

Geospatial Variability of Fluorescent Dissolved Organic Matter in Urban Watersheds: Relationships with Land Cover and Wastewater Infrastructure

Jahir A. Batista-Andrade, Claire Welty, Diego Iglesias Vega, Anna McClain, and Lee Blaney*



Cite This: *Environ. Sci. Technol.* 2024, 58, 7529–7542



Read Online

ACCESS |



Metrics & More



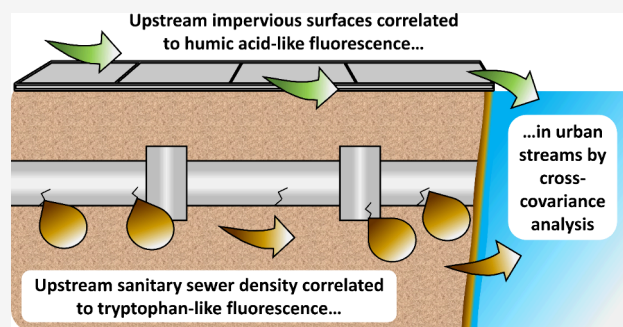
Article Recommendations



Supporting Information

ABSTRACT: We investigated the fluorescent dissolved organic matter (FDOM) composition in two watersheds with variable land cover and wastewater infrastructure, including sanitary sewers and septic systems. A four-component parallel factor analysis model was constructed from 295 excitation–emission matrices recorded for stream samples to examine relationships between FDOM and geospatial parameters. The contributions of humic acid- and fulvic acid-like fluorescence components (e.g., C1, C2, C3) were fairly consistent across a 12 month period for the 27 sampling sites. In contrast, the protein-like fluorescence component (C4) and a related ratiometric wastewater indicator (C4/C3) exhibited high variability in urban tributaries, suggesting that some sites were impacted by leaking sewer infrastructure. Principal component analysis indicated that urban areas clustered with impervious surfaces and sanitary sewer density, and cross-covariance analysis identified strong positive correlations between C4, impervious surfaces, and sanitary sewer density at short lag distances. The presence of wastewater was confirmed by detection of sucralose (up to 1,660 ng L⁻¹) and caffeine (up to 1,740 ng L⁻¹). Our findings not only highlight the potential for C4 to serve as an indicator of nearby, compromised sanitary sewer infrastructure, but also suggest that geospatial data can be used to predict areas vulnerable to wastewater contamination.

KEYWORDS: humic substances, fluorescence, contaminants of emerging concern, geostatistical analysis, sewer, septic system



1. INTRODUCTION

While a large fraction of dissolved organic matter (DOM) is derived from decomposition of plant and animal debris,¹ anthropogenic activities also contribute DOM to urban watersheds. The introduction of untreated wastewater, wastewater effluent, and stormwater runoff to streams can significantly influence the amount and composition of DOM.^{2–4} As untreated wastewater stems from sanitary sewer leaks and overflows, we hypothesized that the presence of wastewater-derived DOM in urban streams that do not directly receive wastewater effluent was correlated to sewer and septic system density. Stormwater runoff, which is related to impervious area, facilitates transport of DOM deposited on urban surfaces. To differentiate DOM sources, some have adopted parallel factor analysis (PARAFAC) of excitation–emission matrices (EEMs) to identify unique components that can be collectively used to model the fluorescence signatures of environmental samples.⁵ Using EEM-PARAFAC, previous studies have evaluated the composition of fluorescent DOM (FDOM) in aquatic environments influenced by agriculture,⁶ forests,⁷ wastewater effluent,⁸ sewer leaks,^{9–11} and septic systems.¹²

Land cover and land use affect the concentrations of organic and inorganic contaminants in the aquatic environment.^{13,14} For

example, percent urban area, population density, and wastewater discharge locations were associated with antibiotic levels in Twin Cities rivers (MN, USA).¹³ A field sampling campaign in Baltimore (MD, USA) demonstrated that chloride, sulfate, and nitrate concentrations in streams were related to impervious surfaces, sanitary sewer infrastructure, and urban fill.¹⁴ Because DOM coexists with the aforementioned contaminants, we hypothesized that EEM-PARAFAC components could be used in conjunction with geospatial land cover and wastewater infrastructure data to identify urban streams that do not receive wastewater effluent but are impacted by sanitary sewer leaks and overflows. Other aspects, including nutrient concentrations,¹⁵ soil properties,¹⁶ and light availability,¹⁷ can also affect the FDOM pool in aquatic systems.

Received: September 24, 2023

Revised: April 6, 2024

Accepted: April 9, 2024

Published: April 22, 2024



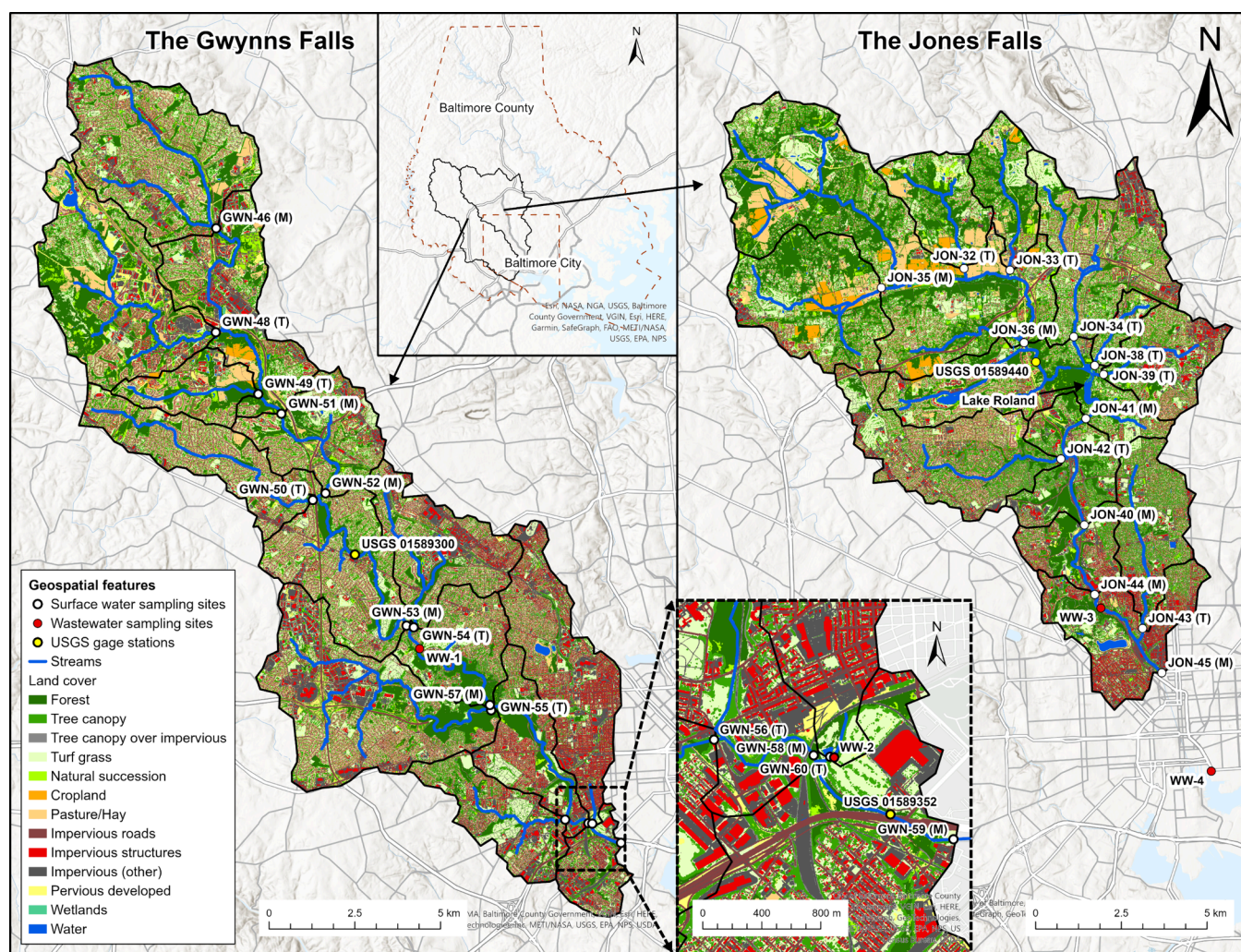


Figure 1. Sampling locations in the Gwynns Falls and Jones Falls watersheds of Baltimore, MD, USA. Streams are indicated by blue lines. Sites located in the main stem and tributaries have "(M)" and "(T)" in their names, respectively. Subcatchments associated with tributaries are outlined in bold lines. Land cover data were obtained from the Chesapeake Conservancy Conservation Innovation (CCCI) Center high-resolution database.²⁷

Previous work demonstrated that ratiometric parameters related to EEM-PARAFAC components with tryptophan- and fulvic acid-like fluorescence could be employed to identify the impacts of failing sewer infrastructure on water quality in low-order streams.⁹ However, relationships were not explicitly examined between FDOM composition, land cover (e.g., cropland, impervious surfaces), and wastewater infrastructure (e.g., sewer density, septic system density). Such relationships could enable *a priori* identification of areas vulnerable to wastewater contamination and improve on-the-ground efforts to maintain wastewater infrastructure. Conventional approaches to correlate land cover features with FDOM composition include the Wilcoxon signed-rank¹⁸ and Kruskal–Wallis tests,¹⁹ redundancy analysis,⁴ and partial least-squares,²⁰ generalized least-squares,²¹ and autoregressive²² models. These techniques do not account for the upstream-to-downstream flow direction or the spatial cross continuity or variability of geospatial attributes. Cross-covariance analysis between sampled sites and upstream land cover characteristics can address this gap.

The objective of this study was to assess spatiotemporal patterns of FDOM in urban watersheds with variable land cover and wastewater infrastructure to identify areas impacted by sanitary sewer leaks and overflows and septic systems. To

achieve this objective, we collected 295 surface water samples from 27 sites in two watersheds over a one year period, measured EEMs for each sample, developed a global EEM-PARAFAC model to explain FDOM composition, and applied principal component analysis (PCA) and cross-covariance functions to relate FDOM composition, impervious surfaces, septic system density, and sewer density. Because the sampling sites were located in adjacent watersheds characterized by similar geology, land cover, land use, and climate, we posited that the FDOM profiles for natural organic matter and, separately, raw wastewater would be similar in each watershed. However, we hypothesized that the two watersheds would exhibit different relationships with FDOM due to their variable land cover patterns (e.g., impervious surfaces) and wastewater infrastructure (e.g., sewer density, septic system density). The inclusion of 27 sampling sites across two watersheds was meant to strengthen the outcomes of this work by confirming changes in urban water quality, regardless of differences in the spatial patterns of land cover and wastewater infrastructure present in the drainage areas.

