

FROM LAND TO STREAM: AN ASSESSMENT OF WATERSHED-SCALE
BIOGEOCHEMICAL INTERACTIONS AT THE STREAM-GROUNDWATER INTERFACE

By

Joseph Albert Lee-Cullin

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Geological Sciences —Doctor of Philosophy

2020

ABSTRACT

FROM LAND TO STREAM: AN ASSESSMENT OF WATERSHED-SCALE BIOGEOCHEMICAL INTERACTIONS AT THE STREAM-GROUNDWATER INTERFACE

By

Joseph Albert Lee-Cullin

The stream-groundwater interface (SGI) is typically studied at scales <1000 m, whereas watershed management needs to understand outcomes of stream-groundwater interactions at scales of tens of kilometers. As a ubiquitous reactive ecotone, the SGI plays a critical role in biogeochemical cycling across stream networks. Process-based models have examined these needed larger scales, but there is a distinct absence of field data to validate modeling efforts. Due to the paucity of these data, this dissertation sets out to be among the first efforts to sample across the SGI at the watershed scale. The three primary goals are to 1) identify the grain of measurements needed to assess the SGI across a stream network, 2) determine whether landscape biogeochemical signals are modified by the SGI at the watershed scale, and finally 3) to investigate whether the SGI acts consistently to modify biogeochemical inputs from the landscape. This work was done through the lens of common solutes found in streams, specifically focusing on dissolved organic carbon, an important driver of many stream biogeochemical reactions.

In Chapter 1, I evaluate how to sample the SGI across a stream network. I test two fundamental sampling schemes, focusing on local heterogeneity (i.e., features or plots) compared to longitudinal heterogeneity (i.e., stream reaches). There was previously no clear guidance as to which kind of sampling scheme would be most appropriate. This was necessary because sampling in the SGI is time and labor consuming, and one must determine how to distribute a finite number of sampling points. These data were collected in two synoptic sampling campaigns

in a third-order stream network in southwest Michigan. Here, I found that longitudinal sampling accounted for similar stream network variance as localized heterogeneity. Therefore, it may be useful to focus on longitudinal sampling as local sampling becomes redundant.

In Chapters 2 and 3, I investigate, first, how different watershed delineations are used to understand landscape contributions to the biogeochemical signal of the stream water. I compared surface watershed and novel groundwatershed delineations to evaluate which areal delineation of the landscape would best predict stream biogeochemistry. Both delineations were then used to investigate how the biogeochemical signal propagated from the land into the SGI, and whether this signal was modified as it entered the SGI by way of spatially lagged linear models. I found that both watersheds were comparable, and therefore the groundwatersheds may be appropriate for lowland watersheds, with strongly upwelling groundwater. Further, I found that the landscape signal found in surface waters through linear modeling was modified as models were propagated into the SGI, given the decreasing performance of linear models in the stream subsurface.

In Chapter 4, I evaluated the SGI effects from multiple watersheds on various sources of dissolved organic carbon and its molecular components. I used mixed-effects linear models to test if there was a consistent modification of dissolved organic carbon across a multitude of SGIs as compared to stream water alone. I found that for most optical properties tested, the interaction between the specific carbon source and the SGI sediments was important, functionally obscuring the effects of the sediment alone. The one exception to this was a proxy for humic substances called Peak T, for which the SGI sediment had a significant, identifiable effect. These results indicate that it may be difficult to make broad generalizations about the function of SGIs, where local heterogeneity might be an important consideration.

This dissertation is dedicated to Mordred the hedgehog. May he receive mealworms to his tiny heart's content across the Rainbow Bridge

ACKNOWLEDGMENTS

Pursuing the PhD is so much more than writing a dissertation. Many, many people have helped to shape me into an entirely new person through these last several years. There is not enough room to thank every individual who made a difference throughout this process, but I will try my best.

First and foremost, I must send a very appreciative ‘thank you’ to my advisor, Jay Zarnetske. I learned so much under his tutelage and have become a better researcher because of it. I would not have made it to the end of this program without him pushing me to be better but also his understanding of what it means to be a person balancing research, teaching, and personal life. Jay is producing really cool research and shaping really cool researchers, and I am glad to be counted among his mentees and peers.

Thank you also to each of my committee members. Steve Hamilton aided with establishing field sites, developing sampling regimes, and analyzing biogeochemical data. I would not have known where to start without Steve’s guidance. Anthony Kendall, who is a fountain of knowledge and an expert at GIS, helped me to ask new, exciting questions. And Phani Mantha, whose wealth of knowledge regarding streams and groundwaters, provided a perspective I would never have reached on my own.

Thank you to the Watershed Science and Hydroecology Lab, who were there as friends and colleagues throughout this process, from venting sessions to editing manuscripts. This includes, in no particular order, Sydney Ruhala, Tyler Hampton, Emma Haines, Stephen Plont, Rachel Geiger, Evan Wiewiora, Abraham Downer, Chenxi Li, Megan Duda, Sam Cairns, Elizabeth Tripp, Christian Poelstra, Ben Abbott, and Arial Shogren,

Beyond my lab group, the graduate students across MSU had a profound influence on me throughout my program through their dedication to Earth science, organizing social events, and generally being the most supportive peers I could ask for. A very special thank you to the Earth and Environmental Sciences Graduate Student Organization and especially its initiators, Alex Kuhl, Ben Brugman, Caitlin Kirby, Erin Haacker, Mary Sabuda, and Sydney Ruhala. The GSO helped to facilitate such an enormous growth in the sense of community in the department, which in turn made us all better researchers and teachers. A BIG thank you to the Graduate Employees Union, particularly to Jordan Lindsay, Jolyse Race, Sara Bijani, Nick Rowe, Cathleen Fry, Kathryn Frens, and Kevin Bird, whose combined leadership created a more equitable and just environment for graduate employees.

There is such a hodgepodge of people at MSU that don't fit neatly into any one category, so this paragraph is for them. Thank you to Dr. Lee Wang at MSU Counseling and Psychiatric Services, who prescribes the drugs that keeps the anxiety and depression at bay. I have so much gratitude for Andrew Deinhardt and Hope Akaeze, who were graduate student employees in Center for Statistical Training and Consulting, both of whom made significant contributions to my development as a researcher heavily reliant on statistical modeling. Thank you to Osvaldo Hernandez and all the ISP203L TAs and students for helping me to find and refine my passion as an instructor. And of course, to the Earth and Environmental Sciences office staff past and present, who are arguably the most important people in the department, Brittany Walter, Pam Robinson, Dallas Coryell, Elizabeth McElroy, and Ami McMurphy.

Finally, thank you to my friends and family throughout the world. Thank you to my parents, Randy and Nina Cullin, who always encouraged me to ask questions and delve deeper into whatever was piquing my curiosity. Thank you to my dearest friends Matt McCune, Alex

and Sarah Schomers, Becca Good, and Ron and Sarah Reynolds, who always built me up and instilled the confidence one needs to pursue a PhD. Thank you to my cat Harley, who has been instrumental in providing affection over the last two years. Last, and certainly not least, thank you to my loving spouse Jei who has been a pillar of emotional and intellectual support throughout the last decade.

PREFACE

The surface water-groundwater (SGI) interface, where surface and sub-surface waters mix, is an ecotone that encapsulates much of the metabolic activity controlling the fate and transport of reactive solutes across stream networks (Grimm and Fisher 1984). The unique chemical and physical gradients observed in this region produce metabolic rates that are regularly orders of magnitude greater than in surface waters alone (Findlay and Sobczak 1996). Because of the high level of reactivity in the SGI, many researchers have dedicated time to understanding when and where this zone exists and have worked towards elucidating whether it has an overall function across catchments. This has produced interesting observations such as the overall ubiquity of this interface and its role as a control point across watersheds (e.g., McClain et al. 2003; Bernhardt et al. 2017). Others have proposed that contact time of stream water within the SGI will be the dominant factor regarding net changes of pore water chemistry (e.g., Valett et al. 1996; Pinay et al. 2009; Zarnetske et al. 2012; Briggs et al. 2014).

Over the last 30+ years, several models of water and solute transport in streams have included terms for water temporarily stored in the SGI or surface storage zones, collectively referred to as transient storage (e.g., Bencala et al. 1983; Runkel et al. 1998; Wörman and Wachniew 2007; Ward et al. 2017). The goals of these models are diverse. Some work towards separating the surface and stream-groundwater compartments (e.g., Briggs et al. 2009; Jackson et al. 2012), while others are working towards addressing metabolically active transient storage zones, allowing us to hone in on reactive characteristics in these transient storage zones (Haggerty 2009; Argerich et al. 2011). Still other models have been used to identify physical parameters that control biogeochemical cycling within the SGI (Gomez et al. 2012; Zarnetske et al. 2012). While these models are useful in furthering our understanding of certain solutes (e.g.,

denitrification of NO_3^- across large river networks, Kiel and Cardenas 2014; Gomez-Velez et al. 2015), they have 1) largely ignored solutes like dissolved organic carbon (DOC), because the fate and function of DOC is poorly understood and 2) not been validated with field data at the scales of interest.

Despite tremendous modeling work and the collective effort of researchers to understand what SGIs are and their function in the landscape, there is a paucity of data from this ecotone at larger scales. Most studies within the SGI have been conducted at scales smaller than 1 km (Ward 2016), meaning that the primary focus across all studies is at points, features, or reaches. In contradiction to where our knowledge of SGIs is focused, modelers, land managers, and society at-large need an increased understanding at scales exceeding 1 km, particularly at the watershed scale (Krause et al. 2011).

There is evidence that aquatic respiration, fueled by organic carbon, tends to decrease from headwaters to downstream locations (Battin et al. 2008). This evidence falls in line with proposed conceptual frameworks such as the River Continuum Concept (Vannote et al. 1980) and the Reactive Pipe Model (Cole et al. 2007). Given this respiration, SGIs have been estimated to account for as much as 88% of ecosystem respiration in some stream systems (Kaplan and Newbold 2000). Although researchers have identified scales of various process controls on surface water biogeochemistry (e.g., McGuire et al. 2014), there is a paucity of information on the scales of various process controls on SGI biogeochemistry (Boano et al. 2014), due, in part to the absence of field data reflecting this scale.

This dissertation is constructed as three distinct, but complementary parts, with the goal of addressing three critical questions: 1) How do we sample at the SGI to gain meaningful insight at the watershed scale? 2) What is the role of the SGI as a biogeochemical reactor at this

watershed scale? 3) Is there consistency across SGIs in processing unique DOC sources derived from the surrounding landscape? I have conceptualized this as a progression addressing how we measure aquatic biogeochemistry at a relatively understudied scale of the SGI, once we have these data how do we gain insight into the processes occurring at this scale, and finally I flip these studies on their head and scale back down to investigate how individual DOC sources interact with the SGI. The first two parts investigate SGI controls on landscape inputs while the last part investigates landscape input controls on the SGI.

In this dissertation, I rely largely on empirically based statistical analyses of field and laboratory data. The data used in these statistical models were collected from studies that I designed in conjunction with my laboratory group. Across all four chapters, there is a consistent theme of using this method of analysis to develop a conceptual framework of pushing the boundaries of the SGI community's understanding of the SGI at a larger scale than we have historically studied.

The first chapter, "Toward measuring biogeochemistry within the stream-groundwater interface at the network scale: an initial assessment of two spatial sampling strategies," sets out to resolve an issue of the spatial resolution needed when sampling the SGI at the watershed scale. Collecting samples from the SGI is time and labor consuming, and it is therefore important to know whether one must focus on local heterogeneity or if a coarser scale is permissible. Through a comparison of two synoptic SGI studies across the Augusta Creek watershed, this chapter investigates the variance of measurements when focusing on fewer overall sites across the watershed but with more samples at any given point as compared to fewer samples at any given point but with greater focus on longitudinal sampling. This chapter was published in the journal *Limnology and Oceanography: Methods* and is cited as Lee-Cullin et al. (2018)

The second part, composed of chapters two and three, assesses first 1) the methods by which landscape units are delineated within a watershed and 2) how the biogeochemistry of those landscape units contributes to the stream water quality, and are subsequently modified by the SGI. Chapter two assesses the use of a typical surface watershed (i.e., the watershed defined by surface topography) compared to the use of a groundwatershed (i.e., the watershed defined by groundwater level topography) as a method of landscape delineation across the synoptic sampling campaign from chapter one. The use of a groundwatershed is novel and potentially useful delineation tool in lowland watersheds, where groundwatersheds are often distinct as compared to surface watersheds (Boutt et al. 2001).

This second study was done through the lens of biogeochemistry, where landscape units within each delineation are used in statistical models to predict biogeochemical parameters of the stream surface water (after e.g., Soranno et al. 1996; Dillon and Molot 1997; Hollister et al. 2008). This entailed the implementation of a mixed regressive-spatial autoregressive model to account for spatial autocorrelation (Anselin 1988; Overmars et al. 2003), or as a geographer might tell a person, “things that are close together are more similar than things that are far apart.” Sampling points in the stream that are near to each other, especially those in the same stream reach, have similar controls and could lead to improperly weighted predictions in a statistical model. This chapter ultimately compares the predictive power of surface watershed and groundwatershed delineations.

The third chapter builds upon the second. Here the statistical models of surface water biogeochemistry are used as representing the landscape biogeochemical ‘fingerprints.’ The models are then propagated into six discrete depths within the SGI to evaluate whether this landscape fingerprint is preserved in the interface, or if it is effectively removed, creating a new,

unique subsurface fingerprint. This was done simply by using the selected model parameters and seeing if the predictive power in the subsurface at 2.5 through 20 cm was comparable to the predictive power in the surface water.

