Proceedings of Business and Economic Studies, 2022, Volume 5, Issue 6
http://ojs.bbwpublisher.com/index.php/PBES
ISSN Online: 2209-265X
ISSN Print: 2209-2641

# Volatility Prediction via Hybrid LSTM Models with GARCH Type Parameters

**Mingyu Liu[1], Jing Ye[2]*, Lijie Yu[1]**

[1]College of Business and Public Management, Wenzhou-Kean University, Wenzhou 325060, Zhejiang Province, China

[2]College of Science and Technology, Wenzhou-Kean University, Wenzhou 325060, Zhejiang Province, China

****Corresponding author:*** Jing Ye, yejin@kean.edu



**Abstract:** Since the establishment of financial models for risk prediction, the measurement of volatility at risky market has improved, and its significance has also grown. For high-frequency financial data, the degree of investment risk, which has always been the focus of attention, is measured by the variance of residual sequence obtained following model regression. By integrating the long short-term memory (LSTM) model with multiple generalized autoregressive conditional heteroscedasticity (GARCH) models, a new hybrid LSTM model is used to predict stock price volatility. In this paper, three GARCH models are used, and the model that can best fit the data is determined.

**Keywords:** Time series; Exchange rate forecast; GARCH model; Stock market volatility; Error

***Online publication:*** December 9, 2022

## 1. Introduction

Since the establishment of KOSPI 200 options market on July 7, 1997, there has been a significant increase in the volume of trade, alerting investors and financial institutions of the risks brought about by the increased volatility of KOSPI 200. Accurate volatility prediction is an important aspect of risk management, especially when allocating assets to various portfolios to effectively hedge the risks of these portfolios. It has been shown that a low level of financial management expertise, a large amount of financial assets, and a high time opportunity cost could increase the perceived value of information intermediaries, thereby raising the possibility of using information intermediaries. We have also found that the use of information intermediaries is positively correlated with the overall scope of information search and affects the likelihood of using other sources of information. In section 2, we explain the generalized autoregressive conditional heteroskedasticity (GARCH) financial time series model and the hybrid long short-term memory (LSTM) model involving the artificial neural network (ANN) and multiple GARCH models. We explain how the experiment is conducted and present the experimental results and analysis in sections 3 and 4, respectively. Section 5 provides several conclusions to this study and the forecast of future research.

## 2. Literature review

The volatility of asset market, like the stock market, is a measurement of the degree to which asset prices fluctuate and the degree of uncertainty in forecasting. Investment firms and private investors measure the instabilities of underlying asset prices [1]. In order to determine the risk of investment profile, it is essential to determine the volatility of constituent assets. Volatility is significant not only to evaluate such intangible derivative goods as stock index options but also in price detection. In such cases, volatility is a deductive topic within different types of financial time series models. In order to deduce the future trends with the heteroscedastic condition of past information, an autoregressive conditional heteroskedasticity (ARCH) model that uses the present error term as a function of the previous time has been proposed [2]. By using the proposed GARCH model, it is feasible to reduce the estimated number of parameters [3]. When the volatility is high, a high volatile state is more likely to maintain to some extent, but when the volatility is low, a low volatility state is maintained up to a certain threshold. In addition, Morgan and Reuters have put forward the exponentially weighted moving average (EWMA) model based on GARCH(1,1) [4]. The former emphasizes recent data more than other models. It clearly forecasts recent changes and is less influenced by the quantity of data. A number of scholars have used time series [5-7] and deep learning models [8,9] to estimate stock prices and stock return movements.

