IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 15, 2019, accepted May 3, 2019, date of publication May 10, 2019, date of current version May 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2916000

A Hybrid Vibration Signal Prediction Model Using Autocorrelation Local Characteristic-Scale Decomposition and Improved Long Short Term Memory

HUI-XIN TIAN[1,2], DAI-XU REN[1,2], AND KUN LI[3]

[1]School of Electrical Engineering and Automation, Tianjin Polytechnic University, Tianjin 300387, China
[2]Key Laboratory of Advanced Electrical Engineering and Energy Technology, Tianjin Polytechnic University, Tianjin 300387, China
[3]School of Management, Tianjin Polytechnic University, Tianjin 300387, China

Corresponding author: Kun Li (lk_neu@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 71602143, Grant 61403277, and Grant 51607122, in part by the Tianjin Natural Science Foundation under Grant 18JCYBJC22000, in part by the Tianjin Science and Technology Correspondent Project under Grant 18JCTPJC62600, and in part by the Tianjin High School Innovation Team Training Program under Grant TD13-5036 and Grant TD13-5038.

ABSTRACT Accurate state monitoring and the fault prediction model is very important for the smooth running of a reciprocating compressor. Vibration signal is a sensitive characteristic parameter for fault prediction of a reciprocating compressor. Thus, it is necessary to develop an accurate and stable vibration signal prediction model. However, it is difficult to predict using a simple model for its nonlinear and nonstationary characteristics. Aiming at the characteristics of the vibration signal, a hybrid prediction modeling strategy called ACLCD-PSOLSTM is proposed by combining autocorrelation local characteristic-scale decomposition (ACLCD) and improved long short term memory (LSTM) neural network. To reduce the complexity of modeling, the original vibration signal is decomposed into many intrinsic scale components (ISCs) and a residue item by ACLCD. Then, each of the ISCs is predicted using the particle swarm optimization LSTM (PSOLSTM) model. And all the predicted results are accumulated as the final predicted result of the vibration signal, where the autocorrelation characteristics of the signal are considered to overcome the end effect of traditional LCD. For better performance of the LSTM prediction model, a multiobjective optimization model is established that balanced the prediction ability of the LSTM (RMSE) with the model complexity (hidden neurons and time lags). And the model is solved by the PSO algorithm. To validate the predicting capacity of the proposed hybrid ACLCD-PSOLSTM model, four different predicting models are implemented on the vibration signal series. The results of experiments show the superiority of the hybrid model over other models in improving the predictive performance.

INDEX TERMS Vibration signal predicting, signal autocorrelation, local characteristic-scale decomposition, long short term memory neural network, particle swarm optimization, hybrid model.

I. INTRODUCTION

Reciprocating compressor is the core equipment in petrochemical production. Once the compressor breaks down, it will affect the whole production process, reduce production efficiency and even cause environmental pollution. Thus, accurate fault prediction is crucial to the smooth running of the reciprocating compressor. Vibration signal is a sensitive characteristic parameter for fault prediction of reciprocating compressor [1]. Most of the faults, such as air valve damage, piston rod sinking, crosshead bolt loosening, connecting rod wear, are accompanied by abnormal vibration signal [2], [3]. Therefore, accurate prediction of vibration signal is the basis of compressor fault prediction. However, due to the complexity of the production environment, the vibration signals often present nonlinear and nonstationary characteristics, which

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

brings great difficulties to modeling. Therefore, it is still a challenge to establish an accurate vibration signal prediction model.

The vibration signal produced in the running process of compressor has the characteristics of time series. Time series prediction models can be divided into three categories: linear models, intelligent models and hybrid models. The linear model is a very widely used time series model because of its simple structure. The general linear time series model is composed of Auto-Regressive (AR) model, Moving Average (MA) model and the Auto-Regressive Moving Average (ARMA) [4]–[7]. These models only apply to linear systems, describing linear relationships between variables. Intelligent models [8]–[12], such as Back Propagation neural network (BP) can process complex nonlinear data and capture the inherent characteristics of data, which has been widely studied in recent years. However, the traditional intelligent model cannot solve the timing problem in the data due to the limitation of structure. The Recurrent Neural Network (RNN) [13] improves the connection mode of neurons in the traditional neural network, connecting neurons at different moments. Hence, the information of the present moment is relevant not only to the present moment but also to the past. In recent years, RNN has been widely used in speech recognition and machine translation [14]–[16]. However, RNN is easily affected by the vanishing gradient problem, which makes it difficult to deal with the long-term dependencies problem [17]. Long short-term memory(LSTM) [18] is a special type of RNN architecture, to solve the vibration signals predicting problem. LSTM performs well in dealing with long-term dependencies problems and can suppress the vanishing gradient problem. Li *et al.* [19] and Hu and Chen [20] predicted wind speed series used LSTM network. The results show that LSTM network is superior to other traditional neural networks in the prediction of wind speed series. Lin *et al.* [21] predicted the operation state of the transformer with LSTM network and achieved good results.

However, LSTM will encounter various problems in the use process, such as slow convergence, overfitting and parameter selection and so on. Gregory and Gauvain [22] proposed a modified version of the LSTM neural network in which direct links are added between the three gates of a LSTM cell. This modification aims to prevent that convergence rate from being too slow when the LSTM activation function gets stuck in a saturated state. Zhao *et al.* [23] proposed an improved LSTM network. This method introduced batch normalization procedure that aims to reduce the internal covariate shift of LSTM. ElSaid *et al.* [24] proposed an improved LSTM model that changed the way neurons are fully connected. This method speeds up the convergence and avoids the overfitting problem. As there is no scientific guidance to choose the number of neurons in hidden layer and the number of time lags of LSTM network, an improved model PSOLSTM is presented in this paper which PSO is introduced to solve the optimal number of neurons and lags by minimizing a multiobjective function.

Due to the limited information that a single model can process, more and more scholars use hybrid model to improve the prediction accuracy in recent years. To improve the prediction accuracy, Jatin and Toshniwal [25] use Empirical Mode Decomposition (EMD) to decompose the signal before establishing the LSTM prediction model. EMD [26], [27] can decompose complex signals into multiple simple components to extract local characteristics of signals. The advantage of EMD over wavelet transform is that there is no need to define base function in advance. Yu *et al.* [28] developed a diagnosis model using EMD and Artificial neural network (ANN) to improve the accuracy of fault diagnosis. Sharma *et al.* [29] presented a novel signal classification method using EMD and least squares support vector machine. However, there are two obvious drawbacks in EMD: iteration time is long and mode mixing. Recently, Zheng *et al.* [30] presented a new signal decomposition method called local characteristic-scale decomposition (LCD) and demonstrated that LCD is better than EMD in processing complex signals because LCD can reduce invalid components and suppress mode mixing. Liu *et al.* [31] proposed a bearing fault diagnosis method based on LCD-Teager Energy Operator to improve accuracy. However, in the process of decomposition, LCD tends to produce end effect, which leads to the large error of the decomposed signal at the end. Some methods to solve end effect are proposed in the literature [32]–[36], such as image continuation and extremum continuation. However, it does not work well for nonstationary vibration signals. In this paper, an improved signal decomposition method called Autocorrelation-LCD (ACLCD) is proposed which consider the autocorrelation of signal. In summary, to reduce the complexity of modeling, the complex vibration signal is decomposed into some ISCs and a residue item. Then the prediction model is established for each component. Hence, a hybrid model ACLCD-PSOLSTM is proposed to enhance the prediction performance of vibration signals.

The structure of this paper is as follows. Section 2 describes the proposed model and relevant theoretical knowledge in detail. Then, the validity of the proposed model is verified through different experiments and the experimental results are analyzed in detail in Section 3. Finally, the conclusion of the paper is summarized in Section 4.

II. METHODOLOGY

In Section 2, the proposed model and relevant theoretical knowledge are described as follows, including LCD, ACLCD, LSTM network, PSOLSTM.

A. LOCAL CHARACTERISTIC-SCALE DECOMPOSITION (LCD)

Local Characteristic-scale Decomposition (LCD), proposed by Zheng *et al.* [30] in 2013. The LCD approach assumes that any complex signal can be decomposed into many independent ISC components. Any ISCs must satisfy the following two conditions:

(1) In the whole data segment, there is monotonicity between any two adjacent local maxima and local minima.

