Heliyon 8 (2023) e12423

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

In-situ monitoring of nitrate in industrial wastewater using Fourier transform infrared attenuated total reflectance spectroscopy (FTIR-ATR) coupled with chemometrics methods

Fangqun Gan [a], Ke Wu [a], Fei Ma [b], Cuilan Wei [a], Changwen Du [b,c,*]

[a] College of Environment and Ecology, Jiangsu Open University, Nanjing, 210017, China
[b] The State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science Chinese Academy of Sciences, Nanjing, 210008, China
[c] College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China

ARTICLE INFO

Keywords:
Industrial wastewater
Nitrate
Fourier transform infrared attenuated total reflectance spectroscopy
Deconvolution
Chemometrics

ABSTRACT

Quantitative prediction of nitrate contents in different industrial wastewater was carried out using Fourier transform infrared attenuated total reflectance (FTIR-ATR) spectroscopy. The algorithm of Gaussian deconvolution was applied in the spectral range of 1500–1200 cm^{-1} to eliminate the background interferences on target information of nitrate, and partial least squares regression (*PLSR*) model and support vector machine (*SVR*) model were developed for the prediction of nitrate. The results showed that the *PLSR* model ($Rv^2 = 0.921$, $RMSEv = 0.351$ mg/L, $RPDv = 3.56$) and *SVR* model ($Rv^2 = 0.856$, $RMSEv = 0.473$ mg/L, $RPDv = 3.15$) reached excellent prediction accuracy and robustness for electroplating wastewater, and for metallurgical wastewater the *SVR* model ($Rv^2 = 0.916$, $RMSEv = 1.38$ mg/L, $RPDv = 3.26$) showed a better prediction performance. The *PLSR* and *SVR* models exhibited poor prediction accuracy of nitrate for pesticide wastewater and dyeing wastewater due to the strongly interference by carbonate. The spectra pretreatment by deconvolution dramatically improved the prediction models. Therefore, combined with deconvolution spectra pretreatment and chemometrics methods, FTIR-ATR could achieve a fast and effective in-situ monitoring of nitrate in industrial wastewater.

1. Introduction

Quality monitoring of industrial wastewater is one of the necessary ways for enterprises or wastewater treatment plants to practically realize the design, optimization, and control of wastewater treatment processes (Drolc et al., 2010). Nitrate is the most abundant form of inorganic nitrogen, and nitrate in wastewater is not only produced by the industrial production; but also generated in the nitrification reaction of wastewater treated by aeration. The discharge of excessive nitrate wastewater has severe impacts on the environment and human health. Therefore, in-situ determination of nitrate content in wastewater is of great significance.

The conventional methods for determining nitrate-nitrogen (NO_3^--N) in water include spectrophotometry, cadmium column reduction, ion chromatography, ion electrode method, and ultraviolet absorption spectrometry (Kelly et al., 2007; Gal et al., 2004; Lopez-Moreno et al., 2016; Morie et al., 1972; Singh et al., 2019). However, The process of spectrophotometry is tedious and time-consuming; the cadmium column reduction method generates toxic wastes such as cadmium in the process of quantitative determination, thereby resulting in secondary pollution; ion chromatography has the highest measurement accuracy among various nitrate-nitrogen detection methods, but the instrument is expensive and the maintenance is rather complicated, and the method is time-consuming; the ion electrode method can be easily interfered by other ions, and the detection accuracy is rather poor; ultraviolet absorption spectrometry shows a simple measurement process, however, the organic matter and turbidity can interfere with this method's determination. Recently, various sensors have been widely developed and used for nitrate detection in water bodies. A planar interdigital sensor used for nitrate detection in lake, stream, river and canal water has been reported (Alahi and Mukhopadhyay, 2018; Alahi et al., 2018). The sensor's material was developed by an imprinting polymer technique, and had a detection range from 1–10 mg/L. The plasmonic sensors exhibit excellent plasmonic properties and have been applied for selective and label-free detection of very low concentrations of aquatic pollutants (Brahmkhatri et al., 2021). A rapid method for detection of nitrate (NO_3^-) was developed using cysteamine functionalized AuNP (CyAuNPs) as

* Corresponding author.
E-mail address: chwdu@issas.ac.cn (C. Du).

https://doi.org/10.1016/j.heliyon.2022.e12423
Received 25 August 2022; Received in revised form 3 November 2022; Accepted 9 December 2022

plasmonic sensor (Mura et al., 2015). Rapid and ultra-sensitive detection of nitrate ions can be achieved by using amino group modified graphene oxide sensors. The proposed method exhibited remarkable sensitivity, yielding a very low limit of detection of 5 nm, and required only 20 μL of nitrate solution to complete the detection within 75 min (Ren et al., 2015). In addition, sensitive conductometric enzyme biosensor has also been used for nitrate determination (Wang et al., 2006). Although many studies are still going on to develop robust sensors which could be utilized continuously, repeatability remains a major challenge for sensors (Alahi et al., 2018).

In recent years, Fourier transform infrared attenuated total reflectance spectroscopy (FTIR-ATR) has shown unique advantages of determining nitrate. The nitrate in soil and solution was quantitatively determined by FTIR-ATR (Shaviv et al., 2003; Shao et al., 2017a, 2017b), and in-situ monitoring of nitrate content in vegetables was also achieved (Ma et al., 2021). Additionally, typical nitrate absorption of FTIR-ATR was isolated for natural aquatic water sample by deconvolution, and a rapid and accurate quantitative determination was achieved combined with *PLSR* model (Gan et al., 2020; Wu et al., 2022). FTIR-ATR has the advantages of fast, non-destructive, reliable, and in-situ monitoring, which is time- and cost-saving, thereby providing a possible strategy for monitoring nitrate in different types of industrial wastewater. However, industrial wastewater has a rather complex composition with various organic and inorganic materials, and the molecular bonds between different substances are not independent but interrelated, resulting in mutual interference with the absorption peaks of these molecularity bonded groups. Peak-differentiating of typical absorption band is a strategy to extract the target information from the overlapping peaks, which can remove the interferences, and Gaussian deconvolution is a useful algorithm to reach this purpose.

Herein, electroplating wastewater, pesticide wastewater, dyeing wastewater and metallurgical wastewater were collected, and were involved in the determination of nitrare using FTIR-ATR, and two algorithms (partial least squares regression, *PLSR* and support vector regression, *SVR*) were applied to optimize the quantitative prediction of nitrate, which provides an alternative method for rapid monitoring of nitrate in industrial wastewater.

2. Experimental section

2.1. Wastewater sample collection

Four types of industrial wastewater, i.e., electroplating wastewater (12 samples), pesticide wastewater (12 samples), dyeing wastewater (12 samples) and metallurgical wastewater (20 samples), were collected from different enterprises. All the samples were kept in a refrigerator at 4 °C for use.

2.2. Chemical assay

The nitrate contents in water samples were determined by Smartchem200 automatic discrete chemical analyzer (AMS Alliance, Frepillon, France), three replicate measurements per sample. A pH meter (Orion Star A211, Thermo Fisher, USA) was used to determine the pH of wastewater. The total organic carbon (TOC) contents were measured using a TOC-VCPH Analyzer (Shimadzu, Japan). The carbonate contents in the water samples were analyzed by acid-base indicator titration (State Environmental Protection Administration 2002).

