

## Research Article

# Prediction of Financial Time Series Based on LSTM Using Wavelet Transform and Singular Spectrum Analysis

Qi Tang <sup>1</sup>, Ruchen Shi <sup>1</sup>, Tongmei Fan <sup>1</sup>, Yidan Ma <sup>2</sup> and Jingyan Huang <sup>1</sup>

<sup>1</sup>School of Economics and Management, Southeast University, Nanjing, China

<sup>2</sup>School of Cyber Science and Engineering, Southeast University, Nanjing, China

Correspondence should be addressed to Qi Tang; [regantangqi@seu.edu.cn](mailto:regantangqi@seu.edu.cn)

Received 11 March 2021; Revised 6 May 2021; Accepted 23 May 2021; Published 9 June 2021

Academic Editor: Jie Chen

Copyright © 2021 Qi Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to further overcome the difficulties of the existing models in dealing with the nonstationary and nonlinear characteristics of high-frequency financial time series data, especially their weak generalization ability, this paper proposes an ensemble method based on data denoising methods, including the wavelet transform (WT) and singular spectrum analysis (SSA), and long-term short-term memory neural network (LSTM) to build a data prediction model. The financial time series is decomposed and reconstructed by WT and SSA to denoise. Under the condition of denoising, the smooth sequence with effective information is reconstructed. The smoothing sequence is introduced into LSTM and the predicted value is obtained. With the Dow Jones industrial average index (DJIA) as the research object, the closing price of the DJIA every five minutes is divided into short term (1 hour), medium term (3 hours), and long term (6 hours), respectively. Based on root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE), and absolute percentage error standard deviation (SDAPE), the experimental results show that in the short term, medium term, and long term, data denoising can greatly improve the stability of the prediction and can effectively improve the generalization ability of LSTM prediction model. As WT and SSA can extract useful information from the original sequence and avoid overfitting, the hybrid model can better grasp the sequence pattern of the closing price of the DJIA.

## 1. Introduction

As the worldwide largest economy, the US has advanced statistical constitutions and a mature financial supervision system, whose financial data are comprehensive, accurate, and credible. At the same time, the US stock market cooperates with other markets in an efficient way and plays an important role in the US financial system, and all these characteristics make the market a good model. On the one hand, global stock markets react quickly to the tendency of this market, especially in the case of unusually high market volatility. On the other hand, most economic theories and assumptions are based on the study of a developed financial system with a larger and more active stock market, a more mature economy, and a more effective financial supervision system. As a representative of developed markets, the US market is also the most favorable object for empirical or

theoretical propositions in academic research. The three major stock indexes in the United States are the Dow Jones Index [1], the Standard and Poor's 500 Index, and the Nasdaq Composite Index. The most famous of these indexes is the Dow Jones Index. The importance of the Dow Jones Index has been further recognized in global markets beyond its role in the domestic market. The 30 companies that make up the Dow, such as Citigroup, Coca-Cola, General Motors, and Intel, are prestigious multinational corporations. They cover a wide range of large industries and their performance is behind the global economy. Therefore, forecasting the Dow Jones index is of great significance to the entire financial system.

At present, there are two categories of prediction models that are suitable for financial time series: parametric model and nonparametric model. Autoregressive (AR), moving average (MA), autoregressive moving average (ARMA), and

autoregressive comprehensive moving average (ARIMA) are typical models of parameter types [2]. However, all of the above models can only be used if the predicted time series conforms to the statistical assumptions they have. Therefore, the parametric model has limitations to some degree. Additionally, the parametric model is more functional for the time series with linear characteristics, but the financial time series of DJIA price is characterized with nonlinearity and is highly fluctuating, which will cause so many estimated parameters and increase the complexity of the model. Due to the limitations of the parametric model and features of DJIA price, the parametric model is unsuitable for the Dow Jones index price forecast.

Stock price prediction is a pertinent and tricky problem that has caught the interest of many scholars. In order to adapt to the characteristics of DJIA and ensure high accuracy, nonparametric models are used. Specifically, machine learning (ML) and deep learning (DL) models are used for DJIA prediction. Ahmed [3] found that machine learning methods can perform better than traditional econometrics methods. Sun et al. [4] have compared the accuracy of the echo state network (ESN) and the long short-term memory model (LSTM) on Kweichow Moutai's stock price prediction. The empirical results found that the ESN model can alleviate the problems of the low accuracy of the deep learning model, the slow convergence speed, and the complex network structure. At the same time, it has higher prediction accuracy. Cao et al. [5] used the convolutional neural network (CNN) model to predict the stock index and finally established the CNN-support vector machine (SVM) stock index prediction model. The empirical results show that the effect of using a neural network to predict the financial time series is better than the traditional measurement method. Ustali et al. [6] have used Artificial Neural Network (ANN), Random Forest (RF) algorithm, and XGBoost algorithm to estimate the future price of company stocks listed on the Istanbul Stock Exchange (BIST) United Joint Stock Company (BIST) 30 Index. The empirical results show that although the results of XGBoost and Random Forest algorithms are similar, the prediction results of XGBoost are slightly better. Moreover, the performance of both models is better than ANN.

LSTM (Long Short-Term Memory) network is one of the cyclic neural networks (RNNs). Sepp Hochreiter and Jürgen Schmidhuber first presented this algorithm in Neural Computing [7]. It has better performance than ordinary RNN in data processing and prediction. Considering the excellent performance of the LSTM network in time series, Jiang et al. [8], taking the daily data of Shanghai Composite Index and Dow Jones Index as the research object, respectively, uses RNNs and LSTM to build the model. Then they found through experiments that the LSTM model prevails over the RNNs model for the neural network model. However, this model is still not that well applicable to Dow Jones Index. Actually, to ensure the prediction accuracy of the Dow Jones Industrial Average, more potential factors, such as the influence of policy information, should be taken into consideration. However, the quantification of these parameters is exceedingly difficult. Therefore, another

method is to further mine the information of the sequence itself through the lag term.

In order to make up for the deficiency of a single forecasting model, a new forecasting model-hybrid model is presented. In this category, decomposition methods, such as empirical mode decomposition (EMD) [9] and singular spectrum analysis (SSA) [10], are usually combined with ML and DL based on financial time series prediction models [11–13]. In recent works [14, 15], the superiority of hybrid models has been verified because of their preponderance in identifying time series patterns. As there is a nonlinear relationship between the predicted price of agricultural products and the influencing factors, Jia et al. [16] have designed a neural network model of LSTM-DA (Long Short-Term Memory-Double Attention) which combines the convolutional attention network, the LSTM network, and the attention mechanism. Compared with the traditional signal model, this model can improve the prediction accuracy, and the predicted price index can accurately describe the overall trend of vegetable products in the next week.

However, hybrid models are rarely used to forecast DJIA prices. Meanwhile, lots of previous studies have proved that denoising the high-frequency time series can significantly raise the ability to extend the model and dramatically optimize the prediction results. At present, the empirical data decomposition and noise reduction methods mainly include ensemble empirical mode decomposition (EEMD), singular spectrum analysis (SSA), and wavelet transform decomposition (WT). Although the integrated empirical mode decomposition (EEMD) can suppress the mode aliasing problem to some extent, it may increase the complexity of the sequence. Jung et al. [17] have integrated wavelet transforms and recurrent neural network (RNN) based on artificial bee colony (ABC) algorithm (called ABC-RNN) to establish a system for the purpose of stock price prediction, and it turns out that the performance of the presented model is the best in TAIEX. However, it still has some insufficiencies. For example, this system lacks a solution that includes a feature selection function and addition parameter information provision function to achieve a simplified system organization.

This paper proposes an integrated method to establish a prediction model, which utilizes data denoising methods including the wavelet transform (WT) and singular spectrum analysis (SSA), and long-term short-term memory neural network (LSTM). At the same time, we compare the results of different models. The second part of this paper will introduce the model formula used in this paper, and the third part will introduce the data, model prediction accuracy, and stability results comparison. The fourth part is a summary.

