IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 25, 2019, accepted September 27, 2019, date of publication September 30, 2019, date of current version October 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2944755

The Forecasting of PM2.5 Using a Hybrid Model Based on Wavelet Transform and an Improved Deep Learning Algorithm

WEIBIAO QIAO[1,2], WENCAI TIAN[1], YU TIAN[1], QUAN YANG[1], YINING WANG[1], JIANZHUANG ZHANG[1]

[1]School of Environmental and Municipal Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450046, China
[2]Sinopec Zhongyuan Petroleum Engineering Company Ltd., Zhengzhou 450046, China

Corresponding author: Jianzhuang Zhang (jzz112496@163.com)

This work was supported by the High-Level Talents Start-Up Project of North China University of Water Resources and Electric Power under Grant 40691.

ABSTRACT In recent years, the haze has caused serious troubles to people's lives, with the continuous increase of PM2.5 emissions. The accurate prediction of PM2.5 is very crucial for policy makers to make predictive measures. Due to the nonlinearity of the PM2.5 time series, it is difficult to predict accurately. Despite some studies about PM2.5 being proposed, the problem of the LSTM (long short-term memory) gradient disappearance and random selection of wavelet orders and layers isn't still solved. In this study, a novel model based on WT (wavelet transform)-SAE (stacked autoencoder)-LSTM is proposed. Firstly, six study sites from China are taken as examples and WT is used to decompose PM2.5 time series into several low-and high- frequency components based on different samples. Secondly, the decomposed components are predicted based on SAE-LSTM. Finally, the predicted results are reconstructed in view of all low-and high-frequency components and the predicted results are obtained. The results imply that: (1) the forecasting performance of SAE-LSTM is better than that of other models (e.g., BP (back propagation)) used for comparison; (2) for six different PM 2.5 samples, four orders five layers, five orders six layers, five orders seven layers, three orders six layers, five orders seven layers, and five orders six layers are the most appropriate. The conclusion that such a novel model may help to enhance the accuracy of PM 2.5 prediction can be drawn.

INDEX TERMS PM 2.5 time series, wavelet transform, stacked autoencoder, long short-term memory, prediction.

I. INTRODUCTION

With the frequent occurrence of the smog in recent years, FPM (fine particulate matter) has attracted wide widespread attention [1]–[4]. PM 2.5 whose equivalent diameter is less than or equal to 2.5 μm can be suspended in the air for a long time [5]. The higher the concentration of PM 2.5 in the air, the more serious the air pollution is. And, compared with the coarser ambient air particulate matter, PM 2.5 has a smaller particle size, stronger activity, which is easy to be accompanied by toxic and harmful substances (e.g., heavy metals, microorganisms) [6], [7]. Furthermore, PM 2.5 has a long residence time in the atmosphere, which has a great impact on human health and the quality of the atmospheric environment [8]. Therefore, accurate prediction of PM 2.5 concentration is of great significance for the protection of public health and the formulation of preventive measures.

However, the accurate prediction of PM 2.5 has become a challenging task, because of the volatility characteristics of PM 2.5. Last several years, some scholars have established some models to try to predict PM 2.5. In addition, these results can be roughly divided into two categories: (1) conventional prediction models; (2) artificial intelligence prediction models. What is more, some research results on the conventional forecasting models are listed in Table 1.

It can be seen from Table 1 some conventional prediction models have been used to forecast the PM2.5. However, due to the volatility characteristics of PM2.5 in view of the

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

TABLE 1. Forecasting of the PM 2.5 based on the conventional models in recent years.

Model name	Results and conclusions	Reference
Nonlinear regression	The results indicate that the viability of the NLR models for O3 and PM2.5 forecasting in China.	Lv et al., 2016,[9]
Empirical nonlinear regression models	The conclusions imply that novel models have high-precision results to forecast PM2.5 compared with the linear model.	Ausati and Amanollahi, 2016,[10]
Classification and regression trees-linear model-Kalman filter-analog combination	The proposed model has the best Global performance for different lead times.	Lyu et al., 2017,[11]
Autoregressive integrated moving average	The forecasting results are in agreement with the raw data.	Ni et al., 2017,[12]
Autoregressive integrated moving average	The proposed combination method is superior to single models.	Wang et al., 2017,[13]
Autoregressive integrated moving average	The results indicate that the PM2.5 has important positive correlations with NO2, SO2, and PM10.	Zhang et al., 2018,[14]
Exponential smoothing with drift model	The results imply that 90 % of the stations have an error less than 1.5 μg/m3.	Mahajan et al., 2019,[15]
The combination model	The combination model outperformed MLR due to the consideration of the residuals of the MLR model.	Chelani, 2019,[16]

TABLE 2. Forecasting of the PM 2.5 based on some artificial intelligence models in recent years.

Model name	Results and conclusions	Reference
Support vector regression and grey wolf optimizer	The study implies that the proposed hybrid model is superior to the comparative models.	Niu et al., 2016,[17]
Support vector machine and flower pollination algorithm	The study results demonstrate that the proposed combination model is superior to the comparative models.	Li et al., 2017,[18]
Support vector regression and Elman	The proposed hybrid model exhibits the best performance.	Chen et al., 2017,[19]
Recursive neural network	The comparison shows that the recursive neural network can predict the PM2.5 with much accuracy	Biancofiore et al., 2017,[20]
Second order self-organizing fuzzy neural network	In conclusion, the proposed model provided accurate results for the hourly distribution of PM2.5.	Qiao et al., 2017,[21]
Adaptive fuzzy neural network	This hybrid model indicates the potential to be an administrative and political administrative method.	Jiang et al., 2017,[22]
Artificial neural network	According to the two experiments, the industrial factor and stock farming factor are items that may influence the PM2.5 concentration change.	Chang et al., 2017,[23]
Convolutional neural network and long short-term memory	In the future, this study can be applied to the control and prevention of PM2.5.	Huang and Kuo., 2018,[24]
Adaptive back propagation neural network	The model establishes the correlation of the PM2.5 and AOD.	Chen, 2018,[25]
Long short-term memory neural network	It is demonstrated that the E-LSTM obtained better forecasting performance than that using the single LSTM and feed-forward neural network.	Bai et al., 2019,[26]
Long short-term memory	The results show that the proposed model gives a better predictive performance.	Zhao et al., 2019,[27]

different samples, conventional prediction models have some limitations. In recent years, artificial intelligence forecasting models have been applied to the forecasting of PM2.5, in view of its strong fitting ability. These study results on artificial intelligence prediction models are shown in Table 2.

By means of summarizing Table 2, artificial intelligence forecasting models are widely used for PM2.5 forecasting (e.g., NN (neural network)), but NN has the disadvantage of local extremum. So, some scholars have tried to combine wavelet transform with artificial intelligence prediction model to obtain more information about the original PM2.5 and improve the prediction accuracy of PM2.5. These studies are shown in Table 3.

To make a long story short, the combination of the artificial intelligence forecasting models and wavelet transform are applied to the forecasting of PM2.5. However, when the wavelet transform is adopted to decompose PM2.5 time series, wavelet orders and layers are randomly determined. In addition, LSTM solves the gradient disappearance problem of RNN (recurrent neural network) to some extent. So, to solve these two scientific problems, some novel research work is carried out in this paper:

(1) To improve the problem of LSTM gradient disappearance, the combination of SAE and LSTM is proposed. Furthermore, to test the effectiveness based on the proposed model, some advanced forecasting models are adopted for comparisons, e.g. SAE-BP (SAE-back propagation), SAE-ELM (SAE -extreme learning machine), SAE-BiLSTM (SAE - bi-directional), LSTM, BP, and ELM;

(2) Coiflets is adopted to decompose the PM2.5, into several high- and low-frequency components. In addition, SAE-LSTM is used to predict the decomposed components. Lastly, the forecast results obtained by SAE-LSTM are reconstructed. Thereby, the optimal wavelet layers and orders are determined by comparing the evaluating indicators for different samples.