While 66 contaminants of emerging concern (CECs) were analyzed, sucralose (artificial sweetener) and caffeine (stimulant) were the most frequently detected and, therefore, used as

indicators to confirm the presence of wastewater and validate conclusions from FDOM analysis. The novelty and impact of this study stem from (1) the cross-covariance analysis of F_{max} values for EEM-PARAFAC components and geospatial data related to land cover (i.e., impervious surfaces) and wastewater infrastructure (i.e., sewer density, septic system density), (2) deployment of geostatistical analyses to differentiate the influence of natural- and wastewater-related sources of FDOM in urban watersheds that do not receive treated wastewater effluent, and (3) the reported correlations between a particular EEM-PARAFAC component, sewer density, and impervious surfaces, which enable smart design of future monitoring studies to identify the locations of sewer leaks. Overall, this study pioneers the adoption of EEM-PARAFAC and cross-covariance tools to identify wastewater contamination in urban streams that do not receive wastewater effluent, a challenging task that requires highly resolved data sets and intricate data analysis compared to rivers that directly receive wastewater or wastewater effluent from point sources. The corresponding approaches can be applied to other urban watersheds to extend the impact of our work to predict and confirm the effects of sanitary sewer leaks and overflows on water quality.

2. MATERIALS AND METHODS

2.1. Sampling Sites. This study was based in the Gwynns Falls and Jones Falls watersheds of Baltimore, MD, USA (Figure 1). Neither watershed has permitted discharges for wastewater effluent; therefore, any observed wastewater signatures would stem from sanitary sewer leaks or overflows, illegal wastewater discharges, or partially treated septic system effluent. From September 2020 to August 2021, monthly surface water samples were collected from 14 sites in Gwynns Falls and 13 locations in Jones Falls. For both watersheds, seven sampling sites were located in tributaries and the remainder were in the main stem. More details on site classification, hydrological connectivity, stream order, subcatchment area, distance from the outlet, and land cover are presented in Table S1. Sampling sites were classified as urban, suburban, and rural in Table S1 according to holistic analysis of proximity to the Baltimore Urban Rural Demarcation Line,²³ percent impervious surface area (i.e., < 20% for rural, 20–40% for suburban, > 40% for urban sites), and observations from field visits. Discharge data (Table S2) were retrieved from the United States Geological Survey (USGS) National Water Information System database for stream gages located in the Jones Falls (USGS 01589440)²⁴ and Gwynns Falls (USGS 01589300, USGS 01589352).^{25,26}

Due to age and insufficient maintenance, the 5000 km Baltimore sanitary sewer system is subject to leaks and overflows. In fact, 33% of the sanitary sewer lines in the Gwynns Falls and Jones Falls watersheds were built before 1953,^{28,29} increasing the likelihood of infrastructure issues that introduce raw wastewater into shallow groundwater and, ultimately, urban streams; however, some of these sewers have been lined to prevent leaks. To assign FDOM signatures observed in the stream network to natural processes or wastewater inputs, raw wastewater samples were collected from two manholes in the middle (WW-1) and near the outlet (WW-2) of the Gwynns Falls watershed and at upstream (WW-3) and downstream (WW-4) pumping stations in Jones Falls (Figure 1).

2.2. Sample Collection. For EEM analysis, surface water samples were collected by immersing 100 mL, amber-glass bottles 8–10 cm below the water surface until the bottles were filled. The samples were stored on ice and transported to the lab.

Due to logistical constraints that precluded immediate analysis, surface water samples were frozen at $-20\text{ }^{\circ}\text{C}$ to prevent biodegradation of FDOM before EEM analysis. Previous studies have reported contrasting effects of freeze–thaw processes on fluorescence intensity.^{30–34} Spencer et al.³⁴ reported changes in fluorescence after one year of storage, whereas Otero et al.³⁰ did not observe differences in fluorescence after 1 month. In this study, EEMs were recorded within 48 h of collection for all surface water samples. A total of 295 samples were analyzed during the 12 month campaign; note, samples were not collected from some sites due to unsafe weather conditions in particular months. Field blanks containing deionized water were prepared each month, transported to the sites, transferred to another bottle, and analyzed for quality assurance and quality control. EEMs were recorded within 48 h of sample collection.

To confirm the presence of wastewater, 11 sites were selected for analysis of 66 contaminants of emerging concern (CECs), including 43 antibiotics, 9 hormones, 13 UV filters, and sucralose, during the full campaign. Caffeine analysis was included for the last four months of sampling. More details of the CECs are provided in Table S3. The sites selected for CEC analysis were situated in rural (i.e., JON-32, JON-33), suburban (i.e., JON-39, GWN-46, and GWN-48), and urban (i.e., JON-42, JON-43, JON-45, GWN-55, GWN-58, and GWN-60) areas. Samples were collected by immersing 1-L, amber-glass bottles 8–10 cm below the water surface. Field blanks composed of deionized water were included for quality assurance and quality control. Samples were stored at $4\text{ }^{\circ}\text{C}$ for up to 24 h before preprocessing for CEC analysis by liquid chromatography with triple quadrupole tandem mass spectrometry (LC-MS/MS).

2.3. EEM Analysis and EEM-PARAFAC Modeling. The 100 mL surface water samples were thawed to room temperature, and then a 10 mL aliquot was passed through a $0.45\text{-}\mu\text{m}$ polyvinylidene fluoride syringe filter. EEMs were recorded for the filtrates using an Aqualog fluorometer (Horiba Scientific; Edison, NJ, USA) in accordance with prior protocols.⁹ Following collection, wastewater samples were immediately analyzed after 4 $\times$ dilution to decrease inner-filter effects; these samples were not frozen. No appreciable fluorescence was measured in field or laboratory blanks.

Dissolved organic carbon (DOC) concentrations were measured as nonpurgeable organic carbon by a Shimadzu TOC-L analyzer (Shimadzu; Kyoto, Japan). Six other spectral properties were evaluated to inform DOM composition: absorbance at 254 nm (A_{254}); biological index (BIX); fluorescence index (FI); humification index (HIX); ratio of absorbance at 250 nm to that at 365 nm ($E2/E3$); and spectral slope at 275–295 nm ($S_{275-295}$). A_{254} is directly related to DOM aromaticity.³⁵ $E2/E3$ and $S_{275-295}$ are inversely proportional to the molecular weight distribution of DOM.^{35,36} FI denotes the relative contributions of terrestrial and microbial sources to the DOM pool.³⁷ BIX and HIX are indicators of autotrophic productivity³⁸ and humic substances,³⁷ respectively. The DOC content and spectral properties are reported for each sample in Table S4.

With the 295 EEMs from surface water samples, a four-component EEM-PARAFAC model was developed and validated using the drEEM toolbox (version 0.6.4) in Matlab.⁵ The EEM-PARAFAC components (Figure S1) were compared to previous models in OpenFluor.³⁹ Component 1 (C1) had maximum excitation ($\lambda_{ex,max}$) and emission ($\lambda_{em,max}$) wavelengths of 250 and 395 nm, respectively, similar to terrestrial humic acid-like fluorescence.^{40,41} Components 2 (C2) and 3

(C3) exhibited similar $\lambda_{\text{ex,max}}$ at 262–265 nm, but the emission peaks for C2 (i.e., $\lambda_{\text{em,max}}$ at 346 nm, 435 nm) were at lower wavelengths than those for C3 (i.e., 382 nm, 490 nm). Based on those characteristics, C2 was similar to humic acid-like fluorescence derived from biological processes,^{42,43} and C3 corresponded to fulvic acid-like fluorophores from terrestrial and microbial sources.^{44,45} Component 4 (C4) demonstrated tryptophan-like fluorescence^{46,47} with $\lambda_{\text{ex,max}}$ at 277 nm and $\lambda_{\text{em,max}}$ at 335 nm. The components were remarkably similar to those in the EEM-PARAFAC model from a previous sampling campaign in the same watersheds⁹ (Figure S2). The C4/C3 ratiometric parameter was calculated and used as a wastewater indicator.⁹ The maximum fluorescence intensity of each component ($F_{\text{max},i}$) and fractional $F_{\text{max},i}$ values ($F_{\text{max},i}/F_{\text{max}}$) were used to describe differences in FDOM composition between samples. While the $F_{\text{max},i}$ values indicated the magnitude of each component in streamwater samples, the $F_{\text{max},i}/F_{\text{max}}$ ratios reflected the relative abundance of each component in a sample (eq 1).

$$F_{\text{max},i}/F_{\text{max}} = \frac{F_{\text{max},i}}{\sum_i F_{\text{max},i}} \quad (1)$$

2.4. CEC Analysis. The 1 L samples were passed through 1.2 μm glass-fiber filters, and the filtrates were acidified to pH ≤ 3.0 with 3 M HCl. The solid-phase extraction and LC-MS/MS protocols followed previously reported methods⁴⁸ with minor modifications to enable caffeine quantitation (Text S2). CEC analysis was completed within 7 days of sample collection. No CECs were detected in field and laboratory blanks.

2.5. Cross-Covariance Analysis. A series of pour points was created at 500 m intervals along the stream networks of the two watersheds using ArcGIS Pro v3.0.2 (Environmental Systems Research Institute (ESRI); Redlands, CA, USA), as shown in Figure S3. The pour points were user-designated locations in the stream network. For each sampling site and pour point, the upgradient drainage areas were delineated with the fill, flow direction, flow accumulation, snap pour point, and watershed functions in the ESRI Hydrology toolbox using a LiDAR-based digital elevation model with 5 m resolution.⁴⁹ The distances between each sampling site and all upstream pour points were designated as lag distances and measured as one-dimensional curvilinear functions of stream length (Tables S5 and S6). In the Jones Falls, the 13 sampling sites and 164 pour points resulted in 722 lag pairs. The 14 sampling sites and 202 pour points in Gwynns Falls generated 959 lag pairs. Short lag distances occurred throughout the watersheds and mostly corresponded to locations with similar geospatial features. Long lag distances were more likely to span rural, suburban, and urban areas and therefore exhibited greater heterogeneity in land cover and wastewater infrastructure.

For the drainage area of each sampling site and pour point, the following land cover and wastewater infrastructure attributes were calculated: (1) fraction of impervious surface area (i.e., impervious area divided by total drainage area); (2) sewer density (i.e., sum of sewer pipe length divided by drainage area, km^{-1}); and (3) septic system density (i.e., number of septic systems divided by drainage area, km^{-2}) (Tables S7 and S8). Land cover data were obtained from the CCCI database at 1 m resolution.²⁷ The impervious surface category was redefined to include impervious roads, impervious structures, tree canopy over impervious surfaces, and pervious developed surfaces classifications in the CCCI database. The locations of sanitary sewer pipes and septic systems were obtained from the

Baltimore County Department of Environmental Protection and Sustainability²⁸ and Baltimore City Enterprise Geographic Information Services.²⁹

Cross-covariance is a measure of the spatial continuity or variability of two attributes separated by a known lag distance.^{50–52} In this case, one-dimensional cross-covariance analysis was conducted along the stream network using (i) $F_{\text{max},i}$ values of EEM-PARAFAC components measured at a sampling site and (ii) geospatial land cover and wastewater infrastructure features associated with the drainage areas of upstream pour points. Following best practices,^{14,53} cross-covariance was calculated using eq 2 for 20 lag classes with each containing 36 lag pairs for the Jones Falls and 10 lag classes containing 96 lag pairs for the Gwynns Falls.

$$C(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} x_i y_i - m_{-h} m_{+h} \quad (2)$$

In eq 2, $C(h)$ is the cross-covariance, $N(h)$ is the number of data pairs in a lag class, x_i is the value of the attribute at the tail of the data pair, y_i is the value of the attribute at the head of the data pair, h is the average lag distance between the head and tail, m_{-h} is the mean of the tail attribute values (i.e., $\frac{1}{N(h)} \sum_{i=1}^{N(h)} x_i$), and m_{+h} is the mean of the head attribute values (i.e., $\frac{1}{N(h)} \sum_{i=1}^{N(h)} y_i$).

