PAPER

View Article Online
View Journal | View Issue

Cite this: *Environ. Sci.: Processes Impacts*, 2014, **16**, 2451

Organic removal assessment at full-scale treatment facilities using advanced organic characterization tools†

Ruyuan Jiao,^a Christopher W. K. Chow,^{abc} Hui Xu,^a Xiaofang Yang^{*a} and Dongsheng Wang^{*ac}

The character of dissolved organic matter (DOM) in source water determines its treatability in drinking water treatment. In order to understand and improve drinking water treatment process operations, advanced characterization tools, three-dimensional excitation emission matrix technology (3D-EEM) and high-performance size exclusion chromatography (HPSEC), have been used. The DOM characteristics in both source and treated waters from 52 full-scale drinking water treatment plants of 23 cities in China were investigated. The dissolved organic carbon (DOC) concentration was similar as compared with other countries; however, differences in the compositions of DOM between Chinese water sources and those of other countries were observed. The assessment confirmed that the origin of DOM is the key factor influencing its removal. Source water with the lowest anthropogenic source had the best DOC removal. According to the principal component analysis result of DOM characteristics, the source waters were found to be divided into two groups based on geographic location. Source waters in the east, north and center of China where there are larger populations tend to cluster together; while those of the southwest, northwest and south of China tend to cluster together. These densely populated areas in China should face a bigger challenge of DOC removal during drinking water treatment. Moreover, it was concluded that more attention should be paid to advanced coagulation or hybrid treatment in the actual water treatment train because of the significant increase of the middle molecular weight fraction (1000–2000 Da) after coagulation treatment.

Received 18th April 2014
Accepted 20th June 2014

DOI: 10.1039/c4em00227j

rsc.li/process-impacts

Environmental impact

The existence of dissolved organic matter (DOM) in drinking water treatment processes raises many issues. Therefore, the assessment and control of DOM are two of the major issues for securing drinking water safety. The characteristics of DOM, including the concentration, chemical property (composition) and physical property (molecular weight distribution), in both source and treated waters from 52 full-scale drinking water treatment plants of 23 cities in China were revealed for the first time, and the key factors of the numerous parameters which influence DOM removal were revealed. Moreover, according to the source water characteristics, various areas of China were grouped.

1. Introduction

The quality of drinking water with respect to public health is an important issue worldwide. In order to ensure the supply of safe and high-quality potable water to all consumers, the treatment operators are facing a challenging task to manage and improve

current treatment processes. In the last decade, more attention has been paid to the existence of dissolved organic matter (DOM), for it raises either aesthetic concerns like color, odor and taste or health issues associated with disinfection by-products.^{1,2} In addition, DOM affects treatment operations. For example, DOM controls coagulant and disinfectant demand,³ blocks activated carbon pores,⁴ and is one of the major membrane foulants.^{5,6} Therefore, DOM is a key focus of drinking water research, and the changes in DOM composition and concentration over time will affect various treatment processes. Characterization of DOM in water and linking it to treatability are crucial to optimize water treatment processes.

In previous researches, indicators such as dissolved organic carbon (DOC), UV₂₅₄, SUVA (defined as UV absorbance/DOC) and hydrophobicity/hydrophilicity were usually used to evaluate

^aState Key Laboratory of Environmental Aquatic Chemistry, Research Center for Eco-Environmental Sciences, CAS, POB 2871, Beijing 100085, China. E-mail: wgds@rcees.ac.cn; xfyang@rcees.ac.cn; Fax: +86 10 62849138; Tel: +86 10 62849138

^bSouth Australian Water Corporation, Australian Water Quality Centre, 250 Victoria Square, Adelaide, South Australia 5000, Australia

^cSA Water Centre for Water Management and Reuse, University of South Australia, Mawson Lakes, South Australia 5095, Australia

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c4em00227j

the character of DOM. Among them, SUVA, as a comprehensive index of DOC and UV_{254} , usually represents the treatability of DOM by conventional approaches.^{7–9} SUVA values beyond $3\text{ L mg}^{-1}\text{ m}^{-1}$ were linked to the high aromatic and hydrophobic natural organic matter (NOM) fractions, whereas a SUVA value of less than $3\text{ L mg}^{-1}\text{ m}^{-1}$ suggests that the NOM contains largely non-humic and hydrophilic fractions.¹⁰ However, for the non-conjugated and/or non-aromatic structural organic matters, SUVA cannot provide accurate information of organic matter character owing to its low ultraviolet response. In recent studies, the molecular weight (M_w) of organic matter was considered as another important factor to determine the treatability of water by conventional means.^{11,12} The M_w shows good correspondence with hydrophobicity/hydrophilicity of DOM.¹³ Recently, high-performance size exclusion chromatography (HPSEC) has received increasing attention because it can determine M_w efficiently due to its simple operation.^{14,15} This technique separates NOM constituents based on different retention times of varied molecular weight in the porous solid phase. Wang *et al.* found that HPSEC had good correlation to resin adsorption of chemical fractionation.⁶