2.3. Spectra determination and pretreatment

The water samples were detected by Nicolet6700 FTIR spectrometer (Thermo Fisher Scientific, USA) attached with ZnSe at 45 °C. During the determination, an appropriate amount of solution sample was placed in the sample cell, then the sample cell was completely blow-dried with pure nitrogen following each determination. the wavenumber range was set at 4000–500 cm^{-1}, the mirror velocity was 0.32 cm s^{-1}, and the spectral resolution was 4 cm^{-1}. Each spectrum was scanned for 32 consecutive repetitions with three replications.

FTIR-ATR spectra were prepossessed by Savitzky–Golay smoothing filter to eliminate baseline floating and noise to improve the signal-to-noise ratio (Savitzky and Golay, 1964; Du et al., 2019). According to the absorption characteristics of nitrate nitrogen, the spectrum in the 1500–1200 cm^{-1} region was subjected to smoothing, baseline correction, and Gaussian deconvolution with Peak-fit 4.12 software. The deconvolution process has been elaborated in detail in the supporting information (Figure S1) (Zou and Unbehauen, 1995).

The deconvolution of the spectral curve assumes that several single-peak spectral bands superimpose the experimental spectrum Y(x). The purpose of fitting is to find the single-peak spectral band $F_i(x)$ $(i = 1,2..., n)$ (Buslov et al., 1997; Buslov et al., 2002). The principle is as follows:

$$Y\left(x\right) = \sum Fi(x) \tag{1}$$

where Y is the spectrum, x is the wave number, i $(1, 2, 3..., n)$ is the number of independent peaks, and F is the kernel function of the expansion or deconvolution. The Gaussian function is taken as the kernel function:

$$y = \frac{a_0}{\pi\sqrt{\pi a_2}} \exp\left[-\frac{1}{2}\left(\frac{x - a_1}{a_2}\right)^2\right] \tag{2}$$

where a_0, a_1, and a_2 represent peak amplitude, position, and width, respectively, while x and y represent wave number and absorption intensity.

2.4. Prediction models

2.4.1. Partial least squares regression (PLSR) model

PLSR model is one of the most commonly used stoichiometry algorithms in spectral data analysis. It was a bilinear model where a matrix X, containing the variables (spectra wave-number), and matrix Y, a function of the variables of X (nitrate contents), were used to predict the smallest number of latent variables. Cross-validation is a statistically sound method for choosing the number of components in *PLSR* model. It avoids over-fitting data by not reusing the same data to both fit a model and estimate the prediction error. Thus, estimate of the prediction error is not optimistically biased downwards. *PLSR* has an option to estimate the root-mean-squared-error (*RMSE*) by cross-validation. When the model has the first lowest *RMSE*, the corresponding number of factor is optimal. In this study, the model's optimal factor number of latent variables was determined based on the minimal *RMSE* of cross-validation by leave-one-out cross-calibration (Ma et al., 2019).

2.4.2. Support vector regression (SVR) model

SVR is a machine learning algorithm based on statistical learning theory, aiming to maximize the generalization ability by using the structural risk minimization principle (Filgueiras et al., 2014a). According to the definition of the error function, there are two *SVR* models: (1) Epsilon-*SVR* (*ε-SVR*), which optimizes the model using the adjustable parameters epsilon (upper tolerance on prediction errors) and C (cost of prediction error is more remarkable than epsilon); (2) *Nu-SVR*, which optimizes a model using the adjustable parameter *Nu* (Jimenez-Carvelo et al., 2017). Detailed information on the *SVR* algorithm can be found in relevant literature (Devos et al., 2009; Li et al., 2009).

In this study, the *ε-SVR* was selected for modeling. Nonlinear radial basis function (RBF) was used as kernel function. The RBF kernel, which is widely used in many problems, eventually was applied for this study.

Table 1. Physicochemical properties of four types of industrial wastewater.

Wastewater types	Nitrate contents (mg/L)			pH			TOC contents (mg/L)			Carbonate contents (mg/L)		
	MIN	MAX	AVG	MIN	MAX	AVG	MIN	MAX	AVG	MIN	MAX	AVG
Electroplating wastewater	2.62	14.5	4.93	1.72	2.52	2.10	47.5	73.6	56.6	0.52	1.32	0.80
Pesticide wastewater	3.82	12.8	5.48	4.61	9.95	7.50	62.2	79.7	73.5	15.7	25.5	23.0
Dyeing wastewater	4.81	19.9	9.43	8.53	12.31	10.6	212	262	232	485	986	731
Metallurgical wastewater	2.45	46.3	18.4	1.25	2.26	1.70	41.5	49.2	45.2	0.46	0.98	0.70

Notes: TOC, total organic carbon; MIN, minimum; MAX, maximum; AVG, average.

The RBF is also a simple function and can model systems of varying complexity. The different *SVR* parameters needed for each model were optimized with leave-one-out-cross-validation for the training set in terms of root mean squared error (*RMSE*). During this process, keeping the value of ε equal to 0.1, for C and γ a broad range of parameter settings was investigated with large steps and after identifying a promising region, this region was searched in detail. A new prediction model according to the *SVR* is being recommended in this paper to predict nitrate in industrial wastewater. Aiming to create a powerful *SVR* model, the *SVR* parameters were set up carefully. *SVR* looks for to minimize the generalization error to gain generalized performance rather than minimizing the training error.

2.5. Calibration and validation datasets

Spectral data sets of each type of wastewater were randomly divided into a calibration data set containing 75% spectra and a validation data set containing the remaining 25% spectra.

2.6. Model evaluation

R^2, *RMSE,* and *RPD* were used to evaluate the model's predictive power. The following formula can define these evaluation parameters:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{3}$$

$$RPD = \frac{SD}{RMSE} \tag{4}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{5}$$

where y_i and $\hat{y}_i$ respectively denote the measurement value and predicted value of nitrate in the e_{th} sample; $\bar{y}$ is the mean value of the measured nitrate content; *SD* is the standard deviation. $RMSE_C$ and $RMSE_P$ represent the root mean square error (*RMSE*) in the

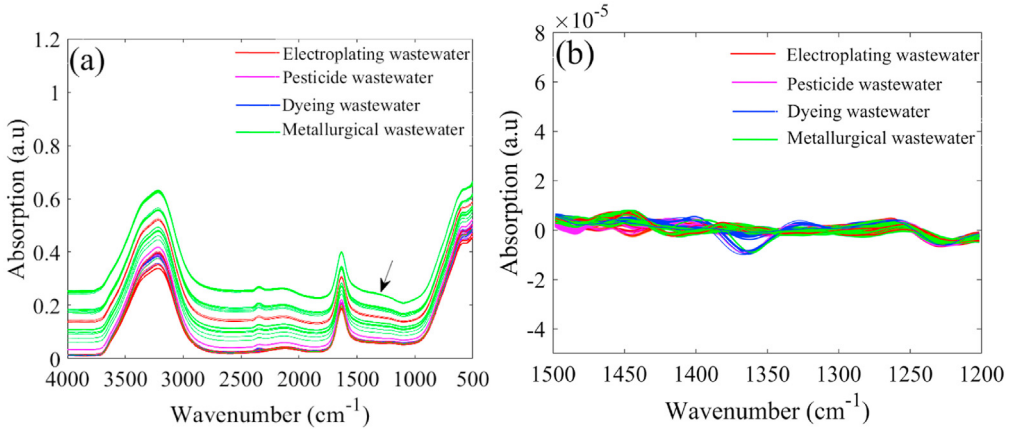

Figure 1. Typical FTIR-ATR absorption spectra of wastewater (a), the second derivative nitrate spectra range from 1500 to 1200 cm^{-1} (b).