The coordinate of the tail in the lag separation vector was defined as the sampling site location, and the 12 month average F_{max} values of the C1, C2, C3, and C4 EEM-PARAFAC components were used as the attribute values (Table S9). Cross-covariance analysis is not appropriate for evaluating dynamic, event-driven relationships between variables. Therefore, the 12 month average F_{max} values were employed to determine the consistent, long-term influence of upstream land cover and wastewater infrastructure on FDOM composition. The head of the lag separation vector was located at an upstream pour point, and the fraction of impervious area, sewer density, or septic system density of the upgradient area draining to the pour point were defined as the attribute values (Tables S7 and S8); these attributes do not influence each other in cross-covariance analysis. Spatial cross-covariance analysis was carried out between each of the four EEM-PARAFAC components and each of the three land cover and wastewater infrastructure parameters, yielding 12 sets of independent cross-covariance calculations for each watershed. This approach captures the directional influence of upstream land use and land cover on downstream sampling points. Because of the disparity in units between attributes, dimensionless z-scores were calculated for use in eq 2. Positive cross-covariance values indicated positive correlations, and larger numbers implied stronger relationships.

2.6. Statistical Analysis. A paired sample Wilcoxon signed-rank test was conducted to compare FDOM composition at the main stem sites to that in the tributaries. PCA and Spearman correlations were assessed using land cover and wastewater infrastructure data for the drainage areas of sampling sites (Table S1), F_{max} values for the EEM-PARAFAC components, caffeine and sucralose concentrations, and spectral properties of DOM. For PCA, the principal components were obtained for eigenvalues greater than 1. Biplots of the PC1 and PC2 components were created using the loadings for each variable and the scores for each sample. All statistical analyses were performed in OriginPro 2023 (Northampton, MA, USA), and differences were considered significant when $p < 0.05$.

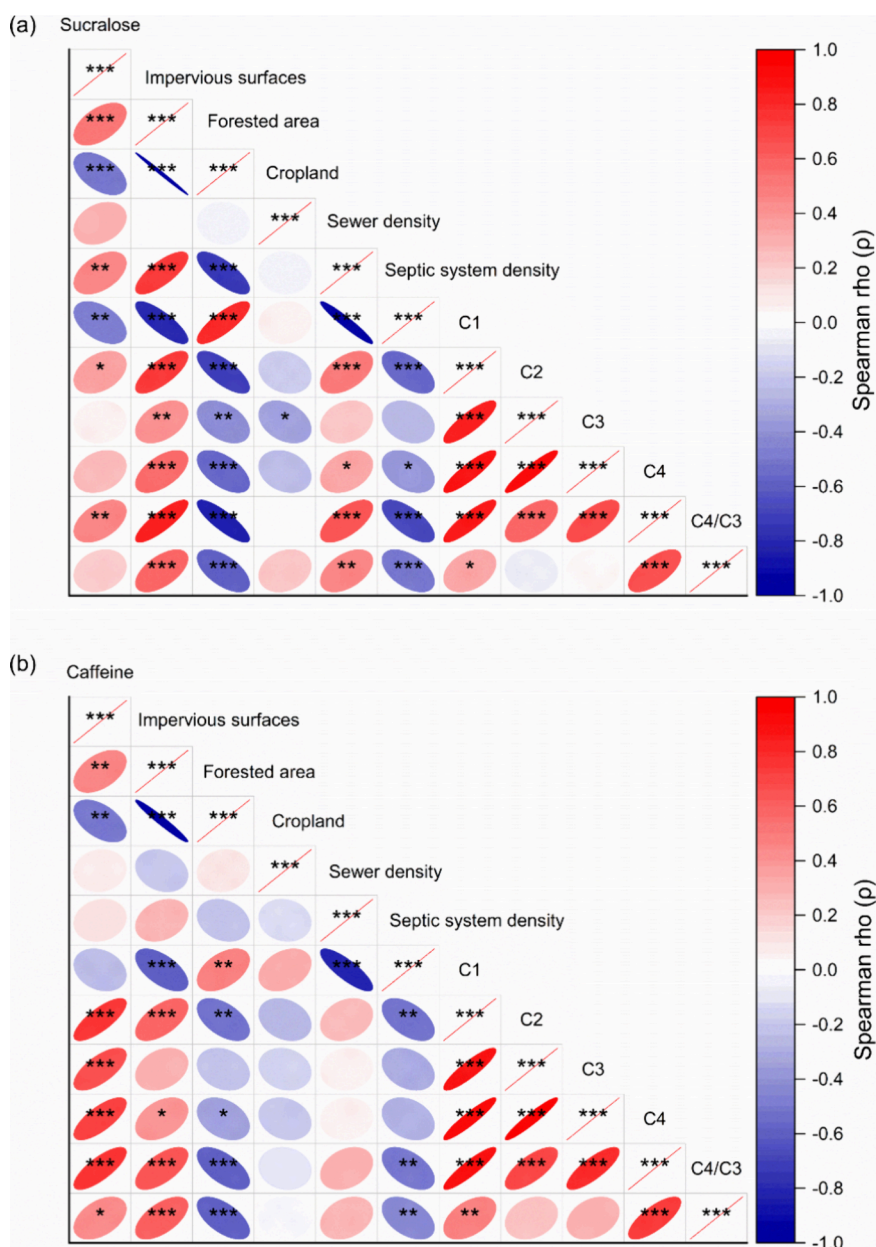


Figure 2. Spearman correlations (color) between measured (a) sucralose and (b) caffeine concentrations, $F_{\max}$ values for EEM-PARAFAC components, land cover features, and wastewater infrastructure in the Jones Falls and Gwynns Falls watersheds. Correlations were evaluated only for sites and months with quantifiable detections of sucralose ($n = 43$) and caffeine ($n = 32$). The *, **, and *** symbols indicate $\rho \leq 0.05$, $\rho \leq 0.01$, and $\rho \leq 0.001$, respectively.

3. RESULTS AND DISCUSSION

3.1. Relationships between Sucralose, Caffeine, EEM-PARAFAC Components, Land Cover, and Wastewater Infrastructure. Sucralose and caffeine were the most detected CECs with frequencies of 34.7% (43/124) and 78.0% (32/41), respectively. The concentrations ranged from <33 to $1,660 \text{ ng L}^{-1}$ for sucralose and <3 to $1,740 \text{ ng L}^{-1}$ for caffeine; full concentration summaries are available in Table S10. These two wastewater indicators were detected in every sample from the GWN-60 urban tributary. Sucralose was primarily found in urban areas, but 19% of detections (8/43) were recorded at rural sites with high septic system density (i.e., JON-32, JON-33), where concentrations ranged from <100 to 753 ng L^{-1} . Caffeine was only detected in one rural sample (i.e., JON-33), and the lower prevalence was attributed to its biodegradable nature.⁵⁴

Other studies have employed these two CECs as wastewater indicators. For example, similar sucralose (e.g., $590\text{--}3,360 \text{ ng L}^{-1}$) and caffeine (e.g., $40\text{--}1,650 \text{ ng L}^{-1}$) levels were reported in the San Diego River watershed, and their presence was attributed to sanitary sewer leaks and overflows.¹¹ The wastewater-impacted Simeo and San Leonardo Rivers in Italy contained $32\text{--}1032 \text{ ng L}^{-1}$ sucralose and $24\text{--}4,478 \text{ ng L}^{-1}$ caffeine.⁵⁵ However, the CEC levels were lower than those reported in Florida, where up to $35,666 \text{ ng L}^{-1}$ sucralose and $30,293 \text{ ng L}^{-1}$ caffeine were measured in surface water and groundwater impacted by wastewater effluent and septic systems, respectively.⁵⁶

Figure 2 reports Spearman correlations between sucralose levels, caffeine concentrations, $F_{\max}$ values for the EEM-PARAFAC components, land cover, and wastewater infra-

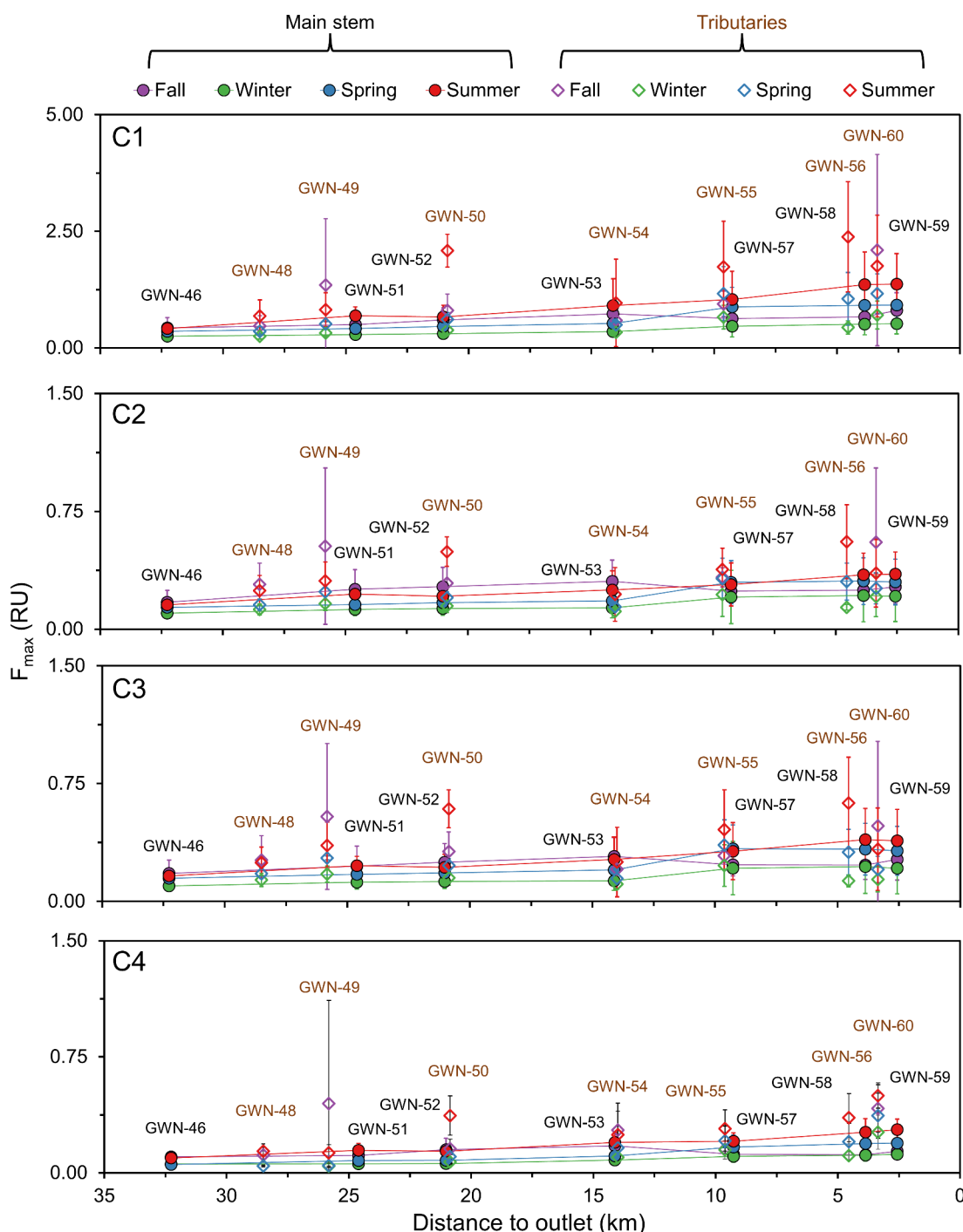


Figure 3. F_{max} values for the C1, C2, C3, and C4 EEM-PARAFAC components as a function of distance from the outlet of the Gwynns Falls. The F_{max} values are reported as seasonal average $\pm$ standard deviation ($n = 3$) according to the following categories: fall (i.e., September, October, November); winter (i.e., December, January, February); spring (i.e., March, April, May); and summer (i.e., June, July, August). The filled, circular symbols connected by lines represent data from the main stem sites, whereas the unfilled diamonds are data from tributaries. Data are reported from upstream (left) to downstream (right).

