


PARAFAC model as an innovative tool for monitoring natural organic matter removal in water treatment plants

Sikelelwa N. Ndiweni, Michael Chys, Nhamo Chaukura,
Stijn W. H. Van Hulle and Thabo T. I. Nkambule 

ABSTRACT

The increase of fluorescent natural organic matter (*f*NOM) fractions during drinking water treatment might lead to an increased coagulant dose and filter clogging, and can be a precursor for disinfection by-products. Consequently, efficient *f*NOM removal is essential, for which characterisation of *f*NOM fractions is crucial. This study aims to develop a robust monitoring tool for assessing *f*NOM fractions across water treatment processes. To achieve this, water samples were collected from six South African water treatment plants (WTPs) during winter and summer, and two plants in Belgium during spring. The removal of *f*NOM was monitored by assessing fluorescence excitation–emission matrices datasets using parallel factor analysis. The removal of *f*NOM during summer for South African WTPs was in the range 69–85%, and decreased to 42–64% in winter. In Belgian WTPs, *f*NOM removal was in the range 74–78%. Principal component analysis revealed a positive correlation between total fluorescence and total organic carbon (TOC). However, TOC had an insignificant contribution to the factors affecting *f*NOM removal. Overall, the study demonstrated the appearance of *f*NOM in the final chlorinated water, indicating that *f*NOM requires a customised monitoring technique.

Key words | excitation–emission matrices, fluorescent natural organic matter, parallel factor analysis, protein to humic ratio, water treatment processes

Sikelelwa N. Ndiweni

Nhamo Chaukura

Thabo T. I. Nkambule  (corresponding author)

Nanotechnology and Water Sustainability Research Unit, College of Science, Engineering and Technology, University of South Africa, Florida Campus, 1709, Johannesburg, South Africa
E-mail: nkambtt@unisa.ac.za

Michael Chys[†]

Stijn W. H. Van Hulle

LIWET, Department of Green Chemistry and Technology, Ghent University, Campus Kortrijk, Graaf Karel de Goedelaan 5, B-8500 Kortrijk, Belgium

[†]Present address: VEG-i-TEC, Ghent University, Campus Kortrijk, Graaf Karel De Goedelaan 5, B-8500 Kortrijk, Belgium

INTRODUCTION

Natural organic matter (NOM) in water sources reduces the performance of water treatment plants (WTPs), downgrades water quality, and increases operational costs (Sillanpaa *et al.* 2018). Fluorescent natural organic matter (*f*NOM) is largely made up of humic-like, fulvic-like, and protein-like components (Shao *et al.* 2014). Currently, parameters used for tracking recalcitrant *f*NOM are turbidity, total organic carbon (TOC), UV₂₅₄, pH, and conductivity (Ndiweni *et al.* 2019). However, these monitoring tools do not have capabilities to reveal enough information to fast track the removal of problematic *f*NOM fractions. For instance, online turbidity measurements for raw water were proven to be unreliable compared to tryptophan-like (tryp-like) and humic-like fluorescence (FL) measurements (Sorensen *et al.* 2018).

A conventional WTP commonly uses different water treatment stages such as coagulation, sedimentation, sand filtration (SF), and chlorination. The effects of *f*NOM on water treatment include fouling of filtration media such as membranes and activated carbon (AC) (Shao *et al.* 2014). Moreover,

during coagulation, flocculation, filtration, ozonation and chlorination, *f*NOM fractions can transform to form secondary products, which are more toxic and complex compared to the parent compounds (Brumer *et al.* 2019). Secondary product refers to compounds formed when ozone breaks down large molecules to small molecules that exhibit varying characteristics. Secondary product can also be used to refer to product formed when *f*NOM fractions react with chlorine disinfection by-products (DBPs), which are potentially formed (Chaukura *et al.* 2018). When traces of biodegradable *f*NOM are present in the chlorination stage, there is a possibility of bacterial regrowth and corrosion in the water distribution system (Baghoth & Amy 2012). These DBPs such as trihalomethanes and haloacetic acids stimulate cancer cells in humans; hence when optimising WTPs it is important to ensure that water is safe for consumption (Haarhoff *et al.* 2010).

Analysing unknown heterogeneous *f*NOM, some of which can occur at low relative concentrations or below detection limit, is a challenge (Liu *et al.* 2017). In order to identify

*f*NOM fractions in WTPs, fluorescence excitation-emission matrices (FEEM) have been investigated to measure fluorescent intensities of *f*NOM in water samples (Murphy *et al.* 2018). The technique is user-friendly, highly selective, sensitive, and time- and cost-saving (Markechová *et al.* 2013). FEEM reveals the spectral properties of *f*NOM fractions and the data can be further processed by various applications such as fluorescence regional integration, and multi-way data analysis methods such as parallel factor (PARAFAC) analysis to detect the presence of different *f*NOM fractions and perform predictions of *f*NOM removal throughout the water treatment plant (Cuss *et al.* 2016; Sun *et al.* 2017). The relatively large datasets, resulting from fluorescence measurements, can further be analysed and categorised based on the most dominating functional substituents of *f*NOM fractions such as humic acid-like (HA-like), fulvic acid-like (FA-like), and protein-like materials (Baghoth & Amy 2012; Murphy *et al.* 2018). The PARAFAC model assists to interpret *f*NOM fractions denoted as components, as well as their scores, which are directly proportional to the concentration of *f*NOM fractions (Sun *et al.* 2017).

Fluorescence spectroscopy, and the use of PARAFAC, has proven to be reliable because of its high sensitivity and selectivity on most dominant *f*NOM fractions (Sun *et al.* 2017). Moreover, PARAFAC sensitivity increases with an increase in the size of the datasets. It is therefore suitable to use FEEM-PARAFAC in developing a model that can reveal the actual concentration of fluorophores provided the quantum yields or specific absorption coefficients are known (Sun *et al.* 2017). The interpretation of *f*NOM fractions in WTPs can be supported using protein compounds to humic compounds (P/H) ratio, which monitors the removal efficiency in the treatment stages (Arango *et al.* 2017). In previous research, this ratio revealed that in spring and summer the occurrence of HA-like material was higher compared to autumn, which could be due to observed high volume of leaves falling and flourishing algae (Arango *et al.* 2017). In this study, PARAFAC analysis was used to monitor the removal of *f*NOM in WTPs. This study is a follow-up to the previous study which focused on characterisation the properties of WTP source water. The presence of HA-like and protein-like compounds was monitored using P/H ratios. The objectives of the study were to: (1) determine the composition of the different *f*NOM fractions within the drinking water treatment process using FEEM datasets and PARAFAC model derived data; (2) use previous generated knowledge to monitor the removal of *f*NOM fractions throughout the WTPs; and (3) trace major *f*NOM fractions in the final stage for drinking water.

