



Development of chlorine dioxide-related by-product models for drinking water treatment

Caroline Korn^a, Robert C. Andrews^{b,*}, Michael D. Escobar^c

^a Acres and Associated Environmental Limited, Suite 525, 21 Four Seasons Place, Toronto, Ont., Canada M9B 6J8

^b Department of Civil Engineering, University of Toronto, 35 St. George Street, Toronto, Ont., Canada M5S 1A4

^c Department of Public Health Sciences, University of Toronto, Canada

Received 1 January 1999; received in revised form 1 April 2001; accepted 1 April 2001

Abstract

Factorial experiments were conducted using source waters from seven drinking water treatment plants in Ontario, Canada to develop statistically based model equations capable of predicting chlorine dioxide consumption and chlorite and chlorate formation upon chlorine dioxide application. The equations address raw water quality and operational parameters including pH, temperature, chlorine dioxide concentration, reaction time and water organic content (as described by non-purgeable organic carbon \times ultraviolet absorbance measured at 254 nm, NPOC \times UV₂₅₄). Terms describing two-factor interaction effects were also included, improving the accuracy of the predictive equations in fitting measured response concentrations as evaluated through internal and external validations. Nearly 80% of the predictions for chlorine dioxide consumption and chlorite formation were observed to be within 20% of the measured levels. Over 90% of the predicted chlorate levels were within ± 0.1 mg/L of the measured levels. Chlorine dioxide concentration and NPOC \times UV₂₅₄ were key parameters when developing the predictive models. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Chlorine dioxide; Chlorite; Chlorate; Disinfection; Modeling; Water treatment

1. Introduction

The use of chlorine as a primary disinfectant has historically been viewed as common practice. However, during treatment, chlorine reacts with naturally occurring organic material to produce numerous by-products, some of which are suspected carcinogens. Alternative disinfectants such as chlorine dioxide (ClO₂) could be considered as viable alternatives to chlorine use. Chlorine dioxide has been reported to be effective in the inactivation of pathogenic organisms including *Cryptosporidium parvum* [1]. Further, unlike chlorine,

chlorine dioxide is not known to react with humic substances to form trihalomethanes (THMs) [2–6]. While employing chlorine dioxide as a disinfectant presents many advantages, concern exists over the formation of by-products including chlorite (ClO₂⁻) and chlorate (ClO₃⁻) which may lead to hemolytic anemia at low levels of exposure; while higher levels of exposure can result in an increase in methemoglobin [7].

Tradeoffs between achieving a desired disinfection level and forming disinfection by-products (DBPs) encourage the development and use of forecasting models [8]. In view of proposed regulatory limits for DBPs such as chlorite ion [9,10] as well as potential limits to be set for chlorate [11,12], the development of equations which are capable of predicting DBPs associated with chlorine dioxide treatment will become increasingly useful. Moreover, the fact that a maximum residual disinfectant level (MRDL) of 0.8 mg/L has been

*Corresponding author. Tel.: +1-416-978-5399; fax: +1-416-978-6813.

E-mail addresses: kornc@toronto.aae.on.ca (C. Korn), andrews@civ.utoronto.ca (R.C. Andrews), escobar@utstat.toronto.edu (M.D. Escobar).

proposed for residual chlorine dioxide [9], may require plants to further optimize their treatment processes.

While previous research into the development of predictive by-product formation models related to chlorination has been conducted by others [13–19], similar model equations are lacking in the literature for chlorine dioxide-related by-product formation.

This study presents disinfection by-product formation results (chlorite and chlorate) obtained from statistically designed experiments using natural waters. Equations were developed in terms of water quality and chlorine dioxide disinfection-related parameters including temperature, pH, chlorine dioxide dose, contact time and organic content. Significant two-factor interaction effects were also examined and included in the equations. Care was taken to develop equations capable of predicting chlorine dioxide consumption, chlorite and chlorate formation within boundary conditions of the examined parameters, while maintaining reasonable chemical justification and “statistical soundness”. Model validations involved comparing response levels as predicted from the equations with “internal” (levels employed as part of the original model data base) as well as with “external” (levels reported in the literature and observed upon testing waters that were not part of the database) response levels. Sensitivity analyses consisted mainly of evaluating parameter contributions to the “explanatory power” of the relevant model equations.

2. Experimental design

Experiments were conducted at bench-scale using 250 mL amber bottles serving as batch reactors. The selected “factorial design” approach differs from the conventional experimental design of varying “one-

factor-at-a-time” (OFAT). Factorially designed experiments allow the simultaneous variation of individual parameters within a given experimental run [20]. Using a factorial design allowed the identification of not only main parameter effects (such as pH or temperature), but also two parameter-interaction effects (such as pH-chlorine dioxide concentration) on DBP formation, in both a time- as well as resource-efficient manner.

To “screen” for significant main parameter and interaction effects on chlorite and chlorate formation, fractional factorially designed experiments were conducted using “synthetic” waters. The parameters that were examined included pH, temperature, ClO_2 :NPOC ratio, chlorine concentration, chlorate concentration and reaction time. During these preliminary experiments, sufficient data was provided to show that two-factor interactions may be important in chlorite and chlorate formation when employing chlorine dioxide. As such, two-factor interactions were considered when developing subsequent DBP predictive model equations.

Subsequent to the screening experiments, individual 2^4 (2 levels, 4 factors) full factorial experiments were conducted using raw waters (7 surface waters and one groundwater). All possible parameter combinations were examined, thus allowing the determination of main parameter and two-factor interaction effects on DBP formation. The experiments consisted of “two-level” factorials where the effect of each of the investigated parameters was assessed by varying the parameter between a pre-selected “low” and “high” level (Table 1). Predictive equations were developed from experimental data pooled from all eight factorially designed experiments. Hence, unless otherwise noted, each equation was developed from a database of 128 different responses (i.e. 16 experimental conditions \times 8 waters).

Table 1
Experimental factors and levels investigated in raw water experiments

Factor	Parameter values	
	Lowest level examined ^a	Highest level examined ^a
pH	6	9
Temperature (°C)	4	25
Chlorine dioxide concentration (mg/L)	1.0	6.0
Reaction time (h)	1	24
NPOC (mg/L)	0.60 (groundwater) 2.03 (Toronto) ^b	5.96
UV ₂₅₄ (1/cm)	0.009 (groundwater) 0.039 (Toronto) ^b	0.212
NPOC \times UV ₂₅₄ (mg/L/cm)	1.264	0.0054

^a Some factors were adjusted to their specified low or high levels.