Finally, the fourth chapter focuses on specific landscape inputs into the SGI, and how DOC from these inputs is ultimately broken down across several different watersheds. This was a laboratory bioassay study (after Dahm 1981) assessing whether watershed-specific DOC would break down similarly across watersheds when exposed to SGI sediments as compared to the same DOC when exposed only to stream water. There were several watersheds selected and for each of these watersheds, two to four DOC sources were selected. They were then allowed to decompose in a controlled setting of treatments with only stream water and treatments of stream water and sediment from their specific watershed. This decomposition across different watersheds with different DOC sources was assessed using a mixed-effects linear model (Baayen et al. 2008a; Bates et al. 2015) which uses model parameters that do not vary concomitant with model parameters that vary randomly. This modeling effort was meant to address whether changes in the molecular composition of DOC were consistent across watersheds and across DOC sources derived from these watersheds.

TABLE OF CONTENTS

LIST OF TABLES	xvi
LIST OF FIGURES	xvii
KEY TO ABBREVIATIONS	xxi
 CHAPTER 1: TOWARD MEASURING BIOGEOCHEMISTRY WITHIN THE STREAM- GROUNDWATER INTERFACE AT THE NETWORK SCALE: AN INITIAL ASSESSMENT OF TWO SPATIAL SAMPLING STRATEGIES	 1
Abstract	1
Introduction.....	2
Materials and Methods.....	8
Site description.....	8
Sampling schemes.....	9
Sample and data collection	12
Data analysis	12
Results.....	15
Local Sampling scheme results.....	16
Longitudinal Sampling scheme results	20
Discussion.....	20
Guiding future sampling	20
Different sampling resolution concerns	23
A need for more assessment of sampling schemes	24
Conclusions.....	27
Acknowledgments	28
 CHAPTER 2: GROUNDWATERSHED VERSUS A SURFACE WATERSHED APPROACH TO LANDSCAPE MODELS OF STREAM CHEMISTRY: AN INITIAL ASSESSMENT OF EACH APPROACH IN A LOWLAND STREAM NETWORK.....	 29
Abstract	29
Introduction.....	30
Materials and Methods.....	35
Site Description.....	35
Synoptic Sampling Campaign.....	36
Watershed delineations and land use	37
Results.....	42
General stream network surface water biogeochemistry conditions	42
Watersheds and land use results.....	43
Model results.....	47
Discussion.....	49
Surface watershed and groundwatershed differences and potential implications.....	49
Stream Biogeochemistry: Surface watershed linear models vs. groundwatershed linear models	52

Linear model performance for subsets of LULC and stream-order.....	56
Limitations of this study	58
Model Performance vs. Stream Order and Dominant Land Use	60
Conclusion	62
Acknowledgments	63
 CHAPTER 3: EVIDENCE FOR STREAM WATER-GROUNDWATER INTERFACE MODIFICATION OF LANDSCAPE BIOGEOCHEMICAL SIGNATURES.....	
Abstract.....	64
Introduction.....	64
Materials and Methods.....	67
Site Description.....	67
Synoptic Sampling Campaign.....	68
Watershed Delineation & Land Use	70
Linear Models to Assess Landscape Controls on Stream and SGI Chemistry	72
Results.....	75
Stream Network Biogeochemistry	75
Surface versus Groundwater Watersheds	76
Linear Models	77
Discussion.....	78
Conclusion	82
Acknowledgments	83
 CHAPTER 4: EVALUATING THE SURFACE WATER-GROUNDWATER INTERFACE FUNCTION IN TRANSFORMING AND DEGRADING OF DISSOLVED ORGANIC CARBON ACROSS SIX MIDWESTERN STREAMS.....	
Abstract.....	84
Introduction.....	84
Materials and Methods.....	88
Field Sites and Sample collection/preparation.....	88
Laboratory bioassays	90
DOC analysis	90
Mixed effects model	92
Results.....	93
DOC optical properties and Principal Components Analysis.....	93
Mixed effects models	96
Discussion.....	97
DOC Optical Properties	97
Mixed Effects Models.....	98
Conclusion	102
Acknowledgements.....	103
APPENDICES	105
APPENDIX A: Chapter 2 Supplemental Tables and Figures.....	106
APPENDIX B: Chapter 4 Supplemental Table	112

REFERENCES	114
------------------	-----

LIST OF TABLES

Table 1 – Outline of site locations and the types of DOC sources used across each site.	88
Table 2 - Results of each mixed effects model for individual optical properties. An 'x' denotes significance ($p < 0.05$) of model parameters.	97
Table 3 - Information on groundwatersheds (GWshed) and surface watersheds (SWshed) across all 39 sampling points in the Augusta Creek watershed	106
Table 4 - Best fit models for all measurements of surface watersheds (S) and groundwatersheds (G). A value in a column indicates that a the landcover category—Barren (B), Cultivated Crops (C), Emergent Herbaceous Wetland (EHW), Evergreen Forest (EF), Hay/Pasture (HP), Deciduous Forest (DF), Herbaceous (H), Open Water (OW), Shrub Scrubs (SS), Woody Wetland (WW), and Developed (D)—indicates coefficient value.....	107
Table 5 - Number of models passing our criteria for each biogeochemical measurement.	108
Table 6 - Average optical properties for each medium/DOC source combination.....	112

LIST OF FIGURES

- Figure 1 - Simplified plan view of stream network reaches illustrating the main conceptual differences for Local Sampling (L) and Longitudinal Sampling (R) sampling schemes. Local Sampling represents high characterization of local heterogeneity with low characterization of longitudinal heterogeneity, while Longitudinal Sampling has low characterization of local heterogeneity and high characterization of longitudinal heterogeneity. Note that each MINIPOINT sample location includes up to six depths of porewater samples in the present study..... 7
- Figure 2 - Map illustrating sediment porewater sampling locations for the Local Sampling and Longitudinal Sampling campaigns, where the large, green circle symbols are the Local Sampling scheme locations, and small, yellow circle symbols are the Longitudinal Sampling scheme locations. Stream orders are identified by color, where first order streams are purple, second order streams are orange, and third order streams are green. 9
- Figure 3 - Discharge conditions at the downstream USGS gaging station on Augusta Creek (04105700) for water years 2015 (red) and 2016 (blue), with shading corresponding to the Local Sampling (red) and Longitudinal Sampling (blue) efforts described in this study..... 11
- Figure 4 - Field example of the division between "points" and "plots." A point representing a single MINIPOINT array at a site and a plot representing all three MINIPOINT arrays at a site under Local Sampling scheme, whereas there would only be one MINIPOINT array point in a plot under Longitudinal Sampling scheme. 14
- Figure 5 - Illustration of the distinction between point variance (on left) and plot variance (on right) in this study. Point variance represents the variance of 6 discrete depths of a single MINIPOINT, while plot variance represents the variance between all 18 measurements of the three MINIPOINTs at a site. 15
- Figure 6 - Point and whisker plots representing the mean (points) and range (whiskers) of observed porewater dissolved organic carbon (DOC) concentrations in Augusta Creek (all depths included) for both Local Sampling and Longitudinal Sampling schemes across the all of the network and grouped by stream order. 16
- Figure 7 - Box and whisker plots illustrating the distribution of variance for Local Sampling (i.e., high local characterization) and Longitudinal Sampling (i.e., low local characterization, but greater longitudinal characterization) for measurements of dissolved organic carbon (DOC; a-c), NO_3^- (d-f), and Cl^- (g-i) at points (a single MINIPOINT at a site) and plots (three MINIPOINTs at a site) across first, second, and third-order reaches of the Augusta Creek system. Distributions of the same sampling type are all significantly different per a Wilcoxon Rank Sum Test ($p < 0.05$), as noted with an * or otherwise stated with the specific p-value. 18

Figure 8 - Point (single MINIPOINT) to plot (three MINIPOINTs) variance ratios across stream orders during the Local Sampling (high local characterization) sampling campaign. The box and whiskers represent the quartiles at each stream order for the Local Sampling scheme, with the solid line indicating median values. The red diamonds represent mean values. Ratio values less than 1 indicate point variability is greater than plot variability, values greater than 1 indicate that point variability is less than plot variability, and values equal to 1 indicate point variability is equal to plot variability. The red dashed line represents a value of 1. 19

Figure 9 - The Augusta Creek watershed (black outline) and stream network (blue lines), with its location highlighted in the inset map of Michigan, USA (upper right). Sampling sites (yellow points) are distributed across all orders in the stream network. 35

Figure 10 - The observed Cl^- , NO_3^- , DOC, and SUVA_{254} conditions across the entire Augusta Creek stream network from the surface water synoptic sampling campaign (n=39 sample sites). Box and whiskers represent interquartile ranges, while red diamonds represent mean values and black circles represent outliers. 43

Figure 11 - Violin plots showing the distribution of (a) watershed areas (b) and unshared area percentage for groundwatersheds and surface watersheds (defined via Equations 1 and 2). The curvilinear outline and color fill of the violin plot shows a mirrored kernel density plot of values, while the internal boxplots show interquartile ranges. It is evident in (a) that the majority of areas for both surface watersheds and groundwatersheds are concentrated in smaller (<30 km²) areas. While the overall distributions of areas vary, including a greater maximum area in groundwatersheds, median values are similar between both. Note that the distributions of unshared areas differ substantially between groundwatersheds and surface watersheds. 45

Figure 12 - LULC in the August Creek stream network for both surface watershed (black border) and groundwatershed (white border) extent. Note that the groundwatershed includes large open water body, wetlands, and additional agricultural lands that are not captured in the surface watershed. 46

Figure 13 - An overall comparison of the groundwatershed versus surface watershed model performance is shown in the AIC values for all best-fit models across biogeochemical measurements Cl^- (a), NO_3^- (b), DOC (c), and SUVA_{254} (d). Each model is represented by a point, red circles are groundwatershed models and blue triangles are surface watershed models. All points that fall below the dashed line are considered to be comparable to the best-fit model. Models above the dashed line are considered to be worse than the best-fit model. 47

Figure 14 - A comparison of how groundwatershed and surface watershed models compare across solutes. Pseudo R^2 values for best-fit models based on either groundwatersheds or surface watersheds across all biogeochemical measurements. Red circles are groundwatersheds and blue triangles are surface watersheds. Values closer to 1 indicate the variance of predicted values is closer to the variance of observed values. 48

Figure 15 - Comparison of groundwatersheds (red) and surface watersheds (blue), as well as their overlap (purple) in Augusta Creek. The three insets present examples of (a) a low degree of overlap for surface and watersheds delineated to a point, (b) a high degree of overlap for the groundwatershed, and (c) high degree of overlap for both surface watersheds and groundwatersheds. 50

Figure 16 - A comparison of groundwatershed and surface watershed model performance as it relates to dominant land cover class and placement in the stream network is shown in the Pseudo R^2 values for sampling subsets related to predominant land use within an area, cultivated crops (a), deciduous forest (b), and hay/pasture (c); and stream order, first-order (d), second-order (e), and third order (f). 61

Figure 17 - Map of Augusta Creek including (a) with the watershed boundaries in which the double gray line is the groundwatershed and the solid black line is the surface (topographic) watershed, the blue line is the stream network, yellow points represent the sampling points, and the green triangle represents the USGS Gaging Station (04105700); (b) displays the National Land Cover Database 2011 with categories found in the Augusta Creek watershed indicated in the legend. 68

Figure 18 - Comparison of surface watershed (black outline) and groundwatershed (red double outline) for a particular sampling point on a tributary of Augusta Creek, showing a case of particularly large disagreement between the two watershed areas. The image displays the different land use categories (various colors defined in Figure 17) that are included depending on which watershed was selected. 71

Figure 19 - Comparison of AIC values for surface water spatial lag models, modified from Lee-Cullin et al. (*Chapter 2*). Red circles indicate groundwatersheds were used for drainage area delineation while blue triangles indicate that surface watersheds were used for drainage area delineation. The dashed line indicates the threshold at which models below the line are considered ‘best fit’ models. In AIC, lower values mean a more parsimonious model. 74

Figure 20 - Biogeochemical results of surface water and at various depths beneath the sediment surface in the SGI across Augusta Creek. Measurements include DOC, SUVA₂₅₄, S_R , Cl^- , and NO_3^- . Box and whiskers represent interquartile range, red diamonds represent mean values, and black circles represent outliers. 76

Figure 21 - Pseudo R^2 results for spatial lag models across all depths for each biogeochemical measurement across all depths within the Augusta Creek stream network. Solid red lines represent groundwatershed models while dashed blue lines represent surface watershed models. Pseudo R^2 values close to one indicate a higher proportion of the overall variance of modeled values representing the overall variance of observed values. 77

Figure 22 - Conceptual figure representing the proposed mechanisms by which DOC is consistently processed in the SGI. Here DOC enters the SGI as a unique, diverse pool of

constituents. Once in the SGI processes such as sorption (e.g., Nelson et al. 1993), enzymatic hydrolysis (e.g., Volk et al. 1997), biological metabolism (e.g., Cammack et al. 2004), and redox transformation (e.g., Lovley et al. 1996) all occur, leaving only the less reactive components of the DOC pool to exit the SGI back into the surface stream water. 88

Figure 23 - Principal components 1 and 2 for all optical measurements and all DOC sources included in this study. Data include all time points in the bioassays. Each axis represents a principal component, which condenses multi-dimensional space into two-dimensional space. Arrows represent each DOC optical property that went into the PCA. Each unique DOC source is given a unique color to show its placement in this reduced dimensionality..... 95

Figure 24 - PCA results of all optical measurements for all DOC sources included in this study, plotted the same as Figure 23 but with symbol colors indicating time of incubation in the bioassays, with initial samples appearing as navy blue and end point samples appearing as red. 96

Figure 25 - Location of groundwater wells in the Augusta Creek watershed from which groundwater levels were determined. The pink to red gradient represents depth to water table ranging from 0 feet down to 75 feet or greater. Well data available through the State of Michigan Environment, Great Lakes & Energy Wellogic. 109

Figure 26 - AIC values for all best-fit models across each biogeochemical measurement. Each model is represented by a point, red circles are groundwatershed models, blue triangles are surface watershed models, and purple circles are combined watersheds. All points that fall below the dashed line are considered to be comparable to the best-fit model. Models above the dashed line are considered to be worse than the best-fit model. 110

Figure 27 - Pseudo R2 for all biogeochemical measurements for surface watersheds (blue square), groundwatersheds (red triangle), and combined watersheds (purple circles). 111

KEY TO ABBREVIATIONS

AIC	Akaike's Information Criterion
DOC	Dissolved Organic Carbon
HZ	Hyporheic Zone
OLS	Ordinary Least Squares
SGI	Stream-Groundwater Interface
KBS	Kellogg Biological Station
LULC	Land Use Land Cover
RCC	River Continuum Concept
VIF	Variance Inflation Factor
USGS	United States Geological Survey

CHAPTER 1: TOWARD MEASURING BIOGEOCHEMISTRY WITHIN THE STREAM- GROUNDWATER INTERFACE AT THE NETWORK SCALE: AN INITIAL ASSESSMENT OF TWO SPATIAL SAMPLING STRATEGIES

A version of this chapter appeared in the journal *Limnology & Oceanography: Methods* (16:11, pp. 722-733).