The composition of DOM is complex. It originates from a variety of materials and is mixed/modified by natural physico-chemical processes.^{16,17} DOM has distinctive characteristics associated with its origin.¹⁸ Thus DOC sources, which are believed to be crucial to evaluate DOM reactivity and treatability, vary according to their origins. Three-dimensional excitation emission matrix technology (3D-EEM) has been widely used to obtain specific analytical data of DOM components such as humic acid- and fulvic acid-like materials, protein-like material, and soluble microbial byproduct-like material. These data are useful information to identify the DOM origin.^{19–21}

The aim of this investigation was to compare several selected raw and conventionally treated drinking water sources from 52 full-scale drinking water treatment plants (DWTPs) of 23 cities in China with regard to NOM character, in order to obtain a better understanding of the impact of source water on water treatment. These source water samples form a good network of information to cover all the predominant river systems in China, including the Yellow, Yangtze and Pearl rivers and the coastal areas. This study provides useful information and guidance for the design, operation and optimization of water treatment processes.

2. Materials and methods

2.1 Water sample collection

In this study, water samples were collected from 52 full-scale DWTPs of 23 cities in China, and the specific sampling sites are shown in Fig. S1.† The samples were collected in amber glass bottles and delivered under cooled conditions ($4\text{ }^{\circ}\text{C}$ in cooling boxes) without pH control to the lab in 48 h. Once arrived at the lab, the samples were stored in the dark at $4\text{ }^{\circ}\text{C}$ until use. The locations of these water samples are downstream of four main river systems, namely the Yellow river system (YR), Yangtze river

system (YA), Pearl river system (PR) and coastal areas system (CA).

2.2 Treatment process of drinking water treatment plants

All the DWTPs surveyed in this paper use conventional coagulation technology, *i.e.*, coagulation–sedimentation–filtration, as the main water treatment process. Almost all of them used a polyaluminum chloride type of coagulant except several plants that used a composite iron and aluminum coagulant in order to control residual aluminum concentrations. The dosages of coagulants ranged from 26 mg L^{-1} to 31 mg L^{-1} , and the specific information is shown in Table S1.† All plants operated to meet the Chinese water quality standards of residual turbidity and residual aluminum concentration. The selection of different operational conditions, such as coagulant dose and coagulation conditions, was based on the specific operational requirements of each plant. From our observations, these are considered as slight differences, especially in such a large-scale survey. Therefore, in this paper, the differences among plants are ignored and the characteristics of the source water are considered as the sole factor for the treatability assessment.

2.3 Analyses of water quality

Analyzed parameters included pH, turbidity, DOC and UV_{254} . pH was determined by using a pH meter (MP220, Mettler-Toledo, Switzerland). Turbidity was measured by a turbidity meter (2100P, Hach, USA). DOC was determined using a total organic carbon analyzer (TOC-VCPH, Shimadzu, Japan). UV_{254} was determined using a UV/visible spectrophotometer (UV-VIS 8500, Tianmei, China) with a 1 cm quartz cell. All parameters were determined by the methods described in the Standard Methods.²²

2.4 Three-dimensional excitation emission matrix (3D-EEM) fluorescence spectroscopy

The compositions of the DOM were characterized using 3D-EEM methods. In this work, 3D-EEM spectra were employed to analyze DOM using a Hitachi F-4500 fluorescence spectrophotometer with an excitation range from 200 to 400 nm at 5 nm sampling intervals and an emission range from 280 to 550 nm at 0.5 nm sampling intervals. The spectra were recorded at a scan rate of $12\,000\text{ nm min}^{-1}$. The fluorescence in different spectral regions is associated with different types of functional groups.²³ Therefore, DOM from different water samples could have unique and characteristic EEM spectra. The fluorescence regional integration (FRI) method was used to quantify the EEM spectra by delineating the EEM landscape into five regions, and the integrated volume under each region was calculated in order to characterize the DOM.²⁰ The FRI technique gives the specific percent fluorescence response in each region by using the calculus method. The volume (Φ_i) beneath region i of the EEM can be calculated by

$$\Phi_i = \sum_{\text{EX}} \sum_{\text{EM}} I(\lambda_{\text{ex}}, \lambda_{\text{em}}) \Delta\lambda_{\text{ex}} \Delta\lambda_{\text{em}} \quad (1)$$

in which EX is the excitation and EM is the emission. $\Delta\lambda_{\text{ex}}$ and $\Delta\lambda_{\text{em}}$ are the excitation wavelength interval and emission wavelength interval respectively, and $I(\lambda_{\text{ex}}, \lambda_{\text{em}})$ is the fluorescence intensity at each excitation–emission wavelength pair. All Φ_i values were normalized to a DOC concentration of 1 mg L^{-1} for comparison of EEM spectra from different sources of DOM. In order to reduce the effects on the overall interpretations of dominance by shoulders in different EEM regions, Φ_i needs to be normalized by means of dividing by relative regional areas. The normalized excitation–emission area volumes $\Phi_{i,n}$ were calculated as

