Front. Environ. Sci. Eng. 2019, 13(5): 75
https://doi.org/10.1007/s11783-019-1159-7

RESEARCH ARTICLE

Optimization and modeling of coagulation-flocculation to remove algae and organic matter from surface water by response surface methodology

Ziming Zhao[1,2], Wenjun Sun (✉)[1,3], Madhumita B. Ray[2], Ajay K Ray[2], Tianyin Huang[4], Jiabin Chen[4]

1 School of Environment, Tsinghua University, Beijing 100084, China
2 Department of Chemical and Biochemical Engineering, Western University, London Ontario N6A 5B9, Canada
3 Research Institute for Environmental Innovation (Suzhou), Tsinghua University, Suzhou 215163, China
4 School of Environmental Science and Engineering, Suzhou University of Science and Technology, Suzhou 215009, China

HIGHLIGHTS

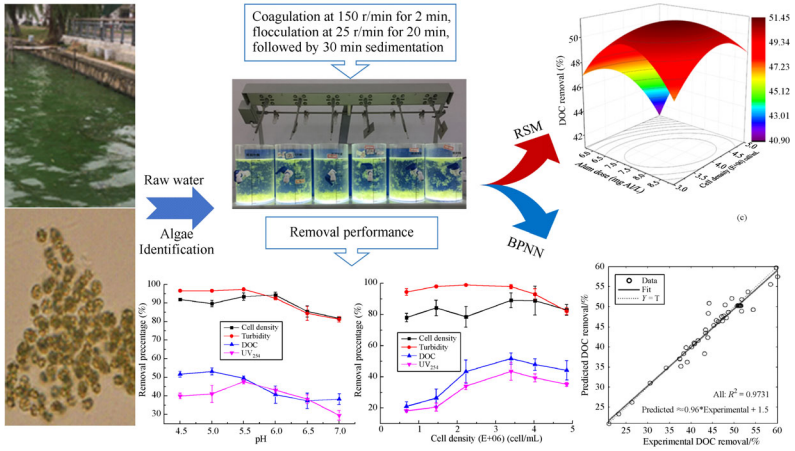
- Charge neutralization and sweep flocculation were the major mechanisms.
- Effect of process parameters was investigated.
- Optimal coagulation conditions were studied by response surface methodology.
- ANN models presented more robust and accurate prediction than RSM.

ARTICLE INFO

Article history:
Received 7 May 2019
Revised 22 August 2019
Accepted 23 August 2019
Available online 13 October 2019

Keywords:
Algae
Coagulation-flocculation
Response surface methodology
Artificial neural networks

GRAPHIC ABSTRACT



ABSTRACT

Seasonal algal blooms of Lake Yangcheng highlight the necessity to develop an effective and optimal water treatment process to enhance the removal of algae and dissolved organic matter (DOM). In the present study, the coagulation performance for the removal of algae, turbidity, dissolved organic carbon (DOC) and ultraviolet absorbance at 254 nm (UV_{254}) was investigated systematically by central composite design (CCD) using response surface methodology (RSM). The regression models were developed to illustrate the relationships between coagulation performance and experimental variables. Analysis of variance (ANOVA) was performed to test the significance of the response surface models. It can be concluded that the major mechanisms of coagulation to remove algae and DOM were charge neutralization and sweep flocculation at a pH range of 4.66–6.34. The optimal coagulation conditions with coagulant dosage of 7.57 mg Al/L, pH of 5.42 and initial algal cell density of 3.83×10^6 cell/mL led to removal of 96.76%, 97.64%, 40.23% and 30.12% in term of cell density, turbidity, DOC and UV_{254} absorbance, respectively, which were in good agreement with the validation experimental results. A comparison between the modeling results derived through both ANOVA and artificial neural networks (ANN) based on experimental data showed a high correlation coefficient, which indicated that the models were significant and fitted well with experimental results. The results proposed a valuable reference for the treatment of algae-laden surface water in practical application by the optimal coagulation-flocculation process.

✉ Corresponding author
E-mail: wsun@tsinghua.edu.cn

1 Introduction

A persistent worldwide concern in drinking water treatment is the proliferation of algae and the resultant metabolites in source water. Rivers, lakes and reservoirs as the important freshwater reserves in China are facing increasing threat of eutrophication (Wang et al., 2017). According to the "China's Ecological Environment Statements Bulletin of 2017", about 30% of China's lakes and reservoirs suffer from mild to moderate level of eutrophication (Li, 2018). Lake Taihu is one of the largest freshwater lakes in eastern China affected by algal bloom periodically due to non-point nutrient run-off sources, while Lake Dianchi, a heavily polluted lake in Yunnan Province shows both algal and fungi pollution. Algae bloom not only affects the ecology and aesthetic value of the aquatic system (Lee et al., 2017), it also derives multiple problems and poses many challenges to drinking water production, such as increasing coagulant demand, clogging filters (Campinas and Rosa, 2010; Zhang et al., 2011), taste and odor issues (Westrick et al., 2010) and disinfection byproduct formation (Engelage et al., 2009; Shen et al., 2011; Gough et al., 2015) in drinking water treatment plants. Blooms involving toxin-producing algal species even can pose serious threats to human health (Oyama et al., 2015; Zamyadi et al., 2015). The increasingly water eutrophication caused by cyanobacteria outbreak, resulted in several serious threats to local residential, commercial, industrial and agricultural production. Many countries and the World Health Organization (WHO) have established a guideline 1–1.5 μg/L) for microcystin-LR (MC-LR), which is one of the most toxic cyanotoxin produced by *Microcystis aeruginosa* (Merel et al., 2013). In 2007, the large-scale cyanobacteria outbreak in Lake Taihu from May to June caused serious safety threat to drinking water and led to the disruption of water supplies to millions of people in Wuxi and its surrounding areas. Lake Yangcheng is the third largest freshwater lake on the Taihu Plain. It is also the main drinking water source for Suzhou and Kunshan urban areas. As the second drinking water source for the city of Wuxi, Lake Yangcheng has been affected by severe eutrophication (Wang et al., 2018), which has resulted in the cyanobacteria blooms. Therefore, it is of vital importance to enhance the removal of algae and dissolved organic matter in water treatment process. Among the conventional water treatment processes, coagulation-flocculation is one of the economical methods to deal with "algal bloom" caused by the outbreak of microalgae (Zheng et al., 2015). As the primary barrier for the algal removal in conventional drinking water treatment, several studies have focused on the investigations of coagulation with/without pre-treatment for algae cells and the metabolites removal from raw water. It has been found that the removal of algal cells is easier than the removal of dissolved algal organic matter (AOM) (Henderson et al., 2010; Ma et al., 2016). More than 98% of algal cells could be removed with aluminum chloride dosage of 13 mg/L when the initial cell density less than 1.0×10^6 cell/L (Shen et al., 2011). However, for the combined coagulation and peroxidation processes, a poor removal of *Microcystis aeruginosa* cells and larger amount of trihalomethane (THM) formation occurred due to the release of AOM after peroxidation (Lin et al., 2017). Considering the maximum of algae removal and avoiding the lysis of algal cells to release AOM, it is necessary to enhance the coagulation conditions for maximum removal of algal cells and AOM without causing cell lysis. The success of this process implementation depends on how precisely pH and coagulant dosage are chosen with respect to the specific initial water quality.

Response surface methodology (RSM), as a combination of mathematical and statistical methods, has been widely applied for solving multivariable problems to optimize the process parameters with a less number of experimental runs and analyzing the interaction between the parameters. The objectives of RSM are: 1) to develop approximating functions for predicting responses, and 2) to optimize the responses based on the factors of interests (Javadi et al., 2014). The advantages of RSM method include low number of tests, high precision of regression equations, and continuous analysis of various levels of test factors. It has been widely applied in engineering fields such as biology, medicine and environment (Wang et al., 2007; Trinh and Kang, 2011; Halder et al., 2015; Li et al., 2015; Kim, 2016). The most commonly used RSM method is the central composite design (CCD), which includes center points, factorial points and axial points. From the CCD design, a quadratic approximation can be employed to develop a second-order response surface model for predicting the optimal point for a certain set of variables as follows (Eq. (1)):

$$\hat{Y} = \beta_0 + \Sigma_{i=1}^{3}\beta_i x_i + \Sigma_{i=1}^{3}\beta_{ii} x_i^2 + \Sigma_{i<j}^{3}\beta_{ij} x_i x_j + \varepsilon, \quad (1)$$

where $\hat{Y}$ is the predicted response; β_0, β_i, β_{ii} and β_{ij} are the coefficients for the intercept, linear, square, and interaction term of regression, respectively, which can be derived from ordinary least squares (OLS) or multiple linear regression (MLR) method, x_i and x_j represent the coded values of independent variables, ε indicates the statistical error.

