Water Treatment Analysis Based on Support Vector Machine Regression Algorithm

XiaoMan Guan*
Shandong Institute of Commerce and Technology Shandong Jinan 250103 China

*Corresponding author e-mail: gxm1949@126.com

Abstract. In the current sewage treatment process, the cost of sewage treatment is very high. From the practical problems encountered in the water treatment process, starting with the typical amount of scale buildup prediction, an intelligent prediction algorithm of support vector machine regression algorithm is introduced in this paper as the basis of modeling to analyze the prediction of the typical amount of scale buildup issue encountered in the water treatment field. Through empirical analysis, the specific data obtained from the single-factor experiment is introduced into the support vector machine regression algorithm for prediction operation. The experimental results show that the prediction accuracy is significantly improved compared with that of the traditional method.

Keywords: Power Plant Chemistry, Support Vector Machine Regression Algorithm, Amount of Scale Buildup, Water Treatment

1. Introduction

Human beings are inseparable from water. Water is closely related to our lives, which is our most precious natural resource and affects all aspects of our lives [1-2]. However, with the rapid development of modern society, people's demand for water is increasing. China is a country with inferior freshwater resources, which is not only reflected in the shortage of water resources, but also in the geographical distribution of freshwater is highly uneven [3-4]. According to relevant data, currently, there are more than 300 cities in 684 cities in China whose residents are affected by freshwater insufficiency, and most of these cities are distributed in the north. Meanwhile, with the continuous acceleration of the pace of industrialization in China, the phenomenon of severe water pollution has made the situation of water resources in China worse [5]. According to statistics, the total discharge of urban wastewater in China is about 60 billion m$^3$ / year, of which industrial sewage is the main one[6], and the proportion is as high as 70070. Currently, the treatment rates of urban sewage and industrial sewage in China are only 1020% and 79070, respectively, and the discharge rate of industrial wastewater is only 55070. A
large part of industrial sewage and urban wastewater are discharged when they are not up to standard or treated Rivers cause severe pollution to the environment and threaten the production of industry and agriculture and the health of the people. It is estimated that the annual economic loss caused by water pollution in China has reached 40 billion yuan. With the decrease of the total amount of freshwater and the rising cost of water, how to solve the contradiction between the shortage of water resources and the demand for people's life is one of the significant challenges our country faces in the new century.

In the water treatment process, scale buildup prediction is often encountered. How to ensure the prediction accuracy and provide a reliable guarantee for the stable operation of the power plant? Traditional methods use qualitative analysis methods such as judgment index to estimate from the perspective of experience. Some scholars have also proposed the method of neural network prediction, but the neural network is prone to fall into the local minimum value. Therefore, a support vector machine regression algorithm is introduced in this paper, which can effectively overcome the shortcomings of the neural network and ensure the accuracy of prediction.

2. Support vector machine regression algorithm model
SVM method is a useful practice of statistical learning theory (SLT). It is based on the VC (vapnikchervonenkis) theory and structural risk minimization (SRM) principle of SLT. According to the limited sample information, it seeks the best fit between the complexity of the model and learning ability, in order to obtain better generalization ability, which is obviously superior to the traditional experience-based risk minimization. The neural network method based on the principle of transformation can ensure that the solution obtained is the optimal global solution, and can better solve practical problems such as small sample size, non-linear, high-dimensional data, etc. It has the advantages of simple structure, global optimization, and good generalization ability.

The basic idea of SVM is through nonlinear mapping $\varphi(*)$ The input vector is mapped to the high-dimensional feature space $H$, the optimal decision function is established by using the principle of structural risk minimization, and the kernel function of the original space is used to replace the point product operation in the high-dimensional feature space. The optimal regression function constructed by SVM in high dimensional space $H$ is as follows:

$$f(x) = w^* \varphi(x) + b$$  \hspace{1cm} (1)

Where $W$ stands for the weight vector, $w \in \mathbb{R}^k$; 
$B$ standards for the intercept, $b \in \mathbb{R}$.