structure. Sucralose levels were positively correlated to impervious surface area ($\rho = 0.528$, $p < 0.001$), sewer density ($\rho = 0.472$, $p < 0.01$), and F_{max} values for C4 ($\rho = 0.479$, $p < 0.01$); conversely, negative associations were found between sucralose concentrations and forested area ($\rho = -0.503$, $p < 0.001$) and septic system density ($\rho = -0.482$, $p < 0.01$) (Figure 2a). The C4/C3 parameter was positively correlated to impervious surfaces ($\rho = 0.584$, $p < 0.001$) and sewer density ($\rho = 0.477$, $p < 0.001$). These outcomes reinforce the use of C4

and C4/C3 as wastewater indicators and suggest that higher impervious surface area and sewer density lead to greater impacts on water quality. While caffeine concentrations were positively correlated to impervious surface area ($\rho = 0.480$, $p < 0.01$) and the F_{max} values for all EEM-PARAFAC components ($\rho \geq 0.680$, $p < 0.001$), no significant relationships were identified between caffeine, sucralose, and sewer density ($p > 0.05$) (Figure 2b).

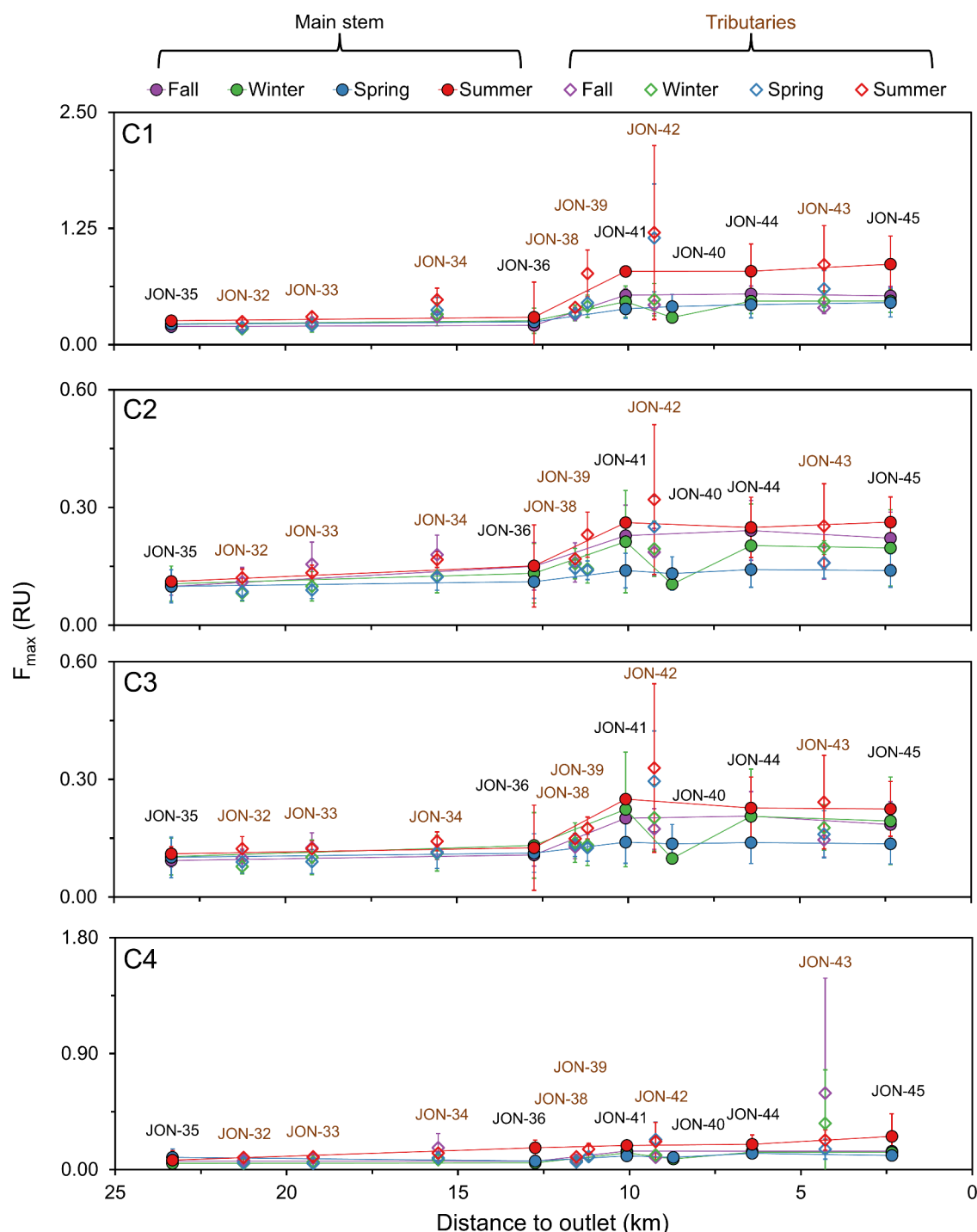


Figure 4. F_{max} values for the C1, C2, C3, and C4 EEM-PARAFAC components as a function of distance from the outlet of the Jones Falls. The F_{max} values are reported as seasonal average $\pm$ standard deviation ($n = 3$) according to the following categories: fall (i.e., September, October, November); winter (i.e., December, January, February); spring (i.e., March, April, May); and summer (i.e., June, July, August). The filled, circular symbols connected by lines represent data from the main stem sites, whereas the unfilled diamonds are data from tributaries. Data are reported from upstream (left) to downstream (right).

3.2. Distribution of F_{max} Values across the Watersheds.

In Figure 3, the F_{max} values for the four EEM-PARAFAC components are plotted as a function of season and upstream distance for the Gwynns Falls watershed. The prevalence of all four components in the main stem increased moving downstream, highlighting FDOM mobilization from upstream drainage areas. Greater and more variable F_{max} values were recorded in the tributaries. Notably, tributaries with small

drainage areas (i.e., $< 15 \text{ km}^2$), such as those that feed GWN-49, GWN-50, GWN-56, and GWN-60, displayed prominent fluorescence intensity in the fall and summer months, suggesting greater vulnerability to intermittent inputs of FDOM. The differences between F_{max} values in the main stem and tributaries were significant ($p < 0.05$) for each EEM-PARAFAC component (Figure S4). This outcome was attributed to the (i) proximity of tributaries to small drainage areas with high

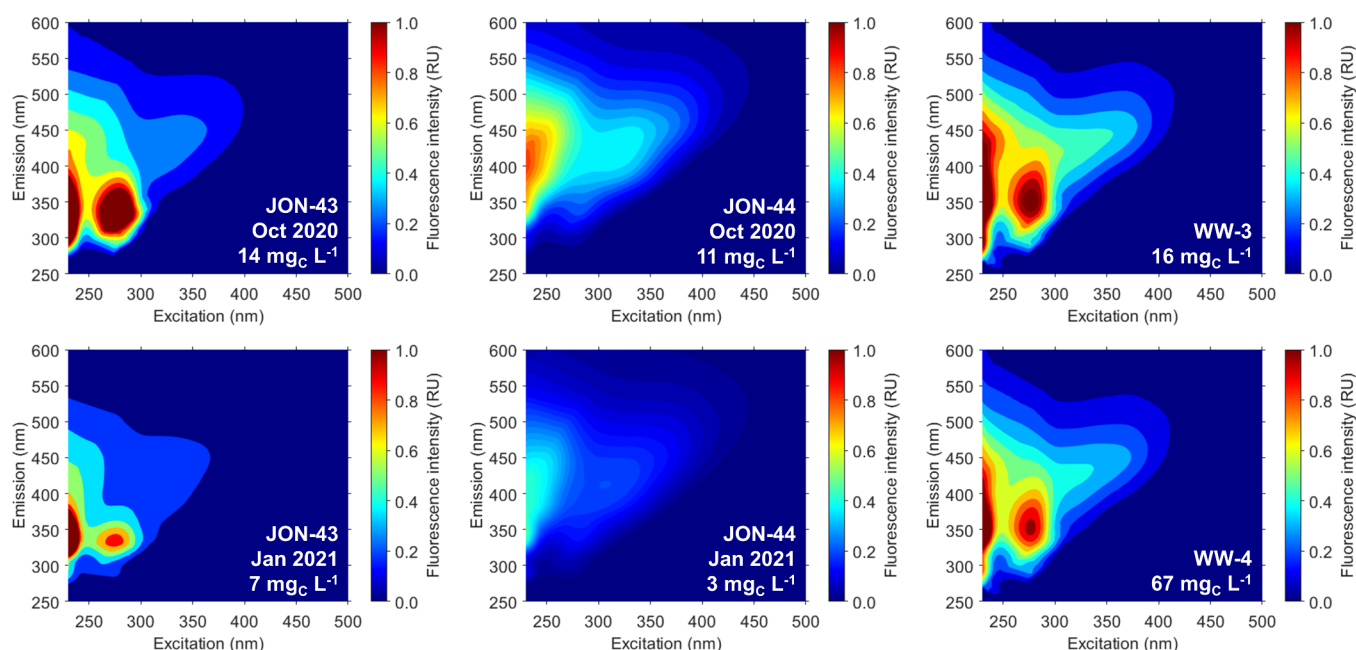


Figure 5. Fluorescence EEMs and DOC content for samples collected at JON-43 (tributary), JON-44 (main stem), and the WW-3 and WW-4 pumping stations.

impervious surface area and sewer density and (ii) main stem sites being more affected by dilution from the large upstream drainage areas. The proximity explanation in (i) was evaluated by PCA (Section 3.3) and cross-covariance analysis (Section 3.4). The dilution explanation in (ii) was supported by the 157% lower 24 h geometric mean discharge measured at the USGS 01589300 stream gage, which is near the upstream GWN-50 and GWN-52 sites, compared to that measured at USGS 01589352, located near the downstream GWN-58 site (Table S2, Figure S5).

