

Research Article

Data-Driven Optimal Control for Pulp Washing Process Based on Neural Network

Wenjuan Shan ^{1,2} and Wei Tang²

¹School of Light Industry Science and Engineering, Shaanxi University of Science & Technology, Xi'an 710021, China

²School of Electrical and Information Engineering, Shaanxi University of Science & Technology, Xi'an 710021, China

Correspondence should be addressed to Wenjuan Shan; swj0230@126.com

Received 25 September 2019; Revised 18 December 2019; Accepted 17 January 2020; Published 22 February 2020

Academic Editor: Francesc Pozo

Copyright © 2020 Wenjuan Shan and Wei Tang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Pulp washing process has the features of multivariate, time delay, nonlinearity. Considering the difficulties of modeling and optimal control in pulp washing process, a data-driven operational-pattern optimization method is proposed to model and optimize the pulp washing process in this paper. The most important quality indexes of pulp washing performance are residual soda in the washed pulp and Baume degree of extracted black liquor. Considering the difficulties of modeling, online measurement of these indexes, two-step neural networks, and multivariate logistic regression are used to establish the prediction models of residual soda and Baume degree. The mathematical model of the washing process can be identified, and the indexes can meet the production requirements. In the target of better product quality, low cost, and low energy consumption, a multiobjective problems is solved by ant colony optimization algorithm based on the optimized operational-pattern database. It shows that the theoretical analyses are correct and the practical applications are feasible, optimization control system has been designed for the pulp washing process, and the practical results show that pulp production increased by 20% and water consumption decreased by nearly 30%. This method is effective in the pulp washing process.

1. Introduction

Pulp washing is one of the important parts in the pulp-making process. Its main purpose is to wash soluble inorganic and organic substances in pulp and obtain high concentration of black liquor [1]. The most important quality indexes to evaluate pulp washing performance are residual soda for the washed pulp and Baume degree for the yielded thick black liquor. The conductivity of the residual soda extracted from the washed pulp is a measurement of dissolved solid content; It is expected that the residual soda in the washed pulp should be as little as possible. The Baume degree is used to denote the consistency of the thick liquor obtained in the first stage of the pulp washer [2]; it is required to be as high as possible in production. It can be seen that the two indexes are incompatible. From the perspective of pulp washing, it is required much more water to wash the pulp, but the black liquor will be diluted by the increase of water. From the perspective of drug recovery, in order to

reduce the steam consumption of black liquor during evaporation and condensing procedure [3], a higher concentration and temperature of black liquor is required, so the washing water should be used as less as possible. For these two contradictory requirements, the key point is that a balance must be maintained between the amount of washing water and desired pulp cleanliness.

To address these issues, several classical control schemes are produced. Black liquor consistency control, residual soda loss control, dilution factor optimization control, multi-component control, and model-based optimization control were produced. As the loss of residual soda is not taken into account, the method of Baume degree online measurement is only implemented in some small production mills [4]. Aiming at the disadvantage of single control, a residual soda predictor is constructed based on the predictive control [5]. As it is limited to the measurement and control of a certain parameter, more attention should be given to the modeling and optimization control for the whole washing process.

Model-based optimal control and multicomponent control [6] play a great role in the optimization of operation and control. Model-based optimization control can analyze the relationship between variables according to the model. In recent years, data-driven modeling has been widely used. In this method, based on online learning and calculating the control quantity, the current state is matched by a large amount of process data. Then various static qualities required by the system can be obtained. The characteristics of sample data are the main criterion in data-driven modeling, which means that data speak for themselves [7–9]. It can convert high-dimensional data into low-dimensional data without losing important information. Pulping enterprises produce and store a large amount data about production parameters, equipment, and process data every day; those data imply the process change, equipment operation, working condition fluctuation, and other information. In the washing process, there is a wealth of online and offline measurement data, such as temperature, washing drum pressure difference and vacuum degree, pulp layer thickness, pulp concentration, pulp species, pulp hardness, the amount of fresh water added, and the times of washing.

In view of this, aiming at the characteristics of the pulping process, a data-driven operational-pattern optimization method to model and optimize the pulp washing process is proposed in this paper. Based on mechanism analysis, the basic concepts about data-driven operational pattern for the pulp washing process are described. The data-driven prediction models of residual soda and Baume degree are established by PCA-BP (principal component analysis backpropagation) neural network and multivariate logistic regression. Based on the modeled indexes, an overall evaluation model of washing quality is proposed to analyze the large number of industrial operation data. Then, an optimized operational-pattern database is constructed by the criterion of fuzzy cluster and pattern matching algorithm. Finally, with the target of better product quality, low cost, and low energy consumption, the optimum operational pattern is obtained by ant colony optimization algorithm from the optimized operational-pattern database. Based on the hardware of Siemens S7-400 PLC and software of WinCC6.0 & Step7, an optimization control system has been designed for the pulp washing process. A practical application in a paper mill in Shandong province of China shows that pulp production increased by 20% and water consumption decreased by nearly 30%. It proves the effectiveness of this method.

The main contribution of this paper is to solve the soft sensor modeling problem and optimize the pulp washing process. The soft sensor model of the residual soda and the Baume degree were obtained by two-step neural network. The multiobjective optimization of high yield, low cost, and low consumption in the pulp washing process was realized. A systematic and targeted optimization scheme of the pulp washing process was provided. The rest of the paper is organized as follows. Section 2 illustrates the crafts of the pulp washing process. The two-step neural network identification and mathematical model of residual soda and Baume degree based on least square

method are given in Section 3. Multiobjective optimization for the washing process is given in Section 4. Optimization results and the application of this study in a paper mill are described in Section 5. Section 6 concludes the paper.

2. Crafts and Mode Description of the Pulp Washing Process

2.1. Crafts of the Pulp Washing Process. Currently, multi-stage countercurrent washing has been widely used in multiple series washing machines (vacuum washing machines) or one separation into multiple washing machines (horizontal belt washing machines). In this device, the pulp is discharged from the last section, and washing water is opposite to the last section. The reverse washing process can be briefly described as shown in Figure 1.

In this sequence of washers, the pulp medium and the washer are generally arranged to flow counter current to each other. Fresh water is typically used to wash the pulp sheet on the last stage washer. The filtrate was pulled through pulp on the proceeding washer. This aids in minimizing dilution of the liquor which is separated from the pulp and facilitates the formation of recyclable cooking chemicals.. Hence, it is desirable that the washing losses be kept as low as possible while using a minimum of washing water.

2.2. Operation Mode Description of the Pulp Washing Process. In the pulp washing control system, six different washing stages are required to obtain the necessary pulp cleanliness. Subsystems of consistency, level, flow, temperature, drum rotation speed, and multiobjective optimization are involved. The former five subsystems are used as a basic control hierarchy to guarantee the steady running of the counter current washing process, meanwhile, which ensures the washed pulp with a residual soda as low as possible and thick black liquor in the first stage filtrate tank with a Baume degree as high as possible.

The data of complex industrial processes mainly include input conditions, state parameters, operation parameters, and technical index [10, 11]. For the pulp washing process, the process data can be defined according to the operation mode description as follows:

- (i) The input condition of the pulp washing process at time t is $r(t) = [r_1(t), r_2(t), r_3(t), r_4(t), r_5(t)]$, where $r_1(t)$, $r_2(t)$, $r_3(t)$, $r_4(t)$, and $r_5(t)$ indicate pulp species, pulp thickness, pulp hardness, cooking method, and washing times, respectively.
- (ii) The state parameters of the pulp washing process at time t are $s(t) = [s_1(t), s_2(t)]$, where $s_1(t)$ and $s_2(t)$ indicate the Baume degree of black liquor and residual soda, respectively.
- (iii) The operation parameters of the pulp washing process at time t are $q(t) = [q_1(t), q_2(t), q_3(t)]$, where $q_1(t)$, $q_2(t)$, and $q_3(t)$ indicate pulp concentration, pulp flow, and water flow, respectively.

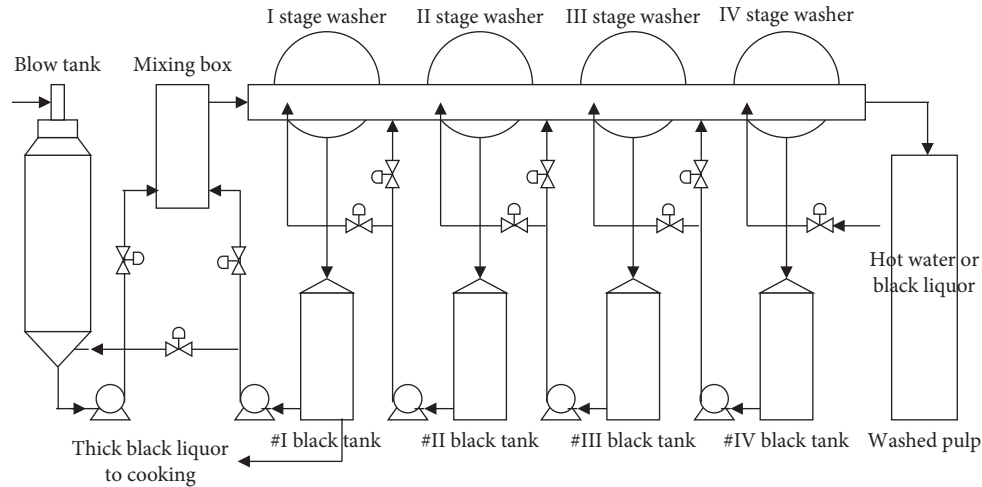


FIGURE 1: Washing process craft of straw brown stock.

- (iv) The technical index of the pulp washing process at time t is $o(t) = [o_1(t), o_2(t), o_3(t)]$, which indicates the target required to be achieved in the production process, where $o_1(t)$, $o_2(t)$, and $o_3(t)$ indicate pulp washing cleanliness, pulp output, and consumption cost, respectively.

Finally, the operation mode of the pulp washing process is established, and it can be described as $p = [r^T, s^T, q^T]^T = [r_1, \dots, r_5, s_1, s_2, q_1, \dots, q_3]^T$. It is construed by input condition, state parameters, and operation parameters.

