Proceedings of the 6th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/68/20241466

A study of European wheat price series forecasting based on multiple variants of Holt-Winters and LSTM models

Houhan Li

Central University of Finance and Economics, School of Statistics and Mathematics, Beijing, 102206, China

houhan2003@gmail.com

Abstract. Given the importance of wheat in global food security and agricultural economy, accurate price forecasts are of significant practical significance for growers, the food chain, and government macro-control. Four different time series forecasting models were used in this study, including the traditional Holt-Winters three-parameter exponential smoothing method and three improved models based on the Long Short-Term Memory Network (LSTM): the CNN-LSTM, the GRU, and the VMD-LSTM. By comparing the forecasting effects of these models under the same training and test sets, this study aims to select the best models using the best model for predicting European wheat prices from 2010 to 2023 after training. The results show that the CNN-LSTM model performs the best in predicting the European wheat price series, and the future wheat price will show a slow downward trend, and the CNN-LSTM model combines the feature extraction ability of convolutional neural network and the long-term dependency learning ability of LSTM, which can effectively capture the fluctuation characteristics of the price series and the trend characteristics. The GRU model also shows a better prediction ability, especially in dealing with long-term dependencies. in dealing with long-term dependencies. However, the VMD-LSTM model has certain shortcomings in its prediction effect when facing time series data with large fluctuations in the later period. In addition, the Holt-Winters model predicts that after a period of slow decline, wheat prices may rebound, which may be related to the adjustment of market supply and demand and changes in external factors.

Keywords: Time series prediction, Wheat, LSTM

1. Introduction

The study of wheat price forecasting is of practical significance for several types of economic agents [1]. In China, wheat cultivation accounts for 20% of the country's total arable land- 30% [2]. Wheat ranks second among the three major cereals and is also the main raw material for a variety of common staple foods such as bread and biscuits. Its price factor has a wide range of impacts, from wheat growers to the cereal chain. For the national government, wheat price forecasts can adjust the planned import volume of wheat and the subsidy to wheat farmers; for farmers, wheat price forecasts can be used to adjust the types of crops to be planted; and for the wheat futures market, wheat price forecasts can provide risk protection to mitigate the impact of price fluctuations. Considering the wide range of wheat price impacts mentioned above, the main objective of this paper is to try to forecast the price volatility trend of European wheat using two forecasting methods, namely, the traditional time series model and

DOI: 10.54254/2755-2721/68/20241466

the machine learning model, and to select a better model for wheat price forecasting by comparing the accuracy of the forecast.

There are four main models used in this paper. The first is the Holt-Winters three-parameter time series approach, which is a traditional time series forecasting model. It considers level, trend and seasonal variations and is suitable for data such as wheat prices where there is an implied element of cyclical variation. Next is a series of improved models based on Long Short Term Memory (LSTM) networks. In this paper, we first consider a hybrid model using a combination of Convolutional Neural Networks (CNNs) and LSTMs as the second model. Sequential features are extracted by CNNs, and the resulting matrix is then fed into LSTMs to deal with temporal correlations and long-term dependencies in the sequential data. The third model uses gated recurrent units (GRU) as the main sequence prediction method, which achieves faster predictions by optimising the gating mechanism of the LSTM. Finally, This paper will consider the modal decomposition approach, which will use the variational modal decomposition (VMD) The VMD method will be used to decompose the sequences and then apply the LSTM to predict the results of the decomposed sequences, and finally accumulate the sequence prediction results. The research results in this paper are not only of guiding significance for the European wheat market, but also provide a scientific basis for the management and regulation of the Chinese wheat market.

2. Literature review

Within the scope of this paper, the previous studies are divided into two categories. One category is related papers that focus on the discussion of forecasting methods, and the other category is related papers that focus on the factors influencing wheat prices as well as wheat price forecasting and early warning studies. In Yang Yang et al.'s study, the ARIMA model was used as the basis to predict the future trend of GDP in the cities of Sichuan Province, and to provide suggestions for the order of public transport projects landing in each city [3]. Similarly, in this paper, the Holt-Winters three-parameter exponential smoothing method of the traditional statistical model was selected as a representative of the traditional statistical model, and the forecasting effect of the Holt-Winters model was used to represent the performance of the traditional statistical model forecasting method on the time series data of wheat price. And in the study of Zhang Jin and Zhang Ruibin, BP neural network was used to predict the price trend of pork [4]. Similarly, this paper predicts the same bulk commodity of wheat, using a more complex neural network LSTM than BP to complete the wheat price prediction. In terms of methodology, not only can the prediction be done from the perspective of studying the linear relationship of the series and factors such as cycles and seasons, but also consider the possibility of improving the accuracy of the model provided by the method of modal decomposition. In the study by Xu Weishuai et al., a combination of VMD and LSTM was used to complete the prediction of the Kuroshio front strength series in the East China Sea [5]. This paper will verify that for the wheat price series in the two regions of Central Europe, it will test whether the prediction results are better after using the modal decomposition. In summary, a more complete system of sequence forecasting methods already exists, and in this paper, we will use four different models to apply to the price series of commodity wheat for forecasting, and compare and discuss the results of the four models.

