

Received June 10, 2017, accepted July 1, 2017, date of publication July 11, 2017, date of current version July 31, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2723610

Elman Neural Network Soft-Sensor Model of Conversion Velocity in Polymerization Process Optimized by Chaos Whale Optimization Algorithm

W. Z. SUN¹, (Member, IEEE), AND J. S. WANG^{1,2}, (Member, IEEE)

¹School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan 114051, China

²National Financial Security and System Equipment Engineering Research Center, University of Science and Technology Liaoning, Anshan 114051, China

Corresponding author: J. S. Wang (wang_jiesheng@126.com)

This work was supported in part by the Project by National Natural Science Foundation of China under Grant 21576127, in part by the Program for Research Special Foundation of the University of Science and Technology of Liaoning under Grant 2015TD04, and in part by the Opening Project of National Financial Security and System Equipment Engineering Research Center under Grant USTLKFJ201502.

ABSTRACT According to the requirement of real-time monitoring of the conversion rate of vinyl chloride in the production process of polyvinyl chloride polymerization and the nonlinearity of the industrial data, the Elman neural network with strong nonlinear performance is chosen to build the soft-sensor model. However, because of the early stage of the Elman neural network to train the connection weights between the layers, the training effect is difficult to guarantee with the connection weights. So the whale optimization algorithm (WOA) is adopted to optimize the Elman neural network, to avoid it falling into the local optimum. At the same time, to solve the problem that the position of the search agent is randomly distributed in the initialization process of the WOA algorithm, and to introduce the idea of chaos, a chaos WOA (CWOA) based on the idea of chaos is proposed to improve the diversity of all search agents and egocentricity of agent search by utilizing the chaotic features. At the end of this paper, considering that the input vector dimension is too large, the neural network topology is very large, which will lead to the complexity of the training process. Therefore, the locally linear embedding method is introduced to reduce the dimension of high-dimensional input vectors. The simulation results show that the chaotic whale algorithm can significantly improve the prediction accuracy of economic and technical indexes of PVC polymerization process, which especially has a significant improvement in the prediction effect in the early stage and meets the requirements of real-time control of the production process of the polymerization reactor.

INDEX TERMS PVC polymerization process, Elman neural network, Soft sensor, Whale optimization algorithm, Locally linear embedding method.

I. INTRODUCTION

Polyvinyl chloride (PVC) resin is one of the plastic varieties which is the first industrialization realization of the plastics in the world and is one of the most widely used polymers [1]. By using the vinyl chloride monomer (VCM) as raw materials, the production of PVC resin by suspension polymerization method is a typical intermittent chemical production process. The different VCM conversion has a certain impact on the molecular weight of PVC resin, thermal stability, porosity, the residues of VCM, the absorption of plasticizers and processing liquidity [2]. With the development

of the large-scale polymerization and in-depth study on the polymerization process of vinyl chloride, the conversion velocity in single tank is further improved in order to increase production capacity and reduce production cost. A large number of uncertain information, diversification of massive data led to the use of traditional control methods and mathematical models for local control often cannot get satisfactory results. When recycling the VCM, a low conversion rate will obtain the resin with higher porosity. That is to say a high conversion rate will get the resin with lower porosity. However, the absorption rate of the plasticizer decreases with the increase

of the VCM conversion rate. Due to the restrictions by on-site conditions and the lack of mature measurement devices, it is hard to measure the real-time VCM conversion rate in the actual production process. Therefore, it is difficult to achieve the direct closed-loop control.

Elman NN is a typical local recursive neural network with time delay feedback, which introduces the feedback signal and stores the internal states to realize the dynamic mapping function based on the BP neural network [3]. But because most of the industrial production process is dynamic, the use of the static neural network, such as BP neural network or RBF neural network, to be used to build the neural network model, often cannot obtain the desired predictive results. For example, it is necessary to prioritize the nonlinear autoregressive model of the hypothetical system. It is necessary to determine the order of the model. When the system order is unknown or too high, the network structure will become complicated and affect the convergence speed of the network. The increase in the input node will also make the system more sensitive to noise. The Elman NN model was used for the transient estimation and a good prediction result was obtained [4]. The genetic algorithm (GA) was used to optimize the weights of Elman NN to realize the detection of the motor fault [5]. The pressure of the emulsion pump was predicted by using the Elman NN [6]. Through the design of the self-feedback connection with fixed gains and the feedback of nodes in the output layer, the improved Elman NN has faster convergence speed and generalization ability [7]. Elman NN was integrated with the quantum computation to improve the precision based on the approximation and information processing ability of quantum computation [8]. The whale optimization algorithm (WOA) is a global swarm intelligent optimization algorithm inspired by the hunting behavior of the Humpback whales by Mirjalili and Lewis [9], which is applied in production optimization problems [10] and the sizing optimization of skeletal structures [11]. The Elman NN soft-sensor model optimized by chaos whale optimization algorithm is proposed to predict the VCM conversion velocity in polymerization process. Simulation results proved the effectiveness of the proposed strategy.

The paper is organized as follows. In section II, the technique flowchart of the PVC polymerization process is introduced. The Elman NN is described in section III. In section IV, the Elman NN soft sensor model based on CWOA is introduced in details. The simulation experiments and results analysis are discussed in section V. The conclusion illustrates the last part.

II. TECHNIQUE FLOWCHART OF PVC POLYMERIZATION PROCESS

PVC production process includes polymerization, stripping, drying and packaging process. The techniques of the PVC polymerization production process has four kinds: suspension method, bulk method, emulsion method and micro suspension polymerization [2]. The suspension method accounts for 80%, the bulk method accounts for 10%, and the emulsion

and micro suspension polymerization accounts for 10%. Most of the countries adopts the suspension method. This method makes the liquid VCM under its saturated pressure stirring and dispersing. The dispersing agent is adopted to make VCM liquid beads under protection and dispersion. Then the oil soluble initiator is adopted to produce the polymerization reaction in monomer phase. The advantages of suspension polymerization are described as:

- 1) The system has low viscosity, easy control of heat transfer and temperature and the molecular weight and its distribution are stable.
- 2) The molecular weight of the product is higher than that of solution polymerization;
- 3) The post-treatment process is simpler than the emulsion polymerization and solution polymerization, and the production cost is low.

The disadvantage of suspension polymerization is that the product has a small amount of desparation residues. So in order to produce transparent and excellent insulation products, these residues need to be removed.

In this paper, the PVC production process is a typical intermittent chemical process to produce SG5 resin products, whose technique is very complex, flammable, explosive and easy to poison. In order to ensure the safe production and stable product quality, the higher requirement is needed on the degree of automation control system. The main raw material used in the PVC production process is VCM and PVC is produced by VCM polymerization, whose technique flowchart of polymerization kettle production process is shown in Fig. 1.

PVC production process has a certain production cycle, which includes materials feed, polymerization, recycling, packaging and other steps. The first stage is to feed materials, where various materials are fed into the polymerization kettle in sequence according to a certain proportion. In the feeding stage, the wall coated agent is first fed into the polymerization kettle so as to remove residues and reduce the impurity content in the finished product. Then the quantitative buffer, monomer and water, desparation and initiator are sequentially fed into the kettle. When the monomer and water are added into the polymerization kettle, the stirrer starts to stir and the polymerization reaction begins. The polymerization process is the core of PVC batch production process, is a key step to guarantee the production safety and product quality. In this stage, the process parameters of polymerization reactor (such as temperature and pressure) changes according to a certain time trajectory. When the pressure drop down to 0.05 Mpa and the conversion rate reached about 80-82%, the terminator are added and the reaction is terminated. The slurry is discharged from the bottom of the polymerization kettle. After a series of processing for slurry, such as recovery, stripping, condensation, drying and so on, the finished products are packed. A typical PVC polymerization kettle technological process is shown in Fig. 2 [12].

The PVC polymerizing process is a typical batch production process. In PVC polymerizing process, all kinds of raw

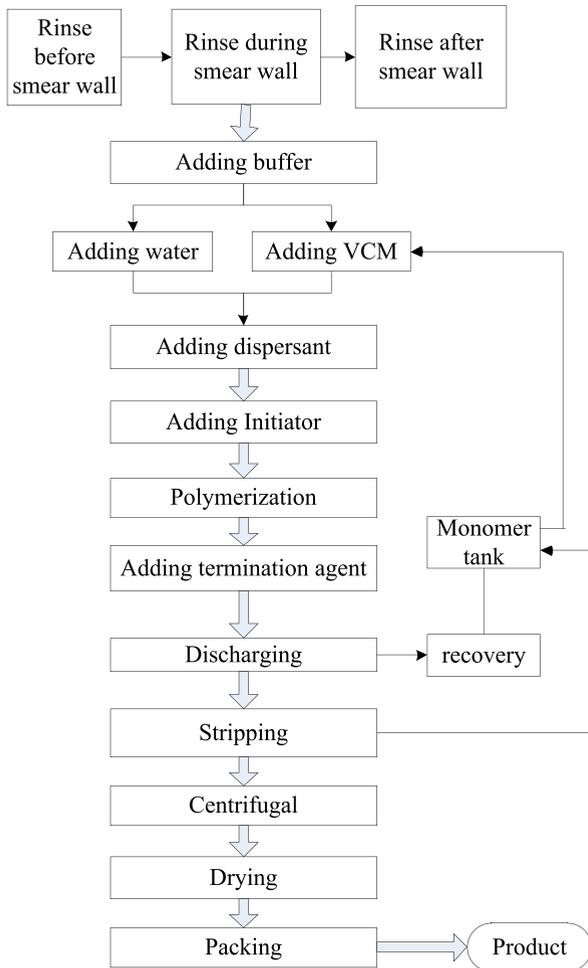


FIGURE 1. Polymerization kettle production process flowchart.

materials and auxiliary agents are placed into the reaction kettle. They are fully and evenly dispersed under the function of stirring. Then, we begin to ventilate the cooling water to the clip set of the reaction kettle and baffle plate constantly in order to remove homo-polymer. When the conversion rate of VCM reaches a certain value, the reactions are terminated and the finished product is got. The polymerization degree decreases with an increase in temperature. Therefore, polymerization degree only has a relationship with temperature for the VCM polymerization reaction. Ultimately, the accuracy of the conversion velocity predictive model has a direct influence on the product quality and type of PVC. According to the characteristics of a polymerization process, 10 process variables related with the conventional rate and velocity of VCM are identified as the secondary variables of the soft-sensor model, which are listed in the Table 1.

