IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 16 August 2023, accepted 7 September 2023, date of publication 11 September 2023, date of current version 15 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3314329

RESEARCH ARTICLE

Soybean Futures Price Prediction Model Based on EEMD-NAGU

JINDIAN LIU[1,*], **BO ZHANG**[2,*], **TIANHU ZHANG**[1], **AND JINGYANG WANG**[1]
[1]Hebei University of Science and Technology, Shijiazhuang, Hebei 050018, China
[2]Hebei Polytechnic Institute, Shijiazhuang, Hebei 050091, China

Corresponding authors: Bo Zhang (yatouzhangbo@163.com) and Jingyang Wang (jingyangw@hebust.edu.cn)

This work was supported in part by the Foundation of Science and Technology Research Project of Colleges and Universities in Hebei Province under Grant QN2023140, and in part by the Defense Industrial Technology Development Program under Grant JCKYS2022DC10.

*Jindian Liu and Bo Zhang contributed equally to this work.

ABSTRACT The soybean futures market in China occupies an important position in the agricultural product futures market. The research on the fluctuation of soybean futures price and the prediction of the future price trend has always been the focus of extensive attention in the field of agricultural economics. This paper proposes an EEMD-NAGU hybrid prediction model for soybean futures price based on Ensemble Empirical Mode Decomposition (EEMD) and New Attention Gate Unit (NAGU). The model first uses EEMD to process soybean futures price data into multiple Intrinsic Mode Functions (IMFs) and residual sequence. Then calculates the sample entropy of each IMF and reconstructed the IMF into three components of low-frequency, medium-frequency, and high-frequency, according to the size of the sample entropy. NAGU is formed by embedding the Attention mechanism (Attention) into the Gate Recurrent Unit (GRU) structure, further improving the learning capability of the model. NAGU splits the original reset gate and update gate, sets up two-stage respectively, and uses different activation functions to capture the information in historical data better. Soybean futures price data is complex nonlinearities and contain more noise. In this model, EEMD plays denoises the time series data and fixes the model input. NAGU can perform differential learning on data and finally produce prediction results. EEMD-NAGU is compared with thirteen other prediction models (Support Vector Regression (SVR), LSTM, GRU, NAGU, EEMD-LSTM, EEMD-GRU, EEMD-NGU, Attention-LSTM, Attention-GRU, Attention-NGU, EEMD-Attention-LSTM, EEMD-Attention-GRU, and EEMD-Attention-NGU). The evaluation indexes of the experiment are Mean Absolute Error (MAE), Mean Square Error (MSE), and R Squared (R^2). The experimental results show that the EEMD-NAGU model outperforms other models with better prediction performance. The model can be widely used to predict the price of wheat, corn, gold, oil, and other time series data.

INDEX TERMS EEMD, NAGU, EEMD-NAGU, attention mechanism, time series prediction.

I. INTRODUCTION

The earliest appearing in the futures market of China is agricultural product futures which occupy a high proportion in the domestic futures market. Soybean, as an important chemical raw material for production, its price fluctuations have a significant impact on governments, enterprises, and investors. There is often a linkage between soybean futures price and spot price, and futures have the function of price discovery. Therefore, effective prediction of soybean futures market trading price can help the government stabilize the market, enable soybean-related enterprises and farmers to reasonably avoid market price fluctuations, increase or decrease planting areas to ensure market circulation, and participate in rational investment, thereby ensuring my country's food security and market stability.

As an advanced method and technology, deep learning has shown strong capabilities in handling numerous complex prediction problems in recent years, and its inherent principles are suitable for predicting nonlinear time

The associate editor coordinating the review of this manuscript and approving it for publication was Xinyu Du.

series [1], [2], [3], [4]. In the face of noise and abnormal factors in time series prediction, deep learning can effectively handle and provide more accurate prediction results. With the increasing use of neural networks, Recurrent Neural Network (RNN), suitable for time series prediction, has been proposed. The RNN neurons are not isolated but have a structure of back-and-forth circulation. RNN predicts the next step's information based on the previous moment's information. But the original RNN also has obvious shortcomings, namely ''vanishing gradient'' and ''exploding gradient.'' It will not be reflected when the length of the time series is short, but for the long time series data, in the process of neural network back-propagation, the weight adjustment loses its effect [5]. RNN is only suitable for short-term memory.

The neuron part of Long Short-Term Memory (LSTM) adds an input variable based on ordinary RNN neurons and introduces three gate units, which control the transmission of information through the gate unit, selectively forgetting unimportant information, solving the shortcomings of RNN in processing long time series [6]. But the LSTM unit is composed of four fully connected layers internally. If there are a large number of neurons, it increases the training time and the number of parameters that need to be trained.

The original LSTM controls the flow of information through gate and more accurately explores the impact of more distant data on the current data. Still, the LSTM structure is more complex, and the training time is extended. On this basis, the GRU integrates the gate structure, changing from the original three gate units to two gate units, named reset gate and update gate, respectively, and merges the hidden layer and cell state in the LSTM. Chung et al. [7] conducted speech signal prediction and polyphonic music prediction, respectively and found that for RNN, LSTM, and GRU neural networks, the parameters of GRU were updated faster, and the training time is shorter during the training process to achieve the same prediction effect.

LSTM and GRU ease the gradient problem of RNN to a certain extent. Soybean futures data features are not consistent in their importance, and they still fail to provide a deeper understanding of the relationships in the data when dealing with more extended series. The ability to extract the internal correlation of non-serial data is weak, and they cannot accurately capture feature combinations that are important to the predicted target, limiting the model's predictive accuracy.

This paper proposes the NAGU, which embeds the Attention mechanism into the gate unit to focus on specific features adaptively and enhances the model's ability to extract the data's internal correlation. Based on the captured sequence information, further feature combinations which is important to the output target are captured. Meanwhile, NAGU has improved the reset gate and update gate based on GRU with a two-stage reset gate and two-stage update gate. Different activation functions are used between reset gate at each stage. The first stage reset gate uses the Sigmoid function, and the second stage uses the Leaky ReLU function. The results of the two-stage reset gate are combined as the output of the synthetic reset gate. The synthetic update gate is handled similarly. Through two filters, NAGU can better extract effective information in the input sequence from historical data. Although NAGU makes up for the shortcomings of the existent gate technology that cannot adaptively focus on specific features and the learning is insufficient, the historical data contains a large amount of random noise that needs to be processed. We use EEMD to deal with the random noise. EEMD is an improved version of EMD, which not only realizes partial denoising of data but also avoids the problem of mode aliasing.

Based on this, a hybrid EEMD-NAGU prediction model is proposed. This article uses soybean futures closing price data as the main research data and takes the Dow Jones Industrial Index (DJIA), S&P Dow Jones Indices Index (S&P 500), and National Association of Securities Dealers Automated Quotations Index (NASDAQ) as impact factor. By introducing 13 individual basic comparison models, the prediction results are overall analyzed using MAE, RMSE, and R^2. Through experiments, the MAE of EEMD-NAGU is 22.739, the RMSE is 28.500, and the R^2 is 0.95755. Compared with other comparable models, the prediction results of EEMD-NAGU are the best.

The main contributions of this paper are as follows:

(1) By studying the internal structure and principle of the GRU model, the NAGU model is proposed. NAGU has improved the reset gate and the updated gate based on GRU. It has two-stage reset gate and two-stage update gate and uses different activation functions to better learn from historical data.

