Use of pH, contact time, chlorine dose and temperature on the formation of trihalomethane and some predictive models

I. García[1] & L. Moreno[2]
[1]*Department of Chemical Engineering and Environment,*
National University of Engineering, Managua, Nicaragua
[2]*Department of Chemical Engineering and Technology,*
Royal Institute of Technology, Stockholm, Sweden

Abstract

An empirical model for predicting trihalomethane (THMs) formation, using linear and nonlinear regression, was carried out using a database built when the influence of enhanced coagulation on the formation of THMs was studied. These experiments were done under different conditions of pH, temperature, chlorine dosages and contact time; and two types of coagulation (enhanced and conventional). This database was first used to compare the formation of trihalomethane obtained in these experiments with some existing predictive models. Because the values measured experimentally did not agree enough with the values from the predictive models for both coagulation types, new ones were developed using the results of enhanced coagulation experiments. In order to test the predictive capacity of the models, both linear and non-linear models, were used to predict the formation of THMs using the database obtained from the conventional coagulation experiments. The nonlinear model gave best results in comparison with the linear model. The results were statistically validated by analysis of variance and other statistics parameters.
Keywords: anova, empirical model, natural organic matter, statistic, trihalomethane.

1   Introduction

Empirical and chlorination decay models have been used to describe the Chlorination By-products (CBPs) behaviour. Amy [1] developed in 1987 THMs

prediction models for chlorination of raw and untreated water based in statistical approach using high chlorine dosages. However, efforts to model chlorine decay in natural waters started before 1950 (McClellan *et al* [7]). The objective of the models is to predict the formation of CBPs for different operating conditions such as pH, temperature, chlorine dose, bromine concentration, contact time and natural organic matter concentration; in addition, these models may reduce time and operation cost at drinking water plants.

In 1998, Amy *et al* [2] presented new models based on lower chlorine doses for raw or conventional coagulated water by alum and ferric salts. These models can predict total plus four individual trihalomethanes; total and six haloacetic acids; and chloral hydrate. The models are based in two statistic approaches: linear and nonlinear multiple regressions. The concentration of CBPs compound is the dependent variable and the water quality parameters are the independent variables. The disadvantage of those models was that they were applicable only to temperature of 20°C and pH of 7.5. Later, in 2004 Yoon *et al* [13] presented a correction to the models including the effect of pH, temperature and either alum or ferric salts were used as coagulant.

Other types are the chlorine decay models, pointed out as more flexible than the empirical model. The limitation in this type of model is the complexity of the organic matrix present in the water. At the moment, the study of the reaction pathways between natural organic matter and chlorine is limited; this makes difficult the formulation of models for the reaction between chlorine and NOM. Amy *et al* [2] also built up a two step chlorine decay model, which represents rapid and slow decay. However, Yoon *et al* [13] reported that the Amy *et al* [2] model is less accurate in chlorine residuals prediction; and the two phase models developed by Haas and Karra [5] are better predicting chlorine residuals. Nevertheless, Yoon *et al* [13] also tested the two phase models to predict THMs and HAAs formation, but no significant improvements were observed with respect to the empirical power function model of Amy *et al* [2].

Rodriguez and Serodes [9] explained that few models had been developed because analysis complexity and impossibility of precise seasonal variations evaluations. These difficulties do not permit to develop multivariate models for estimating CBPs like trihalomethanes, thus including all the parameters affecting their formation.

Information about the presence of CBPs is lacking in Nicaragua. THMs were measured by the first time in the rainy season (García and Moreno, [6]). Four drinking water plants that use conventional treatment were evaluated at that time. The results showed concentration of THMs slightly exceeding the maximum contaminant level of 80 µg/l of (USEPA, [10]) but never overpassed the maximum allowable value of 460 µg/l of the Nicaragua guideline (CAPRE, [3]). Nevertheless, when the raw water of those drinking water plants were treated with enhanced coagulation, the concentration of natural organic matter was reduced in a 50%. As a consequence, the trihalomethanes formation decreased considerably and never exceeded the guidelines values of USEPA and CAPRE.

