

# Analysis of Ammonia Nitrogen Content in Water Based on Weighted Least Squares Support Vector Machine (WLSSVM) Algorithm

Jinwu Ju, Lanying Wang

School of Computer Science of Sichuan University of Science and Engineering, Zigong, China  
Email: tougaozhoubianji@126.com

Received 6 January 2016; accepted 15 February 2016; published 18 February 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY).  
<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Determination of ammonia nitrogen content in water is the basic item of the environmental water pollution, and is the key index to evaluate the water quality. This article designs a water quality monitoring system based on the on-line automatic ammonia nitrogen monitoring system, and establishes a forecasting model based on the weighted least squares support vector machine algorithm. The weighted least squares support vector machine algorithm increases the weight parameter setting, improves the speed and accuracy of prediction learning, and improves the robustness. In this article, a comparison between neural network model and weighted least square support vector machine model is made, which shows that the weighted least squares support vector machine model has better prediction accuracy.

## Keywords

Support Vector Machine, Water Quality, Ammonia Nitrogen, Forecasting Model

---

## 1. Introduction

With the rapid development of social economy, water pollution has become an important environmental monitoring content. The determination of ammonia nitrogen content in water quality is the basic item of water quality monitoring, and is the key index to evaluate water quality. At present, the method of on-line water quality and ammonia nitrogen content determination system [1] is often used in electrochemical method or the method of spectro photometry. Electrochemical method is based on the electrode sensor, the determination of the size of

the ammonia sensor electrode, the indirect determination of ammonia nitrogen content. Nessler's points spectro photometric method based on light chromaticity determination, mercuric iodide and potassium iodide solution will react with ammonia and will generate red brown compound, chroma and water content of ammonia and nitrogen in a proportional relationship. The main disadvantage of electrochemical detection is that it is easy to interfere with the concentration of ions in water samples, and in the long run, the sensor can easily be contaminated, and the sensitivity of the sensor is obviously decreased. The disadvantages of the method of spectro photometry [2] are that the reagent cost is relatively high, the determination time is long, the determination data are small, and the determination process is easy to be influenced by the environmental changes and other factors.

The frequently used method to solve the data reliability is to establish a data prediction model. Neural network [3] [4], support vector machine [5] and other forecasting methods can be used to establish data prediction model, but the neural network requires a large number of training sample data, not suitable for the determination of ammonia nitrogen content analysis. Support vector machine is less than the sample data required, suitable for the establishment of data prediction model. Least squares support vector machine (LS-SVM) is the improvement of basic support vector machine (SVM), which can solve the problem of its generalization ability, and it has a better prediction data bias, which has better prediction learning speed, but its prediction robustness is obviously decreased. Weighted least squares support vector machine (WLS-SVM) [6] based on the standard least squares support vector machine based on the increase of weight components, can get good results and improve the robustness of the algorithm.

This article aims to design the on-line automatic ammonia nitrogen monitoring system. The accuracy of the data fluctuates during measurement of ammonia nitrogen content by spectro-photometry. The weighted least squares support vector machine (WLS-SVM) is applied to the analysis of ammonia nitrogen content, and the model is established to improve the accuracy of the data.

## 2. Algorithm Principle

Support vector machine is proposed by Vapnik *et al.* This algorithm is based on the learning theory of limited sample, which is followed by the principle of structural risk minimization. The disadvantage of standard support vector machine is that it is relatively large in the prediction of abnormal data. The least squares support vector machine is based on the standard support vector machine, which is proposed by Suykens [7] *et al.* By changing the constraint condition and risk function of SVM, it can be used to improve the shortcomings of standard SVM, but its robustness is reduced. Weighted least squares support vector machine can improve the robustness of the least square support vector machine by adding weights to the error components.