(2) Integrating Attention into the GRU model elevates the justifiability of parameter allocation learning, enabling the model to adaptively take notice of and remember significant parts during the learning process.

(3) Compared with LSTM and GRU, under the same dataset, NAGU's MAE, MSE, and R^2 are better than these of LSTM and GRU in predicting soybean futures price.

(4) Use the EEMD method to reconstruct the low-frequency, medium-frequency, and high-frequency of the decomposed IMF components, which improves the accuracy of the model prediction. And a new hybrid model EEMD-NAGU for soybean futures price prediction is proposed based on EEMD and NAGU.

II. RELATED WORK

For the past few years, machine learning algorithms have achieved some results in time series forecasting [8]. Tong et al. [9] used Decision Forest (DF) to predict the affinity of chemical substances and receptors, and the prediction accuracy was significantly improved compared with the single Decision Tree (DT) model. Kim [10] used Support Vector Machines (SVM) to forecast stock price indices and concluded that SVM provides a promising alternative for financial time series forecasting. However, the traditional single SVM model is often unsuitable for training on large-scale datasets due to insufficient computing power. Wang [11] used the AutoRegressive Integrated Moving Average (ARIMA) model to predict the price trend of soybean futures. The results showed that the ARIMA model could better simulate and predict the closing price trend of soybean futures contracts, but it is also concluded that the ARIMA model can only be used for short-term prediction. With the increase of time, the change of data will increase and the prediction error will also increase. Babu and Reddy [12] predicted the time series using the combined model of differential ARIMA and Artificial Neural Network (ANN) and concluded that the mixture model has higher accuracy than a single model. Yao et al. [13] used AutoRegressive (AR) to predict the price of eggs and found that the price of eggs was significantly affected by external effects. This shows that external factors have a certain impact on futures price.

In recent years, predicting high-density, nonlinear, and non-stationary time series data through deep learning has become increasingly popular [14]. RNN has advantages in processing time series data. Huarng and Yu [15] used RNN to predict stock indices. Coulibaly and Baldwin [16] used a dynamically driven model to predict non-stationary hydrological time series and used RNN for this model, significantly improving prediction results. However, in the training process of RNN, the issues of ''vanishing gradient'' and ''exploding gradient'' will happen due to the long-term data dependence on the model. Ouyang et al. [17] proposed Short-term Time Series Network (LSTNet) to predict agricultural product futures price. The agricultural product futures price data set involved a mixture of long-term and short-term information, and a mixture of linear and nonlinear structures. The results show that LSTNet outperforms ARIMA and RNN. Wang and Gao [18] used LSTM to predict the price of soybean futures. Compared with other models, LSTM had sufficient time memory and was better at extracting time features. The above research shows that LSTM is more accurate in time series prediction. However, although LSTM can alleviate the problems of ''vanishing gradient'' and ''exploding gradient'', the training of a single LSTM is not stable. An et al. [19] constructed a two-stage long-short-term mixed memory TSH-LSTM model and found that adding social media text features can help improve prediction performance. Huang and Yong [20] used Variational Mode Decomposition (VMD) and LSTM to predict daily and monthly crude oil price, and used the ISE method to select VMD hyperparameters. The result indicated that the appropriate value of hyperparameters is significant. Amalia et al. [21] used LSTM and GRU to predict the agricultural products' price. The experiment indicated that GRU has a higher accuracy than LSTM, and has an advantage over LSTM in speed. Peng et al. [22] employed the CNN-GRU-Attention combination model to forecast the quality of water sources, which is superior to other decomposition techniques and machine learning algorithms in terms of prediction accuracy. Yang et al. [23] combined GRU and Attention to predict individual stocks. The results show that the combination of GRU and attention has a better effect than using attention or GRU alone.

EEMD is developed from EMD. EEMD decomposes the original data into several IMFs with different periods and a residual sequence. Still, unlike EMD, EEMD introduces Gaussian white noise into each sequence to be decomposed so that the original signal is at different characteristic time scales continue to reduce the degree of mode aliasing. Yu et al. [24] proposed a new neural network ensemble learning paradigm based on EMD to predict crude oil price. Based on the Generalized AutoRegressive Conditional Heteroskedasticity (GARCH), Zhang et al. [25] performed EEMD decomposition on the original data and then combined the particle swarm optimization algorithm and the least squares support vector machine to predict the nonlinear and time-varying parts of crude oil price, respectively. The final results showed that the prediction effect of the model was excellent. Zhang et al. [26] used the EMD algorithm to preprocess the data, and based on this, LSTM was used to build a prediction model. Wang and Wang [27] introduced GRU to train the model based on data preprocessing using EMD.

III. MODELS

A. EEMD

Huang et al. [28] proposed EMD, an adaptive signal decomposition algorithm. The IMF components and residual sequence after EMD decomposition, the IMF components represent the original data fluctuation, and the residual sequence represents the trend. The oscillation frequency of the IMF components decreases successively, describing the fluctuation characteristics of different periods of the original data, and the residual sequence has obvious linear features, representing the trend part of the original sequence. However, there is a problem of mode aliasing in the EMD decomposed sequence, where data of different frequencies may exist in the same IMF component or data of the same frequency may simultaneously appear in other IMF components [29]. Wu and Huang [30] proposed an improvement to EMD in 2009 and proposed EEMD, which solved the mode aliasing problem of the EMD algorithm.

Since EEMD decomposes historical data, adding future data may lead to changes in the number of decomposition results, so the trained model may need to process future input data. In addition, as the number of components increases, the processing time also increases. In order to reduce training time, each component's ZCR is calculated according to equation (1) to reflect the fluctuation frequency of the data by

$$\mathrm{ZCR} = \frac{1}{2N-1} \sum_{n=1}^{N-1} |sgn\,[x\,(n)] - \mathrm{sgn}[x(n-1)]| \qquad (1)$$

where $x\,(n)$ is the n-th data value, N is the total length of a frame of data, and sgn (n) is denoted by:

$$\mathrm{sgn}[x(n)] = \begin{cases} 1, & x(n) \geq 0 \\ -1, & x(n) < 0 \end{cases} \qquad (2)$$

Based on previous research findings [31], [32] and combined with the characteristics of soybean futures data. Finally, the closing price of soybean futures is decomposed into 10 IMF components and a residual sequence, as shown in Figure 1.

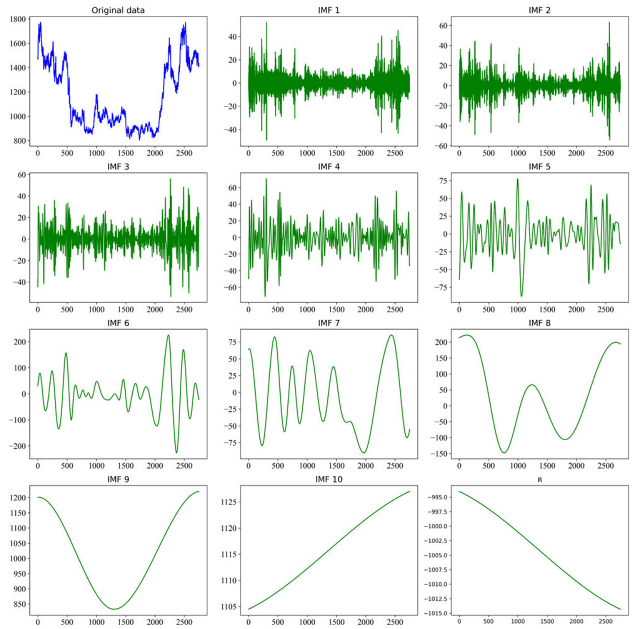

FIGURE 1. EEMD decomposition process.

