Forecasting Of Time Series Data Using Hybrid ARIMA Model With The Wavelet Transform.

Jigneshkumar R. Valvi[1], Prashant K. Shah[2].

[1]*M.Tech. Student, Department of Electronics Engineering, SVNIT Surat, Gujarat, India*
[2]*Associate Professor, Department of Electronics Engineering, SVNIT Surat, Gujarat, India*

---***---

Abstract - *To do accurate forecasting of any process, we should have to understand the process and according to that, we can develop the model. This can help to decide the best possible values of forecast. To do the best forecast, we can use past values as well as present values. We can divide available data information into data points with respect to appropriate time, that data is called as time series data. For predicting time series data various models are available that can help to forecast, i.e Stochastic, Neural networks and SVM based models. All these models have their own merits and demerits. In this paper, we will talk about stochastic model (ARIMA) modeling with help of wavelet transform decomposition; the designed hybrid model is applied in the application related to finance. For that we have taken previous six months data (daily basis) of S&P BSE information technology, based on this available data we have predicted the 10 days step ahead future values, and comparative analysis has been done between ordinary model and proposed models.*

***Key Words*: ARIMA Model, Wavelet transform, Time series data analysis.**

1. INTRODUCTION

The finance is the field that deals study of investments; it is also called as management of money. A stock exchange is one subset of finance, where stockbrokers and traders can buy and sell their shares, in the stock exchange where and when to invest is a most difficult task, because proper investment may bring a huge profit, the same chances of losses, so prediction of accurate future value is more necessary.

There are various non-linear and linear models are available in stochastic modeling where, ARCH (autoregressive conditional heteroskedasticity) GARCH (generalized autoregressive conditional heteroskedasticity), are non-linear models and AR (Autoregressive), MA (Moving Average) ARMA (Autoregressive Moving Average), ARIMA (Autoregressive Integrated Moving Average), are examples of some linear models [1]. As discussed above forecasting models can be divided into three categories Stochastic, ANN (Artificial neural networks), SVM (Support vector machine) based. Among all these models stochastic models have been used more frequently due to their best forecasting capability and modelings steps [2]. They help to analysis various types of time series data. The ARIMA model is used in this paper due to its fabulous performance for short-term periods compare to others models [2]. To use these kinds of linear regression models we have to stationary test. Only stationary data can be used for the forecast, and if data is not stationary, we should have to make stationary by differencing it. In ARIMA 'I' suggest integrated which is same as differencing. In stationary models the statistical properties mean and variance remains constant while in non-stationary, these properties are time-dependent. Linear models are the models which focus on conditional means while non-linear models focus on conditional variance.

In the literature, these types of models have been used not only for economics but it applies for various purposes, Some of their applications i.e. [3] has used regression forecasting models to improve the location information, which helps the user to get accurate information provided by mobile networks or wireless networks. [4] has used ARMA model for biomedical signal applications. [5] has used these models for network traffic modeling. [6]-[7] has used ARMA model for application to audio coding.[8] has used ARMA model to avoid disturbances of non-Gaussian signals from the non-stationary vibration signals. In [9] classification of a various signal has done by using different stochastic models at different scale, ARMA model models the signal at a different scale, and the features at coarse scales are extracted from the model without performing expensive filtering operation, which increases the accuracy of radar signal classification by exploiting features at different scales.

We have used hybrid ARIMA model with wavelet decomposition technique in this paper. Wavelet decomposition which has been used in many research area such as image processing, signal processing etc. Usually, all the time series data are not deterministic in nature; we assume in many cases the series to be stationary time series. One way to model any time series to consider it as a deterministic function plus white noise. Denoising is a method which helps to minimize unwanted white noise in the time series data, where some mathematical models such as Fourier transform and wavelet transform can be applied. As discussed in [10] wavelet transform is the best method for decomposing any signal (data) in time series. The wavelet transform decomposes the data in terms of both time and frequency, which helps to do accurate and detailed filtering. There are many more methods available to do prediction, filtering, and smoothing, among them state space models are also a good area to do modeling due to Kamal filtering types approaches; these models are more accurate and trustable. Later, which can be modeled as ARIMA, ARMA, etc. in state space domain and various state space domain features can be applied [11].

In this paper, the wavelet decomposing methods with the hybrid ARIMA model. The data which we have used are from S&P BSE information technology, data belongs to last six months closed data (daily). Data belong to range from 01-Dec-2017 to 31-May-2018 and analysis has been done for that data. In between some days are missing to close, total 124 days analysis has been done; and the best model has been designed to forecast ten days ahead of value, with 95% surety range.

