RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2021, **11**, 28476

Source identification and characteristics of dissolved organic matter and disinfection by-product formation potential using EEM-PARAFAC in the Manas River, China

Xinlin Wang, Yanbin Tong, Qigang Chang,* Jianjiang Lu,* Teng Ma, Fangdong Zhou and Jiaqi Li

Dissolved organic matter (DOM) is ubiquitous in natural water and reacts with disinfectants to form disinfection by-products (DBPs). The analysis of DOM in raw water is helpful in evaluating the formation potential of DBPs. However, there is relatively little research on the DOM identification of raw water in northern China. In this study, the sources and characteristics of DOM were investigated in the samples collected from the Manas River. Dissolved organic carbon (DOC), UV_{254}, specific ultraviolet absorbance, and fluorescence indices (fluorescence index, humification index, and biological index) were measured to characterize the DOM, and trihalomethanes (THMs) were quantified following formation potential tests with free chlorine. The maximum amount of total trihalomethane formation potential (THMsFP) was 225.57 μg L^{-1}. The DOM of the Manas River consisted of microorganisms and soil resources. The excitation–emission matrix combined with parallel factor analysis (EEM-PARAFAC) identified microbial humus (C1, 54%) and tryptophan-like protein (C2, 46%). PARAFAC components were evaluated as the precursor surrogate parameters of THMsFP. Additionally, the linear THMsFP correlation was stronger with C1 + C2 (r = 0.529, p < 0.01) than with C1 (r = 0.485, p < 0.01). Thus, C1 + C2 is an accurate THMsFP precursor surrogate parameter for the Manas River, and the use of fluorescence spectroscopy may be a robust alternative for predicting DOC removal.

Received 5th May 2021
Accepted 28th July 2021

DOI: 10.1039/d1ra03498g

rsc.li/rsc-advances

1. Introduction

Dissolved organic matter (DOM) is defined as the fraction of organic matter that passes through a 0.45 mm filter; it is ubiquitous in surface water bodies and poses a challenge to the effectiveness of water treatment processes. Until now, disinfection has been effective for inactivating pathogenic microorganisms in drinking water treatment.[1–4] However, commonly used disinfectants can react with the organic matter in water to generate diverse disinfection by-products (DBPs), such as trihalomethanes (THMs) and haloacetic acids (HAAs), some of which have been reported to have adverse effects on human health, inducing cytotoxicity, genotoxicity, teratogenicity, and carcinogenicity.[5–7] Therefore, a comprehensive understanding of DOM composition and effective removal of DBP is essential to ensure the success of a water treatment plant and to achieve the desired quality for intended water use.

DOM is the main precursor of DBP. Given its complex composition and low concentration, surrogate parameters were developed to predict the removal of DOM through water treatment and estimate its reactivity toward DBP formation.[8,9] These surrogate parameters include dissolved organic carbon (DOC), ultraviolet (UV) absorbance, and specific ultraviolet absorbance (SUVA).[10,11] However, they provide limited information and are time consuming.

Excitation–emission matrix (EEM) fluorescence spectroscopy is a rapid, sensitive, and informative method, which has been widely applied to trace the quality and quantity of DOM. It has been proven to have the potential to assess DBP formation potentials (FPs).[12–14] An EEM spectrum is constructed using emission wavelength, excitation wavelength, and fluorescence intensity. However, it contains a vast amount of fluorophore information, and this is difficult to unscramble and characterize because of the influence of fluorescence overlap and matrix effect. On the other hand, a fluorescence EEM coupled with parallel factor analysis (EEM-PARAFAC) can efficiently extract the information contained in EEMs and eliminate interference among components. Therefore, it can provide better insights into the DOM variation in the water treatment processes,[15] and has been applied to a variety of natural and engineered systems. The PARAFAC model can be applied to analyze large data sets of hundreds or thousands of EEMs,

School of Chemistry and Chemical Engineering, Shihezi University, Key Laboratory of Environmental Monitoring and Pollutant Control of Xinjiang Bingtuan, Xinjiang, 832003, China. E-mail: qgchang@gmail.com; lujianjiang2015@163.com

making statistical analysis more reliable. It makes it possible to measure the fluorescence characteristics of various water sources and helps to identify the effective fluorescent groups in organic substances, including humic acid, proteins, fulvic acid, *etc.* EEM-PARAFAC method divides fluorescent organic compounds into several independent components according to their unique properties and structures, which provides a basis for further understanding the dynamic changes of fluorescent components.[16–20] The intensities and the relative distributions of PARAFAC components are representative of their levels and the chemical composition of fluorescent DOM, and can be developed as surrogates for conventional water quality parameters, treatability of DOM, DBPsFPs, and performance of treatment processes.[11,21–24] For example, humic-like components identified by PARAFAC were linked to different sources of DOM (*e.g.*, terrestrial and microbial) and were shown to have dissimilar treatability during water and waste treatment.[21,25,26] Some studies have indicated a high correlation between trihalomethane formation potential (THMsFP) and some PARAFAC components in natural aquatic environments[27,28] and in laboratory-scale water treatments.[29] In another study, five EEM-PARAFAC components, including four terrestrially derived humic-like substances (C1, C2, C3, and C4) and one tryptophan-like substance (C5), were identified in the sub-alpine Lake Tiancai.[25] Although the application of the EEM-PARAFAC in a full-scale drinking water treatment plant of Australia has already been reported, it is noteworthy that the ambient climate (*e.g.*, flow rate, rainfall pattern, and temperature) largely governs the relative chemical composition of DOM.[30] This indicates that the specific water quality differs from source to source, and the composition of DOM may be quite different even though similar fluorescence levels may be present on the maps, thus resulting in varying correlations. As such, the values of EEM-PARAFAC for tracking DOM, DBPsFP, and treatment system performance require further discussion.

The main objects of this study were (1) to investigate the variation in DOM characteristics and DBP precursors in the water source of drinking water plants by fluorescence PARAFAC and (2) to determine the relationships between DOM spectroscopic properties and DBPFPs to assess the application potential of fluorescence spectroscopy for monitoring DBPs in raw water, providing valuable information for realtime monitoring of DBP, and optimization of the drinking water treatment process.

2. Methods

2.1. Study area and sample collection

The Manas River (43°24′–45°12′ N, 85°41′–86°32′ E) is located in the northern part of Xinjiang, China. It originates in the Tian Shan Mountains and is an inland river. Its ecological status plays a key role in the development of the economic belt of the northern slope of Xinjiang and the industrial, agricultural, and trade development of the Urumqi, Changji, and Shihezi urban agglomerations. This study used a section of the river that flows through Shihezi City, which is approximately 30.4 km long. As shown in Fig. 1, 12 sampling sites (M01–M12) were established in this section of the river, which were divided into upstream (M01–M04), midstream (M05–M08), and downstream (M09–M12). Samples were collected from October to December in 2020 and once per month. The samples were collected from surface water (0–0.5 m) with a sampling device, each sampling cite close to the edge of the river. Three replicates were collected

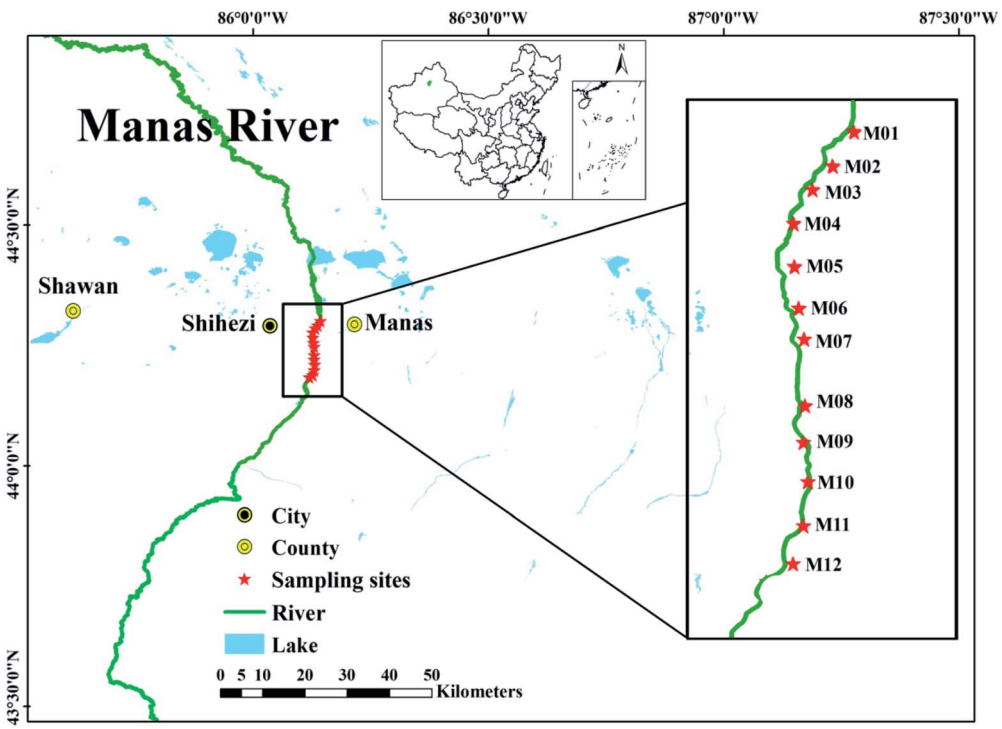

Fig. 1 Sampling map showing the Manas River Basin, China.

in 1.25 L clean plastic bottles from each sampling site, and immediately transferred to the laboratory. All the samples were stored at 4 °C before further treatment and analysis. Pretreatment of the samples included filtration through a pre-washed 0.45 μm cellulose acetate membrane filter. All analytical characterizations were performed within 24 h of sample collection.

