LIMNOLOGY and OCEANOGRAPHY

ASLO

Limnol. Oceanogr. 9999, 2021, 1–15

doi: 10.1002/lno.11862

Dynamic changes in size-fractionated dissolved organic matter composition in a seasonally ice-covered Arctic River

Hui Lin,[1] **Huacheng Xu**,[2] **Yihua Cai**,[3] **Claude Belzile**,[4] **Robie W. Macdonald**,[5,6] **Laodong Guo**[1*]

[1]School of Freshwater Sciences, University of Wisconsin-Milwaukee, Milwaukee, Wisconsin
[2]Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing, China
[3]College of Ocean and Earth Sciences, Xiamen University, Xiamen, China
[4]Institut des Sciences de la Mer de Rimouski, Université du Québec à Rimouski, Rimouski, Québec, Canada
[5]Institute of Ocean Science, Fisheries and Oceans Canada, Sidney, British Columbia, Canada
[6]Centre for Earth Observation Science (CEOS), Department of Environment and Geography, University of Manitoba, Winnipeg, Manitoba, Canada

Abstract

Arctic rivers are sensitive to climate and environmental change, but the biogeochemical response remains poorly understood. Monthly size-fractionated dissolved organic matter (DOM) samples from the lower Yukon River were characterized using UV–visible, fluorescence, and Fourier transform-infrared (FT-IR) spectroscopy techniques. The EEM-PARAFAC analysis revealed three major fluorescent DOM components, including two humic-like components (C_{480} and C_{400}) and one protein-like component (C_{310}), with their relative importance following the order of $C_{480} \geq C_{400} > C_{310}$ in the high-molecular-weight DOM (1 kDa–0.4 μm) and $C_{400} > C_{480} > C_{310}$ in the low-molecular-weight DOM pool (< 1 kDa). Transformation in DOM and change in sources were manifested in major fluorescent components and optical properties, including biological index (BIX), humification index (HIX), spectral slope ($S_{275-295}$) and specific UV absorbance at 254 nm ($SUVA_{254}$). These changes occurred within different DOM size-fractions and among ice-covered, spring freshet, and open seasons. Joint analysis of EEM and FT-IR spectra using a data fusion technique showed that humic-like DOM is mostly associated with C—H, C═C, and C—O bonds, while protein-like DOM is correlated more with C—N and N—H related structures. DOM aromaticity and the ratios of HIX to BIX and protein-like to humic-like components may be used as a compelling proxy to measure change in source waters and to infer permafrost dynamics. Our results provide insight into the seasonal variation in DOM composition for different size-fractions in the lower Yukon River, and a baseline dataset against which future changes can be understood in the context of arctic basin biogeochemical cycling.

Arctic terrestrial and aquatic ecosystems have experienced rapid climate and environmental change over the last several decades, resulting in evident impacts on hydrological cycles, permafrost dynamics, biogeochemical processes, and ecosystems as a whole. Of particular concern is the biogeochemical response of Arctic ecosystems to amplified warming, increased discharge, and accelerated permafrost degradation (McGuire et al. 2009; Abbott et al. 2016; Turetsky et al. 2020). Indeed, changes in sources, export fluxes, and chemical composition of dissolved organic matter (DOM) and other chemical species in Arctic rivers have attracted considerable attention (Frey and Smith 2005; Finlay et al. 2006; Guo et al. 2007). While recent studies have advanced our understanding of the sources, composition, and fluxes of DOM in Arctic rivers, little is known about the seasonal variations in DOM dynamics and cycling pathways especially when there is an ice cover (Shogren et al. 2020; O'Donnell et al. 2021).

The Yukon River, one of the largest northern rivers, has a prolonged seasonal ice cover and drainage basins that contain large areas of continuous and discontinuous permafrost (Brabets et al. 2000). Although there has been relatively little anthropogenic disturbance within the drainage basin itself, the position of this river just below the Arctic Circle places it

*Correspondence: guol@uwm.edu

Additional Supporting Information may be found in the online version of this article.

Author Contribution Statement: L.G. designed the study. Y.C. and C.B. collected samples and participated in sample processing and analysis. H.L. and H.X. analyzed the samples and data. H.L. and L.G. wrote the first draft. All authors contributed to data interpretation and the writing of the manuscript.

in a state of high vulnerability to amplified global warming and significant impacts have already been observed (Striegl et al. 2005; O'Donnell et al. 2021). Most of the previous studies on the dynamics of DOM in the Yukon River basin have relied on data collected during open seasons between spring freshet and early fall (Guo and Macdonald 2006; Striegl et al. 2007; Spencer et al. 2008), partly due to the extremely harsh sampling conditions during winters and partly due to cost. Although a few studies have accumulated a sparse set of winter data for organic carbon (Spencer et al. 2008; Stedmon et al. 2011; Aiken et al. 2014), detailed seasonal variations in DOM compositions/functionalities, optical properties, and PARAFAC-derived fluorescent DOM components in the Yukon River, especially during cold seasons under the ice, remain poorly observed.

In addition to bulk dissolved organic carbon (DOC), measurements of chromophoric DOM (CDOM), including UV-absorbance and fluorescence DOM (FDOM) using excitation-emission matrices (EEM), have been reported for the Yukon River (Belzile and Guo 2006; Spencer et al. 2009; Aiken et al. 2014). Over the last two decades, EEM coupled with parallel factor (PARAFAC) analysis have been widely used to characterize natural DOM in both freshwater and saltwater ecosystems (Stedmon and Bro 2008; Murphy et al. 2013; Yang et al. 2019). To the best of our knowledge, there is not a single study thus far reporting the characteristics and fluorescent components of the DOM in the lower Yukon River using these recently developed techniques. Likewise, reports characterizing CDOM and FDOM in size-fractionated DOM, such as the high-molecular-weight (HMW-) and low-molecular-weight (LMW-) DOM fractions are also scarce for the Yukon River.

In the upper Yukon River, elemental compositions (C and N) and fatty acid contents have been shown to differ significantly between the < 1 kDa LMW-DOM and the > 1 kDa HMW-DOM during ice open season from May to September (Zou et al. 2006). In the lower Yukon River, Belzile and Guo (2006) observed contrasting DOM absorbance and fluorescence signatures between the LMW- and colloidal DOM size-fractions, but these authors did not further investigate the differences using EEM-PARAFAC analysis. Specific differences in functional groups, fluorescent-DOM components, and other optical properties between the LMW- and HMW-DOM fractions, as well as the temporal variation from growing season to winter ice-covered periods, remain to be discovered.

Fluorescence spectroscopy, a highly sensitive and cost-effective method, has been widely used in the optical characterization of DOM in aquatic environments (Coble 2007; Murphy et al. 2014; Osburn et al. 2019). Based on fluorescence EEM, aquatic DOM consists predominantly of two broad groups of chemical species: protein-like and humic-like. However, there are limitations in using only EEM data given that they do not provide detailed molecular level information. Similarly, Fourier transform-infrared (FT-IR) spectroscopy is another cost-effective analytical technique for DOM characterization (Abdulla et al. 2010; Yang et al. 2015). It provides molecular level information, such as functional groups, which can be highly complementary to fluorescence spectroscopy.

Data fusion, the process of joint analysis of multiple datasets to generate more useful information than what a single dataset can provide (Acar et al. 2011), has been applied to the characterization of DOM in aquatic environments. For example, Wünsch et al. (2018) used advanced coupled matrix tensor factorization (ACMTF) to integrate two datasets of fluorescence EEM, and Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS) to improve the understanding of DOM biogeochemistry. Similarly, joint analysis of the FT-IR spectra and fluorescence EEM datasets should provide more comprehensive information about DOM dynamics.

In this study, monthly water samples from the lower Yukon River were collected to size-fractionate the bulk DOM into the LMW- (< 1 kDa) and HMW-DOM (1 kDa-0.4 μm) size-fractions using cross-flow ultrafiltration. The size-fractionated DOM samples were characterized using UV–visible absorption, fluorescence, and FT-IR spectroscopy techniques, as well as EEM-PARAFAC modeling and data fusion. Our major goals were to (1) provide the first baseline dataset on EEM-PARAFAC derived fluorescent DOM components in the lower Yukon River, (2) investigate biogeochemical characteristics of size-fractionated DOM and determine seasonal variations in DOM abundance, size distributions, optical properties, and chemical composition, and (3) apply a state-of-the-art data fusion technique (ACMTF) to elucidate relationships between fluorescent components decomposed from EEM and organic functional groups revealed by FT-IR spectra. Our results show that selected DOM optical properties have a promising potential to provide proxies for change in hydrological conditions and permafrost dynamics consequent to warming, as reflected by the composition of DOM fractions carried by the Yukon River.

Methods and materials

Study area

The Yukon River is fed by three major classes of tributaries (Fig. 1), which are commonly categorized as blackwater, clearwater, and meltwater rivers (Brabets et al. 2000; Striegl et al. 2007). Blackwater rivers, such as the Porcupine River, exhibit highly colored river water with high DOC concentrations and relatively low dissolved inorganic carbon (DIC) loadings because they drain through permafrost-dominated wetlands and peatlands. Clearwater rivers are supplied by headwater lakes with low DOC and suspended sediment concentrations. Glacier-fed rivers, such as the Tanana River, are dominated by meltwater with low DOC concentrations, but high DIC and sediment loadings. Our sampling station, which is located in the lower Yukon River (Fig. 1), is supplied by a mix of water from all major tributaries of the Yukon River and therefore integrates the varied water chemistry, vegetation cover, DOM and other biogeochemical characteristics of the drainage basin (Brabets et al. 2000).

