Environmental Pollution 254 (2019) 113028

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Use of multiple regression models for predicting the formation of bromoform and dibromochloromethane during ballast water treatment based on an advanced oxidation process☆

Xiaoye Zhang [a], Yiping Tian [b, *], Xiaofang Zhang [b], Mindong Bai [c], Zhitao Zhang [a, c]

[a] Marine Engineering College, Dalian Maritime University, Dalian, 116026, China
[b] Environmental Engineering Institute, School of Science, Dalian Maritime University, Dalian, 116026, China
[c] Key Laboratory of Ministry of Education for Coastal and Wetland Ecosystems, College of the Environment and Ecology, Xiamen University, Xiamen, 361005, China

ARTICLE INFO

Article history:
Received 20 May 2019
Received in revised form
29 July 2019
Accepted 4 August 2019
Available online 7 August 2019

Keywords:
Multiple regression model
Disinfection byproducts
Concentration prediction
Independent variables
Ballast water treatment

ABSTRACT

Disinfection byproducts (DBPs) generated by ballast water treatment have become a concern worldwide because of their potential threat to the marine environment. Predicting the relative DBP concentrations after disinfection could enable better control of DBP formation. However, there is no appropriate method of evaluating DBP formation in a full-scale ballast water treatment system (BWTS). In this study, multiple regression models were developed for predicting the dibromochloromethane (DBCM) and bromoform (TBM) concentrations produced by an emergency BWTS using field experimental data from ballast water treatments conducted at Dalian Port, China. Six combinations of independent variables [including several water parameters and/or the total residual oxidant (TRO) concentration] were evaluated to construct mathematical prediction formulas based on a polynomial linear model and logarithmic regression model. Further, statistical analyses were performed to verify and determine the appropriate mathematical models for DBCM and TBM formation, which were ultimately validated using additional field experimental data. The polynomial linear model with four variables (temperature, salinity, chlorophyll, and TRO) and the logarithmic regression model with seven variables (temperature, salinity, dissolved oxygen, pH, turbidity, chlorophyll, and TRO) exhibited good reproducibility and could be used to predict the DBCM and TBM concentrations, respectively. The validation results indicated that the developed models could accurately predict DBP concentrations, with no significant statistical difference from the measured values. The results of this work could provide a theoretical basis and data reference for ballast water treatment control in engineering applications of emergency BWTSs.

1. Introduction

Ballast water is essential for ship stability and navigation security during voyages, but the discharge of exotic ballast water is the main route by which non-indigenous species are introduced (Casas-Monroy et al., 2016). To minimize the negative impacts of non-indigenous species on local marine environments, the International Maritime Organization (IMO) has enforced the "International Convention for the Control and Management of Ships Ballast Water and Sediments" on September 8, 2017 (Zhang et al., 2017a,b), which requires that ballast water must be treated before discharge to meet the D-2 discharge standard: Phytoplankton/zooplankton ≥ 50 μm, $<$ 10 viable organisms per m^3; Phytoplankton/zooplankton 10−50 μm, $<$ 10 viable organisms per mL; and *Vibrio cholera* $<$ 1 cfu/100 mL, *Escherichia coli* $<$ 250 cfu/100 mL, *Intestinal Enterococci* $<$ 100 cfu/100 mL (IMO, 2008a).

Chemical-based ballast water treatment systems (BWTSs) using chlorine, ozone, hydrogen peroxide, or other strong oxidants (Werschkun et al., 2012; Shah et al., 2015) have been widely used to effectively inactivate harmful non-indigenous species in ballast water. Especially, the BWTSs based on advanced oxidation processes (AOPs) have received increasing attention because they provide rapid and highly efficient treatment. AOPs are defined as those processes that use •OH radicals for oxidation (Kornmueller,

☆ This paper has been recommended for acceptance by Charles Wong.

* Corresponding author.
E-mail address: yiping_tian@dlmu.edu.cn (Y. Tian).

https://doi.org/10.1016/j.envpol.2019.113028
0269-7491/

2007; Wang and Xu, 2012; Stefan, 2017; Miklos et al., 2018), such as photo-Fenton processes (Zhao et al., 2014; Fang et al., 2015), which were reported to efficiently degrade organic pollutants in wastewater. Several types of AOPs, such as those involving ozonation and UV irradiation, or electrochemical AOPs have also been used in BWTSs and received final approval for environmental safety by the IMO (IMO, 2008b).

However, when these chemical-based BWTSs are used, a series of genotoxic and carcinogenic DBPs (Stehouwer et al., 2015; Min and Min, 2016; Zeng et al., 2016) can easily be formed. It has been well demonstrated that the formation of DBPs during ballast water treatment is strongly influenced by the characteristics of the water (such as temperature, pH, salinity, and type and concentration of precursor), type of oxidant, oxidant dosage, and reaction time (Hao et al., 2017; Li et al., 2017; Ziegler et al., 2018). Thus, some onboard chemical-based BWTSs might unpredictably generate DBPs while treating ballast water in various ports worldwide (Shah et al., 2015), even though they have been approved as environmentally safe according to the G9 procedure of the IMO (IMO, 2008b). Moreover, more and more emergency BWTSs using active substances have been developed for those ships that do not install BWTSs or in which the treated ballast water still cannot meet IMO emission standards (Zhang et al., 2018). These BWTSs also easily produce excess DBPs during treatment under various ballast water conditions, which could potentially pollute the marine environment.

To control DBP formation during ballast water treatment, it is essential to accurately and quantitatively predict the DBP concentrations generated by chemical-based BWTSs before the treatment by chemical oxidation process. Many studies (Zhang et al., 2013a; Ged and Boyer, 2014) have discussed in broad terms the effects of the water characteristics on the production of DBPs in artificial seawater. Zhang et al. reported that the highest TBM concentration was found at pH = 7, and the increase in temperature enhanced the formation of trihalomethanes (THMs) during a simulated ballast water treatment process. Ged and Boyer found that increasing the bromide concentration of ballast water from 38 $\mu g\,L^{-1}$ to 974 $\mu g\,L^{-1}$ in a chlorination treatment process, resulted in the increase of concentration of THMs and haloacetic acids (HAAs), from 43 to 206 $\mu g\,L^{-1}$ and from 39 to 75 $\mu g\,L^{-1}$, respectively. However, quantitative prediction of DBP generation in full-scale ballast water treatment is still limited owing to the complexity of natural seawater and the lack of complete information.

