Forecast Volatility based on Realized GARCH and deep LSTM Neural Network

Yingzhi Zhang*
School of Statistics and Mathematics
Zhejiang Gongshang University
Hangzhou, China
zyz199737@163.com

Ping Huang
School of Statistics and Mathematics
Zhejiang Gongshang University
Hangzhou, China

Ping Zhou
School of Statistics and Mathematics
Zhejiang Gongshang University
Hangzhou, China

Yanfang Wu
School of Statistics and Mathematics
Zhejiang Gongshang University
Hangzhou, China

2020 International Conference on Public Health and Data Science (ICPHDS) | 978-1-7281-8571-2/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICPHDS51617.2020.00028

***Abstract*—In order to overcome the defects of the existing models which can not truly reflect the long-term memory, non-stationary and non-linearity characteristics. This study combines Realized GARCH (RGARCH) model with the Long short-term Memory (LSTM) Neural Network to predict volatility. The prediction effect of different single models and mixed models on volatility in SZSE Component Index (SZSE) shows that the new hybrid model has the smallest error, and the hybrid model uses generalized realized measure is significantly improved. It confirms that the hybrid model based on generalized realized measure has high application value in the prediction of stock volatility.**

Keywords-Realized GARCH; generalized realized measures; deep LSTM model; volatility prediction

I. INTRODUCTION

Volatility is a key indicator for the financial time series. It is particularly important to be accurately predicted. GARCH models usually use low-frequency information to explore volatility, but the stock market fluctuates frequently, the GARCH models are likely to omitting important intraday information, especially the stock price changes significantly in a very short time. Hansen et al. introduce a measurement equation to combine daily log returns and realized measures to construct Realized GARCH model[1]. Huang et al. use the RGARCH model to analyze the option pricing S&P 500 index, demonstrate that RGARCH is more robust than the traditional GARCH models in fitting and predication[2]. Yu and Wang extend the RGARCH model to a mixed frequency Realized GARCH model (M-Realized GARCH) that includes intraday return, daily return, and realized volatility to compare the volatility of different models on the CSI 300 index, the evidence shows that M-Realized GARCH has higher prediction accuracy[3]. Jiang and Gu introduce four generalized realized measures to improve the performance of RGARCH model in fitting and volatility prediction[4]. However, the above models have many assumptions, and the description of non-stationary and non-linearity is not specific enough, so RGARCH models still have some limitations.

The deep learning has fewer assumptions than traditional models and is more useful in dealing with non-stationary, non-linear and highly autocorrelated financial data. Hajizadeh et al. compared EGARCH model and the model addition of Artificial Neural Network (ANN) in prediction the volatility of S&P 500 index[5]. Zhang and Zhang mix GARCH model with support vector machine (SVM), and the prediction of return rate on the model is more consistent with the actual situation[6]. Kim and Won use a hybrid model integrating LSTM with multiple-GARCH-type models to improve the accuracy of forecasting the KOSPI 200 index volatility[7]. Yang and Wang use deep LSTM, SVM, MLP and ARIMA in predicting 30 global stock over long, medium and short periods, demonstrates that the LSTM model has superior accuracy and stability[8]. Li et al. use five models to forecast the closing prices of CSI 300 index, it shows that adding LSTM to the ARIMA model can significantly improve the prediction accuracy of the model[9]. Ouyang et al. compare the prediction of closing price of DJIA index by LSTM, GARCH, MLP, K-means and SVM, it shows that LSTM is more accurate in predicting the trend of the stock price[10].

The above researches confirm that combining deep learning with traditional models can improve the prediction performance and describe the future trend more accurately. However, most of the studies combine the low-frequency GARCH models with the LSTM, and do not consider the impact of intraday high-frequency price changes. Therefore, compared with the original literature, this study adds the generalized realized measure RVaR into the RGARCH model, and constructs the hybrid model combined with the deep LSTM model, then compares the prediction performance of single and hybrid models on the volatility of SZSE index. The result shows that the new high-frequency LSTM+RGARCH model has strong applicability and superiority in the prediction of volatility.