ZCR denotes the proportion of sign change in each component data, reflecting the fluctuation frequency of the data. According to Table 1, IMF1 is a high-frequency component, and its ZCR is greater than 40%; IMF2 to IMF3 is a medium-frequency component, and its ZCR is between 10% and 40%; IMF4 to IMF10 are low-frequency components, and its ZCR is less than 10% [33]. The four reconstructed components are as shown in Figure 2.

TABLE 1. Zero crossing rate of each IMF component.

Components	IMF1	IMF2	IMF3	IMF4	IMF5
ZCR	62.5%	26.9%	12.4%	5.2%	2.1%
Components	IMF6	IMF7	IMF8	IMF9	IMF10
ZCR	0.9%	0.4%	0.1%	0%	0%

B. GRU

GRU is a kind of RNN composed of repetitive neural units. The processing nodes of a normal RNN are just simple neurons formed of activation functions, but the processing nodes of a GRU are units that select and control the flow of information through gate technology. The hidden layer can be one or more layers. The output of the GRU unit will be transmitted to the next unit in the same hidden layer, and at

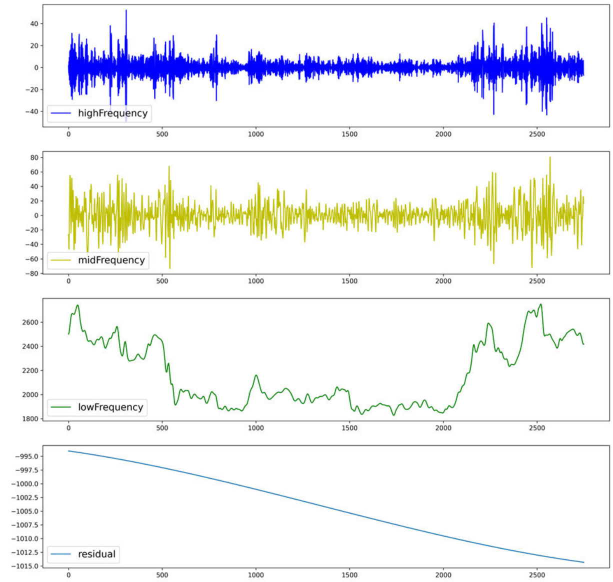

FIGURE 2. The four reconstructed components.

the same time, it will be shipped as input to the next hidden layer. The GRU model architecture is shown in Figure 3.

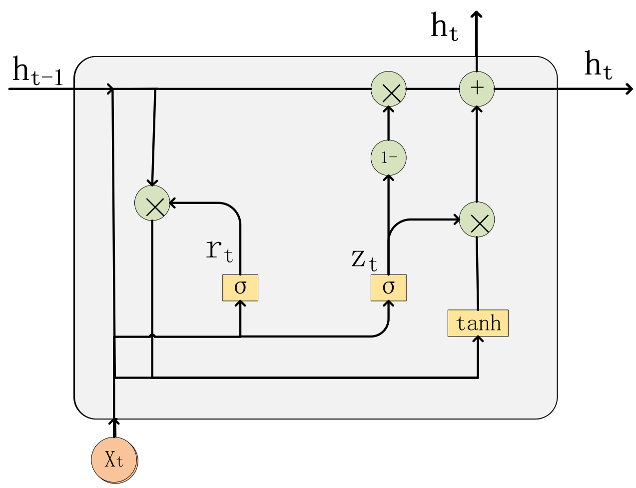

FIGURE 3. GRU model architecture.

C. NGU

Compared with GRU, NGU splits the reset gate and update gate, sets two-stage respectively, and introduces the Leaky ReLU function, and the activation function changes from the original Sigmoid to a part of Sigmoid and a part of Leaky ReLU. Sigmoid function has a similar nonlinear expression with biological neurons, but it has the characteristic of bi-directional saturation, the gradient update is slow in the saturated region, and there may be the problem of "vanishing gradient" and "exploding gradient" during the process of multiplying back propagation gradient. Leaky ReLU is a piecewise function that gives the input value a small slope when the input is negative. On the basis of solving the problem of zero gradient in the case of negative input, it alleviates the possibility that some neurons may not be activated, resulting in the problem that the corresponding parameters have not been updated. When the input is not negative, its gradient is 1, and there is no saturation region, so there is no problem of "vanishing gradient" and "exploding gradient" when the backpropagation gradient is multiplied. The NGU model architecture is shown in Figure 4.

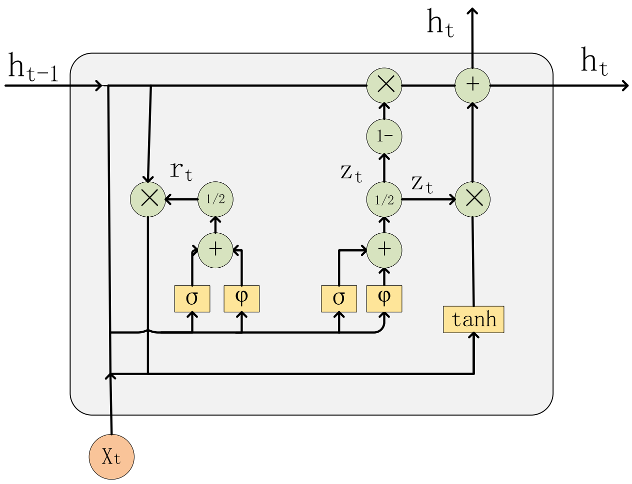

FIGURE 4. NGU model architecture.

D. NAGU

Incorporating Attention into the model modifies the method of computing candidate hidden states. The NAGU model architecture is shown in Figure 5.

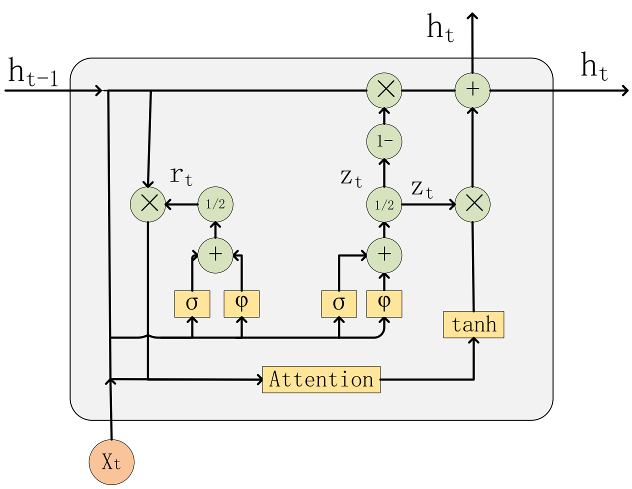

FIGURE 5. NAGU model architecture.