The aim of this paper is to present empirical models that were developed using a database from one of the drinking water plant evaluated. The data base was generated from water previously treated by conventional and enhanced coagulation. The chlorination was carried out at different pH values, chlorine doses, temperatures, and contact times. Analysis of variance was performed to the database in order to determine which of the parameters evaluated have more influence in the formation of trihalomethanes. The database was first used to test the Amy *et al* [2] models for coagulated water with the correction of Yoon *et al* [13]. The hypothesis, the mean of the difference (μd) of Measured-Predicted THMs equal to 0.0 was used to validate the model. Other parameters like R-square, adjusted R-square, correlation coefficient were also estimated.

2 Material and methods

2.1 Sampling

Water samples were taken from the drinking water plant of Boaco. The conventional treatment used in the Boaco facility consists of mixing with aluminium sulphate as coagulant and calcium hydroxide to adjust the pH, followed by flocculation, sedimentation, rapid sand filtration and disinfection with chlorine (Figure 1).
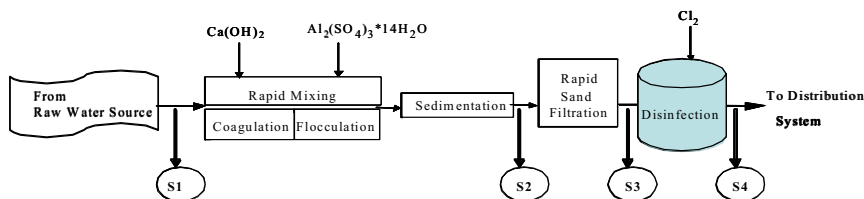

Figure 1:   Diagram of the conventional treatment at the drinking water plant and the location of the sample points.

Sample 1 (S1) was used to determine the characteristics of the raw water used in the treatment plant. Water from this point was also used to analyze the NOM surrogate parameters as total organic carbon (TOC), dissolved organic carbon (DOC), ultraviolet absorbance at 254 nm (UV_{254}) and specific ultraviolet absorbance (SUVA); and for enhanced coagulation tests at the laboratory using the procedures described by USEPA [10]. The same coagulant type, ($Al_2(SO_4)_3*14H_2O$), as that used at the drinking water plant was applied in these experiments. The effectiveness of the enhanced coagulation was evaluated in terms of the removal of precursors of THMs, and the decrease in THM level.

Samples 2 (S2) and 3 (S3) were taken to determine the removal of NOM and the chemical, organic, and physical characteristics of the water after conventional coagulation and after rapid sand filtration respectively. Sample 4 (S4) was taken

for analysis to determine the presence of THMs in the facility after disinfection with chlorine.

The samples for TOC and DOC were taken using amber glass containers of 120 ml having a screw cap with a teflon septum. They were preserved with ice and were analyzed immediately after its arrival to the laboratory. The samples for DOC were filtered using a 0.45 µm fibre filter (Whatman). THM samples were taken using 40 ml glass bottle sealed with teflon lined septa caps. The bottles were filled slowly to overflowing so that no air was included with the samples. The Persulphate Oxidation Method (HACH 10129) was used for TOC and DOC analyzes. THMs were analyzed by HACH method 10132. The accuracy of both HACH methods was determined by using the standard addition method. Besides, some samples were sent to Sweden to be compared with HACH methodology.

2.2 THMs and the influences of pH, chlorine dosage, contact time and temperature

The influences of pH, chlorine dosage, contact time and temperature on the formation of trihalomethane were investigated for raw water treated with conventional and enhanced coagulation. According to the factorial design, 81 set of parameters were considered with three samples in each case. This yields a total de 243 experiments which make up the database for the elaboration of the empirical models. The sampling period was from May to October 2004, that is rainy season in Nicaragua, the sampling were performed in May, July, and October. The description of the experiments is presented below.

For water treated with conventional treatment, the samples were taken in the facility in point S3 after conventional coagulation and rapid sand filtration. For water treated by enhanced coagulation at the laboratory, the samples were taken after filtration with paper filter No. 1 (Whatman). Both water samples were treated in the laboratory using different chlorine doses, pH values, contact times, and temperatures. The ranges were chosen so that they included the values used at the facility (Table 1).

Table 1:   Parameters to be evaluated.

Parameter	Level		
A: Chorine Doses (mg/l)	1	3	5
B: pH	5	7	10
C: Temperature (ºC)	20	25	35
D: Reaction Time (h)	24	50	100

Reagent grade sodium hypochlorite was used as chlorine source. The pH was adjusted by addition of HCl or NaOH. A constant temperature was maintained using thermostats (LAUDA, M40). 300 ml of water was used for each test. For the determination of TOC, DOC and THM, samples were extracted using glass vials with a lined septum cap and analyzed immediately.

