient Market Hypothesis [3] is a widely discussed theory that says that market prices reflect all available information at any point of time. Still the response of traders and investors to the same information varies widely leading fluctuations. Sometimes fluctuations are caused by uncertain economic climate due to global and national events. Although fluctuations in prices are accepted as natural, excessive price fluctuations known as market volatility severely hinders investment and economic growth [4]. Market Volatility has a strong impact on capital costing and asset allocation [5]. The investors are interested in quantifying market risks to avoid unexpected financial shocks and be better prepared for eventualities. In times of extreme uncertainty it has been observed that investors stay out of the markets leading to declining valuations and capital erosion in the economy. Therefore it is vital to understand market volatility and be able to predict and manage it using risk management techniques.

The fundamental analysis of stock prices is the prediction of future prices through analysis of the business parameters of the company and the macro economic factors [6]. The subjective nature of fundamental analysis and the difficulty in accurate prediction has led to technical analysis [7]. In technical analysis, the historical price movements are used to predict the future based on certain patterns. Although both approaches have their usefulness, successful investors and traders always employ prudent risk management techniques. The ability to gauge and understand market volatility is essential in risk management.

The time series approach to predicting volatility was utilized by Stephen Brown [8], Engle [9] and Abdurrahman Aydemir [10]. Volatility in Currency market was predicted using Implied Standard deviations (ISD) by Philippe Jorion [11]. It was observed that ISDs outperformed classical time series models in accuracy. However ISDs were biased and had statistical errors. Implied volatility was found to have upward biased forecast but it also contains information about future volatility [12]. The time series models included simple moving average, exponential smoothing, exponential weighted moving average. James Taylor proposed a more flexible exponential smoothing [13]. Auto regression is also used to forecast volatility such as ARMA, ARIMA and ARFIMA models [14]. Another class of regression methods are the ARCH class of conditional volatility models which allow heteroscedasticity in the regression [15]. Heteroscedasticity is the condition in which the regression residual is not independent of the predictors. The Exponential GARCH known as EGARCH was introduced by Nelson in 1991 which specifies conditional variance in logarithmic form [16]. Threshold GARCH or TGARCH models were introduced by Lawrence Glosten, Ravi Jagannathan and David Runkle in 1993 [17]. Philip Franses and Dick Van Dijk reviewed other nonlinear GARCH such as Quadratic GARCH (QGARCH) [18]. Stephen Gray in 1996 [19] and Franc Klassen in 2002 [20] used a generalized form of regime switching model known as RS-GARCH (1,1). Several authors [21], [22], [23] have used stochastic volatility models based on derivatives. Derivatives are risk management securities traded in the stock market with asymmetric payoff conditional on the performance of the underlying assets. The Black-Scholes model [24] relates option pricing with the volatility of the market.

This study focuses on the volatility of NSE measured as Indiavix which is a tradable index itself [25]. The dataset used is from the NSE website. The dependence of Indiavix on past Indiavix as well as the fundamentals represented by the Price to Earnings ratio (P/E) is analyzed.

## 2. IndiaVIX Volatility Index

India VIX is a volatility index based on the index option prices of NIFTY 50 index. It is calculated using the bid and ask prices of all NIFTY options [26]. Options are of two types. A PUT option gives the right to the buyer to sell a specified quantity of the underlying stock at a specified price. A CALL option gives the right to buy at a specified price. In both cases, the seller of the option is obliged to the exchange. This asymmetric nature of options allow the investors to hedge their investments effectively transferring the risk to others willing to take them for a premium price. At any point in time, any option price is the sum of intrinsic value as well as implied volatility. Higher the prices, higher the anticipation of volatility in the participants. Figure 1 displays the plot of India VIX from 2011 to 2016 considered in this study. Some of the essential statistics of India VIX are given in Table 1. The above average volatility has been observed only for 40% of the days. However very high volatility days of above average plus one standard deviation is more than the very low volatility days of average minus one standard deviation. This implies that although markets are calm more days, when the volatility spikes it is prolonged and intense. This can be attributed to the effect of fear which is a strong emotion on market decisions. Fear can feed on itself to become panic and escalate much quicker than calmness.

