Hindawi
International Journal of Chemical Engineering
Volume 2022, Article ID 4133703, 9 pages
https://doi.org/10.1155/2022/4133703

Research Article

Research on Soft-Sensing Methods for Measuring Diene Yields Using Deep Belief Networks

Xiangwu Deng,[1] **Zhiping Peng**,[2] **and Delong Cui**[1]

[1]*College of Electronic Information Engineering, Guangdong University of Petrochemical Technology, Maoming 525000, China*
[2]*Jiangmen Polytechnic, Jiangmen 529000, China*

Correspondence should be addressed to Delong Cui; delongcui@gdupt.edu.cn

Received 21 November 2021; Revised 28 March 2022; Accepted 3 April 2022; Published 29 April 2022

Academic Editor: Jitendra Kumar

As an important raw material for the chemical industry, ethylene is one of the surest indicators that measure the development level of a country. The diene yield is an important production quality index parameter of ethylene units, and it is very important to detect and control them in real time. Due to the limitations of online analytical instrumentation technology, diene yields are difficult to measure online. Motivated by this, this article has studied soft-sensing technology for measuring diene yields. A diene yield prediction method based on a deep belief network algorithm network is proposed, and the regularity of historical diene yield data is fully explored by the method. First, the data feature vectors are fused and normalized. Then, the data are fed into a DBN consisting of two layers of restricted Boltzmann machines for unsupervised training, and finally, a DBN model is used to predict the diene yield. The experimental results show that the mean squared error of the test set with historical data is 1.15%, and the mean absolute percentage error of the measured data is 2.79%. The experimental results are provided to show the effectiveness of the proposed method.

1. Introduction

Ethylene was an important raw material for the chemical industry, and the industrial products made using ethylene already account for more than 75% of petrochemicals. Ethylene was one of the surest indicators that the petrochemical industry used as a tool to measure the development level of a country. In recent years, the Chinese ethylene industry had already made much progress. However, a gap in the technical level of ethylene existed between China and developed countries. For petrochemical production businesses, strengthening the quality was a long-term strategic task.

Productive enterprises implement active information technology to improve product quality, and advanced control strategies that had been applied in production plants have been introduced [1]. Under real-time monitoring of the critical parameters and effective control of the key manipulated variables, safety in the production process can be guaranteed, and the production efficiency of enterprises was improved. In most cases of actual industrial processes with cracking furnaces, the process involved complex physicochemical processes and conversion and delivery of energy and materials. Therefore, the actual production process had coupling, uncertainty, nonlinearity, hysteresis, and other characteristics. The result of all these factors was that some key process parameters were difficult to detect online. For ethylene production units, it was difficult to measure the yields of ethylene products and propylene products.

Ethylene and propylene were the main products of ethylene units. The diene yield was the sum of the yield of an ethylene product and a propylene product and was an important production and operational index of ethylene plants. In order to improve the competitiveness of China's ethylene industry and achieve product quality control and stabled production operations, it was necessary to develop the real-time detection and optimization control of the diene yield, an important production quality index of ethylene plants. At present, due to the limitations of online analytical instrument technology, it was difficult to measure diene yields online.

Based on the research and analysis of the cracking furnace process mechanism, the main factors affecting diene yields could be obtained. The model of diene yields in a cracking furnace could be established based on the main factors of the soft measurement online prediction. First, it was convenient to establish a diene yield prediction model based on a soft measurement method for industrial implementation. Different from online analysis instruments, the model did not require careful maintenance, and the online maintenance costs were lower. Second, the online estimation value could be given in real time, which overcomes the time delay of online analysis instruments. Therefore, the establishment of a diene yield soft measurement model was feasible and timely. The soft measurement model was gradually improved as the amount of data increases, and the model had a wide range of applications, small overall investment costs, easy online maintenance, large-scale promotion, and feasible applicability.

In order to realize the effective online measurement of some variables and parameters that were difficult to measure [2], such as ethylene and propylene yields in ethylene plants, many scholars had conducted in-depth research on software measurement technology. Soft measurement technology could estimate key quality indicators that could not be directly measured. Compared with traditional hard measurement methods, soft measurement had many advantages such as quick response, low costs, diversified methods, convenient maintenance, and safe operations.

Regarding building suitable soft measurement models, deep learning must be mentioned. In 2006, deep learning and the revolutionary layer-by-layer greedy training method were proposed, marking a new level of realizability of neural networks [3]. Developed from artificial neural networks, deep learning was experiencing its third wave in the twenty-first century. Deep learning had made remarkable breakthroughs in many fields such as computer vision, provided new feasible tools for solving old problems, and provided scholars with new hope and foci. However, in the process of using deep learning technology to build soft measurement models, many problems still need to be explored and solved.

Soft measurement was essentially a type of mathematical model. Therefore, the techniques and methods of signal processing, mathematical statistics, expert systems, and other research fields had been applied by many scholars and engineers in this field. After summarizing the previous work, the industrial process modelling methods were roughly divided into three categories: mechanism-based methods, knowledge-based methods, and data-driven methods [4, 5].

Mechanism-based methods, as the earliest developed process modelling methods, depicted and reflected the nature and mechanism of industrial processes by using accurate and analytical mathematical models [6–8]. Therefore, the mechanism model usually had high accuracy and reliability and was widely used in aerospace, automotive, precision instruments, and other fields where it was easy to accurately describe the process mechanism; however, for the process industry and other instances where it was difficult to obtain accurate mechanism knowledge, the model's effect was greatly reduced. Typical examples of such methods included the state estimation method, parameter estimation method, equivalent space method.