3. Two-Step Neural Network Modeling

The core idea of the operation optimization of the pulp washing process is to optimize the pulp washing quality. As the two contradictory pulp washing quality indexes (residual soda and the Baume degree of black liquor) are difficult to measure directly, a prediction model should be established. Neural network has a very prominent effect on modeling, classification, and prediction of nonlinear systems [12]. The neural network modeling method was used to establish the network structure between the main components and quality variables, and the soft measurement model of the pulp washing process was obtained.

Generally speaking, the result of washing is a function of many variables, such as the amount of wash water, feed consistency, air entrancement, sheet formation, wash water distribution, and discharge consistency. Most of these variables are interrelated, and an improvement in one variable may well have a favorable or unfavorable impact on the others. The primary objectives for achieving the best washing result are the lowest feed consistency, optimum mat formation with a uniform basis weight, a uniform shower liquor distribution, the highest discharge consistency, and minimal air content in the feed to the washer [13]. Pulp washing is a function of many variables, such as the amount of wash water, feed consistency, air entrancement, sheet formation, wash water distribution, and discharge consistency. And most of these variables are interrelated. The PCA

method was used to preprocess the process variables involved in noise reduction, dimension reduction, and negative correlation elimination [14]. It helps to reduce the complexity of neural network.

3.1. Input Variable Selection Based on PCA. A large number of experiments have shown that the factors influencing the washing quality of pulp include sizing concentration sizing flow rate, water addition, thickness of pulp layer, vacuum degree, pulp species, washing water temperature, pulp hardness, and pulping method. The dynamic PCA method was used to screen out these variables for establishing the mathematical model of the pulp washing process. The calculation steps are as follows:

Step 1. Initialize the sample database.

In order to fully reflect the dynamics of the process, some appropriate process data are added and a new sample database is set up. Standardize the new sample database and calculate the correction variance every time, until the mean variance no longer changes; the correction value of each variable can be obtained, and then the initialized sample database is established.

Step 2. PCA for the sample database.

The load matrix and eigenvalue vector were obtained by PCA of pulp concentration, flow, dilution water, pulp thickness, water temperature, and pulp hardness. Find and eliminate the variable corresponding to the largest absolute value of coefficient in the first load vector. Repeat the process until just three variables are left (the number of principal variables was determined by significant correlation [15]; given the significance level $\alpha = 1\%$, the critical value of correlation coefficient significance is 0.487; find the largest absolute value of each column from the component matrix and compare with 0.487; when the largest absolute value in the $k + 1$ column is less than 0.487, k principal components can be extracted; in this paper, $k = 3$).

Step 3. According to multiple analyses, the contribution of the three components reaches 98% and inlet pulp consistency, inlet pulp flow, and hot clean water input flow are considered to be the most important variables in pulp washing.

The amount of residual soda in the final washed pulp and Baume degree in the first stage filtrate tank have typically been used as output indicators in pulp washing.

Therefore, the input and output of the pulp washing process can be described as $X = [x_1, x_2, x_3]$ and $Y = [y_1, y_2]$.

3.2. Two-Step Neural Network. The pulp washing process is not a steady-state model which can be accurately identified. In this paper, a two-step identification method is employed. The basic idea of this method can be illustrated as follows.

Firstly, by means of the dynamic information of a process (for example, the step response), a dynamic model is obtained by a multilayer forward neural network. Secondly, with the obtained dynamic model, a steady-state data collection is produced as the training sample for the final steady model. Finally, another multilayer forward NN is trained to approximate the steady characteristic of the process and thus acquire the final steady NN model of the process. The schematic diagrams of this method are shown in Figures 2 and 3. The abbreviation of TDL (tapped delay line) [16] means a delay unit according to sampling period.

3.2.1. Neural Network Dynamic Model Identification. The dynamic NN mathematical model describing the pulp washing process is

$$\begin{aligned} \bar{y}_1 &= h_1(x_1(k), x_1(k-1), \dots, x_1(k-d), x_2(k), x_2(k-1), \dots, x_2(k-d), x_3(k), x_3(k-1), \dots, x_3(k-d), y_1(k), \\ &\quad y_1(k-1), \dots, y_1(k-r_1)), \\ \bar{y}_2 &= h_2(x_1(k), x_1(k-1), \dots, x_1(k-d), x_2(k), x_2(k-1), \dots, x_2(k-d), x_3(k), x_3(k-1), \dots, x_3(k-d), \\ &\quad y_2(k), y_2(k-1), \dots, y_2(k-r_2)), \end{aligned} \quad (1)$$

where d is the length of input samples ($d = 3$) and r_1 and r_2 are delay factors of residual soda and Baume degree, respectively. According to a one-year field survey conducted in the pilot paper mill, the two delay factors are determined as $r_1 = 8$ and $r_2 = 6$.

Substitute ($d = 3, r_1 = 8, r_2 = 6$) into equation (1). Then, the number of independent variables for \bar{y}_1 is 21 and the number of independent variables for \bar{y}_2 is 19. According to the actual pulp washing process, residual soda has more dynamic influence factors than Baume degree. By training the neural networks, the number of dynamic independent variables for residual soda and Baume degree is 24 and 16; they would be the input layers for the both neural networks. The structure parameters of the dynamic NN model for residual soda and Baume degree are given in Table 1.

The identifying procedures are depicted as follows:

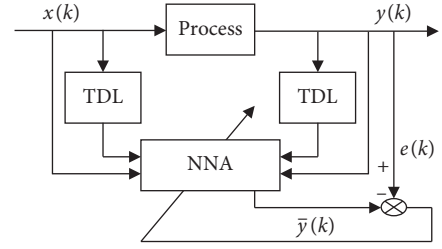


FIGURE 2: NN dynamic model identification.

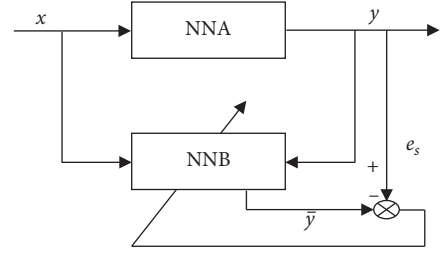


FIGURE 3: NN stationary model identification.

TABLE 1: Neural network dynamic model parameters of residual soda (RS) and Baume degree (BD).

	Input layer	Hidden layer	Output layer
Node number of RS	24	27	1
Node number of BD	16	19	1
Activating function	$1/(1 + e^{-s})$	$1/(1 + e^{-s})$	$(1 - e^{-s})/(1 + e^{-s})$

Step 1. Initializing the weights in the dynamic NNs where usually the random numbers distributed evenly in the interval of $[-1, 1]$ are employed for weight initialization.

Step 2. Constructing the input vector $X(k)$ according to r and d and figuring out $\bar{y}(k)$ and $e(k) = y(k) - \bar{y}(k)$.

Step 3. Modifying the weights using a suitable learning algorithm.

Step 4. Shifting the elements of $x_1(k), x_2(k), x_3(k)$ and $y(k)$ and terminating the training process if $|e(k)| < \varepsilon$ or turning to step (2) if $|e(k)| < \varepsilon$, where ε is the expected permitted error.

To obtain a relatively accurate mathematical model, the standard BP (backpropagation) algorithm, the momentum and adaptive learning rate-based BP algorithm, and L-M

(Levenberg and Marquardt) optimization-based BP algorithm are used to train the NNs mentioned above. It is found that the dynamic model trained by L-M optimization-based BP algorithm is the optimal strategy. The training parameters of NN dynamic models for residual soda and Baume degree are listed in Table 2.

The error curve in the training process of residual soda and Baume degree is shown in Figures 4 and 5. The simulation results show that the model has high convergence velocity; it is not suitable for local optimization. Based on the training of yielded data, the learning curves are shown in Figures 6 and 7. The true curve reflects the dynamic characteristics of the pulp washing process. The identification curve obtained by neural network reflects a good self-learning ability of neural network. The generalizing performance on validation data is shown in Figures 8 and 9, which describes the performance of the dynamic neural network models.

3.2.2. Neural Network Stationary Model Identification. After fully learning the pulp washing process by the dynamic neural network, the simulation data can be generated as the sample of the steady-state model, and only the dominant inputs are trained in the steady-state neural network model.

The steady NN mathematical model of the pulp washing process can be described as follows:

$$\begin{cases} \bar{y}_1 = f_1(x_1, x_2, x_3), \\ \bar{y}_2 = f_2(x_1, x_2, x_3), \end{cases} \quad (2)$$

where x_1 is inlet pulp consistency (kg/m^3), x_2 is inlet pulp flow (m^3/h), x_3 is hot clean water input flow (m^3/h), \bar{y}_1 is residual soda in the final washed pulp (g/L), and \bar{y}_2 is the Baume degree in the first stage filtrate tank (Be). $f_1(x_1, x_2, x_3)$ and $f_2(x_1, x_2, x_3)$ are steady models of residual soda and Baume degree trained by BP neural network.

A three-layer BP NN is also employed for steady model identification. The structure of the neural network model is shown in Figure 10.

The corresponding model parameters of steady NN for residual soda and Baume degree are listed in Table 3. Two steady data collections are yielded based on the NN dynamic models of residual soda and Baume degree by continually changing the values of input variables. Then, two data collections with 150 sets of sample set for neural network are obtained, respectively; 150 samples are composed of 30 sets of actual production simulation data and 120 sets of field data are collected by DCS (distributed control system).

Those sets of data are used as samples for identifying NN steady models. The 150 samples are divided into two parts: 100 sets of data are used to train the steady NN (30 sets of simulation samples are included) and another 50 sets of data are used to test the generalizing ability of the trained NN. For the training of NN steady models, the same three methods as above are employed. It is found that the L-M optimization-based BP algorithm is the best. The training parameters of NN steady models for residual soda and Baume degree are listed in Table 4. BP neural networks were learned through their classified samples, which led to improved learning

efficiency. The yielded data learning curves after neural network training are described in Figures 11 and 12. It reflects that the employed BP neural network has wonderful function and robust property. The generalizing curves of neural network are shown in Figures 13 and 14. The results show that the L-M training function works effectively and accurately in data correction and has good generalization performance.

3.2.3. Data Preprocessing Based on Pattern Clustering. As is well known, the learning and training of the neural network depend on sample data. In the actual pulping process, due to the influence of measurement device, measurement environment, measurement method, and human factors, the original measurement data inevitably have errors in the pulping process [17].

