Available online at www.isas.org.in/jisas
JOURNAL OF THE INDIAN SOCIETY OF
AGRICULTURAL STATISTICS 74(3) 2020 187–194

Application of Machine Learning Techniques with GARCH Model for Forecasting Volatility in Agricultural Commodity Prices

Tanima Das, Ranjit Kumar Paul, L.M. Bhar and A.K. Paul
ICAR-Indian Agricultural Statistics Research Institute, New Delhi

Received 20 January 2020; Revised 16 July 2020; Accepted 22 July 2020

SUMMARY

Food price forecasting is very useful for farmers, consumers, policy makers and industrialists too. In the recent era, crucial management of food security in agriculture has a great value in India. Volatility forecasting is an integral part of commodity trading and price analysis. In many literatures, the inefficiency of single parametric model in capturing volatility in a series has been strongly proved. In this context non-parametric nonlinear models like Support Vector Regression (SVR) and Neural Network (NN) may be used to improve forecasting performance. Hence, in search of improved alternatives to the classical econometric methods, machine learning techniques viz. SVR, NN is applied along with its combination with GARCH model. The outperformance of this new approach has also been established by means of Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and R^2 log.

Keywords: GARCH, Hybrid model, Neural network, Onion price, SVR.

1. INTRODUCTION

Various important agricultural commodities mainly the perishable products depict volatility *i.e.* all possible outcomes of an uncertain variable. Among those, onion is very important fresh vegetable consumed all over the world around the year. India ranks first in onion acreage in the world covering 21% of the world area and second in production next to China. Onion production in India is about 24.45 million ton but in India, onion is produced only in some concentrated pockets like Maharashtra, Gujrat and Madhya Pradesh. Onion, produced in those states finds major markets in Delhi, Bhopal, Lucknow, Kolkata etc. Erratic weather and fluctuation in production is the major factors causing volatile market price, also result in irregular and excess supply and demand. The well-known Box-Jenkins Autoregressive Integrated Moving Average (ARIMA) model is unable to capture volatility in price (Box *et al.*, 2008) and to handle this type of uncertainty where variance changes over time, "nonlinear time-series models" should be taken under consideration. The most commonly used parametric nonlinear time-series model is Autoregressive Conditional Heteroscedastic (ARCH) model (Engle, 1982). To cope up with its limitations like rapid decaying unconditional autocorrelation function (acf) of squared residuals, Bollerslev (1986) proposed Generalized ARCH (GARCH) model. Ghosh *et al.* (2010) employed GARCH nonlinear time series model to describe data sets depicting volatility and they also studied its estimation procedure. To overcome the deficiencies of any single model, use of various hybrid models have been proposed by combining different time-series models together (Pagan *et al.*, 1990; Khashei *et al.,* 2012; Anjoy *et al.*, 2017; Paul *et al.*, 2020). It is well established that, nonparametric models can be applied in various situations with higher efficiency than a parametric one. With such background "Nonparametric Nonlinear Time-Series Modelling" has created entire interest of the researchers. In this class, "Machine Learning Techniques" can be applied. Many nonlinear processes with unknown functional relationship can be modelled by "Machine Learning". The superiority of "Machine Learning Technique" such as Support Vector (Boser, 1992; Cortes, 1995)

Corresponding author: Ranjit Kumar Paul
E-mail address: ranjitstat@gmail.com

and Artificial Neural Network (NN) over traditional time-series models has been proved in many research papers (Wedding *et al.*, 1996; Zhang *et al.*, 1998; Zhang, 2003; Kim, 2003; Mitra *et al.*, 2017; Paul *et al.*, 2019). Donaldson *et al.* (1997) developed a semi-nonparametric nonlinear GARCH model by the use of the Artificial Neural Network (ANN) and also evaluated its forecasting ability of stock return.

In this paper, some brief introduction of various models has been discussed. Then description of "Support Vector Regression Technique" is given along with the form of its application in time-series modelling and forecasting. Our purpose is to study GARCH model and its combination with SVR and NN hybrid model and also to compare these models by considering the monthly maximum market price of onion in Kolkata market during the period of January 2004 to December 2018.

