

Oil Price Forecasting Model Based on GARCH-LSTM Model

Xiaojing Fu^{1, *}

¹Institute of Problem Solving, Hunan Normal University, Changsha, 410081, China

*Corresponding author's e-mail: fxj_7222@163.com

Abstract: Based on the importance of oil in daily life and economic development, the price prediction of oil has become a pressing problem. Previously, the methods for oil price prediction were broadly divided into two categories, respectively, statistical-based prediction, which has more reliable mathematical principles, but the prediction effect is not very ideal, and deep learning methods, which use neural networks for prediction, although the prediction effect is more ideal, but there are still many mathematical features in the data have not been completely extracted. Therefore, this paper combines the statistical method with the deep learning method, the statistical method selects the GARCH family model, and the deep learning method selects the LSTM model, and establishes the GARCH-LSTM hybrid model for oil price prediction. The experimental data are selected from Brent crude oil from January 2000 to December 2022. After fitting the data using different models, by calculating their corresponding fitting accuracy, it can be found that the fitting accuracy of the GARCH-LSTM model is significantly better than other models, and its RMSE corresponds to 0.148% and MAE to 1.378%. The model can be trained both by the mathematical features in the data through the GARCH model and in depth using the LSTM neural network to obtain more excellent training results.

Keywords: Oil price forecasting, GARCH family model, LSTM model.

1. Introduction

Oil is known as the "blood of industry" and is an extremely important industrial raw material and power source. Oil and its derivatives play an important role in the country's industry, military and people's food, clothing and housing, so oil price fluctuations can have a certain impact on the social economy and even cause political instability.

In recent years, oil price fluctuations have been a hot topic of discussion, especially since the outbreak of the epidemic in 2020, international oil prices have continued to rise sharply, once approaching the peak of international oil prices during the financial crisis in 2008. The current oil price trend has brought about a great impact on the economies of major oil producing and consuming countries around the world, and further affects the recovery process of the global economy.[1]

Therefore, it is important to forecast oil prices, and only by mastering the fluctuations of international crude oil prices can we ensure the continuous and stable development of economic markets. At the same time, the price of oil also affects the prices of its derivatives, and making forecasts of oil prices can also better analyze the direction of the economic market and reduce certain losses.[2]

The existing methods for predicting oil prices have gradually been divided into two schools of thought with the development of technology, which are taking the classical statistical approach and the more recently developed deep learning approach.

Ding, Jingzhi used ARIMA model to model the Platts spot quotes of Brent crude oil in the international oil market from July 1999 to May 2005 and validated the model with indicators to find a good fit [3]. Since crude oil prices are affected by major international events, variance values are generated. Xun Zhang used the BP multi-breakpoint detection algorithm proposed by Bai and Perron to detect the points in

time when structural changes in crude oil prices occurred from 1946 to the present [4]. Wu, Hong et al. proposed a time series forecasting model based on a combination of ARIMA and SVM based on a comprehensive analysis of the linear and nonlinear conforming characteristics of oil prices [5], and found that the combined model has higher forecasting accuracy compared to the single model.

Before the popularity of deep learning methods, the more widely used forecasting methods were all corresponding within classical statistics, and after 2017 deep learning methods were gradually applied to price forecasting. Most studies comparing price forecasting have mostly used a single method, such as the LSTM method or the GARCH model, but both methods have their corresponding limitations. In this paper, we take into account the time-series nature of the data, and at the same time, we are able to extract the characteristics of the data to a great extent, and build a hybrid model of LSTM deep neural network and traditional time series model with multiple GARCH models to forecast international oil prices.