## 2. Model Formulation

In this section, based on the previous work [18], we provide an overview of the main models used in this study, including the LSTM, WT, hybrid WT-LSTM, SSA, and hybrid SSA-LSTM models.

**2.1. Long Short-Term Memory.** LSTM neural network was first proposed by Hochreiter and Schmidhuber, which is widely used to process sequence information owing to its advantages in discovering long-term dependencies. Therefore, it is theoretically feasible to establish an LSTM neural network model for financial high-frequency time series data. The structure of each neuron in LSTM is shown in Figure 1, and its internal structure is constituted of a cell and three gates. Cell records the state of neurons, and the function of input gate and output gate is to receive, output, and modify parameters. Forget gate controls the forgotten degree of the previous state of neurons; that is, it determines the information to be removed from the cell. The selection of activation function is an important part in the process of training a neural network, which can make the neural network learn the nonlinear factors in the data. The activation function used in this paper is the traditional sigmoid activation function and tanh activation function. It mainly includes the following stages.

The first stage is eliminating part of the information from the cell through the forget gate.

$$f_t = \sigma(b_f + W_f x_t + U_f h_{t-1}), \quad (1)$$

where  $x_t$  represents the input of the current cell,  $h_{t-1}$  represents the output of the previous cell, and  $\sigma$  represents the sigmoid activation function which reads the information of  $x_t$  and  $h_{t-1}$  and outputs a value between 0 and 1.

The second stage is updating the status of information in the cell.

$$\begin{aligned} i_t &= \sigma(b_i + W_i x_t + U_i h_{t-1}), \\ \tilde{C}_t &= \tanh(b_c + W_c x_t + U_c h_{t-1}), \\ C_t &= f_t \times C_{t-1} + i_t \times \tilde{C}_t, \end{aligned} \quad (2)$$

where  $C_{t-1}$  indicates the status of old cell information.

The third stage is outputting the information controlled by output gate. Firstly, we run the sigmoid layer to determine which part of the cell state outputs. Secondly, we process the cell state through the tanh activation function (get a value between -1 and 1) and multiply it with the output of the sigmoid layer. In the end, we will only output the part of the controlled information.

$$\begin{aligned} O_t &= \sigma(b_o + W_o x_t + U_o h_{t-1}), \\ h_t &= O_t \times \tanh(C_t). \end{aligned} \quad (3)$$

For the purpose of verifying the effectiveness and versatility of WT and SSA filtering scheme, the most common LSTM neural network structure is adopted in this research. The characteristic quantity is selected as the most basic highest price in the last five minutes, lowest price in the last five minutes, closing price in the last five minutes, and opening price in the last five minutes. Specifically, the main structure of the LSTM neural network in this paper includes a 150-node LSTM layer, a 50-node LSTM layer, and a fully connected layer. Besides, a dropout layer is introduced to this model in order to compare the performance of the dropout layer and data

denoising methods. The calculation diagram structure of the LSTM neural network constructed in this paper is shown in Figure 2. The dotted box represents the neural network structure.

**2.2. Wavelet Transform Analysis.** Dow Jones Index is susceptible to a large number of factors such as economic development, policy changes, and investor sentiment. They usually contain lots of noise and are characterized by nonlinearity. To raise the ability to extend of the model, the noise data should be filtered out when the deep neural network is used to process the nonlinear data. Wavelet analysis can carry out multiscale refined analysis of signals through operation functions such as stretching and shifting, effectively eliminating noise contained in the data and retaining the characteristics of original signals [19]. Accordingly, this paper intends to use wavelet decomposition and reconstruction for data preprocessing of financial time series, which is shown in Figure 3, and adopts “wavelet denoising” to eliminate the high-frequency components of noise in time series, so as to weaken the influence of short-term noise disturbance on neural network structure and improve the prediction performance of the model.

Wavelet decomposition decomposes each input signal into a signal of low frequency and a signal of high frequency and merely decomposes the part of low frequency. Assume that  $C_0$  is the original financial time series signal;  $C_1, C_2 \dots C_l$  and  $D_1, D_2 \dots D_l$  are the first, second, and L-layer low-frequency and high-frequency signals. Then it can be mathematically expressed as follows:

$$C_0 = C_l + D_l + D_{l-1} + \dots + D_2 + D_1 \quad (4)$$

In order to denoise, the Mallat wavelet is used to be reconstructed on the basis of the coefficients of the first N layer with low frequency and the coefficients of 1-N layer with high frequency of wavelet decomposition, and the high-frequency part of zero. The low frequency part of wavelet decomposition of financial sequential data reports the general tendency of the series, and the high-frequency part reports the short-term stochastic disturbance of this financial time series. Therefore, on the one hand, setting the high-frequency part to zero can eliminate the noise and smooth the signal. On the other hand, it can also obtain the approximate signal of the primary financial sequential data so as to prevent the excessive learning of the neural network structure caused by short-term stochastic disturbance factors and raise the extrapolation and generalization ability of the model.

**2.3. WT-LSTM.** In order to obtain higher accuracy, LSTM and WT are combined to predict the price of DJIA. Hybrid WT-LSTM consists of the following three phases and the modeling process is shown in Figure 4.

The first stage is WT decomposition: sym wavelet is chosen as the wavelet basis for DJIA closing price forward fractional solution.

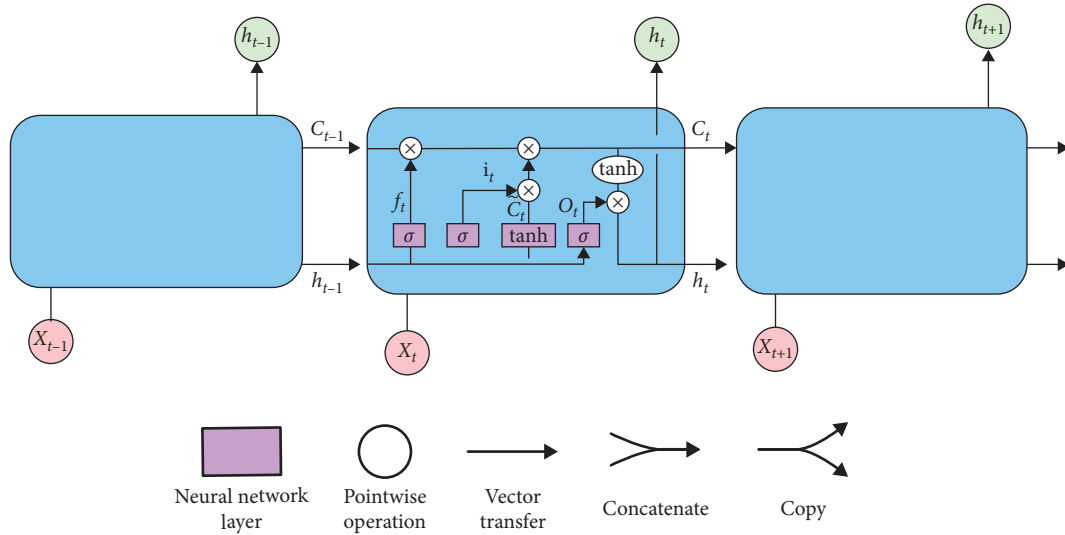


FIGURE 1: LSTM neuron structure.