^b Water obtained from the Toronto R.L. Clark Treatment Plant had the lowest NPOC and UV₂₅₄ of the seven surface waters examined in this study.

Table 2
Raw water quality parameters

Parameter	Water treatment plant (water source) ^a							
	Brantford (Grand river via Homedale canal)	Ottawa Britannia (Ottawa river)	Windsor (Detroit river)	Metro Toronto R.L. Clark (Lake Ontario)	Peterborough (Otonabee river)	Belleville (Bay of Quinte)	Mannheim Surface water ^b (Grand River)	Mannheim Groundwater ^b (Waterloo groundwater supply, Well K-26)
NPOC (mg/L)	4.29	5.96	2.06	2.03	4.82	4.96	5.71	0.60
UV ₂₅₄ (l/cm)	0.118	0.212	0.042	0.034	0.136	0.140	0.147	0.009
Specific abs. (/cm/mg/L)	0.028	0.036	0.020	0.017	0.028	0.028	0.026	0.015
pH	8.01	7.21	7.98	8.04	7.80	8.11	7.95	7.57
Alkalinity (mg/L as CaCO ₃)	223.3	33.3	93.7	98.2	80.5	99.5	170.5	265.8
Turbidity (NTU)	2.75	3.91	24.0	2.30	0.73	3.84	6.03	<0.1
Colour	15.4 TCU ^c	33.1 TCU ^c	5.40 TCU ^c	3.92 TCU ^c	15 TCU ^d	20 app. col. units ^d	21 TCU ^d	no app. colour ^d
Bromide (mg/L)	0.030	0.020	0.022	0.041	<0.020	<0.020	0.034	0.027
Iron (mg/L)	0.118	0.254	0.290	0.196	<0.100	<0.100	<0.100	<0.100
Manganese (mg/L)	<0.100	<0.100	<0.100	<0.100	<0.100	<0.100	<0.100	<0.100
Ammonia (mg NH ₃ -N/L)	0.45	<0.20	<0.20	<0.20	<0.20	<0.20	0.20	<0.20

^a All plants located in Ontario, Canada.

^b Located in Kitchener, Ontario.

^c Measured at the Ottawa Britannia plant.

^d Measured at the respective treatment plants.

Simple linear regression analyses were conducted on chlorine dioxide consumption as well as chlorite and chlorate formation data with respect to different precursor-related parameters (NPOC, UV_{254}) and combinations thereof of $UV_{254}/NPOC$ and $NPOC \times UV_{254}$. This was done to determine which terms could best describe precursor content and thus, be included in the predictive models. The linear regressions were assessed by examining various statistical data: R^2 -values (coefficient of estimation which provides an indication as to the percentage of the response level variation explained by a given correlation), SEE (standard error of the estimate), F -statistic (which tests the hypothesis that both regression coefficients have values of zero) and associated P -values (level of significance at which the given hypothesis can be rejected). An analysis of variance (ANOVA) was employed to assess the statistical significance of the predictive models as well as the individual parameters making up the model equations. The design and analyses of the “synthetic” water factorial experiments were conducted using Design Ease[®], a statistically based software package [21]. Using Minitab[®] for Windows Student Version, Release 9 statistical package [22], the development of model equations was accomplished by performing stepwise multiple linear regressions on the various parameters to obtain term coefficients fitting a generalized linear model (GLM).

While duplication of every trial run would be difficult and time-consuming during the raw water experiments, two trial runs for each of the eight experiments were duplicated to provide an estimate of pure error. The data obtained from these runs were then “pooled” and pure error estimates for chlorine dioxide, chlorite and chlorate were individually calculated.

3. Materials and methods

The eight raw water samples were obtained from seven different treatment facilities (one plant utilizing both ground and surface sources) located in Ontario, Canada. These waters were selected to cover a range of levels for the different water quality parameters, most notably NPOC (Table 2). The low and high levels reported in Table 1 for non-purgeable organic carbon (NPOC; mg/L) and the ultra-violet (UV) absorbance at 254 nm, correspond to the lowest and highest levels observed for any of the eight test waters. Manganese and iron concentrations were also measured since they may exert a chlorine dioxide demand.

Raw water samples were shipped in 18 L carboys by an overnight courier from the various treatment plants to the laboratory at the University of Toronto, Canada. Upon their arrival, the carboys were stored at 4°C until testing, typically conducted within 2–4 days of receipt.

pH adjustments were achieved through the addition of phosphate buffers as well as the addition of either reagent water:concentrated HCl (on a 1:1 volume basis) or 50% (weight/weight) NaOH. The added buffer volumes rarely exceeded 1% of the total water sample volume in order to minimize dilution effects while attaining the desired pH levels. For consistency purposes, buffer solutions were all “phosphate-based”. Minimization of phosphate addition was also important to reduce potential interferences which could hinder chlorite and chlorate analyses while using ion chromatography. While the initial pH levels were generally within ± 0.05 units of the desired value, pH checks showed that pH changes which occurred during experimentation were typically within ± 0.2 units. Following temperature and pH adjustments, samples were decanted into 250 mL-amber chlorine dioxide-demand free bottles with minimal headspace (typically less than 1 mL). The minimal headspace provided the necessary volume to provide proper mixing. To prevent volatilization losses of the stock chlorine dioxide solution, screwtop vials were fitted with Teflon[®]-lined Mininert valves (Supelco, Toronto, ON). Chlorine dioxide was added using a gas-tight glass syringe.

Chlorine dioxide was generated using a combination of methods proposed by *Standard Methods* ([23], Method 2350E), Sterling Pulp Chemicals, Ltd. (Toronto, ON) and the Pulp and Paper Research Centre at the University of Toronto (ON). As reported by Andrews et al. [24], by incorporating certain aspects and conditions of the previously mentioned generation techniques, the resulting chlorine dioxide generation method culminated in a safer and more efficient procedure. Using a bench-scale apparatus, the method involved cautiously mixing a 25% sodium chlorite solution with a 12 N sulfuric acid solution. Generally, the generated chlorine dioxide stock solutions' concentrations never exceeded 4 g/L as assessed by thiosulfate titration [23] and were found to be essentially free of chlorine, chlorite and chlorate contaminants (i.e., purity of approximately 99% or greater). Chlorine dioxide solutions (diluted to concentrations of approximately 1000 mg/L) were then stored, headspace free in 40 mL-amber vials.