Abstract

It is important to understand how point measurements across spatially heterogeneous ecosystems are scaled to represent the system of interest. Stream biogeochemistry presents an illustrative example because water quality concerns within stream networks and recipient water bodies motivate heterogeneous watershed studies. Measurements of the stream-groundwater interface (SGI, i.e., the shallow subsurface of streams) are well-documented for small, point-scale sampling density measurements (i.e., $\text{cm}^2\text{-m}^2$ features), but poorly characterized for larger, watershed scale sampling density measurements (i.e., km^2 ; stream reaches and networks). Further, sampling the SGI is more time- and labor-intensive than surface water sampling, meaning sample point selection must be made with care when attempting a network-scale analysis. In this study, we endeavor to determine which of two common spatial sampling schemes is appropriate for characterizing the biogeochemistry of the SGI across a temperate, third-order stream network, focusing on dissolved organic carbon. The first scheme, called here Local Sampling, focuses on characterization of the small-scale ($< 10 \text{ m}^2$) variability produced by the local physical and biogeochemical heterogeneity, with fewer points across the stream network. The second scheme, called here Longitudinal Sampling, has approximately the same number of measurements distributed over many more points across the stream network with less characterization of local variability. This comparison reveals that selection of a Local Sampling

versus a Longitudinal Sampling scheme influences the interpretation of biogeochemical patterns at the stream network scale. Additionally, this study found an increase in observation efforts at the local scale added limited information for reach- to network-scale biogeochemical patterns, suggesting that emphasis should be placed on characterizing variability across broader spatial scales with the Longitudinal Sampling approach.

Introduction

At what spatial resolution do we make measurements and observations to characterize patterns and processes across stream networks? It is well-established in terrestrial landscape ecology that measurements made at a certain spatial sampling density (i.e., resolution or grain size) can be extrapolated to different scales of spatial extent (e.g., Schneider 1998; Wu and Li 2006). However, the best practices for extrapolating between scales are continually evolving, including many methods that have been developed to upscale or downscale observations to different resolutions (Turner and Gardner 2015). Few studies have presented best sampling practices and methods across scales for aquatic ecosystems. Streams and their interfaces with groundwater are particularly challenging for choosing the most appropriate sampling resolution due to the inherent effects of directionality in flowing water and logistical challenges of measuring surface water and groundwater parameters. To determine the ecological conditions and functioning of the stream-groundwater interface (SGI) that are relevant to landscape biogeochemical budgets, watershed management, and ecosystem theories at the reach to network scales (Krause et al. 2011; Bernhardt et al. 2017), we must address how best to measure SGI interactions across spatial and temporal scales.

At stream network scales, the River Continuum Concept (RCC; Vannote et al. 1980) was a key step in starting to address the landscape ecology of stream networks, including the effects of

directional flow through streams. The RCC postulated a gradient, moving from headwaters to higher-order streams, to explain the downstream movement and transformation of organic matter by physical and biological processes. Although aspects of the RCC are still debated (see, e.g., Creed et al. 2015; Rosi-Marshall et al. 2016), the general conceptual model of a gradient wherein the biogeochemistry of stream reaches changes systematically from upstream to downstream in networks is central to contemporary literature of stream ecosystems (e.g., Poole 2002; Thorp and Bowes 2017). The RCC is raised here not to debate its strengths and weaknesses in explaining how conditions may change through a river network, but because it did not specify what scale of measurements is needed to assess the ecological hypotheses of the RCC. Hence, there is still uncertainty in how to assess the RCC. However, in a review of stream ecology and biogeochemistry, Fisher et al. (2004) identified a broad understanding that streams are largely influenced by longitudinal (i.e., upstream/downstream) changes, and are composed of a multitude of parallel flowpaths leading to a high degree of heterogeneity. Other studies have stressed that nearby sampling points of stream chemistry were very similar but were able to maintain a broader heterogeneous trend (Dent and Grimm 1999).

Generally, spatial biogeochemical variation in surface waters decreases with increasing stream order (Temnerud and Bishop 2005), but there is evidence in some streams that comparable variability can be found at all scales depending on the sampling density (Zimmer et al. 2013; Abbott et al. 2018). This spatial biogeochemical variability suggests that the role of study design, especially the spatial resolution of sampling, can introduce bias or confusion in our understanding of stream ecology. Furthermore, the biogeochemical variability in SGIs across stream networks is virtually unknown and almost never documented in the literature (but see Ruhala et al. 2018). This is particularly true for assessing the structure and dynamics of the

hyporheic zone (HZ), the ecotone where stream water readily interacts and exchanges properties with groundwater (Boulton et al. 1998). The HZ is a known biogeochemical control point in watersheds influencing ecosystems and water quality (McClain et al. 2003; Bernhardt et al. 2017). The primary limitation with sampling sediment pore water in the SGI is that it can be time- and labor-intensive, given that porewater must be drawn out slowly to avoid disrupting stream, hyporheic, and groundwater flow fields (i.e., typically $< 5 \text{ ml min}^{-1}$) (e.g., Duff et al. 1998). The SGI is also known to exhibit large spatiotemporal heterogeneity in physical and biological conditions (Boano et al. 2010).

Our understanding of SGI biogeochemistry at stream network scales has been limited by a lack of understanding of how to best allocate sampling efforts in space and time. The local scale (i.e., the within-reach scale) over which SGI data are typically collected does not match the stream network scale at which many environmental problems need to be addressed (Krause et al. 2011). In fact, most SGI studies do not make direct measurements in the SGI, and instead use indirect measurements (i.e., tracer studies) that span a longitudinal scale of 10-1000 m (Ward 2016). These indirect measurements are often rife with model uncertainty and interpretation, especially for quantifying SGI exchange (Kelleher et al. 2013). Despite the lack of direct measurements, significant advances in process-based modeling of SGI processes at the stream network scale have proceeded, including the transport and fate of nutrients (Kiel and Cardenas 2014; Gomez-Velez et al. 2015). Unfortunately, there is still a paucity of data sets of the SGI at the network scale available to validate these types of models.

In thinking about a sampling scheme of the SGI across an entire stream network, one must consider the effort spent for an individual sampling point while ensuring that the limited available number of sampling points reasonably represent the entire network. Generally, there

are two stream network-scale sampling schemes that appear in the literature (Figure 1): 1) high-resolution characterization of local-scale variability at few sites across the network (e.g., Zimmer et al. 2013; hereafter Local Sampling), wherein effort is focused on taking many samples at specific local-scale features in a watershed instead of fewer samples at more locations, or 2) low-resolution characterization of local-scale variability at many sites across the network, (e.g., McGuire et al. 2014; hereafter Longitudinal Sampling), wherein effort is focused on taking samples at more locations across the entire network instead of more samples at specific local-scale features. The schemes are either deliberately or arbitrarily selected to investigate properties relevant to stream network biogeochemistry. Local Sampling is often applied for investigations of specific SGI processes, while there are very few examples of Longitudinal Sampling studies for any type of SGI processes (Ward 2016). However, it is unknown whether one of these two sampling schemes is more appropriate for research questions dealing with characterization of SGI biogeochemistry at the network-scale. Our objectives in this paper are to raise awareness regarding SGI sampling design unknowns and to begin addressing these unknowns in our investigation of network-scale SGI interactions by comparing the two common sampling schemes across a stream network. Determining which scheme, Local Sampling or Longitudinal Sampling, best characterizes the overall stream network will help advance SGI investigations and thus guide best sampling practices (Krause et al. 2011). To direct these main objectives, we developed the following hypotheses:

H1: A single point profile is representative of multiple point profile measurements of SGI biogeochemistry, because inter-reach variability will be greater than intra-reach variability. This hypothesis will assess whether sampling of the SGI should focus on fewer points at more sites in

a network or if it is necessary to have many points to characterize each individual site, which, in turn, will guide sampling design for future SGI studies.

H2: Variance in SGI biogeochemistry profiles will decrease with increasing stream order, because the effects of upstream processes are integrated downstream due to directional flow. This hypothesis will help inform the development of network continuum concept in the SGI, such as the continuum concepts of the RCC.

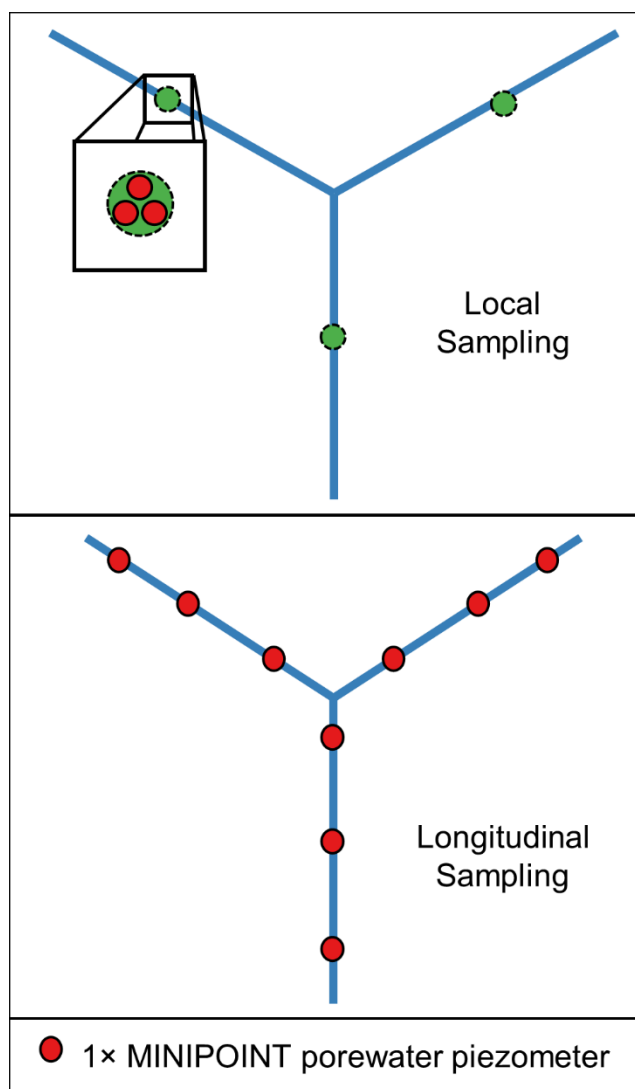


Figure 1 - Simplified plan view of stream network reaches illustrating the main conceptual differences for Local Sampling (L) and Longitudinal Sampling (R) sampling schemes. Local Sampling represents high characterization of local heterogeneity with low characterization of longitudinal heterogeneity, while Longitudinal Sampling has low characterization of local heterogeneity and high characterization of longitudinal heterogeneity. Note that each MINIPOINT sample location includes up to six depths of porewater samples in the present study.

To evaluate these objectives and hypotheses, we analyzed a spatially intensive sampling of SGI biogeochemistry (as compared to other SGI studies in the literature) in a stream network that spans the two study sampling schemes (Ruhala et al. 2018). Specifically, we focus on the surface water and SGI pore-water concentrations of dissolved organic carbon (DOC) in a lowland, third-

order, mixed land use watershed. DOC was selected as the focus for this initial assessment because it is a fundamental control on water quality and ecosystem ecology of freshwaters due, in part, to its role in nutrient and metal cycling, ability to influence pH, effect on net carbon balances, and control of photochemistry (Aiken 2014). In addition to DOC, we include analyses for select anions, including chloride (Cl^-) and nitrate (NO_3^-) to represent nonreactive and reactive solutes, respectively (e.g., Triska et al. 1993; Barber et al. 2005; Zarnetske et al. 2011; Bernhardt et al. 2017).

Materials and Methods

Site description

The data sets used in this study were generated by Ruhala et al. (2018) in Augusta Creek (Figure 2), which is a low gradient, third-order watershed draining 98 km² in southwest Michigan, USA. The watershed is composed of glacial till, and flows through a mixed-use landscape that includes wetlands, lakes, agriculture, and upland forests. The stream is primarily groundwater-fed, gaining water along much of its length, and the low overland runoff as well as abundant wetlands and lakes along its course buffer the stream discharge response to storm events (Poff et al. 1997; Hamilton et al. 2018). Stream reaches included in this study range from first- to third-order, with variable origins including lake outflows, wetland outflows, and forested headwater streams (Figure 2). Located near the W.K. Kellogg Biological Station of Michigan State University (KBS), Augusta Creek is a historically important site for freshwater biogeochemical and ecological research. For example, it was a site in the seminal RCC and Natural Flow Regime papers (Vannote et al. 1980; Poff et al. 1997), is part of the KBS Long Term Ecological Research site activities, and has an active, long-term (>50y) United States Geological Survey (USGS) gaging station (04105700).