The knowledge-based method was applied to the existing production experience and expert knowledge where it qualitatively described each link and internal structure in the production process after inference and deduction, and then applied specific models to solve problems [9–11]. Therefore, in order to achieve a better effect, experts were often required to be very meticulous in place of understanding and rich practical experience. Additionally, considerable time and money must be spent to set up and update the maintenance expert knowledge base; nevertheless, the accuracy and reliability were still not effectively guaranteed, especially in complex industrial process modelling scenarios.

Due to the increasing complexity of industrial production equipment and related processes, it was very increasingly difficult to define process mechanisms and obtained reliable expert knowledge. Furthermore, with the advent of Industry 4.0 and the industrial big data era, a large amount of data could be obtained, and the modelling gains and advantages brought by data were becoming increasingly more obvious. Therefore, data-driven industrial process modelling methods were becoming mainstream [12–14]. The data-driven modelling method usually needs to reasonably select and improve the corresponding modelling method according to the actual data characteristics or process characteristics. The method used sufficient data to learn the parameters involved in the mathematical model and then obtained a practical soft measurement model. Classical data-driven methods included support vector machines [15], K-means clustering algorithm [16], and artificial neural networks.

Different from shallow learning, deep learning had an excellent nonlinear fitting ability and could fit complex functions. Common neural networks included stacked autoencoders (SAEs), deep belief networks (DBNs) [17], convolutional neural networks (CNNs) [18], and recurrent neural networks (RNNs). This type of method had also been gradually applied to industrial process monitoring and quality prediction. For example, a network structure was proposed based on self-organizing mapping (SOM) and compared it with the SVR and other traditional methods [19]. The results highlighted the superiority of the neural network method. The DBN was used to estimate the polymer melt in an industrial polymerization process [20, 21], which is suitable for applications to nonlinear data-driven models.

Due to the rapid dynamic response of soft measurement technology, continuous online measurement displays could be used without the aid of the analysis. In this article, a diene yield model for ethylene plants was established based on soft-sensing technology to realize the online measurement of diene products. Specifically, it included the following aspects. First, aiming at the production process of ethylene from naphtha pyrolysis as a raw material, this article proposed a software measurement method for diene yields in cracking furnaces based on the DBN algorithm. On the basis of fully mining the regularity in historical load data, data feature vectors were input into a DBN composed of a two-layer RBM for feature fusion. The DBN model was

used to predict the diene yield in a cracking furnace, and the model was pretrained by unsupervised training. Finally, the BP algorithm was used to calculate the diene yield in a cracking furnace.

2. Material and Methods

2.1. Ethylene Cracking Process. There were many ways to produce ethylene by cracking petroleum fractions, but pipeline cracking furnaces, which had the advantages of being mature technology, possessing a simple structure and good operating stability and providing high olefin yields, were currently the main way to produce ethylene in the world [22]. The entire structure of the tube cracking furnace was divided into two parts: a radiation chamber and a convection chamber. In the radiation chamber of a tube cracking furnace, many metal tubes were arranged in sequence. Uniformly distributed burners were designed outside the tube (including bottom burners and sidewall burners). The raw material passes through the metal tube quickly, and the heat was transferred to the material by the tube wall. The material in the tube cracking reaction occurred at a high temperature. The radiation chamber was arranged at the bottom of the cracking furnace, and it was mainly for cracking raw materials. The convection chamber was arranged above the radiation chamber, and it mainly used waste heat in flue gas to heat raw materials, steam, and other media. The feedstock entered the convection section from the top of the cracking furnace and was mixed with steam as a diluent after some preheating. After further preheating to the required temperature, the feedstock entered the radiation section furnace tube for high-temperature cracking. A schematic diagram of a cracking furnace structure is shown in Figure 1.

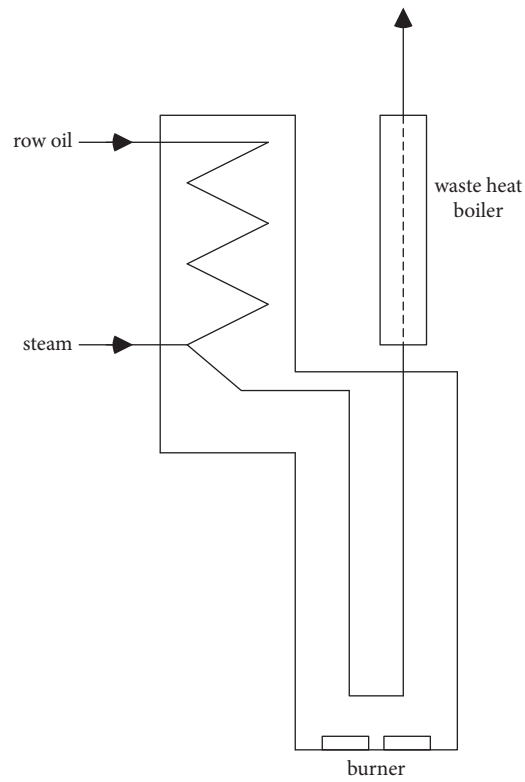

Figure 1: Schematic diagram of the structure of a cracking furnace.