The THMs formation during chlorination carried out at different operating conditions was analyzed by multifactorial analysis of variance (ANOVA) using the program STATGRAPHIC PLUS version 5.1. Four parameters were assumed to have major effects and are called with capital letters: A for chlorine doses; B for pH; C for temperature; and D for time as is shown in Table 1.

2.3 Comparison with existing empirical model

Waters treated by conventional and enhanced coagulation were compared with the THM concentrations obtained form the Amy models. Equations (1) and (2) were used to determine the THM concentrations predicted by these models for water treated with alum.

For Alum Coagulated Water (Amy et al [2]*):*

$$TTHM = 23.9(DOC*UV)^{0.403}(Cl_2)^{0.225}(Br^-)^{0.141}(t)^{0.264}$$

(1)

For waters treated with alum, the correction for pH and temperature of Yoon *et al* [13] was added,

$$TTHM = \left(TTHM_{pH=7.5,Temp=20^oC}\right)*(1.156)^{(pH-7.5)}(1.0263)^{(Temp-20)}$$ (2)

The comparison was statistically validated through the hypotheses test, in which the null hypothesis (*Ho*) is the mean of the difference (μ_d), equal to 0.0 (Montgomery, [8]); versus the alternative hypothesis (H_1) where the mean of the difference is distinct to 0.0. The statistical test for this hypothesis is t_o. *Ho*: $\mu_{d=0}$ would be rejected if $|t_o| > t_{\alpha/2n-1}$. Because the observations from the levels factor are paired on each experiment unit, this procedure is usually called the *paired t-test*. $t_{\alpha/2n-1}$ is 1.645 for 243 paired data when the probability (α) value is 0.05 to the 95% of confidence level.

Also the comparison was done using the recommended criteria of Yoon *et al* [13]: that a perfect model simulation corresponds to an intercept of 0, a slope of 1 and R^2 of 1. R-Squared indicates the percentage of variability represented by the model after accounting for the number of parameters that were estimated. A high R^2 indicates a high accuracy in the model. The adjusted R^2 is a variation of the ordinary R^2 statistic that reflects the number of factors or parameters in the model. Correlation Coefficients values explain the strength of the relationship between the variables.

2.4 Development of empirical models for THM formation

Statistical modelling was conducted in order to identify the most important parameters responsible for the variations of THMs in this research, and to develop a statistical model that can be used to predict concentrations of THMs by using linear and nonlinear models.

3   Result and discussion

3.1   Natural organic matter (NOM) at the raw water

The surrogate parameters of NOM as TOC, DOC, UV_{254}, SUVA and Colour were measured to the raw water of the facility in order to determine the type of organic matter present. TOC mean is 16.4±9.0 mg/l, value which is in the range of 1-40 mg/l (Croué et al [4]) for surface raw water but it is considerable a high value and indicative of erosion problems at the watershed. It is reinforced with the high average value found for colour (97.7 mg/l). DOC value (7.6±5.1) was within the typical range of surface waters (Croué et al [4]). The high SUVA (4.1±2.0) value indicates that mostly organic matter is hydrophobic of high molecular weight and can be removed easily by enhanced coagulation because the amount of NOM typically requires a greater coagulant demand than the amount of inorganic particles present in the raw water. UV on the other hand (0.27±0.16) was relatively low compared with water samples with similar DOC values (Croué et al [4]). A lower UV per mass of DOC would likely result in less CBP formation since UV and CBP are strongly correlated.

3.2   Natural organic matter after conventional (CC) and enhanced coagulation (EC)

Enhanced coagulation remove significantly more NOM that conventional coagulation (Table 2), it is explained by the fact that high coagulant doses (20-70 mg/l) were used to remove natural organic matter. On the contrary, with conventional coagulation, doses used are lesser (20-30 mg/l) and the main objective is turbidity reduction instead NOM removal. However, some NOM remained in the treated water after the water treatment process using alum salts, it means that both techniques are able to reduce some NOM, but not all of it, and THMs can still be formed at the disinfection step.

Table 2:   NOM parameters after coagulation processes.