2.2. Analytical characterization

Electrical conductivity (EC), temperature, and pH of the samples were measured on site using portable meters (2100Q and HQ40d, HACH, Japan). The NH_3–N and total dissolved solids (TDS) concentrations were determined using the standard examination methods for drinking water parameters, China (GB/T 5750.5-2006). The DOC concentrations were measured using a TOC-VWP total organic carbon analyzer (Shimadzu, Japan). Samples were acidified before measurement to purge inorganic carbon. Therefore, the amount of non-purgeable organic carbon (NPOC) was equal to the DOC. Furthermore, UV absorbance at 254 nm (UV_{254}) was measured using a Varian Cary 50 UV-Vis spectrophotometer and by consulting Standard Method 5910. $SUVA_{254}$ was calculated as the 100-fold ratio of UV_{254} to DOC, which correlated positively with the aromaticity of DOM.

2.3. EEM and EEM-PARAFAC analysis

EEM fluorescence spectroscopy was conducted using a Hitachi F-4700 luminescence spectrometer equipped with a xenon excitation source. The scan rate was set at 1200 nm min^{-1}, and the excitation/emission slits were 5 nm band-pass. The scanning field was set at emission (Em) spectra ranging from 280 to 550 nm and excitation (Ex) spectra from 200 to 400 nm, with an interval of 5 nm. Milli-Q water was used as a control and was subtracted to eliminate the effect of Raman scattering.

Fluorescence data were modeled in MATLAB software using the DOMFluor toolbox.[31] Outlier samples, which were identified using leverage comparison, were removed, and PARAFAC models with two to eight components were generated. Split-half analysis was used to determine the number of components. The fluorescence intensity of each component (*i.e.*, F_{max} of C1, C2, or C3) was used to represent the relative concentration. Additionally, three common DOM quality indices were obtained from the EEMs: (1) biological index (BIX), which is indicative of the relative significance of microbial or biological DOM, (2) humification index (HIX), which is an indicator of humification degree or humic character, and (3) fluorescence index (FI) to distinguish algal and microbial sources from terrestrial DOM sources. The BIX was determined as the ratio of Em ranging between 380 nm and 430 nm with Ex at 310 nm; HIX was obtained by dividing the integrated intensity from Em between 435 and 480 nm by that of Em ranging between 300 and 345 nm at Ex 254 nm; FI is the ratio of the fluorescence intensity at Em 450 nm to that at Em 500 nm, given that Ex = 370 nm.[13,32,33] Correlation analyses were conducted using SPSS 22.0. Correlation coefficients (r), rather than R^2, were used to differentiate negative and positive correlations.

2.4. DBP analysis

Chlorine disinfection experiments were performed in dark brown bottles at 25 °C for three days. Phosphate solution (KH_2PO_4/NaOH) was used to maintain the pH of the sample at 7.0 ± 0.1. To ensure the presence of measurable residual chlorine in each sample, chlorine was added at a dose of 5 mg Cl_2 per mg C.[14] Sodium hypochlorite (NaOCl, Sinopharm Chemical Reagent Co., China) was used as a disinfectant. The residual chlorine was quenched using ascorbic acid, and the samples were collected for DBP analysis.

Typical DBPs include chloroform (TCM), dichlorobromomethane (DCBM), dibromochloromethane (DBCM), and tribromomethane (TBM). These were extracted and measured according to the U.S. EPA Method 551.1.[34] All DBPs were determined using a gas chromatography/electron capture detector (GC-ECD; 7890A, Agilent, USA).

3. Results and discussion

3.1. DOM quantity and quality

Raw water absorbance (UV_{254}), DOC, pH, chloride, conductivity, TDS, NH3–N, and TN (total nitrogen) are summarized in Table 1. Between October and December, Manas River had a pH

Table 1 Summary of water characteristics of surface water samples collected from Manas River

Sample/index	Absorbance (UV_{254})	DOC (mg L^{-1})	pH	Chloride (mg L^{-1})	Conductivity (ms m^{-1})	TDS (mg L^{-1})	NH_3–N (mg L^{-1})	TN
M01	0.030 ± 0.05	4.698 ± 2.29	7.40 ± 0.12	5.33 ± 0.54	38.6 ± 1.03	302.5 ± 2.11	0.66 ± 0.07	0.077 ± 0.07
M02	0.027 ± 0.11	6.069 ± 3.35	7.45 ± 0.08	5.38 ± 1.13	37.5 ± 0.98	277.5 ± 1.09	0.62 ± 0.13	0.129 ± 0.11
M03	0.031 ± 0.03	6.707 ± 2.14	7.39 ± 0.18	4.85 ± 0.66	37.6 ± 1.24	276.2 ± 1.26	0.75 ± 0.06	0.031 ± 0.14
M04	0.028 ± 0.04	6.229 ± 1.37	7.50 ± 0.21	4.75 ± 1.47	36.0 ± 2.07	268.7 ± 2.57	0.68 ± 0.22	0.147 ± 0.12
M05	0.027 ± 0.05	6.263 ± 0.16	7.39 ± 0.16	5.08 ± 1.15	37.6 ± 1.63	251.2 ± 1.94	0.64 ± 0.14	0.048 ± 0.15
M06	0.025 ± 0.13	5.314 ± 2.85	7.43 ± 0.17	5.17 ± 0.84	37.2 ± 1.09	321.3 ± 1.57	0.63 ± 0.06	0.041 ± 0.23
M07	0.026 ± 0.06	6.000 ± 2.73	7.54 ± 0.25	5.25 ± 0.61	37.4 ± 0.77	321.3 ± 3.02	0.64 ± 0.17	0.082 ± 0.17
M08	0.031 ± 0.16	4.489 ± 2.66	7.56 ± 0.09	5.48 ± 0.57	37.1 ± 0.85	253.8 ± 4.18	0.62 ± 0.09	0.030 ± 0.05
M09	0.027 ± 0.13	5.247 ± 1.57	7.56 ± 0.23	5.50 ± 1.29	46.0 ± 1.25	240.0 ± 1.04	0.67 ± 0.27	0.036 ± 0.04
M10	0.026 ± 0.04	5.558 ± 1.44	7.60 ± 0.19	5.30 ± 0.79	34.4 ± 2.44	228.7 ± 1.19	0.67 ± 0.19	0.151 ± 0.19
M11	0.025 ± 0.23	5.287 ± 3.08	7.51 ± 0.13	5.02 ± 1.22	34.7 ± 1.72	202.5 ± 0.99	0.66 ± 0.04	0.075 ± 0.03
M12	0.027 ± 0.19	5.294 ± 0.79	7.56 ± 0.26	5.38 ± 1.01	38.3 ± 1.88	238.7 ± 1.28	0.64 ± 0.05	0.175 ± 0.21
Mean	0.030 ± 0.14	5.596 ± 1.08	7.49 ± 0.77	5.34 ± 1.86	38.6 ± 2.03	310.0 ± 3.01	0.52 ± 0.21	0.085 ± 0.37