At present, empirical and mathematical models are commonly used in DBP prediction. Engerholm and Amy (1983) tried to develop empirical models to predict the production of total trihalomethanes (TTHMs) and chloroform in drinking water. Zhang's group using empirical models had well explored the formation kinetics of overall Cl-DBPs, Br-DBPs, and I-DBPs during chlor (am) ination of drinking water (Zhai et al., 2014; Zhu and Zhang, 2016). However, most of the empirical models were developed from bench-scale experiments in site-specific cases, and the simulation results were reported to inadequately predict full-scale data (Siddiqui and Amy, 1993; Sohn et al., 2004). Mathematical modeling, typically that using multiple regression between a dependent variable and several independent variables, is another promising method with potential for accurate prediction based on field data (Owen et al., 2011). More than 150 mathematical models have been developed to predict TTHM formation, but most of them are used for drinking water treatment (Golfinopoulos and Arhonditsis, 2002; Chowdhury et al., 2009; Chen and Westerhoff, 2010). There is no proposed prediction model in the literature for the evaluation of DBPs generated in ballast water.

In this study, mathematical multiple regression models were established to evaluate the concentrations of DBPs produced by an emergency BWTS based on an AOP that composed of strong ionization discharge and hydrodynamic cavitation. The AOP-based emergency BWTS was developed on the basis of previous studies (Bai et al., 2012; Zhang et al., 2013a,b; Tian et al., 2015; Yu et al., 2017) and installed in a port to perform field experiments on ballast water treatment. On the basis of 21 sets of field experimental data, possible relationships between the DBP concentration and experimental variables [including water parameters and the total residual oxidant (TRO) concentration] were analyzed. Further, multiple regression models with different combinations of independent variables were constructed on the basis of a polynomial linear model and a logarithmic regression model. Then statistical analyses were performed to evaluate the models and identify appropriate models for DBP prediction. Finally, the accuracy of the appropriate models was validated using 29 sets of additional field experimental data. The research results can provide theoretical guidance for efficient, safe, and practical application of AOP-based emergency BWTSs in different sea areas and further help to reduce the risk of DBPs to the marine environment.

2. Material and methods

2.1. Field tests of AOP-based emergency BWTS

2.1.1. Experimental area

Field tests of ballast water treatment using the AOP-based emergency BWTS were conducted at the fourth wharf of Dalian Port, China. Dalian Port is the largest multipurpose port in Northeast China; as shown in Fig. 1, it is located at the south end of Liaodong Peninsula. Large amounts of ballast water from various oceans could be discharged from ships, leading to quite complex water characteristics. Additionally, because the port is surrounded on three sides by breakwaters, natural water purification in this sea area is very limited. As a result, there is a high risk of invasion by harmful non-indigenous aquatic organisms and pathogens at Dalian Port. For this reason, this sea area was selected as a representative site for the field tests.

2.1.2. AOP-based emergency BWTS

The AOP-based emergency BWTS used in the field tests was developed for a treatment capacity of 30 tons per hour (t/h); it consists mainly of units for plasma reaction, hydrodynamic cavitation, filtering, and neutralization. A schematic diagram of the treatment system is shown in Fig. 2. The core technologies in the

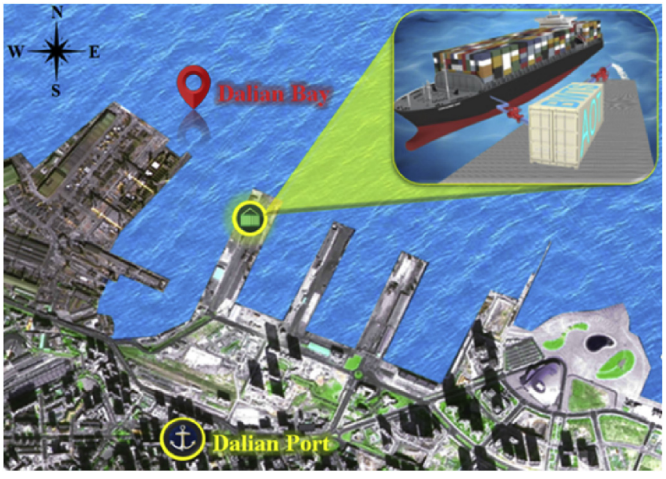

Fig. 1. Topographical map of Dalian Port showing location of emergency BWTS.

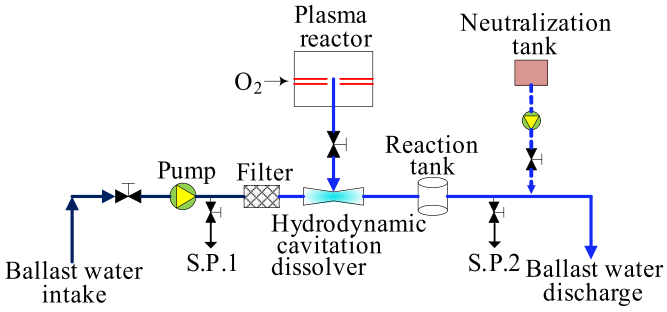

Fig. 2. Schematic diagram of AOP-based emergency BWTS.

system are strong ionization discharge (plasma reaction) and hydrodynamic cavitation, by which large quantities of •OH radicals can be formed for ballast water treatment (Bai et al., 2016; Bai et al., 2018). A schematic of ballast water treatment by strong ionization discharge in conjunction with hydrodynamic cavitation is shown in Fig. S1. The plasma reactor unit (details are shown in Fig. S2) is used to generate the strong ionization discharge used to produce reactive oxygen species (ROSs, including O_2^+, O, $O_2^{\bullet-}$, O_2 ($a^1\Delta_g$), O (1D), and O_3) (Zhang et al., 2011; Bai et al., 2016; Bai et al., 2018). Under the effect of hydrodynamic cavitation, the ROSs can react with water to simultaneously form •OH radicals via a series of plasma chemical reactions (Zhang et al., 2018).

During treatment, the ballast water is first filtered to remove particles and organisms larger than 50 μm. Then the pretreated water is conveyed into the hydrodynamic cavitation unit, into which high concentrations of ROSs produced by the plasma reactor unit are also injected to mix with the inflow of ballast water; large numbers of •OH radicals are also generated (Bai et al., 2010; Zhang et al., 2014). The ballast water containing •OH radicals is then delivered to the reaction tank for adequate mixing, during which the harmful algae and bacteria in the ballast water can be efficiently inactivated. Sodium thiosulfate ($Na_2S_2O_3$) is used to neutralize the remaining residual oxidants before discharge if the residual oxidant value is higher than 0.2 mg L^{-1}.

2.1.3. Sample collection and analysis

Twenty-one sets of ballast water treatments (four or five tests per month) were conducted using the AOP-based emergency BWTS at Dalian Port from July to November 2016. For each test, ballast water samples were collected as controls from S.P.1 (see Fig. 2), and treated water samples were collected from S.P.2 immediately after treatment. To obtain accurate DBP concentrations, $Na_2S_2O_3$ (3 mg/40 mL) was added to the treated ballast water samples to neutralize the residual oxidants and stop the reactions. All samples were immediately stored at 4 °C and transported to laboratory for further characterization.