III. ELMAN NEURAL NETWORK

A. OVERVIEW OF THE ELMAN NEURAL NETWORK

Elman NN is a kind of typical dynamic neural network, which contains dynamic link, and does not need to use more system state as inputs, so the number of input layer units is reduced. Moreover, the approximation ability is better than

TABLE 1. Process variables of the discussed polymerization.

Process variable	Variable symbol	Unit	Measurement range
Temperature inside kettle	TIC-P101	°C	0-100
Pressure inside kettle	PIC-P102	MPa	0-1.2
Water flow rate of baffle	FIC-P101	m ³ /h	0-500
Water flow rate of clip set	FIC-P102	m ³ /h	0-500
Water feed flow rate	FIC-P104	m ³ /h	0-500
Seal water flow rate	FIC-P105	m ³ /h	0-500
Water inlet temperature of cooling	TI-P107	°C	0-100
Water outlet temperature of clip set	TI-P109	°C	0-100
Water outlet temperature of damper	TI-P110	°C	0-100
Outlet temperature of cold water tank	TIC-WA01	°C	0-100

the general static network, and the convergence speed is faster. For the dynamic system identification on the unknown model order, the Elman network does not appear the problem of network structure expansion. However, it is limited to the identification of the first order system, so it is difficult to be used in the high order system, so an improved Elman network is proposed. Elman NN belongs to a kind of feedback neural network. In order to make the network has the memory function, the feedback is added between the hidden layer and the structure layer [13]. The structure of Elman NN is shown in Fig. 3.

It has a special structural unit besides the input layer, hidden layer and the output layer. Feedback network is an important type of neural network. The addition of the feedback link among layers or within the layers makes it possible to have a delay between the input and the output. So the dynamic equations are used to describe this kind of model. However, the feed-forward neural network can only achieve the nonlinear mapping. But it is also because of this feedback, this network has the memory ability so that it is widely used in a lot of research areas, such as time series prediction, system identification and process control, and so on. In addition to the input layer, the hidden layer and the output layer, it also has a special structural unit. This structural unit is used to remember the output value of the hidden layer unit at the previous moment, which can be considered as a delay operator. So the weights of the feed-forward connection can be revised, while the recursive part is fixed and the corresponding weights cannot be revised [14], [15].

It can be seen from Fig. 3 that the output in structural layer at k moment is equal to the output in the hidden layer at $k-1$ moment plus α times of the output in the structural unit

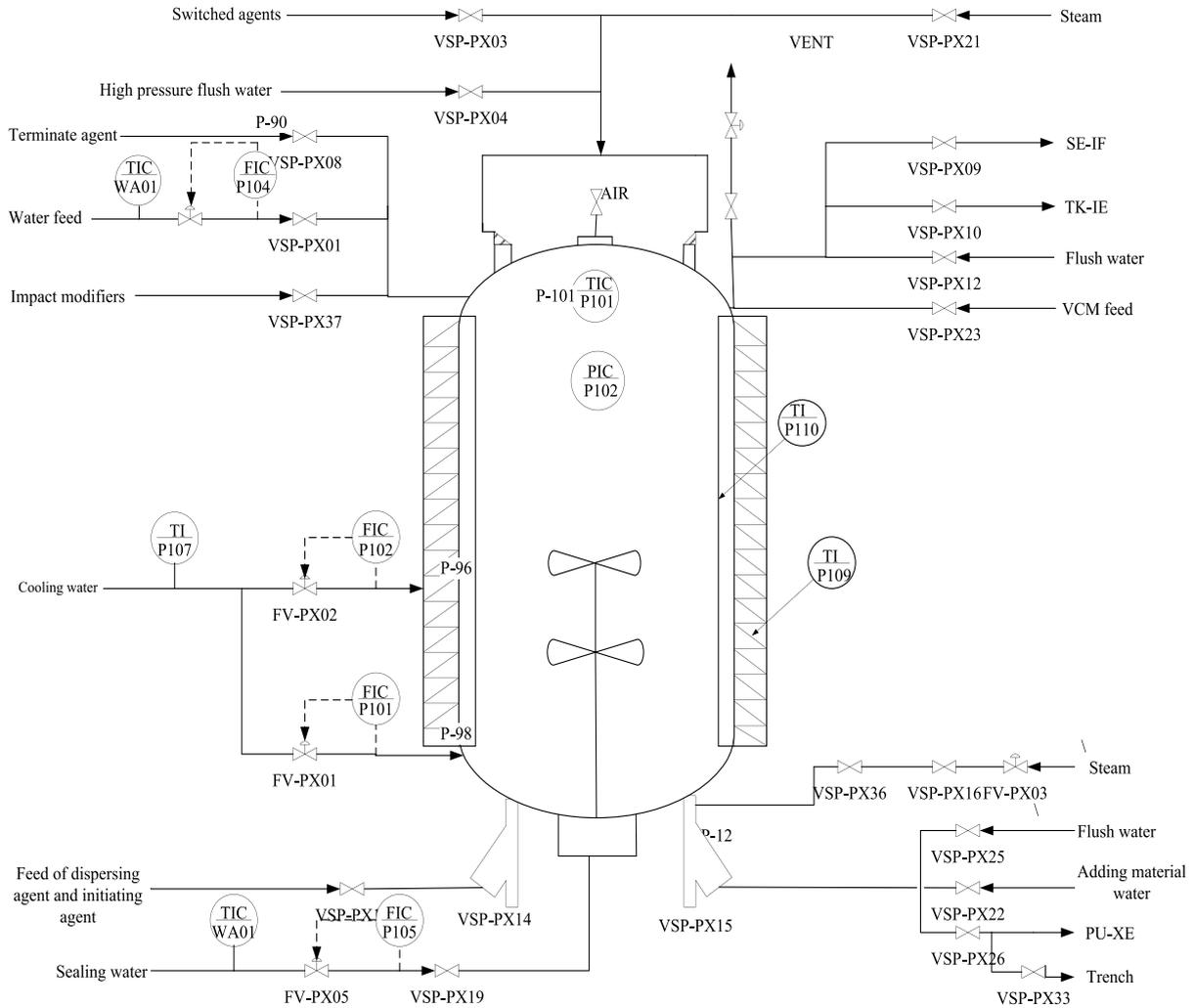


FIGURE 2. Technique flowchart of polymerization kettle.

at $k-1$ moment.

$$x_{c,l}(k) = \alpha \cdot x_{c,l}(k-1) + x_l(k-1) \quad l = 1, 2, 3, \dots, n, \quad (1)$$

where $x_{c,l}$ is the output of the l th structure layer, $x_l(k)$ is the output of the l th hidden layer, α is a self-connected feedback gain factor. So the mathematical model of OSF-Elman NN is shown as follows.

$$x(k) = f(W^1(x_c(k) + y(k)) + W^2u(k-1)) \quad (2)$$

$$x_c(k) = \alpha \cdot (x_c(k-1) + y(k-1)) + x(k-1) \quad (3)$$

$$y_k = g(W^3x(k)), \quad (4)$$

where, is W^1 the weights between structural layer and hidden layer, W^2 is the weights between the input layer and hidden layer, W^3 is the weights between the output layer and hidden layer, $f(x)$ is the Sigmoid activation function of hidden layer neurons, $g(x)$ is a linear activation function of output

layer neurons. Their expressions are listed as follows.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (5)$$

$$y_k = W^3x(k), \quad (6)$$

where, y_k is the output of the network model.

B. LEARNING ALGORITHM OF ELMAN NEURAL NETWORK

The error function E is defined as:

$$E = \frac{1}{2}(y_d(k) - y(k))^T(y_d(k) - y(k)), \quad (7)$$

where, $y_d(k)$ is expected output and $y(k)$ is the expected output. The calculation expression $y(k)$ is fed into (7) to obtain:

$$\frac{\partial E}{\partial w_{ij}^3} = -(y_{d,i}(k) - y(k)) \frac{\partial y_i(k)}{\partial w_{ij}^3} = -(y_{d,i}(k) - y(k)) g'_i(\cdot) x_j(k), \quad (8)$$

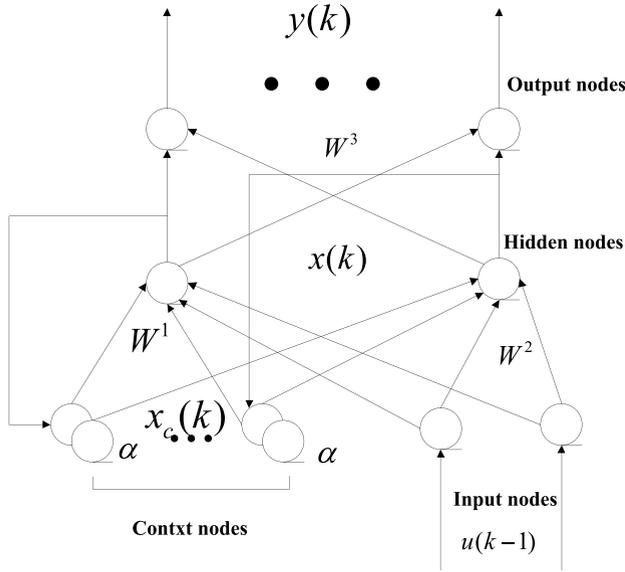


FIGURE 3. Structure of Elman neural network and OSF Elman neural network.

where w_{ij}^3 is the weight between the hidden layer and the output layer, $y_{d,i}$ represents the target values and y_i is the predicted output of the model.