2. Related Work

2.1. GARCH family model principle

2.1.1. GARCH model

The GARCH model, known as the generalized autoregressive conditional heteroskedasticity model, was proposed by Bollerslev[6] based on the ARCH model. The core idea of the model is that, on the basis of the ARCH model, each lagged value of the conditional variance σ_t^2 is introduced to obtain the generalized ARCH model, which makes the model more flexible, and the corresponding model equation is shown below.

$$\begin{cases} y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \dots + \alpha_p \varepsilon_{t-p}^2 + \beta_1 \sigma_{t-1}^2 + \dots + \beta_q \sigma_{t-q}^2 \\ \dots = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^q \beta_i \sigma_{t-i}^2 \\ \varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2) \end{cases} \quad (1)$$

where α_0 denotes the constant parameter that $\alpha_i \geq 0 (i=1, 2, \dots, q)$ is the ARCH effect parameter for the short-term fluctuations to be estimated, and $\beta_i \geq 0 (i=1, 2, \dots, p)$ is the GARCH benefit parameter for the long-run fluctuations to be estimated, $q > 0, p \geq 0$ and $\sum_{i=1}^q \alpha_i + \sum_{i=1}^p \beta_i = 1$.

In practice, the most widely used is the GARCH (1, 1) model, which corresponds to a model expression of the form

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (2)$$

2.1.2. EGARCH model

Nielsen created the EGARCH model in 1991[7], the core idea of which is to combine ε_t the conditional variance of σ_t^2 . The model is based on the logarithmic treatment of the conditional variance of the GARCH model, which can effectively overcome the asymmetry and leverage effects of the GARCH model in financial markets, as follows

$$\begin{cases} y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + \varepsilon_t \\ \log \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \frac{|\varepsilon_{t-i}|}{\sigma_{t-i}} + \sum_{i=1}^r \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} + \sum_{i=1}^q \beta_i \log(\sigma_{t-i}^2) \\ \varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2) \end{cases} \quad (3)$$

Generally speaking, the most widely applicable EGARCH model is the EGARCH(1,1) model, which corresponds to the following specific expressions:

$$\log \sigma_t^2 = \alpha_0 + \alpha_1 \frac{|\varepsilon_{t-1}|}{\sigma_{t-1}} + \gamma_1 \frac{\varepsilon_{t-1}}{\sigma_{t-1}} + \beta_1 \ln(\sigma_{t-1}^2) \quad (4)$$

where $\sum_{i=1}^r \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} > 0$, i.e., it means that the series exhibits an inverse leverage effect, with positive news having a greater volatility shock on financial markets than negative news, and there is a $(\alpha_1 + \gamma_1)$ -fold shock; $\sum_{i=1}^r \gamma_i \frac{\varepsilon_{t-i}}{\sigma_{t-i}} > 0$, i.e., indicates that the series exhibits the phenomenon of the leverage effect, where the volatility of the financial market is hit more by positive news than by good news, and there is a $(\alpha_1 - \gamma_1)$ -fold shock.

2.1.3. TGARCH model

In response to the limitations of the ARCH and GARCH models, which do not accurately describe the leverage effects of financial time series, i.e. the asymmetry problem, Zakoian created the TGARCH model in 1990. In 1990, Zakoian created the Threshold ARCH model, or TGARCH model, which can effectively address these issues by distinguishing between positive and negative shocks to conditional volatility, as follows:

$$\begin{cases} y_t = \beta_0 + \beta_1 x_{1t} + \dots + \beta_k x_{kt} + \varepsilon_t \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^r \gamma_i \varepsilon_{t-i}^2 d_{t-i} + \sum_{i=1}^q \beta_i \sigma_{t-i}^2 \\ \varepsilon_t | I_{t-1} \sim N(0, \sigma_t^2) \end{cases} \quad (5)$$

Generally speaking, the most widely applicable TGARCH

model is the TGARCH(1,1) model with the following expression:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 d_{t-1} + \beta_1 \sigma_{t-1}^2 \quad (6)$$

where $\gamma_1 \varepsilon_{t-1}^2 d_{t-1}$ is the TGARCH term, and if $\gamma_1 = 0$ then asymmetry does not exist and if $\gamma_1 \neq 0$ then asymmetry exists. When $\varepsilon_{t-1} \geq 0, d_{t-1} = 0$ when, i.e., it means that the good news has α_1 multiplier shock; when $\varepsilon_{t-1} \leq 0, d_{t-1} = 1$ when, i.e. representing bearish news has a $(\alpha_1 + \gamma_1)$ times the shock