Table 2. Statistics of the *PLSR* and *SVR* models used in the calibration and validation sets for the prediction of nitrate contents in different types of wastewaters using original FTIR-ATR spectra.

Wastewater types	Datasets for modeling	Calibration			Validation			Bias
		R_C^2	$RMSE_C$	RPD_C	R_V^2	$RMSE_V$	RPD_V	
Electroplating wastewater	*PLSR*	0.663	1.64	1.67	0.198	2.49	0.591	−0.136
	SVR	0.712	1.27	1.68	0.190	2.67	0.625	−0.116
Pesticide wastewater	*PLSR*	0.507	1.73	0.921	0.408	2.13	0.761	0.127
	SVR	0.515	1.62	0.965	0.347	2.31	0.635	−0.096
Dyeing wastewater	*PLSR*	0.921	0.253	3.23	0.104	6.13	0.749	0.023
	SVR	0.652	3.67	1.47	0.412	5.26	0.812	0.193
Metallurgical wastewater	*PLSR*	0.735	7.86	1.52	0.519	11.5	1.26	0.178
	SVR	0.703	8.37	1.39	0.374	12.6	0.563	0.133

Notes: *PLSR*: partial least squares regression; *SVR*: support vector regression; *RMSE*: the root mean square error; *RPD*: the residual prediction deviation.

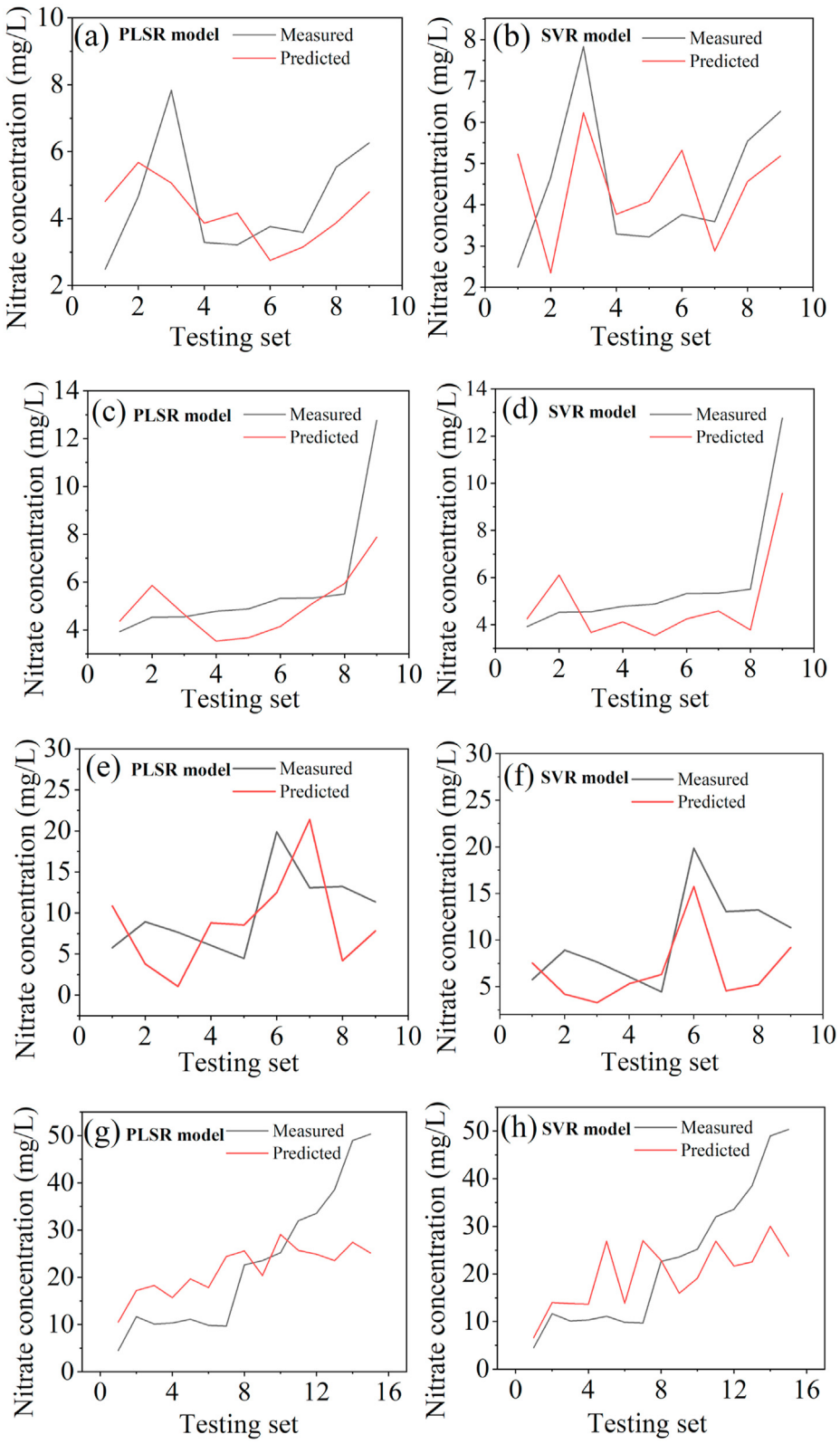

Figure 2. Measured and predicted nitrate concentration of testing set in electroplating wastewater (a, b), pesticide wastewater (c, d), dyeing wastewater (e, f) and metallurgical wastewater (g, h) based on original FTIR-ATR spectra by *PLSR* and *SVR* models.

calibration and validation dataset models. High R^2, *RPD*, and low *RMSE* define the robustness and accuracy of the models. *RPD* less than 1.8 is considered unsuitable for quantitative measurement; *RPD* between 2 and 2.5 indicates good model prediction performance; while an *RPD* value higher than 3 indicates excellent model prediction performance.