The ballast water parameters, including temperature (*T*), salinity (Sal), dissolved oxygen content (DO), pH, turbidity (Tur), and chlorophyll concentration (Chl), were determined at the experimental site before treatment by multiparameter water quality sondes (YSI-6600, YSI, US). The TRO, including $HOCl/OCl^-$, $HOBr/OBr^-$, and other active compounds such as NH_2Br, $NHBr_2$, NBr_3, etc., formed during treatment (Zhang et al., 2012, 2013a,b), were monitored by the *N*,*N*-diethyl-p-phenylenediamine colorimetric method based on the US EPA's standard method 330.5 (US EPA, 1978). Dibromochloromethane (DBCM) and bromoform (TBM) were analyzed as representative DBPs after treatment by a purge-and-trap sample concentrator (Tekmar Stratum 3100, TEKMAR, US) connected to a gas chromatograph with an electronic capture detector (Agilent 7890A, Agilent Technologies Inc., US) according to US EPA method 524.2 (US EPA, 1995).

2.2. Development of multiple regression models

The polynomial linear model and logarithmic regression model, two general approaches that are typically used to model DBP formation (Amy et al., 1987; Serodes et al., 2003) were selected to develop the prediction models in this study. The concentration of DBCM or TBM was used as the dependent variable, and the other variables [including water parameters and TRO concentration] represent candidate independent variables. To determine the regression equations, the linear or nonlinear relationship between the dependent variables (or their transformed logarithmic forms) and the independent variables (or their transformed logarithmic forms) was analyzed using statistical analysis software (SPSS, Version 19.0) (Amy et al., 1987). The analysis results indicate that DBP formation under different water parameters and TRO values could be accurately modeled by transforming the dependent variable into the log 10 form or without transformation. Thus, the polynomial linear equation ($Y = a_0 + X_1a_1 + X_2a_2 + \ldots X_pa_p$) and logarithmic regression equation ($\log_{10}Y = a_0 + X_1a_1 + X_2a_2 + \ldots X_pa_p$) were used to develop the prediction models of DBP formation during ballast water treatment, where *Y* (the dependent variable) is the DBP concentration in the treated ballast water, X_i ($i = 1$ to p) is the independent variable, a_i ($i = 1$ to p) represents the statistical coefficients to be estimated, and a_0 is a constant term (Chen and Westerhoff, 2010).

For model development, the measured data for the variable parameters obtained in field tests, including TRO concentrations, DBP concentrations, and ballast water parameters such as Sal, DO, and Chl, were first analyzed statistically using SPSS. Then, the correlations between the experimental parameters and DBP formation were verified using the Pearson correlation matrix at the 95% significance level ($\alpha < 0.05$). On the basis of the results, several independent variables were selected to construct the multiple regression models (Golfinopoulos and Arhonditsis, 2002). By using different combinations of the selected variables, the reproducibility of the logarithmic regression model and polynomial model were evaluated by statistical analyses. Further, the models meeting the criteria of maximum R^2 and minimum standard error (SE) were selected for further statistical analyses.

The predicted data calculated using the selected models were compared with the measured values using the F-test, Student's T-test, and analytical variance (AV) (Domínguez-Tello et al., 2017). If the F-test value was >0.5, the Student's T-test with equal variance was conducted; if the F-test value was <0.5, the Student's T-test with unequal variance was conducted. When the Student's T-test result was >0.5, uncertainty analyses including SE, AV, and R^2 were conducted (Chen and Westerhoff, 2010). The model with a higher *t* value, higher R^2, and lower AV and SE according to the statistical analysis was identified as the appropriate model. Moreover, the statistical significance of the selected model was checked using the *F* value and Durbin–Watson estimate (Chowdhury et al., 2009).

2.3. Validation of multiple regression models

To validate the developed multiple regression models, additional 29 sets of ballast water treatments were performed using the AOP-based emergency BWTS at Dalian Port from May to July 2018. The concentrations of DBPs generated during treatment were predicted using the developed models. Then, the R^2 and R^2 adj were obtained to estimate the difference between the predicted and measured values. Finally, a T-test was performed to calculate the t_{value} of the model predictions, and the difference of the predicted values and measured values was determined by comparing t_{value}

with the $t_{critical}$ value obtained from references. If $t_{value} < t_{critical}$, the difference is considered to be insignificant (Hong et al., 2016; Domínguez-Tello et al., 2017).

3. Results and discussion

3.1. Data analysis

3.1.1. Statistical analysis

The ballast water taken from Dalian Port in the 21 sets of field tests had various characteristics. The ballast water parameters *T*, Sal, DO, pH, Tur, and Chl were statistically analyzed during the field tests, and the results are shown in Table S1. The salinity of the ballast water used in the experiments was close to that of high-salinity seawater, with an approximate salinity of 31.5 PSU, and the average values of DO and pH were 6.89 mg L^{-1} and 7.5, respectively. From July to November 2016, the temperature of the ballast water ranged between 11.5 and 23.7 °C. The higher temperature in summer, accompanied by the growth of algae, resulted in obvious changes in the Tur and Chl values, which ranged from 0.61 to 5.24 NTU and 3.50–58.90 μg L^{-1}, respectively.

During treatment of the ballast water by the AOP-based emergency BWTS, the active substances (such as •OH radicals) could react with halogen ions in the seawater to form mainly HOCl/OCl^-, HOBr/OBr^- (Zhang et al., 2012, 2013a,b), and other active compounds. These chemical oxidants then reacted with natural organic matter in water to form various halogenated DBPs such as THMs, etc (Jiang et al., 2017; Jiang et al., 2018; Shao et al., 2018), while brominated byproducts were the major generated DBPs because HOBr is more susceptible than HOCl to electrophilic substitution reactions with methyl ketones or organic structures (Ichihashi et al., 1999; Westerhoff et al., 2004; Sun et al., 2009; Zhang et al., 2017a,b). In the field experiments, DBCM and TBM were found to have the much higher concentrations than those of other THMs in each treatment, the highest concentration of which were 6.72 μg L^{-1} and 281.00 μg L^{-1}, respectively. Thus, DBCM and TBM were selected as the representative DBPs for analysis in this study. The measured concentrations of DBCM and TBM for model development are shown in Table S2.

