Water Research 227 (2022) 119314

Contents lists available at ScienceDirect

Water Research

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

NOM fractionation by HPSEC-DAD-OCD for predicting trihalomethane disinfection by-product formation potential in full-scale drinking water treatment plants

Meritxell Valenti-Quiroga [a], Pepus Daunis-i-Estadella [b], Pere Emiliano [c], Fernando Valero [c], Maria J. Martin [a,*]

[a] LEQUIA. Institute of the Environment, Universitat de Girona, Carrer Maria Aurèlia Capmany, 69, Girona E-17003, Spain
[b] Department of Computer Science, Applied Mathematics and Statistics, Universitat de Girona, Carrer Universitat de Girona, 6, Girona E-17003, Spain
[c] Ens d'Abastament d'Aigua Ter-Llobregat (ATL), Sant Martí de l'Erm 2, E-08970 Sant Joan Despí, Barcelona, Spain

ARTICLE INFO

Keywords:
Disinfection by-products
Formation potentials
Natural organic matter
Multiple linear regression
NOM fractionation
Size exclusion chromatography

ABSTRACT

Chlorination is a common method for water disinfection; however, it leads to the formation of disinfection by-products (DBPs), which are undesirable toxic pollutants. To prevent their formation, it is crucial to understand the reactivity of natural organic matter (NOM), which is considered a dominant precursor of DBPs. We propose a novel size exclusion chromatography (SEC) approach to evaluate NOM reactivity and the formation potential of total trihalomethanes-formation potentials (tTHMs-FP) and four regulated species (i.e. $CHCl_3$, $CHBrCl_2$, $CHBr_2Cl$, and $CHBr_3$). This method combines enhanced SEC separation with two analytical columns working in tandem and quantification of apparent molecular weight (AMW) NOM fractions using C content (organic carbon detector, OCD), 254-nm spectroscopic (diode-array detector, DAD) measurements, and spectral slopes at low ($S_{206–240}$) and high ($S_{350–380}$) wavelengths. Links between THMs-FP and NOM fractions from high performance size exclusion chromatography HPSEC-DAD-OCD were investigated using statistical modelling with multiple linear regressions for samples taken alongside conventional full-scale as well as full- and pilot-scale electrodialysis reversal and bench-scale ion exchange resins. The proposed models revealed promising correlations between the AMW NOM fractions and the THMs-FP. Methodological changes increased fractionated signal correlations relative to bulk regressions, especially in the proposed HPSEC-DAD-OCD method. Furthermore, spectroscopic models based on fractionated signals are presented, providing a promising approach to predict THMs-FP simultaneously considering the effect of the dominant THMs precursors, NOM and Br^-.

1. Introduction

Drinking water treatment facilities (DWTPs) provides safe drinking water to consumers. This seemingly straightforward day-to-day service results from many complex procedures. The most ubiquitous health concern in surface water sources is related to microbial contamination, making it vital to properly control. To counteract this risk, several treatment units act as chemical and physical barriers against pathogens and pollutants, of which chlorination disinfection plays a key role.

DBPs are the unintended outcome of disinfecting water with chlorination and are chemical risks to drinking water. Over 800 DBPs have been reported (Jiang et al., 2020), which can be classified as aromatic and aliphatic. Neither all DBPs account for the same toxicological effects nor are they found in the same amounts, which can differ notably. Nitrogenous DBPs are reported to be more toxic than carbonaceous (Plewa et al., 2017); however, trihalomethanes (THMs) and haloacetic acids (HAAs) are some of the most common DBPs found at higher concentrations globally (Krasner et al., 2006; Zhang et al., 2020) and are under regulations.

NOM is the major precursor for DBPs, although its molecular composition and reactivity are not yet well understood (Wang et al., 2017). NOM treatability and reactivity toward DBPs formation varies significantly because each water source has unique features, with seasonal influences dependant on catchment characteristics and climate/weather-related effects, or whenever bromides occur. Surrogate parameters, such as dissolved organic carbon (DOC), ultraviolet absorbence at 254 nm (UV_{254}), and specific ultraviolet absorbence

* Corresponding author.
E-mail address: maria.martin@udg.edu (M.J. Martin).

https://doi.org/10.1016/j.watres.2022.119314
Received 25 June 2022; Received in revised form 4 October 2022; Accepted 29 October 2022
Available online 31 October 2022

Abbreviations

AMW	Apparent molecular weight
ANOVA	Analysis of variance
ASI	Absorbance slope signal
BB	Building blocks
DAD	Diode-array detector
DBPs	Disinfection by-products
DOC	Dissolved organic carbon
DOM	Dissolved Organic Matter
DWTP	Drinking water treatment facilities
EDR	Electrodialysis reversal
EEM	Fluorescence
FP	Formation potentials
HPSEC	High performance size exclusion chromatography
IEX	Ion exchange
JT	Jar testing
LMW	Low molecular weight
MLR	Multiple linear regression
MW	Molecular weight
NOM	Natural organic matter
OCD	Organic carbon detector
PSS	Polystyrene sulfonate
ROI	Region of interest
SC	Standardised model coefficients
SEC	Size exclusion chromatography
SUVA	Specific ultraviolet absorbance
tTHMs	Total trihalomethanes
TOC	Total organic carbon
UV	Ultraviolet

(SUVA), are frequently used to quantify NOM reactivity toward DBPs formation, although some water utilities find these techniques ineffective year-round due to seasonal allochthonous inputs caused by climate and catchment changes (Ritson et al., 2014; Rosario-Ortiz, 2014; Wang et al., 2017).

Considering the implications of NOM in the production of safe drinking water, more extensive characterization methodologies are required to gain a better understanding of its responsiveness to the generation of DBPs linked to the inefficiency of treatment operations under changing circumstances (Health Canada, 2020; Rosario-Ortiz, 2015; Sillanpää, 2014).

Size exclusion chromatography (SEC) has gained attention as a separation technique for NOM weight fractionation and is useful for tracking NOM removal effectiveness (Brezinski and Gorczyca, 2019). UV spectra can estimate activation of NOM moieties (Ignatev and Tuhkanen, 2019). Single-wavelength bulk measurements at 254 nm have been a direct benchmark for conjugated double bonds and aromatic structures, which are closely related to DBPs formation. Limiting the information to a single wavelength and bulk measurements may miss relevant elements related to DBPs formation. Mechanistic approaches based on photocatalytic production probe the formation of products with a lower degree of aromaticity and molecular weight (MW), presenting a higher absorbence at shorter wavelengths (Liu et al., 2010). Following this multi-wavelength approach, several spectroscopic calculations (e.g., ratios, slopes, and indices) have been proposed as novel DBPs tracking indicators (Chen et al., 2020). A diode-array detector (DAD) may benefit data acquisition at multiple wavelengths, which is useful for further calculations. Overall, coupling spectroscopy with SEC could provide a usable method for exploring the reactivity of different AMW NOM fractions.

The main limitation of these spectroscopic approaches is their unique ability to detect chromophores, leaving behind other non-absorbing NOM constituents, such as biopolymers. Universal detectors should be an end-to-end solution; however, their universality does not provide insights into the chemical nature of NOM fractions.

The SEC combined with an organic carbon detector (OCD) overcomes these spectroscopic limitations by encompassing all organics in a sample. LC—OCD, a custom-made setup developed by Huber et al. (2011), is the most widespread approach, in which OCD and UV_{254} detectors are attached to provide adjunctive information, discriminating between aromatic and aliphatic characteristics, except for MW. This further qualitative information bridges the gap between spectroscopic and high-compound identification potential techniques (e.g., MS detection-based) and high-mass balancing potential techniques (e.g. bulk DOC) (Allpike et al., 2005; Kawasaki et al., 2011).

Despite widespread application of SEC (Brezinski and Gorczyca, 2019), the effect of NOM-specific molecular weight fractions on DBPs formation under standardised chlorination conditions—formation potentials (FPs)—has rarely been investigated. Korshin et al. (2009) defined the absorbence slope signal (ASI) to explore the correlation between AMW fractions and THMs yields. Awad et al. (2016) modelled the relationships between NOM, characterised by DOC, UV–vis, fluorescence (EEM), and apparent molecular weight (AMW), using HPSEC-UV and THMs-FP, and encompassed the effect of bromide.

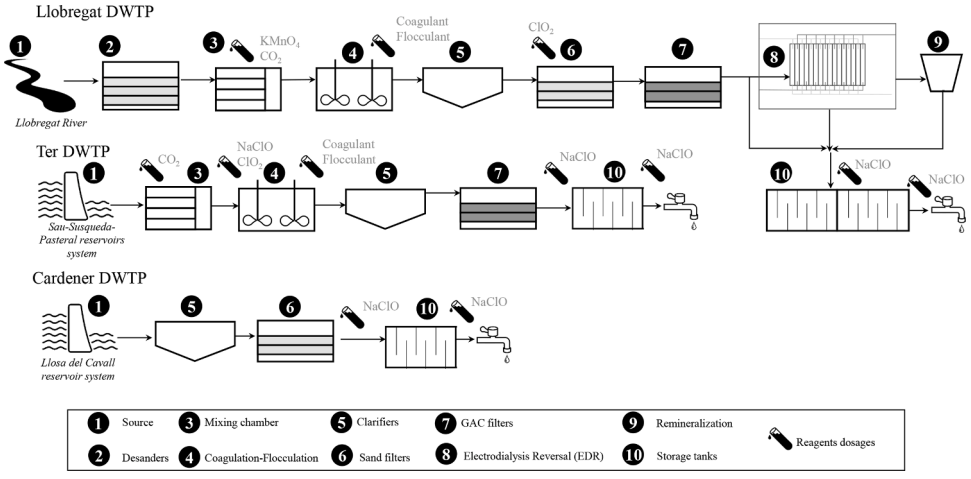

Fig. 1. Treatment process and points of reagent additions at the Llobregat, Ter, and Cardener drinking water treatment plants.

