Dynamic soft sensing of organic pollutants in effluent from UMIC anaerobic reactor for industrial papermaking wastewater

Yajuan Xing[1], Zhong Cheng[2],*, Shengdao Shan[3].

With the rapid development of paper industry, the pressure of environmental pollution is going more and more serious. Recently, resource utilization of wastewater by anaerobic digestion has become a feasible way to solve this problem. In order to maintain the safe and efficient production of the process, a novel adaptive soft sensor model was developed to infer the chemical oxygen demand (COD) of paper mill effluent in this paper. First, the principal component analysis technique was performed in this model so as to eliminate the col-linearity between the process variables and accordingly obtain the low-dimensional feature principal component. Then, the least square support vector machine method was used to construct a quantitative regression model between principal component and the effluent COD. Along with it, particle swarm optimization was implemented to search for the best value of the LSSVM model parameters, namely the kernel parameters and the regularization factor. Finally, an online calibration strategy was designed to adapt to the process dynamic changes in an adaptive iterative manner. When the constructed model tested for performances in a full-scale factory, the average relative deviation and maximum deviation are 1.80% and 6.26%, respectively. The experimental results show that this proposed soft sensor model is featured with high accuracy and strong dynamic stability, and it can provide good guidance for COD prediction and optimal control of paper mill wastewater treatment.

Contact information:
[1]Key Laboratory of Recycling and Eco-treatment of Waste Biomass of Zhejiang Province, School of Environment and Resources, Zhejiang University of Science and Technology, Hangzhou 310023
[2]School of Biological and Chemical Engineering, Zhejiang University of Science and Technology, Hangzhou 310023
[3]Key Laboratory of Recycling and Eco-treatment of Waste Biomass of Zhejiang Province, School of Environment and Resources, Zhejiang University of Science and Technology, Hangzhou 310023

Dynamic soft sensing of organic pollutants in effluent from UMIC anaerobic reactor for industrial papermaking wastewater

Yajuan Xing[1], Zhong Cheng[2,*], Shengdao Shan[3]

[1]Key Laboratory of Recycling and Eco-treatment of Waste Biomass of Zhejiang Province, School of Environment and Resources, Zhejiang University of Science and Technology, Hangzhou 310023
[2]School of Biological and Chemical Engineering, Zhejiang University of Science and Technology, Hangzhou 310023
[3]Key Laboratory of Recycling and Eco-treatment of Waste Biomass of Zhejiang Province, School of Environment and Resources, Zhejiang University of Science and Technology, Hangzhou 310023

*Corresponding author e-mail: chengzhong@zust.edu.cn

Abstract. With the rapid development of paper industry, the pressure of environmental pollution is going more and more serious. Recently, resource utilization of wastewater by anaerobic digestion has become a feasible way to solve this problem. In order to maintain the safe and efficient production of the process, a novel adaptive soft sensor model was developed to infer the chemical oxygen demand (COD) of paper mill effluent in this paper. First, the principal component analysis technique was performed in this model so as to eliminate the col-linearity between the process variables and accordingly obtain the low-dimensional feature principal component. Then, the least square support vector machine method was used to construct a quantitative regression model between principal component and the effluent COD. Along with it, particle swarm optimization was implemented to search for the best value of the LSSVM model parameters, namely the kernel parameters and the regularization factor. Finally, an online calibration strategy was designed to adapt to the process dynamic changes in an adaptive iterative manner. When the constructed model tested for performances in a full-scale factory, the average relative deviation and maximum deviation are 1.80% and 6.26%, respectively. The experimental results show that this proposed soft sensor model is featured with high accuracy and strong dynamic stability, and it can provide good guidance for COD prediction and optimal control of paper mill wastewate treatment.

1. Introduction

The paper-making industry is a major water consumer and also a major wastewater discharger. According to the statistics of the Ministry of Ecology and Environment, In 2015, the total water consumption of the paper-making industry and the paper product industry (4,180 enterprises involved in the statistics) was 11.835 billion tons, and the wastewater discharge was 2.367 billion tons,