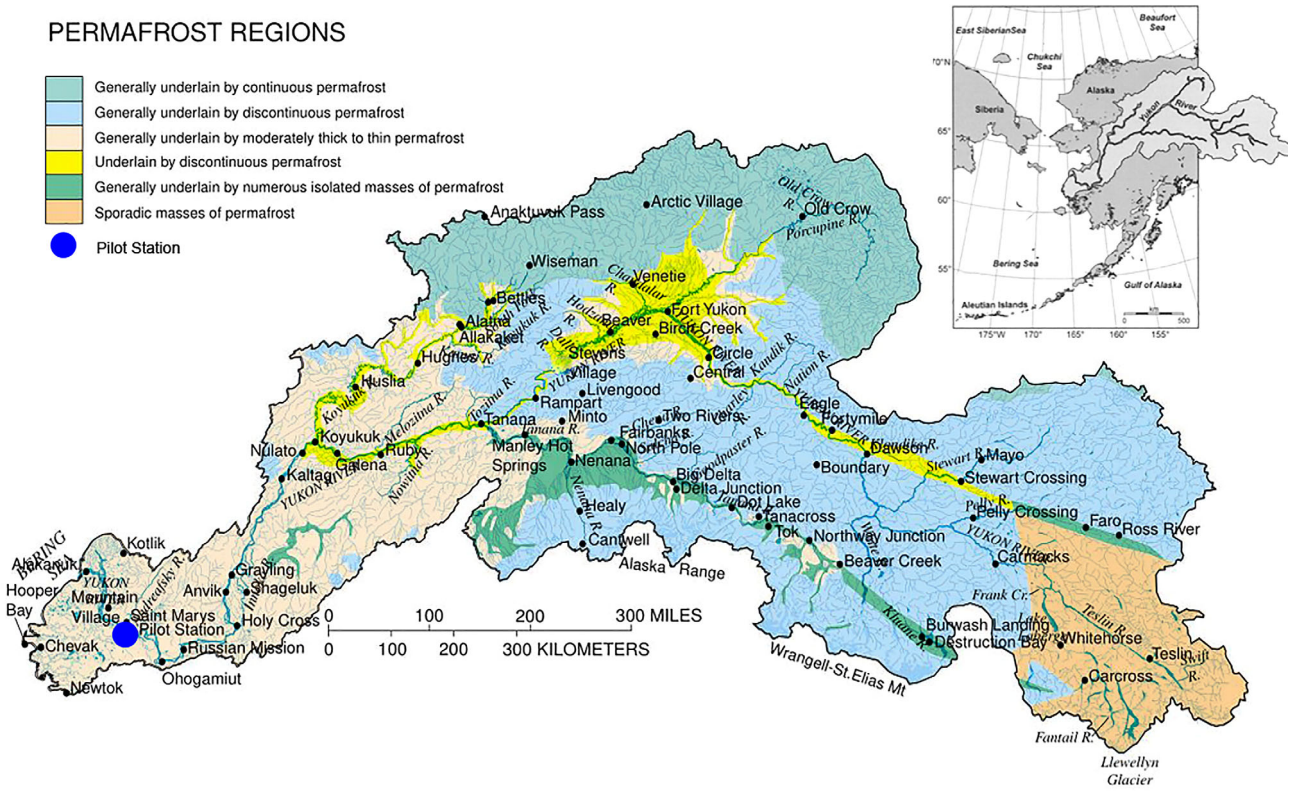

Fig. 1. Map of the Yukon River basin showing the distribution of permafrost, classified from continuous to sporadic, and the sampling location at Pilot Station, Alaska, (blue dot) in the lower Yukon River. Modified from Brabets et al. (2000) and Guo and Macdonald (2006).

Sampling procedures

Nine water samples from the lower Yukon River were collected monthly or bi-monthly for DOM size-fractionation using ultrafiltration between September 2004 and September 2005 at Pilot Station, Alaska (65°56′10″ N, 162°53′0″ W, Fig. 1). Other sampling details and measurements of water quality parameters have been described elsewhere (Guo et al. 2012) and the instantaneous discharge and selected hydrographic parameters during the sampling time period are given in Fig. S1. Briefly, using a peristaltic pump with acid-cleaned polypropylene tubing, water samples were pumped from the river through a 0.45 μm polycarbonate filter cartridge (Osmonics) into acid-cleaned 20 L HDPE containers. A manual ice auger was used to drill through the ice before sampling when the river was ice-covered between November and May. Nutrients (N, P, and Si), bulk DOC, and other hydrographic data for these water samples have been presented in Guo et al. (2012).

The < 0.45 μm filtrate was further ultrafiltered using a spiral-wound 1 kDa cartridge (regenerated cellulose, Amicon, S10Y1) to size-fractionate the bulk DOM into HMW- (1 kDa-0.45 μm) and LMW- (< 1 kDa) DOM size-fractions. A total of 40 L of pre-filtered water sample was used for ultrafiltration to isolate sufficient amounts of HMW-DOM for other chemical and isotopic characterization (e.g., Guo and Macdonald 2006; O'Connor et al. 2020). The first 20 L was ultrafiltered to quantify the abundance of HMW-DOC using the ultrafiltration permeation model and time series DOC concentrations of permeate samples (Belzile and Guo 2006; Guo and Santschi 2007). The final retentate volume was reduced to 500 mL, corresponding to a concentration factor of 80. The cleaning and calibration of the ultrafiltration system are described in previous studies (Guo and Santschi 2007). Right after ultrafiltration, the retentate or HMW-DOM samples were further diafiltrated to remove residual salts and LMW-DOM. The HMW-DOM and aliquots of the LMW-DOM and bulk DOM samples, including 9 HMW-, 9 LMW-, and 9 bulk DOM samples (a total of 27), were freeze-dried immediately after sample processing each month for further optical/fluorescence and chemical/molecular characterization.

Measurements of DOC, UV–vis absorption, and fluorescence EEM

Concentrations of DOC were measured on a TOC-Analyzer (Shimadzu TOC-L) using the high-temperature combustion method (Guo et al. 1995). Since one of our major objectives was to integrate both EEM and FT-IR datasets using ACMTF modeling, freeze-dried DOM samples were consistently used for both optical characterization and FT-IR analyses. For the measurements of UV–vis absorbance and fluorescence EEM

spectra, predetermined amounts of the freeze-dried DOM samples, including HMW-, LMW-, and bulk-DOM, were redissolved in boric acid-borate buffer solution with a pH = 8 similar to the pH range (7.2–8.4) of Yukon River waters (Guo et al. 2012) and shaken for 2 h (on a VWR Mini Shaker) followed by ultrasonication for 2 h (using a VWR Symphony Ultrasonic cleaner). Absorption spectra were obtained using a UV–visible spectrophotometer (Agilent 8453) with a 1-cm quartz cuvette. Specific UV Absorbance at 254 nm ($SUVA_{254}$), which has been used as an indicator of DOM aromaticity (Weishaar et al. 2003), was calculated as: $SUVA_{254} = A_{254}/[DOC]$, where A_{254} is the UV-absorbance at 254 nm (m^{-1}), and [DOC] is the DOC concentration (mg-C L^{-1}). The spectral slope between 275 and 295 nm ($S_{275–295}$), which is a proxy for apparent DOM molecular weight, was calculated using a nonlinear exponential model based on $a_\lambda = a_0 e^{S(\lambda-\lambda_0)}$ (Helms et al. 2008).

Fluorescence EEM were measured using a spectrofluorometer (Fluoromax-4, Horiba). Samples in a 1 cm path-length quartz cuvette were scanned at excitation wavelengths ranging from 250 to 480 nm with 5 nm intervals, and emission wavelengths ranging from 240 to 600 nm with 2 nm intervals. Humification index (HIX) is the ratio of fluorescence signals over the emission range of 435–480 nm to those over the range of 300–345 nm with excitation at 254 nm (Zsolnay et al. 1999). Biological index (BIX), an indicator of autochthonous fluorescent DOM (Huguet et al. 2009), is the ratio of fluorescence intensity at 380 nm over that at 430 nm under the excitation wavelength of 310 nm.

Fourier transform infrared spectroscopy

The freeze-dried DOM samples were further characterized using FT-IR spectroscopy (IRTracer-100, Shimadzu). To increase the reproducibility, FT-IR absorbance spectra were obtained by collecting 30 scans at each wavenumber to reduce errors and subtracted from the air background scan. These FT-IR spectra ranging from 4000 to 400 cm^{-1} had a resolution of 1 cm^{-1}. The spectra in the 4000–1800 cm^{-1} region contain one broad peak from O—H stretching, while the 900–400 cm^{-1} region contains mostly noise (Fig. S2). Therefore, only the spectral region between 1800 and 900 cm^{-1} was selected for data fusion modeling. This region covers major bands from natural organic matter, such as amide, carboxylic, ester and carbohydrate functional groups (Abdulla et al. 2010) although O—H and N—H stretching at the > 3000 cm^{-1} has been observed in treated wastewater samples (Yang et al. 2015).

PARAFAC analysis

PARAFAC analysis was conducted on all EEM data using the drEEM toolbox (v.0.5.0) in MathWorks® Matlab (R2020a) software platform (Murphy et al. 2013). Rayleigh and Raman scattering peaks were first eliminated for conformity to the linear assumptions required for PARFAC models. Fluorescence intensities were corrected for the inner-filter effect using the UV–vis absorbance spectra and calibrated into quinine sulfate equivalent (QSE). A total of 18 samples, including nine HMW-DOM and nine LMW-DOM samples, were assembled and normalized. The models were constrained to nonnegative values. Split-half validations were performed based on a randomized split of the dataset and the comparisons of Tucker Congruence Coefficient (TCC) between split models. It is worth noting that although HMW-DOM and LMW-DOM are derived from their corresponding bulk samples, marked differences in optical properties and fluorescence composition between the HMW- and LMW-DOM indicate that they are independent samples. To enhance the robustness and confidence of the PARAFAC model, comparisons between our results and known fluorescent components taken from online OpenFluor dataset (Murphy et al. 2014) were conducted, and the results are discussed.

ACMTF model

To explore the potential connections between EEM and FT-IR spectra, ACMTF analysis was conducted to analyze these spectroscopic data jointly. To potentially increase variances, EEM data from nine bulk DOM samples were incorporated into the ACMTF analysis. Before ACMTF analysis, EEM spectra were corrected, calibrated, normalized, and assembled as described above. Data of FT-IR were scaled to the maximum absorption of each FT-IR spectra. The dimension of the assembled EEM tensor and FT-IR spectra matrix are $27 \times 181 \times 47$ (sample × emission × excitation) and 27×935 (sample × wavenumber), respectively.

ACMTF analysis was performed on both EEM and FT-IR datasets using the CMTF toolbox (Acar et al. 2011), Tensor toolbox (Bader and Kolda 2007), and SNOPT toolbox (Gill et al. 2005) in Matlab. ACMTF analysis is based on three assumptions, similar to PARAFAC models, including variability, additivity, and linearity (Murphy et al. 2013). For the variability assumption, all components have distinct excitation/emission/FT-IR spectral loadings. Second, fluorescence spectroscopy and FT-IR spectroscopy are both based on the Beer–Lambert Law, which satisfies the additivity assumption. Lastly, fluorescence EEM have tri-linearity in terms of component concentrations, excitation spectra, and emission spectra. The FT-IR absorption is linearly related to FT-IR spectral loadings and its invariant spectral wavenumber.