Recently, Carra et al. (2021) presented multiple linear regressions amongst AMW fractions (LC—OCD) and DBPs-FPs, revealing promising results regarding the capability of fractionation techniques for potential rapid risk assessment of water sources.

In this study, the addition of commercially available equipment to the HPSEC-DAD-OCD method merged their deep multiwavelength analysis and universal detection capabilities to evaluate the possible correlations between these signals and THMs-FP. The HPSEC-DAD-OCD configuration was used to clarify the effect of NOM SEC fractions on the regulated DBPs-FP model for three drinking water treatment plants operated by Ens d'Abastament d'Aigua Ter-Llobregat (ATL) serving 4.5 million inhabitants. For these facilities, THMs are the primary concern because HAAs are systematically below the regulated limits (Postigo et al., 2018).

Possible correlations between surrogate parameters and molecular weight fractions of DOM obtained by this novel HPSEC-DAD-OCD approach and THMs-FP were unveiled, considering four regulated species (i.e. $CHCl_3$, $CHBrCl_2$, $CHBr_2Cl$, and $CHBr_3$) and their sum.

Correlation-based predictions from these metrics are thought to be relevant and effective tools for DWTPs to predict THMs-FP and tailor treatments.

2. Materials and methods

2.1. Drinking water samples: origin and characterization parameters

The samples were collected through the treatment processes of three DWTPs in Catalonia: Llobregat (PTLL), Ter (PTT), and Cardener (PTC) (Fig. 1). These DWTPs supply water to the Barcelona metropolitan area, serving over 4.5 million inhabitants, and are managed by ATL (Valero and Arbós, 2010). By direct diversion, PTLL catches raw water from the Llobregat River and has a treatment capacity of 3.2 m^3/s; owing to salt deposit mining activities in the upper part of the river basin, high salinity issues are prevalent, which requires an electrodialysis reversal (EDR) desalination unit with a total capacity of 2.2 m^3/s in the treatment train (Postigo et al., 2018; Valero and Arbós, 2010), which operated at 64% during the sampling.

PTT receives water from a system of water reservoirs connected in series (Sau-Susqueda-Pasteral) along the Ter River through a 56-km pipeline, with a maximum treatment capacity of 8 m^3/s (Postigo et al., 2018). The fluctuating riverine regimes in both reservoirs make water characterization an ever-changing task. The PTC has a treatment capacity of 0.35 m^3/s of water from the Llosa del Cavall reservoir.

Ion exchange (IEX) efficiency was evaluated in PTLL and PTT using samples taken before and after flocculation (i.e., unit operation 5 in Fig. 1). The IEX resin LEWATIT® S5128 was conditioned with repeated MilliQ washings until the A_{254} was less than 0.2 abs/m. For each litre of the sample, 0.40 mL of resin was added, and the mixture was agitated at 120 rpm for 30 min. Following IEX testing, catchment samples were subjected to coagulation-flocculation jar testing (JT) with PAX-XL10 (25 mg/L) and PoliDADMAC (0.8 mg/L) as coagulant and flocculant, respectively, in PTLL samples, and PAX-XL10 (22 mg/L) and starch (0.3 mg/L), respectively, in PTT samples.

Two more samples completed the PTLL set, which were taken from the running EDR pilot plants testing different membrane configurations.

The samples were collected in glass containers and stored at 4 °C to avoid exposure to light. Rutinary DWTPs measurements of total organic carbon (TOC), conductivity, turbidity, pH, and ion chromatography concentrations were performed for each sample according to standard protocols (Supporting Information Tables SI 1–3). The concentrations of trichloromethane ($CHCl_3$), bromodichloromethane ($CHBrCl_2$), dibromochloromethane ($CHBr_2Cl$), tribromomethane ($CHBr_3$), and total THMs, determined as the sum of the four species (tTHMs), following the formation potential tests (THMs-FP) were analysed according to standard procedures (Standard Methods 5710 A) by HS-GC-ECD, as detailed in the Supporting Information.

2.2. DOM fractionation by size exclusion chromatography

To enable inter-procedure comparisons, samples were analysed using two SEC fractionation methods. For LC—OCD, the samples were shipped to the HET Waterlaboratorium (Haarlem, The Netherlands), where they were analysed following the method described by Huber et al. (2011). The chlorinated samples were quenched with sodium thiosulfate before shipment.

2.2.1. DOM fractionation by HPSEC-DAD-OCD

An HPSEC-DAD-OCD system was established to study DOM fractionation using an Agilent 1260 Infinity II HPLC with a 1260 vial sampler coupled to a DAD, enabling UV detection during the whole sweep from 190 to 600 nm, from which 206, 240, 254, 350, and 380 nm were selected for further analysis. The spectroscopic slopes between 206 and 240 nm ($S_{206–240}$) and 350–380 nm ($S_{350–380}$) were calculated as follows:

$$S_{\lambda_1-\lambda_2} = \frac{(A_{\lambda 1} - A_{\lambda 2})}{|\lambda_1 - \lambda_2|} \quad (1)$$

The hyphenated Sievers M9 SEC OCD was operated in turbo mode, acquiring data every 4 s with 5 μL s^{-1} acid (6 M H_3PO_4) and 2 μL s^{-1} oxidiser (15% $(NH_4)_2S_2O_8$) dosages, and was calibrated with potassium hydrogen phthalate.

A 6.8 pH phosphate buffer with 0.1 M of ionic strength (0.2880 g/L NaH_2PO_4 + 0.2864 g/L $Na_2HPO_4 \cdot 2\,H_2O$ + 3.5 g/L Na_2SO_4, from Sigma Aldrich) was pumped in the mobile phase at a flow rate of 0.75 mL/min (0–45 min), and 1 mL/min (45–60 min). MW separation was conducted using PL Aquagel-OH (7.5 × 300, 5 μm) × PL Aquagel-OH MIXED-M (7.5 × 300 mm, 8 μm) analytical columns from Agilent after 400 μL sample injections. The combination of columns covered a MW cut-off from 100 Da to 500 kDa and provided a baseline of approximately 60 ppb DOC. The autosampler temperature was 4 °C and the column oven temperature was 25 °C.

The samples were filtered through 0.20 μm nylon membrane filters, with no ionic strength adjustments made because additional buffer would be exchanged throughout the separation, eliminating high salt concentrations that could increase hydrophobic interactions. Triplicates of each sample were injected with a variation coefficient of 3.5%.

System void volume, determined with polystyrene sulfonate (PSS) 77 KDa (effective MW of 80,100 Da, PSS Polymer Standards Service GmbH), was approximately 12.35 mL (elution time 16.60 min), and the permeation volume determined with acetone was 27.27 mL (elution time 36.36 min). The size exclusion columns were calibrated with PSS standards, individual standards of humic and fulvic acid standards of the Suwannee River (3S101H and 3S101F, International Humic Substance Society), amino acids and organic acids covering MW values from 204 Da to 77 kDa (see Table SI 4).

2.2.2. HPSEC-DAD-OCD peak deconvolution and fraction quantification

DOM fractions in the HPSEC-DAD-OCD chromatograms were quantified by combining direct mathematical integration and peak deconvolution (see Fig. SI 1). The amounts of biopolymers (MW > 77 KDa) and low molecular weight (LMW) substances (MW < 200 Da acids and neutrals) were calculated from the area underneath the HPSEC—OCD spectra using the trapezoidal rule. The region of interest (ROI) of integration was as follows: biopolymers (10.8–17 min), LMW acids (29–33 min), and LMW neutrals (33–58.5 min).

Between 16 and 29 min (10 kDa to 200 Da), the DOC signal was deconvoluted to differentiate peaks within the fractions of humic substances (HS I at 19 min and HS II at 20 min; approximately 10 kDa to 700 Da) and building blocks (BB I 25.3 min and BB II 27.7 min; 700 to 200 Da) (Fig. SI 1). Deconvolution of the signals was performed by finding the peak through first derivatives, fitting Gaussian peaks with a chi-square tolerance value of 10^{-15} and R^2_{adj} of greater than 0.95. Retention times were fixed to reproduce the same deconvolution