(2) In the whole data segment, assume that all extreme point for x_k, the corresponding moment is t_k, $(k = 1, 2, \ldots, m)$. A line that connects any two adjacent local maxima (minima) (t_{k-1}, x_{k-1}) and (t_{k+1}, x_{k+1}). The intermediate local minima (maxima) (t_k, x_k) that corresponding point (t_k, A_k) is denoted as:

$$A_k = x_{k-1} + \frac{t_k - t_{k-1}}{t_{k+1} - t_{k-1}}(x_{k+1} - x_{k-1}) \tag{1}$$

More generally, the proportions of x_k and A_k remain constant, as follows:

$$aA_k + (1 - a)x_k = 0, a \in (0, 1) \tag{2}$$

where a is a proportional coefficient. Generally, a is set as 0.5.

The single component signal satisfying the above conditions (1) (2) is called intrinsic scale component (ISC). The LCD of arbitrary signal $x(t)$ is performed and decomposed into many ISC components and a residue item. The algorithm is as follows:

Where θ is iteration termination threshold. Traditional LCD iteration termination condition is that $r(t)$ is a constant function. However, the condition is difficult to satisfy in the actual situation. Hence, the iteration is terminated when the variance of $r(t)$ is less than the threshold θ in this paper. SD is the termination criterion, as follows:

$$SD = \sum_{t=0}^{T} \left[\frac{|(h_{k-1}(t) - h_k(t)|^2}{h_{k-1}^2(t)} \right] \tag{3}$$

If SD <= 0.3 for the kth loop, it is considered as ISC component; otherwise, the loop continues until the variance of $r(t)$ is less than the threshold θ. Finally, $x(t)$ will be decomposed into the sum of N ISCs and a monotone function$r(t)$, as follows:

$$x(t) = \sum_{i=1}^{n} ISC_i(t) + r(t) \tag{4}$$

For clarity, the partial decomposition process of LCD is shown in Fig. 1.

From the above LCD decomposition process can be noted that the value of L_k is from 2 to $m - 1$. It is necessary to estimate the values of L_1 and L_m at ends. This is the reason for the end effect in the process of LCD decomposition. End effect is one of the main factors that affect the accuracy of this method in signal decomposition. There are two main ways to deal with the end effect: one is to extend the original signal and bring the original end into the interior of the new signal, thus eliminating the uncertainty of the extreme at the end. Another is to extend the sequence of signal extremum. However, it does not work well for nonstationary vibration signals. Based on the above two ideas, an improved signal decomposition method called Autocorrelation-LCD (ACLCD) is proposed which considers the autocorrelation of signal.

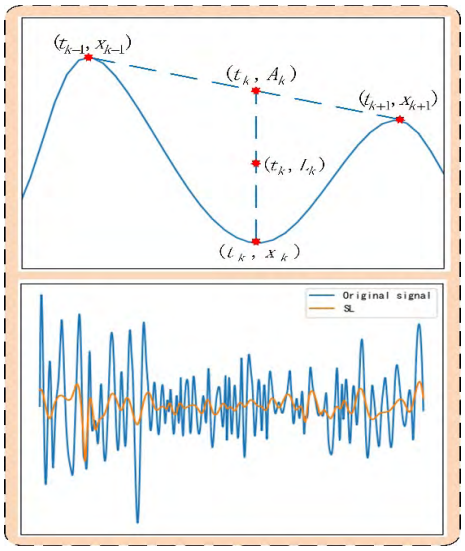


FIGURE 1. The partial decomposition process of LCD.

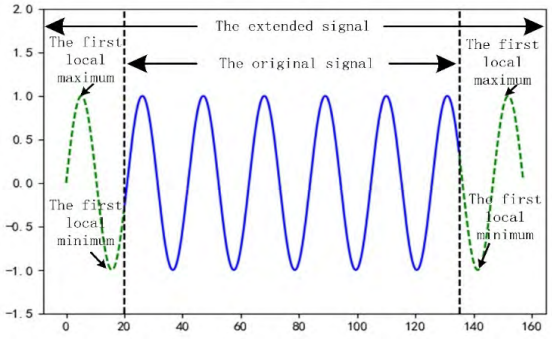


FIGURE 2. The process of signal continuation.

B. AUTOCORRELATION LOCAL CHARACTERISTIC-SCALE DECOMPOSITION (ACLCD)

In this section, based on Radial Basis Function (RBF) neural network prediction method, two ends of signal are predicted separately. RBF can approximate arbitrary nonlinear signals, but its premise is that RBF neural network can accurately predict the signal. However, when the signal components are complex or have more interference (such as the vibration signal in this paper), the fitting ability of RBF neural network is poor, and sometimes even the minimum point can be fitted at the position where the maximum value should be, which obviously has a huge impact on the whole algorithm.

To solve this problem, an extremum correction algorithm is designed. The basic idea is: According to the autocorrelation characteristics of signals [37], the sum of all local maxima or local minima can affect the trend of signals for a certain scale of signals. The RBF neural network was used to fit the extended signals containing both local maximum and local minimum at the front end and the back end of the signals, as shown in Fig. 2. Next, the sum of all local maxima and local

minima of the original signal is calculated, and the local maxima and local minima in the extension interval are corrected to obtain the complete extension signal. Finally, the extended signal is decomposed to LCD, and the decomposed signal is intercepted to the original signal interval, to obtain the final decomposed signal.

The algorithm process is as follows:

(1) A signal is described as follow:

$$\{x(t)|t = 1, 2, \ldots, N\}$$

(2) Locate all the local maximum of $x(t)$:

$$m_1, m_2, \ldots, m_i(i = 1, 2, \ldots, m)$$

and all the local minimum of $x(t)$:

$$h_1, h_2, \ldots, h_j(j = 1, 2, \ldots, n)$$

Then the mean of the local maximum and local minimum can be obtained:

$$\overline{m} = \frac{1}{m}\sum_{i=1}^{m} m_i, \quad \overline{h} = \frac{1}{n}\sum_{j=1}^{n} h_j \tag{5}$$

(3) The training data P and T are shown as follows:

$$P = \begin{bmatrix} x_1 & x_2 & x_3 \\ x_2 & x_3 & x_4 \\ \vdots & \vdots & \vdots \\ x_{t-3} & x_{t-2} & x_{t-1} \end{bmatrix}, T = \begin{bmatrix} x_4 \\ x_5 \\ \vdots \\ x_t \end{bmatrix} \tag{6}$$

Here P is the input and the T is the output of RBF. Then the next period signal is predicted a $(x_{t+1}, x_{t+2}, \ldots, x_{t+k}, x_{t+k+1}, \ldots, x_{t+l}, x_{t+l+1})(l > k > t)$, x_{t+k} is the local minimum and x_{t+l} is the local maximum.

(4) Set

$$P = \begin{bmatrix} x_t & x_{t-1} & x_{t-2} \\ \vdots & \vdots & \vdots \\ x_5 & x_4 & x_3 \\ x_4 & x_3 & x_2 \end{bmatrix}, T = \begin{bmatrix} x_{t-3} \\ \vdots \\ x_2 \\ x_1 \end{bmatrix}$$

Here P is the input and the T is the output of RBF. Then the previous period signal is predicted as $(o_1, o_2, \ldots, o_r, o_{r+1}, \ldots, o_q, o_{q+1})(q > r > 1)$, o_r is the local minimum and o_q is the local maximum.

(5) To calculate:

$$\max\_\arg 1 = \alpha_1 \times \overline{m} + \beta_1 \times x_{t+l}, (\alpha_1 + \beta_1 = 1) \tag{7}$$

$$\min\_\arg 2 = \alpha_2 \times \overline{m} + \beta_2 \times o_q, (\alpha_2 + \beta_2 = 1) \tag{8}$$

$$\min\_\arg 1 = \alpha_3 \times \overline{h} + \beta_3 \times x_{t+k}, (\alpha_3 + \beta_3 = 1) \tag{9}$$

$$\min\_\arg 2 = \alpha_4 \times \overline{h} + \beta_4 \times o_r, (\alpha_4 + \beta_4 = 1) \tag{10}$$

Here α, β is a custom parameter.

(6) Correct extremum:

If $\max\_\arg 1 > x_{t+l}$ and $1 \leq \max\_\arg 1/x_{t+l} \leq \gamma$, Replace x_{t+l} with $\max\_\arg 1$, if it doesn't meet the conditions, it stays the same;

If $\max\_\arg 2 > o_q$ and $1 \leq \max\_\arg 2/o_q \leq \gamma$, Replace o_q with $\max\_\arg 2$, if it doesn't meet the conditions, it stays the same;

If $\min\_\arg 1 \leq x_{t+k}$ and $\varepsilon \leq \min\_\arg 1/x_{t+k} \leq 1$, Replace x_{t+k} with $\min\_\arg 1$, if it doesn't meet the conditions, it stays the same;

If $\min\_\arg 2 \leq o_r$ and $\varepsilon \leq \min\_\arg 2/o_r \leq 1$, Replace o_r with $\min\_\arg 2$, if it doesn't meet the conditions, it stays the same;

Here ε, γ is a correction factor, preventing false substitutions. Generally, γ is set as 1.2 and ε is 0.8.