2.2. Selection of Auxiliary Variables. Variables or parameters that could be easily measured were called auxiliary variables, and variables and parameters that could not be measured or were difficult to measure were called dominant variables. For the soft measurement model of diene yields, the dominant variable was the diene yield, which was the sum of the ethylene and propylene yields. The auxiliary variables were some influencing variables and parameters directly related to the diene yield.

In order to ensure the accuracy and validity of the soft sensor model, the information content of selected data covered a wide range when selecting production operating data. Generally, the input parameters of a model fall into two categories: one is the physical parameters of cracking raw materials, and the other is the process parameters of the cracking furnace. These two types of parameters commonly used in industry could be used as input parameters after screening and simplification, and the output parameters were the diene yield data of the cracking furnace that were hoped to obtain through the model.

2.2.1. Properties Indexes of Cracking Materials. Ethylene cracking furnace is mainly used to process natural gas, refinery gas, crude oil, and naphtha and other raw materials into cracking gas, and the study object is naphtha in this article. The quality of naphtha, feeding flux, and pyrolysis temperature are the major factors on the yield of naphtha pyrolysis products. The aim product of naphtha pyrolysis products is ethylene and the coproduct of that is propylene, so the diene yield of cracking furnace was the sum of the ethylene and propylene yields.

Quality of naphtha will affect the reaction process and the purity of product in cracking furnace, and when the hydrogen content of naphtha increased, the olefin of small molecules is produced more. The C/H ratio is an important index of the naphtha quality. Also, the lower the proportion of aromatic compounds, the higher the diene yield is. The main physical constants of raw material are density, distillation, and chemical admixture. The parameter of chemical admixture was sulfur content. The distillation of raw material includes initial boiling point (IP), 10% distillate temperature (t10), 30% distillate temperature (t30), 50% distillate temperature (t50), 70% distillate temperature (t70), 90% distillate temperature (t190), and end point (EP).

The main characterizations of cracking raw material quality are characterization factor K, PONA (paraffin, olefin, naphthene, aromatics), and Bureau of Mines Correlation Index (BMCI). The petroleum hydrocarbons can be scientifically classified into five groups: n-paraffins, i-paraffins, olefin naphthene, and aromatics. The characterization factor K represents chemical composition characteristic of crude oil and the fraction oil, and the computation methodology of K is shown in the following formula:

$$K = \frac{\sqrt[3]{t^3\,(^0R)}}{\rho}, \tag{1}$$

where ^{0}R represents the Rankine scale and ρ represents the density of naphtha, and the computation methodology of t is shown in the following formula:

$$t = \frac{\left(t_{10\%} + t_{30\%} + t_{50\%} + t_{70\%} + t_{90\%}\right)}{5}. \tag{2}$$

The parameter BMCI is the combinations of the relative density and the boiling point, which is shown in the following formula:

$$BMCI = \frac{48460}{t_1 + 273} + 473.6 \times d_{15.6}^{15.6} - 456.8, \tag{3}$$

where for a single hydrocarbon, t1 represents its boiling point, and for mixed hydrocarbons, t1 represents their average boiling point. $d_{15.6}^{15.6}$ represents the proportion of oils at 15.6°C (600F) to water at 15.6°C (600F).

2.2.2. Working Conditions of Cracking Furnace. If the raw material is determined in the cracking reaction, the main factors affecting the product yield are temperature, residence time, and hydrocarbon partial pressure. High cracking temperature, short residence, and low hydrocarbon partial pressure are conducive to the formation of diene products. With the increase in the pyrolysis temperature, the yield of propylene increased first and then decreased. The yield of propylene should be increased as much as possible while ensuring the yield of ethylene, so as to maximize the profit of the enterprise and obtain more profits. The furnace tube outlet temperature (COT) is usually replaced with cracking temperature. The residence time refers to the time when the cracking raw material passes through the furnace tube of the radiation section. At present, the residence time of tubular cracking furnace is generally between 0.1 s and 1 s. The secondary reaction and coking reaction will occur if the residence time of raw materials in the radiation section is too long. Therefore, the residence time should be shortened as far as possible. From the perspective of chemical equilibrium, the cracking reaction is a chemical reaction with an increase in the number of molecules. Reducing the hydrocarbon partial pressure of reactants is conducive to the formation of products. The furnace tube outlet pressure (COP) represents hydrocarbon partial pressure of tubular cracking furnace. Actually, adding steam into the raw material has many advantages. It can not only reduce the partial pressure and prevent the coking of the cracking furnace tube, but also protect the furnace tube, stabilize the cracking temperature, and remove the coking. The disadvantage of adding too much steam into the raw material is that it will make the subsequent quench operation more difficult and require greater heat load. The parameter of water oil ratio varies with different cracking raw materials, and the principle is to minimize the use of dilution steam under the requirement of preventing coking.

Seventeen physical properties of naphtha, including the density, distillation range, characterization factor K, PONA, C/H ratio, Bureau of Mines correlation index (BMCI), and sulfur content, were selected in this article. The distillation process contained the following 7 indicators: IP, t10, t30, t50, t70, t90, and EP. The group structure of petroleum hydrocarbons could be divided into several categories, namely, n-paraffins, i-paraffins, olefin naphthene, and aromatics. The above showed that there were 17 material property parameters selected by the model. The process conditions that had a significant influence on pyrolysis products and could be adjusted and controlled include COT, COP, water-oil ratio, and residence time, and these four industrial parameters were selected as input parameters. Therefore, a total of 21 input parameters, including 17 physical parameters and 4 technological parameters, were selected for the model.