### 2.1. Support Vector Machine Theory

The training set  $D$  with a given sample number is  $N$ , and  $D$  region is expressed as:  $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ , in which  $x$  is the input data,  $y$  is the output data. Using a nonlinear mapping  $\phi(\cdot)$ , the sample data from the original space to high dimensional feature space  $\phi(x) = (\phi(x_1), \phi(x_2), \dots, \phi(x_N))$  is mapped. In this space, the optimal decision function is constructed:  $y(x) = \omega\phi(x) + b$ , the parameters  $\omega$  &  $b$  are to be determined,  $\omega$  is weight vector,  $b$  is threshold value.

The objective optimization function of support vector machine is:

$$\begin{aligned} \text{Min } J(\omega, \xi) &= \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \xi_k \\ \text{s.t. } y_k \left[ \left( \phi(x_k) \omega^T + b \right) \right] &\geq 1 - \xi_k \end{aligned} \quad (1)$$

The error component in above objective optimization function is  $\xi_k \geq 0$ ,  $k = 1, 2, \dots, N$ ; the parameter  $C$  is regularization parameter and can be used to balance error and complexity.

### 2.2. Weighted Least Square Support Vector Machine Theory

The objective optimization function of least squares support vector machine is based on the principle of different structural risk selected.

$$\begin{aligned} \text{Min } J(\omega, \xi) &= \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \nu_i \xi_k^2 \\ \text{s.t. } y_k &= \phi(x_k) \omega^T + b + \xi_k \end{aligned} \quad (2)$$

In order to improve the robustness of the least squares support vector machine algorithm, on the basis of the above optimization function, weights are added  $\nu_i$  [8] [9]. Lagrange method is used to solve the above optimization function,

$$L(\omega, b, \xi, \alpha) = \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^N \nu_i \xi_k^2 - \sum_{i=1}^N \alpha_k \left[ (\omega^T \phi(x_k) + b + \xi_k) - y_k \right] \quad (3)$$

In above function,  $\alpha_k (k=1, 2, \dots, N)$  is the multiplier, solving partial derivative  $\omega$ ,  $b$ ,  $\xi$ ,  $\alpha$ , Setting it to zero, the optimization problem can be converted into linear equation as follows:

$$\begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & k(x_1, x_1) + 1/C & \cdots & k(x_1, x_1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & k(x_N, x_1) & \cdots & k(x_N, x_1) 1/C \end{bmatrix} \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} 0 \\ Y_1 \\ \vdots \\ Y_N \end{bmatrix} \quad (4)$$

Setting  $\Omega$  is  $N \times N$  symmetric matrices, such as (5) formula:

$$\Omega = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \cdots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \cdots & k(x_2, x_n) \\ \vdots & \vdots & \ddots & \vdots \\ k(x_n, x_1) & k(x_n, x_2) & \cdots & k(x_n, x_n) \end{bmatrix} \quad (5)$$

Defined kernel function  $k(x_i, x_j)$ , conform to (6) formula *Mercer* condition.

$$\Omega_{ij} = k(x_i, x_j) = \phi(x_i) \phi(x_j) i, j = 1, \dots, N \quad (6)$$

Obtained the model of weighted least squares support vector machine is:

$$f(x) = \sum_{k=1}^N \alpha_k k(x, x_k) + b \quad (7)$$

The selected of kernel function  $k(x_i, x_j)$  can be constructed by different support vector machine, and the radial basis function (RBF) has the characteristics of fast learning speed, and its formula is (8).

$$K(x_i, x_j) = \exp\left(-\frac{|x_i - x_j|^2}{2\sigma^2}\right) \quad (8)$$

The regularization parameter  $C$  and kernel width parameter  $\sigma$  are two adjustable parameters of the algorithm, and their values are determined by the generalization performance and learning ability of the algorithm.

We can use the search method to choose the parameters, and then use the cross validation method to find the optimal parameters of the objective function to find the best parameters.