(7) Get the extended signal:

$$x'(t) = [o_1, o_2, \ldots, o_r, o_{r+1}, \ldots, o_q, o_{q+1}, \ldots, x_1, x_2, \ldots, x_t, x_{t+1}, \ldots, x_{t+k}, x_{t+k+1}, \ldots, x_{t+l}, x_{t+l+1}] \tag{11}$$

(8) Decomposition $x'(t)$ using LCD:

$$x'(t) = \sum_{i=1}^{n} ISC_i'(t) + r'(t) \tag{12}$$

(9) Intercept the signal to the original signal interval and get the final decomposition result:

$$x(t) = \sum_{i=1}^{n} ISC_i(t) + r(t) \tag{13}$$

The RBF neural network fitting ability and autocorrelation characteristics of signal are fully considered in improved boundary extension method. The extreme of new end and neural network fitting are limited in the algorithm execution. The algorithm avoids the effect of signal mutation points on the end fitting and provides optional parameters to select different key points according to different signals. When the complexity of signal is high and the fitting effect of neural network is poor, α can be increased and β is reduced properly, whereas the signal is more regular and predictable, α can be reduced and β is increased.

C. LONG SHORT TERM MEMORY NETWORK (LSTM)

In the prediction of vibration signal, the vibration signal at the current moment is often related to the signal at the previous moment, and the value of the signal before and after is not independent. The traditional neural network cannot deal with this kind of time series problem well because of its limited structure. LSTM was proposed by Hochreiter and Schmidhuber [18]. in 1997 as a special kind of the recurrent neural networks (RNNs). LSTM model has widely used the field of time series prediction and has a strong ability in solving long-term dependence problems. The basic architecture of LSTM network is shown in Fig. 3. Because of three gates: input, forget and output, the LSTM network can selectively receive and delete information of cell state. The calculation formula of each gate and the update of cell state can be expressed

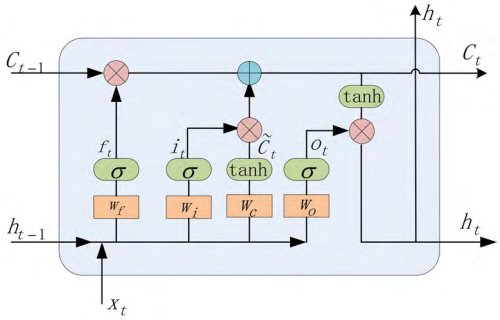


FIGURE 3. The architecture of the LSTM memory block.

as follows:

$$f_t = \sigma(w_f[h_{t-1}, x_t] + b_f) \tag{14}$$
$$i_t = \sigma(w_i[h_{t-1}, x_t] + b_i) \tag{15}$$
$$\tilde{c}_t = \tanh(w_c[h_{t-1}, x_t] + b_c) \tag{16}$$
$$o_t = \sigma(w_o[h_{t-1}, x_t] + b_o) \tag{17}$$
$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{18}$$
$$h_t = o_t \odot \tanh(c_t) \tag{19}$$

where x_t is the input at the current moment and h_{t-1} is the output of the hidden layer at the previous moment. i_t, f_t, and o_t denote the input gate, forget gate and output gate. $\tilde{c}_t$ denotes candidate information of input gate. And c_t denotes state value at current moment while h_t means the output value of LSTM memory cell at current moment. σ and tanh denote activation function. $\odot$ denotes the dot product of the vectors. Additionally, w_f, w_i, w_c and w_o are the weight to be learned; b_f, b_i, b_c and b_o are the corresponding bias vectors.

D. PSO-BASED OPTIMIZATION OF LSTM NETWORK (PSOLSTM)

In the design process of LSTM network, as the parameters such as the number of hidden layer neurons and time lags are custom parameters, there is a lack of clear knowledge to guide researchers. Many researchers choose to use trial-and-error method or experience, which is time consuming and cannot be used effectively. So how to obtain the number of hidden layer neurons and the number of time lags automatically and efficiently is the key to using LSTM successfully. Here, we proposed an improved LSTM by designing an optimization strategy for the best performance of prediction.

In theory, the more neurons in the hidden layer and the longer the time lag within a certain range, the better the prediction capacity of the model. But at the same time, the model becomes more complex and computational speed slower. These can be described as a multiobjective optimization problem. To balance the prediction performance and the complexity of the model, the objective function and constraint are described as follows:

$$\min \begin{cases} f_1 = N_n + N_t \\ f_2 = RMSE = \sqrt{\dfrac{1}{N}\sum_{i=1}^{N}(predicted_i - observed_i)^2} \end{cases}$$
$$\text{s.t.} \begin{cases} LN_n \le N_n \le UN_n \\ LN_t \le N_t \le UN_t \end{cases} \tag{20}$$

where N_n is the number of hidden layer neurons and N_t is the number of time lags. Moreover, The LN_n and UN_n represent the lower and upper bounds of N_n while LN_t and UN_t represent the lower and upper bounds of N_t. RMSE is Root Mean Squared Error; $predicted_i$ is predicted value and $observed_i$ is observed value. Number of hidden layer neurons will be tested from 5 to 80 and number of time lags from 1 to 99. In the other word, the LN_n and UN_n are set as 5 and 80, and the LN_t and UN_t are set as 1 and 99 in this paper.

In order to calculate conveniently, we transform the multi-objective problem into a single-objective problem, the objective function is considered as follows:

$$f = \omega_1 f_1^{'} + f_2$$
$$f_1^{'} = h(N_n) + h(N_t) \tag{21}$$

where ω_1 is the weight coefficient to balance the predicting performance and algorithm complexity. $h(\cdot)$ denote Normalized function that prevents ignoring the change of f_2 because of different orders of magnitude.

PSO is a popular swarm intelligence algorithm for its simple and effective characters and has been widely used in many kinds of model optimization. Therefore, PSO algorithm was adopted to solve the above optimization problem, and the objective function (21) was taken as the Fitness function in the PSO algorithm. The Fitness function as follow:

$$\min \ Fitness = f(N_n, N_t) \tag{22}$$

To elaborate further, the main steps of the optimization of PSO-based LSTM network are described as following:

Step1: Initialization parameters of PSO algorithm.

Step2: The particle is randomly initialized according to the range of particle position and velocity.

Step3: The Fitness function value of each particle was calculated according to (22).

Step4: Update the optimal position of each particle according to the Fitness function value of the current moment. The smallest value among all particles is denoted as global optimal position.

Step5: Update the velocity and the position of each particle according to (23) -(24)

$$v_{ij}(t+1) = \omega v_{ij}(t) + c_1\gamma_1(pBest_{ij}(t) - x_{ij}(t)) + c_2\gamma_2(gBest_j(t) - x_{ij}(t)) \tag{23}$$
$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t) \tag{24}$$

where i denotes the number of particles, j denotes the jth dimension of the search space and t is iterations. Moreover, v denote particle velocity. x is particle position and denotes