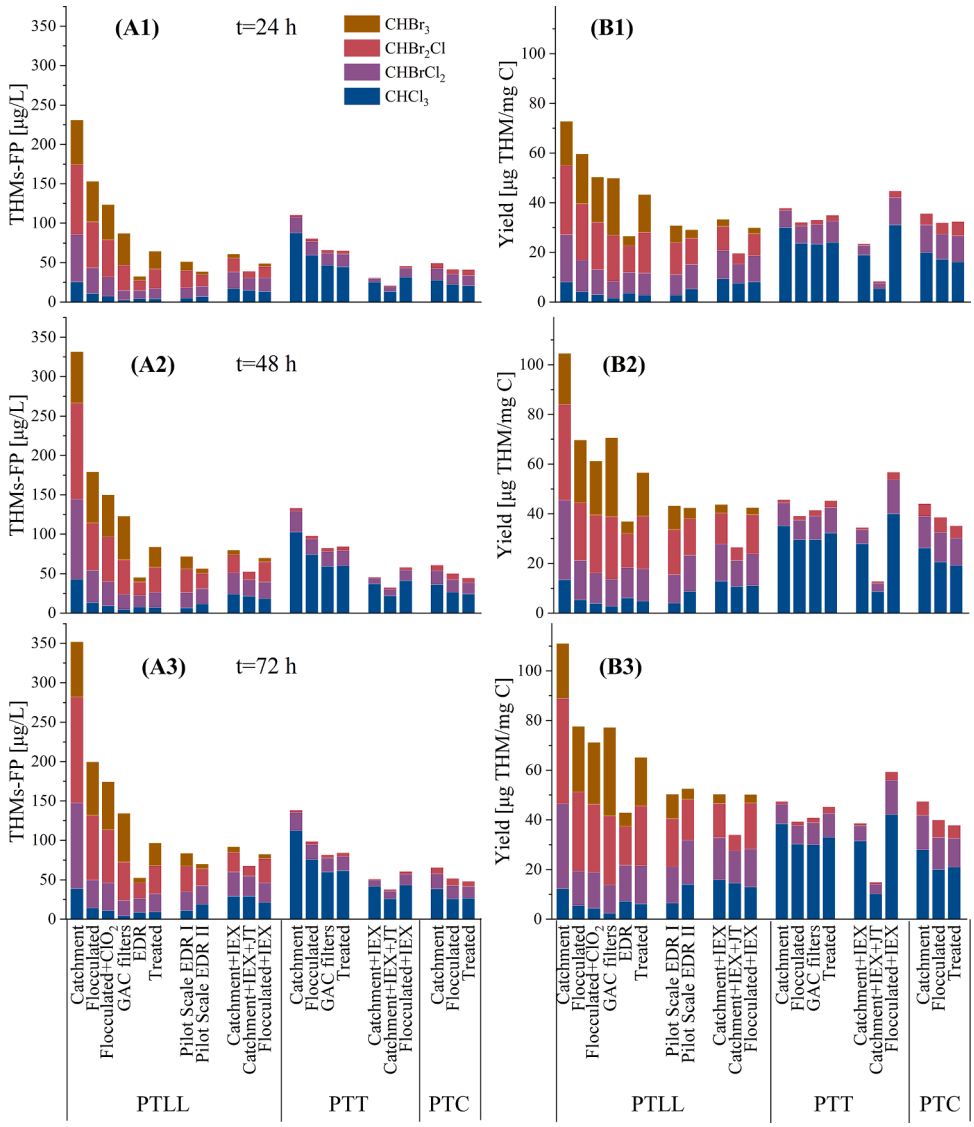

Fig. 2. THMs-FP (A1–3) and yields (B1–3) of analysed samples from Ter DWTP (PTT), Cardener DWTP (PTC), and Llobregat DWTP (PTLL) as a function of reaction time.

procedure for all samples, with a maximum tolerance of ± 0.5%.

Spectroscopic signal fractions (A_{254} and spectroscopic slopes $S_{350–380}$ and $S_{206–240}$) were computed by deconvoluting the spectra (from 10.8 to 55 min) using the same criteria previously stated. Five representative peaks were found for the A_{254}, $S_{350–380}$ signals (i.e., F1–F5), and six in the $S_{206–240}$ signals (i.e., F1–F6) (Table SI 5).

Numerical integration and signal deconvolution were performed using Origin (Pro) Version 2021b (OriginLab Corporation, Northampton, MA, USA).

2.3. THMs-FP modelling and statistical analysis

Possible correlations between DOM fractions, tTHMs-FP, and individual species were statistically assessed by evaluating multiple linear regression (MLR), as follows:

$$FP = \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n \quad (2)$$

where β_i is the coefficient corresponding to the different DOM fractions (x_i). To evaluate the models, the adjusted determination coefficient R^2_{adj} was calculated as follows:

$$R^2_{adj} = 1 - \left[\frac{(1 - R^2)(n - 1)}{n - k - 1}\right] \quad (3)$$

where R^2 represents the coefficient of determination, n is the number of sample points, and k is the number of variables used in the model. Analysis of variance (ANOVA) test probability was also analysed, with $F < 0.05$ being statistically significant.

Furthermore, the models were simplified using a stepwise multiple regression algorithm and independent variables (i.e., DOM fractions) that did not contribute significantly were eliminated. To allow for a better understanding of the contribution of the DOM fractions to THMs-FP, the β coefficients of the simplified linear models were standardised (SC), as follows:

$$SC_i = \beta_i \frac{\text{Stnd. deviation of } (x_i)}{\text{Stnd. deviation of THMs} - \text{FP}} \quad (4)$$

To ensure the validity of the predicted models, both the normality and homoscedasticity assumptions of the residuals were tested using the Shapiro–Wilk and Brausch–Pagan tests. MLR adjustments, model simplifications, and statistical tests were performed using R (version 4.1.1; 10/8/2021).

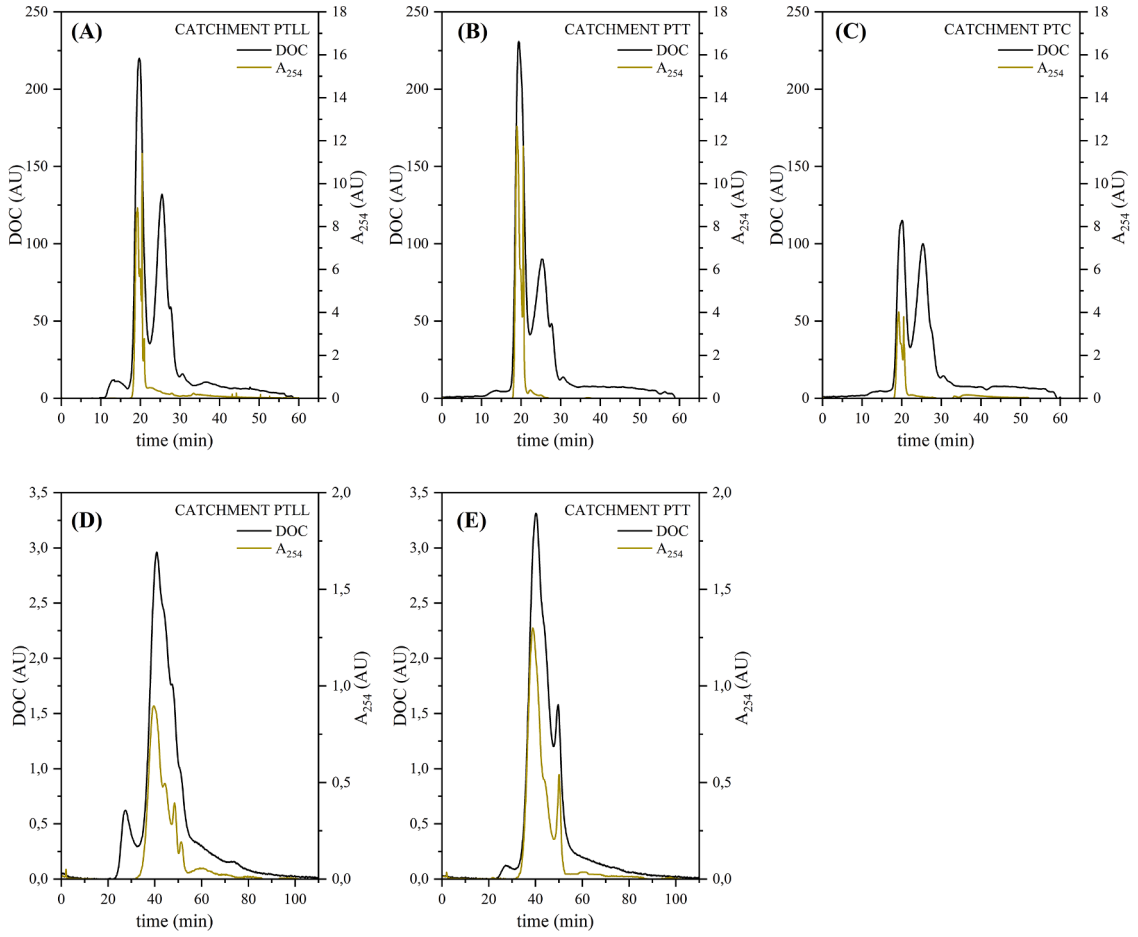

Fig. 3. Comparison of overlapped chromatograms of the DOC and A_{254} spectroscopic responses of the catchment samples obtained by HPSEC-DAD-OCD (A, B, C) and LC—OCD (D, E). Samples from the Cardener DWTP (PTC) were not analysed by LC—OCD. The quantifications are detailed in Tables SI 5 and SI 6.

3. Results and discussion

3.1. THMs precursors removal trends

The formation potential of the different THMs congeners during the determination of total THMs-FP for the 21 analysed samples was summarised (Fig. 2 and Table SI 5) to find distinctions amongst source water characteristics.

PTLL samples, presenting higher concentrations of DOC (3.17 mg/L) and salinity (554 μg/L of bromide and 189.3 mg/L chloride) exhibited higher THMs-FP values than the other source waters. While high-chlorinated THMs were prevalent and high-brominated compounds were not noticeable in PTT and PTC samples, $CHCl_3$ was the minority compound in PTLL, where highly brominated DBPs predominate ($CHBr_2Cl > CHBrCl_2 > CHBr_3 > CHCl_3$).

The FPs increased with the reaction time. Major increases were observed, particularly in the PTLL catchment sample, where tTHMs-FP increased approximately 1.5 times over 48 h, reaching the highest concentration after 72 h. Furthermore, the increases in tTHMs-FP for PTT and PTC were much more gradual.

As expected, a decrease in the THMs-FP was observed concurrently with the decrease in DOC concentration, from either the concatenation of unit operations within the treatment process or by IEX at the bench scale. Overall, samples subjected to IEX had higher DOC removal and Br^- reduction, resulting in lower THMs-FP (Table SI 3).

Furthermore, EDR led to a notable decline in tTHMs-FP in PTLL, with final levels approaching those of IEX-treated samples (32 μg/L compared to 39 μg/L and 49 μg/L when coupled with flocculation). Although the full-scale EDR performed the best, the pilot-scale EDRs also yielded substantial reductions, which were comparable to the IEX tests.

