

The research of least squares support vector machine optimized by particle swarm optimization algorithm in the simulation MBR prediction

Weiwei Li^{1, a}, Chunqing Li^{2, b}, Jingyun Nie^{3, c}, Tao Wang^{4, d}

^{1,2,3}School of Computer Science and Software Technology, Tianjin Polytechnic University. China

⁴School of Environmental and Chemical Engineering, Tianjin Polytechnic University. China

^aliweiwei0725@163.com, ^b franklcq@163.com, ^c 1712862492@qq.com, ^d 1208215959@qq.com

Keywords: MBR; Membrane flux; PCA; LSSVM; PSO

Abstract. This paper proposes an intelligent algorithm to predict the MBR membrane flux. The algorithm applies the least squares support vector machine (LS-SVM) to the research of MBR simulation prediction, optimize the penalty factor and kernel parameters of LS-SVM model by particle swarm optimization (PSO) for avoiding the blindness of artificial selection parameter. Due to the complexity and cross-cutting of the factors that affect MBR membrane fouling, first of all, we analyze the factors by principal component analysis (PCA), extract the important factors as the LS-SVM input layer, MBR membrane flux as output layer, and then create PSO-LSSVM prediction simulation model. In the end, we get predictive results with the model. By comparing the predicted results with experimental data, the algorithm has higher prediction accuracy for MBR membrane flux. To further verify the effectiveness of the algorithm, we also compare the model with BP neural network model, the results show that the prediction model of PSO-LSSVM has a higher prediction accuracy.

Introduction

MBR, combining membrane separation technology with biotechnology, is a new technology which is applied in the field of sewage treatment. Large number of experimental studies have shown that MBR membrane fouling severely affect the performance of MBR technology and caused much energy consumption. The most direct consequence of membrane fouling is caused by decrease in flux, so the size of the membrane flux is important to characterize the degree of membrane fouling. Therefore, the study of membrane pollution factors, and the describing membrane fouling process with intelligent simulation model have become a hot topic in today's MBR simulation system.

Only when the number of samples tends to infinity, the forecasting methods based on the traditional statistics have good performance in theory. In most practical cases, the number of samples is limited, and even very small. Therefore, these methods are difficult to achieve the desired results. Compared with the traditional statistics, Vapnik et al established statistical learning theory (SLT), which is a theory of specialized research in machine learning rule with the small samples, and provides a unified machine learning problem framework when the sample is limited. Support Vector Machine (SVM) that proposed by Vapnik in 1995 is kind of machine learning methods based on statistical learning theory. With the in-depth study of Vapnik, it is extended to function regression estimation aspects and demonstrate excellent learning performance. SVM parameter selection problem has not been resolved, so the further application of SVM is severely limited. To this end, we propose a method that combines PSO with SVM. We search parameters of SVM with PSO to achieve optimal prediction model. The method is used to establish a least squares support vector machine (LS-SVM) forecasting model. The model is applied on MBR membrane flux predictions.

The Prediction Model of MBR Membrane Fouling

Principal Component Analysis(PCA).

Principal component analysis (PCA) method is usually used to simplify the process of solving multivariate high-dimensional complex system, it has been effectively applied to the statistical analysis, pattern recognition, image processing and communication technologies. Typically, there are many variables that affect MBR Membrane Fouling, these variables associate with each other, so there will be a measurement of the noise carried, and these variables have different degrees of importance. The workload can be greatly reduced when PCA is used in simulation prediction of MBR, and this also directly improves the effectiveness and reliability of the forecast. Specifically described as follows:

- (1) Select 6 contamination factors to compose the original input variable matrix:

$$X = \{MLSS, \text{Resistance}, \text{Pressure}, \text{COD}, \text{PH}, \text{Temperature}\}$$

- (2) Normalize X matrix with Standardized methods of mean zero to obtain a new matrix A:

$$A_{ij} = (X_{ij} - \bar{X}_j) / s_{ij}, i = 1, 2, \dots, I, j = 1, 2, \dots, n.$$

Where i is the number of samples, j is the number of components of samples, s_{ij} is the standard deviation of variable X_j

- (3) Find covariance matrix S of A matrix, arrange the characteristic roots λ in descending order to obtain the corresponding eigenvalues matrix V of the S, and the corresponding eigenvectors U.
- (4) Decompose the matrix: $A = TU^T$, Where T is the main component of the matrix.
- (5) Analyze eigenvalues matrix V to get minimum number of main element: $n = 3$ when element cumulative contribution rate is more than 89% .
- (6) Select properly 3 main elements in the matrix T as the input vector of least squares support vector regression machine: $A = \{MLSS, \text{Resistance}, \text{Pressure}\}$

The build of Prediction Model of MBR Membrane Fouling.