MATERIALS AND METHODS

Sampling

Water samples were collected during the winter (August) of 2017, and the summer (January and February) and spring (June) of 2018 from six WTPs in South Africa, namely Plettenberg Bay (A), in the Western Cape; Umzinto (B), in the KwaZulu Natal; Reitvlei (I) in Gauteng; Midvaal (C and D) in the North West, and Ebenezer (E), and Flag Bushiello (F), both in Limpopo. In Belgium, samples were collected during the spring (June) of 2018, and the selected WTPs (operated by De Watergroep) were De Gavers (G) in Harelbeke and Dikkebus (H) in Ieper. The treatment stages are as indicated in Table 1. The samples were analysed straight away to prevent NOM degradation over time. The following parameters were measured onsite in triplicate: pH, turbidity, and conductivity for South African and Belgian treatment plants.

Further analysis of water samples

Total organic carbon was measured in triplicate using a Fusion TOC analyzer (Tekmar, Teledyne Instruments, USA) for South African samples, and using method 10129 and a TOC analyser (Hach, USA) for Belgian samples. To ensure they were representative of the organic matter in the WTPs, the samples were not treated further in the laboratory. For South African water samples, UV₂₅₄ absorbance was measured concurrently with fluorescence measurements, while a UV-Vis spectrophotometer (Shimadzu, 1601, Japan) was used for Belgian samples. An Aqualog spectrofluorometer with built-in UV-Vis (HORIBA Jobin Yvon AquaLog, USA) was used to measure *f*NOM using Raman water as a blank in a quartz cuvette with 1 cm width light path length. The EEMs datasets for South African samples were obtained at 1 s integration time for excitation wavelengths 239.00–800.00 nm at 3.00 nm intervals, and emission wavelengths of 246.70–825.13 nm at an average interval of 4.54 nm. For Belgium samples, fluorescence spectra were obtained using a fluorimeter (Shimadzu, RF-5301, Japan) at 0.25 s integration time for excitation wavelengths 220.00–450.00 nm at 5 nm intervals, and emission wavelengths of 280.00–600.00 nm at 1.00 nm increments. The excitation and emission slits width were 5 nm for all measurements. The Raman scan was performed at excitation and emission slit width of 5 nm and a response time of 0.25 s using demineralised water at an excitation wavelength of 350.00 nm and an

Table 1 | Monitoring parameters (turbidity, TOC, UV₂₅₄ and PARAFAC outputs) removal from raw to the final water in plants in South Africa and Belgium

WTPs	Turbidity removal from raw to the final water (%)		TOC removal from raw to the final water (%)		UV ₂₅₄ removal from raw to the final water (%)		PARAFAC components removed from raw water to final water (%)				
	Measured	Literature reference	Measured	Literature reference	Measured	Literature reference	TF	Fmax1	Fmax2	Fmax3	Literature reference
Plant A: raw (surface water) – coagulation – settling – SF (sand filtration) – final	59.30		46.44	NOM 50% for raw water with 3.8 mg/L DOC (Callegari <i>et al.</i> 2017); TOC 47% for raw water with 5.6 mg/l (Matilainen <i>et al.</i> 2002); 48.70 (Sillanpaa <i>et al.</i> 2018)	76.67		72.25	75.38	100.00	9.12	
Plant B1: raw – coagulation – settling – SF – final			26.62		71.80	70.00 (Sillanpaa <i>et al.</i> 2018)	59.44	47.59	56.03	89.45	
Plant F: raw (surface water) – coagulation – settling – SF – final	99.05		60.72	TOC 68% (Nissinen <i>et al.</i> 2001)	81.82		89.22	86.62	100.00	80.04	Humic fractions 91% (Nissinen <i>et al.</i> 2001). TF 86% (HA-like, tyrosine-like, protein-like, tryptophan-like) (Markechová <i>et al.</i> 2013)
Plant I: raw (surface water) – coagulation – DAF – final	winter 75.66, summer 97.98		winter 0.51, summer 29.59		winter -, summer 38.10	36.10 (Sillanpaa <i>et al.</i> 2018)	41.77	44.32	27.34	45.37	Humic substances 44% (Volk <i>et al.</i> 2005)
Plant C and D: raw (surface water) – pre-ozonation – coagulation – DAF – post-ozonation – SF – final	winter 99.40, summer 99.68						69.30	59.81	58.38	83.34	TF ~50–70% (HA-like and protein-like), (Baghoth & Amy 2012). humic substances 65% (Sillanpaa <i>et al.</i> 2018)
Plant E: raw (surface water) – aeration – coagulation – SF – final	winter 89.67, summer 98.70		winter, summer 32.10				50.68	60.27	52.34	31.81	
Plant F: raw (surface water) – coagulation – settling – SF – final	winter 74.74, summer 74.98		winter 0.29, summer		winter -, summer 33.33		49.94	61.61	46.90	56.28	
Plant G: raw (surface water) – nitrification pond – SF – ultra-filtration – AC filtration – final	94.33	Turbidity 95% for raw water with 5.5 FTU (Matilainen <i>et al.</i> 2002)	100.00	NOM 90% for raw water with TOC > 10 mg/L (Volk <i>et al.</i> 2005)	90	90.00 (Sillanpaa <i>et al.</i> 2018)	77.98	84.33	87.30		
Plant H: raw – coagulation – DAF – SF – AC – final	99.75		100.00		78.95		74.34	79.21	79.37	20.54	

DAF, dissolved air flotation; DOC, dissolved organic carbon; TF, total fluorescence.

emission wavelength of 365.00–450.00 nm using 0.20 nm increments.