Yield representations of the different THMs-FP indicate the sample reactivities. Amongst conventional working operations, flocculation reduced the reactivity most. The reactivity of PTT and PTC was generally unaffected by the rest of the operational treatments; however, in PTLL, EDR also produced notable reductions in TOC (62%) and bromide (82%) (Table SI 2 and Table SI 3). The combined IEX and coagulation process reduced the reactivity of NOM by approximately 30% towards the formation of tTHMs (30.6% for PTLL and 31.3% in PTT). These final values are comparable to previously reported results in reservoir waters of 18 μg/mg DOC (Finkbeiner et al., 2018).

3.2. Characterization of DOM fractions by deconvoluted HPSEC-OCD signals

Compared to the LC—OCD method, the proposed method changed the chromatographic separation. shows that a single semi-preparative SEC column in LC—OCD does not provide sufficient resolution for the DOM fractions in the raw and treated waters to separate successfully. In LC—OCD separation, the first excluded fraction corresponds to the higher apparent molecular weight (i.e. biopolymers) which is clearly resolved from the rest of the primary DOC fractions (i.e. humics, building blocks, and LMW neutrals and acids). Peaks from this unresolved region are identified according to their physicochemical behavior and apparent molecular weight (in contrast to standard injections), and then the DOC signals are mathematically deconvoluted and quantified (Huber et al., 2011).

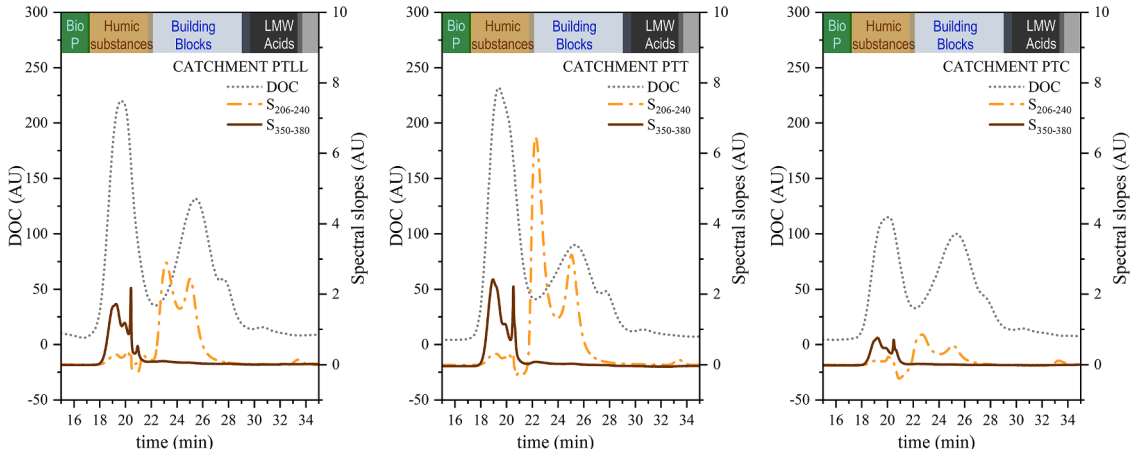

Fig. 4. Chromatograms of the HPSEC-DAD-OCD spectroscopic slopes between absorbence wavelengths of 206–240 nm and 350–380 nm, and overlapped DOC profiles for the river catchment sample (PTLL) and reservoirs (PTT and PTC).

Although the different analytical techniques provided the same elution sequence of standard compounds (Table SI 4) (Huber et al., 2011), that is, separation is ruled by the same mechanisms, the resolution of humic substances and building blocks was enhanced in HPSEC-DAD-OCD.

These separation differences were observed in the DOC signal profiles (Fig. 3) for the catchment samples (the full set of profiles is shown in Figs. SI 3 and 4 for HPSEC-DAD-OCD separation and Figs. SI 5 and 6 for LC—OCD). While biopolymers can be easily identified in both separations, humics, and building blocks were recognized in the HPSEC-DAD-OCD chromatograms. This new identification approach allows for a more accurate profile interpretation and comparison, as evidenced by the different DWTP catchment samples, where changes in these primary components were visible.

Although an in-depth analysis of the removal of the different DOM fractions within the treatment process is beyond the scope of this study, further insight can be uncovered into the efficiency of each unit operation. Conventional operation treatments, such as flocculation, mostly affected higher MW fractions (humics and biopolymers), reducing a 29% and 17% their reactivity towards tTHMs-FP in PTLL and PTT, respectively (see Section 3.1). However, EDR accounted for a 54% of tTHMs-FP decrease due to the reduction of humics (3%) and bromide (82%) whereas IEX eased the removal of lower-MW compounds such as building blocks and anions (Andersson et al., 2020).

The main issue is whether there are any possible correlations between the molecular weight fractions of DOM obtained by this novel HPSEC-DAD-OCD approach (both in terms of DOC and spectroscopic measurements) and THMs-FP of the four regulated species (i.e., CHCl3, CHBrCl2, CHBr2Cl, and CHBr3) and their sum (see Section 3.5).

3.3. *Characterization of DOM fractions by deconvoluted HPSEC spectroscopic signals*

Following a review of the literature (Rodríguez et al., 2016; Yan et al., 2012) and the evaluation of single-wavelength spectra, absorbence profiles at 254 nm (A_{254}) and spectral slopes at short ($S_{206–240}$) and long ($S_{350–380}$) wavelengths were further considered to test their THMs-FP modelling capabilities.

Whereas $S_{206–240}$ presumably provides information about the non-aromatic fractions, $S_{350–380}$ could correlate with the aromatic components of DOM, allowing both slopes to cover a wide range of spectroscopic characteristics. The lower range wavelengths (i.e. those less than 250 nm and associated with the presence of aromatic carbon) provide information about non-aromatic compositions associated with DOM, including the contributions of fulvic acids and microbial biopolymers (Huang et al., 2016; Ignatev and Tuhkanen, 2019; Korshin et al., 2009). According to Chen et al. (2020), relatively long wavelengths (i.e., those greater than 300 nm) correlate strongly with the yields of some DBPs species due to long conjugated chains reflecting the engagement of slow chromophores.

Furthermore, differences in separation methods can be observed when evaluating chromatographic A_{254} profiles. For samples analysed by HPSEC-DAD-OCD, most chromophores absorbed at 254 nm, commonly attributed to aromatic structures in humic substances, and coeluted at 18–22 min. In PTLL, PTT, and PTC, this DOM fraction represented 85, 92, and 88% of the total area of A_{254}, respectively. Compared to LC—OCD separation, the A_{254} signal was present in the entire region of the DOC signals, excluding biopolymers. The $S_{350–380}$ signal, according to the HPSEC—OCD fraction assignment, encompassed a range more similar to that of A_{254}, matching the chromophore zone of humic substances (HS I and HS II, 18–23 min).

Differences in absorbence profiles were found between the PTLL and PTT-PTC samples, particularly in both A_{254} and $S_{350–380}$ at approximately 20.93 min, where a peak was identified in the PTLL samples but was undetectable in the PTT-PTC samples (Fig. SI 2).

The lower wavelengths showed a distinctive absorption profile, with spectra spanning a wider time range (Fig. 4). The $S_{206–240}$ signal was weakest in the zone of aromatic chromophores of humic substances and was strongest in the range correlated with medium and low molecular weight fractions, including the final tail of HS II ($F4_{S206–240}$ at 23 min), building blocks ($F5_{S206–240}$ at 24.5 min), and LMW compounds (F6 at 33.5 min).

Inorganic compounds may affect spectroscopic signals at low wavelengths; therefore, they should be examined. In the absence of secondary interactions between the solute and stationary phases, SEC retention is governed by the molecular size; however, anions eluted within 23.2 min and 25.0 min (Fig. SI 7) in HPSEC-DAD-OCD, which corresponds to previous findings for nitrate (Szabo and Tuhkanen, 2010). The early elution of a divalent anion is also evident in the LC—OCD A_{254} profiles (see SI Figs. 5 and 6 for the anomalous A_{254} spectra of sodium thiosulfate quenched samples). Thus, SEC elution of anions under suitable conditions for NOM assessment (pH around neutral and low ionic strength) occurred sooner than expected because of electrostatic repulsion forces. Although these secondary ionic interactions may be undesirable in the SEC, simultaneous information on organic and inorganic contents may be beneficial as an indirect measurement for assessing THMs precursors.

