Importance of correcting for fluorescence quenching in fluorescence-based prediction of trihalomethane formation potential

K. Saipetch and C. Yoshimura

ABSTRACT

Fluorescence excitation–emission matrix (EEM) spectroscopy is often used to determine the levels of trihalomethane (THM) precursors in natural organic matter. However, humic substances are known to quench the fluorescence of amino acids and proteins. To date, none of the EEM-based models for predicting THM formation potential (THMFP) have explicitly accounted for these quenching effects. Thus, we investigated the importance of correcting for fluorescence quenching during THMFP prediction. Fluorescence titration experiments revealed that the correction improved the accuracy of THM prediction. EEM-based models using the corrected fluorescence intensity displayed the highest accuracy ($R^2 > 0.99$; mean absolute error 8.1 μg/L and 13.9 μg/L for chloroform and bromoform, respectively) among models using individual parameters of EEM intensity, dissolved organic carbon (DOC), ultraviolet absorbance at 254 nm (UV_{254}), specific UV_{254} ($SUVA_{254}$) and differential ultraviolet absorbance at 272 nm (ΔUV_{272}). Thus, EEM-based models require both the fluorescence intensity of a humic-like component and the corrected fluorescence intensity of a protein-like component for accurate THMFP prediction, for both chlorination and bromination processes. We also found it to be unnecessary to combine DOC with EEM intensity in terms of prediction accuracy, as long as the fluorescence quenching correction is applied.

Key words | fluorescence excitation–emission matrix, fluorescence quenching, natural organic matter, trihalomethane predictive model

K. Saipetch (corresponding author)
C. Yoshimura
Department of Civil and Environmental Engineering,
Tokyo Institute of Technology,
2-12-1, M1-4, Ookayama, Meguro-ku, Tokyo 152-8550,
Japan
E-mail: *saipetch.k.aa@m.titech.ac.jp*

INTRODUCTION

Disinfection byproducts (DBPs) are formed by the reaction of chlorine with natural organic matter (NOM) during disinfection processes. In the case of bromide-containing water, additional DBPs are often produced because chlorination generates aqueous bromine species ($HOBr/OBr^-$) that further react with NOM (Langsa *et al.* 2017). Such DBPs are a major drawback of chlorination because they are often carcinogenic (Richardson *et al.* 2007). DBPs include various classes of compounds, among which trihalomethanes (THMs) are a major component, accounting for approximately 10% of the total halogenated byproducts in water after chlorination (Krasner *et al.* 2006). Commonly regulated THMs include chloroform (trichloromethane, TCM), bromodichloromethane (BDCM), dibromochloromethane (DBCM) and bromoform (tribromomethane, TBM) (e.g. United States Environmental Protection Agency 2009).

For operational control during disinfection processes, THM concentrations are commonly monitored using gas chromatography (GC) coupled with either an electron capture detector or a mass spectrometer. Although GC affords sufficient measurement accuracy, it is expensive and has limited potential as an online monitoring technique. To supplement or even replace such costly standard methods,

doi: 10.2166/ws.2019.039

various predictive models for THMs have been developed based on water quality parameters indicating the levels of THM precursors (Chowdhury *et al.* 2009; Beauchamp *et al.* 2018). Bulk water quality parameters, such as dissolved organic carbon (DOC), ultraviolet absorbance at 254 nm (UV_{254}), specific UV_{254} ($SUVA_{254}$) and differential ultraviolet absorbance at 272 nm (ΔUV_{272}), have been employed as the predictors in these models (Chowdhury *et al.* 2009; Beauchamp *et al.* 2018). Robust models using these bulk parameters to predict the total concentration of four THMs (TTHM) displayed coefficients of determination (R^2) ranging from 0.75 to 0.85 and standard error (SE) values ranging from 69 to 328 μg/L for a relatively large data set including various water sources (Ged *et al.* 2015). However, these models must be further improved because their SE values are similar to or higher than typical regulatory limits for TTHM (e.g. 80 μg/L).

The chemical structures of NOM compounds govern their reactivities with chlorine and bromine during the formation of THMs (Westerhoff *et al.* 2004) and also the types and concentrations of THMs formed (Liang & Singer 2003). Fluorescence excitation–emission matrix (EEM) spectroscopy is an effective method for characterizing the chemical properties of NOM specimens. The EEM peak intensities are also correlated with the concentration of TTHM or TCM, which is attributed to the common aromatic structure between fluorophore and THM precursors. Consequently, TTHM and TCM concentrations are more closely correlated with EEM peak intensities than with the bulk parameters considered in other THMFP tests (Pifer & Fairey 2012; Pifer *et al.* 2013). Thus, the integration of EEM data has been hypothesized to improve the prediction accuracy of models using bulk parameters for THM prediction (Peleato & Andrews 2015; Awad *et al.* 2016).