Artificial Neural Networks (ANN) are computing systems with learning algorithms and architectures inspired by the working and structure of the human brain. Although there is a considerable amount of researches on various scenarios using both RSM and ANN techniques in the literature (Khayet et al., 2011; Bingöl et al., 2012), only a few studies on the coagulation-flocculation process were presented with the methods of both RSM and ANN techniques. Gadekar developed an

artificial neural networks to predict color removal using aluminum-based coagulant to remove color from a disperse dye solution, the performance of the model was found with correlation coefficient (R^2) values greater than 0.90 (Gadekar and Ahammed, 2016). To minimize settled water turbidity, It was reported that artificial neural networks (ANNs) can been applied to predict both the optimum carbon dioxide and coagulation dosages with correlation (R^2) values of 0.68 and 0.90, respectively. (McArthur and Andrews, 2015).

Hence, the key motivation behind this study was to develop an approach to evaluate and predict coagulation process efficiency for the removal of turbidity, cell density, DOC and UV_{254} absorbance of algae and organic matter using both RSM and ANN techniques. A two-level, three-factors central composite design (CCD) was applied to investigate the correlation between experimental variables and responses as the removals of microalgae, turbidity and dissolved organic carbon (DOC) in a real surface water body to provide solutions for the treatment of algae and algal matter-rich raw water.

2 Materials and methodologies

2.1 Study site and sample collection

Lake Yangcheng (31°25′N, 120°48′E), located between Lake Tai and the Yangtze River, has a surface area of about 20 square kilometers and mean depth of 1.9 m with an annual average temperature of 16°C–18°C. As the third largest freshwater lake on the Taihu Plain, Lake Yangcheng is the major drinking water source in Suzhou and Kunshan urban areas, and the second drinking water source of Wuxi City. The water samples from the Lake Yangcheng were collected twice per month in a 25 L plastic container from July 15 to September 10, 2017. All samples were preserved in the fridge before use within two weeks. The characteristics of raw water during the test period are shown in Table 1.

Table 1 Water quality characteristics of Lake Yangcheng

Parameters	Range	Mean
pH	7.08–8.45	7.48
Temperature (°C)	25–28	27
Cell density (10^6 cell/mL)	4.2–5.8	4.6
Turbidity (NTU)	198–252	223
DOC (mg/L)	10.2–13.5	12.41
UV_{254} absorbance (m^{-1})	0.083–0.094	0.089

Quantitative characterization of algae species in water was carried out using alga counter (Algae C model from Wansheng Ltd., China) and an automatic identification software.

2.2 Coagulation-flocculation

Coagulant aluminum sulfate hydrate ($Al_2(SO_4)_3 \cdot 18H_2O$), sodium hydroxide and hydrochloric acid for pH adjustment, were all of analytical grade and commercially available from Shanghai Lujie Chemical Reagent Co., Ltd., China. Coagulation tests were conducted using model ZR4-6 joint coagulation experiment mixer (Shenzhen Zhongshui Co., Ltd, China). Hemocytometer (Dark Line (0650010), Paul Marienfeld GmbH & Co., Germany) and microscope (CX-23, Olympus Co. Japan) were used for counting the algal cell before and after coagulation. The turbidity was measured using a Turbidity meter, (Hach 2100Q, Hach Company, USA). The DOC and UV_{254} absorbance of water samples were measured using a Shimadzu TOC-L analyzer (CPH TOC , Shimadzu Scientific Instruments Ltd., Japan) and the UV-Vis spectrophotometer (Model V-1200, Shanghai Meipuda instrument Ltd., China), respectively.

Coagulation-flocculation experiments of 2 L algae-laden water were performed at room temperature, various pH, alum doses (mg Al/L) and initial cell densities. The pH of the test solution was adjusted by adding pre-determined amount of 0.1 mol/L hydrochloric acid or 0.1 mol/L sodium hydroxide solution prior to the coagulation. The algal suspension was mixed at the agitation speed (150 r/min) for 2 min followed by a low mix/flocculation of 25 r/min for 20 min, and finally a 30 min settling. The supernatant was taken from 2 cm below the water surface for analysis of remaining cell density and turbidity. The DOC and UV_{254} absorbance were measured after filtering the supernatant through 0.45 μm membrane filter (Tianjin JINTENG Co., Ltd., China). The error analysis of duplicate experiments indicated an error $\leqslant 5\%$.

2.3 Response surface methodology with central composite design

Preliminary experiments indicated that three major variables affected coagulation-flocculation performance, i.e. coagulant dosage, pH and initial algal cell density. Preliminary experiments with single factor investigations narrowed the range of variables prior to experimental design. Based on those results, a complete set of three-factor central composite design (CCD) shown in Table 2 was applied to investigate the effects of individual variables and their interactions on the removal of algal cell, turbidity, DOC and UV_{254} absorbance to determine the response pattern and optimum combination of variables. Fourteen experiments were augmented with six replications at the center values (zero level) to evaluate the experimental error. The significance of each variable's effect on responses can only be compared with coded pattern because of their different units and limits of variation. For statistical calculations, the variable X_i was coded as x_i according to the following equation (Eq. (2)):

Table 2 Analytical factors and levels for RSM experimental design

Independent variables		Coded and actual levels				
		$-\alpha/-1.682$	-1	0	1	$+\alpha/1.682$
X_1	Alum dose (mg Al/L)	4.57	5.67	7.29	8.91	10.02
X_2	pH	4.66	5.00	5.50	6.00	6.34
X_3	Initial cell density (10^6 cell/mL)	2.32	3.00	4.00	5.00	5.68

$$x_i = (x_i - x_0)/\delta X, \quad (2)$$

where X_i is the uncoded value of the i^{th} independent variable, X_0 is the value of i^{th} variable at the center point of the experimental range and δX is the step change (Wang et al., 2014).

Analysis of variance (ANOVA) was applied for data analyses to obtain the interactions between the variables and the responses. The fit quality of polynomial regression models were demonstrated by the coefficient of determination R^2, and F-test and p-value (probability) evaluation were applied to check statistical significance with 95% confidence level.

2.4 Artificial neural network model

A feed-forward backpropagation neural network algorithm (BPNN) with three layers was developed by neural network tool box of MATLAB software version 9.2.0 (The Mathworks Inc., USA). Mathematically, the structure of a 3-layer ANN with n, m, and p the number of input, hidden and output nodes respectively, is shown as Fig. 1:

Where Y_k are the output values (responses) and X_i are the input values (variables) of the network; W_{ij} are the connection weights between the input layer and the hidden layer; W_{jk} are the connection weights between the hidden layer and the output layer; S is a transfer function. At each node, the weighted input signals are summed with a bias value (W_j). The combined input (H_i) then passes through the transfer function (S) to produce the output node (Y_k) as demonstrated in Fig. 1 (Al-Abri et al., 2010). The Levenberg- Marquardt back propagation algorithm was used for ANN model training. The proposed neural networks had two transfer functions, of which the first transfer function was tansig and the second one was linear transfer function (purelin) (Gadekar and Ahammed, 2016).