Table 1. Statistics of India VIX from Jan 2011 to Aug 2016

Statistic	Value
Mean Value ( $\mu$ )	19.21
Maximum Value	37.71
Minimum Value	11.57
Standard Deviation ( $\sigma$ )	4.76
% Days Above $\mu$	38.72
% Days Below $\mu$	61.28
% Days Above $\mu + \sigma$	17.60
% Days Below $\mu - \sigma$	11.71

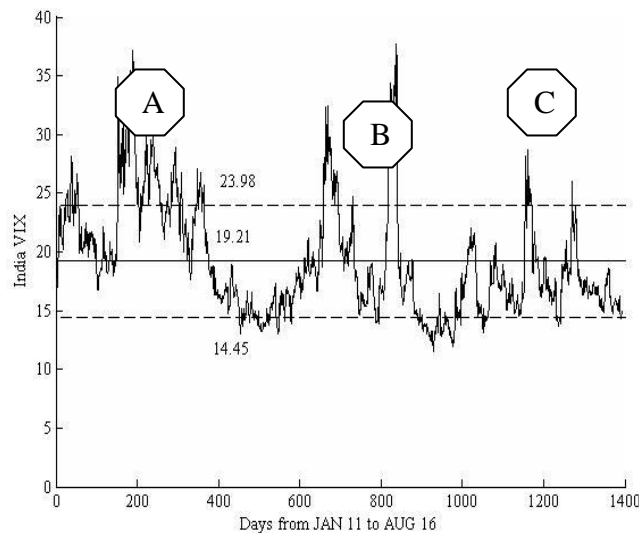


Figure 1. India VIX from Jan 2011 to Aug 2016 The mean, mean + sd, mean - sd levels are marked by lines

The distinct highly volatile periods are marked in Figure 1. The period of August to October 2011 marked A is the global panic due to European Sovereign debt crisis to Spain and Italy, downgrading of credit rating of France and the slow growth in the United States. It was a lingering effect of 2008 financial meltdown. Due to interference from central banks, the volatility slowly subsided. The period of August to September 2013 marked B is the effect of investors anticipating a cut to the bond buying stimulus program of United States Federal Reserve also known as Tapering. Indian stocks lost nearly 11% of their value. The period of May 2014 marked C also saw high volatility due to the historic verdict of Indian general elections resulting in change of leadership at the helm of government. The prices dropped drastically in periods A and B, but increased significantly during C.

### 3. Price to Earnings Ratio (P/E)

P/E ratio is the fundamental parameter driving markets [27]. If all other news, speculation and forces of change in macro economic climate are discounted, share buying and selling is done purely on the basis of P/E ratio. It is the ratio of a company's share price to its earnings per share (EPS).

$$P/E \text{ ratio} = \frac{\text{Share Price}}{\text{Earnings per Share}} \quad (1)$$

The P/E ratio of NIFTY 50 largely remains within a band of 15 to 25 and exhibits wavelike patterns with different periods. It is a useful measure of the profitability of a company. Lower the value higher are the returns of investing in the company. However the real success of investing is in identifying companies with low priced poor earning company which has the potential for growth and profitability in the future. The P/E ratio is not expected to have a

straightforward relationship with volatility but the characteristics of volatility may be dependent on the current range of P/E values. In times of panic, all fundamental parameters are usually ignored but during the upper range of P/E, the markets may assimilate panic differently than in the lower range. Figure 2 shows the P/E ratio of NIFTY 50 during the period January 2011 to August 2016.