Chlorine dioxide residual measurements were obtained using the ACVK method reported by Masschelein et al. [25] and Hofmann et al. [26]. A Hewlett-Packard[®] 8452A diode array UV-visible spectrophotometer with a 5 cm path length quartz cell was used. Chlorite and chlorate were measured in accordance with EPA Method 300.0 [27] with modifications, using a Dionex[®] DX 500 ion chromatograph. Instead of using a carbonate eluent, a borate eluent (120 mM H_3BO_3 /30 mM NaOH) solution was employed. Hautman and Bolyard [28] reported that either eluent could be used for chlorite and chlorate detection. Typically, water quality

parameter measurements for NPOC, UV_{254} , pH, turbidity, alkalinity, ammonia, iron and manganese followed standard methods [23] and/or recommendations were provided by instrument manufacturers. Bromide measurements were obtained following EPA Method 300.0 [27]. Sample storage times and preservation methods followed those suggested in the respective analyses methods.

Unless otherwise specified, reagents employed during experimental testing conformed at least to the minimum reagent grade specified by the analytical method of interest. Reagent water was produced by passing distilled water through a Millipore Milli-Q[®] UV-plus filtration system. The electrical resistance of the reagent water was never below 18.2 M Ω cm. Glassware (including the 250 mL-amber reactor bottles) employed in chlorine dioxide testing was rendered chlorine dioxide demand-free by soaking the articles in a strong chlorine dioxide solution (≈ 500 –1000 mg/L).

4. Results and discussion

4.1. Model formulation

Since only limited information regarding the modeling of chlorine dioxide-related organic and inorganic DBP formation has been presented in the literature, concepts arising from modeling organic DBPs upon chlorination were considered.

While examining THM formation, Amy et al. [13] transformed both dependent (response variables) and independent variables to natural-log forms based on previous work [29,30]. Log-transformation of the data collected in this study was useful in meeting the assumption of constant variation, which underlies linear regression. The advantage of applying this type of transformation is shown in Figs. 1 and 2 where chlorite formation is plotted with respect to $NPOC \times UV_{254}$ in both untransformed and log-transformed formats,

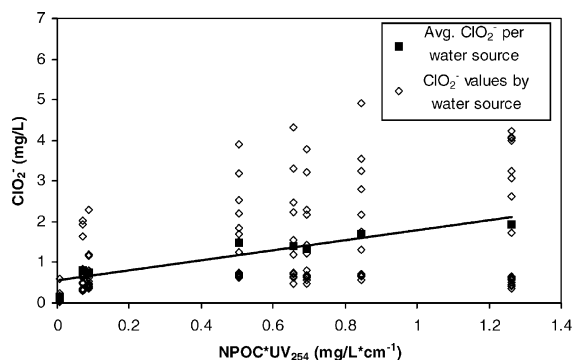


Fig. 1. Chlorite Formation in terms of $NPOC \times UV_{254}$.

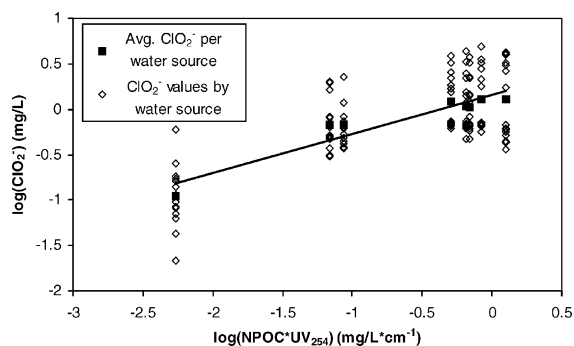


Fig. 2. Log(chlorite formation) in terms of $\log(N-$

respectively, for the different source waters. The eight “columns of data points” as shown in these two figures represent the chlorite formation observed for each of the eight water sources over the range of investigated parameter levels. For the log-transformed case (Fig. 2), the variation in chlorite concentration levels between the individual water sources is homogeneous. As such, response variable concentrations (of chlorine dioxide consumed as well as chlorite and chlorate formed) and all parameter values have been log-transformed (base 10).

It was noted that a log-transformation of the low parameter level for chlorine dioxide concentration and reaction time (1.0 mg/L and 1.0 h, respectively) would have resulted in a value of zero. To avoid this, these parameters have been expressed as $\log(\text{chlorine dioxide} + 1)$ and $\log(\text{reaction time} + 1)$. A similar procedure was described by Amy et al. [13] for the bromide concentration parameter in their predictive equations.

$NPOC \times UV_{254}$ was selected to define precursor-related material in the predictive equations for all the three response variables (chlorine dioxide, chlorite and chlorate). This selection was based on the chemical significance of NPOC and UV_{254} , statistical analyses showing significant correlations between $NPOC \times UV_{254}$ and the three response variables, as well as the “added” descriptive information which may be assumed while using two surrogates rather than one (average of both effects based on a geometric mean).

Equations formulated by Amy et al. [13] for predicting trihalomethane formation potential (THMFP) have shown that the parameter which best served as a precursor surrogate was $NPOC \times UV_{254}$. The same authors discussed the chemical relevance of this parameter combination whereby NPOC relates to precursor concentration, while UV_{254} provides evidence as to the reactivity of the precursor material. Statistical data shown in Table 3 pertain to regressions of organic-precursor surrogates performed on “all the individual response variable levels observed for the different water sources”. Based on the results, a better correlation was

Table 3

Correlation results between various precursor surrogate parameters and response variables^{a,b}

Response variable	Statistic	Precursor surrogate parameter			
		log(NPOC)	log(UV ₂₅₄)	log(NPOC × UV ₂₅₄)	log(UV ₂₅₄ /NPOC)
log (chlorine dioxide consumed)	y-intercept	−0.4925	0.9585	0.3515	4.097
	Slope	1.1086	0.7995	0.4661	2.488
	R ²	0.37 (0.90)	0.36 (0.86)	0.37 (0.88)	0.27 (0.66)
	SEE	0.448	0.454	0.451	0.483
	F-statistic	74.18	69.01	71.53	46.44
	P-value	<0.001	<0.001	<0.001	<0.001
log (chlorite)	y-intercept	−0.6147	0.7085	0.1541	3.6083
	Slope	1.010	0.7303	0.4252	2.2951
	R ²	0.52 (0.93)	0.50 (0.89)	0.51 (0.91)	0.37 (0.67)
	SEE	0.313	0.321	0.317	0.358
	F-statistic	136.90	123.91	130.07	75.53
	P-value	<0.001	<0.001	<0.001	<0.001
log (chlorate)	y-int	−1.2022	−0.5888	−0.8511	0.967
	Slope	0.4578	0.3431	0.1967	1.1974
	R ²	0.52 (0.96)	0.14 (0.98)	0.14 (0.98)	0.13 (0.92)
	SEE	0.369	0.368	0.368	0.370
	F-statistic	20.31	20.84	20.69	19.21
	P-value	<0.001	<0.001	<0.001	<0.001

