Water Resources Research

Supporting Information for

Effect of decreasing biological lability on dissolved organic matter dynamics in streams

Angang Li[1*], Jennifer D. Drummond[2*], Jennifer Bowen[3], Rose M. Cory[3], Louis A. Kaplan[4], Aaron I. Packman[1]

[1] Civil and Environmental Engineering, Northwestern University, Evanston, IL, USA

[2] School of Geography, Earth & Environmental Sciences, University of Birmingham, Edgbaston, Birmingham, UK.

[3] Earth and Environmental Sciences, University of Michigan Ann Arbor, Ann Arbor, MI, USA

[4] Stroud Water Research Center, Avondale, PA, USA

Correspondence to j.drummond@bham.ac.uk

*Both authors contributed equally to this work

Contents of this file

Introduction

The supporting information includes a description of the bioreactor laboratory experiments (Text S1) and the particle tracking model (Text S2). Also included is supporting information on the fluorescence characteristics of White Clay Creek bioreactor inflow (Table S1), and model fits and tracer injection data (Table S2-S5). Figure S1 shows the fluorescence signatures of the five PARAFAC components identified in the White Clay Creek data set. Figures S2-S4 show the bioreactor and stream data. Figure S5-S6 compare stream data to simulations.

Text S1. Bioreactor laboratory experiments

The design and operation of the bioreactors were previously described in detail (Kaplan and Newbold, 1995). The use of bioreactors for laboratory experiments has also been described in Kaplan et al. (2008), Cory and Kaplan (2012), and Sleighter et al. (2014). Briefly, the bioreactor experiments followed the same procedure, with influents of natural streamwater amended with ^{13}C-DOC tracer prepared from tulip poplar tree tissues (*Liriodendron tulipifera)* on Oct 2 2002 followed on Oct 4 2002 with unamended natural streamwater collected from White Clay Creek near the Stroud Water Research Center (39.8594N, 75.78381W). Natural streamwater from White Clay Creek was the influent on Aug 4, 2016, Nov 3, 2016, Jan 26, 2017, and May 23, 2017 and was collected under baseflow conditions (64.0, 52.3, 82.9, and 81.2 L s^{-1} respectively, with streamwater temperatures of 18, 11.2, 6.2, and 13 ºC respectively). The streamwater was filtered using a three-stage Balston glass-fiber cartridge system of 75, 25, and 0.3 μm filters in series, which removed large particles and allowed 95% of streamwater bacteria to pass (Kaplan and Newbold, 1995). The filtered streamwater was then pumped from the bottom up into bioreactors at a constant flow rate of 4 mL min^{-1}. The bioreactors, kept in dark at controlled temperature (20 °C) are chromatography columns filled with sintered glass beads (Siran, Jaeger Biotech Engineering) and maintained at a range of volumes corresponding to a range of residence times (0.5, 1.5, 3, 6, 18, 37, 150, and 271.5 min).

The fluorescence signatures of the five PARAFAC components were identified in White Clay Creek (Figure S1) and the fluorescence characteristics of the bioreactor inflow have been compared with previously identified components (Table S1), presented and discussed in more detail in Cory and Kaplan (2012). Figure S2 shows the Oct 2002 bioreactor results of concentration of ^{13}C-DOC as a function of residence time in the bioreactor and the apparent first-order reaction rate constant of ^{13}C-DOC as a function of residence time in the bioreactor. Figure S3 shows FDOM intensity for each component (C1-C5) and DOC concentration as a function of residence time in the bioreactor. Figure S4 shows a seasonal comparison of the apparent first-order reaction rate for total DOC and total FDOM (sum of C1-C5) as a function of residence time in the bioreactor.

As a first approximation, our study assumed similar uptake of DOM between the bioreactors and the bioactive region of the streambed, although bioreactors differ substantially from the stream. Here we discuss some of the assumptions that were made and how improvements can be made in future studies, however it is important to note that the lability measurements in the bioreactors cannot be easily measured in the stream because of the multiple simultaneous reactions that occur within a stream that either produce or consume the DOC. Therefore, to isolate this process in order to use it as a model input, the bioreactors were the best available data to use within our model-data synthesis approach.

The bioreactors were not designed to generate uptake rates, but rather they are a bioassay tool designed to provide a direct measure of biodegradable dissolved organic carbon concentrations. Through the use of empty bed contact time, we have extended the bioreactor measurements to generate lability profiles for components of stream water DOC. However, the source of microbes and carbon do not differ between the streambed and the bioreactors. The organisms colonizing the bioreactors from the stream water over a period of months to years were most likely released from the streambed and measurements based on biochemical analyses of phospholipid fatty acids (Wiegner et al. 2015) and 16S-rRNA (Mosher, unpublished

data) show similar microbial communities in the bioreactors and the streambed. Therefore, the C source is the stream water. Thus, the bioreactors colonized by bacteria in suspension provide valid measurements of biodegradable DOC and its lability fractions for a stream dominated by benthic activity.