These econometric methods are theoretically described based on statistical charts and data. However, if qualitative variables are used, the stability of model-based predictions will be significantly reduced. In order to avoid restrictions and assumptions on the model, ANN, which is more stable and connected, is considered. The comparison of the two models shows that the density estimation neural network without specific target distribution has better performance than the multi-layer perceptron. Therefore, in volatility prediction, the nonlinear characteristics that cannot be captured by econometric models can be clearly extracted using the neural network-based method, which has been proven to be successful. Therefore, the LSTM model can be used by the recurrent neural network (RNN) to forecast exchange rates and foreign exchange. LSTM is obviously superior to the feedforward neural network model as a functioning financial time model. This paper studies the model combining neural network with econometric model and the single neural network model for forecasting volatility. By taking into account of these characteristics and analyzing the volatility of S&P 500 index futures options, it has been found that ANN can surpass the financial time series model [10]. By forecasting the volatility of financial time series, the feasibility of using ANN has been proven [11].

## 3. Materials and methods

### 3.1. Data

This study takes the volatility of KOSPI 200 index in South Korea as the research subject, tracking the market value of 200 stocks represented by South Korea as the fundamental analysis target. The data is derived from the Data Guide and consists of 4922 data points from January 1, 2000, to August 30, 2020. It is estimated that there will be 3315 data points during this process. The interest rate of three-year Korean Treasury Bond (KTB) and three-year AA corporate bond (CB) is based on the daily data provided by the Korean Asset Management Corporation. The daily closing prices of gold and crude oil are derived from Bloomberg during the same period. We integrate the hybrid GARCH-exponential (E)GARCH-EWMA model and the hybrid GARCH-EGARCH model into the LSTM model, respectively, and compare the error and volatility between them to determine which of those two is a better model.

### 3.2. Financial time series models

#### 3.2.1. Generalized autoregressive conditional heteroscedasticity (1,1) model

In order to understand the GARCH model, we must first look at the ARCH model. The ARCH model plays a significant role in predicting risks in the short-term period and aims to identify economic variables that can be used to predict volatility. It is assumed that the conditional variance is constant, and the estimator is unbiased; however, it cannot be used for valid estimation. Therefore, it cannot be used to acquire the settings of the confidence interval and commonly used tests. In 1982, Engle proposed the ARCH(p) model, which is a conditional heteroscedasticity model, to model the volatility of the conditional distribution. The ARCH model is as follows:

$$y_t = \mu_t + \sigma_t \eta_t,\ \eta_t \sim N\ (0,1) \tag{1}$$

$$\varepsilon_t = \sigma_t \eta_t,\ \varepsilon_t | \chi_{t-1} \sim N(0, \sigma_t^2) \tag{2}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 \tag{3}$$

Equation (1) comprises $\mu_t$, which can be predicted by the average equation for the $y_t$ time series, and the unpredictable error term $\varepsilon_t$. Equation (2) states that the error $\varepsilon_t$ follows a normal distribution when time (*t*–1) is known. Equation (3) states that the conditional variance depends on the past square of the errors.

ARCH is circumscribed because it cannot avoid too many parameters when the p-value is too large nor explain the leverage effect on financial time series. In 1986, Bollerslev established the GARCH model. The GARCH(p, q) model is as follows:

$$y_t = \mu_t + \sigma_t \eta_t,\ \eta_t \sim N\ (0,1) \tag{4}$$

$$\varepsilon_t = \sigma_t \eta_t,\ \varepsilon_t | \chi_{t-1} \sim N(0, \sigma_t^2) \tag{5}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2 \tag{6}$$

The GARCH model is the extension of the ARCH model. The first two equations of the GARCH model are as the same as those in the ARCH model. Equation (6), however, indicates the square of residuals and conditional variance in the GARCH model. This equation is significant, as it has helped to solve the problem of excessive parameters when the p-value is too large.

When p = 1 and q = 1, the GARCH(1,1) model is as follows:

$$\alpha_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{7}$$

We can then infer the following:

$$\sigma_t^2 = \frac{\alpha_0}{(1-\beta_1)} + \alpha_1 \sum_{j=1}^{\infty} \beta_1^{j-1}\, \varepsilon_{t-j}^2 \tag{8}$$

The GARCH model enhances the ARCH model so that there will be less parameters; thus, it will be easier to calculate. It predicts the volatility of the current period of KOSPI 200 index returns from the past, but it still has its limits. It has to satisfy the condition that every coefficient must be nonnegative. In addition, it cannot measure the leverage effect in financial pricing. However, the EGARCH model can explain the leverage effect on financial series.

