

Time Series Analysis Of Nifty Bank Index - A Measure Of Banking Industry

Aastha Khera* Prinkle**
Research Scholars
Department of Commerce
Kurukshetra University, Kurukshetra

ABSTRACT

This paper would employ Box Jenkins Methodology (ARIMA) for modeling and forecasting Bank Nifty. Bank Nifty index was launched in 2003 in order to capture movements of banking industry. Time series data of Bank Nifty is taken from the website of National Stock exchange. The data was first made stationary in order to arrive at the conclusions. An appropriate model is selected on the basis of evaluation and diagnosis of data. The selection criteria for finalizing the model is AIC and SC and the final model came out to be (2,1,0). Residuals of the series were found white noise in nature. Model explained that series could be forecasted on the basis of auto regression of two lags.

Key words: Box Jenkins, ARIMA, Bank Nifty, AIC, forecasted

INTRODUCTION

Time series analysis is an imperative component of statistics. It is compilation of observations obtained at equal and distinct intervals. It has varied applications in diversified fields. Time series and forecasting models are being applied in various fields ranging from commerce, finance to healthcare and tourism. Traditional definition explains time series as data observed in sequential order and in the modern times, time series is defined as the realization of stochastic process where the events happen in an uncertain manner. It is a vital tool for analyzing characteristics for creating future adjustments. It consist of trend, seasonal, cyclical and irregular components. Historical patterns are explained and analyzed in order to predict the movements. Future values are predicted on the basis of past values. It is playing an important role in finalizing policies to be used in taking economic and financial decisions. Forecasting stock returns has always been a exigent task for researchers.(Wang et al. 2012). Stock prices are affected by innumerable factors. Still, Researchers are delving deep into this

area and coming up with finest models for forecasting. Outcome of this hard work is impressive and attractive.

An appropriate modeling technique is identified for forecasting Bank Nifty index. Stock price prediction is useful for all types of investors. There are two types of model for forecasting the series- Statistical and artificial intelligence model. Hybrid models have also been developed to capture the characteristics in the series. ARIMA is a type of statistical model. It is an appropriate model regarding the predictions to be made for future. It was proffered by Box and Jenkins in the year 1970, thus known as Box-Jenkins Methodology. This acronym stands for Auto regressive Integrated Moving averages. Previous values are examined and scrutinized in this model.

The importance of stock markets is undeniable. Stocks markets are the coordinators for organized trading of stock and shares by its buyers and sellers. It comprises of Primary market and Secondary market. Primary market is the one where the securities are being issued for the first time. It provides liquidity in the market by allowing the companies to raise the capital from the market and facilitating the investors to trade in secured manner. (Ho and Iyke, 2017). Secondary market facilitates trading for existing securities. Stock exchanges are the mediums through which exchange of securities becomes possible. Stock market development is influenced by both macroeconomic and as well institutional factors. Macroeconomic factors comprise of banking sector development, inflation rates, exchange rates and capital flows. Stock market indexes are the barometers of the economy stating the economic condition of nation's growth.

In India, NSE and BSE are the two important stock exchanges operating in India. Sensex and Nifty are the two important indices of our country. Stock markets are becoming advanced and technical day by day, introducing efficient changes. Various sectoral and thematic indices were introduced to cover the specific conditions of the sector. One of the important index is Bank Nifty that captures the performance of Indian banking sector.

Banking sector has always been indispensable part of economic system converting idle resources into productive usage. Its importance could be correlated as lifeblood of financial activities. Banking stock returns are significantly related to future economic growth. (Cole, 2008). Indian banking sector was made profitable and efficient after financial liberalization of 1991. Considering the importance of banking sector, it has become a attractive option for investors. Investors track the performance of banking sector through this Bank Nifty Index. Bank Nifty index was launched in 2003. This index consists of the most liquid and large

banking stocks. It facilitates the passage of information to investors and intermediaries regarding the performance of Indian banking industry. Investors could track the banking industry by including Bank nifty index in their portfolio. 12 Banking companies comprise this index. It is based on free-float market capitalization method.

This paper employed ARIMA technique in forecasting this index for the period of 1 year. E-views software is used for calculation purposes.

REVIEW OF LITERATURE

In this section, efforts have been done to collect the findings of previous studies. A large number of empirical studies have been carried out for forecasting of various variables i.e. stock prices, GDP, agricultural prices. Application of ARIMA is not confined only to stock market; it has varied applications in diversified areas.

