Application of Response Surface Method as an Experimental Design to Optimize Coagulation Tests

Thuy Khanh Trinh, Lim Seok Kang[†]

Department of Environmental Engineering, Pukyong National University,
559-1, Daeyeon-3-dong, Nam-gu, Busan 608-739, Korea

Abstracts

In this study, the response surface method and experimental design were applied as an alternative to conventional methods for the optimization of coagulation tests. A central composite design, with 4 axial points, 4 factorial points and 5 replicates at the center point were used to build a model for predicting and optimizing the coagulation process. Mathematical model equations were derived by computer simulation programming with a least squares method using the Minitab 15 software. In these equations, the removal efficiencies of turbidity and total organic carbon (TOC) were expressed as second-order functions of two factors, such as alum dose and coagulation pH. Statistical checks (ANOVA table, R^2 and R^2_{adj} value, model lack of fit test, and p value) indicated that the model was adequate for representing the experimental data. The p values showed that the quadratic effects of alum dose and coagulation pH were highly significant. In other words, these two factors had an important impact on the turbidity and TOC of treated water. To gain a better understanding of the two variables for optimal coagulation performance, the model was presented as both 3-D response surface and 2-D contour graphs. As a compromise for the simultaneously removal of maximum amounts of 92.5% turbidity and 39.5% TOC, the optimum conditions were found with 44 mg/L alum at pH 7.6. The predicted response from the model showed close agreement with the experimental data (R^2 values of 90.63% and 91.43% for turbidity removal and TOC removal, respectively), which demonstrates the effectiveness of this approach in achieving good predictions, while minimizing the number of experiments required.

Keywords: Coagulation tests, Drinking water treatment, Experimental design, Optimizing coagulation, Response surface methodology

1. Introduction

Coagulation has been widely investigated for the removal of turbidity and natural organic matter (NOM) in water treatment. The effectiveness of coagulation influences the efficiency of subsequent sedimentation and filtration processes in water treatment [1]. The performance of coagulation is affected by many factors, not only the characteristics of raw water, such as the amount of particulate material, the amount and nature of NOM present and the chemical/physical properties, but also the conditions of coagulation, such as coagulant type, dose, and pH [2].

For individual water, the performance of a coagulant depends on two effective factors; the coagulant dose and coagulation pH. To seek the optimal conditions for these factors, jar tests using a conventional multifactor method, also known as one factor at a time (OFAT) method, was employed. In this approach, optimization is usually carried out by varying a single factor, while keeping all other factors fixed at a specific set of conditions. The OFAT method is not only time and energy consuming, but also usually incapable of reaching the true optimal conditions because

it ignores their interactions [3]. On the other hand, the statistical method using response surface methodology (RSM) has been proposed to determine the influences of individual factors and the influence of their interactions. RSM is a technique for designing experiments, building models, evaluating the effects of several factors, and achieving the optimum conditions for desirable responses with a limited number of planned experiments [4]. RSM helps to demonstrate how a particular response is affected by a given set of input variables over some specified region of interest, and what input values will yield a maximum (or minimum) for a specific response. RSM was initially developed for the purpose of determining optimum operation conditions in the chemical industry, but it is now used in a variety of fields and applications, not only in the physical and engineering sciences, but also in biological, clinical, and social sciences [5]. The response surface design was used in this study to: 1) find how jar tests comprising of several levels of coagulation factors can be simplified, 2) determine how turbidity and total organic carbon (TOC) (as responses) are affected by changes in the level of alum dose and coagulation pH (as factors), 3) determine the optimum combination of dose and pH that yields the best removal of tur-

bidity and TOC, and 4) quantitatively measure the significance and interaction between factors with respect to the optimum turbidity and TOC removals.

2. Materials and Methods

2.1. Synthetic Waters

In this study, synthetic water was used as a raw water to maintain a consistent water quality while conducting the experiments. Commercial humic acid and kaolin (Sigma Aldrich, Milwaukee, WI, USA) were used to simulate humic substances and particles in surface water. The synthetic water was prepared by mixing prescribed amounts of kaoline clay and stock solution of humic acid in tap water. The humic stock solution was made by dissolving 1.5 g humic acid in 1,000 mL of 0.01 M NaOH with 6 hours continuous stirring, filtered through 0.45 μm membrane filter and stored in refrigerator for later use. The characteristics of the synthetic water are given in Table 1.