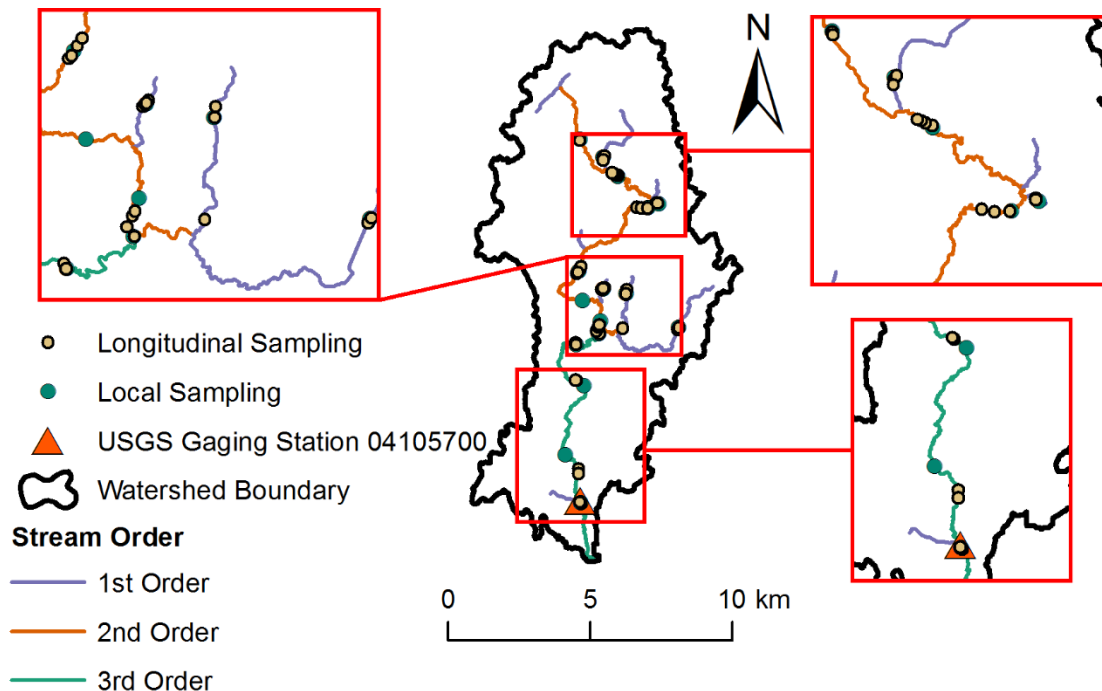


Figure 2 - Map illustrating sediment porewater sampling locations for the Local Sampling and Longitudinal Sampling campaigns, where the large, green circle symbols are the Local Sampling scheme locations, and small, yellow circle symbols are the Longitudinal Sampling scheme locations. Stream orders are identified by color, where first order streams are purple, second order streams are orange, and third order streams are green.

Sampling schemes

Ruhala et al. (2018) collected data that span the Local Sampling and Longitudinal Sampling schemes, and importantly, each sampling date represented roughly the same field sampling effort (~10 field work days for 4 researchers), the same sampling techniques and equipment, and a comparable total number of SGI biogeochemical sample locations ($n \approx 40$). However, the team distributed these sampling points differently across the stream network, stratifying the sampling to capture most subwatersheds and all stream orders in the Augusta Creek watershed. The sampling scheme roughly corresponded to the two study scheme types, Local and Longitudinal (Figure 2).

In the data set, Local Sampling samplings characterized the local heterogeneity of a limited number of sites across the network and were carried out from 10-17 August 2015. In the Local Sampling scheme, 16 locations, stratified by stream order (first through third) were selected across the network (Figure 2). Within each location, 3 MINIPPOINT porewater piezometers (Duff et al. 1998) were deployed close to each other (<3 meters apart), and hereafter the group of three samplers will be referred to as a plot (Figure 1). The MINIPPOINT porewater piezometers are relatively non-invasive and allow sampling of pore water profiles from six discrete depths in the SGI (Duff et al. 1998), set between 2.5 and 20 cm as detailed in the next section and Ruhala et al. (2018). Thus, there were 18 SGI samples collected at each of the 16 plots for a total of 288 unique SGI biogeochemical sample locations from the Local Sampling approach. In Augusta Creek, most of the stream sediment is unconsolidated sandy and gravelly sediments, which is compatible with the MINIPPOINT technology. However, the exact MINIPPOINT porewater piezometer location at a selected site depended on the capacity to physically insert all the piezometers the specified depth into the sediment (i.e., sites with cobble or armored sediments could not be sampled).

The Longitudinal Sampling scheme represented a coarser characterization of local heterogeneity but increased the total number of plots across the stream network and thus was meant to capture the spatial variability across the stream network. This sampling was carried out from 16-22 August 2016 during similar seasonal, stream DOC conditions, and daily discharge conditions as the Local Sampling campaign (Figure 3), though 2016 data was collected during discharge recession from a preceding high flow event. For Longitudinal Sampling, a similar field effort yielded 39 points across the network. At each location, a single MINIPPOINT porewater

piezometer was sampled, optimally collecting six porewater samples per point for a total of 230 unique SGI biogeochemical sampling locations from the Longitudinal Sampling.

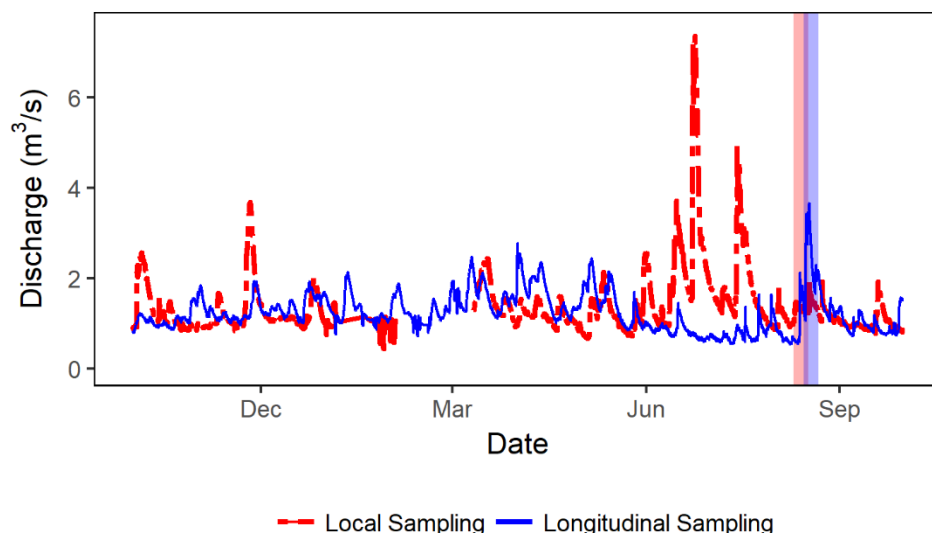


Figure 3 - Discharge conditions at the downstream USGS gaging station on Augusta Creek (04105700) for water years 2015 (red) and 2016 (blue), with shading corresponding to the Local Sampling (red) and Longitudinal Sampling (blue) efforts described in this study.

Furthermore, given that we are specifically interested in the biogeochemistry with respect to DOC at larger spatial scales, we also analyzed data grouped by stream order similar to the RCC (Vannote et al. 1980). Stream order acts as a proxy for the physical hydrography of stream reaches, which in turn is fundamental to ecological patterns and processes (Harvey and Gooseff 2015). It is a simple method to discretize the network that allows for quick analysis of how an ecological variable related to DOC varies from upstream to downstream through a stream network (e.g., Creed et al. 2015). In the Local Sampling scheme there were 6 first-order, 5 second-order, and 5 third-order locations, while the Longitudinal Sampling scheme was composed of 16 first-order, 14 second-order, and 9 third-order plots. This enables an assessment of how the biogeochemistry changes with different hydrological characteristics distributed from headwaters to mainstem outlet (as addressed by H2 above).

Sample and data collection

To illustrate the procedure and effort involved in collecting SGI samples, here we briefly review the sampling protocol from Ruhala et al. (2018). Each MINIPPOINT porewater piezometer was deployed to collect six discrete samples at 2.5, 5, 7.5, 10, 15, and 20 cm depth. The MINIPPOINTs were attached to a Masterflex peristaltic pump (Cole-Parmer) using L/S Tygon tubing, and water was drawn from the SGI at a rate of 2.5 ml min^{-1} . They collected 80 mL of water from each depth. They used 20 mL of sample as a rinse through the filter (Whatman GF/F, $0.7 \text{ }\mu\text{m}$ nominal pore size) to remove particulate matter. The remaining 60 mL was filtered through the $0.7 \text{ }\mu\text{m}$ filter to remove particulates and larger microbes the placed in acid-rinsed HDPE amber bottles and stored on ice. At the end of the sampling day, 10 mL were first used to rinse through a filter (Sartorius Stedim cellulose acetate, $0.2 \text{ }\mu\text{m}$ nominal pore size), then the remaining 50 mL were filtered and stored in the dark at 4°C and analyzed within 28 days. Each filtered sample was analyzed for non-purgeable organic carbon using a TOC-L total organic carbon analyzer (Shimadzu) with Pt-catalyzed oxidation at 680°C . Concentrations for Cl^{-} and NO_3^{-} were analyzed on a Dionex ICS-2100 Ion Chromatography System (ThermoScientific).

Data analysis

The Local Sampling data were divided into points, representing a single MINIPPOINT with six samples at vertically distributed depths, and plots representing three MINIPPOINTs with eighteen samples, varying in depth, at a single site (Figure 4). The Longitudinal Sampling data was simply divided into points, as there was only a single MINIPPOINT with six vertically distributed samples deployed at each individual site. We calculated variance for a point as the variance across the six individual depths from a MINIPPOINT sampling, and variance for a plot

as variance across all eighteen samples (6 depths at 3 points) from the clustered MINIPOINTs (Figure 5) as:

$$\sigma^2 = \sum \frac{(X-\mu)^2}{N} \quad (1)$$

where X is a biogeochemical concentration value at one discrete piezometer (within a MINIPOINT array for point variance and within the three MINIPOINT arrays for plot variance), μ is the mean of all concentration measurements (again, within a single MINIPOINT array for point variance and for all three MINIPOINT arrays for plot variance), and N is the number of observations ($N=6$ for point variance, $N=18$ for plot variance).

For the Local Sampling data, to assess the relative utility of a single MINIPOINT as compared to three MINIPOINTs we took the ratio of the plot variance to point variance (F), shown as:

$$F = \frac{\sigma^2_{\text{plot}}}{\sigma^2_{\text{point}}} \quad (2)$$

where σ^2 is the variance (Equation 1) and the subscripts represent the plot and points. Finally, to compare the full distributions of point and plot measurements across stream orders we used a non-parametric Wilcoxon Rank Sum Test (Wilcoxon 1945) implemented in the software R v 3.4.2 (R Core Team 2017). The Wilcoxon Rank Sum Test allows us to assess whether the distribution of samples within orders are increasing or decreasing across first, second, and third orders. This assessment is used to determine if similar patterns emerge when comparing point and plot measurements and when comparing Local Sampling to Longitudinal Sampling.

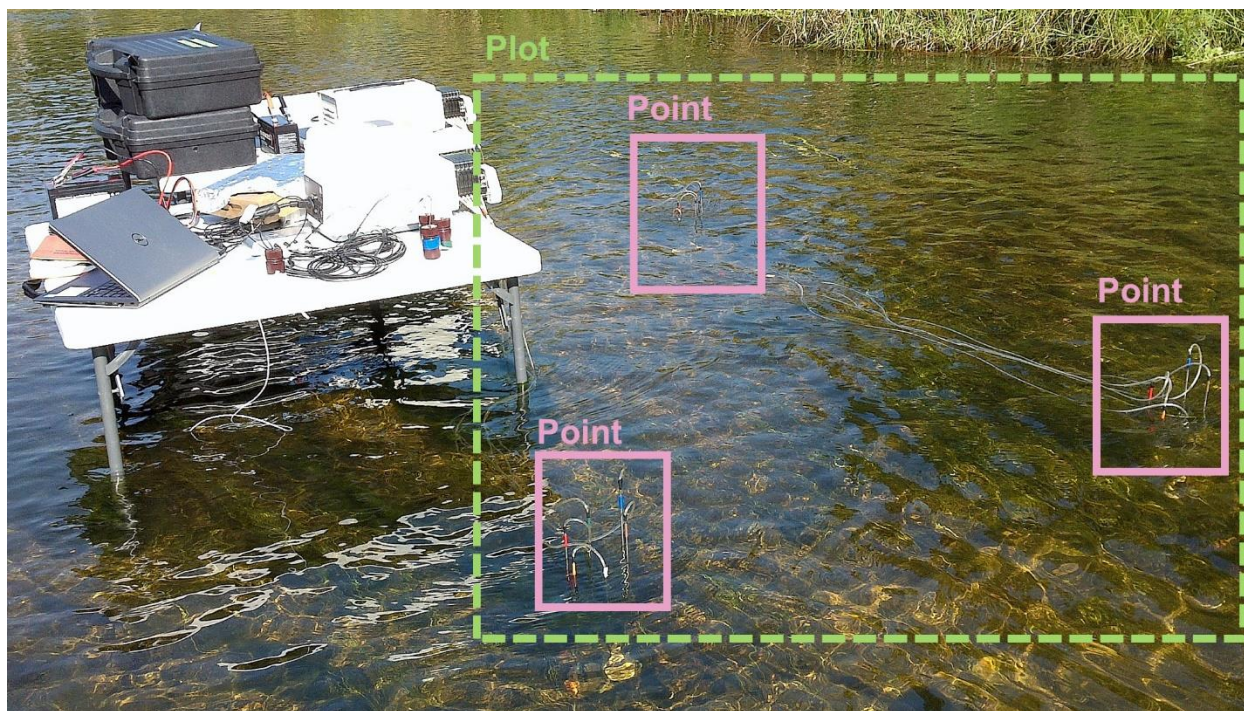


Figure 4 - Field example of the division between "points" and "plots." A point representing a single MINIPOINT array at a site and a plot representing all three MINIPOINT arrays at a site under Local Sampling scheme, whereas there would only be one MINIPOINT array point in a plot under Longitudinal Sampling scheme.