However, the inclusion of EEM data in predictive models for THMs remains challenging, primarily owing to the large volumes of data generated. To address this problem, researchers have explored various interpretation methods for EEM data, such as peak picking, fluorescence regional integration (FRI) and parallel factor analysis (PARAFAC), in an effort to obtain the greatest prediction accuracy (Peleato & Andrews 2015). However, the limitations of the EEM technique itself have often been overlooked, not only in the application of DBP prediction but also in other applications involving NOM components (Baghoth *et al.* 2011; Hur & Cho 2012). One major limitation is that the fluorescence intensity does not reveal the actual concentrations of amino acids (AAs) and proteins owing to fluorescence quenching (Wang *et al.* 2015), even though AAs and proteins are also DBP precursors in addition to humic substances (Bond *et al.* 2012). These components are always present in water sources (Meng *et al.* 2013; Lavonen *et al.* 2015) and the inter-component interactions between humic substances and AAs or proteins are responsible for the fluorescence quenching. Approximately 35%–52% of the fluorescence intensity of AAs can be quenched by the interaction with humic substances (Wang *et al.* 2015), which may significantly reduce the accuracy of EEM-based models for DBPs. However, this limitation has never been considered during the application of EEM data in predictive models for DBPs and its influence on prediction accuracy remains unknown.

In contrast to previous studies, which directly used EEM data in their predictive models, we aimed to investigate whether correcting for inter-component fluorescence quenching could improve the performance of EEM-based models for predicting the THMFP. We conducted two fluorescence titration experiments, namely, titration of the protein bovine serum albumin (BSA) against a known concentration of humic substances and titration of humic substances against a known concentration of BSA, to investigate the influence of correcting for fluorescence quenching during THMFP prediction. The importance of correcting for fluorescence quenching was then evaluated by comparing the fitness of EEM-based models using the measured and corrected fluorescence intensities as well as that of other models based on bulk parameters. The importance of this correction was investigated for two aqueous halogens (chlorine and bromine) to obtain a fundamental understanding of the individual contribution of correcting for fluorescence quenching in the case of each halogen.

METHODS

Fluorescence titration

Suwannee River natural organic matter (SWNOM, 2R101N) was obtained from the International Humic

Substances Society (IHSS) and was used to represent the end-member of humic substances, and BSA (≥96%, Sigma Aldrich) was used as a representative protein. A protein was considered as the end-member in this study rather than other protein-related species such as AAs owing to the occurrence of proteins as the major fraction in water sources (Aiken 2014). The preparation of the NOM stock solutions is described in Section S1 of the Supplementary Material (available with the online version of this paper). Two fluorescence titration experiments were conducted according to the procedure described by Wang *et al.* (2015). In the first experiment, various concentrations of SWNOM (0 to 15 mg/L) were added to a fixed concentration of BSA (10 mg/L). In the second experiment, various concentrations of BSA (0 to 15 mg/L) were added to a fixed concentration of SWNOM (10 mg/L). The final volume of the titrated samples was fixed at 200 mL in all cases. To ensure complete interaction, the titrated samples were shaken for 24 h using a mechanical shaker at room temperature in the dark. We confirmed that the shaking procedure itself did not lead to any loss of fluorescence intensity for either SWNOM or BSA. The concentrations of the components in the titrated solutions were chosen to afford a similar level of fluorescence intensity as wastewater-impacted source water.

THMFP tests

THMFP tests were conducted for all of the samples obtained from the fluorescence titration experiments, and phosphate buffer solution was used as a negative control. These water samples were divided into two series for chlorination and bromination experiments. The tests were conducted according to the 22nd edition of *Standard Methods for the Examination of Water and Wastewater* (APHA/AWWA/WEF 2012) using a headspace-free reactor. An aqueous stock solution of chlorine (5,000 mg Cl_2/L) was prepared from sodium hypochlorite (>5%, Kanto Chemical), and an aqueous stock solution of bromine (5,000 mg Br_2/L) was prepared from pure liquid bromine (99.9%, Sigma Aldrich). To perform the THMFP tests, these stock solutions were separately added to the reactor containing 100 mL of a water sample. Two series of THMFP tests were conducted, using each of the two aqueous halogen solutions at a final concentration of 1 mmol/L, which was selected based on the halogen demand corresponding to a maximum DOC concentration of 6.67 mg C/L. The reaction mixtures were incubated for 24 h under controlled laboratory conditions at 25 °C and pH 7.0 ± 0.2. For each sample, we analysed the EEM, DOC, UV_{254}, $SUVA_{254}$, ΔUV_{272}, halogen concentration and THM concentrations both before and after the test. Analytical procedures and the instruments used for water quality parameters, halogen concentration and THM concentrations are presented in Sections S2 and S3 of the Supplementary Material (available online).