II. COMPONENT MODELS

A. GARCH models and Realized GARCH model

1) GARCH-type models

GARCH-type models can effectively filter out the effects of autocorrelation caused by excess peaks and conditional heteroscedasticity. Therefore, it is effectively applied in financial

978-1-7281-8571-2/20/$31.00 ©2020 IEEE
DOI 10.1109/ICPHDS51617.2020.00028

time series. The structure of the GARCH (1,1) model is defined as follows:

$$\begin{cases} r_t = \sigma_t \varepsilon_t, \varepsilon_t \sim i.i.d \\ \sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{cases} \quad (1)$$

where $\omega \geq 0, \alpha_i \geq 0, \beta_j \geq 0, i = 1,2,\cdots,p, j = 1,2,\cdots,q$, σ_t^2 is the conditional variance, ε_t is the innovation sequence. The first equation is called the mean equation, and the second equation is called the conditional variance equation.

The TGARCH model developed on the basis of the GARCH model, that can reflect the asymmetric effects of positive and negative shocks on the conditional fluctuation of financial data. The variance equation of the TGARCH (1,1) model is:

$$\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \gamma \sigma_{t-1}^2 I_{t-1}(\varepsilon_{t-1} < 0) \quad (2)$$

where I_{t-1} is the indicative function. If $\varepsilon_{t-1} \geq 0$, $I_{t-1} = 0$. If $\varepsilon_{t-1} < 0$, $I_{t-1} = 1$. The leverage effect exists if $\gamma > 0$.

2) Realized GARCH model

The Realized GARCH model (Hansen et al., 2012) combines the logarithmic rate of return, conditional variance with the realized measures by introducing a measurement equation. The RGARCH (1,1) model is given by:

$$\begin{cases} r_t = \sigma_t z_t,\ z_t \sim i.i.d. \\ \ln \sigma_t^2 = \omega + \beta_1 \ln \sigma_{t-1}^2 + \gamma_1 x_{t-1} \\ \ln x_t = \zeta + \phi \ln \sigma_t^2 + \tau(z_t) + u_t, u_t \sim i.i.d.N(0, \sigma_u^2) \end{cases} \quad (3)$$

where r_t is the log return at time t , σ_t^2 is the conditional variance, x_t is the realized measure, z_t and u_t are independent sequence. $\tau(z_t) = \eta_1 z_t + \eta_2 (z_t^2 - 1)$ is the leverage function, η_1 and η_2 are opposite sign, which are used to reflect the asymmetric impact of rising and falling yields.

In order to consider the impact of market microstructure noise and price jump behavior on the tail risk estimation of the model, we adopt two common realized measures, RV and BPV, and a more generalized realized loss measure RVaR.

Realized Volatility (RV):it estimates the volatility in a certain period of time by the sum of squares of logarithmic returns.

$$RV_t = \sum_{j=1}^{n} r_{t,j}^2 \quad (4)$$

where $r_{t,j}$ is the log return within the interval j at time t , $n = 48$ is the number of daily samples. RV calculation is the most widely used in high-frequency data modeling. However, as the sampling interval expands, it is affected by structural noise and price jumps. It is a biased and inconsistent estimator. Realized Bipower Variance (BPV) it estimates the volatility by the sum of the product of the absolute values of two near returns, so it can reduce the influence of microstructure noise caused by the discontinuity of the extraction interval

$$BPV_t = \frac{\pi}{8}\left(\sum_j^n |r_{(k)t,j} + r_{(q)t,j}||r_{(k)t,i-1} + r_{(q)t,j-1}|\sigma_t^2 - \left|r_{(k)t,j} - r_{(q)t,j}\right|\left|r_{(k)t,i-1} - r_{(q)t,j-1}\right|\right) \quad (5)$$

where $r_{(k)t,j}$ is the k-th log return of $r_{t,j}$, BPV can correct price jump.

Realized Value at Risk (RVaR) is derived from the risk loss evaluation index VaR, and is a broadly realized loss measurement.

$$RVaR = \sqrt{n} z_p^t \quad (6)$$

where z_p^t is the p-th order quantile of the return sequence at time t . RVaR reduces the impact of intraday microstructure noise, and can more accurately find the extreme price and improve the predication performance.