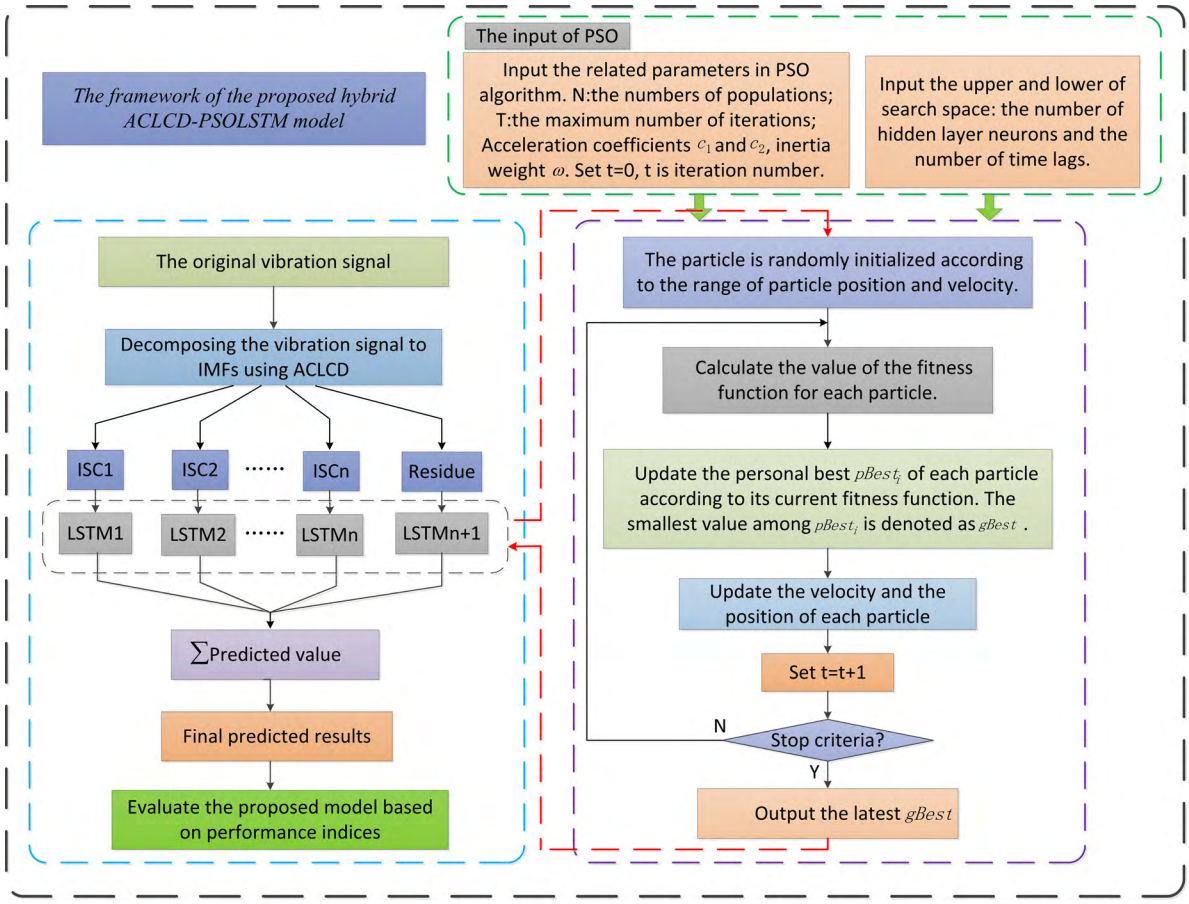


FIGURE 4. The framework of the proposed hybrid ACLCD-PSOLSTM model.

optimized numbers of neurons for LSTM network and time lags in this paper. *pBest*, *gBest* denote individual optimal position and global optimal position. ω is inertia weight, decides the range of motion of particles. Generally, ω is set as 0.5. c_1 and c_2 are acceleration coefficients and are usually set to 2. γ_1 and γ_2 are random numbers in range (0,1).

Step6: Repeat Steps 3-5 until the maximal number of iterations is reached.

Step7: The latest *gBest* is the optimized numbers of neurons for LSTM network and time lags.

E. THE NOVEL HYBRID ACLCD-PSOLSTM MODEL

Due to the complexity of the production environment, the vibration signal of reciprocating compressor often presents nonstationary characteristics, which brings great difficulties to accurate prediction. In this section, a new hybrid model is presented to enhance the predictive performance of the vibration signal based on ACLCD, LSTM network and PSO algorithm. The ACLCD is firstly used to decompose the vibration signal into many ISCs and a residue item to reduce the complexity of modeling. Then, each of the components is predicted using the PSOLSTM. Finally, the final result is obtained by summing the predicted results of each component. The whole process of the proposed hybrid ACLCD-PSOLSTM model is presented in Fig. 4. The main steps of the ACLCD-PSOLSTM are as follows:

Step1: Vibration signal decomposing. The original vibration signal is decomposed into some ISCs and a residue item using ACLCD.

Step2: ISCs and residue item modeling and predicting. PSO algorithm is used to determine the optimal number of hidden layer nodes and the number of time lags. Then the LSTM model is established for each component. Finally, many predicted ISCs and residue item results are obtained.

Step3: Obtain the final predicted result. The final result is obtained by summing the predicted results of each component.

Step4: Evaluate the proposed model. Several evaluation indicators are used to verify the performance of the proposed model.

In section 2, the proposed model and related theories are introduced in detail. First, aiming at the end effect problem in LCD decomposition process, ACLCD method is proposed according to the autocorrelation characteristics of the signal. Second, PSOLSTM model is presented which introduced PSO algorithm to optimize the number of neurons in hidden

layer and the number of time lags of LSTM network. Finally, to reduce the complexity of modeling and improving the accuracy of prediction model, a novel hybrid model ACLCD-PSOLSTM is proposed. In the next section, experiments will be conducted to verify the effectiveness of the proposed model.

III. EXPERIMENTS AND ANALYSIS

In this section, the performance of ACLCD-PSOLSTM hybrid model is to be validated on two experiments. The first experiment is to compare ACLCD with conventional decomposition methods including LCD and EMD. This experiment can prove that the improvement of LCD in computing speed and the superiority of signal autocorrelation in solving the problem of end effect. Another experiment is to test the vibration signal predicting capacity of the proposed hybrid model. According to the different forms of fault, compressor fault can be divided into gradual fault and abrupt fault. For the gradual fault, if the sampling interval of the signal is too short, the fault cannot be predicted effectively. Hence, two cases were studied including short-term vibration signal prediction and long-term vibration signal prediction in the second experiment. Four other predicting models are used for comparative experiments. The involved models consist of Back Propagation Neural Network (BP), LSTM, PSOLSTM and LCD-LSTM. It should be noted that all experiments run in the Python 3.6 environment on 2.80 GHz PC with process i5-7440HQ and 16G RAM.

In addition, three commonly used evaluation indices are employed to evaluate the performance of the hybrid ACLCD-PSOLSTM model. They are the mean absolute error (MAE), the root mean square error (RMSE) and the mean absolute percentage error (MAPE) that defined as follow:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|predicted_i - observed_i| \tag{25}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(predicted_i - observed_i)^2} \tag{26}$$

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{predicted_i - observed_i}{observed_i}\right| \times 100\% \tag{27}$$

where $predicted_i$ and $observed_i$ denote the predicted value and observed value of the $i$sample respectively. N denotes sample size.

A. EXPERIMENT I: COMPARISON RESULTS OF MODEL ACLCD, LCD AND EMD

To make the experimental results clearer, we use the simulation signal to verify the model. The signal as follow:

$$x(t) = x_1(t) + x_2(t) = (1 + 0.5\sin 5\pi t)\cos(250\pi t + 20\pi t^2) + 4\sin 40\pi t \quad t \in [0, 1] \tag{28}$$

The simulation signal consists of an AM-FM signal and a sinusoidal signal. The time domain waveforms are shown in Fig. 5.

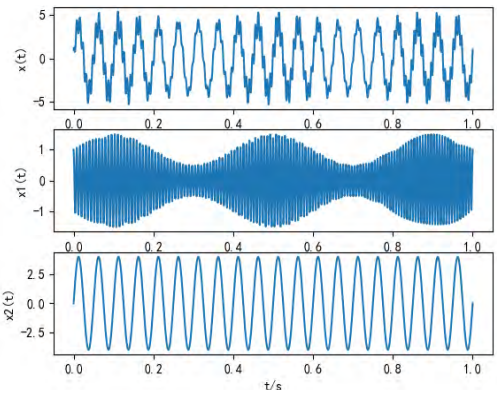


FIGURE 5. The waveforms of simulation signal and its components.

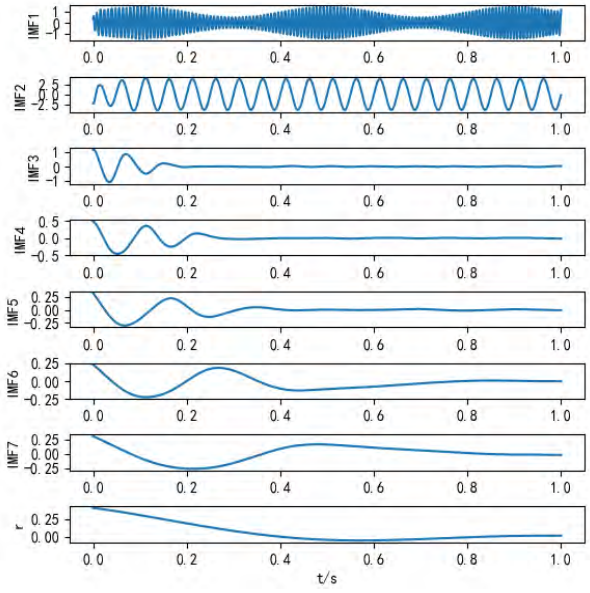


FIGURE 6. The result of EMD.

TABLE 1. Times and iterations times of each method.

Methods	Times/s	Iteration times
EMD	0.2005	17
LCD	0.0830	4
ACLCD	0.0674	4

First, EMD is used to decompose the signal. The termination criterion is standard deviation (SD) method, and the iteration was terminated when SD<0.3. The decomposed result is shown in Fig. 6.