Fluorescence measurements and calculation of unquenched fluorescence intensity

Three-dimensional fluorescence spectra were measured at controlled room temperature (25 °C) using a spectrofluorometer equipped with a xenon lamp as the light source (RF5300, Shimadzu, Japan) and a quartz cuvette (1 cm path length). The EEM spectra of the samples and phosphate buffer solutions (blanks) were scanned over an excitation (Ex) range of 240–500 nm with an increment of 5 nm and an emission (Em) range of 280–550 nm with an increment of 2 nm. The raw EEM data were pretreated according to the procedure described by Murphy *et al.* (2010) to correct for the inner filter effect and fluorescence signal fluctuation (Raman unit calibration). The pretreatment was performed using MATLAB software (MathWorks Inc., USA) with the fdomcorrect script of the FDOMcorr1.6 toolbox, as described by Murphy *et al.* (2010). Following this pretreatment, the fluorescence intensities were analysed by determining the change in fluorescence intensity for five specific Ex/Em pairs (Table 1) representing different components present in NOM (Coble 1996). These Ex/Em pairs were selected based on the peak locations of PARAFAC components that are typically present in various standard NOM samples from the IHSS (Ateia *et al.* 2017).

The fluorescence intensities of BSA at the Ex/Em pairs corresponding to protein-like components (i.e. peaks T and B) in the absence of SWNOM were considered to represent the unquenched fluorescence intensities of BSA in the fluorescence titration experiment involving a fixed concentration of BSA and various concentrations of SWNOM. The average quenching ratios (Q) were obtained from titration

Table 1 | Ex/Em pairs used in this study

		This study		Reference peak regions (Coble 1996)	
Component	**Peak name**	**Excitation (nm)**	**Emission (nm)**	**Excitation (nm)**	**Emission (nm)**
Humic-like	A	240	444	237–260	380–480
Tyrosine-like	B	275	306	270–280	300–320
Humic-like	C	350	460	330–350	420–480
Marine humic-like	M	300	444	310–320	380–420
Tryptophan-like	T	275	340	270–280	320–350

series for the calculation of the unquenched fluorescence intensities of BSA in another series. The unquenched fluorescence intensities were calculated according to Equation (1):

$$I_0 = \frac{I_q}{1 - Q} \quad (1)$$

where I_q and I_0 are the fluorescence intensities of BSA in the presence and absence of SWNOM, respectively. The error in the calculation of the unquenched fluorescence intensities using this equation was determined to be 7.0% ± 0.2% for both of the Ex/Em pairs corresponding to protein-like components.

THMFP predictive models

Mathematical models were developed using simple and multi-linear regression to examine the relationships between the water quality parameters (EEM, DOC, UV_{254}, $SUVA_{254}$ and ΔUV_{272}) and THM concentrations. In the case of EEM, stepwise multi-linear regression (forward and backward, $p < 0.05$) was applied to determine the suitable combinations of Ex/Em pairs, including all of the peaks listed in Table 1, with both measured and corrected fluorescence intensities. Model fitness was evaluated using the fitness indicators of R^2 and mean absolute error (MAE). In addition, the models were compared with the EEM-based model recently reported by Awad *et al.* (2016). In terms of prediction accuracy, the contribution of the addition of DOC was determined using multi-linear regression. The DOC was added to the EEM-based models both before and after the correction for fluorescence quenching, as well as the models based on the other water quality parameters, to cover both non-aromatic and aromatic organic matter fractions for THMFP prediction.