For DOM uptake to occur, molecules need to come into contact with microbial cells. The effective depth in the bioreactors (water volume divided by surface area) is likely in the range of a μm or less, while in the stream effective depth ranges from cm to m. Thus, the bioreactors may overestimate the (bioactive) zone (the region over which DOM is in contact with microbial cells). The simulations presented here assume that all of the subsurface is bioactive, which led to a 25- to 108-fold overestimation in the uptake rates previously reported for White Clay Creek (Kaplan et al., 2008). In addition, DOM in streams can be exposed to sunlight that can alter the lability of DOM to microbial respiration (Cory et al., 2014; Wetzel et al., 1995) although this was not a focus for the current study. The bioreactors are maintained in the dark under constant temperatures and are fed with streamwater that has been filtered to remove particles larger than the size of bacteria. By removing the influence of processes other than microbial respiration in the bioreactors, it is possible to measure changes in the concentrations of labile and semi-labile DOM between inflow and outflow waters pumped through the bioreactors. However, the cost of isolating (dark) microbial respiration from other in-stream processes is the inability to address the effects of variable temperature, physical disturbance, microbial composition, spatial variability, detrital burial, autochthonous production and excretion, or sunlight exposure.

Text S2. Particle tracking model

The key processes included in this particle tracking framework are advection, mixing, and microbial uptake. Advection and mixing occur in the entire two-dimensional domain, and vary as a function of the vertical position in the stream-subsurface continuum (Li et al., 2017). Microbial uptake of DOM in small streams occurs mostly in the streambed near the water / sediment interface [*Fischer and Pusch*, 2001; *Battin et al.*, 2016]. We made the simplifying assumption that the entire streambed functioned as a bioreactor identical to the bioreactor column experiments, i.e., we assumed the entire streambed to be bioactive. This assumption was necessary given the unknown distribution of microbial biomass or activity in the streambed. We defined the bioactive residence time as the time DOM spends in the bioactive region of the streambed. Bioactive residence time represents the timescale over which DOM may be taken up by microbes in the streambed.

The model uses a discrete representation of the classic advection-dispersion equation based on random walk theory [*Kinzelbach*, 1988; *Delay et al.*, 2005; *Fischer et al.*, 2013]:

$$x(t+\Delta t) = x(t) + u_x \Delta t + \xi\sqrt{2D_x \Delta t} \qquad (1)$$
$$y(t+\Delta t) = y(t) + \frac{\partial D_y}{\partial y}\Delta t + \xi\sqrt{2D_y \Delta t}$$

where x is downstream position, y is vertical position, t is time, Δt is incremental time step, $u_x(y)$ is downstream velocity, ξ is a random variable with standard normal distribution. $D_y(y)$

is the vertical mixing coefficient, defined as the vertical turbulent diffusion coefficient in the water column and the vertical dispersion coefficient in the subsurface. Variability in $u_x(y)$ and $D_y(y)$ inherently captures longitudinal dispersion caused by longitudinal advection, vertical shear, and vertical mixing. Vertical shearing does not fully describe longitudinal dispersion in real systems, because of the transverse variability of velocity and mixing across the stream [*Fischer*, 1966; 1967]. Therefore, we use D_x to account for additional longitudinal mixing not inherently captured by the variability in $u_x(y)$ and $D_y(y)$. Following standard assumptions for reach-scale analysis, we assume negligible variability of flow and reaction parameters in the longitudinal direction.

Within the random-walk framework, first-order uptake is represented by the probability that a virtual particle is removed during a time step [*Prickett et al.*, 1981]:

$$p = 1 - e^{-k\Delta t} \approx k\Delta t \tag{2}$$

where k is the temperature-adjusted apparent first-order reaction rate constant, and $k\Delta t \ll 1$ to ensure that the approximation in (2) holds. The random walk in (1) and (2) defines that, at each time step, a virtual particle is subject to motion governed by velocity and mixing profiles, as well as probabilistic removal governed by the apparent first-order reaction rate constant. The uptake rate constant k was determined from bioreactor experiment data, as described in Section 2.2. DOM is assumed to be metabolized only in the subsurface ($y < 0$), so k = 0 for particle positions $y \geq 0$. To represent the decrease in lability over time as more labile components are preferentially metabolized, the effective DOM uptake rate constant k was made a function of bioactive residence time for each numerical particle, i.e., the total period of time that each particle spent in the bioactive region. We calculated bioactive residence time of each numerical particle as the cumulative time that it was in the subsurface ($y < 0$).
Following our prior work [*Li et al.*, 2017], we represented the velocity distribution across the stream-subsurface continuum, $u_x(y)$, using the classic log-law profile in the open channel [*Beavers and Joseph*, 1967; *Mendoza and Zhou*, 1992; *Manes et al.*, 2011] and an exponential transition to Darcy flow in the subsurface [*Nagaoka and Ohgaki*, 1990; *Zhou and Mendoza*, 1993]

$$\begin{aligned} u_x(y \geq 0) &= \frac{u_*}{\kappa} ln\left(\frac{y + y_0}{y_0}\right) + u_s \\ u_x(y < 0) &= u_d + (u_s - u_d)e^{My} \end{aligned} \tag{3}$$