II. METHODS

In this part, some methods are used in this paper, including WT [28], SAE [34], LSTM [27], the combination process of

TABLE 3. Combination of wavelet and neural network to forecast PM2.5 in recent years.

Model name	Results and conclusions	Reference
Artificial neural networks and wavelet transformation	The results obtained by the proposed model indicate that the potential to be applied in other countries' air quality forecasting systems.	Feng et al., 2015,[28]
Back propagation neural network and wavelet transformation	The results demonstrate that the proposed model outperforms all the other considered models in this paper.	Wang et al., 2017,[29]
Autoregressive integrated moving average and wavelet transformation	The proposed model could be efficiently and successfully applied to the PM forecasting field.	Zhang et al., 2017,[30]
Multi-layer perceptron and wavelet packet decomposition	The results indicate the proposed model has excellent forecasting performance.	Liu et al., 2019,[31]
Back propagation neural network and wavelet packet decomposition	The proposed model has satisfactory performance in forecasting PM2.5.	Liu et al., 2019,[32]
Empirical wavelet transform and stacking ensemble methods	It has been proved in the study that the model proposed in the study has better accuracy and wide applicability comparing to the existing models.	Liu et al., 2019,[33]

SAE and LSTM, and BiLSTM [35]. Furthermore, statistical evaluation indexes and forecasting framework are given in detail.

A. WAVELET TRANSFORM

WT inherits and develops the idea of short-time Fourier transform localization, and overcomes the shortcomings of window size not changing with frequency. Furthermore, WT is an ideal tool for signal analysis and processing, because it can provide a "time-frequency" window that varies with frequency.

In practical applications, because most of the computer processing is a discrete equation, the continuous wavelet transform is often discretized. The Mallat algorithm is adopted, which can be expressed as:

$$\mathrm{a}_j = \mathrm{a}_{j+1} h_1; \quad \mathrm{d}_j = \mathrm{d}_{j+1} l_1, \quad (j = 0,1,\cdots,n-1) \qquad (1)$$

where h_1 and l_1 are low-pass filters and high-pass filters respectively.

Mallat algorithm is used for wavelet decomposition. After each decomposition, the low- and high-frequency component are twice as much as the signal points before decomposition. The reduction of points is disadvantageous to prediction. In order to overcome this disadvantage, the decomposed components can be reconstructed by the reconstruction algorithm. The reconstruction algorithm is described as follows:

$$\mathrm{a}_j = \mathrm{a}_{j+1} h_2 + \mathrm{d}_{j+1} l_2, \quad (j = n-1,\cdots,1,0) \qquad (2)$$

where h_2 and l_2 are dual operators of h_1 and l_1, respectively. The process of WT is shown in FIGURE 1.

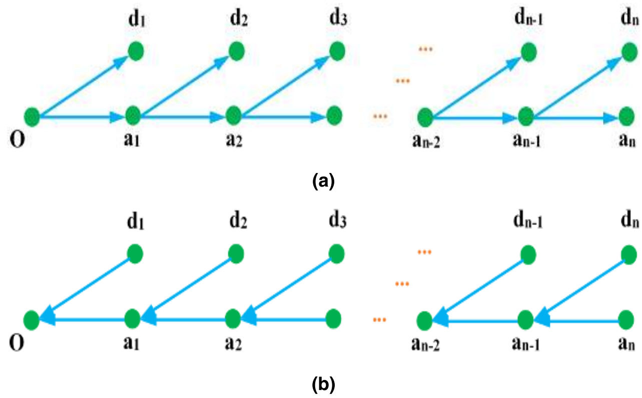

FIGURE 1. Diagrammatic sketch of wavelet transform: (a) decomposition process; (b) reconstruction process.

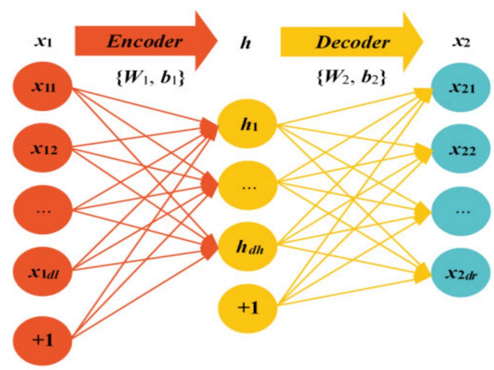

FIGURE 2. Model structure of AE.

B. MACHINE LEARNING ALGORITHM

1) STACKEN AUTOENCODER

Autoencoder is a kind of unsupervised one hidden layer neural network, in which the output layer is set to be equal to the input layer. FIGURE 2 shows the basic structure of an AE model.

AE is composed of an encoder and decoder, and their mapping functions are defined as follows.

$$\boldsymbol{h} = f_1(\boldsymbol{x}_1) = s_{f1}(\boldsymbol{W}_1 \boldsymbol{x}_1 + \boldsymbol{b}_1) \qquad (3)$$

$$\boldsymbol{x}_2 = f_2(\boldsymbol{h}) = s_{f2}(\boldsymbol{W}_2 \boldsymbol{h} + \boldsymbol{b}_2) \qquad (4)$$

where $\boldsymbol{x_1} = [x_{11}, x_{12}, \cdots, x_{1dl}]^{\mathrm{T}} \in \mathrm{R}^{1dl}$ is the inputs of the AE; $\boldsymbol{h} = [h_1, h_2, \cdots, h_{dh}]^{\mathrm{T}} \in \mathrm{R}^{dh}$ is the join vector between $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$; $\boldsymbol{x_2} = [x_{21}, x_{22}, \cdots, x_{2dr}]^{\mathrm{T}} \in \mathrm{R}^{2dr}$ is the inputs of the AE; $1dl$ is the dimension of the inputs; dh is the dimension of the hidden variable vector; $2dr$ is the dimension of the outputs; $\boldsymbol{b}_1 \in \mathrm{R}^{1dl}$ is the bias vector; $\boldsymbol{b}_2 \in \mathrm{R}^{2dr}$ is the bias vector; the nonlinear activation function of s_{f1} can be chosen as the sigmoid function, or others like the tanh function the rectified linear unit function; the activation function s_{f2} of the decoder can be either the sigmoid function or other functions.

Stacked autoencoder, deep belief network, and deep convolutional neural networks are three typical deep learning algorithms, which is a hierarchical deep neural network structure composed of multilayer AEs. The model structure of SAE based on multiple AEs is shown in FIGURE 3.

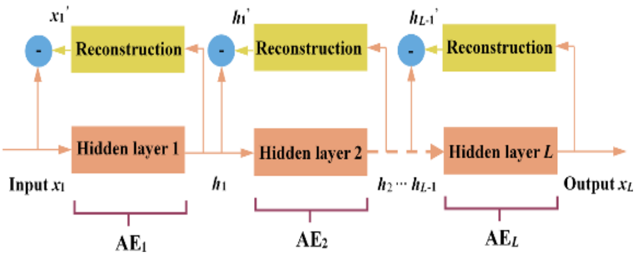

FIGURE 3. Structure of AE.

2) LONG SHORT-TERM MEMORY

The structure of the basic neural network includes input layer, hidden layer, and output layer. The output is controlled by the activation function, and the weights are used to connect the layers. Recently, on the basis of the basic neural network, a new type of neural network has been developed, which is called RNN. The biggest difference between RNN and basic neural network is that RNN also establishes weighted connections between neurons. However, RNN has the problem of gradient disappearance. Therefore, in order to solve this problem, some RNN variants such as LSTM have been proposed. LSTM adds three gates based on RNN to control information transmission and final result calculation. The three gates are forgetting gate, input gate, and output gate. The structure of the LSTM processor unit is shown in FIGURE 4.

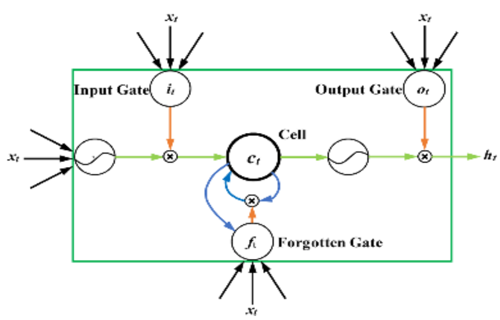

FIGURE 4. LSTM processor unit.

