

Simultaneous Rapid Analysis of Multiple Nitrogen Compounds in Polluted River Treatment Using Near-Infrared Spectroscopy and a Support Vector Machine

Jian Huang^{1,2}, Xiong Zhang², Qingye Sun^{1*}, Hua Zhang²,
Xiaokun Yu², Zhaoliang Wu²

¹School of Resources and Environmental Engineering, Anhui University, Hefei, 230601, China

²Anhui Jianzhu University, Key Laboratory of Anhui Province of Water Pollution Control and Wastewater Reuse, Hefei, 230601, China

Received: 19 October 2016

Accepted: 29 March 2017

Abstract

An intermittent aerobic process has been developed to effectively remove nitrogen from polluted rivers. In addition, a chemometric model was developed to achieve simultaneous rapid analysis of total nitrogen, ammonia nitrogen, and nitrite nitrogen based on near-infrared spectroscopy data combined with a support vector machine. An intermittent aeration process showed that total nitrogen decreased from 54.25 mg·L⁻¹ to 0.64 mg·L⁻¹. Ammonia nitrogen decreased significantly in the aeration stage, but increased in the non-aeration stage. Eventually, ammonia nitrogen decreased from 45.04 mg·L⁻¹ to 0.57 mg·L⁻¹. Nitrite nitrogen and nitrate nitrogen increased in the aeration stage, but decreased in the non-aeration stage. The concentration ranges of nitrite nitrogen and nitrate nitrogen were, respectively, 0.05~31.40 mg·L⁻¹ and 0~0.38 mg·L⁻¹. The 138 water samples were collected during the intermittent aeration process, of which 116 samples were used as the calibration set and the remaining 22 samples were used as a test set in modeling. The actual concentration values and the near-infrared spectroscopy data were used as input of the models. Then the corresponding calibration values and predication values were output by the models. All the samples were scanned with near-infrared spectroscopy from 4,000~12,500 cm⁻¹ and measured by chemical methods. Principal component analysis of raw near-infrared spectral data showed that the matrix dimension of spectral data was significantly reduced, which decreased from 2,203×10⁶ to 6×10⁶. Support vector machine models of total nitrogen, ammonia nitrogen, and nitrite nitrogen showed that the calibration correlation coefficient (R²) of calibration values and actual values were, respectively, 0.9561, 0.9661, and 0.9702, with the root mean square error of cross validation (RMSECV) being 0.09372, 0.04749, and 0.03187. The test results of support vector machine models of total nitrogen, ammonia nitrogen, and nitrite nitrogen showed that the predication correlation coefficient (R²) of prediction values and actual values were, respectively, 0.9616, 0.9410, and

*e-mail: sunqingye@ahu.edu.cn

0.9284, with the root mean square error of prediction (RMSEP) being 0.09420, 0.08227, and 0.06770. This study indicated that nitrogen in a polluted river can be removed through the intermittent aerobic process. Moreover, simultaneous rapid determination of total nitrogen, ammonia nitrogen, and nitrite nitrogen may be achieved with near-infrared spectroscopy and a support vector machine. The results showed that the proposed methods provided effective treatment and detection technology for a polluted river.

Keywords: near infrared spectroscopy, support vector machine, principal component analysis, intermittent aeration

Introduction

Excessive nitrogen is a major cause of black and odorous rivers [1-3]. The release of endogenous nitrogen in sediment is an important source of nitrogen in overlying river water when exogenous nitrogen is controlled [4-6]. Recently, continuous aeration technology was usually used to treat black and odorous rivers [7-8]. However, the removal of nitrogen can hardly be achieved because denitrification was neglected in continuous aeration [9]. Therefore, it is of great importance to develop new processes to remove nitrogen in a polluted river. Nitrification-denitrification is an important method for removing nitrogen from polluted rivers [10-12]. Usually, nitrification-denitrification is achieved through the intermittent aerobic process, according to the principle of the biological removal of nitrogen. In the aerobic phase, ammonia nitrogen is converted into nitrate and nitrite by nitrification of nitrifying bacteria. In the non-aerobic phase, nitrate and nitrite are converted into nitrogen by denitrification of denitrifying bacteria. However, there are few studies on nitrogen removal with intermittent aerobic processes in a polluted river.