RESULTS AND DISCUSSION

Fluorescence quenching

The Ex/Em pairs were used to track the changes in fluorescence intensity during the fluorescence titration experiments. In this method, the overlap of fluorescence spectra between the two end-members was neglected because it contributed less than 5% of the fluorescence peak intensity in each case. At a fixed BSA concentration, the presence of SWNOM reduced the fluorescence intensities of BSA at both of the Ex/Em pairs corresponding to protein-like components (i.e. peaks T and B; Figure 1(a)). The fluorescence intensities of BSA at the Ex/Em pairs of peaks T and B were quenched in the presence of SWNOM with average quenching ratios (Q) of 0.89 ± 0.01 and 0.69 ± 0.01, respectively. The higher value of Q for peak T than for peak B was possibly attributable to the greater initial fluorescence intensity of BSA at the Ex/Em pair of peak T in the absence of SWNOM. In the second titration experiment at a fixed SWNOM concentration, the addition of BSA did not quench the fluorescence intensities of SWNOM at any of the Ex/Em pairs corresponding to humic-like components (i.e. peaks A, C and M; Figure 1(b)). These results are in accordance with a previous report that described the quenching of the fluorescence of AAs and proteins by humic substances and the lack of quenching of the fluorescence of humic substances by AAs and proteins

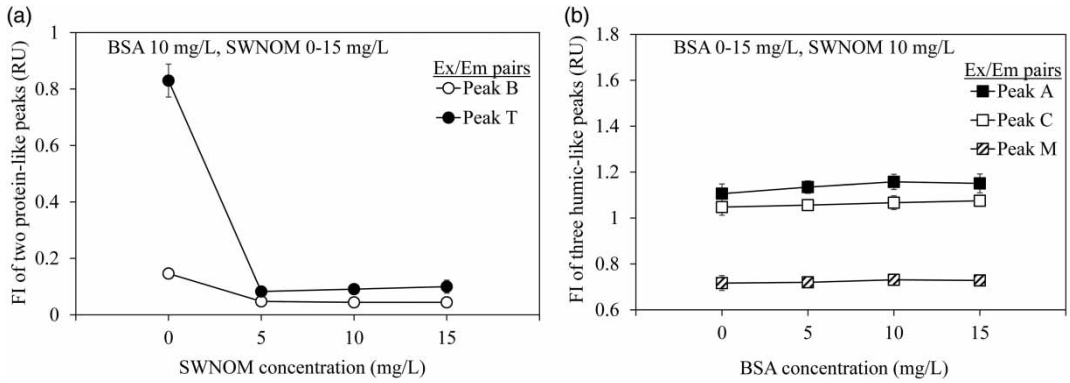

Figure 1 | Changes in fluorescence intensity during the two fluorescence titration experiments (mean ± standard deviation): (a) changes in fluorescence intensities of BSA at the Ex/Em pairs of peaks T and B during the first titration experiment using a fixed BSA concentration (10 mg/L) and various SWNOM concentrations (0–15 mg/L); (b) changes in fluorescence intensities of SWNOM at the Ex/Em pairs of peaks A, C and M during the second titration experiment using a fixed SWNOM concentration (10 mg/L) and various concentrations of BSA (0–15 mg/L).

(Wang *et al.* 2015). Further details regarding the fluorescence characteristics of the end-members and their mixtures as well as an example of an EEM spectrum of natural water are presented in Section S4 of the Supplementary Material (available with the online version of this paper).

THMFP tests

In the THMFP tests, TCM was the only THM detected in the chlorination experiments and it was present in the samples at concentrations in the range of 130–480 μg/L. In the bromination experiments, TBM was the dominant product among the THMs and occurred in the range of 350–1,326 μg/L, while the concentrations of other species were <0.5% of the TBM concentration. The concentrations of both TCM and TBM increased upon increasing the concentration of either BSA or SWNOM. Therefore, both of the end-members reacted with both chlorine and bromine to serve as THM precursors.

Influence of correcting for fluorescence quenching on model fitness

The use of the corrected fluorescence intensities improved the model fitness in terms of both R^2 and MAE, as indicated by the dashed lines shown in Figure 2(a) and 2(b). For the combination of the fluorescence intensities at peaks A and T, the use of the corrected fluorescence intensity of peak T (i.e. peak T_c) for TCM and TBM prediction increased the value of R^2 by 25.0% and 0.5%, respectively, and decreased the value of MAE by 45.5% and 30.0%, respectively. For the combination of the fluorescence intensities at peaks A and B, the use of the corrected fluorescence intensity of peak B (i.e. peak B_c) for TCM and TBM prediction increased the value of R^2 by 13.0% and 4.2%, respectively, and decreased the value of MAE by 59.5% and 79.9%, respectively. The use of the corrected fluorescence intensities also reduced the deviation between the actual and modelled concentrations of THMs in the case of the models showing the best fitness for TCM and TBM prediction (Figure 3(a) and 3(b)). When this correction was not applied, the predicted TCM concentrations exhibited a greater deviation from the 1:1 line than the predicted TBM concentrations. The different magnitudes of the deviation for TCM and TBM prediction were presumably attributable to the greater quenching ratio for the fluorescence intensity at peak T compared with that at peak B upon the interaction with SWNOM, as discussed earlier. Furthermore, the EEM-based model using the corrected fluorescence intensities displayed the highest accuracy, with $R^2 > 0.99$ and MAE values of 8.1 μg/L and 13.9 μg/L for TCM and TBM, respectively, compared with the models based on individual water quality parameters such as EEM peaks, DOC, UV_{254}, $SUVA_{254}$ and ΔUV_{272} (Figure 2(a) and 2(b)). The examples of the data set used for calculation of R^2 and MAE of each individual water quality parameter are shown in Figure 3(c) and 3(d) for TCM and TBM, respectively. Therefore, correcting for fluorescence