Therefore, the traditional mathematical statistics method is difficult to preprocess.

In view of this, in this paper, the pattern clustering method is used to preprocess the initial sample training sets. This method can detect the error of fault and carry out weighted average of the measurement data in the same mode according to the measurement time, which reflects the time variability of the washing process and reduces the random error to a certain extent. The principle of the process data preprocess method based on pattern clustering is shown in Figure 15, which is divided into three parts: pattern clustering, error elimination data, and weighted average filtering. For the original measurement data, it is divided into input space and output space. In the data sample, X_s is in the input space, Y_s is in output space, and the original data imply the mapping information from the input space to the output space.

Based on this method, the sample data are in the left Table 5, and the clustering results are displayed in the right of Table 5. It can be seen that 40 sample data are aggregated into 34 sample data (see Table 6 for data distribution after pretreatment).

The dataset tests the clustering effect of the improved algorithm. Cluster analysis data are applied to neural network, and 150 sets of data are obtained as the samples of neural network steady-state model identification. The 150 sets of data were randomly divided into two groups. The first group used 100 sets of data to train the stable neural network, and the second group used 50 sets of data to test the generalization ability of the training neural network. The learning curve and generalization curve after training are shown in Figures 16 and 17. Algorithm analysis and experimental results show that the improved algorithm has better detection performance, higher detection rate of learning performance, and stronger generalization ability.

The steady-state data generated by the steady-state neural network model provide a reliable data source for the specific mathematical model of Baume degree of residual alkali black liquor. By the least square fitting of 200 sets of data in the steady-state model, the mathematical models for

TABLE 2: Neural network dynamic model training parameters of residual soda and Baume degree.

Dynamic model	Learning rate	Initial value of μ	Increment of μ	Decrement of μ	Learning epochs	Error
RS	0.3	0.001	10	0.1	56	0.0019
BD	0.11	0.001	10	0.1	67	0.002

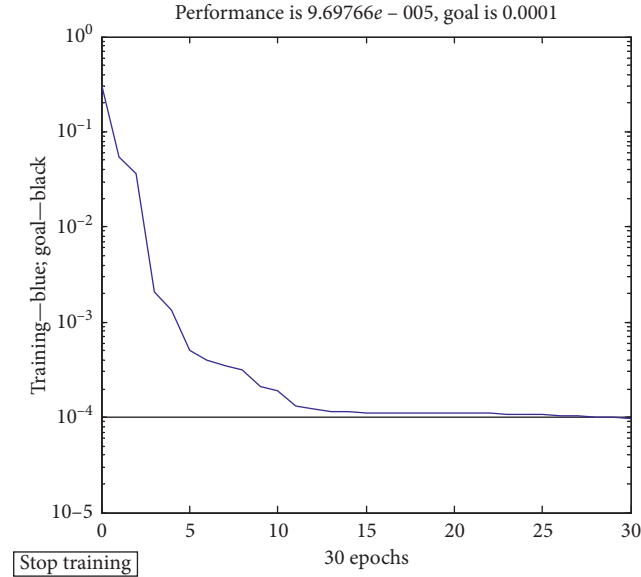


FIGURE 4: Residual soda error curve in training process.

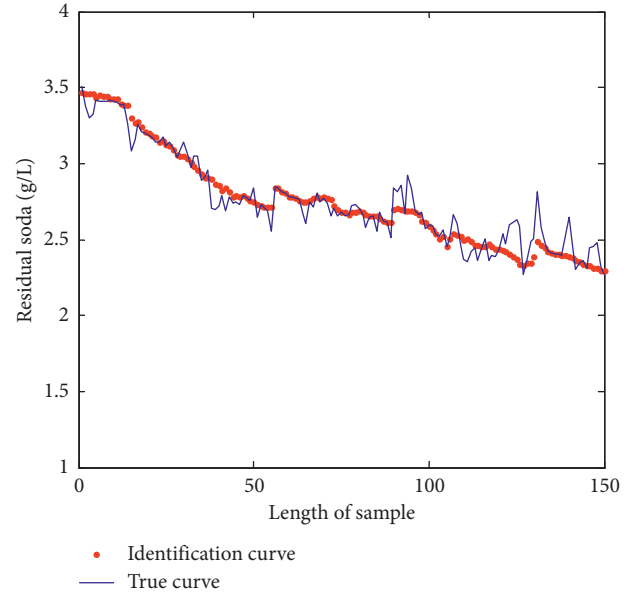


FIGURE 6: Neural network dynamic model learning curve of residual soda.

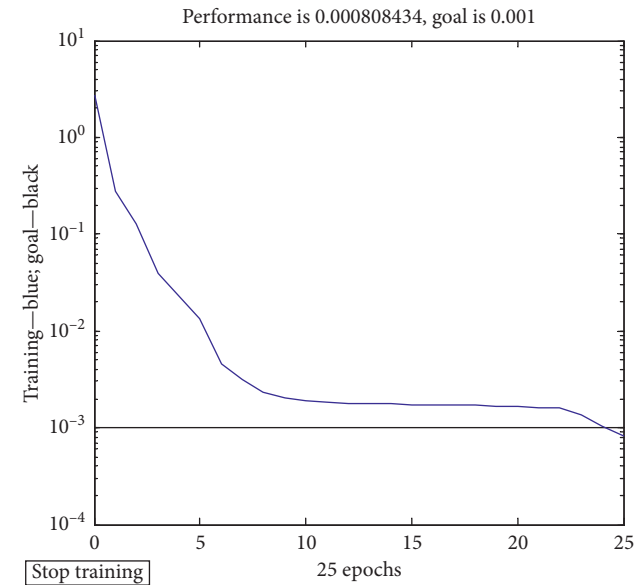


FIGURE 5: Baume degree error curve in training process.

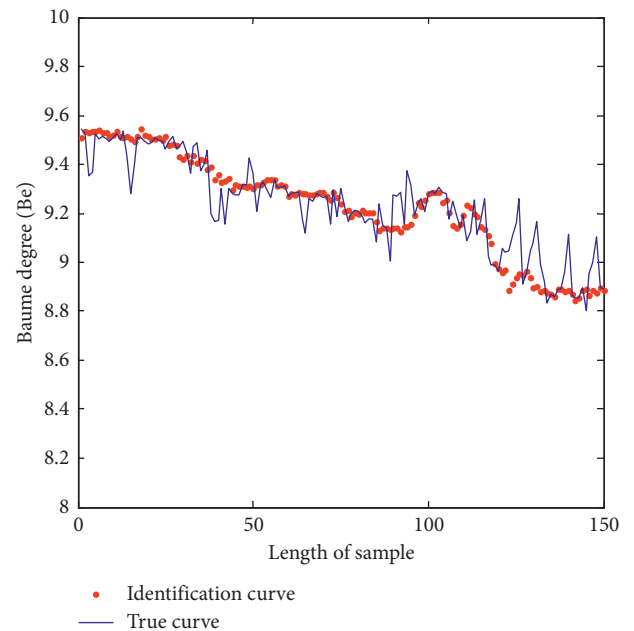


FIGURE 7: Neural network dynamic model learning curve of Baume degree.

the residual soda and Baume degree are established as shown in the following equation:

$$\begin{cases} y_1 = 0.27x_1 - 0.0013x_2 - 0.0078x_3 + 1.923, \\ y_2 = 0.77x_1 - 0.0118x_3 + 7.59. \end{cases} \quad (3)$$

Regression analysis is carried out for the parameters estimation, and the result is shown in Tables 7 and 8.

Regression analysis shows that 87% and 85% of residual soda and Baume degree are explained by the linear models, and the fitting deviation approximately follows the normal distribution. The predicted value can accurately reflect the

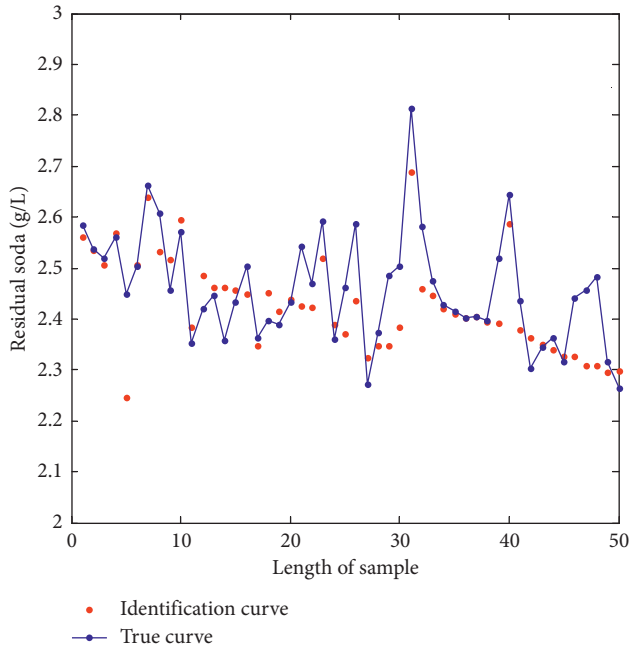


FIGURE 8: Model validation of residual soda.

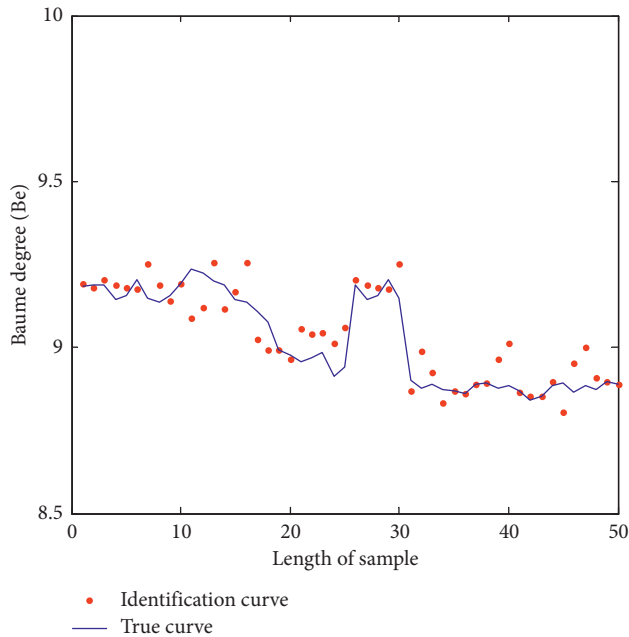


FIGURE 9: Model validation of Baume degree.