2. WAVELET TRANSFORM

A wavelet-transform is transform which provides the time-frequency representation.it was developed as an alternative to the STFT (Short time Fourier Transform) [12]. We pass the time-domain signal from the various high pass and low pass filters, which filters out either high frequency or low frequency portions of the signal. This procedure is repeated, every time some portion of the signal corresponding to some frequencies being removed from the signal. Suppose we have a signal with 500 Hz frequency, in the first stage we split up the signal into two parts by passing a signal from the high pass and low pass filter. (Filters should satisfy some certain conditions, so-called admissibility condition) which results in two different versions of the same signal, portion of the signal corresponding to 0-250 Hz (low pass portion), and 250-500 Hz (high pass portion). Then, we take either portion (usually low pass portion) or both and do the same thing again. This operation is called decomposition. Wavelet transform gives higher frequencies are better resolved in time, and lower frequencies are better resolved in frequency. This means that a certain high-frequency component can be located better in time (with less relative error) than a low-frequency component. On the contrary, a low-frequency component can be located better in frequency compared to high-frequency component. The wavelet transform was used to adopt a wavelet prototype function (mother wavelet), which utilized the basic function $\psi \epsilon L_2(\mathbb{R})$ that is stretched and shifted to capture features that are local in time and local in frequency. The above function satisfied following property:

(I)
$$C_\psi = \int_0^\infty \frac{|\hat{\psi}(w)|}{w}\, dw < \infty$$

Where $\hat{\psi}$ is the Fourier transform of ψ this condition is called admissibility conditions, ensures $\hat{\psi}(w)$ goes to zero quickly as $w \to 0$ In fact to guarantee that C_ψ<∞ <∞, it is necessary $\hat{\psi}(w) = 0$ which is equivalent to

$$\int_{-\infty}^{\infty} \psi(t)dt = 0 \qquad (2.1)$$

(II) Wavelet function must have unit energy. That is

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1 \qquad (2.3)$$

Above both equations (2.1) and (2.2) imply that at least some coefficient of wavelet transform must be different from zero and these departures from zero must cancel out. By consolidating a few mixes of moving and extending of the mother wavelet, the wavelet change can catch all the data in the time arrangement and connect it with particular specific horizon time and location in time. [15] Wavelet can be defined as pairs of filters here we will define father wavelet and mother wavelet, where father wavelet works as low pass filter and mother wavelet as high pass filter as discussed above. father and mother wavelet can be defined as below.

$$\varphi_{J,k}(t) = 2^{\frac{J}{2}}\varphi(2^J t - k); k\epsilon\mathbb{Z}, J\epsilon\mathbb{Z}^+$$

$$\int_{-\infty}^{\infty} \varphi(t) = 1$$

And

$$\psi_{j,k}(t) = 2^{\frac{j}{2}}\psi(2^j t - k); j = 1,2,3 \dots J$$

$$\int_{-\infty}^{\infty} \psi(t) = 0$$

Where $\psi_{j,k}$ is mother wavelet and $\varphi_{J,k}$ is father wavelet, by given this family of basis function, we can define a sequence of coefficients.

$$c_{J,k} = \int_{-\infty}^{\infty} f(t)\varphi_{J,k}(\mathrm{t})\mathrm{dt}$$

And

$$d_{j,k} = \int_{-\infty}^{\infty} f(t)\psi_{J,k}(\mathrm{t})\mathrm{dt}; j = 1,2,3 \dots J$$

Where $c_{J,k}$ is the coefficient of father wavelet (smooth coefficients), and $d_{j,k}$ are coefficient for mother wavelet (detail coefficient).

So from the coefficients, the function $f(.)$ can be represented by

$$f(t) = \sum_{k\in\mathbb{Z}} c_{J,k}\,\varphi_{J,k}(t) + \sum_{k\in\mathbb{Z}} d_{J,k}\psi_{J,k}(t) + \cdots + \sum_{k\in\mathbb{Z}} d_{1,k}\psi_{1,k}(t)$$

of $f(t)$ can be represented as

$$f(t) = C_J + D_J + D_{J-1} + \cdots + D_1$$

Where $C_J = \sum_{k\in\mathbb{Z}} c_{J,k}\,\varphi_{J,k}(t)$ and $D_J = \sum_{k\in\mathbb{Z}} d_{j,k}\psi_{j,k}(t); J = 1,2,3 \dots, J$.

The Discrete wavelet transform (DWT) represents time series data in terms of the coefficient that are associated with particular scales, so it is an effective tool for time series analysis. DWT decomposed signal (data) into a different scale of resolutions, by using inverse DWT we can reconstruct a signal from its wavelet coefficient. We have used MATLAB software to all the procedure for a prediction here also we used MATLAB to do DWT on signal means obtained wavelet coefficient. We will get two parts of wavelet coefficient one is smoothed part (low frequency) and other is detailed part (high frequency) part. for decomposition of the signal, a specific wavelet is required. There are various wavelets are available i.e Daubechies, Symlet, Meyer, etc. choice of mother wavelet is to depend on characteristics of data. we have used Symet wavelet to do decomposition.