2. DESCRIPTION OF MODELS

2.1 GARCH Model

Bollerslev (1986) proposed the Generalized ARCH (GARCH) model. The process $\{\epsilon_t\}$ is Autoregressive Conditional Heteroscedastic model (ARCH) (q), if the conditional distribution of $\{\epsilon_t\}$ given ψ_{t-1} (denotes the available information up to time $t-1$) is:

$$\epsilon_t|\psi_{t-1} \sim N(0, h_t) \text{ and } \epsilon_t = \sqrt{h_t}\varepsilon_t,\ \varepsilon_t \sim iid(0,1) \quad (1)$$

$$h_t = a_0 + \sum_{i=1}^{q} a_i\, \epsilon_{t-i}{}^2 \quad (2)$$

where $a_0 > 0, a_i \geq 0\ \forall i$ and $\sum_{i=1}^{q} a_i < 1$ are necessary conditions to be satisfied to ensure the non-negativity and finite unconditional variance (h_t).

The GARCH (q,p) process has the following form provided $a_0 > 0, a_i \geq 0\ \forall i;\ b_j \geq 0\ \forall j$, as:

$$h_t = a_0 + \sum_{i=1}^{q} a_i\, \epsilon_{t-i}{}^2 + \sum_{j=1}^{p} b_j h_{t-j} \quad (3)$$

The GARCH (q,p) process will be weakly stationary if and only if

$$\sum_{i=1}^{q} a_i + \sum_{j=1}^{p} b_j < 1 \quad (4)$$

To estimate the parameters of GARCH model, "Method of Maximum Likelihood" is used. The log likelihood function of a sample of T observations (apart from constant), is

$$L_T(\theta) = T^{-1} \sum_{t=1}^{T} (log h_t + \epsilon_t^2 h_t^{-1}) \quad (5)$$

Some of the applications of GARCH models in agriculture can be found in Paul *et al.* (2009, 2014 and 2016) and Paul (2015).

2.2 Support Vector Regression

Considering a data set $S = \{(x_i, y_i)\}$, $x_i \in R^n$ and $y_i \in R, i = 1,2,\ldots,N$, where x_i is the input vector, y_i is output (scalar) and N is the size of S. The general form of the "Nonlinear SVR" is:

$$f(x) = u^T g(x) + a \quad (6)$$

where $g(.)\colon R^n \to R^{n_h}$, which is a non-linear mapping function from input space to a higher dimensional feature space which may be of infinite dimensions, $u \in R^{n_h}$ is the weight vector and a denotes bias. Basically, SVR minimizes the following function:

$$R(C) = \frac{1}{2}\|u\|^2 + C\left[\frac{1}{N}\sum_{i=1}^{N} L_\varepsilon(y_i, f(x_i))\right] \quad (7)$$

where $x_i \in R^n$ and both C and ε are user-determined hyper-parameters. The 'Empirical error' which is estimated by the "Vapnik's ε-insensitive Loss Function". Where, the Vapnik's Loss function is given by:

$$L_\varepsilon(y_i, f(x_i)) = \begin{cases} |y_i - f(x_i)| - \varepsilon & |y_i - f(x_i)| \geq \varepsilon \\ 0 & |y_i - f(x_i)| < \varepsilon \end{cases}$$

where y_i denotes actual value and $f(x_i)$ is the estimated value at period i^{th} period.