Shakira (2011) used Box-Jenkins approach (ARIMA models) for projecting time series using varied sampling intervals. Two intervals were taken one related to 1st day of the trading month and 2nd is the 15th day of the trading month. Varied industries were taken into consideration and it was noticed that individual stocks are not relied on the industry they are involved in. The outlook of the behavior is showing individual characteristics. Residual plots were analyzed to measure the accuracy of forecasts. Mean average percentage error (MAPE) was studied to decide the duration of timing interval. Banerjee (2014) applied ARIMA modeling in forecasting Sensex data for the year 2013. In order to foresee the future, one must analyze the present and scrutinize it. Monthly stock index data was collected from NSE website from 2007 to 2012. Model selected for forecasting is (1,0,1) and forecasting evaluation criteria's used are Root mean squared error (RMSE), Mean absolute percentage error(MAPE), Mean absolute error(MAE), Ljung-Box statistic.

Jarret and Kyper(2011) employed ARIMA to model stock price Index in China. Stock index prices of Shanghai index were forecasted. ARIMA intervention technique was used for calculating the effect of world financial crisis in Chinese stock markets. A log likelihood ratio was employed for comparing the two models. It was found that crisis got negative impact resulting in decline of stock prices. Manufacturing sector was badly impacted due to financial crisis.

Angad and Kulkarni(2015) used this data mining technique in R- software. Data was derived from google finance and with the help of Quant-mod package series was forecasted. Intra-day INFY data was taken from 2007 to 2015. Series was forecasted. ARIMA (0, 1,0) was used for the prediction of the series.

Amin (2000) conducted study to predict monthly prices of potato crop in Bangladesh. Seasonal differencing was undertaken in order to eliminate seasonality component and make data ready for computations. Appropriate model selected was of (1, 1, and 0). Single equation ARIMA model for Y_t was constructed. Prices were forecasted after completing the required examination of residuals.

Dwivedi et. al. (2017) forecasted mean temperature of Junargah district for the year 2016 with help of Seasonal ARIMA (SARIMA). These models eliminate the seasonality component in the series. Time period analysed was from 1984 to 2015 and series was found stationary at level. Seasonal differencing was undertaken and appropriate model was (1,0,1) (1,1,1) (12). Residuals were duly checked using serial correlation test, histograms Q-Q plots.

Bakar & Rosbi(2017) forecasted Bitcoin exchange rate using ARIMA modeling. Parameters are finalized on the basis of auto-correlation and Partial auto-correlation functions. Bitcoin is differenced to attain the stationary. Finalized model is (2,1,2) . Value of R- squared comes out to be 44.44%. Past historical data and error terms have the power to explain the variation by 44%. MAPE error analysis came out to be 5.36%.

Ahmed et. al (2017) predicted KIBOR(Karachi inter-bank offered rate) with help of Box-Jenkins methodology. This study has utilized 6-moths average interest rate for the time period of four years from 2012 to 2016. The results attained level of significance at 1%. The model was found useful and fit for forecasting with the confirmation of AIC criteria. ARIMA (1,1,1) model was selected for predicting. Error statistical tools further enhanced the reliability of the model.

Nwaigwe et al. (2018) made an attempt to analyze the daily exchange rate of Nigerian Naira for US Dollar in terms of selling rate, buying rate and central rate. Best model was found to be ARIMA (1,1,0) for all the 3 series after converting the nature of the series. No presence of serial correlation was found in the residuals series. During the diagnostic checking, it was found that model computing the selling rate has lowest error measures. Future prediction involved depreciation of naira.

RESEARCH METHODOLOGY

One could witness five approaches for forecasting time series patterns. These are single-equation regression models, simultaneous-equation regression models; exponential smoothing methods autoregressive integrated moving average models (ARIMA), and vector auto-regression (VAR).

Auto regressive integrated moving average model presents a parsimonious description of stationary series. It is standard statistical model for prediction and forecasting in time series analysis. ARIMA is getting more popular and its application is increasing in various fields. It explains the variations for data by analyzing its stochastic properties.

It was introduced by Box and Jenkins in 1970 better known as Box-Jenkins approach. It comprises of AR, I and MA. Here AR stands for the Autoregressive model, I represents the Integration indicating the order of single integer, and MA stands for the Moving Average model. If the series is non-stationary, then first it is converted to stationary for further analysis. ARIMA is the extension of ARMA model. The variable ie. Y_t is described by its own past or lagged values and error terms involved in it. It does not consider any external factor that could impact the variable.