3.1.2. Correlation analysis

On the basis of the 21 sets of field experimental results, the relationships between the DBPs and the experimental parameters (including the water parameters and TRO concentrations) were investigated using the Pearson correlation test (as shown in Table 1). The Pearson correlation coefficient (*r*) is one of the most widely used measure of the linear dependence between two random variables, which is formally defined as the covariance of the two variables divided by the product of their standard deviations, while *r* close to 1 indicates a positive relationship; *r* close to −1 indicates a negative relationship; and *r* close to 0 indicates the absence of relationship between the two variables (Rodgers and Nicewander, 1988; Di Lena and Margara, 2010).

Both the DBCM and TBM concentrations exhibited the highest correlations with the TRO, with significant correlations of $r = 0.911$ and $r = 0.926$, respectively, indicating that the TRO dose was the key factor directly affecting the concentrations of DBCM and TBM in the ballast water treated using the AOP-based emergency BWTS. The relationship between the TRO dose and the generation of DBCM and TBM is shown in Fig. S3. As the TRO concentration increased from 0.4 to 4.2 mg L^{-1}, the concentration of generated TBM increased significantly from 15.84 to 281.00 μg L^{-1}, whereas that of DBCM increased from 0.82 to 6.72 μg L^{-1}.

At the same TRO concentration, more TBM was generally formed than DBCM. The reason is that TBM and DBCM were formed by the reaction of dibrominated reaction intermediates with HOBr and HOCl, respectively, and the reaction rate constant for HOBr (k_{HOBr}) and the intermediates was approximately 15 times that for HOCl (k_{HOCl}) (Nokes et al., 1999). Thus, TBM was more likely to form during the treatments.

The Pearson test results also indicate that the water temperature had a positive significant correlation with the formation of DBCM ($r = 0.655$) and TBM ($r = 0.694$). A higher ballast water temperature could increase the rate of DBP formation during the oxidation process (Elshorbagy et al., 2000; Zhang et al., 2013a,b), as well as increasing the concentrations of DBP precursors (mainly algae) (Babaei et al., 2015; Huang et al., 2019). The field test results also demonstrated that the concentrations of DBCM and TBM in summer were higher than those in winter under the same ballast water treatment conditions.

The correlations *r* between Chl and the concentrations of DBCM and TBM were 0.531 and 0.565, respectively. The Chl value generally reflects the concentration in seawater of algae cells, which contain diverse and numerous organic nitrogen compounds such as proteins, peptides, amino sugars, and other organic acids, which could greatly affect DBP formation during treatment by providing one source of precursors (Li et al., 2012; Hua et al., 2018). The potential role of algae cells in DBP formation has also been considered in several studies in previous studies (Lee et al., 2015; Tomlinson et al., 2016).

In addition, the correlation between the Tur value and DBP formation was similar to that of the Chl value, reflecting their similar behavior in the field tests. The stable parameters such as the pH and DO showed less marked correlations with DBP formation. However, although Sal was a stable parameter without obvious changes during the treatments, it was found to have a strong relationship with the concentrations of DBCM ($r = 0.745$) and TBM ($r = 0.762$). That is because salinity is positively correlated with the bromide and chloride concentrations in seawater, which had significant effects on the contribution factors of the reactions in DBP formation. (Ding et al., 2013; Pan and Zhang, 2013; Shi et al., 2013).

3.2. Model development

To develop the prediction models for the evaluation of DBP formation in the AOP-based emergency BWTS, six possible

Table 1
Pearson correlation matrix.

Chemical	Pearson correlation	TRO	Ballast water parameters					
			T	Sal	DO	pH	Tur	Chl
DBCM	r	0.911**	0.655**	0.745**	0.411	0.232	0.580**	0.531*
	p	0.000	0.001	0.000	0.064	0.311	0.006	0.013
TBM	r	0.926**	0.694**	0.762**	0.513*	0.077	0.602**	0.565**
	p	0.000	0.000	0.000	0.017	0.740	0.004	0.008

* Significant value at 0.05. ** Significant value at 0.01.

combinations (A, B, C, D, E, G) of the independent variables (Chl, *T*, Sal, DO, pH, Tur, TRO) were constructed according to the Pearson test results and were assessed using the polynomial linear model and logarithmic regression model. As shown in Table 2, the combination A included three independent variables (Chl, *T*, TRO); combination B, four independent variables (Chl, *T*, Sal, TRO); combination C, four independent variables (Chl, DO, pH, Tur); combination D, four independent variables (Chl, *T*, Sal, Tur); combination E five independent variables (Chl, *T*, Sal, Tur, TRO); and combination G, seven independent variables (Chl, *T*, Sal, DO, pH, Tur, TRO).

For the polynomial linear model with the combinations including the variables (Chl, *T*, Sal, Tur, TRO) that had significant correlations with DBPs (combinations A, B, E, and G), the correlations (R^2, R^2adj) between the measured and calculated values of the DBCM concentration were higher. Combination D included almost the same variables as combination E except for TRO; however, the correlation results using these two combinations for DBCM differed greatly, indicating that TRO was a critical variable in the polynomial linear model. Correspondingly, for combinations including mainly the relatively weakly correlated variables (pH, DO, Tur), such as combination C, the result for DBCM had a relatively low correlation ($R^2 = 0.403$, R^2 adj $= 0.253$), and the standard error (SE $= 1.85$) was nearly twice that of the other combinations. As a result, the polynomial linear model with combination B, which exhibits good reproducibility ($R^2 = 0.873$, R^2 adj $= 0.841$) and the lowest standard error (SE $= 0.86$), was considered to be suitable for predicting the DBCM value. In addition, for prediction of the TBM value, the polynomial linear model with combination G was found to have acceptable reproducibility ($R^2 = 0.964$, R^2 adj $= 0.945$, SE $= 20.67$), making it suitable for further evaluation.

Logarithmic regression models with the six combinations of variables were also proposed to predict the formation of DBCM or TBM. The results for DBCM formation obtained using the logarithmic regression model were similar to those obtained using the polynomial linear model, and combination A exhibited good reproducibility ($R^2 = 0.836$, R^2 adj $= 0.807$, SE $= 0.15$). For TBM formation, the logarithmic regression model with combination G had the best reproducibility ($R^2 = 0.926$, R^2 adj $= 0.886$) and the lowest standard error (SE $= 0.16$).

On the basis of the evaluation of different combinations, combinations B (which includes four variables: *T*, Sal, Chl, and TRO) and G (which includes all the independent variables: *T*, Sal, DO, pH, Tur, Chl, and TRO) were selected to construct the polynomial linear models for DBCM [$P_{(DBCM)}$] and TBM [$P_{(TBM)}$], respectively. Further, combinations A (which includes T, Chl, and TRO) and G were selected to construct the logarithmic regression models for DBCM [$L_{(DBCM)}$] and TBM [$L_{(TBM)}$], respectively. The constructed polynomial linear models and logarithmic regression models for DBCM and TBM are summarized in Table S3.