$$\Phi_{i,n} = \text{MF}_i \Phi_i \quad (2)$$

The MF_i values are equal to the inverse of the fractional projected excitation–emission area, and they are constant values for the fixed excitation wavelength, emission wavelength and intervals. The values for the five regions are 20.4, 16.4, 4.81, 8.76 and 1.76, respectively. The cumulative volume beneath the EEM ($\Phi_{T,n}$) and percent fluorescence response ($P_{i,n}$) were calculated as follows:

$$\Phi_{T,n} = \sum_{i=1}^5 \Phi_{i,n} \quad (3)$$

$$P_{i,n} = \Phi_{i,n} / \Phi_{T,n} \times 100\% \quad (4)$$

2.5 Molecular weight distribution

The apparent molecular weight (AMW) distribution profile of DOM was obtained by the HPSEC method, using a high-performance liquid chromatography system (Waters 1525, Waters, USA) coupled with a size exclusion chromatography column (Shodex Protein KW-802.5, Shoko, Japan) at 25°C . The mobile phase was prepared using 0.02 mol L^{-1} (M) orthophosphate and 0.1 M NaCl (pH 6.8), followed by degassing for at least 15 minutes prior to use. The flow rate of the mobile phase was 0.8 mL min^{-1} . Polystyrene sulfonate standards with M_w ranging from 1800 Da to 35 000 Da were used to calibrate the retention time for AMW. Before determination, the DOM samples had to be filtered through a $0.22 \mu\text{m}$ membrane filter. The UV absorbance of the samples was monitored at 254 nm. HPSEC chromatograms were analyzed using Peakfit (Version 4.12, Systat Software Inc.) to resolve the overlapped peaks. The peak fitting parameters were chosen as in Chow *et al.*¹¹ Chromatography log normal 4-parameter (area) was selected as the peak type to represent the resolved component peaks. The percentage of each peak area was regarded as the percentage of each physical fraction.

3. Results and discussion

3.1 Geographical locations and water quality

The general water quality parameters (pH, turbidity and type of water source) of the 52 source water samples are detailed in Table S1.† The pH of source water varied in the range of 6.39 to

8.23. There was a prominent trend that the pH values decreased from northern to southern China. In previous studies, water alkalinity of the north has been proven to be higher than that of the south.^{6,24,25} Therefore, the variation of pH might be attributed to the alkalinity differences. The average pH values for north and south were 7.89 and 7.55, respectively. The variations of DOM (DOC and UV_{254}) in the source and treated waters are summarized in Fig. 1. The DOM contents of the source waters were similar to those of waters of the United States, which ranged from 1.1 mg L^{-1} to 8.2 mg L^{-1} ,²⁶ but relatively low compared to those of Australia, which ranged from 5.8 mg L^{-1} to 12.8 mg L^{-1} .¹⁸ The DOC and UV_{254} values varied within the ranges of $0.924\text{--}5.94 \text{ mg L}^{-1}$ and $0.003\text{--}0.354 \text{ cm}^{-1}$, respectively.

SUVA as an important parameter has been used to determine the conjugation and aromaticity of DOM by many researchers.^{7–9} The SUVA value indicates to some extent the quality of the water and the treatability of DOM by coagulation.²⁷ SUVA values beyond $3 \text{ L mg}^{-1} \text{ m}^{-1}$ indicate the presence of hydrophobic, humic-like and high M_w organic matter, which can be easily removed by advanced coagulation. For water with SUVA values below $3 \text{ L mg}^{-1} \text{ m}^{-1}$, this material cannot be removed easily. In this survey, the SUVA of source water varied considerably. As shown in Fig. 1, the SUVA values of most source waters were less than 3, indicating a low proportion of humic substances which leads to low average remove rate of DOC ($\frac{\text{DOC}_{\text{source water}} - \text{DOC}_{\text{treated water}}}{\text{DOC}_{\text{source water}}} \times 100\%$) with a value of 14.5%.

3.2 The composition analysis of DOM

Aquatic DOM can be classified roughly into two groups.¹⁸ One group is humified material, containing humic-like material and fulvic-like material, and the other group is non-humified material, consisting of protein-like materials. In order to improve our understanding of the water quality of different regions, 3D-EEM spectra were used to characterize the classification of DOM in different regions. The excitation and emission boundaries were defined into five regions according to the supporting literature.²⁸ Based upon the study by Wen Chen *et al.*,²⁰ peaks at $\text{EX} < 250 \text{ nm}$ and $\text{EM} < 350 \text{ nm}$ belong to regions I and II, which represent simple aromatic proteins (AP) such as tyrosine, and the DOM is easily biodegraded. Peaks at $\text{EX} < 250 \text{ nm}$ and $\text{EM} > 350 \text{ nm}$ representing fulvic acid-like materials (FA) are related to region III. Peaks at $\text{EX} = 250\text{--}280 \text{ nm}$ and $\text{EM} < 380 \text{ nm}$ representing soluble microbial byproduct-like material (SMP) are related to Region IV. And Region V, which corresponds to $\text{EX} > 280 \text{ nm}$ and $\text{EM} > 380 \text{ nm}$, represents humic acid-like organics (HA). In order to better understand the similarities and differences among the four river systems, FRI analysis was also conducted, and the FRI distribution of the source water samples is illustrated in Fig. 2.