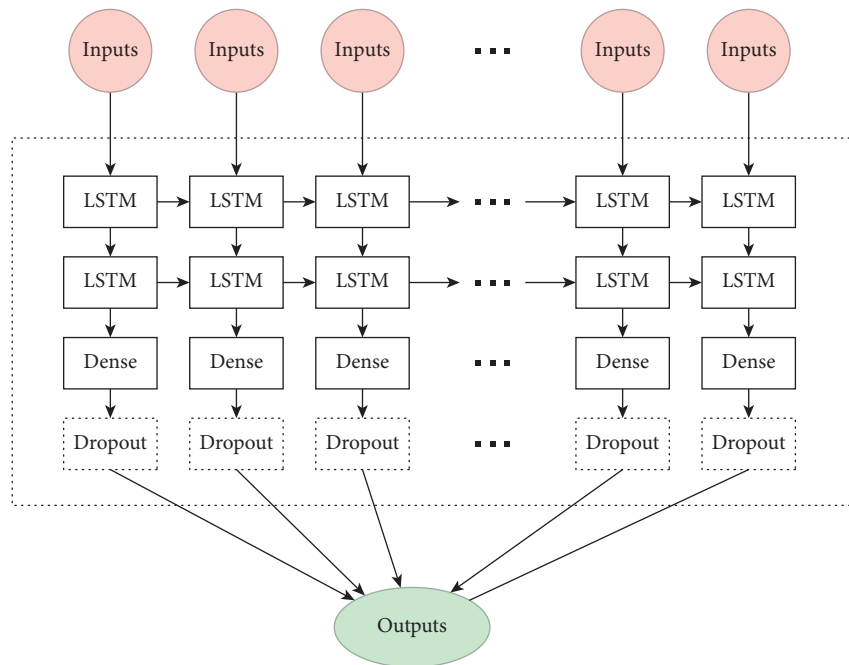


FIGURE 2: LSTM deep neural network structure.

The second stage is time series reconstruction: reconstruct the financial time series data and raise the ability to extend of the prediction model.

The third stage is LSTM prediction: smoothed series  $\tilde{x}_t$  and volume ( $V_t$ ) are the input characteristics of LSTM. According to the partial autocorrelation function (PACF) of  $x_t$ , the time lag of  $\tilde{x}_t$  and  $V_t$  is determined. Then, the final prediction result is obtained.

The sym wavelet is an approximate symmetric orthogonal wavelet function of db wavelet, and it has better symmetry [20]. The more the layers of wavelet decomposition, the better the stability of detail signal and approximate signal. However, it will lead to greater errors in

the decomposition process, so the number of layers should not be too much or too little. Under the circumstance of four layers of decomposition, the effect of denoising is remarkable without eliminating much valid information. Therefore, in this paper, the sym4 wavelet basis is firstly used to divide the closing price of DJIA into four layers, so as to reconstruct the time series data of the gold melting and raise the ability to extend the prediction model. On this basis, the general tendency and market volatility information in the primary data are pre-processed. As a result, the hybrid WT-LSTM model proposed can avoid overfitting and outperform the single LSTM model.

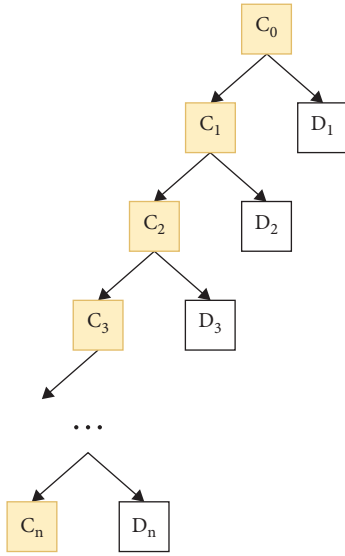


FIGURE 3: Financial time series wavelet decomposition.

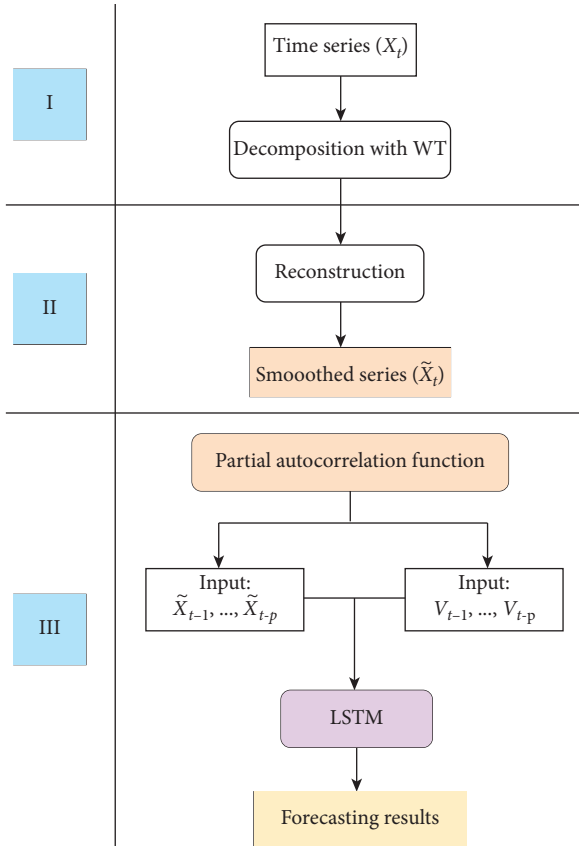


FIGURE 4: Hybrid WT-LSTM model processing process.

**2.4. Singular Spectrum Analysis and SSA-LSTM.** In order to study the effects of the prediction accuracy of out-of-sample data and the prediction ability of future dynamic trend after decomposing and reconstructing financial sequential data by wavelet transform (WT), we also carry out a controlled experiment to study the model prediction after financial time series data decomposition and

reconstruction using singular spectrum analysis (SSA). The effect of SSA can construct a trajectory matrix from the observed financial time series, decompose and reconstruct the trajectory matrix, and extract the different parts of the signal, thus effectively eliminating the noise of the financial sequential data and retaining the features of the primary signal.

Suppose there is a one-dimensional sequence  $x(i)$  ( $i = 1, 2, \dots, n$ ). Given that the embedding dimension is  $m$  ( $m < (n/2)$ ), a time-delay matrix  $X$  can be obtained, and its dimension is  $m \times k$  ( $k = n - m + 1$ ),

$$X = \begin{bmatrix} x_1 & \cdots & x_{n-m+1} \\ \vdots & \ddots & \vdots \\ x_m & \cdots & x_n \end{bmatrix}. \quad (5)$$

Let  $S$  be the  $m \times m$  dimensional covariance matrix of the delay matrix, then

$$S = \begin{bmatrix} s(0) & \cdots & s(m-1) \\ \vdots & \ddots & \vdots \\ s(m-1) & \cdots & s(0) \end{bmatrix}. \quad (6)$$

Singular spectrum analysis is used to decompose the covariance matrix  $S$  to obtain  $m$  singular values  $\lambda_i$  ( $i = 1, 2, \dots, m$ ). Then arrange the obtained  $m$  singular values in descending order. The magnitude of the singular value represents the relative relationship between the signal and the noise. The singular value points with larger values are regarded as signal points, and the points with smaller values are regarded as noise points. The eigenvector  $E^k$  corresponding to  $\lambda_k$  is called the empirical orthogonal projection function. The orthogonal projection coefficient of the sampled signal  $x(i)$  on the eigenvector  $E^k$  is the  $k$ -th principal component:

$$a_i^k = \sum_{j=1}^m x_{i+j} E_j^k \quad (0 \leq i \leq n - m). \quad (7)$$

If each principal component and the empirical orthogonal function are known, the process of reversing the original sequence is as follows:

$$x_{i+j} = \sum_{j=1}^m a_i^k E_j^k \quad (1 \leq j \leq m). \quad (8)$$

So as to reach a comparison model with the WT-LSTM model, we combined LSTM and SSA to predict the DJIA. The hybrid SSA-LSTM consists of the following three stages and the modeling process is shown in Figure 5.

The first stage is SSA decomposition: SSA technique is used to decompose the primary sequential data ( $x_t$ ) into tendency, market volatility, and noise.

The second stage is time series reconstruction: reconstruct the smooth sequence based on the tendency and market volatility signal ( $\tilde{x}_t$ ).