Method development for online monitoring of fluorescent natural organic matter

A modelling technique, PARAFAC, was used to mathematically separate *f*NOM fractions into a trilinear setup using an alternating least square algorithm. The PARAFAC model was developed in Matlab R2016a environment and using the DOMFlour tool and method (Cuss *et al.* 2016). The EEMs datasets correction was achieved by blank subtraction using the decomposition routines for the excitation–emission matrices (drEEM) toolbox to remove the influence of sample handling and instrument measurement instability. The measurements of UV absorbance were used to correct for the inner filter. Pre-processing of spectral corrected EEMs was limited to non-negative values and PARAFAC model running at no particular order. The fluorescence intensities of the samples were normalised using the area given by the Raman measurements and thereafter the units of sample fluorescence intensity were Raman units (RU) to ensure consistency (Sun *et al.* 2017). Robust PARAFAC models were developed by eliminating outliers and visually inspecting residuals to ensure they only consisted of noise. Prior PARAFAC model validation of the loadings of *f*NOM spectra was examined for accuracy and outliers removed. Inappropriate spectra were improved by adjusting iteration repetition and convergence from the default of 10^{-6} to a stricter 10^{-8} and 10^{-10} . Thereafter, the split half analysis technique was performed to validate that the same PARAFAC model can be achieved through grouping differently EEMs datasets. The validated components were recorded in the drEEM toolbox (Markechová *et al.* 2013). The *f*NOM fractions were identified based on documented emission and excitation wavelength peaks. The HA-like fractions were located at emission and excitation wavelengths within the range 380–540 nm and 250–340 nm, FA-like range 380–540 nm and 200–250 nm, aromatic protein I range 280 nm–330 nm and 200 nm–250 nm, aromatic protein II range 330–380 nm and 200–250 nm, and soluble microbial products range 280–380 nm and 250–300 nm (Ndiweni *et al.* 2019). The P/H ratios (Equation (1)) were calculated using fluorescence intensities at specific wavelengths, whereby for protein-like compounds excitation at 275.00 nm and emission at 340.00 nm, and for humic material excitation at 350.00 nm and emission at 480.00 nm were used

(Arango *et al.* 2017).

$$\frac{P}{H} = \frac{\text{Intensity at } \lambda_{(ex275 \& \text{em}340)}}{\text{Intensity at } \lambda_{(ex350 \& \text{em}480)}} \quad (1)$$

where λ_{ex} is the excitation wavelength and λ_{em} is the emission wavelength.

RESULTS AND DISCUSSION

Quality parameters for water samples

The pH measurements for all the WTPs investigated in this study were consistently in the range 5.84–9.27, which is a huge variation, but for final disinfection stage the pH was within the SANS 241 acceptable level. Notably, the final stage for plant A had a high pH (9.27 ± 0.04). Although this pH is not harmful to human health, it causes aesthetic problems such as alkaline-like taste. The variation of turbidity and its removal do not indicate huge seasonal differences. In previous studies, turbidity removal was 50.00–90.00% and 95.00% for a four-stage conventional WTPs with raw water turbidity measured at 4.80 NTU and 5.50 NTU respectively (Matilainen *et al.* 2002; Callegari *et al.* 2017). In this study, the four-stage conventional WTPs' raw water turbidity was in the range 3.00–6.50 NTU and removal percentage was within the range 74.81–99.05% (Table 1). Turbidity substances facilitate transportation and protection and provide food for microbial species, and should, therefore, be kept to a minimum (Farrell *et al.* 2018).

In plant L, the nitrification process removes ammonium, which is an indicator of fecal pollution and the presence of nitrogen-based fertilisers (Westgate 2009). A reduction of 94.33% in the turbidity was observed for plants K and L, respectively (Table 1). However, the filtration stages were effective in reducing the turbidity to less than 1.00 NTU. The removal of the TOC for WTPs A to F from raw water to the final stages were within the range 21.00–73.00% during summer, whereas WTPs G, H, I, and J were within the range 13.00–24.00%. However, the removal of the TOC by WTPs G, H, I, and J during winter from raw water to the final stage was within the range 1.72–32.09%. Plant F TOC removal corresponded to previous studies (Nissinen *et al.* 2001; Callegari *et al.* 2017). The contamination of water sources originates mainly from wastewater treatment plants and informal settlements with poor waste management of sewage (Chaukura *et al.* 2018). During summer (January and February), South Africa experiences

high levels of rainfall that generally keep dams and rivers full, potentially diluting TOC concentrations. The removals of TOC in Belgium WTPs using SF and ultra-filtration (UF) membranes were 52 and >98.00%, respectively. A previous study indicated that UF is more effective for the removal of large molecular weight NOM (Haarhoff *et al.* 2010). In another study, UF showed maximum efficiency of 90.00% (Prisciandaro & di Celso 2016). However, the efficiency of the membrane is compromised by fNOM fractions that cause fouling (Sillanpaa *et al.* 2018). The values recorded in South African winter (June and July) of 0.12–0.48 cm⁻¹, in South African summer (January and February) of 0.05–0.13 cm⁻¹, and in Belgian spring (May) of 0.01–0.04 cm⁻¹ of UV₂₅₄ for the final water indicate traces of light-absorbing aromatic compounds which are fNOM. It is deduced that there is a persistence of fNOM based UV₂₅₄ measurements in the final disinfection stage. Although it reveals limited information, UV₂₅₄ is currently a valuable tool to monitor NOM on-line.

A monitoring tool for fluorescence natural organic matter in WTPs

The EEM datasets were processed in PARAFAC analysis to calculate the values of the three validated major contributing components per plant at the maximum intensity, also referred to as relative concentrations. The wavelengths and the three-dimensional fluorescence spectra of the isolated major fNOM fractions in this study, classified as Fmax1, Fmax2 and Fmax3, are described in detail in a previous publication (Ndiweni *et al.* 2019). The maximum intensity (Fmax) during the South African winter (June and July) for Fmax1 raw water and final water was in the range 4.89–0.68 and 2.16–0.20 RU, respectively. During the South African summer (January and February) Fmax1 was in the range 1.62–0.05 and 0.60–0.02 RU in the raw water and final water, respectively. However, during the Belgian spring (May) Fmax1 was in the range 1.36–0.76 and 0.28–0.13 RU in the raw water and final water, respectively. The Fmax2 outcomes during winter were in the range 1.11–0.28 and 1.38–0.14 RU for raw water and final water, respectively, while during summer they were in the range 0.61–0.11 and 0.24 RU to less than the detection limit for raw water and final water, respectively. However, during spring Fmax2 was 0.86–0.47 and 0.11–0.10 RU for raw water and final water, respectively. The Fmax3 values during winter were in the range 0.54–0.23 and 0.66–0.08 RU for raw water and final water, respectively. In summer Fmax3 was in the range 0.57–0.02 and

0.37–0.01 RU for raw water and final water, respectively, while in spring it was in the range 0.46–0.10 and 0.31–0.03 RU for raw water and for final water, respectively. It is observed that this study is in agreement with a previous study in terms of sensitivity (Murphy *et al.* 2018). Samples with Fmax as low as 0.01 RU were detected in the final water.