A cross-check of the principal inorganic species in the analysed waters (Table SI 3) and their spectroscopic properties between 206 and 240 nm wavelengths (Birkmann et al., 2018) narrows the possible $S_{206–240}$ profile interferences in the DOM assessment to nitrate and bromide. According to the retention times of the bromide and nitrate standards

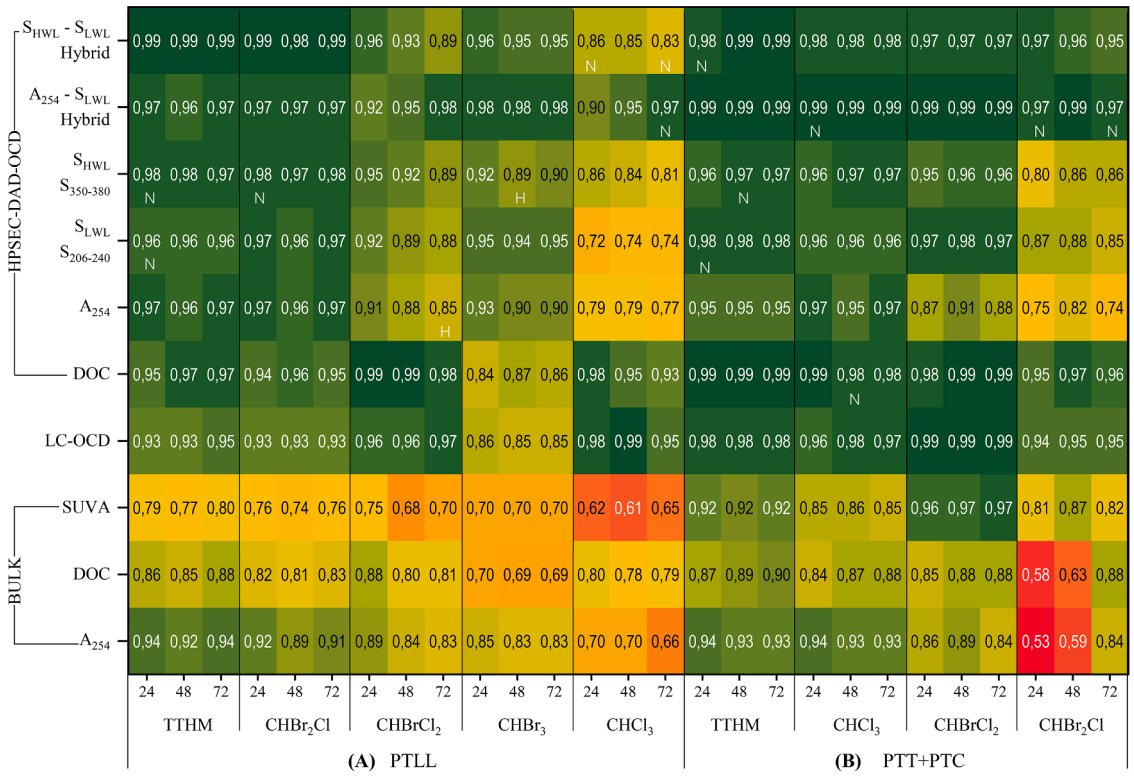

Fig. 5. R^2_{adj} coefficients heatmap of the MLR simplified models fitting bulk measurements and DOM fractions obtained by LC—OCD and HPSEC-DAD-OCD. Results are shown for Llobregat (PTLL, A) and Ter–Cardener (PTT+PTC, B) DWTPs. Species are depicted on the x-axis in order of abundance. The tags H (homoscedasticity) and N (normality) refer to the failure of statistical assumptions. LC—OCD models did not include samples from PTC.

(Fig. SI 7), the fractions $F4_{S206–240}$ and $F5_{S206–240}$ were influenced by their effects; however, the benefits and drawbacks of this simultaneous information on the organic content and anions (i.e. bromide) to model THMs-FP must be further investigated.

3.4. THMs-FP modelling with DOM fractions

Analysis of variance (ANOVA) and MLR were used to explore the statistical significance and possible relationship between THMs-FP, including individual congeners, bulk surrogated parameters, and different DOM fractionation strategies, such as LC—OCD and HPSEC-DAD-OCD.

As a first attempt, the feasibility of modelling the full set of data from the three DWTPs was examined; however, the simplified models obtained from the SEC signals did not satisfy homoscedastic assumptions (Table SI 7), likely because an explanatory variable accounting for unequal variance was omitted. Considering the raw water characteristics (e.g., DOC and bromide levels, see Table SI 3) and prevalent THMs, a split of samples was proposed: samples from the PTLL, where brominated THMs species dominated, were analysed independently of the PTT and PTC samples, where chlorinated THMs were prominent.

Besides the first attempt, 270 tests were conducted, which involved bulk measurements for DOC, A_{254}, and SUVA and fractionated DOC and spectroscopic signals. The SEC models were simplified, statistical assumptions were tested, and the fitting quality was evaluated using adjusted determination coefficients (Fig. 5). R^2_{adj} penalizes independent variables that do not contribute to the dependant variable, preventing over-parameterization and resulting in high R^2 from unnecessarily well-fitted data. When comparing the variations in the coefficients (for R^2_{adj} vs. R^2 in Figs. SI7 and SI8), increasing the number of predictors can be rejected as being responsible for fit quality. The models that did not pass the statistical test of residuals (i.e., normality and homoscedasticity) were scattered (Fig. 5). Most failures were due to residual non-normality, which could be overcome by increasing the number of samples in the training set ($n = 11$ for PTLL and $n = 10$ for PTT+PTC) or by adjusting a non-linear model.

Although the influence of FP kinetics was regarded using three reaction times (i.e., 24, 48, and 72 h), minimal differences were observed.

Generally, bulk A_{254} presented closer ties to FPs than other bulk parameters, whereas SUVA had the weakest associations. Amongst all, the best adjustments for tTHMs-FP were found regardless of speciation, which reflects the contribution of the most dominant THMs in each DWTP ($CHBr_2Cl$ in PTLL and $CHCl_3$ in PTT+PTC). The R^2_{adj} values for the bulk parameters were comparable to those documented (Table SI9).

R^2_{adj} improved when modelling FPs with DOC DOM fractions. A similar behavior was recently discussed by Carra et al. (2021) when assessing an MLR analysis for tTHMs and MW fractions, reporting better correlations with fractionated DOM following LC—OCD separation compared to bulk DOC. In this study, HPSEC-DAD-OCD showed enhanced regression compared to LC—OCD correlations.

Breakdown analyses of distinct THMs-FPs were conducted, showing the distinctive contribution of each species. Overall, the tTHMs models provided better adjustments than those of individual species, despite the SEC approach accounting for a more refined speciation compared to bulk predictions. Spectroscopic models exhibited closer linkages than DOC fractionated models, except for the less prevalent species.

Consequently, fractionated A_{254} signal improved tTHMs and minority species correlations compared to bulk A_{254}. Furthermore, the correlations between $S_{206–240}$ and $S_{350–380}$ presented a noticeable improvement compared to the bulk A_{254} predictions, slightly enhancing correlations for individual THMs. The three signals account for different information: although A_{254} and $S_{350–380}$ overlap in retention times and are both related to aromatic moieties, higher wavelengths (i.e., 350–380 nm) could reflect the engagement of slow chromophores involved in the formation of tautomeric ketones that are adducts implicated in DBPs generation (Chen et al., 2020). Therefore, differences

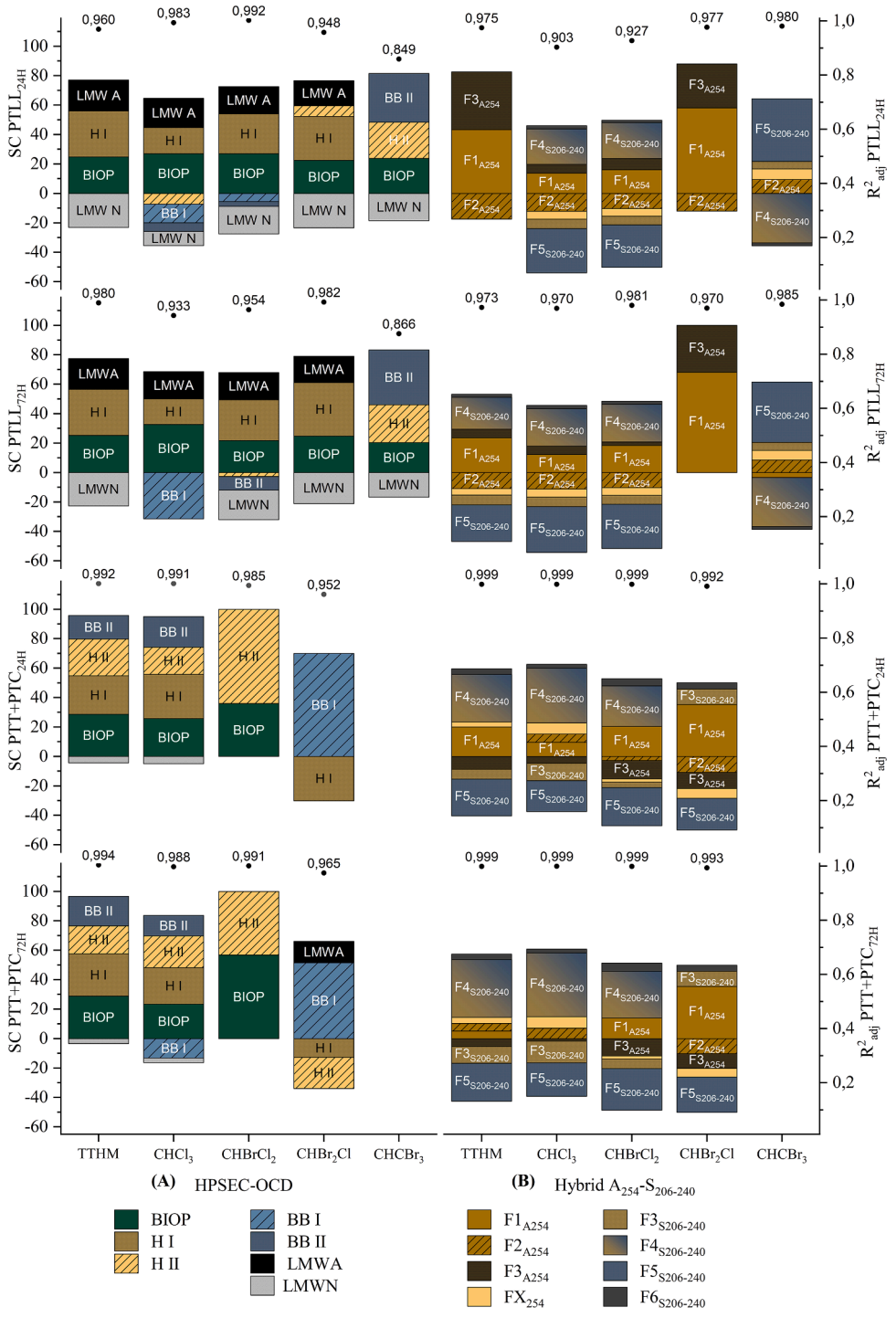

Fig. 6. Standardised coefficients (SC) and R^2_{adj} values as a function of reaction time (24 and 72 h) for modelling the total (tTHMs) and individual THMs-FPs from DOC fractionation (A) and spectroscopic hybrid signals A_{254}–$S_{206-240}$ (B). Note that FX_{254} refers to $F4_{254}$ for PTLL samples and F5 A_{254} for PTT-PTC (see Section 3.5).

in correlations could be anticipated because of the different mechanistic implications that the two signals would reflect. In contrast, $S_{206-240}$ encompasses both organic and inorganic contributions, as discussed in Section 3.3.