3. Research methodology

3.1. Research ideas

In this paper, we hope to forecast the time series data of European wheat prices by using a variety of forecasting models, to give the forecasting effectiveness and characteristics of each model for the European wheat price series, and ultimately to give a suitable methodology for conducting the forecasting of the European wheat price series. There are two types of models used in this paper: the traditional statistical forecasting model Holt-Winters and variants of three machine learning models based on long- and short-term temporal memory networks, respectively, and the latter This paper obtains the model with more accurate predictions by comparing the results of the models under the same training and test sets.

3.2. Data sources and description

The European wheat prices used in this article belong to the European Agricultural Data website: https://agridata.ec.europa.eu. This website provides agricultural prices from representative retail chains in each European country. Due to the lack of data, this paper selects the wheat prices of France and Germany, the two countries with the largest wheat planting areas in Europe, and takes the mean value to get the time series data representing the European wheat price series, which is collectively referred to as the European wheat price data later. In the European wheat exporting countries in accordance with the import and export volume in descending order, France and Germany respectively occupy the first two; in the European wheat import and export volume value, France and Germany accounted for more than 50%, which can represent the European wheat price series. The time span of data selection in this paper is from 2010 to 2023, totalling 13 years. The data frequency is weekly sampling, i.e., the length of the series is $13 \times 52 = 676$. In Figure 1, the trend of wheat prices in Europe is shown, and it can be seen that wheat has shown a decrease and then an increase in the last 13 years followed by a sudden drop to 250 euros per tonne. The outbreak of the COVID-19 in 2020 had a large impact on the European wheat market, even reaching a price of more than €400 per tonne at one point in late 2022.

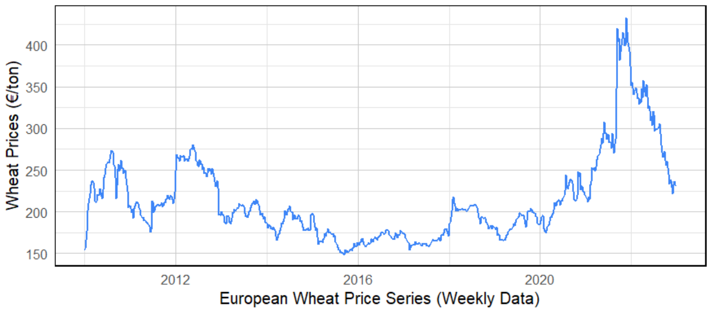

Figure 1. Chart of the European wheat price series

3.3. Modelling

3.3.1. Traditional time series models--Holt-Winters. In this paper, forecasts for the European wheat price series, the use of three-parameter exponential smoothing Holt-Winters can be better applied to time series forecasts with trend and seasonality [6]. The model better captures the pattern of change in the time series by exponentially smoothing the trend and seasonality as well as weighting the series. In this paper, the least squares method is used to estimate the three parameter values.

The European wheat price series are substituted into the Holt-Winters model for parameter estimation and forecasting, and the results are obtained in Figure 2, where the red dashed line represents the model's fitting effect in the training set, and the light blue solid line represents the prediction of the future results. The Holt-Winters method predicts small fluctuations in the future, with an overall downward trend.

Figure 2. European wheat price forecasts - Holt-Winters model

DOI: 10.54254/2755-2721/68/20241466

3.3.2. Three variants of the LSTM model. LSTM is commonly used as a forecasting tool when dealing with time series data, learning patterns and regularities in the sequence and predicting future values through information from historical time steps. There are other neural network models in the field of time series that enable prediction, such as BP neural networks, RNN neural networks, etc., and there are two unique advantages of LSTM over other neural networks :

A. LSTM is capable of capturing long-term dependencies in time series data [7].

B. LSTM is capable of handling variable length sequence data [7].

In this paper, the three models CNN-LSTM, GRU, and VMD-LSTM will be used to complete the prediction of the European wheat price series, and the optimal model will be derived for the wheat price series environment.

