



Seasonal variability of dissolved organic matter in the Columbia River: In situ sensors elucidate biogeochemical and molecular analyses

Urban Johannes Wünsch^{1,2}, Boris Peter Koch^{2,1}, Matthias Witt³, Joseph Andrew Needoba⁴

5 ¹ University of Applied Sciences Bremerhaven, An der Karlstadt 8, D-27568 Bremerhaven, Germany

² Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Am Handelshafen 12, D-27570 Bremerhaven, Germany

³ Bruker Daltonik GmbH, Fahrenheitstraße 4, D-28359 Bremen, Germany

10 ⁴ Oregon Health and Science University Institute of Environmental Health, 3181 Southwest Sam Jackson Park Road, Portland, Oregon 97239-3098, United States of America

Correspondence to: Urban J. Wünsch (urbw@aqu.dtu.dk)

Abstract. The in situ detection of fluorescent dissolved organic matter (FDOM) at high temporal resolution is a powerful proxy to follow dissolved organic matter (DOM) dynamics and DOM flux to coastal oceans when FDOM measurements and dissolved organic carbon (DOC) are highly correlated. Here, we investigated the relationship between FDOM sensors and
15 DOC concentration in the lower Columbia River, USA in spring and summer 2013. Furthermore, we studied the seasonal variability of FDOM and chromophoric DOM (CDOM) optical indices, as well as the seasonal and spatial variability for the molecular characteristics of DOM using ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). The fieldwork was conducted concurrently with the operation of in situ sensor platforms that recorded physical and biogeochemical data at hourly intervals. In situ FDOM and DOC concentration was
20 highly correlated and the relationship was used to quantify the river's DOC flux between March and August 2013. The average flux was 0.93 Gg d⁻¹, which included over three-fold temporal variability (0.45 to 1.62 Gg d⁻¹) associated with seasonal biogeochemical variability. Spectrofluorometry measurements demonstrated that FDOM parameters correlated with major seasonal biogeochemical shifts in the river associated with phytoplankton blooms and river discharge and thus revealed predictable seasonal patterns in DOM quality. FT-ICR-MS analyses elucidated these shifts on the molecular level:
25 the relative abundance of 561 formulas, most of which contained N and S, correlated significantly with chlorophyll a, while 417 formulas (mostly CHO) correlated with CDOM absorbance at 254 nm.

Keywords

Colored dissolved organic matter, Fluorescent dissolved organic matter, CDOM, FDOM, dissolved organic carbon flux, ultrahigh resolution mass spectrometry, FT-ICR-MS



1 Introduction

Riverine organic matter is a large source of organic carbon to the oceans (approximately 0.37×10^{15} g C yr⁻¹, Schlesinger and Melack 1981). Globally, riverine dissolved organic carbon (DOC) supplies the largest quantity of reduced carbon (0.25×10^{15} g C yr⁻¹), whereas the flux from individual rivers can vary by several orders of magnitude, depending on watershed characteristics (Hansell and Carlson, 2002; Spencer et al., 2013).

Chromophoric, or colored, dissolved organic matter (CDOM) is the light absorbing fraction of DOM and its properties have been used as indicators of the chemical composition of DOM with regards to molecular weight and aromaticity (Helms et al. 2008; Weishaar et al. 2003). CDOM typically includes a portion that release photons after absorbance of light as fluorescence (FDOM) and this portion of DOM has been widely studied using spectrofluorometry techniques. Important biogeochemical processes are also reflected in changes in CDOM and FDOM properties; these include allochthonous vs. autochthonous carbon sources, and microbial degradation (Cory and McKnight, 2005; Parlanti et al., 2000). Furthermore, DOM absorbance and fluorescence facilitate in situ measurements using submerged sensors (Comeau et al., 2007; Gardner et al., 2005; Jannasch et al., 2008). Recently, the relationship between CDOM and DOC has been characterized for a large number of rivers in order to improve estimates of terrestrial carbon fluxes at the continental scale (Spencer et al., 2012; Stedmon et al., 2011).

With its high mass resolving power ($> 400,000$; i.e. ratio of ion mass to smallest separable ion mass difference), ultrahigh resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) is a powerful analytical method for resolving single mass formulas in the complex mixture of DOM (Stenson et al., 2003). The application of this technique has led to an increased understanding of the processing of DOM in the ocean (e.g. Koch et al., 2014; Kujawinski et al., 2004; Lechtenfeld et al., 2014) and in estuaries (e.g. Lechtenfeld et al., 2013; Sleighter and Hatcher, 2008). Other studies linked optical and chemical DOM properties (Kellerman et al., 2015; Stubbins et al., 2014).

With an annual freshwater discharge of $7,790 \text{ m}^3 \text{ s}^{-1}$, and an annual DOC flux of 4.04×10^{11} g, the Columbia River is one of the major North American sources of DOC to the Pacific Ocean (Hinck et al., 2006; Spencer et al., 2013). Hopkinson et al. (1998) compared DOM of numerous U.S rivers and concluded that Columbia River has a higher contribution of aliphatic DOM, which resulted in increased organic carbon bioavailability (Hopkinson et al., 1998). Further characterization of the chemical composition and seasonal dynamics of this carbon flux are needed to better evaluate the ultimate fate of this DOC source in the estuary and the coastal ocean. Since 2009 sensors deployed in the lower Columbia River estuary have measured FDOM along with a suite of other biogeochemical parameters (Baptista et al., 2015) that provide unprecedented spatial and temporal resolution for this large ecosystem (e.g. Gilbert et al. (2013)).

However, the correlation between CDOM and DOC has previously been described as weak in comparison to other U.S. rivers (Spencer et al., 2012) and therefore the goal of this research was to improve our understanding of the correlations between FDOM and DOC in the context of variability in related biogeochemical cycles and hydrologic conditions. Our specific goal was to investigate if the existing high resolution in situ sensors can be used to quantify the seasonal variability



of the DOC flux to the coastal zone. Moreover, to better understand the influence of biogeochemical events (i.e. phytoplankton blooms and the spring freshet) on the composition of DOM we examined spectrofluorometric properties and used FT-ICR-MS to further characterize Columbia River DOC on the molecular level.



2 Materials & Methods

2.1 Study area and sampling

Samples for laboratory measurements were collected between 26 March and 20 August 2013 at two locations in the lower Columbia River (Fig. 1) that are sites of biogeochemical sensor platforms operated by the Center for Coastal Margin
5 Observation and Prediction (Baptista et al., 2015). With a total of 61 samples (including replicates), station SATURN-08 (river mile 122) was the focus of our sampling campaign due to its accessibility (marina breakwater). Sampling at SATURN-08 included three time series efforts where hourly samples were collected (16 April – 17 April, 29 May – 30 May, 15 July – 16 July). Due to limited access by boat, SATURN-05 (river mile 53) was sampled less frequent ($n = 4$), and primarily served as a site for the downstream comparison of molecular properties of DOM.

10 1 L amber glass bottles with PTFE-lined caps (Thermo Fisher Scientific, cleaned with 10 % HCl for 24 h) were filled with sample water, and stored at 4 °C. Filtration was performed immediately after return to the laboratory, using a 1 L glass filtration unit (Merck Millipore) and GF/F filters (Whatman, 47 mm, pre-combusted at 450 °C for 5 h). To create the necessary vacuum (< 200 mbar), a hand operated vacuum pump (Thermo Fisher Scientific) was used. Since combusted GF/F-filters can adsorb organic compounds (Maske and Garcia-Mendoza, 1994), the filtration unit was pre-rinsed with 200–
15 300 mL of sample to saturate the filter, and rinse the system. Between samples, the filtration unit was thoroughly rinsed with ultrapure water (MilliQ® Plus 185® water purification system, conductivity 18.2 MΩ, total organic carbon < 5 pbb, 22 – 29 °C) before applying the next sample.

DOC samples were stored in 60 mL HDPE bottles at – 20 °C; filtrates for spectroscopic measurements were maintained at 4 °C in 60 mL amber glass bottles (VWR International, pre-combusted at 500 °C for 5 h). The filtrate used for solid-phase
20 extraction was immediately acidified (pH 2) using hydrochloric acid (Suprapur grade®, Merck), and stored at 4 °C until further processing.

2.2 Solid-phase extraction of dissolved organic matter

On four occasions, replicate solid-phase extractions were carried out at both sampling sites ($n = 8$, $n_{\text{total}} = 16$, Tab. 1). Solid-phase extraction (SPE) was performed using PPL cartridges (200 mg, Agilent Technologies; (Dittmar et al., 2008)). Four
25 extractions were carried out simultaneously using a Supelco VisiPrep® unit (Sigma Aldrich). Cartridges were equilibrated with 6 mL of methanol (LC-MS grade, Sigma Aldrich) and 6 mL of acidified ultrapure water (pH 2, hydrochloric acid, Suprapur® grade, Merck). 300 mL of sample filtrate was processed at a flow rate of 2.5 mL min⁻¹ using a peristaltic pump. The cartridges were rinsed with 6 mL of ultrapure water (pH 2) to remove salts, followed by drying with nitrogen (purity grade 5.0). The cartridges were stored at – 20°C until elution with ~1.4 mL methanol (LiChrosolv® grade, Merck). The
30 precise extraction volume was determined by weight.



2.3 Absorbance spectroscopy

UV/Vis spectra were obtained using a spectrophotometer (Tidas I, J&M Analytik AG) equipped with both Tungsten and Deuterium light sources (DH-2000-S, World Precision Instruments). Absorbance through a 1 cm quartz cuvette was scanned in the range of 200 – 723 nm. Ultrapure water was used as a reference; electronic noise was subtracted by measuring signal intensities with the light shutter closed. The mean signal intensity above 700 nm was subtracted from sample absorbance to account for instrument offsets due to particle scattering and differences in refractory indices (Green and Blough, 1994). Absorbance units were converted to naperian absorbance coefficients as follows:

$$a = \frac{2.303 * A}{l} \quad (\text{Eq. 1})$$

where a was the naperian absorption coefficient, A the measured absorbance, and l the measurement path length. The specific ultraviolet absorbance (SUVA) was determined by dividing the UV absorbance at 254 nm (in m^{-1}) by the DOC concentration measured in mg L^{-1} (Weishaar et al., 2003).