Two hybrid models were proposed by combining the major intensity peaks of different signals to provide a more sensitive prediction. The first hybrid model that combined the main absorbence signals in A_{254} and $S_{206-240}$ (A_{254}–$S_{206-240}$) included $F1_{A254}$ to $F4_{A254}$ from the A_{254} chromatograms and $F3_{S206-240}$ to $F6_{S06-240}$ from the $S_{206-240}$ profiles (Table SI5). The first A_{254} peaks correlated with the main humic substances and building blocks (10 kDa to 700 Da) chromophores, whereas the $S_{206-240}$ peaks primarily covered the chromophores of the LMW range (< 200 Da). The second hybrid model was performed by replacing A_{254} with $S_{350-380}$ ($S_{350-380}$–$S_{206-240}$), considering fractions $F1–F3_{S350-380}$, $F5_{S350-380}$, and $F3–F6_{S206-240}$. While $S_{206-240}$ in that range (21–33.5 min) represented non-aromatic and inorganic compounds, neither $S_{350-380}$ nor A_{254} included non-organic contributions as input.

Although the slope spectroscopic models presented better correlation coefficients than A_{254}, the hybrid models explained all the outlooks best.

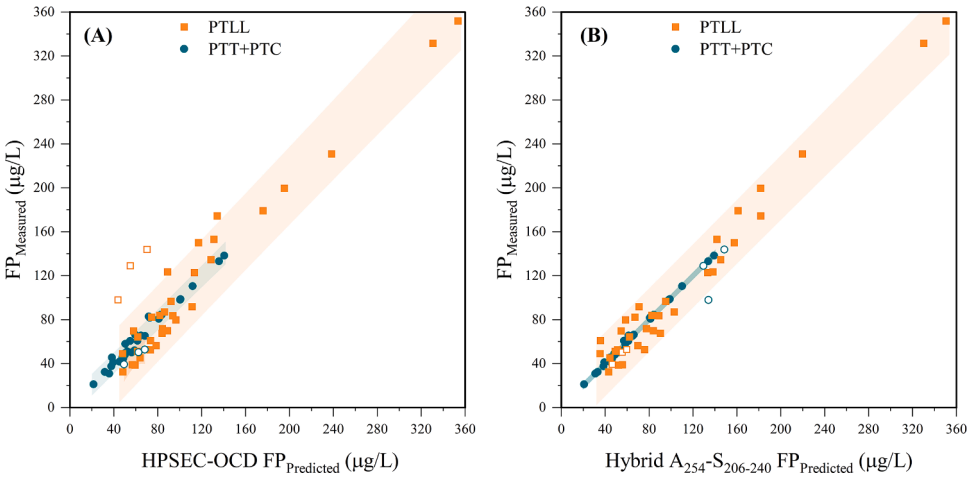

Fig. 7. Measured vs. predicted tTHMs-FP estimated with the HPSEC—OCD model (A) and the spectroscopic hybrid model A_{254}–$S_{206–240}$ (B) including the representation of 95% prediction bands and validation samples (void symbols).

The A_{254} -$S_{206–240}$ hybrid model for tTHMs and the individuals presented the strongest coefficients ($R^2_{adj} > 0.90$), especially in PTT and PTC. The $S_{350–380}$-$S_{206–240}$ hybrid model also presented high R^2_{adj} coefficients; however, adjustments for PTT and PTC slightly decreased comparatively.

3.5. *Analysis of DOM fractions contributing to DOC and spectroscopic THMs-FP models*

The analysis of standardised model coefficients (SC) for the best-fitted models from DOC and spectroscopic signals can reveal the role of primary DOM fractions in the FP of tTHMs and THMs species. This A_{254}–$S_{206–240}$ spectroscopic model was selected to proceed further due to their main improvements regarding speciation.

Fig. 6 presents the SCs of the simplified models obtained for the HPSEC—OCD and the A_{254}–$S_{206–240}$ models. Fig. 6 comprises a stacked bar chart representation, with higher contributions of a given DOM fraction being expressed by greater SC bar lengths. Although the information extracted from this modelling is far from a mechanistic approach, it can provide empirical clues about the most available reactive fractions that could enhance formations of the different THMs.

DOM compounds eluted within humics and biopolymers were the most significant precursors of halogenated THMs in drinking water samples. Spectroscopy revealed a strong contribution from the A_{254} fractions, especially for $CHBr_2Cl$, which further supports the role of aromatic moieties during THMs formation (Brezinski and Gorczyca, 2019). Although no significant differences were observed between 24 and 72 h in the hybrid spectroscopic method, a significant contribution of fractions in the aliphatic range ($F4_{S206–240}$–$F5_{S206–240}$) was observed in PTLL samples.

The contributions of DOM moieties to FPs appear to be halogen-specific, which is a factor linked to speciation. The dominant high-brominated THMs, $CHBr_2Cl$ for the PTT-PTC cluster and $CHBr_3$ for PTLL, presented a different trend compared to the other modelled THMs-FP. Furthermore, the coefficient significances flipped with other congeners, and the building block fractions in the DOC PTT-PTC models exhibited a negative contribution to tTHMs and $CHCl_3$, but not for $CHBrCl_2$, and are primarily responsible for explaining $CHBr_2Cl$ (Fig. 9). The spectroscopic model for $CHBr_3$ in PTLL samples followed a similar trend, with reverse contributions for F3, F4, and F5 from the $S_{206–240}$ signal. This observation supports the indirect contribution of the anions discussed in Section 3.3.

These patterns in high-brominated THMs match earlier reports (Brezinski and Gorczyca, 2019). Furthermore, the bromination degree may be influenced by factors such as the bromide/chloride ratio and possible electrophilic substitutions in electron-rich structures, such as aromatic rings, phenolic hydroxyls, or double conjugated bonds typical of high-MW hydrophobic fractions (Chowdhury et al., 2009; Ichihashi et al., 1999). In the first stage, disinfectants activated moieties, and progressive bromide substitutions may occur depending on the availability and reactivity of these moieties. Therefore, these phenomena could be explained by the behavior of these flipped fractions.

The negative contributions in the models could be explained by their specific role in enhancing the formation of particular THMs. For example, the BB fraction did not promote the formation of $CHCl_3$. These fractions appeared to be a crucial precursor in the formation of $CHBr_2Cl$ in low-brominated water samples from PTT-PTC. Generally, these negative contributions could also be related to the formation of intermediates converted into unmonitored DBPs (Carra et al., 2021).

3.6. *Validity of the THMs-FP models for future observations*

For each individual THMs and tTHMs-FP, the measured and predicted values were represented to evaluate the goodness of the predicted capacity of the suggested models (Figs. 7 and 8). Furthermore, to cover the possibility of future observation predictions using the proposed models, prediction bands were calculated assuming a probability of 95%.

The differences in bandwidth between DOC and A_{254}–$S_{206–240}$ suggest differences in the prediction accuracy of each model. In particular, the hybrid model, which presented narrower bands, appeared to be more accurate than the DOC model. Mathematically, this is biased in part by the limitations of the analytic process itself, the sensitivity and accuracy of the measurements and deconvolutions, and in part by the variability of the sample used to develop the models. A comparison of the bandwidths for the same model on both datasets revealed heterogeneity differences. This is reflected in the prediction models for DOC and A_{254}–$S_{206–240}$, where the adjustment bands for PTLL were broader, meaning the analysed samples presented higher variability between them, compared to PTT-PTC, which had narrower bands.

To further test these models and cross-validate our results, two new samples of storage tanks from PTLL and treated water from PTC (see Tables SI6 and SI7) were collected a year after the initial set test. The time span between sampling campaigns led to a change in hydrological conditions, which resulted in a redefinition of the scenarios. Although the mean values of the water quality parameters did not vary significantly (see Table SI11), during the validation set sampling, the end of the drought pre-warning scenario was experienced.