Compared with traditional chemical analysis, near-infrared spectroscopy has the advantages of rapid detection, simple analysis, and non-pollution [13-14]. Recently, some research has indicated that near-infrared spectroscopy in combination with chemometrics was used in food, medicinal material, and water analysis [15-18]. The support vector machine algorithm has unique advantages in solving problems of small samples, and nonlinear and high-dimensional pattern recognition. In addition, the support vector machine can effectively avoid the curse of dimensionality and overlearning. Therefore, the quantitative model based on the support vector machine has better prediction performance and stability [19-20]. Near-infrared spectroscopy, when combined with the support vector machine algorithm, had better generalization ability and stability, which was widely used in rapid identification of species, sewage treatment condition, and COD concentration [21-25].

Although the support vector machine is very useful for solving small sample, nonlinear, and high-dimensional pattern recognition problems, the models based on the support vector machine are susceptible to the background and the noise in spectra. The influences of the background and the noise will lead to a very complicated model calculation and reduce the predicting ability of the

model. Therefore, in order to reduce the influence of the background and the noise, spectral data are often preprocessed before quantitative models are established. Principal component analysis is a popular preprocessing method of spectral data. It can effectively reduce the background and the noise of spectral data, and keep useful information at the same time. In addition, principal component analysis can effectively reduce the dimensions of spectral data [26-28]. Therefore, principal component analysis was usually selected as a preprocessing method of spectral data before the establishment of quantitative models of analysis.

In this study, the intermittent aeration process in overlying water was developed to achieve nitrogen removal in polluted river. Near-infrared spectroscopy combined with a support vector machine was used to establish support vector machine models of total nitrogen, ammonia nitrogen, and nitrite nitrogen in overlying river water. Principle component analysis was used to preprocess the raw spectral data and reduce the dimensions of spectral data. Based on infrared spectroscopy, support vector machine models were established to rapidly analyze total nitrogen, ammonia nitrogen, and nitrite nitrogen. In addition, the predicting ability of support vector machine models was further tested. The proposed methods can provide effective treatment technology for a polluted river and rapid detection of total nitrogen, ammonia nitrogen, and nitrite nitrogen in water.

Material and Methods

Sample Collection and Experimental Control

The Shiwuli River runs through southwest Hefei, China. It flows through residential and industrial areas, and is heavily polluted with nitrogen, and is a main source of pollution of Chaohu Lake, which is one of the five largest freshwater lakes in China [29]. Sediments and overlying water were collected on 27 March 2015. Sediments were collected at a surface depth of 30 cm and stored away from light. At the same time, the overlying water was also collected and taken back to the laboratory.

A rectangular organic glass reactor was adopted for this experiment. The length, width, and height of the reactor were, respectively, 21 cm, 20 cm, and 51 cm. The depth of sediment and water were 9 cm and 27 cm, respectively. The reactor was covered by dark film to three-fourths of the

height. Experiment lasted for 138 d, and dissolved oxygen was kept at $5.5 \text{ mg}\cdot\text{L}^{-1}$ in aerobic phase. The aerobic phase was respectively 11~40 d, 69~71 d, 80~90 d, 99~101 d, 114~117 d, 123~125 d, and 130~135 d. The pH of the reactor was 7.02~8.33. The temperature in the reactor was 19.2~22.8. Overlying water samples were collected every day. The 138 water samples were centrifuged and filtered by a $0.45 \text{ }\mu\text{m}$ filter membrane. After preprocessing, the 138 water samples were analyzed with a near-infrared spectrometer. The concentrations of total nitrogen, ammonia nitrogen, nitrate nitrogen, and nitrite nitrogen of water samples were also measured by chemical methods. Determining total nitrogen was the alkaline potassium persulfate digestion UV spectrophotometric method; determining ammonia nitrogen was Nessler's reagent spectrophotometry; determining nitrate nitrogen was ultraviolet spectrophotometry; and for determining nitrite nitrogen we used the spectrophotometry method.

Experimental Method and Equipment

Spectral data were determined with a near-infrared spectrometer (Bruker, Germany). The scanning range of wave number was $4,000\text{--}12,500 \text{ cm}^{-1}$, and the times of scanning were 32. The resolution of the near-infrared spectrometer was 8 cm^{-1} . An ultraviolet-visible spectrophotometer (Shimadzu, Japan) was used to measure total nitrogen, ammonia nitrogen, nitrate nitrogen, and nitrite nitrogen.