^aNote: SEE = Sum of squares of the errors.^bAll reported statistical data pertain to linear regressions drawn through all individual response variable concentrations with the exception of reported R²-values shown in () which pertain to linear regressions drawn through average response variable concentration per water source only.

obtained between $\text{NPOC} \times \text{UV}_{254}$ and the response variables (chlorine dioxide, chlorite and chlorate) when compared to $\text{UV}_{254}/\text{NPOC}$. Further, $\text{NPOC} \times \text{UV}_{254}$ correlations were typically equivalent to those solely involving either NPOC or UV_{254} .

4.2. Model equations

Models for predicting chlorine dioxide consumption, chlorite formation and chlorate formation are presented, respectively, as Eqs. (1)–(3). The significance level of the main parameter and two-factor interaction terms included in each equation are 5%, 5% and 1%, respectively:

$$\begin{aligned}
 &\log(\text{chlorine dioxide consumed, mg/L}) \\
 &= -0.482 + 0.338\log(\text{pH}) - 0.0934\log(\text{Temp.}) \\
 &\quad + 0.455\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad - 0.0288\log(\text{Rxn time} + 1) + 0.162\log(\text{NPOC UV}_{254}) \\
 &\quad + 0.361\log(\text{Temp.})\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad + 0.258\log(\text{ClO}_2 \text{ conc.} + 1)\log(\text{Rxn time} + 1) \\
 &\quad + 0.336\log(\text{NPOC} \times \text{UV}_{254})\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad - 0.114\log(\text{NPOC} \times \text{UV}_{254})\log(\text{Temp.}), \quad (1)
 \end{aligned}$$

$$\begin{aligned}
 &\log(\text{chlorite conc., mg/L}) \\
 &= -0.346 - 0.070\log(\text{pH}) - 0.0253\log(\text{Temp.}) \\
 &\quad - 0.597\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad - 0.136\log(\text{Rxn time} + 1) - 0.0038\log(\text{NPOC UV}_{254}) \\
 &\quad + 0.293\log(\text{Temp.})\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad + 0.393\log(\text{ClO}_2 \text{ conc.} + 1)\log(\text{Rxn time} + 1) \\
 &\quad + 1.27\log(\text{pH})\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad + 0.670\log(\text{NPOC} \times \text{UV}_{254})\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad - 0.161\log(\text{NPOC} \times \text{UV}_{254})\log(\text{Rxn time} + 1), \quad (2)
 \end{aligned}$$

$$\begin{aligned}
 &\log(\text{chlorate conc., mg/L}) \\
 &= -1.99 + 0.621\log(\text{pH}) - 0.090\log(\text{Temp.}) \\
 &\quad + 0.698\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad - 0.104\log(\text{Rxn time} + 1) + 0.0460\log(\text{NPOC} \times \text{UV}_{254}) \\
 &\quad + 0.389\log(\text{Temp.})\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad + 0.346\log(\text{ClO}_2 \text{ conc.} + 1)\log(\text{Rxn time} + 1) \\
 &\quad + 0.486\log(\text{NPOC} \times \text{UV}_{254})\log(\text{ClO}_2 \text{ conc.} + 1) \\
 &\quad - 0.119\log(\text{NPOC} \times \text{UV}_{254})\log(\text{Rxn time} + 1). \quad (3)
 \end{aligned}$$

In accordance with Peixoto [31] and Griepentrog [32], to obtain well-formulated polynomial regression

models, terms which are “hierarchically inferior” (i.e. lower-order) to significant two-factor interaction terms have also been included. For instance, if the “chlorine dioxide dose–temperature” interaction was determined as statistically significant, then the “chlorine dioxide dose” and “temperature” factors were also included in the model, regardless of whether these parameters were found to have significant effects by themselves, on chlorite or chlorate formation or chlorine dioxide consumption. The aforementioned researchers suggest that not doing so may lead to incorrect statistical analysis of the data. Pure error estimates calculated on the basis of duplicated samples for chlorine dioxide consumption, chlorite and chlorate are reported in Table 4. Since model development was based on the log scale, pure error assessments for these variables were also performed using the same basis.

These equations, which are based on surface water data only, were found to reasonably fit the measured levels and displayed some of the highest R^2 -values and lowest sum of squares of the errors (SSE) and standard error of the estimate (SEE) values (Table 5). It was observed (data not shown) that, including results from the groundwater experiment in the model database “drove” the linear regressions downward at the lower $\text{NPOC} \times \text{UV}_{254}$ levels, thereby resulting in steeper regression lines. This provided some indication that the groundwater exhibited high leverage thereby having an important influence on the model regressions. As a result, model equations incorporating groundwater data tended to overpredict at high $\text{NPOC} \times \text{UV}_{254}$ values and underpredict at low. It is unclear whether this situation arose due to the lower concentration, or reactive nature of the organic content associated with the groundwater.

The equations presented in this study differ from those developed by previous researchers that examined chlorination [13,14,16,18], in that they incorporate two-factor interactions. It is proposed that two-factor interaction terms may “refine” parameter effects on DBP formation or disinfectant consumption. For instance, while considering Eq. (2), the effect of temperature on chlorite formation has been captured not only singularly, but also through the temperature–chlorine dioxide concentration interaction. Including two-factor interactions involving $\text{NPOC} \times \text{UV}_{254}$ in the equations aided in improving the model fits. This suggests that these terms will “adjust”, to some degree, the response prediction level in accordance with a site-specific water organic content.

4.3. Internal model validations

For well-formulated equations developed from linear regressions, three main assumptions should be met [33]: the error constituent should be normally distributed, have a constant deviation (constant variation) and should be independently distributed. Based on diagnostic plots of the residual data, the reported equations did not significantly violate any of the three linear regression assumptions. As such, they were assumed to be well-formulated and statistically sound.