ARIMA process consist of identifying the parameters (p,d,q). p refers to the number of autoregressive terms, d refers to the order of differentiation. q refers to the number of moving averages. Initially autocorrelation and partial autocorrelation is studied between the items of the series for the identification of p,d and q. If d becomes 0 then it becomes an ARMA model.

Auto regressive(p)

It means associating the values with one self. This term is connected with regressing the values with themselves. AR (p), here p denotes the number of times value is being linked with oneself.

$$Y_t = \hat{a} + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + e_t$$

ϕ_1, ϕ_2 and ϕ_p are the model parameters and e_t is the error terms.

Moving Average (q)

It is used for identifying error terms involved in the model. Here, the term q refers to the preceding number of error terms intricate in the model.

$$Y_t = e_t - \Theta_1 e_{t-1} - \Theta_2 e_{t-2} - \dots - \Theta_q e_{t-q}$$

$\Theta_1, \Theta_2, \dots, \Theta_q$ are the parameters of error terms.

ARIMA Model

It is fashioned as a blend of past values and error terms which is depicted in this equation

$$Y_t = \hat{a} + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \phi_3 Y_{t-3} + \dots + \phi_p Y_{t-p} + e_t - \Theta_1 e_{t-1} - \Theta_2 e_{t-2} - \Theta_3 e_{t-3} - \dots - \Theta_q e_{t-q}$$

where Y_t are the series values and e_t are the error terms ; $\phi_i (i=1,2,\dots,p)$ and $\Theta_j (j=1,2,\dots,q)$ are the model parameters. p and q are integers denoting orders of autoregressive and moving average polynomials.

Detailed Steps of ARIMA Modeling are

1. Recognizing the stationarity of the series. It is analyzed on the basis of line plots, correlograms showing autocorrelation and partial autocorrelation. The ADF test is commonly used in capturing this property.
2. If the series is non stationary then the next step is to identify the level of differention required. It is represented by d in this technique. Adequate transformation is done for this step ie taking log, differences or square root. If series is stationary then direct proceeding to the 3rd step will be done.
3. Autocorrelation coefficient (ACF) and Partial autocorrelation (PACF) are generated through Correlograms. Appropriate autocorrelation order p and moving average of order q are identified. Parameter estimation is done in this step.
4. Diagnostic checking of series is conducted in this step. Residuals are checked whether they are white noise or not. Correlograms of finalized models is investigated. A flat model is the best fit model.
5. Forecasting of the series is conducted. E-views provides static and dynamic forecasts for the series. With the help of the forecast menu, prediction of series is done. Comparison graphs of actual and forecasted value are made.

ANALYSIS AND DISCUSSIONS

Closing prices of Bank Nifty index is taken from the website of NSE from the time period 2008 to 2019. Table 1 shows the descriptive statistics of this index. Fig1 is depicting the graphical representation of the index.

Table 1
Descriptive Statistics regarding Bank Nifty Index

Mean	13405.50
Median	11552.30
Maximum	27379.45
Minimum	3892.400
Std. Dev.	5689.416
Skewness	0.526100
Kurtosis	2.452940
Jarque-Bera	7.031999
Probability	0.029718
Sum	1608660.
Sum Sq. Dev.	3.85E+09
Observations	120

