Hindawi
Journal of Sensors
Volume 2017, Article ID 7872030, 13 pages
https://doi.org/10.1155/2017/7872030

Research Article

Soft-Sensor Modeling of PVC Polymerizing Process Based on F-GMDH-Type Neural Network Algorithm

Wei-zhen Sun,[1] Jie-sheng Wang,[1,2] and Shu-zhi Gao[3]

[1] *School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan City, Liaoning Province, China*
[2] *National Financial Security and System Equipment Engineering Research Center, University of Science and Technology Liaoning, Anshan, Liaoning Province, China*
[3] *College of Information and Engineering, Shenyang University of Chemical Technology, Shenyang, Liaoning Province, China*

Correspondence should be addressed to Jie-sheng Wang; wang_jiesheng@126.com

Received 14 December 2016; Accepted 30 January 2017; Published 19 February 2017

Academic Editor: Fanli Meng

For predicting the conversion velocity of the vinyl chloride monomer (VCM) in the polymerization process of polyvinylchloride (PVC), an improved Group Method of Data Handling- (GMDH-) type neural network soft-sensor model is proposed. After analyzing the technique of PVC manufacturing process, the auxiliary variables for setting up the soft-sensor model are selected and the experimental data are normalized. Because the internal standard of the original GMDH-type neural cannot solve the problem of multiple-collinearity problem and the useful variables tend to be prematurely eliminated in the modeling process, a hybrid method combining the regression analysis method and the least squares method is proposed to solve the multiple-collinearity problem. On the same time, by adopting some optimization experiences in genetic algorithm (GA), the generational crossover combination variables method is proposed to solve the shortcoming of useful variable being eliminated prematurely. The simulation results show that the proposed soft-sensor model can significantly improve the prediction accuracy of economic and technical indicators in the PVC polymerization process and can meet the real time control requirements of polymerization reactor production process.

1. Introduction

Polyvinyl chloride (PVC) is one of the most widely used polymers in the world, and it is also one of the first polymers to realize industrial production [1]. The production of polyvinyl chloride (PVC) is a kind of typical batch chemical production process by using vinyl chloride monomer (VCM) as raw material and using the suspension polymerization process to produce polyvinyl chloride (PVC) resin. The quality index of PVC is closely related to its processing process, but it is easily influenced by many factors, such as raw materials, additives, and technique parameters. Among them, the conversion rate and conversion velocity of VCM are the most important factors affecting the quality of PVC. The different VCM conversion has a certain impact on the molecular weight of PVC resin, thermal stability, porosity, the residues of VCM, the absorptivity of plasticizers, and processing liquidity [2]. The conversion rate of VCM is a key factor affecting the thermal aging property of the resin. The higher the conversion rate, the higher the yield of PVC resin. When the conversion rate reaches 80%, if the proportion of the unstable structure in the polymer PVC continues to increase, the thermal aging property of PVC resin will decreased. However, due to the limitation of the industrial field conditions and lack of mature measurement equipment, the conversion rate and conversion velocity of vinyl chloride in the actual production process are very difficult to obtain in real time, so it is also difficult to achieve the direct and efficient closed-loop control [3]. So it is very important to establish the soft-sensor model of the conversion rate and the conversion velocity of VCM.

Group method of data handling (GMDH) is a family of inductive algorithms for computer-based mathematical modeling of multiparametric datasets that features fully automatic structural and parametric optimization of models [4]. Inductive GMDH algorithms give possibility of finding

automatically interrelations in data, to select an optimal structure of model or network, and to increase the accuracy of existing algorithms [5]. GMDH is a heuristic self-organization method and its essence is parameter estimation. Based on the dynamic analysis of the target process, the heuristic method is used to search the model structure, and then the model parameters are estimated according to the model results. The best GMDH model is indicated by the minimum of the external criterion characteristic. Multilayered procedure is equivalent to the artificial neural network (ANN) with polynomial activation function of neurons. Therefore, the algorithm with such an approach usually referred to as GMDH-type neural network (NN) or Polynomial Neural Network [6]. GMDH-type neural network is also a feed-forward network; the key is that it has three advantages over other forward neural networks [7]. (1) It can obtain the explicit function analytic expression of the model. That is to say it solves the problem that the model structure can be used to reveal the interaction and dependence among all variables, which cannot be achieved by the traditional neural network in the past. (2) The modeling process of the model is self-organized without any initial assumptions. Because the algorithm is based on the data driven to find the input items that have a substantial impact on the explanatory variables. (3) It has the optimal complexity and high precision prediction. It reduces the impact of small samples or bigger noise on the system and ensures the system's generalization ability.

GMDH-type neural network has been applied in a great variety of areas for data mining and knowledge discovery, forecasting and systems modeling, optimization, and pattern recognition. GMDH neural network predictive model combining Harmony Search (HS) algorithm was proposed to predict the pullout capacity of suction caissons in clay [8]. Neurofuzzy method was combined with GMDH network so as to establish the NF-GMDH forecast model to predict the local scour depth around pile groups under clear-water conditions [9]. GMDH-type NN was used to realize the short-term prediction and the prediction accuracy was very stable [10]. The wavelet transformation (WT) algorithm was introduced into GMDH-type NN. The time series of significant wave height (SWH) were decomposed into some subseries using WT and then the decomposed time series were imported to the GMDH NN model to forecast the SWH in different time periods [11]. The particle swarm optimization (PSO) algorithm and Neurofuzzy theory were introduced into the GMDH network to realize the NF-GMDH-PSO algorithm, which was used to predict the longitudinal dispersion coefficient of the river. In this paper, an improved Group Method of Data Handling- (GMDH-) type neural network soft-sensor model is proposed for predicting the conversion velocity of the Vinyl Chloride Monomer (VCM) in the polymerization process of polyvinylchloride (PVC). The paper is organized as follows. In Section 2, the technique flowchart of the PVC polymerization process is introduced. The GMDH-type neural network is described in Section 3. In Section 4, the improved GMDH-type neural network soft-sensor model is introduced in detail. The simulation experiments and results analysis are discussed in Section 5. The conclusion illustrates the last part.

2. Polymerization Production Process of PVC

In the resin industrial production industry, the following four kinds of polymerization patterns are generally used: suspension polymerization, noumenon polymerization including gas phase polymerization, emulsion polymerization including microsuspension polymerization, and solution polymerization [1]. The suspension polymerization production technology is the main production mode of PVC resin because it is easy to adjust the product variety, the production process is simple, and it is easy to be controlled and realize the mass production. The typical PVC polymerization process is shown in Figure 1 [3].

PVC is polymerized by VCM. The general production process of the PVC resin based on suspension method is firstly to clean the polymerization reactor, which includes the cleaning before and after nurikabe; then the vinyl chloride monomer, water and suspending agent, and antioxidant are added in the polymerization kettle. These materials form a suspension in the polymerization reactor under strong agitation. The PVC monomer was polymerized into PVC particles at the elevated temperature and with the addition of initiator. When the polymerization proceeds to a certain extent, these particles will form PVC slurry. This is the polymerization process of polyvinyl chloride. In addition, the PVC production process also includes monomer recovery, PVC slurry stripping, PVC drying, and the packaging of the products. The production flow chart of polymerization kettle is shown in Figure 2.