Set $\delta_i^0 = (y_{d,i}(k) - y(k))g'_i(\cdot)$ to obtain:

$$\frac{\partial E}{\partial w_{ij}^3} = -\delta_i^0 x_j(k) \quad i = 1, 2, \dots, m \quad j = 1, 2, \dots, n, \quad (9)$$

where $g'_i(\cdot)$ is the derivative of activation function of the output layer neurons. On the same time, the partial derivative of W^2 is described as follows.

$$\frac{\partial E}{\partial w_{ij}^2} = \frac{\partial E}{\partial x_j(k)} \frac{\partial y_i(k)}{\partial w_{ij}^2} = \sum_{i=1}^m (-\delta_i^0 w_{ij}^3) f'_j(\cdot) u_q(k-1), \quad (10)$$

where w_{ij}^2 is the weight between the input layer and the hidden layer, $u_q(k-1)$ is the input values of the input layer at previous moment and $f'_j(\cdot)$ is the derivative of activation function of the hidden layer.

Alike, set $\delta_j^h = \sum (\delta_i^0 w_{ij}^3) f'_j(\cdot)$ to obtain:

$$\frac{\partial E}{\partial w_{ij}^2} = -\delta_j^h u_q(k-1) \quad (j = 1, 2, \dots, n; q = 1, 2, \dots, r). \quad (11)$$

At last, the connection weight W^1 from the following layer to the hidden layer is carried out the partial derivation to obtain:

$$\frac{\partial E}{\partial w_{jl}^1} = \sum_{i=1}^m (\delta_i^0 w_{ij}^3) \frac{\partial x_j(k)}{\partial w_{jl}^1} \quad (j = 1, 2, \dots, n; l = 1, 2, \dots, n) \quad (12)$$

$$\begin{aligned} & \frac{\partial x_j(k)}{\partial w_{jl}^1} \\ &= \frac{\partial}{\partial w_{jl}^1} \left[f_j \left(\sum_{i=1}^h w_{jl}^1 (x_{c,i}(k) + y_i(k-1)) + \sum_{i=1}^r w_{jl}^2 u_i(k-1) \right) \right] \\ &= f'_j(\cdot) \left[x_{c,j}(k) + y_i(k-1) + \sum w_{jl}^1 \frac{\partial x_{c,j}(k)}{\partial w_{jl}^1} \right], \quad (13) \end{aligned}$$

where w_{jl}^1 is the weight between the input layer and the hidden layer.

It can be seen from the structure diagram that the basic relationship between $x_c(k)$ and w_{jl}^1 can be ignored. So when the dependence of $x_c(k)$ on the connection weight w_{jl}^1 is not considered, the following equation can be obtained.

$$\frac{\partial x_j(k)}{\partial w_{jl}^1} = f'_j(\cdot) (x_{c,l}(k-1) + y_l(k-1)), \quad (14)$$

where $x_{c,j}(k)$ is the output of the j th structural unit and $x_j(k)$ is the output of the j th hidden layer unit.

$$f'_j(\cdot) x_{c,j}(k) = f'_j(\cdot) (x_c(k-1) + y_l(k-1)) + \alpha f'_j(\cdot) x_{c,j}(k). \quad (15)$$

Eq. (15) is fed into (14) to obtain:

$$\frac{\partial x_j(k)}{\partial w_{jl}^1} = f'_j(\cdot) (x_c(k-1) + y_l(k-1)) + \alpha \frac{\partial x_j(k-1)}{\partial w_{jl}^1}. \quad (16)$$

Based on $\Delta W = -\eta \frac{\partial E}{\partial w}$, the learning algorithm of Elman NN can be described as:

$$\Delta w_{ij}^3 = \eta_1 \delta_i^0 x_j(k) \quad i = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (17)$$

$$\Delta w_{jq}^2 = \eta_2 \delta_j^h u_q(k-1) \quad q = 1, 2, \dots, r; j = 1, 2, \dots, n \quad (18)$$

$$\Delta w_{ij}^1 = \eta_3 \sum_{i=1}^m (\delta_i^0 w_{ij}^3) \frac{\partial x_j(k)}{\partial w_{jl}^1} \quad l = 1, 2, \dots, m; j = 1, 2, \dots, n \quad (19)$$

where η_1, η_2 and η_3 are the learning steps of W^1, W^2 and W^3 , respectively.

$$\delta_i^0 = (y_{d,i}(k) - y(k)) g'_i(\cdot) \quad (20)$$

$$g_j^h = \sum (\delta_i^0 w_{ij}^3) f'_i(\cdot) \quad (21)$$

$\partial x_j(k) / \partial w_{jl}^1$ can be deduced from (16). The flowchart of the learning algorithm of the Elman NN is shown in Fig. 4.

IV. ELMAN NEURAL NETWORK SOFT-SENSOR MODEL BASED ON CWOA

A. STRUCTURE OF SOFT-SENSOR MODEL

The discussed process variables are the input and the VCM conversion velocity is the output. The nonlinear relationship $f\{\cdot\}$ between input and output variables are fitted by using Elman neural network so as to establish the soft-sensor model of the VCM conversion velocity. The structure of the proposed soft-sensor model is shown in Fig. 5.

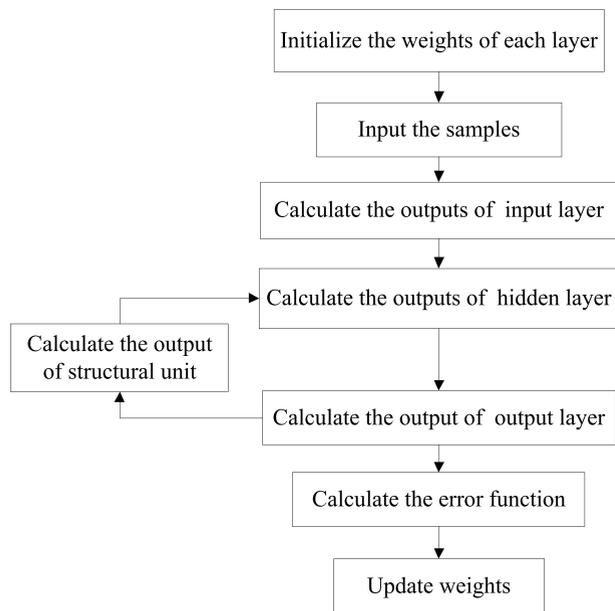


FIGURE 4. Flowchart of learning algorithm of the Elman neural network.

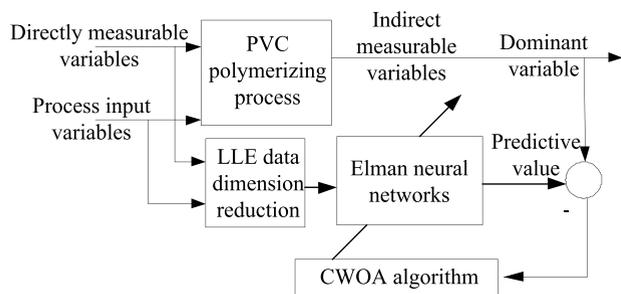


FIGURE 5. Structure of the proposed soft-sensor model.

B. DATA DIMENSION DEDUCTION BASED ON LLE METHOD

The data with high dimension will make the sample size required by the whole data analysis reach the level that the system is far from bearing. However, there will be some redundant data in the data with high dimension, which may influence the performance of many real-time algorithms. In order to solve this problem, the idea of dimension reduction is proposed. There are three main types of dimension reduction methods: linear dimension reduction, nonlinear dimension reduction and manifold learning dimension reduction. The linear dimension reduction algorithm is easy to compute and understand. But in many cases the linear exists only in the theoretical assumptions and the data structural characteristics in real life are often nonlinear. This leads to the emergence of nonlinear dimension reduction algorithms.

The main principle of nonlinear dimension reduction method based on kernel idea is to map the nonlinear structure data in the low dimensional space to the high-dimensional space by using the kernel technique. It is hoped that the linear non-separable data in the low dimensional space can

be represented as a linear structure in the high-dimensional space. Its limitations are:

- 1) The problem of how to selected kernel function.
- 2) The algorithm is derived to focus more on the formula, so people can not recognize specific expression of the original data in the kernel feature space through observation, which led to the non-linear dimension reduction algorithm based on kernel theory is very obscure.

Manifold learning dimension reduction, simply speaking, is a method of mapping high-dimensional data into low dimensional space. Its main principle is to observe the structural characteristics of the data in the high dimensional space, and to establish a one-to-one mapping relationship from the high dimension to the low dimensional space by the mathematical model. The original data obtained in the low dimensional space can retain the information of the original high dimensional space and reduce the data dimension to the maximum extent. Locally linear embedding (LLE) algorithm is a kind of manifold learning dimension reduction methods, which mainly calculates the samples with enough dense. Under the condition that the relative position between each sample point and the adjacent point does not change, according to the method that the local topology is unchanged to solve the low dimensional coordinates, the data dimension reduction is realized. This algorithm is to calculate the local linear combination, so it is equivalent to cut the data into a smaller areas from a high dimensional manifold. Then these areas in the low dimensional space according to the characteristics of the invariant relative position are integrated into a complete manifold in order to achieve the dimension reduction.

According to the characteristics of the sample data and the characteristics of LLE, it is concluded that the choice of LLE can reduce the dimension of the data to achieve the best effect. The reasons are described as follows:

- 1) The data manifold used in this simulation is not completely closed.
- 2) The sampling is not difficult so that the sample points with very high dense can be obtained.
- 3) Because the chemical reaction process can be considered that the change of the data is linear in a short period of time, it complies with the characteristics of LLE reduction method.

For soft-sensor model by using neural networks, that the input vector dimension is too large will make the network topology structure and training become very complex. So the local linear embedding (LLE) algorithm, one of the nonlinear dimension reduction method based on manifold learning, is adopted to reduce the dimension of high-dimensional data [18], [19]. Its basic idea is to assume that each data point and its neighboring point lie in a linear or nearly linear region of the manifold. The global nonlinear was transformed into the local linear and the overlapping local neighborhoods can provide information about the global structure. The specific procedure is described as follows.