accounting for 13.0% of the total industrial wastewater discharge. The chemical oxygen demand (COD) in the discharged wastewater is 335,000 tons, accounting for 13.1% of the total industrial COD emission. In recent years, with the increasing shortage of water resources, production water has become a problem that restricts the development of paper-making enterprises. At present, in order to solve the environmental pollution due to paper-making wastewater and realize resource utilization, biogas production through anaerobic digestion has become a main method. The anaerobic digestion process under the action of microorganisms is featured as multi-factor influence, dynamic variability, complex nonlinearity (Yang Hao et al., 2016), etc. and the mechanism model thereof is difficult to construct, so the real-time operation control and optimization and calibration that affect safe production and effluent water production conditions cannot be realized. The production effectiveness of the industrialization process of anaerobic digestion for paper-making wastewater is often measured by the effluent COD. However, the current COD testing of enterprises is mostly realized by timed manual sampling and laboratory analysis. The test results cannot be obtained till several hours later, so the real-time performance is poor (Xu Lisha et al., 2012). In case that a COD on-line analyzer is installed on site, failure often occurs, resulting in loss of data. And also, the maintenance is difficult and the instrument is expensive(Langergraber et al., 2004; Bourgeois et al., 2010). With the improvement of enterprise automation as well as the deep integration of informationization and industrialization, the methods like pivot element regression, partial least squares regression, neural network, support vector machine and fuzzy logic have been used for the data modeling and operational control of the performance indicators including COD concentration, volatile fatty acid (VFA), dissolved oxygen, suspended solids (SS) concentration and gas production in the process of paper-making wastewater treatment (Bourgeois et al., 2010; Haimi et al., 2013) Choi et al., 2001; Wan et al., 2011; Huang et al., 2015 Dürenmatt et al., 2012; Zhou Hongbiao et al., 2017; Liu Lin et al., 2017; Tang Wei et al., 2017). With respect of the method selection, Wan et al. (2011) designed an adaptive fuzzy inference system integrating fuzzy subtractive clustering and PCA technologies, of which the fuzzy subtractive clustering is used to identify the model structure, and PCA is used to reduce the complex collinearity between variables as well as the dimensionality. The model accuracy with this integrated method is higher than that with the BP neural network method in the performance test about the COD and SS concentration prediction of paper-making wastewater. Wang Yao et al. (2017) chose the LSSVM method to predict the COD and SS concentrations. The results show that the soft-sensor model created by optimizing the LSSVM method parameters via the PSO algorithm has a higher prediction accuracy. The LSSVM method based on minimum structural risk is widely used in soft-sensor modeling because of its features of low dependence on sample data, less parameters to be estimated, and strong generalization ability (Souza et al., 2016; Wang et al., 2015; Fortuna et al., 2007; Liu Bo et al., 2015; Zheng Rongjian et al., 2017). However, the prediction accuracy of the soft-sensor model based on the offline sample data architecture, will gradually decline in the face of dynamic changes in continuous production processes. In order to solve the above problem, this paper proposes an OCS-PCA-PSO-LSSVM soft-sensor method integrating data analysis technology and regression modeling, which can eliminate the complex collinearity between variables and achieve dimensionality reduction via PCA technology; then, implement the LSSVM method to establish the nonlinear relationship between input and output variables, and realize the optimization of LSSVM model parameters by means of PSO; and finally, initiate the online calibration strategy (OCS) in case the prediction deviation of the new sample individual exceeds the set error limit, iteratively updating the soft-sensor model in an adaptive manner.

2. Materials and Methods

2.1. Process and Data Collection

With the wastewater anaerobic treatment system of a papermaking mill as the test object of application, the production process is shown in Fig. 1, in which the ascending multistage internal circulation anaerobic reactor UMIC is the main device. The UMIC reactor works based on the principle of granular sludge (Ruggeri et al., 2015; Zhang Yi et al., 2014), namely, the papermaking wastewater is thoroughly mixed with anaerobic microbial sludge after being pumped into the reactor by a lift pump,

and the organics in the mixture are chemically converted into the gases like methane and carbon dioxide, as well as microbial bacterial plastids under the action of the microorganisms.

Based on the analysis of production behavior and process mechanism of the UMIC device, with the combination of experts' experience and knowledge as well as the sensitivity analysis of field data, 8 process variables that affect the COD of the treatment system were selected as the input variables of the model, and they are: influent COD/mg•L^{-1}, influent SS/mg•L^{-1}, influent pH, influent flow/m^3, influent temperature/°C, circulating pool level/%, effluent pH and effluent temperature/°C, while the output variable of the model is effluent COD/mg•L^{-1}. Two sample data collection methods were adopted, one of which was that the mill's distributed control system DCS was used to collect 8 process variables, and the other was that the on-site sampling laboratory obtained effluent COD through offline test (Sun Jun et al., 2017). After the collection of the mill's field operation data from July 2016 to February 2017 was completed, the missing data was directly removed, then the abnormal data was identified and deleted, and finally the initial sample matrix set containing 175 sample individuals was obtained.
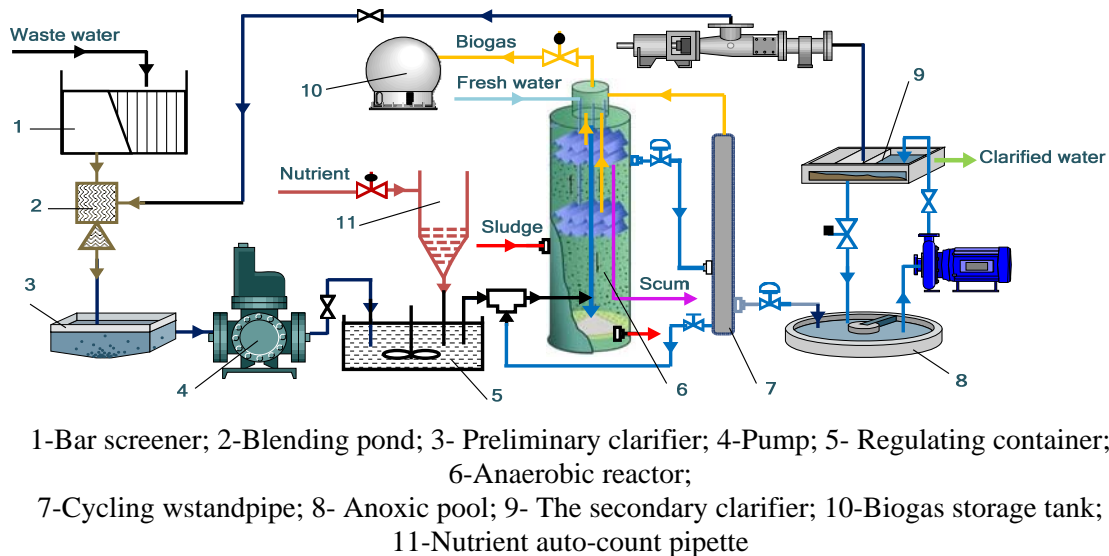