NAGU's memory cell receives the output value h_{t-1} at the previous moment and the current input value x_t into the reset gate and update gate. The synthetic reset gate controls how much past cognition is forgotten. The synthetic reset gate algorithm of NAGU is denoted by:

$$\hat{r}_t = \sigma\left(W_{\hat{r}}\left[h_{t-1}, x_t\right] + b_{\hat{r}}\right) \tag{3}$$

$$\bar{r}_t = \varphi\left(W_{\bar{r}}\left[h_{t-1}, x_t\right] + b_{\bar{r}}\right) \tag{4}$$

$$r_t = \frac{\hat{r} + \bar{r}}{2} \tag{5}$$

where σ is the Sigmoid activation function, φ is the Leaky ReLU activation function, x_t is the input vector at the current moment, h_{t-1} is the hidden state passed down from the moment antecedently, $W_{\hat{r}}$ is $b_{\hat{r}}$ signals the weight matrix and bias term of the first stage reset gate, $W_{\bar{r}}$ and $b_{\bar{r}}$ is the weight matrix and bias term of the second stage reset gate. $\hat{r}_t$ is the first stage reset gate results, $\bar{r}_t$ is the second stage reset gate results, and r_t is the synthetic reset gate results.

The synthetic update gate z_t controls how much information the model should transmit to the future. The synthetic update gate algorithm of NAGU is denoted by:

$$\hat{r}_t = \sigma\left(W_{\hat{r}}\left[h_{t-1}, x_t\right] + b_{\hat{r}}\right) \tag{6}$$

$$\bar{r}_t = \varphi\left(W_{\bar{r}}\left[h_{t-1}, x_t\right] + b_{\bar{r}}\right) \tag{7}$$

$$r_t = \frac{\hat{r} + \bar{r}}{2} \tag{8}$$

where $W_{\hat{z}}$ and $b_{\hat{z}}$ is the first stage update gate weight matrix and bias term, $W_{\bar{z}}$ and $b_{\bar{z}}$ is the second stage update gate weight matrix and bias term. $\hat{z}_t$ is the first stage update gate results, $\bar{z}_t$ is the second stage update gate results, and z_t is the synthetic update gate results.

Attention receives the input sequence x_t and the hidden state h_{t-1} to calculate the matching score α_t denoted by:

$$\alpha_t = \sigma(W_\alpha\left[h_{t-1}, x_t\right] + b_\alpha) \tag{9}$$

where W_α is the weight of the input gate, scale α_t through softmax function to obtain the attention probability distribution U_t denoted by:

$$U_t = \operatorname{softmax}(\alpha_t) \tag{10}$$

In calculating the candidate hidden state, the synthetic reset gate r_t is spliced with the input x_t at the current moment, and the spliced results are weighted and summed. At this point, the weight matrix for calculating the candidate hidden states is the attention probability distribution U_t. The output gate algorithm of NAGU is denoted by:

$$\tilde{h}_t = \tanh(Wx_t + U_t(r_t \times h_{t-1})) \tag{11}$$

$$h_t = (1 - z_t)\, h_{t-1} + z_t \tilde{h}_t \tag{12}$$

where tanh is the tanh activation function. Wx_t is the weight of the output gate.

E. EEMD-NAGU

The overall structure diagram of EEMD-NAGU is shown in Figure 6. This paper uses soybean futures price data as the research object, and the EEMD-NAGU model is separated into six parts. The first part is the input layer. This paper takes the soybean futures data as the research object and DJIA, S&P 500, and NASDAQ as the influencing factors.

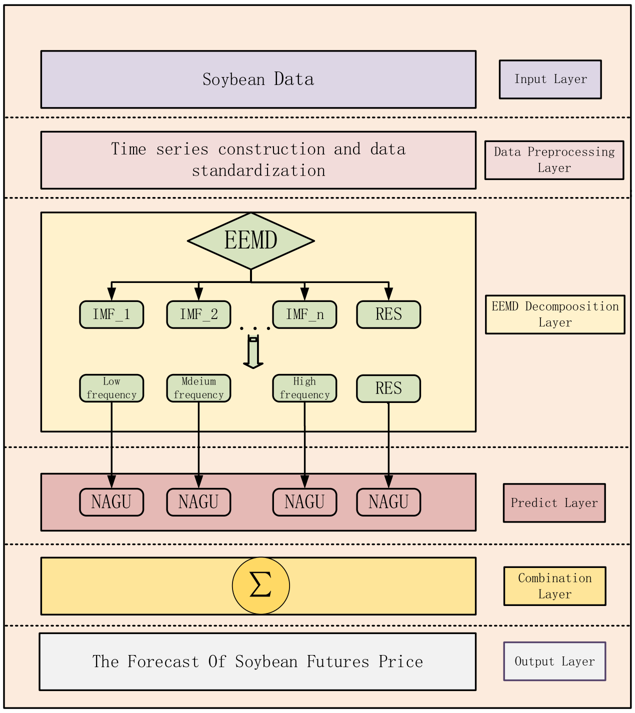

FIGURE 6. EEMD-NAGU structure.

The second part is the data preprocessing layer. In order to guarantee the reliability of the prediction results and enhance the accuracy of the prediction results, and preprocessing operations such as normalization of the raw data and three-dimensional time series construction are required. The third part is the EEMD decomposition layer. EEMD decomposes the data into IMF components of each frequency and residual sequence, then IMF components are reconstructed into three components of low-frequency, medium-frequency, and high-frequency for subsequent input into the neural network. The fourth part is the prediction layer, using NAGU to forecast the four components separately. The fifth part is the combination layer. The prediction results of each component are summarized. Finally, the output layer scale the summarized results to obtain the final prediction result.

IV. EXPERIMENTS

A. EXPERIMENTAL ENVIRONMENT

The experimental environment is Windows 10, the CPU is Intel(R) Core(TM) i5-10400F, NVIDIA GTX2060, and the memory is 24.00GB. The compilation environment is Python 3.8, the compiler is PyCharm 2021.1.1 × 64.

B. DATA COLLECTION AND ANALYSIS

The data used in this experiment is obtained from the https://tushare.pro/ website. Tushare is an open financial big data platform. In this experiment, DJIA, S&P 500, U.S. Dollar Index (USDX), and NASDAQ are selected as impact factors. This paper uses the Maximum Information Coefficient (MIC), which is better for nonlinear data, to analyze the correlation of influencing factors. For a set of two-dimensional variables (X, Y), divide them into m and n points in their respective directions, and the two-dimensional variables form a lattice diagram of mn size on the plane. Under the two-dimensional variable sample data set D, the calculation process of MIC is denoted by:

$$M(D)_{m,n} = \frac{MI^*(D, m, n)}{\log_2 \min\{m, n\}} \tag{13}$$

$$M_{\mathrm{MIC}}(X, Y) = \max_{mn<B(N)} \{M(D)_{m,n}\} \tag{14}$$

where $M_{\mathrm{MIC}}(X, Y)$ is the maximum information coefficient value of the two-dimensional variable (X, Y), $MI^*(D, m, n)$ is the maximum mutual information of two-dimensional variables under the data set D, N is the number of samples, $B(N)$ is a function of the number of samples, and its value is usually set to $\mathrm{N}^{0.6}$ [34].

The value of $M_{\mathrm{MIC}}(X, Y)$ is on [0,1]. The closer the matter is to 1, the stronger the correlation between the two variables is, and vice versa. According to the above MIC theory, Table 2 shows the results. Finally, we selected the DJIA, S&P 500, and NASDAQ as influencing factors. From Figures 7, 8, and 9, it can be seen that there is a strong correlation between them.

TABLE 2. Calculation results of MIC.