ACMTF decomposed the EEM and FT-IR spectra into a set of trilinear fluorescence components and bilinear molecular formula components. Briefly, the ACMTF model can be expressed as:

$$x_{ijk} = \sum_{f=1}^{F} \mu_f a_{if} b_{jf} c_{kf} + \varepsilon_{ijk} \quad (1)$$

$$y_{im} = \sum_{f=1}^{F} \sigma_f a_{if} v_{mf} + \eta_{im} \quad (2)$$

where $i = 1, \ldots, I$; $j = 1, \ldots, J$; $k = 1, \ldots, K$; $f = 1, \ldots, F$; $m = 1, \ldots, M$. The x_{ijk} is the fluorescence signal intensity of sample

i measured at $\lambda_{Em} = j$ and $\lambda_{Ex} = k$. f is the underlying component (factor in statistics) with a range from 1 to F (the total number of factors). μ_f and σ_f are penalty terms, which are the differences between ACMTF and PARAFAC models. These terms determine whether one specific component, f, is shared between the EEM and FT-IR spectra (Wünsch et al. 2018). For example, in a case where the μ_f is 0.9, but the σ_f is 0.03, component f might be unshared between these datasets. a_{if}, b_{jf}, and c_{kf} are model-derived parameters representing the fluorescence maximum (F_{max}, which could be used to approximate fluorescence concentration) of fth component in sample i, excitation loadings of the fth component in $\lambda_{Em} = j$, and emission loadings of the fth component in $\lambda_{Ex} = k$, respectively. The final terms, ε_{ijk} and η_{im}, are residuals representing noise and variation uninvolved in the model. y_{im} is the FT-IR absorption of sample i measured at wavenumber m. The v_{mf} is the FT-IR spectral loading of the fth component at wavenumber m. Overall, ACMTF models two datasets using Eqs. 1 and 2, fitting the equation by minimizing the sum of squares of the residuals (ε_{ijk} and η_{im}). All models were constrained to nonnegative values. Model comparisons and validations are quantified using Tucker Congruence Coefficient (TCC) similar to those in Murphy et al. (2018).

Statistical analysis

Statistical data analyses, including t-test and F-test for comparisons of variances and correlations, were conducted with a Matlab statistics toolbox (MathWorks, R2020a). A significant level of 0.05 was selected for F-test and 0.01 for t-test.

Results

Environmental parameters in the lower Yukon River

In the lower Yukon River, the discharge during the spring freshet of 2005 (up to 33,428 m^3 s^{-1}, Fig. S1) at the downstream USGS hydrological station at Pilot Station is almost double the maximum discharge in spring 2004 (at 18,406-m^3 s^{-1}), showing high variability in river discharge between hydrological years. Water isotopic composition ($\delta^{18}O$ and δ^2H) exhibited similar patterns, with higher values during the ice-covered season due to a larger relative contribution of isotopically heavier groundwater input but lower values during snowmelt and river open seasons (Fig. S1c). Intense precipitation during 2005 (including snow and rainwater) resulted in lower stable isotope values compared with summer 2004. In 2004, [DOC] increased slightly between September and November (Fig. S1b), likely due to the exclusion of DOC from ice during freeze-up, followed by a steady decrease from 428 to 182 μM during the ice-covered season, reaching the lowest concentration in April, 2005 (Guo et al. 2012). During the spring freshet (May 2005), [DOC] increased abruptly to as high as 1682 μM, dropped rapidly after freshet, and then decreased monotonically between June and September (Fig. S1b). In contrast to [DOC], $SUVA_{370}$ showed elevated values during both ice-covered (November 2004) and ice-free (September 2005) periods with the lowest values observed between February and May 2005 (Fig. S1d).

Variations in $SUVA_{254}$ and $S_{275–295}$ between HMW- and LMW-DOM

For the > 1 kDa HMW-DOM pool, the $SUVA_{254}$ value, an indicator of DOM aromaticity (Weishaar et al. 2003), varied from 3.30 ± 0.60 L mg-C^{-1} m^{-1} in the ice-covered season (with a water temperature at or below 0°C) to 3.95 ± 0.36 L mg-C^{-1} m^{-1} in the ice-free season (post-May 2005, Fig. 2a). The abrupt increase in $SUVA_{254}$ values during the freshet period is concurrent with the peak [DOC] during spring freshet with heavy loads of terrestrial DOM. After the spring freshet, $SUVA_{254}$ declined from June to September, similar to the trend of [DOC], but to a relatively smaller extent. On average, the $SUVA_{254}$ value of the HMW-DOM is higher during the summer growing season than the ice-covered season ($p < 0.05$), indicating overall higher aromaticity in the HMW-DOM pool during summer. In the ice-covered season, similar to changes in [DOC] (Fig. S1b), the initial river-ice formation and thus the exclusion of DOM from ice (Belzile et al. 2002) also resulted in a slight increase of the HMW-DOM's $SUVA_{254}$ value in November, followed by a general decrease between November and April (Fig. 2a).

For the LMW-DOM pool (Fig. 2a), $SUVA_{254}$ values are relatively less variable compared to those of the HMW-DOM ($p < 0.025$). However, they showed a general decrease during the ice-covered season from 2.57 L mg-C^{-1} m^{-1} in September 2004 to 1.86 ± 0.03 L mg-C^{-1} m^{-1} during December–April, followed by an increase during the spring freshet with a $SUVA_{254}$ value of 2.12 L mg-C^{-1} m^{-1} in May 2005. After snowmelt (May 2005), $SUVA_{254}$ values changed little, ranging from 2.12 to 2.19 L mg-C^{-1} m^{-1} with an average of 2.15 ± 0.05 L mg-C^{-1} m^{-1} in the lower Yukon River.

Overall, $SUVA_{254}$ values of the LMW-DOM samples are systematically lower than those of the HMW-DOM (2.06 ± 0.23 vs. 3.31 ± 0.49 L mg-C^{-1} m^{-1}, $p < 0.01$), indicating that the HMW-DOM pool is of higher aromaticity than the LMW-DOM pool. $SUVA_{254}$ values observed here are consistent with previous measurements for the bulk DOM pool at Pilot Station (Striegl et al. 2007; Spencer et al. 2008; Aiken et al. 2014). Compared to $SUVA_{254}$, the values of $SUVA_{370}$ showed greater temporal variability in the HMW-DOM pool (Fig. S3), although both the HMW- and LMW-DOM pools had similar trends in their variability and were highly correlated with one another ($r^2 = 0.979$, $p < 0.01$).

Spectral slope ($S_{275–295}$), a proxy for DOM apparent molecular weight (Helms et al. 2008), showed a general temporal change opposite to the pattern of $SUVA_{254}$ (Fig. 2b). For both HMW- and LMW-DOM, the $S_{275–295}$ increased gradually when $SUVA_{254}$ declined during the ice-covered season, suggesting a general decrease in the apparent DOM molecular weight with decreasing aromaticity. In particular, the $SUVA_{254}$ of the HMW-DOM sharply increased during spring freshet, accompanied by a simultaneous decline in the $S_{275–295}$ values of the

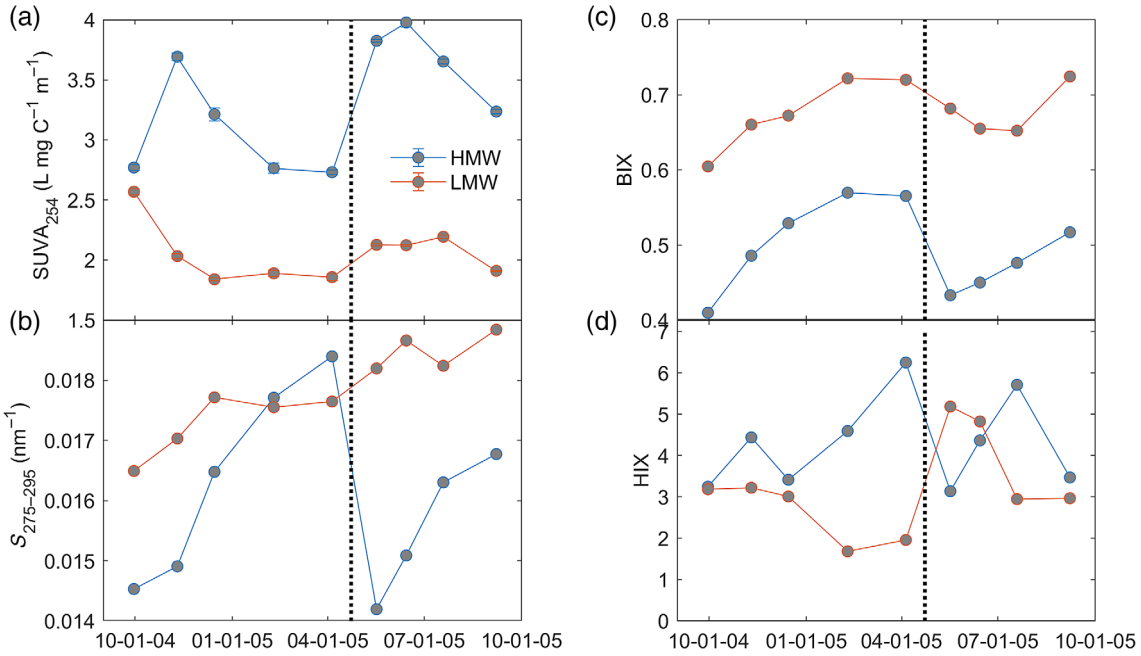

Fig. 2. Temporal variations in (**a**) $SUVA_{254}$, (**b**) $S_{275–295}$, (**c**) BIX, and (**d**) HIX between HMW-DOM and LMW-DOM pools in the lower Yukon River with dashed lines denoting the beginning of spring freshet.

HMW-DOM (Fig. 2b). After freshet, the decrease in $SUVA_{254}$ of the HMW-DOM pool was accompanied by an increase in its $S_{275–295}$ values, indicating a decrease in the DOM molecular weight. These relationships suggest that terrestrial HMW-DOM during spring freshet in the lower Yukon River was of both higher aromaticity and higher molecular weight, while DOM under the ice was of lower $SUVA_{254}$ with lower molecular weight or higher spectral slope values.