1-Bar screener; 2-Blending pond; 3- Preliminary clarifier; 4-Pump; 5- Regulating container;
6-Anaerobic reactor;
7-Cycling wstandpipe; 8- Anoxic pool; 9- The secondary clarifier; 10-Biogas storage tank;
11-Nutrient auto-count pipette
Fig 1. Flow chart of anaerobic degradation for paper mill wastewater

2.2. OCS-PCA-PSO-LSSVM Soft-sensor Method

2.2.1. PCA Technology

The independent variable matrix $X_{n\times p}$ of the obtained initial sample data was recorded, where n is the number of sample individuals, i.e. the sample size, and p is the number of process variables. PCA technology (Jolliffe et al., 2002) was namely to project X from the Euclidean space to the latent vector space of the pivot element.

$$X = TQ^{\mathrm{T}} + E = \sum_{k=1}^{d} t_k q_k^{\mathrm{T}} + E \qquad (1)$$

Where, t_k is the kth extracted pivot element, q_k is the load vector used to extract the pivot element, and E is the final residual matrix. In essence, the construction of the PCA latent vector space is to represent most of the dynamic information in the initial process variables in the sample data by extracting d pivot elements ($d \le p$) (Sun Jun et al., 2017), of which, the information contribution of the kth pivot element can be calculated according to Formula (2).

$$\eta_k = \lambda_k / \sum_{k=1}^{p} \lambda_k \qquad (2)$$

2.2.2. Soft-sensor Optimization Model PSO-LSSVM

The least square support vector machine (LSSVM) is an extension of the standard SVM method (Cristianini et al., 2000), and the main ideas of the algorithm are summarized as follows:

Suppose the modeling sample data set is $\{(\boldsymbol{t}_i, y_i)\}_{i=1}^{n}$, where, $\boldsymbol{t}_i \in R^d$ is the input vector of the ith-dimension pivot element in the latent vector space expanded by the dth-dimension pivot element, and $y_i \in R$ is the target output variable of effluent COD of the papermaking wastewater. In the high-dimension feature space constructed by the nonlinear mapping function $\varphi(t)$, the model establishment between the output variable and the input variable is to find the best fitting function:

$$y(\boldsymbol{t}) = \boldsymbol{w}^{\mathrm{T}} \varphi(\boldsymbol{t}) + b \qquad (3)$$

Where, $\boldsymbol{w}$ is the weight coefficient vector to be estimated in the high-dimension feature space, b is the constant deviation term. For the LSSVM method, the parameter estimate in the above formula can be transformed to satisfy the constraint of Formula (4):

$$y_i = \boldsymbol{w}^{\mathrm{T}} \varphi(\boldsymbol{t}_i) + b + \xi_i, i = 1, 2, \cdots, n \qquad (4)$$

The minimization optimization problem was solved as below:

$$\min_{\mathbf{w}, b, \xi} J(\boldsymbol{w}, b, \boldsymbol{\xi}) = \frac{1}{2} \boldsymbol{w}^{\mathrm{T}} \boldsymbol{w} + \frac{1}{2} \gamma \sum_{i=1}^{n} \xi_i^2 \qquad (5)$$

In the formula, γ is a penalty factor, used to balance the complexity and approximation accuracy of the model, ξ_i is the training error of the ith sample point. The Lagrange multiplier α_i is now introduced to transform the above-mentioned constraint optimization problem of the formula into an unconstrained optimization problem:

$$L(\boldsymbol{w}, b, \boldsymbol{\xi}, \boldsymbol{\alpha})$$
$$= J(\boldsymbol{w}, b, \boldsymbol{\xi}) - \sum_{i=1}^{n} \alpha_i (w^T \varphi(\boldsymbol{t}_i) + b + \xi_i - y_i) \qquad (6)$$