Data	DJIA	S&P 500	USDX	SSE Index	NASDAQ
MIC	0.8288	0.8716	0.4616	0.4402	0.8498

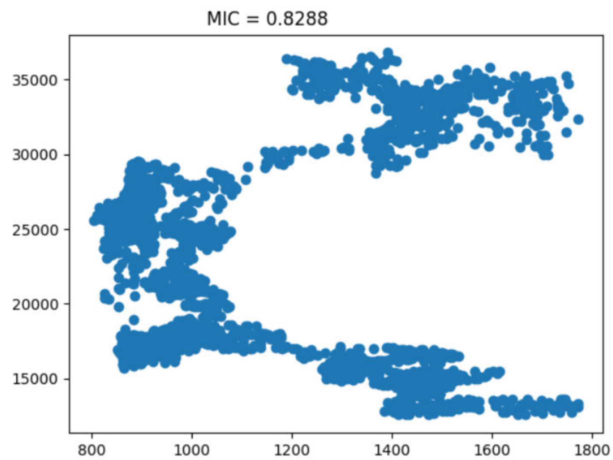

FIGURE 7. DJIA index.

This experiment selects the soybean futures trading data of the Dalian Commodity Exchange from June 26, 2012, to May 10, 2023, and the data of influencing factors in the same time range as the experimental data. There are 2750 records in total; 2200 records are used as training sets, 275 as verification sets, and 275 as test sets.

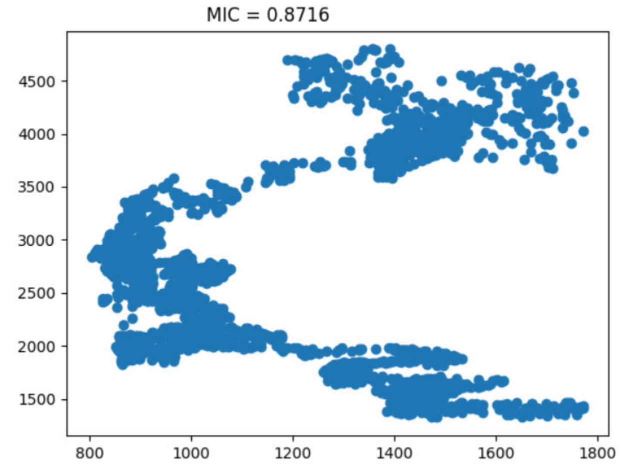

FIGURE 8. S&P 500 index.

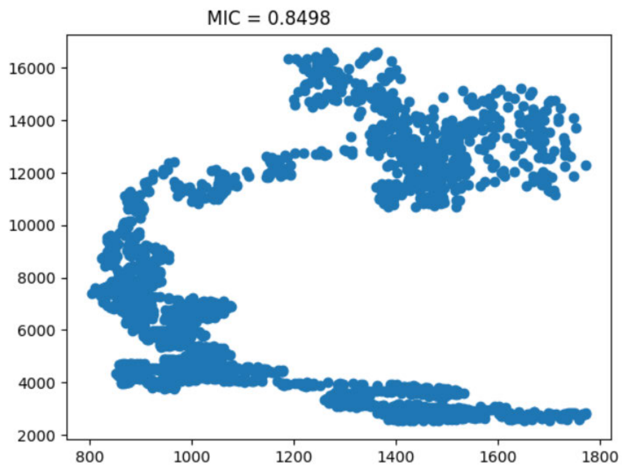

FIGURE 9. NASDAQ index.

C. DATA PREPROCESSING AND STANDARDIZATION

Due to differences in opening and closing dates between domestic and foreign markets, it is necessary to fill in the missing parts of the original data. The process of data standardization is denoted by:

$$V_t = \frac{V_{t-1} + V_{t-2}}{2}, \tag{15}$$

where V_t is the missing value at time t, V_{t-1} is the data of the previous trading day at time t, V_{t-2} is the data of the first two trading days at time t. Finally, the influencing factors and soybean futures data will be combined.

When the values of different features in the data set are greatly different, it is easy to cause the algorithm to not fully learn the data features. So, in this experiment, MinMaxScaler is selected to process the original data [35], and feature scaling is scaled to a specific interval [0,1].

D. BUILDING TIME SERIES DATA

As shown in Figure 10, the step size in this experiment is 1, and the sequence length is 7. Assuming that the data length is 2750, the experimental data is three-dimensionally constructed, and the final data dimension is (2744, 7, 12), which means the data is divided into x-6 data groups with seven rows and eight columns. Divide the first data to the seventh data into a group according to the time series, then form another data group from the second to the eighth data, and so on. Each data set has twelve columns representing the opening price, highest price, lowest price, settlement price, closing price, DJIA Index, S&P 500 Index, NASDAQ Index, low-frequency, medium-frequency, high-frequency, and residual sequence of soybean futures. Finally, the constructed time series data is divided into training, verification, and test sets. The prediction model takes the first three pieces of data in each group as features and the closing price of the fifth piece as a label for training and evaluation.

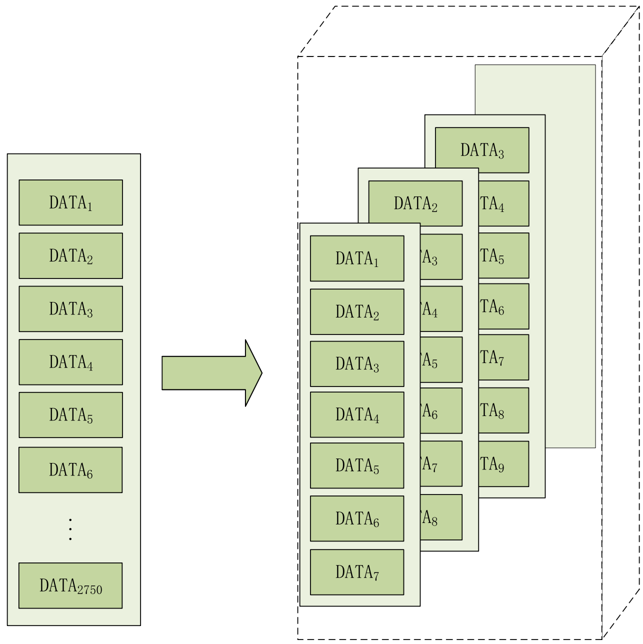

FIGURE 10. Time Series Construction Process.

E. MODELS ADJUSTMENT AND VALIDATION

Generally, the generalization performance of neural networks exceeds that of linear models. But many parameters of the neural networks need to be fine-tuned. We need to find the optimal combination of hyperparameters to ensure the best generalization. In this experiment, we first use the experience to find a value range of the optimal hyperparameter and then use the grid search optimization algorithm to optimize the parameters further. Within a given value range, try various combinations of all hyperparameters until the optimal result is found. The final model parameters are detailed in Table 3.

F. EXPERIMENTAL ANALYSIS

To verify the accuracy of the soybean futures price prediction model based on EEMD-NAGU, using SVR, LSTM, GRU, NAGU, Attention-LSTM, Attention-GRU,

TABLE 3. Model Parameter.

