Comparing Time Series Model for Future Prediction of Stock Market

A. Sheik Abdullah and K. Shalini
PG and Research Department of Statistics
Salem Sowdeswari College
Salem – 636 010.
Tamil Nadu, India.

ABSTRACT

Stock market volatility is the important factor for investment, option pricing and financial market regulation. In recent years, stock market analysis and prediction have the greatest significance for many fields of stock market. There are so many time series models available in the literature to solve the problem of future prediction. In this paper, comparative study is carried out of Single Exponential Smoothing (SES) model, Auto Regressive Integrated Moving Average (ARIMA) model and Generalized Autoregressive Conditional Heteroscedasticity (GARCH). Values are forecasted based on the above mentioned three time series models. Forecasted values for different α values are calculated from SES method, ARIMA (0, 1, 0) and GARCH(1,1). Also, Mean Square Error (MSE), Root Mean Square (RMSE), Mean Absolute Deviation (MAD) and Mean Absolute Percentage Error (MAPE) are calculated individually for the three methods.

Keywords: *Time series, single exponential smoothing, ARIMA, GARCH.*

1. INTRODUCTION

Stock markets play an important role for allocation of savings to investment and provide a variety of assets to savers in various forms. At the same time, the financial market is integral part of the financial system which makes significant contribution to the countries economic development.

Time series analysis is carried out for identifying the nature of the data represented by the sequence of observations and forecasting. Exponential smoothing methods are one of the most efficacious method adopted for forecasting the time series in seasonal patterns; moreover, it is also easy to adjust for past errors and several different practices are used depending on the presence of trend or cyclical variations. Meanwhile, it is an averaging technique that uses unequal weights where weights are applied to past observations decline in an exponential manner.

The construction of forecast model based on discounted past observations is most commonly carried out by exponential smoothing procedures. These procedures are very vital and attractive in time series forecasting, which they allow the forecast function to be updated very easily and every time a new observation become popular. They are very easy to implement and it can be quite effective for forecasting, which is implemented without respect to a properly defined statistical model.

Generalized Autoregressive Conditional Heteroscedasticity (GARCH) models have become important tools in the analysis of time series data, particularly in stock market related applications. These models are especially useful when the goal of the study is to analyze and forecast volatility. Financial decisions are generally based upon the tradeoff between risk and return; the econometric analysis of risk is therefore an integral part of asset pricing, portfolio, optimization, option pricing and risk management. GARCH processes exhibit heavy tails. Therefore, when GARCH models are adopted it can model both the conditional heteroscedastic and the heavy-tailed distributions of financial markets data.

2 LITERATURE SURVEY

Related literature for various forecasting methods based on SES, ARIMA and GARCH are discussed in this section. Landsman and Damodaran (1989) have estimated the James-Stein ARIMA parameter estimator and which improves the forecast accuracy compare with other methods, under MSE loss criterion. Kim (2003) estimated in the model parameter and the forecasting of AR models for small samples and found that (bootstrap) bias-corrected parameter estimators provide more accurate forecasts than the least square estimator.

Brown (1950) extended single exponential smoothing and developed the methods for trends and seasonality in discrete data. The developed model was applied in forecasting the demand for spare parts in Navy inventory systems. Gardner (1985) started that exponential smoothing should be disregarded because it was either a special case of ARIMA modeling. Hyndman et. al., (2002) have provided a new approach to automatic forecasting based on an extended range of exponential smoothing methods. Taylor (2003) introduced a new damped exponential smoothing method follows the multiplication trend formulation also, suggested that the assumption of a multiplicative trend is not difficult to forecast the values.

Blair et al., (2002) compared the volatility of the S&P 100 index and all its constituent stocks by estimating Autoregressive Conditional Heteroscedasticity (ARCH) and Threshold ARCH model. They concluded that a majority of stocks have a greater volatility response to negative returns than to positive returns and the asymmetry is higher for the index than for most stocks.