3.3.2.1. CNN-LSTM. Convolutional Neural Network (CNN) is a feed-forward neural network. When CNNs are used to process time series data usually the temporal information and the sequence values are considered as corresponding two dimensional information. Multi-scale feature extraction can be achieved by combining CNN and LSTM. CNN is used to extract local features while LSTM is used to capture global features in the sequence. By combining the two, multi-scale features can be extracted and integrated at different levels for a more comprehensive description of time series data [8].

In this paper, the CNN layer is set to accept the length of 52, and output 52-dimensional column vectors into the LSTM network, and the process is shown in Fig. 3. In the LSTM parameter setting, this paper selects the parameter that can make the prediction effect reach the optimal under the environment of the same training set and test set as the final model setting. Specifically: the number of LSTM hidden layers is 3 and the number of nodes is 64.

With such a setting, the original data is divided into training and test sets according to 12:1, and the results of performing the data on the test set are shown in Fig. 4.

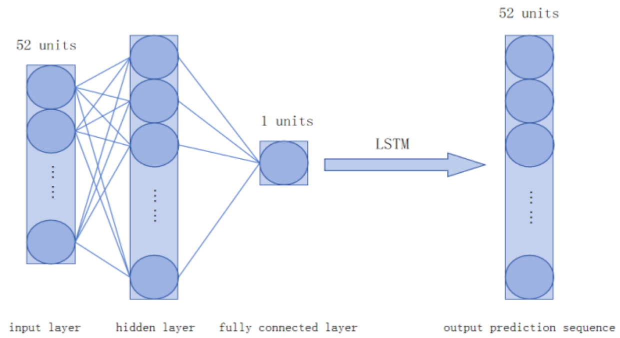

Figure 3. CNN-LSTM process

Figure 4. CNN-LSTM prediction effect on the test set

With this prediction effect, the European wheat price series is retrained using the full European wheat price series and the European wheat price series is predicted for the coming year. The prediction effect for the next 52 periods is plotted in Figure 5.

Figure 5. CNN-LSTM prediction of future

From Figure 5, under the prediction of the CNN-LSTM model, European wheat will show a slow decline in the future and a stationing point in mid-2023.

3.3.2.2. GRU. Based on the LSTM gating mechanism, the gating is changed to reset gate and update gate. In comparison with LSTM, GRU updates the gating mechanism so that GRU requires less memory while having higher computational efficiency.

The parameter settings of GRU are slightly different from those of LSTM. This paper found that setting the hidden layer nodes to 32 and the input window to 52 can get relatively good prediction results on the test set as shown in Fig. 6, and used the model to predict the future 52 periods, and the prediction results were obtained as shown in Fig. 7. Similar to the prediction of the CNN-LSTM, the GRU model was also less accurate than the CNN-LSTM model in the test set when the two models were compared, which is related to its simple gating mechanism. model comparisons, it is found that the accuracy of the GRU model is weaker than the CNN-LSTM model on the test set, which is related to its simple gating mechanism.

