

Modelling of sludge discharge based on PSO-RVM

Luo Long ^{1, 2}

¹Mechanical & Electrical Department , Guangzhou Institute of Technology,
No. 465, HuanShi Road, Yuexiu Dist., Guangzhou 510075, P. R. China

²College of Railway Transportation, Guangzhou Railway Polytechnic
No. 100, Qinglong Road, Baiyun Dist., Guangzhou 510430, P. R. China
gtluolong@163.com

Keywords: sludge discharge; PSO; RVM

Abstract. The sludge discharge model in wastewater biochemical treatment is highly nonlinear. Accurate prediction of sludge discharge will provide basic data for the precise control of subsequent sludge dewatering process. In the research based on this method, due to the relevance vector machine (RVM) has high sparseness and uses probability factor in predict, which is superior to the support vector machine (SVM) in the sludge discharge. However, the kernel function parameters of RVM are estimated by experience. Therefore, a kind of RVM method based on the particle swarm optimization (PSO) algorithm is proposed, which adopts the PSO algorithm to determine the kernel parameter of RVM, then builds RVM model and uses forecast the final sludge discharge.

Introduction

The activated sludge biological treatment system will produce some activated sludge every day, which will increase the total amount of sludge in the system. To keep the total amount of sludge is balanced basically, on the one hand, it must be regularly discharge part of residual activated sludge, biochemical pool mixture at the same time in order to maintain stable sludge concentration, need to return part of the sludge into the biochemical pool, so the excess sludge discharge in the activated sludge process control is important for an operation. By regulating the amount of mud, can change in the activated sludge microbial species and growth speed, can change the oxygen demand, can improve sludge settling property, at the same time, it will directly affect the subsequent sludge dewatering process, thus changing the function of the system. However, the sludge discharge system is a highly nonlinear system with multiple variables and random influence factors, which brings great difficulties to the modeling and prediction of the system.

In order to realize the wastewater treatment of the nonlinearity, large time delay object measurement and control, experts at home and abroad a variety of intelligent modeling method is proposed: fuzzy model, neural network and fuzzy neural network modeling, modeling of the main object is the aeration tank or the activated sludge system, which based on neural network modeling tools are the most active method based on knowledge of the research. Neural network based on the principle of empirical risk minimization, the sample is limited, the learning process easily through learning, dimension disaster, local minima, and the relevance vector machine has the strict theoretical and mathematical basis, based on structural risk minimization principle, generalization ability is superior to the former, algorithm has global optimality, and is in view of the small sample statistic theory. Especially suitable for solving small sample, nonlinear, high dimension, local minimum, etc. In this paper, the residual sludge discharge of PSO+RVM is proposed and the model of residual sludge emission is established.

Relevance vector machine (RVM)

Relevance vector machine (RVM), proposed by M. E. Tipping, is a sparse probabilistic model similar to support vector machine (SVM). It is based on the active decision theory, to obtain sparse

model by removing irrelevant points from the structure of prior parameters under the framework of Bayesian. Compared with support vector machine (SVM), relevance vector machine(RVM) has the following advantages: (1) subjective setting error parameters is avoided; (2) the relevant vector is less than the SVM; (3) kernel function need not meet the Mercer condition, resulting in a greater range of options⁷.

For a given data set $\{x_i, t_i\}_{i=1}^l$, $x_i \in R^d$, $t_i \in R$, d is the number of dimension of vector data set. In the experiment, the target value is affected by noise, so it is defined as $t = y(x, w) + e$, in which, e is the noise that obeys Gaussian distributions with 0 mean error and variance S^2 , and its probability is $p(t_i | y(x_i), \sigma^2) = N(y(x_i, w), \sigma^2)$. It is defined that $y(x, w) = \sum_{i=1}^l w_i j(x, x_i) + w_0$, in which, l is the size of data set,

$t = (t_1, t_2, \dots, t_l)$, $x = (x_1, x_2, \dots, x_l)$, w is weight vector, $j(x, x_i)$ is kernel function. In this paper, common radial basis kernel function is used. Assume independent identically distributed, the likelihood function of the whole data set can be expressed as:

$$p(t | w, S^2) = (2\pi S^2)^{-\frac{l}{2}} \exp \left\{ -\frac{1}{2S^2} \|t - \Phi w\|^2 \right\} \quad (1)$$