Establishment and Test of Support Vector Machine Model and Evaluation Index

The 138 overlying water samples were collected from the intermittent aerobic process. All the samples were filtered through $0.45 \text{ }\mu\text{m}$ filter membrane. Each sample was measured by chemical method (an ultraviolet-visible spectrophotometer) to get actual concentration values of

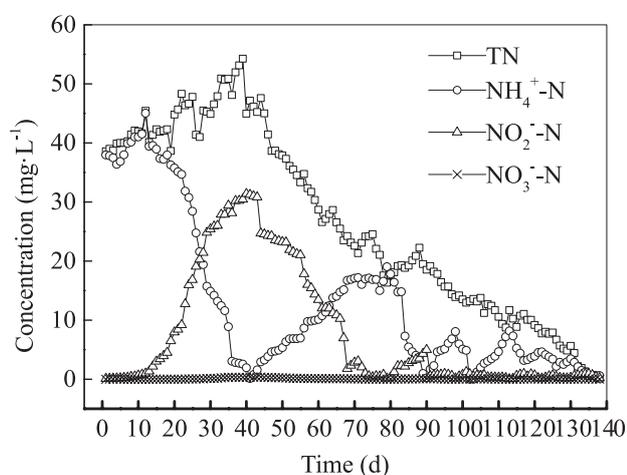


Fig. 1. Changes of nitrogen during the intermittent reaction process.

total nitrogen, ammonia nitrogen, and nitrite nitrogen. In addition, every sample was measured by a near-infrared spectrometer to get spectral data. All the samples were divided into two groups. One group includes 116 samples (calibration set), and the other includes 22 samples (test set). The 116 samples were used to build support vector machine models. The root mean square error of cross validation (RMSECV) and calibration correlation coefficient (R^2) were used to assess the accuracy of support vector machine models.

The 22 samples were used to test the predicting ability of support vector machine models. These models were tested by comparing prediction values with the actual values. The root mean square error of prediction (RMSEP) and prediction correlation coefficient (R^2) were used to estimate the prediction performance of support vector machine models.

Results and Discussion

Changes of Nitrogen Concentration in the Intermittent Aeration Process

The intermittent aeration process showed that the release of nitrogen in sediment was promoted. In addition, effective nitrification and denitrification were achieved during the intermittent aeration process. The nitrogen in a polluted river was effectively removed through nitrification-denitrification in the intermittent aeration process. The changes of total nitrogen, ammonia nitrogen, nitrate nitrogen, and nitrite nitrogen are shown in Fig. 1.

Fig. 1 shows that the total nitrogen gradually increased with the release of nitrogen in sediment in the first 38 d. After 38 d, the total nitrogen gradually decreases with nitrification and denitrification. The total nitrogen finally decreases to $0.64 \text{ mg}\cdot\text{L}^{-1}$. Ammonia nitrogen gradually increased in the first 12 d with the release of ammonia nitrogen in the sediment. After 12 d, ammonia nitrogen

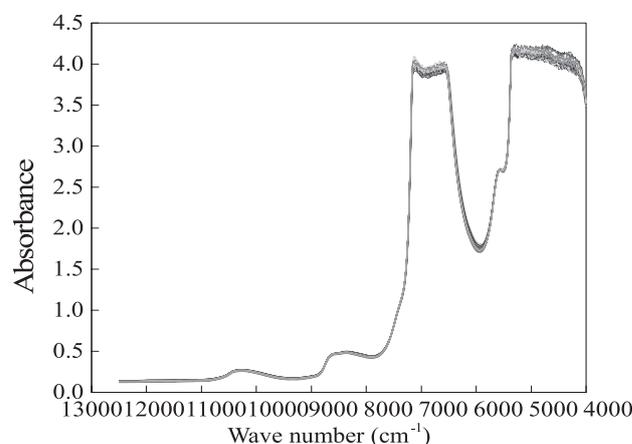


Fig. 2. Raw near-infrared spectra of the calibration set in the intermittent aeration process.