Fmax1 mainly consist of fNOM fractions with large molecular weight and high hydrophobicity such as HA-like material, which is rich in terrestrial materials. HA-like material is dominant in water bodies due to decomposition of dead animals and plants (Shao *et al.* 2014; Ndiweni *et al.* 2019). Notably, Fmax2 consists of fNOM fractions that are a mixture of both high and low hydrophilicity and hydrophobicity as well as large and small molecular weight material. However, Fmax3 consists of protein-like substances, soluble microbial by-products (SMPs), which possess both low hydrophobicity and low aromaticity, and small molecular weight fractions (Zhu *et al.* 2017). Hence monitoring the removal of fNOM fractions throughout the WTPs (Figure 1) is important for fast-tracking changes in composition of fNOM fractions. This study tracked evolution of Fmax throughout the treatment process as a follow-up to a previous validated model for treatment plant water from sources such as rivers, dam and borehole (Ndiweni *et al.* 2019). Using PARAFAC analysis, a previous study isolated three major fNOM fractions from a diverse reservoir of heterogeneous polydisperse carbon-based species (Murphy *et al.* 2018). The fNOM total fluorescence (TF) removal during summer for South African WTPs ranged 69–85% and decreased to 42–64% in winter. In Belgian WTPs, fNOM removal ranged 74–78% (Table 1). A comparison of datasets with literature showed comparable fNOM removals (Table 1). In aquatic systems, photodegradation and biodegradation processes contribute to heterogeneous properties of fNOM (Hansen *et al.* 2016). For plants A–F, Fmax2 consisting of trypt-like and protein-like substances containing traces of SMPs and HA-like substances was efficiently removed by the filtration step (Figure 1(a)).

Furthermore, Fmax2 removal was possibly enhanced by variation in its composition such as aromaticity, molecular weight, and charge. Plant A and B Fmax1, which consists of HA-like materials, had removal range 58–79%, while Fmax2 had >98% removal. In the treatment process primary stages such as coagulation and flocculation target Fmax1 based on its characteristics such as high hydrophobicity, large molecular weight and its particle charge. In the succeeding treatment stages both Fmax2 and Fmax3 fNOM removals increase as they are targeted by final stages which are SF, AC filtration and UF (Markechová *et al.* 2013). The humic material removal was in agreement with previous studies for plant B and I

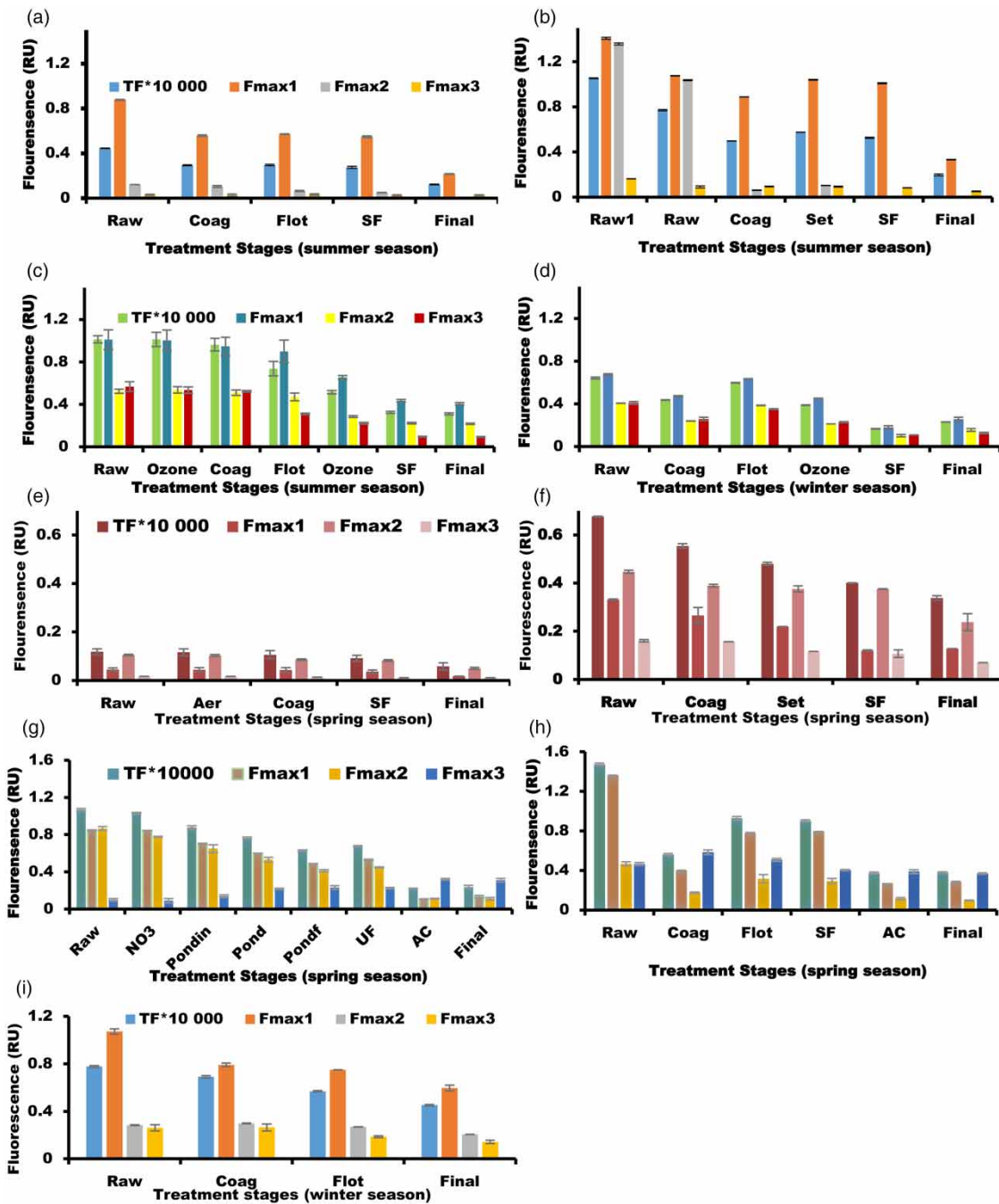


Figure 1 | The removal of fluorescence natural organic matter in water treatment plants monitored using total fluorescence and maximum fluorescence. TF is total fluorescence divided by 10,000 to make data handling much easier. (a)–(i): Plants A to I, respectively. Aer: aeration; Coag: coagulation; Set: settling tank; Flot: flotation; Nitri: nitrification stage; Pondin: pond inlet; Pondf: pond outlet.