Deb et al., (2003) forecasted the monthly volatility of market indices (Sensex & S&PCNX-Nifty) of Indian capital markets using eight different univariate models. Out-of-sample forecasting performance of these models shows that GARCH (1, 1) model outperforms the other models. Karmakar (2005) used conditional volatility models to estimate the volatility of fifty individual stocks of Indian stock market and observed that the GARCH (1, 1) model provides reasonably good forecast. Further, applied for Indian stock exchange and observed that the GARCH model provides reasonably good forecast. Similarly, Magnus and Fosu (2006) for the Ghana Stock Exchange Market, Liu and Morley (2009) for the Hong Kong stock market established that GARCH technique offers a superior performance than historical averaging models.

3. MATERIALS AND METHODS

3.1 Data Source

Time series is a sequence of observations which are taken sequentially in time.

The daily closing values (Rupees) of S&P BSE and Nifty 50 are obtained during the period 1st January 2015 to 31st December 2018 has been collected through yahoofinance.com. Basic statistics of the selected data are given in the table 1.

Statistics	S&P BSE	Nifty 50
Mean	30019.56	9181.82
S.D	3855.45	1172.76
Maximum	38896.63	11738.5
Minimum	22951.83	6970.60
Skewness	0.403	-1.010
Kurtosis	0.277	-1.178

3.2 Single Exponential Smoothing

SES calculates the smoothed series as a damping coefficient period the actual values. The extrapolated smoothed series is a constant, equal to the last value of the smoothed series during the period when actual data on the underlying series are available. The simple moving average method is a special case of the exponential smoothing; the exponential smoothing is more in its data usage. Figure 1, shows the flow chart for selecting SES model. In exponential smoothing, a new estimate is the addition of the estimate for the present time period and a portion of the error $(x_t - \hat{x}_t)$ generated in the present time period, that is

$$\hat{x}_{t+1} = \hat{x}_t + \alpha(e_t) \qquad (1)$$

This equation is usually written as $S_t = S_{t-1} + \alpha(x_t - S_{t-1})$ (2)

where S_t = the new estimated value or forecasted value for the next time period (made in the present time period);

S_{t-1} = the estimated or forecasted value for the present time period (made in the last time period);

x_t = the actual data value in the present time period;

$x_t - S_{t-1}$ = estimating or forecasting error for the present time period;

α = weight value or discount ranges between 0 and 1.

The smoothing equation can be written as

$$\hat{x}_{t+1} = S_t + \alpha x_t + (1-\alpha)S_{t-1} \qquad (3)$$

or in another way of smoothing equation can be written as follows:

Next period forecast value = weight (present period observations)
+ (1- weight) (present period forecast value)

The smoothing equation is based on averaging (smoothing) of past values of a series in decreasing (exponential) manner. The observations are weighted, with the more recent observations are given more weight. The weights are 'α' used for the most recent observation, '$\alpha(1-\alpha)$' for the next most recent, '$\alpha(1-\alpha)^2$' for the next, and so on. At each time for producing a new forecasted value, the weighted observation along with the weighted estimate for the present period is combined.

$$\hat{x}_{t+1} = \hat{S}_t = \alpha x_t + \alpha(1-\alpha)x_{t-1} + \alpha(1-\alpha)^2 x_{t-2} + \cdots + \alpha(1-\alpha)^{t-1} x_1 \quad (4)$$

$$\text{or } \hat{x}_{t+1} = \hat{S}_t = \alpha \sum_{j=0}^{t} (1-\alpha)^j x_{t-j} \quad (5)$$