The MBR membrane flux is an important indicator of the degree of contamination of the membrane, so we take the flux as the output of the model. There are a lot of factors that affect flux. Firstly, we use principal component analysis method to extract the important factors, then train experimental data with PSO-LSSVM model. Finally, we get membrane pollution simulation prediction model that is mainly composed of two parts .As is shown in Figure 1.

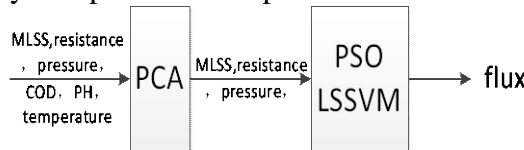


Fig. 1: The Prediction Model of MBR Membrane Fouling

The theory of squares support vector machine regression

LS-SVM is a kind of SVM algorithm based on regularization theory that is proposed by Suykens et al. It introduces the squared term in the standard SVM optimization problem, chooses the two-norm in the optimization target as the loss function and changes the inequality constraints into the equation constraints. LS-SVM takes least-squares linear system as the loss function and transforms the classic quadratic programming optimization problems into solving linear equations, which greatly reduces the computational complexity of SVM, improves the efficiency of solving SVM and reduces the SVM learning difficulty with fast processing speed. For nonlinear regression problem, given a training set of M data points

$(x_1, y_1), (x_2, y_2), \dots, (x_i, y_i), \dots, (x_m, y_m)$ $x_i \in R^d, y_i \in R$, where x_i is the input vector, y_i is the target value. m is the number of samples set. Map samples into a high-dimensional space through a

nonlinear function ϕ , then use linear regression, regression function is:

$$f(x_i) = \omega^T \phi(x_i) + b \quad (1)$$

Where $\omega = (\omega_1, \omega_2, \dots, \omega_m)$ is weight coefficient vector, b is the basis term. The use of structural risk minimization principle (SRM) minimizes risk, the optimization problem of function approximation for least squares support vector machines is:

$$\min J(\omega, \xi) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^m \xi_i^2 \quad (2)$$

$$s.t. y_i = \omega \phi(x_i) + b + \xi_i, i=1, 2, \dots, m$$

Where C is the error penalty function; ξ_i is slack variables. Lagrange function L is constructed:

$$L(\omega, b, \xi, a) = \frac{1}{2} \omega^T \omega + \frac{1}{2} C \sum_{i=1}^m \xi_i^2 - \sum_{i=1}^m a_i \left\{ \omega^T \phi(x_i) + b + \xi_i - y_i \right\} \quad (3)$$

Where a_i is lagrange multiplier, because all the optimization problem must meet KKT (Karush-Kuhn-Tucker) conditions in any constraints, we can get the following equations:

$$\begin{cases} \frac{\delta L}{\delta \omega} = 0 \rightarrow \sum_{i=1}^m a_i \phi(x_i) \\ \frac{\delta L}{\delta b} = 0 \rightarrow \sum_{i=1}^m a_i = 0 \\ \frac{\delta L}{\delta a_i} = 0 \rightarrow a_i = C \xi_i \\ \frac{\delta L}{\delta \xi_i} = 0 \rightarrow \omega^T \phi(x_i) + b + \xi_i - y_i = 0 \end{cases} \quad (4)$$

ω and ξ_i are erased, the results are as follows:

$$\begin{pmatrix} 0 & Q^T \\ Q & K + C^{-1} I \end{pmatrix} \begin{pmatrix} b \\ A \end{pmatrix} = \begin{pmatrix} 0 \\ Y \end{pmatrix} \quad (5)$$

$$\text{where } Q = (1, \dots, 1)^T, A = (a_1, a_2, \dots, a_m)^T, Y = (y_1, y_2, \dots, y_m)^T$$

According to the conditions of Mercer, the kernel function can be determined:

$$K(x_i, x_j) = \phi(x_i)^T \phi(x_j) \quad (6)$$

The function estimation of LSSVM as follows:

$$f(x) = \sum_{i=1}^m a_i K(x, x_j) + b \quad (7)$$

There are some kernel functions that are used as follows:

- (1) Linear kernel: $K(x, y) = x \cdot y$
- (2) Polynomial kernel: $K(x, y) = (x \cdot y + 1)^d$
- (3) RBF kernel: $K(x, y) = \exp\{-\|x - y\|^2 / 2\sigma^2\}$

In this paper, we take radial basis function (RBF) as the kernel function, where σ is kernel width. According to LSSVM regression theory, we know that its main parameters are the kernel function parameter σ and the penalty parameter C , these two parameters can influence greatly learning and generalization of LSSVM. So we apply PSO algorithm on optimizing choice process of the parameters to reduce blindness of subjective experience choice in some degree.