Figure 6. GRU prediction effect on the test set

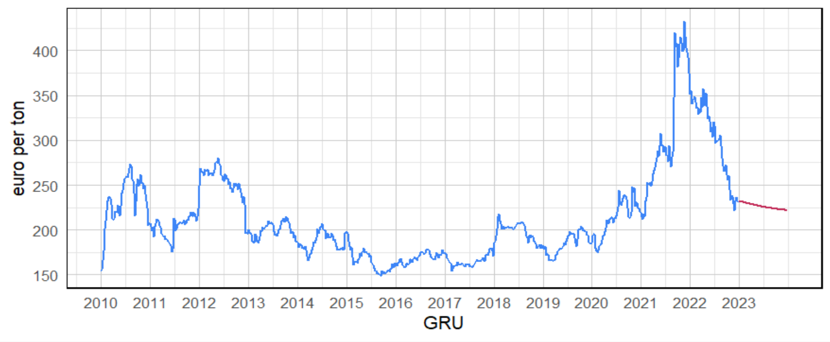

Figure 7. GRU projections for the future

3.3.2.3. VMD-LSTM. Among the traditional methods of time series analysis, series decomposition is a class of classical solutions. There are two perspectives on time series decomposition: one is the decomposition of time series according to long-term trend, seasonality, periodicity and residual terms proposed by Yule. The second is the decomposition of time series according to their frequency. In 2014, Dragomiretskiy et al. proposed the variational modal decomposition (VMD) method [9]. VMD is a modal decomposition method based on variational inference, which decomposes a signal into a set of FM-AM components through an iterative optimisation process. The basic idea of the method is to represent the signal as a set of locally FM oscillatory modes, each with a different frequency and amplitude. The VMD method not only adaptively determines the number of modes and their centre frequencies, but also retains a better resolution and extracts the local features of the signal more accurately.

Since the research object of this paper is the European wheat price sequence with high-frequency noise, and we hope that the final decomposition of the sequence does not have spectral overlapping region, so this paper uses the VMD method as the basic method of sequence decomposition, and set the parameter K to 4, i.e., the sequence is decomposed into a high-frequency sequence, a medium-frequency sequence, a low-frequency sequence, and a random error, as shown in Fig. 8. After discarding the random error IMF_3, we implement LSTM prediction for the remaining three sequences, and sum up the prediction results to get the results. LSTM prediction, the results are obtained by accumulating the predictions, the effect of sequence prediction on the test set and the future prediction results are shown in Fig. 9 and Fig. 10.