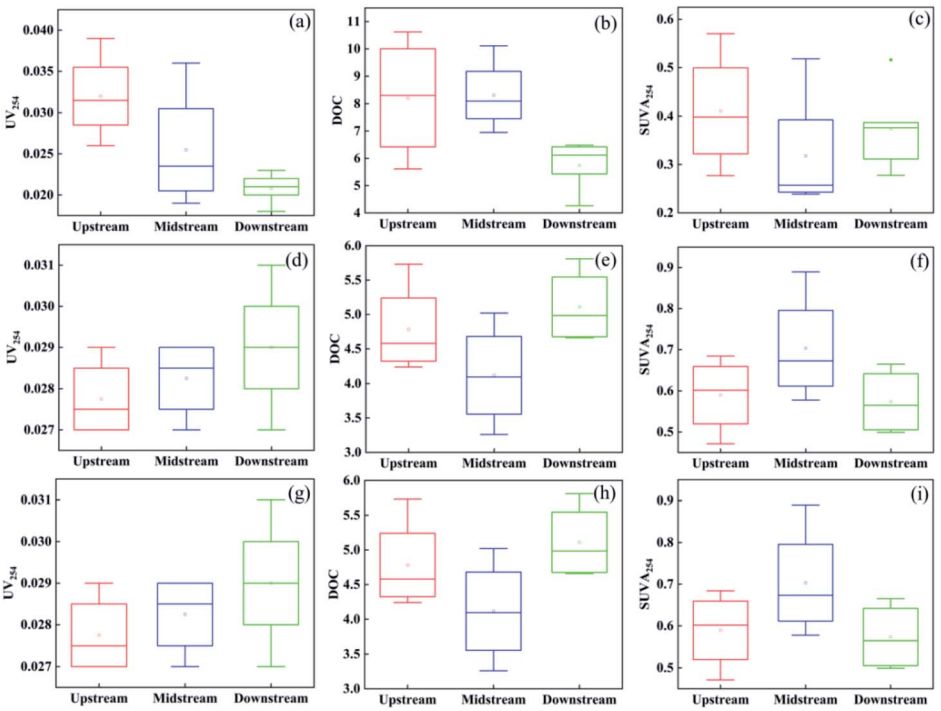

Fig. 2 Box plot of UV_{254}, DOC, and SUVA in the Manas river (a–c) October; (b–f) November; (g–i) December. Upstream: sample site 1–4; midstream: sample site 5–8; downstream: sample site 8–12.

ranging from 7.39 to 7.60, indicating a weakly alkaline environment. It had an aromatic content of 0.030 as measured by UV_{254}, a mean DOC of 5.596 mg L^{-1}, a chloride value of 5.34 mg L^{-1}, and a conductivity of 38.6 ms m^{-1}.

Fig. 2 and 3 shows the DOC content and UV_{254} values of surface water samples collected from Manas River. In October, the average DOC levels in the upstream, midstream, and downstream of Manas River were 8.212, 8.315, and 5.820 mg L^{-1}, respectively, and the average values of UV_{254} were 0.032, 0.026, and 0.021 AU cm^{-1}, respectively. The highest SUVA content was upstream (0.411), while the SUVA content midstream and downstream were 0.318 and 0.370, respectively.

In November, the DOC content downstream of the Manas River was 5.11, which was similar to the value estimated in October, while the DOC levels of upstream and midstream decreased to 4.783 and 4.118, respectively. The average UV_{254}

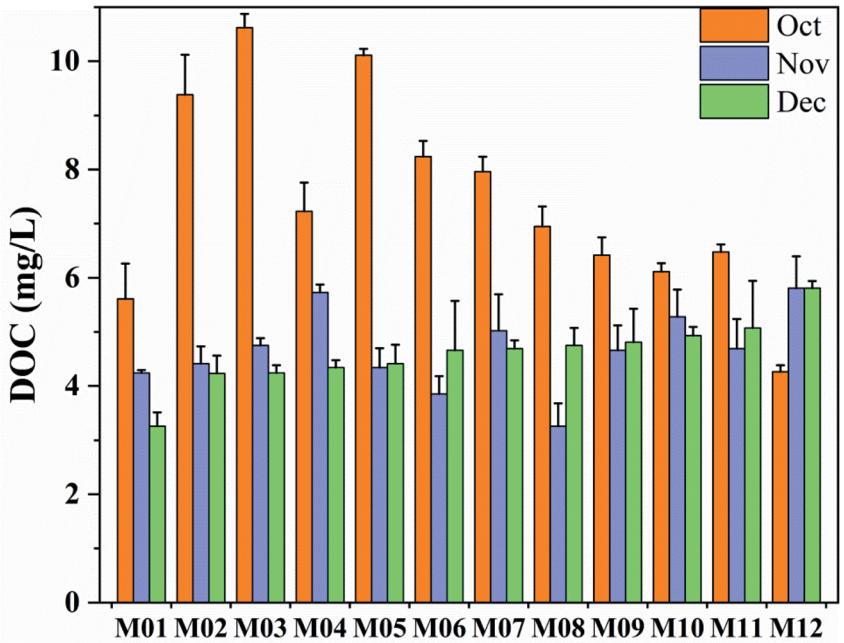

Fig. 3 Seasonal changes of DOC content in the Manas River.

values upstream, midstream, and downstream were 0.028, 0.028, and 0.029 AU cm^{-1}, respectively. In December, the average DOC levels upstream, midstream, and downstream were 4.02, 4.63, and 5.16 mg L^{-1}, respectively, and the average UV_{254} values were 0.027, 0.028, and 0.030 AU cm^{-1}, respectively. Furthermore, in December, the maximum SUVA upstream was 0.68, followed by midstream (0.61) and downstream (0.57).

The average DOC content in Manas River (5.596 mg L^{-1}) was higher than that of Mississippi River (3.55 mg L^{-1}), Yellow River (2.35 mg L^{-1}), and Poyang Lake (3.5 mg L^{-1}).[35–37] In terms of spatial trend, the DOC content displayed an expected decreasing trend from upstream to downstream because the downstream regions of the Manas River are mainly rural and agricultural areas, whereas the urban and industrial areas are located in the upstream regions.[38] The values upstream may also be higher owing to the contribution of domestic sewage and industrial wastewater to the DOC content in the upstream, as well as the water flow from the Jiahezi Reservoir upstream. Owing to the low drift and long retention time, the water quality of the reservoir undergoes complex physical, chemical, and biological changes; it usually contains a high DOC content.[39] There are a few villages and towns near the middle and lower reaches. Although they contribute to the organic matter content in the Manas River, they have little impact. In terms of temporal changes, the DOC was significantly low in November and December, as compared to October (Fig. 3). This could be because the impact of precipitation on the DOC in the water body is weakened in peak winter, that is, as the river reaches a withered water period, the input of external DOC is greatly reduced.[40,41] Meanwhile, as water temperature ranges from 0.9 °C to 6.7 °C in winter, this low temperature further reduces the DOC produced on site. All these results indicate that the Manas River contains a higher amount of DOC, which is the main component of naturally DOM.[42,43] Therefore, the material structure and content of DOM exhibit spatial and temporal differences. In other words, higher DOC levels were observed in October and upstream of the river, and domestic activities frequently affected the DOC content in Manas River.

The SUVA index is an indicator of molecular weight and is used to characterize the aromaticity of DOM.[44] It has been reported that an SUVA value less than 3 generally indicates that the majority of DOM consists of more aliphatic compounds, while SUVA greater than 3.5 suggests more complex aromaticity.[41] As shown in Fig. 2(c), (f) and (i), all the measured SUVA values between October and December were less than 3, that is between 0.239 and 0.890. This suggests that the major components of DOM in Manas River were low molecular weight aliphatic compounds.[45] It can be seen that as the SUVA decreases from upstream to downstream, the aromaticity decreases. This may be due to the reduction of aromatic components by sediment adsorption or microbial decomposition in the water flow process. It is also evident that the higher frequency of human activities leads to stronger aromaticity of DOM in upstream. The overall SUVA demonstrated a gradual increase from October to December, indicating that the organic matter was more aromatic in winter and could contain more aromatic compounds with unsaturated carbon–carbon bonds. In general, the Manas River contained aliphatic compounds that did not absorb at 254 nm. Furthermore, the aromaticity gradually decreased with the direction of the river (from upstream to downstream), and was stronger in winter.

3.2. PARAFAC analysis

PARAFAC analysis was performed on all samples by selecting two to seven components. Half split analysis was also performed to find a suitable number of components. Only the models containing four and five components could be split-half validated.[32] Tucker's congruence coefficients (r_c) were applied to identify the similarity between two pairs of Ex and Em loadings.[14] Some of the fluorescent components extracted by PARAFAC indicated the presence of specific organic substances, which could also represent organic compounds with similar fluorescent properties. The EEM measurements were performed at a constant temperature (25 ± 1 °C) and pH to minimize variation. The maximum fluorescence intensity of each component was obtained to illustrate the quantitative and qualitative differences between the samples.