Changes in FDOM composition across the Gwynns Falls watershed were also investigated as $F_{\max, \text{frac}, i}$ values (Figure S6). C1 was the dominant fraction and accounted for $51.7 \pm 5.1\%$ of the summed $F_{\max}$ values, with C2, C3, and C4 comprising $18.1 \pm 3.4\%$, $18.4 \pm 3.6\%$, and $11.7 \pm 4.4\%$, respectively. The C1, C2, and C3 components exhibited relatively stable $F_{\max, \text{frac}}$ values, suggesting consistent inputs of natural organic matter. At most main stem sites, the highest $F_{\max}$ values for the four EEM-PARAFAC components were observed in summer (Figure 3). The high $F_{\max}$ values were especially influenced by the samples from June, when high discharge was recorded at upstream ($1.003 \text{ m}^3 \text{ s}^{-1}$) and downstream ($3.442 \text{ m}^3 \text{ s}^{-1}$) locations (Table S2). The lowest $F_{\max}$ values for the EEM-PARAFAC components occurred in winter and were greatly influenced by samples collected in January, when discharge was low both upstream ($0.725 \text{ m}^3 \text{ s}^{-1}$) and downstream ($1.428 \text{ m}^3 \text{ s}^{-1}$). These outcomes suggest that stormwater runoff, which transports natural organic matter, was an important driver of naturally-derived FDOM mobilization.

The $F_{\max}$ and $F_{\max, \text{frac}}$ parameters for C4 in the Gwynns Falls were 0.03–1.22 RU and 3.2–29.5%, respectively. C4 has been identified as an autochthonous component of natural organic matter in pristine environments^{57,58}; however, this tryptophan-like component has also been employed as an indicator in wastewater-impacted streams.⁹ High $F_{\max, \text{frac}}$ values were recorded for C4 in November, April, and July at GWN-54 and in most months at GWN-60, where caffeine and sucralose were

also detected in every sample, suggesting that these tributaries contained elevated wastewater content due to sanitary sewer leaks or overflows in the densely populated drainage areas. Moreover, these locations were estimated to contain 1–35% wastewater based on multilinear regressions developed from EEMs for samples collected between April 2019 and March 2020.⁵⁹ During the current campaign, we observed active repairs to the sanitary sewer system adjacent to GWN-54 in January 2021 (Figure S7); however, no notable differences were observed in the $F_{\max}$ and $F_{\max, \text{frac}}$ profiles for C4 before and after that date, suggesting that further improvements to the sewer system are needed at or upstream of this location to improve water quality.

Figure 4 plots the $F_{\max}$ values of each EEM-PARAFAC component according to season and distance from the outlet of Jones Falls. Major differences were noted between the $F_{\max}$ profiles in the Gwynns Falls and Jones Falls watersheds. First, high $F_{\max}$ values were more prevalent in tributaries of Gwynns Falls. Second, $F_{\max}$ values for C1, C2, and C3 demonstrated abrupt changes between JON-36 and JON-41 for most seasons. Downstream of JON-36, tributaries from densely populated subwatersheds that feed JON-38 and JON-39 join the main stem, flow into Lake Roland, and drain to the JON-41 sampling site. Except for C1, the $F_{\max}$ values were significantly different between the main stem and tributaries (Figure S4); however, the $F_{\max}$ values at JON-38 and JON-39 were lower than those at JON-41, suggesting that biological processes produce natural FDOM in Lake Roland and cause the observed changes between JON-36 and JON-41.

The dominant EEM-PARAFAC component in the Jones Falls was C1, which contributed $48.2 \pm 6.5\%$ of the summed $F_{\max}$, with C2, C3, and C4 accounting for $19.5 \pm 3.3\%$, $18.0 \pm 3.0\%$, and $14.2 \pm 7.3\%$, respectively (Figure S8). While the ranges of $F_{\max}$ (i.e., 0.02–1.62 RU, Table S4) and $F_{\max, \text{frac}}$ (i.e., 4.6–65.2%) for C4 were wider and more variable in the Jones Falls than in the Gwynns Falls, these findings were greatly influenced by measurements at the JON-43 tributary in October and

January. In fact, EEMs measured for those samples demonstrated fluorescence spectra similar to those of raw wastewater (Figure 5), although the C4 peak shifted from excitation/emission at 277/335 to 275/350 nm. At JON-43, the C4/C3 ratio metric parameter, a proposed wastewater indicator,⁹ was 8.35 in October and 7.29 in January (Table S4). Wastewater collected from the WW-3 and WW-4 pumping stations exhibited C4/C3 values of 2.27–4.00, reinforcing the presence of wastewater at JON-43. Spatiotemporal trends for the C4/C3 parameter are presented in Figure S9 for both watersheds, and the data underscore the role of tributaries in delivering wastewater-derived FDOM to the main stem.

3.3. PCA of FDOM Parameters and Geospatial Information. The PCA identified four principal components that explained 80.8% of the variance. The first (PC1) and second (PC2) components explained 40.0% and 21.9% of the variance, respectively; the third (PC3) and fourth (PC4) components explained only 12.4% and 6.5% of the variance, respectively. Figure 6 shows the biplot for PC1 and PC2 generated from the

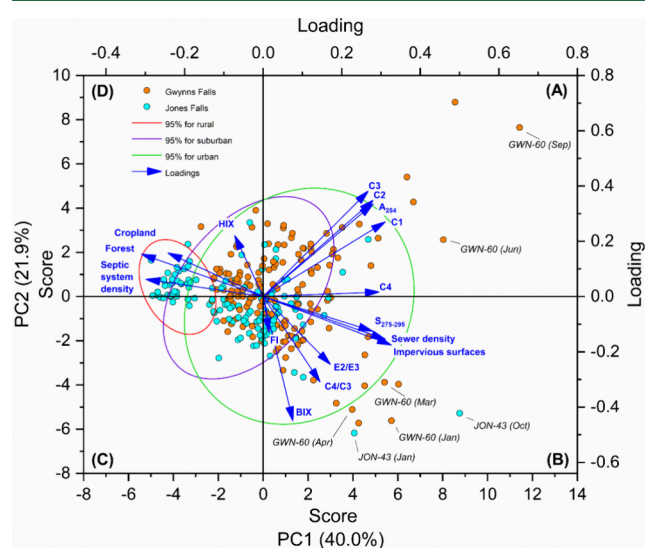


Figure 6. Biplot of PC1 and PC2 from the PCA of land cover, sewer density, septic system density, F_{max} values, and spectral indices. The data (colored circles) were separated by watershed, and the ellipses represent 95% confidence intervals of the data according to rural, suburban, and urban land use. Only sites with more than one outlier are labeled.

F_{max} values for EEM-PARAFAC components, spectral indices (i.e., A_{254} , BIX, E2/E3, FI, HIX, and $S_{275-295}$), land cover features (i.e., cropland, forest, and impervious surfaces), and septic system and sewer densities. The loadings for A_{254} , C1, C2, C3, and C4 were collocated in Quadrant A, whereas Quadrant B contained loadings for BIX, C4/C3, E2/E3, impervious surfaces, sewer density, and $S_{275-295}$. C4 was close to the boundary between Quadrants A and B, which were mostly associated with urban sites. These findings suggested that fresh, labile, low molecular weight FDOM and the C4/C3 wastewater indicator were related to sewer density. Loadings for cropland, forested land, septic system density, and HIX were in Quadrant D and associated with sites in rural subwatersheds.

Certain sites contained more outliers in the PCA (Figure 6). For example, JON-43 data were outside of the 95% confidence domain for urban sites in October and January, when FDOM composition was similar to wastewater (Figure 5). The

aforementioned data were present in Quadrant B, which included loadings for sewer density, impervious surface area, and the C4/C3 wastewater indicator. The C4/C3 values at JON-43 were 8.35 in October and 7.29 in January, the highest for the entire data set (Figure S9). Outliers were also apparent in Quadrant B for the urban tributary at GWN-60 in January, March, and April, and the corresponding C4/C3 parameters were 2.20–2.69, similar to the 2.24–4.26 recorded for wastewater samples from WW-1 and WW-2. These outcomes supported the occurrence of sanitary sewer leaks or overflows at JON-43 and GWN-60. Two PCA outliers at GWN-60 were present in Quadrant A for September and June, which were the only months that this site exhibited below-average C4/C3 (Figure S9). The abundance of C1 was high in those samples (Figure S6), suggesting the possible influence of natural processes. To further explore these phenomena, the cross-covariance of F_{max} values for the EEM-PARAFAC components was evaluated with impervious surfaces, sewer density, and septic system density.

3.4. Cross-Covariance Analysis of EEM-PARAFAC Components, Impervious Surfaces, and Wastewater Infrastructure. The cross-covariance data are plotted as a function of lag distance for the Jones Falls watershed in Figure 7. The F_{max} values of C1 were positively correlated to impervious surface area (+0.85 to +0.15) and sewer density (+1.05 to +0.18) at all lag distances. The correlations were strongly positive for lag distances of 0–3.0 km and minimal for lag distances of 8.5–10.0 km. These results suggest that this terrestrial humic acid-like fluorophore was preferentially mobilized in areas with high impervious surface area and sewer density. Similar trends were observed for C2 and C3, although the cross-covariance with impervious surfaces and sewer density was lower than that for C1 at most lag distances. In fact, C2 and C3 exhibited weak negative correlations (−0.18 to −0.15) with impervious surface area at lag distances of 8.5–10.0 km. The lower cross-covariance for the C1–C3 parameters with impervious surface area and sewer density was explained by the distribution of sampling sites and pour points. At lag distances of 8.5 km, 56% of the lag pairs were composed of main stem sites in suburban and urban areas (i.e., JON-40, JON-41) and pour points located in suburban and rural areas with 7.9–59.7% of impervious surfaces and no sanitary sewer infrastructure (Table S7). The correlations became moderately positive (up to +0.45) for lag distances of 10.5–12.0 km, where more lag pairs involved sampling sites in urban areas (i.e., JON-44, JON-45) and pour points draining suburban areas with 13.3–72.7% impervious surfaces and 0.60–15.1 km^{−1} sewer density. Overall, the cross-covariance profiles of C1, C2, and C3 were quite similar, suggesting common sources of these natural FDOM components throughout the watershed.

The F_{max} values of C4 demonstrated strong, positive correlations with impervious surface area (+0.92 to +0.52) and sewer density (+1.30 to +0.78) at lag distances of 0–3.0 km, for which 75% of lag pairs were situated in urban areas. The correlations at short lag distances were strongest for C4, suggesting that C4 was more closely associated with impervious surfaces and sewer density than C1, C2, and C3. At lag distances of 3.0–5.0 km, the cross-covariance of C4 with impervious surfaces and sewer density was moderately positive (+0.72) to negligible (+0.01). The strong positive correlations at short distances suggested that (i) C4 was derived from nearby sanitary sewer leaks or overflows and (ii) regions with high impervious surface area were more vulnerable to wastewater contamination.