And the forgotten gate can be computed as:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{5}$$

where f_t is the vector of the input gate; W_f and b_f is the weight and bias vector of forgotten gate; $[h_{t-1}, x_t]$ means connecting two vectors into a longer vector; σ which is the sigmoid function used in this study is activation function. The expansion of $W_f \cdot [h_{t-1}, x_t]$ is as follows:

$$\begin{aligned} W_f \cdot [h_{t-1}, x_t] &= [W_f] \cdot \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} \\ &= \begin{bmatrix} W_{fh} & W_{fx} \end{bmatrix} \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} = W_{fh} h_{t-1} + W_{fx} x_t \end{aligned} \tag{6}$$

The input and output gate can be computed as:

$$\begin{aligned} i_t &= \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) & (7) \\ c_t &= f_t \cdot c_{t-1} + i_t \cdot \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) & (8) \\ o_t &= \sigma\left(W_0 \cdot [h_{t-1}, x_t] + b_0\right) & (9) \\ h_t &= o_t \cdot \tanh\left(c_t\right) & (10) \end{aligned}$$

where i_t, o_t and c_t are the vectors for input gate, output gate, and cell activations, respectively; h_t is the output vector; W_i, W_c, and W_o are the weight of the corresponding gate; b_i, b_c, and b_o are the bias vectors of the corresponding gate.

3) BI-DIRECTIONAL LONG SHORT-TERM MEMORY

In timing processing, standard RNN and LSTM often ignore future information, while BiLSTM can take advantage of future information. The basic structural idea of BiLSTM is that the front and back layers of each training sequence are two LSTM networks, respectively, and the LSTM networks are both connected to one input layer and one output layer. The output layer can obtain past information of each point in the input sequence, and can also get future information of each point through this structure. FIGURE 5 shows a BiLSTM that expands along time. Increased neural network update equation can be computed as:

$$\begin{aligned} h_{tr} &= H\left(W_1 x_t + W_2 h_{(t-1)r} + b_r\right) & (11) \\ h_{tl} &= H\left(W_1 x_t + W_2 h_{(t-1)l} + b_l\right) & (12) \\ y_t &= W_4 h_{tr} + W_6 h_{tl} + b_y & (13) \end{aligned}$$

where h_{tr}, h_{tl}, y_t are respectively the vectors forward propagation, backward propagation and output layer; W_1, W_2, W_3, W_4, W_5, and W_6 are respectively the corresponding weight coefficients; b_r, b_l, b_y are the corresponding bias vectors.

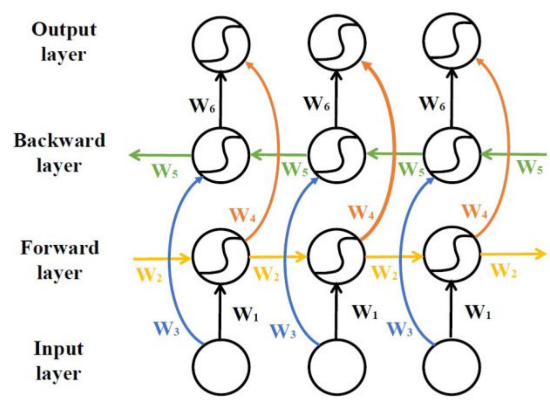

FIGURE 5. Expansion structure of BiLSTM.

4) THE COMBINATION PROCESS OF SAE AND LSTM

The combination of SAE and LSTM is actually a process of data transfer. The specific calculation process is as follows:

Step 1: The PM2.5 time series is divided into training samples, testing samples, and prediction samples.

Step 2: Set the parameters of SAE and LSTM;

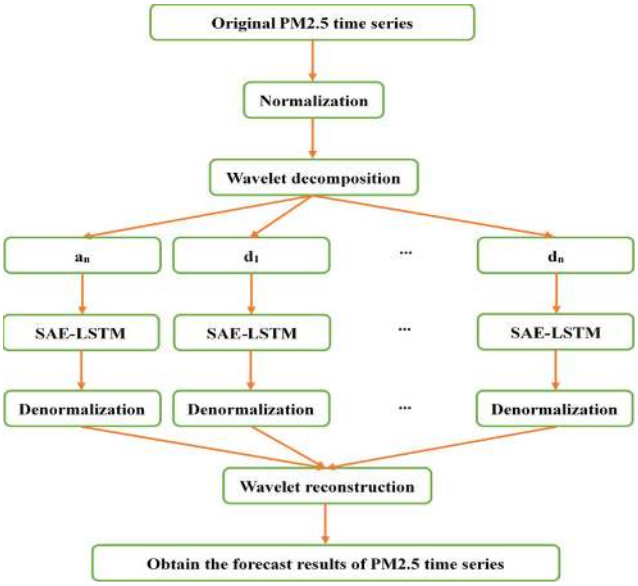

FIGURE 6. Forecasting framework using novel WT-SAE-LSTM for PM 2.5 time series in this paper.

Step 3: Train SAE network;

Step 4: The trained SAE network is used to predict the training samples, and the prediction results are used as the input of LSTM;

Step 5: Based on the output results of SAE, LSTM network is trained;

Step 6: The training samples, test samples, and prediction samples are predicted by the trained LSTM network. Also, if the set error precision is satisfied, the output result is exported or returned to Step 3.

C. STATISTICAL EVALUATION

In order to comprehensively assess the characteristics of different prediction models, seven commonly used and mean absolute error (MAE) [36]–[39] is applied in this subsection. The definition of this index is shown in EQUATION (14).

$$\text{MAE} = (1/n) \sum_{t=1}^{n} |A_t - F_t| \qquad (14)$$

where n represents the number of training or test set; A_t and F_t represent the raw and forecasting value.

D. PREDICTIVE FRAMEWORK

The predictive framework in this study is given in FIGURE 6. Furthermore, the detailed prediction process is as follows:

To eliminate the effect of the PM 2.5 magnitude on the forecasting results, the PM 2.5 is normalized based on the normalization method whose interval is from - 1 to 1.

Furthermore, to get more information about PM 2.5 time series, it is broken down into several low- and high-frequency components by wavelet decomposition algorithm. In addition, the low- and high-frequency components are forecasted by SAE-LSTM, and the forecasting results are gotten. After reconstructing the forecasting results, the final prediction results are denormalized.

III. RESULTS ANALYSIS AND DISCUSSION

A. SAMPLE COLLECTION AND PREPROCESSING

In order to verify the generality of the forecasting model proposed in this paper, six groups of PM2.5 time series are selected from Jiayuguan, Datong, Fushun, Qiqihar, Weinan, and Xuchang. They are located in China, as shown in FIGURE 7(A). These data are from China air quality online monitoring and analysis platform (https://www.aqistudy.cn/), which are shown in FIGURE 7(B). In addition, in order to understand the data differences of different PM2.5 time series, some statistical indicators (e.g., Mean, S.D., min, and max) are calculated, as shown in Table 4.

The normalization method is adopted to normalize PM 2.5 time series, as depicted in FIGURE 7(C). Here, one-step-ahead forecasting is adopted in all experiments.

TABLE 4. Statistical results of PM 2.5 based on the different study sites.

Study site	Sample type	Index				
		Size	Mean	S.D.	Min	Max
Jiayuguan	All	2007	25.0264	16.1871	0	134
	Training	1610	25.6360	16.3081	0	134
	Test	397	22.5542	15.4616	0	110
Datong	All	2007	36.2611	23.2933	0	160
	Training	1610	37.4118	24.0305	0	160
	Test	397	31.5945	19.3640	0	111
Fushun	All	2007	47.1794	31.5628	0	267
	Training	1610	49.0540	32.0832	0	267
	Test	397	39.5768	28.1412	0	187
Qiqihar	All	2007	34.8037	35.6520	0	536
	Training	1610	37.0143	34.4273	0	346
	Test	397	25.8388	39.0201	0	536
Weinan	All	2007	65.8052	55.6347	0	499
	Training	1610	68.7025	57.0129	0	499
	Test	397	54.0554	47.9553	0	279
Xuchang	All	2007	67.8062	53.9406	0	506
	Training	1610	70.1578	53.9727	0	506
	Test	397	58.2695	52.8128	0	348

B. EXPERIMENTAL DESIGN AND PARAMETER SETTINGS

In this paper, to ensure the fairness of the comparison of the experimental results, all experiments are calculated on the same computer. And the detailed configuration of the computer is shown in Table 5.