A total of 44 data points, including the results from CCD experiments and single variable (alum dose, pH and initial cell density) and validation investigations, were used in ANN modelling. These data points were split randomly into training (70%), validation (15%), and test (15%) subsets. All experimental variables data were normalized in the limits from -1 to $+1$ using the following equation (Eq. (3)) (Zhao et al., 2010; Piuleac et al., 2013):

$$\text{Normalized data} = \frac{2X_{\text{AC}} - (X_{\min} + X_{\max})}{X_{\max} - X_{\min}}, \quad (3)$$

To match the tangent sigmoid function applied in ANN modeling, where X_{AC}, $X_{\min}$ and $X_{\max}$ are the actual, minimum, and maximum data, respectively. A minimum mean squared error (MSE) is shown as the following Eq. (4), where Y_i and $\hat{Y}_i$ are the i^{th} experimental and predicted values were computed. The ANN model and the

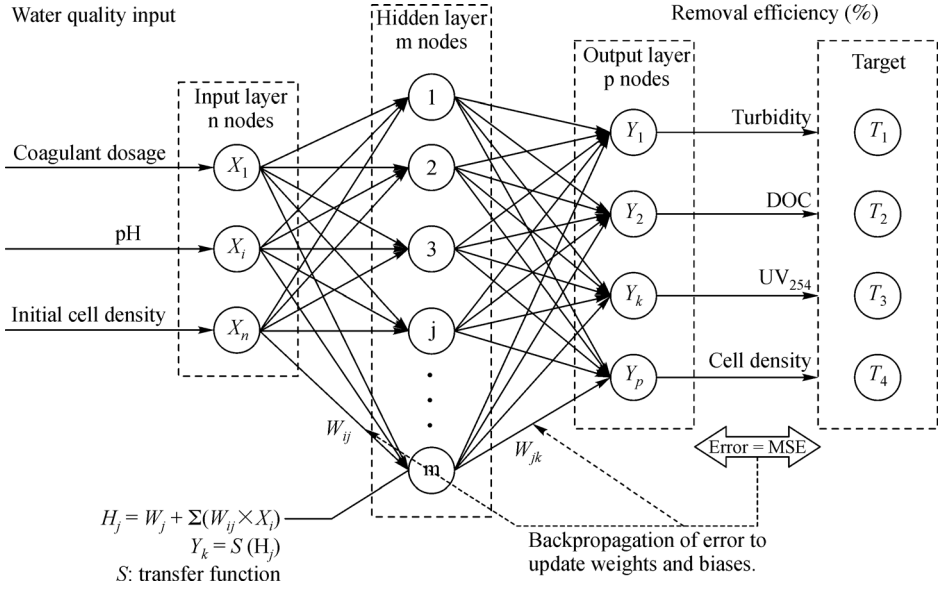


Fig. 1 Architecture of the three layers backpropagation artificial neural network (BPNN).

variation of parameters were evaluated based on the minimum value of the MSE of the training and prediction set.

$$\mathrm{MSE} = \frac{1}{n}\Sigma_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2, \quad (4)$$

The training parameters were used with three input nodes, 8 to 10 hidden neurons and one output node with respect to one response each time, learning rule: Levenberg–Marquardt, number of epochs: 1000, error goal: 0.0001 in this study.

3 Results and discussion

3.1 Algae species distribution

As shown in Fig. 2, more than 98% of the microalgae in the investigated water body were cyanobacteria (mainly *microcystis*), only 0.1% and 0.34% of algae belonged to *oscillatoria* in the two samples, while the concentration of *Nitzschia palea*, a diatom, was 1.55% and 1.67%, respectively in the two samples. The amount of *protosiphon*, a *chlorophyta* detected in sample 2 shown in Fig. 2(b), was only 0.03%. Considering the average cell density of 4.6×10^6 cell/mL, water sample was seriously contaminated by cyanobacterial bloom, which may be due to the surrounding municipal and industrial wastewater discharge containing high total phosphorus and total nitrogen into water under mild hydrological and weather condition (Ma et al., 2013). Therefore, as shown in Fig. 2, *microcystis* dominated the phytoplankton community and produced enormous biomass. The average specific UV absorbance (SUVA) of 0.72 L/(m·mg C) indicated that the dissolved organic matter in algae-laden water was predominately hydrophilic, with low SUVA value (0.34–1.7 L/(m·mg C)) (Goslan et al., 2017).

3.2 Effect of alum dose on the coagulation performance

Aluminum sulfate (alum) is one of the most commonly applied coagulants in drinking water plants, due to its cost-effectivity and widespread availability (Carty et al. 2002). The dosage of coagulant is the most vital parameter for algae and dissolved organic matter removal. The effects of alum dosed from 3.24–8.10 mg Al/L on the coagulation performance under initial cell density of $(4.85\pm0.26) \times 10^6$ cell/mL were presented in Fig. 3.

The coagulation performance increased with the increasing of coagulant dose for the removal of turbidity, DOC and UV_{254} absorbance, which was consistent with previous results of increased DOM removal with the increasing alum dose to a certain point (Lanciné et al., 2008). However, higher alum dosage will contribute to relatively high aluminum residuals causing possible health hazard, although this can be remediated, even be avoided by pH control in the finished water (Matilainen et al., 2010). It can be noted that the cell removal efficiency reached a plateau at dosage$\geqslant$4.86 mg Al/L with the maximum algal cell removal of 81.56%±5.81%. With the increase of alum dose, the removal efficiency of turbidity, DOC and UV_{254} absorbance increased up to 97.02%± 1.38%, 56.25%±3.08% and 42.96%±0.08%, respectively, which can be explained by the higher charge neutralization ability with the increase of alum dose (Aktas et al., 2013). However, at a higher dose, charge reversal may occur and result in a reduction of the removal efficiency. Considering the potential health risk of high alum dosage, 7.3 mg Al/L was chosen as the appropriate alum dose for further experiments.

3.3 Effect of pH on the coagulation performance

The effect of pH on the coagulation performance was tested at different pH between 4.5 and 7.0 with the same initial algal cell density and coagulant dosage of 7.3 mg Al/L (Fig. 4). It can be noted that higher removal of all four responses occurred at lower pH of 4.5–6.0. The maximum algal cell removal of 94.22%±1.57% occurred at pH 6.0, however, for turbidity, DOC and UV_{254} absorbance removal, the maximum coagulation performance of 97.36%±0.21%, 53.20%±1.81%, and

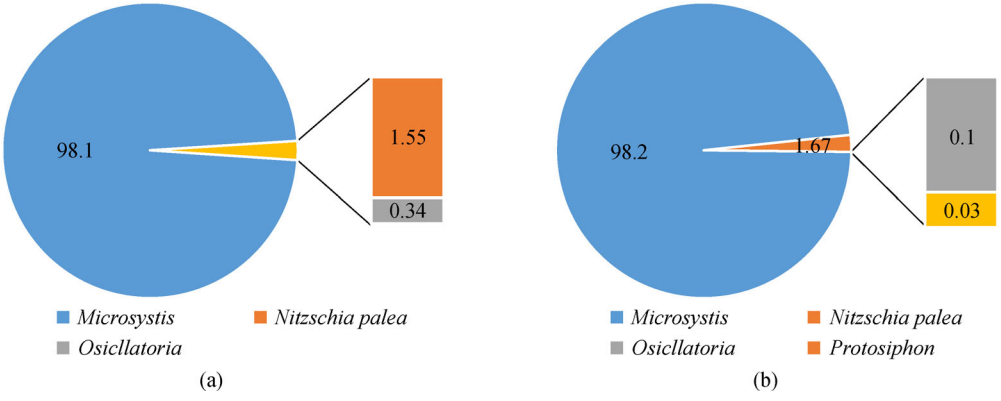


Fig. 2 Algal species distribution on the raw water.

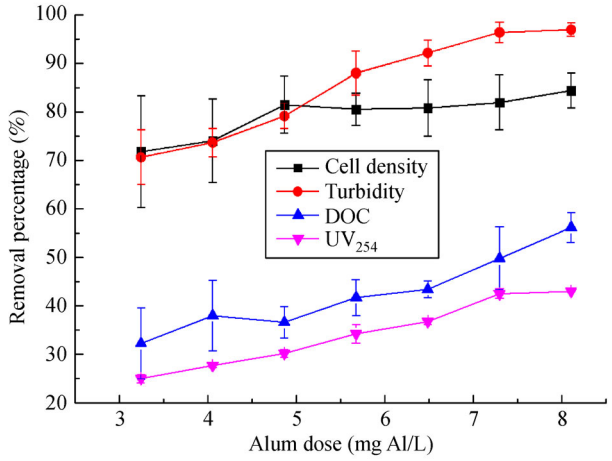


Fig. 3 Effect of alum dosage on coagulation performance for the water samples with cell density of 4.55×10^6 cell/mL without pH adjustment (error bars represent the standard deviation from duplicate experiments)

47.71%±1.20% occurred at pH 5.5, 5.0 and 5.5, respectively. At pH lower than 5.5, positive hydrolyzates, such as $Al(OH)^{2+}$, $Al_2(OH)_2{}^{4+}$ were formed by alum, which neutralize the exterior negative charges of cell and colloids to promote the floc growth by physical or chemical adsorption of destabilized cell and DOM colloids (Yang et al., 2010). At pH⩾6.0, $Al(OH)_4^-$ formed which was not beneficial for negative charge neutralization of the cells (Zhang et al., 2008).