2.3. Algorithm Flowchart. A static model DBN was trained by 120 sets of training data and tested by 30 sets of testing data. The prediction DBN model of the diene yield was trained by training data, and the prediction model was tested using testing data. The algorithm flow is shown in Figure 2.

(1) There were 21 auxiliary variables used to collect cracking furnace material characteristic parameters and industrial operating parameters as DBN input variables. The output variables were the sum of ethylene and propylene yields.

(2) In order to ensure the scale consistency of all data, feature vectors must be normalized, and all data after normalization were between [0, 1], as shown in the following formula:

$$x_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}}, \tag{4}$$

where $x_{\min}$ and $x_{\max}$ were the minimum and maximum eigenvectors in the feature set, respectively.

(3) The 4-layer DBN structure was trained by using 120 sets of historical operating data of a cracking furnace, and the fast-learning algorithm for contrastive divergence was used to conduct learning.

(4) The weight and bias obtained in the DBN training process were used to test the test set. According to the distribution of the RBM, the errors in the sample obtained by Gibbs sampling and thirty groups of measured data of a cracking furnace were evaluated, and the test results were obtained. In the regression prediction, the mean squared error was usually used as the loss function, as shown in the following formula:

$$MSE = \frac{1}{N}\sum_{n=1}^{N}\left(y - h(x)\right)^2, \tag{5}$$

where N represents the number of sample data points, y represents the expected value, and h(x) represents the estimates of the DBN model.

2.4. The Diene Yield Prediction Model of a Cracking Furnace. DBN could be regarded as a process of abstracting high-level features by combining various meaningful low-level features

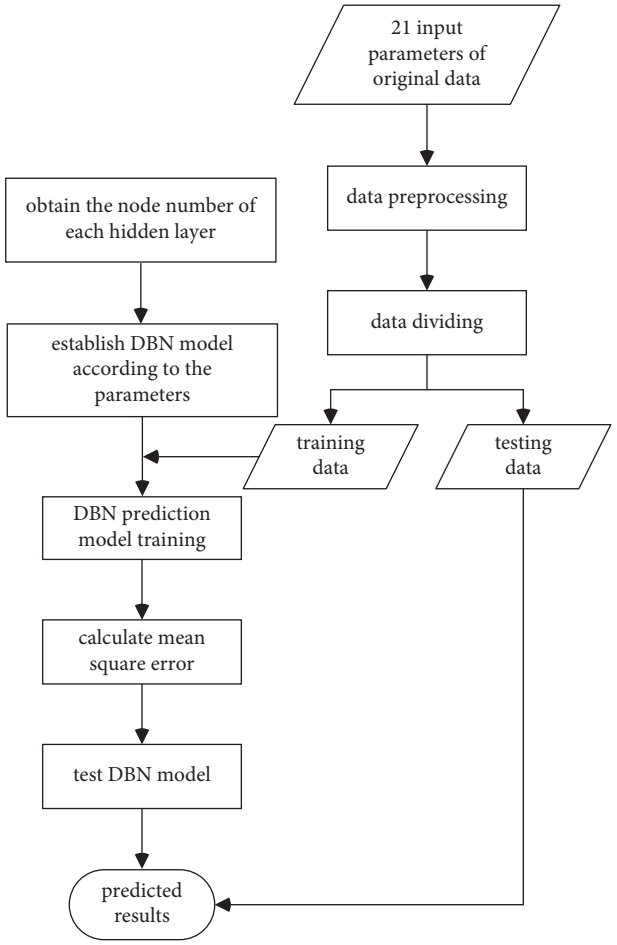

Figure 2: The algorithm flowchart.

and discovering distributed characteristics of data through deep learning. In fact, DBNs were probabilistic generative depth structure models, which had better effects in the probability estimation of prior data or posterior data. The core idea of a DBN was that each layer of a restricted Boltzmann machine extracted and abstracted input data from bottom to top and retained as much important information as possible.

2.4.1. Restricted Boltzmann Machine Model. A restricted Boltzmann machine (RBM) was used in a deep belief network [23], and understanding how an RBM works allowed one to understand how a DBN works. An RBM was a deep learning algorithm with a visual layer (denoted by v) and hidden layer (denoted by h), and the simplified model is shown in Figure 3. In an RBM, there was no correlation between nodes in the visual layer and hidden layer, and the conditions were independent of each other. Therefore, the total probability distribution of the visible layer and the hidden layer satisfied the following formula:

$$p\left(\frac{h}{v}\right) = p\left(\frac{h_1}{v}\right) = \cdots = p\left(\frac{h_n}{v}\right). \tag{6}$$

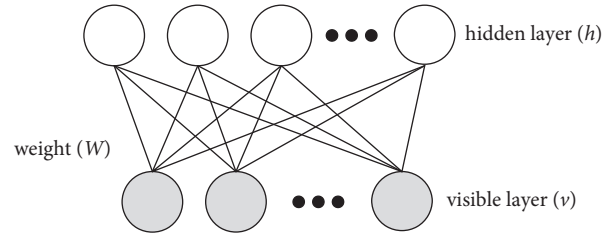

Figure 3: Restricted Boltzmann machine model.