Compared with the HMW-DOM, values of $S_{275–295}$ in the LMW-DOM pool changed little ($p < 0.05$) but showed a general increase over the entire sampling period from September 2004 to September 2005 (Fig. 2b). $SUVA_{254}$ in both HMW- and LMW-DOM decreased to the lowest values during the ice-covered season, whereas the corresponding $S_{275–295}$ values increased, suggesting that the average DOM molecular weight decreased in both HMW- and LMW-DOM pools during winter under the ice. After the spring freshet, the $S_{275–295}$ values in the LMW-DOM pool remained relatively high but stable.

Seasonal variations in fluorescence indices between HMW- and LMW-DOM

Values of BIX in the HMW-DOM pool increased consistently during the ice-covered season from September 2004 up until May 2005 (Fig. 2c), indicating an increase in autochthonous DOM sources under ice. However, during the ice melt season, the BIX dropped abruptly from 0.57 in April under the ice to 0.43 in May (spring freshet). Values of BIX in the LMW-DOM pool also had a variation trend similar to the HMW-DOM, increasing from September (0.60) to early April (0.72) during the ice-covered season, declining slightly during the spring freshet with a slight decrease from May (0.68) to July (0.65), and an elevated value in September 2005 (Fig. 2c). During the spring freshet, values of BIX in the HMW-DOM pool dropped to their lowest value (0.42), while the LMW-DOM pool always exhibited a higher BIX value than the HMW-DOM pool ($p < 0.01$).

Similar to BIX, the values of HIX in the HMW-DOM pool also increased generally from September 2004 to early May 2005 during the ice-covered season (Fig. 2d). The highest HIX value of the HMW-DOM was found in April under the ice, followed in mid-May by the lowest HIX during the spring freshet. For the LMW-DOM pool, the HIX values during the winter season were, on average, lower than those during the open season ($p < 0.05$, Fig. 2d). HIX reached its maximum during spring freshet and then decreased monotonically from May to September 2005. Overall, HIX values in the HMW-DOM pool were slightly higher than those in the LMW-DOM (4.30 ± 1.11 vs. 3.22 ± 1.15, $p < 0.10$), consistent with the difference in $SUVA_{254}$ between HMW- and LMW-DOM pools in the lower Yukon River.

Major functionalities of DOM derived from FT-IR spectra

FT-IR spectroscopy was applied to identify major functional groups in size-fractionated DOM and their seasonal variations in the lower Yukon River. Samples from February and May were used to represent the ice-covered and ice-free seasons, respectively, and examples of FT-IR spectra of both HMW- and LMW-DOM samples, along with examples of their EEM

spectra, are shown in Fig. S4. There are three main peaks with locations at 1620–1640 cm^{-1}, 1400–1410 cm^{-1}, and 1040–1120 cm^{-1}, respectively. These peaks correspond to functional groups of C=O stretching of amide, quinone, or ketones at 1620–1640 cm^{-1}, O—H deformation, C—O stretching of phenolic OH^- or COO^- asymmetric stretching at 1400–1410 cm^{-1}, and C—O stretching of polysaccharide or polysaccharide-like substances at 1040–1120 cm^{-1} (Abdulla et al. 2010). In general, major functional groups are broadly similar, especially in the HMW-DOM between the winter and spring-freshet samples. For the LMW-DOM, the absorbance for each major functional group was lower in the winter sample compared to the spring-freshet sample (Fig. S4). In addition, specific peak locations in wavenumbers between 1040–1120 cm^{-1} were slightly different between the HMW-DOM and the LMW-DOM pools, with a lower wavenumber in the LMW-DOM samples (Fig. S4).

Major fluorescent DOM components derived from EEM-PARAFAC analysis

Examples of EEM spectra of size-fractionated HMW- and LMW-DOM samples are shown in Fig. S4. Based on all EEM data, PARAFAC analysis revealed three major fluorescent DOM components (Fig. S5), including two humic-like components (C_{480} and C_{400}) and one protein-like component (C_{310}), whose characteristics are given in Table S1. Component C_{480} (Ex/Em = 250/480 nm) is one of the most widely distributed humic-like fluorescent components in aquatic environments

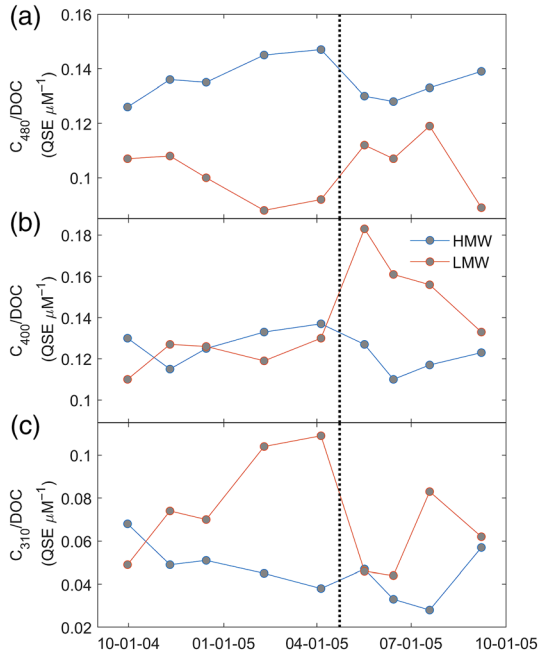

Fig. 3. Temporal variations in DOC-normalized major fluorescent components, including C_{480}/DOC (**a**), C_{400}/DOC (**b**), and C_{310}/DOC (**c**) in both HMW- and LMW-DOM pools (dashed lines denote the beginning of spring freshet).

(Zhou et al. 2013; Wünsch et al. 2019; Lin and Guo 2020). Its EEM contour covers Peak A and Peak C areas defined in Coble (2007), which are UVC and UVA humic fluorophores. Component C_{400} (Ex/Em = 260/400 nm) is the second most abundant fluorescent DOM component, with characteristics of Peak M and part of Peak A (Wünsch et al. 2019). This C_{400} component was previously recognized as a marine-derived humic-like fluorophore (Coble 2007), but has recently been found widely distributed in freshwater environments (DeVilbiss et al. 2016; Zhou et al. 2016). Component C_{310} (Ex/Em = 275/310 nm) is a protein-like component covering tryptophan-like Peak T (Ex/Em = 275/340 nm) and tyrosine-like Peak B (Ex/Em = 275/305 nm). Component C_{310} was found to be strongly related to microbial activities and autochthonous DOM production (Coble 2007; Zhao et al. 2017).

Online comparisons to the OpenFluor dataset (Murphy et al. 2014) show that these three fluorescent DOM components are similar to PARAFAC-derived components identified in previous studies ($TCC_{ex,em} > 0.95$). For example, component C_{480} has 64 pairs of match-ups, including components found in fresh waters to salt waters in coastal seas (e.g., Gonçalves-Araujo et al. 2016; Osburn et al. 2016). Within the 64 model pairs, $TCC_{ex,em}$ for 5 pairs are larger than 0.99, 28 pairs between 0.97 and 0.99, and 31 pairs between 0.95 and 0.97. The five highly similar pairs include samples from drinking water (Shutova et al. 2014), wetlands (Yamashita et al. 2010), and Japan Sea (Tanaka et al. 2014). Match-ups for components C_{400} and C_{310} are also found in 58 and 27 previous studies, respectively, from both natural waters (Yamashita et al. 2010; Tanaka et al. 2014; Kulkarni et al. 2018) and wastewaters/drinking waters (Shutova et al. 2014). However, both components did not seem to exhibit a high matchup (i.e., $TCC_{ex,em} > 0.99$) with any previous reported components in the OpenFluor database. Apparently, the higher the threshold of $TCC_{ex,em}$, the stronger the support for the identification of fluorescent components between different models (Wünsch et al. 2019). In fact, samples in the OpenFluor database are from different aquatic environments, and data were obtained under different instrumental specifications (such as, sensitivity, signal/noise ratio, etc.) and handling procedures. Thus, the probability is likely low for $TCC_{ex,em}$ to reach the highest values during model comparisons between different studies; nevertheless, a stringent $TCC_{ex,em}$ criterion is highly recommended for PAFARAC model validations.

To compare these DOM components on an intensive-property basis, fluorescence intensities (quinine sulfate equivalent, QSE) of each fluorescent component were normalized to DOC concentration (Fig. 3). The DOC-normalized C_{480} (C_{480}/DOC) in the HMW-DOM (Fig. 3a) increased in general from September 2004 to April 2005, decreased during the spring freshet, and then increased again from June to September 2005. In contrast, the C_{480}/DOC in the LMW-DOM (Fig. 3a) decreased during the winter season between September 2004

and April/May 2005, but increased during the spring freshet, and decreased again in September 2005. The C_{400}/DOC (Fig. 3b) was less variable in the HMW-DOM over the entire sampling period but was elevated during spring freshet and then decreased from May to September 2005 in the LMW-DOM pool ($p < 0.01$). The seasonal variation in the C_{310}/DOC (Fig. 3c) differs distinctly between the HMW- and LMW-DOM pools. For example, this DOC-normalized protein-like component, C_{310}/DOC, in the HMW-DOM pool generally decreased over the whole sampling time period except for a slightly higher value during ice opening and an elevated value in September 2005, while values in the LMW-DOM increased consistently below the ice, attaining its lowest values during spring freshet (May and June), and increasing thereafter in July and September.

To compare the relative contribution of each component to the total fluorescence, percentages were calculated as $C\% = C_i / \Sigma C_i \times 100$, where C_i is the F_{max} of the ith component derived from EEM-PARFAC analysis. Both C_{480}% and C_{400}% are approximately equally dominant fluorescence components, comprising $44.4 \pm 2.7\%$ and $40.6 \pm 1.5\%$ in the HMW-DOM pool and $33.0 \pm 3.8\%$ and $44.3 \pm 5.4\%$ in the LMW-DOM pool, respectively, while C_{310} contributes a smaller component, making up $15.0 \pm 3.4\%$ of the fluorescence in the HMW-DOM and $22.8 \pm 7.0\%$ in the LMW-DOM (Fig. S6). The percentage of each component varied less in the HMW-DOM pool than in the more dynamic LMW-DOM pool. For example, in the LMW-DOM pool, C_{480}% decreased consistently, whereas C_{310}% increased from September to April, while C_{400}% changed little. These compositional changes resulted in a large increase in the C_{310}% in the LMW-DOM during winter (Fig. S6). During the ice-free season, the C_{480}% in the LMW-DOM pool changed little, and the C_{400}% became higher or predominated, while the C_{310}% was lower than it was in the winter season. The main implication of these data in terms of monitoring is that dynamic changes occur between ice-covered and ice-free seasons in the DOM components for both HMW- and LMW-DOM pools, and these two pools respond differently.