### 3.2.2. Exponential generalized autoregressive conditional heteroscedasticity model

The EGARCH model, which was innovated by Nelson in 1991, considers the influence of the leverage effect on financial series. More importantly, it does not have to satisfy the condition in the GARCH model that "every coefficient must be nonnegative" because this model also defines the logarithm of conditional variance. Unlike the ARCH and GARCH models, the EGARCH model cannot solve the problems of negative future volatility. The EGARCH model is as follows:

$$r_t = X_t M + \varepsilon_t \tag{9}$$

$$ln\sigma_t^2 = \alpha_0' + \beta ln\sigma_{t-1}^2 + \omega\left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right) + \gamma\left|\frac{\varepsilon_{t-1}}{\sigma_{t-1}}\right| \tag{10}$$

In equation (9), $X_t$ is the explanatory variable, $M$ is the parameter, and $\varepsilon_t$ is the residual term. Equation (10) states the logarithm of conditional variance. This model assumes the conditional variance is positive even if the parameters are negative.

### 3.2.3. Exponentially weighted moving average model

The EWMA model is a model that can reduce the influence of actual data in the past. The model is as follows:

$$\sigma_t^2 = \rho\sigma_{t-1}^2 + (1-\rho)\varepsilon_{t-1}^2 \quad (0 < \rho < 1) \tag{11}$$

In equation (11), we can see that the variable $\rho$ has significant influence on the autoregressive moving average (ARMA) model. When $\rho$ is infinitely close to one, there will be almost no influence from past information. This is not significant because it will only reflect recent information. However, if $\rho$ is too small, there will be excessive past information and the residual term will be large. Hence, it will be better if $\rho$ is large and not too close to 1. In that way, the accuracy is higher, and the data in the past would not be avoided.

### 3.2.4. Long short-term memory

LSTM is a type of recurrent neural network. Traditional RNNs use backpropagation through time. Hence, over a long period of time, the residual term will decrease exponentially; thus, the long-term memory effect of RNNs will not be reflected. In view of that, Hochreiter and Schmidhuber proposed LSTMs that can store memories and catch necessary information, while ignoring unnecessary information.

$$g_t = \sigma(U_g x_t + W_g h_{t-1} + b_f) \tag{12}$$

$$i_t = \sigma(U_i x_t + W_i h_{t-1} + b_i) \tag{13}$$

$$\widetilde{c_t} = \tanh\left(U_c x_t + W_c h_{t-1} + b_c\right) \tag{14}$$

$$c_t = g_t * c_{t-1} + i_t * \widetilde{c_t} \tag{15}$$

$$o_t = \sigma(U_o x_t + W_o h_{t-1} + b_o) \tag{16}$$

$$h_t = o_t * \tanh\left(c_t\right) \tag{17}$$

LSTM consists of memory blocks. As shown in **Figure 1**, LSTM includes a memory cell and three gates: an input gate ($i_t$), a forget gate ($g_t$), and an output gate($o_t$). Equation (13) is an input gate equation. When the sigmoid function is 0, there is no input getting through the gate; when it is 1, all information will pass through the gate. Equation (12) demonstrates the weighted average of $x_t$ and $h_{t-1}$, which will be too large if the input gate is 1. When the input value is large, the gradient of tanh and sigmoid function will basically disappear. Hence, a forget gate is added to eliminate the information in memory. The main innovation of LSTM is its storage unit $c_t$, which essentially acts as the accumulator of state information.

The unit is accessed, written, and cleared by several self-parameterized control gates. Whenever a new input arrives, if the input gate $i_t$ is activated during *t*, its information will be accumulated into the unit. In addition, if the forget gate is on, then the past cell status $c_{t-1}$ may be "forgotten" in the process. When the latest unit outputs $c_t$, it will be propagated to the final state $h_t$, and further controlled through the output gate $o_t$. One advantage of using memory cells and gates to control the information flow is that the gradient will be captured in the cell (also known as the constant error turntable) and prevented from disappearing too quickly, which is a key problem of ordinary RNN models.