### 3. Water Quality Online Automatic Ammonia Nitrogen Monitoring System

Water quality online automatic ammonia nitrogen monitoring system, based on the method of nano-determination of ammonia nitrogen content of the measurement, the overall structure of the system is shown in **Figure 1**. The system is shown in **Figure 2**. The system is composed of industrial computer, PLC controller, LCD touch screen, spectro-photometer, reactor, metering device, injection pump, multi solenoid valve and all kinds of reagent. Through LCD touch screen display software interface, through the PLC controller, on the other hand, the measurement of the reactive color data of the reactor is used to calculate the content of ammonia nitrogen. The basic working process is as follows:

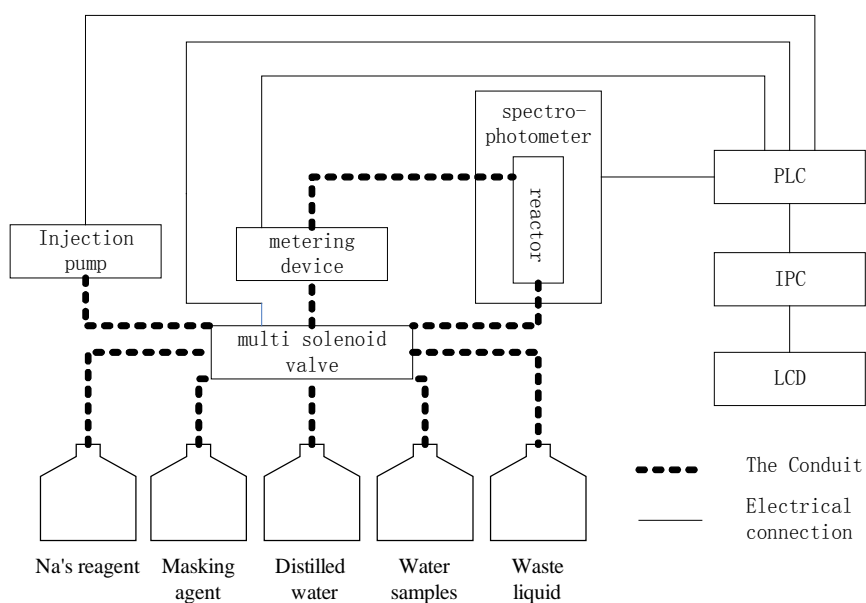


Figure 1. Overall structure system diagram.



Figure 2. Overall structure of the system.

A) Control of multi pass solenoid valve, access to water samples, drive the injection pump, pump the water samples into the metering device.

B) Once achieve the specified capacity of the metering device, switch to the solenoid valve, pump the water sample from the metering device into the reactor.

C) Switch the multi pass solenoid valve, the masking agent is pumped into the metering device, to achieve the specified capacity, pumping into the reactor.

D) Switch to the solenoid valve, the nano's reagent is pumped into the metering device, to achieve the specified capacity, pumped into the reactor.

E) Standing 10 minutes, then measure the output of the spectrophotometer, and get the intensity of the measurement.

F) Switch to the solenoid valve, the reactor solution to the liquid bottle.

G) Switch to the solenoid valve, the distilled water is pumped into the metering device, to achieve the specified capacity, pumping into the reactor. Again, the output of the spectrophotometer is measured, and the reference intensity is obtained.

H) After the completion of the reactor, and then the waste water is pumped into the waste liquid bottle. Complete the whole measurement process.