B. Deep Long short-term Memory Neural Network

LSTM model is an improved model of recurrent neural network (RNN). It is an unsupervised learning method. LSTM has the advantages of standard RNN in processing nonlinear, non-stationary, and highly autocorrelated financial data. It overcomes the defects of gradient disappearance or explosion in the original model. In the process of forecasting time series, it not only improves the forecast accuracy of the model but also reduces the probability of over fitting in the forecast process. The standard LSTM model is composed of a memory cell (Cell) and three gates: Input Gate, Output Gate, and Forget Gate. The first two gates are used to receive, output, and modify parameters. The LSTM not only has an external loop between hidden layers, but also has an internal self-loop between memory units, which is more efficient for the extraction and updating of historical data, so as to improve the forecast performance of the volatility.

III. EXPERIMENTS WITH REAL DATA

A. Single model construction and empirical analysis

In this study, we select the representative index of SZSE from January 1, 2014 to July 24, 2020.A total of 1599 closing price samples, and choose 5-minute high-frequency data, so there are 76752 intraday samples, which is obtained from the JoinQuant database, the software used for analysis is R and Python.

The closing price series is non-stationary, so we calculate it by taking logarithmic difference. Set t as a specific trading day, p_t is the closing price of day t , then the log return of day t can be expressed as:

$$r_t = 100 * (\ln p_t - \ln p_{t-1}) \quad (7)$$

In this study, we use the 5-minute high-frequency data to estimate the real volatility.

$$\sigma_t = \sqrt{\sum_{i=1}^{48} r_{t,i}^2} \quad (8)$$

1) Data description

Table 1 is the description statistics of the log return and three different realized measures. From the perspective of financial time series, we can find that the skewness of the log return is less than zero, the kurtosis is 4.304, so it does not satisfy the normal distribution, and the J-B statistic also shows that the log return has the peak and heavy-tail characteristic of financial data, so it has practical significance to assume that the residuals obey the t distribution. Ljung-Box and Arch tests indicate the log return series has significant autocorrelation and heteroskedasticity, so using the GARCH-type models and the RGARCH model is reasonable. The three realized measures series also do not satisfy the characteristics of normal distribution, they have right-skewed and heavy-tail. The standard variance, kurtosis and skewness of the generalized realized measure RVaR are significantly smaller than the other two realized measures.

TABLE I. DESCRIPTIVE STATISTICAL ANALYSIS

	Mean	***Std***	***Skew***	***Kurt***	***J-B***	***LB***	***LM***
Logret	0.029	1.727	-0.924	4.304	<0.01	<0.01	<0.01
RV	2.494	5.227	7.231	73.837	<0.01	<0.01	<0.01
BPV	1.906	3.729	6.947	65.146	<0.01	<0.01	<0.01
RVaR	1.8	1.331	3.042	13.565	<0.01	<0.01	<0.01

Note: Logret is the log return series, J-B (Jarque-Bera) is the normal test, LB (Ljung-Box) is the autocorrelation test, and LM denotes the heteroskedasticity test.

2) Model Fitting

Table 2 shows the parameter estimation results of fitting the low-frequency GARCH, TGARCH model and the high-frequency RGARCH based on RV, BPV, RVaR.

From Table2, it can be seen that under the condition of t-distribution, all the parameters are significant at the level of 0.05 except for the coefficient reflecting the asymmetry in the TGARCH model. The β -value of the two low-frequency models is greater than 0.9, but it is lower in the RGARCH models. Because the volatility in the RGARCH model is not only influenced by the previous volatility, but also influenced by the previous realized measures, and the model $\gamma < \beta$ and $\gamma + \beta \approx 1$, indicating that the volatility is affected by the previous volatility more than the previous realized measures. The leverage coefficients in both the TGARCH and RGARCH models verify that the stock market reacts more strongly to return declines. The standard deviation σ_u of the residuals of the measurement equation is estimated to be 0.296 under the generalized realized measures RVaR, which is smaller than the residuals under RV and BPV, indicating that the measurement equation fits better under RVaR.