Compared to the samples collected from the other river systems, the samples collected from the Yellow river showed higher intensity in regions I, II and III, but much lower intensity in regions IV and V, with average values of 21.91%, 34.66%,

Fig. 1 DOM parameters for water samples: (a) Yellow river, (b) Yangtze river, (c) Pearl river and (d) coastal areas.

Fig. 2 Fluorescence regional integration results for source water and treated water.

29.22%, 5.96% and 8.25% for the five regions, respectively. This is indicative of a higher proportion of proteinaceous material and a lower proportion of HA. It has been observed that the

fluorescence intensity in regions I and II would be high in wastewater-impacted rivers due to the microbiologically derived organic matter.²⁹ The samples collected from the Yangtze and Pearl rivers gave similar results to each other. The two systems were both characterized by higher intensities in regions II, IV and V (27.45%, 24.39%, 24.51% for the Yangtze river and 22.55%, 28.05%, 30.87% for the Pearl river), and lower intensity in region III (9.91% and 8.49%), compared with the other systems. Compared with the Yellow river, these two systems showed a distinctively different level of region IV, which represents the presence of SMP. The high temperature and subtropical monsoon climate of YA system and PR system accelerate algal blooms and hence create a strong bias of autochthonous-originated organic materials. Samples from coastal areas showed low intensities in region IV (15.36%), indicative of low SMP, but relatively high intensities in regions II, III and V (17.69%, 30.38% and 28.24%), characteristic of AP, FA and HA. This indicated that allochthonous sources and anthropogenic sources are all dominating factors in the CA system. However, CA water had the lowest AP proportion, which is an indication of an anthropogenic source. The absolute values of DOM in different fractions (the sum of regions I and II was classified as AP) are compared in Fig. 3. It can be seen that DOM existing as AP was the greatest, and the amount of humified material in Chinese source water was a much smaller proportion compared with source water from the United States.³⁰ This indicated that although, in China, DOC in source

Fig. 3 DOC values in different fractions.

water is not higher than in other countries, it is affected seriously by human activities. On the whole, there were slight variations between source and treated waters (Fig. 2 and 3).

In order to get the correlation between different DOC fractions (AP, FA, SMP and HA) and their treatability, data sets for DOC and DOC removal rate were plotted against the five regions, and least squares regression with analysis of variance (ANOVA) was applied using SPSS 19.0 software (Table 1). For acceptance of a valid relationship, both a high correlation factor (R^2) and a probability (p) value of less than 0.05 (95% confidence of a real correlation) were deemed to be necessary. By analyzing the five different regions in relation to DOC of the source water, FA showed the best correlation, with R^2 of 0.707 and $p < 0.01$ (Fig. S4†). In addition, it also showed a relatively high correlation when linking it to the DOC removal rate. In this study, source water with high DOC value generally showed a high removal rate, and *vice versa*. Therefore, the statistical results here proved that a large proportion of the DOC was mainly related to the high-molecular-weight or hydrophobic fraction, which is easily removed by traditional treatment processes. The coagulation treatability of source water could not be predicted simply from the DOC concentration of samples.

3.3 The analysis of M_w distribution

HPSEC has been reported to be an effective method for determining the M_w of NOM.^{2,16} It separates NOM constituents by different elution times simply and quickly.³¹ Therefore, it has been widely used for DOM characterization and is particularly

well received by water treatment operators. In this paper, water samples were analysed by HPSEC. The results are shown in Fig. S2.†

The M_w of DOM in source water in northern China is generally below 3000 Da, lower than that of water in southern China, which is usually in the range of 1000 Da to 10 000 Da. This indicates that the impact of location on M_w distribution is prominent, even for the same river. For example, the Yellow river, the second longest river in China, goes through the north of China and the northwest of China. The M_w distributions of the Yellow river were separated into two regions depending on location. Those values for the part of the Yellow river in the northwest are mainly in the region of 1000–10 000 Da, whereas for the part in the north the M_w distribution is mainly between 300 Da and 3000 Da. Despite the differences in the level of absorbance response due to different water sources (YR4, YR5, YR6, YR8, YR9 and YR13 are underground water and the others are surface water), the source water showed reasonable similarity in molecular weight ranges. The same conclusion was arrived at for coastal areas. However, the variation caused by the location is only noticeable from the north to the south of China. For the Yangtze and Pearl rivers, which are located in the south of China, there is hardly any difference of M_w distribution from the east to the west of China. These results confirmed the

Table 1 Correlation of source water DOM composition with DOC and DOC removal rate

The composition of DOM		DOC		DOC removal rate	
		R^2	p	R^2	p
Source water	AP ₁	0.499	<0.01	0.202	0.152
	AP ₂	0.512	<0.01	0.179	0.203
	FA	0.707	<0.01	0.361	<0.01
	SMP	0.692	<0.01	0.309	0.026
	HA	0.530	<0.01	0.476	<0.01

Fig. 4 Characteristics of averaged molecular weight between source water and treated water.