Currently, SVM for regression analysis mainly includes $\varepsilon$-SVR and $\nu$-SVR. The two algorithms are the same and the same, which is illustrated by $\nu$-SVR in this paper. $\nu$-SVR transforms regression analysis into the following optimization problems:
\[
\min_{w,b,\xi^-_i,\xi^+_i} \frac{1}{2} w^T w \left[ C \left( \sum_{i=1}^{l} (\xi^-_i + \xi^+_i) \right) \right]
\]
\[
y'(x') \leq y_i + \xi^-_i
\]
\[
f(x') - y' \geq \varepsilon + \xi^+_i
\]
\[
\xi^-_i, \xi^+_i \geq 0 \quad i = 1, L, l
\]

Where \( C \) represents the penalty factor to balance the maximum classification boundary and the minimum training error; \( V \) represents the number of control support vector machine regression algorithms; \( l = 1, L, l \) represents the insensitive loss function.

The dual problem of the quadratic programming problem is as follows:

\[
\min_{\alpha, \alpha^*} \frac{1}{2} (\alpha - \alpha^*)^T Q (\alpha - \alpha^*) + y^T (\alpha - \alpha^*)
\]
\[
e^T (\alpha - \alpha^*) = 0
\]
\[
\alpha^T (\alpha + \alpha^*) \leq Cv
\]
\[
0 \leq \alpha_i, \alpha^*_i \leq C/l \quad i = 1, L, l
\]

\( Q_{ij} = K(x_i, x_j) = \varphi(x_i)^T \varphi(x_j) \), \( K(x, y) \) kernel function, \( \varphi(*) \) represents the projection function of the input vector in the high dimensional space.

3. Application of support vector machine regression algorithm to water treatment

The activated sewage treatment process is currently widely used in China. The process can be described as follows: Firstly, the sewage is put into the primary sedimentation tank, filtered and precipitated to remove large volume visible pollutants. Subsequently, it enters the aeration tank. The
aeration tank is the core part of the activated sewage treatment method. It contains not only the sewage from the primary sedimentation tank but also the sewage from the hydrolytic acidification device, the activated sludge from the secondary sedimentation tank and the compressed air from the blower end mix thorough in the aeration tank. The organic pollutants are separated and degraded under the action of activated sludge for the purification of the sewage.

As a new intelligent algorithm, the support vector machine regression algorithm is more accurate and applicable than the traditional neural network method in simulating the working model of the human brain, parallel distributed processing model, and self-learning capacity. It can accurately and quickly deal with the relevant evaluation and prediction problems in the field of non-linear and complex power plant chemistry.

The comparative analysis of monthly inflow and outflow of a sewage treatment plant is as follows, as shown in Figure 2.

The analysis of the above chart shows that the monthly difference between the inflow and outflow of a sewage treatment plant is about 60000 tons, which is mainly due to the backwater, including the following:

1. Backflow of domestic water in the plant;
2. Backwater of sand water separator of the fine grid;
3. Backwash backwater of filter and dynamic flow sand filter;
4. The mud water separator in the desliming room and part of the backwater in the PAM process.

![Figure 2. Comparison of annual inflow and outflow of a sewage treatment plant](image)

In this paper, the prediction of scale buildup amount of hydraulic ash flushing pipeline often encountered in water treatment is taken as an example to establish a prediction model based on support vector machine regression algorithm and further improve the prediction accuracy, to provide data support and guarantee for the safe and stable operation of the power plant. The specific model is as follows. The support vector machine regression algorithm is applied to the prediction of scale buildup amount in the hydraulic ash flushing pipeline of water treatment, which breaks through the
shortcomings that the existing scale buildup judgment index of ash flushing pipeline can only reflect the scale buildup trend qualitatively, but cannot judge the scale buildup amount quantitatively, and breaks through the shortcomings that the traditional neural network method is easy to sink into the local minimum value and the prediction error is excessively large. The factors affecting the scale buildup of the ash flushing pipeline mainly include the nature, particle size, time, and flow pattern of the ash sample. The main influencing factors are screened out according to the actual situation and the single-factor simulation experiment is performed for each influencing factor, to obtain the action mode and law of each influencing factor on scale buildup in the water treatment process.

