Sludge Conditioning Soft Measurement Based on GA-RVM

Luo Long
Mechanical & Electrical Department, Guangzhou Institute of Technology, Guangzhou 510075, P. R. China
luolonggdut@163.com

Abstract. To solve the highly non-linear problem of sludge conditioning, the Relevance Vector Machine (RVM), was used to softly measure the sludge conditioning process. The parameters of the kernel function seriously affect the comprehensive performance of RVM. Using genetic algorithm (GA) to obtain the optimal kernel function parameters is an effective solution. The simulation results indicated that the filtering performance was basically consistent with the actual value of dosage and the model prediction result, and the target value could be well predicted. It showed that the model had a certain value for industrial application.

1. Introduction
Municipal sewage sludge has the characteristics of large production, high moisture content, and contains a large number of organic matters, heavy metals, pathogenic bacteria, etc. Sludge dewatering is a necessary step for subsequent disposal (landfill or incineration). At present, chemical, physical and biological ways are used to improve the effect of sludge dewatering frequently, combining with advanced dewatering equipment. Compared with physical and biological methods, chemical conditioning is much more simple and feasible.

There have been many studies and practices on sludge chemical conditioning in the past. The main factors affecting sludge sedimentation, flocculation and dewatering performance are the size of microbial flocs, surface charge, organic matter, pH value and zeta potential. According to the characteristics of sludge, the conditioning reagents with different functions and properties are selected. At present, the main conditioning reagents for application are flocculants, bridge-erecting reagents and wall-breaking reagents. The main functions of these conditioning reagents are neutralizing sludge potential, destroying the colloid structure of sludge, building close bridges between sludge particles, increasing the strength of sludge particles and improving the dewatering performance of sludge, thus the dewatering efficiency of sludge pressing will be improved. By using the combination of different conditioning reagents, the natural drying efficiency will also be increased, which reduces the energy consumption of the system.

Therefore, it is of great significance to improve the process and effect of municipal sludge dewatering by researching and optimizing a more efficient combination of flocculant with quicklime and fly ash, and building an exact model that can indicate the quantitative relationship between the performance of sludge dewatering and any other factors.

2. 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 options.

For a given data set \( \{x_i,t_i\}_{i=1}^l \), \( x_i \in \mathbb{R}^d \), \( t_i \in \mathbb{R} \), \( d \) is the number of dimension of vector data set. In the experiment, the target value is affected by noise, so it is defined as \( t = y(x,w) + \epsilon \), in which, \( \epsilon \) is the noise that obeys Gaussian distributions with 0 mean error and variance \( \sigma^2 \), and its probability is \( p(t | y(x), \sigma^2) = \mathcal{N}(y(x,w), \sigma^2) \). It is defined that \( y(x,w) = \sum_{i=1}^l w_i \varphi(x,x_i) + w_0 \), in which, \( l \) is the size of data set, \( t = (t_1,t_2,...,t_l) \), \( x = (x_1,x_2,...,x_l) \), \( w \) is weight vector, \( \varphi(x,x_i) \) 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:

\[
p(t | w, \sigma^2) = (2\pi \sigma^2)^{-\frac{l}{2}} \exp\left\{-\frac{1}{2\sigma^2} \| t - \Phi w \|^2 \right\} \tag{1}
\]

Where, \( w = (w_0,w_1,...,w_l) \) \( \Phi \) is inputted kernel function mapping, \( \Phi = [\varphi(x,x_1) \varphi(x,x_2) ... \varphi(x,x_l)]^T \). 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:

\[
p(w | \alpha) = \prod_{i=1}^l \mathcal{N}(w_i | 0, \alpha_i^{-1}) \tag{2}
\]

in this formula: \( \alpha = (\alpha_1,\alpha_2,...,\alpha_l) \) 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:

\[
p(w | t, \alpha, \sigma^2) = \mathcal{N}(\mu, \Sigma) \tag{3}
\]

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

\[
p(t | \alpha) = (2\pi)^{-\frac{l}{2}} \| \sigma^2 I + \Phi A^{-1} \Phi^T \|^\frac{l}{2} \exp\left\{-\frac{1}{2} t^T (\sigma^2 I + \Phi A^{-1} \Phi^T)^{-1} t \right\} \tag{4}
\]

By solving the maximum likelihood distribution, hyper-function optimal value \( \alpha_{MP} \) and \( \sigma^2_{MP} \) can be obtained. So far, the model of target value \( t \) is constructed. For input value \( x_0 \), the probability distribution of its corresponding output is:

\[
p(t_* | t_0, \alpha_{MP}, \sigma^2_{MP}) = \mathcal{N}(t_0, y(x_0,w), \sigma^2_{MP}) \tag{5}
\]

where, the vector \( t_* \) is the predicted value of \( x_0 \), and its mean value is \( y(x_0,w) = \mu \varphi(x_0) \), variance shows its uncertainty, and its formula is \( \sigma^2_0 = \sigma^2_{MP} + \varphi^T(x_0) \Sigma \varphi(x_0) \).