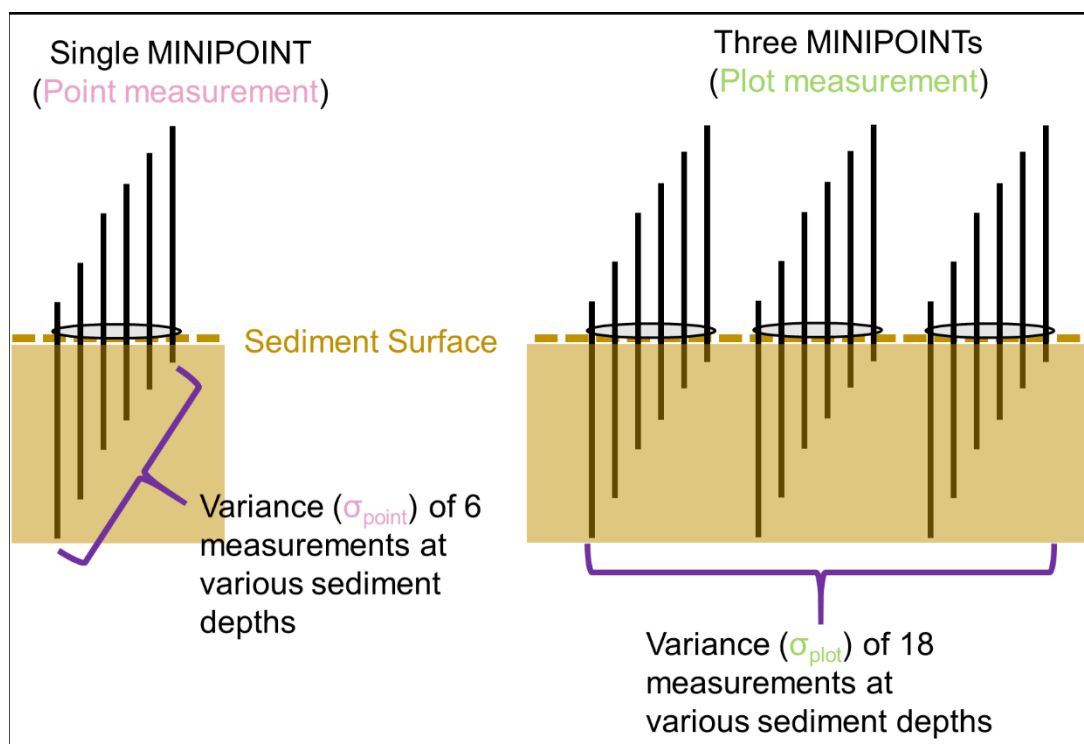


Figure 5 - Illustration of the distinction between point variance (on left) and plot variance (on right) in this study. Point variance represents the variance of 6 discrete depths of a single MINIPOINT, while plot variance represents the variance between all 18 measurements of the three MINIPOINTs at a site.

Results

Concentrations of DOC in the SGI were comparable between the Local Sampling and Longitudinal Sampling schemes across the network and across samplings grouped by stream order (Figure 6). Minimum and maximum SGI DOC concentration values for the Local Sampling were 1.50 and 15.70 mg L⁻¹, respectively, while minimum and maximum SGI DOC concentration values for the Longitudinal Sampling were 1.34 and 17.04 mg L⁻¹, respectively (Figure 6).

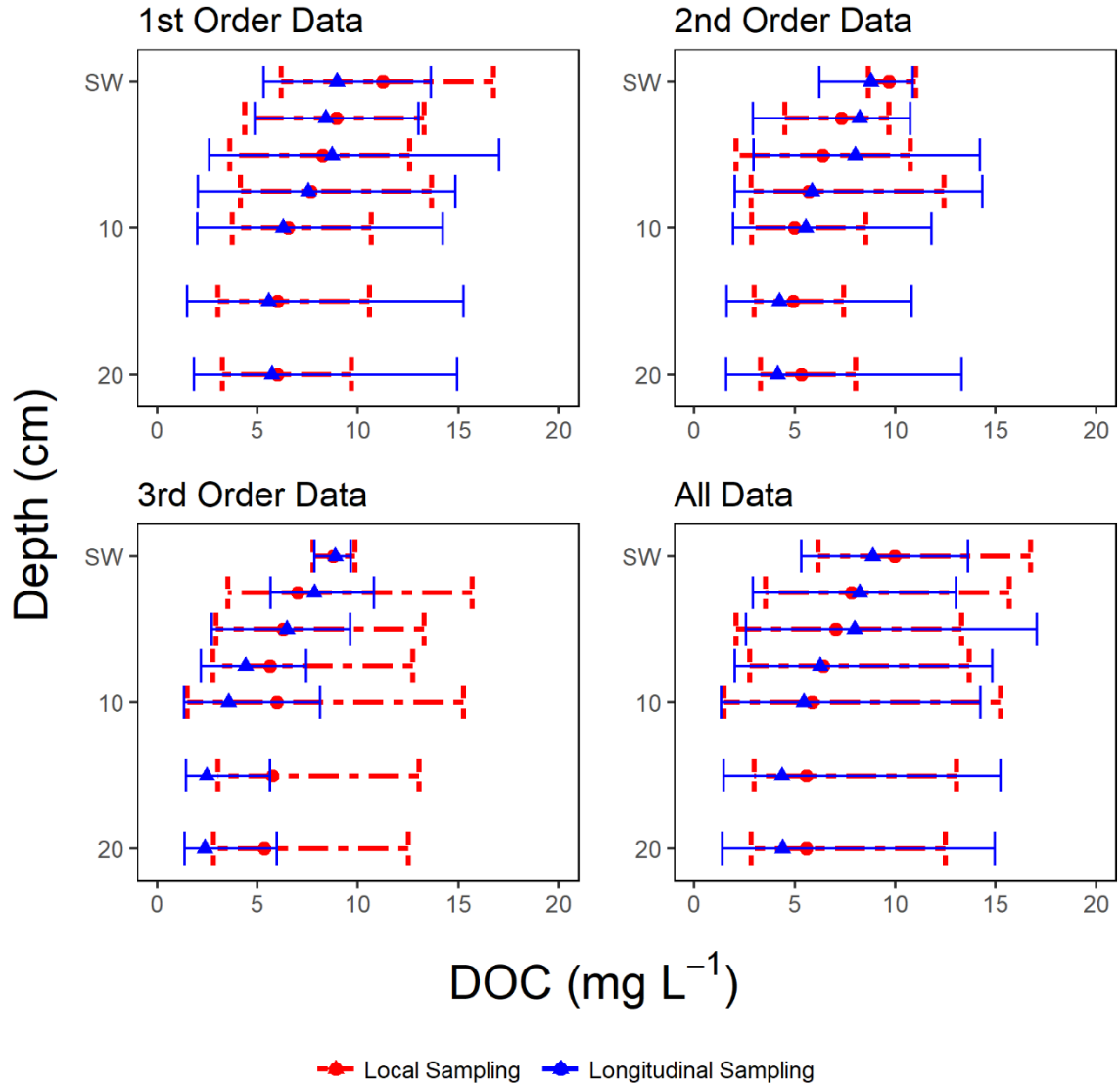


Figure 6 - Point and whisker plots representing the mean (points) and range (whiskers) of observed porewater dissolved organic carbon (DOC) concentrations in Augusta Creek (all depths included) for both Local Sampling and Longitudinal Sampling schemes across the all of the network and grouped by stream order.

Local Sampling scheme results

Point measurements of DOC exhibited a general decrease in variance from first- to third-order (Figure 7a), where there are significant differences among first- to third-order variances ($p < 0.05$). Plot measurements of DOC also exhibited decreasing variance from first- to third-order (Figure 7b) with significant differences noted ($p < 0.05$). The DOC variance ratio, F from

equation 2, ranged from 0.4 to 4.3, 0.5 to 2.3, and 0.4 to 2.4 for first-, second-, and third-order streams, respectively (Figure 8a). The corresponding median ratio values for first-, second-, and third-order streams were 1.2, 1.0, and 1.2, respectively.

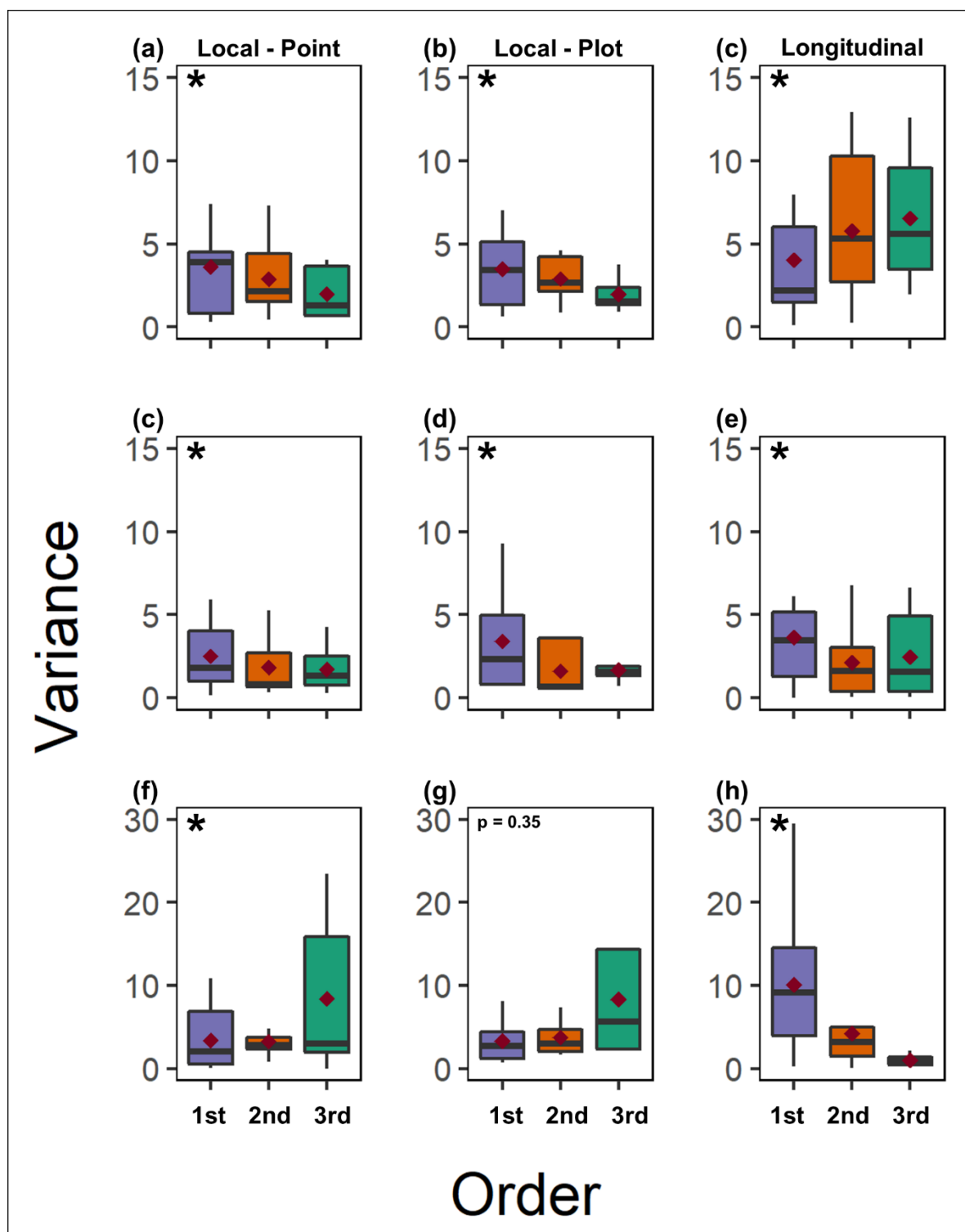


Figure 7 - Box and whisker plots illustrating the distribution of variance for Local Sampling (i.e., high local characterization) and Longitudinal Sampling (i.e., low local characterization, but greater longitudinal characterization) for measurements of dissolved organic carbon (DOC; a-c), NO_3^- (d-f), and Cl^- (g-i) at points (a single MINIPPOINT at a site) and plots (three MINIPPOINTS at a site) across first, second, and third-order reaches of the Augusta Creek system. Distributions

Figure 7 (cont'd) of the same sampling type are all significantly different per a Wilcoxon Rank Sum Test ($p < 0.05$), as noted with an * or otherwise stated with the specific p-value.