Two fluorescent components (C1 and C2) were identified from the samples using PARAFAC (Fig. 4, Table 2). Component C1 had Ex/Em maxima at (235, 335)/390 nm, which resembled the traditionally defined peak A (at 260/380–460 nm) and peak M (290–310/370–430 nm),[21,46] assigned to a humic-like component associated with fulvic acids. Component C2 showed Ex/Em maxima at (≤200, 275)/300 nm, which was similar to the tryptophan-like peak T[47] observed in farmland environments. Accordingly, C1 represented terrestrial humus, which is primarily observed in surface water. However, it is also found in sewage and agricultural environments; it mainly originates from land-derived and microbial organic matter.[48] C2 is an amino acid that is free or bound to proteins. It is similar to free tryptophan, which is widely present in aquatic environments and is an indication of recent biological activity.[24] Thus, a protein-like component may be associated with refractory

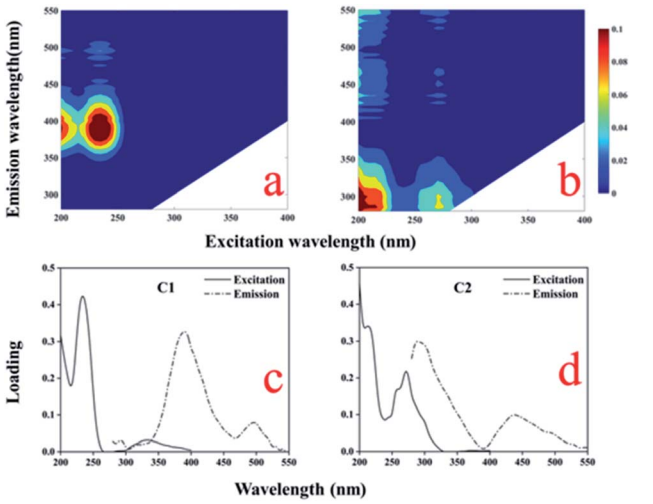

Fig. 4 EEM contours (a, b) and excitation and emission loadings (c, d) of the two fluorescent components identified using PARAFAC.

Table 2 Description of the PARAFAC components and their comparison with previous studies[a]

Component	Ex/Em (nm)	DOM groups and descriptions by previous studies	This study
C1	(235, 335)/390	C2: <250 (320)/390, humic-like materials[26] C1: 240(320)/404, humic-like materials produced from organic matter as a result of microbial activities[50]	Humic-like components
C2	(≤200, 275)/300	C7: 275/340, tryptophan-like components[14] C3: 230(280)/319, tryptophan-like components[38] C4: 260–280/250–340, protein-like components[47]	Protein-like components

[a] Ex/Em represents the excitation and emission maxima.

DOM, which is mainly affected by autogenous sources produced by phytoplankton and microbial metabolism in water bodies. Humus components are produced by terrestrial and microbial sources, and are mainly affected by domestic sewage, agricultural irrigation drainage, and microbial degradation. Studies have shown that DOM in the water body primarily contains humic acid, hydrophilic organic acids, proteins, carbohydrates, and lipids.[49] Its main source is the sewage from domestic activities, industrial wastewater, and the natural decomposition of dead organisms or the by-products of bacterial decomposition. The city of Shihezi, which lies the northwest region of the Manas River, has airports, factories, and small villages and towns near the middle and lower reaches. Due to factory sewage, domestic sewage, and farmland irrigation discharged into Manas River, as well as the degradation of microorganisms and the self-growth of phytoplankton, terrestrial humus (C1), and protein-like substances (C2) were detected in the river.

The mean fluorescence intensity of the two components and their respective contribution to total DOM fluorescence intensity (the percentage of fluorescence maximum score for each component to the total fluorescence maximum scores of all components) using the PARAFAC model differed by location (Fig. 5). In the entire Manas River, the proportion of humus substances (C1, 54%) was higher than that of protein substances (C2, 46%), which indicated that the humus components had a greater impact on the organic matter content. The data revealed that from upstream to downstream, the humus and protein-like components showed a downward and upward trend, respectively, indicating that the land source gradually decreased along the river direction, while the self-generating source gradually increased. However, sampling site M11 had a higher proportion of the humus substances component, which was different from the other downstream sites. This may be due to the influence of man-made inputs or the influence of a small iceberg meltwater runoff at this site. This means that DOM contained more contributions from terrestrial humus upstream of the river, while aquatic systems and microorganisms contributed more downstream. This was consistent with the results of the above.

3.3. Fluorescence indices

The ratios among fluorescence peak intensities can also be useful fluorescence indicators, including the fluorescence index (FI), biological source index (BIX), and humification index (HIX).The FI was used to indicate the source of humus components in DOM, where FI < 1.4, indicated land and soil origins, FI > 1.9 reflected a microbial source, and FI between 1.3 and 1.9 demonstrated that the water body was a mixture of terrestrial and self-generating source.[50] As shown in Fig. 6, the average FI value (3.26) upstream in October was greater than 1.9, indicating that the DOM mainly consisted of microbes or aquatic resources upstream. However, this may be because there were more plants near the upper reaches, and the decaying of their leaves in autumn probably increased the microbial activities. In the midstream and downstream regions, the FI values were lower, that is 1.62 and 1.41, respectively. Furthermore, the DOM may be affected by microorganisms and terrestrial sources. In contrast, in November, the upstream and midstream FI values were 1.52 and 1.06, respectively, while downstream had a higher FI of 4.91, especially M10. This showed that there was a high external source input, which was

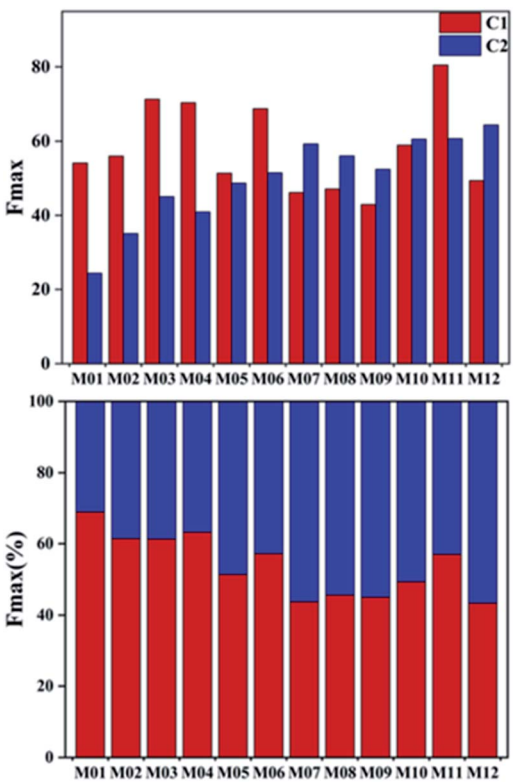

Fig. 5 Parallel factor component (PARAFAC) maximum fluorescence intensity (F_{max}) values organized by sampling location.

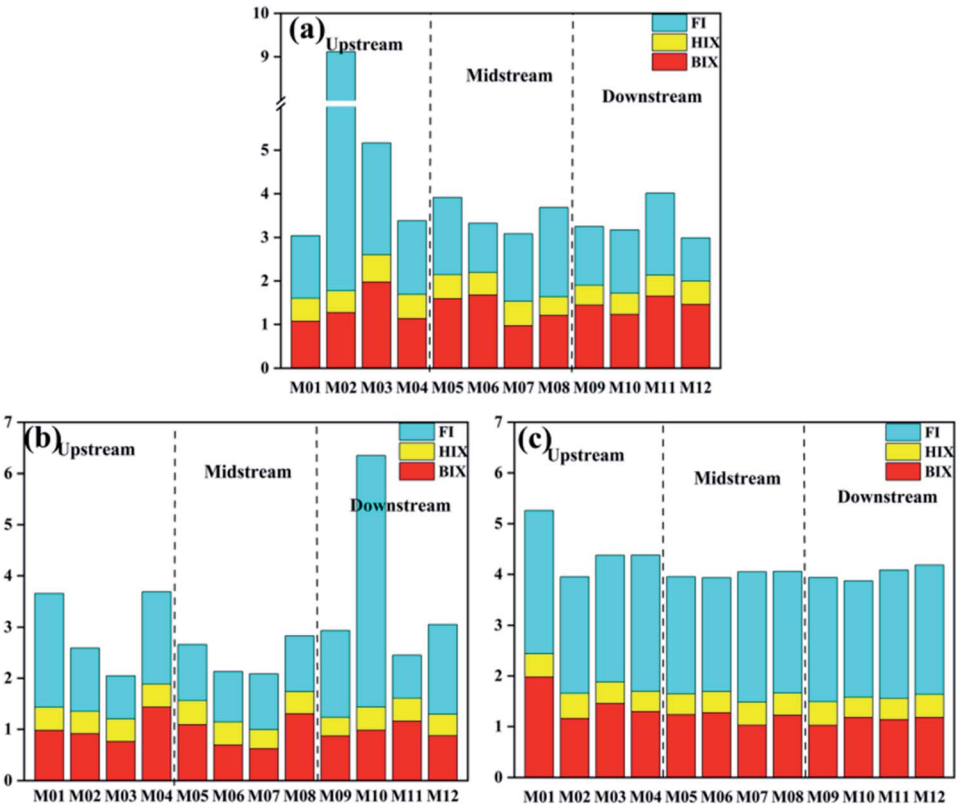

Fig. 6 FI, HIX, and BIX values of Manas River: (a) October; (b) November; (c) December.