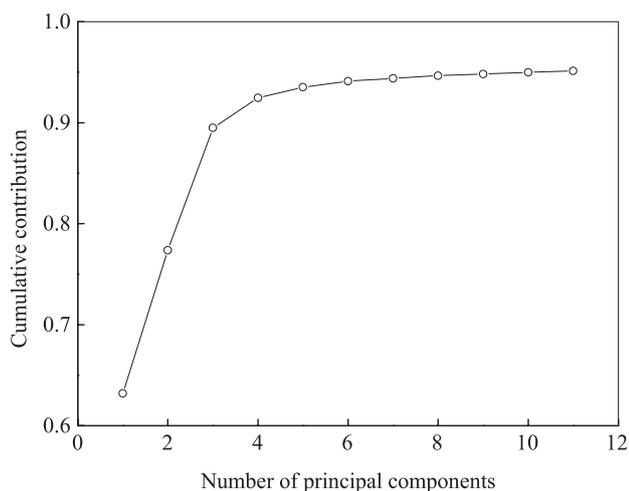


Fig. 3. Cumulative contribution of calibration set versus the number of principal components.

decreases under aeration conditions with nitrification. However, ammonia nitrogen increases under non-aeration conditions with the release of ammonia nitrogen in the sediment. After repeated intermittent aerations, ammonia nitrogen decreases to $0.57 \text{ mg}\cdot\text{L}^{-1}$. Nitrite nitrogen increases under aeration conditions with nitrification and decreases under non-aeration conditions with denitrification. The concentration of nitrite nitrogen is $45.04\sim 0.57 \text{ mg}\cdot\text{L}^{-1}$. However, the change of nitrate nitrogen is not obvious and a lower concentration ($0\sim 0.37 \text{ mg}\cdot\text{L}^{-1}$) persists. The results show that effective nitrogen removal may be achieved by the intermittent aeration process.

Preprocessing of Near-Infrared Spectra

The near-infrared spectra of 116 samples were obtained with a near-infrared spectrometer (Fig. 2). The scanning wavenumber is $4,000\sim 12,500 \text{ cm}^{-1}$ with 2,203 spectral data. As seen from Fig. 2, all the samples have similar near-infrared spectra. However, there are some differences in the wave number of $4,000\sim 5,100 \text{ cm}^{-1}$ and $6,500\sim 7,500 \text{ cm}^{-1}$. The differences are mainly due to the background and the noise of a near-infrared spectrometer [30-31].

To reduce background, noise, and dimension of spectral data, the preprocessing of raw spectral data was needed before the establishment of support vector machine models. Principle component analysis was used to preprocess raw spectral data to improve computational efficiency and prediction performance of the model. The results of principle component analysis, as shown in Fig. 3, indicates that the cumulative contribution changes with the principle component number. When the principle component number reaches 6, the cumulative contribution reaches stability. Therefore, six principle component numbers were identified in modeling. The main information of raw spectral data is included in the six principle components. The matrix of spectral data greatly decreased from $116\times 2,203$ to 116×6 .

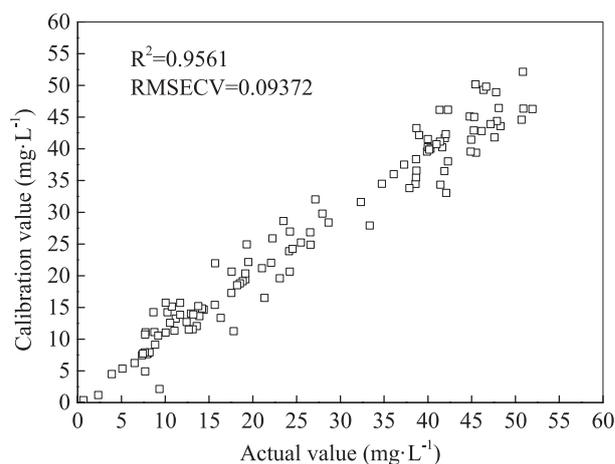


Fig. 4. Establishment of support vector machine model of total nitrogen.

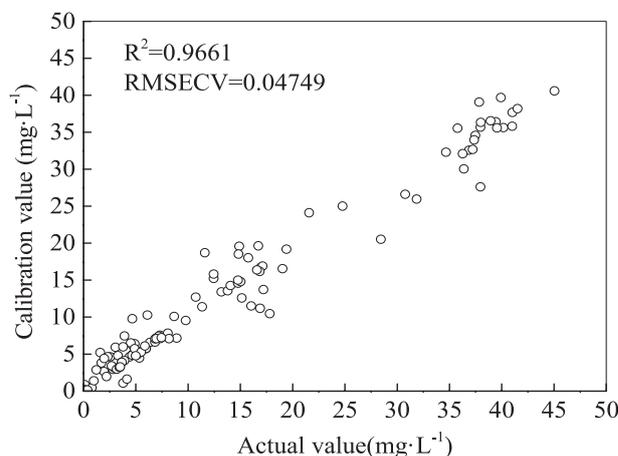


Fig. 5. Establishment of support vector machine model of ammonia nitrogen.