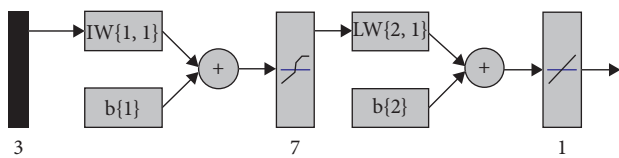


FIGURE 10: Structure of the stationary neural network mode.

TABLE 3: NN steady model parameters of residual soda (RS) and Baume degree (BD).

	Input layer	Hidden layer	Output layer
Node number of RS	3	7	1
Node number of BD	3	7	1
Activating function	$1/(1 + e^{-s})$	$1/(1 + e^{-s})$	$(1 - e^{-s})/(1 + e^{-s})$

changes of these two indexes, and the accuracy is conducive to the actual production.

4. Operation Mode Optimization of Pulp Washing Process

4.1. Condition Judgment. The idea of the comprehensive optimal control of the washing process is to take the stability of the comprehensive working condition of the washing process as the control objective and find out the best operating parameters, i.e., the pulp concentration, the pulp flow rate, and the amount of water, by adopting the optimal control method. Therefore, a new rigorous optimization model must be established.

The residual soda and black liquor Baume degree reflect the working state of the pulp washing process. In order to control the dynamic process of pulp washing, the prediction models of residual soda and black liquor Baume degree are applied to judge the working state of pulp washing quality. The working condition index S is shown in the following equation:

$$S = k_1 \left(\frac{y_1 - 1.923}{1.923} \right)^2 + k_2 \left(\frac{y_2 - 7.59}{7.59} \right)^2, \quad (4)$$

where 1.923 and 7.59 are the target values of residual soda and Baume degree of black liquor and k_1 and k_2 are the weights; the values of k_1 and k_2 are generally 0.4 and 0.6. According to the calculated S value, the comprehensive operating condition index can be divided into four intervals: excellent, good, medium, and poor. If it is optimal, the current parameter is maintained; if it is not optimal, adjust the operation parameters to optimize the model.

4.2. Comprehensive Optimization Control of Pulp Washing Process. The comprehensive optimization control framework of the pulp washing process is shown in Figure 18, which can be summarized as follows:

- (a) Collect a large number of experimental data under different conditions and then establish the model of the washing process quality index based on neural network and online correction.
- (b) According to judgment of condition S , establish the optimization mode library based on the condition parameters and operation modes in the interval.

TABLE 4: NN steady model training parameters of residual soda (RS) and Baume degree (BD).

Steady model	Learning rate	Initial value of μ	Increment of μ	Decrement of μ	Learning epochs	Error
RS	0.01	0.001	10	0.1	14	0.005
BD	0.01	0.001	10	0.1	11	0.0019

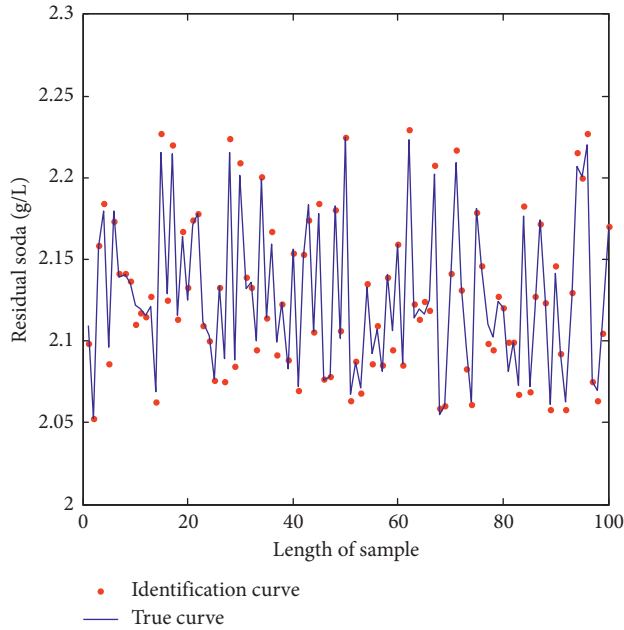


FIGURE 11: Learning curve of residual soda.

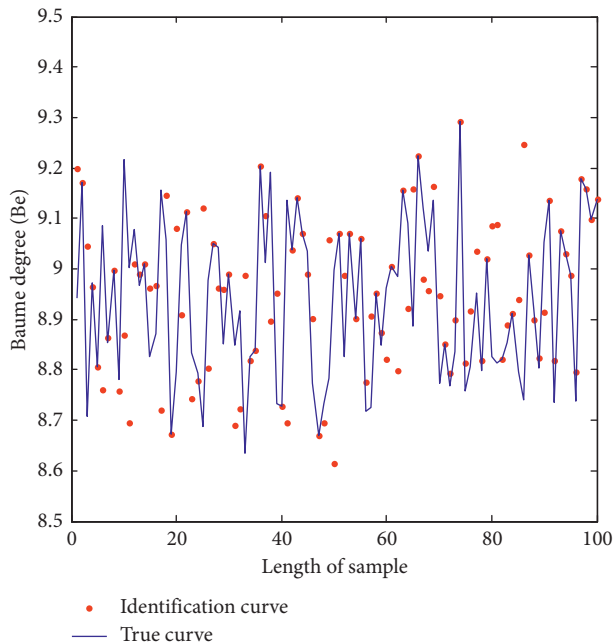


FIGURE 12: Learning curve of Baume degree.

Optimize the mode library when adding the condition samples and then recombine the samples [18]. Fuzzy matching of samples and clustering centers under current working conditions by similarity coefficient.

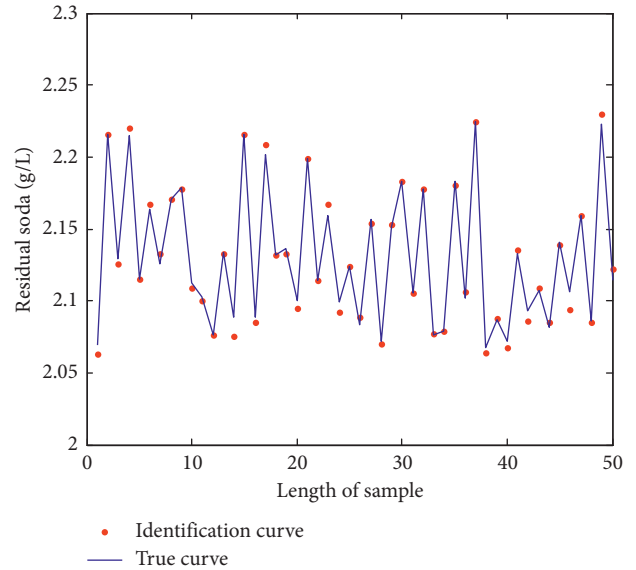


FIGURE 13: Generalizing curve of residual soda.

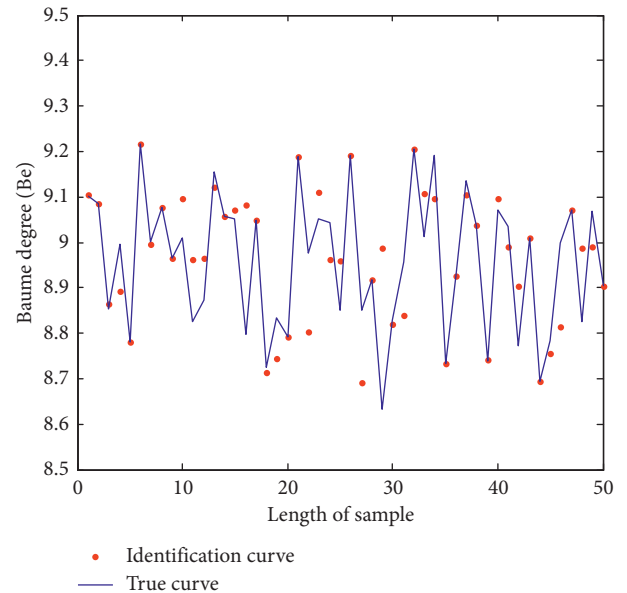


FIGURE 14: Generalizing curve of Baume degree.

(c) Search the similar samples and predict y_1 and y_2 by the neural network. If the operating condition index S_{is} is judged to be optimal, the operating parameters x_1, x_2 and x_3 will be added to the optimization mode database, and the current operation parameters will be maintained. If nonoptimal, reduce the number of samples for operation optimization, adjust the

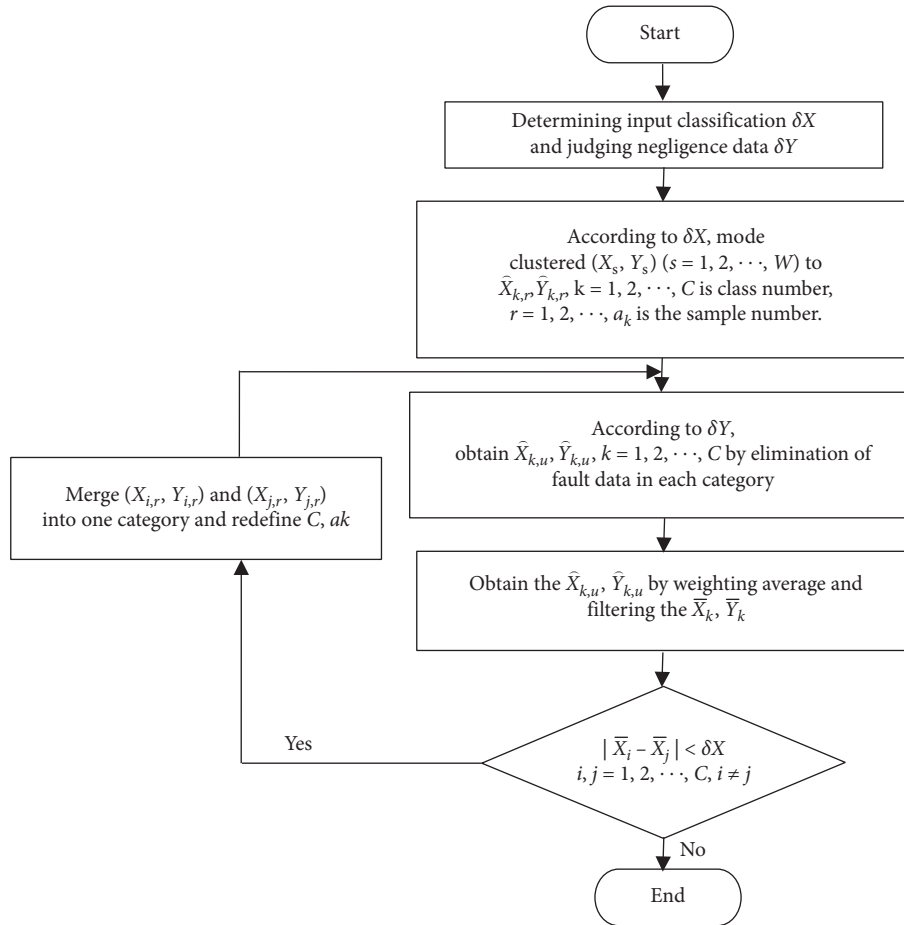


FIGURE 15: Flow of data preprocessing based on pattern clustering.

working conditions according to the mechanism process, and establish a large-scale optimization mode library.