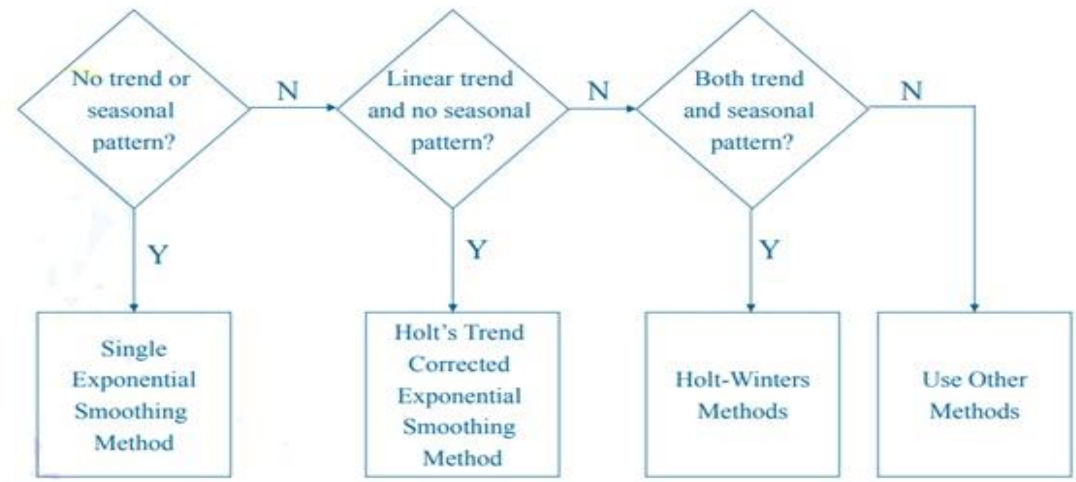

Figure: 1 Selection Procedure for Exponential Smoothing method

Since, the double exponential smoothing can evaluate in linear trends and triple exponential smoothing can handle both trend and seasonal pattern in time series data. Figure 1 shows the selection procedure of different exponential smoothing method is defined.

3.3 ARIMA Model

Autoregressive (AR) model can be efficiently coupled with moving average (MA) model to form a general and useful class of time series models called autoregressive moving average ARMA (p, q) models. However, it can be used only when the time series is stationary. When a time series is studied based on the dependence relationship among the time lagged values of the forecast variable and the past error terms, an autoregressive integrated moving average (ARIMA) model is more appropriate. It can be used when the time series is non-stationary. Applying the ARIMA model step by step procedure is given in the figure 2. The general form of the ARIMA (p, d, q) model is

$$Y_t = c + b_1 Y_{t-1} + b_2 Y_{t-2} \cdots + b_p Y_{t-p} + e_t - \varphi_1 e_{t-1} - \varphi_2 e_{t-2} - \cdots - \varphi_q e_{t-q} \quad (6)$$

where p, d and q represent respectively the order of an autoregressive part, the degree of difference involved in the stationary time series which is usually 0,1 or at most 2 and the order of the moving average part.

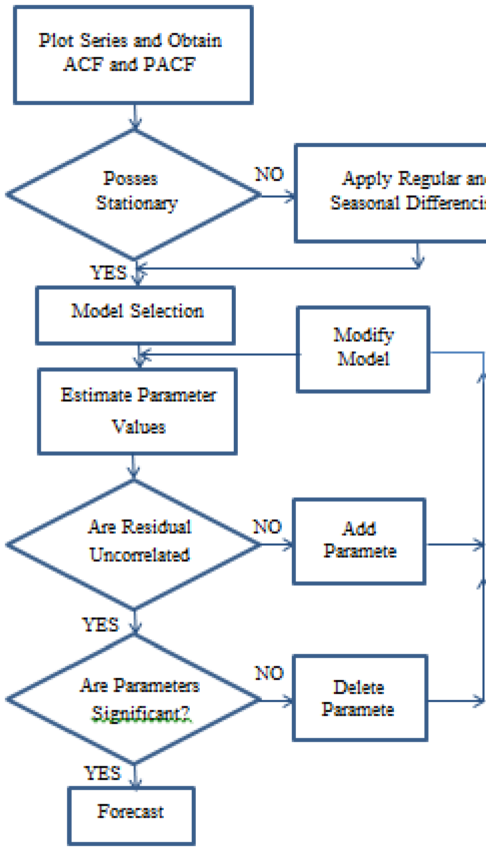

Figure 2: Step by Step Procedure of ARIMA Model

An ARIMA model can be obtained by estimating its parameters. The values of p and q can be determined from the patterns in the plotting of the values of ACF and PACF. The spikes falling above the time axis are used to estimate the value of p. the spikes falling below the time axis are used to estimate the value of q. For an AR (p) model, the spikes of ACF decay exponentially or there is a sine wave pattern and the spikes of PACF are close to zero beyond the time lag q whereas the spikes of PACF decay exponentially or there is a sine wave pattern.