TABLE 5. The specific configuration of the computer.

Name	Settings
Hardware	
CPU	Intel(R) Core (TM) I7-8550U
Frequency	1.99 GHz
RAM	16.0GB (15.9 GB Available)
Hard drive	1TB
Software	
Operating system	Windows 10
Language	MATLAB R2018a

The goal of this study is to improve the gradient disappearance of LSTM and to determine the optimal wavelet layers and orders for different PM2.5 samples. According to these two goals, two experiments are designed, which

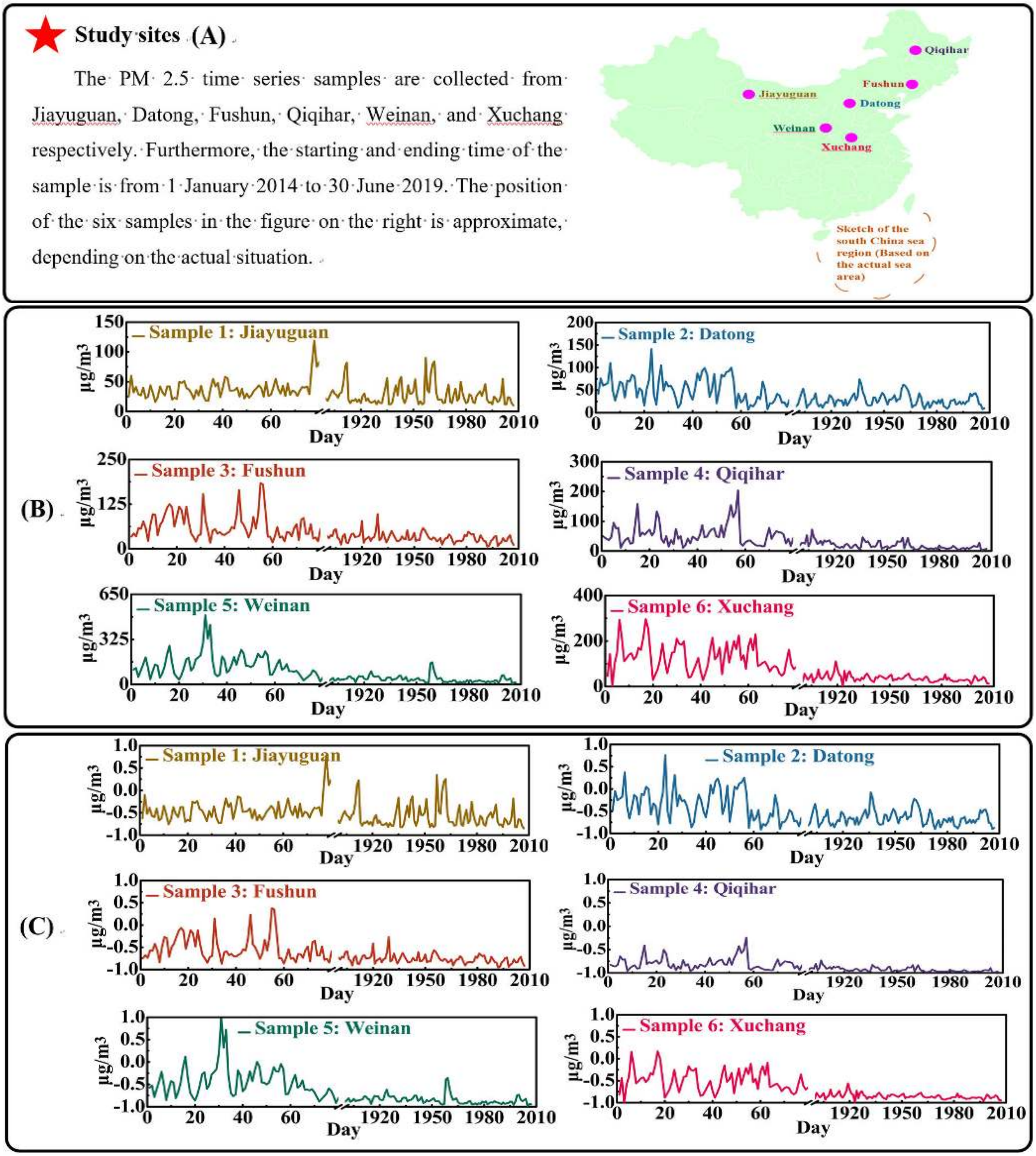

FIGURE 7. Research site and sample data.

are Experiment I: comparison of forecasting efficiency and accuracy based on the proposed model and four models considered for comparison and Experiment II: determination of the optimal wavelet layers and orders based on six different samples.

In Experiment I: the proportion of the test sample and the training sample is 0.2 and 0.8, respectively. In addition, the length of the sliding time window is 20 and the experiment is repeated 10 times. The parameters of the Experiment II are the same as those of the Experiment I. Furthermore, the

TABLE 6. The specific parameter settings in two experiments.

Experiments	Model	Settings
I	SAE-BP	ETF=satlin, DTF=purelin, L2WR=4, SP=0.6; ME=1000, GT=1.00e-07, HL1=20, HL2=20. AF=S
	SAE-ELM	ETF=satlin, DTF=purelin, L2WR=4, SP=0.6; ME=1000, AF=S
	SAE-BiLSTM	ETF=satlin, DTF=purelin, L2WR=4, SP=0.6; ME=250, GT=1, ILR=0.01, LRDP=125, LRDF=0.1, V=0
	LSTM	ME=250, GT=1, ILR=0.01, LRDP=125, LRDF=0.1, V=0
	SAE-LSTM	ME=250, GT=1, ILR=0.01, LRDP=125, LRDF=0.1, V=0, ETF=satlin, DTF=purelin, L2WR=4, SP=0.6
	BP	ME=1000, GT=1.00e-07, HL1=20, HL2=20. AF=S
	ELM	ME=1000, AF=S
II	Coiflets wavelet	MINON=1, MAXON=5, MINLN=1, MAXLN=8,
	SAE-LSTM	ME=250, GT=1, ILR=0.01, LRDP=125, LRDF=0.1, V=0, ETF=satlin, DTF=purelin, L2WR=4, SP=0.6

detailed parameter settings of the two experiments are listed in Table 6.

C. EXPERIMENT I: COMPARISON OF FORECASTING EFFICIENCY AND ACCURACY BASED ON THE PROPOSED MODEL AND FOUR MODELS CONSIDERED FOR COMPARISON

To know the forecasting efficiency and accuracy of the proposed model, six models including SAE-BP, SAE- ELM, SAE-BiLSTM, LSTM, BP, ELM are considered for comparison.

The parameters in this experiment are shown in Section II. B. In addition, the results are described in FIGURE 8 and Table 7.

The following crucial findings are listed by analyzing FIGURE 8 and Table 7.

(1) It can be seen that the results gained by SAE-LSTM and the raw value are the closest based on the six test samples, comparing with other forecasting models considered for comparison from FIGURE 8.

(2) From Table 7, the MAE value of SAE-LSTM is 0.3094, 0.4291, 0.0527, 0.0325, 0.1304, 0.0665, 0.3733, 0.3059, 0.1511, 0.2514, 0.2125, 0.1073, 0.7248, 0.4030, 0.0604, 0.1222, 0.1113, 0.1446, 0.9039, 3.8966, 0.0757, 0.0752, 1.1352, 0.7127, 0.7040, 1.1887, 0.0724, 0.4541, 0.4476, 0.4418, and 0.8935, 1.4574, 0.1723, 0.6040, 0.6549. 0.6378 lower than the that of SAE-BP, SAE-ELM, SAE-BiLSTM, LSTM, BP, ELM for Jiayuguan, Datong, Fushun, Qiqihar, Weinan, and Xuchang.