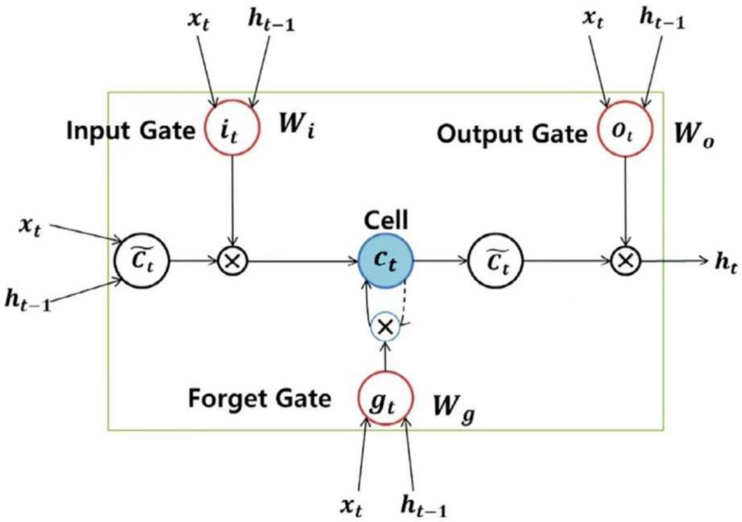


**Figure 1.** Process of long short-term memory

### 3.3. Measurement and statistical test

#### 3.3.1. Realized volatility

Realized volatility is also historical volatility. Realized volatility measures the volatility of stock prices in one day. Realized volatility at time *t* is calculated as follows:

$$RV_t = \sqrt{\frac{1}{\rho_t}\sum_{t=1}^{\rho_t}(S_t - \bar{S}_t)^2} \tag{18}$$

In equation (18), $\rho_t$ is the days remaining after *t*, $s_t$ is the logarithmic return at time *t*, and $s_t$ is the average return of the log return at time *t* days, during the time period of $\rho_t$ after time *t*.

#### 3.3.2. Loss functions

There are four types of loss functions. Mean absolute error (MAE) is the mean of all absolute values of all errors and is unaffected by outliers. Mean square error (MSE) is the average of the square of the difference between the real value and the predicted value. These two are basic, and the equations of MAE and MSE are as follows:

$$MAE = \frac{1}{T}\sum|\widehat{v}_t - RV_t| \tag{19}$$

$$MSE = \frac{1}{T}\sum(\widehat{v}_t - RV_t)^2 \tag{20}$$

Heteroskedasticity adjusted MAE (HMAE) and heteroskedasticity adjusted MSE (HMSE) are forms of MAE and MSE adjusted through heteroscedasticity, respectively. They are non-linear loss measurements. The functions of HMAE and HMSE are as follows:

$$HMAE = \frac{1}{T}\sum|1 - \widehat{v_t}/RV_t| \tag{21}$$

$$HMAE = \frac{1}{T}\sum(1 - \widehat{v_t}/RV_t)^2 \tag{22}$$

$\widehat{v_t}$ is the predicted volatility at time $t$, $RV_t$ is the realized volatility at time $t$, and $T$ is of population predictions.

## 4. Experiment

The steps of our experiment are as follows: first, the hybrid model consisting of three models (GARCH, EGARCH, and EWMA) integrated into LSTM is analyzed to predict volatility; second, the hybrid model of GARCH-EGARCH integrated into LSTM is used. These two models help us to compare and predict volatility. There are four types of errors that are crucial, including MAE, MSE, HMAE, and HMSE. In the experiment, we also use the Durbin Watson (DW) and Wehner Schulze (WS) tests, whose purpose is to ensure a more accurate prediction.