Integrating EEM and FT-IR datasets using data fusion

The ACMTF modeling on both EEM and FT-IR datasets identified three model-derived components, C_{480}, C_{400} and C_{310}, named after their characteristic emission wavelengths (Table 1). Their scores, excitation loadings, and emission loadings are highly congruent with those generated from the EEM-PARAFAC model (Fig. S7). $TCC_{ex,em}$ values for C_{480}, C_{400} and C_{310} are 0.9993, 0.9998, and 0.9999, respectively, indicating that both the ACMTF model and the EEM-PARAFAC model highly resemble each other (Fig. S7). Online comparisons between our model-derived components and those previously identified in the OpenFluor database also show identical results ($TCC_{ex,em} > 0.95$), further validating the ACMTF and PARAFAC models.

Similar to those derived from EEM-PARAFAC analysis, the ACMTF model-derived C_{480} and C_{400} are humic-like components, while the C_{310} is a protein-like component (Table 1). In addition, both C_{480} and C_{400} fluorescent components are associated with the functional groups that are composed of C—O bonds, C—H bonds, and C=C double bond (Fig. 4). The protein-like C_{310} component matches two unique nitrogen-related bonds, including C—N bending and N—H bending. These results imply that humic-like DOM components represent mostly carbon-related bonds, such as C—O, C—H, and C=C, while the protein-like fluorescent component comprises more nitrogen-related molecular structures, such as C—N and N—H (Fig. 4). Similar results have been reported by Maie et al. (2007), in which a protein-like fluorophore (Peak T) was positively correlated with nitrogen concentration, and Wünsch et al. (2018), who found that C_{310} (similar component with the emission maximum at 300 nm) contained the highest N/C ratio among all six DOM components.

Discussion

Details on ACMTF model validations

The validation of the ACMTF model includes explained variance, the randomness of residuals, chemical coherence of component loadings, and reproducibility. Three models, including a two-component, a three-component, and a four-component model, were tested. The two-component model explained 95.66% and 96.67% of total variances for EEM and FT-IR spectra, respectively, which was considered to be a low explanation of total variance. The three-component model (98.69% EEM, 96.76% FTIR) and four-component model (99.21% EEM, 96.82% FT-IR) have very close explained percentages of total variances, which also implies that increasing a model's component number from three to four does not significantly improve the fit for the FT-IR spectra.

As shown in Fig. S8, the three-component and the four-component models for EEM have similar sum of squared error (SSE) values and distributions, and both of them have much smaller SSE compared to the two-component model. For the FT-IR spectra, however, increasing the component number in the ACMTF models did not yield a smaller SSE. As for the distributions of residuals between all three models (Fig. S9), the EEM contour from the two-component model poorly resembled the original EEM contour, and most of the residuals are larger than $\pm$ 10% of the original signals. Residuals from both the three-component and four-component models are smaller and are within $\pm$ 10% of the original signals, with major losses in the UVA areas (Ex 300–400 nm, Em > 400 nm) and major gains in the UVC areas (Ex 250–300 nm), having residuals mostly around Ex = 250 nm. For the FT-IR spectra (Fig. S10), the residuals of all three models are within $\pm$ 5%. The negative residuals are mostly in the regions between 1700–1650 cm^{-1} and 1000–900 cm^{-1}, representing double-bond stretching (majorly C=C and C=O) and bending (C=C), respectively. On

Table 1. Characteristics and properties of the three ACMTF model-derived components. μ_f and σ_f are factors representing the sharedness of components for EEM and FT-IR spectra, respectively.

Component	C_{480}	C_{400}	C_{310}
Ex/Em (nm)	250/480	260, 310/400	275/310
μ_f/σ_f	$20.35 \pm 1.03/9.36 \pm 0.65$	$21.28 \pm 1.08/9.15 \pm 0.67$	$21.63 \pm 0.96/9.20 \pm 0.68$
Peaks in EEM	A and C	A and M	T and B
Peaks in FT-IR (cm^{-1})	1060, 1230, 1400, 1470	1085, 1190, 1400, 1650	1114, 1417, 1621
Major bonds	C—H bending, C—O stretching	C—H bending, C—O stretching, C=C stretching	C—N stretching, C—H bending, N—H bending

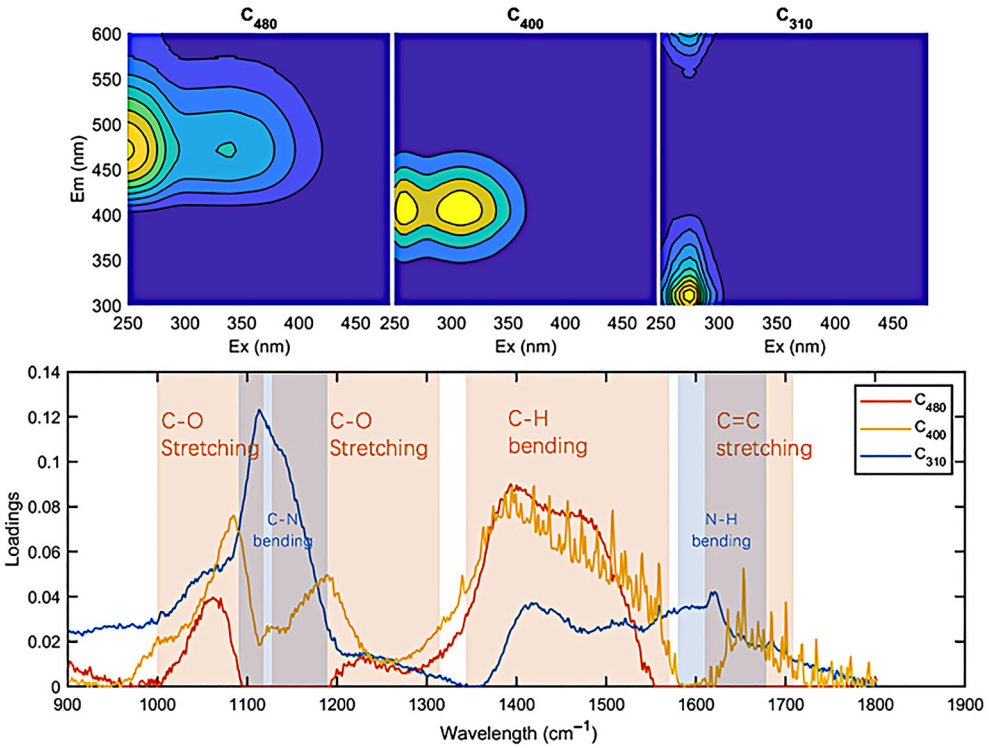

Fig. 4. Integrated EEM fluorescent components and FT-IR spectra using the data fusion technique, showing major functional groups associated with each fluorescent component (C_{480}, C_{400}, and C_{310}).

the other hand, the positive residuals are mostly within the 1650–1000 cm^{-1} area, especially the 1500–1400 cm^{-1} and 1100–1000 cm^{-1}, representing bonds with C—H bending and C—O stretching, respectively. Overall, the model seemed to miss small signals of double bonds and added some single-bond signals, but the model results are all within ± 5%, which is similar to errors associated with most analytical methods.

Compared to the three-component model, the four-component model seems to perform better in residuals, errors, and the percentage of explained variances. However, the most important reason for adopting the three-component model is that some derived FT-IR spectra in the four-component model exhibited several unexpected sharp peaks. Thus, a three-component ACMTF model was adopted as more chemically realistic even though the explained variance was slightly lower compared to the EEM-PARAFAC model where > 99% of the variance can be explained.

Uniqueness is validated using random initialization validation, which was run 100 times to obtain an optimized model least square error (Fig. S11). Loadings of excitation, emission and FTIR spectra are compared in Figs. S12 and S13. Due to the constraint of computational resources, the criterion for the convergence is set to 10^{-3}. The average λ and σ values for the three model-derived components with standard deviations (Table 1 and Fig. S11) indicate that all components are shared.

The penalty terms μ_f and σ_f were used to gauge whether a specific component is shared between the EEM and FT-IR datasets in the ACMTF model. Theoretically, values of both μ_f

and σ_f should be either 0 or 1, representing unshared or shared components, respectively. However, when the model is applied on field data instead of a randomly-distributed experimental dataset, the μ_f and σ_f are influenced by several distinct factors or parameters. For example, fluorescence quantum yield and ionization efficiency have been identified as potential factors controlling the sharedness of model-derived components between EEM and FT-ICR-MS datasets (Wünsch et al. 2018), while centering across the subject mode could have significant effects on these terms of a model integrating multiple instrument datasets (Acar et al. 2019). Thus, the μ_f and σ_f in this case could be regarded as weights or contribution of one component to the total variances to EEM and FT-IR spectra. As shown in Table 1, weights of all components for EEM (μ_f) are significantly higher ($p < 0.01$) than those for FT-IR spectra (σ_f), indicating that FT-IR spectra are potentially unshared (Fig. S11). However, considering that the percentage of total explained variances for FT-IR spectra is larger than 96%, lower contributions from FT-IR spectra (σ_f) might result from a different instrumental response since fluorescence is generated from excited electrons, while FT-IR spectra are based on rotational and/or vibration energy of chemical bonds. Similarly, Wünsch et al. (2018) found that C_{310} had significantly different component-weights for EEM and FT-ICR-MS and was regarded as a shared component. There is likely a small possibility of unsharedness between EEM and FT-IR spectra, which should be further evaluated using larger datasets in future studies. In addition, the more chemically meaningful results also render a strong support for the sharedness of C_{310} between EEM and FTIR spectra. Overall, although values of the penalty terms, μ_f and σ_f, are significantly different, all three components are regarded as shared components.