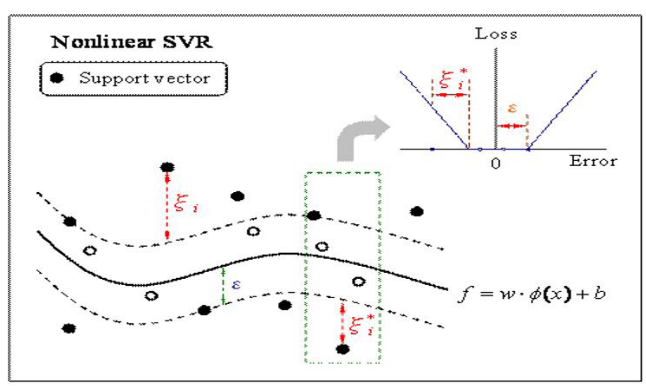

Fig. 1. Schematic diagram of Nonlinear Support Vector regression

2.3 Neural Network

Another non-linear approach to deal with such series is a class of mathematical model, viz. Neural Network (NN). Moreover, the non-parametric, data-driven and self-adaptive nature of NN has made it more easily approachable method. ANN forecasting has two basic steps *i.e.* Training and Learning. A Neural Network is a series of algorithms that endeavors to recognize underlying relationships in a set of data through a process that mimics the way the human brain operates by making the right connections. Like the structure of neuron, ANN comprises of several layers namely: input layer that receives external information; one or more hidden layer that performs mathematical operations on the data and an output layer that produces the results. All the layers are connected through an acyclic arc (Khashei *et al.,* 2010). NN can adapt to changing input and it generates the best possible result without redesigning the output criteria. These NNs are more flexible computing system for modelling a wide variety of nonlinear problems. There are two Neural Network topologies namely, feed-forward and feedback. In feed-forward topology, the flow of information is unidirectional and there is no feedback path where as in feedback topology, feedback paths are there.

The application of neural network structure for solving a particular time-series problem involves determination of number of layers and total number of nodes in the structure which is done on experimentation basis. Single hidden layer with sufficient number of nodes at the hidden layer and adequate data for initialization are well established. In neural network determination of number of input nodes which are lagged observations of same variable plays an important role in model building.

The relationship between one output (y_t) and the p inputs $(y_{t-1}, y_{t-2}, \dots, y_{t-p})$ is represented by

$$y_t = \alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij} y_{t-i}\right) + \varepsilon_t \quad (8)$$

where $\alpha_j (j = 0, 1, 2, \dots, q)$ is the weight attached to the connection from the jth hidden node to the output node and $\beta_{ij} (i = 0, 1, 2, \dots, p;\ j = 0, 1, 2, \dots, q)$ represents the weight attached to the connection between ith input node and jth node of hidden layer; p is the number of input nodes and q is the number of nodes at hidden layer and g is the activation function which is usually a nonlinear function.

3. FITTING OF HYBRID MODEL COMBINING GARCH-SVR AND GARCH-NN

For univariate time-series forecasting problems, dimension of the input vectors is supposed to be the past lagged observations. Both SVR and NN model perform a functional mapping from past observations to future value as:

$$y_t = f(y_{t-1}, y_{t-2}, \dots, y_{t-p}) + \varepsilon_t$$

here y_t is the output value and ε_t is error term and p is the size of input vectors. The volatility of a time-series can be supposed to be comprised by two parts as:

$$\sigma_t^2 = h_t + R_t$$

This $\hat{h}_t$ can be first obtained by GARCH, then the residual at time t, *i.e.* R_t will be:

$$R_t = \sigma_t^2 - \hat{h}_t$$

R_t can be model by SVR or NN as follows:

$$R_t = \varphi(R_{t-1}, R_{t-2}, \dots, R_{t-p}) + \epsilon_t^*$$

where p is the no. lagged observations and $\varphi(.)$ is the function estimated by SVR or NN. $\hat{R}_t$ can be obtained by applying any of these models. Therefore, the combined forecast is

$$\hat{\sigma}_t^2 = \hat{h}_t + \hat{R}_t$$

So, the procedure to forecast volatility comprises the following steps:

1. Actual data is seasonally adjusted then percentage log-return series (r_t) is calculated.
2. Best suitable mean model (ARIMA) is fitted on r_t series.
3. ARCH - LM test is conducted on the residuals obtained from the mean model to check the presence of conditional volatility.
4. A specific GARCH model is fitted to model the conditional volatility $(\hat{h}_t)$.
5. Residuals are calculated from the fitted GARCH: $R_t = (r_t^2 - \hat{h}_t)$, where actual volatility can be considered as the squared percentage price log-return (r_t) (Li *et al.*, 2013), which can be calculated as:

$$r_t = 100 * (\ln(Y_t) - \ln(Y_{t-1}))$$

6. The residual series is diagnosed, if those are correlated or not. If correlated, the method proceeds to step 7. Otherwise the method proceeds towards final result.
7. SVR or NN is applied to forecast the residual series $(\hat{R}_t)$.
8. Again residuals $(R_t - \hat{R}_t)$ are obtained from the fitted improved technique and diagnosed.
9. If they are again correlated, then the method should be reviewed from step 7; otherwise it proceeds to step 10.
10. Final forecast of the volatility can be obtained by the forecasted value of original volatility plus the residual obtained by SVR or NN *i.e.* $(\hat{h}_t + \hat{R}_t)$.