Table 1. Characteristics of the synthetic water [6]

Parameter	Turbidity (NTU)	UV_{254} (cm^{-1})	TOC (mg/L)	SUVA (L/mg.m)	pH	Alkalinity (mg CaCO$_3$/L)
Range	9.0~11	0.09~0.10	2.6~2.7	3.8~4.1	7.4~8.0	75~80

TOC: total organic carbon, SUVA: specific ultraviolet absorbance.

2.2. Jar Tests

Alum was used as a coagulant in this work. The alum was in the powder form, with the formula: $Al_2(SO_4)_3.16H_2O$, and supplied by Sigma. The coagulation pH was adjusted using 0.1 M HCl or 0.1 M NaOH just before dosing of the coagulant.

To investigate the removal efficiencies of turbidity and TOC, jar tests were used to perform the coagulation-flocculation process. The experiments were carried out using 2 L square jars, with six paddle stirrers (Phipps and Bird, Richmond, VA, USA). The time and speed for rapid and slow mixing were set with an automatic controller as following: Rapid mixing at 250 rpm (G = 550 sec^{-1}) for 1 min after coagulant addition, followed by slow mixing at 30 rpm (G = 22 sec^{-1}) for 30 min, and then settling for 30 minutes. The supernatant was then analyzed for turbidity and TOC following the Standard Methods [6] and manufacturer's instructions for the instrument. The turbidity and TOC removals (%) were used as responses requiring optimization.

2.3. Design of Experiments

When process factors (independent variables) satisfy an important assumption that they are measurable, continuous, and controllable by experiments, with negligible errors, the RSM procedure was carried out as follows:

1) A series of experiments were performed for adequate and reliable measurement of the response of interest.
2) A mathematical model of the second-order response surface with the best fit was developed.
3) The optimal set of experimental parameters producing the optimum response value was determined.

4) The direct and interactive effects of the process parameters (factors) were represented through two and three-dimensional plots.
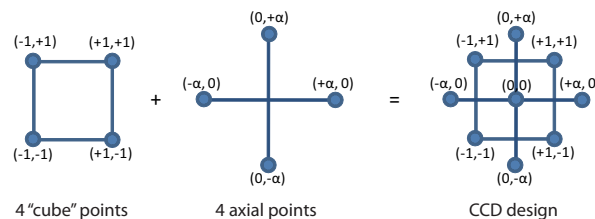

Fig. 1. Central composite design (CCD) design: 4 "cube" points, 4 axial points, and 5 replicates at the center point (0,0).

The first requirement of the RSM, as mentioned above, involves the design of experiments to achieve adequate and reliable measurements of the response of interest. A central composite design (CCD), which is a very efficient design tool for fitting second-order models [7], was selected for use in this study. The number of tests required for a CCD include: the standard 2^k cube points, 2^k axial points fixed axially at a distance, say α, from the center to generate quadratic terms, and replicate tests at the center of experimental region; where k is the number of variables. Replicates of the test at the center are very important as they provide an independent estimate of the experimental error. Generally, three to five center runs are recommended [7]. A CCD for 2 factors (alum dose and coagulation pH), with 5 replicates at the center resulting in total $2^2 + 2^2 + 5 = 13$ runs, is illustrated in Fig. 1. A CCD is made rotatable by the choice of α. A value of α = $[2^k]^{1/4}$ assures rotation of the CCD [8]. In this study k was 2 factors (alum dose and coagulation pH), therefore α became 1.414.

In order to define the experimental domain explored, preliminary experiments were carried out to determine narrower, more effective ranges of alum dose and pH prior to designing the experimental runs. It was found from the preliminary tests that coagulation was most effective in the range of alum doses from 12 to 68 mg/L and pH in the range from 5.5 to 8.5. Once the desired value ranges of the variables had been defined, they were coded to lie at ±1 for the factorial points, 0 for the center points, and ±α for the axial points. The codes were calculated as functions of the range of interest of each factor, as shown in Table 2. A CCD with 4 factorial points, 4 axial points and 5 additional experimental trials (run numbers 10~13) as replicates of the center point are given in Table 3.