Dynamic changes in composition, source and sink of DOM across ice-covered and ice-free seasons

Arctic terrestrial ecosystems have undergone the effects of amplified warming, which have been manifested in the Arctic by changes including increasing river discharge, accelerated permafrost degradation, and changes in sources and fluxes of fluvial organic carbon (McGuire et al. 2009; Schuur et al. 2015). Despite the importance of such changes, the detailed biogeochemical responses of Arctic rivers to climate and environmental change remains poorly characterized (Pokrovsky et al. 2011; Shogren et al. 2020; O'Donnell et al. 2021). Arctic rivers generally have a strong seasonality in discharge and DOM fluxes compared to other world rivers (Finlay et al. 2006; Guo et al. 2012; Shogren et al. 2020). Time-series sampling that includes all seasons is thus indispensable to a better understanding of river DOM biogeochemistry and a more robust way to design long-term monitoring. However, sampling and characterization of DOM during ice-covered seasons remain scarce, due partly to the difficulty of sampling during a prolonged ice-covered season and harsh weather conditions during winter, and partly to the remoteness of most locations. The results of the year-long time-series study, presented here, clearly elucidate a dynamic change in DOM aromaticity, composition, fluorescent components, and optical properties during the transition from low productivity, ice-covered conditions, to spring freshet and higher productivity seasons. Given the resolution implied by nine samples collected monthly/bimonthly during the 2004–2005 sampling year, the data set presented here will likely not capture events of short duration, especially during spring freshet. Future studies with more frequent sampling, producing higher temporal resolution during periods of seasonal change are needed.

Other than changes in water and DOM sources, we speculate that physicochemical and microbial processes under the ice might significantly change the molecular structure of DOM and transform aromatic components to less aromatic DOM (Lee et al. 2018; Murphy et al. 2018), resulting in both a aromaticity decrease and an increase in $S_{275–295}$ or decrease in DOM molecular weight (Helms et al. 2008). Bacterial transformation and production would also induce higher and increased BIX values during the ice-covered period (Frenette et al. 2008; Huguet et al. 2009; Birdwell and Engel 2010). The concurrently high HIX and BIX values are seemingly contradictory, but the higher HIX values in the HMW-DOM pool under the ice may signal a shift in DOM sources from mostly surface soils during the open seasons to increasing groundwater and deeper soil during winter (Striegl et al. 2007; O'Donnell et al. 2010) compounded by the preservation and accumulation of humic-like DOM components under the ice due to their relative resistance to biological degradation (Chen and Jaffé 2016; Xu and Guo 2018). In addition, the assembly of less aromatic LMW-DOM into HMW-DOM components (Xu and Guo 2018) may be another factor contributing to the $SUVA_{254}$ decrease in the HMW-DOM, as supported by the changes in spectral slope or apparent DOM molecular weight (Fig. 2b).

During the spring freshet, BIX in the HMW-DOM pool dropped to its lowest value (0.42), while $SUVA_{254}$ and HIX increased abruptly, owing to the increased contribution from terrestrial DOM input resulting in a decrease in the proportion of autochthonous DOM. Interestingly, the LMW-DOM pool always exhibited a higher BIX value than the HMW-DOM pool ($p < 0.01$), indicating the former LMW-DOM pool contained a relatively higher proportion of autochthonous components in this terrestrial DOM-dominated river, probably due to the preferential decomposition of autochthonous HMW-DOM (Benner and Amon 2015) and the transformation and production of DOM from HMW- to LMW-DOM during decomposition either under the ice or from a deeper active layer. After spring freshet, the decrease in $SUVA_{254}$ may reflect increased photodegradation due to decrease in turbidity and prolonged sun irradiation and the change in DOM sources/composition due to a deepened active layer, permitting leaching of older soils (Guo and Macdonald 2006; Zou et al. 2006; Aiken et al. 2014).

Compared to other aquatic ecosystems, DOM from the Yukon River is generally characterized as having intermediate to high HIX (between 1 and 7) and low BIX values (< 0.6–0.8), consistent with a terrestrial DOM-dominated river system. For example, DOM from the Milwaukee River and Green Bay, both in Wisconsin and somewhat hypereutrophic, had medium HIX (< 8) and BIX (0.55–0.8) values (DeVilbiss et al. 2016; Teber 2016), whereas DOM from the Laurentian Great Lakes with predominantly autochthonous DOM (Zhou et al. 2016) and Veterans Park Lagoon, a highly eutrophic ecosystem, both were characterized by lower HIX (< 4), especially the open Great Lakes samples, but high BIX (> 0.8) values (Fig. 5). There is generally an inverse relationship between HIX and BIX, with a nonlinear transition along a trophic gradient from high HIX-low BIX in terrestrial DOM-dominated river systems to medium BIX (with a wide range in HIX) in hypereutrophic ecosystems, to low HIX-high BIX values in eutrophic, mostly autochthonous ecosystems (Fig. 5). Thus, a tandem plot of both HIX and BIX indices and/or their ratios may be used as a compelling proxy to illustrate relative changes in DOM sources and trophic states as responses to a changing climate.

Heterogeneity of DOM in properties and source between HMW- and LMW-DOM pools

Natural DOM has been shown to be highly heterogeneous in chemical properties, composition, and reactivities (Benner and Amon 2015; Xu and Guo 2017). In the upper Yukon River, the HMW-DOM was found exclusively contemporary while the LMW-DOM had an average ^{14}C age of 976 ± 212 y-BP during river open seasons (Guo and Macdonald 2006). In addition, fatty acids and biomarker composition were also distinctly different between the HMW- and LMW-DOM pools (Zou et al. 2006). In the lower Yukon River, the HMW-DOM is mostly associated with aromatic DOM compounds with much higher $SUVA_{254}$ values compared to its LMW-DOM counterpart (Fig. 2a), deriving from either high vascular plant-derived DOM or leaching of surface organic-rich soil layers (Guo et al. 2007; Striegl et al. 2007). Similarly, higher $SUVA_{254}$ values and higher contents of hydrophobic organic acid were found for the HMW-DOM pool in the lower Mississippi River (Cai et al. 2015). During the cold season, the decrease of $SUVA_{254}$ in the HMW-DOM (November 2004–April 2005) suggests a change in DOM sources: a decrease in surface soil leached DOM and increase in groundwater DOM, consistent with increased ^{14}C ages between November and April (Aiken et al. 2014; Guo and Lin 2017).

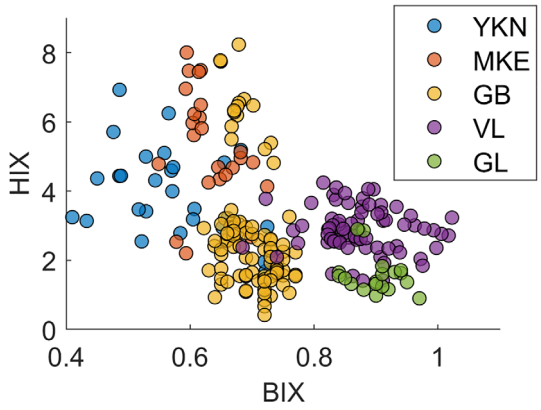

Fig. 5. A scatter plot between HIX and BIX combining data from the Yukon River (YKN), the Milwaukee River (MKE, data from Teber 2016), Green Bay (GB, data from DeVilbiss et al. 2016), the Veterans Park Lagoon (VL), and the Laurentian Great Lakes (GL, data from Zhou et al. 2016).

In the lower Yukon River, the HMW-DOM is dominated by terrigenous organic components. In contrast, the LMW-DOM is mostly derived from microbial activity and degraded organic matter and has a higher BIX and protein-like component (C_{310}) during the ice-covered season (Figs. 2c, 3c). River water below the ice is generally much warmer than the air temperature during winter, and not so different from water temperature in summer, which promotes microbial activity during early spring. Psychrophilic bacteria can thrive in such conditions and contribute to the production of LMW-DOM through DOM transformation from either HMW-DOM or particulate organic matter (Deming 2002). Indeed, the LMW-DOM in Yukon River water contained abundant protein-like DOM components during the late winter season (Fig. 3), which may have a priming effect on the degradation of terrigenous DOM in open seasons although the normalized contents of these protein-like components are lower during spring freshet. Later in the summer season, both LMW- and HMW-DOM had a higher abundance of the protein-like C_{310} component (Fig. 3c) due to in situ production and transformation processes. Our results show that DOM characteristics of the LMW-DOM and HMW-DOM pools are distinct from one another, attesting to the need to characterize not just the bulk DOM, but also the DOM associated with different molecular weight- or size-fractions.

Previous studies have shown that DOM during the spring freshet is mostly terrigenous with high bio-lability. For example, Spencer et al. (2008) reported a maximum concentration of DOM in the Yukon River during spring freshet with high aromaticity and biological lability. Holmes et al. (2008) also reported higher DOM biological reactivity during spring freshet in Alaskan Arctic rivers, with a loss of labile DOM as high as 20%–40% in samples incubated for 90 d. This rate of loss is much higher than the actual loss in DOM in the Yukon River during summer. High biological lability of DOM during spring freshet is likely associated with its contemporary or young ^{14}C age in Arctic rivers (Guo and Macdonald 2006; Raymond et al. 2007). However, old DOC released from ancient permafrost in Arctic River basins can also be labile (Vonk et al. 2013; Gao et al. 2018). Together, these observations imply that natural DOM is highly heterogeneous and contains a variety of components with diverse functional

groups distributed across a wide range of molecular size-fractions. Furthermore, DOM of varied composition and molecular size would likely have varying degradation potentials or varying photochemical and biological reactivities (Benner and Amon 2015; Xu and Guo 2018; Yang et al. 2019). Thus, understanding changes in DOM sources, composition, optical properties, and molecular weight distributions is essential to constrain the response of DOM biogeochemistry to seasonal processing and to understand what drives long-term environmental changes in Arctic River basins.

Implications

The data presented here show that dynamic changes in DOM composition and spectral properties occur throughout the year (ice covered, freshet, open season) in the lower Yukon River, and such changes differ among HMW- and LMW-DOM size fractions. Physicochemical and microbial processes, changes in sources of river water (especially groundwater vs. surface runoff), climate factors (e.g., temperature, precipitation, hydrology), and permafrost dynamics likely all play a role in regulating river DOM quantity and composition, and the response of river biogeochemistry to climate and environmental change. The results presented here provide a baseline dataset against which future changes in the context of a warming climate may be assessed. We strongly suggest that multiyear observations incorporating additional organic biomarkers and organic molecular structural data are needed in future studies to make further progress in establishing specific linkages between the forcing by climate change, and the biogeochemical response occurring in Arctic rivers basins.