3. ARIMA MODEL

The ARIMA word made of three other names AR-auto regressive, I–integrated, MA-moving average. When these three models are combined simultaneously it called as ARIMA model, this model is very famous due to its applicability for stationary and non-stationary time series data, in MATLAB ARIMA is defined as ARIMA (p,d,q) where p is order (lag) for AR term (autoregression),d is order of integrated term (difference), and q is for moving average part.

$$x_t = \emptyset x_{t-1} + \epsilon_t \;\Rightarrow\; AR(1)$$

above is for AR(1 order) which suggest number of lag elements in equation. For p order it can as below

$$x_t = \emptyset_1 x_{t-1} + \emptyset_2 x_{t-2} + \cdots + \emptyset_p x_{t-p} + \epsilon_t$$

Where x_t is data series and $\emptyset$ constant parameter which we can estimate by different methods i.e Yule walker, least square estimation etc.,$\epsilon_t \sim WN(0,\sigma^2), \epsilon_t$ is white Gaussian noise and p is call order of series.

Same for moving average

$$x_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$

Moving average part depends on noise part of current values as well as past values; it is working as low pass filter. When we combine both models with integrated is model ARIMA, Which can define as

$$\emptyset_p(B)(1-B)^d(x_t - \mu) = \theta_q(B)\epsilon_t$$

where $\emptyset_p(B) = 1 - \sum_{i=1}^{p} \phi_i B^i, \theta_q(B) = 1 - \sum_{j=1}^{q} \theta_i B^i$, are polynomial terms of B of degree p and q,$\nabla = (1-B)$, B=shift operator, i.e $B^j Z_t = Z_{t-j}, \mu$ =mean.

ARIMA is model which use Box-Jenkins methodology. We can determine order of AR,MA using ACF (autocorrelation function) and PACF (Partial autocorrelation function) but to determine ARMA and ARIMA order is challenging task for researcher do to exact order to fit model (in some cases), even though we can take help of AIC (Akaike information criterion), BIC (Bayesian information criterion) techniques to find optimum order for ARIMA.

Table 1 gives an idea to choose suitable models to among AR and MA, ARMA as well as suggesting a best fit order to forecast future value, suppose for AR (2), PACF plot will give cuts off after 2 lags and ACF plot will tails of gradually, same opposite for MA. Random distribution means ARMA.

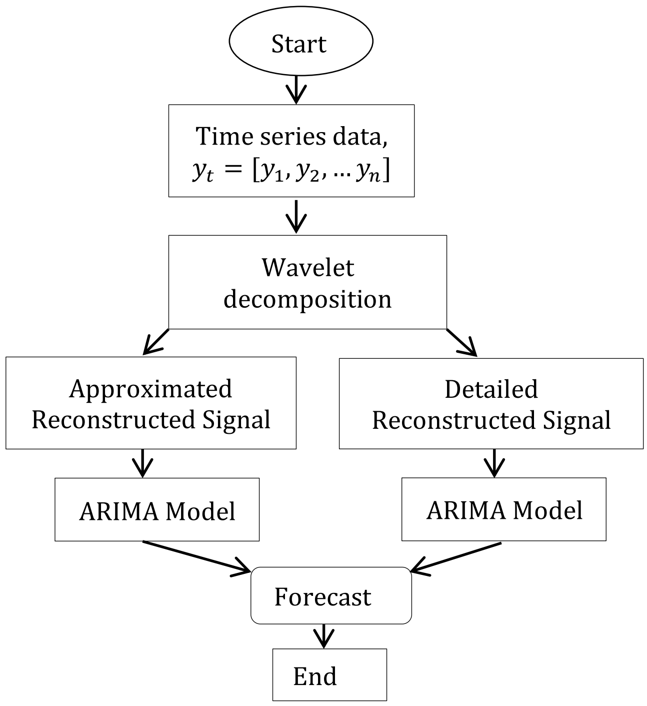
Table 1

Model	Autocorrelation	Partial Autocorrelation
AR(p)	Tails off gradually	Cuts off after p lags
MA(p)	Cuts off after q lags	Tails off gradually
ARMA(p, q)	Tails off gradually	Tails off gradually

4. HYBRID ARIMA-WAVELET

The block diagram 1 gives the flow of hybridization using two ARIMA model with wavelet decomposition method. In this method, first of all, we had decomposed time series using wavelet transform and reconstructed it in two parts (after inverse wavelet) these two parts categorized as approximated signal and detailed signal, we have applied separate ARIMA model to these parts then forecasting was applied. Moreover, we can use GARCH model as used in [13], if residues of hybrid ARIMA is heteroscedastic, we can use any vocality model for it

Block Diagram 1: Procedure for Hybrid ARIMA-Wavelet



4. MODELING AND METHODOLOGY FOR S&P BSE INFORMATION TECHNOLOGY

Basics steps for modeling

1. ***Plotting the data,***
2. ***Possibly transforming the data,***
3. ***Identifying the dependence orders of the model,***
4. ***Parameter estimation,***
5. ***Diagnostics, and***
6. ***Model choice.***

We have taken as only-ARIMA model as reference and explained the steps for it, first step in the modeling is plot data to check whether time series data is stationary or not, as well as to do various analysis if data is stationary we can use AR, MA, or ARMA model as by using Box-Jenkins methodology, which uses ACF and PACF to select order and check stationarity, if ACF will not decay to zero fast as graph of data increases. Thus, a slow decay in ACF is an indication that differencing may be needed, or we can use various stationary test i.e kpsstest, for which model is fit for our data we can refer Table 1.