(1) Select local neighbor. For a given data set $X = \{x_1, x_2, \dots, x_N\} (x_i \in R^D)$, D is the euclidean distance matrix, $D = \{d_{ij}\}_{n \times n}$. the distance between a sample point x_i in a high dimensional space and its $k (k < N)$ th nearest neighbor point is calculated by the following equation.

$$d_{ij} = \left[\sum_{k=1}^D |x_{ik} - x_{jk}|^2 \right]^{\frac{1}{2}} \quad (22)$$

(2) Calculate the local reconstructed weight matrix of the sample points. Define the error function as:

$$\min \varepsilon(W) = \sum_{i=1}^N \left\| x_i - \sum_{j=1}^N w_{ij} x_j \right\|_2^2 \quad (23)$$

where $x_{ij} (j = 1, 2, \dots, k)$ is the j th nearest neighbor of x_i (There is no x_{ij} in this formula, x_{ij} just to illustrate the meaning of w_{ij}), w_{ij} is a weight between x_i and x_{ij} . In order to ensure that the reconstruction is independent of the translation, rotation and scaling of coordinates, $\sum_{j=1}^N w_{ij} x_j = 1$. Combining the restrictive conditions, Eq. (23) can be rewritten as:

$$\begin{aligned} \min \varepsilon(W) &= \sum_{i=1}^N \left\| x_i - \sum_{j=1}^N w_{ij} x_j \right\|_2^2 \\ &= \sum_{i=1}^N \left\| (x_i - x_{ij}) \right\|_2^2 = \sum_{i=1}^N (w_i)^T Z_i w_i, \end{aligned} \quad (24)$$

where $Z_i = (x_i - x_{ij})^T (x_i - x_{ij})$ is the local covariance matrix of the i th sample location and $w_i = \{w_{i1}, w_{i2}, \dots, w_{ik}\}^T$ is the local reconstruction weights of the i th sample location. The Lagrange-multiplier is introduced to solve this constraint problem.

$$\begin{aligned} L(W) &= \sum_{i=1}^N (w_i)^T Z_i w_i + \lambda \left(\sum_{j=1}^k w_{ij} - 1 \right) \\ \Rightarrow \frac{\partial L}{\partial w_{ij}} &= 2Z_i w_i + \lambda \times 1 \Rightarrow Z_i w_i. \end{aligned} \quad (25)$$

Usually a simple solution is used. Set $Z_i w_i = 1$ and then re-adjust the weights in order to make the corresponding sum equal to 1 and obtain w_i .

(3) The weight matrix W is used to find the low dimension embedding Y of the sample set.

Minimizing the reconstruction error sum function $\min \Phi(Y) = \sum_{i=1}^N \left\| x_i - \sum_{j=1}^N w_{ij} x_j \right\|_2^2$ can achieve this purpose. In order to fix Y and avoid the data set collapsing to the origin of coordinates, we can give Y a limit: $\sum_{i=1}^N y_i = 0, \frac{1}{N} \sum_{i=1}^N y_i y_i^T = I$. I is N dimensional unit matrix [20].

Then the corresponding optimization problem can be transferred into the following constraint optimization problem.

$$\begin{cases} \min \Phi(Y) = \sum_{i=1}^N \|YI_i - YW_i\|_2^2 \\ = \sum_{i=1}^N \|YI_i - YW_i\|_2^2 = \min tr(YMY^T) \\ s.t \ YY^T \end{cases} \quad (26)$$

where $M = (I - W)^T (I - W)$, $tr(YMY^T)$ is the trace of matrix YMY^T . By using the Lagrange multiplier method, obtain $MY^T = \lambda Y^T$. Then the eigenvalues of M are the embedded coordinates.

The feature vectors of the d nonzero eigenvalues corresponding to the minimum value of the M are used as the low dimensional coordinates. Usually the minimum eigenvalue is almost zero, so the characteristic vector corresponding to the eigenvalues located $2 \sim (d + 1)$ is used as the output result [21].

TABLE 2. Accuracy rate under different dimension d for training and testing data sets.

d	Accuracy rate for training data set	Accuracy rate for testing data set
2	73%	71.86%
3	99.50%	89.00%
4	100%	93.43%
5	100%	96.57%
6	100%	94.86%
7	100%	95.00%
8	100%	94.56%
9	100%	94.83%
10	100%	96.77%
11	100%	95.62%

In this paper, one thousand sets of data mentioned above were randomly selected to two groups. 50% of the sample points are the training data and the remaining 50% of the data are the testing data. The LLE algorithm is used to reduce the dimension of the data and the maximum number of neighbors is $k_{max} = 10$. For the selection of dimension d , there are no effective method to determine the proper value before the experiments so that all dimensions must be carried out the experiments form 2 to 10 in order to obtain the best dimension. The training accuracy and the test accuracy under different dimension d obtained by simulation experiments are shown in Table 2. It can be seen from Table 2 that the correctness of the testing set is optimal in the case of $d = 5$ and $d = 10$. Therefore, in order to meet the premise of the accuracy, $d = 5$ is selected as the number of the model inputs.

C. WHALE OPTIMIZATION ALGORITHM

The whale optimization algorithm (WOA) is a global optimization algorithm based on swarm techniques which is inspired by the hunting behavior of the Humpback whale proposed by Mirjalili and Lewis [9]. In order to establish the mathematical model of the humpback whale searching path, each humpback whale is set as a search agent. At the same time, the algorithm is divided into three steps: the search for prey, encircling prey and bubble-net attacking method (exploitation phase). Then these three kinds of behavioral mathematical model are established. In the process of modeling the bubble-net attacking method, the shrinking encircling mechanism and the spiral updating position method are proposed. Among them, the searching for prey can be regarded as the exploration stage, whose main purpose is to find a better solution. The bubble-net attacking method can also be called the exploitation phase, whose main purpose is to use this better solution more fully.

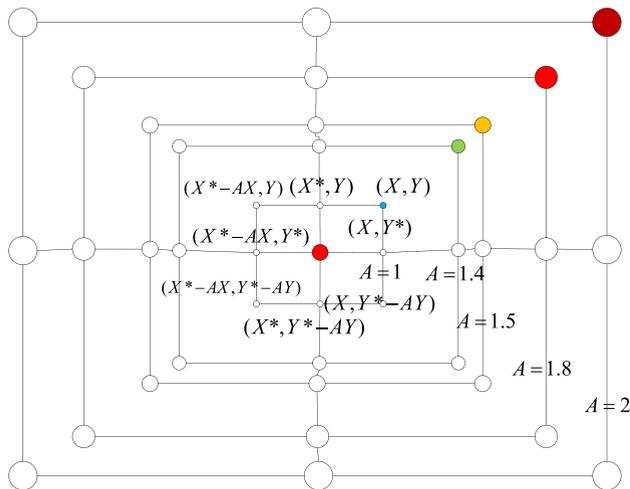


FIGURE 6. Theoretical basis of two-dimensional plane.

In the exploration phase (searching for prey), in order to make the search more extensive, the search agents are pushed away from each other by the positive and negative of a random vector \vec{A} . At the same time, the position of the search agent is replaced by a randomly selected search agent to replace the current optimal search agent. The mathematical model is described as follows:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}_{rand} - \vec{X} \right| \tag{27}$$

$$\vec{X}(t + 1) = \vec{X}_{rand} - \vec{A} \cdot \vec{D}, \tag{28}$$

where \vec{X}_{rand} is a random location vector selected from the current population. Fig. 6 shows a possible location of a solution when $\vec{A} > 1$. When $|\vec{A}| > 1$, the search agent selects a random agent, and when $|\vec{A}| < 1$, the search agent is selected as the current best agent to replace the location of the search agent.

At the initial stage, the optimal location in the search space is unknown when the prey is surrounded. In this algorithm, the best candidate is regarded as the target prey or the best target. After that, the best search agent will be defined, and other agents will try to update their location with the best agent. The mathematical model of this behavior can be described as follows:

$$\vec{D} = \left| \vec{C} \cdot \vec{X}^*(t) - \vec{X}(t) \right| \tag{29}$$

$$\vec{X}(t + 1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D}, \tag{30}$$

where t is the current iteration number, \vec{A} and \vec{C} are the coefficient vector, X^* is the known optimal location vector, \vec{X} is the location vector of other search agents. It is proposed that, in each iteration, if a better solution occurs, the X^* will be replaced with a better one.

The calculated expressions of \vec{A} and \vec{C} are described follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \tag{31}$$

$$\vec{C} = 2 \cdot \vec{r} \tag{32}$$

The \vec{a} is linearly reduced in the interval (0, 2) in the iterative process (including the whole exploration and development phase), and the vector \vec{r} is a random vector in interval [0,1].

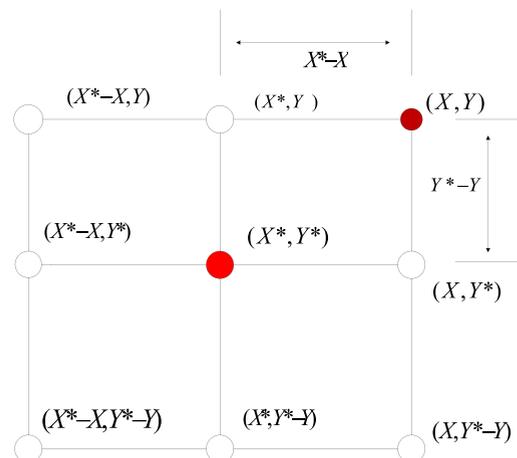


FIGURE 7. Theoretical basis of two-dimensional plane.

Fig. 7 illustrates the theoretical basis of (30) in two dimensional plane. The search agent (X, Y) can update its location based on the location of the current best agent (X^*, Y^*) . It can be seen from this point that if the choice of (X^*, Y^*) is not good, it will be easy to make the algorithm fall into the local optimum. It can be seen from Fig. 7 that anywhere near the best agent can be achieved by changing the value of the \vec{A} and \vec{C} .

The third stage is the bubble-net attacking method. In order to realize the model in this stage, two ideas are introduced.

