

## Developing a model for disinfection by-products based on multiple regression analysis in a water distribution system

H. Shimazu, M. Kouchi, Y. Sugita, Y. Yonekura, H. Kumano, K. Hashiwata, T. Hirota, N. Ozaki and T. Fukushima

### ABSTRACT

The characteristics of disinfection by-product (DBP) formation in the distribution system of Hiroshima City were investigated. Then a model was developed for predicting DBPs based on multiple regression analysis. Potassium permanganate consumption value ( $\text{KMnO}_4$ ), free chlorine residual concentration ( $\text{Cl}_2$ ), pH, water temperature (WT) and residence time (RT) were confirmed to have an influence on DBP formation in the distribution system and used as independent variables.

The correlation coefficients between the logarithmic values of the predicted DBPs and those of the observed DBPs were higher than 0.7. In addition, the seasonal change patterns of the predicted DBPs were very similar to those observed in some DBPs. The model is helpful for managing the DBPs in water distribution systems.

**Key words** | DBP quality modelling, disinfection by-products (DBPs), water distribution system

**H. Shimazu** (corresponding author)  
School of Science & Engineering,  
Kinki University,  
3-4-1 Kowakae, Higashiosaka-shi 577-8502,  
Japan  
Tel: +81-6-6730-5880(Ext.4275)  
Fax: +81-6-6730-1320  
E-mail: hshimazu@civileng.kindai.ac.jp

**M. Kouchi**  
**Y. Sugita**  
**Y. Yonekura**  
**H. Kumano**  
**K. Hashiwata**  
**T. Hirota**

Water Quality Control Division,  
Hiroshima Waterworks Bureau,  
6-1-1 Ochiaiminami, Asakita-ku,  
Hiroshima-shi 739-1732,  
Japan

**N. Ozaki**  
Faculty of Engineering,  
University of Hiroshima,  
1-4-1 Kagamiyama,  
Higashihiroshima-shi 739-8527,  
Japan

**T. Fukushima**  
Graduate School of Life and Environmental  
Sciences,  
University of Tsukuba,  
1-1-1 Tennoudai,  
Tsukuba-shi 305-8572,  
Japan

### INTRODUCTION

Almost all municipal water supply systems in Japan use a form of chlorine for drinking water disinfection. Natural organic substances can react with the added chlorine and hydroxide ions in the distribution systems to form disinfection by-products (DBPs), such as trihalomethanes and haloacetic acids. Some DBPs are of public health concern because of their suspected carcinogenicity as well as developmental, reproductive and hepatic toxicity (Hasegawa *et al.* 1999; Pontius 2001). Therefore, it is important to understand the mechanisms of the DBP formation and to lower DBP concentrations in distribution systems.

In order to take proper countermeasures to lower DBPs, it is very important to develop models for predicting DBPs in distribution systems. Many kinds of study have indicated dominant factors and developed models on DBP formation using experiments in laboratories (Urano *et al.* 1983; Reckhow *et al.* 1990; Adin *et al.* 1991). In addition, some models simulating DBP changes in real distribution systems have been proposed during the last several years (El-shafy & Grunwald 2000; Sung *et al.* 2000; Westerhoff *et al.* 2000). However, the precision of the models in field studies is slightly less than that of the model in laboratory

studies. Higher precision in the models in field studies is required to control DBPs.

In the present study, the DBP concentrations were measured at six locations in the north Hiroshima water distribution system during a 1-year period. Previous studies have shown the importance of many parameters for DBP formation, such as type and amount of organic substances, amount of chlorine residual, pH, water temperature and residence time in the distribution system. The relationship between the parameters as stated above and DBPs in the distribution system are examined in this study. Statistical models based on multiple regression analysis that input the parameters are then developed and validated with the monitoring data.

## METHODS

### Distribution system and sampling locations

The pipelines under study are supplied with drinking water from the Koyo purification plant in the city of Hiroshima, Japan. The plant, which has a capacity of  $200,000 \text{ m}^3 \text{ day}^{-1}$ , takes its raw water from the Ota River. The purification processes in this plant comprise coagulation with aluminium sulfate, flocculation, sedimentation, disinfection and rapid filtration. Gaseous chlorine is used as a residual disinfectant. Caustic soda solution is added after rapid filtration to adjust pH.

This study focused on the pipelines that transport drinking water from the plant to north Hiroshima because this pipeline network was relatively simple. The total length of these pipelines, which were almost all constructed of cast iron, was about 30 km and the pipe diameter ranged from 150 to 1,200 mm.

The samples for this study were taken at six locations along the pipeline once a month from April 2000, to March 2001 (Figure 1). The first and second locations were at the effluent of the rapid filtration basin and distribution reservoir in the Koyo purification plant (locations A and B), respectively, the third and fourth locations were at the effluent of the Karuga first reservoir and the Mita reservoir (locations C and D), respectively, and the fifth and sixth locations were at the Fureai Park and the Shiraki fire department office (locations E and F), respectively.