In order to determine if a relationship existed between the factors and the responses investigated, the collected data was

Table 2. Relationship between the coded and actual values of a factor

Code	Actual value of factor
−α	X_{min}
−1	$\dfrac{(\alpha-1)X_{max}+(\alpha+1)X_{min}}{2\alpha}$
0	$\dfrac{X_{max}+X_{min}}{2}$
+1	$\dfrac{(\alpha-1)X_{min}+(\alpha+1)X_{max}}{2\alpha}$
+α	X_{max}

X_{max} and X_{min}: maximum and minimum value of X.

analyzed statistically using regression analyses. A regression design is employed to model a response as a mathematical function (either known or empirical) of a few continuous factors and 'good' model parameter estimates are desired [7]. Each response of Y can be represented by a mathematical equation that correlates the response surface. The responses can be expressed as second-order polynomial equations, according to Eq. (1):

$$Y = f(x) = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j + \sum_{i=1}^{k} \beta_{ii} x_i^2 \qquad (1)$$

where Y is the predicted response (turbidity removal and TOC removal) used as a dependent variable; k the number of independent variables (factors), x_i ($i = 1, 2$) the input predictors or controlling variables (factors); β_0 the constant coefficient, and β_i, β_{ij} and β_{ii} the coefficients of linear, interaction and quadratic term, respectively. The coefficient parameters were estimated using a multiple linear regression analysis employing the software Design-Expert(version 8.0.1). Design-Expert was also used to find the 3-D surface and 2-D contour plots of the response models.

3. Results and Discussion

3.1. Model Fitting

Thirteen observed responses were used to compute the model using the least square method. The two responses (% turbidity and TOC removal) were correlated with the two factors (alum dose and coagulation pH), using the second-order polynomial, as represented by Eq. (1). From the experimental data, quadratic regression models were obtained, as shown in Eqs. (2) and (3):
Turbidity removal (%):

$$Y_1 = -175.929 - 0.197X_1 + 69.358X_2 - 0.017X_1^2 - 5.020X_2^2 + 0.216X_1X_2 \qquad (2)$$

$$(R^2 = 90.63\%, R^2_{adj} = 83.93\%)$$

TOC removal (%):

$$Y_1 = -19.805 + 0.063X_1 + 16.997X_2 - 0.0002X_1^2 - 1.303X_2^2 + 0.022X_1X_2 \qquad (3)$$

$$(R2 = 91.43\%, R2adj = 85.3\%)$$

where X_1 and X_2 are the alum dose and coagulation pH, respectively. The coefficients with one factor (the ones in front of X_1 or X_2) represent the effects of that particular factor, while the coefficients with two factors (the ones in front of X_1X_2) and those with second-order terms (the ones in front of X_1^2 or X_2^2) represent the interaction between the two factors and the quadratic effects, respectively. The positive sign in front of the terms indicates a synergistic effect, while the negative sign indicates an antagonistic effect.

3.2. Validation of the Models

It is usually necessary to check the fitted model to ensure it provides an adequate approximation to the real system. Unless the model shows an adequate fit, proceeding with investigation and optimization of the fitted response surface is likely to give poor or misleading results. Graphical and numerical methods, as a primary tool and confirmation for graphical techniques were used to validate the models in this study [9]. The graphical method characterizes the nature of residuals of the models. A residual is defined as the difference between an observed value Y and its fitted $\hat{Y}$. The first plot, residuals versus the fitted values, as shown in Fig. 2, was used to exam the sufficiency of the functional part of the model. Then, the second plot, residual versus order, as shown in Fig. 3, each residual is plotted against an index of observation orders of data, which was used to check for any drift in the process. As previously shown in Figs. 2 and 3, the graphical residual analysis indicated no obvious pattern, implying the residuals of the models were randomly distributed. Lastly, in the normal probability plot, shown in Fig. 4, the data

Table 3. CCD and the results obtained

| Run | Experimental design | | Results | | | |
| | Alum dose (X_1) | pH (X_2) | Y_1, turbidity removal (%) | | Y_2, TOC removal (%) | |
			Experimental	Predicted	Experimental	Predicted
1	−1	−1	73.27	74.78	37.64	38.14
2	+1	−1	60.91	65.28	37.64	38.01
3	−1	+1	85.45	81.57	36.77	36.51
4	+1	+1	90.36	89.33	38.50	38.11
5	−1.414	0	74.91	76.69	37.64	37.49
6	+1.414	0	77.73	75.47	38.50	38.54
7	0	−1.414	72.55	68.49	38.50	37.91
8	0	+1.414	86.73	90.30	36.34	36.82
9	0	0	87.73	89.44	39.37	39.97
10	0	0	87.00	89.44	39.80	39.97
11	0	0	88.00	89.44	40.23	39.97
12	0	0	92.55	89.44	40.23	39.97
13	0	0	91.92	89.44	40.23	39.97

CCD: central composite design, TOC: total organic carbon.