(1) Shrinking encircling mechanism. By reducing the value of \vec{a} in (5), the value of is limited in $[-1,1]$. So the new search agent can be defined any \vec{A} location between the initial

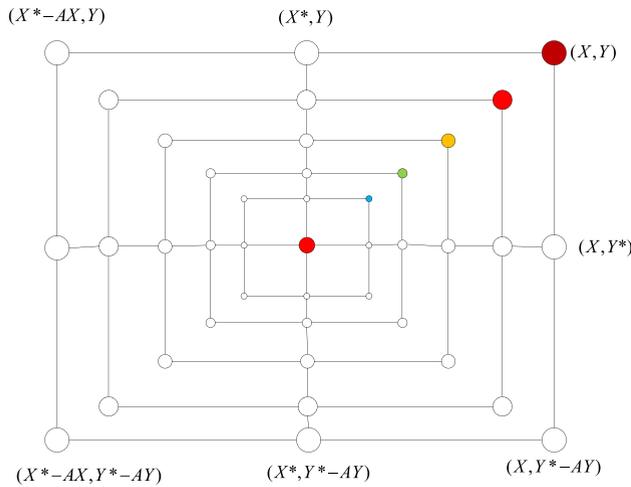


FIGURE 8. Theoretical basis of shrinking encircling mechanism.

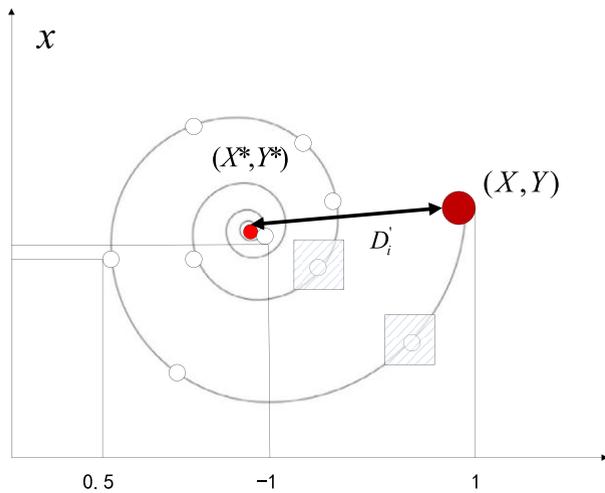


FIGURE 9. Sketch map of spiral renewal position method.

position and the current search agent’s best position. The theoretical illustration of the shrinking encircling mechanism is shown in Fig. 8. It can be seen that any position between (X, Y) and (X^*, Y^*) in the two dimensional plane can be reached by adjusting the value of A between $[0,1]$.

(2) Spiral updating position method. The sketch map of spiral renewal position method is shown in Fig. 9, which is the path of the search agent proposed by the original WOA. It calculates the distance between whale (X, Y) and (X^*, Y^*) prey. The equation imitating the humpback whales spiral moving mode is described as follows.

$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t), \quad (33)$$

where $\vec{D}' = |\vec{X}^*(t) - \vec{X}(t)|$ represents the distance between the i th whale and the prey, b is used to define the constant to limit the logarithmic spiral and l is a random number between the interval $[-1,1]$.

To be mentioned, the humpback whales swim around the prey in a gradual contraction of the circle and a spiral shape.

In order to facilitate the establishment of the model, it is stipulated that each of the fifty percent of the whales may choose to surround the contraction path or spiral model to update their locations. The mathematical model is described as follows.

$$\vec{X}(t+1) = \begin{cases} \vec{X}^*(t) - \vec{A} \cdot \vec{D} & p < 0.5 \\ \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) & p \geq 0.5, \end{cases} \quad (34)$$

where p is a random number between $[0,1]$.

D. CHAOS WHALE OPTIMIZATION ALGORITHM

In the analysis of the original WOA, the shortcomings of WOA can be found.

(1) The initialization process is random. Generally speaking, the stochastic process can make the initial solutions be random distribution, but the quality of the individual cannot be guaranteed. Some solution in the population is far from the optimal solution. If the initial population has good quality, it will help to improve the optimization efficiency and the quality of the solution. That is to say that the random of initial location for each search agent will therefore affect the convergence rate of the training process.

(2) WOA adopts (30) and (31) to update the position of each search agent. Its essence is to use the information of itself and the information of the global optimal value to guide the next iteration of the search agent. This is actually a positive feedback process, which is easy to make the algorithm fall into the local optimal solution.

By making full use of the advantage of the WOA’s fast convergence and the egocentricity of chaotic motion, a chaos whale optimization (CWOA) algorithm is proposed so that the ability of the WOA to get rid of the local extreme points is improved and the convergence speed and precision of the algorithm are improved. The simulation results show that the performance of chaotic whale optimization algorithm is superior to the initial whale optimization algorithm.

1) CHAOS AND ITS CHARACTERISTICS

Generally, the random state of motion obtained by the deterministic equations is called chaos. The chaotic state is widespread in natural phenomena and social phenomena. It is a common phenomenon in nonlinear systems and its behavior is complex and similar to random. It seems that a chaotic changing process is not completely chaotic and there is a fine inherent regularity. A chaotic variable in a certain range has the following characteristics: Randomization (its performance is as random as random variables), ergodicity (it can go through all the states in the space without repetition) and regularity (the variable is derived from the determined iterative equation).

The chaos optimization method is a novel optimization method, which uses the unique ergodicity of chaotic system to achieve global optimization, and it does not require the objective function to be continuous and differentiated. Chaotic motion can be used to optimize the search capabilities.

Its basic idea is to generate a set of chaotic variables with the same number of optimization variables firstly. Chaos is combined with the optimization variables using a similar carrier in order to make the optimization variable appear chaotic state. Let the range of the chaotic motion be the same as the range of the optimal variable. Then, the optimal value of the chaotic variable is used to search the optimal value. Because the chaotic motion has the characteristics of randomness, ergodic and sensitivity to initial conditions, it is no doubt that the chaos based search technology will be more superiority than other random search. The following Logistic equation is a typical chaotic system.

$$z_{n+1} = \mu z_n(1 - z_n) \quad n = 0, 1, 2, \dots \quad (35)$$

where μ is the control parameter. When $\mu = 4$ and $0 \leq z_0 \leq 1$, the system (35) will be completely in a confused state. From any initial value $z_0 \in [0, 1]$, a deterministic time series may be iterated.

2) BASIC PRINCIPLE OF CWOA

The basic idea of CWOA is mainly embodied in two aspects:

(1) The position of search agent is initialized by chaotic sequence, which does not change the stochastic nature of the WOA in initialization, and improves the diversity of search agents and ergodicity of agents by adopting chaos principle. Based on the produced initial agents, a set of better initial agents are selected as the initial population.

(2) The chaotic sequence is generated based on the optimal location of all the current search agents. The search agent which is the best position in the chaotic sequence replaces the position of a search agent in all search agents. The search algorithm which introduces the chaotic sequence can generate many neighborhood points around the local optimal solution in the iterative process, which can help some search agents that are easy to fall into the local optimum to escape from the local minima and quickly search for the optimal solution.

3) CHAOTIC ALGORITHM TO SEEK INITIAL POSITION

The distance between search agents and the prey is defined as the objective function:

$$f(t) = \left| \vec{X}^*(t) - \vec{X}(t) \right|, \quad (36)$$

where $\vec{X}^*(t)$ is the position of the prey and $\vec{X}(t)$ is the position of a certain search agent.

The algorithm procedure of CWOA is described as follows.

Step 1: Initialize the maximum number of allowed iterations or the fitness error limits, and the corresponding parameters of CWOA.

Step 2: Generate the positions of individuals by using chaos idea.

(1) Randomly generating a n vector, whose each component is located in the scope $(0,1)$, $z_1 = (z_{11}, z_{12}, \dots, z_{1n})$, n is the number of variables in the objective function. According to the Eq. (35), N vectors $z_1, z_2, z_3, \dots, z_n$ can be obtained.

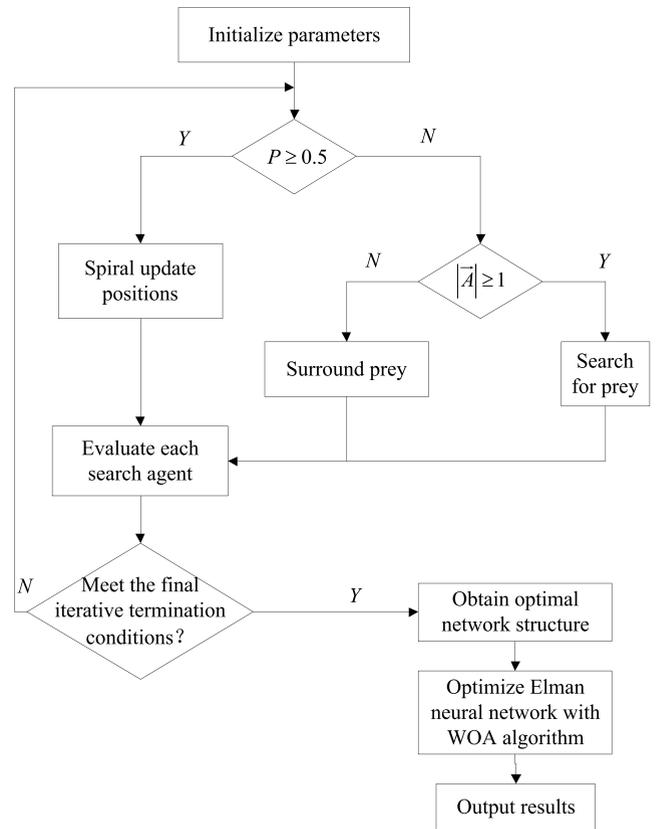


FIGURE 10. Flowchart of WOA-Elman algorithm.

(2) The various components of the z_i carrier to the corresponding variable range.

(3) The fitness value of the search agent is calculated, and the M solutions with better performance are selected as the initial solution from N initial agents.

Step 3: If the search agent is better than the current global extreme, the search agent is set to the optimal search agent. The value is set to the optimal value, where the smart body position is set to the best position

Step 4: According to the (30) and (31), the positions of the search agents are updated.