The weight between the visible layer v and the hidden layer could be obtained by solving the minimization loss function, and the process was as follows. First, convert RBM data from the visual layer to the hidden layer. Then, set the visual layer node to vi, and set the hidden layer node to hj. The total number of visible layer nodes was *i*, the total number of hidden layer nodes was *j*, and the weight between the visible layer node and hidden layer node was Wij. The following equation also represents the energy of the entire combined configuration [24]:

$$E(v,h,\theta) = -\sum_{ij} W_{ij} v_i h_j - \sum b_i v_i - \sum a_j h_j, \theta = \{W, a, b\}, \tag{7}$$

where $\theta = \{W, a, b\}$ represents the parameters limiting the Boltzmann machine model. The joint probability distribution of the configuration could be determined by the Boltzmann distribution, where $Z(\theta)$ represents the distribution function and e^* represents the potential function:

$$\begin{aligned} p_\theta(v,h) &= \frac{1}{Z(\theta)} \exp\left(-E(v,h,\theta)\right) \\ &= \frac{1}{Z(\theta)} \prod_{ij} e^{W_{ij} v_i h_j} \prod_i e^{b_i v_i} \prod_j e^{a_j h_j}, \\ Z(\theta) &= \prod_{h,v} \exp\ \Delta - E\Delta v, h, \theta \Delta\Delta, \end{aligned} \tag{8}$$

For each node in the hidden layer, the related conditions of the joint probability distribution did not affect each other and existed independently, as shown in the following formula:

$$p\left(\frac{h}{v}\right) = \prod_j p\left(\frac{h_j}{v}\right). \tag{9}$$

As shown in formula (10), on the premise of a given visual layer *v*, the probability that the jth hidden layer node was 1 or 0 could be obtained by the factorization of Equation (9). As shown in the following formula, similarly, for a given visual layer *h*, the probability that the ith hidden layer node was 1 or 0 could be obtained [25]:

$$\begin{aligned} p\left(\frac{h_j = 1}{v}\right) &= \frac{1}{1 + \exp\left(-\sum_i W_{ij} v_i - a_j\right)}, \\ p\left(\frac{v}{h}\right) = \sum_i p\left(\frac{v_i}{h}\right), p\left(\frac{v_i}{h}\right) &= \frac{1}{1 + \exp\left(-\sum_i W_{ij} h_j - b_i\right)}. \end{aligned} \tag{10}$$

If the sample set $Test = \{v^1, v^2, \ldots v^N\}$ satisfied the independent distribution condition, the parameter $\theta = \{W, a, b\}$ could be obtained by learning. Here, the maximum-likelihood estimation was used to select parameters to maximize the observed sample probability, as shown in the following formula:

$$L(\theta) = \frac{1}{N}\sum_{n=1}^{N} \log P_\theta\left(v^{(n)}\right) - \frac{\lambda}{N}\|W\|_F^2. \quad (11)$$

Take the derivative of the above formula to obtain the parameter value W with the maximum sample probability, as shown in the following formula:

$$\frac{\partial L(\theta)}{\partial W_{ij}} = E_{P_{data}}\left[v_i h_j\right] - E_{P_0}\left[v_i h_j\right] - \frac{2\lambda}{N} W_i j, \quad (12)$$

where $E_{P_{data}}$ presents the expected value of the training sample set and E_{P_0} presents the expected value defined in the model.

In summary, the above equations gave the general process of the RBM algorithm. The ability to initialize the weight W between the visible layer and the hidden layer was one of the advantages of the restricted Boltzmann machine algorithm. In addition, in order to minimize the local minimum probability, the gradient descent method was used to train the relevant parameters of the experimental data.

2.4.2. Deep Belief Network Model. A deep belief network was a structural model composed of multiple RBMs; therefore, the deep belief network had the same structure as the restricted Boltzmann machine. The weight layer W was used to associate layers with each other, and the nodes within the layer were independent of each other. According to the weights generated from bottom to top, the connection mode of the deep belief network structure model could be determined, which was conducive to learning the weight W. This operation was called the "contrastive bifurcation method," and the weight W of the model was obtained first. Unsupervised greedy learning was conducted on training layer by layer, and then, the BP algorithm was used to adjust and update the weight of each layer. This method was also proven to be effective by Hinton.

In the process of data training following Gibbs, the extracted feature data were used as the input data of the visual layer in order to obtain parameter vector v. The hidden layer connected to the visual layer would obtain the vector values passed in. The hidden layer reorganized the visible layer node through the obtained information and then returned the updated visible layer node to the first hidden layer node. Thus, the hidden layer nodes were updated. The difference between the input data of the visual layer and the data of the hidden layer was compared as the updating condition of the connection weight.

The practice proved that the above training methods could not only greatly reduce the time spent on sample training but also improved the feature learning ability of training data when the number of relevant layers in the neural network increases. However, a large number of layers meant considerable training and learning time. Therefore, the necessary condition for obtaining the prediction accuracy of the optimal performance ratio for the deep belief network structural model was to verify the optimal number of levels through a large number of experiments. The important step of the deep belief network after data pretraining learning was to update the weight of the training sample set with the BP.

When using BP algorithm in the DBN, the learning ability and prediction performance were improved, and the training time was shortened.

2.4.3. Diene Yield Prediction Model Based on a DBN. Considering the accuracy and time complexity of the prediction, the model could guarantee the high accuracy and timeliness of the prediction of diene yields. Combined with the actual experimental conditions, a deep belief network structure model containing two RBM layers was adopted. The first layer was the data input layer, and the fourth layer was the data output layer. Since 21 auxiliary variables were entered, the number of visible layer nodes was set to 21. The number of hidden layer nodes in RBM1 and RBM2 was set to 44 and 22, respectively. Therefore, the number of nodes in the last layer was set to 1, which was the diene yield. The deep belief network structure is shown in Figure 4.