The third stage is LSTM prediction: smooth series  $\tilde{x}_t$  and volume ( $V_t$ ) are the input characteristics of LSTM. In light of the partial autocorrelation function (PACF) of  $x_t$ , the time



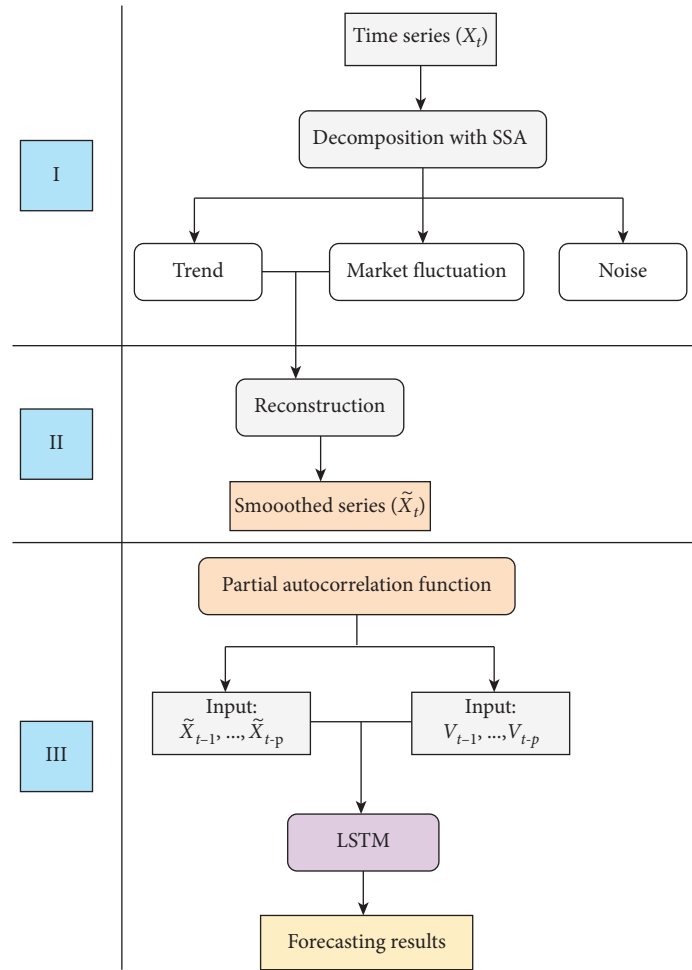


FIGURE 5: Hybrid SSA-LSTM model processing process.

lag of  $\tilde{x}_t$  and  $V_t$  is determined. Then, the final prediction result is obtained.

SSA is used as a pretreatment method to extract effective information of overall tendency and market volatility from primary sequential data.

This paper chooses  $m = 10$ , that is, the original series is decomposed into ten layers. According to the singular values shown in Table 1, it is found that, for this financial high-frequency time series, the first layer already contains more than 99.99% of the sequence information. Therefore, the first layer is selected as the reconstruction of the sequence, and the rest are regarded as noise.

**2.5. Training Method and Optimizer Selection.** The goal of this article is to compare the prediction effect of the closing price of DJIA, so mean square error (MSE) is chosen as the loss function. As for the optimizer, since the Adam algorithm has advantages over other adaptive learning rate algorithms in convergence speed and learning effect, this article uses the Adam optimizer (Adaptive Moment Estimation) for optimization training. Epochs are set to be 10. This article is based on the *Python* language environment and uses TensorFlow as the deep

TABLE 1: Singular values of singular spectrum analysis.

$i$	$\lambda_i$
1	0.999991285112503
2	$5.18930589813891e-06$
3	$1.42616024536684e-06$
4	$6.95916781349973e-07$
5	$4.21779000682764e-07$
6	$2.87333215462852e-07$
7	$2.15658198719760e-07$
8	$1.77801416057950e-07$
9	$1.55528979155002e-07$
10	$1.45403762506482e-07$

learning framework for training, prediction, and comparison.

### 3. Data and Results

**3.1. Data.** In order to research the feasibility and effectiveness of the denoising methods to forecast the actual financial sequential data, this section compares the prediction results of the RNN and LSTM models with a dropout layer and the LSTM model using data denoising methods

with the prediction results of the original LSTM model. The data sample has selected the latest data available in the past years. That is, the time interval is from 00:00 on January 1, 2020, to 23:55 on December 31, 2020, with a time interval of 5 minutes and a total of 19,666 pieces of data. And descriptive statistics of the data are shown in Table 2. The data come from the Wind database.

**3.2. Test Set Prediction Effect Evaluation Index.** To verify the validity of the model, the prediction and verification have been carried out from the three dimensions of short term (1 hour), medium term (3 hours), and long term (6 hours). Root mean square error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE) are used as the prediction accuracy indexes to evaluate the predictive effect of the test set. The smaller the values of the above three indicators are, the higher the prediction accuracy is. The prediction stability is evaluated by the standard deviation of absolute percentage error (SDAPE). The lower the SDAPE value, the higher the reliability of prediction.

$$\begin{aligned} \text{RMSE} &= \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}, \\ \text{MAE} &= \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|, \\ \text{MAPE} &= \frac{1}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100, \\ \text{SDAPE} &= \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{y_i - \hat{y}_i}{y_i} - \text{MAPE} \right)^2}. \end{aligned} \quad (9)$$

Here,  $y_i$  is the actual value and  $\hat{y}_i$  is prediction derived from the forecast model.  $N$  is the number of predictions.

**3.3. Comparative Analysis of Short-Term Forecasting Effects.** For the short-term prediction effect, according to the results shown in Table 3, in terms of prediction accuracy, RNN with a dropout layer is superior to the LSTM model, and RMSE, MAE, and MAPE are decreased by 11.63%, 1.96%, and 1.89%, respectively. Besides, the dropout layer [21] can optimize the prediction effect of the original LSTM model, and RMSE, MAE, and MAPE are decreased by 34.81%, 37.95%, and 37.95%, respectively. SSA-LSTM can optimize the prediction effect of the original LSTM model, and RMSE, MAE, and MAPE are reduced by 70.11%, 65.73%, and 65.72%, respectively. WT-LSTM can optimize the prediction effect of the original LSTM model, reducing RMSE, MAE, and MAPE by 79.55%, 76.91%, and 76.91%, respectively. In terms of prediction stability, RNN with a dropout layer is also superior to the LSTM model and the SDAPE is reduced by 28.24%. Besides, the dropout layer can improve the prediction stability of the original LSTM model and reduce the SDAPE by 30.41%. SSA-LSTM can improve the

TABLE 2: Descriptive statistics of DJIA closing price every 5 minutes.

Obs	Mean	Std. dev	Variance	Skewness	Kurtosis
19666	26876	2513.763	6319004	-0.8633175	3.213287

TABLE 3: 1-hour DJIA closing price forecast results.

	RMSE	MAE	MAPE	SDAPE
LSTM	5.8516916	4.5195833	0.0001481	0.0001218
RNN-dropout	5.1708886	4.4310000	0.0001453	0.0000874
LSTM-dropout	3.8146496	2.8042500	0.0000919	0.0000848
SSA-LSTM	1.7488158	1.5490332	0.0000508	0.0000266
WT-LSTM	1.1966503	1.0434276	0.0000342	0.0000192

prediction stability of the original LSTM model and reduce the SDAPE by 78.14%. WT-LSTM can improve the prediction stability of the original LSTM model and reduce the SDAPE by 84.23%. Through the analysis of short-term prediction, it is found that the generalization effect of filtering to prevent overfitting is better than the effect of the dropout layer to improve the accuracy. At the same time, wavelet transform has a good filtering effect. In terms of prediction accuracy, WT-LSTM can improve the prediction effect of SSA-LSTM model, and RMSE, MAE, and MAPE are reduced by 31.57%, 32.64%, and 32.64%, respectively. In terms of predictive stability, WT-LSTM can improve the predictive stability of the SSA-LSTM model, and SDAPE can be reduced by 27.85%. The prediction results of the four methods are shown in Figures 6–10.