Step 5: The optimal position $P_w = (p_{w1}, p_{w2}, p_{w3}, \dots, p_{wi})$ is optimized by chaos optimization. $p_{wi}(i = 1, 2, \dots, m)$ is mapped to the definition domain $[0,1]$ of Logistic equation (34) is defined in the domain $[0,1]$, that is to say $z_i = (p_{wi} - a_i)/(b_i - a_i)$. Then, the Logistic equation is used to generate the chaotic sequence $z_i^{(k)}(k = 1, 2, \dots)$. The generated chaotic variables are returned to the original space by inverse mapping $P_{wi}^{(k)} = a_i + (b_i - a_i)z_i^{(k)}$ to obtain:

$$P_w^{(k)} = (P_{w1}^{(k)}, P_{w2}^{(k)}, P_{w3}^{(k)}, \dots, P_{wi}^{(k)}). \quad (37)$$

In the original solution space, the fitness value of each feasible solution $P_w^{(k)}(k = 1, 2, \dots)$ is calculated to obtain the best feasible solution P^* .

Step 6: Replacing the position of any search agent in the current group with P^* .

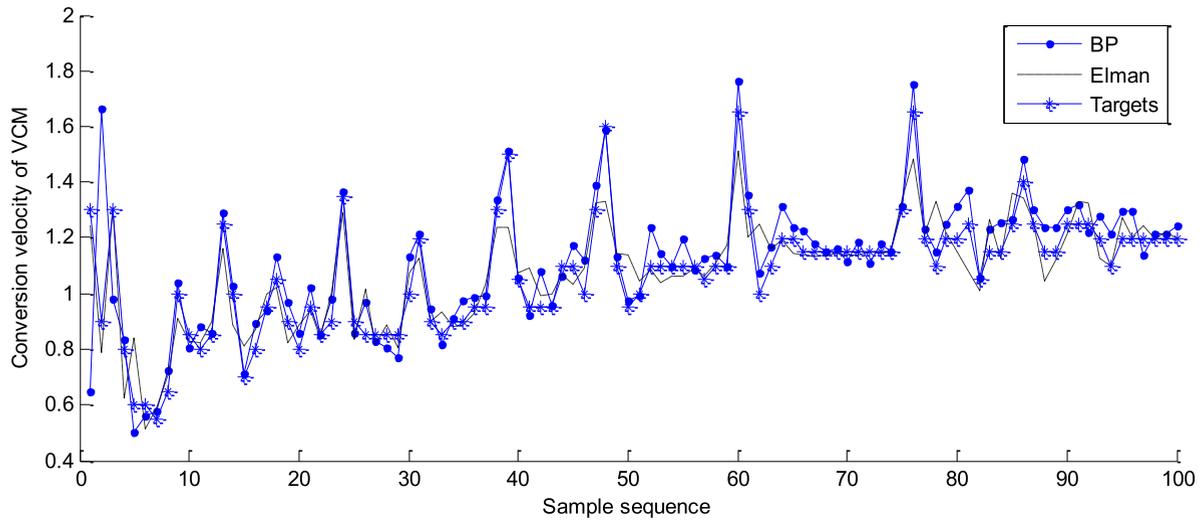


FIGURE 11. Predictive results of VCM conversion velocity under BP and Elman model.

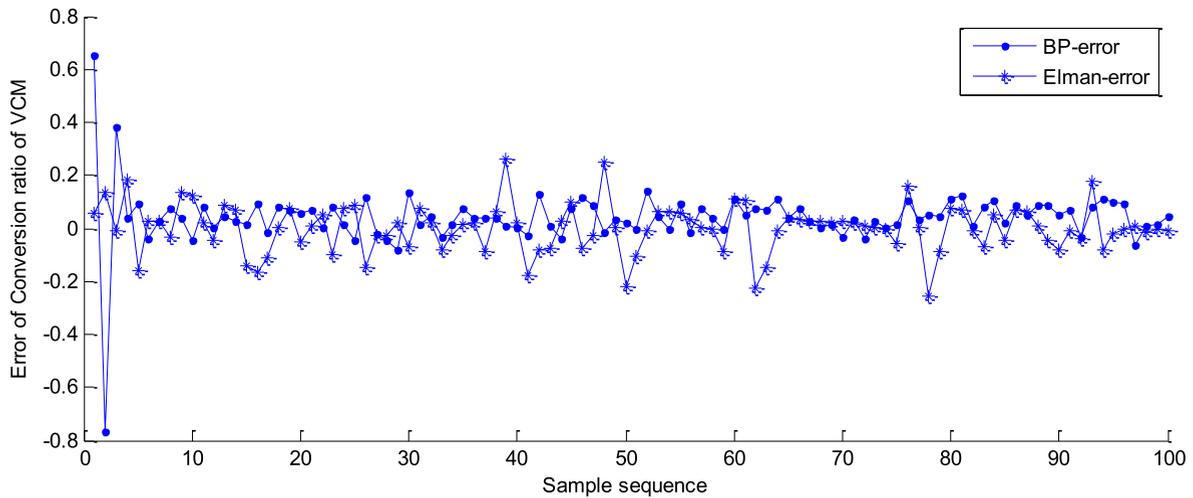


FIGURE 12. Predictive error of VCM conversion velocity under BP and Elman model.

Step 7: If the termination condition required by the algorithm is satisfied, the search stops and the global optimal position is outputted. Otherwise return Step 3.

4) ALGORITHM PROCEDURE OF CWOA TO OPTIMIZE ELMAN NEURAL NETWORK

Seen from the basic principle of Elman neural network, the main parameters include the weights W^3 from the hidden layer to the output layer, the weights W^2 from the input layer to hidden layer and the weights W^1 from the undertake layer to the hidden layer. So, optimizing the Elman neural network is to optimize the above three matrices. In the training phase of the Elman neural network, the best intelligent agents' positions of WOA are in accordance to the connection weights of the Elman neural network. The algorithm flowchart of CWOA to optimize Elman neural network is shown in Fig. 10.

(1) Initialization. Initialize the number of searching agents N , the maximum number of iterations num .

(2) Calculate the fitness value. The degree of fitness used in this paper is also the reciprocal of the previous objective function $1/f(t)$, which is shown as:

$$g(t) = \frac{1}{|\vec{X}^*(t) - \vec{X}(t)|} \tag{38}$$

(3) According to the random value p , the implementation of the surrounding shrinkage mechanism or spiral location updating strategy for the search agents is determined. In fact, this judgment can be a whole, because when the humpback whale spit bubbles, the radius of the spiral path is gradually increased.

(4) According to the value of the coefficient vector \vec{A} , the implementation of the surrounding prey or the search for

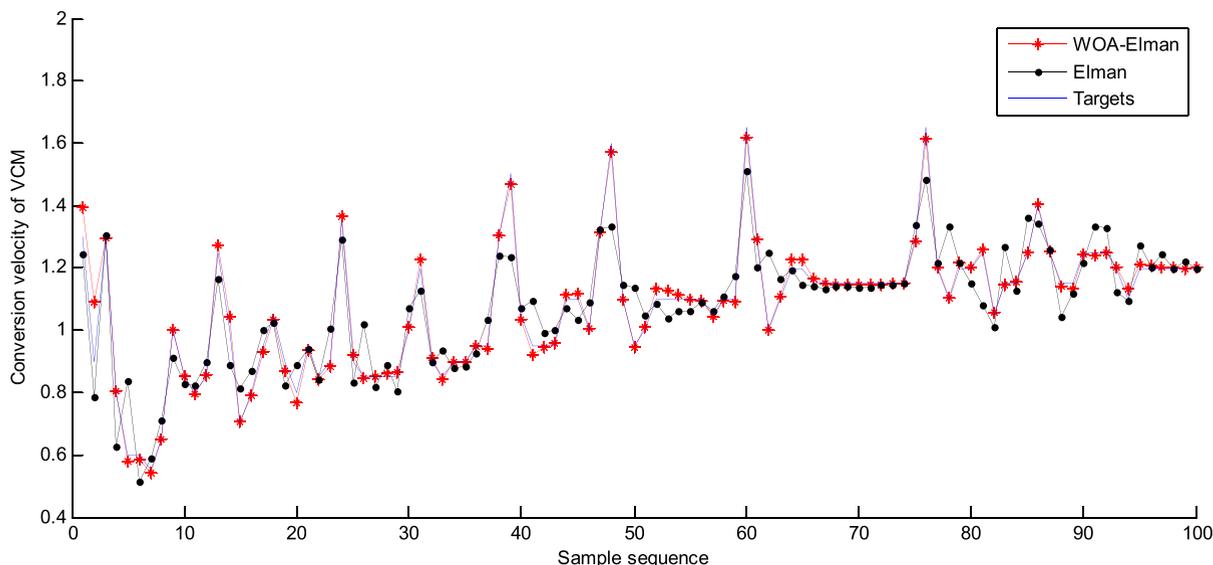


FIGURE 13. Predictive results of VCM conversion velocity under WOA-Elman model and Elman model.