2.4 Fluorescence spectroscopy

Filtered samples were analyzed within 5 days after filtration. Spectra were obtained with a spectrofluorometer equipped with a twin channel beam (Fluoromax 4, Horiba Jobin Yvon). Emission of fluorescence was measured at 300–600 nm (increments of 2nm) with excitation wavelengths of 240 – 450nm (increments of 10 nm). Sample temperatures were adjusted to 20°C during analysis. Instrument-specific correction factors were applied using manufacture supplied correction files. The validity of the instrument calibration was checked on a daily basis using the water Raman peak at excitation / emission 350 / 397 and did not vary more than ± 0.5 nm. The optical immaculacy of the quartz cuvette was checked by ensuring a flat emission profile of ultrapure water at an excitation of 280 nm. A daily blank (ultrapure water) was used to measure the Raman peak area at 350 nm for signal normalization.

To obtain comparable fluorescence spectra, also called excitation-emission matrices (EEM), the raw data was subject to post-processing. Observed fluorescence values were corrected for the inner filter effect by applying a correction factor (Parker, 1968):

$$F_{cor} = F_{obs} * 10^{OD_{ex} + OD_{Em}} \quad (\text{Eq. 2})$$

where F_{obs} , and F_{cor} were the raw and corrected fluorescence. The necessary absorption values for the correction (OD_{ex} and OD_{em}) were obtained from the corresponding UV/Vis spectrum. A daily blank (ultrapure water) was subtracted, and fluorescence counts were normalized to the Raman peak area. The fluorescence index (FI) was calculated as the ratio of emission signals at 470 nm and 520 nm obtained at 370 nm excitation (Cory and McKnight, 2005). The apparent fluorescence quantum yield at 350 nm was investigated by dividing the fluorescence emission integral from 300 to 600 nm at an excitation of 350 nm by absorbance at 350 nm (procedure as described previously (Wünsch et al., 2015; Parker and Rees, 1960)). However, since our spectrofluorometer was not calibrated using e.g. quinine sulfate, we were unable to report quantum yield percentages. The biological index (BIX, an index indicating a recent contribution of autochthonous material)



was calculated as the ratio of fluorescence emission at 380 nm and 430 nm at an excitation of 310 nm (Huguet et al., 2009). The Humification Index (HIX) was calculated with the Raman normalized fluorescence emissions I at 254 nm excitation, as follows (Ohno, 2002):

$$HIX = \frac{\sum I_{435-480}}{\sum I_{300-345} + \sum I_{435-480}} \text{ (Eq. 3)}$$

- 5 Since the dataset resolution of EEMs differed from that necessary to calculate HIX, a linear interpolation was performed to obtain the necessary excitation and emission values. HIX values range from 0-1, and indicate the degree of DOM humification.

2.5 Dissolved organic carbon measurement

- DOC was measured using high temperature catalytic oxidation. Organic carbon was oxidized to carbon dioxide, then measured with a non-dispersive infrared detector (TOC/V_{CPN}, Shimadzu). A 6.5 mL sample was poured into a pre-combusted glass vial, then acidified (30 μ L 1 mol L⁻¹ hydrochloric acid, *pro analysi*, Merck) in an autosampler (ASI-V, Shimadzu). Acidified samples were sparged with oxygen to remove inorganic carbon. 50 μ L acidified samples were injected onto the catalyst (heated to 680 °C) using the instrument's autosampler unit. DOC concentrations were averaged over three measurements.
- 15 The instrument was calibrated using laboratory standards (20 – 240 μ mol L⁻¹, potassium hydrogen phthalate, and potassium nitrate (Wako Pure Chemical Industries). For quality control, the deep sea reference (DSR, Hansell laboratory, Miami) was used. The instrument's detection limit was 7 μ mol L⁻¹; DOC was measured with an accuracy of \pm 5 % (Schmidt et al., 2009). The DOC concentration in solid-phase extracts was determined by drying 50 μ L of each methanol extract at 40°C with N₂, and subsequent redissolving in ultrapure water.

20 2.6 In situ sensor

- At SATURN-08, a Satlantic LOBO platform (Jannasch et al., 2008) was operated by the Center for Coastal Margin Observation and Prediction (Portland, Oregon). The in situ data from the sensor package was obtained from instruments that included an optical nitrate sensor (Satlantic SUNA; (Johnson and Coletti, 2002)); a physical-chemical sensor measuring conductivity, temperature, pressure, dissolved oxygen, chlorophyll a, and turbidity (WET Labs Inc. WQM; (Orrico et al., 25 2007)); a sensor for FDOM (WET Labs ECO); and a wet chemistry instrument for phosphate measurements (WET Labs Cycle Phosphate Sensor; (Barnard et al., 2009)). Sensors were cleaned and maintained monthly, following manufacturer-recommended protocols.

- Periodic discreet samples were collected as a quality control measure for chlorophyll a and dissolved nutrients. Acid-washed 1 L Nalgene bottles were rinsed three times with sample water, then filled and transported back to the laboratory where they were analyzed using standard protocols (Welschmeyer 1994, Gilbert et al, 2013). No corrections were made to the in situ nitrate and phosphate measurements, however chlorophyll a was adjusted using linear regression to account for the disparity



between the in situ measurement and the laboratory measurement according to the relationship $y = 3.55x - 0.181$ ($R^2 = 0.93$, $n = 13$).

The in situ fluorescence was measured at an excitation wavelength of 370 nm and an emission wavelength of 460 nm. Since sensor electronics are affected by temperature and the emission of fluorescence is temperature dependent, a temperature correction was performed as described in (Watras et al., 2011). The sensor was placed in a dark PVC flow-through cell and submerged in 4.5 mg/L Suwannee River natural organic matter (Lot: 2R101N) at temperatures ranging from 4 to 35 °C. The temperature correction coefficient ρ was found to be $-0.0297^\circ\text{C}^{-1}$.

2.7 Hydrological and metrological data

10 River discharge at Bonneville Dam (river mile 146) and Beaver army terminal (river mile 53) was measured and provided by the US Army Corps of Engineers and the U.S. Geological Survey National Streamflow Information Program, respectively. The data was accessed through the CMOP online data access interface (www.stccmop.org/datamart/observation_network/dataexplorer).

The meteorological data was retrieved from the United States National Climatic Data Center (<http://www.ncdc.noaa.gov/>).
15 Data for SATURN-08 was collected at station GHCND:USC00358634 (Troutdale, OR, 45.5533 °N, -122.3886 °W). The 30-year climate normal data (1981 - 2010) was obtained from station GHCND:USW00024229 (Portland International Airport, OR, 45.5958 °N, -122.6093 °W). Snowfall data was retrieved from station COOP:353402 (Government Camp, OR, 45.3014 °N, -121.74167 °W, product “TSNW - Total snow fall”). It should be noted that these three weather stations are insufficient to describe precipitation and snowfall events of the whole Columbia River watershed. However, for the purposes
20 of this study, it was assumed that they provide a suitable indication of lateral inputs to the river at the sampling stations.

2.8 Fourier transform ion-cyclotron resonance mass spectrometry

For FT-ICR-MS analysis, DOM extracts were adjusted to an identical concentration of $4,850 \mu\text{mol L}^{-1}$ by dilution with methanol (LiChrosolv[®], Merck). Samples were ionized with electrospray ionization (Apollo II electrospray source, Bruker Daltonik GmbH) at an infusion flow rate of $2 \mu\text{L min}^{-1}$ in negative ion mode.

25 Analyses were performed on a Bruker Solarix XR FT ICR-MS (Bruker Daltonik GmbH) equipped a ParacellTM and a 12 T-refrigerated, actively shielded superconducting magnet (Bruker Biospin).

Initial mass spectra calibration was performed with arginine clusters using a linear calibration. Ion accumulation time was set to 0.03 s, and 200 scans were added to one mass spectrum ranging from m/z (mass to charge ratio) 100 – 1,000. Mass spectra were acquired with 8 Megawords data points, resulting in a resolving power of 790,519 at the highest peak (m/z 381.11932).

30 After analysis, mass spectra were recalibrated internally with a set of marine DOM molecules (Koch et al., 2014). Molecular formulas were calculated by considering the following isotopes: $^1\text{H}_{0-120}$, $^{12}\text{C}_{0-50}$, $^{13}\text{C}_{0-1}$, $^{16}\text{O}_{0-35}$, $^{14}\text{N}_{0-2}$, $^{32}\text{S}_{0-1}$, and $^{34}\text{S}_{0-1}$ (m/z range of 200 - 600). Formulas were restricted to integer double bond equivalent (DBE) values (Stenson et al., 2003); only



formulas with a mass tolerance of ± 0.5 ppm were considered. Peak intensities were normalized to the highest ion peak in each spectrum, and the dataset was restricted to a relative peak intensity of ≥ 0.3 %. Potential anthropogenic surfactants that are listed in a surfactant database (www.terrabase-inc.com), and other contaminations (detected in the process blanks) were removed. The dataset was restricted to formulas with abundances in respective process duplicates, i.e. a formula was
5 disregarded if only detected in one of two replicates. This approach yielded up to 4,286 formulas per sample. It should be considered that every assigned molecular formula most likely implies an immense structural diversity of isomers (Hertkorn et al., 2008). Since FT-ICR-MS alone is unable to distinguish between such structural isomers, we cannot draw conclusions as to the extent to which this applies to an individual molecular formula.

Intensity-weighted average (wa) molecular masses and element ratios were calculated from the normalized peak magnitudes.
10 The DBE for identified molecular formulas was calculated with the number of atoms of an element i (N_i) and its valence V_i , as follows (Koch et al., 2005):

$$DBE = 1 + \frac{\sum_i^{i_{max}} N_i * (V_i - 2)}{2} \text{ (Eq. 5)}$$

The relative peak intensities of ubiquitous formulas were correlated with various parameters (in situ chlorophyll a, DOC, and in situ FDOM), using the Pearson product-moment correlation coefficient (r) to obtain information about the quality and
15 orientation of dependencies. The significance threshold was $p < 0.005$.

To investigate seasonal trends for individual molecular formulas, the relative peak magnitudes of ubiquitous molecular formulas were transformed: Each molecular formula's relative peak magnitude was divided by the maximum relative peak magnitude across samples obtained at that station ("normalized molecular formula abundance").