The algorithm of particle swarm optimization (PSO)

PSO algorithm based on groups and fitness is kind of optimization algorithm that has the ability of global optimization. In the PSO, Any possible solution of the optimization problem is regarded as

a position in the search space, referred to "particles". All of the particles have two characters of position and speed and have a fitness that is decided by optimization function .PSO is initialized to a group of random particles (random solutions), and then find the optimal solution by iteration. In every iteration, the particles update themselves by tracking the two extreme values. The first one is the optimal solution found by the particle itself, this solution is called individual extreme value p_{best} . The other extreme value is the optimal solution of the entire population, this solution is a global extreme called p_{gbest} . After the two extreme values are calculated, each particle updates its speed and position according to following formula:

$$v_i(t+1) = \omega v_i(t) + c_1 rand_1(p_{best}(t) - x_i(t)) + c_2 rand_2(p_{gbest}(t) - x_i(t)) \quad (8)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (9)$$

where t represents the t-th iteration, ω is the inertia weight, c_1 and c_2 are learning factors, also called acceleration coefficient. $rand_1$ and $rand_2$ are random numbers in the range of [0,1]. The right of formula (8) consists of three parts: The first one is the inertial part, reflects the exercise habits of particles, which means the particles have a tendency to maintain their previous speed. The second one is the cognitive part, reflects the memories of their own historical experience of particles, which means the particle has the tendency to approach its own historical best position. The third one is the social part, reflects the group historical experience of collaboration and knowledge sharing between the particles, which means particle has the tendency to approach best position. According to Formula (9), we can calculate the new position of the particle i.

The LSSVM flux prediction model based on PSO optimization

In this paper, the steps of algorithm to predict the membrane flux with PSO-LSSVM as follows:

- (1) Determine the ranges of penalty coefficient C and kernel parameter σ of the least squares support vector regression model. Where penalty parameter ranges [0.1,1000], kernel parameters σ of radial basis function ranges [0.01,100].
- (2) Take $\{C, \sigma\}$ as the parameters of PSO, initialize PSO algorithm parameters, initialize randomly the particle to form a group of the particles, randomly generates the initial velocity of the particles, where the number of particles is N = 30.
Fitness function as follows:

$$f_i = \frac{1}{len} \sum_{j=1}^{len} (y_{ij} - y_j)^2 \quad (10)$$

- (3) Use the least squares support vector regression procedures to train these particles after they are initialized, get individual fitness values, and update the global optimum value and the optimal value of the individual particles.
- (4) Determine the termination condition, if the maximum number of iterations is reached, then stop; otherwise, produce a new group according to the velocity equation (8), go to step (3), until the termination requirements. At this point, the individual particle in group that has smallest fitness value is optimal solution that we are seeking.
- (5) Use the optimal parameter penalty factor C and kernel parameter σ in least squares support vector regression procedures, use test sample to get predictions.

The comparison and analysis between experimental results and predicted results

We select 75 samples to train with the prediction model, 10 samples taken for testing. The number of particles is 30, the maximum number of iterations is set to 300. Inertia weight ω with iterations t increases linearly decreases from 0.9 to 0.4. Relational expressions for ω and t as follows :

$$\omega=0.9-(t/T_{\max})\cdot0.5 \quad (11)$$

Where T_{\max} is maximum number of iteration .

When acceleration factor $c1$ is very large, the particles will vibration too much in the local area, but when the acceleration factor $c2$ is very large, the particles will lead to premature convergence to a local minimum. In the end , We take the acceleration factor $c1 = c2 = 1.5$.

The program is written by MATLAB. We get two important parameters that are $C=19.9939$ and $\sigma=0.1$ through training train samples in pso-lssvm prediction model. We use the model to forecast membrane flux with test samples. The results are shown in Table 1.

Test curve is shown in Figure 2. As shown in Table 1 and Figure 2, the LS-SVM prediction model that is optimized by PSO has high prediction accuracy .The average absolute error of measurement is 0.776. The average relative error is 2.7%, the maximum relative error is 3.8% and less than 4%.So the experiment achieved relatively good results.

In order to verify the effectiveness of the algorithm, we also establish BP network forecasting model with the same samples to predict the MBR membrane flux. We compare the predicted results with the predicted results of PSO-LSSVM , as shown in Figure 3. We can learn that the prediction model of LSSVM based on PSO can more accurately predict the membrane flux than BP network forecasting model through Figure 3. It indicates PSO-LSSVM can well predict the membrane flux.