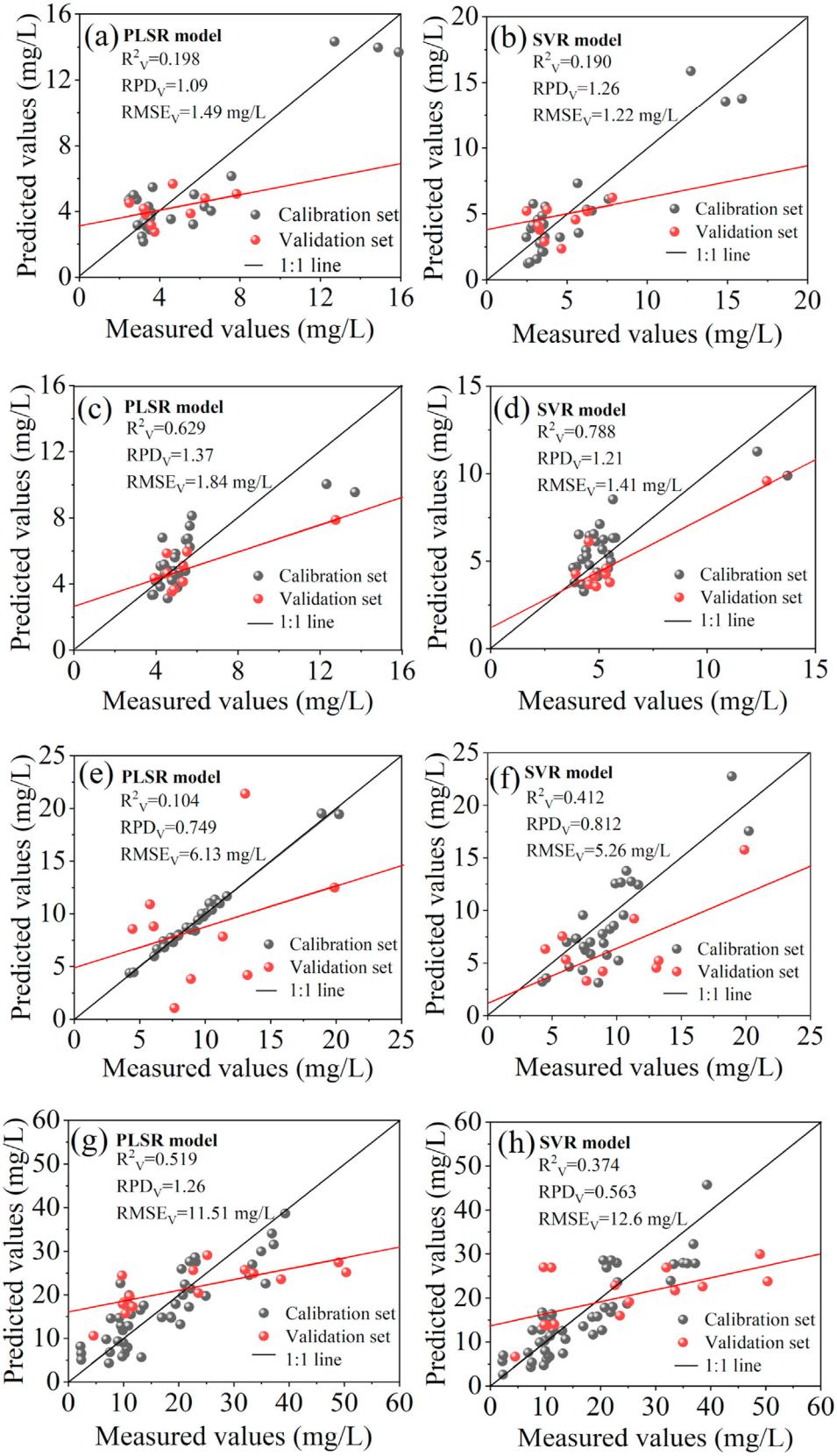

Figure 3. Measured values vs. the predicted values of nitrate contents in electroplating wastewater (a, b), pesticide wastewater (c, d), dyeing wastewater (e, f) and metallurgical wastewater (g, h) based on original FTIR-ATR spectra by *PLSR* and *SVR* models.

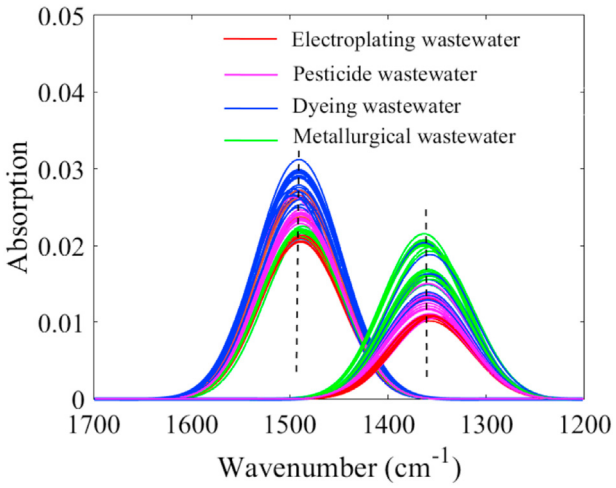

Figure 4. Deconvoluted FTIR-ATR spectra in the range of 1700–1200 cm^{-1}.

2.7. Systematic error assessment

Systematic errors may occur due to the inadequacy of calibration models. The bias is the sum of the differences between the estimated value y_i and the measured value x_i (Filgueiras et al., 2014b; Neto et al., 2017a), which can be expressed as below:

$$bias = \frac{\sum(y_i - x_i)}{n} \quad (6)$$

where n is the number of samples. A bias value close to zero indicates a low systematic error between the measured and predicted values (Neto et al., 2017b).

3. Results

3.1. Properties and characteristics of wastewater

The physicochemical properties of four types of industrial wastewater samples i.e., electroplating wastewater ($n = 12$), pesticide wastewater ($n = 12$), dyeing wastewater ($n = 12$) and metallurgical wastewater ($n = 20$) were summarized in Table 1, representing the minimum, maximum and average values of different parameters. The average nitrate content of electroplating, pesticide, dyeing, and metallurgical wastewater were 4.93, 5.48, 9.43, and 18.4 mg/L, respectively. And the average pH values of electroplating wastewater, pesticides wastewater, dyeing wastewater and metallurgical wastewater were 2.10, 7.50, 10.6 and 1.70, respectively, indicating that the pH values of different types of wastewater varied greatly. The total organic carbon (TOC) contents of dyeing wastewater were much higher than those of other wastewater, reaching 232 mg/L. In addition, the carbonate contents of electroplating, metallurgical, and pesticide wastewater were 0.80, 0.70 and 23.0 mg/L, respectively, while the corresponding content in dyeing wastewater reached as high as 731 mg/L. These results show that there are obvious differences in the background and property of different types of wastewater, which poses a challenge to the determination of nitrate content by spectrometry.

3.2. FTIR-ATR spectra of industrial wastewater

The total spectral shapes and typical bands were similar among the different types of wastewater (Figure 1a). It can be seen that there are two prominent absorption peaks in the range of 3800–3000 cm^{-1} and 1800–1500 cm^{-1}, both of which signify the characteristic absorption peaks of water. The distinct nitrate peak at 1500–1200 cm^{-1} is much weaker than water and is rather difficult to observe directly. As shown in Figure 1b, the second-order derivative spectra ranging from 1500–1200 cm^{-1} were then calculated and plotted. The region did not show any prominent and regular characteristic peaks. There may be two reasons for this observation. First, the infrared absorption peak of nitrate was a small signal, and the nitrate-nitrogen concentration in the wastewater was relatively low, leading to a low intensity of its characteristic absorption peak. Second, the distinct absorption peak of nitrate was easily disturbed by water and other substances (inorganic and organic matter).

3.3. Nitrate prediction model using original spectra

Cross-validation was employed to obtain the optimum number of principal component, and they were 4, 3, 5 and 4 in the electroplating wastewater, pesticide wastewater, dyeing wastewater and metallurgical wastewater respectively, which correspond to the minimum of *RMSECV* (Figure S2). Therefore, the first 4, 3, 5 and 4 principal component were used to establish the *PLSR* models. In addition, quantitative prediction models of nitrate contents in different industrial wastewater were also established by combining original spectra in the range of 1500–1200 cm^{-1} with *SVR* algorithms. The prediction results showed that the validation model was comparable to those from the calibration measurement, indicating that the models were robust and stable (Table 2). However, the values between measured and predicted nitrate concentration did not show good consistence (Figure 2a, b, c, d, e, f, g, h), and the *RPD* values of all models in the validation set were <1.8, implying an unreliable nitrate prediction performance (Figure 3a, b, c, d, e, f, g, h). The original spectra contain a wealth of information due to the complex backgrounds of wastewater; therefore, further preprocess of the original spectra should be conducted to effectively extract the characteristic absorption peaks of nitrate from the complex spectral information.