For the graphical representation, the van Krevelen diagram was used (Kim et al., 2003; Krevelen, 1950). This diagram
20 represents the molecular formula of all molecular formulas with regard to their molecular ratio of oxygen / carbon on the x-axis and hydrogen / carbon on the y-axis. When the z-axis was used to show abundances, we addressed the issue of the graphical representation of molecular formulas with matching molecular ratios by averaging their abundance, rather than plotting superimposed data points.



3 Results

3.1 Seasonal variability of biogeochemical parameters

During the sampling campaign in 2013 (Fig. 2, data with grey background), daily maximum water temperatures ranged from 5.3 °C at the beginning of March to 22.9 °C at the end of August. The total precipitation between March and August 2013 was 257.2 mm. During the winter season (October 2012 – September 2013), the total snowfall at Government camp (station COOP:353402) was 5.7×10^3 mm (data not shown).

A seasonal bloom of phytoplankton was observed as a peak in chlorophyll a during late March 2013 (Fig. 2c). Nitrate was elevated in early spring (until 13 April 2013), and steadily decreased from $30 \mu\text{mol L}^{-1}$ to $6.5 \mu\text{mol L}^{-1}$ by 22 May 2013. Phosphate concentrations were below the sensor detection limit ($0.075 \mu\text{mol L}^{-1}$) from March until mid-May 2013, and never exceeded $0.3 \mu\text{mol L}^{-1}$ for more than one day (data not shown).

The seasonal discharge maximum at SATURN-08 was observed on 15 May 2013, which was followed by the seasonal FDOM maximum on 27 May 2013. With $1000 \text{ m}^3 \text{ s}^{-1}$ at SATURN-08, the seasonal discharge maximum was comparatively low (Fig. 3a). Moreover, spring and summer 2013 were characterized by multiple smaller discharge peaks, whereas other years (e.g. 2011) only show one distinct peak associated with the spring freshet (Fig. 3a). However, the relative discharge contribution of the Willamette River and other regional tributaries to the Columbia River in 2013 followed the typical patterns and magnitudes: With an average of around 15 %, the smallest contribution by volume was observed during July 2013, while reaching up to 65 % during November 2013 (Fig. 3b).

Many of the biogeochemical parameters were comparable at both sensor stations between 2012 and 2013. For example, the seasonal trends of water temperatures were almost identical (Fig. 2d). Moreover, chlorophyll a levels were also comparable with the notable exception of higher levels between May and November 2013, including a small summer phytoplankton bloom during June 2013 at SATURN-05, which was not observed to the same extent at SATURN-08. Discharge was consistently higher at SATURN-05. Water discharge, along with water turbidity during early winter 2012 was proportionally more elevated at the downstream station SATURN-05.

In situ FDOM was a notable exception to the otherwise high similarity between biogeochemical data at both stations. Although showing similar responses to short-term events, e.g. during times of high organic matter concentrations in the river, the instrument at both stations showed different long-term trends. For example, between November 2012 and August 2013, the FDOM sensor at SATURN-05 indicated a noticeable overall decrease in that period, while the signal returned to an apparently constant baseline value at SATURN-08.

3.2 Optical properties of dissolved organic matter

To judge the extent of the seasonal variability of CDOM and FDOM properties, the analysis was focused on samples from station SATURN-08, since the dataset was better suited to investigate correlations ($n = 61$). The highest relative variability was observed in the DOM aromaticity indicator SUVA_{254} , while the FI had the lowest variability (Table 2). Correlation



analyses were conducted to relate the variability of the optical indices to the biogeochemistry of the Columbia River (Fig. 4). With the exception of the SUVA₂₅₄ data, the investigated parameters were normally distributed (Shapiro-Wilk normality test $p < 0.001$). Therefore, Spearman's rank correlation was used to conduct analyses of the relationships between parameters. Seasonal changes in the EEM-derived FI correlated with chlorophyll a (Fig. 4a). BIX, HIX and SUVA₂₅₄ correlated with DOC (Figs. 4b, c, and d). The apparent fluorescence quantum yield did not covary with any of the recorded biogeochemical parameters (data not shown).

The dataset for the linear regression of DOC and FDOM was not normally distributed (Shapiro-Wilk normality test, $p < 0.001$). However, to judge the linearity of relationships, a parametric correlation analysis was conducted and DOC was found to be correlated to in situ FDOM ($R^2 = 0.95$, $p < 0.001$, Fig. 5). Moreover, in situ fluorescence and FDOM fluorescence of filtered samples at excitation / emission 370/460 nm correlated significantly ($R^2 = 0.95$, $p < 0.001$, Fig. 5). It was therefore assumed that in situ fluorescence was sufficiently representative of FDOM fluorescence at the same wavelengths.

As a result of these findings, we used the linear absorbance model to predict DOC using the in situ FDOM data. The average model uncertainty (accounting for uncertainty in the estimated coefficients, and the variance in observations) was $\pm 11.4 \mu\text{mol L}^{-1}$ or $\pm 8.75 \%$.

3.3 Molecular characteristics of dissolved organic matter

The recorded FT-ICR mass spectra displayed very similar distributions of peak magnitudes across samples (Fig. 6a): The highest abundances were observed near m/z 400, in the center of the van Krevelen plot (Fig 6b). On average, we assigned 3994 ± 222 molecular formulas in each sample. The average molecular characteristics for all samples are presented in Table 3.

On a presence or absence basis, the molecular analyses indicated a high degree of similarities between all samples. 3922 out of a total of 5122 molecular formulas were detected at both stations, and 59.4 % (3044) were found in all samples. 85.0 % (3671) and 70.8 % (3345) of the identified molecular formulas were ubiquitous at stations SATURN-05 and SATURN-08, respectively.

However, the dataset also contained unique molecular formulas (444, or 0.8 % at station SATURN-08, Fig. 7a, and 382, or 0.7 % at station SATURN-05, Fig. 7b). At SATURN-08, samples from the spring phytoplankton bloom showed 205 unique molecular formulas that were low in oxygen content and highly saturated ($O/C < 0.5$ and $H/C > 1.2$). However, at SATURN-05, the 90 unique molecular formulas did not show a similar characteristic cluster and were evenly distributed in the compositional space. At both locations, the unique molecular formulas during the spring freshet were dominated by low H/C , but high O/C ratios. Interestingly, samples at SATURN-05 showed high amounts of unique molecular formulas during August 2013 with low O/C and high H/C ratios; a trend that was not equally observed at the upriver station SATURN-08.

Since the degree of similarity between stations and samples was far greater than the amount of unique molecular formulas, we further investigated molecular patterns using the ubiquitous molecular formulas. The degree of similarity between



5 samples, based on the 3044 ubiquitous molecular formulas and their normalized molecular formula abundances (molecular formula abundance divided by the maximum abundance at the respective stations) is shown in the left panel of Fig. 8. The cluster analyses were based on the seasonal abundance changes of each individual molecular formula. Each seasonal event is separated into individual clusters, while process replicates, with the exception of the spring bloom samples, formed their own
5 respective clusters.

To elucidate this cluster analysis, we investigated the average seasonal changes of ubiquitous molecular formula abundances for all sampling times (Fig. 8, right panel). We found a high degree of variability for molecular formulas that can be categorized by their location in the van Krevelen plot. The first group, with H/C (0.9 – 1.3) and O/C ratios (0.4 – 0.6), was high during April 2013 (Fig. 8 top row, right panel). Samples collected during the rain event displayed a broader range of
10 highly abundant molecular formulas extending to the outer margin of the van Krevelen plot.

The second group, with low H/C (< 1.0) and high O/C ratios (> 0.5), had characteristically high relative peak magnitudes during June 2013, when the Columbia River's annual discharge maximum was reached. On the other hand, this group was found to be lower in abundance in all remaining samples. At the same time, most molecular formulas in the first group were found to be comparatively low during June (and also August).

15 To elucidate the peak magnitude variability of ubiquitous molecular formulas (Fig. 8), we combined the in situ dataset at both stations with FT-ICR-MS spectra obtained from samples at this station, and carried out correlation analyses to establish molecular markers for biogeochemical events. Chlorophyll a levels, and the napherian absorbance coefficient at 254 nm were chosen as proxy for the phytoplankton bloom, and the increased organic matter export during the spring freshet, respectively. We found a total of 978 molecular formulas that satisfied the significance threshold ($p < 0.005$) and displayed
20 positive correlations (Fig. 9). With regards to elemental composition, molecular properties of correlated molecular formulas were highly similar between stations (Fig. 9).

Of the 561 molecular formulas that correlated with chlorophyll a levels at both stations, all but 13 contained sulfur or nitrogen as heteroatoms. 94.2 % of correlated molecular formulas at SATURN-08 were also detected at station SATURN-05 (but not necessarily correlated with chlorophyll a there). The elemental characteristics of the chlorophyll-correlated subgroup
25 were close to the average molecular characteristics of all samples (e.g. average H/C ratio of 1.23, and O/C ratio of 0.46, along with an average DBE of 9.2 at SATURN-08).

However, the characteristics of the subgroup of molecular formulas that correlate with CDOM absorbance at 254 nm were different. Only five of the 417 molecular formulas contained sulfur, and none of them contained nitrogen. Compared to the average composition in all samples, these molecular formulas were highly unsaturated.

30



4 Discussion

The Columbia River is well-studied; previous investigations have documented annual patterns in springtime phytoplankton blooms, discharge maxima, and nutrient availability (Sullivan et al., 2001). Since 1995, the U.S. Geological Survey sampled the Columbia River on a monthly basis to collect long-term time series data, including CDOM absorbance (Spencer et al., 2012). However, such constant monitoring relies on a relatively low sampling frequency that is, if at all, usually only supported by basic sensor products, such as river discharge and water temperature. In contrast, this study was supported by a more comprehensive array of in situ sensors, allowing for reactive sampling during different environmental conditions that captured the biogeochemical variability of DOM for optical and molecular studies.