In total, there are 4922 KOSPI 200 index data points between January 5, 2000, and August 21,2020 in the following three parameters: GARCH, EGARCH, and EWMA. The actual train size is 3315, which is taken from this range. We measure the number of classes, number of layers, input size, and hidden sizes, all of which can influence the LSTM model. The train size and the data of predicts affect the errors, including MAE, MSE, HMAE, and HMSE. By identifying the specific errors and learning about the ability of this model, we can predict the time series.

## 5. Results and discussion

**Figure 2** illustrates the fact that the larger the epochs, the lower the loss. The data in **Table 1** correspond to **Figure 2**. We can reach a conclusion that when the loss is 0.00376, the epoch set is approximately 180, suggesting that when we have a larger epoch set, the loss will be nearer to 0.

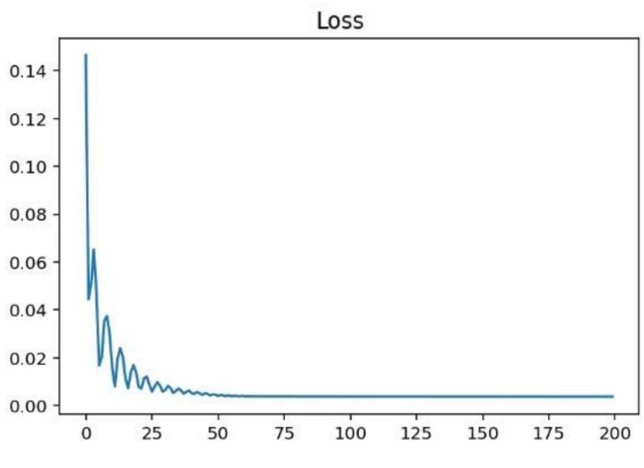


**Figure 2.** Relationship between epochs and loss in the hybrid model of GARCH, EGARCH, and EWMA integrated into LSTM

**Table 1.** Errors of the hybrid model of GARCH, EGARCH, and EWMA integrated into LSTM

| Epoch | Loss |
|---|---|
| 0 | 0.14644 |
| 20 | 0.00779 |
| 40 | 0.00517 |
| 60 | 0.00390 |
| 80 | 0.00381 |
| 100 | 0.00379 |
| 120 | 0.00378 |
| 140 | 0.00377 |
| 160 | 0.00377 |
| 180 | 0.00376 |

**Figure 3** shows the volatility predicted by the hybrid model. This hybrid model consists of GARCH, EGARCH, and EWMA integrated into LSTM. The x-axis and y-axis represent the total days and the realized volatility, respectively; the orange and blue lines represent the training volatility and the time series of realized volatility, respectively. When the days increase, the training volatility is nearer to the time series of volatility.