D. EXPERIMENT II: DETERMINATION OF THE OPTIMAL WAVELET LAYERS AND ORDERS BASED ON SIX DIFFERENT SAMPLES

In this experiment, six cases are used to verify the performance of SAE-LSTM. Furthermore, the parameters of all the cases in this experiment are set to be the same, which are listed in Section II. B in detail.

1) CASE ONE: JIAYUGUAN

The results, in this case, are shown in Table 8. By analyzing Table 8, the following comparisons can be given:

In view of Table 8 and MAE, the MAE of the one order five layers, second orders six layers, three orders eight layers, four orders five layers and five orders seven layers is smaller than that of the other orders and layers. And, compared with one order five layers, second orders six layers, three orders eight layers, and five orders seven layers, four orders five layers is the smallest. Furthermore, the MAE based on SAE-LSTM is 3.0655. And the MAE of four orders five layers is 1.1730 higher than that of SAE-LSTM used individually.

2) CASE TWO: DATONG

The results, in this case, are shown in Table 9. By analyzing Table 9, the following comparisons can be given:

In view of Table 9, the MAE of the one order six layers, second orders four layers, three orders six layers, four orders four layers and five orders six layers is smaller than that of the other orders and layers. And, compared with o one order six layers, second orders four layers, three orders six layers, and four orders four layers, five orders six layers is the smallest. Furthermore, the MAE based on SAE-LSTM is 3.6543. And the MAE of five orders six layers is 1.5427 higher than that of SAE-LSTM applicated individually.

3) CASE THREE: FUSHUN

The results, in this case, are shown in Table 10. By analyzing Table 10, the following comparisons can be given:

In view of Table 11 and MAE, the MAE of the one order six layers, second orders eight layers, three orders six layers, four orders seven layers and five orders seven layers is smaller than that of the other orders and layers. And, compared with one order six layers, second orders eight layers, three orders six layers and four orders seven layers, five orders seven layers is the smallest. Furthermore, the MAE based on SAE-LSTM is 3.8562. And the MAE of five orders seven layers is 1.5559 higher than that of SAE-LSTM applicated individually.

4) CASE FOUR: QIQIHAR

The results, in this case, are shown in Table 11. By analyzing Table 11, the following comparisons can be given:

In view of Table 11 and MAE, the MAE of the one order seven layers, second orders seven layers, three orders six layers, four orders four layers and five orders seven layers is smaller than that of the other orders and layers. And, compared with one order seven layers, second orders seven layers, four orders four layers and five orders seven layers, three orders six layers is the smallest. Furthermore, the MAE based on SAE-LSTM is 3.6819. And the MAE of three orders

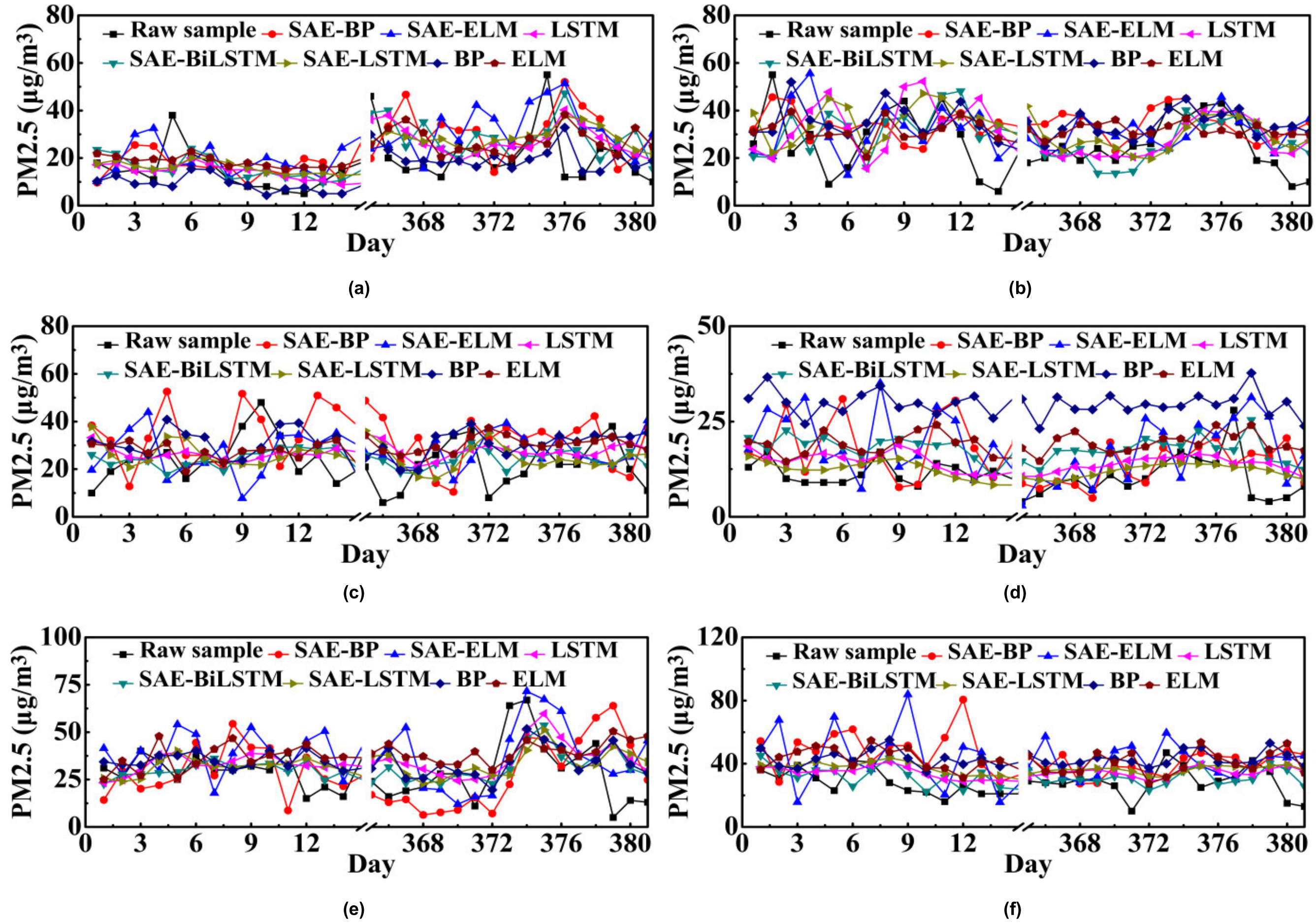

FIGURE 8. Comparison for the proposed forecasting model and other models considered for comparison using the six different samples based on the test set: (a) Jiayuguan; (b) Datong; (c) Fushun; (d) Qiqihar; (e) Weinan; (f) Xuchang.

TABLE 7. Evaluating indicator comparison of the proposed forecasting model and other models considered for comparison using the six different samples based on the test set. The smallest MAE is marked in bold.

Samples	Evaluating indicators	SAE-BP	SAE-ELM	SAE-BiLSTM	LSTM	SAE-LSTM	BP	ELM
Jiayuguan	MAE	3.3749	3.4946	3.1182	3.0980	**3.0655**	3.1959	3.1320
Datong		4.0276	3.9602	3.8054	3.9057	**3.6543**	3.8668	3.7616
Fushun		4.5810	4.2592	3.9166	3.9784	**3.8562**	3.9675	4.0008
Qiqihar		4.5858	7.5785	3.7576	3.7571	**3.6819**	4.8171	4.3946
Weinan		5.2131	5.6978	4.5815	4.9632	**4.5091**	4.9567	4.9509
Xuchang		5.4559	6.0148	4.7297	5.1614	**4.5574**	5.2123	5.1952

six layers is 0.9210 higher than that of SAE-LSTM applicated individually.