where u_* is shear velocity, κ is von Karman constant, y_0 is a length scale characterizing the effect of porous bed on in-stream velocity, u_s is the slip velocity at the sediment-water interface, u_d is the seepage velocity of underlying porewater flow (underflow), and M is the decay rate over depth. Shear velocity was estimated as $u_* = \sqrt{gHS}$, where g is gravitational acceleration, H is the river depth, and S is the river slope [*Fischer et al.*, 2013]. We estimated u_d using Darcy's law $u_d = \frac{KS}{\phi}$, where K is hydraulic conductivity, and ϕ is porosity of the riverbed [*Elliott and Brooks*, 1997a; *Elliott and Brooks*, 1997b]. We assumed D_x to be constant in the open channel and negligible in the subsurface. The assumption of negligible $D_x$in the subsurface is justified because dispersion associated with the vertical profile of downstream

advection has a much greater effect on longitudinal transport than downstream transport. We assigned $D_y(y)$ following conventional parabolic profile in the open channel [*Fischer et al.*, 2013] and an exponential transition to the underflow

$$D_y(y \geq 0) = u_*^2 \left(1 - \frac{y}{H}\right) \frac{dy}{du_x} \quad (4)$$

$$D_y(y < 0) = D_d + (D_s - D_d)e^{My}$$

where $D_d = D_{50}u_d$ is vertical mixing coefficient in the underflow where Darcy's law applies [*Bear*, 1972], and D_s is the vertical mixing coefficient at $y = 0$.

To obtain a reasonable representation of transport conditions for each dataset, the minimum set of fitting parameters consists of y_0, u_s, u_d, M, and D_x. These parameters were constrained using independent measurements and theory. We constrained y_0 and u_s to yield $U > u_p$. We constrained u_d based on reported ranges of K (1.8×10^{-5} m s^{-1} to 1.1×10^{-4} m s^{-1}) and ϕ (20% to 30%) for the study section of White Clay Creek [*Battin et al.*, 2003; *Sawyer et al.*, 2014]. We constrained M <200 m^{-1} based on D_{50} (45-90 mm) at White Clay Creek and the reported ranges of M for gravel-bed streams [*Nagaoka and Ohgaki*, 1990]. To constrain D_x, we first estimated the total longitudinal dispersion $D_{x,tot} = \frac{0.011\langle u\rangle^2 W^2}{Hu^*}$ and the portion of longitudinal dispersion due to shearing of downstream velocity and vertical mixing $D_{x,shear} = 5.93Hu^*$ [*Fischer et al.*, 2013], where W is river width. Laboratory and field measurements suggest that $D_{x,shear}$ accounts for up to 2/3 of $D_{x,tot}$, and that the estimated value of $D_{x,tot}$ is good within a factor of 4 [*Fischer et al.*, 2013]. Based on these conditions, we constrained $\frac{1}{12}D_{x,tot} < D_x < 4D_{x,tot} - D_{x,shear}$ for each study reach. With these constraints, we fitted conservative tracer BTCs by minimizing the weighted mean square error ($WMSE$), given by $WMSE = \frac{1}{N}\sum_{i=1}^{N} w_i\left(C_{i,obs} - C_{i,sim}\right)^2$, where $C_{i,obs}$ and $C_{i,sim}$ are observed and simulated BTC values at time t_i, respectively, N is the total number of observations, and $w_i = \frac{1}{C_{i,obs}}$ is the weight. The $WMSE$ function attributes more weight to lower concentrations, which is important to characterize BTC tails [*Chakraborty et al.*, 2009].

For the Oct 2002 dataset, the best fit Br^- BTC at $x = 51$ m has WMSE of 0.0033. For the Oct 2002 field dataset, we used reach characteristics and the solute BTC measured at $x = 51$ m (Table S4A), since this was the closest site to the injection and therefore concentration values were able to be better distinguished from background compared to sites further downstream, which improved characterization of both the peak and tailing of the BTC. The best-fit transport parameters are summarized in Table S5A. The fitted model is able to match Br^- BTCs measured further downstream (x =15, 426, 1265 m) reasonably well (Figure S5) without additional fitting of the BTCs at those locations. The difference between simulated and measured BTC at $x = 1265$ m can be explained by the different flow conditions between $x = 51$ m and $x = 1265$ m. The discharge measured at $x = 1265$ m (17.6 L s^{-1}) is higher than the discharge simulated based on $x = 51$ m (12.8 L s^{-1}), resulting in an earlier observation of BTC peak than simulated.

The WMSE of the particle tracking model ranged from 0.0024 in Jan and May to 0.0052 in Aug and Nov, compared to the WMSE of the OTIS model that ranged from 0.0055 in Jan and May

to 0.0033 in Aug and Nov. Overall, the WMSE of the particle tracking model was comparable to that of the OTIS model, which indicates that the particle tracking model was able to characterize conservative transport in White Clay Creek in each season (January, May, August, and November).