Fig. 2. Residuals vs. fitted values for (a) turbidity removal and (b) TOC removal.
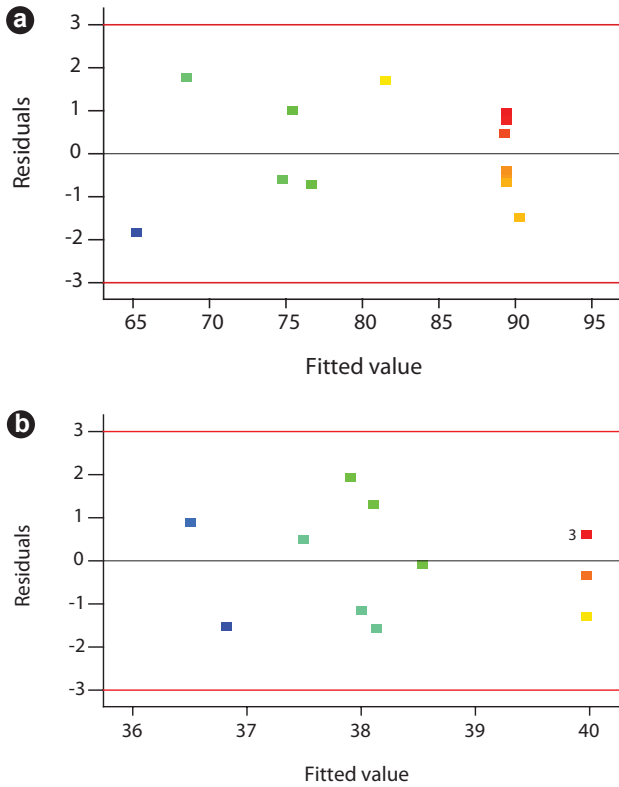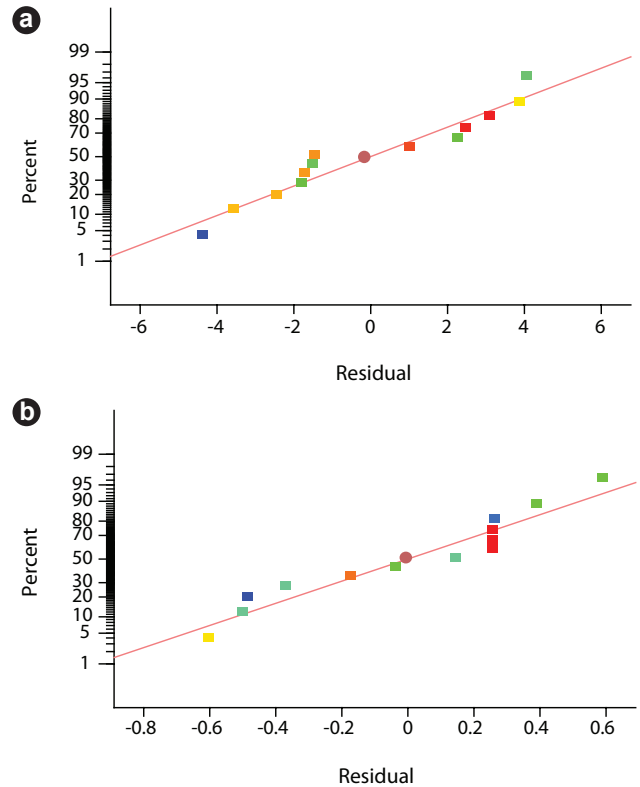

Fig. 3. Residuals vs. observation orders of data for (a) turbidity removal and (b) TOC removal.

Fig. 4. Normal probability plots for (a) turbidity removal and (b) TOC removal.

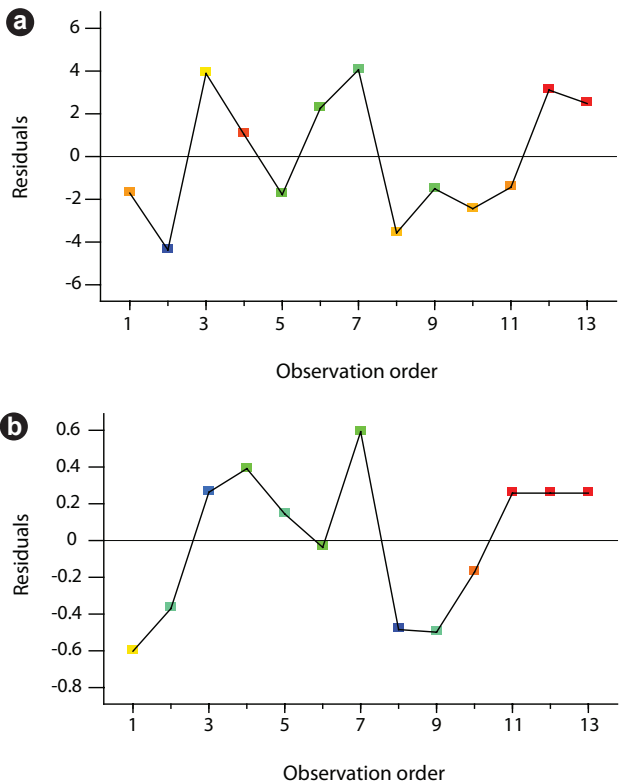were plotted against a theoretical normal distribution in such a way that the points should form an approximate straight line, and a departure from this straight line would indicate a departure from a normal distribution, which was used to check the normality distribution of the residuals. As shown in Fig. 4, it is plausible that the assumptions of normality were satisfied for the data.