Efficiency criteria of forecasting ability

To compare the accuracy of the forecasting performances of the hybrid models and the former GARCH models, mainly three criteria are to be used as:

$$MAE = \sum_{t=1}^{n} \left| r_t^2 - \widehat{r_t^2} \right| / n$$

$$MSE = \sum_{t=1}^{n} \left(r_t^2 - \widehat{r_t^2} \right)^2 / n$$

$$R^2LOG = \sum_{t=1}^{n} \left[ln\left(r_t^2 \Big/ \widehat{r_t^2} \right) \right]^2 / n$$

where, R^2LOG is similar to the R^2 of the Mincer-Zarnowitz regression (Mincer and Zarnowitz, 1969).

4. AN ILLUSTRATION

Agricultural commodity particularly the perishable vegetables often exhibit a noticeable extent of volatility. The seasonal variation, perishable nature and market demand of vegetable specifically of onion are causing this volatile feature of market price. In most of the situations a single parametric model is not sufficiently extensive to capture this volatility. Two hybrid models are to be fitted consisting SVR and NN models along with the GARCH on monthly wholesale onion market price.

4.1 Dataset

For the present investigation, monthly maximum wholesale market price of onion is collected from National Horticulture Research and Development Foundation website (http://nhrdf.org/en-us/) from January, 2004 to December, 2018 for a major market namely Kolkata. Total observation points are 180 of which first 162 datapoints (January, 2004 to June, 2017) are used training purpose and last 18 points (July, 2017 to December, 2018) are used for validation set. Firstly, this actual data has been converted to percentage log-return series (r_t).

4.2 Descriptive statistics

The descriptive statistics are presented in Table 1. The average monthly onion wholesale price is Rs. 1505 per quintal for Kolkata market. The standard deviation is large enough for this market depicting high level of variability.

4.3 Seasonal indices

The actual data series for two markets are first seasonally adjusted to omit the seasonal influence. Table 2 shows the seasonal indices for all the twelve months of those markets. The lowest values of seasonal indices are found in the month of April while the indices are obtaining their maximum value in the month of November.

Table 1. Descriptive statistics of onion price

Markets	Kolkata
Mean(Rs./ q)	1504.93
Standard Deviation(Rs./ q)	993.76
Kurtosis	3.27
Skewness	1.77
Minimum(Rs./ q)	382
Maximum(Rs./ q)	5154
CV(%)	66.03

4.4 Seasonal adjustment

The first and foremost step in a time-series analysis is to plot the data to visualize the presence of several time-series components. Fig. 2 shows the time-series plot of average monthly price of onion for actual series and monthly onion price for seasonally adjusted series from January, 2004 to December, 2018 for Kolkata. An overlook on these figures indicates that the price attains its higher values during the period August to December,

Table 2. Seasonal Indices of onion prices

Months	Kolkata
January	137.62
February	-189.17
March	-598.59
April	-638.68
May	-597.76
June	-365.06
July	-107.16
August	365.57
September	473.46
October	574.56
November	596.07
December	358.15

every year. The highest price has been observed in October and November.

4.5 Stationarity test and ARCH-LM test

Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test has been carried out to test about the stationarity of the series. It was found that the null hypothesis of the test depicts the non-stationarity in seasonally adjusted. Non-rejection of the null hypothesis for the log-return series at 1% level of significance indicates that the log-return series are stationary. The test results are given in Table 3. Later the percentage log-return series is tested for presence of conditional heteroscedasticity by the use of ARCH-LM test (shown in Table 4).

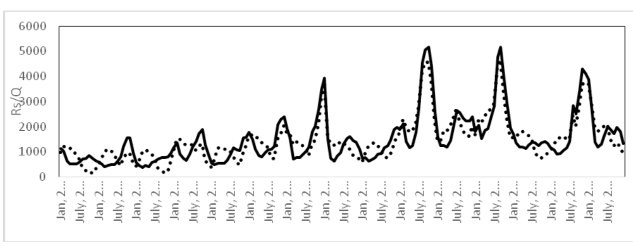

Fig. 2. Original onion price series (solid line) and seasonally adjusted series (dotted line) for Kolkata market

Table 3. KPSS test result for seasonally adjusted and percentage Log-return series