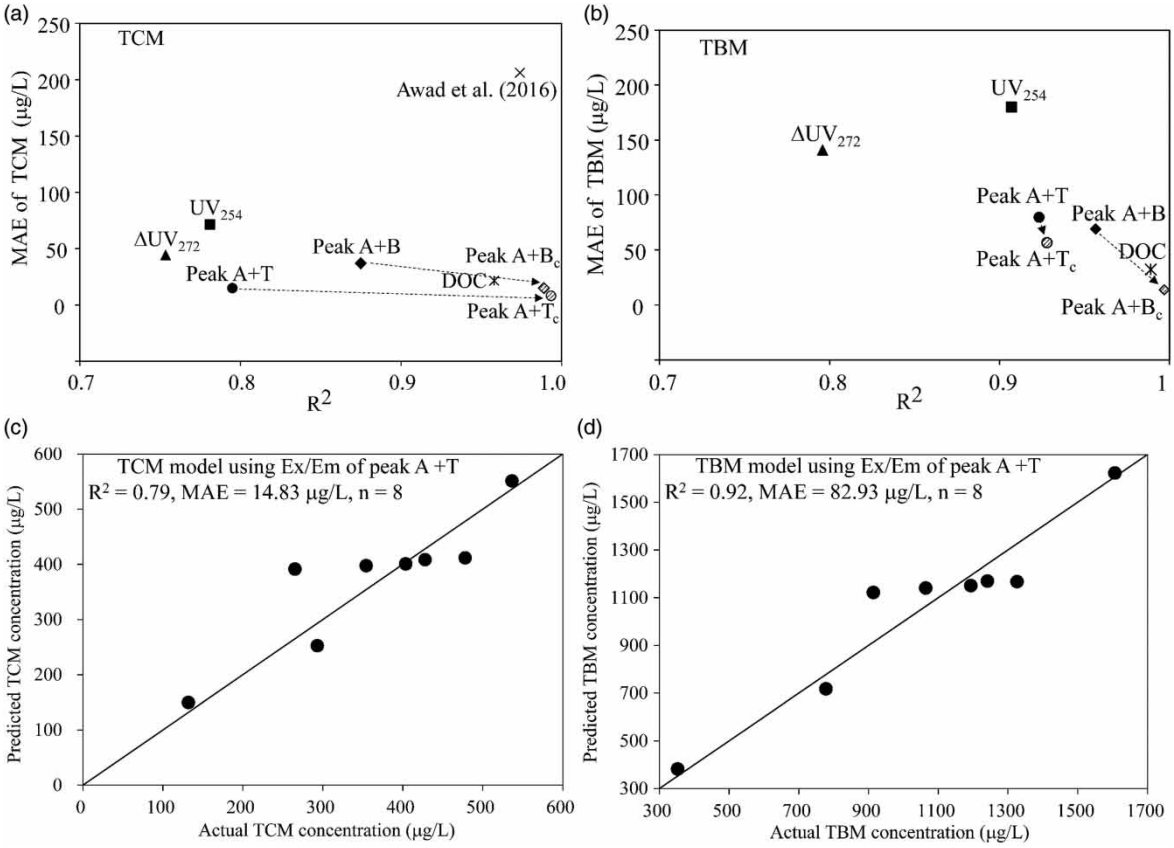

Figure 2 | Comparison of model fitness for models using individual water quality parameters for (a) TCM and (b) TBM. MAE indicates the mean absolute error of the linear regression. The subscript 'c' in the peak names indicates the use of corrected fluorescence intensities for these peaks. The fitness of the model based on $SUVA_{254}$ was outside of the range of these plots (R^2 values of 0.34 and 0.51 and MAE values of 83.8 µg/L and 225.4 µg/L for TCM and TBM prediction, respectively). Panels (c) and (d) show the model performance for TCM and TBM, respectively; *n* indicates the number of samples in each model.