- (d) In the established optimization model database, the dilution factor and operating parameters are optimized with the objective of maximum washed pulp and the minimum water consumption in the pulp washing process.

Through steps a, b, and c, the optimization model library is established. The quality of pulp washing is ensured by the

judgment of the working condition index S. On this basis, the production efficiency is improved, and the comprehensive optimization control of the pulp washing process is formed.

4.3. *Optimization Model of Pulp Washing Process.* In order to optimize the pulp washing process, a multiobjective optimization model of high quality, high yield, and low consumption is established. This objective model can be specified in more detail in the following equation:

$$\min Z = \min [F, Y', C]^T = \min \begin{cases} f_1(DF) = \frac{29}{3} \frac{DF}{(1 + (DF/6))(1 + (DF/5))^4 - 1} + \frac{27}{25} DF + \frac{22}{2}, \\ f_2(x_1, x_2) = x_{1\max} \cdot x_{2\max} - x_1 \cdot x_2, \\ f_3(x_3) = x_3, \end{cases} \tag{5}$$

$$\begin{cases} X_{1\min} \leq x_1 \leq X_{1\max} \\ X_{2\min} \leq x_2 \leq X_{2\max} \\ X_{3\min} \leq x_3 \leq X_{3\max} \\ 0 \leq DF \leq M, \end{cases} \text{ s.t.}$$

TABLE 5: Clustering of sample data.

No.	Original data					Class no.	No.	Clustered data				
	x_1	x_2	x_3	y_1	y_2			x_1	x_2	x_3	y_1	y_2
1	2.33	109.45	33.08	2.21	9.16	34	1	2.33	109.45	33.08	2.21	9.16
2	2.09	116.23	29.60	2.09	9.21	33	3	2.34	112.36	34.11	2.20	8.98
3	2.34	112.36	34.11	2.20	8.98	32	4	2.23	110.23	31.88	2.16	9.00
4	2.23	110.23	31.88	2.16	9.00	31	5	2.31	113.73	28.92	2.21	8.82
5	2.31	113.73	28.92	2.21	8.82	30	6	2.02	123.83	34.67	2.00	8.76
6	2.02	123.83	34.67	2.00	8.76	29	8	2.06	118.29	33.88	2.04	8.90
7	2.30	119.77	31.52	2.16	8.88	28	9	2.29	104.98	34.88	2.19	8.78
8	2.06	118.29	33.88	2.04	8.90	27	10	2.22	120.41	34.76	2.09	8.885
9	2.29	104.98	34.88	2.19	8.78	26	12	2.12	110.22	31.83	2.12	9.02
10	2.22	120.41	34.76	2.09	8.89	25	14	2.01	110.69	27.31	2.09	8.79
11	2.11	110.56	32.25	2.10	8.69	24	15	2.14	121.13	28.92	2.09	9.00
12	2.12	110.22	31.83	2.12	9.02	23	16	2.08	108.58	28.95	2.12	8.92
13	2.33	108.39	32.58	2.22	8.82	22	17	2.28	126.47	29.99	2.12	8.73
14	2.01	110.69	27.31	2.09	8.79	21	18	2.31	107.55	30.03	2.22	9.14
15	2.14	121.13	28.92	2.09	9.00	20	20	2.35	134.76	29.51	2.11	9.155
16	2.08	108.58	28.95	2.12	8.92	19	19	2.08	117.37	32.29	2.06	8.67
17	2.28	126.47	29.99	2.12	8.73	19	21	2.19	111.98	29.73	2.15	8.93
18	2.31	107.55	30.03	2.22	9.14	18	22	2.19	131.19	25.99	2.08	9.12
19	2.08	117.37	32.29	2.06	8.67	17	24	2.17	130.58	34.32	2.02	8.79
20	2.35	134.76	29.51	2.11	9.16	16	25	2.30	108.94	26.34	2.23	9.17
21	2.19	111.98	29.73	2.15	8.93	15	11	2.11	110.56	32.25	2.10	8.69
22	2.19	131.19	25.99	2.08	9.12	15	26	2.25	127.77	26.50	2.21	8.93
23	2.04	102.26	31.40	2.11	8.76	14	23	2.04	102.26	31.40	2.11	8.76
24	2.17	130.58	34.32	2.02	8.79	14	27	2.22	105.80	26.64	2.06	9.20
25	2.30	108.94	26.34	2.23	9.17	13	28	2.30	108.94	26.34	2.23	9.17
26	2.25	127.77	26.50	2.21	8.93	12	29	2.38	121.66	34.84	2.17	9.23
27	2.22	105.80	26.64	2.06	9.20	11	13	2.33	108.39	32.58	2.22	8.82
28	2.30	108.94	26.34	2.23	9.17	11	7	2.30	119.77	31.52	2.16	8.88
29	2.38	121.66	34.84	2.17	9.23	11	2	2.09	116.23	29.60	2.09	9.21
30	2.19	112.28	26.73	2.17	8.78	11	30	2.19	112.28	26.73	2.17	8.78
31	2.21	111.16	29.09	2.17	8.73	10	31	2.21	111.16	29.09	2.17	8.73
32	2.23	124.01	27.12	2.13	9.02	9	32	2.23	124.01	27.12	2.13	9.02
33	2.29	111.37	35.20	2.17	8.82	8	33	2.29	111.37	35.20	2.17	8.82
34	2.14	126.18	34.84	2.03	8.86	7	34	2.14	126.18	34.84	2.03	8.86
35	2.09	133.95	33.77	1.98	8.76	6	35	2.09	133.95	33.77	1.98	8.76
36	2.33	101.93	29.99	2.25	9.05	5	36	2.33	101.93	29.99	2.25	9.05
37	2.03	131.33	35.30	1.96	8.93	4	37	2.03	131.33	35.30	1.96	8.93
38	2.1734	127.40	34.72	2.04	8.93	3	38	2.17	127.40	34.72	2.04	8.93
39	2.25	108.68	27.32	2.20	8.94	2	39	2.25	108.68	27.32	2.20	8.94
40	2.28	102.58	30.98	2.21	8.70	1	40	2.28	102.58	30.98	2.21	8.70

TABLE 6: Data distribution after preprocessing.

Class no.	x_1	x_2	x_3	y_1	y_2	Class no.	x_1	x_2	x_3	y_1	y_2
1	2.28	102.58	30.98	2.21	8.70	18	2.19	131.19	25.99	2.08	9.12
2	2.25	108.68	27.32	2.20	8.94	19	2.14	114.68	31.01	2.12	8.8
3	2.17	127.40	34.72	2.04	8.93	20	2.35	134.76	29.51	2.11	9.155
4	2.03	131.33	35.30	1.96	8.93	21	2.31	107.55	30.03	2.22	9.14
5	2.33	101.93	29.99	2.25	9.05	22	2.28	126.47	29.99	2.12	8.73
6	2.09	133.95	33.77	1.98	8.76	23	2.08	108.58	28.95	2.12	8.92
7	2.14	126.18	34.84	2.03	8.86	24	2.14	121.13	28.92	2.09	9
8	2.29	111.37	35.20	2.17	8.82	25	2.01	110.69	27.31	2.09	8.79
9	2.23	124.01	27.12	2.13	9.02	26	2.12	110.22	31.83	2.12	9.02
10	2.21	111.16	29.09	2.17	8.73	27	2.22	120.41	34.76	2.09	8.885
11	2.23	114.17	30.11	2.16	8.92	28	2.29	104.98	34.88	2.19	8.78
12	2.38	121.66	34.84	2.17	9.23	29	2.06	118.29	33.88	2.04	8.9
13	2.30	108.94	26.34	2.23	9.17	30	2.02	123.83	34.67	2.00	8.76
14	2.13	104.03	29.02	2.09	8.98	31	2.31	113.73	28.92	2.21	8.82
15	2.18	119.17	29.38	2.16	8.81	32	2.23	110.23	31.88	2.16	9.00
16	2.3	108.94	26.34	2.23	9.17	33	2.34	112.36	34.11	2.2	8.98
17	2.17	130.58	34.32	2.02	8.79	34	2.33	109.45	33.08	2.21	9.16