Using the KKT optimization condition to solve the above formula (Zhou Xinran, 2012), that is, to solve the partial derivatives of $\boldsymbol{w}$, b, ξ_i and α_i, we can obtain:

$$\begin{cases} \dfrac{\partial L}{\partial \boldsymbol{w}} = 0 \rightarrow \boldsymbol{w} = \sum_{i=1}^{n} \alpha_i \varphi(\boldsymbol{t}_i) \\[2mm] \dfrac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^{n} \alpha_i = 0 \\[2mm] \dfrac{\partial L}{\partial \xi_i} = 0 \rightarrow \alpha_i = \gamma \xi_i, i = 1, 2, \cdots, n \\[2mm] \dfrac{\partial L}{\partial \alpha_i} = 0 \rightarrow \boldsymbol{w}^{\mathrm{T}} \phi(\boldsymbol{t}_i) + b + \xi_i - y_i = 0, i = 1, 2, \cdots, n \end{cases} \qquad (7)$$

Eliminating the elements from the above equation set, we will obtain the following linear equation set:

$$\begin{bmatrix} 0 & \mathbf{1}_v^{\mathrm{T}} \\ \mathbf{1}_v & \mathbf{K} + \gamma^{-1}\boldsymbol{I} \end{bmatrix} \begin{bmatrix} b \\ \boldsymbol{\alpha} \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix} \qquad (8)$$

Where, $\mathbf{1}_v = [1, 1, \cdots, 1_n]^{\mathrm{T}}$, $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_n]^{\mathrm{T}}$, $\boldsymbol{y} = [y_1, y_2, \cdots, y_n]^{\mathrm{T}}$, $K_{ij}(\boldsymbol{t}_i, \boldsymbol{t}_j) = \varphi(\boldsymbol{t}_i)^{\mathrm{T}} \varphi(\boldsymbol{t}_j)$, $i, j = 1, 2, \cdots, n$, and $\boldsymbol{I}$ is the unit matrix. After solving the parameters of α_i and b in Formula (8) and via the least square method, the LSSVM model will be obtained as below:

$$\hat{y} = f(\boldsymbol{t}) = \sum_{i=1}^{n} \alpha_i K(\boldsymbol{t}, \boldsymbol{t}_i) + b \qquad (9)$$

If the LSSVM model uses the RBF kernel function $K(\boldsymbol{t}, \boldsymbol{t}_i, \sigma) = \exp(-\|\boldsymbol{t} - \boldsymbol{t}_i\|^2 / \sigma^2)$, the different values of the kernel function width σ and the penalty factor γ in Formula (5) will affect the actual performance of the LSSVM model (Zhao et al., 2000). To this end, this paper completes the optimization of the two parameters by taking the minimum of the sum of squared error

$\sum_{l=1}^{n_v}(y_l - \hat{y}_l)^2$ between the experimental value y_l and the predicted value of the model of the effluent COD as the objective function, through the particle swarm optimization (PSO) (Kennedy et al., 1995), based on the validation sample set.

2.2.3.Model Parameter Adaptive Correction

In order to track the dynamic changes of the production process and maintain the prediction performance of the soft-sensor model in real time, an online calibration strategy (OCS) has been designed to iteratively update the soft-sensor model parameters in an adaptive manner. The basic idea of OCS is that if the established soft-sensor model is applied to the prediction of COD for a new sample individual, when the deviation between the experimental value y_{new} of the new sample individual and the predicted value $\hat{y}_{new}$ of the model exceeds the set error limit *maxe*, namely:

$$| y_{new} - \hat{y}_{new} |> maxe \qquad (10)$$

To Initiate the iterative update of the soft-sensor model parameters. The specific method is as follows: firstly, the sample individuals with the largest fitting deviation are retrieved from the training sample set and deleted; then, the sample individuals with the highest ranking in the monitoring sample set are transferred into the training sample set; next, the vacancy of the validation sample set is filled, namely, the sampled individuals with the highest ranking among the accumulated predicted sample individuals are transferred into the validation sample set; and finally, the soft-sensor model is re-established based on the newly formed training sample set and the validation sample set, which is namely the OCS - PCA-PSO-LSSVM model.
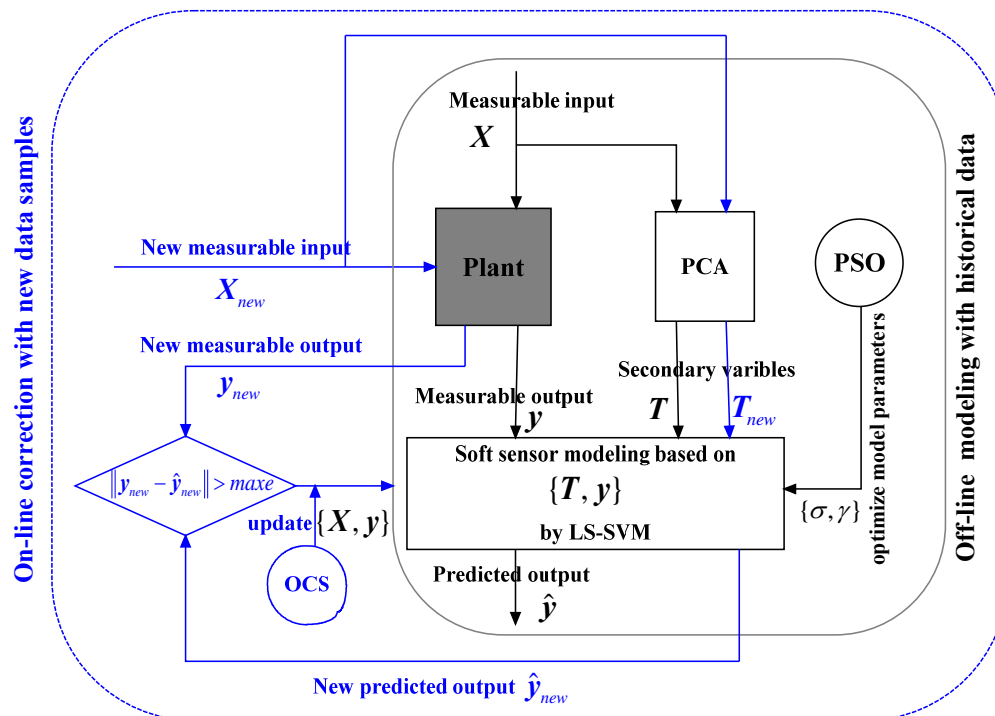

Fig 2. Structure of the OCS-PCA-PSO-LSSVM soft sensor model

With this, the implementation flow of the OCS-PCA-PSO-LSSVM method is shown in Fig. 2. Firstly, the PCA pre-processing of $X_{n\times p} = [x_1, x_2, \cdots, x_p]^T$ was performed and the pivot element matrix $T_{n\times d} = [t_1, t_2, \cdots, t_d]^T$ was obtained after the number d of pivot elements had been selected; then, based on $T_{n\times d}$ and the output variable matrix $y_{n\times1} = [y_1, y_2, \cdots, y_n]^T$ of the effluent COD, the nonlinear mapping relationship between them was established with the LSSVM method, while the values of

model parameters σ and γ were determined by PSO optimization; and finally, the OCS would be initiated to iteratively update the model in case the prediction deviation of the new sample individual was beyond the set error limit.

3.Results and Discussion

3.1. Model Performance Evaluation Indicator

To objectively and independently evaluate the performance of the OCS-PCA-PSO-LSSVM soft-sensor model, the initial sample data set was divided into a training sample set, a validation set, and a test set in time order, of which the training sample set contained 100 sample individuals, used for parameter estimation of the model; the validation sample set contained 50 sample individuals, used for parameter optimization of the model; and the test sample set contained the remaining 25 sample individuals, used for performance evaluation of the model. The performance evaluation indicators include: maximum deviation (MAXE)/mg•L^{-1}, maximum relative deviation (MAXRE)/%, mean absolute deviation (MAE)/mg•L^{-1}, mean relative deviation (MRE)/% ,root mean square error (RMSE) / mg • L^{-1}, standard deviation (STD) / mg • L^{-1}, etc., and their respective definition formula are as follows:

$$\text{MAXE} = \max_{i \in \{1,2,\cdots,n\}} \{|y_i - \hat{y}_i|\} \tag{11}$$

$$\text{MAXRE} = \max_{i \in \{1,2,\cdots,n\}} \{\frac{|y_i - \hat{y}_i|}{y_i}\} \times 100\% \tag{12}$$

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \hat{y}_i| \tag{13}$$

$$\text{MRE} = \frac{1}{n} \sum_{i=1}^{n} \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \tag{14}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2} \tag{15}$$

$$\text{STD} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (e_i - \overline{e})^2} \tag{16}$$

Where, $e_i = y_i - \hat{y}_i, \overline{e} = \frac{1}{n} \sum_{i=1}^{n} e_i$, while y_i and $\hat{y}_i$ denote the experimental value and predicted value of the model regarding COD of the ith sample individual, respectively. Among the above statistical performance indicators, MAXE, MRE, RMSE and STD are absolute accuracy indicators, of which, MAXE measures the limit boundary conditions of the model according to the maximum predicted deviation of the sample individuals, and MRE and RMSE measure the accuracy of the model according to the average prediction accuracy of the sample individuals. while STD measures the stability of the model according to the degree of dispersion of the prediction deviation of the sample individuals. Considering the objective difference between the magnitudes of different physical quantities, MAXRE and MRE are relative accuracy indicators. The former measures the deviation of the prediction results based on a single sample individual, and the latter does the same based on the average of sample individuals. The smaller the values of these statistics are, the better the performance of the model will be indicated.