References

Abbott, B. W., and others. 2016. Biomass offsets little or none of permafrost carbon release from soils, streams, and wildfire: An expert assessment. Environ. Res. Lett. **11**: 034014. doi:10.1088/1748-9326/11/3/034014

Abdulla, H. A. N., E. C. Minor, and P. G. Hatcher. 2010. Using two-dimensional correlations of 13C NMR and FTIR to investigate changes in the chemical composition of dissolved organic matter along an estuarine transect. Environ. Sci. Technol. **44**: 8044–8049. doi:10.1021/es100898x

Acar, E., T. G. Kolda, and D. M. Dunlavy. 2011. All-at-once optimization for coupled matrix and tensor factorizations [accessed 2020 June 2]. Available from https://arxiv.org/abs/1105.3422

Acar, E., C. Schenker, Y. Levin-Schwartz, V. D. Calhoun, and T. Adali. 2019. Unraveling diagnostic biomarkers of schizophrenia through structure-revealing fusion of multi-modal neuroimaging data. Front. Neurosci. **13**: 416. doi:10.3389/fnins.2019.00416

Aiken, G. R., R. G. M. Spencer, R. G. Striegl, P. F. Schuster, and P. A. Raymond. 2014. Influences of glacier melt and permafrost thaw on the age of dissolved organic carbon in the Yukon River basin. Global Biogeochem. Cycles **28**: 525–537. doi:10.1002/2013GB004764

Bader, B. W., and T. G. Kolda. 2007. Efficient MATLAB computations with sparse and factored tensors. SIAM J. Sci. Comput. **30**: 205–231. doi:10.1137/060676489

Belzile, C., and L. Guo. 2006. Optical properties of low molecular weight and colloidal organic matter: Application of the ultrafiltration permeation model to DOM absorption and fluorescence. Mar. Chem. **98**: 183–196. doi:10.1016/j.marchem.2005.08.009

Belzile, C., J. A. E. Gibson, and W. F. Vincent. 2002. Colored dissolved organic matter and dissolved organic carbon exclusion from lake ice: Implications for irradiance transmission and carbon cycling. Limnol. Oceanogr. **47**: 1283–1293. doi:10.4319/lo.2002.47.5.1283

Benner, R., and R. M. W. Amon. 2015. The size-reactivity continuum of major bioelements in the ocean. Ann. Rev. Mar. Sci. **7**: 185–205. doi:10.1146/annurev-marine-010213-135126

Birdwell, J. E., and A. S. Engel. 2010. Characterization of dissolved organic matter in cave and spring waters using UV-vis absorbance and fluorescence spectroscopy. Org. Geochem. **41**: 270–280. doi:10.1016/j.orggeochem.2009.11.002

Brabets, T. P., B. Wang, and R. H. Meade. 2000. Environmental and hydrologic overview of the Yukon River basin, Alaska and Canada. U.S. Geol. Surv. 106. doi:10.3133/wri994204

Cai, Y., L. Guo, X. Wang, and G. Aiken. 2015. Abundance, stable isotopic composition, and export fluxes of DOC, POC, and DIC from the Lower Mississippi River during 2006–2008. JGR Biogeosciences **120**: 2273–2288. doi:10.1002/2015JG003139

Chen, M., and R. Jaffé. 2016. Quantitative assessment of photo- and bio-reactivity of chromophoric and fluorescent dissolved organic matter from biomass and soil leachates and from surface waters in a subtropical wetland. Biogeochemistry **129**: 273–289. doi:10.1007/s10533-016-0231-7

Coble, P. G. 2007. Marine optical biogeochemistry: The chemistry of ocean color. Chem. Rev. **107**: 402–418.

Deming, J. W. 2002. Psychrophiles and polar regions. Curr. Opin. Microbiol. **5**: 301–309. doi:10.1016/S1369-5274(02)00329-6

DeVilbiss, S. E., Z. Zhou, J. V. Klump, and L. Guo. 2016. Spatiotemporal variations in the abundance and composition of bulk and chromophoric dissolved organic matter in seasonally hypoxia-influenced Green Bay, Lake Michigan, USA. Sci. Total Environ. **565**: 742–757. doi:10.1016/j.scitotenv.2016.05.015

Finlay, J., J. Neff, S. Zimov, A. Davydova, and S. Davydov. 2006. Snowmelt dominance of dissolved organic carbon in high-latitute watersheds: Implications for characterization and flux of river DOC. Geophys. Res. Lett. **33**: 2–6. doi:10.1029/2006GL025754

Frenette, J. J., P. Thibeault, J. F. Lapierre, and P. B. Hamilton. 2008. Presence of algae in freshwater ice cover of fluvial Lac Saint-Pierre (St. Lawrence River, Canada). J. Phycol. **44**: 284–291. doi:10.1111/j.1529-8817.2008.00481.x

Frey, K. E., and L. C. Smith. 2005. Amplified carbon release from vast west Siberian peatlands by 2100. Geophys. Res. Lett. **32**: 1–4. doi:10.1029/2004GL022025

Gao, L., Z. Zhou, A. V. Reyes, and L. Guo. 2018. Yields and characterization of dissolved organic matter from different aged soils in northern Alaska. J. Geophys. Res. Biogeo. **123**: 2035–2052. doi:10.1029/2018JG004408

Gill, P. E., W. Murray, and M. A. Saunders. 2005. SNOPT: An SQP algorithm for large-scale constrained optimization. SIAM Rev. **47**: 99–131. doi:10.1137/S0036144504446096

Gonçalves-Araujo, R., M. A. Granskog, A. Bracher, K. Azetsu-Scott, P. A. Dodd, and C. A. Stedmon. 2016. Using fluorescent dissolved organic matter to trace and distinguish the origin of Arctic surface waters. Sci. Rep. **6**: 1–12. doi:10.1038/srep33978

Guo, L., and H. Lin. 2017. Characterization of dissolved organic matter (DOM) from the lower Yukon River. *AGU Fall Meeting*.

Guo, L., and R. W. Macdonald. 2006. Source and transport of terrigenous organic matter in the upper Yukon River: Evidence from isotope (δ13C, Δ14C, and δ15N) composition of dissolved, colloidal, and particulate phases. Global Biogeochem. Cycles **20**: 1–12. doi:10.1029/2005GB002593

Guo, L., and P. H. Santschi. 2007. Ultrafiltration and its applications to sampling and characterisation of aquatic colloids, p. 159–221. *In* K. J. Wilkinson and J. R. Lead [eds.], Environmental colloids and particles: Behaviour, separation and characterisation. John Wiley. doi:10.1002/9780470024539.ch4

Guo, L., P. H. Santschi, and K. Warnken. 1995. Dynamics of dissolved organic carbon (DOC) in oceanic environments. Limnol. Oceanogr. **40**: 1392–1403. doi:10.4319/lo.1995.40.8.1392

Guo, L., C. L. Ping, and R. W. Macdonald. 2007. Mobilization pathways of organic carbon from permafrost to Arctic rivers in a changing climate. Geophys. Res. Lett. **34**: L13603. doi:10.1029/2007GL030689

Guo, L., Y. Cai, C. Belzile, and R. W. Macdonald. 2012. Sources and export fluxes of inorganic and organic carbon and nutrient species from the seasonally ice-covered Yukon River. Biogeochemistry **107**: 187–206. doi:10.1007/s10533-010-9545-z

Helms, J. R., A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber, and K. Mopper. 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. Limnol. Oceanogr. **53**: 955–969. doi:10.4319/lo.2008.53.3.0955

Holmes, R. M., J. W. McClelland, P. A. Raymond, B. B. Frazer, B. J. Peterson, and M. Stieglitz. 2008. Lability of DOC transported by Alaskan rivers to the Arctic Ocean. Geophys. Res. Lett. **35**: L03402. doi:10.1029/2007GL032837

Huguet, A., L. Vacher, S. Relexans, S. Saubusse, J. M. Froidefond, and E. Parlanti. 2009. Properties of fluorescent dissolved organic matter in the Gironde estuary. Org. Geochem. **40**: 706–719. doi:10.1016/j.orggeochem.2009.03.002

Kulkarni, H. V., N. Mladenov, S. Datta, and D. Chatterjee. 2018. Influence of monsoonal recharge on arsenic and dissolved organic matter in the Holocene and Pleistocene aquifers of the Bengal Basin. Sci. Total Environ. **637/638**: 588–599. doi:10.1016/j.scitotenv.2018.05.009

Lee, M.-H., C. L. Osburn, K.-H. Shin, and J. Hur. 2018. New insight into the applicability of spectroscopic indices for dissolved organic matter (DOM) source discrimination in aquatic systems affected by biogeochemical processes. Water Res. **147**: 164–176. doi:10.1016/j.watres.2018.09.048

Lin, H., and L. Guo. 2020. Variations in colloidal DOM composition with molecular weight within individual water samples as characterized by flow field-flow fractionation and EEM-PARAFAC analysis. Environ. Sci. Technol. **54**: 1657–1667. doi:10.1021/acs.est.9b07123

Maie, N., N. M. Scully, O. Pisani, and R. Jaffé. 2007. Composition of a protein-like fluorophore of dissolved organic matter in coastal wetland and estuarine ecosystems. Water Res. **41**: 563–570. doi:10.1016/j.watres.2006.11.006

McGuire, A. D., and others. 2009. Sensitivity of the carbon cycle in the Arctic to climate change. Ecol. Monogr. **79**: 523–555. doi:10.1890/08-2025.1

Murphy, K. R., C. A. Stedmon, D. Graeber, and R. Bro. 2013. Fluorescence spectroscopy and multi-way techniques. PARAFAC. Anal. Methods **5**: 6557–6566. doi:10.1039/c3ay41160e

Murphy, K. R., C. A. Stedmon, P. Wenig, and R. Bro. 2014. OpenFluor—an online spectral library of auto-fluorescence by organic compounds in the environment. Anal. Methods **6**: 658–661. doi:10.1039/C3AY41935E

Murphy, K. R., S. A. Timko, M. Gonsior, L. C. Powers, U. J. Wünsch, and C. A. Stedmon. 2018. Photochemistry illuminates ubiquitous organic matter fluorescence spectra. Environ. Sci. Technol. **52**: 1–12. doi:10.1021/acs.est.8b02648