Kaplan et al. (2008) conducted whole-stream DOC tracer addition in October within a 1.27 km reach at White Clay Creek, and estimated uptake velocity by assuming three homogeneous lability pools (labile, semi-labile, and recalcitrant) of which the uptake velocities were unchanging over downstream distance. They estimated uptake velocity to be 20.3 μm s^{-1} for the labile pool, 1.1 μm s^{-1} for the semi-labile pool, and 0 μm s^{-1} for the recalcitrant pool, resulting in a weighted mean uptake velocity of 2.4 μm s^{-1} (Kaplan et al., 2008). In comparison, the particle tracking model estimated the uptake velocity of the continuum of lability pools in bulk DOC, and obtained an estimate of a weighted mean estimate for November of 34 μm s^{-1}, a velocity far greater than that from Kaplan et al. (2008).

Table S1. From Cory & Kaplan (2012). Fluorescence characteristics of White Clay Creek bioreactor inflow. Primary and (secondary) excitation (l ex) and emission (l em) maxima, compared with previously identified components.

Component no.	Ex : Em (nm : nm)	EEM region (Coble 1996)	Description; likely source(s)	Average F_{max} (RU) (±SD)	Detection limit (RU)
1	250(310)/420	Peak M	Humic-like FDOM; associated with new, biological production of humic substances from autochthonous sources, found in marine and freshwater systems	0.380±0.002	0.010
2	250(420)/500	Peak A	Humic-like FDOM; terrestrial (humic substances) DOM, ubiquitously observed	0.201±0.002	0.006
3	250(365)/466	Peak C	Humic-like FDOM; ubiquitously observed humic-substances, thought to be associated with predominately terrestrial sources	0.163±0.002	0.004
4	(250)285/344	Peak T	Tryptophan-like FDOM; amino acids, proteins	0.0760±0.0009	0.0032
5	275/324	Peak B	Tyrosine-like FDOM; Amino acids, proteins	0.0131±0.0005	0.0018

Table S2. Fitted parameters[a] of the reactivity continuum model that describes reaction rates of (A) total DOC in different seasons and ^{13}C-DOC tracer in October and (B) FDOM components C1-C5 in different seasons. Our data were better described using the reactivity continuum model, compared to linear or exponential models.

(A) DOC and ^{13}C-DOC

	DOC in Jan	DOC in May	DOC in Aug	DOC in Nov	^{13}C-DOC in Oct
α (s)	48	40	26	30	2
ν (unitless)	0.06	0.08	0.06	0.1	0.5

(B) FDOM components C1-C5

	FDOM component	Jan	May	Aug	Nov
α (s)	C1	509	271	226	135
ν (unitless)		0.08	0.07	0.08	0.09
α (s)	C2	401	215	213	113
ν $(unitless)$		0.07	0.06	0.07	0.07
α (s)	C3	980	292	150	223
ν $(unitless)$		0.08	0.07	0.06	0.10
α (s)	C4	745	184	208	62
ν $(unitless)$		0.08	0.05	0.07	0.07
α (s)	C5	44	27	30	81
ν $(unitless)$		0.16	0.16	0.17	0.28

[a] Fitted parameters α and ν were defined in Section 2.2.

Table S3. Matching between seasonal bioreactor and tracer data.

bioreactor id	sampling date	discharge (L/s)	temperature (ºC)	tracer id	relative injection location (m)[a]	date	discharge (L/s)
1	Jan 26, 2017	82.9	6.2	i	602	Jul 23, 2014	81.9
2	May 23, 2017	81.2	13.0	i	602	Jul 23, 2014	81.9
3	Aug 4, 2016	64.0	18.0	ii	-812	Jul 23, 2014	61.7
4	Nov 3, 2016	52.3	11.2	ii	-812	Jul 23, 2014	61.7

[a] Positive and negative values represent distances downstream and upstream of the bioreactor water sampling site, respectively, assuming distance between the bioreactor water sampling site and Spencer road crossing is ~95 m.

Table S4. Reach characteristics of the tracer injection experiments.

(A) Oct 2002 dataset[a]

relative injection location (m)	injection duration (s)	reach length (m)	stream width (m)	stream depth (m)	stream slope	D50 (mm)	discharge (L/s)
-1264	6829	51	3.3	0.10	0.0054	90	12.8

[a] Co-injection of 307 g Br and 53.4 mmol ^{13}C-DOC.

(B) Seasonal datasets (see Table S3)

tracer id	injectate	mass injected (g)	injection duration (s)	reach length (m)	stream width (m)	stream depth (m)	stream slope	D50 (mm)
i	NaCl	756	5	91	1.47	0.18	0.0049	45
ii	NaCl	1000	5	107	3.95	0.13	0.0054	90

Table S5. Summary of best-fit flow parameters.