5) CASE FIVE: WEINAN

The results, in this case, are shown in Table 12. By analyzing Table 12, the following comparisons can be given:

In view of Table 12 and MAE, the MAE of the one order four layers, second orders four layers, three orders six layers, four orders six layers and five orders seven layers is smaller than that of the other orders and layers. And, compared with one order four layers, second orders four layers, three orders six layers, and four orders six layers, five orders seven layers is the smallest. Furthermore, the MAE based on SAE-LSTM is 4.5091. And the MAE of five orders seven layers is 1.7131 higher than that of SAE-LSTM applicated individually.

6) CASE SIX: XUCHANG

The results, in this case, are shown in Table 13. By analyzing Table 13, the following comparisons can be given:

In view of Table 13 and MAE, the MAE of the one order four layers, second orders five layers, three orders six layers,

TABLE 8. Evaluating indicator comparison based on the different layers and orders using SAE-LSTM for the test sample of Jiayuguan PM2.5. The minimum MAE are marked in bold.

Evaluating indicator	One order							
	one	two	three	four	five	six	seven	eight
MAE	2.6401	2.5048	2.4449	2.4502	**2.4258**	2.4263	2.4300	2.4296
Evaluating indicator	Second orders							
	one	two	three	four	five	six	seven	eight
MAE	2.5709	2.2645	2.0831	2.1007	2.0762	**2.0726**	2.0741	2.0758
Evaluating indicator	Three orders							
	one	two	three	four	five	six	seven	eight
MAE	2.4390	2.1691	2.3553	1.9518	1.9563	1.9514	1.9536	**1.9472**
Evaluating indicator	Four orders							
	one	two	three	four	five	six	seven	eight
MAE	2.2744	2.0600	1.9223	1.8945	**1.8925**	1.8947	1.8960	1.8961
Evaluating indicator	Five orders							
	one	two	three	four	five	six	seven	eight
MAE	2.3718	2.0816	1.9567	1.9382	1.9329	1.9343	**1.9311**	1.9565

TABLE 9. Evaluating indicator comparison based on the different layers and orders using SAE-LSTM for the test sample of Datong PM2.5. The minimum MAE are marked in bold.

Evaluating indicator	One order							
	one	two	three	four	five	six	seven	eight
MAE	2.9298	2.7386	2.7657	2.6026	2.6061	**2.5921**	2.5958	2.6296
Evaluating indicator	Second orders							
	one	two	three	four	five	six	seven	eight
MAE	2.9181	3.1517	3.1057	**2.4043**	2.4167	2.4207	2.4201	2.4385
Evaluating indicator	Three orders							
	one	two	three	four	five	six	seven	eight
MAE	3.1067	2.4086	2.2452	2.2164	2.2111	**2.2095**	2.2106	2.2238
Evaluating indicator	Four orders							
	one	two	three	four	five	six	seven	eight
MAE	2.8501	2.3772	2.2817	**2.2339**	2.2468	2.2376	2.2394	2.2373
Evaluating indicator	Five orders							
	one	two	three	four	five	six	seven	eight
MAE	2.7699	2.2870	2.2034	2.1271	2.1174	**2.1116**	2.1178	2.1313

TABLE 10. Evaluating indicator comparison based on the different layers and orders using SAE-LSTM for the test sample of Fushun PM2.5. The minimum MAE are marked in bold.

Evaluating indicator	One order							
	one	two	three	four	five	six	seven	eight
MAE	3.1734	3.0458	3.9437	2.9101	2.9080	**2.9025**	2.9279	2.9043
Evaluating indicator	Second orders							
	one	two	three	four	five	six	seven	eight
MAE	3.1165	2.8414	2.6181	2.6346	2.6009	2.5858	2.5842	**2.5752**
Evaluating indicator	Three orders							
	one	two	three	four	five	six	seven	eight
MAE	3.0036	2.7057	2.5787	2.5643	2.5574	**2.5405**	2.5439	2.6039
Evaluating indicator	Four orders							
	one	two	three	four	five	six	seven	eight
MAE	2.8946	2.5249	2.5366	2.3669	2.3498	2.3445	**2.3434**	2.3814
Evaluating indicator	Five orders							
	one	two	three	four	five	six	seven	eight
MAE	2.8826	2.5515	2.3238	2.3325	2.3079	2.3007	**2.3003**	2.3369

four orders six layers and five orders six layers is smaller than that of the other orders and layers. And, compared with one order four layers, second orders five layers, three orders six layers, and four orders six layers, five orders six layers is the smallest. Furthermore, the MAE based on SAE-LSTM is 4.5574. And the MAE of five orders six layers is 1.6492 higher than that of SAE-LSTM applicated individually.

E. DISCUSSIONS

Precise prediction of PM 2.5 is very crucial for policy-makers to draw up preventive measures. Besides, the goal

TABLE 11. Evaluating indicator comparison based on the different layers and orders using SAE-LSTM for the test sample of Qiqihar PM2.5. The minimum MAE are marked in bold.

Evaluating indicator	One order							
	one	two	three	four	five	six	seven	eight
MAE	3.4693	3.0639	2.9301	2.8928	2.8270	2.8184	**2.8125**	2.8307
Evaluating indicator	Second orders							
	one	two	three	four	five	six	seven	eight
MAE	3.2039	3.1230	3.0853	3.0199	3.0074	2.9692	**2.9627**	3.0776
Evaluating indicator	Three orders							
	one	two	three	four	five	six	seven	eight
MAE	3.1332	2.9749	2.7994	2.8657	2.7615	**2.7609**	2.7660	2.8563
Evaluating indicator	Four orders							
	one	two	three	four	five	six	seven	eight
MAE	3.0241	2.9502	2.8721	**2.7071**	2.7179	2.7130	2.7135	2.7710
Evaluating indicator	Five orders							
	one	two	three	four	five	six	seven	eight
MAE	3.2379	3.0926	2.9313	2.7109	2.7121	2.7122	**2.7077**	2.7799

TABLE 12. Evaluating indicator comparison based on the different layers and orders using SAE-LSTM for the test sample of Weinan PM2.5. The minimum MAE are marked in bold.

Evaluating indicator	One order							
	one	two	three	four	five	six	seven	eight
MAE	4.1334	3.6180	3.5127	**3.4053**	3.3631	3.3798	3.3795	3.3847
Evaluating indicator	Second orders							
	one	two	three	four	five	six	seven	eight
MAE	4.1048	3.6708	4.0322	**3.1455**	3.2097	3.1591	3.1625	3.1643
Evaluating indicator	Three orders							
	one	two	three	four	five	six	seven	eight
MAE	4.3949	3.6199	6.4072	3.0064	2.8706	**2.8564**	2.8567	2.8579
Evaluating indicator	Four orders							
	one	two	three	four	five	six	seven	eight
MAE	4.1186	4.2020	3.5075	3.1222	2.9982	**2.9872**	2.9889	2.9940
Evaluating indicator	Five orders							
	one	two	three	four	five	six	seven	eight
MAE	4.1773	3.4615	4.1119	2.8959	2.7972	2.8012	**2.7960**	2.8044

TABLE 13. Evaluating indicator comparison based on the different layers and orders using SAE-LSTM for the test sample of Xuchang PM2.5. The minimum MAE are marked in bold.

Evaluating indicator	One order							
	one	two	three	four	five	six	seven	eight
MAE	4.6224	3.5053	4.1022	**3.4411**	3.4565	3.4764	3.4955	3.4773
Evaluating indicator	Second orders							
	one	two	three	four	five	six	seven	eight
MAE	4.0643	3.4320	3.2894	3.1173	**3.1086**	3.1169	3.1149	3.1345
Evaluating indicator	Three orders							
	one	two	three	four	five	six	seven	eight
MAE	4.2600	3.3278	3.1375	3.0212	3.0010	**2.9900**	3.0060	3.0229
Evaluating indicator	Four orders							
	one	two	three	four	five	six	seven	eight
MAE	4.2212	3.4971	3.3792	3.2613	3.2463	**3.2403**	3.2514	3.2454
Evaluating indicator	Five orders							
	one	two	three	four	five	six	seven	eight
MAE	4.3805	3.8760	2.9951	2.9242	2.9176	**2.9082**	2.9135	2.9340

of this paper is to modify the problem of LSTM gradient disappearance and to fix the optimal wavelet layers and orders and layers for PM 2.5 from the different study sites. The following study results may be gained, in view of the Experiments I and II.