(Nissinen *et al.* 2001; Volk *et al.* 2005). A relative increase of Fmax3 was observed in the settling and SF steps for plant B denoted as aromatic protein II with traces of SMPs and amino acids. This relative increase was also observed for

plant B settling for TF, Fmax1 and Fmax2; plant F final stage for Fmax1; plant I coagulation; and plant C and D flotation and final stage for Fmax1 and Fmax3 (Figure 1). These process can promote carbon reduction and a shift in

the functional groups of *f*NOM fractions (Khan *et al.* 2019). The Fmax3 for plant K which is aromatic protein I and II with traces of FA-like shows relative increase by 20.54% compared with Fmax1 and Fmax2 in the final stage.

Although the nitrification process for Plant H removes inorganic nitrogen, its removal efficiency for organic nitrogen is limited, and algae growth will therefore persist (Westgate 2009). As was mentioned in a previous study, for plant H the presence of algae that grow is highly possible due to the fertiliser distributed for agricultural reasons and run-off from the nearby land (Ndiweni *et al.* 2019). The removal of aromatic proteins decreased in the SF and AC filtration due to pore blockage and the possibility of a decrease in the adsorption capacity of the AC (Yang *et al.* 2018). Organic compounds such as trypt-like and protein-like substances are indicators of the significant contribution of wastewater effluent to the *f*NOM in water sources (Yang *et al.* 2018). For plant C and D, the removal of the hydrophobic acid containing traces of FA- and HA-like substances was 79.21%, which was mostly removed at the AC stage (Figure 1). The removal of SMPs with traces of aromatic protein was 79.37% and was mostly removed in the AC stage. Generally, the presence of aromatics continued to increase during the treatment process, possibly because the component consisting of proteins was not well defined; thus a small increase lies in the accuracy of the model and measurements. Based on the analysis of the results, the pressing issue for the development of the monitoring tool for discovered *f*NOM fractions is their appearance in the final water. This is an indication that WTPs are not operating optimally (Sillanpaa *et al.* 2018). The presence of *f*NOM fractions presents challenges such as the formation of secondary products due to accumulation of pollutants in the unit processes, which encourages biological activity within the treatment train (Brumer *et al.* 2019). The secondary products differ in chemical composition from their primary *f*NOM fractions, decreasing the removal efficiency and increasing toxicity (Brumer *et al.* 2019). The toxicity and side effects of traces of *f*NOM fractions remain unknown, hence the need to develop a monitoring tool that can rapidly track the occurrence of unknown *f*NOM fractions trace in the final treated water. When the characteristics of *f*NOM are known, it will be possible to optimise treatment stages, chemical dosage, and plant design for optimum *f*NOM fractions removal. Using recent technology, WTPs can improve their efficiency through atomisation, and predictions of the occurrence of *f*NOM fractions due to environmental activities and seasonal changes.

Box 1 | Equations based on measured and calculated drinking water samples parameters

$$\begin{aligned} \text{PC1} = & 0.77[\text{TF}] + 0.75[\text{TF removal}\%] + 0.83[\text{Fmax1}] \\ & + 0.92[\text{Fmax1 removal}\%] \\ & + 0.91[\text{Fmax2 removal}\%] \end{aligned} \quad (1)$$

$$\begin{aligned} \text{PC2} = & 0.67[\text{Fmax3}] + 0.84[\text{Conductivity}] + 0.88 \left[\frac{\text{P}}{\text{H}} \right] \\ & + 0.91[\text{FI}] + 0.62[\beta:\alpha] + 0.77[\text{BIX}] \end{aligned} \quad (2)$$

$$\text{PC3} = 0.53[\text{TF}] + 0.73[\text{Fmax2}] + 0.69[\text{UV254}] \quad (3)$$

Correlation between measured and calculated parameters

To track the correlation the TF, Fmax, H/P ratio, fluorescence index, biological index, freshness index and site-measured parameter, principal component analysis (PCA) was employed. The indices were calculated for influent water as explained in a previous study (Box 1) (Ndiweni *et al.* 2019). The rotated component matrix illustrated main factors influencing the evolution in *f*NOM fractions throughout the treatment plants (Figure 2(a) and 2(b)). To ensure that the output of PCA is reliable a test was conducted using a Kaiser-Meyer-Olkin criterion, which had a value of 0.67, which is acceptable although the closer it is to 1.00 the more reliable it is (Ndiweni *et al.* 2019). The eigenvalues for principal component (PC) factors were 4.48, 3.83, and 1.50 for PC1, PC2, and PC3, respectively (Table 2). PC1 explained that the variance of 35.22% correlated to Fmax1, Fmax1 removal, Fmax2 removal, TF and TF removal. It is observed that the TF and TOC can be correlated; however, TOC had insignificant contribution in PC1 (Figure 2(a)). The main variables influencing the removal of *f*NOM are illustrated in Figure 2(b). PC1 is strongly correlated to Fmax information; therefore it can be used as a measure of fluorescence relative concentration. PC2 explains that 29.42% variation correlated to Fmax3, conductivity, P/H ratio, FI (fluorescence index), HIX (humification index), $\beta:\alpha$ (freshness index; β = recently produced NOM fractions, α = aged NOM fractions), and BIX (biological index); therefore it can be used as a measure of indicators. PC3 explains that 11.54% variation correlated to TF, Fmax2, UV254 be used as measure of UV254 related to Fmax2 *f*NOM fractions.