O'Connor, J. A., K. Lu, L. Guo, B. E. Rosenheim, and Z. Liu. 2020. Composition and lability of riverine dissolved organic matter: Insights from thermal slicing ramped pyrolysis GC–MS, amino acid, and stable isotope analyses. Org. Geochem. **149**: 104100. doi:10.1016/j.orggeochem.2020.104100

O'Donnell, J. A., G. R. Aiken, E. S. Kane, and J. B. Jones. 2010. Source water controls on the character and origin of dissolved organic matter in streams of the Yukon River basin, Alaska. J. Geophys. Res. Biogeosci **115**: 1–12. doi:10.1029/2009JG001153

O'Donnell, J., T. Douglas, A. Barker, and L. Guo. 2021. Changing biogeochemical cycles of organic carbon, nitrogen,

phosphorus, and trace elements in Arctic rivers, p. 315–348. *In* D. Yang and D. Kane [eds.], Arctic hydrology, permafrost and ecosystems. Springer International Publishing. doi:10.1007/978-3-030-50930-9_11

Osburn, C. L., T. J. Boyd, M. T. Montgomery, T. S. Bianchi, R. B. Coffin, and H. W. Paerl. 2016. Optical proxies for terrestrial dissolved organic matter in estuaries and coastal waters. Front. Mar. Sci. **2**: 127. doi:10.3389/fmars.2015.00127

Osburn, C. L., J. D. Kinsey, T. S. Bianchi, and M. R. Shields. 2019. Formation of planktonic chromophoric dissolved organic matter in the ocean. Mar. Chem. **209**: 1–13. doi:10.1016/j.marchem.2018.11.010

Pokrovsky, O. S., L. S. Shirokova, S. N. Kirpotin, S. Audry, J. Viers, and B. Dupré. 2011. Effect of permafrost thawing on organic carbon and trace element colloidal speciation in the thermokarst lakes of western Siberia. Biogeosciences **8**: 565–583. doi:10.5194/bg-8-565-2011

Raymond, P. A., and others. 2007. Flux and age of dissolved organic carbon exported to the Arctic Ocean: A carbon isotopic study of the five largest Arctic rivers. Global Biogeochem. Cycles **21**: 1–9. doi:10.1029/2007GB002934

Schuur, E. A. G., and others. 2015. Climate change and the permafrost carbon feedback. Nature **520**: 171–179. doi:10.1038/nature14338

Shogren, A. J., J. P. Zarnetske, B. W. Abbott, F. Iannucci, and W. B. Bowden. 2020. We cannot shrug off the shoulder seasons: Addressing knowledge and data gaps in an Arctic headwater. Environ. Res. Lett. **15**: 104027. doi:10.1088/1748-9326/ab9d3c

Shutova, Y., A. Baker, J. Bridgeman, and R. K. Henderson. 2014. Spectroscopic characterisation of dissolved organic matter changes in drinking water treatment: From PARAFAC analysis to online monitoring wavelengths. Water Res. **54**: 159–169. doi:10.1016/j.watres.2014.01.053

Spencer, R. G. M., G. R. Aiken, K. P. Wickland, R. G. Striegl, and P. J. Hernes. 2008. Seasonal and spatial variability in dissolved organic matter quantity and composition from the Yukon River basin, Alaska. Global Biogeochem. Cycles **22**: 1–13. doi:10.1029/2008GB003231

Spencer, R. G. M., G. R. Aiken, K. D. Butler, M. M. Dornblaser, R. G. Striegl, and P. J. Hernes. 2009. Utilizing chromophoric dissolved organic matter measurements to derive export and reactivity of dissolved organic carbon exported to the Arctic Ocean: A case study of the Yukon River, Alaska. Geophys. Res. Lett. **36**: 1–6. doi:10.1029/2008GL036831

Stedmon, C. A., and R. Bro. 2008. Characterizing dissolved organic matter fluorescence with parallel factor analysis: A tutorial. Limnol. Oceanogr. Methods **6**: 572–579. doi:10.4319/lom.2008.6.572

Stedmon, C. A., R. M. W. Amon, A. J. Rinehart, and S. A. Walker. 2011. The supply and characteristics of colored dissolved organic matter (CDOM) in the Arctic Ocean: Pan Arctic trends and differences. Mar. Chem. **124**: 108–118. doi:10.1016/j.marchem.2010.12.007

Striegl, R. G., G. R. Aiken, M. M. Dornblaser, P. A. Raymond, and K. P. Wickland. 2005. A decrease in discharge-normalized DOC export by the Yukon River during summer through autumn. Geophys. Res. Lett. **32**: 1–4. doi:10.1029/2005GL024413

Striegl, R. G., M. M. Dornblaser, G. R. Aiken, K. P. Wickland, and P. A. Raymond. 2007. Carbon export and cycling by the Yukon, Tanana, and porcupine rivers, Alaska, 2001–2005. Water Resour. Res. **43**: 2001–2005. doi:10.1029/2006WR005201

Tanaka, K., K. Kuma, K. Hamasaki, and Y. Yamashita. 2014. Accumulation of humic-like fluorescent dissolved organic matter in the Japan Sea. Sci. Rep. **4**: 1–7. doi:10.1038/srep05292

Teber, T. 2016. Seasonal and spatial variations in chemical composition and fluxes of dissolved organic matter and nutrients in the lower Milwaukee River. thesis Univ. of Wisconsin. https://dc.uwm.edu/etd/1423

Turetsky, M. R., and others. 2020. Carbon release through abrupt permafrost thaw. Nat. Geosci. **13**: 138–143. doi:10.1038/s41561-019-0526-0

Vonk, J. E., and others. 2013. High biolability of ancient permafrost carbon upon thaw. Geophys. Res. Lett. **40**: 2689–2693. doi:10.1002/grl.50348

Weishaar, J. L., G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii, and K. Mopper. 2003. Evaluation of specific ultraviolet absorbance as an indicator of the chemical composition and reactivity of dissolved organic carbon. Environ. Sci. Technol. **37**: 4702–4708. doi:10.1021/es030360x

Wünsch, U. J., E. Acar, B. P. Koch, K. R. Murphy, P. Schmitt-Kopplin, and C. A. Stedmon. 2018. The molecular fingerprint of fluorescent natural organic matter offers insight into biogeochemical sources and diagenetic state. Anal. Chem. **90**: 14188–14197. doi:10.1021/acs.analchem.8b02863

Wünsch, U. J., R. Bro, C. A. Stedmon, P. Wenig, K. R. Murphy, R. Bro, and P. Wenig. 2019. Emerging patterns in the global distribution of dissolved organic matter fluorescence. Anal. Methods **11**: 888–893. doi:10.1039/C8AY02422G

Xu, H., and L. Guo. 2017. Molecular size-dependent abundance and composition of dissolved organic matter in river, lake and sea waters. Water Res. **117**: 115–126. doi:10.1016/j.watres.2017.04.006

Xu, H., and L. Guo. 2018. Intriguing changes in molecular size and composition of dissolved organic matter induced by microbial degradation and self-assembly. Water Res. **135**: 187–194. doi:10.1016/j.watres.2018.02.016

Yamashita, Y., L. J. Scinto, N. Maie, and R. Jaffé. 2010. Dissolved organic matter characteristics across a subtropical wetland's landscape: Application of optical properties in the assessment of environmental dynamics. Ecosystems **13**: 1006–1019. doi:10.1007/s10021-010-9370-1

Yang, L., D. H. Han, B. M. Lee, and J. Hur. 2015. Characterizing treated wastewaters of different industries using clustered fluorescence EEM-PARAFAC and FT-IR spectroscopy: Implications for downstream impact and source identification. Chemosphere **127**: 222–228. doi:10.1016/j.chemosphere.2015.02.028

Yang, L., and others. 2019. Characterization and bioavailability of rainwater dissolved organic matter at the southeast coast of China using absorption spectroscopy and fluorescence EEM-PARAFAC. Estuar. Coast. Shelf Sci. **217**: 45–55. doi:10.1016/j.ecss.2018.11.002

Zhao, Z., and others. 2017. Picocyanobacteria and deep-ocean fluorescent dissolved organic matter share similar optical properties. Nat. Commun. **8**: 1–10. doi:10.1038/ncomms15284

Zhou, Z., L. Guo, A. M. Shiller, S. E. Lohrenz, V. L. Asper, and C. L. Osburn. 2013. Characterization of oil components from the Deepwater horizon oil spill in the Gulf of Mexico using fluorescence EEM and PARAFAC techniques. Mar. Chem. **148**: 10–21. doi:10.1016/j.marchem.2012.10.003

Zhou, Z., L. Guo, and E. C. Minor. 2016. Characterization of bulk and chromophoric dissolved organic matter in the Laurentian Great Lakes during summer 2013. J. Great Lakes Res. **42**: 789–801. doi:10.1016/j.jglr.2016.04.006

Zou, L., M. Y. Sun, and L. Guo. 2006. Temporal variations of organic carbon inputs into the upper Yukon River: Evidence from fatty acids and their stable carbon isotopic compositions in dissolved, colloidal and particulate phases. Org. Geochem. **37**: 944–956. doi:10.1016/j.orggeochem.2006.04.002

Zsolnay, A., E. Baigar, M. Jimenez, B. Steinweg, and F. Saccomandi. 1999. Differentiating with fluorescence spectroscopy the sources of dissolved organic matter in soils subjected to drying. Chemosphere **38**: 45–50. doi:10.1016/S0045-6535(98)00166-0

Acknowledgments

We gratefully thank Peter deHard, Lloyd Heckman, Mindy Juliana, Toru Saito, and Chunhao Xu for their assistance during field sampling and sample processing, and Urban Wünsch and two anonymous reviewers for their careful reviews and insightful comments/suggestions, which greatly improved the manuscript. Funding for the Yukon River sampling was provided by the National Science Foundation (EAR-0554781 to L.G.). This work was supported in part by the University of Wisconsin-Milwaukee (DIG-101X405), Freshwater Collaborative of Wisconsin, and UWM's Distinguished Dissertation Fellowship (to H.L.).

Conflict of Interest

None declared.

Submitted 27 October 2020
Revised 07 February 2021
Accepted 18 May 2021

Associate editor: Yunlin Zhang