Further statistical analyses of the developed prediction models for DBCM and TBM were performed, and the results are shown in Table 3. Both the polynomial linear and logarithmic regression models for DBCM and TBM exhibited statistical significance (sig $= 0.000 < 0.05$). For DBCM, the logarithmic regression model with three variables [$L_{(DBCM)}$] had a better AV and SE, but the Student's T-test result (0.74) was lower than that of the linear polynomial model with four variables [$P_{(DBCM)}$]. Thus, the following model $P_{(DBCM)}$ was finally selected to predict the formation of DBCM in the AOP-based emergency BWTS (see Table 4).

$$\text{Model } P_{(DBCM)}: \text{DBCM} = 316.33 + 0.25\text{T} - 10.21\text{Sal} - 0.008\text{Chl} + 1.76\text{TRO} \quad (1)$$

For TBM, the polynomial linear model with all the variables [$P_{(TBM)}$] had a higher *F* value, but the AV and SE were higher than that of the logarithmic regression model [$L_{(TBM)}$]. Consequently, the following $L_{(TBM)}$ model was selected to predict the formation of TBM.

$$\text{Model } L_{(TBM)}: \log_{10}(\text{TBM}) = 19.48 + 0.03T - 0.52\text{Sal} + 0.02\text{DO} - 0.38\text{pH} - 0.09\text{Tur} + 0.006\text{Chl} + 0.34\text{TRO} \quad (2)$$

Table 2
Model options.

Model	Com[a]	var[b]	Chl	T	Sal	DO	pH	Tur	TRO	DBCM			TBM		
										R^2	R^2 adj	SE	R^2	R^2 adj	SE
polynomial linear model	A	3	*	*					*	0.848	0.821	0.91	0.887	0.867	32.19
	B	4	*	*	*				*	0.873	0.841	0.86	0.927	0.908	26.70
	C	4	*			*	*	*		0.403	0.253	1.85	0.466	0.332	72.05
	D	4	*	*	*			*		0.555	0.444	1.60	0.588	0.485	63.27
	E	5	*	*	*			*	*	0.873	0.830	0.88	0.936	0.914	25.80
	G	7	*	*	*	*	*	*	*	0.878	0.812	0.93	0.964	0.945	20.67
logarithmic regression model	A	3	*	*					*	0.836	0.807	0.15	0.900	0.883	0.16
	B	4	*	*	*				*	0.843	0.804	0.15	0.901	0.877	0.16
	C	4	*			*	*	*		0.350	0.188	0.31	0.392	0.24	0.40
	D	4	*	*	*			*		0.567	0.459	0.25	0.656	0.571	0.30
	E	5	*	*	*			*	*	0.845	0.793	0.16	0.906	0.875	0.16
	G	7	*	*	*	*	*	*	*	0.857	0.780	0.16	0.926	0.886	0.16

[a] com: combination.
[b] var: variable.

Table 3
Statistical evaluation of DBCM and TBM models.

Model	*N*	F-test	T-test	R^2	R^2 adj	*F*	$F_{critical}$	SE	AV	Durbin–Watson	sig.
$P_{(DBCM)}$	21	0.76	0.99	0.873	0.841	27.41	3.01	0.86	2.0	2.46	0.000
$L_{(DBCM)}$	21	0.87	0.74	0.836	0.807	28.87	3.20	0.15	0.3	2.23	0.000
$P_{(TBM)}$	21	0.94	0.99	0.964	0.945	50.10	2.83	20.67	86.6	2.35	0.000
$L_{(TBM)}$	21	0.51	0.90	0.926	0.886	23.14	2.83	0.16	0.5	2.33	0.000

Table 4
Abbreviations.

DBCM	bromodichloromethane
TBM	bromoform
BWTS	ballast water treatment system
AOP	advanced oxidation process
DBPs	disinfection byproducts
ROS	reactive oxygen species
TRO	total residual oxidant
SE	standard error
AV	analytical variance
Sal	salinity
DO	dissolved oxygen content
Tur	turbidity
Chl	chlorophyll concentration

Both of the model $P_{(DBCM)}$ and model $L_{(TBM)}$ were valid at the range of following parameters: *T*, 11.5–23.7 °C; Sal, 31.3–31.6 PSU; DO, 6.52–7.68 mg L^{-1}; pH, 7.1–7.7; Tur, 0.61–5.24 NTU; Chl, 3.50–58.90 mg m^{-3}; and TRO, 0.4–4.2 mg L^{-1}.

For $P_{(DBCM)}$ and $L_{(TBM)}$, the Student's T-test results were 0.99 and 0.90, respectively (that is, > 0.5), indicating that there is no statistically significant difference between the measured and predicted values. The *F* values of the two models for DBCM and TBM were 27.41 and 23.14, respectively, which were higher than the $F_{critical}$ values for DBCM and TBM, indicating that the two models were also statistically significant at the probability level of 5%. In addition, some study found that the statistically best models had Durbin–Watson values between 1.5 and 2.5 (Uyak et al., 2007; Kumari and Gupta, 2015). In this study, the Durbin–Watson statistic was found to be 2.40 and 2.33 for the DBCM and TBM models, respectively, indicating that the two models were generally valid for predicting the formation of DBCM and TBM.

The statistical residuals of the model were examined for homoscedasticity, as shown in Fig. S4. If homoscedasticity is present, the residuals should be distributed evenly above and below zero. If it is not, it could be inferred that a calculation error was made or an additional variable should be added to the regression model. As shown in Figs. S4a and S4b, the data showed a normal distribution that fell evenly above and below the zero baseline, indicating that the models can be considered valid to describe the measured data (Baltosser, 1996). Moreover, the residuals from the prediction model of DBCM and TBM were normally distributed (see Figs. S5a and S5b), which demonstrated that the developed models were reasonable.

3.3. Model validation

The predictive ability and accuracy of the developed models for DBCM [$P_{(DBCM)}$] and TBM [$L_{(TBM)}$] were verified. Additional 29 samples were collected in field tests of ballast water treatment from May to July 2018, and the experimental data, including the water parameters and TRO concentrations as independent variables, were used for model validation. The water parameters of the 29 sets of samples were analyzed statistically during the field tests, and the results (Table S1) are summarized as follows: *T*, 19.2–23.7 °C; Sal, 31.5–31.7 PSU; DO, 4.78–7.98 mg L^{-1}; pH, 7.4–7.8; Tur, 1.27–4.91 NTU; and Chl, 6.50–48.80 mg m^{-3}. The TRO concentrations in the 29 sets of ballast water treatments ranged from 1.0 to 4.2 mg L^{-1}. All the variables met the design basis of the developed models and were used for validation.