TABLE II. MODELS PARAMETER ESTIMATION

	GARCH	TGARCH	RG-RV	RG-BPV	RG-VaR
ω	0.026	0.022	0.210	0.334	-0.006
α	0.074	0.099	——	——	——
β	0.922	0.915	0.651	0.562	0.681
γ	——	0.162*	0.288	0.376	0.576
ζ	——	——	-0.654	-0.817	0.063
ϕ	——	——	1.117	1.075	0.502
η_1	——	——	-0.098	-0.132	-0.201
η_2	——	——	0.130	0.088	0.017
σ_u	——	——	0.534	0.503	0.296
n	4.687	4.719	5.335	5.447	5.965
LL	-2827.330	-2823.800	-2795.210	-2789.366	**-2786.671**

Note: RG is RGARCH model; LL is loglikelihood of the model, where the RGARCH model is a semi-loglikelihood; * means insignificant at the 0.05 level.

The loglikelihood value of RGARCH model is composed of the return and the residual of measurement equation, extract the semi loglikelihood value of RGARCH model to compare with the loglikelihood value of GARCH-type models. The loglikelihood values of all three RGARCH models are larger than GARCH-type models, it is verified that the RGARCH models with high-frequency information is more effective. Based on the generalized realized measure RVaR is the largest, so it has the best fitting effect; and the effect on BPV with robust intraday price jumps is second to none; RV performs slightly worse in the three RGARCH models, but is still better than traditional GARCH and TGARCH. TGARCH considers the asymmetric effect of price changing, so it is better than the traditional GARCH model.

3) Single deep LSTM model

Since the volatility of stock index is persistent and long-memory, this paper uses the log return of this stock in the previous 10 trading days and intraday volatility information to predict the next volatility. In order to reduce the gradient disappearance and overfitting problems in deep learning, Dropout technique is used after each layer of LSTM model to accelerate the model learning rate. The LSTM structures were stacked three times. The first 1399 data are used as the training set for the three-layer LSTM model, and 20% of the training interval is used as the validation set for parameter optimization. The data need to be normalized and mapped to (0,1) intervals to improve the prediction performance of the deep learning.

After iterative tests set the number of nodes in the Dense layer as 5 and 1, the number of hidden nodes in each layer of LSTM as 25, 10, 5. Moreover, the parameters of the Dropout layer as 0.2, 0.5, 0.5; select Adam as the optimizer, set the linear as the activation function and the learning rate as 0.004, use MSE as the loss function; set epochs=140, batch size=20. The trained model is used to make predictions about the test set data and to compare the predicted volatility with the realized volatility.

4) Volatility Prediction of the single model

The 1599 data are divided into two parts, the first part is the preceding 1399 data, which is set as the training data. The second part is the last 200 data as the model test data. A single model is trained by the training data, then using the test data to predict the volatility of the trained model. Four different loss functions are utilized as model evaluation. The formulas are as follows:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(f_i - r_i)^2 \quad (9)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|f_i - r_i\right| \quad (10)$$

$$MSPE=\frac{1}{n}\sum_{i=1}^{n}(\frac{f_i - r_i}{r_i})^2 \quad (11)$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{f_i - r_i}{r_i}\right| \quad (12)$$

where f_i and r_i are the predicted and true volatility at time i, and n is the number of predicted days. The smaller the value of the four loss functions, the more accurate in predicting the volatility.

Table 3 compares the volatility prediction results for 200 trading days in the test set under different single models.

TABLE III. COMPARISON OF PREDICTION RESULTS OF SINGLE MODELS

Model	*GARCH*	*TG*	*RG-RV*	*RG-BPV*	*RG-RVaR*	*LSTM*
MSE	0.7538	0.7100	0.4523	0.5047	0.5474	0.6511
	(6)	(5)	(1)	(2)	(3)	(4)
MAE	0.5668	0.5483	0.4248	0.4207	0.4109	0.4119
	(6)	(5)	(4)	(3)	(1)	(2)
MSPE	0.4538	0.4220	0.2276	0.2136	0.1976	0.1587
	(6)	(5)	(4)	(3)	(2)	(1)
MAPE	0.5461	0.5285	0.4087	0.3985	0.3738	0.3113
	(6)	(5)	(4)	(3)	(2)	(1)
Rank	6	5	4	3	1	1

Note: Numbers in parentheses is the model prediction accuracy rankings; Rank is the sample mean of ranks；RG is RGARCH model; TG is TGARCH model.