3.2. Experimental Results and Analysis

As described in Section 2.2.1 above, in order to satisfy that the pivot element under the OCS-PCA-PSO-LSSVM method contain enough initial variable information, and the cumulative information contribution rate of the dth extracted pivot element is now required to be above 85%. Based on the information of the eight latent roots of the correlation matrix for the training sample set

of papermaking wastewater, when the $6t^h$ pivot element is extracted during calculation, namely, $d = 6$, the cumulative information contribution rate is 92.63%. Thus, the six pivot element are determined as the input vectors of the subsequent PSO-LSSVM model.

As described in Section 2.2.2 above, the optimization process for the algorithm of parameters γ and σ PSO under the LSSVM method is shown in Figs. 3 and 4 after the RBF radial basis kernel function was selected, and the population particle number was set at 30, the minimum inertia weight was $w_{min} = 0.01$, the maximum inertia weight was $w_{min} = 0.99$, the particle maximum velocity was $v_{max} = 2$, the particle minimum velocity was $v_{min} = -2$, and the learning factor was $c_1 = c_2 = 2$ and the maximum number of iterations was 100. When $\gamma = 0.3356$ and $\sigma = 2.2026$, the RMSE of the objective function observation sample set reached the minimum, thereby it was determined as the optimal value of the parameter under the LSSVM method.
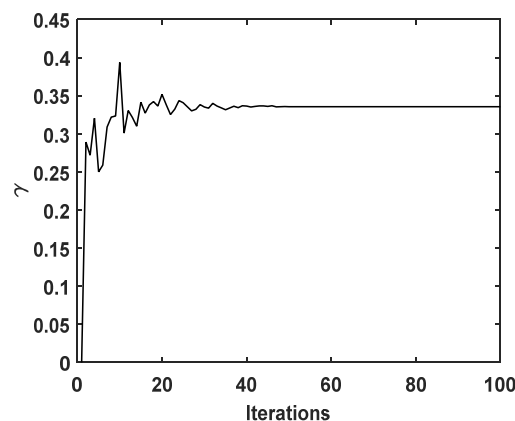

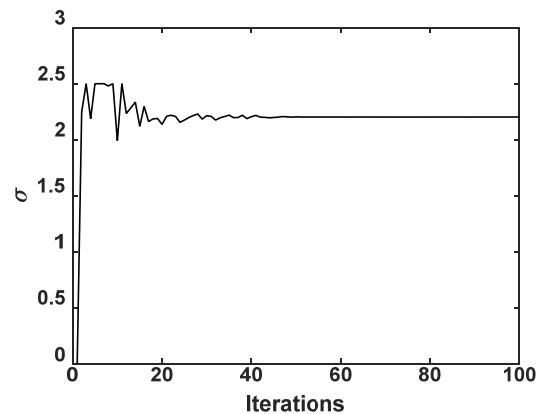**Fig 3.** Regularization factor optimizing curve using PSO

Fig 4. Kernel parameter optimizing curve using PSO

After the optimization for the input variable and parameter of the model was completed, the OCS-PCA-PSO-LSSVM model was applied to the test sample set to detect the model's generalization ability. Table 1 shows the test results of different performance indicators for the three models of OCS-PCA-PSO-LSSVM, PCA-PSO-LSSVM and SVM. It may be observed from the table that the values of the maximum deviation, the maximum relative deviation, the average absolute deviation, the average relative deviation, the root mean square error, and the standard deviation of the OCS-PCA-PSO-LSSVM model are significantly lower than the corresponding results of the PCA-PSO-LSSVM model and the SVM model. Where, compared with the PCA-PSO-LSSVM model, the MAXE of the OCS-PCA-PSO-LSSVM model decreased by 39.15%, the MRE decreased by 25.00%, and the STD decreased by 29.89%. The reason for this is that when the PCA-PSO-LSSVM model predicted the 2nd, 10th, 20th, and 21st sample individuals in the test sample set, their prediction deviations were greater than their respective maximum fitting deviation the training sample set, so the OCS strategy was initiated 4 times to perform the iterative update of the model, therefore, from the 2nd

sample individual of the test sample set containing 25 sample individuals, and the predicted value of the model was different from the predicted value of the PCA-PSO-LSSVM model without the OCS strategy integrated, which was generally reflected as the deviation tends to be small, thus achieving dynamic adjustment and optimization of the model.