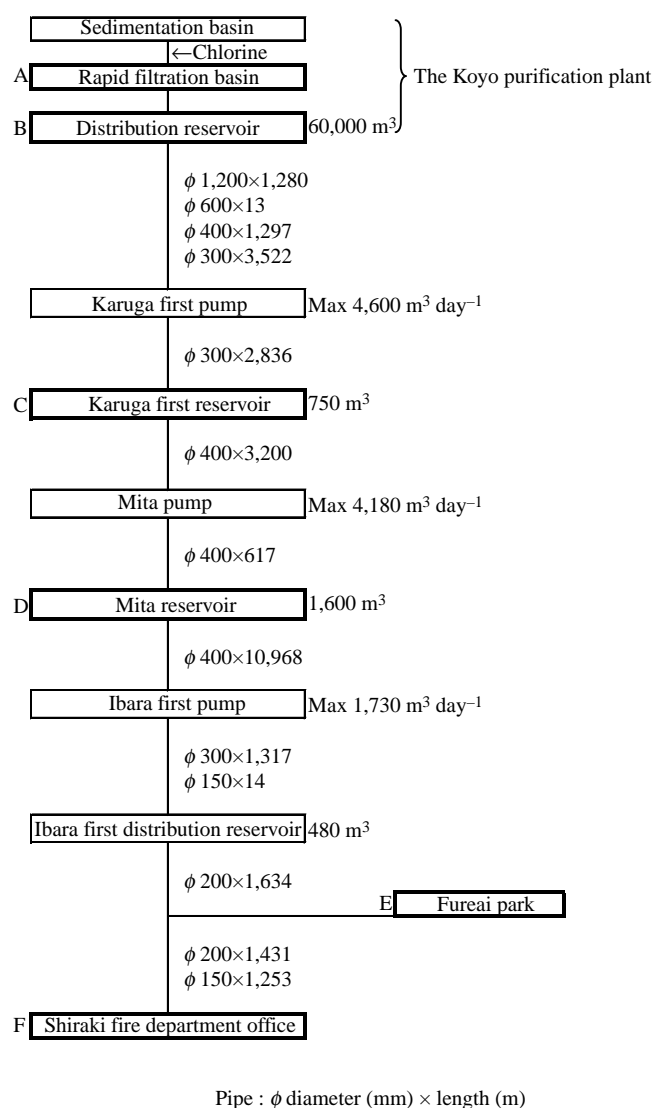


Figure 1 | Sample locations along the pipelines from the Koyo purification plant.

The estimated residence time between the first chlorine injecting point and location F was about 5 days under normal operating conditions. The residence times were estimated by the sum of the volume of the rapid filtration basins, distribution reservoirs and pipelines divided by the monthly mean flow rates provided by the Koyo purification plant and pumps.

### Water analysis

In this study, water temperature and free chlorine residual were measured at the time when the samples were

collected. Free chlorine residual was determined spectrophotometrically with a Hach pocket colorimeter analysis system. pH measurements were carried out with a Toa HM-60G pH meter. Potassium permanganate consumption value was measured by the Japanese Standard Method for the Examination of Drinking Water (Japan Water Works Association 1993).

The DBPs measured in this study were bromoform (BF), bromodichloromethane (BDCM), chloroform (CF), chloral hydrate (CH), dibromochloromethane (DBCM), dichloroacetic acid (DCA), dichloroacetonitrile (DCAN) and trichloroacetic acid (TCA). Samples for trihalomethanes (BF, BDCM, CF and DBCM) analysis were taken from the headspace gas of 22 ml glass vials with PTFE butyl septa and aluminium caps after acidification and removal of the chlorine residual. Trihalomethanes were determined with the headspace technique on a Shimadzu QP-5000 gas chromatograph with mass spectrometer detection (GC/MS). Headspace vials (22 ml) were filled with a 10 ml sample and then were heated for 30 min at 60 °C before the headspace gas was withdrawn and injected into GC/MS. Then 50 ml aliquots of acidified (pH) and salt-containing samples (unfiltered) for DCA and TCA analysis were extracted with 4 ml methyl *tert*-butyl ether (MTBE) after removal of the chlorine residual. The ether layer was separated and dehydrated with Na<sub>2</sub>SO<sub>4</sub>. The samples were analysed on a Hewlett-Packard HP-5890 Series II gas chromatograph with electron capture detector after derivatization with diazomethane. For CH and DCAN analysis, 20 ml aliquots of salt-containing samples (unfiltered) were extracted with 2 ml MTBE after removal of the chlorine residual. The ether layer was separated and dehydrated with Na<sub>2</sub>SO<sub>4</sub>. The samples were then analysed on a Hewlett-Packard HP-6890 gas chromatograph with electron capture detector.

### Multiple regression analysis

Many studies have suggested that type and amount of organic substance, water temperature (WT), pH and chlorine residual (Cl<sub>2</sub>) affected DBP formation (Urano *et al.* 1983; Sugino *et al.* 1986; Miyata *et al.* 1996).

DBPs are produced when some organic substances react with chlorine residual and hydroxide ion in

distribution systems. Frequently, the reaction of DBP formation is characterized as follows:



where OS is the reacting organic substance and a–c are the reaction constants. Then the reaction rate is given by

$$d(\text{DBPs})/dt = k \times [\text{OS}]^a \times [\text{Cl}_2]^b \times [\text{OH}^-]^c \quad (2)$$

Here k is the reaction rate constant that is estimated to be a function of WT. Therefore, OS, Cl<sub>2</sub>, OH<sup>−</sup>, WT and residence time (RT) have a significant effect on the DBP formation. In this analysis, potassium permanganate consumption value (KMnO<sub>4</sub>), free chlorine residual (Cl<sub>2</sub>), pH, WT and RT are chosen as the parameters of multiple regression analysis. The reason KMnO<sub>4</sub> is chosen as a surrogate for amount of organic substances is that this parameter has been used in Japanese drinking water criteria.

Hence, the DBP prediction model in the distribution system was developed as follows:

$$\log(\text{DBPs}) = a + b \times \log(\text{KMnO}_4) + c \times \log(\text{Cl}_2) + d \times \text{pH} + e \times \log \text{WT} + f \times \log \text{RT} \quad (3)$$

where a–f are the constants. Although the data on KMnO<sub>4</sub> and Cl<sub>2</sub> were used at the outlet of the sedimentation basin in the Koyo purification plant, pH and WT data were used at each location for model development. If DBP formation started with the chlorine injection, these data should be used at the outlet of the sedimentation basin. However, pH and WT data were used at each location because these values fluctuate in the distribution system.

The data above detection limits were used for model development. The models were developed for predicting seven DBPs except for bromoform. The reason bromoform was not predicted was that more than half the samples were below the detection limits.