Techniques	TOC (mg/l)	DOC (mg/l)	UV (cm^{-1})	SUVA (l/mg-m)	*Colour (mg/l Pt-Co)*
CC	5.7 ± 3.8	3. 6 ± 3.5	0.06 ± 0.03	2.8 ± 1.9	*11.0 ± 10.5*
EC	*2.7 ± 1.0*	*1.2 ± 0.5*	*0.03 ± 0.02*	*1.4 ± 0.6*	*1.9 ± 1.3*

3.3   Presence of trihalomethanes after conventional and enhanced coagulation

Comparison of the results of both coagulation procedures with the guideline of EPA and CAPRE showed that the mean values of THMs are much less that both guidelines (Table 3). Nevertheless, with conventional coagulation, the MCLs of EPA is exceeding in some opportunities, which never occurred with enhanced coagulation. According to Villanueva et al [11] a mean THM level of 80 µg/l was reported by European survey of surface waters.

Table 3:   Trihalomethanes.

THMs (µg/l)	Maximum Value	Minimum Value	Mean Value	EPA MCL Guideline	CAPRE MCL Guideline
CC	130.0	15.0	57.5 ± 39.8	80	460
EC	35.3	2.4	12.2 ± 13.4	80	460

3.4 Trihalomethanes formation. Impact of the operating parameters

The influence of pH, chlorine dosage, contact time and temperature on the formation of trihalomethane was investigated for raw water treated with conventional and enhanced coagulation (Table 1). The database was generated using coagulant doses of 20-30 mg/l and 20-70 mg/l; DOC concentration of 1.19-2.02 and 0.76-1.23 mg/l; and UV_{254} values of 0.05-0.09 and 0.04-0.09 cm^{-1} for conventional and enhanced coagulation respectively.

One example of the results is shown in Figures 2. The lines in the figures correspond to the mean values of three samples taken at different times in 2004. The red, blue and green characters indicate the experimental data for pH 5, 7 and 10 respectively. In these same figures, can be noticed the effect of enhanced coagulation (right side) for reducing the formation of trihalomethanes in comparison with conventional coagulation (left side figure).

Higher doses of chlorine, pH, temperature and time increase the formation of THMs. Therefore, the maximum contaminant levels of 80 µg/l of USEPA is exceeding several times. Nevertheless, MCLs of 460 µg/l used in Nicaragua is never exceeded because the maximum concentration of THMs found was 176 µg/l to the extreme conditions of 35°C, pH 5, chlorine doses of 5 mg/l and contact time of 100 h (Figure 2). The tendency of the data to increase the concentration of THMs when high values of pH, temperature, contact time and chlorine doses are used has been reported by others researchers as Amy *et al* [2]; Rodriguez and Serodes [9]; and Xie [12]. It is because the high pH creates an alkaline environment which causes a fast chlorination, the high chlorine doses provokes more halogenations and opening of the aromatic structures of the still available NOM after conventional and enhanced coagulation. On the contrary, when pH value of 6 or lower is used, the environment is acid and the reaction is too slow to form end products as THMs, only intermediate products can be formed. To lower pH the dominant chlorine species is HOCl (pH < 7) which is a neutral molecules that penetrates the cell membrane easier than OCl$^-$, therefore HOCl has more disinfection effects than OCl$^-$.

As a consequence of fast substitution and oxidation of the organic matter structures by chlorine, more THMs are formed. Also, the temperature has influenced in the formation of THMs due to the high value of temperature which increase the reactivity between chlorine and NOM, and the rate of production of final end products as THMs. The effect of contact time is notorious on Figure 2, for the first 20 hours of contact time, the spreading of the data is very narrow being the opposite for 50 and 120 h where they are found in a wide interval. It is

explained because THMs is an end product and increasing the time there is more opportunity for the reaction to be totally completed.
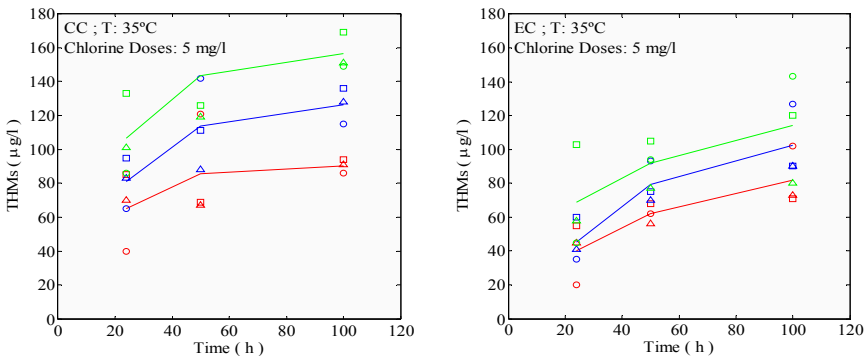

Figure 2:   Influence of temperature, pH and time on THMs for chlorine doses 5 mg/l.