Series	Statistic	p-value
Actual series	1.53	<0.01
Percentage Log-return series	0.02	0.1

Table 4. Result of ARCH-LM test of onion price data

Order	Statistic	p-value
4	98.98	<0.01
8	38.28	<0.01

Table 5. BDS nonlinearity test result of percentage log-return series

Epsilon	13.99		27.98		41.97		55.96	
[2]	4.23	<0.01	4.20	<0.01	3.78	<0.01	3.01	<0.01
[3]	5.45	<0.01	5.21	<0.01	4.24	<0.01	3.18	<0.01

4.6 Nonlinearity test

BDS nonlinearity test has been carried out to know whether the data sets are nonlinear or not. The result of these corresponding tests is shown in Table 5 which clearly indicates the nonlinearity of the dataset.

4.7 Fitting of mean model and GARCH model

ARIMA and GARCH are fitted with suitable orders for the wholesale market price data of onion. The parameter estimates of best fitted ARIMA and GARCH are given in Table 6 and along with their significance level. At first ARIMA and GARCH are fitted on the whole datasets means on all the 179 data points. From this GARCH, conditional variance of whole data is obtained. Actual volatility can be considered as the squared percentage price return (Li *et al.*, 2013) $\mathbf{r_t = 100 * (\ln(Y_t) - \ln(Y_{t-1}))}$. From this, squared percentage log-return series and conditional variance of the residual series can be calculated. It is observed that the coefficients of AR and GARCH effect are highly significant.

4.8 Fitting of GARCH-SVR model

The obtained residual series is divided into two parts namely "training" and "valid" datasets. On the training dataset, Support Vector Regression is fitted considering up two lags as domain.

4.9 Fitting of GARCH model on estimation set

Appropriate GARCH model is fitted only for the training data set consisting 161 data points. The parameter estimates along with their significance level are shown in Table 7. This is true in case of whole set.

4.10 Accuracy checking

While fitting SVR, NN or GARCH models, it is necessary to check the accuracy of their residual series. For this reason, residual series of the GARCH and the combined models fitted on the estimation set are obtained. The acf plots of these residual sets are shown in Fig. 3, 4 and 5. This result depicts that only GARCH model is not sufficient to capture this volatility in time-series data as acf are significant in many lag values.

Table 6. Parameter estimates of ARIMA and GARCH for whole data of Kolkata market

Model	Mean Equation		
		Estimate	p-value
ARIMA (1, 0, 0)	C	-0.18	0.92
	AR	0.31	<0.01
	Variance Equation		
GARCH (1, 1)	C	0.007	1
	ARCH	0.012	0.25
	GARCH	0.98	<0.01

Table 7. Parameter estimates of ARIMA-GARCH for estimation dataset of Kolkata market

Model	Mean Equation		
		Estimate	p-value
ARIMA (1, 0, 0)	C	0.16	0.93
	AR	0.33	<0.01
	Variance Equation		
GARCH (1, 1)	C	0.008	1
	ARCH	0.02	0.34
	GARCH	0.97	<0.01

Hence, other improved techniques can be used to depict the characteristics of the time-series data.