Model	Layer	Parameters
SVR	SVR	kernel='rbf', C=1.0, epsilon=0.1
LSTM	LSTM	activation= 'tanh', units=64 batch size=64, epochs=50
GRU	GRU	activation= 'tanh', units=64 batch size=64, epochs=50
NAGU	NAGU	activation='sigmoid', units=64 batch size=64, epochs=50
Attention-LSTM	Conv1D	filters=64, padding='valid', kernel_size=2, activation= 'tanh'
	MaxPooling1D	padding='valid', pool_size=1
	LSTM	activation= 'tanh', units=64 batch size=64, epochs=50
Attention-GRU	Conv1D	filters=64, padding='valid', kernel_size=2, activation= 'tanh'
	MaxPooling1D	padding='valid', pool_size=1
	NAGU	activation= 'sigmoid', units=64 batch size=64, epochs=50
Attention-NAGU	Conv1D	filters=64, padding='valid', kernel_size=2, activation= 'sigmoid'
	MaxPooling1D	padding='valid', pool_size=1
	NAGU	activation1= 'sigmoid', activation2= 'leaky relu', units=64 batch size=64, epochs=50

TABLE 4. Model Comparison Result.

Model	MAE	RMSE	R^2
SVR	46.432	65.561	0.77538
LSTM	42.992	53.568	0.85004
GRU	40.631	50.305	0.86775
NAGU	32.899	40.877	0.91268
Attention-LSTM	40.079	50.464	0.86692
Attention-GRU	40.366	49.601	0.87143
Attention-NGU	34.051	42.082	0.90746
EEMD-LSTM	36.432	45.232	0.89308
EEMD-GRU	32.573	39.446	0.91869
EEMD-NGU	30.714	35.816	0.93296
EEMD-Attention-LSTM	33.119	40.781	0.91309
EEMD-Attention-GRU	30.151	36.596	0.93001
EEMD-Attention-NGU	25.275	31.323	0.94873
EEMD-NAGU	22.739	28.500	0.95755

Attention-NGU, EEMD-LSTM, EEMD-GRU, EEMD-NGU, EEMD-Attention-LSTM, EEMD-Attention-GRU, and EEMD-Attention-NGU as a comparison model.

This paper uses RMSE, MAE, and R^2 to evaluate the performance of model. Table 4 presents the evaluation results using the test dataset. From the experimental results, the prediction effect of the traditional regression model SVR is the worst; in the deep learning model, NAGU is superior to Attention-GRU, and EEMD-NAGU has the best prediction effect.

It can be seen from Table 4 that compared with Attention-GRU, the prediction and evaluation results of NAGU have reduced MAE by 7.467, reduced RMSE by 8.724, and increased R^2 by 0.04125. So, the NAGU has improved prediction accuracy compared to ordinary GRU. The predicted and real values of GRU, Attention-GRU, Attention-NGU, and NAGU are shown in Figure 11.

It can be seen from Table 4 that based on the prediction models of Attention-GRU and NAGU, the prediction effect is improved after introducing EEMD. Compared with EEMD-Attention-GRU, the prediction and evaluation results of EEMD-NAGU show that MAE is reduced by 7.412, RMSE is reduced by 8.096, and R^2 is increased by 0.02754.

The predicted and real values of Attention-GRU, NAGU, EEMD-Attention-GRU, and EEMD-NAGU are shown in Figure 12.

V. DISCUSSION

The experimental results show that the EEMD-NAGU hybrid model is the best comprehensive evaluation index for predicting soybean futures price. Compared to GRU, the prediction effect of NAGU is improved because of the improvement of the update gate and reset gate. After combining EEMD and NAGU, since EEMD decomposes the original data, it can extract the frequency and time domain characteristics of the signal, which helps to understand the nature and changing laws of the data.

EEMD-NAGU improves the prediction accuracy of soybean futures price by:

(1) The advantages of gate technology and neural network are kept in NAGU. NAGU uses gate technology to learn historical data effectively and reduce the occurrence of "vanishing gradient" and "exploding gradient". Therefore, this model has certain benefits in learning the nonlinear features of time series data.

(2) After EEMD decomposition and reconstruction of the original data, the low-frequency, medium-frequency, high-frequency, and residual sequences are obtained, which fixes the number of inputs of the model.

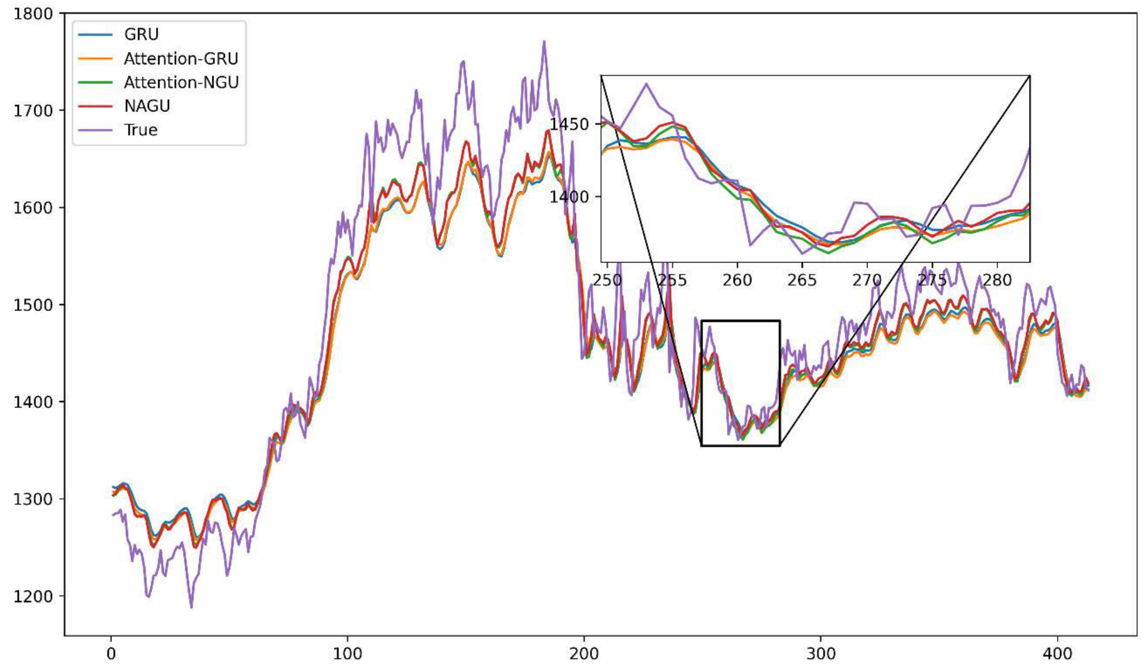

FIGURE 11. GRU, Attention-GRU, Attention-NGU and NAGU predicted and true values.