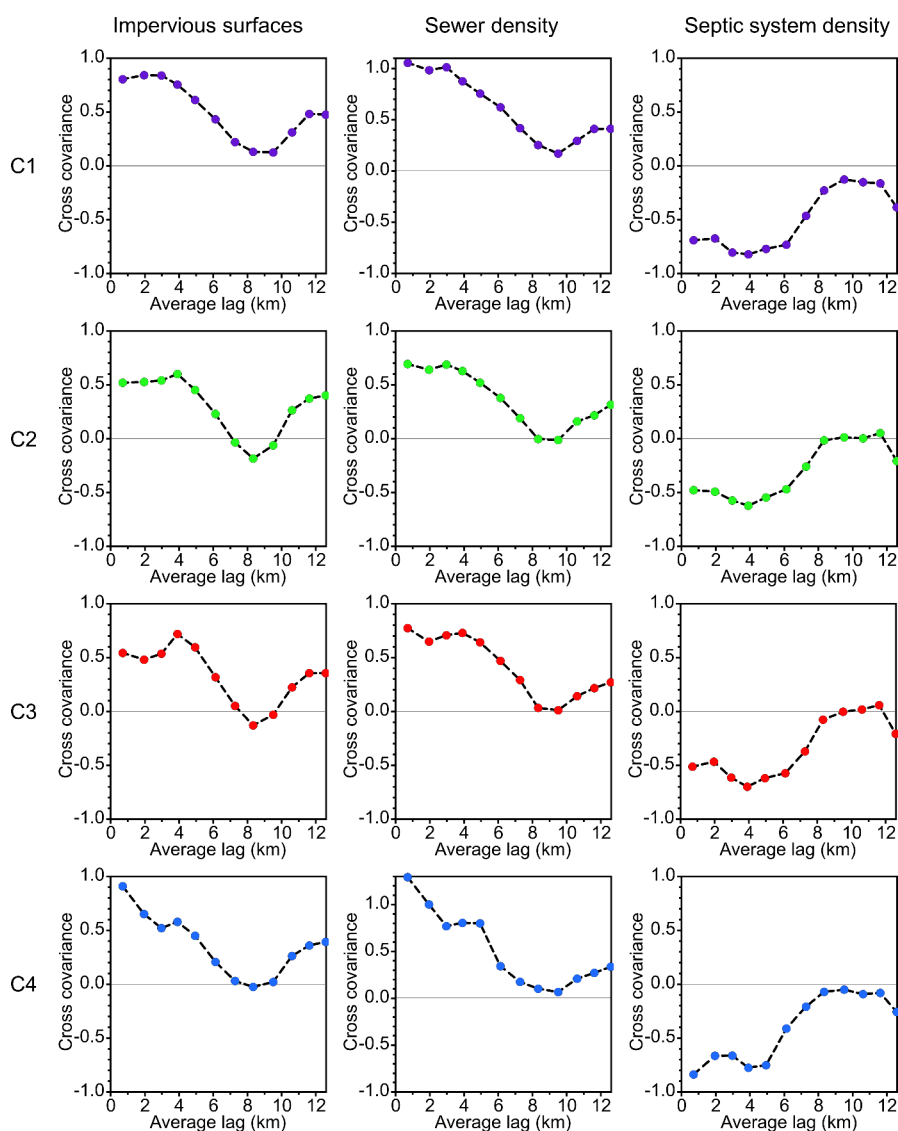


Figure 7. Cross-covariance of F_{max} values for C1, C2, C3, and C4 at the sampling sites with impervious surface area, sewer density, and septic system density at upstream pour points in the Jones Falls. The x -axis accounts for approximately half of the total length of the watershed, in accordance with common practices related to the presence of fewer lag pairs at longer lag distances.

F_{max} values for C1, C2, C3, and C4 were negatively correlated to septic system density, with all four components demonstrating similar trends that involved strong or moderate negative correlations at lag distances of 0–7.2 km. The negative cross-covariance with septic system density was likely influenced by the positive cross-covariance profiles for C1–C4 with impervious surfaces and sewer density; that is, the negative relationship between septic system density and sanitary sewer density in the Jones Falls watershed could have caused the observed cross-covariance trends between the C1–C4 components and septic system density. No apparent correlations were identified at longer lag distances, for which most pour points were located in rural areas with more septic systems. The cross-covariance profiles were similar for all four EEM-PARAFAC components, suggesting that septic systems were not a major contributor of C1–C4 to the rural upstream subwatersheds.

Figure 8 highlights the stronger cross-covariance profiles in the Gwynns Falls. This outcome can be potentially explained by the wider range of impervious surface area ($38.7 \pm 12.9\%$) and sewer density ($8.20 \pm 3.24 \text{ km}^{-1}$) compared to the Jones Falls

(i.e., $22.8 \pm 12.3\%$ and $3.94 \pm 3.92 \text{ km}^{-1}$, respectively). The F_{max} values of C1, C2, C3, and C4 were positively correlated with impervious surface area at all lag distances in the Gwynns Falls; furthermore, the cross-covariance profiles of C1, C2, and C3 were similar, with peaks around 7.0 km. The similar behavior of these components in the two watersheds may derive from fundamental properties that control the fate and transport of natural FDOM. For C4, the correlation with impervious surfaces was strongly positive ($+1.10$) at lag distances of 0–1.0 km, before steadily decreasing at longer distances. The F_{max} values for C1, C2, C3, and C4 were also positively correlated to sewer density at all lag distances. C4 exhibited a particularly strong correlation with sewer density ($+0.95$) at short distances. Wastewater transport from leaking sanitary sewers to the streams may have been facilitated by preferential subsurface flow channels formed along utility pipes and high permeability trenches in urban areas. Kaushal and Belt⁶⁰ suggested that increased interactions between shallow groundwater and underground sanitary sewer pipes in Baltimore can facilitate the transport of carbon and nitrogen to streams. As suggested by

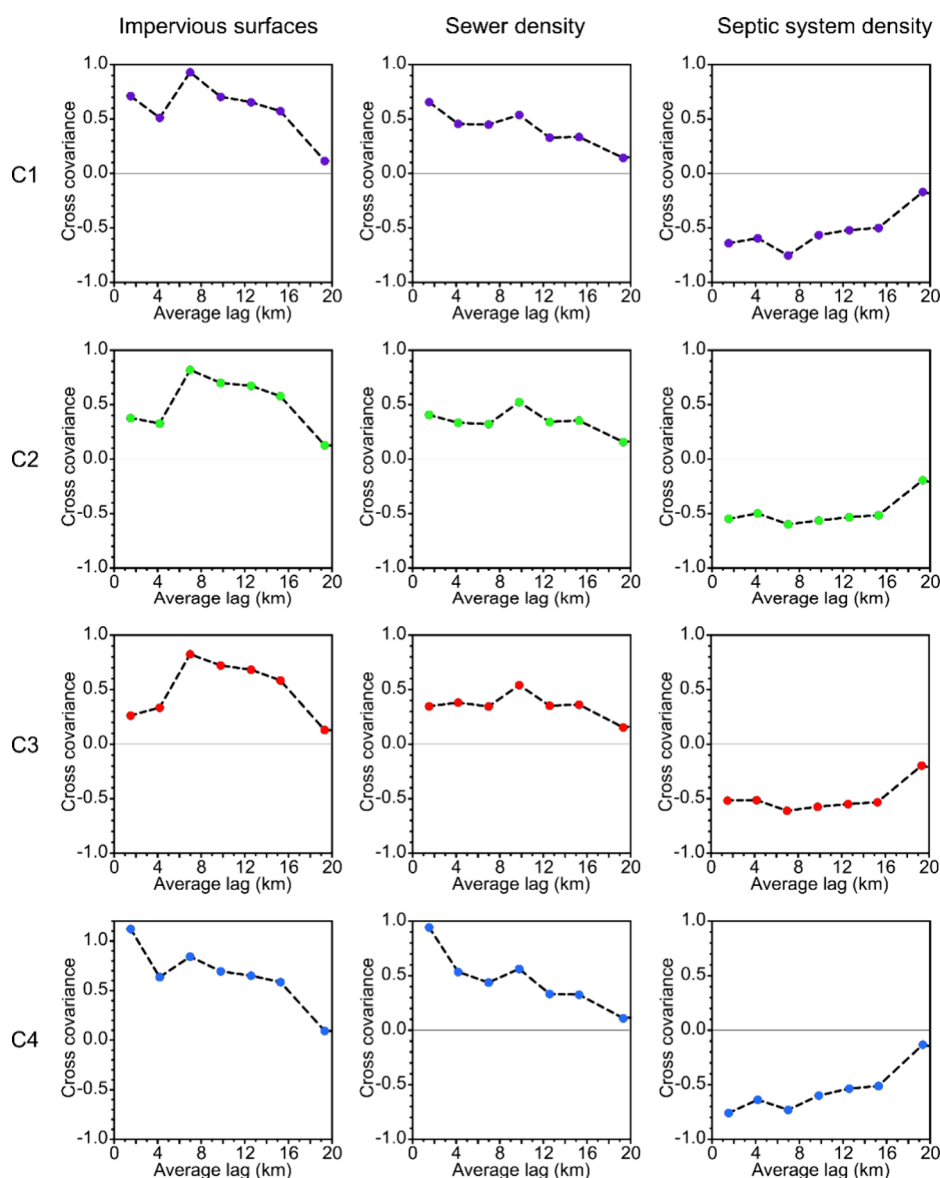


Figure 8. Cross-covariance of F_{max} values for C1, C2, C3, and C4 at the sampling sites with impervious surface area, sewer density, and septic system density at upstream four points in the Gwynns Falls. The x -axis only accounts for approximately half of the total upstream-to-downstream distance of the watershed, in accordance with common practices related to the presence of fewer lag pairs at longer lag distances.

the cross-covariance profiles, sewer density was well correlated to impervious surface area (Spearman $\rho = 0.826$, $p < 0.001$; Figure S10). The aggregate findings confirm that the labile, protein-like C4 component can be used to identify nearby sewer leaks and overflows.

The F_{max} values for C1, C2, C3, and C4 were negatively correlated to septic system density for all lag distances in the Gwynns Falls (Figure 8). A similar result was found for shorter lag distances in the Jones Falls watershed (Figure 7). While the Jones Falls has 3642 septic systems, only 1304 are located in the Gwynns Falls watershed,³⁸ suggesting less potential for downstream influences on FDOM composition. Therefore, the strong-to-moderate negative correlations between F_{max} values and septic system density likely stemmed from inverse relationships between septic system density and the primary FDOM sources. Given the strong positive correlations with impervious surfaces and sewer density, rural sources did not exert any influence on FDOM levels in the Gwynns Falls watershed.

4. ENVIRONMENTAL IMPLICATIONS

Aging wastewater infrastructure is a grand challenge for local and national authorities, especially in urban settings. This study reported novel chemical and geostatistical strategies to identify areas vulnerable to sanitary sewer leaks and overflows. FDOM composition was relatively consistent across the 12-month sampling period for C1, C2, and C3, which displayed fulvic acid- and humic acid-like fluorescence. The intensity of C4, which was composed of tryptophan-like fluorescence and related to signatures observed in municipal wastewater, varied throughout the watersheds. Sites with elevated C4 were generally located in tributaries that drained areas with high impervious surface area and sewer density. This outcome suggests that these features could be employed as measures of failing sewer infrastructure that enable the transport of contaminants and wastewater-derived FDOM to streams via preferential subsurface flow channels. Since wastewater effluent is not discharged into the

Gwynns Falls or Jones Falls watersheds, the C4 signatures were attributed to inputs from sanitary sewer leaks and overflows.