2.2. LSTM model principle

Long Short-Term Memory (LSTM) is a machine learning algorithm based on recursive neural network (RNN) [8], which is characterised by a unique gating design for information transfer and a more sophisticated information transfer mechanism that can fully exploit the long-term dependencies of sequence data. The LSTM neural network model consists of an input layer, a hidden layer, a fully-connected layer and an output layer, where the hidden layer is a chain structure, mainly including input gates, output gates, forgetting gates and memory modules, and its specific formulae are as follows:

$$\begin{aligned} f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \\ \hat{C}_t &= \sigma(W_C \cdot [h_{t-1}, x_t] + b_C) \\ C_t &= f_t * C_{t-1} + i_t * \hat{C}_t \\ o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \\ h_t &= o_t * \tanh(C_t) \end{aligned} \quad (7)$$

Where i_t, f_t and o_t are the input threshold, the forgetting threshold and the output threshold, respectively; σ is the Sigmoid activation function; \tanh is the hyperbolic tangent activation function; and C_{t-1} is the internal state of the cell at time $t-1$; and \hat{C}_t is the cell candidate state at time t ; C_t is the internal state of the cell at time t ; and h_t is the external state of the cell at time t . W_C, W_i, W_f and W_o are the weight matrices of the state cell, the input gate, the forgetting gate and the output gate at time t , respectively; b_C, b_i, b_f and b_o are the bias vectors of the state unit, the input gate, the forgetting gate and the output gate respectively; "*" denotes the vector product.

2.3. Evaluation of predicted effects

In order to better differentiate the fitting accuracy of different forecasting models, i.e. to judge the prediction ability of different models, the loss function method is used to evaluate the prediction accuracy, and the metrics adopted in this paper are: root mean squared error (MSE) and mean absolute error (MAE). The root mean square error (MSE) is obtained by squaring the deviation of the predicted value from the true value, while the mean absolute error (MAE) is obtained by calculating the deviation of the predicted value from the true value and absolutising the value obtained. The formula is shown below:

$$\begin{aligned}
RMSE &= \sqrt{\frac{1}{h} \sum_{t=l+1}^{t=l+h} (y_t - f_t)^2} \\
MAE &= \frac{1}{h} \sum_{t=l+1}^{t=l+h} |y_t - f_t|
\end{aligned} \tag{8}$$

3. GARCH-LSTM

As both the GARCH family model and the LSTM neural network have advantages and disadvantages, this paper combines the information obtained from the GARCH family model with the LSTM neural network for the purpose of improving the model fitting accuracy.

The GARCH-LSTM algorithm flows as follows:

Dividing the sample set into a training set and a test set, fitting the GARCH family model using the data in the training set, and obtaining the prediction results and parameters corresponding to the GARCH family model;

Combining the results and parameters obtained by fitting the GARCH family model with the original data from the training set to obtain the new data set;

The newly obtained dataset is fed into the LSTM neural network and the corresponding fitting accuracy is checked using a test set to obtain the RMSE and MAE of the prediction results.

4. Experimental Results and Analysis

4.1. Data Description

The data set used in this article is Brent crude oil futures data, sourced from Homepage - U.S. Energy Information Administration (EIA). The data for Brent crude oil covers the period from January 2000 to December 2022 and includes the opening, closing, high and low prices and daily flows for each trading day. In the subsequent study, the closing price of each trading day (trading days excluding double days and holidays) is selected for the study.