(A) Case study

$\langle u \rangle$ (m/s)	u_s (m/s)	u_d (m/s)	M (m^{-1})	$\langle D_y \rangle$ ($m^2\ s^{-1}$)	D_d ($m^2\ s^{-1}$)	D_x ($m^2\ s^{-1}$)
0.07	0.017	1.7E-06	80	0.0028	1.6E-07	0.04

$\langle u \rangle$ represents mean in-stream velocity, u_s represents slip velocity, u_d represents underflow velocity, M represents decay rate over depth, $\langle D_y \rangle$ represents mean in-stream vertical mixing coefficient, D_d represents underflow vertical mixing coefficient, and D_x represents longitudinal dispersion coefficient. Note that D_x represents additional longitudinal mixing not inherently captured by the variability in $u_x(y)$ and $D_y(y)$.

(B) Seasonal study

tracer id	$\langle u \rangle$ (m/s)	u_s (m/s)	u_d (m/s)	M (m^{-1})	$\langle D_y \rangle$ ($m^2\ s^{-1}$)	D_d ($m^2\ s^{-1}$)	D_x ($m^2\ s^{-1}$)
i	0.40	0.040	2.0E-06	30	0.0015	8.9E-08	0.01
ii	0.16	0.024	1.3E-06	90	0.0018	7.9E-08	0.02

$\langle u \rangle$ represents mean in-stream velocity, u_s represents slip velocity, u_d represents underflow velocity, M represents decay rate over depth, $\langle D_y \rangle$ represents mean in-stream vertical mixing coefficient, D_d represents underflow vertical mixing coefficient, and D_x represents longitudinal dispersion coefficient. Note that D_x represents additional longitudinal mixing not inherently captured by the variability in $u_x(y)$ and $D_y(y)$.

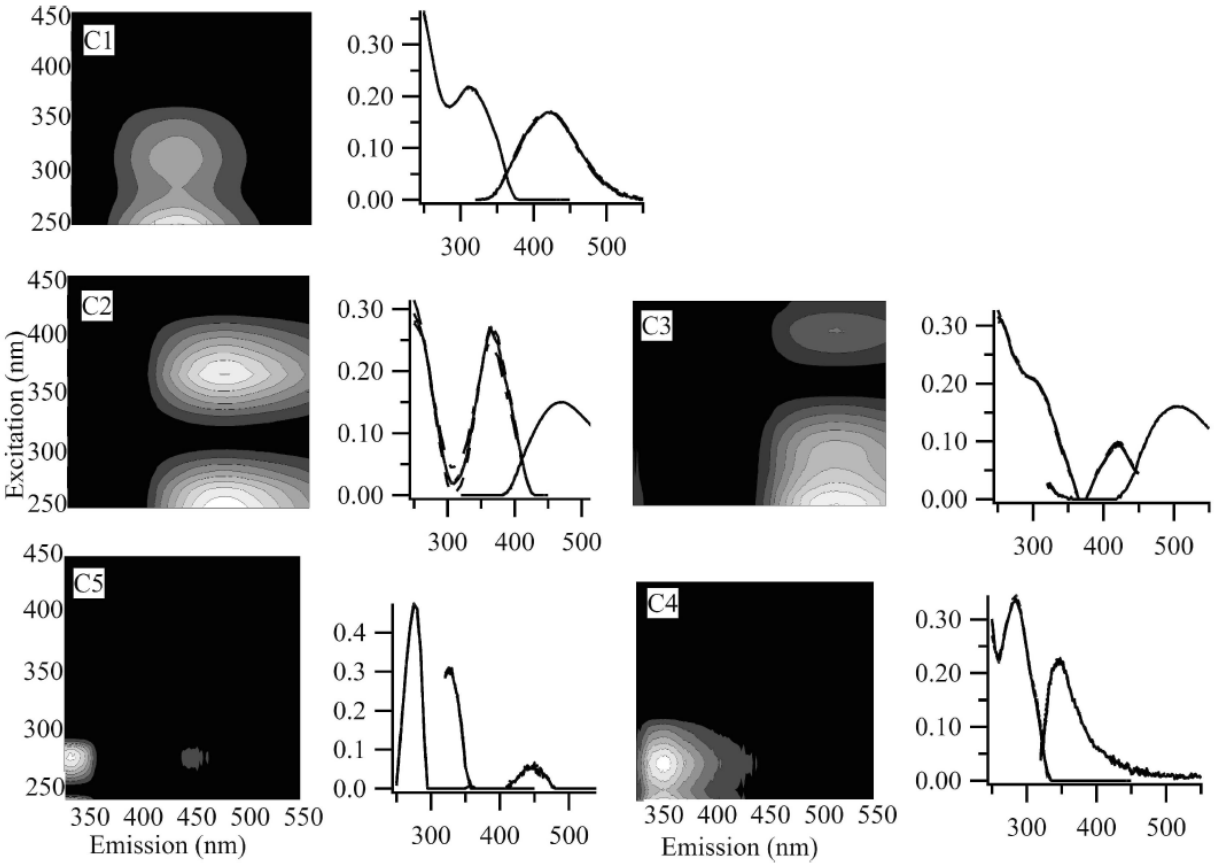

Figure S1. From Cory & Kaplan (2012). Fluorescence signatures of the five PARAFAC components identified in the White Clay Creek data set. Contour plots of components C1–C5 are ordered by decreasing percent explained, with excitation wavelength on the y-axis, emission on the x-axis, and shading representing the relative intensity of emission. Corresponding line plots to the right of each contour plot compare the split-half validation results, in which each component's excitation (left, dashed) and emission (right, solid) spectra are estimated from four independent splits of the data set, to the results from the complete data set. The x-axis in the line plots is excitation or emission wavelength, with relative intensity on the y-axis.