(1) In view of the Experiment I, the forecasting performance of SAE-LSTM is much more outstanding than that of other forecasting algorithms considered for comparison.

(2) In Experiment II, for the different samples from the study sites, four orders five layers, five orders six layers, five

orders seven layers, three orders six layers, five orders seven layers, and five orders six layers are very rightness.

Although this study fixes the optimal wavelet layers and orders of the different samples and improves the problem of LSTM gradient disappearance, there are still some weak points that need to be addressed in future research:

(1) For the different PM2.5 time series, the optimal wavelet layers and orders are fixed, but for other time series, whether these fixed optimal wavelet layers and orders are appropriate or not?

(2) The parameters are set up in this study, which are fixed. In future studies, the optimization algorithms will be adopted to optimize the hyper-parameters in SAE-LSTM, e.g. meta-heuristic algorithms [40].

(3) The algorithm built in this paper has very good performance for PM2.5 prediction. Can this algorithm be applied to other fields, such as [41]–[45]?

IV. CONCLUSION

In this study, for different PM2.5 time series, Coiflets wavelet is adopted to decompose them into 160 high- and low-frequency components, the different neural network models (e.g. ELM) are adopted for comparison. Besides, the comprehensive evaluation indexes are applied to test the performance of SAE-LSTM. At last, some interesting conclusions are drawn:

(1) Comparing with other forecasting models considered for comparison in Experiment I, the forecasting performance of SAE-LSTM is improved. This experimental result implies that SAE-LSTM modifies the problem of the LSTM gradient disappearance to some extent.

(2) The optimal wavelet layers and orders are determined for six kinds of samples based on the SAE-LSTM.

REFERENCES

[1] W. Sun and J. Sun, "Daily $PM_{2.5}$ concentration prediction based on principal component analysis and LSSVM optimized by cuckoo search algorithm," *J. Environ. Manage.*, vol. 188, pp. 144–152, Mar. 2017.

[2] M. Niu, K. Gan, S. Sun, and F. Li, "Application of decomposition-ensemble learning paradigm with phase space reconstruction for day-ahead $PM_{2.5}$ concentration forecasting," *J. Environ. Manage.*, vol. 196, pp. 110–118, Jul. 2017.

[3] K. Gan, S. Sun, S. Wang, and Y. Wei, "A secondary-decomposition-ensemble learning paradigm for forecasting $PM_{2.5}$ concentration," *Atmos. Pollut. Res.*, vol. 9, no. 6, pp. 989–999, 2018.

[4] J. Du, F. Qiao, and L. Yu, "Temporal characteristics and forecasting of $PM_{2.5}$ concentration based on historical data in Houston, USA," *Resour., Conservation Recycling*, vol. 147, pp. 145–156, Aug. 2019.

[5] Y. Qi, Q. Li, H. Karimian, and D. Liu, "A hybrid model for spatiotemporal forecasting of $PM_{2.5}$ based on graph convolutional neural network and long short-term memory," *Sci. Total Environ.*, vol. 664, pp. 1–10, May 2019.

[6] F. Zhao and W. Li, "A combined model based on feature selection and WOA for $PM_{2.5}$ concentration forecasting," *Atmosphere*, vol. 10, no. 4, p. 223, 2019.

[7] Z. Shang, T. Deng, J. He, and X. Duan, "A novel model for hourly $PM_{2.5}$ concentration prediction based on CART and EELM," *Sci. Total Environ.*, vol. 651, pp. 3043–3052, Feb. 2019.

[8] D.-J. Liu and L. Li, "Application study of comprehensive forecasting model based on entropy weighting method on trend of $PM_{2.5}$ concentration in Guangzhou, China," *Int. J. Environ. Res. Public Health*, vol. 12, no. 6, pp. 7085–7099, 2015.

[9] B. Lv, W. G. Cobourn, and Y. Bai, "Development of nonlinear empirical models to forecast daily $PM_{2.5}$ and ozone levels in three large Chinese cities," *Atmos. Environ.*, vol. 147, pp. 209–223, Dec. 2016.

[10] S. Ausati and J. Amanollahi, "Assessing the accuracy of ANFIS, EEMD-GRNN, PCR, and MLR models in predicting $PM_{2.5}$," *Atmos. Environ.*, vol. 142, pp. 465–474, Oct. 2016.

[11] B. Lyu, Y. Zhang, and Y. Hu, "Improving $PM_{2.5}$ air quality model forecasts in China using a bias-correction framework," *Atmosphere*, vol. 8, no. 8, p. 147, 2017.

[12] X. Y. Ni, H. Huang, and W. P. Du, "Relevance analysis and short-term prediction of $PM_{2.5}$ concentrations in Beijing based on multi-source data," *Atmos. Environ.*, vol. 150, pp. 146–161, Feb. 2017.

[13] P. Wang, H. Zhang, Z. Qin, and G. Zhang, "A novel hybrid-Garch model based on ARIMA and SVM for $PM_{2.5}$ concentrations forecasting," *Atmos. Pollut. Res.*, vol. 8, no. 5, pp. 850–860, 2017.

[14] L. Zhang, J. Lin, R. Qiu, X. Hu, H. Zhang, Q. Chen, H. Tan, D. Lin, and J. Wan, "Trend analysis and forecast of $PM_{2.5}$ in Fuzhou, China using the ARIMA model," *Ecological Indicators*, vol. 95, pp. 702–710, Dec. 2018.

[15] S. Mahajan, L.-J. Chen, and T.-C. Tsai, "Short-term $PM_{2.5}$ forecasting using exponential smoothing method: A comparative analysis," *Sensors*, vol. 18, no. 10, p. 3223, 2018.

[16] A. B. Chelani, "Estimating $PM_{2.5}$ concentration from satellite derived aerosol optical depth and meteorological variables using a combination model," *Atmos. Pollut. Res.*, vol. 10, no. 3, pp. 847–857, 2019.

[17] M. Niu, Y. Wang, S. Sun, and Y. Li, "A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term $PM_{2.5}$ concentration forecasting," *Atmos. Environ.*, vol. 134, pp. 168–180, Jun. 2016.

[18] W. Li, D. Kong, and J. Wu, "A new hybrid model FPA-SVM considering cointegration for particular matter concentration forecasting: A case study of kunming and yuxi, China," *Comput. Intell. Neurosci.*, vol. 2017, Aug. 2017, Art. no. 2843651.

[19] Y. Chen, F. Li, Z. Deng, X. Chen, and J. He, "$PM_{2.5}$ forecasting with hybrid LSE model-based approach," *Softw., Pract. Exper.*, vol. 47, no. 3, pp. 379–390, 2017.

[20] F. Biancofiore, M. Busilacchio, M. Verdecchia, B. Tomassetti, E. Aruffo, S. Bianco, S. Di Tommaso, C. Colangeli, G. Rosatelli, and P. Di Carlo, "Recursive neural network model for analysis and forecast of PM10 and $PM_{2.5}$," *Atmos. Pollut. Res.*, vol. 8, no. 4, pp. 652–659, 2017.

[21] J. Qiao, J. Cai, H. Han, and J. Cai, "Predicting $PM_{2.5}$ concentrations at a regional background station using second order self-organizing fuzzy neural network," *Atmosphere*, vol. 8, no. 1, p. 10, 2017.

[22] P. Jiang, Q. Dong, and P. Li, "A novel hybrid strategy for $PM_{2.5}$ concentration analysis and prediction," *J. Environ. Manage.*, vol. 196, pp. 443–457, Jul. 2017.

[23] J.-H. Chang and C.-Y. Tseng, "Analysis of correlation between secondary $PM_{2.5}$ and factory pollution sources by using ANN and the correlation coefficient," *IEEE Access*, vol. 5, pp. 22812–22822, 2017.