**3.4. Comparative Analysis of Medium-Term Forecasting Effects.** For the medium-term prediction effect, according to the results shown in Table 4 in terms of prediction accuracy, RNN with a dropout layer is superior to the LSTM model, and RMSE, MAE, and MAPE are decreased by 10.64%, 4.54%, and 4.53%, respectively. Besides, the dropout layer can optimize the prediction effect of the original LSTM model, and RMSE, MAE, and MAPE are decreased by 34.54%, 36.15%, and 36.14%, respectively. SSA-LSTM can optimize the prediction effect of the original LSTM model, and RMSE, MAE, and MAPE are reduced by 77.67%, 76.49%, and 76.49%, respectively. WT-LSTM can optimize the prediction effect of the original LSTM model, reducing RMSE, MAE, and MAPE by 75.60%, 73.29%, and 73.28%, respectively. In terms of prediction stability, RNN with a dropout layer is also superior to the LSTM model and the SDAPE is reduced by 21.23%. Besides, the dropout layer can improve the prediction stability of the original LSTM model and reduce the SDAPE by 32.03%. SSA-LSTM can improve the prediction stability of the original LSTM model and reduce the SDAPE by 79.20%. WT-LSTM can improve the prediction stability of the original LSTM model and reduce the SDAPE by 79.79%. Through the analysis of medium-term prediction, we can see that the generalization effect of filtering to prevent overfitting is better than the effect of the dropout layer to improve the accuracy. At the same time, singular spectrum analysis also has a good effect on filtering.

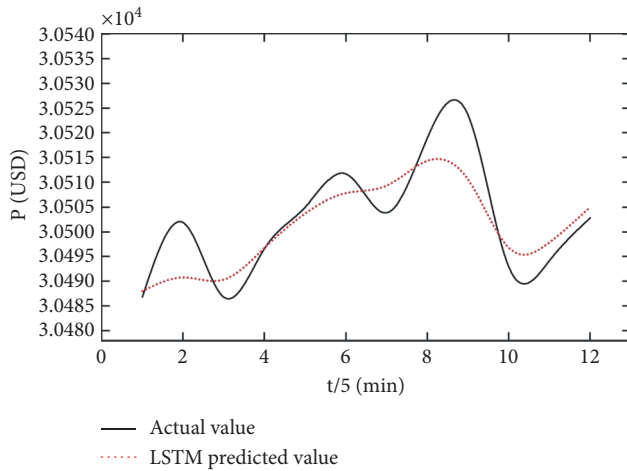


FIGURE 6: Short-term LSTM prediction.

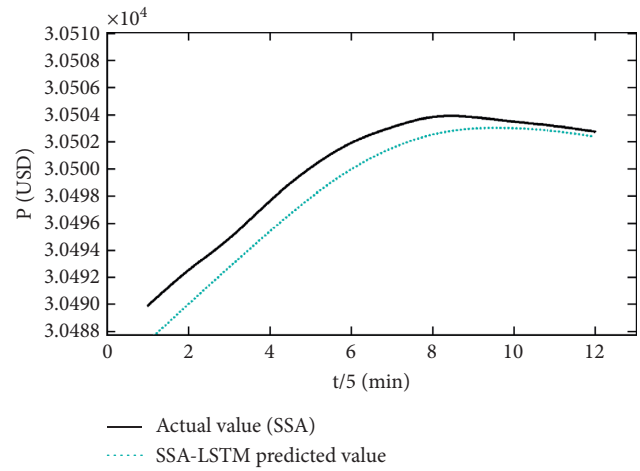


FIGURE 9: Short-term SSA-LSTM prediction.

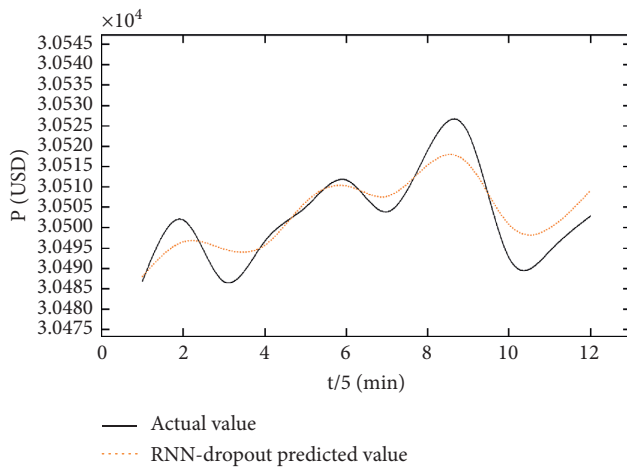


FIGURE 7: Short-term RNN-dropout prediction.

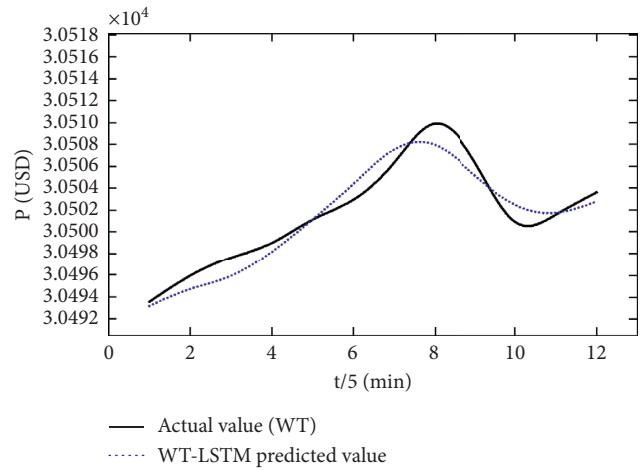


FIGURE 10: Short-term WT-LSTM prediction.

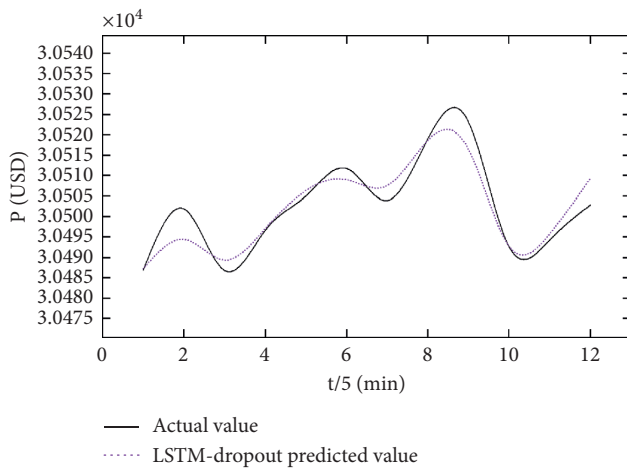


FIGURE 8: Short-term LSTM-dropout prediction.

In terms of prediction accuracy, SSA-LSTM can improve the prediction effect of the WT-LSTM model and reduce RMSE, MAE, and MAPE by 8.46%, 12.00%, and 12.00%,

TABLE 4: 3-hour DJIA closing price forecast results.

	RMSE	MAE	MAPE	SDAPE
LSTM	5.2447743	4.1066389	0.0001347	0.0001069
RNN-dropout	4.6867001	3.9202222	0.0001286	0.0000842
LSTM-dropout	3.4334542	2.6221944	0.0000860	0.0000727
SSA-LSTM	1.1713269	0.9653596	0.0000317	0.0000222
WT-LSTM	1.2796409	1.0970283	0.0000360	0.0000216

respectively. Concerning predictive stability, WT-LSTM can improve the predictive stability of the SSA-LSTM model and reduce SDAPE by 2.84%. The predicted results of the four methods are shown in Figures 11–15.