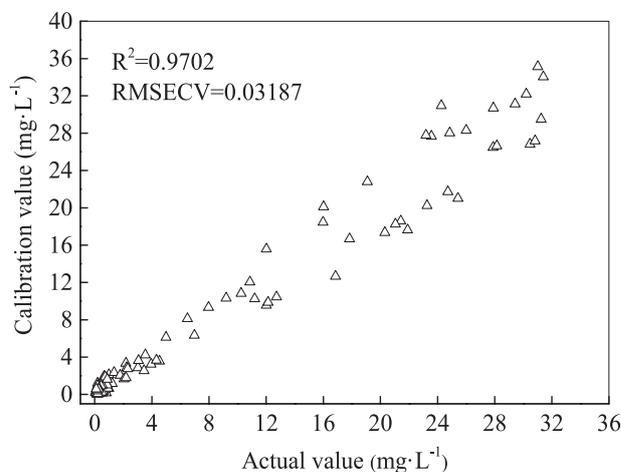


Fig. 6. Establishment of support vector machine model of nitrite nitrogen.

Establishment of Support Vector Machine Models

Preprocessed spectral data matrix (116×6) was used as input of support vector machine. Then support vector machine models of total nitrogen, ammonia nitrogen, and nitrite nitrogen were established by selecting appropriate kernel function and optimizing kernel parameters. Currently, there are four common kernel functions: linear function, polynomial function, radial basis function, and sigmoid function. Radial basis function is most widely used in establishing support vector machine models. It was applicable to the low dimension, high dimension, small sample, large sample, etc., having a wider convergence region [32]. Therefore, radial basis function is suitable for this study. The support vector machine model, based on radial basis function, has excellent learning ability and generalization ability. The optimal penalty coefficient and kernel parameter in modeling were determined through leave-one-out cross validation method. Penalty coefficient and kernel parameter were, respectively, 1.41 and 0.25. Then the support vector machine models of total nitrogen, ammonia nitrogen, and nitrite nitrogen were established through support vector machine regression, as shown in Figs 4-6.

The support vector machine model of total nitrogen shows that the calibration values are in good correlation with the actual values. The calibration correlation efficient and the root mean square error of cross validation (RMSECV) are 0.9561 and 0.09372, respectively (Fig. 4). For the support vector machine model of ammonia nitrogen (Fig. 5), the calibration correlation efficient and the root mean square error of cross validation (RMSECV) between calibration values and actual values are 0.9661 and 0.04749, respectively. For the support vector machine model of nitrite nitrogen, the calibration correlation efficient and the root mean square error of cross validation (RMSECV) between calibration values and actual values are 0.9702 and 0.03187, respectively (Fig. 6). The results indicate that support vector machine models have better calibration results on total nitrogen, ammonia nitrogen, and nitrite nitrogen during the intermittent process.

Test of Support Vector Machine Model

The near-infrared spectral data of the 22 samples are input to the established support vector machine models of total nitrogen, ammonia nitrogen, and nitrite nitrogen. The prediction values of the models and actual values are shown in Figs 7-9. Test of the model of total nitrogen shows that the prediction correlation coefficient and the root mean square error of prediction (RMSEP) are, respectively, 0.9616 and 0.09420 (Fig. 7). A test of the ammonia nitrogen model shows that the prediction correlation coefficient and the root mean square error of prediction (RMSEP) are, respectively, 0.9410 and 0.08227 (Fig. 8). A test of the nitrite nitrogen model shows that the prediction correlation coefficient and the root mean

square error of prediction (RMSEP) are, respectively, 0.9284 and 0.06770 (Fig. 9). As can be seen from Figs. 7-9, the prediction values of support vector machine models are in good correlation with actual values.