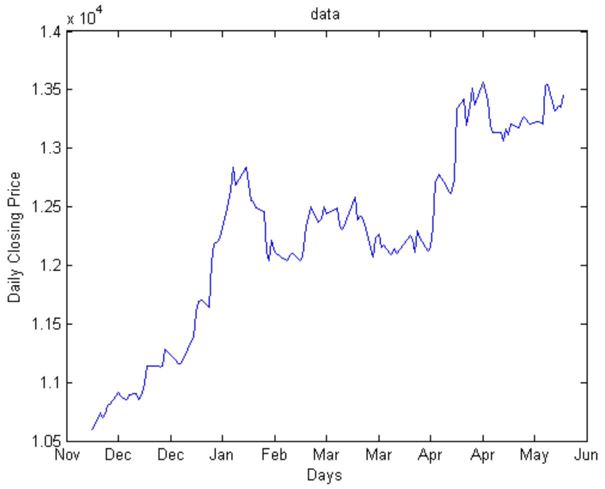


Figure 1: Original Data

Secondly, if data is not stationary we should have to make it stationary by differencing it, up to it would be stationary, a number of difference will be an order of 'I' term in ARIMA model, generally order 1 is enough for all series to best fit [1].

Figure 2 is ACF and PACF for data of Figure 1, after observing fig.2 and 1, we can say that data is not stationary because as data increases in fig 1, there is not decreasing order of lags in ACF (not meets zero quickly). Moreover, we had stationary *kpsstest,* test for it which also tells the series is not stationary, so after differencing we get ACF and PACF as shown in figure 3, where not any lag is crossing confidential area so this types of processes called random walk [1], here we can simply take order as (0,1,0). Moreover, according to [14]. as shown in figure 2, there is lag one in PACF which tells the AR (1) process, AR (1) coefficient (which is the height of the PACF spike at lag 1) will be almost exactly equal to 1, so no need to be differencing for that case. We have tried it but (0,1,0) gives the best result.

The third step is to find out the optimum order for this model which we can decide by the figure 3 is ARIMA(0,1,0), where AR(0), I(1), and MA order is 0. Also, we can use AIC/BIC methods to choose a suitable order for the model we have used BIC for this case which also suggests same values.

For estimating parameters which are constants i.e $\emptyset_1, \theta_1$ to choose best optimum values for it method like maximum likelihood and Bayesian methods can be used, here maximum likelihood is used by the software.

The fifth step is diagnostics, in this step, we will choose the best model by seeing vocality is there or not. if there we can use ARCH family model to remove it or select other best model order which does not have vocality component, also an important thing to check on Residues of series that it should not serial correlated (insignificant autocorrelation, therefore the residuals are uncorrelated in time). Means it should be like random Gaussian noise. We can see it by ACF plot of Residues and normal distribution of it if residues are uncorrelated in time tells the best fit, that order and other same types of order can compare with AIC, BIC methods, also RMSE and MAE, TIC like performance parameters may help to decide the optimum order. For this case the Residues plot and probability distribution plot are available in Figure 4, then the last step after selecting proper order is to use that model for forecast. Figure 5, gives forecasting for 10 days step ahead of data, table 2 list gives estimated parameters for this case. Forecasting can be written directly in difference equation by

$$x_{t+l} = \emptyset_1 x_{t+l-1} + \cdots + \emptyset_{p+d} x_{t+l-p-d} - \theta_1 \epsilon_{t+l-1} - \cdots + \theta_q \epsilon_{t+l-q} + \epsilon_{t+l}$$

where l is number of step ahead to forecast, $l \geq 1$

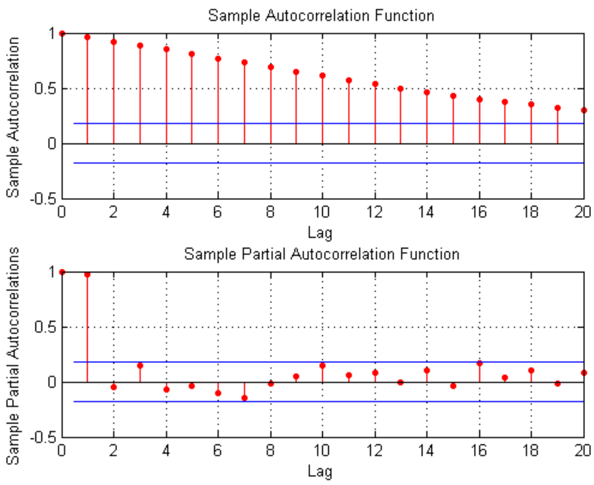


Figure 2: ACF and PACF of Data

Here, the software use the MMSE forecast which minimized the forecasted error between new predicted forecast value i.e suppose, at origin t, that we are to make a forecast $\hat{x_t}$ of x_{t+l} and then MMSE would be taken between them [1].