4. Simulation test

There are 15 groups of test data in this paper, which are the results of the orthogonal test and single-factor experiment respectively. Based on the orthogonal test and single-factor experiment, data samples including six main influencing factors, namely, carbonate content and pH value of raw water for ash washing, fCaO content, ash water flow rate, ash permanent contact time, ash water ALK, particle shape (particle size) are constructed as input vectors, that is, independent variable vector set, with the final amount of scale as input vector For the output vector, i.e., dependent variable, SVM rolling training is carried out, in which RBF inner product function is used as kernel function, parameter $\sigma^2=20$, $C=10$, and scale prediction value is obtained by MATLAB programming using input samples and output samples (Note: All the study and test data of support vector machine regression algorithm come from the results of experimental research).

The experimental results show that the prediction results obtained by the SVM regression algorithm are close to the experimental results, and the prediction relative error is small, which fully shows that the SVM regression algorithm method is feasible for predicting the scale buildup amount, which has significant improvement compared with high prediction deviation in a single prediction test. Where more relevant factors are introduced in the operation process, the prediction accuracy will be significantly improved.

5. Conclusions

Sewage treatment project is not only a basic project in the process of urban development but also plays a vital role in the development of the whole city. For a city, whether in infrastructure construction or in the control of water pollution, sewage treatment-related projects not only play a positive role in construction but also play an essential role in urban operation function. To ensure the smooth implementation of the sewage treatment scheme, specific standards are required in the design process of the sewage treatment project: to adapt to local conditions, effectively combine the concept of the treatment scheme and the characteristics of the sewage, and effectively select a relatively reasonable and scientific treatment process, to help implement the operation with the optimal process scheme. Where the support vector machine regression algorithm is applied to predict the scale of the hydraulic ash washing pipeline in the water treatment process, compared with traditional prediction, it not only requires no complicated reaction modes such as chemical balance and material balance but has also simplified the intermediate deduction links. The results show that the scale buildup regularity identified through the prediction model superimposed budgeting and establishment of the SVM model according to the actual operation condition has implemented the simultaneous interpreting of the ash
washing pipeline, achieved the purpose of accurate prediction of scale buildup, and provided satisfactory results for the safe, stable and reliable operation of water treatment in power plants.

References
[1] Jia, Y. , & Wang, H. . (2014). Image fire detection system based on feature regression analysis and support vector machine. Advances in Intelligent Systems and Computing, 279, 145-153.
[2] Metya, S. , Mukhopadhyay, T. , Adhikari, S. , & Bhattacharya, G. . (2017). Chapter 7 - efficient system reliability analysis of earth slopes based on support vector machine regression model. Handbook of Neural Computation, 127-143.
[3] Wu, Q. , Law, R. , Wu, E. , & Lin, J. . (2013). A hybrid-forecasting model reducing gaussian noise based on the gaussian support vector regression machine and chaotic particle swarm optimization. Information Sciences, 238(7), 96 – 110.
[4] Li, X. , & Ma, S. . (2015). Research on building ecotourism warning sensor system model based on support vector machine and grey relational analysis. Sensor Letters, 13(2), 169-175.
[5] Wang, Y. , Ding, G. , & Liu, L. . (2015). A regression forecasting model of carbon dioxide concentrations based-on principal component analysis-support vector machine. Communications in Computer and Information Science, 482, 447-457.
[6] Al-Dhaifllah, M. , & Westwick, D. T. . (2013). Identification of auto-regressive exogenous hammerstein models based on support vector machine regression. IEEE Transactions on Control Systems Technology, 21(6), 2083-2090.