It also indicated that the cationic H^+ to neutralize the surface charge of algal cell required was less than DOM in water, so that algal cell reached the maximum removal efficiency at a relatively higher pH of 6 compared to DOM (represented by DOC). The pH value of 5.5 was chosen as the most suitable pH for the removal of cell density, turbidity, DOC, and UV_{254} absorbance.

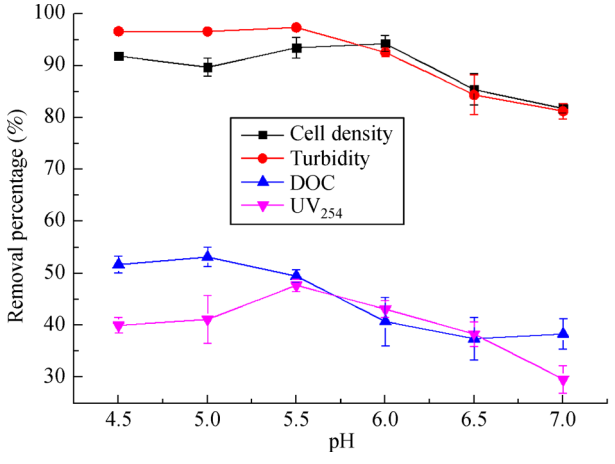


Fig. 4 Effect of pH on coagulation performance for the water samples with cell density of 4.5×10^6 cell/mL under the coagulation dosage 7.30 mg Al/L (error bars represent the standard deviation from duplicate experiments).

3.4 Effect of initial cell density on the coagulation performance

It was noticed that the removal performance at various initial cell densities and the constant coagulant dosage resulted in different removal efficiencies. Thus, the relationship between initial cell density and required coagulant dosages on removal efficiency was further investigated at different cell density with fixed coagulation dosage and initial pH. It was indicated that the four responses of coagulation performance increased initially then decreased with the increase of cell density as shown in Fig. 5.

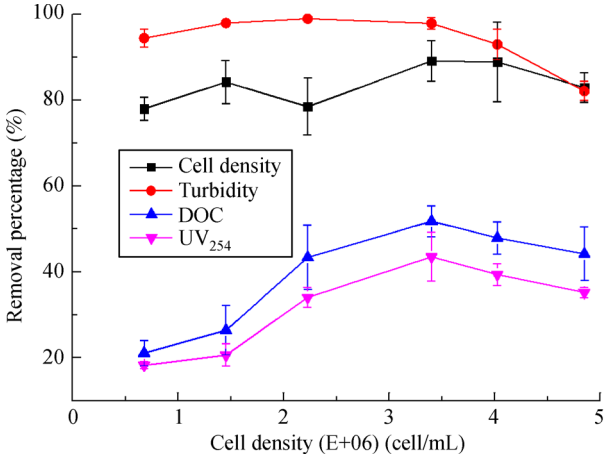


Fig. 5 Effect of initial cell density on coagulation performance under the coagulation dosage 7.30 mg Al/L and pH of 5.5 (error bars represent the standard deviation from duplicate experiments).

For the water sample with low cell density (less than 2×10^6 cell/L) and concentration of DOM, the dosage of 7.3 mg Al/L coagulant was considered an overdose, as the re-stabilization of cell and organic matter occurred and resulted in lower removal efficiency of cell density, turbidity, DOC and UV_{254} absorbance. Once the cell density increased further with the increase of the concentration of DOM in water, the dosage of 7.3 mg Al/L coagulant demonstrated the maximum removal efficiency of the cell density of 3.3×10^6 cell/mL. The removal percentage decreased with the increase of cell density due to relatively insufficient coagulant dosage.

3.5 Response surface model and analysis of variance (ANOVA)

Due to the aforementioned factors, a central composite design (CCD) experiments for optimization of parameters, such as alum dose, coagulation pH, and initial cell density were performed to locate the maximum removal efficiency of cell density, turbidity, DOC and UV_{254} absorbance by Design Expert 7.0 (trial version) from the experimental data shown in Table 3.

Table 3 CCD experimental design and experimental results

Run	Experimental variables			Removal percentage (%)			
	Alum dosage/X_1 (mg Al/L)	pH/X_2	Initial cell density/ X_3(E + 06 cell/mL)	Cell density	Turbidity	DOC	UV_{254}
1	5.67	5	3	93.90	81.25	48.63	28.29
2	8.91	5	3	92.84	92.12	46.35	20.61
3	5.67	6	3	89.50	78.69	41.64	19.24
4	8.91	6	3	91.54	91.26	40.86	28.85
5	5.67	5	5	86.05	85.05	38.88	27.18
6	8.91	5	5	96.86	91.47	44.02	15.01
7	5.67	6	5	83.24	87.06	40.15	24.36
8	8.91	6	5	94.67	92.81	47.12	20.11
9	4.57	5.5	4	91.69	76.71	37.67	25.85
10	10.02	5.5	4	92.17	89.37	47.63	23.99
11	7.29	4.66	4	93.73	97.42	45.88	28.09
12	7.29	6.34	4	90.12	87.68	45.29	25.43
13	7.29	5.5	2.32	93.13	97.25	48.19	26.01
14	7.29	5.5	5.68	92.74	93.69	43.40	23.96
15	7.29	5.5	4	96.92	95.94	50.75	30.43
16	7.29	5.5	4	97.53	95.93	51.55	29.95
17	7.29	5.5	4	97.35	96.21	51.24	30.13
18	7.29	5.5	4	97.78	96.02	51.40	30.42
19	7.29	5.5	4	96.98	96.00	50.96	30.61
20	7.29	5.5	4	97.06	95.94	51.04	30.04

Based on the experimental and ANOVA results, the quadratic regression equations were developed on the basis of CCD experimental sets and input variables, where X_1, X_2, X_3, are the alum dosage (mg/L), coagulation pH and initial cell density (10^6 cell/mL), respectively, only significant items were presented in the regression equation presented in the Eqs. (5)–(8).

$$\begin{aligned}\textbf{Cell removal}/\% &= -168.33 + 6.07X_1 + 89.05X_2 + 1.26X_3 \\ &\quad + 1.64X_1X_3 - 0.79{X_1}^2 - 8.32{X_2}^2 - 1.72{X_3}^2, \quad (5)\end{aligned}$$

$$\textbf{Turbidity removal}/\% = -217.01 + 30.65X_1 + 72.01X_2 - 1.92{X_1}^2 - 6.77{X_2}^2, \quad (6)$$

$$\begin{aligned}\textbf{DOC removal}/\% &= -176.01 + 14.15X_1 + 76.27X_2 \\ &\quad - 16.25X_3 + 1.17X_1X_3 + 4.21X_2X_3 \\ &\quad - 1.21{X_1}^2 - 8.59{X_2}^2 - 2.07{X_3}^2, \quad (7)\end{aligned}$$

$$\begin{aligned}\textbf{UV}_{\textbf{254}}\ \textbf{absorbance removal}/\% &= -137.84 - 3.35X_1 \\ &\quad + 47.32X_2 + 28.18X_3 + 3.89X_1X_2 - 1.42X_1X_3 \\ &\quad - 0.90{X_1}^2 - 6.92{X_2}^2 - 2.36{X_3}^2. \quad (8)\end{aligned}$$