The models were then checked using a numerical method employing the coefficient of determination (R^2), adjusted R^2 (R^2_{adj}), and then calculated as shown in Eqs. (4) and (5) [10]. R^2 indicates how much of the observed variability in the data was accounted for by the model, while R^2_{adj} modifies R^2 by taking into account the number of covariates or predictors in the model.

$$R^2 = 1 - \frac{SS_{residual}}{SS_{model} + SS_{residual}} \tag{4}$$

$$R_{adj}^{2} = 1 - \frac{n-1}{n-p}(1-R^2) \tag{5}$$

where SS is the sum of the squares, n the number of experiments, and p the number of predictors (term) in the model, not counting the constant term. The response surface models were developed in this study with values of R^2 higher than 90%, say 90.63 and 91.63% for turbidity and TOC removal, respectively. Furthermore, an R^2_{adj} close to the R^2 values insures a satisfactory adjustment of the quadratic models to the experimental data. Therefore, the regression models explained the removal efficiency well. Moreover, the ANOVA on these models, as shown in Table 4, demonstrates that the models were highly significant,

as evident from the very low probability of $p > F$ values in the regression = 0.002 and 0.001 for turbidity and TOC removal, respectively.

The ANOVA in Table 4 also shows the results of the lack of fit test for the models. The lack of fit test describes the variation in the data around the fitted model. If the model does not fit the data well, the lack of fit will be significant. The large p values of 0.118 and 0.178 for the lack of fit for the turbidity and TOC removal models, presented in Table 4, illustrate that the lack of fit was not insignificant, implying the models adequately described the data.

3.3. Optimization Analysis

Table 4. ANOVA results for the two responses: turbidity removal and TOC removal

	d.f	Sum of square	Mean square	*F*-value	*p > F*
Turbidity removal					
Regression	5	984.010	196.803	13.54	0.002
Total error	7	101.770	14.539		
Lack of fit	3	75.030	25.009	3.74	0.118
Pure error	4	26.751	6.686		
TOC removal					
Regression	5	19.452	3.890	14.93	0.001
Total error	7	1.824	0.261		
Lack of fit	3	1.225	0.408	2.73	0.178
Pure error	4	0.598	0.150		

TOC: total organic carbon.

Fig. 5 shows the plot of turbidity removal vs. TOC removal data from the coagulation test. As shown in Fig. 5, no clear correlation between these two responses is noted. Therefore, the removal mechanisms of turbidity and TOC were different and the optimum conditions for the removal of each substance would also be different.

Table 5 gives an insight into the linear, quadratic and interaction effects of the parameters. These analyses were done by means of Fisher's 'F' and Student 't' tests. The Student 't' test was used to determine the significance of the regression coefficients of the parameters. The p value is used as a tool to check the significance of each factor and interaction between factors. In general, the larger the magnitude of t and smaller the value of p, the more significant is the corresponding coefficient term [3]. From Table 5, it was found that the variable with the largest effect on turbidity removal was the linear effect of pH, having a p value of 0.001, followed by dose (quadratic), pH (quadratic) and dose×pH (interaction), with p values of 0.002, 0.01, and 0.058, respectively. The linear effect of dose (in range of 12 to 68 mg/L) could be considered to have no affect on the turbidity removal, with p value of 0.727.