**3.5. Comparative Analysis of Long-Term Forecasting Effects.** For long-term forecasts, according to the results shown in Table 5, in terms of prediction accuracy, RNN with a dropout layer is superior to the LSTM model, and RMSE,



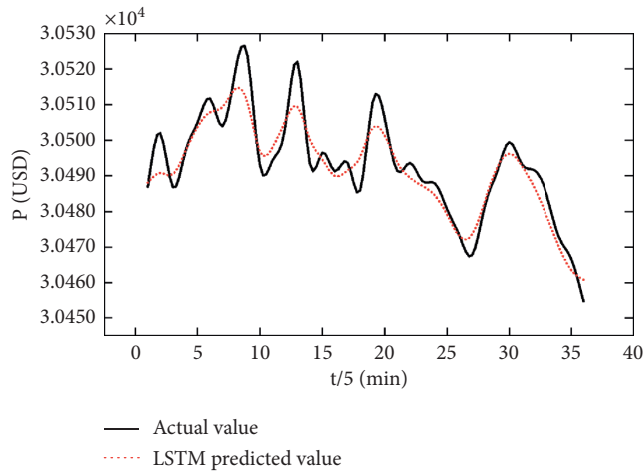


FIGURE 11: Medium-term LSTM prediction.

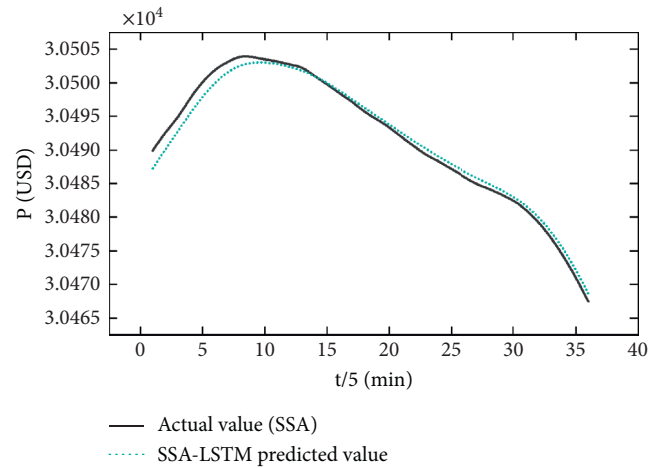


FIGURE 14: Medium-term SSA-LSTM prediction.

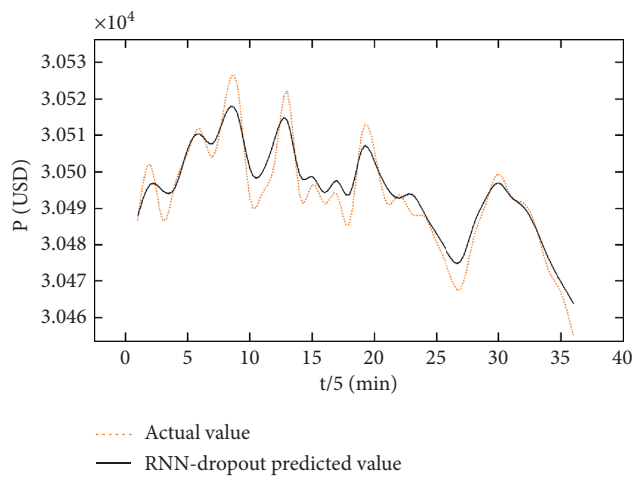


FIGURE 12: Medium-term RNN-dropout prediction.

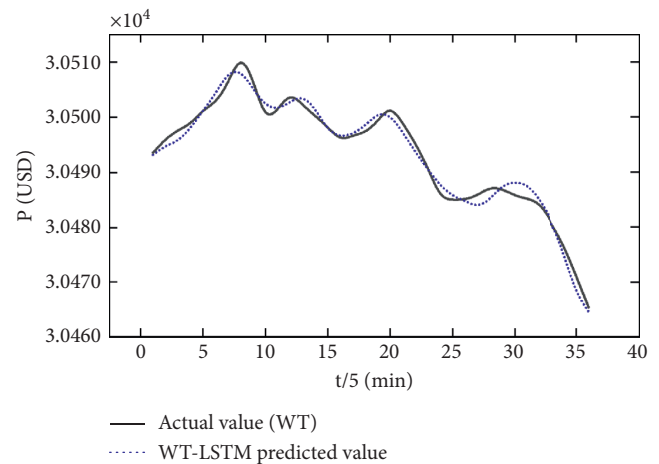


FIGURE 15: Medium-term SSA-LSTM prediction.

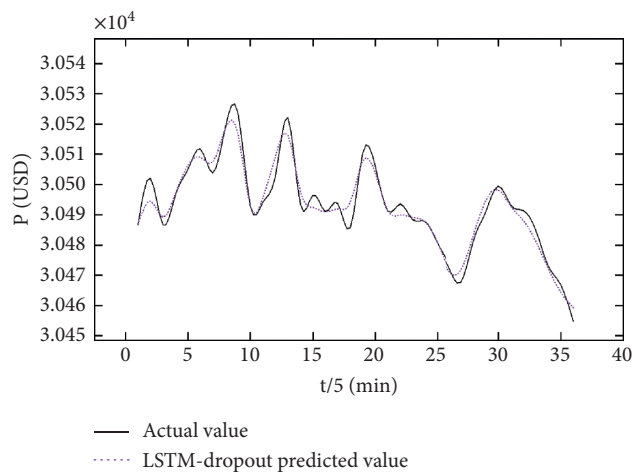


FIGURE 13: Medium-term LSTM-dropout prediction.

MAE, and MAPE are decreased by 13.02%, 6.03%, and 5.99%, respectively. Besides, the dropout layer can optimize the prediction effect of the original LSTM model, and RMSE,

MAE, and MAPE are decreased by 26.99%, 29.56%, and 29.54%, respectively. SSA-LSTM can optimize the prediction effect of the original LSTM model, and RMSE, MAE, and MAPE are reduced by 80.94%, 78.28%, and 78.28%, respectively. WT-LSTM can optimize the prediction effect of the original LSTM model, reducing RMSE, MAE, and MAPE by 68.92%, 69.15%, and 69.14%, respectively. In terms of prediction stability, RNN with a dropout layer is also superior to the LSTM model and the SDAPE is reduced by 22.35%. Besides, the dropout layer can improve the prediction stability of the original LSTM model and reduce the SDAPE by 23.90%. SSA-LSTM can improve the prediction stability of the original LSTM model and reduce the SDAPE by 84.82%. WT-LSTM can improve the prediction stability of the original LSTM model and reduce the SDAPE by 68.57%. According to the analysis of long-term prediction, we can see that the generalization effect of filtering to prevent overfitting is better than the effect of the dropout layer to improve the accuracy. At the same time, singular spectrum

TABLE 5: 6-hour DJIA closing price forecast results.

	RMSE	MAE	MAPE	SDAPE
LSTM	6.1655946	4.5780000	0.0001503	0.0001356
RNN-dropout	5.3630017	4.3020694	0.0001413	0.0001053
LSTM-dropout	4.5014469	3.2249583	0.0001059	0.0001032
SSA-LSTM	1.1753464	0.9942363	0.0000326	0.0000206
WT-LSTM	1.9164739	1.4123594	0.0000464	0.0000426

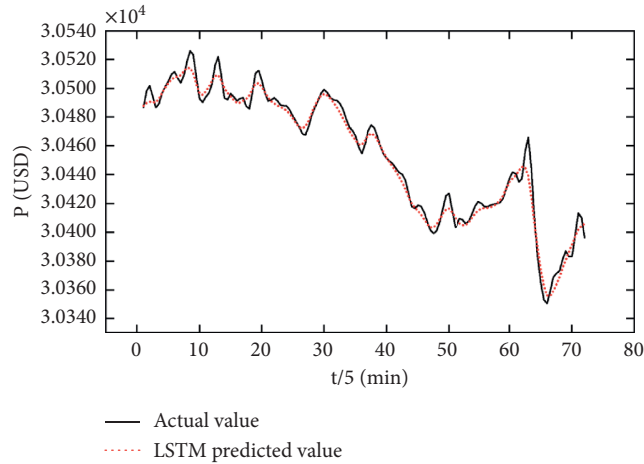


FIGURE 16: Long-term LSTM prediction.