Table 2 Correlation of M_w range with DOC and UV₂₅₄

Apparent M_w range (Da)		DOC		UV ₂₅₄	
		R^2	p	R^2	p
Source water	100–1000	0.406	0.01	0.067	0.685
	1000–2000	0.654	<0.01	0.222	0.174
	2000+	0.132	0.424	0.934	<0.01
Treated water	100–1000	0.413	<0.01	0.556	<0.01
	1000–2000	0.421	<0.01	0.753	<0.01
	2000+	0.248	0.124	0.476	<0.01

important influence of climate factors and human factors on water character. According to the China water resources bulletins (2008–2012),³² the annual average runoff ranges of the Yangtze river and Pearl river were 7713.6×10^9 to $11\,146.1 \times 10^9$ m³ per year and 3676.8×10^9 to 5682.3×10^9 m³ per year, which were covariant with the average rainfall of $16\,603.3 \times 10^9$ to $20\,686.4 \times 10^9$ m³ per year and 7420×10^9 to $10\,438 \times 10^9$ m³ per year, respectively, while that of the Yellow river was just 454.2×10^9 to 660.4×10^9 m³ per year, which were also covariant with the average rainfall of 3443.1×10^9 to 3896.8×10^9 m³ per year. The shortage of rain in the north of China has led to a reduction in the surface runoff. Therefore, autochthonous production and anthropogenic sources may have become the main origin of DOM in the north of China. Although the climate in the northwest is similar to that in the north, there is a considerably larger population in the north. Thus, the

interference of human activity may be the main reason for the differences between source water in northern and northwestern China. This is also consistent with the conclusion that DOM of the Yellow river has the worst coagulation treatability, due to its relatively low SUVA value.

Although HPSEC has been considered to be a useful technique for evaluating various water treatment processes, it can only provide basic qualitative information of water samples. Therefore, in this paper, HPSEC chromatographs were processed by peak fitting software in order to obtain more quantitative information about the NOM constituents. As an example of the physical fractionation process, Fig. S3† shows the AMW distribution profile of the DOM of the YR1 water sample and six individual AMW peaks (fractions) for which the r^2 value is 0.9995 after peak fitting. The six AMW fractions of DOM are distributed around 350 Da (Peak 1), 1000 Da (Peak 2), 2000 Da (Peak 3), 2800 Da (Peak 4), 3900 Da (Peak 5) and 5400 Da (Peak 6) respectively. The six peaks (fractions) can be associated with the chemical groups of low molecular weight acids (Peak 1), low molecular weight humic substances (Peak 2), and high molecular weight humic substances (Peak 3, Peak 4, Peak 5 and Peak 6).¹⁸ The peak areas of the six peaks representing the six individual physical fractions were 0.000145 (Peak 1), 0.000761 (Peak 2), 0.00103 (Peak 3), 0.00144 (Peak 4), 0.00236 (Peak 5) and 0.00125 (Peak 6).

In order to assess the molecular weight variation of DOM before and after treatment, the fraction of each defined AMW of the DOM, *i.e.* high-molecular weight fraction (AMW > 2000 Da), mid-molecular weight fraction (1000 Da < AMW < 2000 Da) and

Fig. 5 PCA plots of source water: (a) PC1 vs. PC2 loading plot, (b) PC1 vs. PC2 score plot, (c) comprehensive qualities in different areas, (d) relationship between the population density and the comprehensive quality values of different areas.

low-molecular weight fraction ($M_w < 1000$ Da), was quantified. From the fitting results shown in Fig. 4, it can be seen that both the high-molecular weight (High- M_w) and low molecular weight (Low- M_w) fractions in source water decrease slightly after treatment, while the mid-molecular weight (Mid- M_w) fraction increases significantly. This is revealed remarkably in the results of individual samples (YR1). Traditional coagulation process in China cannot decrease DOC effectively. In this study, the High- M_w fraction, which was considered as a fraction easily removed by coagulation, seems to be decreased slightly. At the same time, the aggregation between Low- M_w fractions in the coagulation process by the mechanism of adsorption and bridge might contribute to the increase of Mid- M_w fractions. This suggested that more attention should be paid to advanced coagulation in the actual water treatment process or that further treatment is needed to deal with the Mid- M_w fraction in the drinking water treatment.