PVC polymerization process is a typical batch process. In the polymerization process of PVC, all kinds of raw materials and auxiliary agents are put into the reactor. They are fully and evenly dispersed under the function of stirring. Then, the cooling water is ventilated to the clip set of the reaction kettle and baffle plate constantly in order to remove homopolymer. When the conversion rate of VCM reaches a certain value, the reaction terminates, that is to say the finished products are obtained. The degree of polymerization decreases with the increase of temperature, and the degree of polymerization is only related to the reaction temperature of VCM. Ultimately, the accuracy of the conversion velocity prediction model directly affects the quality of the product and the type of polyvinyl chloride. According to the characteristics of the polymerization process, 10 process variables related to the conventional rate and velocity of VCM are identified as the secondary variables of the soft-sensor model, which are listed in Table 1 [2].

3. GMDH Neural Network

The basic idea of GMDH is described as follows. The black box analysis method is used to establish the relationship between input and output. Then the function of the network model is expressed by the description of the relationships among the elements in the network. The establishment of GMDH-type neural network is a process of continuously

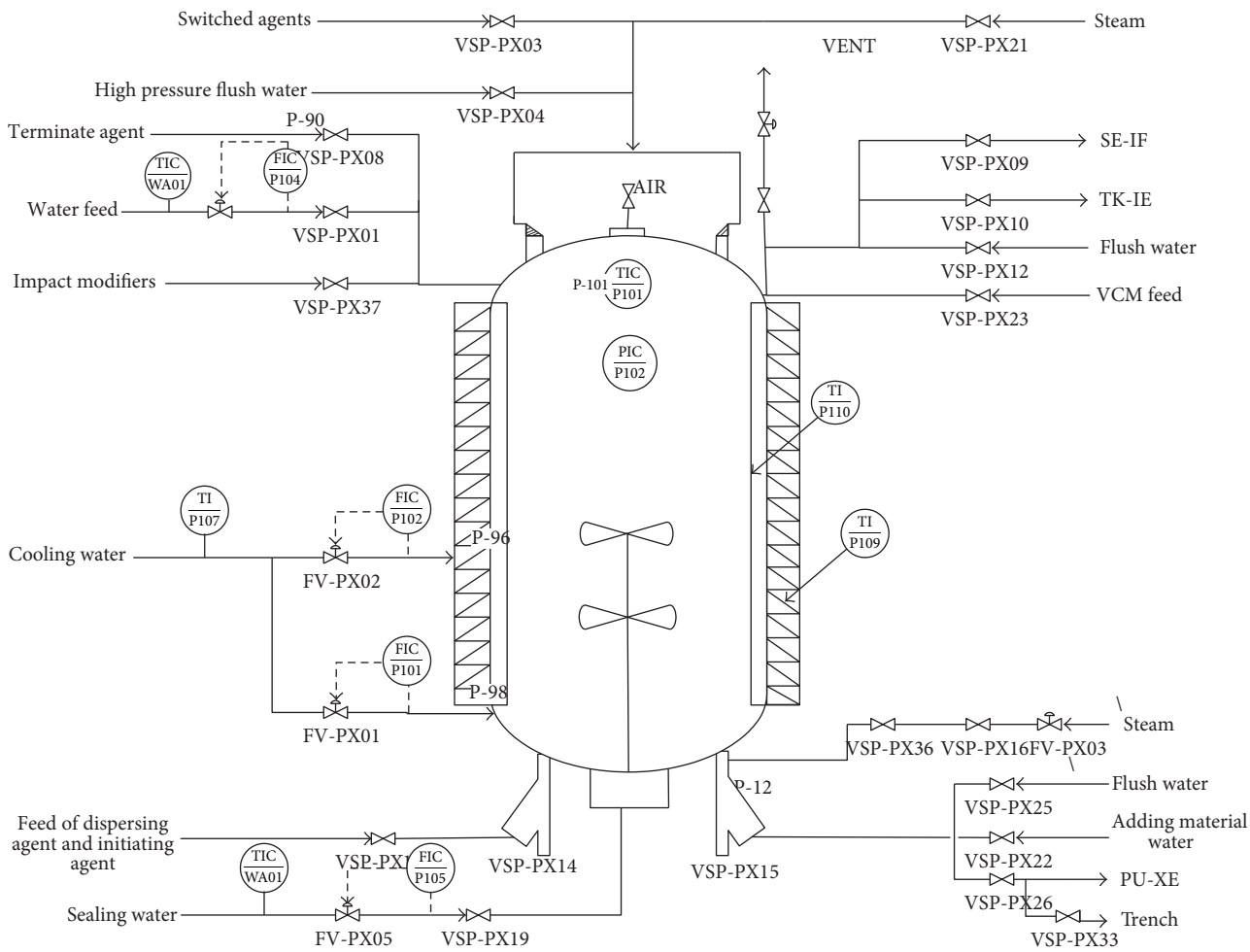

Figure 1: Technique flowchart of polymerization kettle.

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

producing active neurons. Then the external criteria are adopted to screen neurons; the quality of the retained neurons is generally superior to that of the discarded neurons (although some "better" neurons may be prematurely excluded). To combine the retained outstanding neurons is to screen out the better neurons as the neurons of the next layer until that the optimal model is selected.

3.1. Principle and Network Structure of GMDH-Type NN. The algorithm flowchart to generate GMDH-type NN is shown in Figure 3. The algorithm procedure of the standard GMDH-type NN is described as follows.

(1) Divide the training set and test set. The samples data set U is divided into the training set A and the testing set B. The number of samples $N_U = N_A + N_B$, where N_U is the total

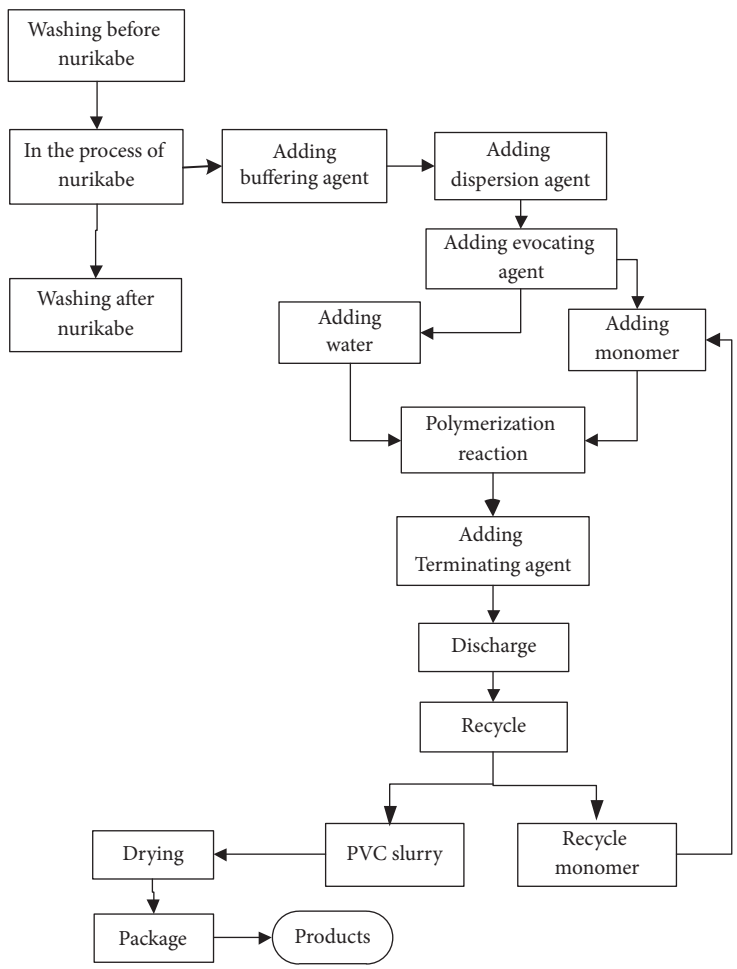

Figure 2: Flow chart of the polymerization reactor process.

number of samples, N_A is the number of training samples, and N_B is the number of testing samples.