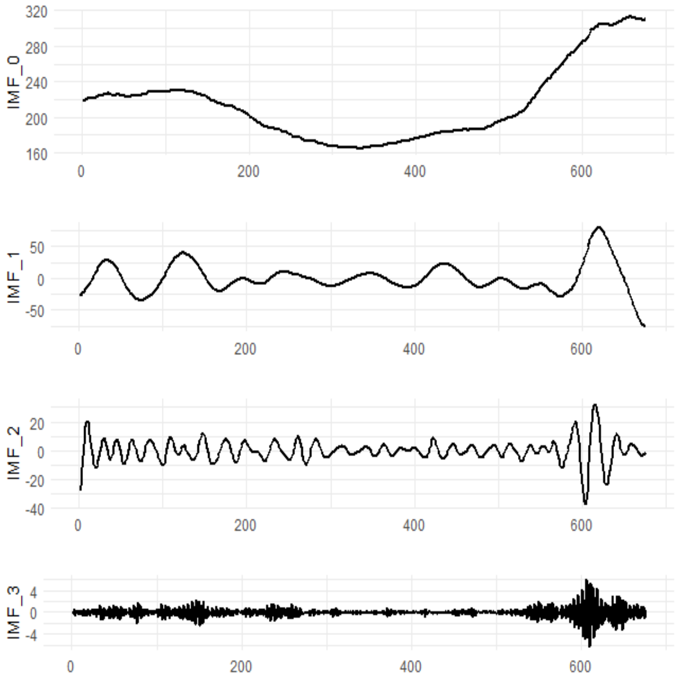

Figure 8. Sequence after VMD decomposition

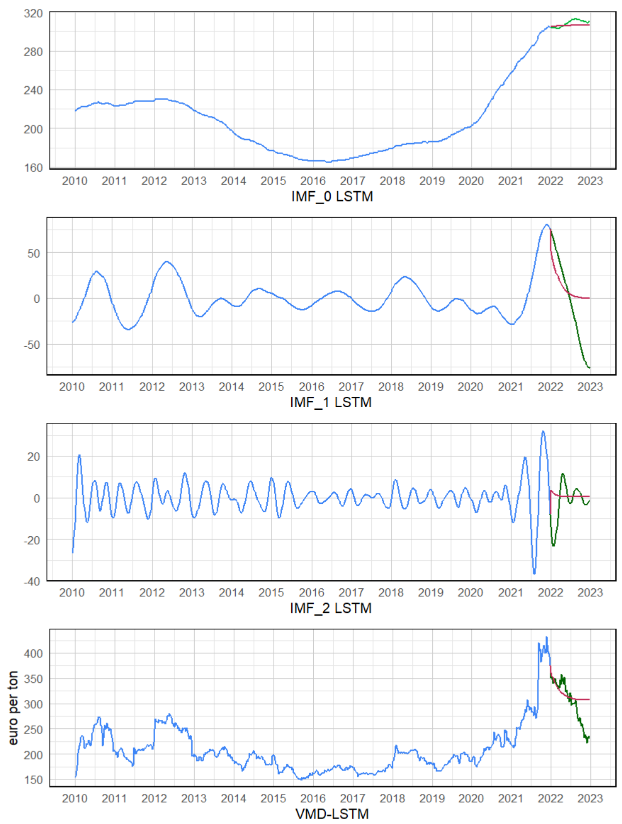

Figure 9. Effect of VMD-LSTM on the test set

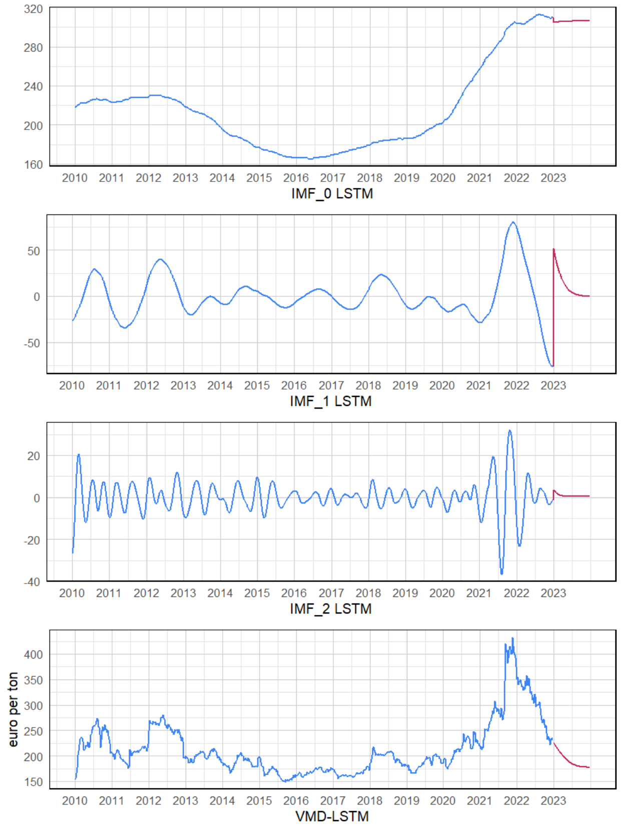

Figure 10. Prediction of the future by VMD-LSTM

In the above LSTM family of models in this paper, the results of setting the model parameters and performance on the test set are shown in Table 1:

Table 1. Neural network model parameter settings

modelling	minimum mean square error	Recurrent neural network parameters		
		batch_size	input_size	epoch
CNN-LSTM	76.35	4	52	2000
GRU	388.41	32	52	100
VMD-IMF0	14.86	32	52	5000
VMD-IMF1	1590.00	32	208	60
VMD-IMF2	102.34	32	208	100
VMD-LSTM	1421.67			

Among the three variants of LSTM, the model with the smallest mean square error applied to the European wheat price series data is the CNN-LSTM model, which uses a convolutional neural network to extract the sequence features to better capture the volatility characteristics of the price series and the trend characteristics, and uses the LSTM neural network as a forecasting method to retain not only the seasonality of the long series, but also the trend fluctuation characteristics. Among the three variants, the model with the worst prediction results from the perspective of mean square error is the VMD-LSTM model. Among the three decomposition sequences of the VMD method, the IMF1 decomposition sequence gives the largest error in the results. After observing the original data, this paper suggests that the reason for the large error in this result is the large difference between the true values of the test set and the true values of the training set for the IF sequence on the selected test set. In terms of trend, the test set data is steeper than the training set data; in terms of values, the domain of values of the test set data is much larger than the training set data.

4. Discussion

The use of the European wheat price series for forecasting in this paper has practical implications for the management of the Chinese wheat market. For the commodity wheat, its price has a strong correlation in all regions of the world, and the forecasting methods and research ideas mentioned in this paper can be applied to the Chinese wheat price series to get the Chinese wheat price forecasting series. Early warning based on wheat price fluctuations can, on the one hand, reduce the losses of wheat farmers, and on the other hand, the state can conduct macro-control through wheat price prediction results to maintain the stability of wheat prices. Among the machine learning methods mentioned in this paper, all three LSTM variants are suitable for single-feature time series forecasting. In the order of presentation of the methodologies, from the original GRU to the proposed LSTM combined with convolutional neural networks to the VMD-LSTM method, which was first proposed in 19 years, these methods are expected to give higher and more accurate judgments. However, in the VMD-LSTM method, this paper draws on the idea of Xu Weishuai et al. to identify the highest frequency sequence as random noise and discard it [5]. When predicting each modal decomposition sequence, it was found that for time series data whose fluctuations in the later period are greater than those in the early period, it is necessary to ensure reasonable cutoffs in the training set and test set, and try to keep the fluctuation range of the test set data does not exceed the fluctuation range of the training set data. When the test set fluctuation in the variable modal decomposition result sequence is too large, the original sequence is not suitable for prediction using the VMD-LSTM method.