Then, LCD is used to decompose the signal. The algorithm parameters are selected as follows: SD $= 0.3$, $\theta = 0.01$ and $a = 0.5$. The decomposed result is shown in Fig. 7.

Finally, ACLCD is used to decompose the signal. The algorithm parameters are selected as follows: SD $= 0.3$, $\theta = 0.01$, $a = 0.5$, $\varepsilon = 0.8$, $\gamma = 1, 2$, $\alpha = 0.35$ and $\beta = 0.65$. The decomposed result is shown in Fig. 8.

To compare the decomposition efficiency, Table 1 gives the time and iteration times required for each method.

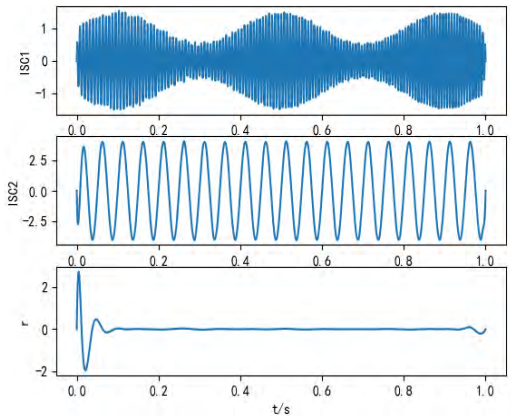


FIGURE 7. The result of LCD.

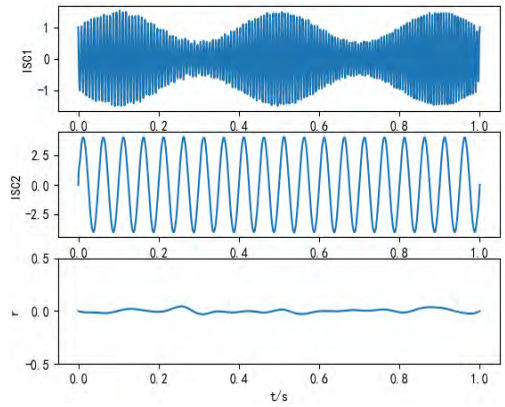


FIGURE 8. The result of ACLCD.

From Table 1 and Fig. 6-8, it can be observed as follows:

1) ACLCD method performs better than other methods. EMD method has the worst performance, whose values of Times and Iteration times are 0.2005 and 17.
2) EMD and LCD have obvious end effect and EMD produces more false components in the process of decomposition.
3) In summary, ACLCD method has obvious advantages in both computing speed and restraining end effect.

B. EXPERIMENT II: COMPARISON RESULTS OF MODEL ACLCD-PSOLSTM, BP, LSTM, PSOLSTM AND LCD-LSTM

1) DATA INFORMATION

In this experiment, the vibration signal data obtained from an oil production platform in Bohai, China, is used. In order to fully prove the superiority of the proposed model, two different time interval datasets that one-minute interval and ten-minute interval are employed to verify the prediction performance of different models. Each data set contains 5,000 data points. The first 3,000 data are used as the train set and the other 2000 data are used as the test set shown in Fig. 9 and Fig. 10. Moreover, in order to clearly display the data characteristics, the detailed information of data series is showed in Table 2 and Table 3.

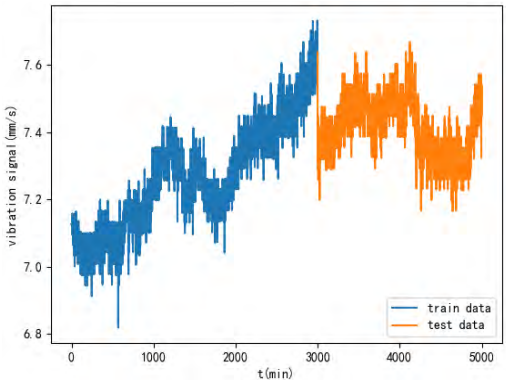


FIGURE 9. Original vibration signal time series with 1 min interval.

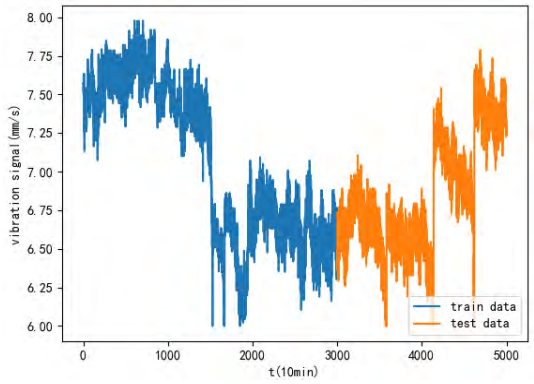


FIGURE 10. Original vibration signal time series with 10 min interval.

TABLE 2. The statistical result of vibration signal data with 1 min interval.

Data set	Max(mm/s)	Min(mm/s)	Mean(mm/s)	Standard deviation
Entire data set	7.7308	6.8195	7.3257	0.1534
Train data set	7.7308	6.8195	7.2670	0.1605
Test data set	7.6665	7.1647	7.4137	0.0851

TABLE 3. The statistical result of vibration signal data with 10 min interval.

Data set	Max(mm/s)	Min(mm/s)	Mean(mm/s)	Standard deviation
Entire data set	7.9790	6.0195	6.9764	0.4651
Train data set	7.9790	6.0195	7.0643	0.5001
Test data set	7.7885	6.0047	6.8445	0.3697

A one-step-ahead vibration signal prediction model is built in this experiment. For example, to signal$\{x(t)|t = 1, 2, \ldots, N\}$, the samples that can be generated are as follows:

$$P = \begin{bmatrix} x_1, x_2, \ldots, x_i \\ x_2, x_3, \ldots, x_{i+1} \\ \vdots \\ x_{N-i}, x_{N-i+1}, \ldots, x_N \end{bmatrix}, T = \begin{bmatrix} x_{i+1} \\ x_{i+2} \\ \vdots \\ x_N \end{bmatrix} \quad (29)$$

TABLE 4. Main parameters of five different models on case 1.

Models	Parameters
BP	40 time lags, 50 hidden layer neurons
LSTM	40 time lags, 30 hidden layer neurons
PSOLSTM	28 time lags, 52 hidden layer neurons
LCD-LSTM	40 time lags, 30 hidden layer neurons
ACLCD-PSOLSTM	28 time lags, 52 hidden layer neurons

where P is input data and T is output data. i is input data dimension, also called time lag in this paper. As can be seen from (29), different amount of lags will lead to different sample size. Since the time lag quantity is the optimization variable in PSO algorithm, the lag quantity of the model with PSO algorithm is different from that of the model without PSO algorithm. In the other word, the sample size of the model with PSO algorithm is different from that of the model without PSO algorithm. Hence, in order to better show the results of model comparison, we divide the model into two cases: the model with optimization algorithm and the model without optimization algorithm. The algorithm group #1 includes BP model, LSTM model and LCD-LSTM model and the algorithm group #2 includes PSOLSTM model and ACLCD-PSOLSTM model. According to the above methods, training samples and test samples are generated. Finally, the algorithm group #1 and algorithm group #2 both select 2000 training samples and 800 test samples to build models.

2) CASE STUDY 1: SHORT TERM VIBRATION SIGNAL PREDICTING

In this case, the one-min interval dataset is used to verify the performance of ACLCD-PSOLSTM model. Four other predicting models are chosen for comparative experiments. The involved models consist of BP Neural Network, LSTM, PSOLSTM and LCD-LSTM.

In order to balance the model complexity and prediction accuracy, the parameters of models without PSO optimization algorithm were determined by trial and error method. The parameters of the model are selected as follows. For BP model, the number of neurons in hidden layer, learn rate and the maximum number of iterations are set as 50, 0.1 and 100, respectively. For LSTM model, the number of neurons in hidden layer and the maximum number of iterations are set as 30 and 100. For ACLCD model, $SD = 0.3$, $\theta = 0.05$, $a = 0.5$, $\varepsilon = 0.7$, $\gamma = 1.2, \alpha = 0.55$ and $\beta = 0.45$. For PSO algorithm, $\omega_1 = 0.9$, $N = 20$, $T = 100$, $c_1 = 1.49, c_2 = 1.49$ and $\omega = 0.5$. The final optimization results of the PSO algorithm on the one-minute interval dataset are shown in Appendix A, where the best performance is highlighted in bold. For clearer display, the main parameters of five different models are shown in Table 4.

The experiment results obtained by five different models are shown in Table 5. Comparison results of different models

TABLE 5. The error evaluation results of different models for the vibration signal series on case 1.