(2) Select the reference function to establish the relationship between the input variables and the output variables. In general, the discrete form of Volterra function or Kolmogorov-Gabor function is used as the reference functions

$$y = a_0 + \sum_{i=1}^{M} a_i x_i + \sum_{i=1}^{M}\sum_{j=1}^{M} a_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M} a_{ijk} x_i x_j x_k + \cdots . \quad (1)$$

(3) Determine the external criteria.

(a) Prediction error sum of squares (PESS)

$$\text{PESS} = \sum_{t=1}^{m} \left[Y(t) - \underline{Y}(t)\right]^2 . \quad (2)$$

(b) An information criterion (AIC)

$$\text{AIC} = m \ln S_k^2 + C + 2k,$$
$$S_k^2 = \frac{1}{m}\sum_{t=1}^{m} \left[Y(t) - \underline{Y}(t)\right]^2 . \quad (3)$$

(c) Average relative error (ARE)

$$\text{ARE} = \frac{1}{m}\sum_{t=1}^{m} \frac{\left|Y_t - \underline{Y_t}\right|}{Y_t}. \quad (4)$$

In the above three criteria, $\underline{Y_t}$ is the estimated output value of the intermediate model on the tth sample, Y_t is the actual output value of the intermediate model on the tth sample, C is a constant, k is a tunable parameter, and m is the number of the observed samples.

(4) Generate variables of the initial layer. Each item in the selected reference function is used as the initial input variable of the algorithm. If the K-G polynomial has been selected, it is shown as follows, where x_1, x_2 are the variables of the input data,

$$f(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1^2 + a_4 x_2^2 + a_5 x_1 x_2. \quad (5)$$

Thus, the input variables of the network can be obtained as follows:

$$v_1 = a_1 x_1,$$
$$v_2 = a_2 x_2,$$

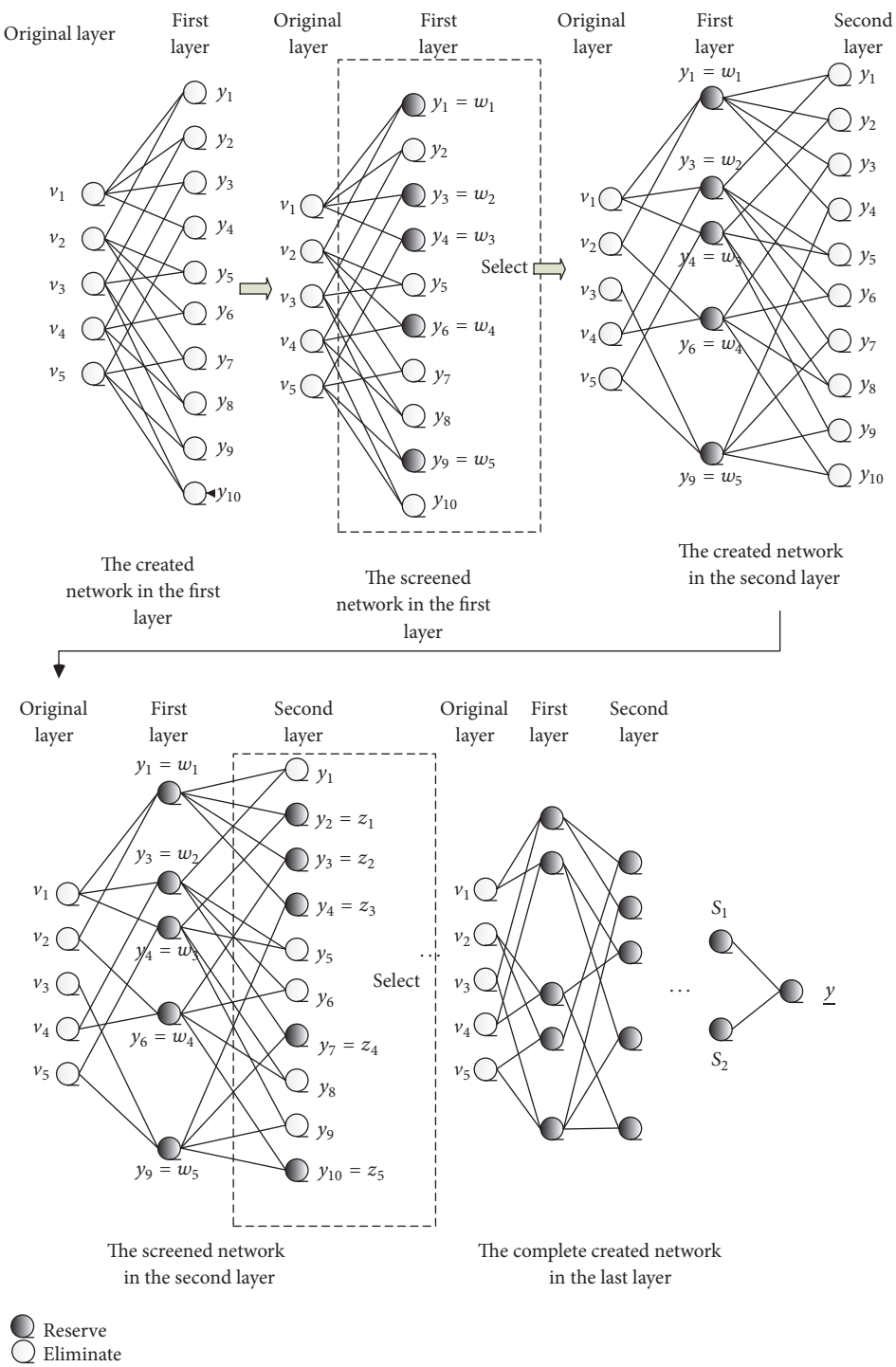

Figure 3: Schematic diagram of network structure of GMDH.

$$
\begin{aligned}
v_3 &= a_3 x_1^2, \\
v_4 &= a_4 x_2^2, \\
v_5 &= a_5 x_1 x_2, \\
v_6 &= a_0.
\end{aligned} \tag{6}
$$

(5) The intermediate model of the first layer is generated by the internal standard. The standard GMDH adopts the least square method. Self-organization strategy is used to generate an intermediate model $y_k = f(x_i, x_j)$ of the first layer. The parameters $(a_1, a_2, a_3, a_4, a_5, a_0)$ of y_k are estimated on the training set A according to the inner criterion.