Data sets for DOC and UV_{254} were plotted against the three fractions in order to obtain the correlation between conventional indicators and the three M_w fractions in China (Table 2). By comparing the three separate M_w fractions with the DOC of the source water, it can be seen that the Mid- M_w fraction has a better correlation (R^2 of 0.654 and $p < 0.01$) than the Low- M_w fraction (R^2 of 0.406 and $p = 0.01$), whereas there is no correlation between the DOC and the High- M_w fraction. In terms of the results in Section 3.2, AP are the most abundant material; the result here corresponds to the general recognition that the molecular weight of AP is relatively low. However, the correlation between DOC and the Mid- M_w fraction in the treated water decreased slightly, with R^2 of just 0.421 ($p < 0.01$), although it was still a better correlation than was found with the other two fractions. It seems the treatment process was not selective for DOC. When the fractions are compared to the UV_{254} of the source water, there is a significant relationship between the UV_{254} and the High- M_w fraction, with R^2 of 0.93 ($p < 0.01$), but there is no relationship with the other two fractions. However, after treatment, the correlations show an obvious increase for the other two fractions, especially for the Mid- M_w fraction (R^2 of 0.75 and $p < 0.01$). This indicates that the treatment process is selective for the substances which have an ultraviolet response at 254 nm. Following coagulation, most high molecular weight organic matter with an ultraviolet response is effectively removed, and thus the best correlation is seen to shift to the lower molecular weight organic matter that is less effectively removed by coagulation.

3.4 Principal component analysis

Principal component analysis (PCA) was applied by using SPSS to classify the source water, based on the DOM characteristics discussed in this paper. Four principal components (PCs), explaining up to 82.93% of total variance, were used in this study. Among them, PC1 explained 49.2% of the total variance. Mid- M_w was the most important factor in PC1, followed by High- M_w , DOC and HA, according to their horizontal axis position (Fig. 5a). PC2 accounted for 23.4% of the total variance. AP1, AP2 and FA were the main factors of PC2.

The PC1 vs. PC2 score plots are presented in Fig. 5b. The percent variance explained by each component was noted on each axis. Although the distribution of the source water collected from different locations overlapped around the zero coordinates, the data still can be gathered into two groups. Namely, the east, the north and the center of China tend to cluster together, and the southwest, the northwest and the south of China tend to cluster together. By calculating the comprehensive score (the accumulation of the products of each variance contribution and principal component score) based on PCA (Fig. 5c),³³ a similar result can be obtained. Using data from the latest census in China,³⁴ the relationship between the population density and the comprehensive score was analyzed. As shown in Fig. 5d, the population density and comprehensive score were almost positively correlated, except for the northwest. For source water in the northwest of China, as discussed in Section 3.3, the dry climate leads to limited surface runoff and further results in the high DOC and UV_{254} values for the source water samples, as shown in Fig. 1 (YR1–YR4). Therefore, the classification may be attributed to the difference in population distributions of the two groups. The areas in the first group have large population density, while the areas in the

Fig. 6 PAC plots of removal rate: (a) PC1 vs. PC2 loading plot, (b) PC1 vs. PC2 score plot.

second group have relatively low population density. Therefore, the PCA results of source water demonstrated the importance of anthropogenic influence on the characteristics of water, especially on DOC and M_w .

For the treated water, the DOC removal rate and the value of each characteristic were used to carry out the PCA. There were also four PCs with 78.23% of total variance. Among them, PC1, with three most important factors HA, SMP and FA (Fig. 6a), accounted for 44.5% of total variance. PC2, with 25.7% of total variance, has two most important factors, AP1 and AP2. HA, FA and SMP are considered mainly to be natural sources, and AP1 and AP2 are recognized mainly as anthropogenic sources. So the results indicate that the origin of DOM is a key factor in DOM removal. From the score plot shown in Fig. 6b, CA exhibited a significant difference because of its high percentage of HA, SMP and FA. The average DOC removal rate of CA water was the highest, with a value of 24.89%, whereas the other systems just had the average removal rate, below 15%.

4. Conclusion

The characteristics of DOM in drinking source waters and treated waters in the main four water systems in China were compared, and the distribution of DOM and the quality of drinking water in China were determined. Although the DOC concentration and UV_{254} value were similar compared with other countries, the DOM in source water displayed a significantly lower removal rate by traditional coagulation due to the relatively low SUVA, low M_w and high content of protein-like materials in source water in China.

The high concentration of protein-like materials in most of the Chinese source water samples indicated the seriousness of anthropogenic influence on the source water. Terrestrially derived DOM is usually considered as easily amenable, whereas anthropogenically derived DOM is the opposite. In addition, the molecular weight of DOM, besides being subject to climatic variation, was also affected by human activities. The results showed that the molecular weight for the south part of China was larger than that for the north, due to the limited terrestrial source input caused by the shortage of rain in the north. However, when comparing the same climatic conditions, human activities became the main factor in the difference between the molecular weight distribution of the north and northwest of China. An area with a larger population usually has a DOM with a lower molecular weight range.