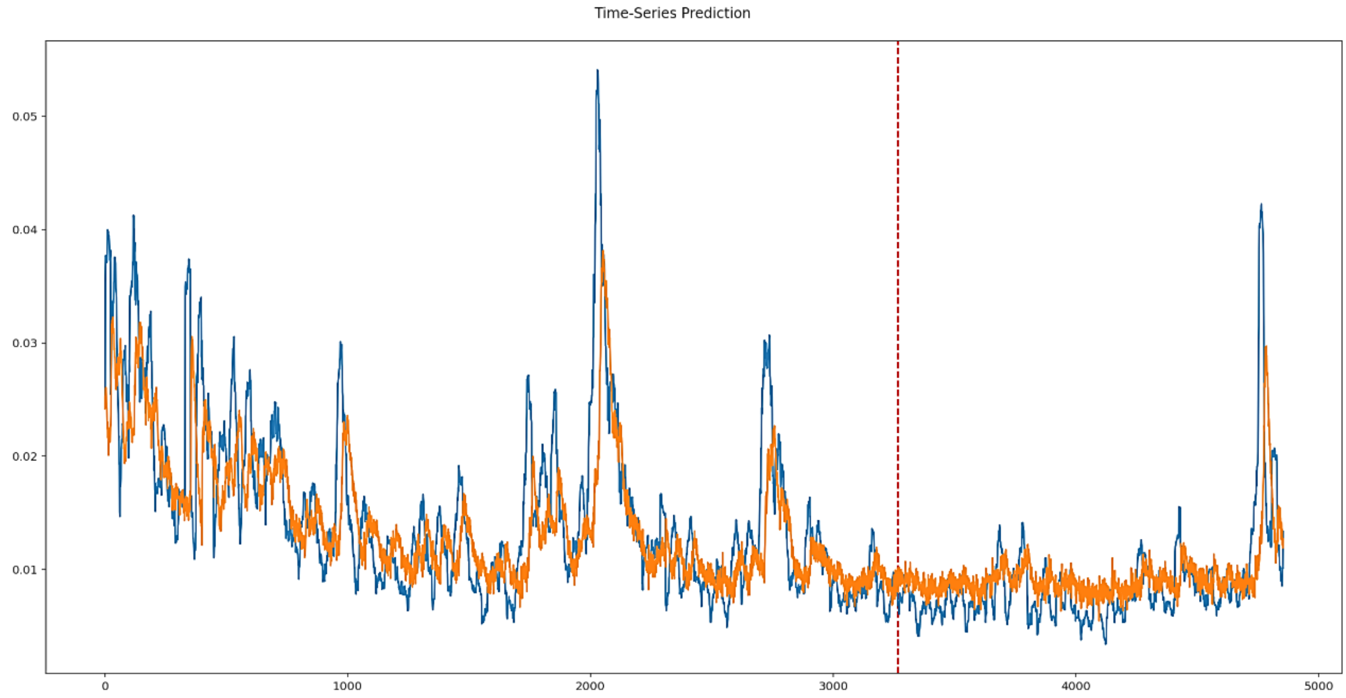


**Figure 3.** Volatility of the hybrid model consisting of GARCH, EGARCH, and EWMA integrated into LSTM

We use four different kinds of loss functions, including MAE, MSE, HMAE, and HMSE. From **Tables 2** and **3**, we can clearly see that the hybrid model of GARCH, EGARCH, and EWMA integrated into LSTM has a larger error compared to the hybrid model of GARCH and EGARCH integrated into LSTM. The overall MAE, MSE, HMAE, and HMSE of the former are 33.71%, 28.03%, 26.05%, and 14.47%, respectively. The latter model is better because its overall MAE, MSE, HMAE, and HMSE are 33.12%, 27.34%, 25.58%, and 14.06%, respectively. Additionally, both the test volatility and the train volatility of MAE, MSE, HMAE, and HMSE in the latter model are better compared to the former. **Tables 2** and **4** are summary tables. Evidently, the latter model has less error in every single term (overall, train, and test) and better performance in terms of accuracy. Hence, we can reach the conclusion that the hybrid model of GARCH and EGARCH integrated into LSTM is a better model that has less errors, suggesting that it has a better capability in predicting time series.

**Table 2.** Errors of the hybrid model of GARCH, EGARCH, and EWMA integrated into LSTM

| | **GARCH + EGARCH + EWMA + LSTM** | | |
|---|---|---|---|
| **Error type** | **Test** | **Train** | **Overall** |
| MAE | 0.25172916 | 0.37858 | 0.33706298 |
| MSE | 0.18135291 | 0.32844272 | 0.28030145 |
| HMAE | 0.24960087 | 0.26586163 | 0.2605397 |
| HMSE | 0.13646944 | 0.14871801 | 0.14470914 |

**Table 3.** Errors of the hybrid model of GARCH and EGARCH integrated into LSTM

| **Epoch** | **Loss** |
|---|---|
| 0 | 0.05432 |
| 20 | 0.00855 |
| 40 | 0.00468 |
| 60 | 0.00381 |
| 80 | 0.00379 |
| 100 | 0.00378 |
| 120 | 0.00378 |
| 140 | 0.00377 |
| 160 | 0.00376 |
| 180 | 0.00376 |