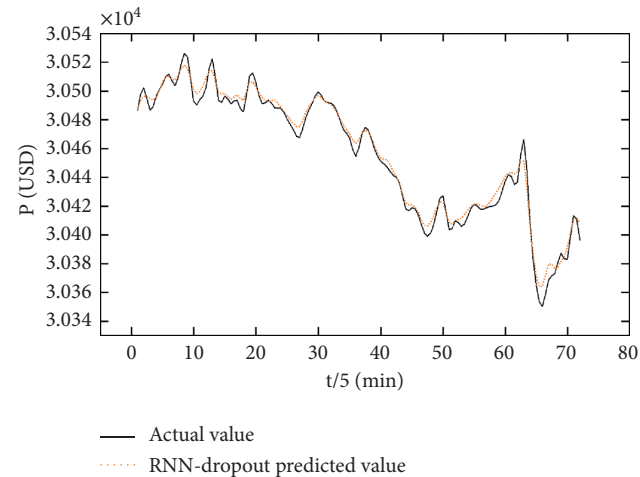


FIGURE 17: Long-term RNN-dropout prediction.

analysis also has a good effect on filtering. In terms of prediction accuracy, SSA-LSTM can improve the prediction effect of the WT-LSTM model and reduce RMSE, MAE, and MAPE by 38.67%, 29.60%, and 29.63%, respectively. In terms of predictive stability, SSA-LSTM can improve the predictive stability of WT-LSTM model and reduce SDAPE by 51.70%. The predicted results of the four methods are shown in Figures 16–20.

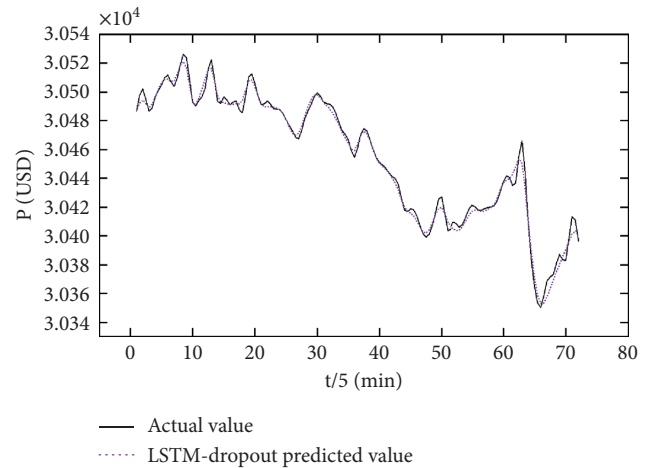


FIGURE 18: Long-term LSTM-dropout prediction.

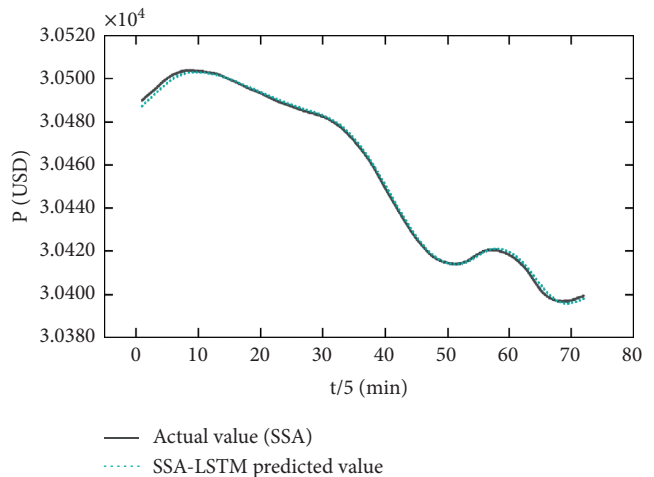


FIGURE 19: Long-term SSA-LSTM prediction.

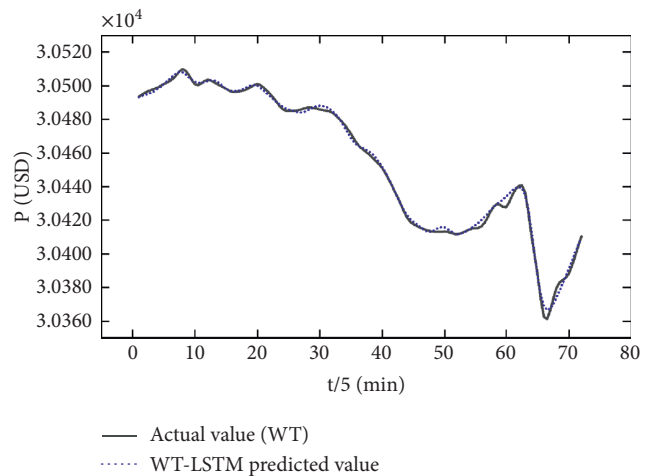


FIGURE 20: Long-term WT-LSTM prediction.

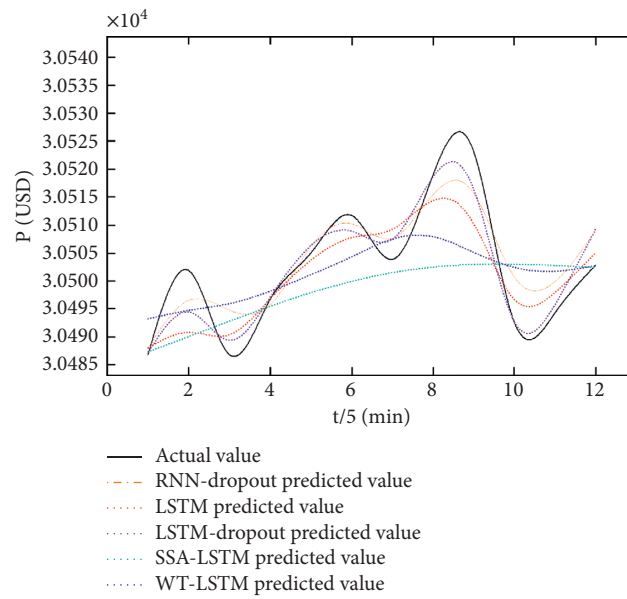


FIGURE 21: Short-term prediction results.

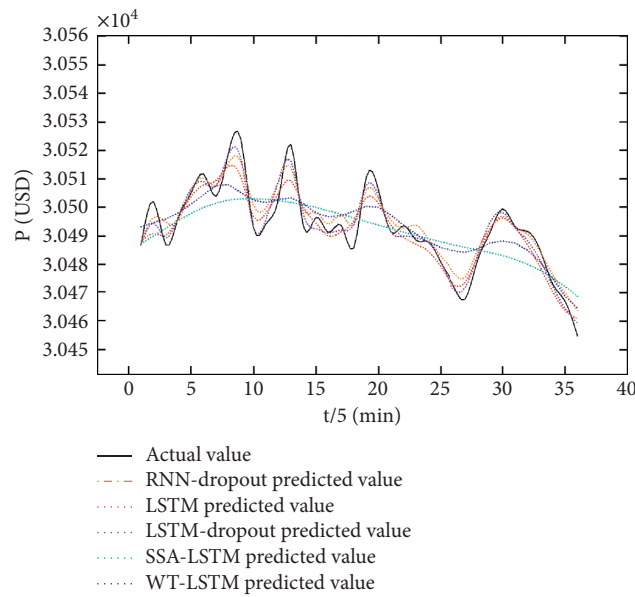


FIGURE 22: Medium-term prediction results.

In summary, both WT-LSTM and SSA-LSTM can significantly enhance the prediction ability of the original LSTM and raise the prediction accuracy and stability, especially the generalization ability, no matter in the short, medium, or long term. Also, in the short term and medium

term, the improvement effect of WT-LSTM is better than that of SSA-LSTM, while in the long term, the improvement effect of SSA-LSTM is better than that of WT-LSTM. A comprehensive comparison of the prediction results of the four methods is shown in Figures 21–23.