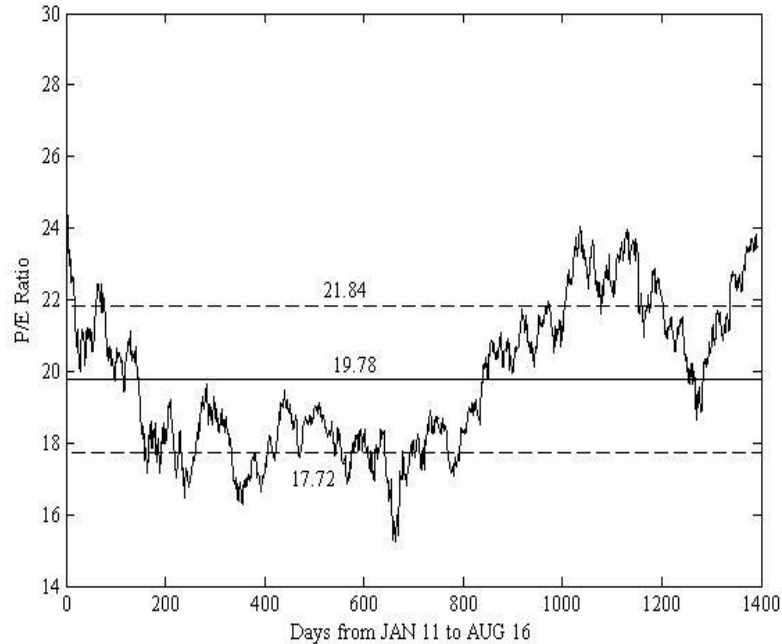


Figure 2. P/E ratio from JAN 2011 to AUG 2016

The index is used to analyze volatility rather than individual stocks. This is due to the fact that individual stock prices are affected by noise like random events pertaining to the company or sector. Further there may be times, when the stocks are not popular for investing and facing liquidity issues. But an index is free from noise. NIFTY 50 is one of most traded indices in the world and faces no liquidity issues. It is a widely watched barometer of economic performance of India and the wider emerging markets.

#### 4. Thin Plate Spline Regression

Cubic splines are a useful mathematical tool for nonlinear regression. Thin plate splines [28] allow smooth nonlinear regression of many independent predictor variables with noisy data.

The advantages of thin plate splines over cubic splines are

- 1) The subjective selection of knot points in cubic spline is eliminated.
- 2) The optimal smoothness results in a mathematical way from the equations.
- 3) The bases can represent any number of predictor variables.

Let the data consist of  $n$  observations  $(y_i, x_i)$ . Then the nonlinear regression problem becomes that of finding a suitable smooth estimate of the function  $g$  in

$$y_i = g(x_i) + \epsilon_i \quad (1)$$

where  $\epsilon_i$  is a random error term and  $x$  is a  $d$ -vector ( $d \leq n$ ). In this work, the case of  $d=2$  i.e., two predictor variables are considered. The thin plate spline smooth estimate to equation 1 is given by the minimizing function  $f$  of the optimization problem

$$\|y - f\|^2 + \lambda J_{md}(f) \quad (2)$$

where  $y$  is the vector of data  $y_i$  and  $f$  is  $\{f(x_1), f(x_2), \dots, f(x_n)\}$ .  $J_{md}(f)$  is a penalty functional that measures the non-smoothness or wiggleness of the function  $f$ .  $\lambda$  is the smoothing parameter. High values of  $\lambda$  will lead to extremely smooth  $f$  which will reduce the problem closer to linear regression. Low values of  $\lambda$  will allow wiggly function and can lead to over fitting problems. The regression solution will fit the sample data accurately but fail when other samples from the same population is used. A value of 1.4 is suggested in the literature for  $\lambda$ .

The wiggleness penalty functional is defined by

$$J_{md} = \int \dots \int \sum_{\mathbb{R}^d} v_1 + \dots + v_d = m \frac{m!}{v_1! \dots v_d!} \left( \frac{\partial^m f}{\partial x_1^{v_1} \dots \partial x_d^{v_d}} \right)^2 dx_1 \dots dx_d \quad (3)$$

For visually smooth results  $m$  is chosen such that  $2m > d + 1$ .

An appropriate choice of value for the parameter  $m$  is 2.

For the case of  $d = 2$  and  $m = 2$  the coefficients  $(v_1, v_2)$  takes the values  $\{(2,0), (1,1), (0,2)\}$  and the penalty functional is given by

$$J_{22} = \int \int \left( \frac{\partial^2 f}{\partial x_1^2} \right)^2 + \left( \frac{\partial^2 f}{\partial x_1 \partial x_2} \right)^2 + \left( \frac{\partial^2 f}{\partial x_2^2} \right)^2 dx_1 dx_2 \quad (4)$$