In this paper, there are 5822 sample data, and the first 4822 data are selected as the training set data and the last 1000 data are used as the test set data to forecast the closing price. The training set data is first used to estimate the parameters corresponding to the GARCH, EGARCH and TGARCH models, such as the parameters representing the magnitude of volatility shocks in GARCH α_1 and volatility representing past persistence β_1 . The LSTM neural network model is then constructed using the other explanatory variables: open price, close price, high price, low price and daily flow as inputs to the LSTM neural network. The models are then tested for predictive performance using a test set to obtain their corresponding fit accuracy and to compare the predictive performance of different models.

4.2. Experimental Procedure

Before fitting the GARCH model, an analysis of the basic characteristics of the volatility of the daily closing prices is

required. The following chart shows the time series of the log difference of the daily closing prices, as shown in Figure 1 below. It can be observed that it has a more pronounced heteroskedasticity, with significant inconsistency in the series fluctuations.

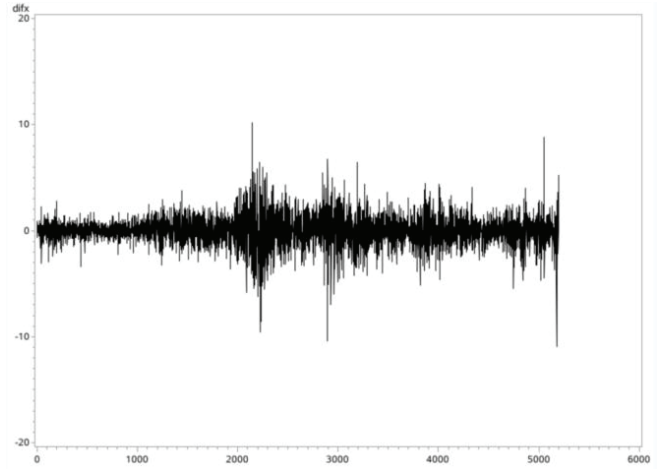


Figure 1. Time-series diagram of the post-log differential daily closing price

In order to draw more precise conclusions to fit the GARCH model, the series needs to be tested for smoothness, normality, serial autocorrelation and heteroskedasticity.

For the smoothness test, the ADF test, the most commonly used test in statistics, is chosen as the method of smoothness test in this paper, and its corresponding test results are shown in the following table:

	t-test statistic	P-value
ADF test statistic	-29.123433	0.0000
Test threshold (0.01)	-4.2431131	-
Test Critical Value (0.05)	-2.8313243	-
Test threshold (0.1)	-2.4810842	-

For the normality test, it can be obtained that the series also has the characteristics of right skewness and spikes and thick tails, so the Student's t distribution is chosen to fit the data in this paper. After the autocorrelation and heteroskedasticity tests, it can be observed that the p-values are less than the significance level, thus inferring that the series is not autocorrelated and has an ARCH effect, satisfying the conditions for fitting the data using the GARCH model.

The data were fitted using the GARCH, EGARCH and TGARCH models respectively, and the parameters obtained from their corresponding fits are shown in the table below.

Table 3. Parameters fitted to each model

Models	Constant term	ε_{t-1}^2	$\varepsilon_{t-1}^2 d_{t-1}$	σ_{t-1}^2	$\frac{\varepsilon_{t-1}}{\sigma_{t-1}}$	$\left \frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right $
GARCH	0.0000642	0.0938304	-	0.9438126	-	-
EGARCH			-	0.9923913	0.0232346	0.2574105
TGARCH	0.000000031	0.1382134	0.0134251	0.9130313	-	-

The parameters obtained from the above table were brought into the equation to obtain the corresponding model for subsequent predictions. And the individual parameters obtained in Table 2 with the prediction results were used as

input to the neural network to build the hybrid model. The figure below shows a plot of the true value versus the predicted value for the 300 data items after August 1, 2021.