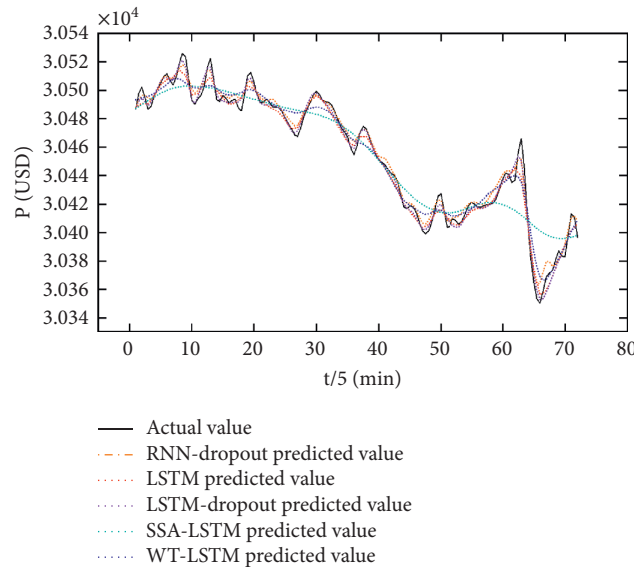


FIGURE 23: Long-term prediction results.

#### 4. Conclusion and Discussions

This paper has discussed the theoretical basis of deep learning and the practical application of LSTM price prediction and has proposed the use of denoising methods to reduce noise on high-frequency financial time series to minimize the effect of random interference noise to raise the prediction generalization of the model for out-of-sample data. The results are also significant enough to prove the improvement of the LSTM predicting model with effective denoising methods, especially wavelet transform and singular spectrum analysis.

In light of the empirical results of the DJIA 5 minutes closing data, the following conclusions can be drawn: firstly, the use of wavelet transform and singular spectrum analysis to denoise data can significantly raise the ability to extend of LSTM neural network and WT's effect is better than SSA's effect in the short term and medium term, but worse than SSA filtering method in the long term. Secondly, with the extension of the time limit, the generalization ability of wavelet transform and reconstructed filter sequence in the prediction of LSTM neural network is weakening, while the generalization ability of singular spectrum analysis decomposition and reconstructed filter sequence in the prediction of LSTM neural network is increasing. But the prediction effect of wavelet transform and singular spectrum analysis reconstruction filter is still significant. Thirdly, the WT-LSTM neural network and SSA-LSTM neural network can converge quickly in a small amount of time and has a good prediction effect under the high-frequency data, which provides a new idea for financial risk management and monitoring under high-frequency trading. These findings can be widely used in the selection of methods for processing time series. To be specific, WT-LSTM is recommended to be chosen when processing relatively short-term time series, while SSA-LSTM is more efficient in processing long-term time series.

In view of the high tunability of neural networks, there are still many technical improvements in future research, such as adding more nonhomogeneous information as input to the neural network and optimizing the structure of the neural network itself. It is worth noting that applying the advantages of big data in the financial high-frequency time series to the investment field can enable investment opportunities to be discovered by investors in a timely manner and facilitate the development of intelligent investment in the financial market. In addition, it can also strengthen risk management, improve the efficiency of risk identification, and effectively maintain the stability of the financial market.

#### Data Availability

The DJIA data used to support the findings of this study were supplied by Wind under license and so cannot be made freely available. Requests for access to these data should be made to Wind, sales@wind.com.cn.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

The authors gratefully acknowledge Southeast University for providing them with rich academic resources. At the same time, thanks are due to Dr. Wei Yin for the inspiration of this work and careful instruction and to Dr. Yawen Han for valuable suggestions. The authors are also grateful to the partners for their full cooperation. In addition, Qi Tang wants to thank Wenye Shi and Yuanyu Wang for their assistance and encouragement, as well as the concern and company from Ciyue Han for years.

## References

- [1] X. Zheng and B. M. Chen, "Analysis of Dow Jones industrial average," *Stock Market Modeling and Forecasting*, vol. 442, pp. 43–51, 2013.
- [2] W. Xu, H. Peng, F. Tian, and X. Peng, "A hybrid modelling method for time series forecasting based on a linear regression model and deep learning," *Applied Intelligence*, vol. 49, no. 8, pp. 3002–3015, 2019.
- [3] N. K. Ahmed, A. F. Atiya, N. E. Gayar, and H. El-Shishiny, "An empirical comparison of machine learning models for time series forecasting," *Econometric Reviews*, vol. 29, no. 5–6, pp. 594–621, 2010.
- [4] G. Sun, J. Lin, C. Yang et al., "Stock price forecasting: an echo state network approach," *Computer Systems Science and Engineering*, vol. 36, no. 3, pp. 509–520, 2021.
- [5] J. S. Cao and J. H. Wang, "Stock price forecasting model based on modified convolution neural network and financial time series analysis," *International Journal of Communication Systems*, vol. 32, 2019.
- [6] N. Koç Ustali, N. Tosun, and Ö. Tosun, "Makine öğrenmesi teknikleri ile hisse senedi fiyat tahmini," *Eskişehir Osmangazi Üniversitesi İktisadi Ve İdari Bilimler Dergisi*, vol. 16, no. 1, pp. 1–16, 2021.
- [7] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [8] Q. Jiang, C. Tang, C. Chen, X. Wang, and Q. Huang, "Stock price forecast based on LSTM neural network," in *Twelfth International Conference on Management Science and Engineering Management*, pp. 393–408, Springer, Berlin, Germany, 2018.
- [9] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical And Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [10] H. Hassani, "Singular spectrum analysis: methodology and comparison," *Journal of Data Science*, vol. 5, pp. 239–257, 2007.
- [11] C.-S. Lin, S.-H. Chiu, and T.-Y. Lin, "Empirical mode decomposition-based least squares support vector regression for foreign exchange rate forecasting," *Economic Modelling*, vol. 29, no. 6, pp. 2583–2590, 2012.
- [12] L. H. Tang, Y. L. Bai, J. Yang et al., "A hybrid prediction method based on empirical mode decomposition and multiple model fusion for chaotic time series," *Chaos Solitons & Fractals*, vol. 141, 2020.
- [13] J. Wang and X. Li, "A combined neural network model for commodity price forecasting with SSA," *Soft Computing*, vol. 22, no. 16, pp. 5323–5333, 2018.
- [14] F. Wen, J. Xiao, Z. He, and X. Gong, "Stock price prediction based on SSA and SVM," *Computer Science*, vol. 31, pp. 625–631, 2014.
- [15] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Economics*, vol. 30, no. 5, pp. 2623–2635, 2008.
- [16] N. Jia and C. J. Zheng, "Short-term forecasting model of agricultural product price index based on LSTM-DA neural network," *Computer Science*, vol. 46, pp. 62–65, 2019.
- [17] H. T. Jung, H. H. Fen, and Y. W. Chang, "Forecasting stock markets using wavelet transforms and recurrent neural networks: an integrated system based on artificial bee colony algorithm," *Applied Soft Computing*, vol. 11, pp. 2510–2525, 2011.
- [18] Q. Tang, T. M. Fan, R. C. Shi, J. Y. Huang, and Y. Y. Ma, "Prediction of financial time series using lstm and data denoising methods," <http://arxiv.org/abs/2103.03505>.
- [19] Z. Y. Pang and L. Liu, "Can agricultural product price fluctuation be stabilized by future market: empirical study based on discrete wavelet transform and GARCH model," *Financial Research*, vol. 11, pp. 126–139, 2013.
- [20] X. L. Liu and X. P. Wang, "High frequency prediction research based on wavelet analysis of stock index futures," *Systems Engineering Theory and Practice*, vol. 35, pp. 1425–1432, 2015.
- [21] N. Srivastava, G. E. Hinton, A. Krizhevsky et al., "Dropout : a simple way to prevent networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.