Under these conditions, the PTLL storage tanks were still partly supplied by an external desalination plant, mixed with in situ produced

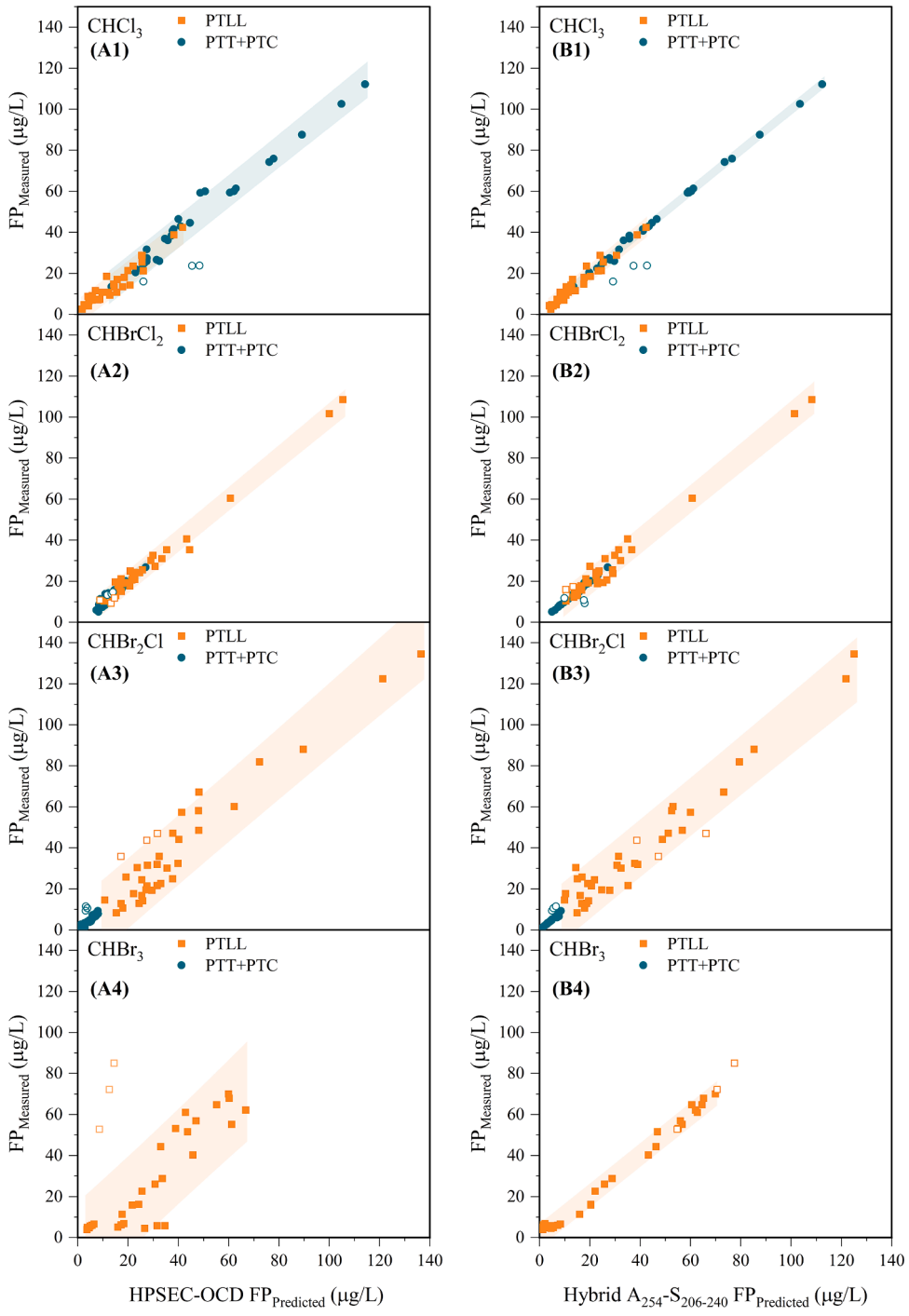

Fig. 8. Measured vs. predicted FP for the distinct THMs congeners estimated with the HPSEC—OCD model (A1–4) and the spectroscopic hybrid model A_{254}–$S_{206–240}$ (B1–4) that includes the representation of 95% prediction bands and validation samples (void symbols).

water, to meet distribution network demands. The desalination facility provides water when the PTLL supply is threatened. During the validation sampling campaign, none of the EDR modules was functional. Operational controls of the EDR modules are based on water quality parameters (e.g., temperature, amount of organic matter and salinity) and ruled by an environmental decision support system that assesses these parameters to determine which working percentage represents the optimum to minimize the formation of THMs (Godo-Pla et al., 2021). Changes in temperatures (e.g., lower temperatures disfavoring THMs formation) were decisive for switching off the EDR modules. In the initial training dataset, the DWTP operated with 64% of the flow diverted to the EDR; however, in the validation sampling, the EDR was shut down and the GAC filters the latest operation before the final disinfection (Fig. 1).

For PTC, this end-of-drought pre-warning translated into a reduction of half of the reservoir level compared to the sampling period (approximately from 65 hm^3 in 2021 to 35 hm^3 in 2022). Despite the low flow rates and simplicity of the treatment train, no major operational changes were observed in the DWTP. The same measurement procedures (both the characterization and THMs-FP tests) were performed as previously described.

The OCD fraction profile and quantification comparisons on the

testing samples with the training dataset showed differences between scenarios (Tables SI5 and SI6). Although the total amount of building blocks and LMW acids fluctuated minimally (< 10% of the variation), biopolymers, humics (I and II) and LMW neutrals increased by 19, 46, and 95%, respectively, in PTLL. This could be correlated with the simultaneous increase in $F1_{A254}$ and $F3_{A254}$–$F5_{A254}$, related to more aromaticity because of the higher content of humic substances. However, building blocks were reduced by 41% and LMW acids, the less variable fraction, decreased by 20% in PTC, whereas biopolymers, LMW neutrals (71%) and humics I increased (205%), again in agreement with the increase in A_{254} signals. Changes in PTLL bromide and nitrate levels, up to 2 and 3 times higher respectively (see Table SI3), were observed in the increasing areas of $F4_{S206–240}$ and $F5_{S206–240}$.

Although the tTHMs-FP of the validation samples were in the order of the training dataset (Table SI2), distinct THMs distributions were observed. For PTLL, chlorinated compounds were less prevalent than in the first scenario, with being $CHBr_3$ the dominant species, followed by $CHBr_2Cl$, $CHBrCl_2$, and non-quantifiable $CHCl_3$. However, in the validation sample for PTC, prevalence was maintained, as the more chlorinated THMs were the compounds most present ($CHCl_3$ > $CHBrCl_2$ > $CHBr_2Cl$). Regarding bromide levels, the PTLL storage tank samples had a concentration of 0.39 mg/L between the GAC filtered sample (0.56 mg/L) and the treated sample (0.18 mg/L) of the training data set. The levels of PTC-treated samples were still below the quantification limit.

After characterization, the THMs were predicted using the proposed methods (see Fig. 7, void symbols).

As discussed in Sections 3.4 and 3.5, the tTHMs-FP models presented better adjustments compared to speciation. However, as shown in Fig. 7 the DOC predictions for PTLL fell outside the prediction bands. Some confounding variables might affect the explicative variable of the models, although they are not included, as it may result from inorganics. A_{254}–$S_{206–240}$ indirectly considers the effects of those inorganics due to their absorbence at lower wavelengths (see Section 3.3), explaining the best predictions for the tTHMs adjustments, particularly in the case of PTLL, where high levels of bromide are a primary DWTP concern.

The bromide effect became clear when contrasting the predictions of $CHBr_3$ (Fig. 8, A4 and B4). DOC predictions did not meet the specified requirements, but A_{254}–$S_{206–240}$ did. However, this effect did not seem relevant for the species in the PTT-PTC. This suggests that, whereas both DOC and spectroscopic methods could be equally suitable for predicting tTHMs, DWTPs about bromide-related issues may selectively benefit from spectroscopic approaches that account for inorganic contributions.

As the models were created using a heterogeneous pool of samples, the predictions remained accurate despite scenario changes. Although large fluctuations in the sample population (e.g., seasonality or climatic episodes, such as droughts or floods) were not included in the original training dataset, modest variability impacts were observed owing to the initial diversity.

Possible methodological modifications, such as refining deconvolutions or adjusting non-linear models, could be explored to improve and increase predictability. In addition, the proposed models are based on FPs, which are determined under controlled conditions (e.g. time and temperature); therefore, they should be considered as approximations. Creating a historical data repository to increase the training data set and include seasonal fluctuations in the models would further strengthen the predictions.

Finally, predictive advantages can be extracted from the proposed models, especially spectroscopic models, but can also account for operational simplification. Predictions based on fractionated spectroscopic signals using HPLC-DAD seem more efficient and accessible than using HPLC coupled to an OCD for highly specialized applications.

4. Conclusions

An in-depth analysis of THMs-FP modelling was performed that uses HPSEC-DAD-OCD for DOM characterization was performed, including diverse typologies of raw and treated waters. Differences in water characteristics make it challenging to find the best pooling to predict THMs-FP. The differences in AMW, the intrinsic properties of DOM, and the inorganic content were crucial characteristics. Although the models did not contain a forthright predictor, such contributions were integrated in the proposed spectroscopic-based model, combining substantial intensities from A_{254} and $S_{206–240}$.

The AMW-fractionated signal, DOC, and spectroscopic models improved correlations compared to single bulk parameters, especially in HPSEC-DAD-OCD separation. The standard coefficients of the models elucidate the contribution and significance of the fractions most involved in THMs-FP. Although some representative fractions converged in the studied DWTPs, such as biopolymers and humic substances; other differences were noted, such as the role of LMW acids in the Llobregat water supply and of building blocks in the Ter and Cardener water supplies. Similar analogies were observed in the spectroscopic model. This relevance allows the creation of a link between precursors and products, opening the door to further targeted optimization of DOM removal.

Although a modest validation of models was performed, proving their prediction ability, a more robust model that includes largest data sets accounting for variable fluctuations and compositions is still needed. Further refined data treatment and, principally, involving deconvolutions could help simplify the predictors and account for other suitable models. However, this study represents a novel approach, including spectroscopic SEC separation, to provide a less tedious alternative to track THMs-FP compared to traditional OCD analysis.

Declaration of Competing 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.

Data Availability

Data will be made available on request.

Acknowledgements

This study was supported by the WATSproof (CTM2017-83598-R) and ShERLOcK projects (PID2020-112615RA-I00), financed by the Ministerio de Ciencia e Innovación (Spain). MVQ thanks AGAUR from the Generalitat de Catalunya for a predoctoral grant under the program FI_SDUR 2020-00330. LEQUIA has been recognized as a consolidated research group by the Catalan government (2017-SGR-1552). Open access funding was provided thanks to the CRUE-CSIC agreement with Elsevier.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.watres.2022.119314.