## DBP FORMATION IN THE DISTRIBUTION SYSTEM

### DBPs in the distribution system

Measured DBP concentrations in the distribution system are summarized in Table 1. The mean concentrations of bromoform (BF), dichloroacetonitrile (DCAN) and chloral hydrate (CH) were calculated after removing data below the detection limit. Other DBPs were detected at all locations

**Table 1** | Measured DBP concentrations in the distribution system ( $\mu\text{g l}^{-1}$ )

	<b>Standard/ Guideline in Japan</b>	<b>Detection limit</b>	<b>A(av. 1.4 h)</b>	<b>B(av. 10 h)</b>	<b>C(av. 27 h)</b>	<b>D(av. 57 h)</b>	<b>E(av. 109 h)</b>	<b>F(av. 113 h)</b>	<b>Mean<sup>*1</sup></b>	<b>Detection rate<sup>*2</sup></b>
Bromodichloromethane	30	0.3	2.2 (1.2 ~ 4.0) <sup>*3</sup>	3.9 (2.3 ~ 6.6)	5.1 (2.8 ~ 9.0)	5.6 (3.3 ~ 8.6)	6.7 (3.7 ~ 9.8)	6.7 (3.6 ~ 10.0)	5.0	100 (72/72)
Bromoform	90	0.3	<0.3	0.33 (<0.3 ~ 0.4)	0.33 (<0.3 ~ 0.4)	0.33 (<0.3 ~ 0.4)	0.34 (<0.3 ~ 0.4)	0.34 (<0.3 ~ 0.4)	0.33	46 (33/72)
Chloral hydrate	30 (provisional)	0.6	1.5 (<0.6 ~ 3.1)	2.3 (<0.6 ~ 4.0)	3.2 (0.7 ~ 6.6)	3.6 (0.9 ~ 5.8)	4.9 (1.4 ~ 8.8)	5.2 (1.5 ~ 9.8)	3.4	94 (68/72)
Chloroform	60	0.3	2.4 (1.0 ~ 5.8)	4.3 (2.0 ~ 10.6)	5.9 (2.3 ~ 15.0)	6.6 (2.8 ~ 12.0)	8.9 (2.9 ~ 15.0)	9.3 (2.9 ~ 16.7)	6.2	100 (72/72)
Dibromochloromethane	100	0.3	1.1 (0.7 ~ 1.9)	2.0 (1.4 ~ 3.3)	2.4 (1.8 ~ 3.9)	2.8 (1.7 ~ 4.2)	3.2 (2.5 ~ 4.4)	3.2 (2.3 ~ 4.6)	2.5	100 (72/72)
Dichloroacetic acid	40	0.9	5.7 (3.3 ~ 11.3)	7.1 (4.0 ~ 13.8)	9.4 (4.6 ~ 20.6)	9.8 (5.5 ~ 13.2)	11.0 (6.5 ~ 14.8)	11.0 (6.7 ~ 13.8)	9.0	100 (72/72)
Dichloroacetonitrile	40 (provisional)	0.6	0.81 (<0.6 ~ 1.4)	0.99 (<0.6 ~ 1.9)	1.2 (<0.6 ~ 2.4)	1.2 (<0.6 ~ 1.8)	1.4 (0.6 ~ 2.1)	1.4 (0.7 ~ 1.8)	1.2	88 (63/72)
Trichloroacetic acid	200	0.9	4.0 (2.7 ~ 6.3)	5.2 (3.6 ~ 8.4)	6.6 (4.0 ~ 11.1)	7.4 (4.7 ~ 9.8)	8.7 (4.8 ~ 13.0)	9.0 (5.0 ~ 13.4)	6.8	100 (72/72)

<sup>\*1</sup>Mean concentrations were calculated after removing data below the detection limits.

<sup>\*2</sup>Detection rate [%] (detected frequency / measured frequency).

<sup>\*3</sup>Mean concentration (minimum concentration ~ maximum concentration).

on every measurement. The means of mole ratio for DBPs were 30.6% for dichloroacetic acid, 19.4% for chloroform, 18.4% for trichloroacetic acid, 12.9% for bromodichloromethane, 7.1% for CH, 5.1% for dibromochloromethane, 4.1% for DCAN and 0.43% for BF, respectively.

Previous studies have reported that DBPs would be produced from similar precursors (Rook 1977, 1980; Ueno *et al.* 1996). The correlation coefficients among DBPs, except BF, are summarized in Table 2. These coefficients were higher than 0.6 ( $p < 0.10$ ), showing that each DBP was strongly related to the other DBPs. This probably indicates that the DBPs detected in the present study were produced from similar precursors. Increases in DBP concentrations have been reported with increasing residence time in the distribution system in both laboratory studies and field surveys (Urano *et al.* 1983; Sugino *et al.* 1986; Kamei *et al.* 1994; Gallard & Gunten 2001). This tendency was also confirmed in this study, i.e. the DBP concentrations at location F (94.0–144.8 h) were two to four times higher than those at location A (1.4–1.7 h). In addition, it was observed that the DBP concentrations from June to August, when water temperature was higher than 20 °C, were much higher than those in other months, showing a tendency similar to previous studies (Chen & Weisel 1998; Rodriguez & Serodess 2001).

### Factors influencing DBP formation

The mean concentrations and the ranges between maximum concentrations and minimum concentrations

for potassium permanganate consumption value ( $\text{KMnO}_4$ ) and free chlorine residual ( $\text{Cl}_2$ ) at the outlet of the sedimentation basin were  $1.7 \text{ mg l}^{-1}$  ( $1.4\text{--}2.2 \text{ mg l}^{-1}$ ) and  $1.5 \text{ mg l}^{-1}$  ( $1.2\text{--}2.0 \text{ mg l}^{-1}$ ), respectively. Those of pH, water temperature (WT) and residence time (RT) in all measured data were 7.3 (6.5–8.4), 16.1 °C (6.1–26.8 °C) and 52.9 h (1.35–145 h), respectively. The correlation coefficient between  $\log(\text{Cl}_2)$  and  $\log(\text{WT})$ , the correlation coefficient between pH and  $\log(\text{RT})$  and the correlation coefficient between  $\log(\text{KMnO}_4)$  and  $\log(\text{WT})$  were 0.82, 0.82 and  $-0.49$  ( $p < 0.05$ ), respectively. The relationship between  $\text{Cl}_2$  and WT was positive because much chlorine was added at the outlet of the sedimentation basin so as not to be diminished at the end of the distribution network in summer when temperature was high. The relationship between pH and RT was positive because pH values at the end of this distribution pipeline tended to be high due possibly to the material quality of the pipes. The relationship between  $\text{KMnO}_4$  and WT was negative because the coagulation rate of organic substances in raw water was so fast in summer that organic substances might be coagulated very quickly.