Models	MAE(mm/s)	RMSE(mm/s)	MAPE(%)
BP	0.1109	0.1299	1.5121
LSTM	0.1022	0.1186	1.3941
PSOLSTM	0.0369	0.0460	0.5009
LCD-LSTM	0.0573	0.0759	0.7557
ACLCD-PSOLSTM	**0.0227**	**0.0286**	**0.3050**

TABLE 6. The error evaluation results of BP and LSTM for the vibration signal series on case 1.

Models	MAE(mm/s)	RMSE(mm/s)	MAPE(%)
BP	0.1109	0.1299	1.5121
LSTM	**0.1022**	**0.1186**	**1.3941**

TABLE 7. The error evaluation results of LSTM and PSOLSTM for the vibration signal series on case 1.

Models	MAE(mm/s)	RMSE(mm/s)	MAPE(%)
LSTM	0.1022	0.1186	1.3941
PSOLSTM	**0.0369**	**0.0460**	**0.5009**

TABLE 8. The error evaluation results of LSTM and LCD-LSTM for the vibration signal series on case 1.

Models	MAE(mm/s)	RMSE(mm/s)	MAPE(%)
LSTM	0.1022	0.1186	1.3941
LCD-LSTM	**0.0573**	**0.0759**	**0.7557**

TABLE 9. The error evaluation results of PSOLSTM and ACLCD-PSOLSTM for the vibration signal series on case 1.

Models	MAE(mm/s)	RMSE(mm/s)	MAPE(%)
PSOLSTM	0.0369	0.0460	0.5009
ACLCD-PSOLSTM	**0.0227**	**0.0286**	**0.3050**

are shown in Tables 6-9. In order to show the results more clearly, we only compare the first 100 test samples. The results of all the test samples are shown in Appendix A. The vibration signal predicting results of the models above are shown in Fig. 11.

From Tables 5-9 and Fig. 11, the following conclusions can be drawn:

1) The ACLCD-PSOLSTM model performs better than four other models, whose values of MAE, RMSE and MAPE are 0.0227, 0.0286 and 0.3050. BP model has the worst performance, whose values of MAE, RMSE and MAPE are 0.1109, 0.1299 and 1.5121. This indicates that the traditional neural network is not suitable for processing complex time series data.
2) According to Table 6, due to LSTM can solve long-term dependencies problem, LSTM provides a better result than traditional BP model.
3) As shown in Table 7, PSOLSTM model has smaller error than LSTM model. It shows that PSO algorithm can improve the accuracy of prediction model

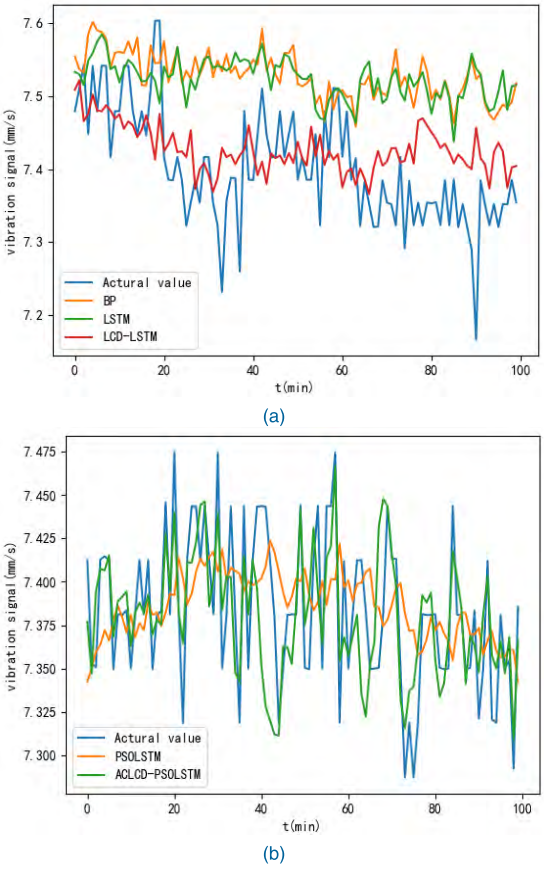


FIGURE 11. The results of the first 100 test samples for the vibration signal series on case 1: (a) the algorithm group #1 and (b) the algorithm group #2.

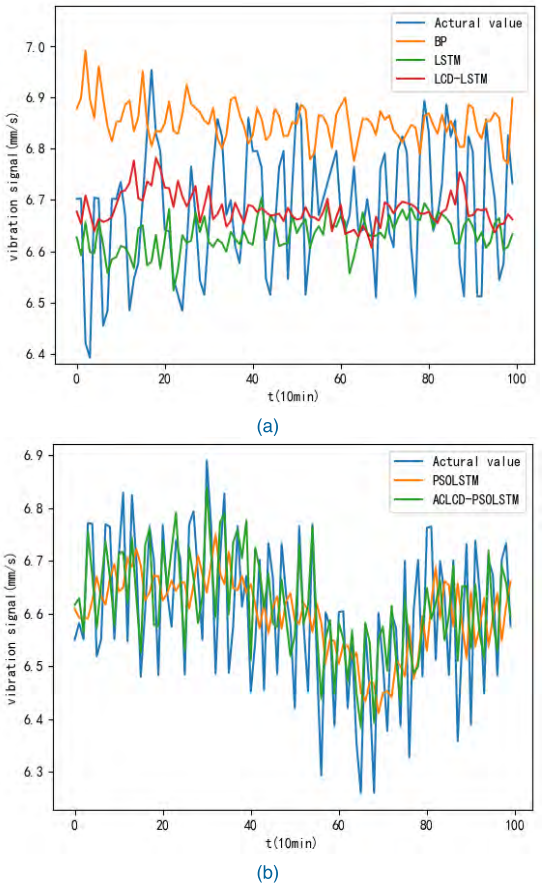


FIGURE 12. The results of the first 100 test samples for the vibration signal series on case 2: (a) the algorithm group #1 and (b) the algorithm group #2.

4) As can be seen from the Table 8, the hybrid LCD-LSTM model has a better performance than the single LSTM model. The LCD is effective in improving the predicting capacity of the LSTM network.
5) According to Table 9, the hybrid ACLCD-PSOLSTM model has a better result than the PSOLSTM model. It shows that the proposed ACLCD method can reduce the complexity of modeling and improve the accuracy of prediction vibration signal.
6) In summary, the ACLCD-PSOLSTM provides significant improvement than single models (i.e., BP, LSTM), hybrid model (i.e., LCD-LSTM) and optimal model (i.e., PSOLSTM).

3) CASE STUDY 2: LONG TERM VIBRATION SIGNAL PREDICTING

In this case study, the ten-min interval dataset is employed to test the performance of ACLCD-PSOLSTM model. As in the case 1, four other prediction models were used for comparison.

The parameters of the model are selected as follows. For BP model, the number of neurons in hidden layer, learn rate and the maximum number of iterations are set as 50, 0.1 and 100, respectively. For LSTM model, the number of neurons in hidden layer and the maximum number of iterations are set as 40 and 100. For ACLCD model, SD $= 0.3$, $\theta = 0.07$, $a = 0.5$, $\varepsilon = 0.7$, $\gamma = 1.2$, $\alpha = 0.7$ and $\beta = 0.3$. For PSO algorithm, $\omega_1 = 0.9$, $N = 20$, $T = 100$, $c_1 = 1.49$, $c_2 = 1.49$ and $\omega = 0.5$. The final optimization results of the PSO algorithm on the ten-minute interval dataset are shown in Appendix A. For clearer display, the main parameters of five different models are shown in Table 4.

TABLE 10. Main parameters of five different models on case 2.

Models	Parameters
BP	40 time lags, 50 hidden layer neurons
LSTM	40 time lags, 40 hidden layer neurons
PSOLSTM	34 time lags, 57 hidden layer neurons
LCD-LSTM	40 time lags, 40 hidden layer neurons
ACLCD-PSOLSTM	34 time lags, 57 hidden layer neurons

The experiment results obtained by five different models are presented in Table 11. In order to show details clearly,

TABLE 11. The error evaluation results of different models for the vibration signal series on case 2.

Models	MAE(mm/s)	RMSE(mm/s)	MAPE(%)
BP	0.2756	0.2288	3.5665
LSTM	0.2030	0.1538	2.3989
PSOLSTM	0.1267	0.1035	1.5646
LCD-LSTM	0.1575	0.1271	1.8183
ACLCD-PSOLSTM	**0.0821**	**0.0646**	**0.9793**

TABLE 12. The final optimization results of PSO algorithm based onv case 1.