3. Results and Discussion

3.1. The Parameter Optimization of Hidden Layer Nodes in the DBN. As the number of hidden layer nodes in the DBN with two hidden layers needed to be set as 2, the number of hidden layer nodes had three changing trends: rising, falling, and constant. Therefore, this article chose the three combination methods for the number of hidden layer nodes of increasing, decreasing, and constant value to optimize the parameters of the DBN model with two hidden layers. A constant value meant that the number of hidden layer nodes was constant, increasing meant that the number of hidden layer nodes increases as the number of hidden layers increases, and decreasing meant that the number of hidden layer nodes decreases as the number of hidden layers increases.

The number of iterations was set to 250, the sample batch size was 20, and the learning rate was 0.001. The average value of 10 operations was taken as the test result. The identification accuracy of the DBN model for the three combination modes of the number of hidden layer nodes is shown in Table 1. The recognition accuracy of the constant and increasing methods was not as good as the decreasing method, which indicated that the decreasing model with a decreasing number of hidden layer nodes could better learn the distributed features of the original feature data. When selecting the number of hidden layer nodes for the decreasing method, too few (8–4) or too many (62–31) hidden layer nodes caused the model parameters to be unable to be fully trained and the recognition accuracy to be low. When the number of hidden layer nodes was within one, the recognition accuracy of the model was improved as the number of hidden layer nodes increases.

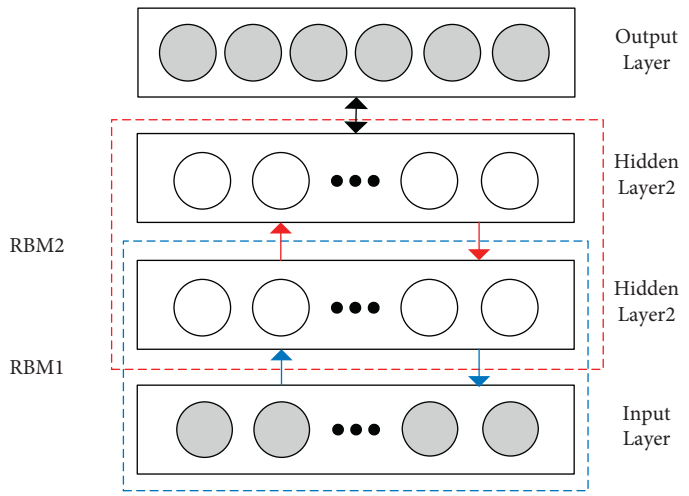

Figure 4: Deep belief network structure.

Table 1: Mean square error of different node types in the hidden layer of double hidden layers.

Group	Constant type	MSE (%)	Rising type	MSE (%)	Descending type	MSE (%)
1	8–8	21.40	8–12	30.16	8–4	9.17
2	14–14	28.78	14–21	21.14	14–7	6.38
3	20–20	22.81	20–30	18.35	20–10	5.41
4	26–26	18.27	26–39	15.27	26–13	3.35
5	32–32	16.32	32–48	18.35	32–16	3.26
6	38–38	13.17	38–57	20.78	38–19	2.29
7	44–44	15.46	44–66	21.14	44–22	1.15
8	50–50	20.25	50–75	23.93	50–25	3.24
9	56–56	22.84	56–84	25.64	56–28	4.19
10	62–62	24.69	62–93	27.18	62–31	8.63

As shown in Table 1, to determine the optimal number of hidden layer nodes of the DBN with two hidden layers, the number of hidden layer nodes in the first layer was successively set at a fixed interval of 6 within the interval [8, 62]. The number of hidden layer nodes in the second layer was calculated as an integer value according to the number of hidden layer nodes in the first layer: it was set the same as the number of hidden layer nodes in the first layer (constant), two-thirds the number of hidden layer nodes in the first layer (increasing), and half the number of hidden layer nodes in the first layer (decreasing). Through the experimental verification of the combination of different numbers of hidden layer nodes for two hidden layers, it was finally obtained that when the structure of the DBN with two hidden layers was [21, 44, 22, 1], the model has the lowest mean squared error of 1.15%.

In this article, the DBN structure included two RBM layers. The input layer and hidden layer constituted RBM1, while the hidden layer and output layer constituted RBM2. The pretraining of the RBM in each layer of the network was completed through unsupervised learning, and the training results were used as the input of the RBM at the higher layer. Finally, the network parameters of all RBM layers were adjusted through supervised learning. As shown in Figure 5, the training error rate decreased significantly in the first 20 iterations of the training of the first RBM layer. However,

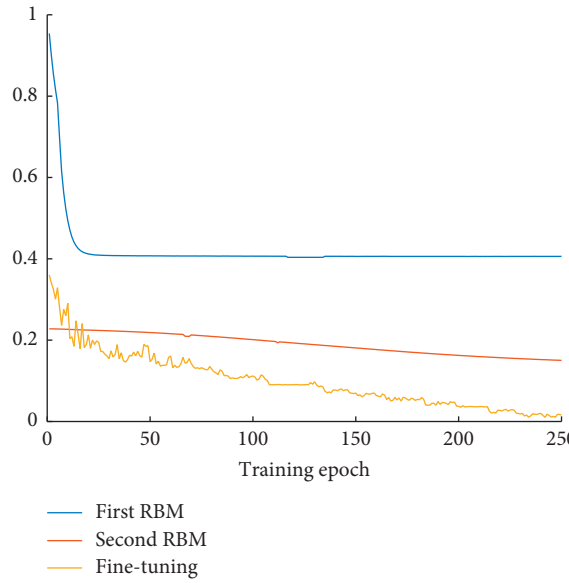

Figure 5: Mean squared errors of the DBN.