### Relationship between DBPs and factors influencing DBP formation

The correlation coefficients between DBP concentrations and water temperature (WT) and those between DBP concentrations and pH were in almost all cases above 0.7, except for bromoform. DBP concentrations increased with

**Table 2** | Correlation coefficients among DBPs

	BDCM	CH	CF	DBCM	DCA	DCAN	TCA
Bromodichloromethane (BDCM)	1.00						
Chloral hydrate (CH)	0.93	1.00					
Chloroform (CF)	0.96	0.93	1.00				
Dibromochloromethane (DBCM)	0.87	0.78	0.73	1.00			
Dichloroacetic acid (DCA)	0.91	0.87	0.90	0.66	1.00		
Dichloroacetonitrile (DCAN)	0.86	0.85	0.87	0.62	0.92	1.00	
Trichloroacetic acid (TCA)	0.94	0.92	0.94	0.77	0.93	0.84	1.00

WT and pH at all locations. The correlation coefficients between DBP concentration and pH at location F were lower than those at the other locations, probably because the interior condition of the pipeline might also influence pH.

The correlation coefficients between DBP concentrations and chlorine residual ( $\text{Cl}_2$ ) concentrations and those between DBP concentrations and potassium permanganate consumption values ( $\text{KMnO}_4$ ) were slightly lower than those between DBP concentrations and WT and those between DBP concentrations and pH.  $\text{Cl}_2$  concentrations have an influence on all DBP formation only during a short residence time (less than about 10 h) shown in Table 3. The DBP concentration increased with  $\text{KMnO}_4$  at all locations, which is similar to the cases of water temperature and pH.

The DBP formation pattern indicated that the DBP formation rate accelerated in the first phase and that the rate slowed gradually and reached equilibrium in the last phase (Urano *et al.* 1983; Gallard & Gunten 2001). Figure 2 shows haloacetic acids formation patterns during a 1-year period of this study. Although June and December did not show the pattern as previously described, almost all months indicated the pattern shown in previous studies.

## MODEL FOR DBPS IN DISTRIBUTION SYSTEM

### Model results

A DBP prediction model in the distribution system was developed by the compulsory input method of multiple

regression analysis. The model's parameters were potassium permanganate consumption value ( $\text{KMnO}_4$ ), chlorine residual ( $\text{Cl}_2$ ), pH, water temperature (WT) and residence time (RT). The data measured in June at the Karuga distribution reservoir and in December at the distribution basin were removed because the DBP formation patterns were different from the others shown in Figure 2.

$R^2$  values,  $R^{2*}$  values and SE for DBP prediction models are shown in Table 4. The  $R^2$  values were more than 0.74. In particular, those for DCA, TCA and CF, whose concentrations were higher than the other DBPs, were more than 0.9. In addition, there was little difference between  $R^2$  values and  $R^{2*}$  values. The reason the  $R^2$  value for DCAN was lower than the others was considered to be that almost all measured DCAN data were around the detection limit.

The relationships between the logarithmic values of the predicted DBPs and those of the observed DBPs are shown in Figure 3. Although the absolute values of the differences between them exceeded 0.2 once in the case of chloroform, once in the case of dichloroacetonitrile and twice in the case of chloral hydrate, the means of the differences between them in all predicted DBPs were within 0.1. In addition, the medians of the ratio (predicted DBPs vs. observed DBPs) were approximately one. The model is fairly useful for the prediction at various locations for which the RT values are very different.

Next, the present model is compared with two previous models (Urano *et al.* 1983; Amy *et al.* 1987). The parameters in the previous models are  $\text{Cl}_2$ , pH, WT and RT. Total

**Table 3** | The correlation coefficients between DBP concentrations and chlorine residual concentrations

	A(av. 1.4 h)	B(av. 10 h)	C(av. 27 h)	D(av. 57 h)	E(av. 109 h)	F(av. 113 h)
Bromodichloromethane (BDCM)	0.74	0.76	0.02	0.23	0.08	0.04
Chloral hydrate (CH)	0.56	0.67	0.08	0.29	0.08	0.08
Chloroform (CF)	0.57	0.59	0.21	0.18	0.015	0.05
Dibromochloromethane (DBCM)	0.82	0.61	0.46	0.30	0.07	0.14
Dichloroacetic acid (DCA)	0.64	0.62	0.20	0.22	0.22	0.10
Dichloroacetonitrile (DCAN)	0.49	0.43	0.22	0.02	0.16	0.11
Trichloroacetic acid (TCA)	0.74	0.70	0.01	0.29	0.36	0.12