The continuous data acquisition at both in situ stations was successfully used to identify biogeochemical events during the sampling campaign in spring and summer 2013. As evident by elevated chlorophyll concentrations, the annual spring bloom developed during late March and early April. The dominant phytoplankton species during that time was *Asterionella formosa* (Maier and Peterson, 2014). The timing of the spring bloom onset matched that observed in previous studies of the Columbia River (Sullivan et al., 2001), and was likely driven by light and nutrient availability. River discharge is known to be a contributing factor for the length and intensity of phytoplankton blooms in the Columbia River (Sullivan et al., 2001). Correspondingly, during early spring 2013, rainfall during the peak of the bloom contributed to increasing river discharge, causing a steep decline in phytoplankton abundance.

Phosphate concentrations were low, compared to historical data, especially during early spring. During this time, historic data (1978-1994, USGS data at river kilometer 227) indicates phosphate concentrations of $> 0.5 \mu\text{mol L}^{-1}$ (Sullivan et al., 2001), whereas during early spring 2013, phosphate was below our sensor's detection limit ($< 0.075 \mu\text{mol L}^{-1}$). Given the molar Redfield ratio of approximately 16N:1P (Redfield, 1934), this indicates that phytoplankton growth was probably phosphate-limited as nitrate to phosphate ratios were never found to be below 33.8. Nonetheless, despite very low phosphate we observed a freshwater diatom bloom at typical levels (approximately $40 \mu\text{g chlorophyll a L}^{-1}$), suggesting that, although low, phosphate was never fully depleted, rather, instantly utilized and rapidly recycled.

The seasonal pattern of nitrate was typical for a temperate environment, and concentrations matched previous studies in the Columbia River (Sullivan et al., 2001). High nitrate concentrations during winter were attributable to reduced nitrogen uptake, and increased terrestrial nitrate runoff (Wall et al., 1998). In contrast, increased primary productivity and low levels of terrestrial nitrate input during spring and summer were likely responsible for lower nitrate concentrations.

As apparent from the meteorological data, spring and summer 2013 had low amounts of rainfall (45 mm or 15 % less rain from March to August, compared to the 30-year climate normal), and less snowfall (25.6 % less, compared to the winter season 2010/2011). This is directly reflected in the comparatively low discharge maximum at Bonneville dam (30 % lower in 2013 compared to 2011; Fig. 3a). Moreover, the existence of multiple comparable discharge peaks during spring, together with coinciding high amount of rainfall during May 2013 and the overall low snowfall during the winter of 2012 / 2013 suggests that the seasonal discharge maximum was not primarily driven by snowmelt. Rather it suggests that, although the



base of the elevated discharge during spring and summer was induced by snowmelt, the discharge peaks were caused by increased lateral runoff due to the rainfall events. Therefore, although the highest seasonal FDOM signals during late May and early June coincided with the discharge maximum, the FDOM peak most likely did not represent a typical snowmelt-dominated spring freshet, but an increased contribution of terrestrial runoff to the river. The influence of such inter-annually varying biogeochemical conditions on the organic matter characteristics of the Columbia River remains unknown, and should thus be subject to further studies.

The comparison data products from both in situ sensors beyond the sampling campaign allowed insights into the variability within the Columbia River. As stated above, many parameters showed significant degrees of overlap (e.g. water temperature) between 2012 and 2013. However, some seasonal patterns also differed between both stations. For example, water discharge and turbidity during the winter months of 2012 showed a disproportionate increase at the downstream location SATURN-05 (Fig. 2a and Fig. 3b). During this period of increased rainfall, the Willamette River receives more terrestrial runoff, which likely contributes to the increased discharge and increased nitrate concentrations at SATURN-05. Contrary to the Willamette watershed, precipitation upstream of SATURN-08 in the Columbia River watershed partially falls as snow and is retained behind multiple hydropower dams and therefore has less direct impact on river discharge.

Interestingly, during July 2013, chlorophyll a levels indicate the presence of a low level phytoplankton bloom at the downstream station SATURN-05, that was not observed upstream, at SATURN-08. During this time, SATURN-05 showed increased nitrate that originated from the Willamette River, as apparent from data of the USGS National Streamflow Information program (data not shown). Since the Willamette River is usually also higher in phosphate (Prahl et al., 1997), increased nutrient availability likely caused a second summer phytoplankton bloom at SATURN-05 that was not observed at SATURN-08.

Most notably, in situ FDOM differs considerably between SATURN-05 and SATURN-08. Since a good agreement between in situ FDOM and DOC was found at SATURN-08, the quality of data acquired from SATURN-05 was further investigated. Unfortunately, due to the low number of samples ($n = 8$), a detailed analysis of the data validity using DOC measurements was difficult at SATURN-05. For the available data, no significant correlation between in situ FDOM readings and DOC was found. This was also true when considering the extended dataset of in situ FDOM and DOC (measured by USGS) between 2009 and 2013. Therefore, the differences in readings between SATURN-05 and SATURN-08 are likely caused by one or several disturbances. Firstly, instrumental issues could have affected readings at SATURN-05. However, it has to be noted that samples discussed in Spencer et al. (2012) were taken at the same location and DOC was also found to be not well correlated to CDOM absorbance of filtered samples. Therefore, other factors, such as local point sources of industrial or urban pollution could have significantly influenced the relationship on a local level since SATURN-05 is located downstream an area affected by industrial activities.



4.1 Chromophoric dissolved organic matter as a proxy for dissolved organic carbon

Previous studies revealed a weak relationship between DOC concentration and CDOM absorbance for the Columbia River (Spencer et al., 2012). The absorbance data acquired at SATURN-08 correlated stronger with DOC compared to a previous study, which highlights the inter-annual variability of the system and highlights the need for continuous monitoring of the DOC-CDOM relationship ($R^2 = 0.70$ in (Spencer et al., 2012) vs. 0.90 in this study). Furthermore, assuming that a negative intercept in the correlation between DOC and CDOM indicates the presence of a non-chromophoric fraction, Spencer et al. (2012) concluded that this fraction is of considerable size in the Columbia River when compared to other US-rivers. Compared to this study, we observed an intercept that was closer to zero (-7.57 in Spencer et al. (2012) vs. -4.7 in 2013). Therefore, the chromophoric portion of DOC was larger, and CDOM represented DOC better than in previous studies. This inter-annual variability in CDOM properties demonstrated that the correlation between DOC and CDOM must be constantly monitored to be able to use in situ systems for the estimation of DOC.

Combining river discharge data from Bonneville dam with the DOC predictions made with the in situ DOM fluorescence data, we were able to estimate DOC fluxes of the Columbia River at river mile 122 (Station SATURN-08) during spring and summer 2013 (Fig. 10a). The average model uncertainty was low with $\pm 8.57\%$, considering that DOC via catalytic combustion is determined with an uncertainty of $\pm 5\%$ (e.g. Schmidt et al., 2009 for the same instrument).

With $0.93 \text{ Gg DOC d}^{-1}$, the average DOC flux during the study period was 15.5 % lower than previously reported ($1.1 \text{ Gg DOC d}^{-1}$ during 2009 and 2010; Spencer et al. (2012)). However the flux and therefore delivery of DOC to the estuary and moreover the costal Pacific Ocean varied significantly; values ranged from 0.46 to 1.62 Gg d^{-1} . Moreover, our flux estimate is based on a year with unusually low river discharge, and is supported by summertime data only. If winter DOC runoff would be included, lower discharge (approx. $3000 - 4000 \text{ m}^3 \text{ s}^{-1}$) along with lower DOC concentration would increase the gap to previous estimates and widen the range of observed DOC flux and lower the average annual DOC discharge. However, during years with high amounts of discharge, the DOC flux estimate would presumably be higher than reported here or in (Spencer et al., 2012). This highlights the need for such in situ sensors for the accurate estimation of riverine DOC flux and delivery to the costal oceans.

Our sensor-based estimates highlighted several advantages. First, besides periodic sampling for validation purposes, our estimates do not rely on extended sampling activities. This eliminates the risk of missing prominent, but short-term phenomena such as the spring freshet (Fig. 10b). Secondly, the capability of in situ DOC prediction provides new possibilities of quantifying the inter-annual variability of the Columbia River DOC discharge. Long-term discharge records from Bonneville dam indicate that the water discharge, and the magnitude and duration of the spring freshet, can vary greatly between years. In situ sensor, such as the systems presented in our work, can help to more accurately describe carbon budgets of the costal margin and reduce uncertainties in large-scale estimates.



4.2 The influence of biogeochemical events on the chemical composition of dissolved organic matter

The optical properties of DOM were successfully used for the cost-effective characterization of Columbia River DOM. With an average HIX of 0.86, Columbia River DOM shows a high degree of humification. The SUVA₂₅₄ values were higher than previously reported (3.1 ± 0.2 in this study vs. $2.6 \pm 0.4 \text{ L mg C}^{-1} \text{ m}^{-1}$ in Spencer et al. (2012)). This implies that DOM was slightly higher in aromaticity compared to previous studies. However, the fact that samples in our study were only taken in spring and summer might explain the slightly different values. Furthermore, the combination of optical indices, namely the FI (mean of 1.38), HIX (mean of 0.86) and BIX (mean of 0.63) indicated that DOM was clearly terrestrially dominated, and low in autochthonous DOM (Cory and McKnight, 2005; Huguet et al., 2009; Ohno, 2002). Overall, we observed relatively low variability in all CDOM- and FDOM-derived parameters. However, changes were correlated to biogeochemical indicators associated with e.g. phytoplankton abundance.

We saw increased SUVA₂₅₄ values, along with a decrease in the EEM-derived BIX, and an increase of HIX during the seasonal discharge maximum. This demonstrates increased input of terrestrial, allochthonous, aromatic DOM during this period.

FT-ICR-MS was able to provide unique molecular formulas for individual sampling time-points. Since these only account for less than 1 % of molecular formulas in the overall dataset, their relative contribution was minor. However, unique formulas can provide valuable insights into temporally and spatially unique characteristics and processes. The relative discharge of the Willamette (and other tributaries) compared to the Columbia River at SATURN-08 varied from ca. 30 % in March to 12 % in August. Consequently, unique molecular features were most similar at both sampling locations during June and August 2013. For example, unique formulas during the discharge maximum were highly unsaturated and high in O/C ratios at both sampling stations. These unique formulas also had highly similar characteristics to the group of molecular formulas that correlated with CDOM abundance. However, samples taken at earlier times in the season (most notably during the phytoplankton bloom) showed noticeably different molecular characteristics of unique molecular formulas. These differences might point at the seasonal differences of the organic matter contribution from the Willamette.