Forecasting Performance Measures:

We have use RMSE (Root Mean Squared Error), MAE (Mean Absolute Error), TIC (Theil Inequality Coefficient).

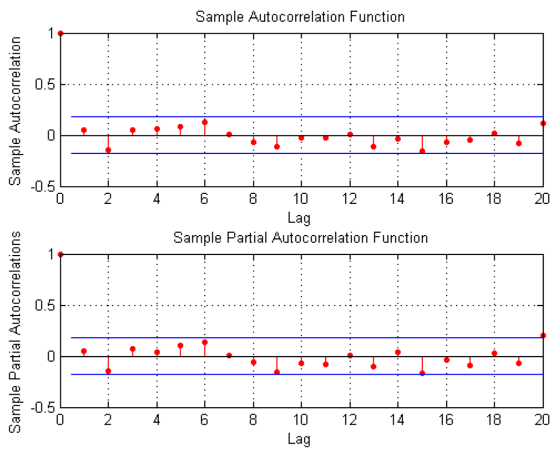


Figure 3: First differenced ACF and PACF of Data

RMSE:

RMSE measure difference between predicted value and true value, (a predicted value which does not step ahead of real value its predicted value at the same time), this gives accuracy about the model forecast. Minimum RMSE value gives goodness of model.

$$\text{RMSE}=\sqrt{\frac{\sum_{t=T+1}^{T+k}(\hat{y}_t-y_t)^2}{n}}$$

MAE (Mean Absolute Error):

Mean Absolute Error is the absolute value of average Residuals; here also minimum value is observed for the best fit model.

$$\text{MAE}=\sum_{t=T+1}^{T+k}\left|\frac{\hat{y}_t-y_t}{n}\right|$$

Theil Inequality Coefficient (TIC):

TIC gives a value between 0 to 1, near to 0 value gives a best model fit suggestion.

$$\text{TIC}=\frac{\sqrt{\sum_{t=T+1}^{T+k}(\hat{y}_t-y_t)^2}}{\sqrt{\sum_{t=T+1}^{T+k}\hat{y}_t^{\,2}}+\sqrt{\sum_{t=T+1}^{T+k}y_t^{\,2}}}$$

Where $\hat{y}_t$=predicted value, y_t=true time series values.

After completing these steps we can use this model for forecast future value. The forecast was done for 10 days ahead values.

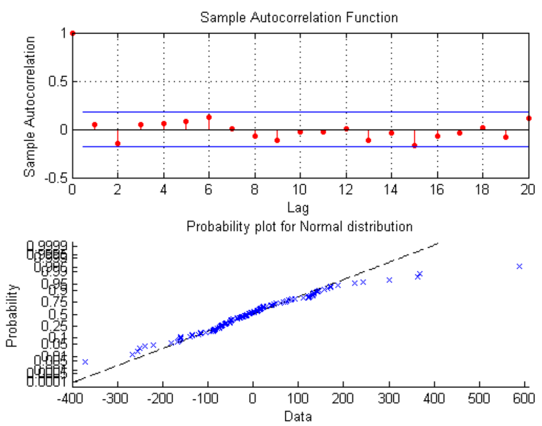


Figure 4: Residues "s ACF and Normal distribution

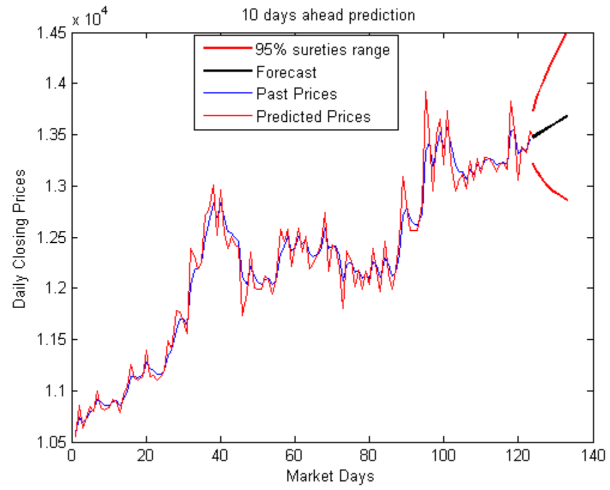


Figure 5: 10 days ahead forecast for Only-ARIMA

5. COMPARATIVE ANALYSIS AND RESULTS DISCUSSION

There are 124 days data are available based on that we have predicted next 10 days data, which we can see as in figure 5.

The same methodology was applied to other two methods

Method-1: In this method, we have directly applied ARIMA model after adding inversed wavelet which contains detailed and smoothed data.

Method-2: Hybrid method, which contains two separate ARIMA model after Wavelet decomposition.