Tab. 1 Model performances comparison on the testing data set

Methods	MAXE /mg·L^{-1}	MAXRE /%	MAE /mg·L^{-1}	MRE /%	RMSE /mg·L^{-1}	STD /mg·L^{-1}
SVM	54.39	7.82	18.62	2.58	22.96	13.71
PCA-PSO-LSSVM	51.09	7.53	17.37	2.40	21.57	13.05
OCS-PCA-PSO-LSSVM	31.09	6.26	12.31	1.80	15.23	9.15

To visually compare the prediction performance of the above three model methods, the experimental values and predicted values of COD of 25 sample individuals in the test sample set are plotted in Fig.5. Through observation of the figure, it can be seen that compared with the PCA-PSO-LSSVM and SVM model methods, the COD results on each sample individual predicted with the OCS-PCA-PSO-LSSVM model method are more closely to their respective experimental values, thereby indicating that the OCS- The PCA-PSO-LSSVM model method has better generalization prediction ability and stronger dynamic stability.
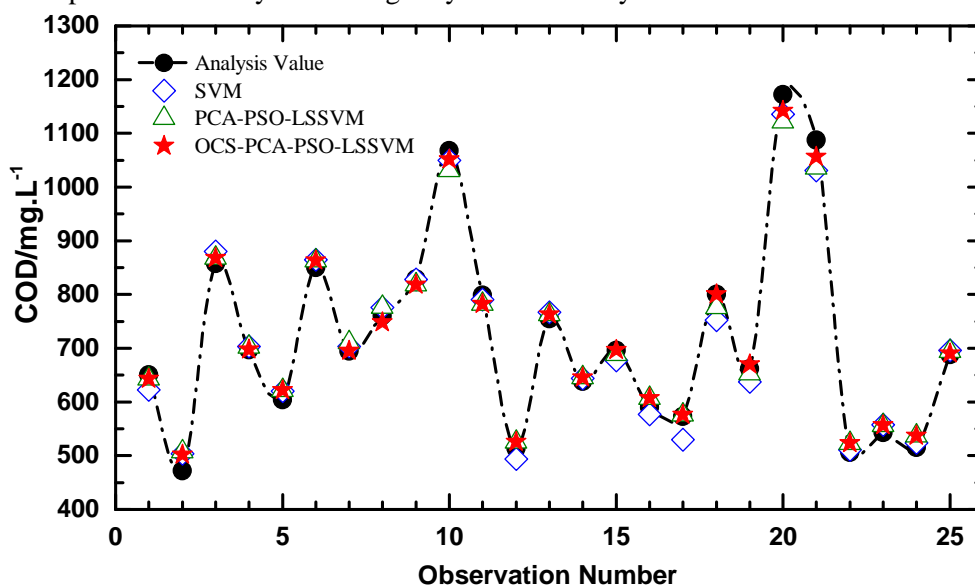

Fig 5. Prediction results of COD on the generalization data set

4. Conclusions

Made in China 2025 clearly pointed out "taking the deep integration of informatization and industrialization as the main line." Based on the safe and healthy management and high-efficiency production requirements in the anaerobic treatment process of papermaking wastewater, this paper focuses on the study of the soft-sensor prediction and dynamic optimization of the model based on the data-driving effluent COD as the water quality indicator, to promote the transformation of the paper industry from extensive development to sustainable development, from the end treatment to the resource utilization, promoting the intelligent management and control of the production process, the main conclusions are as follows:

1) In order to adapt to the structure of anaerobic reactor and the multivariable, nonlinear, time-varying features of the parameters, as well as the special complexity of papermaking wastewater process and the uncertainty of production behavior, the soft-sensor method integrating the modern data analysis technology and intelligent regression model have been developed and designed, which not

only effectively reduces the complex collinearity between variables, but also reduces the spatial dimension of the model, and the prediction accuracy and dynamic stability of the model are significantly improved, achieving the   overall improvement and breakthrough of the model performance by virtue of the integration advantages.

2) Data-driving soft-sensor model method: As the time series data continues to increase, the prediction accuracy of the model based on long-term historical data will decrease. Taking the actual industrial process as the background, combined with the dynamic change characteristics of the process, the method can adaptively iteratively update the model parameters through deviation feedback, and maintain the generalization performance of the soft-sensor model in real time, thus ensuring the continuous efficient and stable operation of the equipment, and monitoring the energy conservation and emission reduction as well as sustainable development of the enterprise.