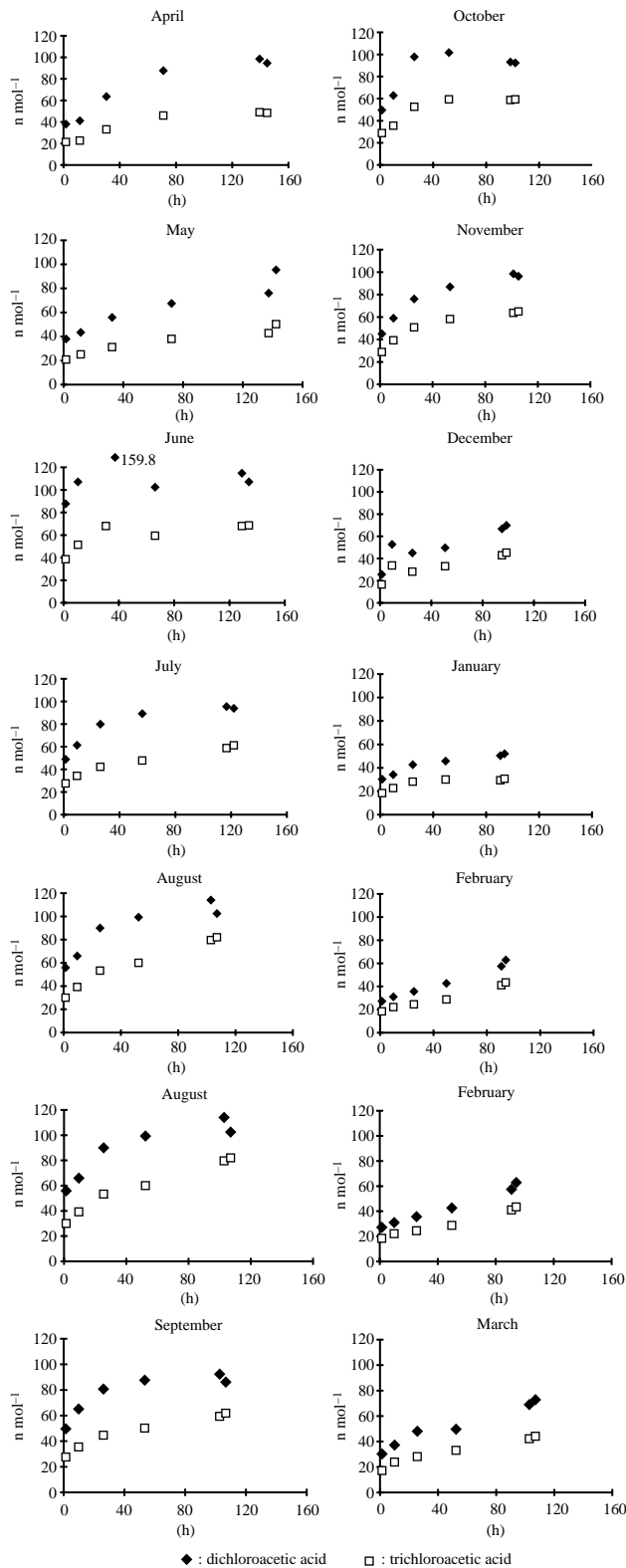


Figure 2 | DBP formation pattern in the distribution system.

Table 4 | Model statistical summary

	N	R <sup>2</sup>	R <sup>2</sup> *	SE
Bromodichloromethane (BDCM)	70	0.97	0.96	0.044
Chloral hydrate (CH)	66	0.92	0.91	0.090
Chloroform (CF)	70	0.94	0.93	0.082
Dibromochloromethane (DBCM)	70	0.87	0.86	0.077
Dichloroacetic acid (DCA)	70	0.91	0.90	0.056
Dichloroacetonitrile (DCAN)	61	0.74	0.72	0.085
Trichloroacetic acid (TCA)	70	0.93	0.92	0.049

N : the number of data for the model development.

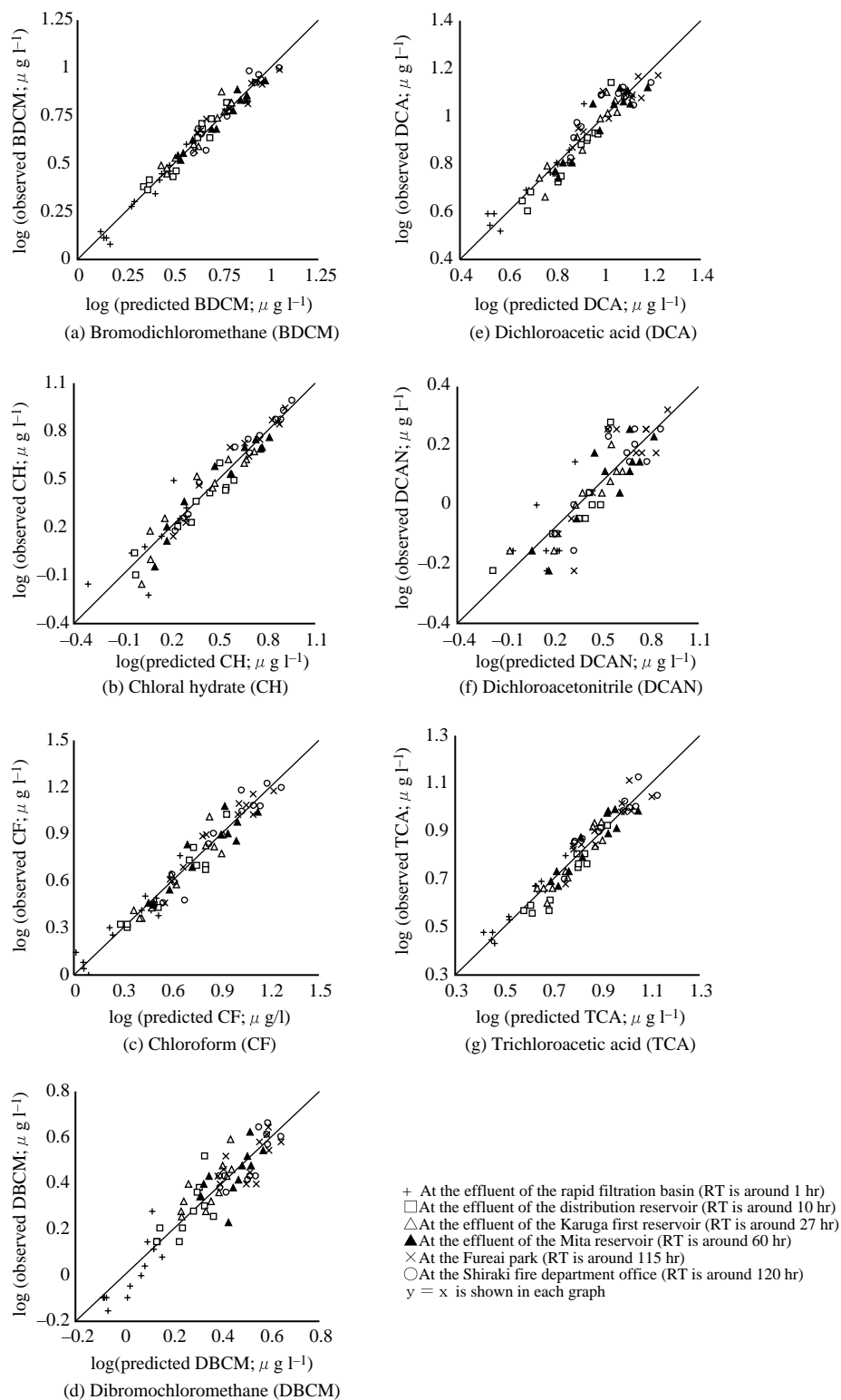
R<sup>2</sup> : R-squared.

R<sup>2</sup>\* : adjusted R-squared.

SE : standard error.

organic carbon (TOC) in the Urano model and the product of nonpurgeable organic carbon and UV absorbance in the Amy model are chosen as a surrogate of organic substances. In addition, the concentration of bromide ion is chosen in the Amy model. Here, the predicted THM (the sum of concentrations for bromoform, bromodichloromethane, chloroform and dibromochloromethane) were calculated assuming that  $\text{KMnO}_4$  data in this study were substituted for TOC or the product of nonpurgeable organic carbon and UV absorbance, and the concentration of bromide ion was zero. The R<sup>2</sup> values were 0.89 in the Urano model, 0.91 in the Amy model and 0.94 in our model. The precision of our model was the same as those of the previous models developed based on laboratory studies.