The function that minimizes equation 2 for this case has the form

$$f(x) = \sum_{i=1}^n \delta_i \eta_{22}(\|x - x_i\|) + \sum_{j=1}^M \alpha_j \phi_j(x) \quad (5)$$

subject to the linear constraints  $T^T \delta = 0$  where  $T_{ij} = \phi_j(x_i)$ . The  $M = \binom{m+d-1}{d} = \binom{3}{2} = 3$  functions  $\phi_i$  are linearly independent polynomials spanning the null space of  $J_{32}$  given by

$$\phi_i = \{1, x_1, x_2, \} \quad (6)$$

The basis function  $\eta_{22}$  is given for the case  $m=2, d=2$  (Woods) by

$$\eta_{22}(r) = \frac{r^2 \log_e r}{8\pi} \quad (7)$$

The thin plate spline fitting problem becomes

$$\text{minimize } \|y - E\delta - T\alpha\|^2 + \lambda \delta^T E \delta \text{ subject to } T^T \delta = 0 \quad (8)$$

with respect to  $\delta$  and  $\alpha$ . The matrix  $E$  is given by  $E_{ij} = \frac{\|x_i - x_j\|^2 \log_e \|x_i - x_j\|}{8\pi}$ . Thin plate regression splines are approximations to thin plate splines which are computationally more efficient. The linear constraints can be incorporated into the minimization function by QR decomposition of the matrix  $T$ . Let  $T = U \begin{bmatrix} P \\ 0 \end{bmatrix}$  be the QR decomposition of  $T$ , where  $U$  is a  $n \times n$  orthogonal matrix and  $P$  is an  $n \times 6$  upper triangular matrix.  $U$  can be partitioned into  $U = (D:Z)$  where  $Z$  is a  $n \times (n-3)$  matrix. Then the minimization problem can be turned into the following unconstrained problem

$$\text{minimize } \|y - EZ\delta_z - T\alpha\|^2 + \lambda \delta_z^T Z^T E Z \delta_z \quad (9)$$

which is solved for  $\delta_z$  and  $\alpha$ . Here  $\delta = Z\delta_z$ . Once  $\delta$  and  $\alpha$  are calculated equation 5 gives the fitting function.

## 5. Experimental Results and Analysis

The result is shown in Figure 3. The upward correlation of volatility is apparently visible in the plot. The effect of P/E ratio shows a nonlinear relation to the volatility. In the upper range of [22-25] and lower range of [15-18], the P/E ratio does not affect the upward correlation of volatility. But in the middle range of [18-22], the volatility becomes more unpredictable. Figure 4

shows the contour plot for better visibility. Index long term trend reversals usually occur in the lower and upper ranges of P/E ratio. During trend reversals volatility is more subdued than in the middle range.