The results in Table 5 also indicate that pH (quadratic) was the most significant factor in determining the optimum TOC removal, with p value of 0.000, followed by dose (quadratic), pH (linear) and dose (linear), with p values of 0.001, 0.072, and 0.08, respectively. The coefficient of interaction between dose and pH
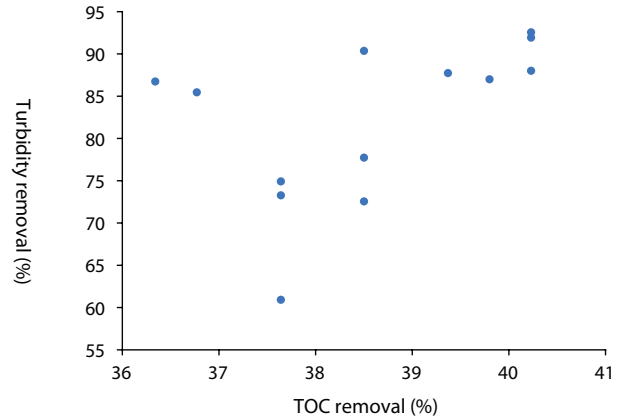

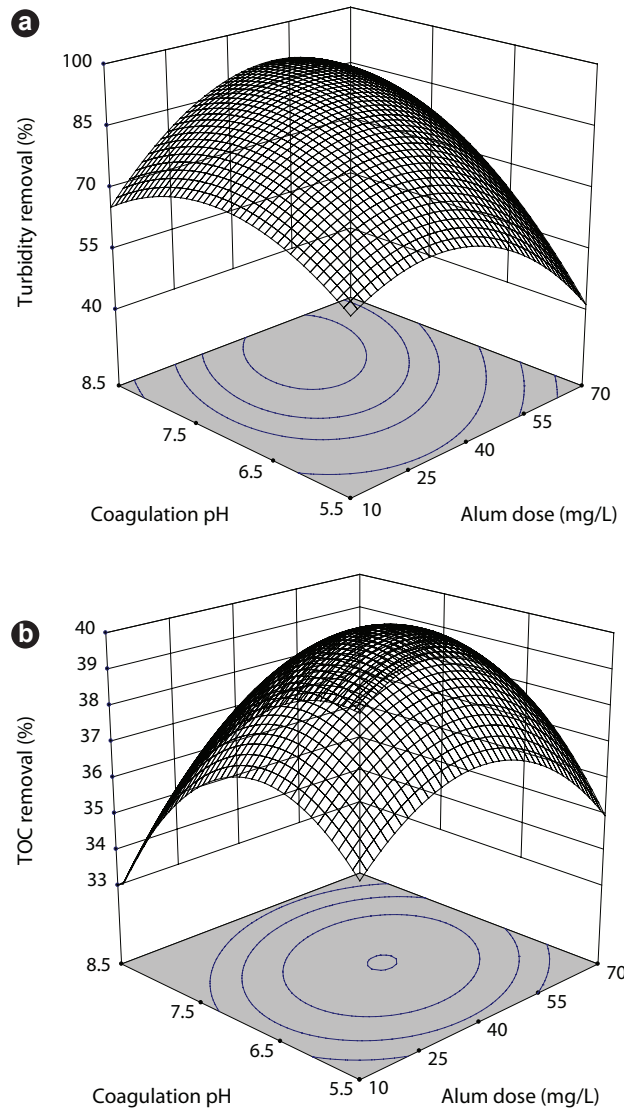**Fig. 5.** Plot of turbidity removal vs. TOC removal.

Fig. 6. Three dimensional surface plots for (a) turbidity removal and (b) TOC removal.

was found to be insignificant, with a *p* value of 0.134.

Except for the pH with respect to turbidity removal, the linear effects of pH and dose were insignificant/less significant in the statistical analyses, and became highly significant only by the quadratic effects. This difference implies that the models were not influenced by dose and pH in a linear manner, but strongly influenced by these factors in a quadratic manner. This trend can be observed in Fig. 6, with both curves showing very prominent changes. Although their linear effects were insignificant, the impacts of pH and dose to the coagulation process were still significant due to the quadratic effects towards turbidity and TOC removals. The roles of the coagulant dose and pH in coagulation are also underlined in other studies [2, 11].

Figs. 6 and 7 show the 3-D surface and 2-D contour plots, respectively. The response surface and contour plots are the graphical representation of the regression equation used to visualize the relationship between the response and experimental levels of each factor. As shown in these plots, increased turbidity removal was observed with increasing alum dose and pH values. However, an increase in both factors beyond the optimum region resulted in a decrease in the removal efficiency. At alum doses higher than 50 mg/L, the turbidity removal began to decrease at all of coagulation pHs, implying re-stabilization of the particles due to overdosing. With respect to TOC removal as a response, an increase in the removal efficiency was observed with decreasing of pH. The plots also showed that higher pH values were favorable for coagulation of turbidity, while lower pH values improved the coagulation of TOC.