The DBCM and TBM concentrations were calculated via the developed models, $P_{(DBCM)}$ and $L_{(TBM)}$, respectively, using the 29 sets of experimental data. Fig. 3 compares the experimental and predicted DBCM and TBM concentrations. Approximately 93% of

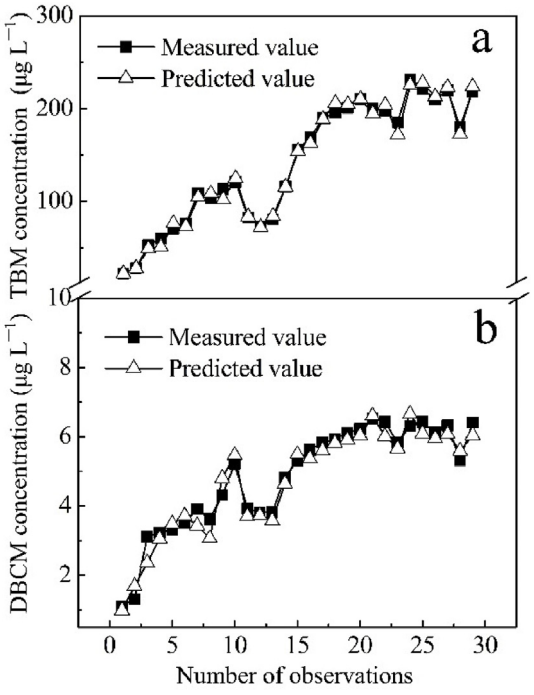

Fig. 3. Validation of the model: predicted and measured values of TBM (a) and DBCM (b) concentrations.

the predicted DBCM values fall within ±20% of the measured ones, and only small fractions of extreme predicted values exhibit a high percentage difference of ±30% of the measured ones. For Model $L_{(TBM)}$, 100% of the predicted values fall within ±20%, and approximately 97% fall within ±10% of the measured ones. Studies have demonstrated that the difference between the values predicted by a model should fall strictly within ±20% of the measured values (Golfinopoulos and Arhonditsis, 2002). The comparison results indicate that the developed models, $P_{(DBCM)}$ and $L_{(TBM)}$, could provide accurate estimations of DBCM and TBM formation.

To better validate the models, the T-test was applied to determinate the differences between the predicted and measured values. For Models $P_{(DBCM)}$ and $L_{(TBM)}$, t_{value} was 1.97 and 0.14, respectively; both of these values were less than $t_{critical}$, indicating that the model differences were not significant. Moreover, the significance levels of $P_{(DBCM)}$ and $L_{(TBM)}$ in the T-test were 0.059 and 0.890, respectively, which were larger than 0.05, indicating that the measured values are not obviously different from the predicted ones.

Furthermore, internal evaluations of Models $P_{(DBCM)}$ and $L_{(TBM)}$ were performed. The results (Table S4 and Fig. 4) indicate satisfactory predictions of DBCM ($R^2 = 0.963$, R^2 adj = 0.962) and TBM ($R^2 = 0.999$, R^2 adj = 0.999). Because an ideal simulation corresponds to an intercept of 0, a slope of 1, and an R^2 value of 1.0 (Sohn et al., 2004), the results demonstrate that the two prediction models have good accuracy.

4. Conclusions

On the basis of 21 sets of field experimental results of ballast water treatments performed using an AOP-based emergency BWTS in Dalian Port, China, a polynomial linear model [Model $P_{(DBCM)}$] with four variables (*T*, Sal, Chl, and TRO) and a logarithmic regression model [Model $L_{(TBM)}$] with seven variables (*T*, Sal, DO, pH, Tur, Chl, and TRO) were developed to predict the formation of DBCM

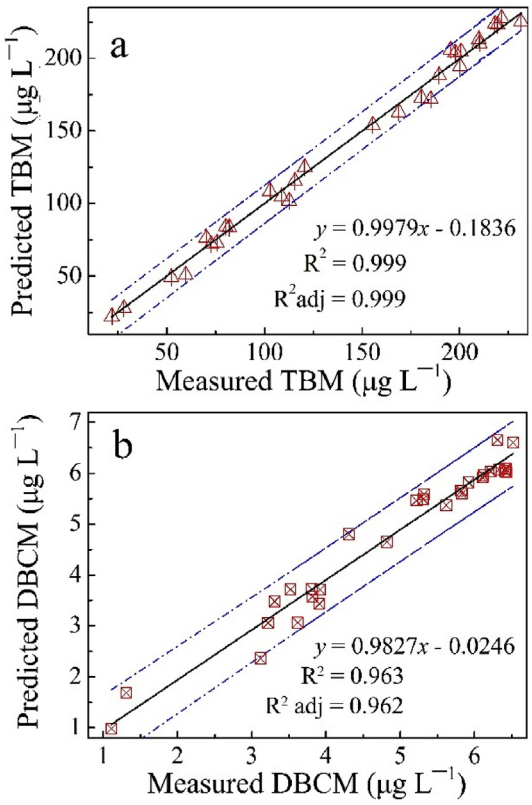

Fig. 4. Relationships between the predicted and measured levels of TBM (a) and DBCM (b).

and TBM, respectively, by the treatment system. The results of a systematic evaluation revealed that the two models had good reproducibility [$R^2 = 0.873$, SE = 0.86 for Model $P_{(DBCM)}$; $R^2 = 0.926$, SE = 0.16 for Model $L_{(TBM)}$], and no statistically significant difference was found between the measured and predicted values. The validation results for Models $P_{(DBCM)}$ and $L_{(TBM)}$ showed that approximately 93% and 100% of the predicted values fell within ±20% of the measured ones, respectively, demonstrating that the two developed models have excellent predictive ability and could be used to accurately and quantitatively predict the concentrations of DBCM and TBM produced by the AOP-based emergency BWTS. However, these models are restricted to field data lying within the ranges of the specific data sets. The relationships between the DBP concentration and the raw water characteristics require further investigation for more general application. These developed models could contribute to engineering applications of AOP-based emergency BWTSs to prevent pollution of the marine environment during seawater treatment in different ports.

Declaration of interests

No conflict of interest exits in the submission of this manuscript.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (61427804, 51877024), the Public Science and Technology Research Funds Projects of Ocean (201305027), and the High Level Talent Innovation Project of Dalian (2016RQ040).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.envpol.2019.113028.