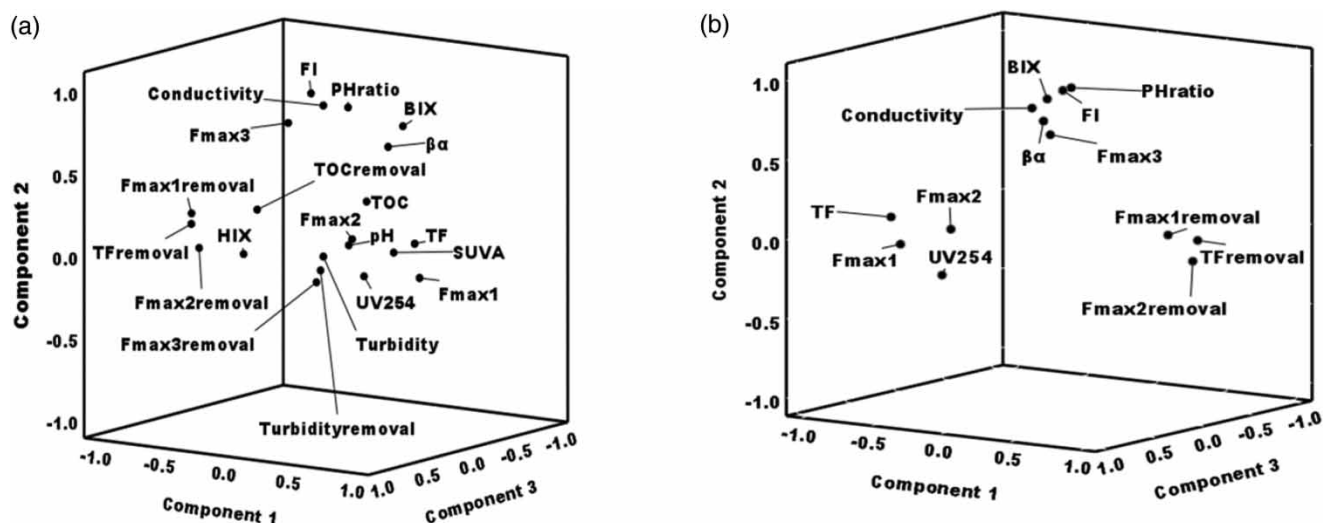


Figure 2 | Principal component analysis for drinking water treatment plants in (a) South Africa, and (b) Belgium.

Table 2 | Principal component analysis for variables and their factors influencing *f*NOM removal throughout the treatment plant

Variables	Factors		
	1	2	3
TF	-0.76 ^a	0.10	0.53 ^a
TF removal%	0.97 ^a	0.03	-0.14
Fmax1	-0.83 ^a	-0.11	0.34
Fmax3	0.17	0.67 ^a	0.23
Conductivity (μS/cm)	0.10	0.84 ^a	0.32
Fmax2	-0.20	0.10	0.73 ^a
Fmax1 removal%	0.92 ^a	0.10	0.09
Fmax2 removal%	0.91 ^a	-0.12	-0.17
UV ₂₅₄	-0.29	-0.20	0.69 ^a
P/H ratio	0.00	0.88 ^a	-0.22
FI	0.12	0.91 ^a	0.04
$\beta:\alpha$	-0.27	0.62 ^a	-0.33
BIX	-0.24	0.77 ^a	-0.33
Eigenvalues	4.48	3.83	1.50
% of explained variance	35.22	29.42	11.54

^aFactors with significant loadings.

It was observed that the overall TF intensity decreased with TOC from raw to final water for plants in South Africa (Figure A1, Supplementary Material). These results are in agreement with a previous study that showed a direct correlation between TF intensity and TOC (Baghoth & Amy 2012). However, the linearity of this relationship can be compromised mainly by high dissolved organic

carbon concentration, and reduced intensity by species such as nitrates (Baghoth & Amy 2012). TOC represents the carbon content that is calculated as dissolved organic and particulate carbon, whereas TF denotes the intensity due to problematic targeted *f*NOM (unsaturated compounds). Using TOC, the carbon content was tracked throughout the WTPs. However, with TF this study revealed that there is a presence of persistent *f*NOM in final water. In the absence of sufficient residual disinfectant, the traces of *f*NOM in the final water may lead to microbial regrowth in the distribution system (Ndiweni *et al.* 2019). The TF data gives better information of the targeted problematic *f*NOM compared to the traditional measurement monitored, which is TOC. The PCA revealed that the TOC throughout the treatment plant had insignificant influence on the removal of *f*NOM fractions. Hence TF should be taken into consideration when evaluating the WTP efficiency and optimisation.

The protein to the humic ratio

In treatment processes the most dominating *f*NOM fractions are humic and protein material; their characteristics are as explained above. Therefore calculation of humic and protein ratio assists in development of *f*NOM fractions monitoring based on the most dominating components. The P/H ratio revealed that for South African WTPs there is an abundance of humic material compared to protein compounds (P/H ratio is <0.6) (Figure 3). However, for Belgian plants (P/H ≥ 1), presence of algae in the influent occurs at higher levels, which causes degradation of humified