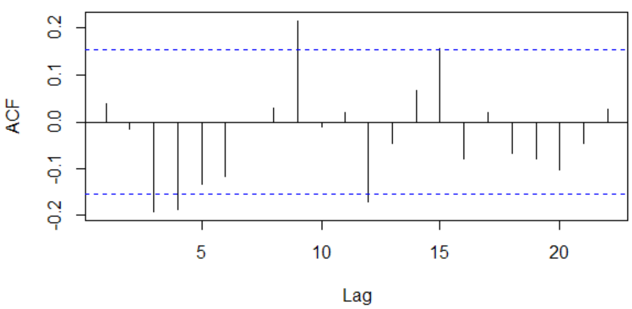

Fig. 3. Acf plot of residuals of the GARCH model for Kolkata

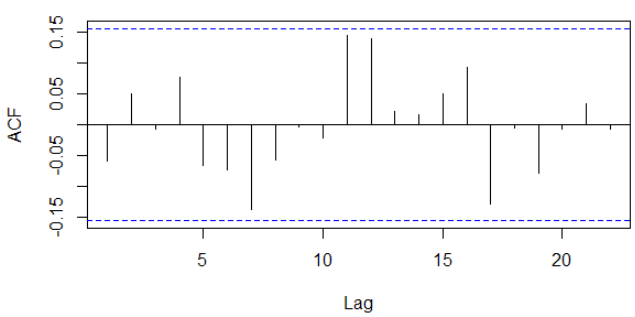

Fig. 4. Acf plot of residuals of the GARCH-SVR model for Kolkata

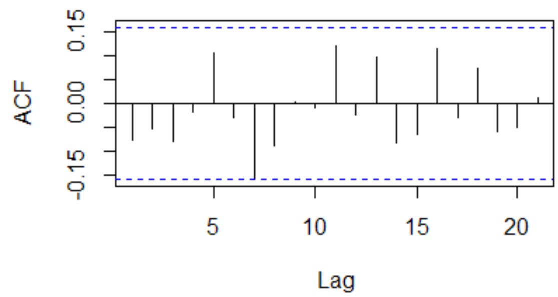

Fig. 5. Acf plot of residuals of the GARCH-NN model for Kolkata

4.11 Validation of results

The traditional GARCH model and the combined techniques *i.e.* GARCH-SVR and GARCH-NN models are compared on the basis of their forecasting performance. Here, as the main intension of the study is to model the volatility, the validation also has been done by seeing accuracy of the models to capture the volatility. To compare these results, the validation set as well as the estimation set are used. This comparison is done by Mean Square Error (RMSE), Mean Absolute Error (MAE) and the R^2 log values. The corresponding results are tabulated in the Table 8. So, there is a precise difference between the RMSE of GARCH and the other improved combined techniques. Considering RMSE, the combined technique GARCH-NN can be taken as the best model. Similarly, considering MAE and R^2 log values, the outperformance of improved combined models over GARCH method can be strongly established. While considering the estimation set, the superiority of hybrid models over the GARCH model can also be well-established. It can also be noticed that, for estimation set GARCH-NN model performs much better than the GARCH-SVR. At last, square log-return series and the predicted series by GARCH, GARCH-SVR and GARCH-NN models are plotted separately

Table 8. Prediction performance of GARCH and GARCH-SVR model

Dataset	Validation Criteria	GARCH	GARCH-SVR	GARCH-NN
Validation set	RMSE	833.07	779.47	728.03
	MAE	517.19	470.81	497.87
	R^2 log	66.47	8.73	9.95
Estimation set	RMSE	1609.94	1465.95	789.37
	MAE	762.17	694.21	518.70
	R^2 log	17.35	7.05	8.78

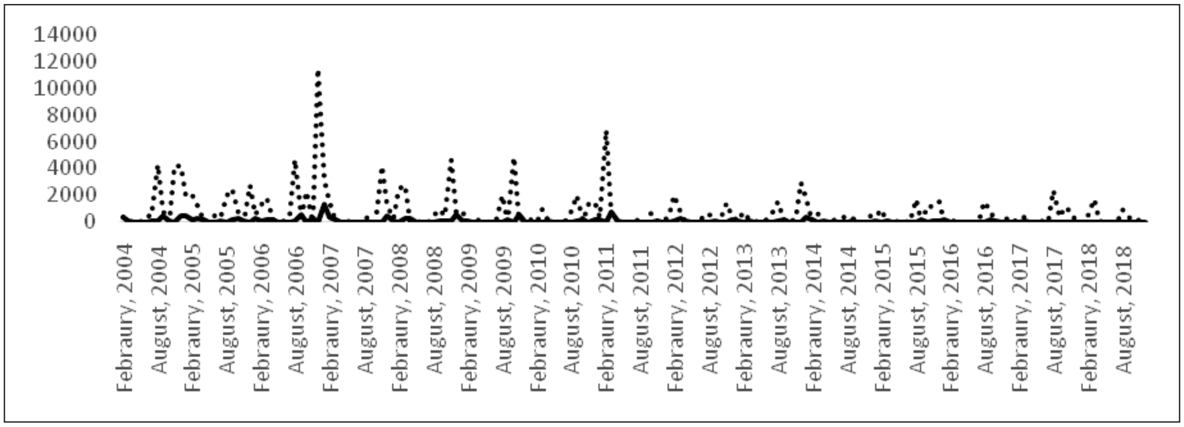

Fig. 6. Observed series (dotted line) and fitted series (solid line) by GARCH for Kolkata