3.4. Spectral Gaussian deconvolution

To extract the characteristic absorption of nitrate from rather complex spectral information, Peakfit 4.12 software was used to conduct Gaussian deconvolution of the spectra in the 1700–1200 cm^{-1} region. In the deconvolution process, all values demonstrating the spectral goodness of fit (R^2) were above 0.95, and all sum of squared errors (*SSE*) was less than 0.116, presenting a good deconvolution effect. The spectrum after deconvolution is shown in Figure 4. Two characteristic absorption peaks were observed among the different types of wastewater. The absorption peak corresponding to nitrate was approximately 1360 cm^{-1}, and the absorption intensity was visually proportional to the nitrate content. The peak around 1485 cm^{-1} was attributed to the absorption of carbonate, which was close to the characteristic absorption peak of nitrate. The position of the characteristic peak of nitrate at different carbonate concentrations remained unchanged.

3.5. Prediction model of nitrate using deconvoluted spectra

Similarly, the optimum number of principal component in the electroplating wastewater, pesticide wastewater, dyeing wastewater and metallurgical wastewater were 6, 3, 3 and 4, respectively (Figure S3). The prediction models of nitrate-based on deconvoluted spectra in the range of 1500–1200 cm^{-1} were established. The predicted nitrate concentration were in good agreement with the measured values in the electroplating and metallurgical wastewater, but the predicted results in the pesticide and dyeing wastewater were not satisfactory (Figure 5a, b, c, d, e, f, g, h). For the electroplating wastewater, the *PLSR* ($Rv^2 = 0.921$, $RMSEv = 0.351$ mg/L, $RPDv = 3.56$) and *SVR* ($Rv^2 = 0.856$, $RMSEv = 0.473$ mg/L, $RPDv = 3.15$) models presented excellent prediction accuracy (Figure 6a, b). All the models for nitrate prediction in the pesticide and dyeing wastewater presented relatively low RPD_V values < 1.8, implying a poor prediction performance (Figure 6c, d, e, f). Carbonate in solution probably interfered with nitrate's spectral characteristic, which further affects the prediction accuracy of the models. This result is

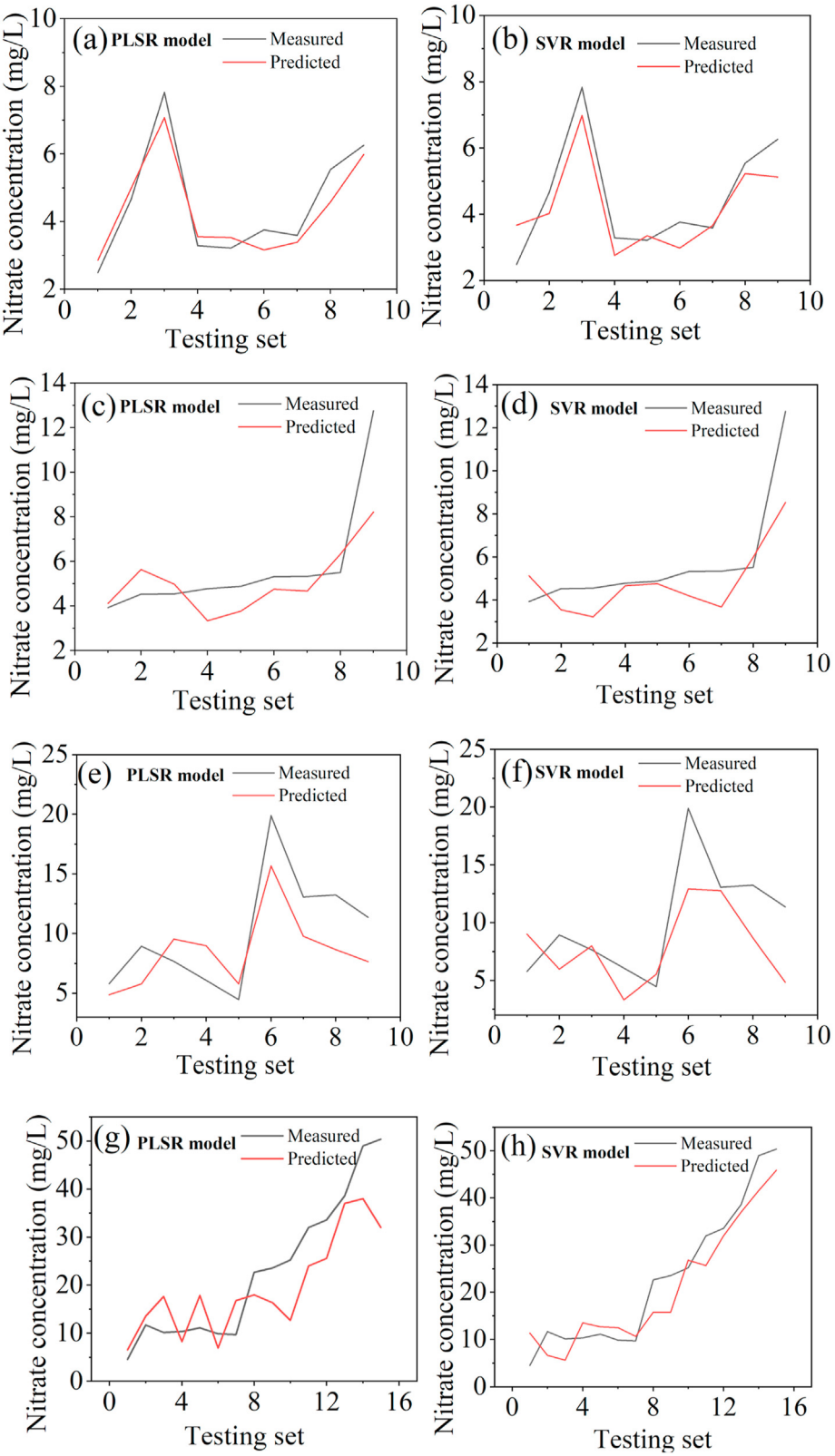

Figure 5. Measured and predicted nitrate concentration of testing set in electroplating wastewater (a, b), pesticide wastewater (c, d), dyeing wastewater (e, f) and metallurgical wastewater (g, h) based on deconvoluted FTIR-ATR spectra by *PLSR* and *SVR* models.

consistent with a previous study (Wu et al., 2022). By adding hydrochloric acid, the carbonate in the pesticide wastewater and dyeing wastewater were eliminated. As a result, the two models' prediction performance was significantly improved, which further verified the carbonate interference (Figure S2). For the nitrate prediction in the metallurgical wastewater, the *SVR* model ($Rv^2 = 0.916$, $RMSEv = 1.38$ mg/L, $RPDv = 3.26$) provided a better prediction accuracy than the *PLSR* ($Rv^2 = 0.785$, $RMSEv = 2.86$ mg/L, $RPDv = 2.17$) model (Figure 6g, h). In