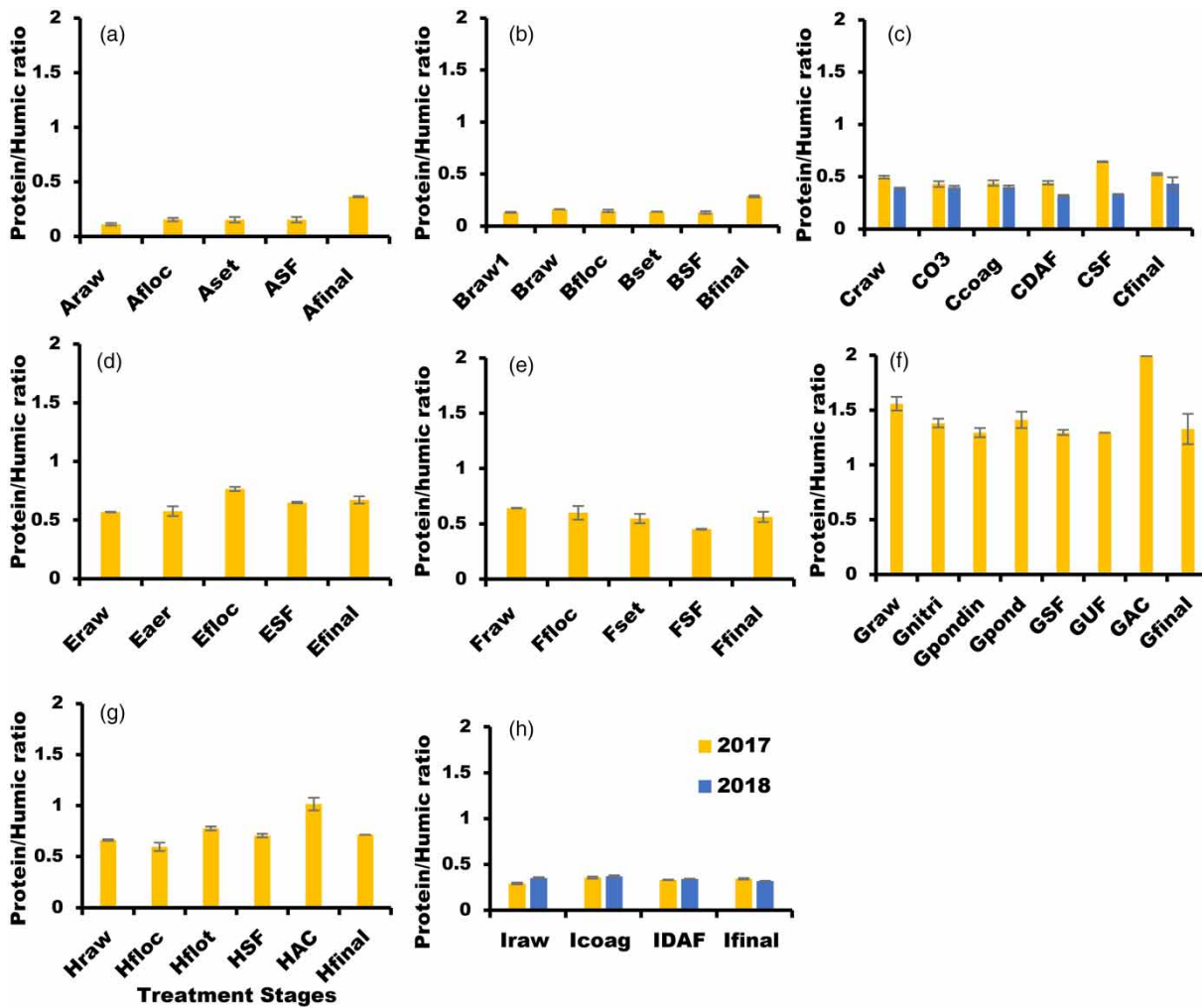


Figure 3 | Tracking the presence of protein-like to the humic-like (P/H) ratio during the different treatment stages for drinking water treatment plants in South Africa and Belgium. In treatment stage abbreviations, the first letter refers to the plant; for explanation of the rest of the abbreviation, please see text and Figure 1 caption.

materials as the $P/H \geq 1$ for plant G and for plant H is >0.6 (Figure 3) (Brumer *et al.* 2019). Furthermore, these findings demonstrated that: (1) the pool of fNOM enormously depends on pollutants in the raw water and (2) the efficiency of WTPs is limited by the composition of fNOM. Previous studies indicate that protein-like compounds dominate due to the influence of algal boom, pollutants resulting from agricultural activities, and poorly treated wastewater effluent (Sorensen *et al.* 2018). Notably humic-like compounds are mostly removed in the coagulation, whereas humified material is not effectively removed prior to the SF or AC stages. However, they tend to compete for active sites with protein-like compounds (Zanacic *et al.* 2016). These results show that there is a significant reduction in the removal rate of the targeted protein-like compounds

and the life span of the filtration medium is compromised. The higher P/H ratio values in the final water compared to raw water confirms that the traces of humified materials are higher than that of proteins compounds. This increase in P/H ratio may also be an indication of the formation of secondary fNOM fractions (Brumer *et al.* 2019). If the traces of humic compounds continue to increase in the final water, it may give rise to carcinogenic compounds which result when fNOM reacts with chlorine in the final water (Chaukura *et al.* 2018). The final water from plant G had higher humic materials in winter compared to summer. In comparison, plant C had water entering the WTP with high microbial activities processing humified materials in winter. However, both plants had higher humic compounds compared to the summer season. A

previous study reported that during summer fNOM was not destroyed; instead there are spectral shifts which indicate transformation of the fNOM (Khan *et al.* 2019).

CONCLUSIONS

The removal of fNOM fractions during summer for South African WTPs was in the range 69–85% and decreased to 42–64% in winter. In Belgian WTPs, fNOM fraction removal ranged 74–78%. The removal or alteration of fNOM fractions is influenced by the chemical structure of fNOM, water treatment regimen, and seasonal changes. Although a low removal percentage does not reflect a high amount of fNOM fractions, the amount in the influent should be taken into consideration. Principal component analysis revealed that TF and TOC can be correlated: however, TOC had insignificant contribution to the factors affecting removal of fNOM. The P/H ratio revealed that for South African WTPs there is an abundance of humic material compared to protein compounds. However, for Belgian plants, it is likely that microbial activity occurs at higher levels, degrading humic materials. This PARAFAC model demonstrated there is appearance of fNOM fractions in the final chlorinated water. Hence further studies can predict amount, toxicity and side-effects of traces of fNOM fractions in effluent.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this paper is available online at <https://dx.doi.org/10.2166/wst.2020.136>.