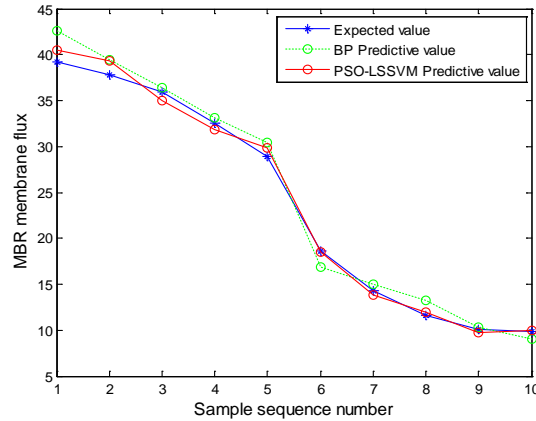


Fig. 2: The Fitting Result of PSO-LSSVM for MBR

Table 1: the Comparison of Prediction Result and MBR Experimental Result

Index	Expected value	Predictive value	Absolute error	Relative error
0	39.2	40.52	1.32	0.034
1	37.8	39.29	1.49	0.038
2	35.9	35.03	0.87	0.024
3	32.5	31.79	0.71	0.022
4	28.9	29.80	0.9	0.031
5	18.6	18.48	1.12	0.007
6	14.3	13.83	0.47	0.033
7	11.6	11.92	0.32	0.028
8	10.1	9.76	0.34	0.034
9	9.8	10.02	0.22	0.022
The mean of relative error				0.027

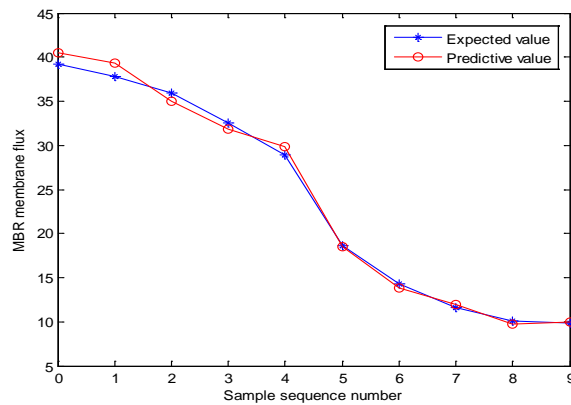


Fig. 3: The Comparison Between PSO-LSSVM Model and BP neural network Model

Conclusion

Because the influence of membrane fouling factors are very complex and cross-cutting between the various factors, firstly use principal component analysis to refine related input variables, get the factors that have greater impact on the MBR membrane flux, then take these important variables as the input layer of least squares support vector machines. Least squares support vector machine has a strong learning ability and good generalization performance, it can solve the limited sample, nonlinear and local minima problems, so this method provides a new route for the forecast of the MBR membrane flux, but its effectiveness depends largely on selecting predictive model parameters. Particle Swarm Optimization as a bionic intelligent algorithm has attracted widespread attention in recent years with good parallelism, robustness and global search and other features. Based on the above analysis, we combine the two algorithms and propose least squares support vector machine flux prediction method (PSO-LSSVM) based on parameter optimization of particle swarm optimization. The comparison between the predicted and experimental results shows that the proposed prediction model can well predict the MBR membrane flux; and the experimental data also show that this method of training SVM has higher prediction accuracy than BP network prediction model.

References

- [1] Han Yongping, Xiao Yan, Song Lei, Wang Xiaolin. Progress of MBR membrane fouling formation and its influence factors[J]. Membrane Science and Technology, 2013, 01: 102-110.
- [2] J.A.K. Suykens, J. Vandewalle, Multiclass least squares support vector machines, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN 99), Washington, DC, 2002, pp. 900-903.
- [3] Kenney J, Eberhart R C. Particle Swarm Optimization[C]. Proc IEEE international conference on Neural Networks. Perth, USA, 1995: 123-125.
- [4] Suykens J.A.K., Vandewalle J. Least squares support vector machine classifiers, Neural Processing Letters, 1999, vol. 9, no. 3, Jun. pp. 293-300.
- [5] Y. Q. Chen, Z. X. Zhou, and Q. J. Chen, "The Research and Application of LS-SVM Based on Particle Swarm Optimization," Proceedings of IEEE International Conference on Automation and Logistics, pp. 1151-1160, Aug. 2007.
- [6] D.F. Shi, N.G. Nabil, "Tool wear predictive model based on least squares support vector machines", Mechanical Systems and Signal Processing, vol. 21, 2007, pp. 1799-1814.