PCA confirmed the associations between EEM-PARAFAC components, land cover features, and wastewater infrastructure. In particular, C4/C3 clustered with sewer density and impervious surfaces in urban areas. Cross-covariance analysis highlighted similar profiles for C1, C2, and C3, indicating the ubiquitous nature and common sources of these FDOM components in the watersheds. However, C4 displayed a markedly different profile, denoting unique sources related to impervious surfaces and sewer density for short lag distances. This outcome suggests that nearby sanitary sewer leaks and overflows could be conveniently detected with C4, which is faster, easier, and cheaper to measure compared to conventional wastewater indicators. As a result, our findings provide a critical tool to differentiate natural organic matter and FDOM derived from raw wastewater that is introduced to urban streams via sanitary sewer leaks and overflows. Sucralose and caffeine were employed as confirmatory metrics, and strong positive correlations were found between these CECs, impervious surfaces, and sewer density. No major trends were observed between FDOM composition and septic system density, but that result may have been influenced by the lower prevalence of sampling sites in areas served by nonsewered sanitation systems; therefore, future studies are recommended to evaluate inputs from septic systems. The overall results of this study reinforce the potential for FDOM-based indicators to serve as effective forensic tools for identifying wastewater inputs to urban streams. Moreover, our reported approach can facilitate the smart design of monitoring studies to *a priori* identify the likely locations of sanitary sewer leaks and overflows in other urban watersheds. Ultimately, this study provided broad new perspectives on wastewater contamination in understudied systems, namely, urban streams that do not receive wastewater effluent, through analysis of highly resolved data sets and adoption of new geostatistical tools to relate FDOM composition to upgradient land cover and wastewater infrastructure. Importantly, the reported methods and outcomes can be extended to inform FDOM composition in other settings.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.est.3c07925>.

Chemical standards; CEC analysis; sampling site characteristics; geospatial and discharge data; lag distances between sampling sites and pour points; 12-month average F_{max} values for each component; sucralose and caffeine concentrations; EEM-PARAFAC component fluorescence spectra and loadings; F_{max} distributions in main stem and tributary sites; spatiotemporal $F_{max,frac}$ data; evidence of sewer repairs; spatiotemporal trends for C4/C3; and correlation between sewer density and impervious surfaces (PDF)

Table S1: names, hydrological details, land cover information, sewer density, and septic system density for the drainage areas associated with the sampling sites; Table S2: 24 h geometric mean discharge at USGS stations in the Jones Falls and Gwynns Falls prior to sample collection; Table S3: method detection limits (MDLs) and method quantitation limits (MQLs) for measured CECs; Table S4: summary of F_{max} and $F_{max,frac}$

values for the EEM-PARAFAC components, spectral data (i.e., A_{254} , E2/E3, FI, BIX, HIX, and $S_{275-295}$), and dissolved organic carbon (DOC) content for each sample; Table S5: lag pairs and distances in the Gwynns Falls watershed; Table S6: lag pairs and distances in the Jones Falls watershed; Table S7: fraction of impervious surfaces, sewer density, and septic system density for upstream drainage areas to each pour point in the Gwynns Falls watershed; Table S8: fraction of impervious surfaces, sewer density, and septic system density for upstream drainage areas to each pour point in the Jones Falls watershed; Table S9: 12 month average F_{max} values for the EEM-PARAFAC components at the 27 sampling sites; Table S10: summary of sucralose and caffeine concentrations (XLSX)

■ AUTHOR INFORMATION

Corresponding Author

Lee Blaney – Department of Chemical, Biochemical, and Environmental Engineering, University of Maryland Baltimore County, Baltimore, Maryland 21250, United States; orcid.org/0000-0003-0181-1326; Phone: +1-410-455-3400; Email: blaney@umbc.edu

Authors

Jahir A. Batista-Andrade – Department of Chemical, Biochemical, and Environmental Engineering, University of Maryland Baltimore County, Baltimore, Maryland 21250, United States; orcid.org/0000-0002-6839-3264

Claire Welty – Department of Chemical, Biochemical, and Environmental Engineering, University of Maryland Baltimore County, Baltimore, Maryland 21250, United States; Center for Urban Environmental Research and Education, University of Maryland Baltimore County, Baltimore, Maryland 21250, United States

Diego Iglesias Vega – Department of Chemical, Biochemical, and Environmental Engineering, University of Maryland Baltimore County, Baltimore, Maryland 21250, United States

Anna McClain – Department of Chemical, Biochemical, and Environmental Engineering, University of Maryland Baltimore County, Baltimore, Maryland 21250, United States

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.est.3c07925>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Science Foundation CAREER Award and Environmental Engineering programs (NSF 1653726). C. Welty's effort was supported by the Urban Critical Zone Network Cluster (NSF 2012340) and The Baltimore Social-Environmental Collaborative Integrated Field Laboratory (DOE DE-SC0023218). We acknowledge the Fulbright foreign student program and National Secretary of Science and Technology (SENACYT) of Panama for the doctoral scholarship to J.A. Batista-Andrade (BEPFUL-2017-2018-002). We also thank Blue Water Baltimore for assistance with sample collection. We acknowledge the anonymous reviewers for their helpful feedback and suggestions.

REFERENCES

- (1) Hudson, N.; Baker, A.; Reynolds, D. Fluorescence analysis of dissolved organic matter in natural, waste and polluted waters—a review. *River Research and Applications* **2007**, *23* (6), 631–649.
- (2) Battin, T. J.; Lauerwald, R.; Bernhardt, E. S.; Bertuzzo, E.; Gener, L. G.; Hall, R. O.; Hotchkiss, E. R.; Maavara, T.; Pavelsky, T. M.; Ran, L.; Raymond, P.; Rosentreter, J. A.; Regnier, P. River ecosystem metabolism and carbon biogeochemistry in a changing world. *Nature* **2023**, *613* (7944), 449–459.
- (3) Quinton, J. N.; Govers, G.; Van Oost, K.; Bardgett, R. D. The impact of agricultural soil erosion on biogeochemical cycling. *Nature Geoscience* **2010**, *3* (5), 311–314.
- (4) Coble, A. A.; Wymore, A. S.; Potter, J. D.; McDowell, W. H. Land Use Overrides Stream Order and Season in Driving Dissolved Organic Matter Dynamics Throughout the Year in a River Network. *Environ. Sci. Technol.* **2022**, *56* (3), 2009–2020.
- (5) Murphy, K. R.; Stedmon, C. A.; Graeber, D.; Bro, R. Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Analytical Methods* **2013**, *5* (23), 6557–6566.
- (6) Hu, X.; Zhou, Y.; Zhou, L.; Zhang, Y.; Wu, L.; Xu, H.; Zhu, G.; Jang, K.-S.; Spencer, R. G. M.; Jeppesen, E.; Brookes, J. D.; Wu, F. Urban and agricultural land use regulates the molecular composition and bio-lability of fluvial dissolved organic matter in human-impacted southeastern China. *Carbon Res.* **2022**, *1* (1), 19.
- (7) Zhang, X.; Yu, H.; Gao, H.; Lu, K.; Liu, D. Explore variations of DOM components in different landcover areas of riparian zone by EEM-PARAFAC and partial least squares structural equation model. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2023**, *291*, No. 122300.
- (8) Wu, Y.-P.; Ji, W.-X.; Liu, F.; Wang, W.-Q.; Cai, M.-H.; Tian, Y.-C.; Zuo, Y.-T.; Shi, P.; Li, Y.; Li, W.-T.; Li, A.-M. Characterizing Molecular Weight Distribution and Optical Properties of Dissolved Organic Matter and Unraveling the Origins of Anthropogenic Fluorophores in Yangtze River and Its Tributaries. *ACS ES&T Water* **2022**, *2* (6), 1056–1064.
- (9) Batista-Andrade, J. A.; Diaz, E.; Iglesias Vega, D.; Hain, E.; Rose, M. R.; Blaney, L. Spatiotemporal analysis of fluorescent dissolved organic matter to identify the impacts of failing sewer infrastructure in urban streams. *Water Res.* **2023**, *229*, No. 119521.
- (10) Mendoza, L. M.; Mladenov, N.; Kinoshita, A. M.; Pinongcos, F.; Verbyla, M. E.; Gersberg, R. Fluorescence-based monitoring of anthropogenic pollutant inputs to an urban stream in Southern California, USA. *Science of The Total Environment* **2020**, *718*, No. 137206.
- (11) Pinongcos, F.; Mladenov, N.; Calderon, J.; Verbyla, M. E.; Kinoshita, A. M.; Gersberg, R.; Batikian, C. M. Chemical and Microbial Markers for Discriminating Sanitary Sewer Contamination in Coastal, Urban Streams. *ACS ES&T Water* **2022**, *2* (10), 1747–1759.
- (12) McClain, A. *Investigating Septic System Impacts on a Rural Subwatershed Using Fluorescent Dissolved Organic Matter and Contaminants of Emerging Concern*. M.S., University of Maryland, Baltimore County, United States -- Maryland, 2023.
- (13) He, H.; Bueno, I.; Kim, T.; Wammer, K. H.; LaPara, T. M.; Singer, R. S.; Beaudoin, A.; Arnold, W. A. Determination of the Antibiotic and Antibiotic Resistance Footprint in Surface Water Environments of a Metropolitan Area: Effects of Anthropogenic Activities. *ACS ES&T Water* **2023**, *3* (2), 387–399.
- (14) Welty, C.; Moore, J.; Bain, D. J.; Talebpour, M.; Kemper, J. T.; Groffman, P. M.; Duncan, J. M. Spatial Heterogeneity and Temporal Stability of Baseflow Stream Chemistry in an Urban Watershed. *Water Resour. Res.* **2023**, *59* (1), No. e2021WR031804.
- (15) Old, G. H.; Naden, P. S.; Harman, M.; Bowes, M. J.; Roberts, C.; Scarlett, P. M.; Nicholls, D. J. E.; Armstrong, L. K.; Wickham, H. D.; Read, D. S. Using dissolved organic matter fluorescence to identify the provenance of nutrients in a lowland catchment; the River Thames, England. *Science of The Total Environment* **2019**, *653*, 1240–1252.
- (16) Li, J.; Yu, Q.; Tian, Y. Q.; Boult, D. F. Effects of Landcover, Soil Property, and Temperature on Covariations of DOC and CDOM in Inland Waters. *Journal of Geophysical Research: Biogeosciences* **2018**, *123* (4), 1352–1365.
- (17) Wasswa, J.; Driscoll, C. T.; Zeng, T. Contrasting Impacts of Photochemical and Microbial Processing on the Photoreactivity of Dissolved Organic Matter in an Adirondack Lake Watershed. *Environ. Sci. Technol.* **2022**, *56* (3), 1688–1701.
- (18) Zhou, Y.; Yu, X.; Zhou, L.; Zhang, Y.; Xu, H.; Zhu, M.; Zhu, G.; Jang, K.-S.; Spencer, R. G. M.; Jeppesen, E.; Brookes, J. D.; Kothawala, D. N.; Wu, F. Rainstorms drive export of aromatic and concurrent bio-labile organic matter to a large eutrophic lake and its major tributaries. *Water Res.* **2023**, *229*, No. 119448.
- (19) Shang, P.; Lu, Y.; Du, Y.; Jaffé, R.; Findlay, R. H.; Wynn, A. Climatic and watershed controls of dissolved organic matter variation in streams across a gradient of agricultural land use. *Science of The Total Environment* **2018**, *612*, 1442–1453.
- (20) Chen, S.; Du, Y.; Das, P.; Lamore, A. F.; Dimova, N. T.; Elliott, M.; Broadbent, E. N.; Roebuck, J. A.; Jaffé, R.; Lu, Y. Agricultural land use changes stream dissolved organic matter via altering soil inputs to streams. *Sci. Total Environ.* **2021**, *796*, No. 148968.
- (21) Granados, V.; Gutiérrez-Cánovas, C.; Arias-Real, R.; Obrador, B.; Harjung, A.; Butturini, A. The interruption of longitudinal hydrological connectivity causes delayed responses in dissolved organic matter. *Science of The Total Environment* **2020**, *713*, No. 136619.
- (22) Howard, D. W.; Hounshell, A. G.; Lofton, M. E.; Woelmer, W. M.; Hanson, P. C.; Carey, C. C. Variability in fluorescent dissolved organic matter concentrations across diel to seasonal time scales is driven by water temperature and meteorology in a eutrophic reservoir. *Aquat. Sci.* **2021**, *83* (2), 30.
- (23) Baltimore County Government Urban Rural Demarcation Line. January, 2023 <https://opendata.baltimorecountymd.gov/datasets/BC-GIS::urban-rural-demarcation-line/explore>.
- (24) USGS Surface-Water Daily Data for the Nation. : USGS 01589440 Jones Falls at Sorrento, MD. https://waterdata.usgs.gov/nwis/dv/?site_no=01589440&agency_cd=USGS&referred_module=sw.
- (25) USGS Surface-Water Daily Data for the Nation. : USGS 01589300 Gwynns Falls at Vila Nova, MD. <https://waterdata.usgs.gov/monitoring-location/01589300/#parameterCode=00065&period=p7D>.
- (26) USGS Surface-Water Daily Data for the Nation. : USGS 01589352 Gwynns Falls at Washington Blvd at Baltimore, MD. https://waterdata.usgs.gov/nwis/dv/?site_no=01589352&agency_cd=USGS&referred_module=sw.
- (27) Chesapeake Bay Program Office. *One-meter resolution land use/land cover dataset for the Chesapeake Bay watershed, 2017/18*. Developed by the Chesapeake Conservancy, U.S. Geological Survey, and University of Vermont Spatial Analysis Lab. <https://www.chesapeakeconservancy.org/conservation-innovation-center/high-resolution-data/lulc-data-project-2022/> 2022.
- (28) Baltimore County Department of Environmental Protection and Sustainability, *Geographic Information System Database for the Jones Falls and the Gwynns Falls*. In Baltimore County, 2021.
- (29) Baltimore City Enterprise Geographic Information Services, *City of Baltimore Geographic Information System Data*. In Baltimore City, 2021.
- (30) Otero, M.; Mendonça, A.; Válega, M.; Santos, E. B. H.; Pereira, E.; Esteves, V. I.; Duarte, A. Fluorescence and DOC contents of estuarine pore waters from colonized and non-colonized sediments: Effects of sampling preservation. *Chemosphere* **2007**, *67* (2), 211–220.
- (31) Kim, E.-A.; Lee, H. K.; Choi, J. H. Effects of a controlled freeze-thaw event on dissolved and colloidal soil organic matter. *Environmental Science and Pollution Research* **2017**, *24* (2), 1338–1346.
- (32) Santos, P. S. M.; Otero, M.; Santos, E. B. H.; Duarte, A. C. Molecular fluorescence analysis of rainwater: Effects of sample preservation. *Talanta* **2010**, *82* (4), 1616–1621.
- (33) Peacock, M.; Freeman, C.; Gauci, V.; Lebron, I.; Evans, C. D. Investigations of freezing and cold storage for the analysis of peatland dissolved organic carbon (DOC) and absorbance properties. *Environmental Science: Processes & Impacts* **2015**, *17* (7), 1290–1301.