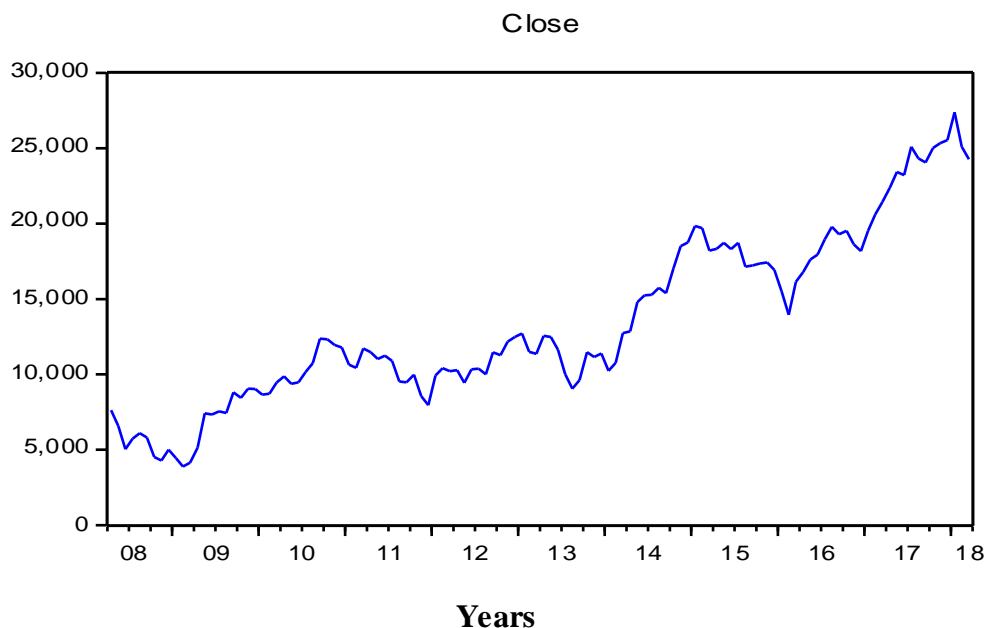


Fig 1: Bank nifty data during 2008 to 2018

**Table2.
Augmented Dickey-Fuller unit root test on Closing prices**

Null Hypothesis: CLOSE has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on AIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-0.143056	0.9412
Test critical values:		
1% level	-3.486064	
5% level	-2.885863	
10% level	-2.579818	

It can be seen that ADF -0.143056 is lower than the critical values at 1%,5% and 10% level of significance in absolute terms thus stating that series is non-stationary. In order to make it stationary, firstly log of closing prices is taken. This process was unable to fill the requirement of stationarity. So differentiation is done of the log of closing prices. ADF test is applied on Dclose (Dclose is series name specifying difference of log of closing prices.) and results are being shown in third table.

Table3
Augmented Dickey-Fuller unit root test on DCIose

Null Hypothesis: DCLOSE has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=12)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-10.09797	0.0000
Test critical values:		
1% level	-3.481217	
5% level	-2.883753	
10% level	-2.578694	

It is clearly seen that t-statistic is greater than the critical values in absolute terms thus rejecting Null hypothesis. This is said that series after transformation is stationary series. This could be verified through the graphical representation of differencing of log closing prices.

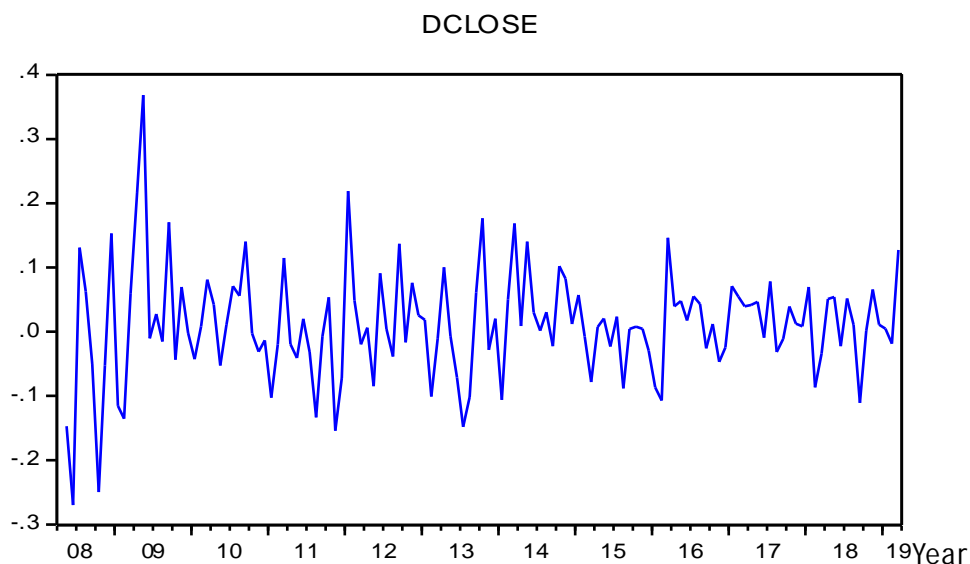


Fig2: Differencing of log of Bank nifty Closing data during 2008 to 2018

For the identification of models, Correlogram of Dclose is generated for autocorrelation and partial functions.