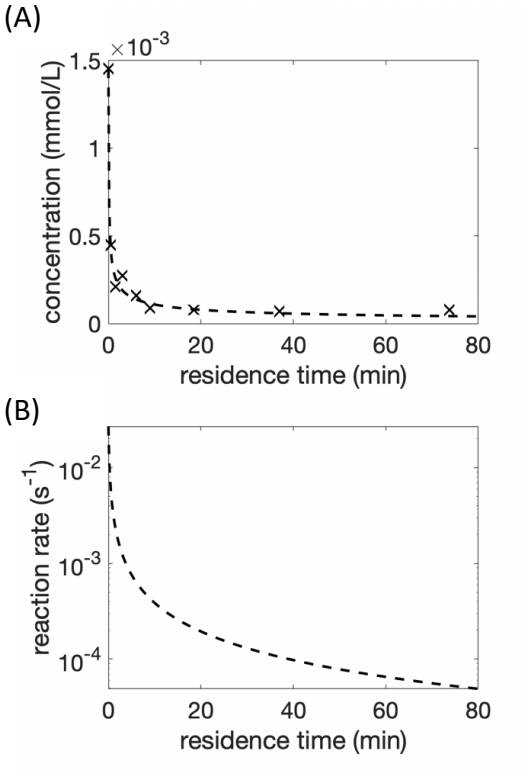

Figure S2. (A) Concentration of ^{13}C-DOC as a function of residence time in the bioreactor, measured in Oct 2002. Markers represent measured data, and lines represent fitted values. (B) Apparent first-order reaction rate constant of ^{13}C-DOC as a function of residence time in the bioreactor.