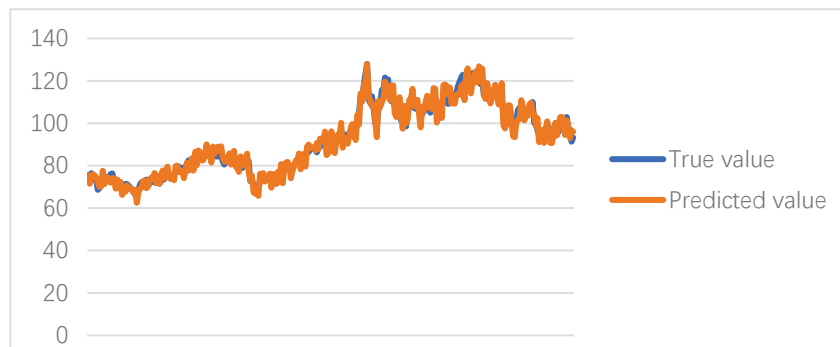


Figure2 Comparison of true and predicted values

By calculating the RMSE and MAE corresponding to the test set and comparing them with the RMSE and MAE corresponding to the test machine after fitting the LSTM and

GARCH models, the results of the comparison are shown in the table below.

Table 4 RMSE and MAE for different models

Models	Root Mean Square Error (RMSE)	Mean Absolute Error (MAE)
GARCH	0.00538	0.01973
TGARCH	0.00453	0.01823
EGARCH	0.00424	0.01792
LSTM	0.00223	0.01575
GARCH-LSTM	0.00148	0.01378

4.3. Data Description

The GARCH-LSTM model was used to improve the prediction accuracy, and the data was predicted under the assumption that the residual terms obeyed a t-distribution, and the corresponding prediction results are shown in the table above. The prediction results using a single model performed poorly, with the GARCH model predicting an RMSE of 0.538% and MAE of 1.973%, the TGARCH model predicting an RMSE of 0.453% and MAE of 1.823%, the EGARCH model predicting an RMSE of 0.424% and MAE of 1.792%, and the LSTM predicting RMSE was 0.223% and MAE was 1.575%. The GARCH-LSTM prediction results were better compared to the GARCH family model and the LSTM model, with significantly smaller prediction errors than the other models, with an RMSE of 0.148% and a MAE of 1.378%.

5. Conclusion

In this paper, data corresponding to Brent crude oil from January 2000 to December 2022 are selected, and after a series of statistical tests: smoothness test, white noise test, ARCH test and heteroskedasticity test, GARCH family models: GARCH, TGARCH and EGARCH are built to fit the data to obtain their corresponding parameters. After the parameters are obtained, the parameters are input as variables together with the data into the LSTM neural network for training, thus building the GARCH-LSTM model. Compared to a single model, it can be observed that the model fits better

and has less prediction error. Based on the experimental results, it can be obtained that the hybrid model is able to compensate for the shortcomings existing in each model and obtain better results compared to the single model studied.

References

- [1] Hu Aimei, Wang Shuping. (2012) Comparative analysis of international oil price forecasting based on ARIMA and GARCH models. *Journal of Economic Research*, 26:196-199.
- [2] Salah A, Hamid B. (2004) On the predictive accuracy of crude oil future prices. *Energy policy*, 32.
- [3] Ding Jingzhi, Min Ti, Lin Yi. (2008) Application of ARIMA model in oil price forecasting. *Logistics Technology*, 27(10): 156-159
- [4] Zhang Xun, Yu Le An, Lai Jian Qiang, et al. (2009) The impact of major emergencies on crude oil prices. *System Engineering Theory and Practice*, 29(3): 11-15
- [5] Wu H, Yin H. (2010) ARIMA and SVM combined model for oil price forecasting. *Computer Simulation*, 5:264-267.
- [6] Engle R F, Bollerslev T. (1986) Modelling the Persistence of Conditional Variances. *Econometric reviews*, 5(1): 1-50.
- [7] Nelson D B. (1991) Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica: Journal of The Econometric Society*, 59 (2): 347-370.
- [8] Graves A, Schmidhuber J. (2005) Framewise Phoneme Classification with Bidirectional LSTM and Other Neural Network Architectures. *Neural Networks*, 18(5-6): 602-610.