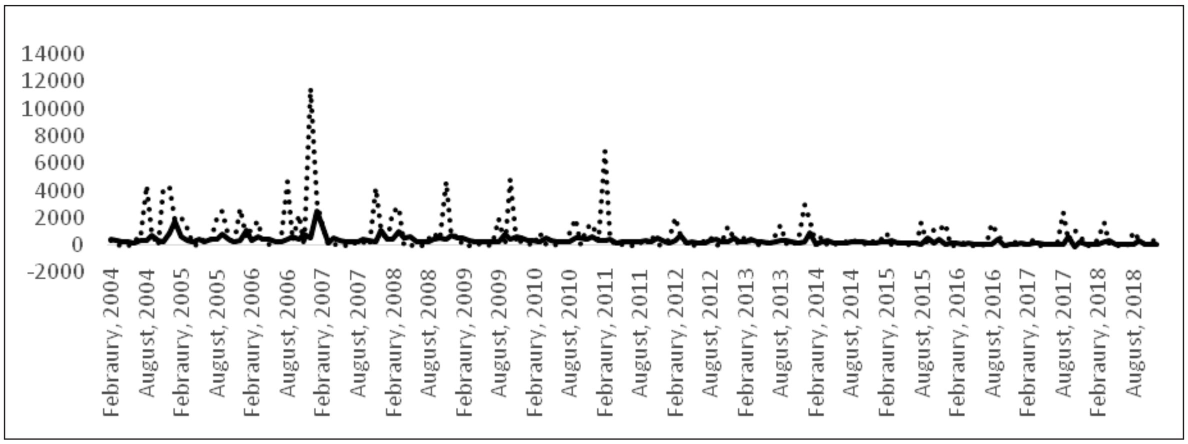

Fig. 7. Observed series (dotted line) and fitted series (solid line) by GARCH-SVR for Kolkata

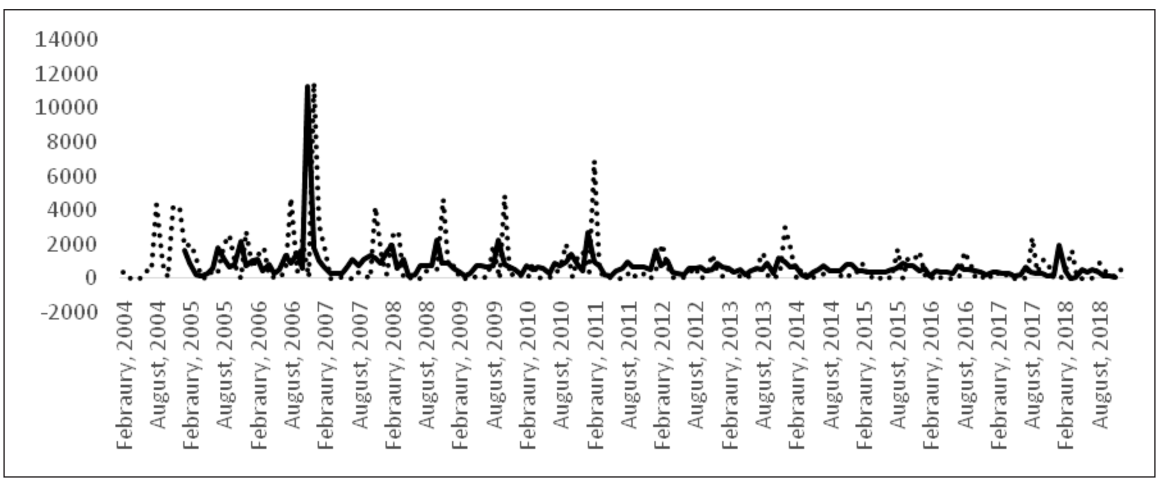

Fig. 8. Observed series (dotted line) and fitted series (solid line) by GARCH-NN for Kolkata

as shown in Fig. 6 to Fig. 8. GARCH model cannot capture the uncertainty with that much accuracy as in case of the other combined approaches. Hence, the accuracy of prediction by the hybrid models is much better than that of GARCH model.