After statistical analysis, the source water samples were classified into two groups according to their characteristics based on the selected parameters. Source waters of the east, north and center of China, which are densely populated, tend to cluster together; while those of the southwest, northwest and south of China tend to cluster together. Meanwhile, when including the DOC results, the statistical analysis proved the importance of DOM origin to the DOM removal. CA water, which showed the lowest anthropogenic source, had the best removal of DOC. Therefore, cities with larger populations in China should face a bigger challenge of DOC removal in drinking water treatment. In addition, after treatment, the

High- M_w and Low- M_w fractions decreased slightly, while the Mid- M_w fraction increased significantly overall. More attention should be paid to advanced coagulation or to including a hybrid process in the actual water treatment train to deal with the Mid- M_w fraction in the drinking water treatment.

Acknowledgements

Financial supports from NSFC (51338010, 51025830, 21107125, and 51179099) and Major Science and Technology Program for Water Pollution Control and Treatment (2009ZX07419-001) are gratefully acknowledged.

References

- 1 P. C. Singer, Humic substances as precursors for potentially harmful disinfection by-products, *Water Sci. Technol.*, 1999, **40**(9), 25–30.
- 2 N. H. Lee, G. Amy, J. P. Croue and H. Buisson, Identification and understanding of fouling in low-pressure membrane (MF/UF) filtration by natural organic matter (NOM), *Water Res.*, 2004, **38**(20), 4511–4523.
- 3 S. Chowdhury, P. Champagne and P. J. McLellan, Uncertainty characterization approaches for risk assessment of DBPs in drinking water: a review, *J. Environ. Manage.*, 2009, **90**(5), 1680–1691.
- 4 G. Newcombe, M. Drikas, S. Assemi and R. Beckett, Influence of characterised natural organic material on activated carbon adsorption: I. Characterisation of concentrated reservoir water, *Water Res.*, 1997, **31**(5), 965–972.
- 5 S. T. Nguyen, F. A. Roddick and J. L. Harris, Membrane foulants and fouling mechanisms in microfiltration and ultrafiltration of an activated sludge effluent, *Water Sci. Technol.*, 2010, **62**(9), 1975–1983.
- 6 D. S. Wang, Y. M. Zhao, J. K. Xie, C. W. K. Chow and J. van Leeuwen, Characterizing DOM and removal by enhanced coagulation: A survey with typical Chinese source waters, *Sep. Purif. Technol.*, 2013, **110**, 188–195.
- 7 A. D. Archer and P. C. Singer, Effect of SUVA and enhance coagulation on removal of TOX precursors, *J. - Am. Water Works Assoc.*, 2006, **98**(8), 97–107.
- 8 C. Bahr, J. Schumacher, M. Ernst, F. Luck, B. Heinzmann and M. Jekel, SUVA as control parameter for the effective ozonation of organic pollutants in secondary effluent, *Water Sci. Technol.*, 2007, **55**(12), 267–274.
- 9 D. S. Wang, L. N. Xing, J. K. Xie, C. W. K. Chow, Z. Z. Xu, Y. M. Zhao and M. Drikas, Application of advanced characterization techniques to assess DOM treatability of micro-polluted and un-polluted drinking source waters in China, *Chemosphere*, 2010, **81**(1), 39–45.
- 10 H. Park, Y. Kim, B. An and H. Choi, Characterization of natural organic matter treated by iron oxide nanoparticle incorporated ceramic membrane-ozonation process, *Water Res.*, 2012, **46**(18), 5861–5870.
- 11 C. W. K. Chow, R. Fabris, J. Van Leeuwen, D. S. Wang and M. Drikas, Assessing natural organic matter treatability