(6) Select the intermediate model. On the testing data set, the intermediate model of the upper layer is selected by using

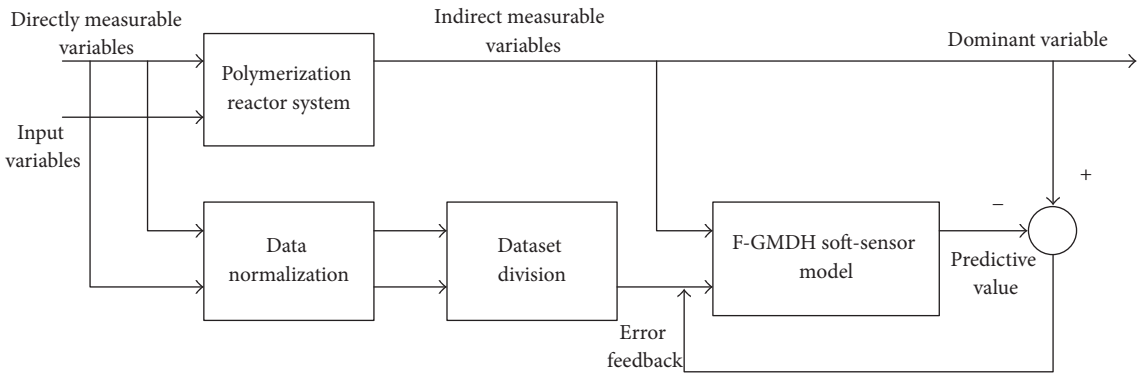

Figure 4: Structure of soft-sensor model.

the external criterion determined in Step (3). Then L models with small external criterion values are selected as the input variables of the next layer. This method is carried out continuously to screen the variables until the optimal results are obtained.

(7) Generate the optimal model. Firstly, the termination rule is established. Normally, the model with the lowest external criterion value or the external criterion value which no longer decreases is selected as the optimal model. The complexity of the model produced in Step (6) is increasing, so the final optimal model can be determined according to the selected termination rule.

3.2. Advantages and Disadvantages of the Standard GMDH-Type NN. In this paper, the standard GMDH-type NN and the classic back propagation (BP) neural network are compared on the following factors: the structure of the network, the connection pattern of neurons, the self-organizing, the estimating method of parameters, the rules of using samples, the interpretability of the model, the convergence performance of the model, and the complexity of the program. The compared results are shown in Table 2.

It can be seen from Table 2 that the most obvious advantages of GMDH are that there is a clear expression, faster convergence rate and ease of controlling the estimated parameters, and so on. But seen from the network structure in GMDH modeling process, GMDH-type NN has certain shortcomings described as follows. (1) It can be seen from the GMDH structure chart, after screening the first layer variables, that the unselected variables will be discarded forever. Because these variables are only filtered once, there are some "useful variables" which will inevitably be prematurely eliminated, which will thereby reduce the quality of the model. (2) For the selection of initial variables, if too many initial variables are selected, the system will become too complex and difficult to understand. But if you choose too few variables, it will lead to some "useful variables" being eliminated in advance. So it is important to choose the initial variable. (3) The internal criterion that the standard GMDH-type NN generates the intermediate model is the least squares estimation method. Because of the limitation of the sample data, the traditional least squares method cannot estimate the coefficient of the regression coefficients when the multiple-collinearity relationship is generated between the regressions, which will reduce the accuracy and reliability of the model constructed by the traditional GMDH-type NN.

4. Improved GMDH-Type Neural Network

In view of the shortcomings of the traditional GMDH-type NN mentioned in Section 3.2, the paper will propose an improve GMDH-type NN to overcome these three shortcomings so as to improve its generalization ability. As the main point of this improvement is to improve the internal criterion, the stepwise regression analysis is introduced to eliminate the multiple-collinearity. Therefore, in order to facilitate the expression, the algorithm is called F-GMDH.

4.1. Structure of the Proposed Soft-Sensor Model. Ten variables described in Section 2 are set as the input variables and the conversion velocity and conversion rate of VCM are the output variables. The improved GMDH-type neural network is used to fit the nonlinear relationship $f\{\cdot\}$ between input and output so as to establish the soft-sensor model of VCM conversion velocity and conversion rate, whose model structure is shown in Figure 4.

4.2. Division Method of Samples Set in Standard GMDH-Type NN

4.2.1. Cross Division Method. The standard GMHD-type NN generally divides the samples into the training data set and testing data set with the manual pattern, but this classification method relies on the personal experience to determine which part of the data is suitable for training data sets and which part of the data for the test data set. At the same time, the intermediate models generated by the different partitioning methods are very different. Such a division strategy is contrary to the idea of self-organization and self-evolution of the GMDH. In order to solve this problem, this paper puts forward the cross division method, which is described as follows.

(1) For a given set of samples $U_{\text{all}} = \{x_1 x_2, \ldots, x_n\}$, randomly generate a positive integer j, $j \in (1, n)$.

(2) The training set and test set are denoted as $A_{\text{train}} = \{x_1 x_2, \ldots, x_j\}$ and $B_{\text{test}} = \{x_{j+1} x_{j+2}, \ldots, x_n\}$.

Table 2: Characteristic comparison between BP neural network and GMDH-type neural network.

Name	BP neural network	GMDH-type neural network
Structure of the Network	Network structure with three or more layers of network	Completely rely on the self-evolvement of network
Connection pattern of the neurons	Full connection among layers	Partial connection among layers
Self-organizing	The number of layers and the number of neurons of the network are predetermined	The number of layers and the number of neurons in the network are self-generated
Estimating method of parameters	Generated by iteration	Estimate the parameters of the nodes according to the internal criterion
Rules of using samples	All the samples are used to train the parameters, which can easily lead to the phenomenon of over fitting	The part samples are used for training the parameters and the rest are used for screening
Interpretability of Model	It is hard to understand	There is an explicit expression
Convergence performance of the model	There is a strong relationship between convergence and initial states	Determine the convergence state according to the termination criterion
Complexity of the program	The calculation time is too long, so it is suitable for using multiprocessor	It is suitable for multiprocessor or single processor

(3) Inspired by the crossover operator in genetic algorithm (GA), at the same time, in order to allow the training set and testing set to be full crossed, in this paper, the data in the two data sets are crossed one by one. Generate a positive integer a, $a \in (1, j)$; extract the data x_a from the training set into the testing set. At the same time, generate a positive integer b, $b \in (j+1, n)$, extract the data x_b from the testing set, and put it into the training set. $(A_{\text{train},1}, B_{\text{train},1}), (A_{\text{train},2}, B_{\text{train},2}), \ldots$ and $(A_{\text{train},p}, B_{\text{train},p})$ are obtained until all the original data in the two data sets are exchanged, where p is the number of exchanges. Then, the intermediate model is generated on the training set A_{train} according to the inner criterion and the searching process for the optimization is carried out by using the testing set B_{test}.

4.2.2. Description of OLS-Frisch Algorithm. When there are multiple-collinearity relationships among the regression coefficients, the traditional least squares method cannot estimate the coefficients of the regression coefficients. This paper introduces a stepwise regression method to eliminate the multiple-collinearity relationship among these variables. Firstly, the standard least squares method is used to establish the model by using one set of variables. Then variables are fed into the model one by one. Each introduction of a variable will results in carrying out a significant test for all variables in the model and extracting no significant variables. The recycling process is gradually carried out (variable introduction-variable elimination-variable introduction) until all the variables in the model are not significant. In order to facilitate the expression, the algorithm is called OLS-Frisch method.