5. CONCLUSIONS

The accuracy of a statistical model is of fundamental interest for selecting that particular model and taking many important decisions. Improved machine learning techniques viz. SVR and Neural Networks are applied along with traditional GARCH model to develop some unique approaches to deal with the volatility. In the present investigation, SVR and NN based hybrid approaches are found to be superior to the traditional GARCH model in modelling and forecasting volatile onion prices. Considering RMSE, MAE and R^2 log value as efficiency criteria, these comparisons between GARCH with GARCH-SVR and GARCH-NN models have been done. So, machine learning based improved techniques should be used in modelling volatility rather than using single GARCH model.

ACKNOWLEDGEMENTS

We would like to express our sincere thanks and gratitude to the anonymous reviewers for the valuable suggestions that helped us a lot in improving the quality of this manuscript.

REFERENCES

Anjoy, P. and Paul, R.K. (2017). Wavelet based hybrid approach for forecasting volatile potato price. *J. Ind. Soc. Agril. Statist.*, **71**(1), 7-14.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *J. Econometrics*, **31**, 307-327.

Boser, B.E., Guyon, I.M. and Vapnik, V.N. (1992). A training algorithm for optimal margin classifiers. *Proceedings of the fifth annual workshop on Computational learning theory,* 144-152.

Box, G.E.P., Jenkins, G.M. and Reinsel G.C. (2008). *Time-Series Analysis: Forecasting and Control,* 3rd edition. Pearson Education, India.

Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, **20**(3), 273-297.

Donaldson, R.G. and Kamstra, M. (1997). An artificial neural network-GARCH model for international stock return volatility. *J. Empirical Finance*, **4**(1), 17-46.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom in1ation. *Econometrica*. **50**, 987–1007.

Ghosh, H., Paul, R.K. and Prajneshu (2010). Nonlinear Time Series Modeling and Forecasting for Periodic and ARCH Effects. *J. Statist. Th. and Practice*, **4(1)**, 27-44.

Khashei, M. and Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Systems with applications*, **37(1)**, 479-489.

Khashei, M. and Bijari, M. (2012). A new class of hybrid models for time series forecasting. *Expert Systems with Applications*, **39(4)**, 4344-4357.

Kim, K.J. (2003). Financial time series forecasting using support vector machines. *Neurocomputing*, **55(1-2)**, 307-319.

Li, Y., Huang, W.P. and Zhang, J. (2013). Forecasting volatility in the Chinese stock market under model uncertainty. *Economic Modelling*, **35**, 231-234.

Mincer, J.A. and Zarnowitz, V. (1969). The evaluation of economic forecasts. *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance,* National Bureau of Economic Research, New York. Mitra, D. and Paul, R.K. (2017). Hybrid time-series models for forecasting agricultural commodity prices. *Model Assisted Statistics and Applications*, **12(3)**, 255-264.

Pagan, G.W. and Schwert, A.R. (1990). Alternative models for conditional stock volatility. *J. Econometrics*, **45**, 267-290.

Paul, R.K., Prajneshu, and Ghosh, H. (2009). GARCH Nonlinear Time Series Analysis for Modeling and Forecasting of India's Volatile Spices Export Data. *J. Ind. Soc. Agril. Statist.*, **63(2)**, 123-131.

Paul, R.K., Ghosh, H. and Prajneshu (2014). Development of out-of-sample forecast formulae for ARIMAX-GARCH model and their application. *J. Ind. Soc. Agril. Statist.*, **68(1)**, 85-92.

Paul, R.K. (2015). ARIMAX-GARCH-WAVELET Model for forecasting volatile data. *Model Assisted Statistics and Application*, **10(3)**, 243-252.

Paul, R.K**.,** Gurung, B., Paul, A.K. and Samanta, S. (2016). Long memory in conditional variance. *J. Ind. Soc. Agril. Statist.*, **70(3),** 243-254

Paul, R.K., Vennila, S., Bhat, M.N., Yadav, S.K., Sharma, V.K., Nisar, S. and Panwar, S. (2019). Prediction of early blight severity in tomato (Solanum lycopersicum) by machine learning technique. *Ind. J. Agril. Sci.*, **89(11)**, 169-175.

Zhang, G.P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, **50**, 159-175.

Zhang, G., Patuwo, B.E. and Hu, M.Y. (1998). Forecasting with artificial neural networks: The state of the art. *Int. J. Forecasting*, **14(1)**, 35-62.