REFERENCES

- Arango, C. P., Beaulieu, J. J., Fritz, K. M., Hill, B. H., Elonen, C. M. & Pennino, M. J. 2017 **Urban infrastructure influences dissolved organic matter quality and bacterial metabolism in an urban stream network.** *Freshwater Biology* **62** (11), 1917–1928. <https://doi.org/10.1111/fwb.13035>.
- Baghoth, S. A. & Amy, G. L. 2012 *Characterizing Natural Organic Matter in Drinking Water Treatment Processes and Trains.* Thesis, Delft University of Technology, The Netherlands.
- Brumer, A. M., Vughs, D., Siegers, W., Bertelkamp, C., Hofman-Caris, R. & Kolkman, A. 2019 **Monitoring transformation product formation in the drinking water treatments rapid sand filtration and ozonation.** *Chemosphere* **214**, 801–811. <https://doi.org/10.1016/j.chemosphere.2018.09.140>.
- Callegari, A., Boguniewicz-Zablocka, J. & Capodaglio, A. G. 2017 **Experimental application of an advanced separation process for NOM removal from surface drinking water supply.** *Separations* **4** (32). <https://doi.org/10.3390/separations4040032>.
- Chaukura, N., Ndlangamandla, N. G., Moyo, W., Msagati, T. A. M., Mamba, B. B. & Nkambule, T. T. I. 2018 **Natural organic matter in aquatic systems – a South African perspective.** *Water SA* **44** (4), 624–635. <http://dx.doi.org/10.4314/wsa.v44i4.11>.
- Cuss, C. W., McConnell, S. M. & Guéguen, C. 2016 **Combining parallel factor analysis and machine learning for the classification of dissolved organic matter according to source using fluorescence signatures.** *Chemosphere* **155**, 283–291. <https://doi.org/10.1016/j.chemosphere.2016.04.061>.
- Farrell, C., Hassard, F., Jefferson, B., Leziart, T., Nocker, A. & Jarvis, P. 2018 **Turbidity composition and the relationship with microbial attachment and UV inactivation efficacy.** *Science of the Total Environment* **624**, 638–647. <https://doi.org/10.1016/j.scitotenv.2017.12.173>.
- Haarhoff, J., Kubare, M., Mamba, B., Krause, R., Nkambule, T., Matsebula, B. & Menge, J. 2010 **NOM characterization and removal at six Southern African water treatment plants.** *Drinking Water Engineering Science* **3**, 53–61. <https://doi.org/10.5194/dwes-3-53-2010>.
- Hansen, A. M., Kraus, T. E. C., Pellerin, B. A., Fleck, J. A., Downing, B. D. & Bergamaschi, B. A. 2016 **Optical properties of dissolved organic matter (DOM): effects of biological and photolytic degradation.** *Limnology and Oceanography* **61** (3), 1015–1032. <https://doi.org/10.1002/lno.10270>.
- Khan, S. I., Zamyadi, A., Rao, N. R. H., Li, X., Stuetz, R. M. & Henderson, R. K. 2019 **Fluorescence spectroscopic characterisation of algal organic matter: towards improved in situ fluorometer development.** *Environmental Science: Water Research & Technology* **5** (2), 417–432. <https://doi.org/10.1039/C8EW00731D>.
- Liu, G., Ya Zhang, Y., Knibbe, W. J., Feng, C., Liu, W., Medema, G. & Meer, W. 2017 **Potential impacts of changing supply-water quality on drinking water distribution: a review.** *Water Research* **16**, 135–148.
- Markechová, D., Tomková, M. & Sádecká, J. 2013 **Fluorescence excitation-emission matrix spectroscopy and parallel factor analysis in drinking water treatment: a review.** *Polish Journal of Environmental Studies* **22** (5), 1289–1295.
- Matilainen, A., Lindqvist, N., Korhonen, S. & Tuhkanen, T. 2002 **Removal of NOM in the different stages of the water treatment process.** *Environment International* **28** (6), 457–465.
- Murphy, K. R., Timko, S. A., Gonsior, M., Powers, L. C., Wunsch, U. J. & Stedmon, C. A. 2018 **Photochemistry illuminates ubiquitous organic matter fluorescence spectra.** *Environmental Science and Technology* **52** (19), 11243–11250. <https://doi.org/10.1021/acs.est.8b02648>.
- Ndiweni, S. N., Chys, M., Chaukura, N., Van Hulle, S. W. H. & Nkambule, T. T. I. 2019 **Assessing the impact of environmental activities on natural organic matter in South Africa and Belgium.** *Environmental Technology* **40** (13),

- 1756–1768. <https://doi.org/10.1080/09593330.2019.1575920>.
- Nissinen, T. K., Miettinen, I. T., Martikainen, P. J. & Vartiainen, T. 2001 Molecular size distribution of natural organic matter in raw and drinking waters. *Chemosphere* **45**, 865–873.
- Prisciandaro, M. & di Celso, G. M. 2016 On the removal of natural organic matter from superficial water by using UF and MF membranes. *Desalination and Water Treatment* **57** (6), 2481–2488. <https://doi.org/10.1080/19443994.2015.1031184>.
- Shao, S., Liang, H., Qu, F., Yu, H., Li, K. & Li, G. 2014 Fluorescent natural organic matter fractions responsible for ultrafiltration membrane fouling: identification by adsorption pretreatment coupled with parallel factor analysis of excitation–emission matrices. *Journal of Membrane Science* **464**, 33–42.
- Sillanpaa, M., Ncibi, M. C., Matilainen, A. & Vepsalainen, M. 2018 Removal of natural organic matter in drinking water treatment by coagulation: a comprehensive review. *Chemosphere* **190**, 54–71. <https://doi.org/10.1016/j.chemosphere.2017.09.113>.
- Sorensen, J. P. R., Vivanco, A., Ascott, M. J., Goody, D. C., Lapworth, D. J. & Read, D. S. 2018 Online fluorescence spectroscopy for the real-time evaluation of the microbial quality of drinking water. *Water Research* **137**, 301–309. <https://doi.org/10.1016/j.watres.2018.03.001>.
- Sun, H. Y., Koal, P., Gerl, G., Schroll, R., Joergensen, R. G. & Munch, J. C. 2017 Water-extractable organic matter and its fluorescence fractions in response to minimum tillage and organic farming in a Cambisol. *Chemical and Biological Technologies in Agriculture* **4** (15), 1–11. <https://doi.org/10.1186/s40538-017-0097-5>.
- Volk, C., Kaplan, L. A., Robinson, J., Johnson, B., Larry, W., Zhu, H. W. & Lechevallier, M. 2005 Fluctuations of dissolved organic matter in river used for drinking water and impacts on conventional treatment plant performance. *Environmental Science Technology* **39**, 4258–4264. <https://doi.org/10.1021/es040480k>.
- Westgate, P. J. 2009 *Characterization of Proteins in Effluents from Three Wastewater Treatment Plants that Discharge to the Connecticut River*. Thesis, University of Massachusetts Amherst.
- Yang, X., Yu, X., Cheng, J., Zheng, R., Wang, K. & Dai, Y. 2018 Impacts of land-use on surface waters at the watershed scale in southeastern China: insight from fluorescence excitation-emission matrix and PARAFAC. *Science of the Total Environment* **627**, 647–657. <https://doi.org/10.1016/j.scitotenv.2018.01.279>.
- Zanacic, E., Stavrinides, J. & Mcmartin, D. W. 2016 Field-analysis of potable water quality and ozone efficiency in ozone-assisted biological filtration systems for surface water treatment. *Water Research* **104**, 397–407. <https://doi.org/10.1016/j.watres.2016.08.043>.
- Zhu, G., Bian, Y., Hursthouse, A. S., Wan, P., Szymanska, K., Ma, J., Wang, X. & Zhao, Z. 2017 Application of 3-D fluorescence: characterization of natural organic matter in natural water and water purification systems. *Journal of Fluorescence* **27**, 2069–2094. doi:10.1007/s10895-017-2146-7.

First received 20 September 2019; accepted in revised form 19 March 2020. Available online 26 March 2020