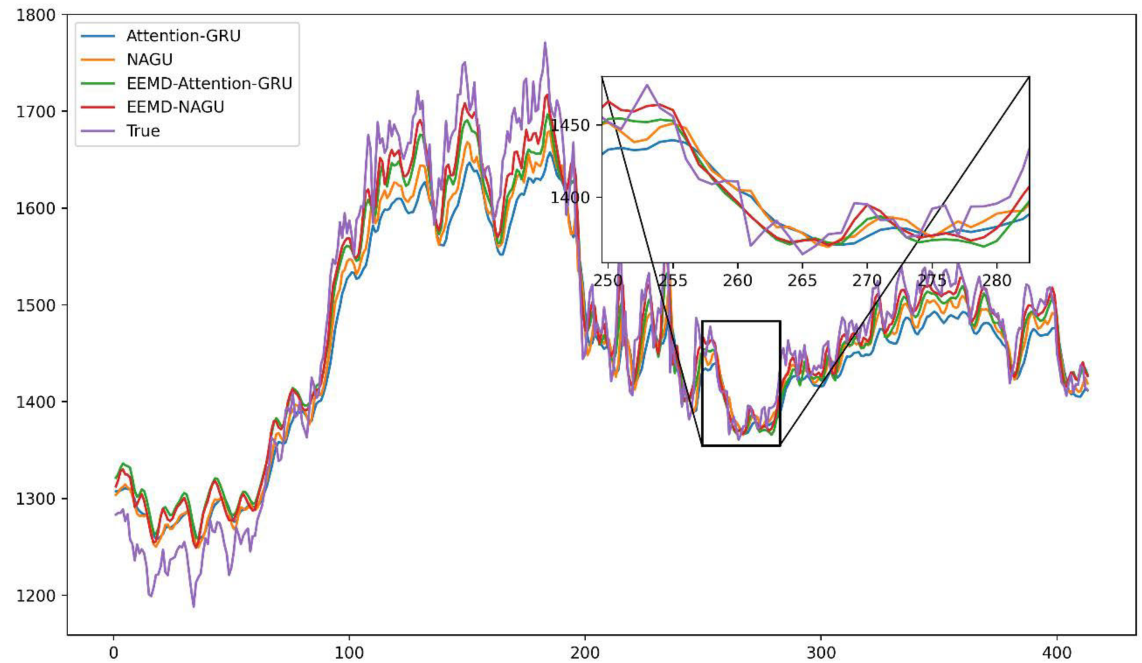

FIGURE 12. Attention-GRU, NAGU, EEMD-Attention-GRU and EEMD-NAGU predicted and true values.

(3) NAGU adds Leaky ReLU function calculations to the reset gate and update gate, respectively, and incorporates the Attention mechanism into the model, which can better capture the connection in historical data so that the model can adaptively focus on and remember essential parts during the learning process, so as to improve the prediction accuracy.

VI. CONCLUSION

This paper proposes the hybrid model of EEMD-NAGU to predict the soybean futures price. The model uses the EEMD to decompose the data, denoise historical data, and fix the number of model inputs; then uses NAGU to extract the time characteristics of the data. NAGU embeds the Attention mechanism in the gate unit to focus on specific features adaptively, and improves the memory ability of the gate unit. Meanwhile, NAGU also improved the gate unit to make the gate unit more sensitive to historical data learning. A comprehensive evaluation is conducted with thirteen comparison models, and the predicted results of EEMD-NAGU are all the best. In summary, the advantages of the model are as follows:

(1) Compared with GRU, NAGU has changed the gate structure. NAGU introduces two-stage reset gate and two-stage update gate in the reset gate and updates gate, respectively, which can improve its performance. And the Attention mechanism is incorporated into the model, which improves the rationality of parameter assignment learning to a certain extent.
(2) Reconstruct the soybean futures price time series data after EEMD decomposition to obtain low-frequency, medium-frequency, high-frequency, and residual sequence, which fixes the number of inputs. It not only reduces the time of model prediction, but also fully retains the information in the data.
(3) Aiming at the deficiency of a single forecasting model, a hybrid model based on EEMD-NAGU for soybean futures price prediction is proposed. Through experimental comparison with other models, the EEMD-NAGU hybrid model has the best prediction accuracy for soybean futures price.

Although the EEMD-NAGU model can improve the accuracy of soybean futures price prediction, it has the disadvantages of a more complex model and a longer training time. Since political events and policies influence soybean futures price, our next work considers using natural language processing techniques to process information such as current news and forum comments, and use them as input terms of the prediction model to improve prediction accuracy further.

EEMD-NAGU can make correct judgments and scientific decisions for relevant practitioners in the soybean industry, help practitioners reduce risks, and play a specific guiding role in stabilizing the soybean futures market. The model of EEMD-NAGU can be widely used for time series data prediction in agriculture, finance, and other fields, such as wheat, corn, gold, oil, etc.

REFERENCES

[1] L.-T. Zhao, K. Liu, X.-L. Duan, and M.-F. Li, ''Oil price risk evaluation using a novel hybrid model based on time-varying long memory,'' *Energy Econ.*, vol. 81, pp. 70–78, Jun. 2019, doi: 10.1016/j.eneco.2019.03.019.

[2] C. Z. Yao, C. Liu, and W. J. Ju, ''Multifractal analysis of the WTI crude oil market, US stock market and EPU,'' *Physica A, Stat. Mech. Appl.*, vol. 550, pp. 124096–124113, Jul. 2020, doi: 10.1016/j.physa.2019.124096.

[3] A. K. Tiwari, Z. Umar, and F. Alqahtani, ''Existence of long memory in crude oil and petroleum products: Generalised Hurst exponent approach,'' *Res. Int. Bus. Finance*, vol. 57, Oct. 2021, Art. no. 101403, doi: 10.1016/j.ribaf.2021.101403.

[4] L. Xu, L. Hou, Z. Zhu, Y. Li, J. Liu, T. Lei, and X. Wu, ''Mid-term prediction of electrical energy consumption for crude oil pipelines using a hybrid algorithm of support vector machine and genetic algorithm,'' *Energy*, vol. 222, May 2021, Art. no. 119955, doi: 10.1016/j.energy.2021.119955.

[5] Y. Bengio, P. Simard, and P. Frasconi, ''Learning long-term dependencies with gradient descent is difficult,'' *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994, doi: 10.1109/72.279181.

[6] J. Sun, L. Di, Z. Sun, Y. Shen, and Z. Lai, ''County-level soybean yield prediction using deep CNN-LSTM model,'' *Sensors*, vol. 19, no. 20, p. 4363, Oct. 2019, doi: 10.3390/s19204363.

[7] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, ''Empirical evaluation of gated recurrent neural networks on sequence modeling,'' 2014, *arXiv:1412.3555*.

[8] H. Sun, D. Gui, B. Yan, Y. Liu, W. Liao, Y. Zhu, C. Lu, and N. Zhao, ''Assessing the potential of random forest method for estimating solar radiation using air pollution index,'' *Energy Convers. Manag.*, vol. 119, pp. 121–129, Jul. 2016, doi: 10.1016/j.enconman.2016.04.051.

[9] W. Tong, H. Hong, H. Fang, Q. Xie, and R. Perkins, ''Decision forest: Combining the predictions of multiple independent decision tree models,'' *J. Chem. Inf. Comput. Sci.*, vol. 43, no. 2, pp. 525–531, Feb. 2003, doi: 10.1021/ci020058s.

[10] K. J. Kim, ''Financial time series forecasting using support vector machines,'' *Neurocomputing*, vol. 55, nos. 1–2, pp. 307–319, Sep. 2003, doi: 10.1016/S0925-2312(03)00372-2.

[11] C. Y. Wang, ''Forecast on price of agricultural futures in China based on ARIMA model,'' *Asian J. Agric. Res.*, vol. 11, no. 9, pp. 12–16, Nov. 2016, doi: 10.22004/ag.econ.253255.

[12] C. N. Babu and B. E. Reddy, ''A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data,'' *Appl. Soft Comput.*, vol. 23, pp. 27–38, Oct. 2014, doi: 10.1016/j.asoc.2014.05.028.