after 25 training iterations, the error rate did not change, and the parameter values in the RBM tended to be stable. The parameters in the second RBM layer were trained on the basis of the parameters in the first RBM layer. The variation

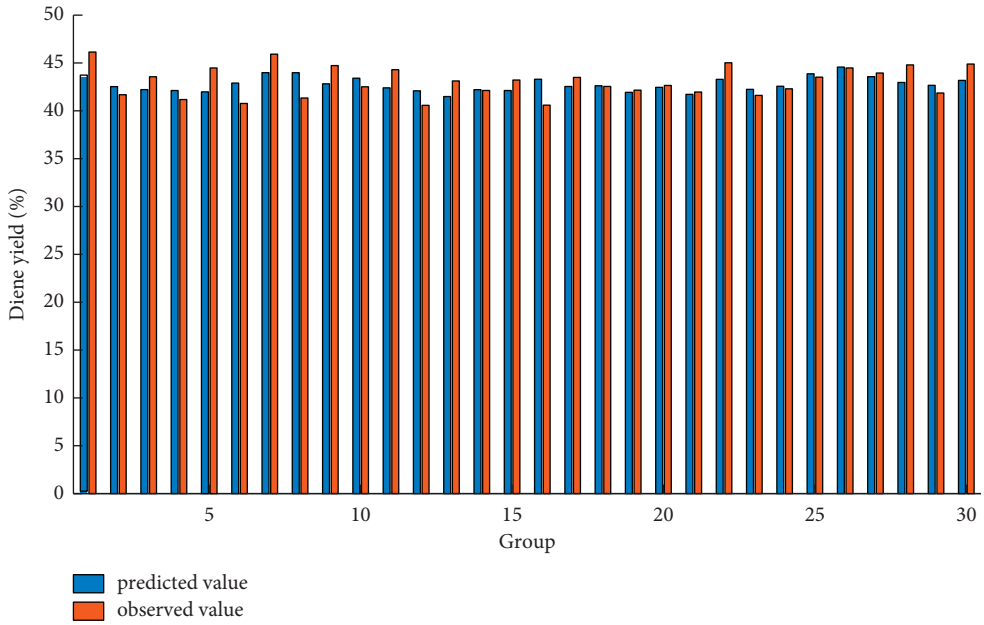

Figure 6: Verification of diene yield model.

in the parameters was relatively gentle, and the model tended to be stable. Finally, the supervised network parameters were fine-tuned and classified to obtain the final training results.

Because the DBN was a data-driven model structure, there was an optimal network structure corresponding to different data samples. In this article, based on the DBN structure with two hidden layers, the sample space of 21 dimensions of auxiliary variables of a cracking furnace was considered. The number of hidden layer nodes was optimized. The experimental results showed that the structure of the DBN with two hidden layers that obtained the lowest mean squared error was [21,44, 22,1]. The mean squared error on the test set with historical data of thirty groups was 1.15%, and the verification of the diene yield model is shown in Figure 6.

3.2. Comparison Models. In this article, the SVM, BP artificial neural network, and DBN prediction models were constructed based on the characteristics of 21 auxiliary variables of a cracking furnace. The models were evaluated and compared using the mean absolute percentage error (MAPE). Because of its stability, the MAPE could be used as a benchmark for many evaluation criteria. The calculation formula is shown as follows:

$$MAPE = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{T - h(x)}{T}\right|, \quad (13)$$

where N denotes the number of sample data points, T represents the actual value, and h(x) is the estimates of the DBN model.

From Table 2, we can see that the prediction error of the diene in a cracking furnace based on the DBN model was lower than that of the SVM model and BP model. Because a DBN was a network structure for mining the distributed characteristics of data, it could better adapt to high-dimensional data structures and mined the distributed characteristics of data. The experimental results showed that the mean absolute percentage error of the measured data was 2.79%. The test times of the SVM model, BP model, and DBN model were 0.0186 s, 0.0201 s, and 0.0237 s, respectively, for the three models of the test samples of the 21 auxiliary variables of the cracking furnace. The test results showed that the test time based on the DBN model was longer than that of the other two models.

Table 2: Prediction time and MAPE of measured data.

Model	Predicted time (s)	MAPE (%)
SVM	0.0186	6.39
BP	0.0201	5.15
DBN	0.0237	2.79

With the consideration of the weakness of the soft-sensing method based on the BP network, SVM, such as uncontrollable convergence speed and local minima, exponential growth of amount of calculation with increasing data samples, the DBN was introduced into soft-sensing methods and then was studied. The DBN uses latent variables to express process variables with high correlation; thus, DBN has a good expression ability. The soft-sensing method based on DBN was applied in a real diene yield production process for the first time. Compared with the soft-sensing method based on BP and SVM, the generalization ability and the accuracy were improved.