Table 4
Auto Correlation and Partial Auto Correlation function graphs of the Dclose series.

Date: 12/11/19 Time: 12:11
Sample: 2008M04 2019M03
Included observations: 131

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.117	0.117	1.8433	0.175
		2	-0.222	-0.239	8.4797	0.014
		3	-0.082	-0.023	9.3868	0.025
		4	0.105	0.071	10.886	0.028
		5	0.163	0.124	14.547	0.012
		6	-0.042	-0.048	14.792	0.022
		7	-0.111	-0.035	16.518	0.021
		8	-0.016	-0.010	16.553	0.035
		9	-0.005	-0.061	16.556	0.056
		10	0.003	-0.012	16.557	0.085
		11	-0.114	-0.115	18.435	0.072
		12	-0.096	-0.059	19.781	0.071
		13	-0.110	-0.151	21.572	0.062
		14	-0.040	-0.050	21.809	0.083
		15	0.024	-0.015	21.892	0.111
		16	0.048	0.056	22.244	0.136
		17	-0.145	-0.153	25.447	0.085
		18	0.012	0.097	25.467	0.113
		19	0.009	-0.084	25.479	0.145
		20	-0.043	-0.073	25.763	0.174
		21	0.003	-0.008	25.765	0.216
		22	-0.041	-0.073	26.032	0.250
		23	-0.071	-0.125	26.846	0.263
		24	0.056	0.020	27.354	0.288
		25	0.076	0.017	28.305	0.294
		26	0.033	-0.017	28.491	0.335
		27	-0.096	-0.073	30.037	0.313
		28	-0.054	-0.057	30.522	0.339
		29	0.008	-0.059	30.531	0.388
		30	-0.016	-0.129	30.574	0.437
		31	0.062	0.052	31.246	0.454
		32	0.042	-0.007	31.553	0.489
		33	0.030	0.040	31.714	0.531
		34	0.045	-0.020	32.081	0.562
		35	-0.080	-0.075	33.238	0.553
		36	-0.009	-0.045	33.253	0.600

Both autocorrelation and partial correlations functions are significant at lag 2. So both p and q could be taken as 2. However they are also little significant at lag 1. So tentative models have to be investigated and appropriate model needs to be identified as different orders could be derived from the correlograms. Table 4 presents the test results for different parameters. Criteria that have been selected for finalizing the model are Adjusted R-squared, AIC value, SC value and S.E. of regression. AIC and SC criteria are used for ranking of the models. Larger the R- squared better is the model along with lowest AIC, SC and SE of regression.

Table 5
Test results of ARMA (p,q).

(p,q)	Adjusted R-squared	AIC	SC	S.E. of regression
(0,1)	0.011288	-1.972897	-1.902835	0.089098
(0,2)	0.037161	-1.990729	-1.897313	0.087925
(1,0)	-0.000111	-1.961656	-1.891594	0.089611
(1,1)	0.018082	-1.971448	-1.878032	0.088792
(1,2)	0.030418	-1.975672	-1.858902	0.088232
(2,0)	0.048857	-2.002653	-1.909237	0.087389
(2,1)	0.041148	-1.986491	-1.869721	0.087743
(2,2)	0.059062	-1.995676	-1.855553	0.086919

Usually the model with lowest AIC and SC values are selected. This condition would not suffice as residual randomness test has to be applied on the models. An appropriate model tends to have flat correlogram and residuals follow white noise distribution. if this condition is not fulfilled then next model has to be picked up with the second lowest AIC and SC. So after due calculations, ARMA (2,0) model was identified. So here is the equation output of the model.

Dependent Variable: DCLOSE
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 12/11/19 Time: 12:17
Sample: 2008M05 2019M03
Included observations: 131
Convergence achieved after 14 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.011146	0.006708	1.661474	0.0991
AR(1)	0.141869	0.086159	1.646585	0.1021
AR(2)	-0.259905	0.081217	-3.200135	0.0017
SIGMASQ	0.006961	0.000761	9.143137	0.0000
R-squared	0.075345	Mean dependent var		0.010560
Adjusted R-squared	0.053502	S.D. dependent var		0.087099
S.E. of regression	0.084737	Akaike info criterion		-2.067316
Sumsquared resid	0.911896	Schwarz criterion		-1.979523
Log likelihood	139.4092	Hannan-Quinn criter.		-2.031642
F-statistic	3.449486	Durbin-Watson stat		1.960162
Prob(F-statistic)	0.018675			
Inverted AR Roots	.07+.50i	.07-.50i		

Final model selected is (2,1,0) and here is the finalized equation for this model

$$Y_t = 0.011146 + 0.141869 Y_{t-1} - 0.259905 Y_{t-2}$$

P value of AR (2) is significant. As p-value is less than 0.05%.