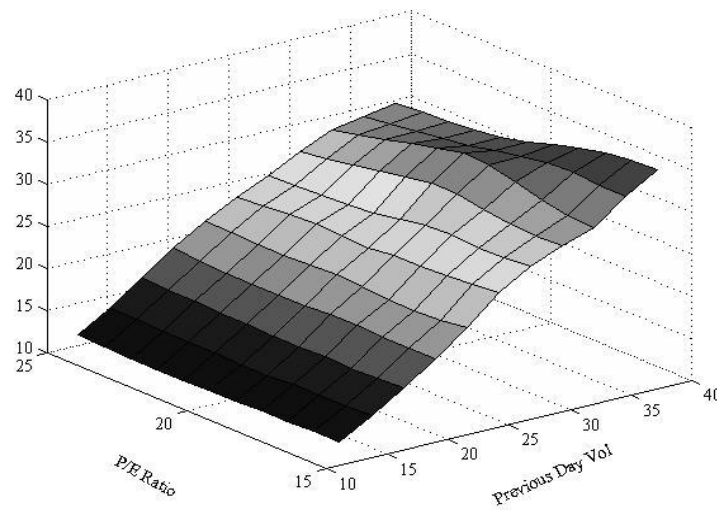


Figure 3. Result of Thin Plate Regression

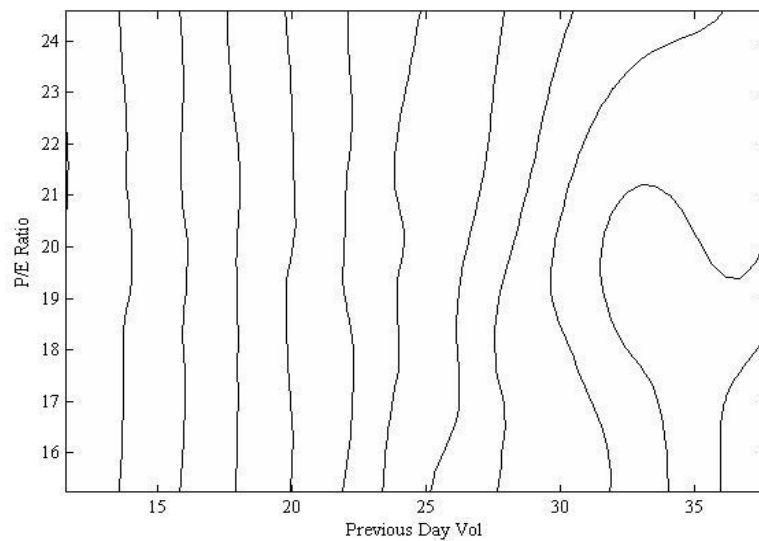


Figure 4. Contour Plot of Thin Plate Regression

The predicted volatility vs. the actual volatility is displayed in Figure 5. Much of the predictive accuracy is due to the upward correlation of volatility. The residuals are more pronounced in highly volatile days as expected.

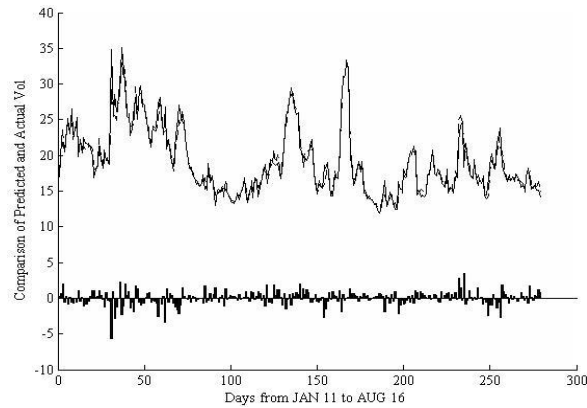


Figure 5. Predicted vs. Actual Volatility, Residuals

To assess the impact of including P/E ratio in the forecast, a simple linear predictor involving only the previous day's volatility is used. The results are shown in Figure 6. The Mean Squared Error of the forecast for the two predictor case is 1.1515 whereas for the single predictor case it is 1.3616.

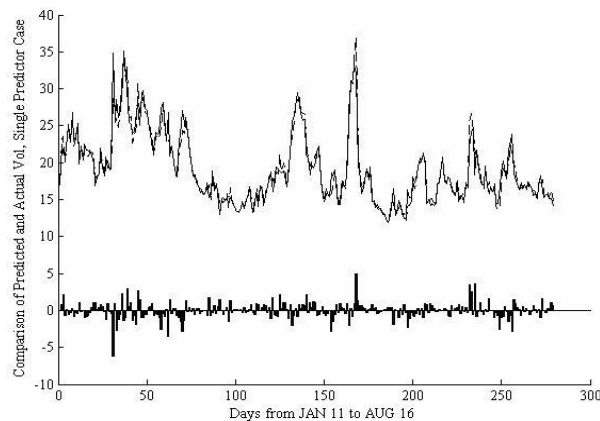


Figure 6. Predicted vs. Actual Volatility, Residuals in Single Predictor Case

## 6. Conclusion

This paper presented a study of volatility of NIFTY 50 index. The predictor variables are the previous day's volatility measured by IndiaVIX index and P/E ratios. Thin plate spline regression was performed and the results indicate that the upward correlation of IndiaVIX values are more pronounced during the middle range of P/E ratios than the upper and lower ranges. Most of the sharp changes in volatility have occurred when P/E ratio is in the range of 18 to 22. During low P/E ratio periods, the prices are low and the market participants are not fully invested. In high P/E ratio periods, the participants are highly invested and are willing to wait for returns. During the middle range of P/E ratios, the participants are not fully committed and are vary of the risk to their investments. This allows the high volatility periods to run longer. In the future more variables such as interest rates of RBI, global factors such as interest rates, economic growth in developed nations could be included in the analysis.

## References

- [1] Kimoto, Takashi, et al. "Stock market prediction system with modular neural networks". *1990 IJCNN International Joint Conference on Neural Networks*, IEEE, 1990.