**Table 4.** Errors of the hybrid model of GARCH and EGARCH integrated into LSTM

| | GARCH + EGARCH + LSTM | | |
|---|---|---|---|
| **Error type** | **Test** | **Train** | **Overall** |
| MAE | 0.24651265 | 0.37245804 | 0.33123732 |
| MSE | 0.17731652 | 0.3201486 | 0.27340057 |
| HMAE | 0.246911 | 0.26009637 | 0.2557806 |
| HMSE | 0.13532783 | 0.14317173 | 0.14060462 |

**Figures 3** and **4** correspond to **Tables 1** and **3**, respectively. The former represents the hybrid model of GARCH, EGARCH, and EWMA integrated into LSTM, while the latter represents the hybrid model of GARCH and EGARCH integrated into LSTM. Comparing **Table 1** with **Table 3**, the similarity is that when epochs increase, loss decreases. This shows that increasing epochs will reduce the errors, thus making it more accurate. As shown in **Tables 1** and **3**, loss decreases from 0.14644 to 0.00376 when epochs increase from 0 to 180 in the former model, whereas in the latter model, loss decreases from 0.05432 to 0.00370 when epochs increase from 0 to 180. Based on these data, it is evident that the loss of the latter model is less than in the former model. Therefore, there is less error in the hybrid model of GARCH and EGARCH integrated into LSTM.