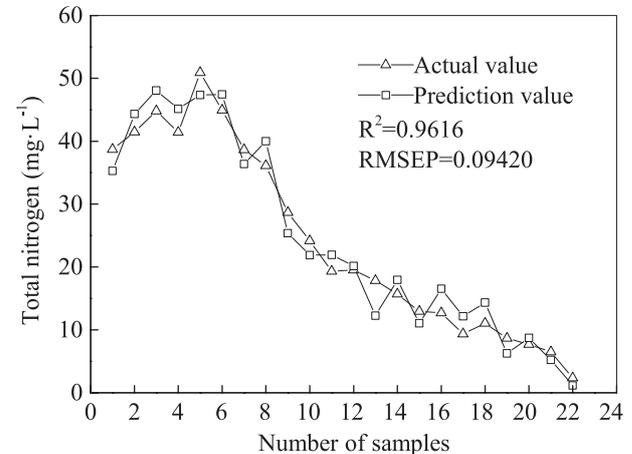


Fig. 7. Test of support vector machine model of total nitrogen.

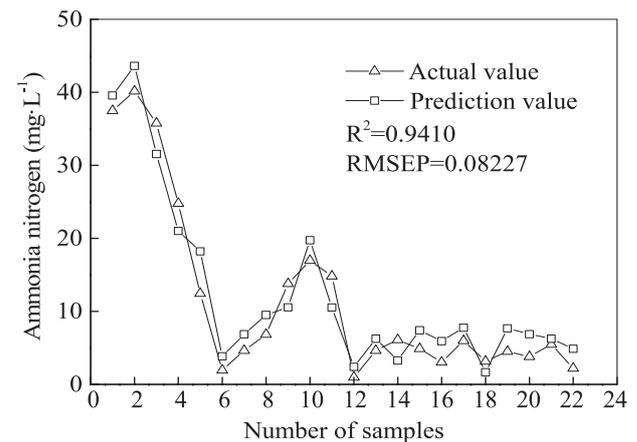


Fig. 8. Test of support vector machine model of ammonia nitrogen.

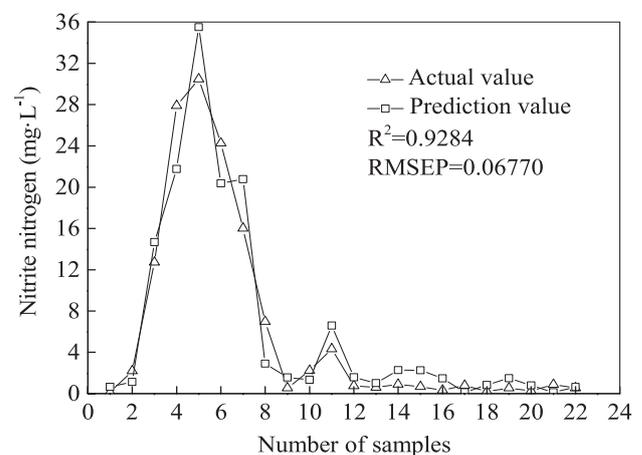


Fig. 9. Test of support vector machine model of nitrite nitrogen.

The results demonstrated that support vector machine models may rapidly determine the concentrations of total nitrogen, ammonia nitrogen, and nitrite nitrogen in overlying water during the intermittent aeration process in a polluted river.

Conclusions

The nitrogen was effectively removed by nitrification-denitrification of the intermittent aeration process in a polluted river. Ammonia nitrogen obviously decreased under aeration conditions by nitrification, and it obviously increased under non-aeration conditions by denitrification. Similarly, nitrite nitrogen obviously increased under aeration conditions and obviously decreased under non-aeration conditions. After repeated intermittent aeration, total nitrogen and ammonia nitrogen were effectively removed by nitrification-denitrification.

The near-infrared spectral data were preprocessed through principle component analysis. Redundant information was greatly eliminated. The calculation efficiency and predication ability of support vector machine models were significantly improved after principle component analysis.

The support vector machine models showed that good correlations were found between the calibration values and actual values of total nitrogen, ammonia nitrogen, and nitrite nitrogen. The support vector machine models were tested by comparing the predicted values with the actual values of total nitrogen, ammonia nitrogen, and nitrite nitrogen. This study demonstrates that support vector machine models based on near-infrared spectroscopy were feasible and practical in determining total nitrogen, ammonia nitrogen, and nitrite nitrogen during the intermittent aeration process in a polluted river.