However, in agreement with the optical properties described earlier, the FT-ICR-MS data demonstrated that the overall chemical variability of DOM in the Columbia River is low, as seen in the high amount of ubiquitous molecular formulas. This high amount of ubiquitous molecular formulas in the FT-ICR-MS dataset enabled us to closely follow changes in temporal abundance of the majority of solid-phase extractable DOM. By referring to each molecular formula's abundance as the ratio of its maximum seasonal abundance (Koch et al., 2014), we gained valuable insights into seasonal trends of DOM dynamics. Another goal was to establish a set of molecular markers that are closely tied to the biogeochemistry of the Columbia River. The combination of optical and biogeochemical data with the FT-ICR mass spectra allowed the identification of molecular formulas that contributed to the terrestrial input during the spring freshet. CDOM-correlated molecular formulas were predominantly CHO containing formulas, and were characterized by low H/C ratios, with above-average numbers of double bonds. The high abundance of such molecular formulas during the spring freshet could therefore



explain shifts in $SUVA_{254}$, BIX, and HIX. These findings agree with previous studies that linked chemical and optical properties of DOM (Kellerman et al., 2015; Stubbins et al., 2014) and indicate the validity of these links across different aquatic environments.

With regard to the molecular analysis, it is important to consider that the ionization technique used (electrospray ionization) varies in efficiency for different DOM compounds (Schmidt et al., 2009). Furthermore, contrasting DOM compound classes may be extracted with varying efficiency. Correlating relative ion intensities with in situ sensor data is a novel approach, allowing for a better interpretation of molecular changes identified in FT-ICR-MS. To ensure the validity of this approach, process replicates were analyzed, and molecular formulas that were not detected in both replicates were disregarded. The cluster analysis suggested that sample preparation and analysis were highly reproducible. However, as stated above, it is important to note that ubiquitous presence of molecular formulas in our FT-ICR-MS dataset does not imply ubiquitous presence of chemically identical compounds as FT-ICR-MS alone is unable to distinguish between isomers.

In addition to changes in DOM characteristics during the spring freshet, molecular changes due to the spring phytoplankton bloom were also observed. Shifts in the fluorescence index indicated increased levels of fresh microbial DOM, demonstrating the link between primary and secondary production in the Columbia River. FT-ICR mass spectra elucidated these findings and enabled us to identify chemical formulas that were tied to the phytoplankton bloom. The fact that the majority of chlorophyll-correlated molecular formulas contained either nitrogen or sulfur indicates a biological source and supports the classification of these formulas as autochthonous. These trends seem to persist across systems, since Kellerman et al. (2015) found molecular formulas with similar average elemental ratios to be correlated to the biological index in 109 Swedish lakes. However, with the available dataset, we were unable to distinguish between DOM derived from primary and secondary production. CHON containing formulas did not directly represent mono- or di-N-atomic peptides or amino acids since their molecular masses were mostly too high. This has previously been observed in a previous study on sediment samples (Schmidt et al., 2014): the authors suggested that CHON containing formulas could have originated from proteins or peptides that were altered by biotic or abiotic processes. However, the fact that the majority of the correlated molecular formulas were found at both sampling locations indicates that their lifetime is sufficiently long or their source is ubiquitous. Hence, these formulas might be suitable molecular tracers to follow the fate of allochthonous and autochthonous DOM in the Columbia River estuary to e.g. elucidate their behaviour in the mixing zone.



5 Conclusions

The results of this study provide new details on the seasonal variability of DOM in the Columbia River. The correlation of FDOM optical indices with biogeochemical indicators such as chlorophyll or DOC demonstrated the applicability of DOM optical properties to track seasonal changes of DOM. Moreover, data from FT-ICR-MS analyses confirmed these observations on a molecular basis. Both optical indices and FT-ICR-MS data demonstrate low overall variability of DOM characteristics for the Columbia River, while also highlighting seasonal shifts within the DOM pool.

The successfully established correlation between in situ FDOM and DOC at station SATURN-08 provides new insights into short- (daily to weekly) and medium-scale (seasonally) DOC flux variability in the Columbia River for the first time. While average DOC flux estimates did not vary significantly from previous estimates, our data provide evidence of high variability on day to week timescales as flux varied up to 3-fold from minimum to maximum estimates. These data highlight the ability of in situ sensors to enable an improved understanding of riverine biogeochemistry in general, and to provide improved estimates of coastal carbon flux budgets in particular.



Code availability

The code used to produce and process data, tables and figures is available upon request directly from the corresponding author Urban J. Wunsch.

Data availability

- 5 During the review process, the data presented in this work is available upon request directly from the corresponding author Urban J. Wunsch. Upon publication of this manuscript, the data will be made publically available at <https://pangaea.de/>.

Author contribution

Urban J. Wunsch: Collection of samples, sample and data processing, drafting of the manuscript.

Boris P. Koch: Processing of mass spectrometry data, respective data evaluation, drafting of the manuscript.

- 10 Matthias Witt: Realization of mass spectrometry measurements, assistance in related data evaluation, drafting of the manuscript.

Joseph A. Needoba: Maintenance of in situ sensors, realization of dissolved nutrient measurements, drafting of the manuscript.

Acknowledgements

- 15 Funding to support the sensor platform at SATURN-08 was provided by the Lower Columbia River Estuary Partnership under its Ecosystem Monitoring Program, and the Bonneville Power Administration. Funding to support the SATURN-05 platform was provided by the National Science Foundation under cooperative agreement OCE-0424602. We thank the U.S. Geological Survey field team for the assistance provided during field sampling of the USGS NASQAN site. We further thank Estefania Garcia Llaneza and Michelle Maier for their support during field sampling and Stephen Fortuna for
- 20 proofreading this manuscript. UJW would like to acknowledge funding provided by the German Academic Exchange Service.



References

- Baptista, A. M., Seaton, C., Wilkin, M. P., Riseman, S. F., Needoba, J. A., Maier, D., Turner, P. J., Kärnä, T., Lopez, J. E., Herfort, L., Megler, V. M., McNeil, C., Crump, B. C., Peterson, T. D., Spitz, Y. H. and Simon, H. M.: Infrastructure for collaborative science and societal applications in the Columbia River estuary, *Front. Earth Sci.*, 9(4), 659–682, doi:10.1007/s11707-015-0540-5, 2015.
- Comeau, A. J., Lewis, M. R., Cullen, J. J., Adams, R. S., Andrea, J., Feener, S., McLean, S. D., Johnson, K. S., Coletti, L. J., Jannasch, H. W., Fitzwater, S. E., Moore, C. and Barnard, A. H.: Monitoring the spring bloom in an ice covered fjord with the Land/Ocean Biogeochemical Observatory (LOBO), *Ocean. Conf. Rec.*, 1–7, doi:10.1109/OCEANS.2007.4449185, 2007.
- Cory, R. M. and McKnight, D. M.: Fluorescence spectroscopy reveals ubiquitous presence of oxidized and reduced quinones in dissolved organic matter, *Environ. Sci. Technol.*, 39(21), 8142–8149, doi:10.1021/es0506962, 2005.
- Gardner, G. B., Chen, R. F. and Berry, A.: High-resolution measurements of chromophoric dissolved organic matter (CDOM) in the Neponset River Estuary, Boston Harbor, MA, *Mar. Chem.*, 96(1-2), 137–154, doi:10.1016/j.marchem.2004.12.006, 2005.
- Gilbert, M., Needoba, J., Koch, C., Barnard, A. and Baptista, A.: Nutrient Loading and Transformations in the Columbia River Estuary Determined by High-Resolution In Situ Sensors, *Estuaries and Coasts*, 36(4), 708–727, doi:10.1007/s12237-013-9597-0, 2013.
- Hansell, D. A. and Carlson, C. A.: *Biogeochemistry of Marine Dissolved Organic Matter*, Academic Press, San Diego., 2002.
- Hinck, J. E., Schmitt, C. J., Blazer, V. S., Denslow, N. D., Bartish, T. M., Anderson, P. J., Coyle, J. J., Dethloff, G. M. and Tillitt, D. E.: Environmental contaminants and biomarker responses in fish from the Columbia River and its tributaries: Spatial and temporal trends, *Sci. Total Environ.*, 366(2-3), 549–578, doi:10.1016/j.scitotenv.2005.11.008, 2006.
- Hopkinson, C. S., Buffam, I., Hobbie, J., Vallino, J., Perdue, M., Eversmeyer, B., Prahl, F., Covert, J., Hodson, R., Moran, M. a, Smith, E., Baross, J., Crump, B., Findlay, S. and Foreman, K.: Terrestrial inputs of organic matter to coastal ecosystems: An intercomparison of chemical characteristics and bioavailability., *Biogeochemistry*, 43, 221–234, doi:10.1023/a:1006016030299, 1998.
- Jannasch, H. W., Coletti, L. J., Johnson, K. S., Fitzwater, S. E., Needoba, J. a. and Plant, J. N.: The Land/Ocean Biogeochemical Observatory: A robust networked mooring system for continuously monitoring complex biogeochemical cycles in estuaries, *Limnol. Oceanogr. Methods*, 6, 263–276, doi:10.4319/lom.2008.6.263, 2008.
- Kellerman, A. M., Kothawala, D. N., Dittmar, T. and Tranvik, L. J.: Persistence of dissolved organic matter in lakes related to its molecular characteristics, *Nat. Geosci.*, 8(May), 454–457, doi:10.1038/ngeo2440, 2015.
- Koch, B. P., Kattner, G., Witt, M. and Passow, U.: Molecular insights into the microbial formation of marine dissolved