Validating the models “internally” with data used in the original database was conducted by following three approaches:

1. Perform a regression of predicted responses against measured levels.

Table 4
Pure error estimates of experimental variation

Compound	No. of duplicated pairs	Pure error ^a estimate	Pure error ^a estimate on a % basis
Chlorine dioxide consumed	15	0.0260 ^b	$10 \pm 0.0260 \approx \pm 6\%$
Chlorite	14	0.0306 ^b	$10 \pm 0.0306 \approx \pm 7\%$
Chlorate	16	0.0296 ^b	$10 \pm 0.0296 \approx \pm 7\%$

^a Pure error reported as standard deviation of the pure error of the measurements.

^b Estimate based on log-transformed data where untransformed data in mg/L.

Table 5
Statistical data associated with reported predictive model equations

Statistical data	Chlorine dioxide consumed model	Chlorite model	Chlorate model
<i>F</i> -statistic; <i>P</i> -value	159.60; <0.001	182.34; <0.001	207.28; <0.001
R^2	0.934	0.948	0.948
Standard error of estimate (SEE)	0.0822	0.0766	0.0902
Sum of squares of the errors (SSE)	0.6886	0.5919	0.8291
Number of trial runs	112	112	112

2. Develop frequency histograms of % recovery distribution (predicted response/measured response).
3. Develop frequency histograms of residual distribution (measured level–predicted level).

Approach 1 was previously applied by Amy et al. [13] and Morrow and Minear [14] while validating their model equations for predicting THM formation and THMFP, respectively, upon chlorination. An analysis of frequency histograms depicting % deviation between predicted (obtained from developed model equations) and measured values has been presented by Greiner et al. [34] with respect to trichloroacetic acid (TCAA) and dichloroacetic acid (DCAA) formation, and by Harrington et al. [8] with respect to organic carbon concentrations.

Internal validation results with respect to the above approaches are reported in Table 6. An example of predicted versus measured chlorite concentrations, based on Eq. (2), is shown in Fig. 3. It is observed that for the chlorine dioxide and chlorite model equations, nearly 80% of the predictions were observed to be within 20% of the measured levels. With respect to chlorate formation, while about 72% of the predicted chlorate levels were within $\pm 20\%$ of the measured values, 99% of the residuals were within 0 ± 0.2 mg/L and 95% were within 0 ± 0.1 mg/L. Comparing these results to those arising from equations (not shown) which incorporated *significant main factors only* (52.4%, 39.8% and 50.0% for chlorine dioxide consumption, chlorite and chlorate formation, respectively) demonstrates that the inclusion of two-factor interaction effects greatly improved model predictions. In addition, the R^2 -values associated with equations incorporating no two-factor interactions for chlorine dioxide (0.871), chlorite (0.814) and chlorate (0.878) were found to be lower than

those associated with model equations including two-factor interactions (Table 6).

On the basis of the performed internal validation approaches, it was concluded that the reported models were complete (in terms of incorporating statistically significant main and two-factor interactions). It is acknowledged that the addition of an increased number of parameters (i.e. heavy metal concentrations, etc.) could ultimately help to explain a greater portion of the variations. However, based on the internal validation results, it is hypothesized that the parameters incorporated into equations presented here do capture a major portion of the variation in the measured levels, resulting in models with very reasonable prediction capabilities.

4.4. External model validations

The final step of model assessment involved performing a model validation using “external” data (data collected outside of the original database). Data published in the literature were limited, whereby level

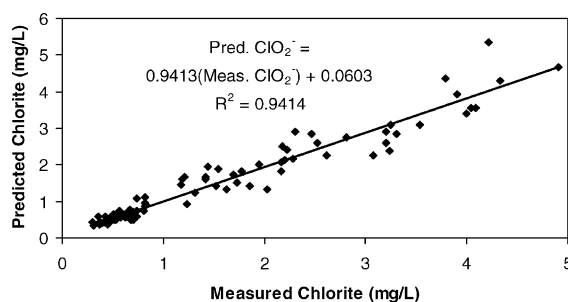


Fig. 3. Predicted and measured chlorite formation levels (Eq. (2), $n = 112$ runs).

Table 6
Internal validation results

Predictive model	Equation for predicted response level (mg/L) ^a	% of original data having a residual ^b within 0 ± 0.2 mg/L	% of original data having a % recovery ^c within $\pm 20\%$ of measured level
ClO ₂ consumed	$= 0.175 + 0.892(\text{MC})$ $R^2 = 0.930$; SEE = 0.3853	70.0	79.5
Chlorite	$= 0.0603 + 0.941(\text{MC})$ $R^2 = 0.941$; SEE = 0.2665	71.4	77.7
Chlorate	$= 0.0133 + 0.90(\text{MC})$ $R^2 = 0.899$; SEE = 0.0476	99.1 (94.6) ^d	72.3

^a Note: For all cases, reported results based on 112 data points.

MC = Measured concentration levels (mg/L), consisting of concentrations measured as part of the factorial trial runs and which served as the original data base for the development of the different models.

^b Residual = Measured level (mg/L) – Predicted level (mg/L).

^c % Recovery = Predicted level (mg/L)/Measured level (mg/L).

^d Values reported in () correspond to % of original database having a 0 to ± 0.1 mg/L residual.

conditions for all required input parameters were typically not reported. To resolve this issue, additional experiments were conducted using water samples that differ from those employed in the initial set of experiments. The “external” water sources used for experimental testing included waters obtained from Michigan (USA), British Columbia, Alberta and Ontario (Canada). Water quality characteristics along with experimental conditions are reported in Table 7.

Model predictions for chlorine dioxide consumption as well as chlorite and chlorate formation fit measured levels well for most cases (Figs. 4–6). While assessing whether measured levels fell within $\pm 20\%$ of the predicted levels was a reasonable means of evaluating chlorine dioxide consumption and chlorite formation, performing such an evaluation with respect to chlorate formation was less effective due to the lower observed concentrations. Nevertheless, differences between measured and predicted chlorate levels were typically small (<0.06 mg/L), indicating good predictive capabilities. It was observed that, for the higher organic carbon-containing waters (Brantford (ON) and Ann Arbor (MI)), the equations displayed a tendency to under-predict chlorine dioxide consumption, and chlorite and chlorate formation. Interestingly, model predictions obtained by extrapolation appeared to be in agreement with measured levels. Eqs. (1)–(3) more than adequately predicted responses for the Vancouver, B.C., water even though it had NPOC and UV_{254} levels which were lower than those observed for the seven surface waters (2.03 mg/L and 0.038/cm, respectively) making up the original database.