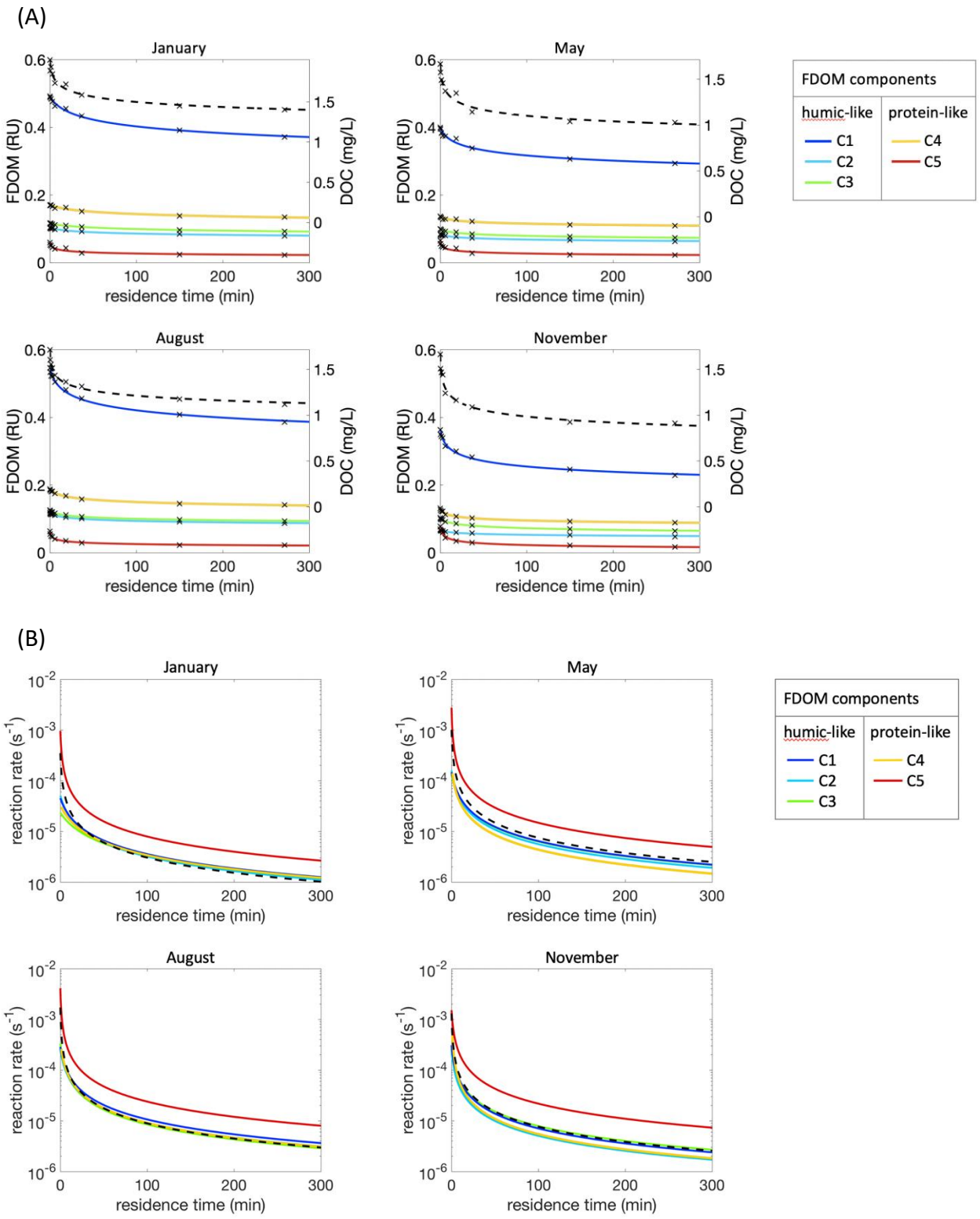

Figure S3. (A) Streamwater DOC concentration (black dashed line) and FDOM (C1-C5) intensity as a function of residence time in bioreactor, measured in Aug 2016, Nov 2016, Jan 2017, and May 2017. Markers represent measured data, and lines represent fitted values. (B) Apparent

first-order reaction rate constant of DOC (black dashed line) and FDOM components (C1-C5) as a function of residence time in bioreactor.

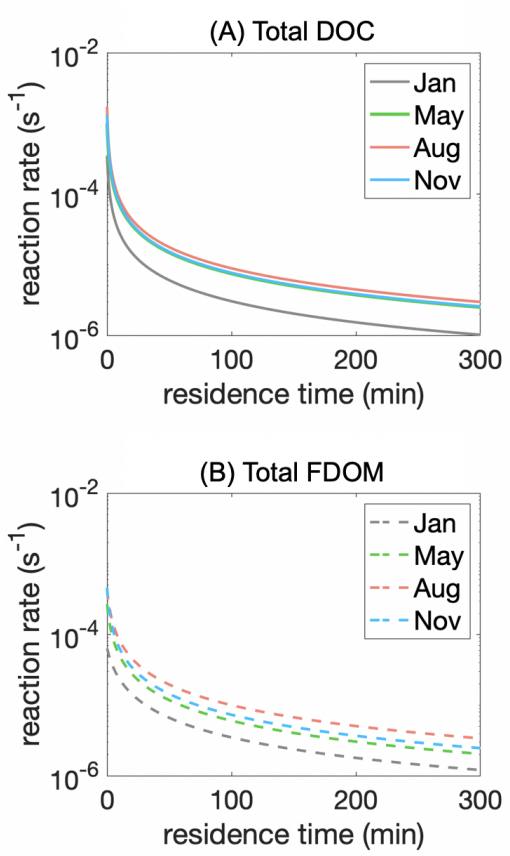

Figure S4. Seasonal comparison (Aug 2016, Nov 2016, Jan 2017, and May 2017) of reaction rate constant as a function of residence time in the bioreactor for (A) total DOC and (B) total FDOM.