consistent with the PARAFAC result. The overall FI value of Manas River in December was relatively high (2.37–3.56) due to the greatly reduced land-source input in winter and the slow flow of the river, which was mainly related to the metabolism of microorganisms in the water body. This indicated that the DOM may not be fully microbiologically derived but could also be the result of terrestrial matter input, as assumed by the groundwater LNE (sample location).[51] Thus, there may be significant differences in the DOM characteristics in different areas of the Manas River. The main sources of DOM in the upper reaches were terrestrial humus and inputs from human activities, followed by the metabolism of microorganisms. This was in stark contrast to the downstream regions. In winter, the river contained higher amounts of DOM.

HIX and BIX were used to infer the degree of humification and the sources of organic matter, respectively. HIX indicates the degree of humification; a high HIX (10–16) suggests a terrestrial origin of DOM and a low HIX (<4) suggests the *in situ* formation of DOM due to biological activities (*i.e.*, a lower humification degree).[52] BIX was used as an indicator of DOM traceability. The BIX values of 0.6–0.7, 0.7–0.8, and 0.8–1.0 indicate weak, intermediate, and strong autochthonous component of DOM, respectively; BIX > 1.0 indicates a biological or aquatic bacterial origin of DOM. The study by Osburn *et al.*[53] found that the trend of BIX values followed that of the FI values. For most groundwater samples, BIX was 0.8–1 and HIX was 4–6, which indicated a weak humic character and strong autochthonous component.[54] In contrast, Manas River had a higher BIX (1.02–1.47) and lower HIX (<4) values, demonstrating that surface water had a more autochthonous component or aquatic biological DOM. Compared to December, the variations in BIX in October and November were more obvious. This showed that in autumn, the river was affected by inputs from different sources, resulting in obvious changes. The HIX in the Manas River showed a downward trend from upstream to downstream, indicating that even under the influence of external sources, the characteristics of the river's self-generating source increased with the migration and transformation of the river.

Based on the above results, organic matter in Manas River primarily originated from the decomposition of microorganisms in the water and changes in the water body. At the same time, the input of humus and human activities also affected the DOM. External inputs affected upstream, and there was an increase the degradation of microorganisms and bacteria in winter.

3.4. Relationship between DBPsFP and DOM characteristics

Natural organic matter reacts with chlorine disinfectants to produce DBPs, which are commonly found in surface water and groundwater. Among them, THMs including chloroform, bromodichloromethane, dibromomonochloromethane, and tribromomethane are the most abundant forms in the chlorination process and pose a threat to human health. According to the regulations of the United States Environmental

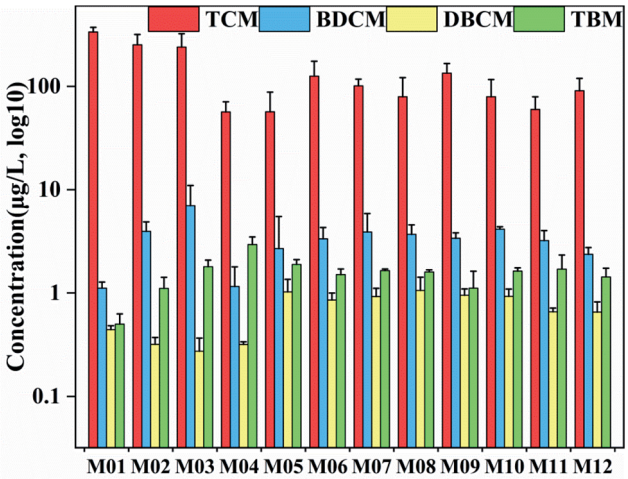

Fig. 7 The formation potential of trihalomethanes in the Manas River. (TCM: trichloromethane; BDCM: bromoxylene; DBCM: dibromomonochloromethane; TBM: tribromomethane).

Protection Agency (US EPA), the highest pollution level of trihalomethane in drinking water is 100 μg L^{-1}.[54] The formation potential of THMs in the Manas River is shown in Fig. 7. The maximum amount of THMsFP in Manas River reached 225.57 μg L^{-1}, which included a chloroform FP (TCMFP) of 220.34 μg L^{-1}. At the same time, upstream contained higher content of THMsFP than the middle and lower reaches. Compared to other studies,[29] the measured DBPsFP in raw drinking water was approximately 150 μg L^{-1}. Ruecker *et al.*[55] collected water samples from the Waccamaw River (a typical blackwater river in the southeastern United States) and measured the THMsFP to be 71–448 μg L^{-1}. Thus, Manas River had a higher THMsFP value, which far exceeded the USEPA pollution standards. This could lead to potential health risks caused by DBPs. However, the method for determining and evaluating DBP precursors through laboratory formation potential is time consuming (chlorination and subsequent extraction and detection takes several days). Therefore, alternative parameters for continuous monitoring of DBP precursors in drinking water treatment plants have been developed, which are essential for optimizing the water treatment process, and effectively reducing and controlling DBPs.

The PARAFAC method is increasingly used in the research on DBPs. The THMsFP, especially TCMFP, which is the most abundant species of THMs, was estimated in most DBP-PARAFAC studies. Some studied reported that the total THMsFP and TCMFP were generally correlated with humic-like components and protein-like components.[56,57] Further studies on the feasibility of PARAFAC components as surrogates for many emerging DBPs are required to unveil the DOM constituents related to DBP formation and to improve the ability to assess the quantity of DBP precursors in drinking water treatment and distribution systems.

Linear correlations were sought between DBPsFP and DOM properties, such as DOC, $SUVA_{254}$, and F_{max} for individual PARAFAC components. There were positive correlations among DOC and PARAFAC humic-like component C1 ($r = 0.535$, $p < 0.01$), HIX ($r = 0.417$, $p < 0.01$) and C1 + C2 ($r = 0.377$, $p < 0.05$), as shown in Table 3. Thus, DOM contained more humus components, which affected the DOC content. There was a significant negative correlation ($r = -0.825$, $p < 0.05$) between the $SUVA_{254}$ and DOC. Scholz *et al.*[39] studied a drinking water reservoir in North Wales and found that the $SUVA_{254}$ had a significant negative correlation ($r = -0.581$, $p < 0.01$) with the removal of DOC (the difference between the water inflow and outflow of the reservoir). Yates *et al.*[58] investigated the catchments of two large rivers in Conwy and Nadder and found that DOC had a significant positive correlation with the SUVA ($r = 0.756$, $p < 0.01$). Although $SUVA_{254}$ and DOC have a good correlation, it differs for different water bodies. Therefore, the $SUVA_{254}$ index cannot be used as a substitute parameter for DOM. However, in this study, both the $SUVA_{254}$ and C1 correctly predicted the DOC content in the Manas River. In addition, the correlation between the fluorescence component was high, which showed the mutual transformation between fluorescence components, meaning that these DOM were from the same organic source. ($r = 0.754$, $p < 0.01$). The protein-like components were easily affected by microbial activities, while the humus and protein-like components in the Manas River had a strong correlation. This indicated that microbial activities was the common source of the two components. Moreover, the correlation between THMsFP, HIX, and C2 was weak ($r = 0.344$, $r = 0.316$, $p < 0.01$), but THMsFP had a strong linear correlation