#### 4. Measurement Results and Analysis of the Monitoring System

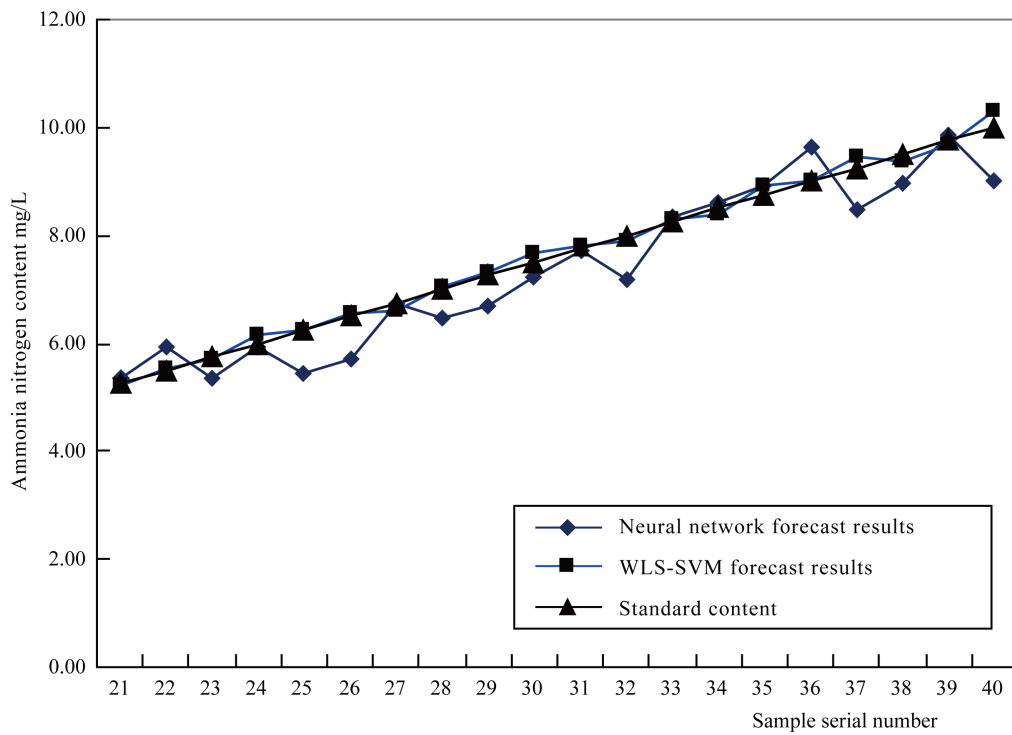
Water quality monitoring system of on-line automatic ammonia nitrogen, based on the measurement data to calculate the final ammonia nitrogen content, and its input data for measuring the zero point of light intensity, reference intensity and measuring intensity, select the 40 groups of ammonia nitrogen standard solution, its ammonia nitrogen content from 0 mg/L, with 0.25 mg/L increments, until 10 mg/L. A set of measurement results of the monitoring system is used as sample data, the total sample data are 40, the first 20 are training samples, and the 20 is the test sample, and the training sample data are shown in **Table 1**. Neural network model and WLS-SVM model were used to predict and analyze the experimental results. The experimental results are shown in **Table 2**.

**Table 1.** Training sample data sheet.

Sample serial number	Reference of light intensity	Measurement of light intensity	Standard content (mg/L)
1	1467.7	2055.3	0.25
2	1468.5	1934.2	0.50
3	1469.8	1820.7	0.75
4	1487.3	1732.9	1.00
5	1494.3	1637.5	1.25
6	1471.4	1516.5	1.50
7	1467.9	1422.9	1.75
8	1482.5	1351.6	2.00
9	1489.9	1277.5	2.25
10	1481.0	1194.4	2.50
11	1488.5	1129.0	2.75
12	1494.4	1066.1	3.00
13	1487.5	998.0	3.25
14	1478.8	933.2	3.50
15	1476.5	876.4	3.75
16	1494.6	834.3	4.00
17	1478.1	776.0	4.25
18	1486.5	734.0	4.50
19	1484.5	689.5	4.75
20	1466.9	640.7	5.00

**Table 2.** Analysis of the experimental results data sheet.

Sample serial number	Neural network forecast results (mg/L)	Neural network prediction error %	WLS-SVM forecast results (mg/L)	WLS-SVM prediction error%	Standard content (mg/L)
21	5.37	2.33	5.20	-0.97	5.25
22	5.94	7.93	5.55	0.83	5.50
23	5.35	-6.98	5.73	-0.29	5.75
24	5.93	-1.23	6.16	2.59	6.00
25	5.46	-12.66	6.25	-0.06	6.25
26	5.73	-11.81	6.56	1.00	6.50
27	6.74	-0.10	6.61	-2.09	6.75
28	6.49	-7.28	7.06	0.83	7.00
29	6.70	-7.62	7.31	0.78	7.25
30	7.22	-3.75	7.68	2.34	7.50
31	7.70	-0.60	7.82	0.90	7.75
32	7.19	-10.17	7.90	-1.24	8.00
33	8.35	1.19	8.30	0.58	8.25
34	8.61	1.34	8.38	-1.47	8.50
35	8.94	2.18	8.90	1.77	8.75
36	9.63	6.95	9.03	0.33	9.00
37	8.46	-8.55	9.47	2.40	9.25
38	8.98	-5.48	9.38	-1.24	9.50
39	9.85	1.04	9.70	-0.50	9.75
40	8.99	-10.12	10.29	2.87	10.00