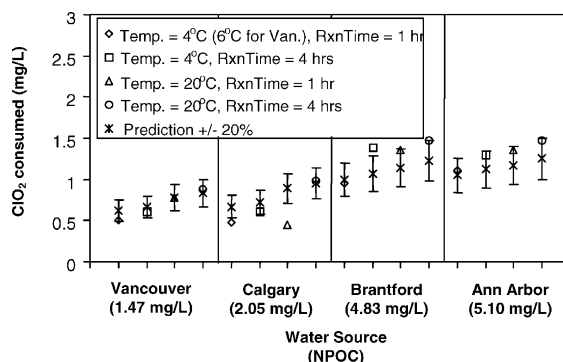


Fig. 4. External validation experimental results (Eq. (1)): comparison of measured and predicted chlorine dioxide consumption levels (chlorine dioxide dose of 1.51 mg/L, ambient pH condition).

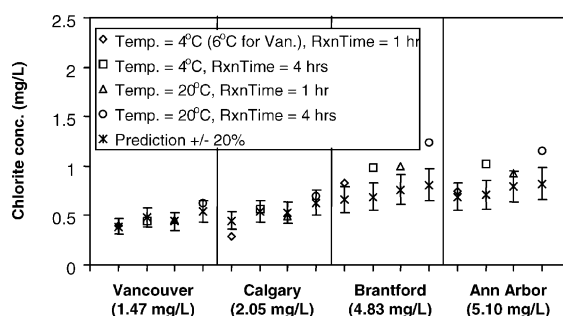


Fig. 5. External validation experimental results (Eq. (2)): comparison of measured and predicted chlorite formation levels (chlorine dioxide dose of 1.51 mg/L, ambient pH condition).

Table 7

Experimental factors and levels investigated during external validation experiments

Water source	Parameter condition			
	pH ^a	Temperature (°C)	ClO ₂ dose (mg/L)	Reaction time (h)
Vancouver (B.C.)	6.57	6	1.51	1
Seymour reservoir	6.57	6	1.51	4
(NPOC = 1.47 mg/L;	6.68	20	1.51	1
UV ₂₅₄ = 0.043/cm)	6.68	20	1.51	4
Calgary (AB)	8.37	4	1.51	1
Glenmore reservoir ^b	8.37	4	1.51	4
(NPOC = 2.05 mg/L;	8.46	20	1.51	1
UV ₂₅₄ = 0.045/cm)	8.46	20	1.51	4
Brantford (ON)	7.79	4	1.51	1
Grand river	7.79	4	1.51	4
(NPOC = 4.83 mg/L;	7.87	20	1.51	1
UV ₂₅₄ = 0.121/cm)	7.87	20	1.51	4
Ann Arbor (MI, USA)	7.78	4	1.51	1
Huron river	7.78	4	1.51	4
(NPOC = 5.10 mg/L	7.70	20	1.51	1
UV ₂₅₄ = 0.141/cm)	7.70	20	1.51	4

^a pH level measured when sample had attained its test temperature.

^b From the Elbow river.

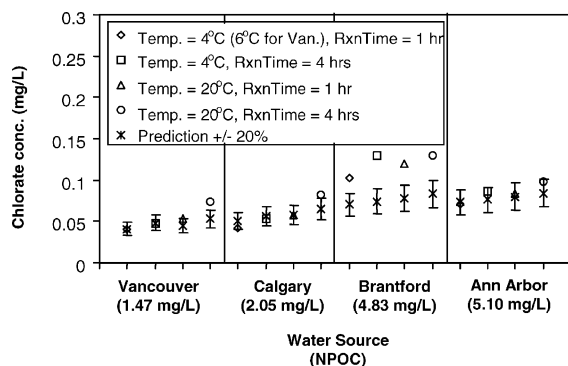


Fig. 6. External validation experimental results (Eq. (3)): comparison of measured and predicted chlorate formation levels (chlorine dioxide dose of 1.51 mg/L, ambient pH condition).

4.5. Sensitivity analyses

Due to the logarithmic form of the predictive equations and the presence of two-factor interactions, a simple incremental parameter increase will result in different predictive response levels depending on the initial selected parameter levels. This makes the evaluation of individual parameter effects on predicted concentrations more involved than when considering model equations based on untransformed data or consisting of main parameter effects only. As such, the percent (%) contribution of each parameter to the “explanatory power” of a given predictive equation has been calculated. The explanatory power of an equation may be defined as the variation in the measured levels (which were part of the original database) explained by the model equations, as determined by the regression sum of squares (SS). All % contributions calculated using Eq. (4) were found to be statistically significant based on calculated *F*-statistics:

$$\% \text{ contribution} = ((A - B)/A) \times 100\%, \quad (4)$$

where *A* is the regression sum of squares (SS) with all terms, and *B* the regression sum of squares (SS) without terms involving a specific factor.

The % contribution of each of the parameters reported in Eq. (2) for predicting chlorite formation is graphically depicted in Fig. 7. As shown, the sum of all the individual parameter contributions (114.7%) is greater than 100%. This is due to the fact that the two-factor interactions have been accounted for twice during the calculations. Based on the results obtained for Eqs. (1)–(3), the parameters have been ordered as to their relative importance in terms of their % contributions to the equation explanatory power.

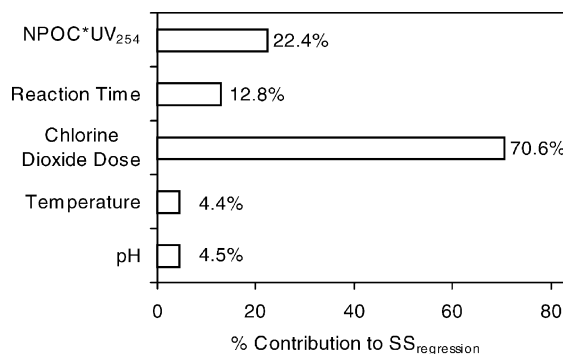


Fig. 7. Chlorite model (Eq. (2)): % contribution of parameter effects (Main and two-factor interactions) to regression SS (Σ contributions = 114.7%).