An optimum level of these factors was obtained by analyzing the response surface-contour and the derivative of Eqs. (2) and (3). The optimum conditions were a set of X_1 (alum dose) and X_2 (pH), where the derivative becomes zero, as shown in Eq. (6):

$$\frac{\partial Y}{\partial X_1} = \frac{\partial Y}{\partial X_2} = 0 \qquad (6)$$

The obvious prominence in the response surfaces indicated that the optimal conditions were located exactly inside the design boundary. The maximum predicted value was indicated by the surface confined in the smallest ellipse in the contour diagram. Moreover, a canonical analysis [7] of the two models resulted in Eigenvalues of λ_1 and λ_2, which were both negative, such as $\lambda_1 = -5.022$ and $\lambda_2 = -0.015$ in the case of turbidity removal, and $\lambda_1 = -1.30$ and $\lambda_2 = -0.002$ for TOC removal, indicating that the stationary point was a single point of maximum response. The model predicted a maximum of 92.2% turbidity removal with an alum dose of 44 mg/L and pH 7.85, while the highest TOC removal was obtained under conditions of 44 mg/L alum and pH 6.90.

The turbidity and TOC removals were the two individual responses, and their optimizations were achieved under different optimal conditions. The optimum turbidity removal might aggravate TOC removal and vice versa [12]. Thus, a compromise between the conditions for the two responses is desirable. The optimum conditions for the simultaneous removals of turbidity and TOC can be visualized graphically by superimposing the contours for various response surfaces in an overlay plot. By defining the limits of the desired turbidity and TOC removals, the stripe portion of the overlain plot, as shown in Fig. 8, defines the permissible factor values. Base on the overlain contour, a compromise for 92.0% turbidity removal and 39.5% TOC removal can be met at 44 mg/L alum and pH 7.6.
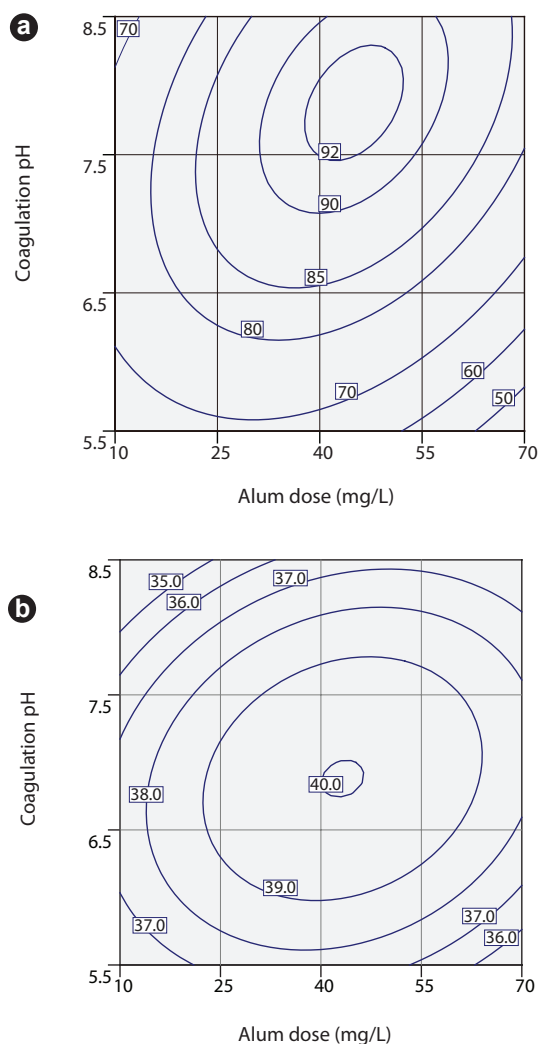
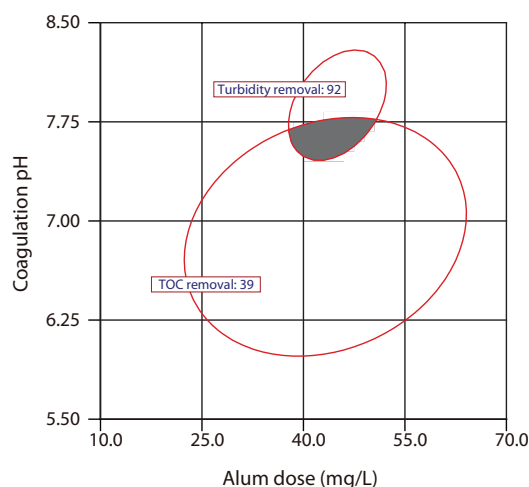**Fig. 7.** Two dimensional contour plots for (a) turbidity removal and (b) TOC removal.

Fig. 8. Overlain contour plot of turbidity and TOC removals.