Where, $w = (w_0, w_1, \dots, w_l)$ Φ is inputted kernel function mapping, $\Phi = [\phi(x, x_1) \ \phi(x, x_2) \ \dots \ \phi(x, x_l)]^T$. In order to improve the generalization ability of the model under the bayesian framework, maximum likelihood method is used to train model weights w . It is defined that each weight obeys Gaussian prior probability distribution, and its expression is:

$$p(w | a) = \prod_{i=1}^l N(w_i | 0, \alpha_i^{-1}) \quad (2)$$

in this formula: $a = (a_1, a_2, \dots, a_l)$ is the prior distribution function of weight w . Based on the combination of formula (3) with formula (4), and according to the rule of Bayesian, the posterior probability distribution of the weight is calculated as:

$$p(w | t, a, \sigma) = N(\mu, \Sigma) \quad (3)$$

where, $\Sigma = (\alpha^{-2} \Phi^T \Phi + A)^{-1}$, $\mu = \alpha^{-2} \Sigma \Phi^T t$, $A = D(\alpha_1, \alpha_2, \dots, \alpha_l)$, among of which, D is the matrix composed of eigenvalue $(\alpha_1, \alpha_2, \dots, \alpha_l)$. The formula (4) indicates that the posterior distribution of weight is determined by mean value μ and Σ . To estimate weight model, the hyper-function optimal value a should be first estimated and determined. Under bayesian framework, the likelihood distribution of hyper-function could be calculated through the following formula:

$$p(t | a, S^2) = (2\pi)^{-\frac{l}{2}} \left| S^2 I + \Phi A^{-1} \Phi^T \right|^{-\frac{l}{2}} \bullet \exp \left\{ -\frac{1}{2} t^T (S^2 I + \Phi A^{-1} \Phi^T)^{-1} t \right\} \quad (4)$$

By solving the maximum likelihood distribution, hyper-function optimal value a_{MP} and S_{MP} can be obtained. So far, the model of target value t is constructed. For input value x_* , the probability distribution of its corresponding output is:

$$p(t_*, | t, a_{MP}, S_{MP}^2) = N(t_* | y(x_*, w), S_*^2) \quad (5)$$

where, the vector t_* is the predicted value of x_* , and its mean value is $y(x_*, w) = \mu^T \phi(x_*)$, variance shows its uncertainty, and its formula is $\sigma_*^2 = \sigma_{MP}^2 + \phi^T(x_*) \Sigma \phi(x_*)$.

Particle Swarm Optimization

The particle swarm is a population-based stochastic algorithm for optimization which is based on social-psychological principles. Unlike evolutionary algorithms, the particle swarm does not use selection; typically, all population members survive from the beginning of a trial until the end. Their interactions result in iterative improvement of the quality of problem solutions over time.

A numerical vector of D dimensions, usually randomly initialized in a search space, is conceptualized as a point in a high-dimensional Cartesian coordinate system. Because it moves around the space testing new parameter values, the point is well described as a particle. Because a number of them (usually $10 < N < 100$) perform this behavior simultaneously, and because they tend to cluster together in optimal regions of the search space, they are referred to as a particle swarm. Each particle keeps track of its coordinates in the problem space which are associated with the best solution (fitness) it has achieved so far. (The fitness value is also stored.) This value is called $pbest$. Another "best" value that is tracked by the particle swarm optimizer is the best value, obtained so far by any particle in the neighbors of the particle. This location is called $lbest$. when a particle takes all the population as its topological neighbors, the best value is a global best and is called $gbest$. In each iteration, the particle updates its speed and position through the individual extremum and the group extremum.

$$v' = w * v + c1 * rand * (pbest - x) + c2 * rand * (gbest - x)$$

$$x' = x + v'$$

$pbest$: The best position that each individual has ever achieved, $gbest$: The best place the whole group has ever been, w : Inertia weight, $c1, c2$: Learning factor.