Acknowledgements

Our work was supported by the Natural Science Project for Colleges of Anhui Province (KJ2016A817), the Cultivation Project for Excellent Young Scholar of Colleges of Anhui Province, and the Major Science and Technology Program for Water Pollution Control and Treatment of China (2014ZX07405-003-03). The authors are grateful to Dr. Chunshan Xu for valuable discussions.

References

1. WANG Z.Q., LI B., LIANG R.J., WANG L.Z. Comparative study on endogenous release of nitrogen and phosphorus in Nansi Lake, China. *Acta Scientiae Circumstantiae*. **33** (2), 487, **2013**.
2. SMALL G.E., COTNER J.B., FINLAY J.C., STARK R.A., STERNER R.W. Nitrogen transformations at the sediment-water interface across redox gradients in the Laurentian Great Lakes. *Hydrobiologia*. **731** (1), 95, **2014**.

3. PAN M., ZHAO J., ZHEN S.C., HENG S., WU J. Effects of the combination of aeration and biofilm technology on transformation of nitrogen in black-odor river. *Water Sci. Technol.* **74** (3), 655, **2016**.
4. ZHANG H.H., BI C.J., CHEN Z.L., WANG X.P. Pollution characteristics of nitrogen and its influence factors in water and sediments of Dishui Lake system. *China Environmental Science*. **34** (10), 2646, **2014**.
5. WANG M., YAN H., JIAO L.X., WANG S.R., LIU W.B., LUO J., LUO Z.Q. Characteristics of internal nitrogen loading and influencing factors in Dianchi Lake sediment. *China Environmental Science*. **35** (1), 218, **2015**.
6. KUWABARA J.S., CARTER J.L., TOPPING B.R., FEND S.V., WOODS P.F., BERELSON W., BALISTRERI L. JR22-Importance of Sediments-Water Interactions in Coeur d'Alene Lake, Idaho, USA-Management Implications. *Environ. Manage.* **32** (3), 348, **2015**.
7. WU Y.P., CHEN J. Investigating the effects of point source and nonpoint source pollution on the water quality of the East River (Dongjiang) in South China. *Ecol. Indic.* **32** (9), 294, **2013**.
8. ZHAO Y., SHAN B.Q., TANG W.Z., ZHANG H. Nitrogen mineralization and geochemical characteristics of amino acids in surface sediments of a typical polluted area in the Haihe River Basin, China. *Environ. Sci. Pollut. R.* **22** (22), 17975, **2015**.
9. HE Y., SHEN S.Y., HUANG M.S., ZHANG B., YAO L. Research of nitrification-denitrification regarding endogenous nitrogen from urban malodorous river sediments: A review. *Ecology & Environmental Sciences*. **1** (6), 1166, **2012**.
10. SHEN S.Y., HE Y., HUANG M.S., GUP.D., RAO Y.F., YAO L.P., ZHAN Y.F. Effects of aerating disturbances on nitrification at sediment-water interface. *Chinese Journal of Environmental Engineering*. **8** (10), 4153, **2014**.
11. CHEN J.Y., XU Z.C., LUO Q.J., LIAO B.H., FU Q.W., HUANG B. Effect of aeration on the Transport and transformation of pollutants from the sediments in the tributary of Dianchi Lake. *Ecology and Environmental Sciences*. **17** (6), 2154, **2008**.
12. RUAN A.D., HE R., XU S.Y., LIN T. Effect of dissolved oxygen on nitrogen purification of microbial ecosystem in sediments. *J. Environ. Sci. Heal. A*. **44** (4), 397, **2009**.
13. BERZAGHI P., RIOVANTO R. Near infrared spectroscopy in animal science production: principles and applications. *Ital. J. Anim. Sci.* **8** (1), 3, **2016**.
14. ROMAN M., BALABIN R.S. Biodiesel classification by base stock type (vegetable oil) using near-infrared spectroscopy data. *Anal. Chim. ACTA*. **689** (2), 190, **2011**.
15. MELENDEZ-PASTOR I., ALMENDRO-CANDEL M.B., NAVARRO-PEDRERO J., GOMEZ I., LILLO M.G., HERNANDEZ E.I. Monitoring urban wastewaters' characteristics by visible and short wave near-infrared spectroscopy. *Water*. **5** (4), 2026, **2013**.
16. MENG Y., WANG S.S., CAI H., JIANG B.H., ZHAO W.J. Discrimination and Content Analysis of Fritillaria Using Near Infrared Spectroscopy. *J. Anal. Methods Chem.* **2015** (1), 101, **2015**.
17. LIU Y.Y., WAN W.B., SHEN J.T., JIANG Y., ZHANG L., JIANG H. Establishment of universal quantitative models for determination of acyclovir tables. *Chinese Pharmaceutical Journal*. **27** (2), 191, **2013**.
18. NIU X.Y., ZHAO Z.L., JIA K.J., LI X.T. A feasibility study on quantitative analysis of glucose and fructose in lotus root powder by FT-NIR spectroscopy and chemometrics. *Food Chem.* **133** (2), 592, **2012**.