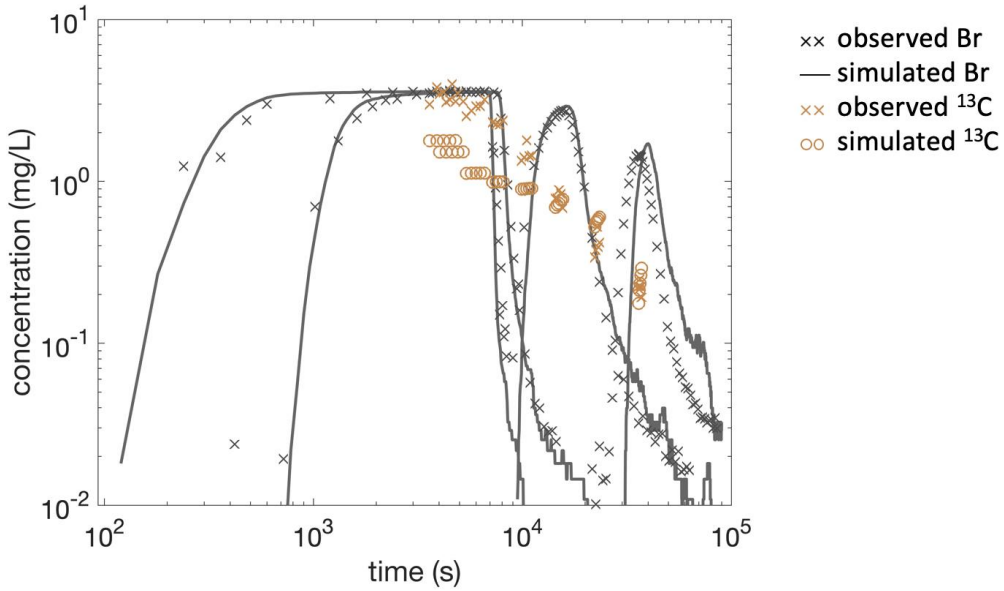

Figure S5. Simulated vs. observed Br^- BTCs and ^{13}C-DOC concentration from the tracer injection experiment of the case study. In-stream BTCs of Br tracer were measured at $x = 15$, 51, 426, and 1265 m; in-stream concentrations of ^{13}C-DOC were measured during the peaks of Br BTCs at $x = 15$, 30, 87, 147, 250, 426, 724, and 1265 m. These concentrations were corrected for background values. For visualization purposes, the unit of ^{13}C-DOC, originally measured as mmol L^{-1}, was rescaled to mg L^{-1}. The unit conversion was calculated as $C_{DOC,\ mg\ L^{-1}} = C_{DOC,\ mg\ L^{-1}} \frac{M_{Br,\ mg}}{M_{DOC,\ mmol}}$, where $C_{DOC,\ mg\ L^{-1}}$ is the ^{13}C-DOC concentration converted to mg L^{-1}, $C_{DOC,\ mmol\ L^{-1}}$ is the ^{13}C-DOC concentration in mmol L^{-1}, $M_{Br,\ mg}$ is the injected Br^- tracer mass in mg L^{-1}, and $M_{DOC,\ mmol}$ is the injected ^{13}C-DOC tracer mass in mmol L^{-1}. The simulated ^{13}C-DOC concentration is consistent with the observed concentration within a factor of 3 (0.6 – 2.8) at all sampling locations.

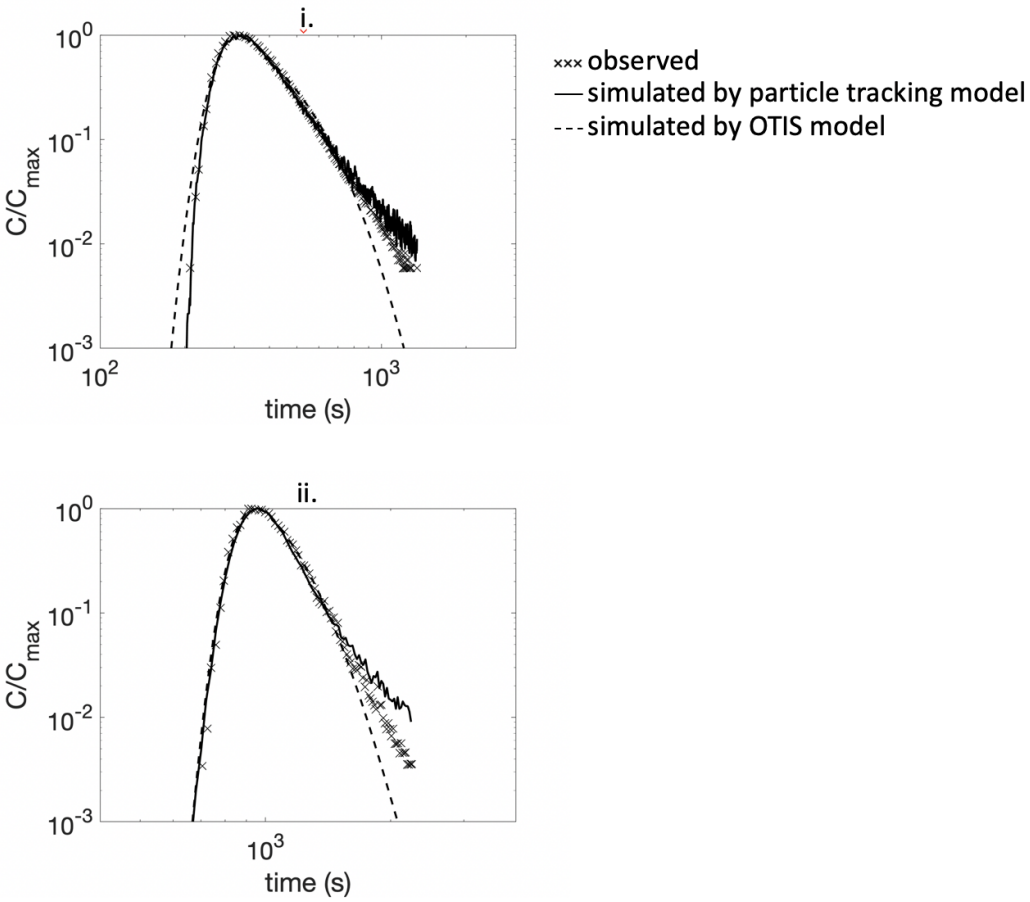

Figure S6. Best fit breakthrough curves (BTCs) of each conservative tracer study (i – ii, see Table S3). The solid line represents BTC predicted by the particle tracking model; the dashed line represents BTC predicted by the conventional OTIS model. The weighted mean square error (WMSE, defined in Section 2.5) of the particle tracking model ranged from 0.0024 in (i) to 0.0052 in (ii), compared to the WMSE of the OTIS model that ranged from 0.0055 in (i) to 0.0033 in (ii).