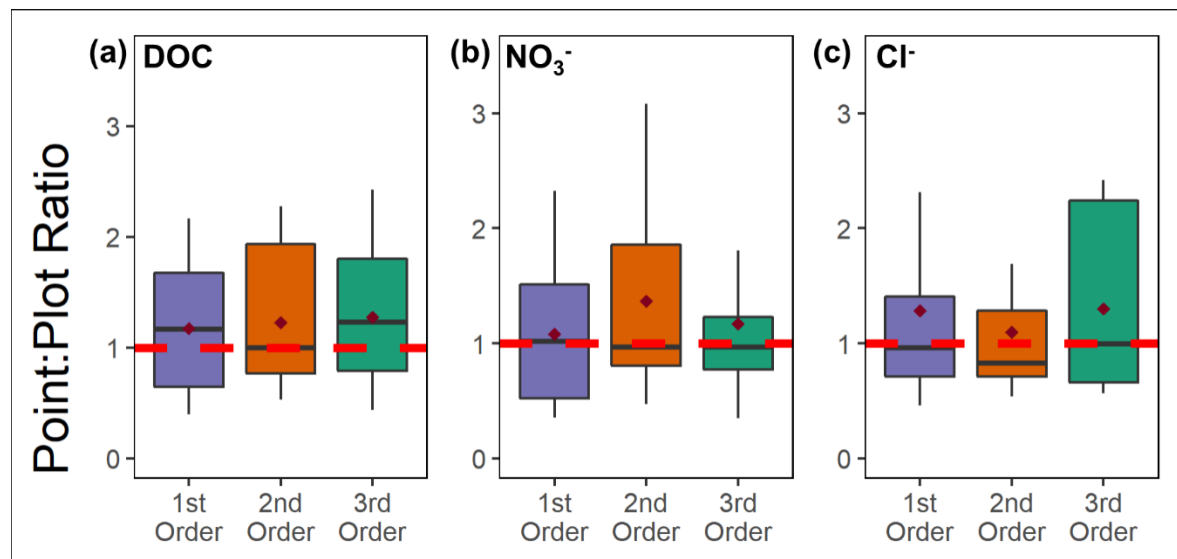


Figure 8 - Point (single MINIPPOINT) to plot (three MINIPPOINTS) variance ratios across stream orders during the Local Sampling (high local characterization) sampling campaign. The box and whiskers represent the quartiles at each stream order for the Local Sampling scheme, with the solid line indicating median values. The red diamonds represent mean values. Ratio values less than 1 indicate point variability is greater than plot variability, values greater than 1 indicate that point variability is less than plot variability, and values equal to 1 indicate point variability is equal to plot variability. The red dashed line represents a value of 1.

Variance of NO₃⁻ point measurements appeared to decrease from first- to second-order and then increase from second- to third-order (Figure 7d) and were significantly different across orders ($p < 0.05$). Plot-scale variance of NO₃⁻ indicates a decrease from first- to second-order and a decrease from second- to third-order (Figure 7e), with first- through third-order exhibiting significant differences ($p < 0.5$). The NO₃⁻ variance ratio ranged from 0.4 to 40.1, 0.5 to 5.9, and 0.4 to 70.9 for first-, second-, and third-order streams, respectively (Figure 8b). The corresponding median values for first-, second-, and third-order streams were 1.3, 1.0, and 1.1, respectively.

Point measurements of Cl⁻ increased from first- to third-order (Figure 7g) and were significantly different ($p < 0.05$). Variances of plot measurements of Cl⁻ were not significantly different from first- to third-order streams (Figure 7h, $p = 0.35$). The Cl⁻ variance ratio ranged

from 0.5 to 22.6, 0.5 to 14.7, and 0.6 to 2.4 for first-, second-, and third-order streams, respectively (Figure 8c). The corresponding median values for Cl^- in first-, second-, and third-order streams were respectively 1.0, 0.8, and 1.0.

Longitudinal Sampling scheme results

The plot variances of DOC had an apparent increase from first- to third-order streams (Figure 7c) and were significantly different among orders ($p < 0.05$). Plot variances of NO_3^- decreased from first- to third-order (Figure 7f) and were significantly different ($p < 0.05$). The plot variances of Cl^- decreased from first- to second-order (Figure 7i) and were significantly different ($p < 0.05$), but a post-hoc Dunn test (Dunn 1964) indicated that there was no significant difference between second- and third-order plot variances ($p = 0.09$).

Discussion

Our analysis of spatial heterogeneity of porewater chemistry from samples throughout the Augusta Creek network reveals several critical insights into how to best collect spatial data from the SGI at the network-scale. Further, this analysis helps demonstrate that SGI investigators must be cognizant of how to sample when interested in larger spatial patterns, especially when considering how stream networks remove or transform reactive biogeochemical solutes.

Guiding future sampling

The results offer an indication of how to best invest our future sampling efforts when a network-scale assessment of SGI biogeochemistry is the goal. Primarily, in Augusta Creek, we find that there is little added value in increasing characterization of the local, plot-scale spatial heterogeneity, particularly for the reactive biogeochemical components DOC and NO_3^- . The point:plot ratio in the Local Sampling scheme generally centered on a value of 1 (Figure 8) for reactive (DOC and NO_3^-) and nonreactive solutes (Cl^-), meaning that a single sampling array at a

site can approximate the variance of a site as well as three separate sampling arrays at a site. In fact, new patterns of variability emerge when focusing on sampling across the stream network as opposed to more detailed local characterization (e.g., Figures 7a and 7b to 7c and Figures 7d and 7e to 7f for DOC and NO_3^- , respectively), wherein the patterns of variance moving from headwaters to downstream locations actually changes when enacting a Longitudinal Sampling scheme as compared to a Local Sampling scheme.

These results indicate that a Longitudinal Sampling scheme may be preferable to a Local Sampling scheme when investigating the biogeochemistry of the SGI at the network-scale. This finding is corroborated by two recent papers that present conceptual and reduced complexity models to understand DOC (Hotchkiss et al. 2018) and NO_3^- (Marzadri et al. 2017) processing as they move from headwater to downstream locations (i.e., from low to high order streams), including the potential differential effects of the SGI across the river network. Our assessment of the two main spatial sampling schemes for SGI and the specific results from Augusta Creek inform how future researchers can attempt to evaluate and validate these new conceptual and modeling frameworks as well as historically important frameworks such as the RCC (Vannote et al. 1980).

The variance ratios observed between the two sampling strategies suggest that point measurements are reasonably representative of plot measurements in Augusta Creek, because median values for all ratios are generally equal to unity (i.e., the ratio of the variance within a plot is close to the variance of each individual point). A Wilcoxon Rank Sum Test of the distributions of variance ratios indicates that point and plot (mean of three points) measurements are similar for most chemistry samples, except for DOC and NO_3^- in third-order reaches. However, the median values of plot to point ratios in third-order streams are still relatively close

to unity (1.23 for DOC, 0.99 for Cl^- , and 1.12 for NO_3^-). Therefore, for this stream network under summer baseflow conditions, the results suggest that the SGI biogeochemistry of first- and second-order streams can be characterized with less focus on the local intra-site heterogeneity, which allows more focus on the inter-reach heterogeneity. In other words, more valuable data about network-scale SGI biogeochemical conditions can be collected using the Longitudinal Sampling scheme as compared to the Local Sampling scheme.

The observed reduction in variances of porewater concentrations moving downstream was dependent upon the biogeochemical species of interest. In the Local Sampling campaign, DOC variance at different sampling densities generally decreased moving from first- to third-order streams (Figure 7a and b). Conversely, Cl^- variance in Local Sampling increased from first- to third-order streams for both sampling densities (Figure 7g and h). NO_3^- variance exhibited an inconsistent trend in Local Sampling, wherein it increased from first- to second-order, then decreased from second- to third-order (Figure 7d and e). In Longitudinal Sampling the NO_3^- variance generally decreased with increasing stream-order (Figure 7c). The reduction in DOC variance with increasing stream order reflects the accumulation and mixing of all upstream inputs (Abbott et al. 2018). Synthesis studies of DOC across stream networks indicate that, indeed, the variability of DOC typically decreases with an increase in disconnection from terrestrial sources (e.g., Creed et al. 2015).

Most stream networks have the majority of total stream length in first- and second-order streams (e.g., first-order = 52% and second-order = 25%, Downing et al. 2012), so the finding that the low-order streams in Augusta Creek can be characterized with less focus on intra-site heterogeneity means that more low-order locations should be sampled (i.e., the Longitudinal Sampling scheme), rather than investing efforts in plot replication at each location. Historically,

SGI research has disproportionately focused on second-, third-, and fourth-order streams (Ward 2016), so more effort should be directed to first and >fifth-order stream SGIs in networks if we are to better represent SGI conditions in future network scale biogeochemical studies and models. Headwaters are demonstrably important in terms of the contribution of biogeochemical processes to downstream nutrient export (Alexander et al. 2007; Boano et al. 2014). Further, smaller networks tend to display the highest variability in water quality (Wolock et al. 1997; Temnerud and Bishop 2005; Abbott et al. 2018).

Different sampling resolution concerns

The Longitudinal Sampling campaign, with low local characterization in favor of higher longitudinal spatial resolution across the stream network, can potentially result in an entirely different interpretation of SGI conditions and DOC stream processing. DOC and Cl^- trends across orders were the opposite as compared to the trends observed in the Local Sampling scheme. Here, DOC variance is generally increasing, while Cl^- is generally decreasing moving from upstream to downstream (Figure 7c, f). While Cl^- fits our hypothesis (H2), DOC does not support it. This is an important revelation given that the same stream system was sampled under similar weather and hydrologic conditions (albeit in a different year) but changing the spatial SGI sampling scheme yielded a completely different apparent pattern across the network. These fundamental differences moving from headwaters to downstream locations have raised concerns particularly for empirical and mechanistic modeling. If data input into a model has a different pattern of variance depending on the sampling scheme, then the results of those models and conclusions that can be drawn from them will be entirely different from one scheme to the next.

Though studies comparing biogeochemistry at different scales are generally absent from the literature for the SGI, several researchers have identified the importance of scale in studies of

SGI processes. The concerns of how sampling resolution will impact attempts to interpret or model the biogeochemical function of SGI interactions is more important than ever now that data users, including modelers, managers, and decision makers, are often thinking at river network scales (Krause et al., 2011). This will lead to an increase in demand for river network scale SGI biogeochemical data, and those seeking to collect that data must grapple with sampling effort and how resolution of sampling can impact the various data users. While SGI biogeochemical investigations at the network scale are limited, there are complementary ecological studies that offer further guidance. Ecological researchers have long known that different processes are scale-dependent and the scale at which one measures should answer the question being asked (e.g., Allen and Starr 1982; Delcourt et al. 1982). River corridor investigators addressing different research questions have observed spatial-resolution and extent dependent patterns, for example, small-scale biotic diversity as compared to larger-scale diversity in the SGI (see review paper by Vinson and Hawkins 1998) or comparing the riparian subsurface flow paths in small vs. large scales (see Dahl et al. 2007). Because it is important to understand all ecological processes at a variety of scales, the present study endeavored to assess how to best measure at an unprecedented network-scale in the SGI. This study helps raise some potential concerns about sampling schemes and their impact on understanding the SGI across spatial scales, and therefore should help guide future research interested in collecting and using data to compare processes across spatial scales. It also underscores that researchers cannot ignore that they must carefully consider what spatial sampling scheme may be best for the SGI question being asked.

A need for more assessment of sampling schemes

This study has a couple notable limitations that must be acknowledged in assessing the key differences between a Local Sampling and Longitudinal Sampling schemes. First and foremost,

both studies from the Ruhala et al. (2018) data sets are snapshots in time. While they sampled at approximately the same time of year and season for Local Sampling and Longitudinal Sampling schemes, they are not capturing any of the potential sub-annual temporal dynamics of the biogeochemistry in the SGI. In SGIs, the biogeochemistry is typically highly variable in relation to seasonal variation in nutrients, organic matter quantity and quality, and flow conditions. For example, Lambert et al. (2013) found that low aromaticity DOC accumulated in the HZ in the summer and was replaced in the wet season by more aromatic DOC, updating earlier research that had concluded that seasonal removal of DOC was relatively stable (Findlay and Sobczak 1996). Others have found that NO_3^- removal in the HZ is highly variable and dependent upon the distribution of precipitation across different seasons, as precipitation controls both productivity and routing of water through the HZ (Rahimi et al. 2015). In part, the biogeochemical variability found in this study may be due to flow variation between the two Ruhala et al. (2018) sampling periods as they observed similar biogeochemical conditions in the surface and groundwaters between sampling periods. Additionally, some variability could be due to the imprecise site selection from one year to the next, where the Longitudinal Sampling samples, while selected to overlap with the Local Sampling sites, were not taken at the exact same locations. However, given that Ruhala et al. (2018) attempted to collect at approximately the same locations both years and the results from the Local Sampling indicating that variability is fairly well-characterized by a single MINIPPOINT at a location as compared to three MINIPPOINTS at the same location, we expect that the variability captured in the Longitudinal Sampling should reflect the specific site from year to year.

This difference in flow conditions raises a second notable limitation to this study in that it is a comparison between two separate years. While Ruhala et al. (2018) attempted to carry out the

study at similar times and seasonal conditions in each year, the hydrologic conditions were not identical, nor will they ever be in most stream systems between different sampling events. In many stream systems, shifts from high to low or low to high flow conditions can weaken or even reverse SGI exchange patterns (e.g., Wroblicky et al. 1998; Boano et al. 2010) as well as change the quantity and quality of solutes delivered to the SGI, such as DOC (e.g., Byrne et al. 2014; Fasching et al. 2015). Many of the limitations listed above are allayed due to the well documented hydrologic stability of Augusta Creek (e.g., Poff et al. 1997). Given that the majority of Augusta Creek stream water arrives in the channel through groundwater flowpaths (Hamilton et al. 2018), the surface water flow fluctuations and impacts on the SGI exchange patterns are buffered and minimized. This is to say, many of the variable flow and storm response effects commonly seen in the SGI of other streams are attenuated by the consistent groundwater inputs in this particular stream system and do not seem to shift the overall biogeochemical conditions of the stream (Figure 8). Consequently, despite these potential limitations with the data, we think that the comparison of the Local Sampling and Longitudinal Sampling data sets is useful and informative for assessing how the two sampling schemes yield different information, especially given that there is a paucity of network-scale SGI biogeochemical assessments available.