Step 1. The least square method is used to establish the model, whose specific calculation process is described as follows [12, 13].

There are m pairs of observational data $(\overline{x_i}, y_i)$, $i = (1, 2, \ldots, m)$, about the variable x and y. Suppose

$$A = \begin{bmatrix} \varphi_1(x_1) & \varphi_2(x_1) & \cdots & \varphi_n(x_1) \\ \varphi_1(x_2) & \varphi_2(x_2) & \cdots & \varphi_n(x_2) \\ \cdots & \cdots & \cdots & \cdots \\ \varphi_1(x_m) & \varphi_2(x_m) & \cdots & \varphi_n(x_m) \end{bmatrix},$$

$$\underline{b} = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_m \end{bmatrix},$$

$$\underline{r} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdots \\ \rho_m \end{bmatrix},$$

$$\underline{x} = \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_m \end{bmatrix}, \tag{7}$$

where A represents the regression matrix, $\underline{b}$ is the desired output, $\underline{r}$ is the default mutual independent residual vector, and $\underline{x}$ is the regression coefficient waiting to be solved and needing to be updated.

The training error can be expressed as follows:

$$E(c_1, c_2, \ldots, c_m) = \sum_{i=1}^{m} \left[y_i - \sum_{j=1}^{n} c_j \varphi_j(x_i) \right]^2 = \sum_{i=1}^{m} \rho_i^2, \tag{8}$$

where, $\varphi_j(x_i)$ is a known function. It can be seen that E is a multivariate function related to $c_1, c_2, \ldots, c_m$. The training error E should be minimized in the training sample under 2-norm sense. So let $\partial E/\partial c_j = 0$ $(j = 1, 2, \ldots, n)$ to obtain $c_1, c_2, \ldots, c_m$. The solving process is described as follows.

Suppose $\varphi_{ji} = \varphi_j(x_i)$ to obtain

$$\begin{aligned} \frac{\partial E}{\partial c_j} &= -2\sum_{i=1}^{m} \left[y_i - (c_1\varphi_{1i} + c_2\varphi_{2i} + \cdots + c_n\varphi_{ni}) \right] \varphi_{ji} \\ &= 2\sum_{i=1}^{m} \left[(c_1\varphi_{1i} + c_2\varphi_{2i} + \cdots + c_n\varphi_{ni}) \right] \varphi_{ji} - 2\sum_{i=1}^{m} \varphi_{ji} y_i. \end{aligned} \tag{9}$$

Suppose $M = (m_{jk})$ is a n-matrix, $\vec{g} = (g_1, g_2, \ldots, g_n)^T \in R^n$, $b_{jk} = \sum_{i=1}^{m} \varphi_{ki}\varphi_{ji}$, $g_i = \sum_{i=1}^{m} y_i \varphi_{ji}$ $(j, k = 1, 2, \ldots, n)$. $\partial E/\partial c_j = 0$ $(j = 1, 2, \ldots, n)$ to obtain

$$\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{bmatrix} = \begin{bmatrix} g_1 \\ g_2 \\ \cdots \\ g_n \end{bmatrix}. \tag{10}$$

According to the above equations, the values of $c_1, c_2, \ldots, c_n$ are obtained, which are the required coefficients. Then the first regression equation is established.

Step 2. The initial correlation coefficient matrix $R^{(0)}$ is obtained from the above sample data set.

Step 3. The stepwise optimization of the variables is to establish the optimal regression equation.

(1) According to the modeling method described in Step 1 to establish a regression model: $\underline{z}_k = d_j^{(1)} z_j$, $j = 1, 2, \ldots, k-1$, where $k = 10$. The test value F is selected, which is used to introduce and eliminate the variables. The principle of the introduction is to make the partial regression square sum largest, which is also called the largest variance contribution. Because the greater the partial regression square sum, the better the effect of the regression equation.

The selection method is a direct method. The direct method is to obtain the maximum partial regression square sum from $R^{(0)}$. That is to say find out the relationship from $R^{(0)} \to R^{(1)}$ based on the inverse compact transformation method. The following results can be obtained from the above transformation relation:

$$d_l^{(1)} = r_{jk}^{(1)} = \frac{r_{jk}^{(0)}}{r_{jj}^{(0)}},$$
$$c_{jj}^{(1)} = r_{jk}^{(1)} = \frac{1}{r_{jj}^{(0)}} = \frac{1}{c_{jj}^{(0)}}, \tag{11}$$

where, r represents the diagonal elements in matrix R and c_{jj} is the diagonal elements of the inverse matrix of R. According to the above deduction, the partial regression square sum of z_j can be expressed as follows:

$$u_j^{(1)} = \left[d_j^{(1)}\right]^2 c_{jj}^{(0)} = \frac{\left[d_j^{(1)}\right]^2}{c_{jj}^{(1)}} = \frac{\left[r_{jk}^{(0)}/r_{jj}^{(0)}\right]^2}{\left[1/r_{jj}^{(0)}\right]} = \frac{\left[r_{jk}^{(0)}\right]^2}{r_{jj}^{(0)}}. \tag{12}$$

It can be seen from the above deduction that $u_j^{(1)}$ can be obtained from $R^{(0)}$. So there is the following relationship:

$$\begin{aligned} R^{(1)} &\longrightarrow u_j^{(2)}, \\ R^{(2)} &\longrightarrow u_j^{(3)}, \\ &\vdots \\ R^{(k-1)} &\longrightarrow u_j^{(k-2)}. \end{aligned} \tag{13}$$

The introduction process of variables is described as follows.

(1) Determine the variables which will be introduced. The direct method is used to calculate all partial regression square sum $u_j^{(0)}$, and then the largest $\max(u_j^{(0)})$ is selected as the first introduced variable z_1; that is to say $\underline{z_k} = d_1^{(1)} z_1$, $j = 1$.

(2) In this paper, the F testing method was used to test the variables. Firstly, the critical value $F_\alpha^{(f_1,f_2)}$ is determined. The size of the value is mainly related to the reliability and degree of freedom, so the selection value must be moderate. If the value is too large, it will cause the introduction of too few variables so as to make the established model not practical. In this paper, $k = 10$. If you choose six variables and each group of data has 361 samples, there is $n = 361$, $f_1 = 6$, $f_2 = n - k - 1 = 354$. Thus $F^{(f_1,f_2)} = F^{(6,354)}$, where α is the corresponding parameter of the F distribution table. When the calculated F is greater than the critical value $F_\alpha^{(f_1,f_2)}$ in F distribution table under the corresponding significant level α ($\alpha = 0.05$) and the degree of freedom f_1, f_2, $F = (\mathrm{ESS}/k)/[\mathrm{RSS}/(n-k-1)]$. That is to say when F is greater than $F_\alpha^{(f_1,f_2)}$, there are significant differences between the two groups of variances; when F is less than $F_\alpha^{(f_1,f_2)}$, there is no significant difference between the two groups of variances.

(3) $R^{(1)}$ can be obtained by solving the inverse compact transformation method on $R^{(0)}$. The process of eliminating variables is described as follows. If it is the first introduction of a variable, you need not eliminate the variable; if the variable is introduced in the Nth sequence, then perform the step of eliminating the variable. The method of eliminating variables is also F testing method.

(4) Introduce new variables. Repeat the above Steps until the optimal model is established.