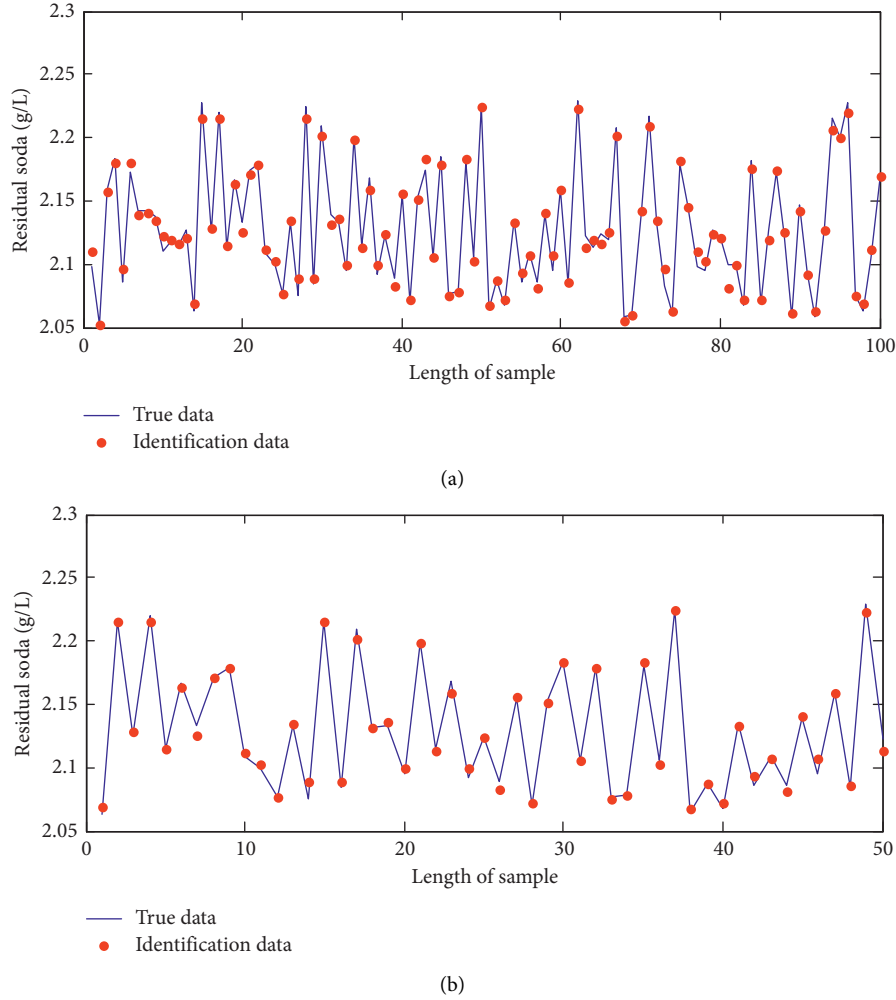


FIGURE 16: (a) Learning curve of residual soda. (b) Generalizing curve of residual soda.

where f_1 (DF) is the consumption cost, which is elated to dilution factor DF, $x_{1\max}$ and $x_{2\max}$ are the maximum inlet pulp consistency and inlet pulp flow, $f_2(x_1, x_2)$ is the deviation of the pulp output, and $f_3(x_3)$ is the water consumption.

4.4. *Solution of the Optimization Model.* Structuring an evaluation function, transform the multiobjective optimization problem [18] into a single objective as follows:

$$\begin{aligned} \min \quad & Z = \omega_1 f_1^2 + \omega_2 f_2^2 + \omega_3 f_3^2, \\ \text{s.t.} \quad & \begin{cases} X_{1\min} \leq x_1 \leq X_{1\max} \\ X_{2\min} \leq x_2 \leq X_{2\max} \\ X_{3\min} \leq x_3 \leq X_{3\max} \\ 0 \leq \text{DF} \leq M, \end{cases} \end{aligned} \quad (6)$$

where $\omega_1 \in [0, 1]$, $\omega_2 \in [0, 1]$, $\omega_3 \in [0, 1]$, and $\omega_1 + \omega_2 + \omega_3 = 1$. It reflects the importance of these objectives.

During operation in a paper mill, the parameters are designed as follows:

$$\begin{cases} X_{1\min} = 2.0 \text{ kg/m}^3, \\ X_{1\max} = 2.4 \text{ kg/m}^3, \\ X_{2\min} = 100 \text{ m}^3/\text{h}, \\ X_{2\max} = 135 \text{ m}^3/\text{h}, \\ X_{3\min} = 25 \text{ m}^3/\text{h}, \\ X_{3\max} = 35 \text{ m}^3/\text{h}. \end{cases} \quad (7)$$

The optimization model is a nonlinear multiobjective optimization problem with linear constraints; ant colony algorithm can be regarded as a distributed multiagent system [19]. In this paper, ant colony optimization algorithm is

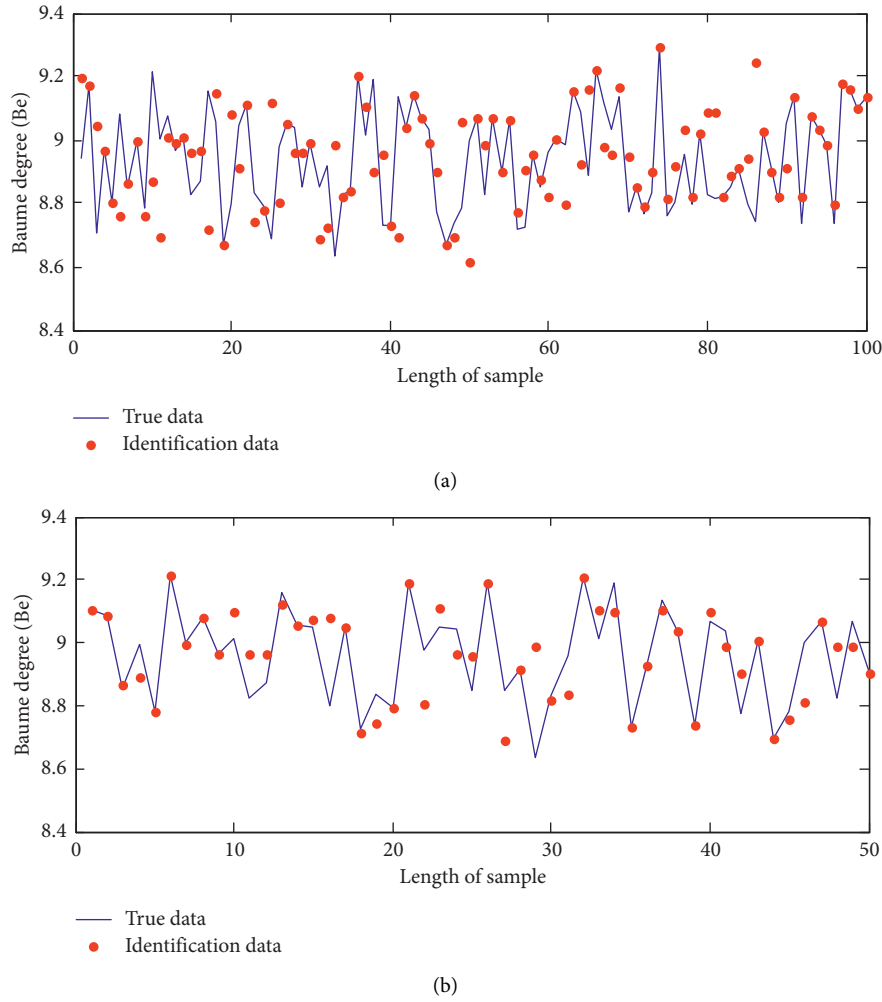


FIGURE 17: (a) Learning curve of Baume degree. (b) Generalizing curve of Baume degree.

TABLE 7: Regression statistics of residual soda.

Multiple R	0.864969
R square	0.848171
Adjusted R square	0.842029
Standard error	0.023671
Observation data	200

TABLE 8: Regression statistics of Baume degree.

Multiple R	0.754913
R square	0.769893
Adjusted R square	0.759403
Standard error	0.03092
Observation data	200

adopted for optimization, and the optimal operation mode is obtained by iteration of 10 steps.

The model of unconstrained optimization based on penalty function is established. The corresponding nonlinear unconstrained optimization model of the pulp washing process is as follows:

$$\min F(x_i, \sigma) = \min \left\{ \omega_1 (x_3 - E^*) + \omega_2 (Y^* - x_1 x_2) + \frac{1}{2} \sigma \sum_{j=1}^9 \{ \min[0, g_j(x_i)] \}^2 \right\}, \quad i = 1, 2, 3, \quad (8)$$

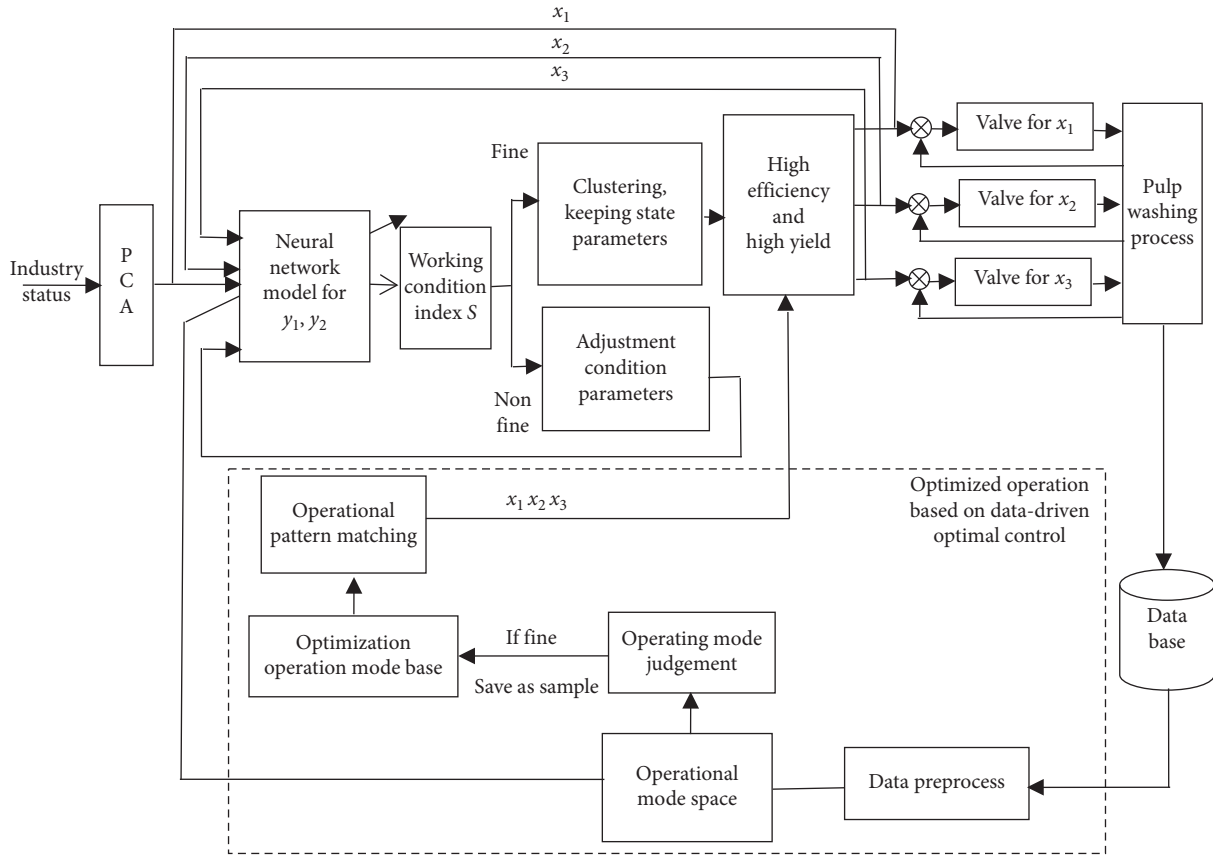


FIGURE 18: Framework of comprehensive optimization control.