Simulation discussion

In this paper, select parameters such as TN, TP, NH_4^+-N , DO in aeration basin, T, PH, ORP, MLSS, NO_3-N and electric conductivity of aeration basin k as the auxiliary variables. Use the 1000 groups of data collected in a wastewater treatment plant in Guangzhou to build soft measurement model, select 350 samples, and get 322 samples after pre-processing through 3σ criterion, from which we choose 200 samples as the training set and the rest 100 samples as the test sample to test the generalization ability of the model. Modelling of sludge discharge based on PSO-RVM mainly includes:

(1) Collect 11 parameters such as the en fluent water flow of original wastewater treatment plant, dissolved oxygen concentration in aeration basin, TN, TP and so on to constitute the signal sample set.

(2) Eliminate abnormal data through 3σ criterion and normalize the data.

(3) The particle location is optimized as the optimal parameter of kernel function

(4) Train and predict regression of the model.

(5) Generalize the model to test the predicted effects.

The simulation results are shown in figures 1.

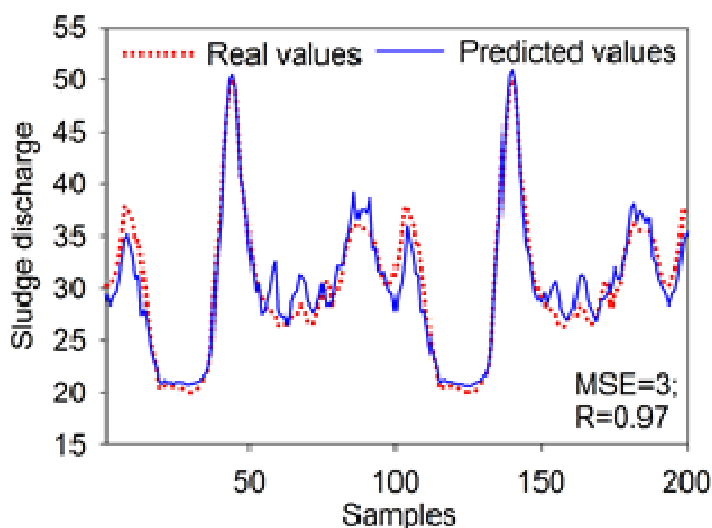


Fig.1. Simulation results of prediction sludge discharge

As can be seen from the Fig.1, the mean square deviation of the algorithm is 3, and the correlation reaches 0.97, which can accurately predict the amount of sludge discharged. Therefore, the online model of PSO-RVM proposed in this paper can better meet the requirements of real-time prediction.

Conclusions

According to the characteristics of wastewater biochemical treatment, we select the important parameter the sludge discharge is selected for modeling , and the 11 parameters such as easily measured in-fluent water flow, TN, TP, NH₄⁺-N, DO in aeration basin, T, PH, ORP, MLSS, NO₃-N and electric conductivity of aeration basin k as the auxiliary variables, build model of the sludge discharge combining with PSO-RVM. The result of the experiment validates that accurate prediction of sludge discharge will provide basic data for the precise control of subsequent sludge dewatering process.

Acknowledgements

This work was financially supported by project of Guangzhou science and technology(201604010032). Guangzhou education planning project(1201534232).The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

- [1] H Brix, Sludge Dewatering and Mineralization in Sludge Treatment Reed Beds, water, 2017 , 9 (3) :160
- [2] Y Xu , L Liu , T Cao ,On-line soft measuring model based on Fast-RVM , CIESC Journal, 2015, 66(11): 4540-4545.
- [3] ME Tipping ,Sparse bayesian learning and the relevance vector machine, Journal of Machine Learning Research, 2001,1 (3) :211-244
- [4] Y Shi , R Eberhart, Modified particle swarm optimizer,IEEE World Congress on IEEE International Conference on evolutionary computation, 1997 , 6 :69-73
- [5] P Ghamisi , MS Couceiro , JA Benediktsson ,A Novel Feature Selection Approach Based on FODPSO and SVM, IEEE Transactions on Geoscience & Remote Sensing , 2015 , 53 (5) :2935-2947
- [6] QH Zhang , WN Yang, HH Ngo,Current status of urban wastewater treatment plants in China, Environment International , 2016 , 92-93 :11
- [7] Y Sakai, T Fukase, H Yasui, M Shibata ,An activated sludge process without excess sludge production, Water Science & Technology , 1997 , 36 (11) :163–170