According to the four loss functions, the deep LSTM model and RGARCH model based on generalized realized loss measure RVaR have the highest prediction accuracy under six different single models, which verifies that LSTM model and RGARCH model with RVaR have outstanding effect on the prediction of financial time series volatility. BPV has the next best prediction accuracy due to its robustness to stock price jumps. The RV is the most deeply affected by market price changes and microstructure, so it has a low prediction accuracy among the three realized measures, but is still stronger than the GARCH-type models. Consider the combination of deep LSTM and high-frequency RGARCH model to improve the prediction accuracy of intraday volatility of stock indices has practical significance.

B. *Volatility Prediction of the hybrid model*

In order to improve the prediction accuracy of the intraday volatility of SZSE index, this study construct a high-frequency LSTM+RGARCH hybrid model by combining RGARCH with LSTM model.

Subsamples of size 1399 to fit RGARCH model and the estimation results related to intraday volatility are taken as new explanatory variables. The parameter estimations of new input variables are shown in Table4.

TABLE IV. PARAMETER ESTIMATION OF NEW INPUT VARIABLES.

Model	$\omega+\gamma\zeta$	$\ln\sigma_t^2$	$\tau(z_t)+u_t$
RGARCH-RV	0.0187	0.9752	0.2881
RGARCH-BPV	0.0215	0.971	0.3489
RGARCH-RVaR	0.0165	0.9776	0.5215

In constructing hybrid models, we take the above RGARCH models parameters estimation as additional feature variables and input them into LSTM model to construct three different hybrid models: LSTM+RG-RV, LSTM+RG-BPV and LSTM+RG-RVaR. These models are used to predict the daily volatility after training. The Table 5 shows the comparison between the hybrid models and the LSTM model which has better prediction effect in single models.

TABLE V. THE PREDICTIONS OF THE HYBRID MODELS.

Model	*LSTM*	*LSTM+ RG-RV*	*LSTM+ RG-BPV*	*LSTM+ RG-RVaR*
MSE	0.6511	0.6181	0.5722	0.563
	(4)	(3)	(2)	(1)
MAE	0.4119	0.3843	0.3817	0.3564
	(4)	(3)	(2)	(1)
MSPE	0.1587	0.128	0.1343	0.0981
	(4)	(2)	(3)	(1)
MAPE	0.3113	0.2875	0.2882	0.2525
	(4)	(2)	(3)	(1)
Rank	4	2	2	1

Note: Numbers in parentheses is the model prediction accuracy rankings; Rank is the sample mean of ranks; RG is RGARCH model; TG is TGARCH model.

From the four loss functions in Table 5, it can be concluded that the performance of hybrid models predict intraday volatility are better than a single LSTM model, and superior to a single RGARCH and GARCH-type models. Among the three high-frequency hybrid models, the LSTM+RG-RVaR model constructed based on the generalized realized measure RVaR has the best average ranking and better prediction performance. And compared to a single deep LSTM neural network, the four loss function values of MSE, MAE, MSPE and MAPE decreased by 13.53%, 13.47%, 38.19% and 18.89%. This suggests that adding the estimated parameters of the RGARCH-RVaR model to the LSTM model provides better access to the asymmetric and persistent characteristics of financial data. RVaR is more sensitive than the other realized measures to predicting extreme situations. Because it considers the impact of extreme price and volatility changes. The sequence of the predicted volatility of the hybrid model and the real volatility are shown in Figure 1.