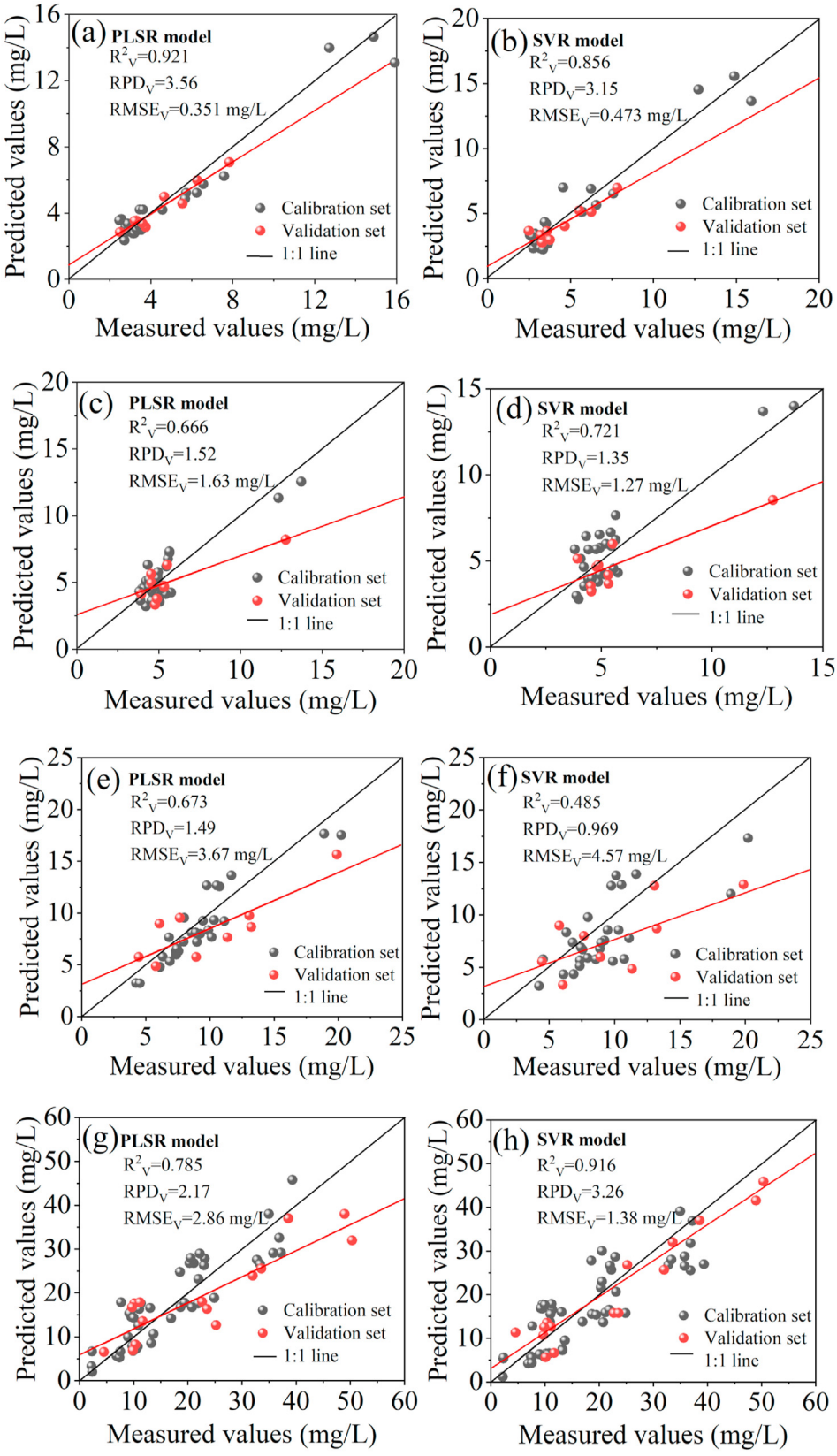

Figure 6. Measured values *vs.* the predicted values of nitrate contents in electroplating wastewater (a, b), pesticide wastewater (c, d), dyeing wastewater (e, f) and metallurgical wastewater (g, h) based on deconvoluted FTIR-ATR spectra by *PLSR* and *SVR* models.

addition, the prediction results of the calibration set and validation set models were similar, indicating that the prediction model was robust and stable (Table 2). In comparison to the models based on original spectra for the nitrate's prediction in the electroplating and metallurgical wastewater, the scatter of the calibration and validation sets in the corresponding model located closer to the 1:1 line, which also suggested a

Table 3. Statistics of the *PLSR* and *SVR* models used in the calibration and validation sets for the prediction of nitrate contents in different types of wastewaters using deconvoluted FTIR-ATR spectra.

Wastewater types	Datasets for modeling	Calibration			Validation			Bias
		R_C^2	$RMSE_C$	RPD_C	R_V^2	$RMSE_V$	RPD_V	
Electroplating wastewater	*PLSR*	0.875	0.476	3.34	0.921	0.351	3.56	−0.093
	SVR	0.821	0.553	2.68	0.856	0.473	3.15	−0.048
Pesticide wastewater	*PLSR*	0.738	1.51	1.67	0.666	1.63	1.52	0.073
	SVR	0.787	1.02	1.58	0.721	1.27	1.35	0.091
Dyeing wastewater	*PLSR*	0.798	2.89	1.73	0.673	3.67	1.49	−0.213
	SVR	0.723	3.22	1.59	0.485	4.57	0.969	0.178
Metallurgical wastewater	*PLSR*	0.835	2.26	2.76	0.785	2.86	2.17	0.126
	SVR	0.879	1.68	3.02	0.916	1.38	3.26	0.059

Notes: *PLSR*: partial least squares regression; *SVR*: support vector regression; *RMSE*: the root mean square error; *RPD*: the residual prediction deviation.

better prediction of nitrate contents. The above results indicated that the high concentration of carbonate ions in pesticide wastewater and dyeing wastewater might impact the nitrate absorption, and lead to a decrease of the model accuracy, which could not be used for the quantitative prediction of nitrate. For electroplating wastewater and metallurgical wastewater, the obtained models still showed excellent prediction accuracy for nitrate due to the low concentration of carbonate ions. In our previous study we have also demonstrated that the nitrate prediction model performed well in condition that the carbonate content in a low level with less than 10 mg/L (Wu et al., 2022).

In the linear multivariate correction, the limit of detection (*LOD*) of the model can reflect the sensitivity of the model detection. The *LOD* of the model can be expressed as $3\sigma/m$, where σ is the standard deviation of the predicted concentration and can be replaced by *RMSEv*; m is the slope of the fitting curve of the model (*X*-axis is the measurement value, *Y*-axis is the predicted value) (Doran et al., 2000). Therefore, the *LOD* of *PLSR* and *SVR* models for nitrate-predicting electroplating wastewater were 1.39 and 2.10 mg/L, respectively. For the metallurgical wastewater, the *LOD* of the *SVR* model was 5.11 mg/L. These results further demonstrated the excellent prediction ability of the models.

Systematic error is a deviation in measurement, which leads to significant differences between the average value of many individual measurements and the actual value of the measurement attribute, which may be caused by imperfect calibration. A bias value close to zero indicates a low systematic error between the measured and the predicted values. The biases of the models are shown in Table 3. According to the absolute value, all the three models for nitrate prediction of the electroplating and metallurgical wastewater showed lower systematic errors than those of the corresponding models based on original spectra (Table 2), indicating that these models possessed the potential for nitrate contents prediction in industrial wastewater.