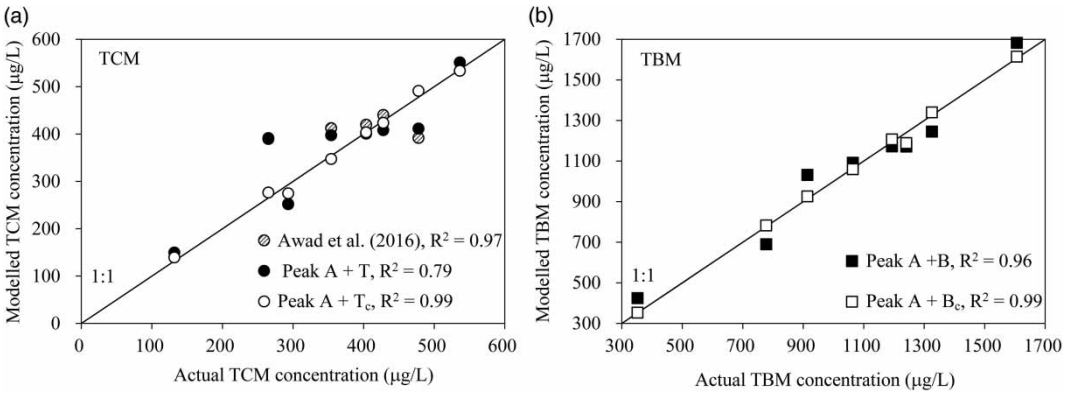

Figure 3 | Actual and modelled concentrations of (a) TCM and (b) TBM.

quenching is necessary for improving the prediction accuracy of THMs, regardless of whether chlorination or bromination is being considered.

The models that displayed the highest fitness are written in Equations (2) and (3) for TCM and TBM prediction, respectively:

$$\text{TCM } (\mu g/L) = 223.7 \times \text{Peak A} + 167.8 \times \text{Peak T}_c \quad (2)$$

$$\text{TBM}(\mu g/L) = 1069.5 \times \text{Peak A} + 1562.9 \times \text{Peak B}_c \quad (3)$$

where the subscript 'c' in the peak names indicates the use of corrected fluorescence intensities for these peaks. The obtained results of the corrected fluorescence intensities of the protein-like components and other water quality parameters are presented in Table S1 of the Supplementary Material (available online).

Interestingly, the stepwise regression resulted in the selection of different Ex/Em pairs corresponding to protein-like components to indicate the TCM and TBM precursors, possibly indicating the specific preferences of chlorine and bromine to react with different parts of the protein molecule. The selection of the fluorescence intensity at peak T rather than that at peak B for TCM prediction was probably attributable to the greater ability of tryptophan to produce TCM (three-fold higher yield in μg TCM/mg C) compared with tyrosine (Hong *et al.* 2009). With respect to the selection of peak B rather than peak T for TBM prediction, further studies are required to explain this observation by comparing the production of TBM from each AA and bromine. In addition, chlorine tended to react more with humic substances than with protein, as indicated by the higher value of the coefficient for the fluorescence intensity at peak A than that at peak T_c. In contrast, bromine tended to react more with protein than with humic substances, as indicated by the higher value of the coefficient for the fluorescence intensity at peak B_c than that at peak A. A similar tendency of chlorine and bromine with different EEM peaks was also reported by Awad *et al.* (2016).

Contribution of DOC in predictive models

The addition of DOC to the EEM-based models with the correction of fluorescence quenching resulted in the same value of R^2 and reduced the MAE value by only 3% compared with the models without DOC, for both TCM and TBM prediction. This result was in contrast to the previous study by Awad *et al.* (2016), which reported that the combination of DOC with UV_{254}, rather than with EEM, afforded the best accuracy for THM prediction. This difference was probably caused by the use of EEM data without correcting for fluorescence quenching in the previous models. Nevertheless, the addition of DOC to the EEM-based predictive model after correcting for fluorescence quenching showed a minor improvement.