TABLE 9: Optimization results of the pulp washing process.

Independent variables	Inlet pulp consistency (kg/m ³)	Inlet pulp flow (m ³ /h)	Inlet water flow (m ³ /h)	DF	Weight
Before optimization	2.1	117.5	34.7	0.5	$\omega_1 = 0.4, \omega_2 = 0.3, \omega_3 = 0.3,$
After optimization	2.4	129.7	25	3.2	$\beta = 0.5$
Dependent variables	Baume degree (Be)	Residual soda (g/L)	Outlet slurry (T/h)	Cost (yuan/ton)	Weight
Before optimization	7.2546	2.398	246.75	13.2	$\omega_1 = 0.4, \omega_2 = 0.3, \omega_3 = 0.3,$
After optimization	9.3003	2.1848	311.28	10.9	$\beta = 0.5$

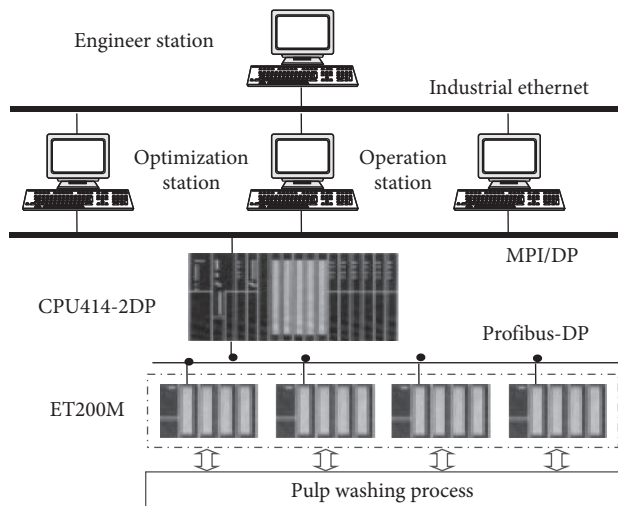


FIGURE 19: DCS structure of straw pulp washing.

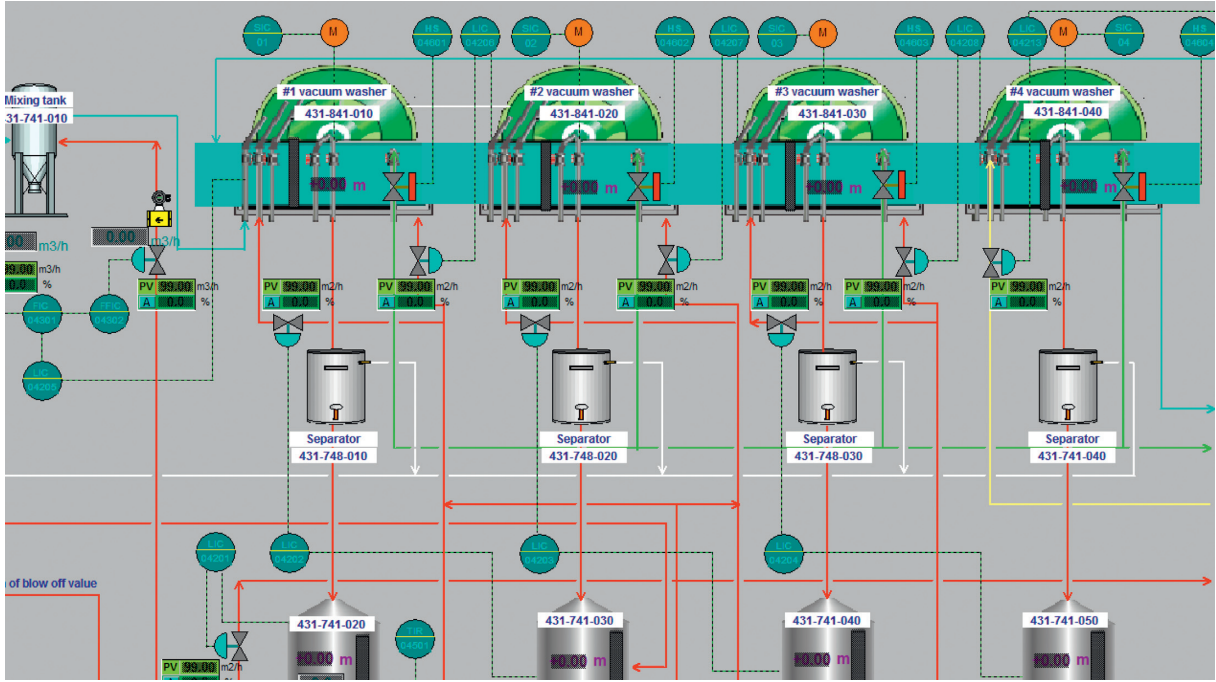


FIGURE 20: Main interface of pulp washing section.

and when x_i is an infeasible point, $\min[0, g_j(x_i)] = 0$. When x_i is a feasible point, $\min[0, g_j(x_i)] = g_j(x_i)$.

Equation (8) is an unconstrained minimization problem. It expresses the constraints of model (6) in the objective function in the form of penalty function [20].

Accordingly, the network state equation of the nonlinear programming is constructed as follows:

$$\begin{cases} C \cdot \frac{du_i}{dt} = -\frac{\partial f(X)}{\partial x_i} - \lambda \sum_{j=1}^9 \left[\sigma \min(0, g_j(X)) \frac{\partial g_j(X)}{\partial x_i} \right], \\ X(t+1) = \min[0, u(t+1)]. \end{cases} \quad (9)$$

The iteration formula of the optimization variable is $x_i(k+1) = x_i(k) + \lambda \nabla F(x_i, \sigma)$.

For the dynamic equation of the neural network,

$$\frac{du_i}{dt} = -\frac{\partial f(X)}{\partial x_i} - \lambda \sum_{j=1}^9 \left[\sigma \min(0, g_j(X)) \frac{\partial g_j(X)}{\partial x_i} \right]. \quad (10)$$

It can be further simplified as follows:

$$\frac{du_i}{dt} = -\frac{\partial f(X)}{\partial x_i} - \lambda \sum_{j=1}^9 \sigma S_j g_j(X) \frac{\partial g_j(X)}{\partial x_i}, \quad (11)$$

where $S_j = \begin{cases} 1 & g_j(x) \leq 0 \\ 0 & g_j(x) > 0 \end{cases}$ and $S_j g_j(x) = \min(0, g_j(x))$.

The calculation process of the dynamic equation is as follows:

$$\begin{aligned} \frac{du_1}{dt} &= \frac{\partial F(x_j, \sigma)}{\partial x_1} = -\omega_2 x_2 + 2 \cdot \sigma \left[(f_1(x_1, x_2, x_3) - 2.5) \cdot \frac{\partial f_1(x_1, x_2, x_3)}{\partial x_1} \right] - 2 \cdot \sigma \left[(7 - f_2(x_1, x_2, x_3)) \cdot \frac{\partial f_2(x_1, x_2, x_3)}{\partial x_1} \right] \\ &\quad + 2 \cdot \sigma (f_2(x_1, x_2, x_3) - 9) \cdot \frac{\partial f_2(x_1, x_2, x_3)}{\partial x_1} + 2 \cdot \sigma [(x_1 - 2.4) - (2.0 - x_1)], \\ \frac{du_2}{dt} &= \frac{\partial F(x_j, \sigma)}{\partial x_2} = -\omega_2 x_1 + 2 \cdot \sigma \left[(f_1(x_1, x_2, x_3) - 2.5) \cdot \frac{\partial f_1(x_1, x_2, x_3)}{\partial x_2} \right] - 2 \cdot \sigma \left[(7 - f_2(x_1, x_2, x_3)) \cdot \frac{\partial f_2(x_1, x_2, x_3)}{\partial x_2} \right] \\ &\quad + 2 \cdot \sigma (f_2(x_1, x_2, x_3) - 9) \cdot \frac{\partial f_2(x_1, x_2, x_3)}{\partial x_2} + 2 \cdot \sigma [(x_2 - 135) - (100 - x_2)], \\ \frac{du_3}{dt} &= \frac{\partial F(x_j, \sigma)}{\partial x_3} = 1 + 2 \cdot \sigma \left[(f_1(x_1, x_2, x_3) - 2.5) \cdot \frac{\partial f_1(x_1, x_2, x_3)}{\partial x_3} \right] - 2 \cdot \sigma \left[(7 - f_2(x_1, x_2, x_3)) \cdot \frac{\partial f_2(x_1, x_2, x_3)}{\partial x_3} \right] \\ &\quad + 2 \cdot \sigma (f_2(x_1, x_2, x_3) - 9) \cdot \frac{\partial f_2(x_1, x_2, x_3)}{\partial x_3} + 2 \cdot \sigma [(x_3 - 35) - (25 - x_3)]. \end{aligned}$$

(12)

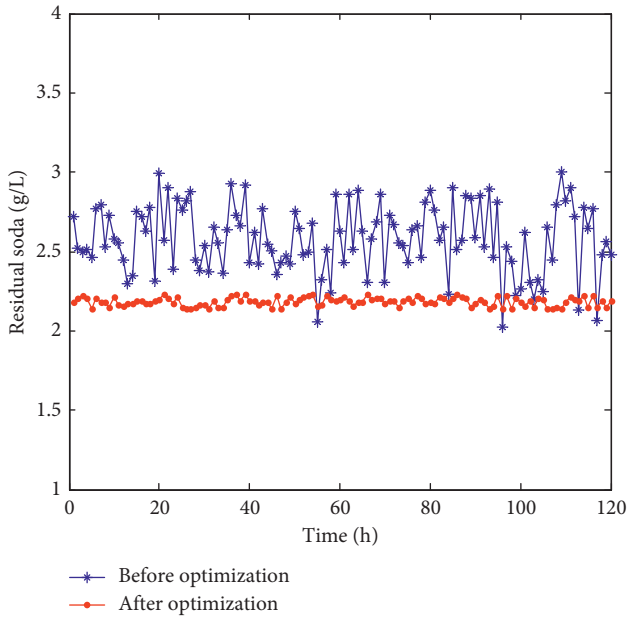


FIGURE 21: Comparison of residual soda before and after optimization.