This model is used to fit the data and the results are presented in figure 5. In this figure actual data is being shown by red line and fitted and residual lines are being shown by green and blue lines respectively.

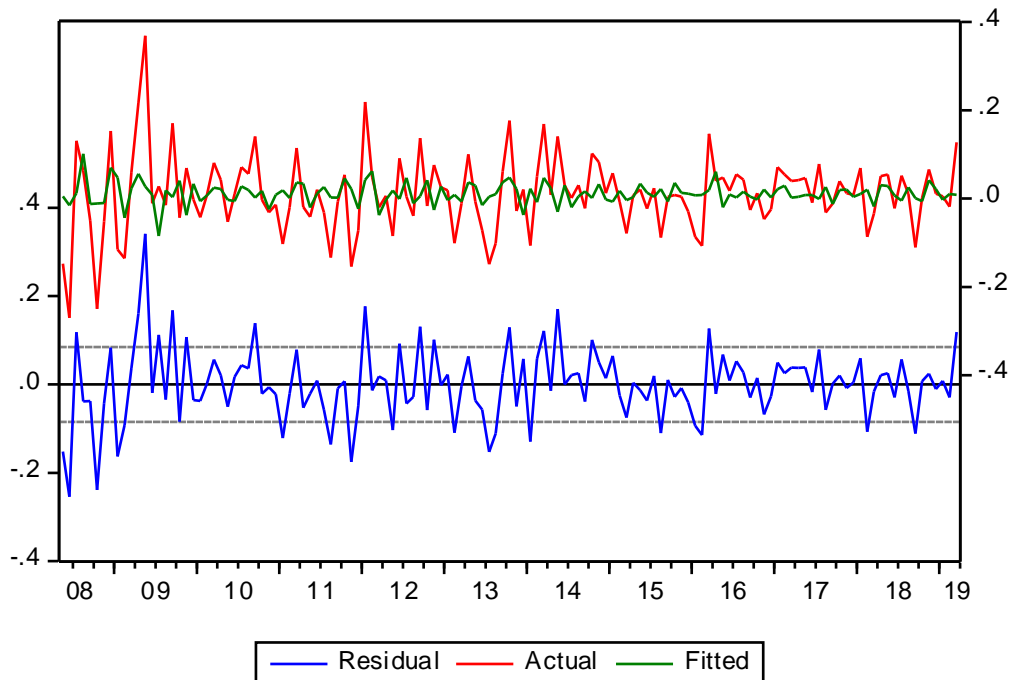


Figure: Actual series, fitted series and residual series of the Dclose sequence

After applying the (2,1,0) model, White noise test was conducted on the residuals. Autocorrelations and partial autocorrelations functions of residual series are depicted in figure. It is seen that residuals are white noise in nature as p-values are greater than 5%. Hence it supports the validity of the model.

Date: 12/11/19 Time: 12:18
Sample: 2008M04 2019M03
Included observations: 131
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.001	-0.001	4.E-05	
		2	0.039	0.039	0.2028	
		3	0.006	0.006	0.2078	0.649
		4	0.031	0.030	0.3411	0.843
		5	0.124	0.124	2.4777	0.479
		6	-0.035	-0.037	2.6478	0.618
		7	-0.078	-0.090	3.5076	0.622
		8	-0.006	-0.005	3.5122	0.742
		9	-0.065	-0.067	4.1092	0.767
		10	-0.007	-0.020	4.1162	0.846
		11	-0.144	-0.127	7.1100	0.626
		12	-0.078	-0.061	7.9998	0.629
		13	-0.129	-0.127	10.472	0.489
		14	-0.045	-0.037	10.776	0.548
		15	-0.040	-0.030	11.015	0.610
		16	0.071	0.104	11.779	0.624
		17	-0.173	-0.167	16.364	0.358
		18	0.043	0.043	16.647	0.409
		19	-0.025	-0.031	16.741	0.472
		20	-0.062	-0.109	17.352	0.499
		21	0.006	-0.033	17.359	0.566
		22	-0.034	-0.029	17.541	0.618
		23	-0.063	-0.116	18.177	0.638
		24	0.062	0.003	18.805	0.657
		25	0.025	0.033	18.910	0.706
		26	0.044	-0.019	19.229	0.740
		27	-0.084	-0.093	20.416	0.725
		28	-0.042	-0.091	20.720	0.756
		29	0.012	-0.017	20.743	0.798
		30	-0.033	-0.118	20.929	0.828
		31	0.077	0.049	21.950	0.822
		32	0.040	0.024	22.237	0.845
		33	0.014	0.011	22.273	0.874
		34	0.071	-0.027	23.172	0.873
		35	-0.078	-0.079	24.271	0.865
		36	0.018	-0.045	24.332	0.889