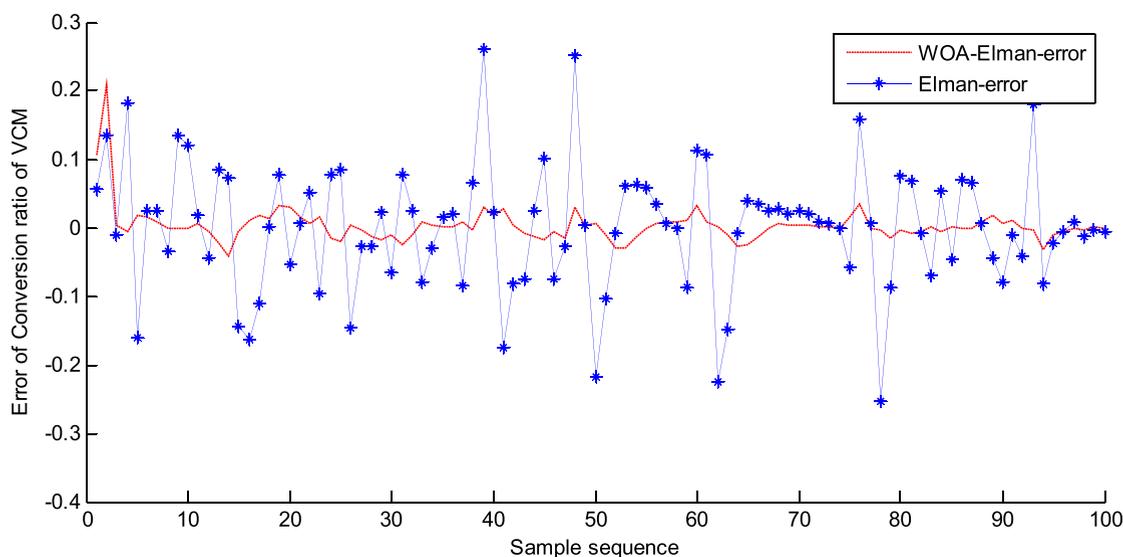


FIGURE 14. Predictive error of VCM conversion velocity under WOA-Elman model and Elman model.

prey is determine. Performing the search for prey is to let the agent as far as possible dispersion so as to achieve the global optimization. Execution surrounds the prey is to allow the more intelligent body concentration.

(5) Judge the termination conditions. Judge whether the number of iterations to reach the maximum or not. If it is satisfied, output the calculation results. Otherwise, $num = num + 1$, turn to Step (2).

V. SIMULATION EXPERIMENTS AND RESULTS ANALYSIS

In this paper, the polymerization industrial process of a chemical factory with 40000 tons/year polyvinyl chloride (PVC) production device is taken as background, whose technology is introduced by America B · F · G company, taking vinyl

chloride monomer (VCM) as raw material, and using suspension polymerization technology to produce polyvinyl chloride (PVC) resin. A soft-sensor model of the VCM conversion velocity in the polyvinyl chloride (PVC) production process based on CWOA-Elman NN is put forward. The parameters when initializing WOA are set as follows: the number of search agents $N = 100$, the maximum number of iterations is 500. In order to reflect the performance of the recurrent neural network with feedback is better than that of the feed-forward neural network, the BP neural network is selected to be compared with the original Elman NN. In order to measure the performances of prediction models, several performance indicators are defined in Table 3 [12], where, $y(t)$ is predicted value and $y_d(t)$ is actual value.

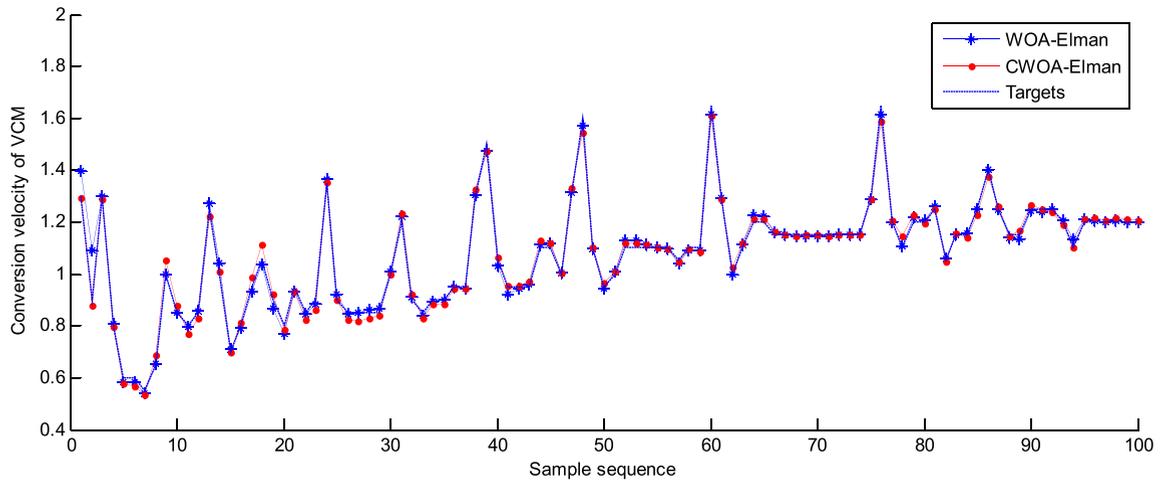


FIGURE 15. Predictive results of VCM conversion velocity under CWOA-Elman and WOA-Elman model.

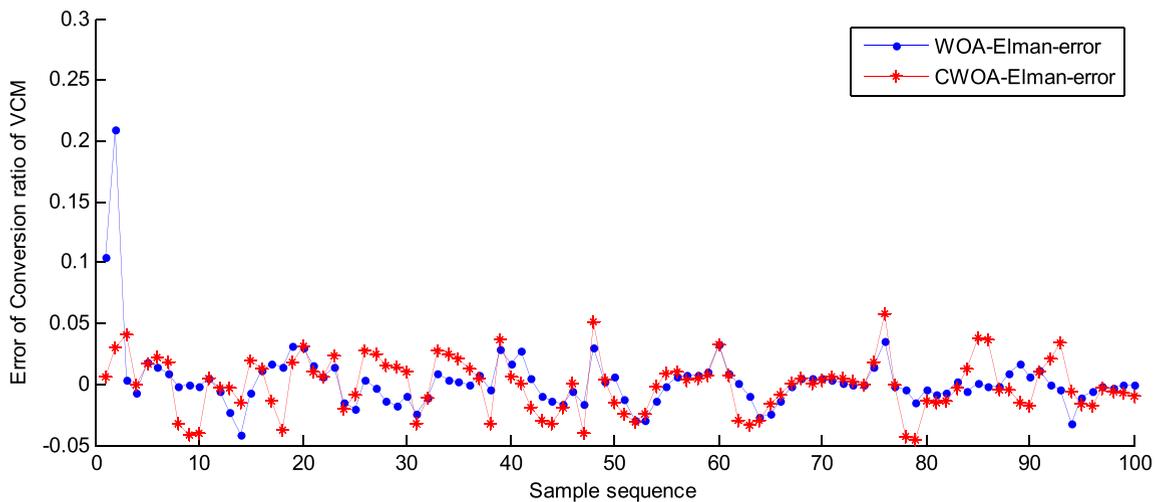


FIGURE 16. Predictive error of VCM conversion velocity under CWOA-Elman and WOA-Elman model.

TABLE 3. Definition of model performance index.

Index	Function
Maximum positive error	$MPE = \max\{(\hat{y} - y), 0\}$
Maximum negative error	$MNE = \min\{(\hat{y} - y), 0\}$
Root mean square error	$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2 \right]^{\frac{1}{2}}$
Sum of squared error	$SSE = \sum_{i=1}^n (\hat{y}_i - y_i)^2$

The production historical data of PVC polymerization process are collected and 2 kettle including 1600 sets historical data with the uniformity and representations are chosen.

Then after data preprocessing the data is divided into two parts, the front 1500 sets data is the training data, and the rest 100 sets data is used to validate the performance of soft-sensor model. Fig. 11 shows the comparison between the original Elman neural network model and BP neural network model for the prediction of VCM conversion rate. Fig. 12 is the error comparison between the BP model and the Elman model. It is found that the Elman NN model is better than the BP neural network model in soft sensing by observing the comparison results. Fig. 11 and Fig. 12 is to compare the forecasting performance of the BP neural network model and the original Elman neural network model. It can be seen from Fig. (11-12) that in the primary stage of prediction, the convergence of the Elman neural network model is better than the BP neural network model. But seen from the whole prediction process, the prediction accuracy of Elman neural network is higher than that of BP neural network.

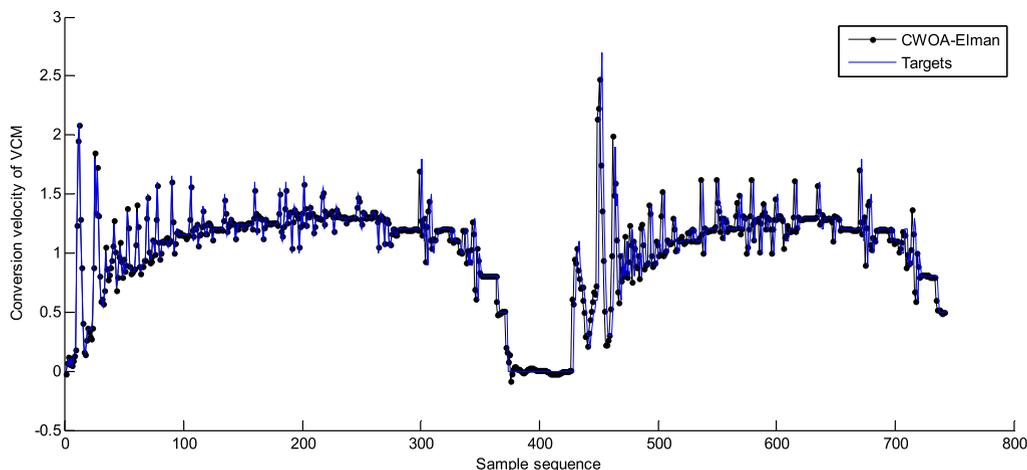


FIGURE 17. Results of continuous prediction for two kettle data under CWOA-Elman model.