[13] M. Yao, Q. Long, D. Zhou, L. Jun, L. Ping, Y. Shi, and Y. Wang, ''AR model based on time series modeling for predicting egg market price in 2021,'' *Agricult. Biotechnol.*, vol. 10, no. 3, pp. 89–93, 2021, doi: 10.19759/j.cnki.2164-4993.2021.03.021.

[14] S. Du, Z. Xu, and J. Lv, ''An EMD- and GRU-based hybrid network traffic prediction model with data reconstruction,'' in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, Jun. 2021, pp. 1–7, doi: 10.1109/ICCWorkshops50388.2021.9473822.

[15] K. Huarng and T. H.-K. Yu, ''The application of neural networks to forecast fuzzy time series,'' *Phys. A, Stat. Mech. Appl.*, vol. 363, no. 2, pp. 481–491, May 2006, doi: 10.1016/j.physa.2005.08.014.

[16] P. Coulibaly and C. K. Baldwin, ''Nonstationary hydrological time series forecasting using nonlinear dynamic methods,'' *J. Hydrol.*, vol. 307, nos. 1–4, pp. 164–174, Jun. 2005, doi: 10.1016/j.jhydrol.2004.10.008.

[17] H. Ouyang, X. Wei, and Q. Wu, ''Agricultural commodity futures prices prediction via long- and short-term time series network,'' *J. Appl. Econ.*, vol. 22, no. 1, pp. 468–483, Jan. 2019, doi: 10.1080/15140326.2019.1668664.

[18] C. Wang and Q. Gao, ''High and low prices prediction of soybean futures with LSTM neural network,'' in *Proc. IEEE 9th Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, Nov. 2018, pp. 140–143, doi: 10.1109/ICSESS.2018.8663896.

[19] W. Y. An, L. Wang, and Y. R. Zeng, ''Text-based soybean futures price forecasting: A two-stage deep learning approach,'' *J. Forecasting*, vol. 38, pp. 549–562, Sep. 2022, doi: 10.1002/for.2909.

[20] Y. Huang and Y. Deng, ''A new crude oil price forecasting model based on variational mode decomposition,'' *Knowl.-Based Syst.*, vol. 213, Feb. 2021, Art. no. 106669, doi: 10.1016/j.knosys.2020.106669.

[21] S. Amalia and A. Dhini, ''Indonesia's food commodity price forecasting using recurrent neural networks,'' in *Proc. Int. Conf. Comput., Commun., Secur. Intell. Syst. (IC3SIS)*, Jun. 2022, pp. 1–6, doi: 10.1109/IC3SIS54991.2022.9885249.

[22] P. Mei, M. Li, Q. Zhang, G. Li, and L. Song, ''Prediction model of drinking water source quality with potential industrial-agricultural pollution based on CNN-GRU-attention,'' *J. Hydrol.*, vol. 610, Jul. 2022, Art. no. 127934, doi: 10.1016/j.jhydrol.2022.127934.

[23] S. Yang, H. Guo, and J. Li, ''CNN-GRUA-FC stock price forecast model based on multi-factor analysis,'' *J. Adv. Comput. Intell. Intell. Inform.*, vol. 26, no. 4, pp. 600–608, 2022, doi: 10.20965/jaciii.2022.p0600.

[24] L. Yu, S. Wang, and K. K. Lai, ''Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm,'' *Energy Econ.*, vol. 30, no. 5, pp. 2623–2635, Sep. 2008, doi: 10.1016/j.eneco.2008.05.003.

[25] X. Zhang, T. Xue, and H. Eugene Stanley, ''Comparison of econometric models and artificial neural networks algorithms for the prediction of Baltic dry index,'' *IEEE Access*, vol. 7, pp. 1647–1657, 2019, doi: 10.1109/ACCESS.2018.2884877.

[26] Y. Zhang, C. Li, Y. Jiang, L. Sun, R. Zhao, K. Yan, and W. Wang, ''Accurate prediction of water quality in urban drainage network with integrated EMD-LSTM model,'' *J. Cleaner Prod.*, vol. 354, Jun. 2022, Art. no. 131724, doi: 10.1016/j.jclepro.2022.131724.

[27] B. Wang and J. Wang, "Energy futures and spots prices forecasting by hybrid SW-GRU with EMD and error evaluation," *Energy Econ.*, vol. 90, Aug. 2020, Art. no. 104827, doi: 10.1016/j.eneco.2020.104827.
[28] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London. Ser. A, Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998, doi: 10.1098/rspa.1998.0193.
[29] Z. Cen and J. Wang, "Forecasting neural network model with novel CID learning rate and EEMD algorithms on energy market," *Neurocomputing*, vol. 317, pp. 168–178, Nov. 2018, doi: 10.1016/j.neucom.2018.08.021.
[30] Z. Wu and N. E. Huang, "Ensemble empirical mode decomposition: A noise assisted data analysis method," *Adv. Adapt. Data Anal.*, vol. 1, no. 1, pp. 1–41, Jan. 2009, doi: 10.1142/S1793536909000047.
[31] Z. Ma, G. Wen, and C. Jiang, "EEMD independent extraction for mixing features of rotating machinery reconstructed in phase space," *Sensors*, vol. 15, no. 4, pp. 8550–8569, Apr. 2015, doi: 10.3390/s150408550.
[32] Y. Guo, X. Cao, B. Liu, and K. Peng, "El Niño index prediction using deep learning with ensemble empirical mode decomposition," *Symmetry*, vol. 12, no. 6, p. 893, Jun. 2020, doi: 10.3390/sym12060893.
[33] A. Pletzer, B. Jamroz, R. Crockett, and S. Sides, "Compact cell-centered discretization stencils at fine-coarse block structured grid interfaces," *J. Comput. Phys.*, vol. 260, pp. 25–36, Mar. 2014, doi: 10.1016/j.jcp.2013.12.020.
[34] D. N. Reshef, Y. A. Reshef, H. K. Finucane, S. R. Grossman, G. McVean, P. J. Turnbaugh, E. S. Lander, M. Mitzenmacher, and P. C. Sabeti, "Detecting novel associations in large data sets," *Science*, vol. 334, no. 6062, pp. 1518–1524, Dec. 2011, doi: 10.1126/science.1205438.
[35] W. Shu and Q. Gao, "Forecasting stock price based on frequency components by EMD and neural networks," *IEEE Access*, vol. 8, pp. 206388–206395, 2020, doi: 10.1109/ACCESS.2020.3037681.

JINDIAN LIU is currently pursuing the master's degree with the Hebei University of Science and Technology. His current research interests include machine learning and deep learning.

BO ZHANG received the B.Eng. degree in computer science and technology from the Hebei Normal University of Science and Technology, China, in 2008, and the M.Sc. degree in computer technology from the Hebei University of Science and Technology, China, in 2015. She is currently an Associate Professor with the Hebei Polytechnic Institute, Shijiazhuang, Hebei, China. Her current research interests include machine learning, deep learning, data mining, software development, and software testing.

TIANHU ZHANG is currently pursuing the master's degree with the Hebei University of Science and Technology. His current research interests include machine learning and deep learning.

JINGYANG WANG received the B.Eng. degree in computer software from Lanzhou University, China, in 1995, and the M.Sc. degree in software engineering from the Beijing University of Technology, China, in 2007. He is currently a Professor with the Hebei University of Science and Technology, Shijiazhuang, Hebei, China. His current research interests include machine learning, deep learning, natural language processing, and big data processing.