[24] C.-J. Huang and P.-H. Kuo, "A deep CNN-LSTM model for particulate matter ($PM_{2.5}$) forecasting in smart cities," *Sensors*, vol. 18, no. 7, p. 2220, 2018.

[25] Y. Chen, "Prediction algorithm of $PM_{2.5}$ mass concentration based on adaptive BP neural network," *Computing*, vol. 100, no. 8, pp. 825–838, 2018.

[26] Y. Bai, B. Zeng, C. Li, and J. Zhang, "An ensemble long short-term memory neural network for hourly $PM_{2.5}$ concentration forecasting," *Chemosphere*, vol. 222, pp. 286–294, May 2019.

[27] J. Zhao, F. Deng, Y. Cai, and J. Chen, "Long short-term memory-Fully connected (LSTM-FC) neural network for $PM_{2.5}$ concentration prediction," *Chemosphere*, vol. 220, pp. 486–492, Apr. 2019.

[28] X. Feng, Q. Li, and Y. Zhu, J. Hou, L. Jin, and J. Wang, "Artificial neural networks forecasting of $PM_{2.5}$ pollution using air mass trajectory based geographic model and wavelet transformation," *Atmos. Environ.*, vol. 107, pp. 118–128, Apr. 2015.

[29] D. Wang, Y. Liu, H. Luo, C. Yue, and S. Cheng, "Day-ahead $PM_{2.5}$ concentration forecasting using WT-VMD based decomposition method and back propagation neural network improved by differential evolution," *Int. J. Environ. Res. Public Health*, vol. 14, no. 7, p. 764, 2017.

[30] H. Zhang, S. Zhang, P. Wang, Y. Qin, and H. Wang, "Forecasting of particulate matter time series using wavelet analysis and wavelet-ARMA/ARIMA model in Taiyuan, China," *J. Air Waste Manage. Assoc.*, vol. 67, no. 7, pp. 776–788, 2017.

[31] H. Liu, Z. Duan, and C. Chen, "A hybrid framework for forecasting $PM_{2.5}$ concentrations using multi-step deterministic and probabilistic strategy," *Air Qual., Atmos. Health*, vol. 12, no. 7, pp. 785–795, 2019.

[32] H. Liu, K. Jin, and Z. Duan, "Air $PM_{2.5}$ concentration multi-step forecasting using a new hybrid modeling method: Comparing cases for four cities in China," *Atmos. Pollut. Res.*, vol. 10, no. 5, pp. 1588–1600, 2019.

[33] H. Liu, Y. Xu, and C. Chen, "Improved pollution forecasting hybrid algorithms based on the ensemble method," *Appl. Math. Model.*, vol. 73, pp. 473–486, Sep. 2019.

[34] Z. Chen and W. Li, "Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1693–1702, Jul. 2017.

[35] U. Lešnik, D. Mongus, and D. Jesenko, "Predictive analytics of PM_{10} concentration levels using detailed traffic data," *Transp. Res. D, Transp. Environ.*, vol. 67, pp. 131–141, Feb. 2019.

[36] W. Qiao, K. Huang, M. Azimi, and S. Han, "A novel hybrid prediction model for hourly gas consumption in supply side based on improved whale optimization algorithm and relevance vector machine," *IEEE Access*, vol. 7, pp. 88218–88230, 2019. doi: 10.1109/ACCESS.2019.2918156.

[37] H. Lu, K. Huang, M. Azimi, and L. Guo, "Blockchain technology in the oil and gas industry: A review of applications, opportunities, challenges, and risks," *IEEE Access*, vol. 7, pp. 41426–41444, 2019. doi: 10.1109/ACCESS.2019.2907695.

[38] Q. Weibiao, L. Bingfan, and K. Zhangyang, "Differential scanning calorimetry and electrochemical tests for the analysis of delamination of 3PE coatings," *Int. J. Electrochem. Sci.*, vol. 14, pp. 7389–7400, Aug. 2019.

[39] W. Qiao and Y. Zhe, "Modified dolphin swarm algorithm based on chaotic maps for solving high-dimensional function optimization problems," *IEEE Access*, vol. 7, pp. 110472–110486, 2019. doi: 10.1109/ACCESS.2019.2931910.

[40] W. Qiao and Z. Yang, "Solving large-scale function optimization problem by using a new metaheuristic algorithm based on quantum dolphin swarm algorithm," *IEEE Access*, to be published. doi: 10.1109/ACCESS.2019.2942169.

[41] W. Liu, Z. Zhang, J. Chen, J. Fan, D. Jiang, Y. Li, and D. Jjk, "Physical simulation of construction and control of two butted-well horizontal cavern energy storage using large molded rock salt Specimens," *Energy*, vol. 185, pp. 682–694, Oct. 2019.

[42] E. Liu, W. Li, H. Cai, and S. Peng, "Formation mechanism of trailing oil in product oil pipeline," *Processes*, vol. 7, no. 1, p. 7, 2019.

[43] Z. Su, E. Liu, Y. Xu, P. Xie, C. Shang, and Q. Zhu, "Flow field and noise characteristics of manifold in natural gas transportation station," *Oil Gas Sci. Technol.*, vol. 74, p. 70, Jul. 2019. doi: 10.2516/ogst/2019038.

[44] W. Qiao and H. Wang, "Analysis of the wellhead growth in HPHT gas wells considering the multiple annuli pressure during production," *J. Natural Gas Sci. Eng.*, vol. 50, pp. 43–54, Feb. 2018.

[45] W. Qiao, H. Lu, G. Zhou, M. Azimi, Q. Yang, and W. Tian, "A hybrid algorithm for carbon dioxide emissions forecasting based on improved lion swarm optimizer," *J. Cleaner Prod.*, 2019, Art. no. 118612. doi: 10.1016/j.jclepro.2019.118612.

WEIBIAO QIAO received the B.S. degree in information and computing science from Northeast Agricultural University, Harbin, in 2009, the M.S. degree in oil and gas storage and transportation engineering from Liaoning Shihua University, Fushun, in 2012, and the Ph.D. degree in oil and gas storage and transportation engineering from China Petroleum University, Qingdao, in 2017.

Since 2017, he has been a Lecturer with the North China University of Water Resources and Electric Power, Zhengzhou. His research interests include the gas consumption forecasting, gas pipeline robot detection, and intelligent scheduling of gas storage group.

WENCAI TIAN was born in Anyang, Henan, in 1998.

He is currently pursuing the bachelor's degree with the North China University of Water Resources and Electric Power. He is involved in swarm intelligence algorithm and machine learning.

He received the National Encouragement Scholarship, in 2017 and 2018. He received the Excellent Award from the Huacai Cup Competition. The title of the award-winning project is the Multifunctional Window Purifier Project.

YU TIAN was born in Puyang, Henan, in 1997.

He is currently pursuing the bachelor's degree with the North China University of Water Resources and Electric Power. He is involved in swarm intelligence algorithm and machine learning.

In the school year 2017 to 2018, he received the three good students award from the school.

QUAN YANG was born in Zhoukou, Henan, in 1999.

He is currently pursuing the bachelor's degree with the North China University of Water Resources and Electric Power. He is involved in swarm intelligence algorithm and machine learning.

He received the third prize of the Green Source Cup from the North China University of Water Resources and Hydropower.

YINING WANG was born in Gongyi, Henan, in 1997.

He is currently pursuing the bachelor's degree with the North China University of Water Resources and Electric Power. He is involved in swarm intelligence algorithm and machine learning.

He received the second prize in the 22nd China Daily 21st century Coca-Cola Cup National English Speech Contest campus selection competition from the North China University of Water Resources and Electric Power.

JIANZHUANG ZHANG was born in Anyang, Henan, in 1996.

He is currently pursuing the bachelor's degree with the North China University of Water Resources and Electric Power. He is involved in swarm intelligence algorithm and machine learning.

In 2018, he received the third prize of the Dahe Seiko Cup Fire Design Competition based on artificial intelligence technology.