Chlorine dioxide consumption:

$\text{ClO}_2\text{dose} > \text{NPOC} \times \text{UV}_{254} > \text{Temperature}^\dagger$
 $> \text{Reaction time}^\dagger > \text{pH}.$

Chlorite formation:

$\text{ClO}_2\text{dose} > \text{NPOC} \times \text{UV}_{254} > \text{Reaction time} > \text{pH}^\dagger$
 $> \text{Temperature}^\dagger.$

Chlorate formation:

$\text{ClO}_2\text{dose} > \text{NPOC} \times \text{UV}_{254} > \text{Reaction time}$
 $> \text{Temperature} > \text{pH}.$

Note: \dagger indicates a pair of parameter % contributions within 1% of each other.

Chlorine dioxide concentration and $\text{NPOC} \times \text{UV}_{254}$ exhibited the greatest contributions while pH had the smallest. This implies that to reduce chlorite and chlorate formation, treatment facilities would likely be best served by focusing their efforts in limiting the applied chlorine dioxide dose, as well as by minimizing the organic matter present in the water.

4.6. Chlorite and chlorate formation in terms of chlorine dioxide consumption

Chlorite and chlorate formation were plotted with respect to chlorine dioxide consumption (Figs. 8 and 9, respectively). As shown by the linear correlation coefficients, chlorite and chlorate formation is strongly dependent on chlorine dioxide consumption: on average, 68% and 9% (as described by the regression slope coefficients, neglecting the *y*-intercept values) of the chlorine dioxide consumed was converted to chlorite

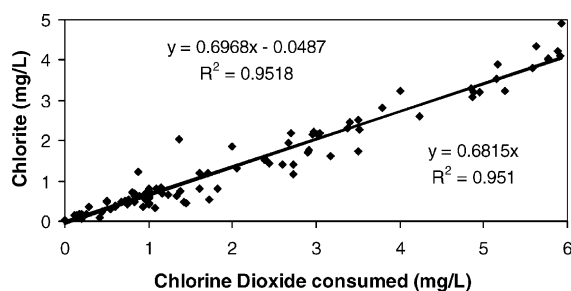


Fig. 8. Chlorite formation with respect to chlorine dioxide consumption for natural test water factorial experiments ($n = 126$ cases).

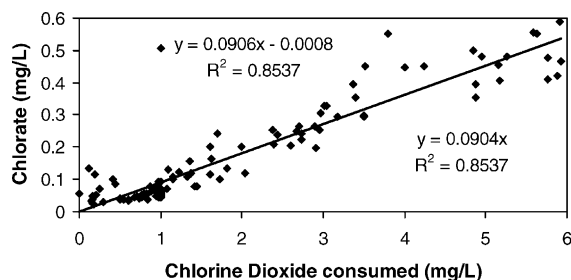


Fig. 9. Chlorate formation with respect to chlorine dioxide consumption for natural test water factorial experiments ($n = 126$ cases).

and chlorate respectively (on a mass basis). These conversion ratios are consistent with those reported in the literature by others [24,5,35].

These 68% and 9% conversion ratios for chlorite and chlorate, respectively, were examined as to whether they differed significantly (on a statistical basis) from conversion ratios calculated from data for a particular water source, or data collected at a specific parameter level (i.e. pH 6 versus 9). This was accomplished by using *F*-statistics [36]. It was observed that, when varying pH, reaction time or water source, statistically significant differences (at the 5% significance level) from the 68% chlorine dioxide:chlorite conversion ratio resulted. For chlorate, statistically significant differences from the 9% conversion ratio were observed when varying chlorine dioxide concentration or water source.

4.7. Considerations and limitations of predictive models

The applicability of the models should ideally be limited to parameter levels within “boundary conditions” as represented by the low and high levels which were investigated for each parameter (Table 1). Resulting predictions based on parameter values beyond this range would be based on extrapolations. It is

recognized that the equations do not include any terms relating to heavy metals or other inorganics, but have focused on physical, chemical and typical operational parameters. Thus, the reader should be cautious while applying the models to waters known to have a composition that may not be representative of typical drinking waters (e.g., high levels of heavy metals). Once again, since the predictive models are log-based and incorporate two-factor interactions, users would be strongly urged to assess “model sensitivity” with respect to chlorine dioxide consumption, or chlorite and chlorate formation, for their own specific treatment conditions.

5. Conclusions

Experiments were conducted using raw waters to develop statistically based model equations capable of predicting chlorite and chlorate formation as well as chlorine dioxide consumption. The following conclusions may be made:

1. The inclusion of two-factor interaction effects in equations improved predictions of chlorine dioxide consumption as well as chlorite and chlorate formation. Inclusion of $\text{NPOC} \times \text{UV}_{254}$ -related two-factor interactions provide an adjustment mechanism to account for site-specific responses.
2. Chlorine dioxide concentration and $\text{NPOC} \times \text{UV}_{254}$ were the most significant parameters in influencing chlorite and chlorate formation as well as chlorine dioxide consumption.
3. Through internal and external model validations, equations based on surface water data only (groundwater data excluded) provided “better-fitting” predictive model equations.
4. The average conversion of chlorine dioxide consumed to chlorite and chlorate (mass basis) was estimated as approximately 68% and 9%, respectively.

Acknowledgements

The authors would like to acknowledge the Natural Sciences and Engineering Research Council (NSERC), Waterloo Centre for Groundwater Research (now CRESTech) and Sterling Pulp Chemicals Ltd. for their financial assistance. Personnel at the water treatment plants who provided raw water samples are also greatly thanked for their cooperation and assistance. The assistance of the Ontario Ministry of Environment and Energy in providing historical raw water data is appreciated. The authors would also like to thank Quanfang Ye, Ron Hofmann and Russell D'Souza of

the University of Toronto, Department of Civil Engineering for providing technical assistance.