3.3 GARCH Model

The GARCH model is one of the techniques based on the assumption that the random component of the model shows changes in variability. It was developed in a simplified form by Engle (1982) and later generalized by Bollerslev (1986). The model was applied successfully in modeling the changing variability (or volatility) of the variable in time series, with the applications being taken in large measure from the area of financial investments. After identifying an asymmetric relationship between conditional volatility and conditional mean value, the econometrists focused their efforts on the design of methods for the modelling of this phenomenon.

$$\sigma_{t|t-1}^2 = \omega + \beta_1 \sigma_{t-1|t-2}^2 + \cdots + \beta_p \sigma_{t-p|t-p-1}^2 + \alpha_1 r_{t-1}^2 + \alpha_2 r_{t-2}^2 + \cdots + \alpha_q r_{t-q}^2 \qquad (7)$$

In terms of the backshift B notation, then the model can be expressed as

$$(1 - \beta_1 B - \cdots \beta_p B^p)\sigma_{t|1-1}^2 = \omega + (\alpha_1 B + \cdots + \alpha_q B^q) r_t^2 \qquad (8)$$

In some of the literature, the notation GARCH(p,q) is written as GARCH(q,p); that is, the orders are switched. Nelson and Cao (1992) and Tsai and Chan (2006) proposed that conditional variances must be nonnegative, the coefficients in a GARCH model are often constrained to be nonnegative.

4 RESULTS AND DISCUSSIONS

For the selected data has no trend and seasonality does not exist so SES method is adopted. Forecasted values are obtained for α in arange from 0 to 1 for both S&P BSE and Nifty 50. Among them α =0.9 gives the forecasted value is very closer actual values compared with the other α values.

In ARIMA model, both data adopting ARIMA(0,1,0). Generally, ARIMA(0,1,0) model is called random walk model. Actual data values does not possess the stationary conditions so first order difference is calculated and it satisfies the stationary, so d=1 is adopted for both the data.

Then the combination of p, d and q values (without changing d=1) found the different BIC values. Among the BIC values ARIMA (0, 1, 0) having the minimum value (11.026) for S&P BSE and (8.673) for Nifty 50 are shown in the Table 2. Finally, GARCH(1,1) model also adopted.

Table 2: Model Selection for ARIMA Model

ARIMA(p,1,q)	**Normalized BIC Values**	
	S&P BSE	**Nifty 50**
ARIMA(0,1,0)	11.026	8.673
ARIMA(1,1,0)	11.030	8.677
ARIMA(2,1,0)	11.038	8.685
ARIMA(0,1,1)	11.030	8.677
ARIMA(0,1,2)	11.038	8.685
ARIMA(1,1,1)	11.038	8.685
ARIMA(1,1,2)	11.044	8.691
ARIMA(2,1,1)	11.045	8.692
ARIMA(2,1,2)	11.053	8.700

Based on the three models error values are presented in the Table 3 for both S&P BSE and Nifty 50.

Table 3: Forecasted Error Values

Errors	**S&P BSE**			**NIFTY 50**		
	SES	**ARIMA**	**GARCH**	**SES**	**ARIMA**	**GARCH**
MSE	61002	61002	210368	5802	5802	18736
RMSE	246.987	246.972	458.6	76.174	76.171	136.88
MAE	186.536	186.428	343.28	57.158	57.127	102.43
MAPE	0.632	0.632	1.123	0.634	0.634	1.086

From the Figure 2, SES and ARIMA model forecasted values are very closer to the actual values. But, forecasted values of GARCH model are deviated from the actual values. So predicting the S&P BSE stock close value, SES and ARIMA (0,1,0) model is appropriate.