5. Conclusion

This paper presents an in-depth analysis and forecasting of the European wheat price series using a variety of time series forecasting models, including the traditional statistical model Holt-Winters and three variants of the Long Short-Term Memory Network (LSTM)-based models: the CNN-LSTM, the GRU, and the VMD-LSTM. After comparing the forecasting effects of the models, the following conclusions are drawn: The CNN-LSTM model shows the best prediction results, which can effectively capture the volatility features and trend characteristics of the wheat price series. The GRU model has the second best performance on the test set, but still shows better prediction ability in dealing with long-term dependence. The VMD-LSTM model has some shortcomings in prediction results, especially when dealing with the time series data with large volatility in the later period. decomposition results of the VMD method may lead to larger prediction errors.

According to the results of this paper, the European wheat price series shows an overall trend of slow decline in the coming period. This trend may be influenced by a variety of factors, including changes in the global wheat supply and demand situation and adjustments in international trade policies. The Holt-Winters three-parameter model, in predicting the future price trend of wheat, points out that after a period of slow decline there may be a rebound in prices. Such a rebound may be due to the re-adjustment of supply and demand in the market, or due to the sudden change of some external factors. This paper is limited to the unavailability of a full dataset of Chinese wheat prices from official platforms and does not provide a forecast of the Chinese wheat market trend. Future research can further explore and optimise the model parameters to improve the model's adaptability and prediction accuracy for complex time series data. Consider introducing more influencing factors and external data, such as climatic conditions, policy changes, etc., to enhance the explanatory power and predictive ability of the model. The application of the VMD-LSTM model on different types of time series data is further investigated, especially its performance in dealing with high-frequency noise and non-stationary time series. The research in this paper not only provides effective tools and methods for the prediction of wheat prices in Europe, but also provides a scientific basis and scientific methods for the management and regulation

DOI: 10.54254/2755-2721/68/20241466

of the Chinese wheat market. By continuously optimising and improving the prediction model, it is expected to play a greater role in the field of agricultural economic management and promote the healthy development of the agricultural market.

References

[1] SHAO Yongtong, WANG Lu, GOU Yifeng. A study on the interactive characteristics of wheat futures and spot price volatility in China[J]. Rural Economy and Technology,2023,34(16):1-3+12.
[2] ONS Bulletin on Grain Production Data for 2022, https://www.stats.gov.cn/sj/zxfb/202302/t20230203_1901673.html
[3] YANG Yang, TIAN Dingsheng, ZHANG Bao'an, et al. Research on urban economy and population forecasting based on ARIMA model[J]. Comprehensive Transport,2023,45(11):79-85+97.
[4] ZHANG Jin,ZHANG Ruibin. BP neural network pork price prediction based on time series[J]. Science and Technology Innovation and Application,2016(20):63.
[5] XU Weishuai, ZHANG Lei, WANG Hua. Temperature front strength prediction of the East China Sea Kuroshio based on front extraction and VMD-LSTM[J/OL]. Advances in Marine Science:1-12[2024-03-30]. http://kns.cnki.net/kcms/detail/37.1387.P.20240201.1231.002.html.
[6] Taylor J W, McSharry P E. Short-term load forecasting methods: an evaluation based on european data[J]. IEEE Transactions on Power Systems, 2007, 22(4): 2213-2219.
[7] Ma X, Tao Z, Wang Y, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data[J]. Transportation Research Part C: Emerging Technologies, 2015, 54: 187-197.
[8] Liang Yuanshun. Research on pork price prediction based on feature selection and LSTM hybrid model[D]. Fuyang Normal University,2024.DOI:10.27846/d.cnki.gfysf.2023.000238.
[9] Dragomiretskiy K. Variational methods in signal decomposition and image processing [D]. UCLA, 2015.