3.5 ANOVA for the results of THMs formation

According with ANOVA, the parameters evaluated have statistically significant effect on THMs formation due *p-values* for each parameter was lesser than 0.05 (Table 4) at the 95.0% confidence level. In the same table, the sums of squares are lower for enhanced coagulation; it means that data variability for enhanced coagulation is lesser than for conventional coagulation. F-Ratios for enhanced coagulation were lower for chlorine doses and pH, in comparison with the same parameters treated by conventional coagulation. However, F-ratios for the parameters evaluated in both coagulation types are higher, and implies that each parameter has great influence on the formation of THMs.

The significant tests for the interactions among parameters for conventional and enhanced coagulation are higher than 0.05, which indicates that the interactions have not statistically significance effect on THMs formation at the 95% of confidence level.

Table 4:   Analysis of variance for THMs.

Main Effects	Conventional Coagulation			Enhanced Coagulation		
	Sum of Squares	F-Ratio	P-Value	Sum of Squares	F-Ratio	P-Value
A: Chlorine Doses	5.4×10^4	83.3	< 0.05	2.6×10^4	59.1	< 0.05
B: pH	4.2×10^4	65.4	< 0.05	2.1×10^4	48.3	< 0.05
C: Temperature	7.1×10^4	108.1	< 0.05	5.9×10^4	135.7	< 0.05
D: Time	5.7×10^4	87.5	< 0.05	4.4×10^4	102.4	< 0.05

3.6 Comparison between predictive and measured THMs

A statistical analysis was carried out to validate the comparison of the THMs measured experimentally with the THMs predicted by Equations (1) and (2). The

hypothesis that the mean of the difference between Measured-Predicted THMs equals to 0.0 versus the alternative hypothesis that the mean of the difference Measured-Predicted THMs is not equal to 0.0 was assumed. For conventional coagulation the mean of the difference was -20.67±3.78 and for enhanced coagulation was -24.80±2.63. t_o were -10.77 and -18.58 respectively. Since absolute t_o value*s* were higher than 1.645 the null hypothesis is not accepted. Therefore, Equations (1) and (2) cannot be used to predict the formation of THMs of Boaco drinking water plant.

Additionally other statistic parameters were used to compare the THMs measured and predicted. The results demonstrated that there is a statistical significance difference between the THMs measured and predicted because the slopes are 0.6 and 0.7, the intercepts are 11.7 and -4.0 for CC and EC respectively. The R^2 statistic indicated that the models present 50 and 60.0% of the variability in THMs measured. A perfect simulation corresponds to an intercept of 0, a slope of 1 and an R^2 of 1. However, the correlation coefficients are 0.7 and 0.8 respectively, indicating a moderately strong relationship between variables.

3.7 Empirical model for predict THMs formation in the drinking water of Boaco

Using the data base generated with the enhanced coagulation experiments; two equations ((3) and (4)) were found which can predict the formation of THMs. Figure 3 presents the curve for linear and nonlinear model which follows a normal distribution that has 95% confidence levels for the coefficients in the models.

At linear model (Equation (3)), the R^2 statistic indicates that the model, as fitted, explains 78% of the variability in THMs. The adjusted R^2 is 77% and the correlation coefficient is 0.88, values that suggest a strong relationship between variables of the model.

$$THMs = 28.05 + 2.49(Cl_2) - 11.85(pH) + 0.27(T) - 0.42(t) - 36.1(DOC)$$
$$+0.06(Cl_2*t) + 0.20(pH*T) + 0.04(pH*t) + 8.15(pH*DOC) + 0.01\ (t*T) \qquad (3)$$

At nonlinear model R^2 statistic indicates that the model as fitted explains 77% of the variability in THMs. The adjusted R^2 is 76.8% and the correlation coefficient is 0.88. The equation of the nonlinear model is 4.

$$THMs=1.35\ [0,015*(Cl_2)^{0,349}*(pH)^{0,708}*(T)^{1,293}*(t)^{0,500}*(DOC)^{0,538}\ ]^{0.94} \qquad (4)$$

Equations (3)-(4) were used to predict the formation of THMs using the database of the experiments with conventional coagulation. The null hypothesis was tested in both cases, the means of the difference Measured-Predicted THMs is 0.0. For the linear model, the mean of the difference was 8.81±3.26 and t_o was 5.33 which is higher than 1.645, therefore the null hypothesis is rejected. It means that the proposed linear model tends to under predict the concentration of THMs.