Population	Number of neurons	Number of time lags	Value of the fitness function
1	51	34	1.2966
2	52	30	1.2605
3	52	30	1.2904
4	51	30	1.2643
5	52	26	1.2967
6	52	28	1.2831
7	52	30	1.2990
8	52	36	1.2783
9	**52**	**28**	**1.2591**
10	52	30	1.2737
11	51	30	1.2684
12	51	30	1.3176
13	52	30	1.3014
14	52	28	1.2978
15	51	31	1.2753
16	51	28	1.2919
17	52	30	1.2755
18	51	28	1.2663
19	52	34	1.2941
20	51	30	1.2684

we only compare the first 100 test samples. The results of all test samples are shown in Appendix A. The vibration signal predicting results of the models above are shown in Fig. 12.

From Table 11 and Fig. 12, the following conclusions can be drawn:

1) The proposed ACLCD-PSOLSTM performs better than four other models. This shows that the proposed hybrid model has a strong advantage in predicting complex vibration signals.
2) PSOLSTM model has better performance than LSTM model. It shows that PSO algorithm can improve the accuracy of prediction model. The hybrid LCD-LSTM model has better performance than the single LSTM model. The LCD is effective in improving the predicting capacity of the LSTM network.
3) As is shown in Fig. 12, the proposed model has better fitting ability for vibration signal than other models.

In summary, from the above two cases, it can be seen that ACLCD-PSOLSTM model has smaller errors in short-term and long-term vibration signal prediction than other models, which indicates that the proposed model is effective in improving the prediction accuracy of vibration signal.

IV. CONCLUSION

In this paper, the ACLCD-PSOLSTM model has been proposed to obtain more accurate vibration signal prediction

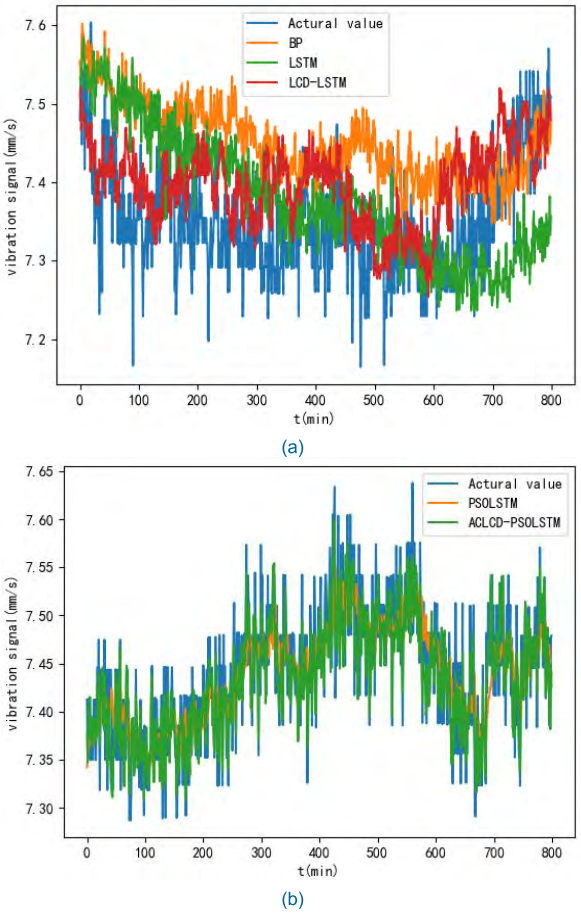


FIGURE 13. The results of all the test samples for the vibration signal series on case 1: (a) the algorithm group #1 and (b) the algorithm group #2.

TABLE 13. The final optimization results of PSO algorithm based on case 2.

Population	Number of neurons	Number of time lags	Value of the fitness function
1	56	30	2.1098
2	56	38	1.9760
3	55	35	2.0342
4	55	36	2.0774
5	**57**	**34**	**1.9261**
6	57	38	2.1475
7	57	38	2.0713
8	57	34	2.0303
9	56	39	1.9427
10	56	39	2.1440
11	55	39	1.9588
12	56	33	2.0201
13	56	31	2.0257
14	57	35	2.0614
15	57	34	2.0122
16	56	38	2.0870
17	57	38	2.1827
18	56	38	1.9338
19	55	36	2.0472
20	56	40	2.0127

results by using ACLCD model, LSTM network and PSO algorithm. The main conclusions of this paper can be summarized as follows: (1) To solve the end effect in LCD, introducing the autocorrelation of signals into RBF continuation

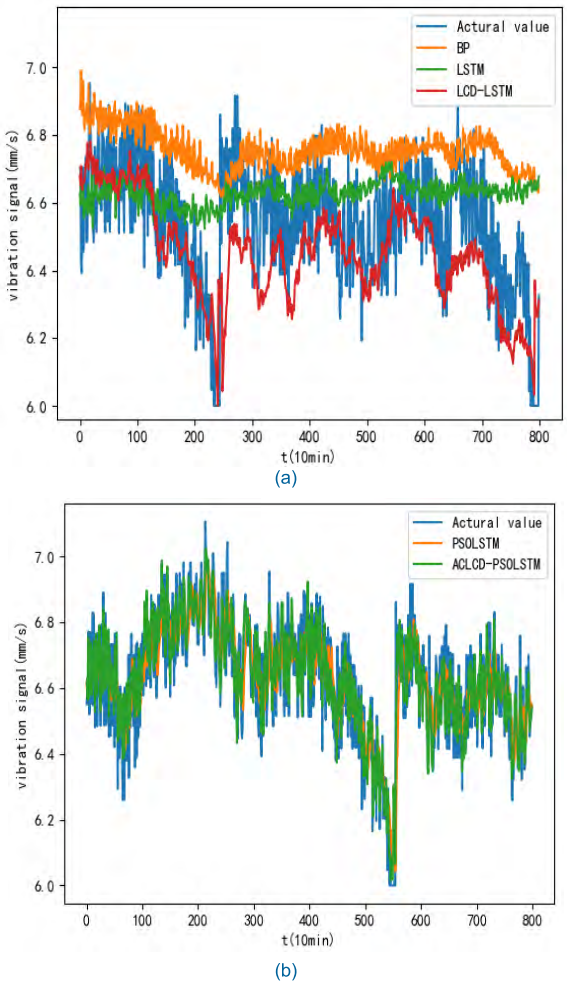


FIGURE 14. The results of all the test samples for the vibration signal series on case 2: (a) the algorithm group #1 and (b) the algorithm group #2.

algorithm, an ACLCD decomposition method is proposed. (2) As there is no scientific guidance to choose the number of neurons in hidden layer and the number of time lags of LSTM network. PSOLSTM model is proposed to search the optimal number of hidden layer neurons and time lags. (3) To reduce the complexity of modeling, the complex vibration signal is decomposed into some ISCs and a residue item with ACLCD model. Then the prediction model is established for each component with PSOLSTM model. Hence, a novel hybrid model aiming at enhancing prediction performance of vibration signals called ACLCD-PSOLSTM is presented in this paper. (4) The validity of ACLCD-PSOLSTM hybrid model is to be validated on two experiments. The first experiment is to compare ACLCD with conventional decomposition methods including LCD and EMD. The results show that ACLCD method has obvious advantages in both computing speed and restraining end effect compared with LCD and EMD. Another experiment is to test the vibration signal predicting capacity of the proposed hybrid model. Three metrics (MAE, RMSE and MAPE) are selected to evaluate the performance of the BP, LSTM, PSOLSTM, LCD-LSTM and ACLCD-PSOLSTM models. The results show that the proposed ACLCD-PSOLSTM performs better than four other models in predicting vibration signals.

The proposed model needs a considerable number of samples for training to ensure the accuracy, so the model needs big data as the basis. For small sample data, the model has no obvious advantage. One-step ahead vibration signal predicting has been studied in this paper. How to accurately predict long-term vibration signals is the focus of the next research. Moreover, how to accurately judge the running state of the compressor through the predicted vibration signal, so as to reduce the occurrence of accidents and improve the production efficiency of the enterprise is the focus of future research.

APPENDIX A

See Figures 13 and 14 and Tables 12 and 13.

REFERENCES

[1] Z. Liang, S. Li, J. Tian, L. Zhang, C. Feng, and L. Zhang, "Vibration cause analysis and elimination of reciprocating compressor inlet pipelines," *Eng. Failure Anal.*, vol. 48, pp. 272–282, Feb. 2015.

[2] L. Zhang, R. Zhu, Y. You, H. You, and Z. Wang, "Research status and Prospect of monitoring and diagnosis technology for reciprocating compressor," *Chem. Prog.*, vol. 10, pp. 1099–1102, 2004.