Method-1: In this method, we have added both parts detailed part and approximated part and then we have applied ARIMA model which gives better results than ordinary ARIMA model. First of we have checked stationarity, where we found that data is not stationary, so we have differenced it once again checked it for stationarity which tells that data is stationary, so here the order of 'I' is selected as 1, Figure 6 tells about first differenced ACF and PACF plot. For selecting order of ARIMA model ACF and PACF observed where ACF cuts down after lag 1 and PACF is random which telling MA effect, so we can select order (0,1,1) but which failed to residues test of using ACF and PACF as well as, Ljung-Box Q-test, in that it is failed but we have checked its nearest value,(0,1,2) which is successfully satisfied, so it is best fit model, but the vocality test not satisfied for it so we can use ARCH family model on residues to remove it. We can see in Figure 7, which tells about a test of a perfect fit model is satisfied. Because its normal distribution flows the cross black line, which suggests residues are random in nature, same in ACF, lags are not broken confident level, these things tell that model is perfectly fit. Table 3 shows chosen parameters for the selected model to perfectly fit.

Figure 8 gives forecated prices for our derived model with 95% surety range.

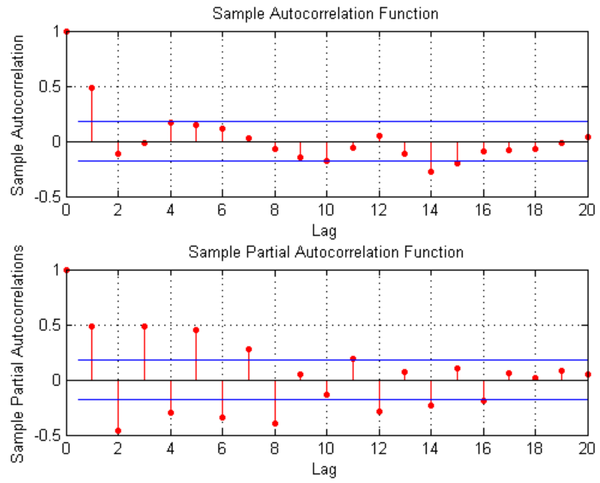


Figure 6: First differenced ACF and PACF of Method-1 (Approx+detailed parts)

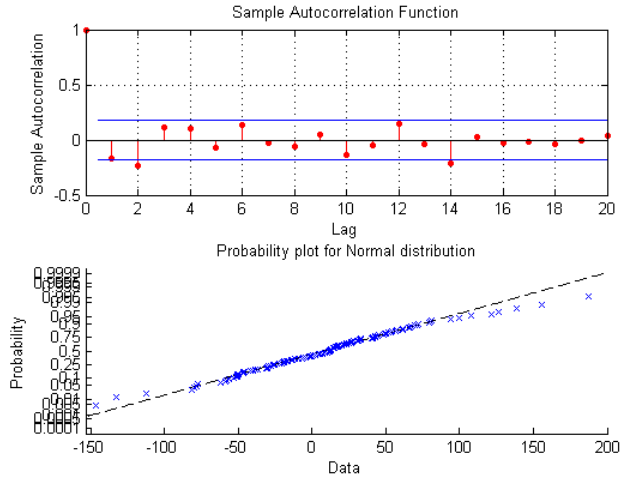


Figure 7: Residues "s ACF and Normal distribution

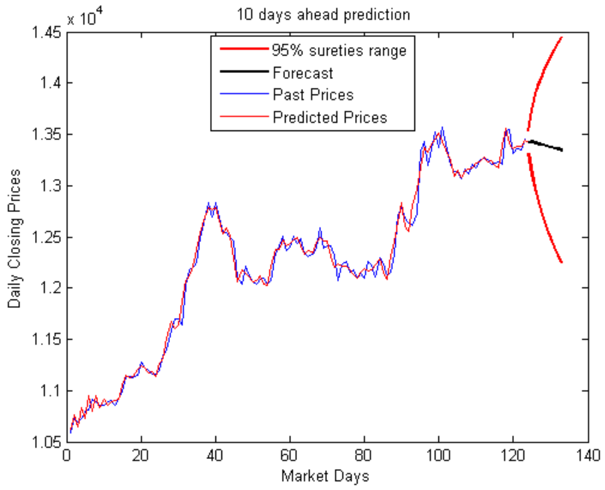


Figure 8: Forecast using ARIMA-Wavelet (Method-1)

Method-2: In this method we have applied two different ARIMA model for two different parts of wavelet data, approximated and smoothed, to determine its order, first of check stationarity test for both the part separately, where gets that approximated part is stationary by two times differencing, for that the ACF and PACF is given in figure 9. The second part of wavelet called detail part is already stationary process, hence do no go ahead, now in both plot, we can see that it suggest ARMA process, where an approximate part needs difference by two, and detail is simple ARMA model. Here its seems too difficult to decide order but using ACF and PACF observation as well as help of BIC method we came to do select best-fitted model is (3,1,3) for approximated part and (1,0,6) for detailed part, here we have used only 1 differenced parameter instead of using 2, because we get good result for 1 and [14] when model is AR(1) or more than it, it makes one differenced by itself so no need to choose higher difference order. After selecting appropriate order, we have done Residues test and vocality using Ljung-Box Q-test. Which suggest that there is not any vocality available, and the model is perfectly fit, for the data. The plot of resides autocorrelation function and normal distribution of same is available in figure 11. Where ACF tells that the values of the correlation coefficient of residuals for various time lag is not significantly different for zero, and its residual value indicates that the there is no significant autocorrelation.