- organic matter: Recalcitrant or labile?, *Biogeosciences*, 11(15), 4173–4190, doi:10.5194/bg-11-4173-2014, 2014.
- Kujawinski, E. B., Del Vecchio, R., Blough, N. V., Klein, G. C. and Marshall, A. G.: Probing molecular-level transformations of dissolved organic matter: Insights on photochemical degradation and protozoan modification of DOM from electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Mar. Chem.*, 92(1-4 SPEC. ISS.), 23–37, doi:10.1016/j.marchem.2004.06.038, 2004.
- 5 Lechtenfeld, O. J., Koch, B. P., Gašparović, B., Frka, S., Witt, M. and Kattner, G.: The influence of salinity on the molecular and optical properties of surface microlayers in a karstic estuary, *Mar. Chem.*, 150, 25–38, doi:10.1016/j.marchem.2013.01.006, 2013.
- Lechtenfeld, O. J., Kattner, G., Flerus, R., McCallister, S. L., Schmitt-Kopplin, P. and Koch, B. P.: Molecular transformation and degradation of refractory dissolved organic matter in the Atlantic and Southern Ocean, *Geochim. Cosmochim. Acta*, 126, 321–337, doi:10.1016/j.gca.2013.11.009, 2014.
- 10 Parlanti, E., Wörz, K., Geoffroy, L. and Lamotte, M.: Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs, *Org. Geochem.*, 31(12), 1765–1781, doi:10.1016/S0146-6380(00)00124-8, 2000.
- 15 Sleighter, R. L. and Hatcher, P. G.: Molecular characterization of dissolved organic matter (DOM) along a river to ocean transect of the lower Chesapeake Bay by ultrahigh resolution electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry, *Mar. Chem.*, 110(3-4), 140–152, doi:10.1016/j.marchem.2008.04.008, 2008.
- Spencer, R. G. M., Butler, K. D. and Aiken, G. R.: Dissolved organic carbon and chromophoric dissolved organic matter properties of rivers in the USA, *J. Geophys. Res. Biogeosciences*, 117(3), 1–14, doi:10.1029/2011JG001928, 2012.
- 20 Spencer, R. G. M., Aiken, G. R., Dornblaser, M. M., Butler, K. D., Holmes, R. M., Fiske, G., Mann, P. J. and Stubbins, A.: Chromophoric dissolved organic matter export from U.S. rivers, *Geophys. Res. Lett.*, 40(8), 1575–1579, doi:10.1002/grl.50357, 2013.
- Stedmon, C. A., Thomas, D. N., Papadimitriou, S., Granskog, M. A. and Dieckmann, G. S.: Using fluorescence to characterize dissolved organic matter in Antarctic sea ice brines, *J. Geophys. Res. Biogeosciences*, 116(3), 1–9, doi:10.1029/2011JG001716, 2011.
- 25 Stenson, A. C., Marshall, A. G. and Cooper, W. T.: Exact Masses and Chemical Formulas of Individual Suwannee River Fulvic Acids from Ultrahigh Resolution Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectra, *Anal. Chem.*, 75(6), 1275–1284, doi:10.1021/ac026106p, 2003.
- Stubbins, A., Lapierre, J.-F., Berggren, M., Prairie, Y. T., Dittmar, T. and Del Giorgio, P. A.: What's in an EEM? Molecular Signatures Associated with Dissolved Organic Fluorescence in Boreal Canada., *Environ. Sci. Technol.*, 48, 10598–10606, doi:10.1021/es502086e, 2014.
- 30



Tables

Table 1. Sampling dates for the extraction of dissolved organic matter. Replicates were collected at each sampling event. All samples were collected between 08:00 and 10:00 PST with the exception of the mid-summer samples at SATURN-08, which were collected at 14:00 PST.

Description	SATURN-05	SATURN-08	Abbreviation
Phytoplankton bloom	01 Apr 2013	01 Apr 2013	SB
Spring rain event	10 Apr 2013	10 Apr 2013	SR
Spring freshet	04 Jun 2013	05 Jun 2013	SF
Mid-summer	14 Aug 2013	20 Aug 2013	ST

5



Table 2. Summary of various optical indices and their relative variability at station SATURN-08. SD: standard deviation, %RSD: relative standard deviation, in percent, FI: Fluorescence index, BIX: Biological index, HIX: Humification index, SUVA₂₅₄: Specific UV absorbance at 254 nm.

Parameter	Range	Average \pm SD	%RSD
FI	1.34 - 1.43	1.37 \pm 0.019	1.4
BIX	0.58 - 0.69	0.63 \pm 0.033	5.3
HIX	0.75 - 0.88	0.85 \pm 0.027	3.2
SUVA ₂₅₄	2.28 - 3.84	3.16 \pm 0.274	8.7



Table 3. Average molecular characteristics of SPE-DOM samples analyzed with FT-ICR-MS. n_{Ion} : Total number of formulas. Weighted averages (wa) of: molecular mass (Mass), double bond equivalent (DBE), oxygen to carbon ratio (O/C), hydrogen to carbon ratio (H/C), nitrogen to carbon ratio (N/C), and sulfur to carbon ratio (S/C). Samples are identified by sampling location (SAT-05 or SAT-08) and sampling date (m/d). All given parameters are averages of process duplicates.

Sample	n_{Ion}	Mass _{wa}	DBE _{wa}	O/C _{wa}	H/C _{wa}	N/C _{wa}	S/C _{wa}
SAT-05 4/1	3798	439.50	9.67	0.50	1.17	0.0067	0.0027
SAT-05 4/10	3761	439.26	9.71	0.50	1.16	0.0066	0.0023
SAT-05 6/4	3752	438.88	9.80	0.51	1.15	0.0050	0.0022
SAT-05 8/20	3952	434.39	9.55	0.50	1.17	0.0060	0.0021
SAT-08 4/1	4271	430.12	9.35	0.49	1.19	0.0070	0.0023
SAT-08 4/10	4286	432.33	9.51	0.50	1.17	0.0068	0.0020
SAT-08 6/5	3963	428.98	9.49	0.50	1.17	0.0049	0.0015
SAT-08 8/14	4169	427.19	9.32	0.50	1.18	0.0060	0.0015