To validate the response surface model from statistical standpoint, the significance of the regression model and the lack-of-fit need to be addressed (Khayet et al., 2011). Generally, *F*-value or *p*-value (also called the Prob>*F* value) are commonly used to evaluate the significance of the models. The larger *F*-value and correspondingly smaller *p*- value, indicate significance of the established regression model. A *p*-value less than 0.05 represents that the design model is statistically significant. The *p*- value for each regression model was less than 0.05 with the lowest values of 0.0006, which indicated that each of the regression model obtained above was significant. The precision of the model can be demonstrated by the determination of the coefficient (R^2) to quantify the strength of the correlation between the observed and predicted values and calculated as the following Eq. 9 (Xiao et al., 2017):

$$R = \frac{\Sigma_{i=1}^{n}\left(Y_i - \overline{Y}_l\right)\left(Y_l - \overline{\hat{Y}}_l\right)}{\sqrt{\Sigma_{i=1}^{n}\left(Y_i - \overline{Y}_l\right)^2}\sqrt{\Sigma_{i=1}^{n}\left(\hat{Y}_l - \overline{\hat{Y}}_l\right)^2}}, \quad (9)$$

where i is the data number, Y_i is observed value, $\hat{Y}_l$ is predicted value, $\overline{Y}_l$ and $\overline{\hat{Y}}_l$ are the means of Y_i and $\hat{Y}_l$, respectively.

The R^2 values for turbidity, DOC, UV_{254} absorbance, and cell removal efficiencies were determined as 0.9089, 0.9521, 0.8931 and 0.8895, respectively. The values of the coefficient of determination ($R^2 \geqslant 0.89$) indicated that more than 89% of the variability in the responses could be explained by the models. The obtained adequate precision (AP) of the models compares the range of the predicted values at the design points with the average prediction error, which indicates the signal-to-noise ratio and a ratio greater than 4 is desirable (Olmez-Hanci et al., 2011). In the present study, the obtained values with the minimum of 10.99 as shown in Table 4 indicated an adequate signal and suggested that the models can explain the relationship of variables and responses successfully.

Table 4 ANOVA results for regression models

ANOVA	Response			
	Turbidity	DOC	UV_{254}	Cell density
R^2	0.9089	0.9521	0.8931	0.8895
p-value	0.0006	< 0.0001	0.0002	< 0.0001
Std. dev.	2.73	1.35	1.90	1.65
Mean	90.89	46.13	25.93	93.29
C.V.%	3.00	2.81	7.31	1.77
PRESS	564.49	104.23	216.66	157.11
AP	11.13	15.82	10.99	11.58

The coefficient of variation (C.V.%) represents the ratio of the standard deviation to the average response value in the model. The smaller the value is, the smaller the dispersion in data. In this study, the maximum C.V.% value of 7.31% was less than 15% removal efficiency of UV_{254} absorbance, which indicated that the reliability of the data was very high, and the experiment had high reproducibility. These findings revealed that the accuracy and ability of the polynomial models obtained for observed responses were appropriate and satisfactory.

Response surfaces for removal efficiency of the coagulation process for algae-laden lake water were created by Design-Expert 8.0 as shown in Fig. 6. Based on ANOVA results, there was no interaction effect of the variables for turbidity removal. It was indicated that alum dose and pH had the dominant effects on turbidity removal, initial cell density had insignificant effects even though turbidity increased with the increase of initial cell density in the experimental range. Figures 6(a) and 6(b) showed the response surface and contour plots for cell density and turbidity removal efficiency as a function of alum dose and pH at an initial cell density of 4.0×10^6 cell/mL. The highest removal efficiency (97.78%, 97.42% for cell density and turbidity, respectively) occurred at the alum dosage of 7.29 mg Al/L and pH of 5.5. The lowest removal occurred at the higher pH of 6.0 and a low coagulant dose of 5.67 mg Al/L. It was found that the cell density and turbidity removal efficiency presented the same pattern, which decreased with increasing pH up to 6.0 at the low coagulant dose of 5.67 mg Al/L.

The interaction surface of DOC removal percentage (Fig. 6(c)) showed a saddle shape, axial steepness and surface curvature increase, which indicated that interaction effect of coagulant dose and pH had a significant response to DOC removal efficiency. The response surface of UV_{254} absorbance removal efficiency was shown in Fig. 6(d); the UV_{254} absorbance removal efficiency decreased significantly with the increase of initial cell density even at a high coagulation dose of 8.91 mg Al/L and pH of 5.5, which indicated that the alum applied could not remove aromatic compounds of water efficiently, and higher cell density competed with aromatics of water for coagulant dosage.

Using the optimization module by Design-Expert software, the optimum parameters of coagulation process were obtained for the removal of cell density, turbidity, DOC and UV_{254} absorbance. With these multiple responses, the overlaid contour plot (Fig. 7) was used to visually demonstrate the optimal conditions range which the required responses can be simultaneously reached.

The optimum parameters of coagulation process were obtained as follows: the dosage of coagulation 7.57 mg Al/L, pH of 5.42 and the initial algae concentration 3.83×10^6 cell/mL. The predicated removal percentage for cell density, turbidity, DOC and UV_{254} absorbance was 97.31%±1.65%, 95.50%±2.68%, 51.18%±1.30% and 30.34%±2.68%, respectively.

The validation test were conducted under the optimized conditions with the coagulant dosage of 7.5 mg Al/L, pH of 5.5 and the initial algae concentration of 4×10^6 cell/mL. The actual removal performances for cell density, turbidity, DOC, and UV_{254} absorbance were 97.27%, 95.43%, 48.65%, and 28.34%, respectively. Although the equivalent alum dosage of 1.97×10^{-9} mg Al/cell is less than that of 4.3×10^{-9} mg Al/cell presented by Gonzalez-Torres et al. (Gonzalez-Torres et al., 2014) who used a higher pH of 7.0 in their study, the optimized condition in the current study for algae-laden water treatment obtained a relatively higher alum dose than usual 2.5–4.0 mg Al/L for drinking water treatment (Trinh and Kang, 2011). This indicated that polymer coagulant or coagulation aid may be needed to reduce the alum dosage or a combination with other treatment, for instance, air flotation may be used.