Table 3 Correlation plot of DOC characters and THMsFP

	DOC	$SUVA_{254}$	FI	BIX	HIX	C1	C2	C1 + C2	THMsFP
DOC	1								
$SUVA_{254}$	−0.825[a]	1							
FI	0.079	−0.057	1						
BIX	0.163	−0.336[a]	0.281	1					
HIX	0.417[b]	−0.385[b]	−0.046	0.301	1				
C1	0.535[b]	−0.516[b]	−0.419[a]	0.141	0.654[a]	1			
C2	0.200	−0.201	−0.636[b]	−0.224	0.288	0.754[b]	1		
C1 + C2	0.377[a]	−0.449[a]	−0.556[b]	−0.079	0.513[b]	0.915[b]	0.911[b]	1	
THMsFP	0.203	0.491[a]	−0.009	0.171	0.344[a]	0.485[a]	0.316[a]	0.529[a]	1

[a] $p < 0.05$. [b] $p < 0.01$.

with other indicators. For example, stronger linear correlations were found between THMsFP and SUVA254 ($r = 0.491$, $p < 0.05$), C1 ($r = 0.485$, $p < 0.01$), and C1 + C2 ($r = 0.529$, $p < 0.01$). THMsFP were correlated strongly with humic-like components than with protein-like components. This result revealed that both humus and protein substances contributed to the formation of trihalomethanes, while humus components contributed more to the formation of THMs. The EEM-PARAFAC approach used in this study was an improvement over DOC and $SUVA_{254}$. Maqbool *et al.*[11] found that the C1/C2 ratio established a direct relationship between seasonal changes and the composition of soil organic matter. They used the ratio of C1/C2 to explore the water quality changes in the treatment units in drinking water plants. In this study, the ratio of the two components of PARAFAC was not correlated with THMsFP, while C1 + C2 had a strong correlation with THMsFP ($r = 0.529$, $p < 0.01$). This suggested that humic-like fluorophores and protein-like fluorophores were important THMs precursors. Moreover, drinking water plants' exposure to trihalomethanes could be evaluated by changing the main components of organic matter in the water treatment process. The removal effect of the precursors indicates that this index may be useful for optimizing the DBP reduction and control processes.

SUVA, C1 + C2, and THMsFP have a high correlation; however, it varies with season. Thus, it is important to understand which index is more representative. Fig. 8 shows that the correlation between THMsFP and C1 + C2 ($r = 0.9828$, 0.9205, 0.8265) was higher than that with SUVA ($r = 0.6105$, 0.7608, 0.6386) from October to December. Throughout the three months of observation, it was found that the correlation between C1 + C2 and THMsFP was more stable, and the prediction of THMsFP was more accurate than that of SUVA.

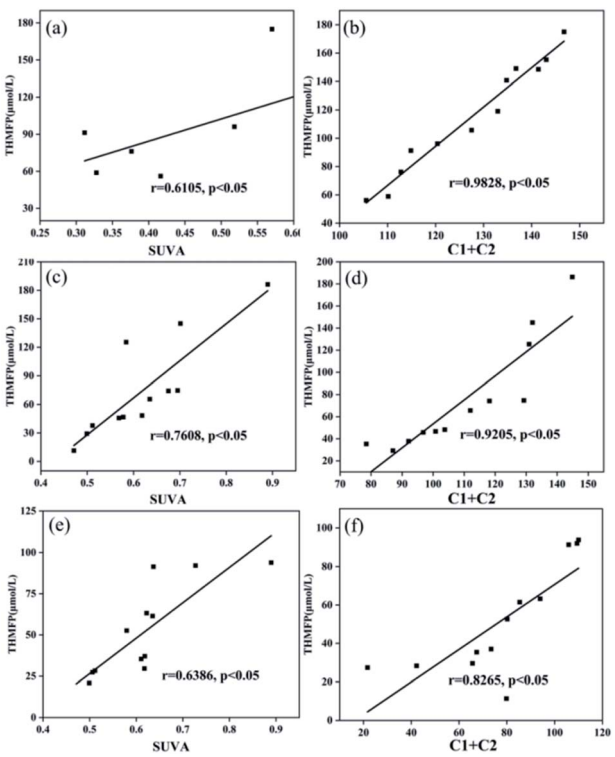

Fig. 8 Correlations between THMsFP and selected DOM properties (a, c, e) THMsFP and SUVA; (b, d, f) THMsFP and C1 + C2 (a, b) October, (c, d) November, and (e, f) December.

As described in Section 3.2 and 3.3, the components of PARAFAC are closely related to the water quality parameters of the Manas River. Therefore, they can predict the origin and composition of DOM and they have the potential to monitor organic matter removal. In this study, the use of PARAFAC analysis proved that the main components of DOM in the Manas River were humus and protein substances. The organic matter in the river was mainly from microbial metabolism and human activities, such as land use, sewage, and wastewater discharge. Thus, the Manas River, a source of drinking water, has a high DOC level, which poses a threat to people's drinking water health. Second, it was proved that the sum of PARAFAC components, C1 + C2, had a strong correlation with the generation potential of DBP ($r = 0.529$, $p < 0.01$), thereby predicting the effect of the actual drinking water purification process on the precursors of DBP removal. Finally, the disinfection technology of Xinjiang drinking water plants is relatively traditional and does not involve advanced monitoring equipment; moreover, the determination of DBPs using the present disinfection technology is troublesome and time consuming. The PARAFAC method avoids these shortcomings and it can conveniently and accurately predict disinfection. Changes in the potential of by-product formation in drinking water plants play an important role in the safety of drinking water for the surrounding residents.

In general, the PARAFAC component was a good substitute for the organic content in the Manas River; the high sensitivity of fluorescence monitoring can detect low-level of organic matter in the water. In addition, the PARAFAC component has high potential to monitor the sub-components of DOC, such as humus and protein.

4. Conclusions

DOC, UV254, SUVA, FI, and F_{max} values of fluorophore components for a diverse group of raw samples from the Manas River provided the following insights:

Due to the surrounding vegetation coverage, the uneven distribution of precipitation throughout the year, the seasonal difference, and the unique hydrological and environmental characteristics of the reservoir during the past three months, the content of organic matter in water changed obviously with the seasons, which was more in October and upstream of the Manas River and less in December and downstream. Meanwhile, the organic matter in the Manas River was dominated by aliphatic compounds. Furthermore, aromaticity gradually decreased along the river and it was higher in winter. Authigenic or aquatic organisms were determined to have a higher DOM and a lower degree of humification. The main sources of DOM were microorganisms and aquatic systems, followed by terrestrial humus and human input. Owing to microbial activity, there are more self-generating sources of DOM downstream, and more external sources upstream, such as domestic

sewage, industrial wastewater, and surrounding agricultural irrigation discharge.

Moreover, EEM and PARAFAC analysis revealed that the main components of organic matter in the Manas River were terrestrial humus (C1) and tryptophan-like protein (C2); the proportion of humus-like components was higher than that of protein-like substances. The upper reaches of the Manas River contained more humus components. The maximum amount of THMsFP in the Manas River was higher than the standard prescribed by the US EPA. Additionally, the upstream contained higher THMsFP than the midstream and downstream. Meanwhile, the THMsFP and PARAFAC components C1 and C1 + C2 had a significant linear correlation. The results of this study suggest that fluorescence spectroscopy has the potential to assess DBPsFP *in situ*, which can aid in the management of drinking water quality by providing a theoretical reference for the removal of DBP precursors in the water purification process of water treatment plants.

Author contributions

Xinlin Wang: validation, conceptualization, investigation, visualization, writing – original draft, writing – review & editing. Yanbin Tong: validation, writing – editing. Jianjiang Lu: validation, resources, ronceptualization, supervision, writing – review & editing. Qigang Chang: validation, supervision. Teng Ma: validation, investigation. Fangdong Zhou: validation, investigation. Jiaqi Li: validation, investigation.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was funded by the National Key Research and Development Program of China, grant number 2016YFC0400704.