In many cases available data do not exist or, in the case of Ruhala et al. (2018), are not ideal for comparing Local Sampling and Longitudinal Sampling schemes. Therefore, in the future, if there were sufficient people and equipment to conduct simultaneous sampling using both Local Sampling and Longitudinal Sampling schemes, it would make for a more robust assessment of the strengths and weaknesses of each sampling scheme as well as tests of our hypotheses. Still, the present study results suggest that this larger investment in testing each study scheme is likely

warranted, because it may illustrate that different network-scale patterns of the SGI biogeochemistry appear depending upon where you sample in the stream network, and inform how researchers and water quality managers can expand methods to conduct SGI studies at larger scales compatible with current watershed management plans and models (Hester and Gooseff 2010; Krause et al. 2011; Harvey and Gooseff 2015).

Conclusions

Based on the findings in this study, we recommend an increased focus on spatial sampling schemes in SGI studies. We also found evidence in the study watershed that longitudinal sampling of the SGI in favor of characterizing local heterogeneity when one is interested in characterizing the SGI across a network. We must find the most efficient means of sampling, because SGI sampling is highly demanding of both labor and costs. From our initial assessment here, we determined that there was not much added value (i.e., detection of biogeochemical variability) with an increased effort in the characterization of local plot-scale heterogeneity (Local Sampling). There were, however, new biogeochemical patterns revealed in the watershed as the sampling scheme shifted to increase the number of plots sampled in longitudinal directions (Longitudinal Sampling), because it allowed the same sampling effort to be distributed across more of the stream network.

Overall, there is a need to investigate what the best practices are for collecting SGI data at watershed scales. Without data from the SGI at the scales of watersheds and across river networks, it may not be possible to assess and upscale the ecological function that SGIs play in network-scale processes, such as nutrient budgets and water quality management (Harvey and Gooseff 2015; Abbott et al. 2016). As highlighted here, a clear, current limitation to assessing the role of SGIs in river corridors is the absence of studies of the SGI attempted at a watershed

scale. Hence, a possible way forward is to collect more SGI data sets at the stream network-scale from different study regions and from a particular stream network across different seasons. Overcoming this data gap will permit future researchers to evaluate if our findings from the Augusta Creek data set are robust in terms of sampling strategy suggestions and, importantly, facilitate assessments of current sampling effort utility and inspire new sampling strategies.

Acknowledgments

Thank you to Michigan State University Department of Earth and Environmental Sciences, Michigan State University Environmental Science and Policy Program, Michigan State University Kellogg Biological Station Long Term Ecological Research site, and the National Science Foundation Award EAR-1446328 for research support. Thank you to S. Hamilton for helpful comments in developing sampling methodology and comments on this manuscript. Thank you to A. Shogren for proofreading the final draft. Thank you to T. Bigg, E. Wiewiora, and T. Hampton for assistance in collecting samples and to the journal editor and reviewers whose constructive comments improved the paper.

CHAPTER 2: GROUNDWATERSHED VERSUS A SURFACE WATERSHED APPROACH TO LANDSCAPE MODELS OF STREAM CHEMISTRY: AN INITIAL ASSESSMENT OF EACH APPROACH IN A LOWLAND STREAM NETWORK

Abstract

It is increasingly important to understand stream biogeochemistry at network- and landscape scales. One common method to investigate stream biogeochemistry at these large scales is to relate characteristics of potential contributing areas of the landscape to stream biogeochemistry at different points across the stream network. This approach has been applied almost exclusively using surface watersheds delineated from surface topography. While surface watersheds are an obvious choice for many types of landscapes, we assert that groundwatersheds—the area of landscape contributing water to the stream via groundwater pathways—may be useful in a variety of settings, especially low-gradient stream networks. In this study, we conduct an initial assessment of this assertion by building linear models to predict stream biogeochemistry using surrounding land cover from both surface watershed and groundwatershed delineated areas, and compare their performance across a third-order, mixed land-use, lowland stream network. We conducted a synoptic sampling campaign across this network to collect biogeochemical data including dissolved organic carbon concentration, SUVA₂₅₄, nitrate, and chloride—five common biogeochemical measurements. First, this study found that groundwatersheds and surface watersheds were substantially different across the lowland stream network. Second, linear models predicting observed stream chemistry using groundwatershed land use often performed as well or better than models using the more traditional surface watersheds approach. Together, these findings suggest that future measurement and modeling efforts of stream network biogeochemistry may benefit from incorporating a groundwatershed approach, but it also

highlights that the availability of groundwater data still limits rigorous assessments of the strengths and weaknesses of a groundwatershed versus surface watershed approach to modeling chemistry in most streams

Introduction

Stream water quality management is enhanced by understanding stream biogeochemistry at network to landscape scales, and how it relates to the surrounding land use. In particular, research over the last two decades has increasingly sought to assess stream biogeochemistry at the network scale, generating a better understanding of the contributions of subsurface and non-channel flow to surface water biogeochemical expressions (e.g., Fisher et al., 2004; Lowe et al., 2006). Network scale analyses in lowland, mixed land use watersheds often present a unique challenge in that researchers must go to many sites across the network, requiring building relationships with landowners or traversing difficult landscapes. Hence, remotely evaluating and predicting stream chemistry is desired as it will allow a greater capacity to assess more streams.

Statistical modeling can be used to predict stream chemistry based on broadly available spatial landscape characteristics or on prevalence of specific land use types that exert controls on chemical sources and sinks (e.g., Dillon and Molot, 1997; Frost et al., 2006; Walker et al., 2012). Commonly, these statistical models are done by determining the surface watershed drainage area (i.e., the topographically determined area that hydrologically drains) to a given point in the stream network and then using the land use characteristics within that drainage area to predict biogeochemical properties of the stream (i.e., as the watersheds relate to surface water chemistry; see e.g., Hollister et al., 2008; Soranno et al., 1996). An alternative drainage area is the groundwatershed, which can be delineated using subsurface drainage information (i.e., the groundwater levels that determine groundwater flow in the subsurface). The groundwatersheds

represent the cumulative groundwater flowpaths contributing to surface water at a point in the stream network. Yet with few exceptions groundwatersheds have almost never been considered in the prediction of surface water biogeochemical conditions, even for lowland watersheds where nearly all stream flow is derived from groundwater inflows (see e.g., Martin et al., 2017). While all the reasons for the omission of groundwatershed considerations from these landscape models is not clear, the main barrier to using a groundwatershed approach is that the spatial and temporal resolution of groundwater conditions is limited in most watersheds. This study serves as an initial assessment to compare use of surface watersheds to groundwatersheds.

Stream biogeochemistry often has a strong relationship to the surrounding land use and land cover (LULC; see e.g., Herlihy et al., 1998; Inwood et al., 2005). In the last few decades, research has revealed that local spatial heterogeneity can alter this relationship, especially at sediment-water interfaces (e.g., riparian and hyporheic zones) that may function as highly reactive control points and cause dramatic changes in surface water concentrations of reactive biogeochemical constituents through various retention and transformation processes (McClain et al. 2003; Bernhardt et al. 2017). However, larger scale patterns are often observed despite these important localized process heterogeneities. McGuire et al. (2014), for instance, found that large scale patterns in stream carbon chemistry were related to landscape scale controls on stream network biogeochemistry.

Landscape ecologists have long understood the utility of statistical models to predict spatial patterns across the landscape (see Turner and Gardner, 2015 for a thorough review of spatial statistics in ecological studies). In watersheds and streams, ecologists have applied these spatial modeling tools to predict surface water biogeochemical constituents as they relate to the landscape features (i.e., LULC attributes). For example, multiple linear regression models have

been used to predict concentrations dissolved organic matter (Mattsson et al. 2005), heavy metals (Xiao and Ji 2007), and nutrients (Dodds and Oakes 2006), all in relation to the surrounding LULC. These models are generally successful in predicting surface water chemistry because landscape interactions have strong influence on the hydrology and chemistry of the landscape that is connected to the freshwater systems (Kratz et al. 1991).

It is well established that in many watersheds, source areas identified through topographic delineation of hydrology (i.e. surface watersheds) do not necessarily match observed chemistry in stream networks (Hinton et al. 1993). To some degree, this mismatch is often attributed to subsurface hydrogeologic heterogeneities driving preferential flow paths that are not represented by surface topography, thus affecting the overall source area for stream waters. This mismatch can be more pronounced in lowland landscapes and river networks with large fluvial plains (e.g., Hinton et al., 1993; Woessner, 2000). For example, a regional modeling effort in Grand Traverse Bay, MI indicated that groundwatersheds in low-relief areas may vary substantially from surface watersheds (Boutt et al. 2001). Further, in some glaciated landscapes groundwater flow does not conform to topographic divides, with the areas of groundwatersheds deviating from those of overlapping surface watersheds by as much as 57% (Hinton et al. 1993).

There is a long history of calls to further integrate our understanding of groundwater into stream ecological studies driven by the idea that most lotic systems are primarily fed by groundwater (Hynes 1983), with lowland streams being almost entirely fed by groundwater (Sear et al. 1999). Given the typically high contributions of groundwaters to lowland stream flows, several studies have since illustrated the biogeochemical importance of groundwater inputs to surface waters. For example, we now know that groundwater inputs can dramatically change the dissolved organic matter composition in the stream (Mulholland 1997) as well as the nitrogen

chemistry in streams (Hill 1990), two important reactive biogeochemical constituents that fundamentally control many water quality and ecosystem conditions in streams. Consequently, Stanford (1998) argued that groundwater upwelling is one of the three dimensions that is often overlooked and must be considered by stream ecologists, with lateral inputs and longitudinal flow being the other two dimensions. Despite the need to incorporate the role of landscape groundwater conditions in stream biogeochemical studies, the actual inclusion of these conditions in stream reach and network studies has been difficult given limitations stemming from a lack of groundwater flow observations and uncertainties in modeling groundwater flow.

There are many potential methods to define the area over which LULC contributes to freshwater systems, including riparian corridors (Allan et al. 1997; Strayer et al. 2003), flow distance from a feature (Brazner et al. 2007), buffered influence zones along the stream network (Floyd et al. 2009), and entire upgradient watersheds (Soranno et al. 1996; Wang et al. 2011; Martin et al. 2017). For this study, we focus on the upgradient watersheds as areas where LULC contribute to the stream network of interest. However, we look at two different ways of delineating upgradient watersheds, either based on topographic (i.e., surface watershed) or water table surface (i.e., groundwatershed). We do this because there remains uncertainty about which delineation method is most appropriate to determine LULC contributing area that affects stream biogeochemistry. The surface watershed area includes the entire potential area that could contribute to any given point in the stream if water exclusively flowed over the landscape surface (i.e., overland flow), however it may only be a fraction of the landscape that is contributing to a point in the stream due to surface and subsurface hydrologic routing (e.g., Hinton et al., 1993). However, in many watersheds overland flow is not the main pathway for water to move from land to the stream network; rather, subsurface flow paths can be dominant. Hence,

groundwatersheds may also be used, provided enough data (measured or modeled) are available to define the groundwater source (recharge) area, which requires knowledge of groundwater levels and flow direction.

This study serves as an initial assessment comparing use of surface watersheds to groundwatersheds in a lowland stream network. Our goals in this study are two-fold. First, in order to improve modeling of how LULC relates to stream water chemistry, we assert that for many streams it is important to assess stream hydrologic and chemical conditions using groundwatersheds as well as surface watersheds. This assertion is timely because groundwater elevation measurements are increasingly available, which facilitates the consideration of groundwatershed approaches. Further, there is a recent precedent supporting the use of groundwatersheds to understand surface water quality as it relates to LULC (Martin et al., 2017), although it was not focused on streams. Our second goal is to directly test this groundwatershed assertion with observations in Augusta Creek, a heavily-researched lowland watershed in the Midwest United States which is predominantly fed by groundwater (e.g., Hamilton et al., 2018; Poff et al., 1997; Vannote et al., 1980) (Figure 9). Hence, in this study, we assess whether surface water biogeochemistry across a stream network is more appropriately modeled using surface watershed drainage areas derived from surface topography or from groundwatersheds derived from regional water table levels. We hypothesized that groundwatersheds would outperform surface watersheds in modeling stream water biogeochemistry of Augusta Creek, given the lowland landscape and predominance of groundwater discharge to the stream network.