3.6 Artificial neural network

The ANN model has been applied extensively to predict

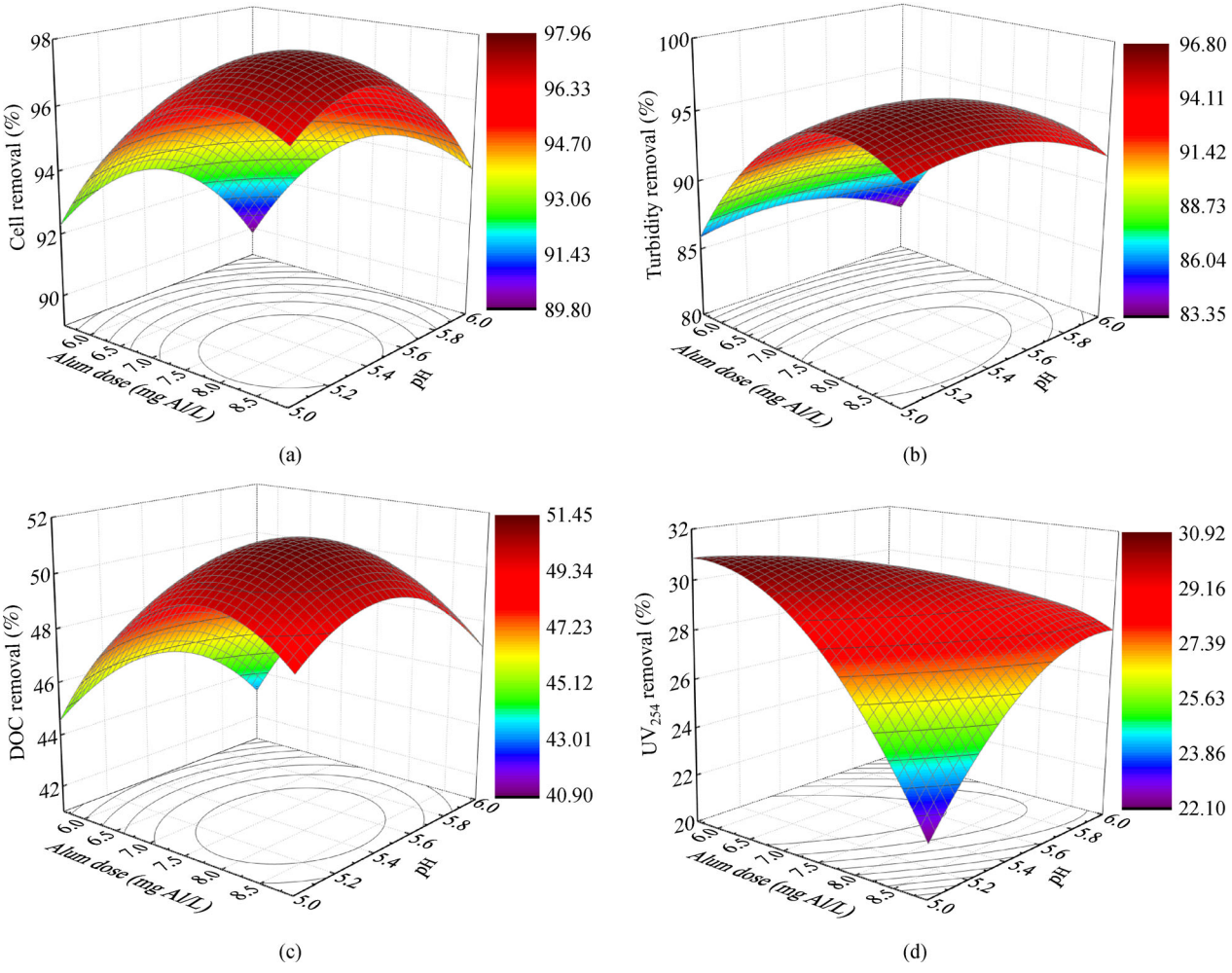


Fig. 6 Surface plots of removal efficiency with the interaction of coagulant dosage and pH with initial cell density of 4.0×10^6 cell/mL, (a) Cell removal; (b) Turbidity removal; (c) DOC removal; (d) UV_{254} removal.

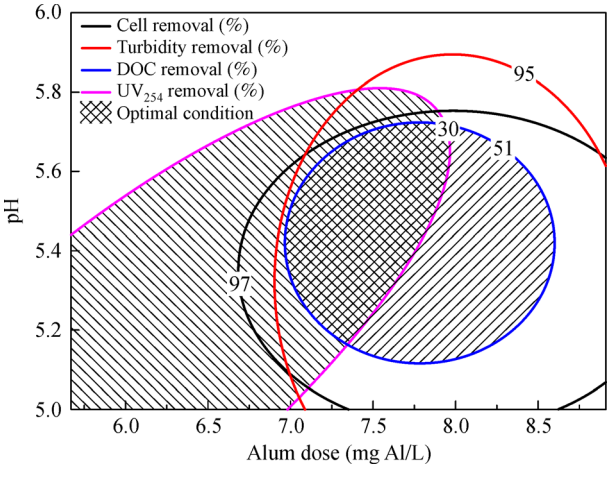


Fig. 7 Overlaid contour plot for cell density, turbidity, DOC and UV_{254} removal percentage by alum coagulation. Data fitted by three-factor central composite design.

nonlinear systems due to the cost and time effectivity and high precision (Moghaddari et al., 2018). Table 5 demonstrated the ANN modelling performance in term of correlation coefficient (R^2) and standard deviation for individual and overall dependent responses with respect to different data section under its specific ANN topological structure.

The topology selected was based on the performance of networks, which gave minimum MSE and R^2 close to one. The training R^2 in all cases of models presented the highest value because the majority of the dataset (70%) were used for training repeatedly several times for adjusting the weights of the network. ANN-predicted values of removal efficiencies for cell density, turbidity, DOC and UV_{254} absorbance versus experimental data were presented in Fig. 8. The linear regression analysis between ANN-predicted and observed individual removal efficiency showed the minimum linear regression coefficient (R^2) of 0.8855 for cell removal. The overall R^2 of the models is

Table 5 Performance of ANN network models

Dependent responses	Topology	Correlation coefficient (R^2)				Std. dev.
		Training	Validation	Testing	All	
Cell density	3: 8: 1	0.9069	0.9188	0.8652	0.8855	1.3106
Turbidity	3: 10: 1	0.9741	0.9579	0.9647	0.9713	1.6387
DOC	3: 10: 1	0.9785	0.9010	0.9938	0.9731	1.2188
UV_{254}	3: 10: 1	0.9469	0.9431	0.8133	0.8980	1.3675
Total	3: 10: 4	0.9903	0.9705	0.9735	0.9814	1.7859

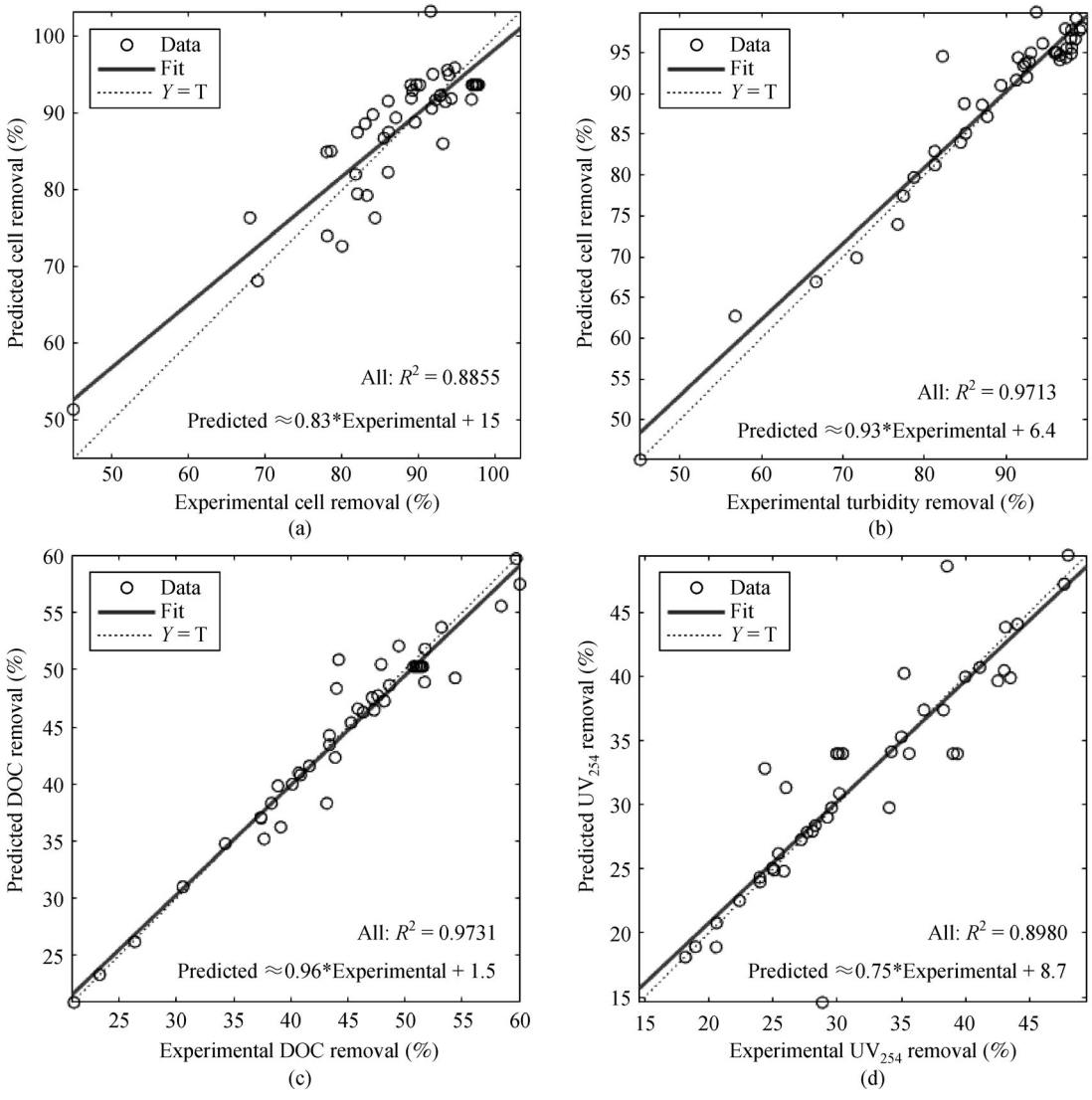


Fig. 8 The plots of predicted vs. actual values of removal efficiency by BPNN: (a) cell density; (b) Turbidity; (c) DOC; (d) UV_{254}.

larger than 0.8 representing that the developed models are robust (Kundu et al., 2013).