4. Discussion

4.1. Application of FTIR spectroscopy in nitrogen determination

Several prediction models based on FTIR-ATR were developed to determine nitrate content in industrial wastewater and demonstrated excellent prediction results. In recent years, FTIR spectroscopy has been applied to determine nitrogen content in many fields. FTIR-ATR spectroscopy combined with the *PLSR* model rapidly determined nitrate content in Chinese cabbage (Yang et al., 2013). FTIR-ATR spectroscopy also showed unique advantages in determining nitrate in soil. Shao et al. quantified nitrate content in soil using FTIR-ATR, and they found that the second derivative of the nitrate characteristic zone was proportional to nitrate content, with a correlation coefficient reaching 0.975, thereby implying that FTIR-ATR combined with the second derivative can be successfully employed for the quantitative determination of soil nitrate (Shao et al., 2017a). An interesting study used FTIR-PAS to study the distribution of nitrogen in Chinese cabbage leaves. The results showed that the distribution pattern of nitrogen in leaves was random even in each treatment repetition, and the light was the main factor contributing to the random distribution of nitrogen (Li et al., 2016). In addition, the ratio of amide II to amide I in rice leaves based on mid-infrared spectroscopy has been used to diagnose nitrogen nutrient levels in rice at different stages of growth (Wu et al., 2019). Therefore, the application of FTIR spectra in agriculture, environment, food and industry can be further explored in the future based on the characteristic absorption peaks of nitrogen.

4.2. Application of deconvolution in spectral data preprocessing

Due to the broadening effect of the instrument response function (also known as point spread function, *PSF*), the spectrum recorded by the spectrometer often exhibits band overlap and signal interference, resulting in a decrease in spectral resolution. Therefore, spectral recovery is extremely important for spectral analysis. As a mathematical operation process, deconvolution is a typical signal extraction and recovery method (Jansson 1984; Zou and Unbehauen, 1995), showing unique advantages in spectral processing. Deconvolution technology can improve the resolution beyond the instrument's limit and significantly improve the signal-to-noise ratio (Du et al., 2019; Wang et al., 2019). To obtain useful, accurate and reliable information, spectral deconvolution can be associated with Gaussian fitting of absorption spectrum to adjust the Gaussian mathematical curve and obtain corresponding characteristic absorption peaks from overlapping peaks of a complex spectrum (Barth et al., 2009). Rapidly detecting nitrogen nutrient levels in rice can be done by combining mid-infrared photoacoustic spectroscopy with Gaussian deconvolution (Wu et al., 2019). Additionally, nitrate contents in solutions were also determined using FTIR-ATR spectra and the characteristic absorption peak of nitrate from the original spectral information was obtained using Gaussian deconvolution. Combination of FTIR-ATR with the *PLSR* model yielded a rapid yet accurate quantitative determination of nitrate concentration in standard nitrate solution and natural water bodies (Gan et al., 2020; Wu et al., 2022). In this study, Gaussian deconvolution was performed on the 1500–1200 cm^{-1} band of the original spectrum. Then two prediction models, namely *PLSR* and *SVR*, were established for determining nitrate contents in different types of industrial wastewater. The results demonstrated that the predictive ability of the models was significantly improved following deconvolution, indicating that deconvolution could effectively extract target spectral signals from complex overlapping spectra.

5. Conclusions

In this study, the nitrate content in electroplating wastewater, pesticide wastewater, dyeing wastewater and metallurgical wastewater were determined by FTIR-ATR spectroscopy. The typical absorption band of

nitrate was extracted by Gaussian deconvolution from the original spectrum. Quantitative prediction models of nitrate in industrial wastewater were established by deconvolution band combined with chemometrics methods. The following conclusions were obtained:

1) The *PLSR* and *SVR* models gave excellent accuracy and robustness for the nitrate prediction of electroplating wastewater. The *SVR* model exhibited a better prediction performance for metallurgical wastewater.
2) The *PLSR* and *SVR* models showed poor nitrate prediction of pesticide wastewater and dyeing wastewater due to the strongly interference of carbonate.
3) The FTIR-ATR spectroscopy combined with Gaussian deconvolution pre-treatment and stoichiometric algorithms could provide a fast and effective method for in-situ monitoring of nitrate content in specific types of industrial wastewater.

Declarations

Author contribution statement

Fangqun Gan; Changwen Du: Conceived and designed the experiments; Wrote the paper.

Cuilan Wei; Ke Wu; Fei Ma: Performed the experiments.

Ke Wu; Fangqun Gan: Analyzed and interpreted the data.

Changwen Du: Contributed reagents, materials, analysis tools or data.

Funding statement

Dr. Fangqun Gan was supported by the National Natural Science Foundation of China [41907154], the National Natural Science Foundation of Jiangsu Province [BK20191110], “Green blue project”of Jiangsu University [2019].

Prof. Changwen Du was supported by the National Natural Science Foundation of China [42077019]. Dr. Fei Ma was supported by the Strategic Priority Research Program of Chinese Academy of Sciences [XDA23030107].

Data availability statement

No data was used for the research described in the article.

Declaration of interest's statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e12423.

References

Alahi, M.E.E., Mukhopadhyay, S.C., 2018. Detection methods of nitrate in water: a review. Sensor Actuat. A-Phys 280, 210–221.

Alahi, M.E.E., Mukhopadhyay, S.C., Burkitt, L., 2018. Imprinted polymer coated impedimetric nitrate sensor for real-time water quality monitoring. Sensor Actuat. B-Chem 259, 753–761.

Barth, A., Haris, P.I., 2009. Biological and Biomedical Infrared Spectroscopy. IOS Press, Amsterdam, The Netherlands.

Brahmkhatri, V., Pandit, P., Rananaware, P., DSouza, A., Kurkuri, M.D., 2021. Recent progress in detection of chemical and biological toxins in water using plasmonic nanosensors. Trends Environ. Analy. Chem 30, e00117.

Buslov, D.K., Nikonenko, N.A., 1997. Regularized method of spectral curve deconvolution. Appl. Spectrosc. 51, 666–672.

Buslov, D.K., Nikonenko, N.A., Sushko, N.I., Zhbankov, R.G., 2002. Analysis of the structure of the bands in the IRspectrum of β-D glucose by the regularized method of deconvolution. J. Appl. Spectrosc. 69, 817–824.

Devos, O., Ruckebusch, C., Durand, A., Duponchel, L., Huvenne, J.P., 2009. Support vector machines (SVM) in near infrared (NIR) spectroscopy: focus on parameters optimization and model interpretation. Chemometr. Intell. Lab 96 (1), 27–33.

Doran, E.M., Yost, M.G., Fenske, R.M., 2000. Measuring dermal exposure to pesticide residues with attenuated total reflectance fourier transform infrared (ATR-FTIR) spectroscopy. Bull. Environ. Contam. Toxicol. 64, 666–672.

Drolc, A., Vrtovsek, J., 2010. Nitrate and nitrite nitrogen determination in waste water using on-line UV spectrometric method. Bioresour. Technol. 101, 4228–4233.