4.2.3. Intelligent Variable Selection Method (Generational Crossover Method). As mentioned in Section 3.2, the variables screening method adopted by the standard GMDH is very rough, which is easy to make some "useful variables" prematurely eliminated and the reliability and accuracy of the final obtained model not ideal. In view of the above problems, this paper puts forward the intelligent variables selection method, which is mainly reflected in the following two aspects. (1) The reasonable choice of the initial variables because it can make the "quality" of the initial variables be guaranteed. (2) Initialize the reasonable rules for the reservation of the neurons in the middle layer so as to ensure that the "useful variables" will not be prematurely eliminated.

(1) Intelligent selection of the number of initial variables: in the division process of the sample set, there are a lot of group data available for training. So in the training of the first data set, the variable is selected from one to N. Then all the models are compared and a model with the best current performance is selected. So the number of initial variables corresponding to this model k is used as the initial variable number in the training of the remaining data set. In order to make the initial variable number more flexible, the number of initial variables is defined as $(k-l, k+l)$, $l < k$ and $k+l$ is less than the maximum number of variables. By adopting this method, a reasonable number of initial variables can be chosen and the method is easier to be programed.

(2) The retained principle of intermediate variables: the retention of intermediate variables has influence on the results of the multilayer iteration, which is described as follows. Suppose a group of independent variables $x_1, x_2, \ldots, x_n$ are the identified object. The output of a linear system is y. The input variables $x_1, x_2, \ldots, x_n$ consist of the vector space $S^n = \{x_1, x_2, \ldots, x_n\}$. Then the projection of the output y on x_i is $\mathrm{Pr}_{si} y$, and $\mathrm{Pr}_{si} y > 0$, $i = 1, 2, \ldots, n$. So the expression of y on the vector space S^n is described as follows:

$$\underline{y_s} n = a_0 + a_1 x_1 + a_2 x_2 + \cdots + a_n x_n, \quad \underline{y_s} n \in S^n. \tag{14}$$

If $S^n \supset S^{n-1} \supset \cdots \supset S^1$, $\mathrm{Pr}_s^n y \succ \mathrm{Pr}_s^{n-1} y \succ \cdots \succ \mathrm{Pr}_s^1 y$, where $\succ$ means "better than."

According to the modeling characteristics of the GMDH algorithm and the principle of orthogonal projection, the necessary and sufficient conditions for the optimal approximation are described as follows.

(1) The necessary condition for the final calculation result with the optimal expression is that the number of

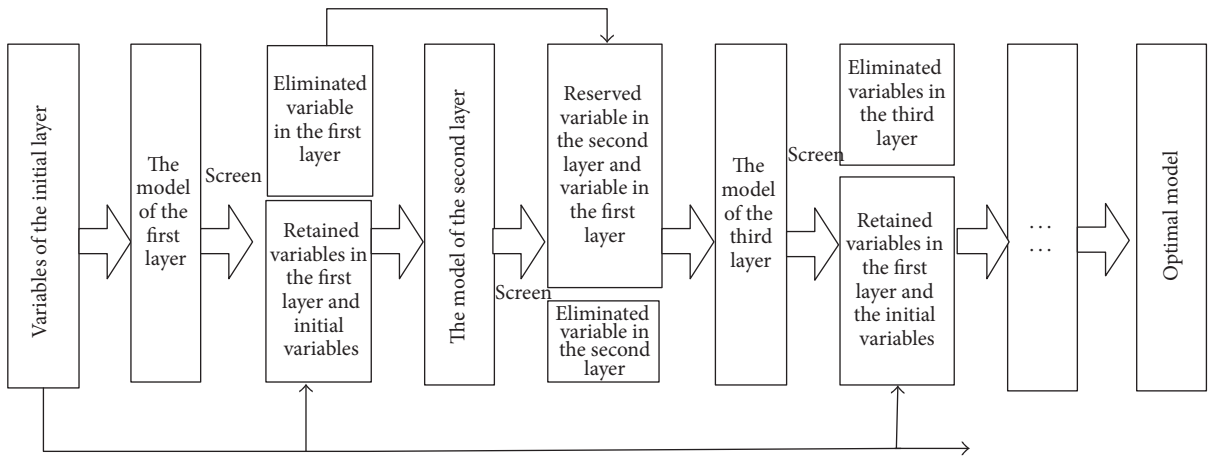

Figure 5: Schematic diagram of the variables combination patterns.

intermediate variables RE_i in the ith layer satisfies the following relationship:

$$\mathrm{RE}_i \geq \frac{n}{2}, \quad n = 2, 3, 4, \ldots,$$
$$\mathrm{RE}_2 \geq \frac{\mathrm{RE}_1}{2} \cdots \mathrm{RE}_i \geq \frac{\mathrm{RE}_{i-1}}{2}. \tag{15}$$

(2) The sufficient condition for the final calculation result with the optimal expression is that the final result of the operation is that the number of intermediate variables RE_i in the ith layer satisfies the following relationship:

$$\mathrm{RE}_1 \geq C_n^2, \quad n = 3, 4, 5, \ldots;$$
$$\mathrm{RE}_2 \geq C_n^{2^2}, \quad n = 5, 6, 7, \ldots; \tag{16}$$
$$\mathrm{RE}_i \geq C_n^{2^i}, \quad n \geq 2^i + 1.$$

The necessary condition is obtained under ideal conditions, and the sufficient condition is obtained under the worst condition. Therefore, if we want to get the optimal model, the above necessary and sufficient conditions need be satisfied and the selection method is designed to obtain the intermediate variables. In view of this problem, the idea of the method proposed in this paper is described as follows.

Because there are too many useful information in the initial variables, the initial variables are used to establish the model as much as possible so as to avoid the useful information prematurely eliminated, while the rest variables in each layer eliminated from the whole variables may have useful information. So, the main idea of this paper is to generate the first initial variable layer model; then the retained variables from the first layer and the initial variables are carried out in the crossover operation to generate the second layer model. Then the eliminated variables from the first layer and the retained variables from the second layer are carried out in the crossover operation to generate a third layer model. On this basis, if there are three consecutive times the variable is not selected, the variable will be completely removed, and so on, until the formation of the optimal model. Its structure is shown in Figure 5.

In the combination of variables, the crossover strategy is used to combine two sets of variables. The method to generate the intermediate model by using the random combination pattern is to replace the enumeration method of the standard GMDH algorithm to generate the intermediate model by using the variables pairwise combinatorial method. This can reduce the number of intermediate models in order to improve the efficiency of the model, reduce the complexity of the model, and effectively eliminate the redundant elements of the model.

4.2.4. Algorithm Flowchart of F-GMDH Method

Step 1. In order to make data have the same quantity rank in the training process for the GMDH-type NN, the normalized method is adopted to deal with the input and output data of neural network.

Step 2. Divide the sample set. The training set and testing set are divided by using the partitioning method mentioned in Section 4.2.1.

Step 3. Select the transfer function. The followed transfer function used in this paper is described as follows:

$$f(x_1, x_1) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1^2 + a_4 x_2^2 + a_5 x_1 x_2. \tag{17}$$

Step 4. Generate the intermediate model in the first layer. In this paper, the OLS-Frisch method is used as the internal standard to train the intermediate model in the first layer.