The partial regression coefficients are shown in Table 5. The coefficients for  $\text{Cl}_2$ , WT and RT in all DBPs were positive as expected. RT has a strong influence on DBP formation because the standard partial regression coefficients were the highest among these parameters. The secondary affecting parameter differed from one DBP to another. The secondary affecting parameter in the case of DCAN was WT and that in the case of other DBPs was  $\text{Cl}_2$ . However,  $\text{KMnO}_4$  and pH have weak effects on DBP formation in the distribution system. The reason  $\text{KMnO}_4$  has a weak influence on DBP formation is considered to be



**Figure 3** | Relationship between the predicted DBPs and the observed DBPs.



**Table 5** | Model coefficients of variables

Name	Variables					
	(Constant)	log(KMnO <sub>4</sub> )	log(Cl <sub>2</sub> )	pH	log(WT)	log(RT)
log(Bromodichloromethane)	-0.169	0.0406 (0.012)	1.31* (0.414)	-0.0102 (-0.016)	0.267* (0.239)	0.254* (0.749)
log(Chloral hydrate)	-1.09*	-0.873* (-0.189)	1.58* (0.373)	0.0570 (0.065)	0.535* (0.344)	0.288* (0.600)
log(Chloroform)	-1.31*	0.105 (0.023)	1.80* (0.419)	0.111* (0.125)	0.455* (0.299)	0.245* (0.531)
log(Dibromochloromethane)	-0.191	-0.0943 (-0.031)	0.786* (0.278)	-0.00954 (-0.016)	0.123 (0.123)	0.256* (0.846)
log(Dichloroacetic acid)	0.727*	0.0374 (0.014)	1.12* (0.465)	-0.0759* (-0.152)	0.238* (0.279)	0.194* (0.751)
log(Dichloroacetonitrile)	-0.532	0.325 (0.134)	0.532 (0.236)	-0.0624 (-0.141)	0.484* (0.544)	0.188* (0.748)
log(Trichloroacetic acid)	-0.245	0.186 (0.073)	0.908* (0.381)	0.0503 (0.102)	0.225* (0.267)	0.159* (0.622)

1) The data of KMnO<sub>4</sub>(mg l<sup>-1</sup>/mg l<sup>-1</sup>) and Cl<sub>2</sub>(mg l<sup>-1</sup>/mg l<sup>-1</sup>) were used at the effluent of the sedimentation basin.

2) The data of pH and WT(water temperature;°C) were used at each sample location.

3) RT(residence time; h).

4) The upper values show the partial regression coefficients (\*; p < 0.05) and the lower values in parentheses show the standard partial regression coefficients.

that  $\text{KMnO}_4$  includes not only DBP precursors but also other organic substances and inorganic substances. The reason pH has a weak influence on DBP formation might be that the purified water has a little buffer ability because free carbon dioxide and sodium are present at around  $2.0 \text{ mg l}^{-1}$  and  $10 \text{ mg l}^{-1}$ , respectively.

### Model validation

We investigate whether the above models are appropriate for predicting DBPs using data obtained from location G, where the Koyo purification plant supplies water, but the pipeline network is different from locations C–F. The estimation of residence time between the first chlorine injecting point and location G was about 15 h under normal operating conditions. The measured data were obtained every month from April 1999 to March 2001 and every three months from May 2001 to November 2002. Then the model for dichloroacetonitrile was not validated because almost all data were around the detection limit. For validation of the model, we used  $\text{KMnO}_4$  data at the effluent of the sedimentation basin by means of weekly laboratory tests and the daily mean  $\text{Cl}_2$  data at the effluent of the sedimentation basin by means of an automatic polarograph measured just one day before the measurement at the G location.

Figure 4 shows the relationships between the predicted and observed DBPs and their seasonal changes. In the cases of chloroform, dichloroacetic acid and trichloroacetic acid whose concentrations were higher than the other DBPs, the correlation coefficients were approximately 0.9 and the patterns of seasonal change in the predicted DBPs were very similar to those of the observed DBPs. In the case of chloral hydrate, the correlation coefficient was 0.74 and the pattern of seasonal change in the predicted was very similar to that of the observed. In the case of bromodichloromethane, the correlation coefficient was 0.8. However, the predicted concentrations were lower than the observed concentrations. In the case of dibromochloromethane, the prediction is not good because the observed concentrations were lower than other DBPs. Although the results differ among DBPs, this model is thought to be appropriate for predicting DBPs in distribution systems.

### Factors influencing model precision

There were four chemicals for which the ratios of the predicted to the observed concentrations were beyond 0.8 in April and August 1999 and May 2001 as shown in Figure 4. In this section, the reason is investigated.

We investigated how the measurement errors or temporal changes of potassium permanganate consumption value ( $\text{KMnO}_4$ ), free chlorine residual ( $\text{Cl}_2$ ), pH, water temperature (WT) and residence time (RT) affect the predicted DBPs. Table 6 summarizes the model precision which resulted from their measurement errors or their temporal changes. The variation ranges of these parameters were determined as follows.

In the case of  $\text{KMnO}_4$ , the measurement error was estimated at  $0.2 \text{ mg l}^{-1}$  because there was considered to be a 0.05 ml reading error. In the case of  $\text{Cl}_2$ , the error was estimated to be around 12% of the  $\text{Cl}_2$  median because the coefficient of variation (CV) of every hour of  $\text{Cl}_2$  data at the outlet of the sedimentation basin was about 6%. In the case of pH, the measurement error was considered to be 0.002 because the precision of the pH analyser was 0.001. In the case of WT, the error was considered to be  $0.2^\circ\text{C}$  because the reading error of the bar type thermometer, whose minimum division was  $0.5^\circ\text{C}$ , was considered to be  $0.1^\circ\text{C}$ . In the case of RT, temporal change of RT was calculated to be approximately 40% because the CV of every hour of distribution volume from the purification plant was around 20%.