- [2] Kara, Yakup, Melek Acar Boyacioglu, and Ömer Kaan Baykan. "Predicting direction of stock price index movement using artificial neural networks and support vector machines: The sample of the Istanbul Stock Exchange". *Expert systems with Applications*. 2011; 38(5): 5311-5319.
- [3] Malkiel, Burton G. "Efficient market hypothesis." *The World of Economics*. Palgrave Macmillan UK. 1991: 211-218.
- [4] Kogan, Leonid. "Asset prices and real investment". *Journal of Financial Economics*. 2004; 73(3): 411-431.
- [5] Gabaix, Xavier, et al. *Institutional investors and stock market volatility*. No. w11722. National Bureau of Economic Research. 2005.
- [6] Fama, Eugene F. "Random walks in stock market prices." *Financial analyst's journal*. 1995; 51(1): 75-80.
- [7] Brock, William, Josef Lakonishok, and Blake LeBaron. "Simple technical trading rules and the stochastic properties of stock returns". *The Journal of finance*. 1992; 47(5): 1731-1764.
- [8] Brown, Stephen J. "Estimating volatility." *Financial Options: From Theory to Practice*. 1990; 516: 537.
- [9] Bollerslev, Tim, Robert F. Engle and Daniel B. Nelson. "ARCH models". *Handbook of econometrics*. 1994; 4: 2959-3038.
- [10] Aydemir, Abdurrahman Bekir. "Volatility modelling in finance". *Forecasting Volatility in the Financial Markets*. 2002: 1.
- [11] Jorion, Philippe. "Predicting volatility in the foreign exchange market." *The Journal of Finance*. 1995; 50(2): 507-528.
- [12] Fleming, Jeff. "The quality of market volatility forecasts implied by S&P 100 index option prices". *Journal of empirical finance*. 1998; 5(4): 317-345.
- [13] Taylor, James W. "Short-term electricity demand forecasting using double seasonal exponential smoothing". *Journal of the Operational Research Society*. 2003; 54(8): 799-805.
- [14] Andersen, Torben G. "Stochastic autoregressive volatility: a framework for volatility modeling". *Mathematical finance*. 1994; 4(2): 75-102.
- [15] Bollerslev, Tim, Robert F. Engle, and Daniel B. Nelson. "ARCH models." *Handbook of econometrics*. 1994; 4: 2959-3038.
- [16] Nelson, Daniel B. "Asymptotic filtering theory for multivariate ARCH models". 1994.
- [17] Glosten, Lawrence R, Ravi Jagannathan and David E Runkle. "On the relation between the expected value and the volatility of the nominal excess return on stocks". *The journal of finance*. 1993; 48(5): 1779-1801.
- [18] Poon, Ser-Huang, and Clive WJ Granger. "Forecasting volatility in financial markets: A review". *Journal of economic literature*. 2003; 41(2): 478-539.
- [19] Bollen, Nicolas PB, Stephen F Gray, and Robert E. Whaley. "Regime switching in foreign exchange rates: Evidence from currency option prices". *Journal of Econometrics*. 2000; 94(1): 239-276.
- [20] Klaassen, Franc. "Long swings in exchange rates: Are they really in the data?" *Journal of Business & Economic Statistics*. 2005; 23(1): 87-95.
- [21] Hull, John, and Alan White. "The pricing of options on assets with stochastic volatilities." *The journal of finance*. 1987; 42(2): 281-300.
- [22] Wiggins, James B. "Option values under stochastic volatility: Theory and empirical estimates". *Journal of financial economics*. 1987; 19(2): 351-372.
- [23] Heston, Steven L. "A closed-form solution for options with stochastic volatility with applications to bond and currency options". *Review of financial studies*. 1993; 6(2): 327-343.
- [24] Black, Fischer, and Myron Scholes. "The pricing of options and corporate liabilities". *The journal of political economy*. 1973: 637-654.
- [25] Shaikh, Imlak, and Puja Padhi. "The information content of implied volatility index (India VIX)." *Global Business Perspectives*. 2013; 1(4): 359-378.
- [26] Bandivadekar, Snehal, and Saurabh Ghosh. "Derivatives and volatility on Indian stock markets". *Reserve Bank of India Occasional Papers*. 2003; 24(3): 187-201.
- [27] Siegel, Jeremy J. "The nifty-fifty revisited: do growth stocks ultimately justify their price?" *The Journal of Portfolio Management*. 1995; 21(4): 8-20.
- [28] Wood, Simon N. "Thin plate regression splines". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*. 2003; 65(1): 95-114.