- (34) Spencer, R. G. M.; Bolton, L.; Baker, A. Freeze/thaw and pH effects on freshwater dissolved organic matter fluorescence and absorbance properties from a number of UK locations. *Water Res.* **2007**, *41* (13), 2941–2950.
- (35) Hansen, A. M.; Kraus, T. E. C.; Pellerin, B. A.; Fleck, J. A.; Downing, B. D.; Bergamaschi, B. A. Optical properties of dissolved organic matter (DOM): Effects of biological and photolytic degradation. *Limnology and Oceanography* **2016**, *61* (3), 1015–1032.
- (36) Peuravuori, J.; Pihlaja, K. Molecular size distribution and spectroscopic properties of aquatic humic substances. *Anal. Chim. Acta* **1997**, *337* (2), 133–149.
- (37) McKnight, D. M.; Boyer, E. W.; Westerhoff, P. K.; Doran, P. T.; Kulbe, T.; Andersen, D. T. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnology and Oceanography* **2001**, *46* (1), 38–48.
- (38) Huguet, A.; Vacher, L.; Relexans, S.; Saubusse, S.; Froidefond, J. M.; Parlanti, E. Properties of fluorescent dissolved organic matter in the Gironde Estuary. *Org. Geochem.* **2009**, *40* (6), 706–719.
- (39) Murphy, K. R.; Stedmon, C. A.; Wenig, P.; Bro, R. OpenFluor—an online spectral library of auto-fluorescence by organic compounds in the environment. *Analytical Methods* **2014**, *6* (3), 658–661.
- (40) Romero, C. M.; Engel, R. E.; D'Andrilli, J.; Chen, C.; Zabinski, C.; Miller, P. R.; Wallander, R. Bulk optical characterization of dissolved organic matter from semiarid wheat-based cropping systems. *Geoderma* **2017**, *306*, 40–49.
- (41) D'Andrilli, J.; Foreman, C. M.; Sigl, M.; Priscu, J. C.; McConnell, J. R. A 21 000-year record of fluorescent organic matter markers in the WAIS Divide ice core. *Climate of the Past* **2017**, *13* (5), 533–544.
- (42) Stedmon, C. A.; Markager, S. Tracing the production and degradation of autochthonous fractions of dissolved organic matter by fluorescence analysis. *Limnology and Oceanography* **2005**, *50* (5), 1415–1426.
- (43) Vines, M.; Terry, L. G. Evaluation of the biodegradability of fluorescent dissolved organic matter via biological filtration. *AWWA Water Sci.* **2020**, *2* (5), No. e1201.
- (44) Osburn, C. L.; Oviedo-Vargas, D.; Barnett, E.; Dierick, D.; Oberbauer, S. F.; Genereux, D. P. Regional Groundwater and Storms Are Hydrologic Controls on the Quality and Export of Dissolved Organic Matter in Two Tropical Rainforest Streams, Costa Rica. *Journal of Geophysical Research: Biogeosciences* **2018**, *123* (3), 850–866.
- (45) Murphy, K. R.; Stedmon, C. A.; Waite, T. D.; Ruiz, G. M. Distinguishing between terrestrial and autochthonous organic matter sources in marine environments using fluorescence spectroscopy. *Marine Chemistry* **2008**, *108* (1), 40–58.
- (46) Cawley, K. M.; Butler, K. D.; Aiken, G. R.; Larsen, L. G.; Huntington, T. G.; McKnight, D. M. Identifying fluorescent pulp mill effluent in the Gulf of Maine and its watershed. *Mar. Pollut. Bull.* **2012**, *64* (8), 1678–1687.
- (47) Murphy, K. R.; Hambly, A.; Singh, S.; Henderson, R. K.; Baker, A.; Stuetz, R.; Khan, S. J. Organic Matter Fluorescence in Municipal Water Recycling Schemes: Toward a Unified PARAFAC Model. *Environ. Sci. Technol.* **2011**, *45* (7), 2909–2916.
- (48) Mitchelmore, C. L.; He, K.; Gonsior, M.; Hain, E.; Heyes, A.; Clark, C.; Younger, R.; Schmitt-Kopplin, P.; Feerick, A.; Conway, A.; Blaney, L. Occurrence and distribution of UV-filters and other anthropogenic contaminants in coastal surface water, sediment, and coral tissue from Hawaii. *Science of The Total Environment* **2019**, *670*, 398–410.
- (49) Maryland Department of the Environment Maryland LiDAR Statewide, DEM Meters. November 15, 2021 <https://data.imap.maryland.gov/datasets/1c6ce663eb3b499b9495010cb8c89df6/explore?location=38.806150%2C-77.265300%2C8.42>.
- (50) Deutsch, C. V.; Journel, A. G. *Geostatistical software library and user's guide*. Oxford University Press 1997, (Second ed.), 369.
- (51) Goovaerts, P. *Geostatistics for Natural Resources Evaluation*. Oxford University Press on Demand: 1997.
- (52) Isaaks, E. H.; Srivastava, R. M. *An Introduction to Applied Geostatistics*. Oxford University Press: 1989.
- (53) Steffy, L. Y.; McGinty, A. L.; Welty, C.; Kilham, S. S. Connecting Ground Water Influxes with Fish Species Diversity in an Urbanized Watershed. *JAWRA Journal of the American Water Resources Association* **2004**, *40* (5), 1269–1275.
- (54) Buerge, I. J.; Poiger, T.; Müller, M. D.; Buser, H.-R. Caffeine, an Anthropogenic Marker for Wastewater Contamination of Surface Waters. *Environ. Sci. Technol.* **2003**, *37* (4), 691–700.
- (55) Sgroi, M.; Roccaro, P.; Korshin, G. V.; Vagliasindi, F. G. A. Monitoring the Behavior of Emerging Contaminants in Wastewater-Impacted Rivers Based on the Use of Fluorescence Excitation Emission Matrixes (EEM). *Environ. Sci. Technol.* **2017**, *51* (8), 4306–4316.
- (56) Henderson, A.; Ng, B.; Landeweer, S.; Quinete, N.; Gardinali, P. Assessment of Sucralose, Caffeine and Acetaminophen as Anthropogenic Tracers in Aquatic Systems Across Florida. *Bull. Environ. Contam. Toxicol.* **2020**, *105* (3), 351–357.
- (57) Logozzo, L. A.; Hosen, J. D.; McArthur, J.; Raymond, P. A. Distinct drivers of two size fractions of operationally dissolved iron in a temperate river. *Limnology and Oceanography* **2023**, *68* (6), 1185–1200.
- (58) Retelletti Brogi, S.; Jung, J. Y.; Ha, S.-Y.; Hur, J. Seasonal differences in dissolved organic matter properties and sources in an Arctic fjord: Implications for future conditions. *Science of The Total Environment* **2019**, *694*, No. 133740.
- (59) Batista-Andrade, J. A.; Iglesias Vega, D.; McClain, A.; Blaney, L. Using multilinear regressions developed from excitation-emission matrices to estimate the wastewater content in urban streams impacted by sanitary sewer leaks and overflows. *Science of The Total Environment* **2024**, *906*, No. 167736.
- (60) Kaushal, S. S.; Belt, K. T. The urban watershed continuum: evolving spatial and temporal dimensions. *Urban Ecosystems* **2012**, *15* (2), 409–435.