19. DING S.F., QI B.J., TAN H.Y. An overview on theory and algorithm of support vector machines. *Journal of Electronic Science and Technology of China*. **40** (1), 2, **2011**.
20. HUANG Y., ZHANG L., LIAN G., ZHAN R., XU R., HUANG Y., MITRAB., WU J., LUO G. A novel mathematical model to predict prognosis of burnt patients based on logistic regression and support vector machine. *Burns Journal of the International Society for Burn Injuries*. **42** (2), 291, **2016**.
21. XIN N., GU X.F., WU H., HU Y.Z., YANG Z.L. Discrimination of raw and processed *Dipsacus asperoides* by near infrared spectroscopy combined with least squares-support vector machine and random forests. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. **89** (4) 18, **2012**.
22. YANG F., TIAN J., XIANG Y.H., ZHANG Z.Y., HARRINGTON P.D.B. Near infrared spectroscopy combined with least squares support vector machines and fuzzy rule-building expert system applied to diagnosis of endometrial carcinoma. *Cancer Epidemiol.* **36** (3), 317, **2012**.
23. MORA C.R., SCHIMLECK L.R. Kernel regression methods for the prediction of wood properties of *Pinus taeda* using near infrared spectroscopy. *Wood Sci. Technol.* **44** (44), 561, **2010**.
24. SHEN T.T., ZOU X.B., SHI J.Y., LI Z.H., HUANG X.W., XU Y.W., CHEN W. Determination Geographical Origin and Flavonoids Content of Goji Berry Using Near-Infrared Spectroscopy and Chemometrics. *Food Anal. Method.* **9** (1), 68, **2016**.
25. MI Y.P., WANG X.P., JIN X. Water COD prediction based on machine learning. *Journal of Zhejiang University (Engineering Science)*. **42** (5), 790, **2008**.
26. ZHANG G.Z., QIAO G.L., WU F.P., YANG H., NIU Y. Application of principal component analysis to an evaluation of water quality in water cellar. *Environ. Sci. Technol.* **37** (4), 181, **2014**.
27. ECKERT-GALLUP A.C., SALLABERRY A.J., DALLDAN A.R., NEARY V.S. Application of principal component analysis (PCA) and improved joint probability distributions to the inverse first-order reliability method (I-FORM) for predicting extreme sea states. *Ocean Eng.* **112** (15), 307, **2016**.
28. DENG X.L., KONG C., WU W.B., MEI H.L., LI Z. Detection of citrus HuangLongBing based on principal component analysis and back propagation neural network. *Acta Photonica Sinica*. **43** (4), 142, **2014**.
29. ZHANG H., WANG K., HUANG J., ZHANG Y., SONG J. Characterization of dissolved organic matter with intermittent aeration by fluorescence. *Analysis Letters*, **49** (12), 1874, **2016**.
30. HUPPERT T.J. Commentary on the statistical properties of noise and its implication on general linear models in functional near-infrared spectroscopy. *Neurophotonics*. **3** (1), **2016**. doi: 10.1117/1.NPh.3.1.010401.
31. HE J.C., YANG X.L., WANG L.R., PAN J.M. Rapid determination of chemical oxygen demand (COD) biochemical oxygen demand and (BOD₅) and pH in wastewater using near-infrared spectroscopy. *Acta Scientiae Circumstantiae*. **27** (12), 2105, **2007**.
32. DERNONCOURT D., HANCZAR B., ZUCKER J.D. Analysis of feature selection stability on high dimension and small sample data. *Comput. Stat. Data An.* **71** (1), 1, **2014**.