**Figure 3.** Comparison of the forecast data with actual measurement results.

Using the neural network model and WLS-SVM model to analyze and forecast the ammonia nitrogen content, the monitoring system of the actual measurement results and the two groups of forecasting data, as shown in **Figure 3**. According to **Table 2**, based on neural network model, the maximum relative error is 12.66%, the average relative error is 5.47%. Based on WLS-SVM model, the average relative error is 2.87%, the average relative error is 1.25%, and the accuracy of WLS-SVM model is significantly higher.

## 5. Conclusion

This article introduces that during on-line monitoring system on the water quality, the WLS-SVM model is applied to the analysis and prediction of ammonia nitrogen content measurement data. The forecast results compared with the actual results show that the WLS-SVM model is a very effective forecasting model, which can be applied to the actual data analysis and forecasting. WLS-SVM model only needs two changeable model parameters, the regularization parameter C and kernel width parameter, which can be used cross validation method to determine the parameters of the model. The model has a fast convergence speed, high forecast accuracy and wide application prospect.

## Foundation Item

Sichuan University of Science and Engineering Incubation Project (2012PY12), Artificial Intelligence Key Laboratory Project of Sichuan Province (2013RZY03), Enterprise Informatization and the Internet of Things Measurement and Control Technology Key Laboratory Project (2013WZY04, 2014WYY01), Luzhou Lao Jiao scholarship (15LJZK04).

## References

- [1] Ma, S.Q. and Liang, B. (2015) The Study on the Method of Using Automatic Analyzer to Detect the Ammonia of Water. *Shandong Chemical Industry*, No. 44, 62-64.
- [2] Yu, J.M. (2013) Research on the Factors Affecting the Determination of Ammonia Nitrogen. *Jiangxi Chemical Industry*, No. 4, 11-14
- [3] Tong, M.M., Zhang, Y. and Qi, M.X. (2006) The Mixed Inflammable Gas Analysis Based on BP Neural Network. *Acta Metrologica Sinica*, **27**, 169-171.
- [4] Hou, D.B., Chen, Y., Zhao, H.F., et al. (2013) Water quality anomaly detection method based on RBF neural network and wavelet analysis. *Transducer and Microsystem Technology*, **32**, 138-141.
- [5] Huang, M. and Cheng, X.P. (2008) Real-Time Tracking on Athletes Based on Support Vector Regression Particle Filter. *Journal of Southwest Agricultural University (Natural Science)*, **30**, 165-169.
- [6] Zeng, Y., Cheng, X.W., Wang, H.Q., et al. (2014) On Analysis of Sewer Combustible Gas Based on Weighted LS-SVM. *Journal of Southwest China Normal University (Natural Science)*, **39**, 94-99.
- [7] Suykens, J.A.K. and Vandewalle, J. (1999) Least Squares Support Vector Machine Classifiers. *Neural Processing Letters*, **9**, 293-300. <http://dx.doi.org/10.1023/A:1018628609742>
- [8] Wen, W., Hao, Z. and Yang, X. (2008) A Heuristic Weight-Setting Strategy and Iteratively Updating Algorithm for Weighted Least Squares Support Vector Regression. *Neurocomputing*, **71**, 3096-3103. <http://dx.doi.org/10.1016/j.neucom.2008.04.022>
- [9] Suykens, J.A.K. and Vandewalle, J. (2000) Recurrent Least Squares Support Vector Machines. *IEEE Transactions on Circuits and System—I*, **47**, 1109-1114. <http://dx.doi.org/10.1109/81.855471>