Figure: Autocorrelation and partial autocorrelation function graphs of the residual series.

Last step comes for forecasting the series. Forecast button in equation tab is used for predicting the series. Predicted value comes under Standard errors in the graph as shown in figure seven. This gives the incomplete picture so comparative plot of actual and forecasted data is created.

It is being shown in figure eight.

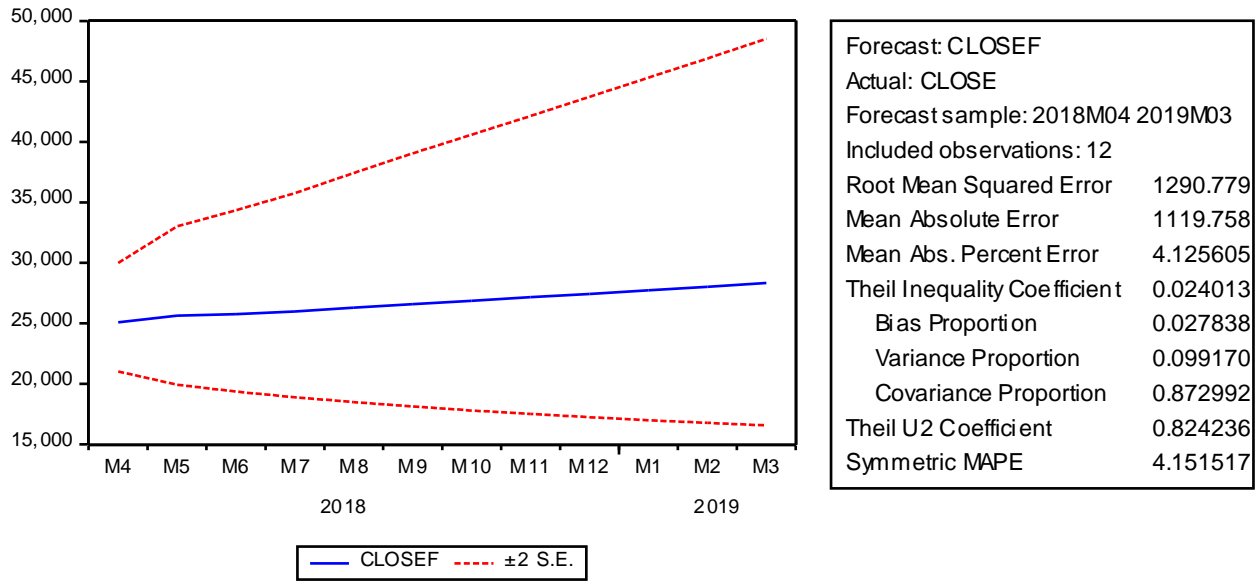


Figure : Forecasted graph from April 2018 to March 2019

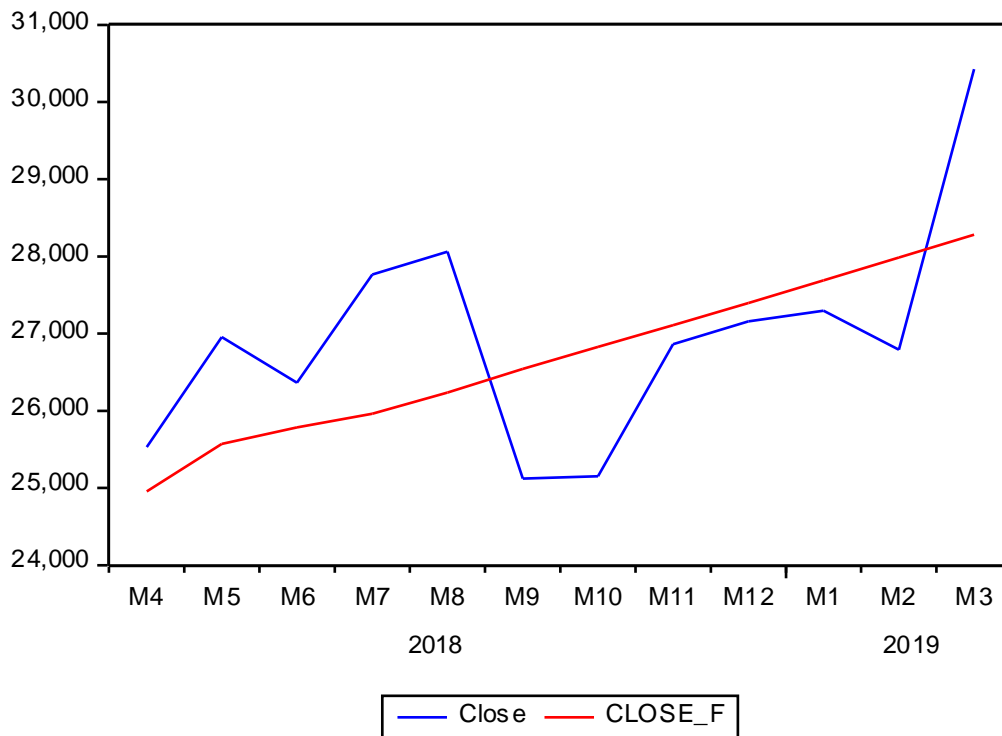


Figure : Comparison of actual and forecasted data

CONCLUSIONS

An endeavor was made for forecasting the stock prices of Bank Nifty index with the help of forecasting model that is based on historical time series data. ARIMA modeling is getting advanced as the time series is becoming more dynamic. This paper got success in forecasting for short term duration. For long time period, variations could occur as there are other different factors that could impact realistically to the series. It must be kept in mind as this Bank nifty series was dynamic in nature and it could be impacted by innumerable factor other than its own past values and error terms.

REFERENCES

- ✚ Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- ✚ Bakar, N. A., & Rosbi, S. (2017). Autoregressive integrated moving average (ARIMA) model for forecasting cryptocurrency exchange rate in high volatility environment: A new insight of bitcoin transaction. *International Journal of Advanced Engineering Research and Science*, 4(11).