References

Amy, G., Chadik, P., Chowdhury, Z., 1987. Developing models for predicting trihalomethane formation potential and kinetics. J. Am. Water Works Assoc. 79, 89–97.
Babaei, A.A., Atari, L., Ahmadi, M., Ahmadiangali, K., Zamanzadeh, M., Alavi, N., 2015. Trihalomethanes formation in Iranian water supply systems: predicting and modeling. J. Water Health 13, 859–869.
Bai, M., Zhang, Z., Xue, X., Yang, X., Hua, L., Fan, D., 2010. Killing effects of hydroxyl radical on algae and bacteria in ship's ballast water and on their cell morphology. Plasma Chem. Plasma Process. 30, 831–840.
Bai, M., Zhang, Z., Zhang, N., Tian, Y., Chen, C., Meng, X., 2012. Treatment of 250 t/h ballast water in oceanic ships using •OH radicals based on strong electric-field discharge. Plasma Chem. Plasma Process. 32, 693–702.
Bai, M., Zheng, Q., Tian, Y., Zhang, Z., Chen, C., Cheng, C., Meng, X., 2016. Inactivation of invasive marine species in the process of conveying ballast water using OH based on a strong ionization discharge. Water Res. 96, 217–224.
Bai, M., Tian, Y., Yu, Y., Zheng, Q., Zhang, X., Zheng, W., Zhang, Z., 2018. Application of a hydroxyl-radical-based disinfection system for ballast water. Chemosphere 208, 541–549.
Baltosser, W.H., 1996. Biostatistical analysis by Jerrold H. Zar. Ecology 77, 2266–2267.
Casas-Monroy, O., Chan, P., Linley, R.D., Vanden Byllaardt, J., Kydd, J., Bailey, S.A., 2016. Comparison of three techniques to evaluate the number of viable phytoplankton cells in ballast water after ultraviolet irradiation treatment. J. Appl. Phycol. 28, 2821–2830.
Chen, B., Westerhoff, P., 2010. Predicting disinfection by-product formation potential in water. Water Res. 44, 3755–3762.
Chowdhury, S., Champagne, P., McLellan, P.J., 2009. Models for predicting disinfection byproduct (DBP) formation in drinking waters: a chronological review. Sci. Total Environ. 407, 4189–4206.
Domínguez-Tello, A., Arias-Borrego, A., García-Barrera, T., Gómez-Ariza, J.L., 2017. A two-stage predictive model to simultaneous control of trihalomethanes in water treatment plants and distribution systems: adaptability to treatment processes. Environ. Sci. Pollut. Res. 24, 22631–22648.
Elshorbagy, W., Abu-Qdais, H., Elsheamy, M., 2000. Simulation of THM species in water distribution systems. Water Res. 34, 3431–3439.
Engerholm, B.A., Amy, G.L., 1983. A predictive model for chloroform formation from humic acid. J. Am. Water works Assoc. 75, 418–423.
Fang, G., Zhu, C., Dionysiou, D.D., Gao, J., Zhou, D., 2015. Mechanism of hydroxyl radical generation from biochar suspensions: implications to diethyl phthalate degradation. Bioresour. Technol. 176, 210–217.
Ged, E.C., Boyer, T.H., 2014. Effect of seawater intrusion on formation of bromine-containing trihalomethanes and haloacetic acids during chlorination. Desalination 345, 85–93.
Golfinopoulos, S.K., Arhonditsis, G.B., 2002. Multiple regression models: a methodology for evaluating trihalomethane concentrations in drinking water from raw water characteristics. Chemosphere 47, 1007–1018.
Hao, R., Zhang, Y., Du, T., Yang, L., Adeleye, A.S., Li, Y., 2017. Effect of water chemistry on disinfection by-product formation in the complex surface water system. Chemosphere 172, 384–391.
Hong, H., Song, Q., Mazumder, A., Luo, Q., Chen, J., Lin, H., Yu, H., Shen, L., Liang, Y., 2016. Using regression models to evaluate the formation of trihalomethanes and haloacetonitriles via chlorination of source water with low SUVA values in the Yangtze River Delta region, China. Environ. Geochem. Health 38, 1303–1312.
Hua, L., Lin, J., Syue, M., Huang, C., Chen, P., 2018. Optical properties of algogenic organic matter within the growth period of Chlorella sp. and predicting their disinfection by-product formation. Sci. Total Environ. 621, 1467–1474.
Huang, R., Liu, Z., Yan, B., Zhang, J., Liu, D., Xu, Y., Wang, P., Cui, F., Liu, Z., 2019. Formation kinetics of disinfection byproducts in algal-laden water during chlorination: a new insight into evaluating disinfection formation risk. Environ. Pollut. 245, 63–70.
Ichihashi, K., Teranishi, K., Ichimura, A., 1999. Brominated trihalomethane formation in halogenation of humic acid in the coexistence of hypochlorite and hypobromite ions. Water Res. 33, 477–483.
IMO, 2008a. International Maritime Organization. Guidelines for approval of ballast water management systems (G8). Annex 4 Resolution MEPC 174 (58), 3–5.
IMO, 2008b. Procedure for approval of ballast water management systems that make use of active substances (G9). Annex 1Resolution MEPC 169 (57), 3–10.
Kornmueller, A., 2007. Review of fundamentals and specific aspects of oxidation technologies in marine waters. Water Sci. Technol. 55, 1–6.
Kumari, M., Gupta, S.K., 2015. Modeling of trihalomethanes (THMs) in drinking water supplies: a case study of eastern part of India. Environ. Sci. Pollut. Res. 22, 12615–12623.
Lee, J., Choi, E.J., Rhie, K., 2015. Validation of algal viability treated with total residual oxidant and organic matter by flow cytometry. Mar. Pollut. Bull. 97, 95–104.
Li, L., Gao, N., Deng, Y., Yao, J., Zhang, K., 2012. Characterization of intracellular &