The addition of DOC to the TCM predictive models using UV_{254}, $SUVA_{254}$ and ΔUV_{272} improved the fitness of these models. However, the model fitness of these combinations remained lower than that of the model using only DOC ($R^2 = 0.96$, MAE = 21.4 μg/L). Thus, the EEM-based model using the combination of the fluorescence intensities at peaks A and T_c was still the best method for TCM prediction ($R^2 = 0.99$, MAE = 8.1 μg/L). For TBM prediction, the addition of DOC to UV_{254} or ΔUV_{272} improved the model fitness in terms of both R^2 and MAE, whereas the addition of DOC to $SUVA_{254}$ increased R^2 but did not decrease MAE. Nonetheless, the model using the combination of the fluorescence intensities at peaks A and B_c afforded superior model fitness for TBM prediction ($R^2 = 0.99$, MAE = 13.9 μg/L) compared with the combined models using DOC and UV_{254} ($R^2 = 0.96$, MAE = 21.25 μg/L), DOC and ΔUV_{272} ($R^2 = 0.99$, MAE = 31.14 μg/L) and DOC and $SUVA_{254}$ ($R^2 = 0.96$, MAE = 21.26 μg/L). Improvements in model accuracy for THM prediction have been observed upon combining DOC with other predictors (Ged *et al.* 2015), although our results suggest that correcting for fluorescence quenching leads to a greater improvement in model accuracy than the addition of DOC to other parameters.

CONCLUSIONS

Fluorescence EEM spectroscopy is often applied to determine the levels of THM precursors in NOM. We investigated the influence of correcting for fluorescence quenching on the accuracy of predictive models for THMs. We found that EEM-based models required both the fluorescence intensity of a humic-like component and the corrected fluorescence intensity of a protein-like component to accurately predict the levels of THMs, for the reactions with both aqueous chlorine and bromine. Correction of the fluorescence intensity of the protein-like components improved the accuracy of THM prediction, and EEM-based models exhibited the highest accuracy among models using individual water quality parameters (i.e. EEM peaks, DOC, UV_{254}, $SUVA_{254}$ and ΔUV_{272}) as the indicator of THM precursors. In addition, the combination of DOC with the corrected EEM data was found to reduce

the MAE by only 3% compared with the models using only the corrected EEM data as the indicator of THM precursors. Thus, the addition of DOC to the EEM peaks was unnecessary in terms of improving the prediction accuracy if the correction for fluorescence quenching was applied. Overall, this study quantitatively investigated the importance of correcting for fluorescence quenching in the development of predictive models for THMs. These results support the potential application of fluorescence measurements for predicting THMs during online monitoring. Further investigations are necessary to elucidate the importance of correcting for fluorescence quenching under various chlorination conditions (e.g. pH, contact time, temperature) and develop correction methods for the fluorescence quenching of unknown samples.

ACKNOWLEDGEMENTS

This study was financially supported by the JSPS Core-to-Core Program: Asia–Africa Science Platforms 'Establishment of Asian Model for Research and Education on Urban Water Resource Management' and SATREPS (JST/JICA). The authors would like to thank the Chiba Prefecture Pharmaceutical Association Inspection Center for their support with THM analysis and Mr Daniel Chang for English proofreading.

REFERENCES

Aiken, G. 2014 Fluorescence and dissolved organic matter: a chemist's perspective. In: *Aquatic Organic Matter Fluorescence* (P. G. Coble, J. Lead, A. Baker, D. M. Reynolds & R. G. M. Spencer, eds), Cambridge University Press, New York, USA, pp. 35–74.

APHA/AWWA/WEF 2012 *Standard Methods for the Examination of Water and Wastewater*, 22nd edn. American Public Health Association/American Water Works Association/Water Environment Federation, Washington, DC, USA.

Ateia, M., Ran, J., Fujii, M. & Yoshimura, C. 2017 The relationship between molecular composition and fluorescence properties of humic substances. *International Journal of Environmental Science and Technology* **14** (4), 867–880.

Awad, J., van Leeuwen, J., Chow, C., Drikas, M., Smernik, R. J., Chittleborough, D. J. & Bestland, E. 2016 Characterization of dissolved organic matter for prediction of trihalomethane formation potential in surface and sub-surface waters. *Journal of Hazardous Materials* **308**, 430–439.

Baghoth, S. A., Sharma, S. K. & Amy, G. L. 2011 Tracking natural organic matter (NOM) in a drinking water treatment plant using fluorescence excitation–emission matrices and PARAFAC. *Water Research* **45** (2), 797–809.

Beauchamp, N., Dorea, C., Bouchard, C. & Rodriguez, M. 2018 Use of differential absorbance to estimate concentrations of chlorinated disinfection by-product in drinking water: critical review and research needs. *Critical Reviews in Environmental Science and Technology* **48** (2), 210–241.

Bond, T., Goslan, E. H., Parsons, S. A. & Jefferson, B. 2012 A critical review of trihalomethane and haloacetic acid formation from natural organic matter surrogates. *Environmental Technology Reviews* **1** (1), 93–113.