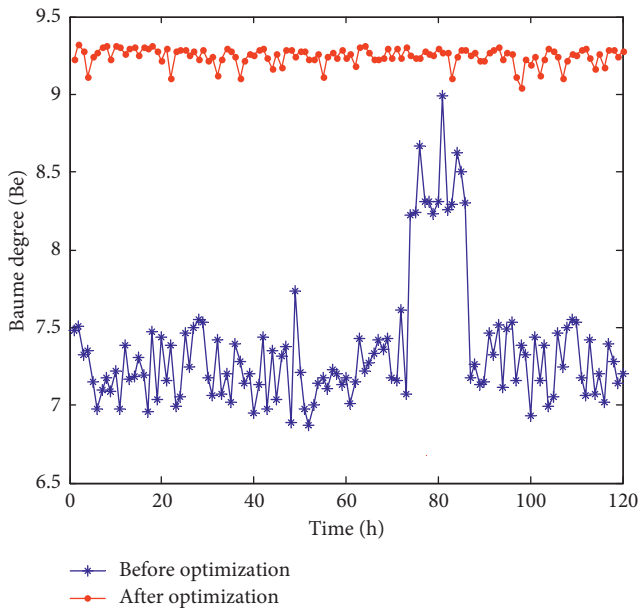


FIGURE 22: Comparison of Baume degree before and after optimization.

The iteration step of λ in the upper median can be obtained either by fixing a constant or by optimizing each part. The penalty factor σ will increase with the number of infeasible solutions. The specific algorithm is as follows:

- (i) Given the initial point $X^{(0)}$, the initial penalty factor $\sigma_1 > 0$ (taken $\sigma_1 = 1000$ for the actual calculation), the magnification factor $C > 1$ (e.g. $C = 2$), and the allowable error $\xi > 0$, $k = 1$.

- (ii) Take $X^{(k-1)}$ as the initial point; the unconstrained minimum of $X^{(k)}$ is obtained by solving the penalty function $\min G(X, \sigma_k)$.
- (iii) If the penalty is smaller than ξ , the calculation is stopped and the approximate minimum of $X^{(k)}$ for the original problem is obtained. Otherwise, $\sigma_{k+1} = C\sigma_k$ and $k = k + 1$, and return to step (ii).

5. Optimization Results and Application

5.1. Optimization Results. A comparison between before and after optimization on residual soda and Baume degree has been done in a paper mill in Shandong province, China. The results are shown in Table 9. After optimization, the residual soda and Baume degree can meet the process requirements (the residual soda is not higher than 2.2 g/L; the Baume degree is between 7.0 and 9.4Be) [21, 22].

The average of residual soda decreased, and the Baume degree is increased to 9.3Be. The concentration and flow of the inlet pulp tend to the upper limit of the production index, the pulp yield is increased by 20%, and the water consumption is decreased by nearly 30%. Meanwhile, by the optimum of DF, the total cost is reduced from 13.2 to 10.9, and it has a certain economic benefit for pulping enterprises.

5.2. Application. Based on the hardware of Siemens S7-400 PLC and software of WinCC 6.0 & Step 7, an optimization control system has been designed for the pulp washing process. The DCS structure is shown in Figure 19. This system is a three-level control system. The soft measurement model and correction model of residual soda and Baume degree are embedded into the DCS in the optimization station. The main interface of the pulp washing process in the engineer station is shown in Figure 20.

Through the trial operation in the pilot paper mill, the curve of one week operation result is given and compared with the curve before optimization. As shown in Figures 21 and 22. It can be seen that before optimization, the residual soda and Baume degree cannot reach the process requirements at the same time and fluctuate greatly. After optimization, both of them are kept in a relatively stable value, and the higher the Baume degree, the lower the residual soda. It solves the incompatibility of the two indexes. In the actual production process, the long-term detection of the dynamic deviation shows that the model has high prediction accuracy.

6. Conclusions

In order to realize the multiobjective optimization of high yield, low cost, and low consumption in the pulp washing process, the operation mode optimization method in the pulp washing process is proposed. The optimal control has been successfully operated in some paper mills of China, and remarkable economic benefits have been achieved. In other words, the multiobjective optimization subsystem can effectively balance the contradiction between the residual soda and Baume degree. On the other hand, by optimizing the

subsystem, a set of recommended optimal operation parameters (including pulp concentration, pulp flow, and hot water flow) can be provided to the engineer, which increases the flexibility of pulp washing scheduling. It is important that the outlet pulp increases substantially and the hot water consumption is decreased to the lower limit. The total cost of the pulp washing process is reduced, which has a certain economic benefit for pulping enterprises.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Shaanxi Provincial Technology Innovation Guidance Project (2018HJCG-10) and Shaanxi Science & Technology Co-ordination & Innovation Project (2016KTCQ01-35). The authors sincerely thank them for funding the project.

References

- [1] C. J. Yang, Y. X. Sun, and B. L. Bao, "The present and future of the pulp washing process control," *China Pulp & Paper*, vol. 25, no. 19, pp. 14–19, 1998.
- [2] W. Tang and S. J. Shi, "Pulp washing process DCS based on double-objective optimization," *Control Theory and Applications*, vol. 4, no. 19, pp. 555–560, 2002.
- [3] W. H. Gui, C. H. Yang, Y. G. Li et al., "Data-driven operational-pattern optimization for copper flash smelting process," *Acta Automatica Sinica*, vol. 35, no. 6, pp. 718–723, 2009.
- [4] A. J. Yan, T. Y. Chai, and H. Yue, "Multivariable intelligent optimizing control approach for shaft furnace roasting process," *Acta Automatica Sinica*, vol. 4, no. 32, pp. 636–640, 2006.
- [5] P. F. Cao and X. L. Luo, "Modeling of soft sensor for chemical process," *CIESC Journal*, vol. 64, no. 3, pp. 789–801, 2013.
- [6] S. M. Xie, X. W. Gao, and T. Y. Chai, "BOF end point prediction based on grey model," *Journal of Iron and Steel Research*, vol. 4, no. 11, pp. 25–29, 1999.
- [7] S. J. Zhao, J. Zhang, and Y. M. Xu, "Monitoring of processes with multiple operating modes through multiple principle component analysis models," *Industrial & Engineering Chemistry Research*, vol. 43, no. 18, pp. 7025–7030, 2004.
- [8] Z. G. Du, G. Q. Wang, Y. G. Zhang et al., "Application of BP neural networks in predicting the product yield of HVGO steam cracking," *Ethylene Industry*, vol. 2, no. 21, pp. 8–11, 2009.
- [9] H. Z. Kun, G. W. Hua, P. X. Qi et al., "Intelligent optimization of optimal operational pattern in the process of copper converting furnace," *Control Theory and Applications*, vol. 2, no. 22, pp. 243–247, 2005.
- [10] P. F. Jiang, X. P. Gu, X. Chen et al., "Parameter estimation of large-scale industrial process model with multiple operating points," *Computers and Applied Chemistry*, vol. 1, no. 27, pp. 46–50, 2010.
- [11] E.-L. Hu and B. Wang, "A new optimization in SDP-based learning," *Neurocomputing*, vol. 365, no. 23, pp. 10–20, 2019.
- [12] K. T. Yao, Z. J. Shao, X. Chen et al., "Data-driven technology and mechanism model based soft sensor modeling in PTA process," *Computers and Applied Chemistry*, vol. 10, no. 27, pp. 1388–1394, 2010.
- [13] W. Tang, M. Wang, L. He, and H. Itoh, "Neural network based double-objective optimization and application to pulp washing process improvement," *Industrial & Engineering Chemistry Research*, vol. 46, no. 14, pp. 5015–5020, 2007.
- [14] V. Cherkassky and Y. Q. Ma, "Practical selection of SVM parameters and noise estimation for SVM regression," *Neural Networks*, vol. 1, no. 17, pp. 113–127, 2004.
- [15] P. F. Zhu, L. Y. Xia, and H. T. Pan, "Multi-model fusion modeling method based on improved kalman filtering algorithm," *CIESC Journal*, vol. 66, no. 4, pp. 1388–1394, 2015.
- [16] F. Han, A. Han, and J. Hao, "Saliency detection method using hypergraphs on adaptive multiscales," *IEEE Access*, vol. 23, no. 2, pp. 245–255, 2018.
- [17] Z. Li and J. J. Cai, "Application of the soft measurement technology to optimization system of boiler burning at power station," *The Chinese Journal of Process Engineering*, vol. 6, no. 4, pp. 49–55, 2004.
- [18] H. Liu, Z. J. Liu, and H. G. Li, "A data-driven approach to chemical process alarm threshold optimization," *CIESC Journal*, vol. 63, no. 9, pp. 2372–2378, 2012.
- [19] X. F. Li, C. Mei, and T. Y. Xiao, "Numerical simulation, analysis of guixi copper flash smelting furnace," *Rare Metals*, vol. 4, no. 21, pp. 260–265, 2002.
- [20] H. Nam and R. Lysecky, "Security-aware multi-objective optimization of distributed reconfigurable embedded systems," *Journal of Parallel and Distributed Computing*, vol. 133, no. 23, pp. 377–390, 2019.
- [21] H. B. Jiang and J. L. Zhang, "Properties and structural optimization of pulverized coal for blast furnace injection," *Journal of Iron and Steel Research*, vol. 3, no. 18, pp. 06–12, 2011.
- [22] W. Tang, M. X. Wang, Y. Y. Chao, L. F. He, and H. Itoh, "Double-objective optimization for a pulp washing process based on neural network," in *Proceedings of the 2007 International Conference on Mechatronics and Automation*, pp. 5–8, Harbin, China, August 2007.