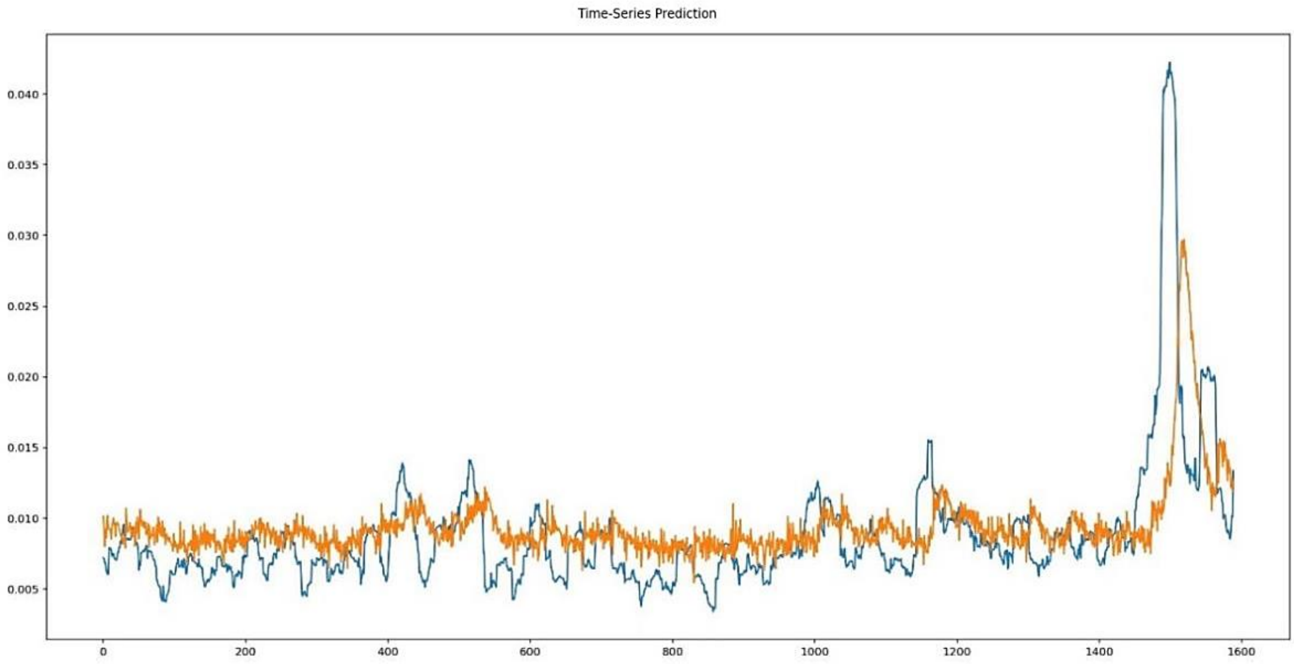


**Figure 4.** Inverse of the time series prediction

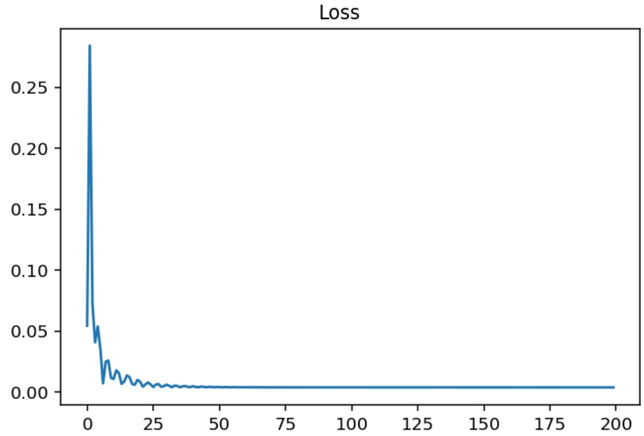


**Figure 5.** Relationship between epochs and loss in the hybrid model of GARCH and EGARCH integrated into LSTM

## 6. Conclusion

By comparing several composite models at different stages, it can be concluded that the hybrid model of GARCH and EGARCH integrated into LSTM is a better model with smaller error as it has better time series prediction ability. With this model, investors will be able to make rational investments based on their actual ability and risk preference, reduce risk as much as possible while ensuring a certain return, and obtain the highest return at the same risk level within the observation range of LSTM model portfolio.

It should be noted that not all securities portfolios are situated on the effective boundary and the risk and return are not always positively correlated as well. There are some differences under the target prediction and actual boundary. Therefore, investors must be prudent and use scientific concepts when choosing securities, judging investment proportion, and analyzing the market environment. In addition, they should avoid blindly pursuing returns, while neglecting the existence of risks. Due to the limitations of this model, this theory should not be applied blindly to investment decisions. While giving full play to investment diversification, investors should make appropriate corrections and adjustments according to the actual situation to avoid unnecessary investment losses.

## Acknowledgement

We would like to express our gratitude to the scholars referenced in this paper and sincerely thank all the experts and editors who participated in the review and revision of this paper.

## Disclosure statement

The authors declare no conflict of interest.

## References

[1] Markowitz H, 1952, Portfolio Selection. The Journal of Finance, 7(1): 77–91.

[2] Engle RF, 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. Econometrica: Journal of the Econometric Society, 50(4): 987–1007.

[3] Bollerslev T, 1986, Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, 31(3): 307–327.

[4] Morgan JP, 1997, Creditmetrics-Technical Document, JP Morgan, New York.

[5] Jin B, Gao S, Tao Z, 2022, ARIMA and Facebook Prophet Model in Google Stock Price Prediction. Proceedings of Business and Economic Studies, 5(5): 60–66.

[6] Weng Q, Liu R, Tao Z, 2022, Forecasting Tesla's Stock Price Using the ARIMA Model. Proceedings of Business and Economic Studies, 5(5): 38–45.

[7] Hu Y, Tao Z, Xing D, et al., 2020, Research on Stock Returns Forecast of the Four Major Banks Based on ARMA and GARCH Model. Journal of Physics: Conference Series, 1616(1): 012075.

[8] Zhu T, Liao Y, Tao Z, 2022, Predicting Google's Stock Price with LSTM Model. Proceedings of Business and Economic Studies, 5(5): 82–87.

[9] Qin J, Tao Z, Huang S, et al., 2021, Proceedings of 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), March 26, 2021: Stock Price Forecast Based on ARIMA Model and BP Neural Network Model. IEEE Publisher, Nanchang, China, 426–430.

[10] Hamid SA, Iqbal Z, 2004, Using Neural Networks for Forecasting Volatility of S&P 500 Index Futures Prices. Journal of Business Research, 57(10): 1116–1125.

[11] Brooks C, 1998, Predicting Stock Index Volatility: Can Market Volume Help?. Journal of Forecasting, 17: 59–80.