Step 5. Screen the intermediate model. The method mentioned in Section 4.2 is used to combine the initial variables, the eliminated variables, and the retained variables in each layer. Then the next layer model is generated based on the provided external criteria.

Step 6. Repeat Step 4 and Step 5 until the optimal model is obtained. The termination condition is that the external

Table 3: Definition of model performance index.

MPE	$\mathrm{MPE} = \max\{(\hat{y} - y), 0\}$
MNE	$\mathrm{MNE} = \min\{(\hat{y} - y), 0\}$
RMSE	$\mathrm{RMSE} = \left[\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2\right]^{1/2}$
SSE	$\mathrm{SSE} = \sum_{i=1}^{n}(\hat{y}_i - y_i)^2$

criterion value is not reduced or the iteration number reaches a certain given value.

5. Simulation Experiments and Results Analysis

In this paper, a chemical group company's 40 thousand tons/year PVC production unit of the polymerization process is selected as the research object, where the VCM is the raw material and the production of PVC resin is based on the suspension polymerization technique. In the process of polymerization, the conversion rate of VCM is an important parameter. The conversion rate of VCM has a great influence on the quality of PVC resin products, such as the molecular weight, porosity, plasticizer absorption rate, VCM residues, and thermal stability. In conclusion, it directly affects the quality of PVC products and the economic and technological efficiency of enterprises. Therefore, it is very important to predict and control the conversion rate and conversion velocity.

In this paper, the simulation experiments data are divided into training set and test set. The definition of model performance index is listed in Table 3, which includes the root mean square error (RMSE), the square sum of error (SSE), the maximum positive error (MPE), and the maximum negative error (MNE). The prediction results of F-GMDH-type NN model NN model and the standard GMDH-type NN model are shown in Figures 6–10 and Table 4. The predictive results of VCM conversion rate under BPNN and GMDH-type NN are shown in Figure 6. The predictive error of VCM conversion rate is shown in Figure 7. It can be clearly seen that the prediction accuracy of the BP network model is not better than the prediction accuracy of the standard GMDH-type NN model. The predictive results of VCM conversion rate under GMDH-type NN and F-GMDH-type NN are shown in Figure 8 and the predictive errors of VCM conversion rate are shown in Figure 9. The performance comparison results of VCM conversion rate under different soft-sensor models are listed in Table 4. It can be seen from the simulation results that the F-GMDH-type NN model has a higher precision than the standard GMDH-type NN model.

Figures 10 and 11 are the comparison results of the prediction results of the conversion velocity of VCM under different soft-sensor models. The predictive results of VCM conversion velocity under BPNN, GMDH-type NN, and F-GMDH-type NN are shown in Figure 10. The predictive errors of VCM conversion velocity under BPNN, GMDH-type NN, and F-GMDH-type NN are shown in Figure 11. The performance comparison results of VCM conversion velocity under different soft-sensor models are listed in Table 5.

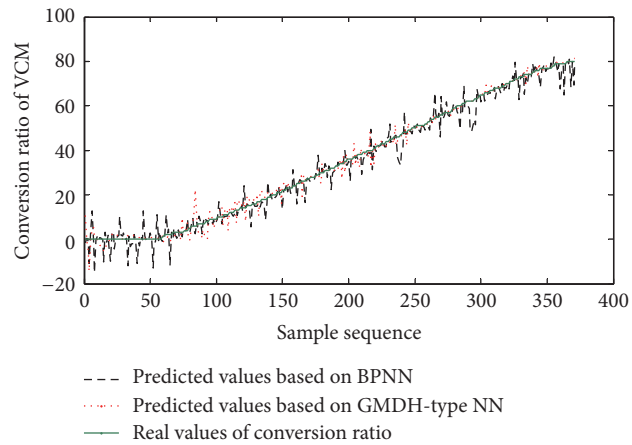

Figure 6: Predictive results of VCM conversion rate under BPNN and GMDH-type NN.

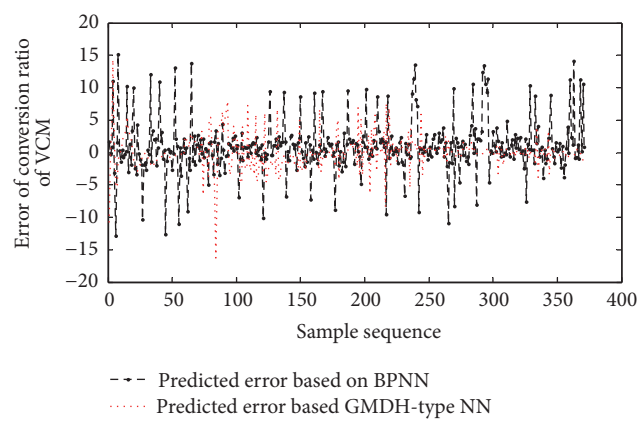

Figure 7: Predictive error of VCM conversion rate under BPNN and GMDH-type NN.

It can be seen from Figure 11 that the accuracy of F-GMDH-type NN model in the prediction of the conversion velocity of VCM is better than the standard GMDH-type NN model and BP NN model. The only shortcoming is that the training time of the proposed F-GMDH-type NN model is long.

6. Conclusions

In order to verify the advantages of the improved GMDH network (F-GMDH), the prediction performance experiments are carried out between the original GMDH neural network and the BP neural network, and the simulation results show that the original GMDH network is better than the BP neural network. Then the original GMDH neural network is compared with the F-GMDH neural network. The simulation results show that the prediction accuracy of F-GMDH neural network is better than the other two neural networks. In addition, the time required for the F-GMDH neural network model is much less than other models. So

Table 4: Performance comparison results of VCM conversion rate under different soft-sensor models.

Name	MPE	MNE	RMSE	SSE	Time/s
BP	14.3965	−16.2978	0.0040	5.0785	13.7 s
GMDH	11.5673	−15.1060	0.0023	3.4687	5.6 s
F-GMDH	3.0617	−4.2171	0.0003	2.9867	6.4 s

Table 5: Performance comparison results of VCM conversion velocity under different soft-sensor models.

Name	RMSE	SSE	MPE	MNE
F-GMDH	0.0621	0.0020	0.0353	−0.0411
GMDH	0.1398	0.0039	0.1308	−0.0474
BP	1.0938	0.0078	0.2610	−0.2527

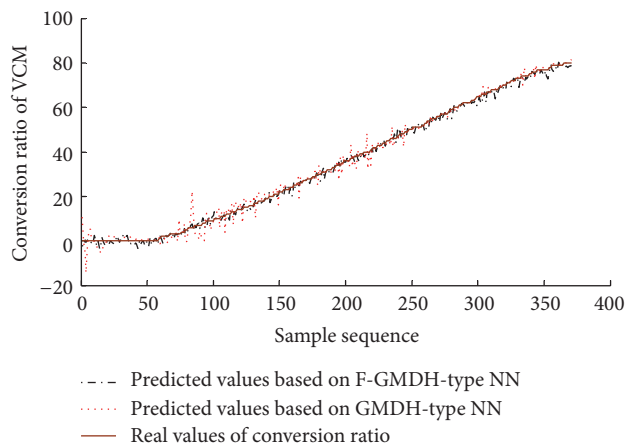

Figure 8: Predictive results of VCM conversion rate under GMDH-type NN and F-GMDH-type NN.