All the values are lied between confidential area which presented with two blue line, which suggests residue is normal, in the same figure probability distribution has been plotted for residual, on the figure we can see that distribution is following the blacked dotted line, which shows normal distribution, because residues are random and normally distributed, which brings to assume noise as a white Gaussian noise, that helps to decide optimum parameters of models for better forecast. Table 4 shows estimated parameters for the selected model. Figure 12 shows the forecasted data for 10 days to the hybrid ARIMA model (second model). moreover, table 5 shows comparative analysis done by comparative parameters.

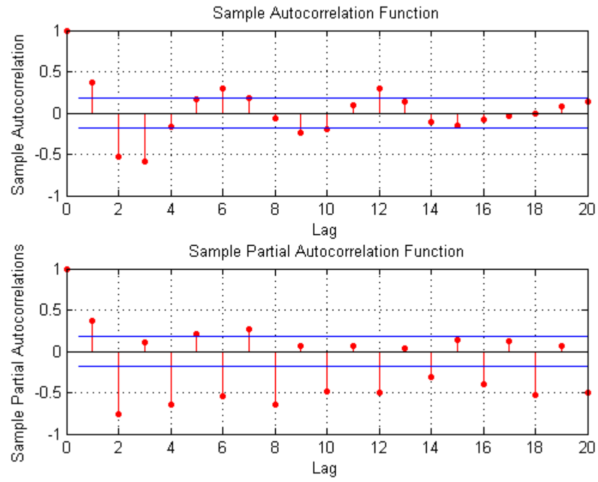


Figure 9: Second differenced ACF, PACF for method-2(approximated part)