Linear regression analysis was carried out between the variables (coagulation dose, pH and initial cell density) and removal performance (cell density, Turbidity, DOC and UV_{254} absorbance) values calculated by ANN and RSM models with their corresponding observed values. The largest standard deviation of these four responses from RSM and ANN was 2.73 and 1.64, respectively, which indicated that RSM model prediction presented a greater deviation than ANN predication. Both models presented stable responses, but the ANN models were better in data fitting and estimation capabilities. In comparison with RSM, ANN was easier to obtain the relatively higher

average regression coefficient of 0.93 than 0.91 from RSM. The modeling results indicated that ANN was slightly more accurate for estimating the values of dependent variables as compared with the RSM models. However, the RSM can be applied to analyze the factor effects (main and interactional) and propose regression equations for responses. Also, RSM can identify the significant main and interaction factors or insignificant terms in the model and thereby can reduce the complexity of the problem with assumption of quadratic non-linear correlation. However, ANN can easily overcome the limitations of RSM, inherently capture almost any form of non-linearity without the requirement of a standard experimental design to build the model (Bingöl et al., 2012).

3.7 Mechanism of algae and DOM removal by alum

It is well-known that four mechanisms of charge neutralization, adsorption, bridging, and sweep flocculation might be involved in coagulation process of natural colloids (Guo et al., 2017). Alum undergoes hydrolysis to form variable mononuclear and polynuclear species depending on pH as shown by Eq. (10).

$$\mathrm{Al(H_2O)_6^{3+}} \leftrightarrow \mathrm{Al(OH)(H_2O)_5^{2+}} \rightarrow$$

$$\mathrm{Al_{13}O_4(OH)_{24}^{7+}} \rightarrow \mathrm{Al(OH)_3(S)} \leftrightarrow \mathrm{Al(OH)_4^-} \qquad (10)$$

Based on the applied dosage of 4.57–10.02 mg Al/L, concentration of $Al_2(SO_4)_3 \cdot 14.3H_2O$ was about 50.7–111.2 mg/L in the experiments. From the results presented in Table 3, it can be seen that the maximum cell removal of 97.78% occurred at pH 5.5. According to the coagulation domain diagram for alum dosage at various pH presented by Amirtharajah and Mills (1982), the region corresponds to the dosage of 50.7–111.2 mg/L at pH 5.5 is in the sweep coagulation zone. In the pH range of 6–8, algal surfaces are negatively charged (Gonçalves et al., 2015). On the other hand, for alum coagulant, the dominant species of aluminum possibly are $\mathrm{Al_{13}O_4(OH)_{24}^{7+}}$ and $\mathrm{Al(OH)_3(S)}$ at this pH range. The optimal coagulation for algae-laden natural water occurred around the pH 5–6, for which both algal cell and DOM are negatively charged, so that electrostatic interaction occurred between cationic aluminum species and cell/DOM. Therefore, both charge neutralization and sweep flocculation are possible mechanisms for the removal of algae and organic matter in this investigation. Guo et al. (2017) also mentioned charge neutralization and sweep flocculation to be the dominant mechanisms for DOM removal. The reduced electrostatic repulsion between DOM colloidal particles/cells may facilitate initial aggregation of colloidal and fine suspended particulate to form microflocs (Agbovi and Wilson, 2017), in addition, the attached polyanions of DOM onto negative cell surface may also favor the agglomeration formation (Baresova et al., 2017).

4 Conclusions

In the present study, the coagulation performance was investigated and optimized for the removal of algal cells and DOM from the eutrophic water sample of Lake Yangcheng. Based on the response surface analysis designed by CCD, the regression models for the coagulation performance were developed. A dosage of 7.57 mg Al/L and pH 5.42 were determined as optimal condition of coagulation for initial algal concentration of 3.83×10^6 cell/mL and an average initial DOC of 12.41 mg/L. Charge neutralization and sweep coagulation were the dominate mechanisms for the treatment of algae-laden natural water.

The variance analysis of regression models and verification tests showed that the regression models were effective in fitting the experimental data. The ANN model were relatively more accurate to estimate the values of the dependent variables. The models developed in this study may provide useful treatment suggestions for water plants to treat surface water affected by algal blooms.

Acknowledgements This work was financially supported by the Ontario-China Research and Innovation Fund (OCRIF) "Ensuring Water Supply Safety in Beijing: Water Diversion from South to North China" (No. 2015DFG71210). This work was also supported by grants from the Mega-projects of Science Research for Water Environment Improvement (Nos. 2012ZX07404-002, 2017ZX07108-003, 2017ZX07502003).

References

Agbovi H K, Wilson L D (2017). Flocculation optimization of orthophosphate with $FeCl_3$ and alginate using the box-behnken response surface methodology. Industrial & Engineering Chemistry Research, 56(12): 3145–3155

Aktas T S, Takeda F, Maruo C, Fujibayashi M, Nishimura O (2013). Comparison of four kinds of coagulants for the removal of picophytoplankton. Desalination and Water Treatment, 51(16–18): 3547–3557

Al-Abri M, Al Anezi K, Dakheel A, Hilal N (2010). Humic substance coagulation: Artificial neural network simulation. Desalination, 253(1–3): 153–157

Amirtharajah A, Mills K M (1982). Rapid-mix design for mechanisms of alum coagulation. Journal- American Water Works Association, 74(4): 210–216

Baresova M, Pivokonsky M, Novotna K, Naceradska J, Branyik T (2017). An application of cellular organic matter to coagulation of cyanobacterial cells (*Merismopedia tenuissima*). Water Research, 122: 70–77

Bingöl DHercan MElevli S, Kiliç E (2012). Comparison of the results of response surface methodology and artificial neural network for the biosorption of lead using black cumin. Bioresource Technology,

112: 111–115

Carty G, O'Leary G, Crowe M (2002). Water Treatment Manuals: Coagulation, Flocculation and Clarification. Washington D. C.: Environmental Protection Agency, 85

Campinas M, Rosa M J O (2010). Evaluation of cyanobacterial cells removal and lysis by ultrafiltration. Separation and Purification Technology, 70(3): 345–353

Engelage S K, Stringfellow W T, Letain T (2009). Disinfection byproduct formation potentials of wetlands, agricultural drains, and rivers and the effect of biodegradation on trihalomethane precursors. Journal of Environmental Quality, 38(5): 1901–1908

Gadekar M R, Ahammed M M (2016). Coagulation/flocculation process for dye removal using water treatment residuals: Modelling through artificial neural networks. Desalination and Water Treatment, 57(55): 26392–26400

Gonçalves A L, Ferreira C, Loureiro J A, Pires J C, Simões M (2015). Surface physicochemical properties of selected single and mixed cultures of microalgae and cyanobacteria and their relationship with sedimentation kinetics. Bioresources and Bioprocessing, 2(1): 21–31

Gonzalez-Torres A, Putnam J, Jefferson B, Stuetz R M, Henderson R K (2014). Examination of the physical properties of Microcystis aeruginosa flocs produced on coagulation with metal salts. Water Research, 60: 197–209

Goslan E H, Seigle C, Purcell D, Henderson R, Parsons S A, Jefferson B, Judd S J (2017). Carbonaceous and nitrogenous disinfection by-product formation from algal organic matter. Chemosphere, 170: 1–9

Gough R, Holliman P J, Cooke G M, Freeman C (2015). Characterisation of algogenic organic matter during an algal bloom and its implications for trihalomethane formation. Sustainability of Water Quality and Ecology, 6: 11–19

Guo T T, Yang Y L, Liu R P, Li X (2017). Enhanced removal of intracellular organic matters (IOM) from Microcystic aeruginosa by aluminum coagulation. Separation and Purification Technology, 189: 279–287