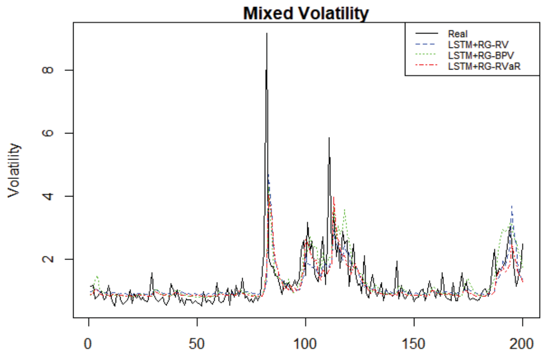

Figure 1. Comparison of hybrid models prediction

From the trends of the curves in Figure 1, we can see that when the stock market is running smoothly and is not affected by a major economic crisis, the error between the predicted volatility and actual volatility of each model is relatively small. However, when the stock market fluctuates sharply, the prediction errors of the hybrid models are increased, and the volatility has the possible to be underestimated, the LSTM+RG-RVaR model has a more consistent trend between the predicted and real volatility, and the distance between the two curves is closer, indicating that the LSTM+RGARCH-RVaR hybrid model can make more accurate predictions when the stock market fluctuates to different degrees.

IV. Conclusions

Based on Jiang and Gu, this study introduces generalized realized measures to improve the fitting and prediction ability of RGARCH model, and then mix the model with the LSTM Neural Network[4]. This model contains intraday high-frequency trading information and can reflect the asymmetry, continuous volatility, and heteroscedasticity of financial time series. Use this model to predict the volatility of the SZSE component index. The final conclusions are as follows:

Compared with the traditional low-frequency GARCH-type models, the RGARCH model with intraday high-frequency information has better fitting and forecasting performance. In terms of the loglikelihood value and the prediction of intraday volatility, the RGARCH model that introduces the generalized realized measure RVaR is more prominent than that of RV and BPV, indicating that the RGARCH model with RVaR is an effective financial sequence fitting and forecasting tool.

In this study, the LSTM+RGARCH-RVaR hybrid model improves the prediction accuracy of the single model in the intraday volatility of the financial market, indicating that the hybrid model has a higher application value in the prediction of the volatility of of financial time series.

References

[1] Hansen, P.R., Huang, Z., Shek, H.H. (2012) Realized GARCH: a joint model for returns and realized measures of volatility[J]. Journal of Applied Econometrics, 27(6):877–906.

[2] Huang, Z., Wang, T.Y., Hansen, P.R. (2017) Option pricing with the realized GARCH model: An analytical approximation approach. Journal of Futures Markets, 37(4): 65–84.

[3] Yu, X.J.,Wang, X.H.(2018) Mix frequency realized GARCH models: the forecast of volatility and measure of VaR. Statistical Research, 35(1): 104–116.

[4] Jiang, W., Gu, Y. (2019) Forecasting VaR: realized GARCH incorporating generalized realized measures. Journal of Quantitative & Technical Economics, 36(7): 156–173.

[5] Hajizadeh, E., Seifi, A., Fazel Zarandi, M.H., Turksen, I.B. (2011) A hybrid modeling approach for forecasting the volatility of S&P 500 index return. Expert Systems with Applications, 39(1): 431–436.

[6] Zhang, G.S., Zhang, X.D. (2016) A SVM-GARCH model for stock price forecasting based on neighborhood mutual information.Chinese Journal of Management Science, 24(9):11–20.

[7] Kim, H.Y., Won, C.H. (2018) Forecasting the volatility of stock price index: A hybrid model integrating LSTM with multiple GARCH-type models. Expert Systems With Applications, 103: 25–37.

[8] Yang, Q.,Wang, C.W. (2019) A study on forecast of global stock indices based on deep LSTM Neural Network. Statistical Research, 36(3): 65–77.

[9] Li, Z.W., Han, J., Song, Y.P. (2020) On the forecasting of high-frequency financial time series based on ARIMA model improved by deep learning. Journal of Forecasting, 39(7):1081–1097.

[10] Ou Yang, H.B., Huang, K., Yan, H.J. (2020) Prediction of financial time series based on LSTM Neural Network. Chinese Journal of Management Science, 28(4): 27–35.