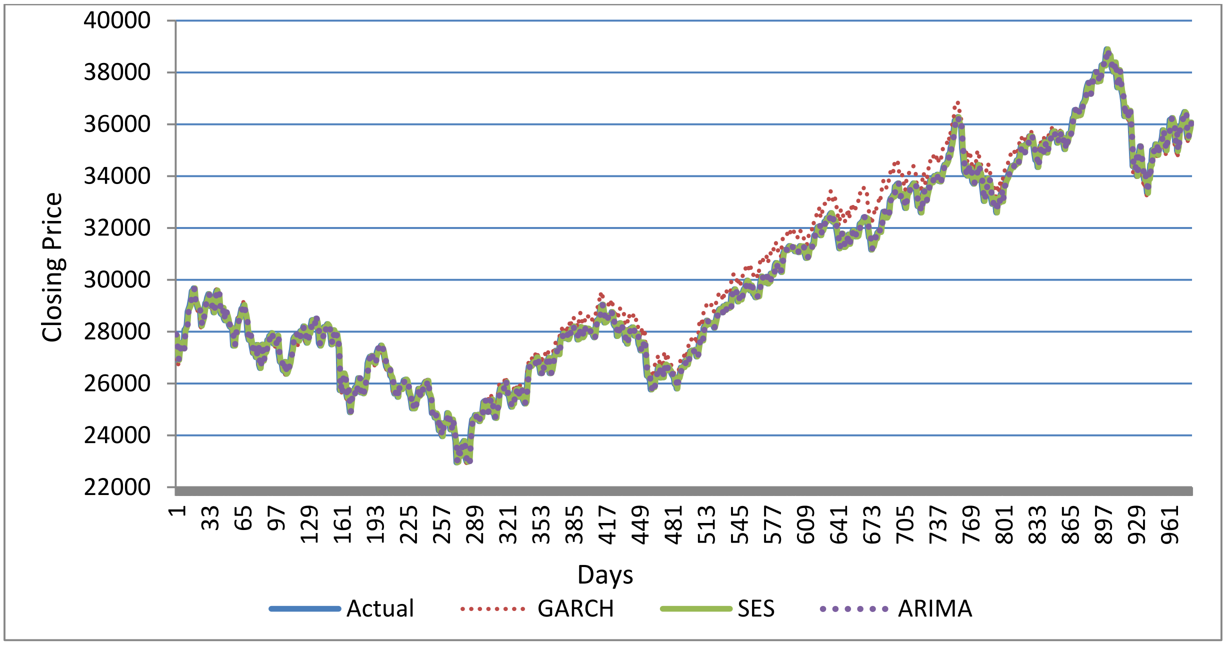

Figure: 3 Comparison of Actual and Forecasted Value of S&P BSE

Similarly, from the Figure 3, the forecasted values of SES and ARIMA models are very closer to the actual values. But forecasted values of GARCH model are deviated from the actual values. So, predicting the Nifty 50 stock close value, SES and ARIMA(0,1,0) model is appropriate.