[3] K. Pichler, E. Lughofer, M. Pichler, T. Buchegger, E. P. Klement, and M. Huschenbett, "Fault detection in reciprocating compressor valves under varying load conditions," *Mech. Syst. Signal Process.*, vols. 70–71, pp. 104–119, Mar. 2016.

[4] X.-H. Zhao and X. Chen, "Auto regressive and ensemble empirical mode decomposition hybrid model for annual runoff forecasting," *Water. Resour. Manage.*, vol. 29, no. 8, pp. 2913–2926, 2015.

[5] C. Ordóñez, F. S. Lasheras, J. Roca-Pardiñas, and F. J. de Cos Juez, "A hybrid ARIMA–SVM model for the study of the remaining useful life of aircraft engines," *J. Comput. Appl. Math.*, vol. 346, pp. 184–191, Jan. 2019.

[6] S. Abbas and M. Ilyas, "Assessing the impact of EI Niño southern oscillation index and Land Surface Temperature fluctuations on dengue fever outbreaks using ARIMAX(p)-PARX(p)-NBARX(p) models," *Arabian J. Geosci.*, vol. 11, no. 24, p. 777, 2018.

[7] T. Xie, G. Zhang, H. Liu, F. Liu, and P. Du, "A hybrid forecasting method for solar output power based on variational mode decomposition, deep belief networks and auto-regressive moving average," *Appl. Sci.*, vol. 8, no. 10, p. 1901, 2018.

[8] C. L. Giles, S. Lawrence, and A. C. Tsoi, "Noisy time series prediction using recurrent neural networks and grammatical inference," *Mach. Learn.*, vol. 44, nos. 1–2, pp. 161–183, 2001.

[9] R. C. Deo and M. Şahin, "An extreme learning machine model for the simulation of monthly mean streamflow water level in eastern queensland," *Environ. Monit. Assessment*, vol. 188, no. 4, p. 90, Feb. 2016.

[10] G. Jiao, T. Guo, and Y. Ding, "A new hybrid forecasting approach applied to hydrological data: A case study on precipitation in northwestern China," *Water*, vol. 8, no. 9, p. 367, 2016.

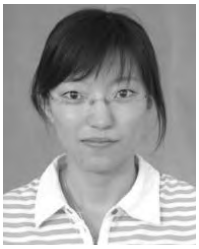
[11] P. T. Nastos, A. G. Paliatsos, K. V. Koukouletsos, I. K. Larissi, and K. P. Moustris, "Artificial neural networks modeling for forecasting the maximum daily total precipitation at Athens, Greece," *Atmos. Res.*, vol. 144, pp. 141–150, Jul. 2014.

[12] V. A. Maksimenko *et al.*, "Artificial neural network classification of motor-related EEG: An increase in classification accuracy by reducing signal complexity," *Complexity*, vol. 2018, Aug. 2018, Art. no. 9385947.

[13] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," *Neural Comput.*, vol. 1, pp. 270–280, Jun. 1989.

[14] B. Chen *et al.*, "Embedding logic rules into recurrent neural networks," *IEEE Access*, vol. 7, pp. 14938–14946, 2019.

[15] A. Cocos, A. G. Fiks, and A. J. Masino, "Deep learning for pharmacovigilance: Recurrent neural network architectures for labeling adverse drug reactions in Twitter posts," *J. Am. Med. Inform. Assoc.*, vol. 24, no. 4, pp. 813–821, Jul. 2017.
[16] M. Morchid *et al.*, "Parsimonious memory unit for recurrent neural networks with application to natural language processing," *Neurocomputing*, vol. 314, pp. 48–64, Nov. 2018.
[17] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.
[18] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
[19] Y. Li, H. Wu, and H. Liu, "Multi-step wind speed forecasting using EWT decomposition, LSTM principal computing, RELM subordinate computing and IEWT reconstruction," *Energy Convers. Manage.*, vol. 167, pp. 203–219, Jul. 2018.
[20] Y.-L. Hu and L. Chen, "A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm," *Energy Convers. Manage.*, vol. 173, pp. 123–142, Oct. 2018.
[21] J. Lin, L. Su, Y. Yan, G. Sheng, D. Xie, and X. Jiang, "Prediction method for power transformer running state based on LSTM_DBN network," *Energies*, vol. 11, no. 7, p. 1880, 2018.
[22] G. Gregory and J.-L. Gauvain, "Minimum word error training of RNN-based voice activity detection," in *Proc. INTERSPEECH*, 2015, pp. 2650–2654.
[23] H. Zhao, S. Sun, and B. Jin, "Sequential fault diagnosis based on LSTM neural network," *IEEE Access*, vol. 6, pp. 12929–12939, 2018.
[24] A. ElSaid, F. El Jamiy, J. Higgins, B. Wild, and T. Desell, "Optimizing long short-term memory recurrent neural networks using ant colony optimization to predict turbine engine vibration," *Appl. Soft Comput.*, vol. 73, pp. 969–991, Dec. 2018.
[25] B. Jatin and D. Toshniwal, "Empirical mode decomposition based deep learning for electricity demand forecasting," *IEEE Access*, vol. 6, pp. 49144–49156, 2018.
[26] Y. Dong, X. Ma, C. Ma, and J. Wang, "Research and application of a hybrid forecasting model based on data decomposition for electrical load forecasting," *Energies*, vol. 9, no. 12, p. 1050, 2016.
[27] B. Huang and A. Kunoth, "An optimization based empirical mode decomposition scheme," *J. Comput. Appl. Math.*, vol. 240, pp. 174–183, Mar. 2013.
[28] Y. Yu, YuDejie, and C. Junsheng, "A roller bearing fault diagnosis method based on EMD energy entropy and ANN," *J. Sound Vib.*, vol. 294, pp. 269–277, Jun. 2006.
[29] R. Sharma, R. Pachori, and U. Acharya, "Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals," *Entropy*, vol. 17, no. 2, pp. 669–691, 2015.
[30] J. Zheng, J. Cheng, and Y. Yang, "A rolling bearing fault diagnosis approach based on LCD and fuzzy entropy," *Mechanism Mach. Theory*, vol. 70, pp. 441–453, Dec. 2013.
[31] H. Liu, X. Wang, and C. Lu, "Rolling bearing fault diagnosis based on LCD–TEO and multifractal detrended fluctuation analysis," *Mech. Syst. Signal Process.*, vols. 60–61, pp. 273–288, Aug. 2015.
[32] Y. Xiao, N. Kang, Y. Hong, and G. Zhang, "Misalignment fault diagnosis of DFWT based on IEMD energy entropy and PSO-SVM," *Entropy*, vol. 19, no. 1, p. 6, 2017.
[33] M. Zhang, T. Wang, T. Tang, M. Benbouzid, and D. Diallo, "An imbalance fault detection method based on data normalization and EMD for marine current turbines," *ISA Trans.*, vol. 68, pp. 302–312, May 2017.
[34] J. Yang, P. Li, Y. Yang, and D. Xu, "An improved EMD method for modal identification and a combined static-dynamic method for damage detection," *J. Sound Vib.*, vol. 420, pp. 242–260, Apr. 2018
[35] H. Li, Y. Hu, F. Li, and G. Meng, "Succinct and fast empirical mode decomposition," *Mech. Syst. Signal Process.*, vol. 85, pp. 879–895, Feb. 2017.
[36] H. Wang and Y. Ji, "A revised Hilbert–Huang transform and its application to fault diagnosis in a rotor system," *Sensors*, vol. 18, no. 12, p. 4329, 2018.
[37] S.-X. Lun, X.-S. Yao, and H.-F. Hu, "A new echo state network with variable memory length," *Inf. Sci.*, vols. 370–371, pp. 103–119, Nov. 2016.



HUI-XIN TIAN received the B.S. degree in automation from the Fushun Petroleum College, Liaoning, China, in 2001, and the M.S. and Ph.D. degrees in control theory and control engineering from Northeastern University, Shenyang, China, in 2005 and 2009, respectively.

She is currently a Professor with Tianjin Polytechnic University. Her main research interest includes modeling, control, and optimization in complex industrial systems.

DAI-XU REN received the B.S. degree in automation from the Xi'an University of Technology, Shaanxi, China, in 2017. He is currently pursuing the M.S. degree with Tianjin Polytechnic University, Tianjin, China.

His main research interest includes modeling optimization in complex industrial systems.

KUN LI received the B.S. degree in mining engineering from the Shandong University of Science and Technology, Shandong, China, in 2002, and the Ph.D. degree in systems engineering from Northeastern University, Shenyang, in 2010.

He is currently an Associate Professor with Tianjin Polytechnic University. He has published over ten papers in international journals. His current research interests include multiobjective optimization, production scheduling, and modeling and optimization in process industries based on process operation optimization.