- using high performance size exclusion chromatography, *Environ. Sci. Technol.*, 2008, **42**(17), 6683–6689.
- 12 M. Ng, S. Liu, C. W. K. Chow, M. Drikas, R. Amal and M. Lim, Understanding effects of water characteristics on natural organic matter treatability by PACl and a novel PACl-chitosan coagulants, *J. Hazard. Mater.*, 2013, **263**, 718–725.
 - 13 J. Van Leeuwen, R. Daly and A. Holmes, Modeling the treatment of drinking water to maximize dissolved organic matter removal and minimize disinfection by-product formation, *Desalination*, 2005, **176**(1–3), 81–89.
 - 14 R. Lewis, J. A. van Leeuwen, C. W. K. Chow, A. Everson and D. M. Lewis, Assessment of coagulated and non-coagulated ASB performance used to treat *Pinus radiata* sulfite pulp and paper mill effluent by resin fractionation and HPSEC techniques, *Chem. Eng. J.*, 2012, **213**, 109–117.
 - 15 R. Gough, P. J. Holliman, N. Willis and C. Freeman, Dissolved organic carbon and trihalomethane precursor removal at a UK upland water treatment works, *Sci. Total Environ.*, 2014, **468**, 228–239.
 - 16 A. Sachse, R. Henrion, J. Gelbrecht and C. E. W. Steinberg, Classification of dissolved organic carbon (DOC) in river systems: influence of catchment characteristics and autochthonous processes, *Org. Geochem.*, 2005, **36**(6), 923–935.
 - 17 F. G. Meng, G. C. Huang, X. Yang, Z. Q. Li, J. Li, J. Cao, Z. G. Wang and L. Sun, Identifying the sources and fate of anthropogenically impacted dissolved organic matter (DOM) in urbanized rivers, *Water Res.*, 2013, **47**(14), 5027–5039.
 - 18 R. Fabris, C. W. K. Chow, M. Drikas and B. Eikebrokk, Comparison of NOM character in selected Australian and Norwegian drinking waters, *Water Res.*, 2008, **42**(15), 4188–4196.
 - 19 J. H. Goldman, S. A. Rounds and J. A. Needoba, Applications of Fluorescence Spectroscopy for Predicting Percent Wastewater in an Urban Stream, *Environ. Sci. Technol.*, 2012, **46**(8), 4374–4381.
 - 20 W. Chen, P. Westerhoff, J. A. Leenheer and K. Booksh, Fluorescence Excitation–Emission Matrix Regional Integration to Quantify Spectra for Dissolved Organic Matter, *Environ. Sci. Technol.*, 2003, **37**(24), 5701–5710.
 - 21 N. Hudson, A. Baker and D. Reynolds, Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters - A review, *River Res. Appl.*, 2007, **23**(6), 631–649.
 - 22 State Environmental Protection Administration of China, *Monitoring and Analytical Method of Water and Wastewater*, ed. F. S. Wei and W. Q. Qi, Environmental and Scientific Press of China, Beijing, 4th edn, 2002.
 - 23 C. M. Sharpless and L. B. McGown, Effects of aluminum-induced aggregation on the fluorescence of humic substances, *Environ. Sci. Technol.*, 1999, **33**(18), 3264–3270.
 - 24 J. F. Yu, D. S. Wang, M. Q. Yan, C. Q. Ye, M. Yang and X. P. Ge, Optimized coagulation of high alkalinity, low temperature and particle water: pH adjustment and polyelectrolytes as coagulant aids, *Environ. Monit. Assess.*, 2007, **131**(1–3), 377–386.
 - 25 I. Lou, S. Y. Gong, X. J. Huang, Y. J. Liu and K. M. Mok, Coagulation optimization using ferric and aluminum salts for treating high algae and high alkalinity source water in a typical North-China plant, *Desalin. Water Treat.*, 2013, **51**(16–18), 3361–3370.
 - 26 L. Liang and P. C. Singer, Factors Influencing the Formation and Relative Distribution of Haloacetic Acids and Trihalomethanes in Drinking Water, *Environ. Sci. Technol.*, 2003, **37**(13), 2920–2928.
 - 27 S. Goel, R. M. Hozalski and E. J. Bouwer, Biodegradation of NOM: Effect of NOM source and ozone dose, *J. - Am. Water Works Assoc.*, 1995, **87**(1), 90–105.
 - 28 D. M. Owen, G. L. Amy, Z. K. Chowdhury, R. Paode, G. McCoy and K. Viscosil, NOM characterization and treatability, *J. - Am. Water Works Assoc.*, 1995, **87**(1), 46–63.
 - 29 R. D. Holbrook, J. Breidenich and P. C. DeRose, Impact of reclaimed water on select organic matter properties of a receiving stream-fluorescence and perylene sorption behavior, *Environ. Sci. Technol.*, 2005, **39**(17), 6453–6460.
 - 30 F. L. Rosario-Ortiz, S. A. Snyder and I. H. Suffet, Characterization of dissolved organic matter in drinking water sources impacted by multiple tributaries, *Water Res.*, 2007, **41**(18), 4115–4128.
 - 31 C. W. K. Chow, R. Fabris, M. Drikas and M. Holmes, A case study of treatment performance and organic character, *J. Water Supply: Res. Technol.-AQUA*, 2005, **54**(6), 385–395.
 - 32 The Ministry of Water Resources of the People's Republic of China, <http://www.mwr.gov.cn/zwzc/hygb/szygb/>, (accessed June 2014).
 - 33 J. Xia, *Evaluate of Regional Water environment and Environment Quality: Multilevel Association Assessment and Application*, Press of Wuhan University of Hydraulic and Electric, Wuhan, 1st edn, 1999, pp. 21–25.
 - 34 The Central People's Government of the People's Republic of China, http://www.gov.cn/gzdt/2011-04/29/content_1854891.htm, (accessed June 2014).