References

Allpike, B.P., Heitz, A., Joll, C.A., Kagi, R.I., Abbt-Braun, G., Frimmel, F.H., Brinkmann, T., Her, N., Amy, G., 2005. Size exclusion chromatography to characterize DOC removal in drinking water treatment. Environ. Sci. Technol. 39, 2334–2342. https://doi.org/10.1021/es0496468.

Andersson, A., Lavonen, E., Harir, M., Gonsior, M., Hertkorn, N., Schmitt-Kopplin, P., Kylin, H., Bastviken, D., 2020. Selective removal of natural organic matter during drinking water production changes the composition of disinfection by-products. Environ. Sci. 6, 779–794. https://doi.org/10.1039/c9ew00931k.

Awad, J., van Leeuwen, J., Chow, C., Drikas, M., Smernik, R.J., Chittleborough, D.J., Bestland, E., 2016. Characterization of dissolved organic matter for prediction of

trihalomethane formation potential in surface and sub-surface waters. J. Hazard. Mater. 308, 430–439. https://doi.org/10.1016/j.jhazmat.2016.01.030.
Birkmann, J., Pasel, C., Luckas, M., Bathen, D., 2018. UV spectroscopic properties of principal inorganic ionic species in natural waters. Water Pract. Technol. 13, 879–892. https://doi.org/10.2166/wpt.2018.097.
Brezinski, K., Gorczyca, B., 2019. Multi-spectral characterization of natural organic matter (NOM) from Manitoba surface waters using high performance size exclusion chromatography (HPSEC). Chemosphere 225, 53–64. https://doi.org/10.1016/j.chemosphere.2019.02.176.
Carra, I., Fernandez Lozano, J., Johannesen, S., Godart-Brown, M., Goslan, E.H., Jarvis, P., Judd, S., 2021. Sorptive removal of disinfection by-product precursors from UK lowland surface waters: impact of molecular weight and bromide. Environ. Sci. Technol. 754, 142152 https://doi.org/10.1016/j.scitotenv.2020.142152.
Chen, B., Zhang, C., Zhao, Y., Wang, D., Korshin, G.V., Ni, J., Yan, M., 2020. Interpreting main features of the differential absorbance spectra of chlorinated natural organic matter: comparison of the experimental and theoretical spectra of model compounds. Water Res. 185. https://doi.org/10.1016/j.watres.2020.116206.
Chowdhury, S., Champagne, P., McLellan, P.J., 2009. Models for predicting disinfection byproduct (DBP) formation in drinking waters: a chronological review. Environ. Sci. Technol. 407, 4189–4206. https://doi.org/10.1016/j.scitotenv.2009.04.006.
Finkbeiner, P., Redman, J., Patriarca, V., Moore, G., Jefferson, B., Jarvis, P., 2018. Understanding the potential for selective natural organic matter removal by ion exchange. Water Res. 146, 256–263. https://doi.org/10.1016/j.watres.2018.09.042.
Godo-Pla, L., Emiliano, P., Poch, M., Valero, F., Monclús, H., 2021. Benchmarking empirical models for THMs formation in drinking water systems: an application for decision support in Barcelona, Spain. Environ. Sci. Technol. 763 https://doi.org/10.1016/j.scitotenv.2020.144197.
Health Canada, 2020. Guidance on natural organic matter in drinking water. On line resource ISBN 9780660336282. https://www.canada.ca/en/health-canada/services/publications/healthy-living/guidance-natural-organic-matter-drinking-water.html.
Huang, H., Sawade, E., Cook, D., Chow, C.W.K., Drikas, M., Jin, B., 2016. High-performance size exclusion chromatography with a multi-wavelength absorbance detector study on dissolved organic matter characterization along a water distribution system. J. Environ. Sci. 44, 235–243. https://doi.org/10.1016/j.jes.2015.12.011.
Huber, S.A., Balz, A., Abert, M., Pronk, W., 2011. Characterization of aquatic humic and non-humic matter with size-exclusion chromatography - organic carbon detection - organic nitrogen detection (LC-OCD-OND). Water Res. 45, 879–885. https://doi.org/10.1016/j.watres.2010.09.023.
Ichihashi, K., Teranishi, K., Ichimura, A., 1999. Brominated trihalomethane formation in halogenation of humic acid in the coexistence of hypochlorite and hypobromite ions. Water Res. 33 (2), 477–483. https://doi.org/10.1016/S0043-1354(98)00227-9.
Ignatev, A., Tuhkanen, T., 2019. Step-by-step analysis of drinking water treatment trains using size-exclusion chromatography to fingerprint and track protein-like and humic/fulvic-like fractions of dissolved organic matter. Environ. Sci. 5, 1568–1581. https://doi.org/10.1039/c9ew00340a.
Jiang, J., Han, J., Zhang, X., 2020. Nonhalogenated aromatic DBPs in drinking water chlorination: a gap between NOM and halogenated aromatic DBPs. Environ. Sci. Technol. 54 (3), 1646–1656. https://doi.org/10.1021/acs.est.9b06403.
Kawasaki, N., Matsushige, K., Komatsu, K., Kohzu, A., Nara, F.W., Ogishi, F., Yahata, M., Mikami, H., Goto, T., Imai, A., 2011. Fast and precise method for HPLC-size exclusion chromatography with UV and TOC (NDIR) detection: importance of multiple detectors to evaluate the characteristics of dissolved organic matter. Water Res. 45, 6240–6248. https://doi.org/10.1016/j.watres.2011.09.021.
Korshin, G., Chow, C.W.K., Fabris, R., Drikas, M., 2009. Absorbance spectroscopy-based examination of effects of coagulation on the reactivity of fractions of natural organic matter with varying apparent molecular weights. Water Res. 43, 1541–1548. https://doi.org/10.1016/j.watres.2008.12.041.
Krasner, S., Weinberg, H., Richardson, S., Pastor, S., Chinn, R., Sclimenti, M., Onstad, G., Thruston, A., 2006. Occurrence of a new generation of disinfection by-products. Environ. Sci. Technol. 40 (23), 7175–7185.
Liu, S., Lim, M., Fabris, R., Chow, C.W.K., Drikas, M., Korshin, G., Amal, R., 2010. Multi-wavelength spectroscopic and chromatography study on the photocatalytic oxidation of natural organic matter. Water Res. 44, 2525–2532. https://doi.org/10.1016/j.watres.2010.01.036.
Plewa, M.J., Wagner, E.D., Richardson, S.D., 2017. TIC-Tox: a preliminary discussion on identifying the forcing agents of DBP-mediated toxicity of disinfected water. J. Environ. Sci. 58, 208–216. https://doi.org/10.1016/j.jes.2017.04.014.
Postigo, C., Emiliano, P., Barceló, D., Valero, F., 2018. Chemical characterization and relative toxicity assessment of disinfection byproduct mixtures in a large drinking water supply network. J. Hazard. Mater. 359, 166–173. https://doi.org/10.1016/j.jhazmat.2018.07.022.
Ritson, J.P., Graham, N.J.D., Templeton, M.R., Clark, J.M., Gough, R., Freeman, C., 2014. The impact of climate change on the treatability of dissolved organic matter (DOM) in upland water supplies: a UK perspective. Sci. Total Environ. 473–474, 714–730. https://doi.org/10.1016/j.scitotenv.2013.12.095.
Rodríguez, F.J., Schlenger, P., García-Valverde, M., 2016. Monitoring changes in the structure and properties of humic substances following ozonation using UV–vis, FTIR and 1H NMR techniques. Environ. Sci. Technol. 541, 623–637. https://doi.org/10.1016/j.scitotenv.2015.09.127.
Rosario-Ortiz, F., 2015. Advances in the Physicochemical Characterization of Dissolved Organic Matter: Impact on Natural and Engineered Systems. Oxford University Press.
Sillanpää, M., 2014. Natural Organic Matter in Water: Characterization and Treatment Methods. Butterworth-Heinemann. doi:10.1016/C2013-0-19213-6.
Szabo, H.M., Tuhkanen, T., 2010. The application of HPLC-SEC for the simultaneous characterization of NOM and nitrate in well waters. Chemosphere 80, 779–786. https://doi.org/10.1016/j.chemosphere.2010.05.007.
Valero, F., Arbós, R., 2010. Desalination of brackish river water using Electrodialysis Reversal (EDR). Control of the THMs formation in the Barcelona (NE Spain) area. Desalination 253, 170–174. https://doi.org/10.1016/j.desal.2009.11.011.
Wang, X., Zhang, H., Zhang, Y., Shi, Q., Wang, J., Yu, J., Yang, M., 2017. New insights into trihalomethane and haloacetic acid formation potentials: correlation with the molecular composition of natural organic matter in source water. Environ. Sci. Technol. 51 (4), 2015–2021. https://doi.org/10.1021/acs.est.6b04817.
Yan, M., Korshin, G., Wang, D., Cai, Z., 2012. Characterization of dissolved organic matter using high-performance liquid chromatography (HPLC)-size exclusion chromatography (SEC) with a multiple wavelength absorbance detector. Chemosphere 87, 879–885. https://doi.org/10.1016/j.chemosphere.2012.01.029.
Zhang, Z., Zhu, Q., Huang, C., Yang, M., Li, J., Chen, Y., Yang, B., Zhao, X., 2020. Comparative cytotoxicity of halogenated aromatic DBPs and implications of the corresponding developed QSAR model to toxicity mechanisms of those DBPs: binding interactions between aromatic DBPs and catalase play an important role. Water Res. 170 https://doi.org/10.1016/j.watres.2019.115283.