- ✚ DWIVEDI, D., SHARMA, G., & WANDRE, S. FORECASTING MEAN TEMPERATURE USING SARIMA MODEL FOR JUNAGADH CITY OF GUJARAT. *International Journal of Agricultural Science and Research (IJASR)*, 7(4), 183-194
- ✚ Ahmed, R., Vveinhardt, J., Ahmad, N., & Streimikiene, D. (2017). Karachi inter-bank offered rate (KIBOR) forecasting: Box-Jenkins (ARIMA) testing approach.
- ✚ Nwaigwe, C. C., Atuzie, G. U., Obinwanne, I. C., & Ogbonna, C. J. (2018) A TIME SERIES ANALYSIS OF DAILY EXCHANGE RATE OF US DOLLAR TO NAIRA FROM 2016-2017 (RECESSION PERIOD). *International Journal of Scientific & Engineering Research* 9(7),
- ✚ Banerjee, D. (2014, January). Forecasting of Indian stock market using time-series ARIMA model. In *2014 2nd International Conference on Business and Information Management (ICBIM)* (pp. 131-135). IEEE.
- ✚ Jarrett, J. E., & Kyper, E. (2011). ARIMA modeling with intervention to forecast and analyze chinese stock prices. *International Journal of Engineering Business Management*, 3(3), 53-58.
- ✚ Green, S. (2011). Time series analysis of stock prices using the box-jenkins approach.

- ✚ Amin, R., & Razzaque, M. A. (2000). Autoregressive integrated moving average (ARIMA) modeling for monthly potato prices in Bangladesh. *Journal of Financial Management & Analysis*, 13(1), 74.
- ✚ Amin, R., & Razzaque, M. A. (2000). Autoregressive integrated moving average (ARIMA) modeling for monthly potato prices in Bangladesh. *Journal of Financial Management & Analysis*, 13(1), 74.