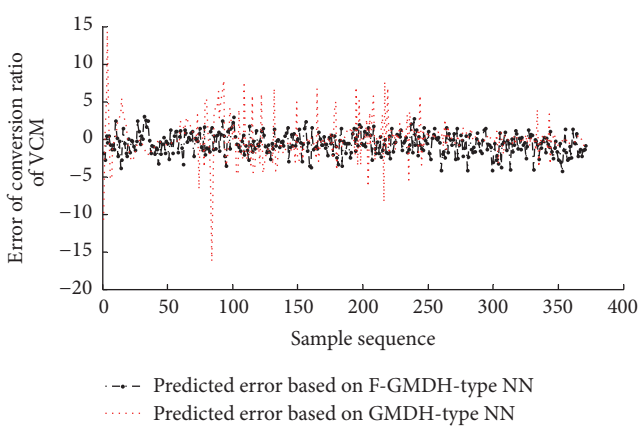

Figure 9: Predictive error of VCM conversion rate under GMDH-type NN and F-GMDH-type NN.

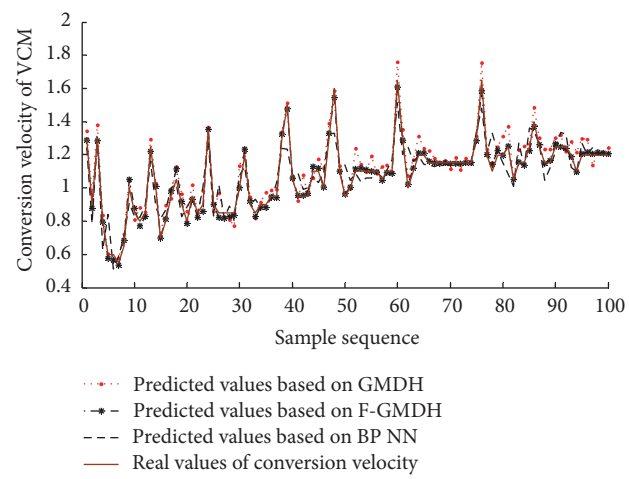

Figure 10: Predictive results of VCM conversion velocity under BPNN, GMDH-type NN, and F-GMDH-type NN.

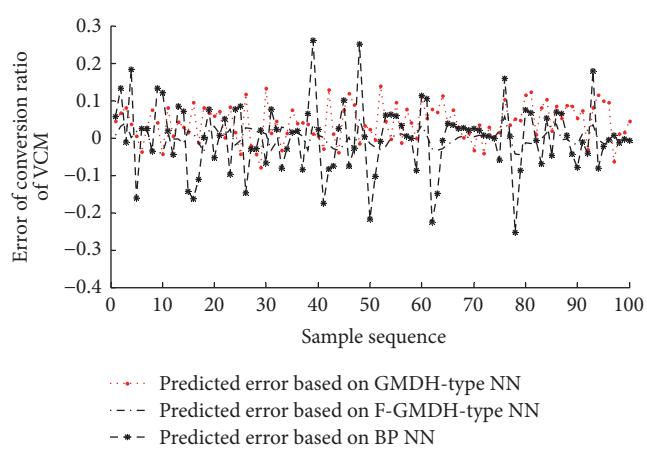

Figure 11: Predictive error of VCM conversion velocity under BPNN, GMDH-type NN, and F-GMDH-type NN.

the simulation results show that the soft-sensor model based on the proposed F-GMDH-type neural network has high prediction accuracy.

Competing Interests

The authors declare no conflict of interests.

Authors' Contributions

Most of Wei-zhen Sun's contributions were in the data collection, analysis, algorithm simulation, and the draft writing. Jie-sheng Wang's contributions were in the concept, design,

interpretation, and comments on the manuscript. Shu-zhi Gao's contribution was in the data collection and analysis of the manuscript.

Acknowledgments

This work is partially supported by the Project by National Natural Science Foundation of China (Grant no. 21576127), the Program for Liaoning Excellent Talents in University (Grant no. LR2014008), the Project by Liaoning Provincial Natural Science Foundation of China (Grant no. 2014020177), the Program for Research Special Foundation of University of Science and Technology of Liaoning (Grant no. 2015TD04), and the Opening Project of National Financial Security and System Equipment Engineering Research Center (Grant no. USTLKFGJ201502).

References

[1] S. Zhou, G. Ji, Z. Yang, and W. Chen, "Hybrid intelligent control scheme of a polymerization kettle for ACR production," *Knowledge-Based Systems*, vol. 24, no. 7, 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 Computing*, vol. 18, no. 3, pp. 457–468, 2014.

[3] W.-H. Cui, J.-S. Wang, and S.-X. Li, "KPCA-ESN soft-sensor model of polymerization process optimized by biogeography-based optimization algorithm," *Mathematical Problems in Engineering*, vol. 2015, Article ID 493248, 10 pages, 2015.

[4] S. Atashrouz, G. Pazuki, and Y. Alimoradi, "Estimation of the viscosity of nine nanofluids using a hybrid GMDH-type neural network system," *Fluid Phase Equilibria*, vol. 372, pp. 43–48, 2014.

[5] M. Najafzadeh and A. Zahiri, "Neuro-fuzzy GMDH-based evolutionary algorithms to predict flow discharge in straight compound channels," *Journal of Hydrologic Engineering*, vol. 20, no. 12, Article ID 04015035, pp. 1871–1881, 2015.

[6] R. Brito, S. Fong, K. Cho et al., "Towards implementation of residual-feedback GMDH neural network on parallel GPU memory guided by a regression curve," *Journal of Supercomputing*, vol. 72, no. 10, pp. 3993–4020, 2016.

[7] A. Sohani, H. Sayyaadi, and S. Hoseinpoori, "Modeling and multi-objective optimization of an M-cycle cross-flow indirect evaporative cooler using the GMDH type neural network," *International Journal of Refrigeration*, vol. 69, pp. 186–204, 2016.

[8] M. Masoumi Shahr-Babak, M. J. Khanjani, and K. Qaderi, "Uplift capacity prediction of suction caisson in clay using a hybrid intelligence method (GMDH-HS)," *Applied Ocean Research*, vol. 59, pp. 408–416, 2016.

[9] A. A. Beheshti and B. Ataie-Ashtiani, "Discussion of 'Neuro-fuzzy GMDH systems based evolutionary algorithms to predict scour pile groups in clear water conditions' by M. Najafzadeh," *Ocean Engineering*, vol. 123, pp. 249–252, 2016.

[10] O. Dag and C. Yozgatligil, "GMDH: an R package for short term forecasting via GMDH—type neural network algorithms," *The R Journal*, vol. 8, no. 1, pp. 379–386, 2016.

[11] S. Shahabi, M. J. Khanjani, and M. Hessami, "Significant wave height forecasting using GMDH model," *International Journal of Computer Applications*, vol. 133, no. 16, pp. 13–16, 2016.

[12] H. Liu, G. Wei, Z. Xu, P. Liu, and Y. Li, "Quantitative analysis of Fe and Co in Co-substituted magnetite using XPS: the application of non-linear least squares fitting (NLLSF)," *Applied Surface Science*, vol. 389, pp. 438–446, 2016.

[13] O. Jović, "Durbin-Watson partial least-squares regression applied to MIR data on adulteration with edible oils of different origins," *Food Chemistry*, vol. 213, pp. 791–798, 2016.