3. Genetic algorithm

Genetic algorithms operate on a set of possible solutions. Because of the random nature of genetic algorithms, solutions found by an algorithm can be good, poor, or infeasible, so there should be a way to specify how good that solution is. This is done by assigning a fitness value to the solution. Chromosomes represent solutions within the genetic algorithm. The two basic components of chromosomes are the coded solution and its fitness value.
Chromosomes are grouped into population on which the genetic algorithm operates. In each step [generation], the genetic algorithm selects chromosomes from a population and combines them to produce new chromosomes. These offspring chromosomes form a new population in the hope that the new population will be better than the previous ones. Populations keep track of the worst and the best chromosomes, and stores additional statistical information which can be used by the genetic algorithm to determine the stop criteria.

A chromosome, in some way, stores the solution which it represents. This is called the representation of the solution. There are a number of probable ways to represent a solution in such a way that it is suitable for the genetic algorithm [binary, real number, vector of real number, permutations, and so on] and they mostly depend on the nature of the problem.

4. Simulation discussion

RVM is a new statistical learning method in Bayesian framework, which has some advantages that SVM does not have. RVM makes use of kernel function to linearize the regression, obtains sparse solution, and avoids over fitting under the current kernel function. Experiments show that the parameters of kernel function seriously affect the comprehensive performance of RVM regression. Genetic algorithm is a kind of stochastic optimization algorithm, which simulates the phenomenon of duplication, crossover and mutation in natural selection and heredity. After iterative evolution, it converges to a group of individuals who are most suitable for the environment, automatically but not at a loss, and efficiently obtains the optimal solution of the problem.

The algorithm flow of the sewage treatment Sludge Conditioning prediction method of GA-RVM is listed below: 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.

(1) Monitoring process parameters and target parameters of Sludge Conditioning process, establishing data samples x,t.

(2) Input the eigenvector matrix after dimension reduction into relevance vector machine to be trained and choose the auxiliary variable and predictor variable of the model.

(3) Train and predict regression of the model.

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

The simulation results are shown in figures 1 and 2.

![Figure 1: Genetic algorithm fitness](image)
Figures 2  RVM predict

5. Conclusion
As shown in the figures above, the actual values of filtration performance and dosing amount are basically consistent with the model prediction results within the effective range of organic matter content, thus the sludge conditioning modeling method based on RVM can predict the target value well, and has certain application value for industrial application. The result of the experiment validates that soft measurement model combining - genetic algorithm with relevance vector machine can tackle the relevant problems effectively between variables with higher degree of prediction process and faster rate of convergence, which provides application value for automatic real-time control of wastewater.

Acknowledgments
This work was financially supported by project of Guangzhou Education Bureau of China (201831786). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References
[1] Feng Guohong, Bai Tiantian, Ma He, Hu Zhi, Guo Yabing, Tan Wei. Establishment of the Permeability Model for Soft Solid Sludge Conditioned with Flocculants.[J]. ACS omega, 2019, 4(20).
[2] Suhua Wu, Ping Li, Jie Xu, Kaiyi Xu, Lan Xiao. (2016). Screening of Flocculants Used in Sludge Deep Dewatering and Optimization of Sludge Conditioning System Applied with Microparticles as Flocculant Aid. Water Purification Technology. 35(2), 64-70.
[3] Zhang Yanan, Yang Huizhong. RVM Soft Sensing Model Based on Optimized Combined Kernel Function[J]. Journal of System Simulation, 2018, 30(1): 272-277.
[4] Gang Li, Guilong Wang, Huifeng Xue. GA Optimizing Method to Kernel Function Parameters of RVM[J]. Control Engineering, 2010, 17(3): 335-338.
[5] Sheng-wei Fei. Kurtosis forecasting of bearing vibration signal based on the hybrid model of empirical mode decomposition and RVM with artificial bee colony algorithm[J]. Expert Systems with Applications (S0957-4174), 2015(42): 5011-5018.