4. Conclusions

According to the physical parameters of the cracking raw material and the characteristics of the cracking furnace process parameters, a total of 21 input parameters, including

17 physical parameters and 4 technological parameters, were selected for the model. The DBN model was established and optimized, and the prediction effect of the DBN model with two hidden layers was explored by selecting different node combinations. The structure of the DBN with two hidden layers that obtained the lowest MSE on the test set of 1.15% was [21, 44, 22,1]. Three prediction models, namely, the SVM, BP, and DBN, were constructed according to the measured sample data of a cracking furnace. By comparing MAPE and prediction times, the DBN-based model had the lowest MAPE but the longest recognition time. The DBN model was the best, and the lowest mean squared error of the measured data was 2.79%.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request. The raw/processed data required to reproduce these findings cannot be shared at this time as the data also form part of an ongoing study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was funded by the National Natural Science Foundation of China (61772145 and 61672174), Key Platform and Scientific Research Project of Guangdong Education Department (2017KTSCX128), Guangdong Basic and Applied Basic Research Foundation (2021A1515012252 and 2020A1515010727), Key Field Special Project of Department of Education of Guangdong Province (2020ZDZX3053), and Maoming Science and Technology Project (mmkj2020008 and mmkj2020033).

References

[1] L. Z. Xia, *Research on Soft-Sensing Modeling Method of Online Estimation Based on Vinyl Acetate Polymerization Rate*, Lanzhou Jiaotong University, Lanzhou, China, 2012.

[2] H. Cui, Y. Guan, H. Chen, and W. Deng, "A novel advancing signal processing method based on coupled multi-stable stochastic resonance for fault detection," *Applied Sciences*, vol. 11, no. 12, p. 5385, 2021.

[3] G. E. Hinton, S. Osindero, and Y. W. Teh, "A fast learning algorithm for deep blief nets," *Neural Computation*, vol. 18, no. 7, pp. 1524–1554, 2006.

[4] V. Venkatasubramamian, R. Rengaswamy, and S. N. Kavuri, "A review of process fault detection and diagnosis: Part 2nd: qualitative models and search strategies," *Computers & Chemical Engineering*, vol. 27, no. 3, pp. 313–326, 2003.

[5] P. M. Frank, "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy," *Automatica*, vol. 26, no. 3, pp. 459–474, 1990.

[6] P. Young, "Parameter estimation for continuous-time models-A survey," *Automatica*, vol. 17, no. 1, pp. 23–39, 1981.

[7] R. Isermann, "Process fault detection based on modeling and estimation methods-A survey," *Automatica*, vol. 20, no. 4, pp. 387–404, 1984.

[8] E. Chow and A. S. Willsky, "Analytical redundancy and design of robust failure detection system," *IEEE Transactions on Automatic Control*, vol. 29, no. 7, pp. 603–614, 1982.

[9] P. V. Suresh, A. K. Babar, and V. V. Raj, "Uncertainty in fault tree analysis: a fuzzy approach," *Fuzzy Sets and Systems*, vol. 83, no. 2, pp. 135–141, 1996.

[10] S. H. Liao, "Expert system methodologies and applications – a decade review from 1995 to 2004," *Expert Systems with Applications*, vol. 10, no. 5, pp. 471–486, 2005.

[11] M. Modarres and S. W. Cheon, "Function-centered modeling of engineering systems using the goal tree-success tree technique and functional primitives," *Reliability Engineering & System Safety*, vol. 64, no. 2, pp. 181–200, 1999.

[12] L. Yao, *Distributed Parallel Modeling and Application for Industrial Processes with Large-Scale Data*, Zhejiang University, University, Zhejiang, China, 2019.

[13] J. L. Zhu, *Robust Monitoring of Industrial Process with Data-Driven Methods*, Zhejiang University, University, Zhejiang, China, 2016.

[14] Q. Sun and Z. Ge, "A survey on deep learning for data-driven soft sensors," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 9, pp. 5853–5866, 2021.

[15] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.

[16] X. Ran, X. Zhou, M. Lei, W. Tepsan, and W. Deng, "A novel K-means clustering algorithm with a noise algorithm for capturing urban hotspots," *Applied Sciences*, vol. 11, no. 23, Article ID 11202, 2021.

[17] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[18] Y. Taigman, M. Yang, M. Ranzato, and W. L. DeepFace, *Closing the Gap to Human-Level Performance in Face Verification. Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE Computer Society, 2014.

[19] V. Gopakumar, S. Tiwari, and I. Rahman, "A deep learning based data driven soft sensor for bioprocesses," *Biochemical Engineering Journal*, vol. 136, pp. 28–39, 2018.

[20] C. H. Zhu and J. Zhang, "Developing soft sensors for polymer melt index in an industrial polymerization process using deep belief networks," *International Journal of Automation and Computing*, vol. 17, no. 1, p. 11, 2020.

[21] C. H. Zhu and J. Zhang, *Developing Robust Nonlinear Models through Bootstrap Aggregated Deep Belief Networks. 25th International Conference on Automation and Computing*, (ICAC), 2019.

[22] Z. X. Wu, *Stress Analysis of Ethylene Cracking Furnace Pope System with Environmental Factors. Master*, China University of Mining & Technology, Xuzhou, China, 2019.

[23] T. Kuremoto, S. Kimura, K. Kobayashi, and M. Obayashi, "Time series forecasting using a deep belief network with restricted Boltzmann machines," *Neurocomputing*, vol. 137, pp. 47–56, 2014.

[24] G. Hinton, "A practical guide to training restricted Boltzmann machines," *Momentum*, vol. 9, no. 1, pp. 926–947, 2010.

[25] G. Hinton, L. Deng, D. Yu et al., "Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.