References

1 M. Wang and Y. Chen, Generation and Characterization of DOM in Wastewater Treatment Processes, *Chemosphere*, 2018, **201**, 96–109.
2 Y. N. Zhang, J. Zhao, Y. Zhou, J. Qu, J. Chen, C. Li, W. Qin, Y. Zhao and W. J. M. Peijnenburg, Combined Effects of Dissolved Organic Matter, pH, Ionic Strength and Halides on Photodegradation of Oxytetracycline in Simulated Estuarine Waters, *Environ. Sci.: Process. Impacts*, 2019, **21**, 155–162.
3 N. Kamjunke, M. R. Oosterwoud, P. Herzsprung and J. Tittel, Bacterial Production and Their Role in the Removal of Dissolved Organic Matter from Tributaries of Drinking Water Reservoirs, *Sci. Total Environ.*, 2016, **548–549**, 51–59.
4 M. Chen and J. Hur, Pre-treatments, Characteristics, and Biogeochemical Dynamics of Dissolved Organic Matter in Sediments: A review, *Water Res.*, 2015, **79**, 10–25.
5 K. Z. Fu, J. Li, S. Vemula, B. Moe and X. F. Li, Effects of Halobenzoquinone and Haloacetic Acid Water Disinfection Byproducts on Human Neural Stem Cells, *J. Environ. Sci.*, 2017, **58**, 239–249.
6 S. Chowdhury, I. R. Chowdhury and M. H. Zahir, Trihalomethanes in Desalinated Water: Human Exposure and Risk Analysis, *Hum. Ecol. Risk Assess.*, 2017, **24**, 26–48.
7 M. A. Mazhar, N. A. Khan, S. Ahmed, A. H. Khan, A. Hussain, F. Changani, M. Yousefi, S. Ahmadi and V. Vambol, Chlorination Disinfection By-Products in Municipal Drinking Water – A Review, *J. Cleaner Prod.*, 2020, **273**, 123159.
8 T. R. Young, W. Li, A. Guo, G. V. Korshin and M. C. Dodd, Characterization of Disinfection Byproduct Formation and Associated Changes to Dissolved Organic Matter During Solar Photolysis of Free Available Chlorine, *Water Res.*, 2018, **146**, 318–327.
9 M. Yang, X. Zhang, Q. Liang and B. Yang, Application of (LC/) MS/MS Precursor Ion Scan for Evaluating the Occurrence, Formation and Control of Polar Halogenated DBPs in Disinfected Waters: A review, *Water Res.*, 2019, **158**, 322–337.
10 L. Li, Y. Wang, W. Zhang, S. Yu, X. Wang and N. Gao, New Advances in Fluorescence Excitation–Emission Matrix Spectroscopy for the Characterization of Dissolved Organic Matter in Drinking Water Treatment: A Review, *Chem. Eng. J.*, 2020, **381**, 122676.
11 T. Maqbool, Y. Qin, Q. V. Ly, J. Zhang, C. Li, M. B. Asif and Z. Zhang, Exploring the Relative Changes in Dissolved Organic Matter for Assessing the Water Quality of Full-Scale Drinking Water Treatment Plants Using a Fluorescence Ratio Approach, *Water Res.*, 2020, **183**, 116125.
12 W. Cai, J. Han, X. Zhang and Y. Liu, Formation mechanisms of emerging organic contaminants during on-line membrane cleaning with NaOCl in MBR, *J. Hazard. Mater.*, 2020, **386**, 121966.
13 J. Yu, K. Xiao, W. Xue, Y. X. Shen, J. Tan, S. Liang, Y. Wang and X. Huang, Excitation–Emission Matrix (EEM) Fluorescence Spectroscopy for Characterization of Organic Matter in Membrane Bioreactors: Principles, Methods and Applications, *Front. Environ. Sci. Eng.*, 2020, **14**, 31.
14 C. Ma, H. Xu, L. Zhang, H. Pei and Y. Jin, Use of Fluorescence Excitation–Emission Matrices Coupled With Parallel Factor Analysis to Monitor C- and N-DBPs Formation in Drinking Water Recovered From Cyanobacteria-Laden Sludge Dewatering, *Sci. Total Environ.*, 2018, **640**, 609–618.
15 S. A. Baghoth, S. K. Sharma and G. L. Amy, Tracking Natural Organic Matter (NOM) in a Drinking Water Treatment Plant Using Fluorescence Excitation–Emission Matrices and PARAFAC, *Water Res.*, 2011, **45**, 797–809.
16 P. E. García, R. D. García, C. S. Cárdenas, M. Gerea, M. Reissig, G. L. Pérez, L. G. D. Stefano, D. Gianello, C. Queimaliños and M. C. Diéguez, Fluorescence Components of Natural Dissolved Organic Matter (DOM)

from Aquatic Systems of An Andean Patagonian Catchment: Applying Different Data Restriction Criteria for PARAFAC Modelling, *Spectrochim. Acta, Part A*, 2020, **229**, 117957.

17 B. F. Chen, H. Wei, S. Z. Ma, M. H. Feng, C. Liu, X. Z. Gu and K. N. Chen, Characterization of Chromophoric Dissolved Organic Matter in the Littoral Zones of Eutrophic Lakes Taihu and Hongze during the Algal Bloom Season, *Water*, 2018, **10**, 861.

18 E. M. Carstea, C. L. Popa, A. Baker and J. Bridgeman, *In Situ* Fluorescence Measurements of Dissolved Organic Matter: A review, *Sci. Total Environ.*, 2020, **699**, 134361.

19 S. K. L. Ishii and T. H. Boyer, Behavior of Reoccurring PARAFAC Components in Fluorescent Dissolved Organic Matter in Natural and Engineered Systems: A Critical Review, *Environ. Sci. Technol.*, 2012, **46**, 2006–2017.

20 Y. Shang, K. Song, P. A. Jacinthe, Z. Wen, L. Lyu, C. Fang and G. Liu, Characterization of CDOM in Reservoirs and Its Linkage to Trophic Status Assessment Across China Using Spectroscopic Analysis, *J. Hydrol.*, 2019, **576**, 1–11.

21 L. Yang, J. Hur and W. Zhuang, Occurrence and Behaviors of Fluorescence EEM-PARAFAC Components in Drinking Water and Wastewater Treatment Systems and Their Applications: A Review, *Environ. Sci. Pollut. Res.*, 2015, **22**, 6500–6510.

22 L. Yang, H. S. Shin and J. Hur, Estimating the Concentration and Biodegradability of Organic Matter in 22 Wastewater Treatment Plants Using Fluorescence Excitation Emission Matrices and Parallel Factor Analysis, *Sensors*, 2014, **14**, 1771–1786.

23 T. Yeh, C. Liao, T. Chen, Y. Shih, J. Huang, F. Zehetner and T. Hein, Differences in N Loading Affect DOM Dynamics During Typhoon Events in a Forested Mountainous Catchment, *Sci. Total Environ.*, 2018, **633**, 81–92.

24 B. Hua, J. Yang, F. Liu, G. Zhu, B. Deng and J. Mao, Characterization of Dissolved Organic Matter/Nitrogen by Fluorescence Excitation-Emission Matrix Spectroscopy and X-Ray Photoelectron Spectroscopy for Watershed Management, *Chemosphere*, 2018, **201**, 708–715.

25 Y. Du, Y. Zhang, F. Chen, Y. Chang and Z. Liu, Photochemical Reactivities of Dissolved Organic Matter (DOM) in a Sub-Alpine Lake Revealed by EEM-PARAFAC: An Insight Into the Fate of Allochthonous DOM in Alpine Lakes Affected by Climate Change, *Sci. Total Environ.*, 2016, **568**, 216–225.

26 S. B. Huang, Y. X. Wang, T. Ma, L. Tong, Y. Y. Wang, C. R. Liu and L. Zhao, Linking Groundwater Dissolved Organic Matter to Sedimentary Organic Matter From a Fluvio-Lacustrine Aquifer at Jianghan Plain, China by EEM-PARAFAC and Hydrochemical Analyses, *Sci. Total Environ.*, 2015, **529**, 131–139.

27 X. Xu, J. Kang, J. Shen, S. Zhao, B. Wang, X. Zhang and Z. Chen, EEM-PARAFAC Characterization of Dissolved Organic Matter and Its Relationship With Disinfection By-Products Formation Potential in Drinking Water Sources of Northeastern China, *Sci. Total Environ.*, 2021, **774**, 145297.

28 H. S. Lee, J. Hur, M. H. Lee, S. R. Brogi, T. W. Kim and H. S. Shin, Photochemical Release of Dissolved Organic Matter From Particulate Organic Matter: Spectroscopic Characteristics and Disinfection By-Product Formation Potential, *Chemosphere*, 2019, **235**, 586–595.

29 A. D. Pifer, S. L. Cousins and J. L. Fairey, Assessing UV- and Fluorescence-Based Metrics as Disinfection By-Product Precursor Surrogate Parameters in a Water Body Influenced by a Heavy Rainfall Event, *J. Water Supply: Res. Technol.–AQUA*, 2014, **63**, 200–211.

30 Y. Shi, L. Zhang, Y. Li, L. Zhou, Y. Zhou, Y. Zhang, C. Huang, H. Li and G. Zhu, Influence of Land Use and Rainfall on the Optical Properties of Dissolved Organic Matter in a Key Drinking Water Reservoir in China, *Sci. Total Environ.*, 2019, **699**, 134301.