The mean of the difference at the non-linear model was 1.26±2.77, and t_o was 0.90. Since t_o was lower than 1.645, the null hypothesis null is accepted. It means that the proposed nonlinear model can be used to predict the THMs concentration of Boaco drinking water plant.

4 Conclusions

Comparison between the water treated by conventional and enhanced coagulation shows that enhanced coagulation drops considerably the THM concentration and never exceeds the existing guidelines.

The study of the trihalomethane formation by the variation of pH, time, temperature, and chlorine dose, using water treated by conventional and enhanced coagulation demostrates that higher values of those parameters increase greatly the formation of THMs.

The experimental results generated in this research were used to test the Equations of Amy to predict THMs formation. Since the agreement was not sufficiently good, it was decided to develop new models.

Two empirical models were done using linear and nonlinear regression. These models were tested using the database of the conventional coagulation experiments; the nonlinear model shows that it can predict the formation of THMs in the Boaco drinking water plant.

Acknowledgement

References

[1] Amy G.L., Minear, R.A., & Cooper, W.J., Testing and validation of a multiple nonlinear regression model for predicting trihalomethane formation potential. *J. Water Research.* 21, pp. 649-659, 1987.
[2] Amy, G., Siddiqui, M., Ozekin, K., Zhu, H.W., & Wang, Ch., Empirically based models for predicting chlorination and ozonation by-products; trihalomethanes, haloacetic acids, chloral hydrate, and bromate. *USEPA Reports CX 819579*, 1998.
[3] CAPRE Guidelines., Regional committee for drinking water institution and sanitation for Central America, Panama and Dominican Republic, pp. 10-3, 2000.
[4] Croué, J.P., Debroux, J.F., Amy, G., Aiken, G.R., & Leenheer, J.A., Natural organic matter: structural characteristic and reactive properties. In formation and control of disinfection by-products in drinking water. *AWWA Publication.* pp. 56-92, 2001.
[5] Haas, C.N., & Karras, S.B., Kinetics of wastewater chlorine demand exertion. *J. Water Pollution con F.* 56(2), pp. 170-173, 2004.

[6]   Garcia, I., & Moreno, L., Presence of Trihalomethanes in four drinking water plant. *J. of Water Supply: Aqua*, 55(3), pp. 221-231, 2006.
[7]   McClellan, D., Reckhow, D., Tobiason. J., & Edzwald, J., A comprehensive kinetic model for chlorine decay and chlorination by-product formation. Department of civil and environmental engineering. University of Massachusetts. 2004.
[8]   Montgomery, D.C., Design and Analysis of Experiments. Fifth Edition. John Wiley & Sons. Inc. USA. 1997.
[9]   Rodriguez, M.J & Serodes, J.B., Spatial and temporal evolution of trihalomethanes in three water distribution system. *J. Water Research.* 35(6)**,** pp. 1572-1586, 2001.
[10]   USEPA., Enhanced coagulation and enhanced precipitative softening guidance manual. Environmental Protection Agency. United States. Office of Water. *EPA 815-R-99-012*. 1999.
[11]   Villanueva, C.M., Kogevinas, M., & Grimalt, J.O., Haloacteic acids and trihalometahnes in finished drinking waters from heterogeneous sources. *J. Water Research*. 37 (2), pp. 953-958, 2003.
[12]   Xie, Y.F., Disinfection by-products in drinking water: formation, analysis and control. *Lewish Publishers. USA*. 2004.
[13]   Yoon, Y., Sohn, H., Amy, G., Cho, J., & Lee, Y., Disinfectant decay and disinfection by-products formation model development: chlorination and ozonation by-products. *J. Water Research.* 38, pp. 2461-2478, 2004.