References
[1]Bourgeois W, Burgess J E, Stuetz R M. 2010. On‑line monitoring of wastewater quality: a review[J]. Journal of Chemical Technology & Biotechnology, 76: 337-348
[2]Choi D J, Park H. 2001.A hybrid artificial neural network as a software sensor for optimal control of a wastewater treatment process[J]. Water Research, 35: 3959-3967
[3]Cristianini N, Shawe T J. 2000. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods[M], Cambridge University Press, 112-120
[4]Dürenmatt D J, Gujer W. 2012. Data-driven modeling approaches to support wastewater treatment plant operation. Environmental Modeling & Software. 30: 47-56
[5]Fortuna L, Graziani S, Rizzo A, et al.2007. Soft Sensors for Monitoring and Control of Industrial Processes[M]. Springer-Verlag: London, 34-45
[6]Haimi H, Mulas M, Corona F, et al. 2013. Data-derived soft-sensors for biological wastewater treatment plants: An overview[J]. Environmental Modelling & Software,   47: 88-107
[7]Huang Mingzhi, Ma Yongwen, Wan Jinquan, et al. 2015. A sensor-software based on a genetic algorithm based neural fuzzy system for modeling and simulating a wastewater treatment process[J]. Applied Soft Computing, 27: 1-10
[8]Jolliffe I T. 2002. Principal Component Analysis (second edition)[M]. Springer-Verlag, 168-176
[9]Kennedy J, Eberhart R C. 1995. Particle swarm optimization[C]. Proceedings of IEEE International Conference on Neural Networks. Perth, Australia,   1942-1948
[10]Langergraber G, Fleischmann N, Hofstaedter F, et al. 2004. Monitoring of a paper mill wastewater treatment plant using UV/VIS spectroscopy[J], Water Science and Technology, 49: 9-14
[11]Liu Bo, Wan Jinquan, Huang Mingzhi, et al. 2015. Online Prediction Model for Effluent VFA from Anaerobic Wastewater Treatment System Based on PCA-LSSVM[J]. Journal of Environmental Sciences, 35(6): 1768-1778
[12]Liu Lin, Ma Yiwen, Wan Jinquan, et al. 2017. Soft-sensor Model of Anaerobic Treatment Process of Wastewater Based on Pso-SVM[J], Journal of Environmental Science, 37(6): 2122-2129
[13]Ruggeri B, Tommasi T, Sanfilippo S. 2015.BioH2 & BioCH4 Through Anaerobic Digestion From Research To Full-Scale Applications[M]. Springer-Verlag, 1-24
[14]Souza F A A, Aráujo R, Mendes J. 2016. Review of soft sensor methods for regression applications. Chemometrics and Intelligent Laboratory Systems, 152: 69-79
[15]Sun Jun, Cheng Zhong, Yang Ruiqin, et al. 2017. PCA-PSO-LSSVM-based Soft-sensor of Effluent COD of the Anaerobic Treatment System for Papermaking Wastewater [J].Computer and Applied Chemistry, 34(9): 706-710
[16]Tang Wei, Bai Zhixiong, Gao Xiang. 2017. Dissolved Oxygen Concentration Control System Based on Adaptive Mutation Differential Evolution Algorithm[J]. Paper China, 36 (6): 49-54

[17] Wan Jinquan, Huang Mingzhi, Ma Yongwen, et al. 2011. Prediction of effluent quality of a paper mill wastewater treatment using an adaptive network-based fuzzy inference system[J]. Applied Soft Computing, 11: 3238-3246

[18] Wang Wenchuan, Liu Xinggao.2015.Melt index prediction by least squares support vector machines with an adaptive mutation fruit fly optimization algorithm[J]. Chemometrics and Intelligent Laboratory Systems, 141: 79-87

[19] Wang Yao, Xu Liang, Yin Wenzhi, et al. 2017. Soft-sensor Modeling of Papermaking Wastewater Treatment Process Based on ANN and LSSVR[J]. Transactions of China Pulp and Paper, 32(1): 50-54

[20] Xu Lisha, Qian Xiaoshan. 2012. A Study on Soft-sensor of Effluent COD Based on MUTATION CPSO ALGORITHM[J]. Chinese Journal of Environmental Engineering, 6: 1455-1458

[21] Yang Hao, Mo Weilin, Xiong Zhixin, et al. 2016. Soft-sensor Modeling of Papermaking Wastewater Treatment Process Based on RPLS[J]. Paper China, 35(10): 31-35

[22] Zhang Wei, Liu Min, Chen Wei, et al. 2014. Influence of External Circulation on the Operation Effect of IC Reactors[J]. Journal of Chemical Industry and Engineering, 65: 2329-2334

[23] Zhao Huiru, Huang Guo, Yan Ning. 2018.Forecasting energy-related CO2 emissions employing a novel SSA-LSSVM model: considering structural factors in China[J]. Energies, 11: 781-802.

[24] Zheng Rongjian, Pan Feng. 2017. Modeling for Concentration Prediction of Glutamic Acid Fermentation Products based on PLS-LSSVM[J]. Journal of Chemical Industry and Engineering, 68: 976-983

[25] Zhou Hongbiao, Qiao Junfei. 2017. Application of Hybrid Multi-objective Particle Swarm Optimization Algorithm in Optimal Control of Wastewater Treatment Process[J]. Journal of Chemical Industry and Engineering, 68: 3511-3521

[26] Zhou Xinran. 2012. Research on online Modeling and Control Method Based on Least Squares Support Vector Machine[D]. Changsha: Hunan University, 19-21