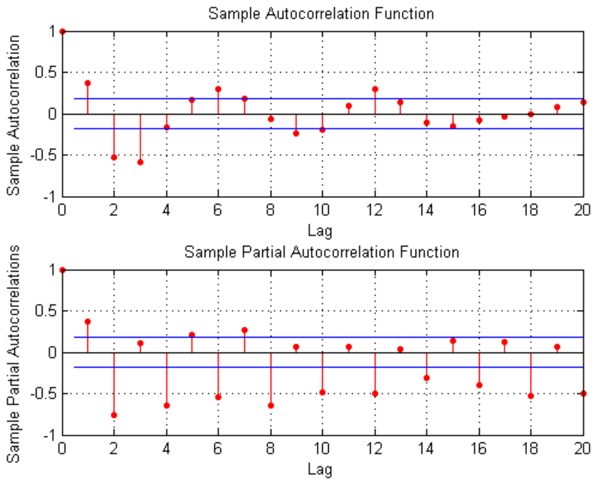


Figure 10: detailed part ACF&PACF for method-2

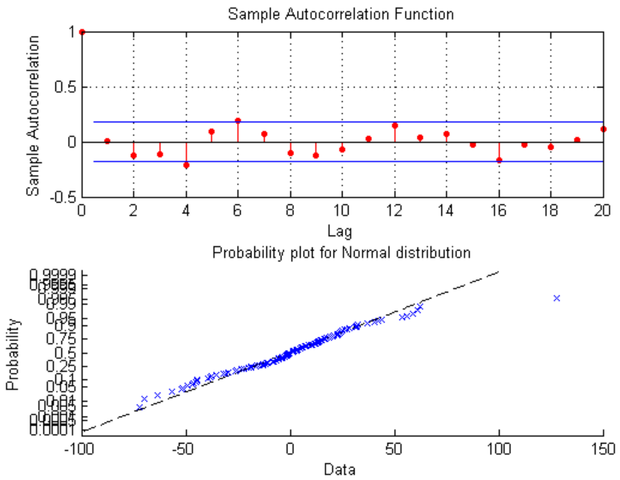


Figure 11: Residue's ACF and normal Distribution

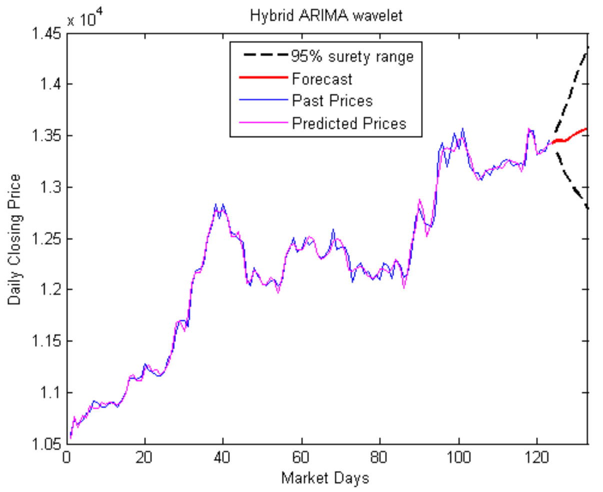


Figure 12: Hybrid ARIMA-Wavelet (Method-2) Forecast

Table 2: Parameter for only ARIMA

Parameters	ARIMA(0,1,0)
Constant	22.698
Variance	17384.4

Table 3: Parameters for method-1

Parameters	ARIMA-Wavelet (0,1,2)
MA	[1.67, 0.752]
constant	-10
variance	3040.33

Table 4: Parameters for method-2

Parameters	ARIMA-Wavelet (3,1,3)	ARIMA-Wavelet (1,0,6)
AR	[0.695,-0.33,0.26]	0.371
MA	[0.968, 0.929, 0.963]	[0.319, -1.5, -1, 0.615, 0.59 ,0.04]
constant	7.690	0.013
Variance	291.263	620.525

Table 5: Forecasting Performance

Model	RMSE	MAE	TIV
ARIMA	131.54	95.42	0.0054
ARIMA-with Wavelet (First method)	77.97	60.40	0.0032
Hybrid Arima-wavelet(second method)	62.76	47.76	0.0026

3. CONCLUSIONS

After doing analysis for S&P BSE Telecom data using ordinary method as well as two unique methods using wavelet transform with ARIMA, among three methods Hybrid ARIMA method (method-2) which used two ARIMA model, gives best results by all the tests as well as RMSE performance is very less compared to both the methods and also removes some vocality parameters due to control on both parts, which suggest the supremacy of this method. While wavelet with ARIMA also gives good performance compared to normal ARIMA model. The disadvantage of these hybrids methods is to find the best order. There might be a chance of difficulty on that in some cases due to filtering and two separated parts, it might vary case by case but use ACF&PACF and AIC/BIC, and trial and error methods help to overcome.

REFERENCES

[1] Box, G.E.P., Jenkins, G.M., REINSEL, G.C., LJUNG, G.M,. : 'Time series analysis forecasting and control' (Holden-Day, San Francisco, 2016, 5th edn.), ISBN: 978-1-118-67502-1

[2] R. Adhikari, R. Agrawal, "An Introductory Study on Time Series Modeling and Forecasting", arXiv preprint arXiv:1302.6613, 2013

[3] N. A. A. Ali and M. Abu-Elkheir, “Regression forecasting model to improve localization accuracy,” 2015 International Wireless Communications and Mobile Computing Conference (IWCMC), 2015.

[4] J. Park, G. Lee and Y. Choi, "Determining ARMA model parameters for biomedical signal applications based on estimation of the equivalent AR model," 2015 12th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI), Goyang, 2015, pp. 456-457.

[5] M. Laner, P. Svoboda and M. Rupp, "Parsimonious Network Traffic Modeling By Transformed ARMA Models," in IEEE Access, vol. 2, pp. 40-55, 2014.

[6] F. Mustière and M. Bouchard, "A novel application of ARMA modelling to audio coding," 2013 IEEE Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE), Napa, CA, 2013, pp. 233-238.

[7] M. Deriche and D. Ning, "Warped ARMA filters in high quality audio coding," Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings., 2003, pp. 153-156 vol.2.

[8] H. Kang, Y. Qi, G. Liu, J. Liu and C. Li, "Application of the ARMA model in non-stationary vibration signals," 2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering, Chengdu, 2012, pp. 751-754.

[9] K. B. Eom, "ARMA modeling in multichannel filter banks with applications to radar signal analysis," 1995 International Conference on Acoustics, Speech, and Signal Processing, Detroit, MI, 1995, pp. 3647-3650 vol.5.

[10] M.T.I.S. Al Wadia, M. Tahir Ismail “Selecting wavelet transforms model in forecasting financial time series data based on ARIMA model''Applied Mathematical Sciences, 5 (7) (2011), pp. 315-326

[11] R. H. Shumway and D. S. Stoffer, Time Series Analysis and its Application. 862 New York, NY, USA: Springer, 2011

[12] Polikar, R., 2001, The Wavelet Tutorial Part I.[Online].Available: web.iitd.ac.in/~sumeet/WaveletTutorial.pdf [Accessed May 23, 2018].

[13] S.R. Yaziz, N.A. Azizan, R. Zakaria and M.H. Ahmad, The Performance of Hybrid ARIMA GARCH Modeling, In: 20th International Congress on Modelling & Simulation 2013 (MODSIM2013), 1-6 December 2013, Adelaide, Australia

[14] Robert Nau,''Identifying the numbers of AR or MA terms in an ARIMA model''.[Online].Available:https://people.duke.edu/~rnau/411arim3.htm#which. [Accessed: May 23, 2018].

[15] Jatinder kumar, Amandeep Kaur, and Pammy Manchanda. "Forecasting the time Series Data Using ARIMA with Wavelet." Journal of Computer and Mathematical Sciences 6.8 (2015): 430-438.