References

- [1] Liyanage LRJ, Gyurek LL, Finch GR, Belosevic M. Modeling cryptosporidium parvum inactivation by chlorine dioxide. In: Proceedings of the Forth Environmental Engineering Speciality Conference, Edmonton, AB, May 29–June 1, 1996. p. 17–25.
- [2] Lafrance P, Duchesne D, Arcouette N, Prévost M. The use of ClO_2 : case study of Laval. In: Chlorine dioxide: drinking water issues. Proceedings of the Second International Symposium, American Water Works Association, Houston, TX, 1993. p. 147–62.
- [3] Lykins Jr BW, Griese MH. Using chlorine dioxide for trihalomethane control. *J Am. Water Works Assoc* 1986;78(6):88–93.
- [4] Muttamara S, Sales CI, Gazali Z. The formation of trihalomethane from chemical disinfectants and humic substances in drinking water. *Water Supply* 1995;13(2):105–17.
- [5] Werdehoff KS, Singer PC. Chlorine dioxide effects on THMPF, TOXFP, and the formation of inorganic by-products. *J Am Water Works Assoc* 1987;79(9):107–13.
- [6] Li J-W, Yu Z, Cai X, Gao M, Chao F. Trihalomethanes formation in water treated with chlorine dioxide. *Water Res* 1996;30(10):2371–6.
- [7] Gallagher DL, Hoehn RC, Dietrich AM. Sources, occurrence and control of chlorine dioxide by-product residuals in drinking water. Denver, CO: AWWA Research Foundation, 1994. 184pp.
- [8] Harrington GW, Chowdhury ZK, Owen DM. Developing a computer model to simulate DBP formation during water treatment. *J Am Water Works Assoc* 1992;84(11):78–87.
- [9] Federal Register (Fed. Reg.). National primary drinking water regulations: disinfectants and disinfection by-products; final rule. vol. 63, No. 241, pp. 69389–69476. Environmental Protection Agency, USA, December 164, 1998.
- [10] Pontius FW. New horizons in federal regulation. *J Am Water Works Assoc* 1998;90(3):38–50.
- [11] Pontius FW. Disinfectant-disinfection by-product rule update. *J Am Water Works Assoc* 1991;83(12):24–115.
- [12] Health Canada, Guidelines for Canadian drinking water quality, 6th ed. Ottawa, ON: Canada Communications Group, 1996.
- [13] Amy GL, Chadik PA, Chowdhury ZK. Developing models for predicting trihalomethane formation potential and kinetics. *J Am Water Works Assoc* 1987;79(7):89–97.
- [14] Morrow CM, Minear RA. Use of regression models to link raw water characteristics to trihalomethane concentrations in drinking water. *Water Res* 1987;21(1):41–8.
- [15] Hutton PH, Chung FI. Simulating THM formation potential in sacramento delta, part I. *J Water Resour Plan Manage* 1992;118(5):513–29.
- [16] Lyn TL, and Taylor JS. Modeling compliance of chlorine residual and disinfection by-products. Proceedings of American Water Works Association Water Quality Technology Conference, Miami, FL, November 7–11, 1993. p. 513–23.
- [17] Hofmann R. Minimising chlorination and chloramination DBPs while ensuring microbiological control. M.A. Sci. Thesis. University of Toronto, 1996.
- [18] Amy GL, Minear RA, Cooper WJ. Testing and validation of a multiple nonlinear regression model for predicting trihalomethane formation potential. *Water Res* 1987;21(6):649–59.
- [19] Adin A, Katzhendler J, Alkaslassy D, Rav-Acha Ch. Trihalomethane formation in chlorinated drinking water: a kinetic model. *Water Res* 1991;25(7):797–805.
- [20] Montgomery DC. Design and analysis of experiments. 4th ed. Toronto, ON: Wiley, 1997. 704pp.
- [21] Stat-Ease Inc., Design Ease[®]. Version 3.0.0. Minneapolis, MN, 1994.
- [22] Minitab Inc., The student edition of Minitab[®] for windows software. Release 9. Reading, MA: Addison-Wesley, 1995.
- [23] American Public Health Association, American Water Works Association and Water Environment Federation (APHA-AWWA-WEF). Standard methods for the examination of water and wastewater, 19th ed. Washington, DC, 1995.
- [24] Andrews RC, Hofmann R, Di Scipio R. Drinking water treatment to minimize disinfection by-products and ensure microbiological control. Report, Environmental Health Centre, Ottawa, ON: Health Canada, 1997. 246 pp.
- [25] Masschelein WJ, Fransolet G, Laforge P, Savoie R. Determination of residual ozone or chlorine dioxide in water with ACVK—an updated version. *Ozone Sci Eng* 1989;11:209–15.
- [26] Hofmann R, Andrews RC, Ye Q. Comparison of spectrophotometric methods for measuring chlorine dioxide in drinking water. *Environ Technol* 1998; 19:761–73.
- [27] United States Environmental Protection Agency (USEPA). Determination of inorganic anions by ion chromatography, Revision 2.1. In: Methods for the determination of inorganic substances in environmental samples, EPA/600/R/93/100, 1993.
- [28] Hautman DP, Bolyard M. Using ion chromatography to analyze inorganic disinfection by-products. *J Am Water Works Assoc* 1992;84(11):88–93.
- [29] Engerholm BA, Amy GL. A predictive model for chloroform formation from humic acid. *J Am Water Works Assoc* 1983;75(8):418–23.
- [30] Engerholm BA, Amy GL. An empirical model for predicting chloroform formation from humic and fulvic acids. In: Jolley RL, Brungs WA, Cotruvo JA, Cumming RB, Mattice JS, Jacobs VA, editors. Water chlorination: environmental impact and health effects, Vol. 4, Bk. 1. Chelsea, MI: Ann Arbor Science Publishers, 1983. p. 243–52.
- [31] Peixoto JL. A property of well-formulated polynomial regression models. *The Am Stat* 1990;44(1):26–30.
- [32] Griepentrog GL, Ryan JM, Smith LD. Linear transformations of polynomial regression models. *The Am Stat*, Part 1 1982;36(3):171–4.

- [33] McKenzie J, Schaefer RL, Farber E. The student edition of MINITAB[®] for windows users manual. Reading, MA: Addison-Wesley, 1995.
- [34] Greiner AD, Obolensky A, Singer PC. Technical note: comparing predicted and observed concentrations of DBPs. *J Am Water Works Assoc* 1992; 84(11):99–102.
- [35] Aieta EM, Roberts PV, Hernandez M. Determination of chlorine dioxide, chlorine, chlorite, and chlorate in water. *J Am Water Works Assoc* 1984;76(1):64–70.
- [36] Myers RH. Classical and modern regression with applications, 2nd ed. In: The Duxbury advanced series in statistics and decision sciences. Boston, MA: PWS-KENT 1990, 488 pp.