Table 6 indicates that the most influential parameter on the precision of DBP prediction is  $\text{Cl}_2$ . The difference between the prediction using  $1.5 \text{ mg l}^{-1}$  of  $\text{Cl}_2$  and the prediction using  $1.7 \text{ mg l}^{-1}$  of  $\text{Cl}_2$  was around 0.05 orders of magnitude (when DBP concentration is  $5 \text{ mg l}^{-1}$ , the difference is calculated to be approximately  $0.6 \text{ mg l}^{-1}$ ). It possibly lowers the precision of DBP prediction using the daily mean  $\text{Cl}_2$  value as the variable in our model.  $\text{Cl}_2$  concentrations at the effluent of the sedimentation basin fluctuate due to the on-site control. The DBPs concentrations in April and August 1999 and May 2001 were therefore predicted using daily mean  $\text{Cl}_2$  data plus  $0.2 \text{ mg l}^{-1}$ . The reason the DBPs concentrations were predicted using daily mean  $\text{Cl}_2$  data plus  $0.2 \text{ mg l}^{-1}$  was that the error of daily mean  $\text{Cl}_2$  data was estimated to be around  $0.2 \text{ mg l}^{-1}$  (the product of  $1.5 \text{ mg l}^{-1}$  and 12%). The

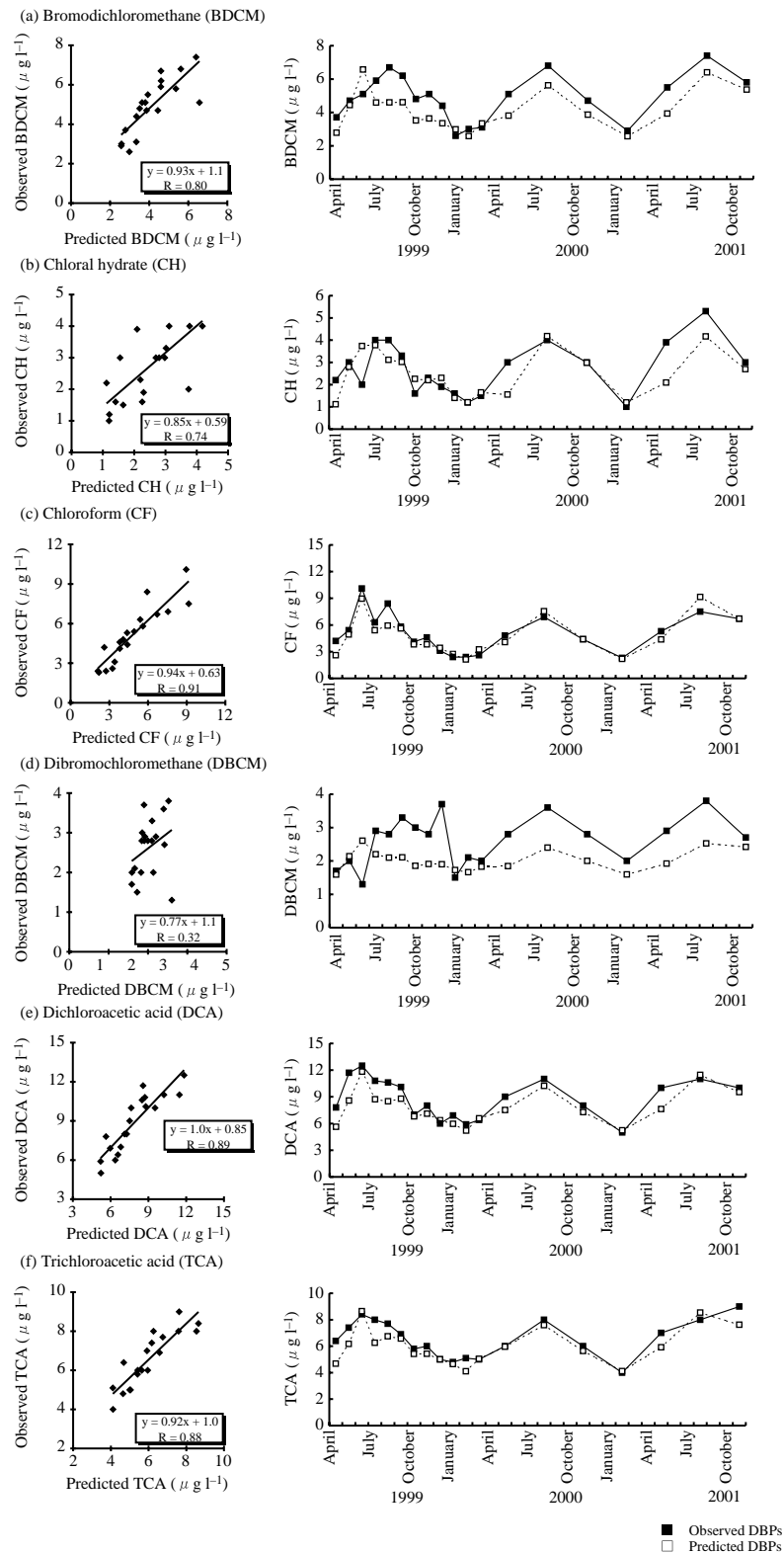


Figure 4 | Relationships between the predicted DBPs and the observed DBPs and their seasonal changes in location G.

**Table 6** | The degree of influence on model precision by measurement errors or temporal changes of variables

KMnO <sub>4</sub> [mg l <sup>-1</sup> /mg l <sup>-1</sup> ]	Cl <sub>2</sub>		pH		Water (WT)	temperature	Residence time (RT)		
	b × log(KMnO <sub>4</sub> )	[mg l <sup>-1</sup> /mg l <sup>-1</sup> ]	c × log(Cl <sub>2</sub> )	pH	d × pH	[°C]	e × log(WT)	[h]	f × log(RT)
1.5	0.018	1.3	0.114	7.298	0.7298	15.8	0.599	11	0.208
1.6	0.020	1.4	0.146	7.299	0.7299	15.9	0.601	14	0.229
1.7	0.023	1.5	0.176	7.300	0.7300	16.0	0.602	17	0.246
1.8	0.026	1.6	0.204	7.301	0.7301	16.1	0.603	20	0.260
1.9	0.028	1.7	0.230	7.302	0.7302	16.2	0.605	23	0.272

Note: When the values of above variables such as KMnO<sub>4</sub> are changed to  $\log(\text{DBPs}) = a + b \times \log(\text{KMnO}_4) + c \times \log(\text{Cl}_2) + d \times \text{pH} + e \times \log(\text{WT}) + f \times \log(\text{RT})$ , the results show how the values such as  $b \times \log(\text{KMnO}_4)$  affect the prediction.

b ~ f are the constants for each parameter (b = 0.1, c = 1, d = 0.1, e = 0.5, f = 0.2) calculated from the experimental results shown in Figure 4.

correlation coefficients between the predicted and observed increased compared with the original ones (Figure 4) from 0.026 to 0.056 orders of magnitude except for bromodichloromethane. The result suggests that the hourly change of Cl<sub>2</sub> concentrations should be considered.