Du, H.Y., Yi, R.G., Dong, L.Q., Liu, M., Jia, W., Zhao, Y.J., Liu, X.H., Hui, M., Kong, L.Q., Chen, X., 2019. Rotating asymmetrical phase mask method for improving signal-to-noise ratio in wave front coding systems. Appl. Opt. 58, 6157–6164.

Filgueiras, P.R., Sad, C.M.S., Loureiro, A.R., Santos, M.F.P., Castro, E.V.R., Dias, J.C.M., Poppi, R.J., 2014a. Determination of API gravity, kinematic viscosity and water content in petroleum by ATR-FTIR spectroscopy and multivariate calibration. Fuel 116, 123–130.

Filgueiras, P.R., Alves, J.C.L., Sad, C.M.S., Castro, E.V.R., Dias, J.C.M., Poppi, R.J., 2014b. Evaluation of trends in residuals of multivariate calibration models by permutation test. Chemometr. Intell. Lab 133, 33–41.

Gal, C., Frenzel, W., Moller, J., 2004. Re-examination of the cadmium reduction method and optimisation of conditions for the determination of nitrate by flow injection analysis. Microchim. Acta 146, 155–164.

Gan, F.Q., Wu, K., Ma, F., Du, C.W., 2020. In situ determination of nitrate in water using fourier transform mid-infrared attenuated total reflectance spectroscopy coupled with deconvolution algorithm. Molecules 25, 5838.

Jansson, P.A., 1984. Deconvolution: with Applications in Spectroscopy. Academic Press, New York, USA.

Jimenez-Carvelo, A.M., Gonzalez-Casado, A., Cuadros-Rodriguez, L., 2017. A new analytical method for quantification of olive and palm oil in blends with other vegetable edible oils based on the chromatographic fingerprints from the methyl-transesterified fraction. Talanta 164, 540–547.

Kelly, R.T., Love, N.G., 2007. Ultraviolet spectrophotometric determination of nitrate: detecting nitrification rates and inhibition. Water Environ. Res. 79, 808–812.

Li, H.D., Liang, Y.Z., Xu, Q.S., 2009. Support vector machines and its applications in chemistry. Chemometr Intell Lab 95 (2), 188–198.

Li, C.Y., Du, C.W., Zeng, Y., Ma, F., Shen, Y.Z., Xing, Z., Zhou, J.M., 2016. Two-dimensional visualization of nitrogen distribution in leaves of Chinese cabbage (*brassica rapa* subsp. *chinensis*) by the fourier transform infrared photoacoustic spectroscopy technique. J. Agric. Food Chem. 64 (41), 7696–7701.

Lopez-Moreno, C., Perez, I.V., Urbano, A.M., 2016. Development and validation of an ionic chromatography method for the determination of nitrate, nitrite and chloride in meat. Food Chem. 194, 687–694.

Ma, F., Du, C.W., Zhou, J.M., Shen, Y.Z., 2019. Investigation of soil properties using different techniques of mid-infrared spectroscopy. Eur. J. Soil Sci. 70 (1), 96–106.

Ma, F., Du, C.W., Zheng, S.L., Du, Y.X., 2021. In situ monitoring of nitrate content in leafy vegetables using attenuated total reflectance−fourier-transform mid-infrared spectroscopy coupled with machine learning algorithm. Food Anal. Methods 14, 2237–2248.

Morie, G.P., Ledford, C.J., Glover, C.A., 1972. Determination of nitrate and nitrite in mixtures with a nitrate ion electrode. Anal. Chim. Acta 60 (2), 397–403.

Mura, S., Greppi, G., Roggero, P.P., Musu, E., Pittalis, D., Carletti, A., Ghiglieri, G., Irudayaraj, J., 2015. Functionalized gold nanoparticles for the detection of nitrates in water. Int. J. Environ. Sci. Technol. 12, 1021–1028.

Neto, A.J., Toledo, J.V., Zolnier, S., Lopes, D.D.C., Pires, C.V., Silva, T.G.F.D., 2017a. Prediction of mineral contents in sugarcane cultivated under saline conditions based on stalk scanning by Vis/NIR spectral reflectance. Biosyst. Eng. 156, 17–26.

Neto, A.J., Lopes, D.C., Pinto, F.A.C., Zolnier, S., 2017b. Vis/NIR spectroscopy and chemometrics for non-destructive estimation of water and chlorophyll status in sunflower leaves. Biosyst. Eng. 155, 124–133.

Ren, W., Mura, S., Irudayaraj, J.M.K.V., 2015. Modified graphene oxide sensors for ultra-sensitive detection of nitrate ions in water. Talanta 143, 234–239.

Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 8, 1627–1639.

Shao, Y.Q., Du, C.W., Shen, Y.Z., Ma, F., Zhou, J.M., 2017a. Evaluation of net nitrification rates in paddy soil using mid-infrared attenuated total reflectance spectroscopy. Anal. Methods 9, 748–755.

Shao, Y.Q., Du, C.W., Zhou, J.M., Ma, F., Zhu, Y., Yang, K., Tian, C., 2017b. Quantitative analysis of different nitrogen isotope labelled nitrate in paddy soil using mid-infrared attenuated total reflectance spectroscopy. Anal. Methods 9, 5388–5394.

Shaviv, A., Kenny, A., Shmulevitch, I., Singher, L., Raichlin, Y., Katzir, A., 2003. Direct monitoring of soil and water nitrate by FTIR based FEWS or membrane systems. Environ. Sci. Technol. 37, 2807–2812.

Singh, P., Singh, M.K., Beg, Y.R., Nishad, G.R., 2019. A review on spectroscopic methods for determination of nitrite and nitrate in environmental samples. Talanta 191, 364–381.

State Environmental Protection Administration, 2002. Monitoring and Analysis Methods of Water and Wastewater. China Environmental Press, Beijing, China, pp. 234–238.

Wang, X.J., Dzyadevych, S.V., Chovelon, J.-M., Renault, N.J., Chen, L., Xia, S.Q., Zhao, J.F., 2006. Conductometric nitrate biosensor based on methyl viologen/Nafion®/nitrate reductase interdigitated electrodes. Talanta 69, 450–455.

Wang, D.Y., Kong, X., Dong, L.Q., Chen, L.H., WangYJ, Wang, X.C., 2019. A predictive deconvolution method fornon-white-noise reflectivity. Appl. Geophys. 16, 101–115.
Wu, K., Du, C.W., Ma, F., Shen, Y.Z., Liang, D., Zhou, J.M., 2019. Rapid diagnosis of nitrogen status in rice based on Fourier transform infrared photoacoustic spectroscopy (FTIR-PAS). Plant Methods 15 (1), 94.
Wu, K., Ma, F., Li, Z.W., Wei, C.L., Gan, F.Q., Du, C.W., 2022. In-situ rapid monitoring of nitrate in urban water bodies using Fourier transform infrared attenuated total reflectance spectroscopy (FTIR-ATR) coupled with deconvolution algorithm. J. Environ. Manag. 317, 115452.
Yang, J.B., Du, C.W., Shen, Y.Z., Zhou, J.M., 2013. Rapid determination of nitrate in Chinese cabbage using fourier transforms mid-infrared spectroscopy. Chin. J. Anal. Chem. 41 (8), 1264–1268.
Zou, M.Y., Unbehauen, R., 1995. A deconvolution method for spectroscopy. Meas. Sci. Technol. 6, 482–487.