Table 5. Estimation of the second-order response surface parameters (coded unit)

Parameter	Turbidity removal (%)					TOC removal (%)				
	d.f	Estimate	Std Err	*T*-value	*p* > *T*	d.f	Estimate	Std Err	*T*-value	*p* > *T*
Constant	1	89.438	1.706	52.449	0.000	1	39.971	0.228	175.098	0.000
Dose	1	−0.434	1.438	−0.322	0.757	1	0.369	0.181	2.046	0.080
pH	1	7.712	1.348	5.720	0.001	1	−0.382	0.181	−2.119	0.072
Dose×dose	1	−6.679	1.446	−4.620	0.002	1	−0.979	0.194	−5.057	0.001
pH×pH	1	−5.202	1.446	−3.437	0.010	1	−1.303	0.194	−6.734	0.000
Dose×pH	1	4.318	1.907	2.265	0.058	1	0.433	0.255	1.695	0.134

TOC: total organic carbon.

Table 6. Confirmation experiments at optimum conditions

	Coagulation condition		Responses (removal, %)	
	Alum dose	Coagulation pH	Turbidity	TOC
Experimental value	44	7.6	90.00	39.37
Predicted value	44	7.6	92.50	39.30
Error (%)			2.78	0.18

TOC: total organic carbon.

To confirm the agreements of the results achieved from the model and experiments, two additional experiments were conducted by applying the alum dose and pH in the optimum region. As shown in Table 6, the turbidity and TOC removals obtained from the additional experiments are very close to those estimated using the model, implying that the RSM approach was appropriate for optimizing the conditions of the coagulation process.

4. Conclusions

This work has demonstrated the application of RSM in seeking optimal conditions for coagulation tests. Simultaneous removals of turbidity and TOC were investigated. In order to gain a better understanding of the two factors for optimal coagulation performance, the models were presented as 3-D response surface and 2-D contour graphs. The following conclusions were obtained:

1) From the statistical analyses, the coagulant dose and coagulation pH have significant effects on the coagulation of turbidity and TOC.
2) A maximum amount of turbidity was removed using 44 mg/L alum at pH 7.8. At the same alum dose of 44 mg/L, reducing the pH to 6.9 resulted in maximum removal of TOC.
3) To simultaneously remove 92.5% turbidity and 39.5% TOC, 44 mg/L alum and pH 7.6 were selected based on the overlain contour.
4) The results of a confirmation experiment were found to be in good agreement with the values predicted by the model. This demonstrates that to obtain a maximum amount of information in a short period of time, with the least number of experiments, RSM and CCD can be successfully applied for modeling and optimizing the coagulation process.

Acknowledgements

References

1. Clearsby JL, Dharmarajah AH, Sindt GL, Baumann ER. Design and operation guidelines for optimization of the high-rate filtration process: plant survey results. Denver, CO: AWWA Research Foundation; 1989.
2. Pernitsky DJ. Coagulation 101. Alberta Water and Wastewater Operators Association (AWWOA) Annual Seminar; 2004; Alberta, Canada.
3. Mason RL, Gunst RF, Hess JL. Statistical design and analysis of experiments with applications to engineering and science. 2nd ed. New York: John Wiley and Sons; 2003.
4. Khuri AI, Cornell JA. Responses surfaces: design and analyses. 2nd ed. New York: Marcel Dekker; 1996.
5. Khuri AI. An overview of the use of generalized linear models in response surface methodology. Nonlinear Anal. 2001;47:2023-2034.
6. Clesceri LS, Greenberg AE, Eaton AD. Standard methods for the examination of water and waste water. 20th ed. Washington, DC: American Public Health Association, American Water Works Association, Water Environmental Federation; 1998.
7. Montgomery DC. Design and analysis of experiments. 5th ed. New York: John Wiley & Sons; 2001. p. 427-510.
8. Box GEP, Hunter JS. Multi-factor experimental designs for exploring response surfaces. Ann. Math. Statist. 1957;28:195-241.
9. NIST/SEMATECH e-handbook of statistical method [Inter-

net]. Available from: http://www.itl.nist.gov/div898/hand-book.

10. Haber A, Runyun RP. General statistics. 3rd ed. Reading, MA: Addision-Wesley; 1977.

11. Faust SD, Aly OM. Removal of particulate matter by coagu-lation. In: Faust SD, Aly OM, eds. Chemistry of water treat-ment, 2nd ed. Florida: CRC Press; 1998. p. 217-270.

12. Kim SH. Enhanced coagulation: determination of control-ling criteria and an effect on turbidity removal. Environ. Eng. Res. 2005;10:105-111.