extracellular algae organic matters (AOM) of Microcystic aeruginosa and formation of AOM-associated disinfection byproducts and odor & taste compounds. Water Res. 46, 1233–1240.
Li, Y., Zhang, X., Yang, M., Liu, J., Li, W., Graham, N.J.D., Li, X., Yang, B., 2017. Three-step effluent chlorination increases disinfection efficiency and reduces DBP formation and toxicity. Chemosphere 168, 1302–1308.
Miklos, D., Remy, C., Jekel, M., Linden, K., Drewes, J., Hübner, U., 2018. Evaluation of advanced oxidation processes for water and wastewater treatment – a critical review. Water Res. 139, 118–131.
Min, J., Min, K., 2016. Blood trihalomethane levels and the risk of total cancer mortality in US adults. Environ. Pollut. 212, 90–96.
Nokes, C., Fenton, E., Randall, C., 1999. Modelling the formation of brominated trihalomethanes in chlorinated drinking waters. Water Res. 33, 3557–3568.
Owen, L.U., Krasner, S.W., Liang, S., 2011. Modeling approach to treatability analyses of an existing treatment plant. Journal 103, 103–117.
Serodes, J.B., Rodriguez, M.J., Li, H., Bouchard, C., 2003. Occurrence of THMs and HAAs in experimental chlorinated waters of the Quebec City area (Canada). Chemosphere 51, 253–263.
Shah, A.D., Liu, Z., Salhi, E., Höfer, T., Werschkun, B., von Gunten, U., 2015. Formation of disinfection by-products during ballast water treatment with ozone, chlorine, and peracetic acid: influence of water quality parameters. Environ. Sci.: Water Res. Technol. 1, 465–480.
Shao, Y., Pan, Z., Rong, C., Wang, Y., Zhu, H., Zhang, Y., Yu, K., 2018. 17 beta-estradiol as precursors of Cl/Br-DBPs in the disinfection process of different water samples. Environ. Pollut. 241, 9–18.
Shi, H., Qiang, Z., Adams, C., 2013. Formation of haloacetic acids, halonitromethanes, bromate and iodate during chlorination and ozonation of seawater and saltwater of marine aquaria systems. Chemosphere 90, 2485–2492.
Siddiqui, M.S., Amy, G.L., 1993. Factors affecting DBP formation during ozone-bromide reactions. J. Am. Water Works Assoc. 85, 63–72.
Sohn, J., Amy, G., Cho, J., Lee, Y., Yoon, Y., 2004. Disinfectant decay and disinfection by-products formation model development: chlorination and ozonation by-products. Water Res. 38, 2461–2478.
Stefan, M., 2017. Advanced oxidation processes for water treatment - fundamentals and applications. Water Intell. Online 16, 493–533.
Stehouwer, P.P., Buma, A., Peperzak, L., 2015. A comparison of six different ballast water treatment systems based on UV radiation, electrochlorination and chlorine dioxide. Environ. Technol. 36, 2094–2104.
Sun, Y., Wu, Q., Hu, H., Tian, J., 2009. Effect of bromide on the formation of disinfection by-products during wastewater chlorination. Water Res. 43, 2391–2398.
Tian, Y., Yuan, X., Xu, S., Li, R., Zhou, X., Zhang, Z., 2015. Biological efficacy and toxic effect of emergency water disinfection process based on advanced oxidation technology. Ecotoxicology 24, 2141–2150.
Tomlinson, A., Drikas, M., Brookes, J.D., 2016. The role of phytoplankton as precursors for disinfection by-product formation upon chlorination. Water Res. 102, 229–240.
US EPA 524.2, 1995. Measurement of purgeable organic compounds in water by capillary column gas chromatography/mass spectrometry. Revision 4.1. Using regression models to evaluate the formation of trihalomethanes and haloacetonitriles via chlorination of source water with low SUVA values in the Yangtze river Delta region, China. Environ. Geochem. Health 38, 1303–1312.
USEPA, 1978. Chlorine, Total Residual (Spectrophotometric, DPD) (CAS No. 7782-50-5), Method 330.5. Environmental Protection Agency (USEPA).
Uyak, V., Ozdemir, K., Toroz, I., 2007. Multiple linear regression modeling of disinfection by-products formation in Istanbul drinking water reservoirs. Sci. Total Environ. 378, 269–280.
Wang, J., Xu, L., 2012. Advanced oxidation processes for wastewater treatment: formation of hydroxyl radical and application. Crit. Rev. Environ. Sci. Technol. 42, 251–325.
Werschkun, B., Sommer, Y., Banerji, S., 2012. Disinfection by-products in ballast water treatment: an evaluation of regulatory data. Water Res. 46, 4884–4901.
Westerhoff, P., Chao, P., Mash, H., 2004. Reactivity of natural organic matter with aqueous chlorine and bromine. Water Res. 38, 1502–1513.
Yu, Z., Zhang, Z., Xu, S., Zhang, Y., Liu, P., Tian, Y., 2017. Partitioned operation method for reactive oxygen species reactor array at atmospheric pressure. Plasma Chem. Plasma Process. 37, 475–487.
Zeng, Q., Zhou, B., He, D., Wang, Y., Wang, M., Yang, P., Huang, Z., Li, J., Lu, W., 2016. Joint effects of trihalomethanes and trichloroacetic acid on semen quality: a population-based cross-sectional study in China. Environ. Pollut. 212, 544–549.
Zhang, N., Zhang, Z., Bai, M., Chen, C., Meng, X., Tian, Y., 2012. Evaluation of the ecotoxicity and biological efficacy of ship's ballast water treatment based on hydroxyl radicals technique. Mar. Pollut. Bull. 64, 2742–2748.
Zhang, N., Ma, B., Li, J., Zhang, Z., 2013a. Factors affecting formation of chemical by-products during ballast water treatment based on an advanced oxidation process. Chem. Eng. J. 231, 427–433.
Zhang, Y., Bai, M., Chen, C., Meng, X., Tian, Y., Zhang, N., Yu, Z., 2013b. •OH treatment for killing of harmful organisms in ship's ballast water with medium salinity based on strong ionization discharge. Plasma Chem. Plasma Process. 33, 751–763.
Zhang, N., Zhang, Y., Bai, M., Zhang, Z., Chen, C., Meng, X., 2014. Risk assessment of marine environments from ballast water discharges with laboratory-scale hydroxyl radicals treatment in Tianjin Harbor, China. J. Environ. Manag. 145, 122–128.
Zhang, R., Zhang, Y., Rong, C., Song, Y., Wang, Y., Pei, J., Tang, X., Yu, K., 2017a. Oxidation of the antibacterial agent norfloxacin during sodium hypochlorite disinfection of marine culture water. Chemosphere 182, 245–254.
Zhang, X., Bai, M., Tian, Y., Du, H., Zhang, Z., 2017b. The estimation for ballast water discharged to China from 2007 to 2014. Mar. Pollut. Bull. 124, 89–93.
Zhang, Y., Tian, Y., Zhang, Z., Lin, S., 2018. Experimental and numerical study of cavitating flow with suction in a mixing reactor for water treatment. Chem. Eng. J. 353, 796–804.
Zhao, C., Arroyo-Mora, L.E., DeCaprio, A.P., Sharma, V.K., Dionysiou, D.D., O'Shea, K.E., 2014. Reductive and oxidative degradation of iopamidol, iodinated X-ray contrast media, by Fe(III)-oxalate under UV and visible light treatment. Water Res. 67, 144–153.
Ziegler, G., Tamburri, M.N., Fisher, D.J., 2018. Long-term algal toxicity of oxidant treated ballast water. Mar. Pollut. Bull. 133, 18–29.