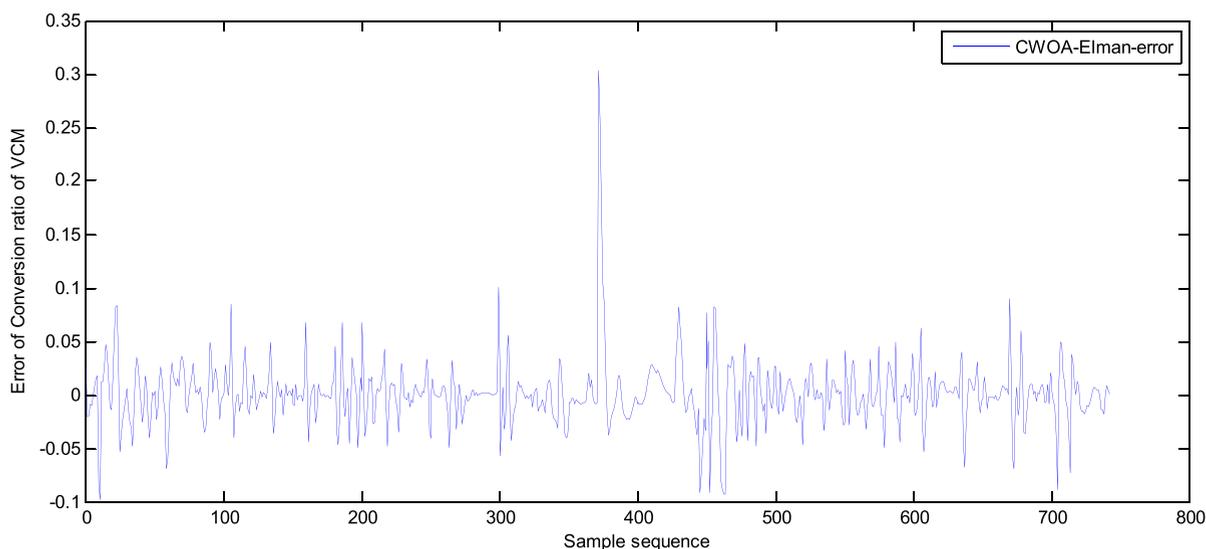


FIGURE 18. Error of continuous prediction of two kettle data under CWOA-Elman model.

However, it still can be seen from Fig. 11, in the initial stage of prediction, the prediction accuracy of Elman neural network model is better than the BP neural network model, but there is still a great error, which prompted to improve the algorithm performance.

Fig. 13 is the predicted output of the Elman NN model and the WOA-Elman NN model. Fig. 14 is the error comparison between Elman model and WOA-Elman NN model. Fig. 15 is the comparison of WOA-Elman NN model and CWOA-Elman NN model. Fig. 16 is the comparison of the prediction error between CWOA-Elman NN model and WOA-Elman NN model. Fig. 17 is the verification of the stability of the system after the introduction of the perturbation. Fig. 18 is a continuous prediction of the error of the two kettle. Table 4 is the performance comparison.

It can be seen from Fig. (13-14) that after introducing the WOA algorithm, the prediction accuracy of Elman neural

TABLE 4. Performance comparison results of different soft-sensor models.

Soft-sensor model	MPE	MNE	SSE	RMSE
Elman NN	0.2610	-0.2527	0.0077	1.0938
BP NN	0.3645	-0.4632	0.0096	1.4538
WOA-Elman NN	0.0629	-0.0432	0.0003	0.0173
CWOA-Elman NN	0.0353	-0.0403	0.0001	0.0144

network has been greatly improved. But in the initial stage of prediction, we can still see that the Elman neural network optimized by WOA algorithm has very good prediction accuracy, but there are still some shortcomings in the early prediction, the error is obvious. After analysis, the error may be due to the initial position of search agents of WOA and the initial

weights of Elman neural network are random. Therefore, the chaos algorithm is adopted to optimize the initial positions of the search agents in the WOA so as to enhance the search ability and the diversity of the search agents.

It can be seen from Fig. 15-16, the prediction performance of the CWOA-Elman neural network model based on chaos theory has been very good, especially in the early stage of prediction. At the same time, the anti-disturbance ability of CWOA-Elman neural network model based on chaos theory is also very strong.

In order to verify the robustness of the CWOA-Elman NN model, this paper argues that after the end of each production, the initial stage of the second production can be considered as a perturbation. Therefore, this paper will train the neural network to predict the two kettle data continuously. It can be seen from the Fig. 17-18, after the end of the first reactor at 371st sample points, the initial stage of the production of the second kettle appeared a large oscillation, but it immediately converges. Therefore, the robustness of the model is worthy of recognition.

In conclusion, the proposed soft-sensor model based on CWOA-Elman NN has higher prediction accuracy than the WOA-Elman NN model and BP NN model for predicting the VCM conversion velocity in the PVC polymerization process. Its application of predicting the VCM conversion velocity in PVC polymerization process has great significance in improving the capacity of equipment and reducing the production cost.

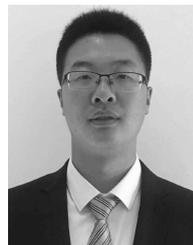
VI. CONCLUSIONS

Based on the good nonlinear approximation ability of Elman NN and the simplicity and easiness to realize and obtain the global extreme, avoid local extreme and have the certain characteristics of the adaptive ability to the search space of WOA, CWOA is proposed to optimize the connection weights of Elman NN among layers. Then a CWOA-Elman soft-sensor model is proposed to predict the conversion velocity of VCM. The simulation results show that the soft-sensor model based on the proposed CWOA-Elman NN has high prediction accuracy.

REFERENCES

- [1] S. Zhou *et al.*, "Hybrid intelligent control scheme of a polymerization kettle for ACR production," *Knowl.-Based Syst.*, vol. 24, no. 3, pp. 1037–1047, 2011.
- [2] J. S. Wang, S. Han, and Q. P. Guo, "Echo state networks based predictive model of vinyl chloride monomer convention velocity optimized by artificial fish swarm algorithm," *Soft Comput.*, vol. 18, no. 3, pp. 457–468, 2014.
- [3] S. Raghu, N. Sriraam, and G. P. Kumar, "Classification of epileptic seizures using wavelet packet log energy and norm entropies with recurrent Elman neural network classifier," *Cognit. Neurodyn.*, vol. 11, no. 1, pp. 1–16, 2017.
- [4] Y. Ge *et al.*, "An indicated torque estimation method based on the Elman neural network for a turbocharged diesel engine," *Proc. Inst. Mech. Eng. Part D, J. Automobile Eng.*, vol. 230, no. 10, pp. 1299–1313, 2016.
- [5] X. Z. Gao and S. J. Ovaska, "Genetic algorithm training of Elman neural network in motor fault detection," *Neural Comput. Appl.*, vol. 11, no. 3, pp. 37–44, 2002.

- [6] C. Tan *et al.*, "A pressure control method for emulsion pump station based on elman neural network," *Comput. Intell. Neurosci.*, vol. 215, no. 3, pp. 1–8, 2015.
- [7] Y.-S. Sun, Y.-M. Li, G.-C. Zhang, Y.-H. Zhang, and H.-B. Wu, "Actuator fault diagnosis of autonomous underwater vehicle based on improved elman neural network," *J. Central South Univ.*, vol. 23, no. 3, pp. 808–816, 2016.
- [8] Z. Zhang and W. Gong, "Short-term load forecasting model based on quantum elman neural networks," *Math. Problems Eng.*, vol. 2016, Mar. 2016, Art. no. 7910971.
- [9] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, Mar. 2016.
- [10] A. Ebrahimi and E. Khamelchi, "Sperm whale algorithm: An effective metaheuristic algorithm for production optimization problems," *J. Natural Gas Sci. Eng.*, vol. 29, pp. 211–222, Feb. 2016.
- [11] A. Kaveh and M. I. Ghazaan, "Enhanced whale optimization algorithm for sizing optimization of skeletal structures," *Mech. Based Design Struct. Mach.*, vol. 130, no. 4, pp. 1–18, 2016.
- [12] W. H. Cui, J. S. Wang, and S. X. Li, "KPCA-ESN soft-sensor model of polymerization process optimized by biogeography-based optimization algorithm," *Math. Problems Eng.*, vol. 2015, no. 3, pp. 1–10, 2015.
- [13] H. Liu *et al.*, "Wind speed forecasting approach using secondary decomposition algorithm and Elman neural networks," *Appl. Energy*, vol. 157, pp. 183–194, Nov. 2015.
- [14] C. Lin and E. A. Boldbaatar, "Fault accommodation control for a biped robot using a recurrent wavelet Elman neural network," *IEEE Syst. J.*, to be published.
- [15] F. J. Lin, K. H. Tan, and C. H. Tsai, "Improved differential evolution-based Elman neural network controller for squirrel-cage induction generator system," *IET Renew. Power Generat.*, vol. 10, no. 3, pp. 988–1001, 2016.
- [16] Y. Wang *et al.*, "Fault diagnosis for manifold absolute pressure sensor(MAP) of diesel engine based on Elman neural network observer," *Chin. J. Mech. Eng.*, vol. 29, no. 3, pp. 386–395, 2016.
- [17] C. Ni and X. Yan, "Elman neural networks with sensitivity pruning for modeling fed-batch fermentation processes," *J. Chem. Eng. Jpn.*, vol. 48, no. 3, pp. 230–237, 2015.
- [18] H. Wang *et al.*, "Application of dimension reduction on using improved LLE based on clustering," *J. Comput. Res. Develop.*, vol. 43, no. 3, pp. 1485–1490, 2006.
- [19] Y. Liu *et al.*, "Incremental supervised locally linear embedding for machinery fault diagnosis," *Eng. Appl. Artif. Intell.*, vol. 50, pp. 60–70, Apr. 2016.
- [20] X. Wang *et al.*, "Bearing fault diagnosis based on statistical locally linear embedding," *Sensors*, vol. 15, no. 3, pp. 16225–16247, 2015.
- [21] G. Wang *et al.*, "Fault diagnosis of supervision and homogenization distance based on local linear embedding algorithm," *Math. Problems Eng.*, vol. 2015, no. 3, pp. 1–8, 2015.



W. Z. SUN is currently pursuing the master's degree with the School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan, China.



J. S. WANG received the B.Sc. and M.Sc. degrees in control science from the University of Science and Technology Liaoning, China, in 1999 and 2002, respectively, and the Ph.D. degree in control science from the Dalian University of Technology, China, in 2006. He is currently a Professor and a Master's Supervisor with the School of Electronic and Information Engineering, University of Science and Technology Liaoning. His main research interest is the modeling of complex industry process, intelligent control, and computer-integrated manufacturing.

• • •