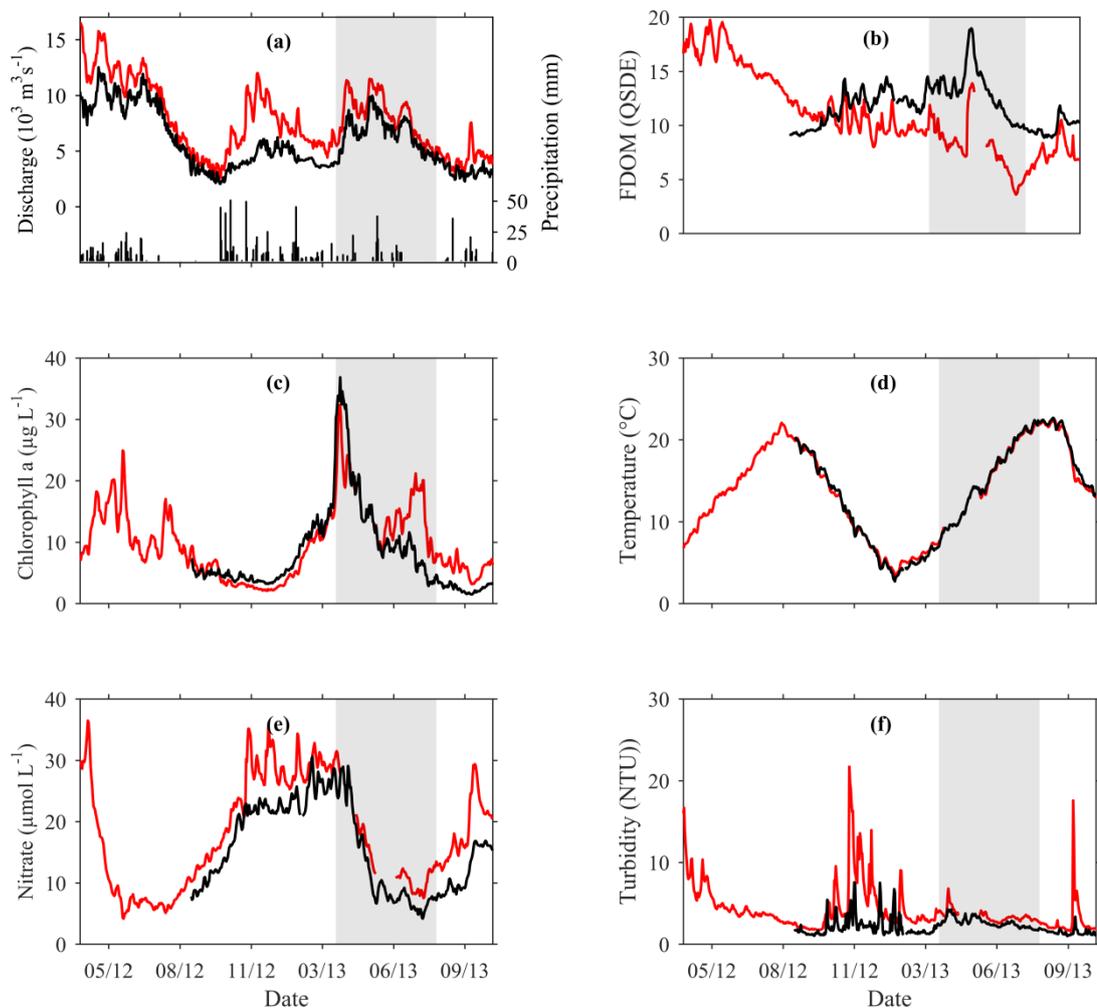


Figures

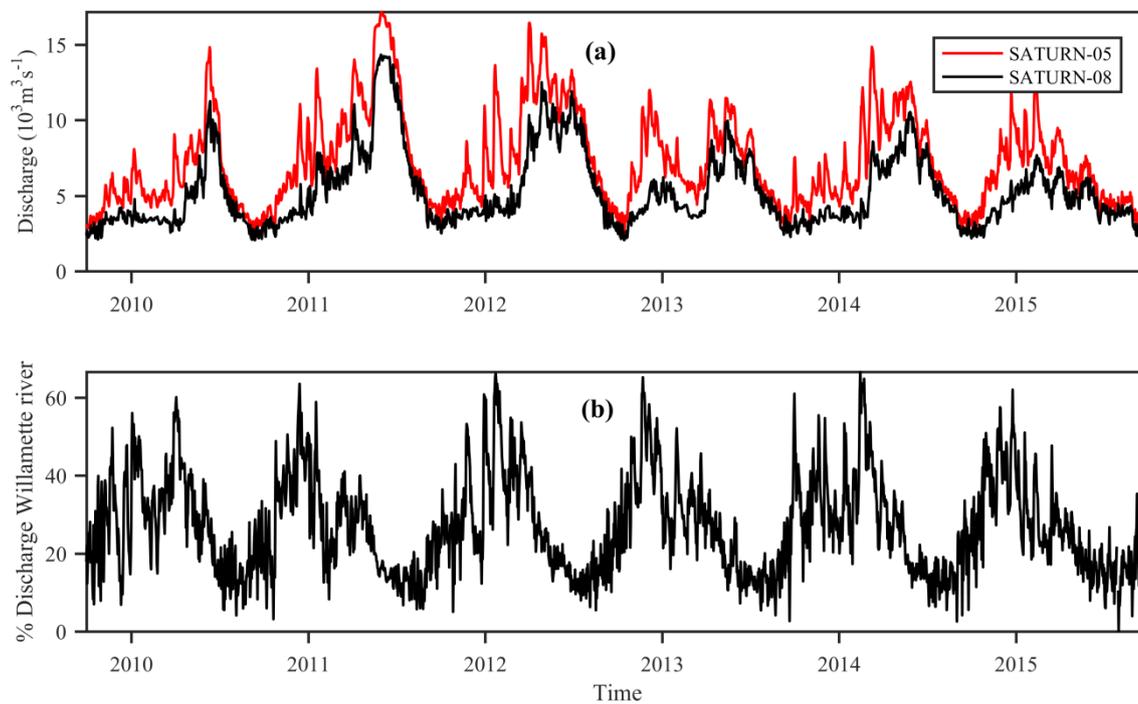


Figure 1. Sampling locations in the lower Columbia River, USA. White lines represent the two rivers, Columbia and Willamette. SATURN-05 (46.184 °N, -123.187 °W) is located at river kilometer 85, SATURN-08 (45.577 °N, -122.379 °W) at river kilometer 196. The Willamette River joins the Columbia River between SATURN-05 and SATURN-08, at river kilometer 139. Map data provided by Google and U.S. Geological Survey.

5



5 **Figure 2.** Daily average meteorological and biogeochemical data from SATURN-08 (solid black lines) and SATURN-05 (solid red lines) from 01 Apr 2012– 01 Nov 2013. Precipitation data (a, right axis, bar-plot) is provided as daily precipitation total and provided for a meteorological station close to SATURN-08 (GHCND:USC00358634). SATURN-08 was installed in September 2012, thus no data prior to that date is available at this station. Other missing data are the result of sensor malfunction or scheduled maintenance.



5 **Figure 3. Multi-year discharge patterns of the Columbia River from 2010 to 2015 at stations SATURN-05 (river mile 53) and SATURN-08 (river mile 122) (a), and relative contribution of the Willamette (b). The relative contribution was calculated with the assumption that the Willamette River is the only significant source of freshwater discharge to the Columbia River between both sampling stations.**

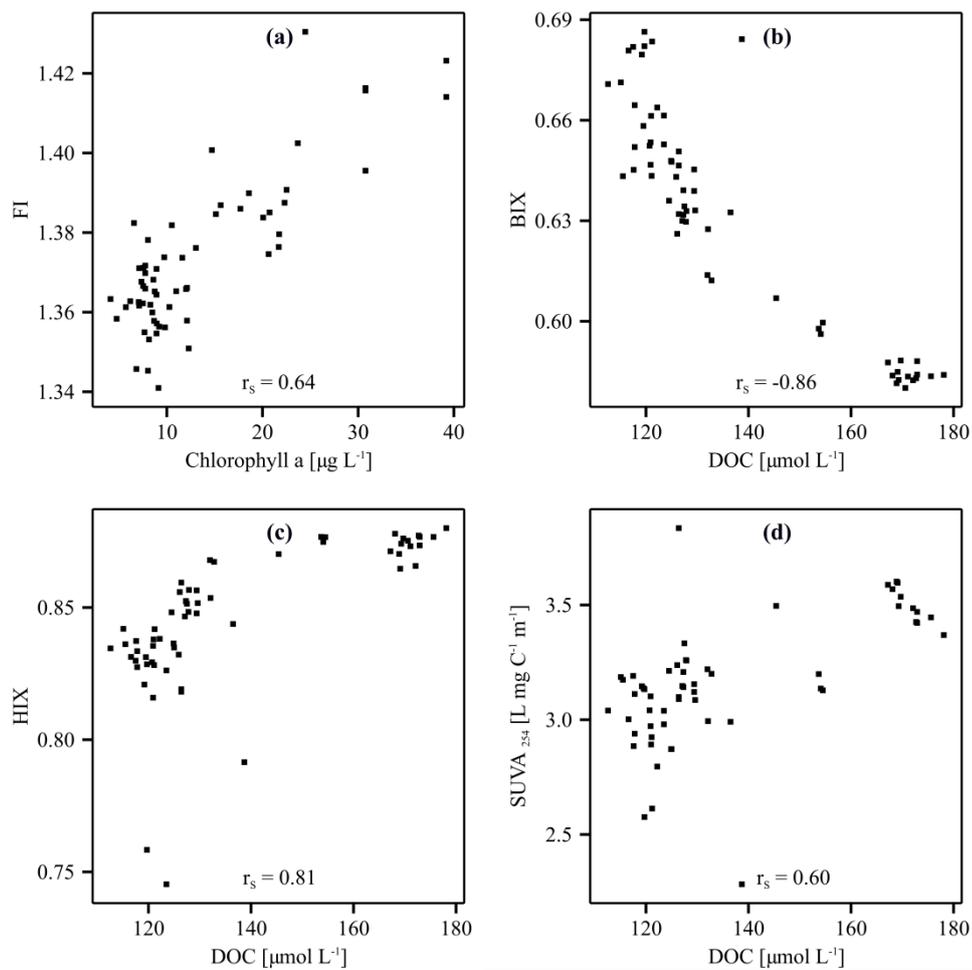


Figure 4. Relationship of optical indices with biogeochemical parameters. r_s : Spearman's rank correlation coefficient. All of the visualized correlations were significant ($p < 0.001$).

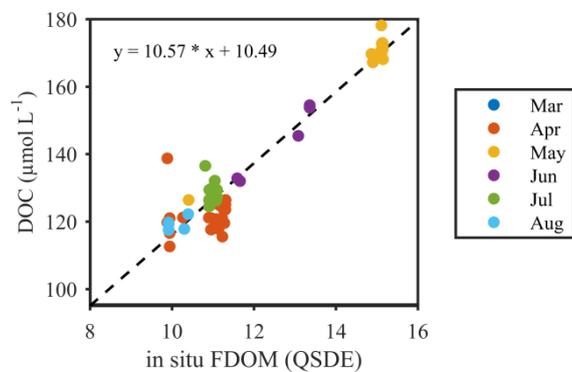


Figure 5. Correlation of DOC and temperature-corrected in situ FDOM (ex/em 370/460 nm). Colored dots represent the month of sampling, the dashed line represents the fitted linear model. FDOM correlated statistically significant ($R^2 = 0.95$). QSDE: Quinine Sulphate Dihydrate Equivalent.

5

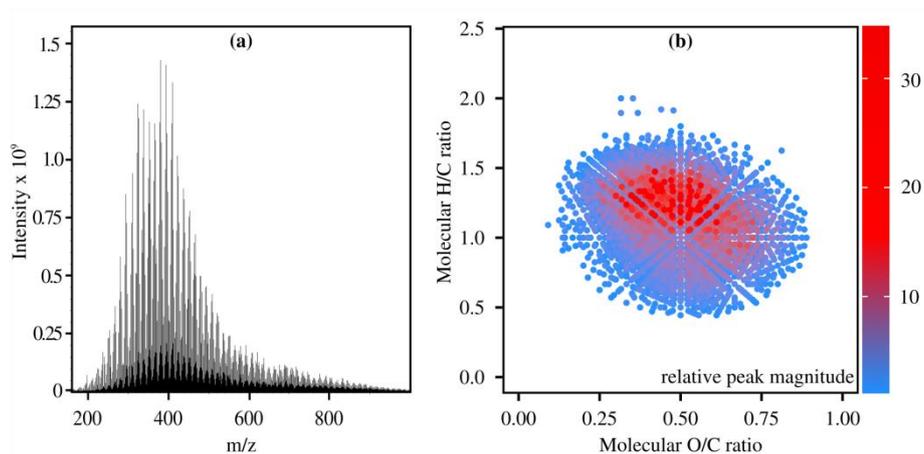


Figure 6. Representative FT-ICR-MS analysis of Columbia River SPE-DOM (August 2013, SATURN 08). (a) ESI-negative spectrum. (b) Molecular ratios of identified formulas (Van Krevelen diagram, colored by relative abundance of individual formulas. Identical ratios are represented by the average relative intensity of the respective formulas.

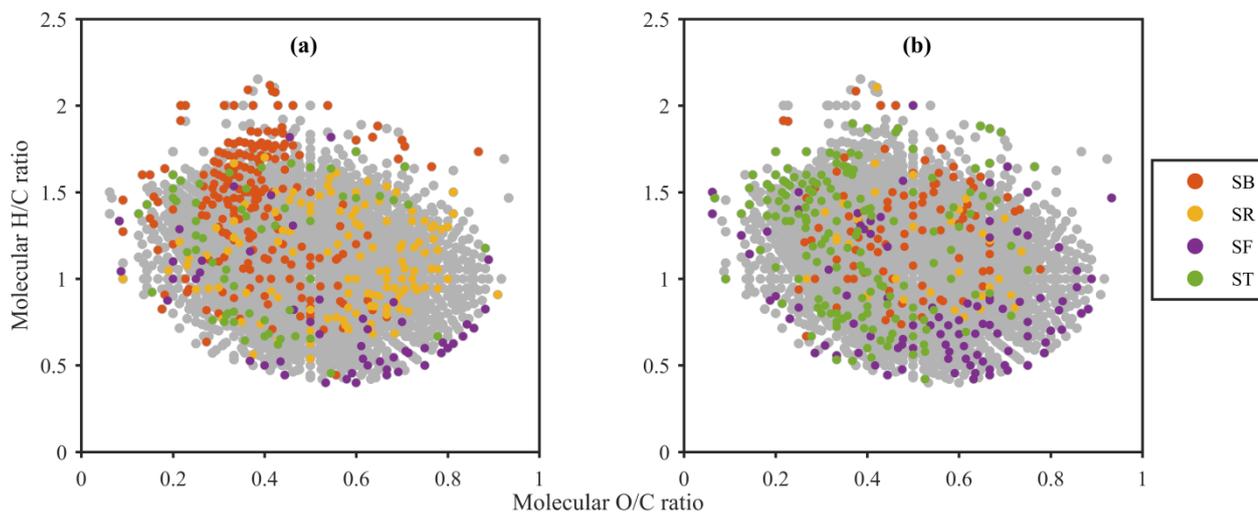


Figure 7. Unique molecular formulas during the four sampling events at stations SATURN-08 (a), and SATURN-05 (b) plotted against all formulas (grey). Formulas were only plotted when present in both process replicates. SB: spring phytoplankton bloom (01 April 2013), SR: Spring rain event 10 April 2013, SF: spring freshet (04 and 04 June 2013), ST: summer time (14 and 20 August 2013).