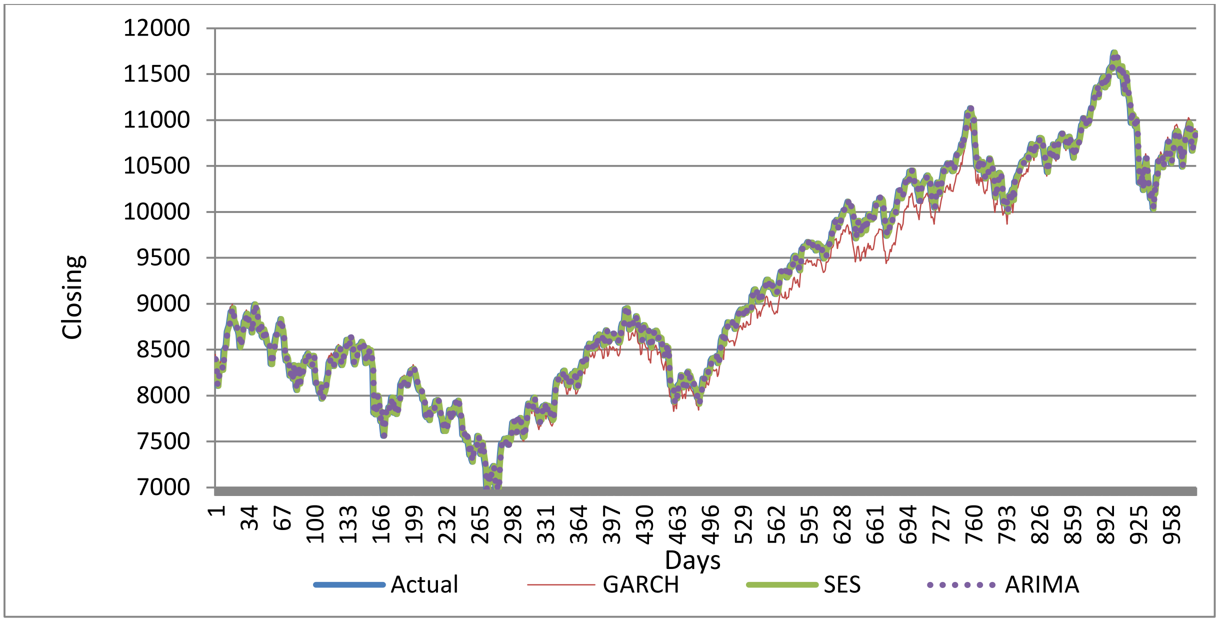

Figure: 4 Comparison of Actual and Forecasted Value of Nifty 50

Figure 4 shows the line chart of actual and forecasted values, which shows that actual values are closely associated with the SES model. The following Table 3 shows, MSE, RMSE, MAD and MAPE error values of three models.

5 CONCLUSIONS

Stock market forecasting is a very tedious job because its price values are non-linear. Over the years, there are several forecasting techniques have tried and used for forecasting. In this paper forecasted S&P BSE and Nifty 50 using SES, ARIMA and GARCH methods. SES is very flexible to use the non-linear models. Because we can adjust the values of α, from which we can reduce the error values. The ARIMA models used when the data are not possess stationary. Normally stock market data are non-stationary, so it is very useful to estimating and validating the model. The GARCH approach is widespread used in where the volatility of returns arises. The error values such as MSE, RMSE, MAE and MAPE are compared for three models. In this case, SES and ARIMA models error values are much closed to each other and also less than the GARCH model. It concludes that SES and ARIMA models are more appropriate for prediction of S&P BSE and Nifty 50.

REFERENCES:

1. Brown, R. G. (1959). Statistical Forecasting for Inventory Control, New York: *McGraw-Hill*.
2. Gardner, E. S. (1985). Exponential Smoothing: the state of the art, *Journal of Forecasting*, 4, 1-28.
3. Hyndman, R. J., Koehler, A. B., Snyder, R. D. and S. Grose. (2002). A state space framework for automatic forecasting using exponential smoothing methods, *International Journal of Forecasting*, 18, 439-454.
4. Kim, J. H. (2003), Forecasting Autoregressive time series with bias corrected parameter estimators, *International Journal of Forecasting*, 19, 493 – 502.
5. Landsman, W. R. and A. Damodaran. (1989). A comparison of quarterly earnings per share forecasts using James-Stein and unconditional least squares parameter estimators, *International Journal of Forecasting*, 5-4, 491-500.
6. Taylor, J. W.(2003). Exponential Smoothing with a Damped Multiplicative Trend, *International Journal of Forecasting*, 19, 715 – 725.
7. Blair, B., Poon, S., and S.J. Taylor (2002). Asymmetric and Crash Effects in Stock Volatility for the S&P 100 Index and Its Constituents, *Applied Financial Economics*, 12 (5), 319-329.
8. Deb, S.S., Vuyyuri, S., and B. Roy (2003). Modeling Stock Market Volatility in India: A Comparison of Univariate Deterministic Models, *The ICFAI Journal of Applied Finance*, 9(7), 19-33.
9. Karmakar, M. (2005). Modeling conditional volatility of the Indian stock markets, *Vikalpa*, 30, 21-37.
10. Magnus, F.J., and O.A.E. Fosu (2006). Modeling and Forecasting Volatility of Returns on the Ghana Stock Exchange Using GARCH Models, *American Journal of Applied Sciences*, 3, 2042-2048.
11. Liu, W., and B. Morley (2009). Volatility Forecasting in the Hang Seng Index using the GARCH Approach, *Asia-Pacific Financial Markets*, 16, 51-63.