31 T. Ohno, Fluorescence Inner-filtering Correction for Determining the Humification Index of Dissolved Organic Matter, *Environ. Sci. Technol.*, 2002, **36**, 742–746.

32 C. A. Stedmon and R. Bro, Characterizing Dissolved Organic Matter Fluorescence With Parallel Factor Analysis: A Tutorial, *Limnol. Oceanogr.: Methods*, 2008, **6**, 572–579.

33 M. Chen, C. Li, R. G. M. Spencer, N. Maie, J. Hur, A. M. McKenna and F. Yan, Climatic, Land Cover, and Anthropogenic Controls on Dissolved Organic Matter Quantity and Quality From Major Alpine Rivers Across the Himalayan-Tibetan Plateau, *Sci. Total Environ.*, 2021, **754**, 142411.

34 D. P. Hautman and D. J. Munch, Development of U.S. EPA Method 551.1, *J. Chromatogr. Sci.*, 1997, **5**, 221–231.

35 Y. Shen, C. G. Fichot and R. Benner, Floodplain Influence on Dissolved Organic Matter Composition and Export From the Mississippi–Atchafalaya River System to the Gulf of Mexico, *Limnol. Oceanogr.*, 2012, **57**, 1149–1160.

36 S. Lü, R. Jiao, F. Wang, Q. Yu, X. Li, L. Zhang and W. Yan, Characteristics and Chemical Compositions of DOC Linking to the Partial Pressure of Carbon Dioxide in the Lake-River Systems of Lower Changjiang River Basin, *Acta Sci. Circumstantiae*, 2018, **38**, 2034–2044.

37 Y. Zhang, D. Zhang and Y. Mao, Study on POC Transport Characteristics in Yellow River Impacted by Runoff and Sediment Control of the Xiaolangdi Reservoir, *Acta Sci. Circumstantiae*, 2015, **35**, 1721–1727.

38 Y. Luo, Y. Zhang, M. Lang, X. Guo, T. Xia, T. Wang, H. Jia and L. Zhu, Identification of Sources, Characteristics and Photochemical Transformations of Dissolved Organic Matter With EEM-PARAFAC in the Wei River of China, *Front. Environ. Sci. Eng.*, 2021, **15**, 96–106.

39 C. Scholz, T. G. Jones, M. West, A. M. Ehbair, C. Dunn and C. Freeman, Constructed Wetlands May Lower Inorganic Nutrient Inputs but Enhance DOC Loadings Into a Drinking Water Reservoir in North Wales, *Environ. Sci. Pollut. Res.*, 2016, **23**, 18192–18199.

40 H. Zhang, S. Ren, J. Yu and M. Yang, Occurrence of Selected Aliphatic Amines in Source Water of Major Cities in China, *J. Environ. Sci.*, 2012, **24**, 1885–1890.

41 T. Lambert, C. R. Teodoru, F. C. Nyoni, S. Bouillon and F. Darchambeau, Along-Stream Transport and Transformation of Dissolved Organic Matter in a Large Tropical River, *Biogeosciences*, 2016, **13**, 2727–2741.

42 D. M. Golea, A. Upton, P. Jarvis, G. Moore, S. Sutherland, S. A. Parsons and S. J. Judd, THM and HAA Formation From NOM in Raw and Treated Surface Waters, *Water Res.*, 2017, **112**, 226–235.

43 S. S. Marais, E. J. Ncube, T. A. M. Msagati, B. B. Mamba and T. T. I. Nkambule, Assessment of Trihalomethane (THM) Precursors Using Specific Ultraviolet Absorbance (SUVA) and Molecular Size Distribution (MSD), *J. Water Process Eng.*, 2019, **27**, 143–151.

44 S. Chowdhury, Trihalomethanes in Drinking Water: Effect of Natural Organic Matter Distribution, *Water SA*, 2013, **39**, 1–7.

45 A. M. Hansen, T. E. Kraus, B. A. Pellerin, J. A. Fleck, B. D. Downing and B. A. Bergamaschi, Optical Properties of Dissolved Organic Matter (DOM): Effects of Biological and Photolytic Degradation, *Limnol. Oceanogr.*, 2016, **61**, 1015–1032.

46 L. Yang, D. Kim, H. Uzun, T. Karanfil and J. Hur, Assessing Trihalomethanes (THMs) and *N*-nitrosodimethylamine (NDMA) Formation Potentials in Drinking Water Treatment Plants Using Fluorescence Spectroscopy and Parallel Factor Analysis, *Chemosphere*, 2015, **121**, 84–91.

47 K. Wang, Y. Pang, C. He, P. Li, S. Xiao, Y. Sun, Q. Pan, Y. Zhang, Q. Shi and D. He, Optical and Molecular Signatures of Dissolved Organic Matter in Xiangxi Bay and Mainstream of Three Gorges Reservoir, China: Spatial Variations and Environmental Implications, *Sci. Total Environ.*, 2018, **657**, 1274–1284.

48 C. A. Stedmon, R. M. W. Amon, A. J. Rinehart and S. A. Walker, The Supply and Characteristics of Colored Dissolved Organic Matter (CDOM) in the Arctic Ocean: Pan Arctic Trends and Differences, *Mar. Chem.*, 2011, **124**, 108–118.

49 Z. Fan, H. Yang, S. Li and X. Yu, Tracking and Analysis of DBP Precursors' Properties by Fluorescence Spectrometry of Dissolved Organic Matter, *Chemosphere*, 2019, **239**, 124790.

50 Z. Wei, X. Wang, X. Zhao, B. Xi, Y. Wei, X. Zhang and Y. Zhao, Fluorescence Characteristics of Molecular Weight Fractions of Dissolved Organic Matter Derived from Composts, *Int. Biodeterior. Biodegrad.*, 2016, **113**, 187–194.

51 A. D. Pifer and J. L. Fairey, Suitability of Organic Matter Surrogates to Predict Trihalomethane Formation in Drinking Water Sources, *Environ. Eng. Sci.*, 2014, **31**, 117–126.

52 A. Huguet, L. Vacher, S. Relexans, S. Saubusse, J. M. Froidefond and E. Parlanti, Properties of Fluorescent Dissolved Organic Matter in the Gironde Estuary, *Org. Geochem.*, 2008, **40**, 706–719.

53 M. H. Lee, C. L. Osburn, K. H. Shin and J. Hur, New Insight Into the Applicability of Spectroscopic Indices for Dissolved Organic Matter (DOM) Source Discrimination in Aquatic Systems Affected by Biogeochemical Processes, *Water Res.*, 2018, **147**, 164–176.

54 H. MacKeown, J. A. Gyamfi, K. V. K. M. Schoutteten, D. Dumoulin, L. Verdickt, B. Ouddane and J. Criquet, Formation and Removal of Disinfection By-Products in a Full Scale Drinking Water Treatment Plant, *Sci. Total Environ.*, 2020, **704**, 135280.

55 A. Ruecker, H. Uzun, T. Karanfil, M. T. K. Tsui and A. T. Chow, Disinfection By-Product Precursor Dynamics and Water Treatability During an Extreme Flooding Event in a Coastal Blackwater River in Southeastern United States, *Chemosphere*, 2017, **188**, 90–98.

56 L. Yang, D. Kim, H. Uzun, T. Karanfil and J. Hur, Assessing Trihalomethanes (THMs) and *N*-nitrosodimethylamine (NDMA) Formation Potentials in Drinking Water Treatment Plants Using Fluorescence Spectroscopy and Parallel Factor Analysis, *Chemosphere*, 2015, **121**, 84–91.

57 C. A. Mash, B. A. Winston, D. A. Meints II, A. D. Pifer, J. T. Scott, W. Zhang and J. L. Fairey, Assessing Trichloromethane Formation and Control in Algal-Stimulated Waters Amended With Nitrogen and Phosphorus, *Environ. Sci. Process. Impacts*, 2014, **16**, 1290–1299.

58 C. A. Yates, P. J. Johnes, A. T. Owen, F. L. Brailsford, H. C. Glanville, C. D. Evans, M. R. Marshall, D. L. Jones, C. E. Lloyd, T. Jickells and R. P. Evershed, Variation in Dissolved Organic Matter (DOM) Stoichiometry in U.K. Freshwaters: Assessing the Influence of Land Cover and Soil C:N Ratio on DOM Composition, *Limnol. Oceanogr.*, 2019, **64**, 1–13.