5

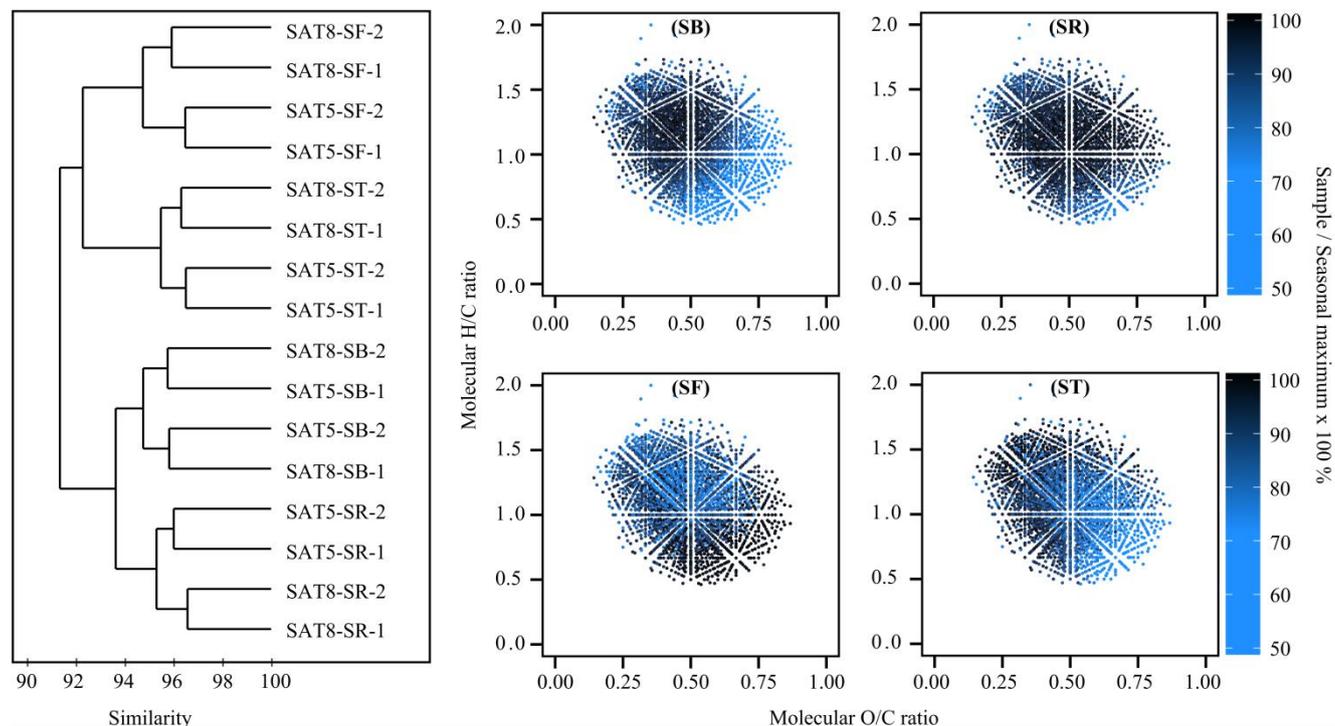


Figure 8. Seasonal changes of normalized molecular formula abundances. Left panel: Cluster analysis (based on Bray-Curtis similarity and group-average clustering) of all 16 samples (including process replicates) based on normalized molecular formula abundances (molecular formula abundance divided by maximum molecular formula abundance at respective station x 100 %). Right panel: Van Krevelen plot representing average abundance changes of individual formulas throughout the study period. SB: spring phytoplankton bloom (01 April 2013), SR: Spring rain event 10 April 2013, SF: spring freshet (04 and 04 June 2013), ST: summer time (14 and 20 August 2013). Formulas with identical H/C and O/C values are represented by their average normalized peak magnitudes.

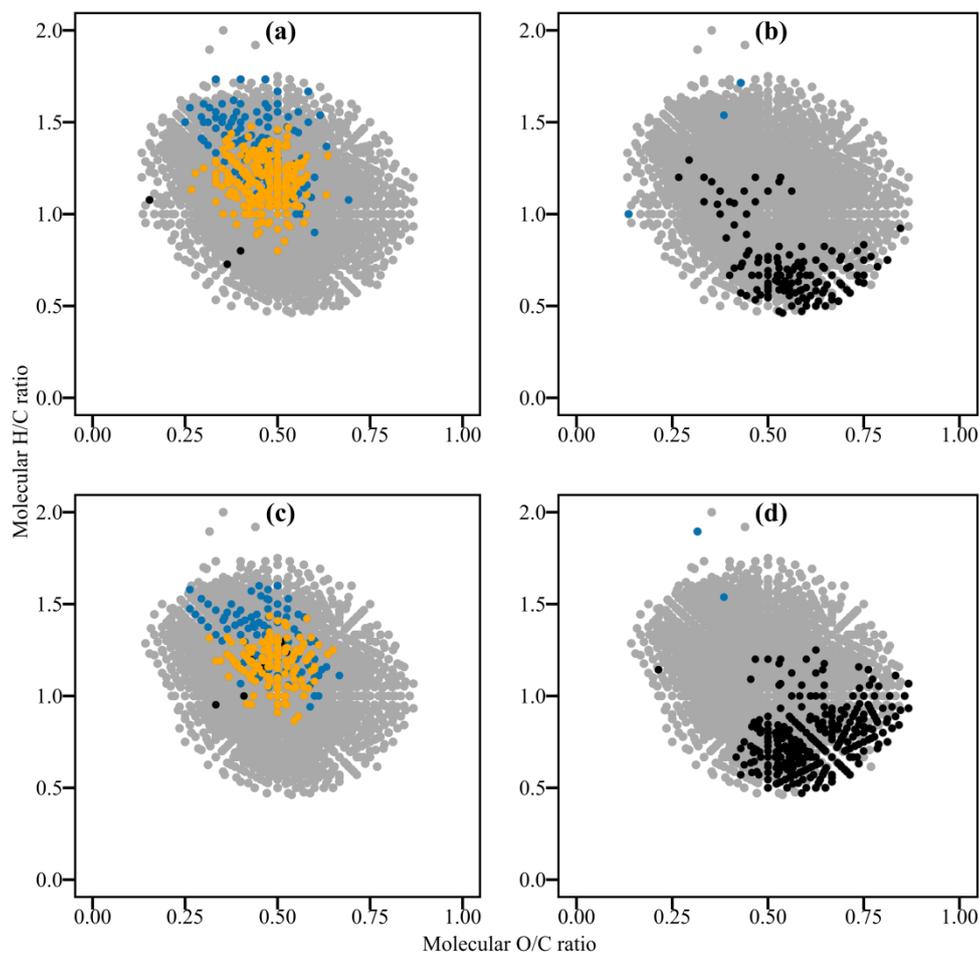
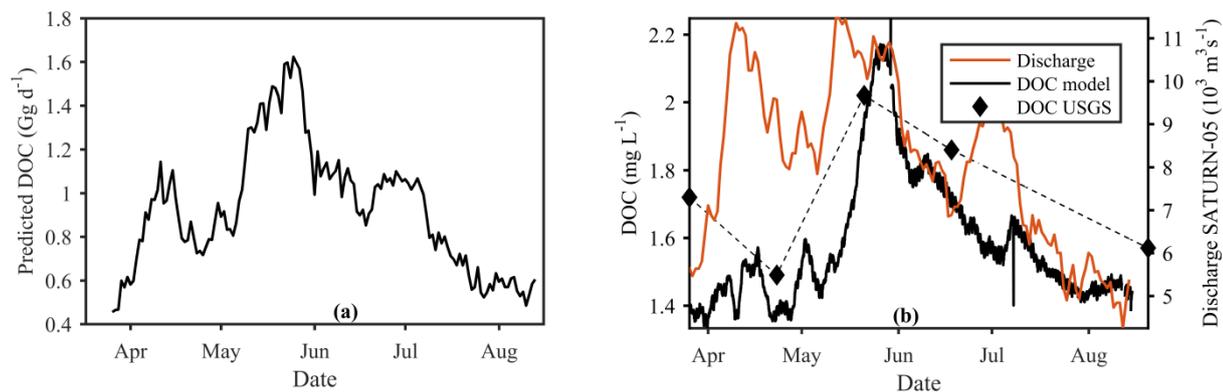


Figure 9. Van Krevelen plot of ubiquitous formulas at SATURN-08 (a-b, top panel) and SATURN-05 (c-d, bottom panel) that correlated with: (a,c) in situ chlorophyll a ($n = 343$, and $n = 218$), (b,d) absorbance at 254 nm ($n = 129$, and $n = 288$). The entirety of ubiquitous formulas is given in grey. The molecular composition is represented as follows: CHO (black), CHOS (blue), and CHON (orange). Formulas were selected with the following thresholds: $r \geq 0.8$, $p \leq 0.005$. Correlation analyses were carried out using the relative peak magnitudes of FT-ICR mass spectra and in situ sensor data at SATURN-08.

5



5 **Figure 10. DOC flux estimates (a) and DOC concentration (b) during spring and summer 2013. The presented fluxes in (a) are daily averages that were calculated with daily average river discharge at Bonneville dam (34 km upstream of SATURN-08). The uncertainty for predictions was 0.07 Gg d⁻¹. Data in (b) provides evidence of the level of detail of the high resolution DOC flux model in comparison to monthly DOC sampling in the same period.**