Halder G, Dhawane S, Barai P K, Das A (2015). Optimizing chromium (VI) adsorption onto superheated steam activated granular carbon through response surface methodology and artificial neural network. Environmental Progress & Sustainable Energy, 34(3): 638–647

Henderson R K, Parsons S A, Jefferson B (2010). The impact of differing cell and algogenic organic matter (AOM) characteristics on the coagulation and flotation of algae. Water Research, 44(12): 3617–3624

Javadi N, Ashtiani F Z, Fouladitajar A, Zenooz A M (2014). Experimental studies and statistical analysis of membrane fouling behavior and performance in microfiltration of microalgae by a gas sparging assisted process. Bioresource Technology, 162: 350–357

Khayet M, Cojocaru C, Essalhi M (2011). Artificial neural network modeling and response surface methodology of desalination by reverse osmosis. Journal of Membrane Science, 368(1–2): 202–214

Kim S C (2016). Application of response surface method as an experimental design to optimize coagulation–flocculation process for pre-treating paper wastewater. Journal of Industrial and Engineering Chemistry, 38(Supplement C): 93–102

Kundu P, Debsarkar A, Mukherjee S (2013). Artificial neural network modeling for biological removal of organic carbon and nitrogen from slaughterhouse wastewater in a sequencing batch reactor. Advances in Artificial Neural Systems, 2013: Article ID 268064

Lanciné G D, Bamory K, Raymond L, Jean-Luc S, Christelle B, Jean B (2008). Coagulation-Flocculation treatment of a tropical surface water with alum for dissolved organic matter (DOM) removal: Influence of alum dose and pH adjustment. Journal of International Environmental Application and Science, 3(4): 247–257

Lee J, Rai P K, Jeon Y J, Kim K H, Kwon E E (2017). The role of algae and cyanobacteria in the production and release of odorants in water. Environmental Pollution, 227: 252–262

Li G (2018). China's Ecological Environment Statements Bulletin of 2017. Beijing: Ministry of Ecology and Environment of China (in Chinese)

Li L, Zhang S, He Q, Hu X B (2015). Application of response surface method in experimeatal design and optimization. Research and Exploration in Laboratory, 34(8): 41–45 (in Chinese)

Lin J L, Hua L C, Hung S K, Huang C (2017). Algal removal from cyanobacteria-rich waters by preoxidation-assisted coagulation–flotation: Effect of algogenic organic matter release on algal removal and trihalomethane formation. Journal of Environmental Sciences, 63:147–155

Ma C X, Hu W R, Pei H Y, Xu H Z, Pei R T (2016). Enhancing integrated removal of Microcystis aeruginosa and adsorption of microcystins using chitosan-aluminum chloride combined coagulants: Effect of chemical dosing orders and coagulation mechanisms. Colloids and Surfaces A—Physicochemical and Engineering Aspects, 490: 258–267

Ma J R, Deng J M, Qin B Q, Long S X (2013). Progress and prospects on cyanobacteria bloom-forming mechanism in lakes. Acta Ecologica Sinica, 33(10): 3020–3030

Matilainen A, Vepsäläinen M, Sillanpää M (2010). Natural organic matter removal by coagulation during drinking water treatment: a review. Advances in Colloid and Interface Science, 159(2): 189–197

McArthur R H, Andrews R C (2015). Development of artificial neural networks based confidence intervals and response surfaces for the optimization of coagulation performance. Water Science and Technology-Water Supply, 15(5): 1079–1087

Merel S, Walker D, Chicana R, Snyder S, Baurès E, Thomas O (2013). State of knowledge and concerns on cyanobacterial blooms and cyanotoxins. Environment International, 59: 303–327

Moghaddari M, Yousefi F, Ghaedi M, Dashtian K (2018). A simple approach for the sonochemical loading of Au, Ag and Pd nanoparticle on functionalized MWCNT and subsequent dispersion studies for removal of organic dyes: Artificial neural network and response surface methodology studies. Ultrasonics Sonochemistry, 42: 422–433

Olmez-Hanci T, Arslan-Alaton I, Basar G (2011). Multivariate analysis of anionic, cationic and nonionic textile surfactant degradation with the H_2O_2/UV-C process by using the capabilities of response surface methodology. Journal of Hazardous Materials, 185(1): 193–203

Oyama Y, Fukushima T, Matsushita B, Matsuzaki H, Kamiya K, Kobinata H (2015). Monitoring levels of cyanobacterial blooms using the visual cyanobacteria index (VCI) and floating algae index (FAI). International Journal of Applied Earth Observation and Geoinformation, 38: 335–348

Piuleac C G, Curteanu S, Rodrigo M A, Saez C, Fernandez F J (2013). Optimization methodology based on neural networks and genetic

algorithms applied to electro-coagulation processes. Central European Journal of Chemistry, 11(7): 1213–1224

Shen Q, Zhu J, Cheng L, Zhang J, Zhang Z, Xu X (2011). Enhanced algae removal by drinking water treatment of chlorination coupled with coagulation. Desalination, 271(1–3): 236–240

Trinh T K, Kang L S (2011). Response surface methodological approach to optimize the coagulation–flocculation process in drinking water treatment. Chemical Engineering Research & Design, 89(7): 1126–1135

Wang J P, Chen Y Z, Ge X W, Yu H Q (2007). Optimization of coagulation–flocculation process for a paper-recycling wastewater treatment using response surface methodology. Colloids and Surfaces. A, Physicochemical and Engineering Aspects, 302(1–3): 204–210

Wang L, Yan X, Ma J, Xu X (2017). Process analysis study on algae removal from eutrophic water in Taihu lake. Journal of Changzhou University, 29(1): 41–45 (in Chinese)

Wang Y, Chen K, Mo L, Li J, Xu J (2014). Optimization of coagulation–flocculation process for papermaking-reconstituted tobacco slice wastewater treatment using response surface methodology. Journal of Industrial and Engineering Chemistry, 20(2): 391–396

Wang Y, Hu W, Peng Z, Zeng Y, Rinke K (2018). Predicting lake eutrophication responses to multiple scenarios of lake restoration: A three-dimensional modeling approach. Water (Basel), 10(8): 994–1012

Westrick J A, Szlag D C, Southwell B J, Sinclair J (2010). A review of cyanobacteria and cyanotoxins removal/inactivation in drinking water treatment. Analytical and Bioanalytical Chemistry, 397(5): 1705–1714

Xiao X, He J, Huang H, Miller T R, Christakos G, Reichwaldt E S, Ghadouani A, Lin S, Xu X, Shi J (2017). A novel single-parameter approach for forecasting algal blooms. Water Research, 108 (Supplement C): 222–231

Yang Z L, Gao B Y, Yue Q Y, Wang Y (2010). Effect of pH on the coagulation performance of Al-based coagulants and residual aluminum speciation during the treatment of humic acid-kaolin synthetic water. Journal of Hazardous Materials, 178(1–3): 596–603

Zamyadi A, Coral L A, Barbeau B, Dorner S, Lapolli F R, Prévost M (2015). Fate of toxic cyanobacterial genera from natural bloom events during ozonation. Water Research, 73: 204–215

Zhang P, Wu Z, Zhang G, Zeng G, Zhang H, Li J, Song X, Dong J (2008). Coagulation characteristics of polyaluminum chlorides PAC-Al_{30} on humic acid removal from water. Separation and Purification Technology, 63(3): 642–647

Zhang Y, Tian J, Nan J, Gao S, Liang H, Wang M, Li G (2011). Effect of PAC addition on immersed ultrafiltration for the treatment of algal-rich water. Journal of Hazardous Materials, 186(2–3): 1415–1424

Zhao J N, Sun L X, Tan Z C (2010). Low-temperature heat capacities and thermodynamic properties of *N*-benzyloxycarbonyl-L-3-phenylalanine ($C_{17}H_{17}NO_4$). Journal of Chemical & Engineering Data, 55(10): 4267–4272

Zheng X Y, Zheng H L, Zhao S Y, Chen W, Yan Z Q, Dong L H (2015). Review on the removal of algae in source water by coagulation technology. Chemical Research and Application, 27(11): 1619–1624