Chowdhury, S., Champagne, P. & McLellan, P. J. 2009 Models for predicting disinfection byproduct (DBP) formation in drinking waters: a chronological review. *Science of the Total Environment* **407** (14), 4189–4206.

Coble, P. G. 1996 Characterization of marine and terrestrial DOM in seawater using excitation-emission matrix spectroscopy. *Marine Chemistry* **51**, 325–346.

Ged, E. C., Chadik, P. A. & Boyer, T. H. 2015 Predictive capability of chlorination disinfection byproducts models. *Journal of Environmental Management* **149**, 253–262.

Hong, H. C., Wong, M. H. & Liang, Y. 2009 Amino acids as precursors of trihalomethane and haloacetic acid formation during chlorination. *Archives of Environmental Contamination and Toxicology* **56** (4), 638–645.

Hur, J. & Cho, J. 2012 Prediction of BOD, COD, and total nitrogen concentrations in a typical urban river using a fluorescence excitation-emission matrix with PARAFAC and UV absorption indices. *Sensors* **12** (1), 972–986.

Krasner, S. W., Weinberg, H. S., Richardson, S. D., Pastor, S. J., Chinn, R., Sclimenti, M. J., Onstad, G. D. & Thruston, A. D. 2006 Occurrence of a new generation of disinfection byproducts. *Environmental Science and Technology* **40** (23), 7175–7185.

Langsa, M., Heitz, A., Joll, C. A., von Gunten, U. & Allard, S. 2017 Mechanistic aspects of the formation of adsorbable organic bromine during chlorination of bromide-containing synthetic waters. *Environmental Science and Technology* **51** (9), 5146–5155.

Lavonen, E. E., Kothawala, D. N., Tranvik, L. J., Gonsior, M., Schmitt-kopplin, P. & Köhler, S. J. 2015 Tracking changes in the optical properties and molecular composition of dissolved organic matter during drinking water production. *Water Research* **85**, 286–294.

Liang, L. & Singer, P. C. 2003 Factors influencing the formation and relative distribution of haloacetic acids and trihalomethanes in drinking water. *Environmental Science and Technology* **37** (13), 2920–2928.

Meng, F., Huang, G., Yang, X., Li, Z., Li, J., Cao, J., Wang, Z. & Sun, L. 2013 Identifying the sources and fate of anthropogenically impacted dissolved organic matter (DOM) in urbanized rivers. *Water Research* **47** (14), 5027–5039.

Murphy, K. R., Butler, K. D., Spencer, R. G. M., Stedmon, C. A., Boehme, J. R. & Aiken, G. R. 2010 Measurement of dissolved organic matter fluorescence in aquatic environments: an interlaboratory comparison. *Environmental Science and Technology* **44**, 9405–9412.

Peleato, N. M. & Andrews, R. C. 2015 Comparison of three-dimensional fluorescence analysis methods for predicting formation of trihalomethanes and haloacetic acids. *Journal of Environmental Sciences* **27**, 159–167.

Pifer, A. D. & Fairey, J. L. 2012 Improving on $SUVA_{254}$ using fluorescence-PARAFAC analysis and asymmetric flow-field flow fractionation for assessing disinfection byproduct formation and control. *Water Research* **46** (9), 2927–2936.

Pifer, A. D., Cousins, S. L. & Fairey, J. L. 2013 Assessing UV- and fluorescence-based metrics as disinfection byproduct precursor surrogate parameters in a water body influenced by a heavy rainfall event. *Journal of Water Supply: Research and Technology – AQUA* **63** (3), 200–211.

Richardson, S. D., Plewa, M. J., Wagner, E. D., Schoeny, R. & DeMarini, D. M. 2007 Occurrence, genotoxicity, and carcinogenicity of regulated and emerging disinfection by-products in drinking water: a review and roadmap for research. *Mutation Research* **636** (1–3), 178–242.

USEPA 2009 *National Primary Drinking Water Regulations*. United States Environmental Protection Agency, USA.

Wang, Z., Cao, J. & Meng, F. 2015 Interactions between protein-like and humic-like components in dissolved organic matter revealed by fluorescence quenching. *Water Research* **68**, 404–413.

Westerhoff, P., Chao, P. & Mash, H. 2004 Reactivity of natural organic matter with aqueous chlorine and bromine. *Water Research* **38**, 1502–1513.

First received 14 September 2018; accepted in revised form 14 February 2019. Available online 4 March 2019