## CONCLUSIONS

In this study, DBPs were measured in the north Hiroshima water distribution system at six locations from April 2000 to March 2001. This study has shown the importance of many parameters for DBP formation, such as potassium permanganate consumption value (KMnO<sub>4</sub>), concentration of residual chlorine (Cl<sub>2</sub>), pH, water temperature (WT) and residence time (RT) in the distribution system. Then a model was developed for predicting DBPs based on multiple regression analysis as follows:

$$\log(\text{DBPs}) = a + b \times \log(\text{KMnO}_4) + c \times \log(\text{Cl}_2) + d \times \text{pH} + e \times \log \text{WT} + f \times \log \text{RT}$$

The model was very useful because the R<sup>2</sup> values between the logarithmic values of the predicted DBPs and those of the observed DBPs were higher than 0.7 for all DBPs.

Next, the model was validated at a location that was not used for the model development. The correlation coefficients between the predicted and observed concentrations

of chloroform, dichloroacetic acid and trichloroacetic acid, whose concentrations were higher than the other DBPs, were approximately 0.9. In addition, these seasonal change patterns of the predicted DBPs were very similar to those of the observed DBPs.

To achieve greater accuracy, it is necessary to investigate the type and amount of organic substances affecting the DBP formation. However, our model will be helpful for managing DBPs because the model was developed based on monitoring data in the distribution system and the input parameters for the model were measured easily.

## REFERENCES

- Adin, A., Katzendler, J., Alkaslassy, D. & Rav-Acha, C. H. 1991 Trihalomethane formation in chlorinated drinking water: a kinetic model. *Wat. Res.* **25**, 797–805.
- Amy, G. L., Chadik, P. A. & Chowdhury, Z. K. 1987 Developing models for predicting trihalomethane formation potential and kinetics. *J. Am. Wat. Wks Assoc.* **79**, 89–97.
- Chen, W. J. & Weisel, C. P. 1998 Halogenated DBP concentrations in a distribution system. *J. Am. Wat. Wks Assoc.* **90**, 151–163.
- El-shafy, M. A. & Grunwald, A. 2000 THM formation in water supply in south Bohemia, Czech Republic. *Wat. Res.* **34**, 3453–3459.
- Gallard, H. & Gunten, U. V. 2001 Chlorination of natural organic matter: kinetics of chlorination and of THM formation. *Wat. Res.* **36**, 65–74.
- Hasegawa, R., Hirose, A., Nishikawa, A., Kobayashi, H., Ema, S. & Kurokawa, Y. 1999 Toxicity evaluation of dichloroacetic acid

- and estimation of the tolerable daily intake through oral route. *J. Japan Soc. Wat. Environ.* **22**, 821–826.
- Japan Water Works Association 1993 *Japanese Standard Method for the Examination of Drinking Water*. Japan Water Works Association, Japan.
- Kamei, T., Tambo, N. & Osako, S. 1994 Prediction of TOX & THM formation by a parallel consecutive first-order reaction model. *J. Japan Wat. Wks Assoc.* **63**(8), 21–32.
- Miyata, M., Umetani, T., Shiode, S. & Hashimoto, M. 1996 Effect on various chemical terms on the formation of disinfection by-products. *J. Wat. Waste* **38**, 1025–1031.
- Pontius, F. W. 2001 Regulatory update for 2001 and beyond. *J. Am. Wat. Wks Assoc.* **93**, 66–80.
- Reckhow, D. A., Singer, P. C. & Malcolm, R. L. 1990 Chlorination of humic materials: byproduct formation and chemical interpretations. *Environ. Sci. Technol.* **24**, 1655–1664.
- Rodriguez, M. J. & Serodess, J. 2001 Spatial and temporal evolution of trihalomethanes in three water distribution systems. *Wat. Res.* **35**, 1572–1586.
- Rook, I. J. 1977 Chlorination reactions of fulvic acids in natural waters. *Environ. Sci. Technol.* **11**, 478–482.
- Rook, I. J. 1980 Possible pathways for the formation of chlorinated degradation products during chlorination of humic acids and resorcinol. In Jolly, R. L. (ed.) *Water Chlorination: Environmental Impact and Health Effects*, vol. 3. Ann Arbor Science Publishers, Michigan, pp. 85–98.
- Sugino, K., Nishi, S., Ohura, R. & Horimoto, Y. 1986 Formation of chloroacetic acids and chloral hydrate from humic substances during aqueous chlorination. *Japan J. Wat. Pollut. Res.* **9**, 437–444.
- Sung, W., Reilley-Matthews, B., O'Day, D. K. & Horrigan, K. 2000 Modeling DBP formation. *J. Am. Wat. Wks Assoc.* **92**, 53–63.
- Ueno, H., Moto, T., Sayato, Y. & Nakamuro, K. 1996 Mutagenicity of volatile chlorination by-products and their occurrence in tap water. *J. Japan Wat. Wks Assoc.* **65**(2), 20–27.
- Urano, K., Wada, H. & Takemasa, T. 1983 Empirical rate equation for trihalomethane formation with chlorination of humic substances in water. *Wat. Res.* **17**, 1797–1802.
- Westerhoff, P., Debroux, J., Amy, G. L., Gatel, D., Mary, V. & Cavar, D. J. 2000 Applying DBP models to full-scale plants. *J. Am. Wat. Wks Assoc.* **92**, 89–102.

First received 21 August 2004; accepted in revised form 26 January 2005