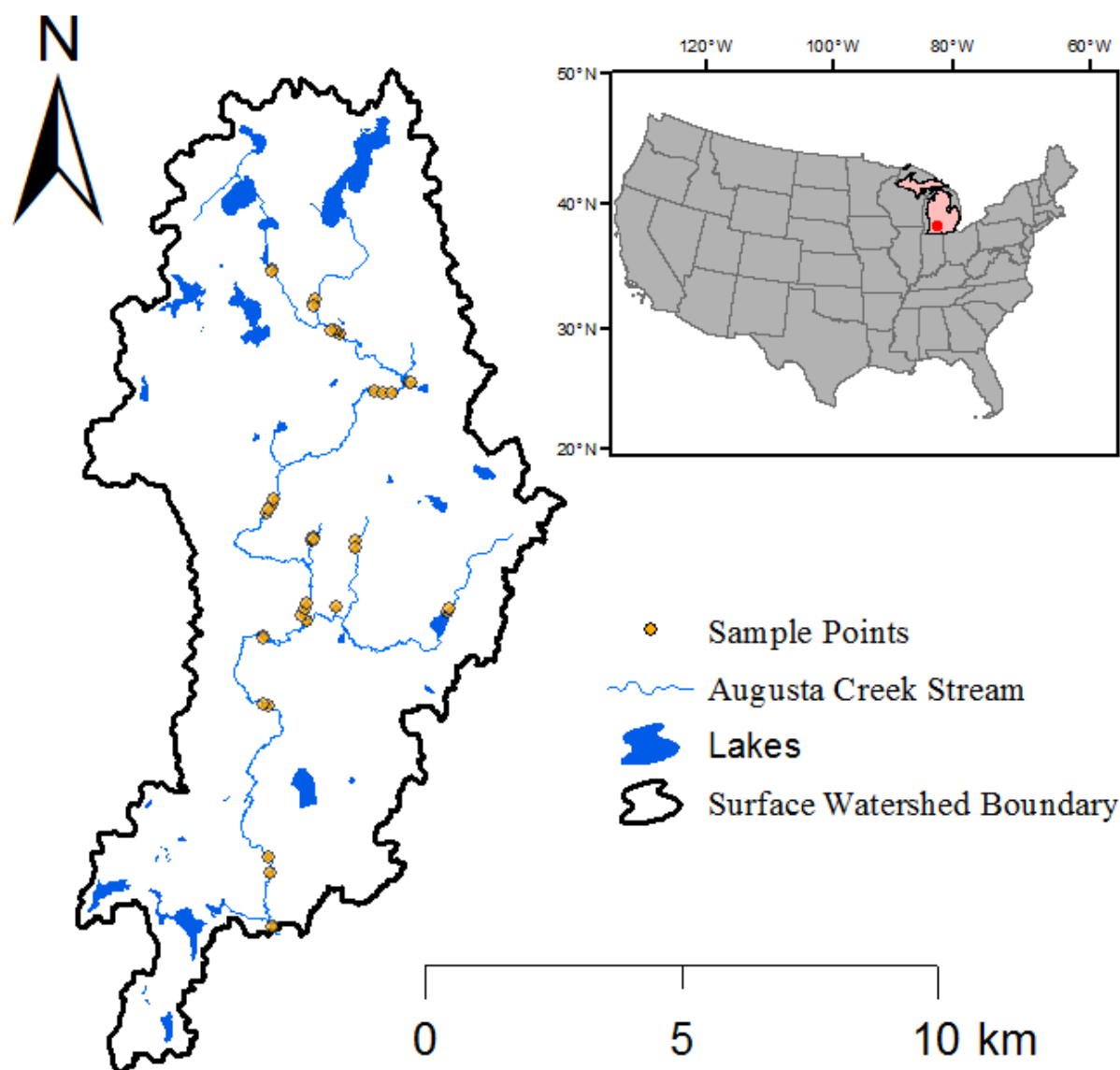


Figure 9 - The Augusta Creek watershed (black outline) and stream network (blue lines), with its location highlighted in the inset map of Michigan, USA (upper right). Sampling sites (yellow points) are distributed across all orders in the stream network.

Materials and Methods

Site Description

The stream chemistry data used in this study were collected in the Augusta Creek watershed, a third-order, lowland, groundwater-fed stream in southwest Michigan (Figure 9)

over the period of August 16-22, 2016. Augusta Creek is composed of mixed-use landscape including wetlands, lakes, pasture, upland forests, and row crops. The stream network of Augusta Creek overlies a surface geology of glacial tills. The surface flow is groundwater dominated with minimal contributions from overland flow (Poff et al. 1997; Hamilton et al. 2018). This data set encompasses sampling points across all three stream orders and includes sampling representative of the different stream headwater origins (e.g., lake outflows, upland forests, wetland outflows). The Augusta Creek network passes through the Kellogg Biological Station Long Term Ecological Research program, possesses an active USGS gaging station (04105700) with more than 50 years of discharge data, and is of historical research significance as part of the Natural Flow Regime (Poff et al. 1997) and River Continuum Concept (Vannote et al. 1980) studies.

Synoptic Sampling Campaign

Stream chemistry was collected via a spatially-extensive synoptic sampling campaign, detailed in Ruhala et al. (2018). Briefly summarized here, stream water was collected across 39 sites that include sites within first-, second-, and third-order stream reaches. For each sample 80 mL of water was collected. First, 20 mL of water were used to prime and rinse two filters in series: a 0.7 μm nominal pore size prefilter (Whatman GF/F filters) followed by a 0.2 μm filter (Sartorius Stedium cellulose acetate). The remaining 60 mL was then filtered into acid-rinsed HDPE amber bottles and stored on ice until the end of the sampling day after which it was stored at 4°C and analyzed within 28 d.

Each filtered sample was analyzed for Dissolved Organic Carbon (DOC) as non-purgeable organic carbon on a TOC-L total organic carbon analyzer (Shimadzu Scientific Instruments, Kyoto, Japan) with Pt catalyzed oxidation at 680°C. Samples were all analyzed for anions, in particular for this study Cl^- and NO_3^- , on a Dionex ICS-2100 Ion Chromatography

System (Thermo Fisher Scientific, Massachusetts, USA). Finally, all samples were assessed for their optically-derived DOC properties through absorbance measurements on a Shimadzu dual-beam ultra-violet 1800 spectrophotometer (Shimadzu Scientific Instruments, Kyoto, Japan). In this study we assessed Specific Ultraviolet Absorbance at 254 nm ($SUVA_{254}$, the absorbance at $\lambda = 254$ nm divided by DOC concentration; Weishaar et al., 2003), ($L\ mg-C^{-1}\ m^{-1}$).

Watershed delineations and land use

We delineated and measured areas of both surface watersheds (the area of drainage based on surface topography) and groundwatersheds (the area of drainage based on groundwater levels) to each individual surface water sampling site. These watershed delineations were completed using ArcGIS 10.6 (ESRI, California, USA) following the approach detailed in Martin et al. (2017). Surface topography was extracted from the 1/3 arc-second (10 meters) USGS 3DEP digital elevation model. Following topographic extraction, we then filled local sinks, determined flow directions using a D8 flow direction algorithm (allowing only the 8 cardinal and intermediate directions), and computed watersheds from those flow directions.

For this study we selected a straightforward approach to delineate groundwatersheds that relies on readily available groundwater level data and groundwater flow principles. For this delineation, we assume that 1) flowpaths within the groundwater system are predominantly two-dimensional (i.e. flow is lateral rather than vertical), and 2) that vertical gradients in hydraulic head within the unconfined glacial aquifer beneath the Augusta Creek watershed are small relative to larger lateral gradients in hydraulic head (Rheaume 1991). The relatively low relief, shallow nature of the aquifer, and relatively high hydraulic conductivity of the glacial outwash sands and coarse-textured tills supports these two assumptions. With these assumptions, groundwater flow direction is entirely dictated by the shape of the water table—the groundwater

topography. To create this surface, we interpolated water tables using static water levels collected at the time of installation from Michigan’s drinking water wells (Appendix A, Figure 25). These measurements are scattered across and beyond the watershed, and through time—with some records in the region dating back to the 1970s.

Our groundwatershed delineation method mirrors traditional surface watershed delineation with one additional step following Haacker et al. (2016) to better define the surficial aquifer water levels given the propensity of lakes in the area to be dominated by groundwater inflows (Robertson and Hamilton 2015). First, we interpolated the elevation of the water table from the static water level measurements using Empirical Bayesian Kriging, a method that removes much of the manual model selection and guesswork from traditional kriging approaches (Krivoruchko, 2012). Second, we then identified surface water features below the interpolated water table and assigned a DEM elevation to these features, assuming that the water table elevation at those locations is given by the surface elevation. These locations can then provide additional “measurements” of water level to better define the groundwater topography. To incorporate the surface water features into the interpolation, we had to reduce the spatial density of measurements by randomly sub-selecting the vertices of these surface water features to match the density of measurements from groundwater wells. Next, we merged the static water levels with the DEM-assigned surface water feature elevations and interpolated the merged set again using Empirical Bayesian Kriging. Finally, we constrained the interpolated water table to lie at or below ground elevation.

We quantified surface water and groundwatershed similarity to facilitate comparisons in each approach, by using two different but complementary metrics for percent unique area (i.e.,

the area of either the surface water or groundwatershed not shared with the other). These metrics are defined as:

$$U_{GW,i} = \frac{A_{GW,i} - A_{shared,i}}{A_{GW,i}} \times 100 \quad (1)$$

$$U_{SW,i} = \frac{A_{SW,i} - A_{shared,i}}{A_{SW,i}} \times 100 \quad (2)$$

where U is the unshared area percentage of either the groundwatershed (GW) or surface watershed (SW) and the A terms are the areas of either the groundwatershed, surface watershed, or that intersection of the two ($shared$) for each watershed, i . A unique percentage of 0 for a watershed indicates that the entire watershed of one approach spatially overlaps with a watershed of the other approach. Likewise, a unique percentage of 100 would indicate that there is no spatial overlap between the surface and groundwater watershed approach.

Augusta Creek land use data were obtained from the National Land Cover Data Base (NLCD) 2011 (U.S. Geological Survey 2014; Homer et al. 2015). The NLCD 2011 is composed of 8 categories each with subcategories: Water, Developed, Barren, Forest, Shrubland, Herbaceous, Planted/Cultivated, and Wetlands. In this study, we combined all Developed subcategories into a single category because Augusta Creek is a rural area with sparse areas of development. For each surface and groundwatershed we determined the relative proportion of land use categories by taking the area of an individual land use within a watershed divided by the total area of the watershed.

We constructed multiple linear regression models in R (R Core Team 2017) using land use to predict surface water biogeochemistry. All 39 subwatersheds were included in the models. The primary function of these models was to compare their performance when using either surface or groundwatershed information per the scope of this initial study, rather than trying to create the best absolute model for predicting Augusta Creek stream biogeochemical conditions.

We recognize that a combination of both surface and groundwatershed derived landscape characteristics may ultimately best describe and predict stream geochemistry or water quality.

The multiple linear regression model construction used the following generalized ordinary least squares (OLS) equation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon \quad (3)$$

where y is the surface water parameter we wish to predict (i.e., DOC, SUVA₂₅₄, Cl⁻, or NO₃⁻) normalized to the maximum value, x is the landscape proportion within the watersheds, β is the linear coefficient of the landscape unit proportion, and ε is the random error. We created separate OLS equations for the LULC for surface watersheds and for the LULC of groundwatersheds.

While linear models are relatively simple and transferable method, when they are built on fractional land use, they can be highly multicollinear necessitating additional procedures. To address this multicollinearity, we implemented two stepwise procedures in sequence that generated a suite of reduced variable models. First, we started with all landscape variables in the model and removed these in a stepwise manner using the Variance Inflation Factor (VIF) to reduce multicollinearity in the models. We calculated VIF values for each independent variable in the model. If any variables had VIF values greater than a threshold, here set to 3 (Zuur et al. 2010), then up to two variables with the highest VIF were individually removed. Each variable removal generated a new model, which was then subjected to this stepwise procedure. Once models passed the VIF criterion (i.e. no remaining variables had VIF values exceeding the threshold), our second stepwise procedure further reduced model complexity by selecting independent variables using stepwise regression, with a 0.05 significance criterion for the independent variables through backwards stepwise regression (after de Koning et al., 1998).

Once a suite of OLS models were selected using the two-step procedure above we developed the OLS models into mixed regressive-spatial autoregressive models to account for any potential spatial autocorrelation in the model residuals (see e.g., Anselin, 1988; Overmars et al., 2003) which may result from both the nested nature of watersheds as well as the close proximity of sampling points within reaches. This was done because a majority of the OLS models were spatially autocorrelated as assessed using the Moran's I test for spatial autocorrelation. The mixed models were formulated as follows:

$$Y = \rho Wy + X\beta + WX\gamma + e \quad (4)$$

where ρ is the autoregression coefficient and fit first, W is the spatial weights matrix based on distance and using nearest neighbors, X is the original predictor variable matrix from Equation 3, β is a vector representing slopes of the original independent variables from the predictor matrix (i.e., $\beta = [\beta_0, \beta_1, \beta_2, \beta_k]$), γ is the autoregression coefficient of the spatially lagged explanatory values (WX) and also a fit parameter, and e represents the spatially independent errors.

Finally, we applied Akaike's Information Criteria (AIC) to choose the most parsimonious models for each constituent, some of which had dissimilar structures (i.e. different numbers of retained predictor variables). The most parsimonious model and all models within a value of AIC = 2 (Burnham and Anderson 2002) of the most parsimonious models (ΔAIC) were selected as "best-fit" models.

To assess model predictive power, in addition to the AIC we used a Pseudo R^2 value. We used Pseudo R^2 because R^2 does not properly account for predictions from spatial lag models, which take spatial auto-correlation into consideration, because one cannot give each individual measurement equal weight in goodness of fit. Therefore, we used a Pseudo R^2 measure of fit equal to the variance of the predicted values over the variance of the observed values (i.e.,

$\frac{\text{Variance}_{\text{predicted}}}{\text{Variance}_{\text{observed}}}$; Anselin, 1992; Overmars et al., 2003). Note, by this definition the Pseudo R^2

matches the coefficient of determination R^2 for OLS regression alone. A value of 1 for the model indicates a perfect model fit, however due to the mixed autoregressive modeling, Pseudo R^2 for subsets of the entire dataset can exceed 1.

After producing the mixed spatial lag models, we investigated model performance as compared to specific subsets of watersheds as an exploration of where our spatial lag models performed well or poorly. These subsets included watersheds based on predominant LULC category and watersheds of a particular stream order (ranging from first- to third-order). We evaluated performance of models within these subsets with the aforementioned Pseudo R^2 values. Ultimately, the surface versus groundwatershed approach to predicting stream chemistry can be quantitatively assessed by comparing the Pseudo R^2 values for the different observed water chemistry conditions (i.e., Cl^- , NO_3^- , and DOC, and SUVA_{254}) across the stream network.

Results

General stream network surface water biogeochemistry conditions

A summary of the surface water biogeochemical measurements collected across the watershed are shown in Figure 10. Concentrations of Cl^- , NO_3^- , and DOC ranged from 6.4 to 12.7, 3.8 to 14.4, and 5.3 to 16.3 mg L^{-1} , respectively. SUVA_{254} values ranged from 2.7 to 4.0 $\text{L mg-C}^{-1}\text{m}^{-1}$. These SUVA_{254} values correspond to an estimated 21% to 30% aromatic content of DOC using a linear regression developed by Weishaar et al. (2003).

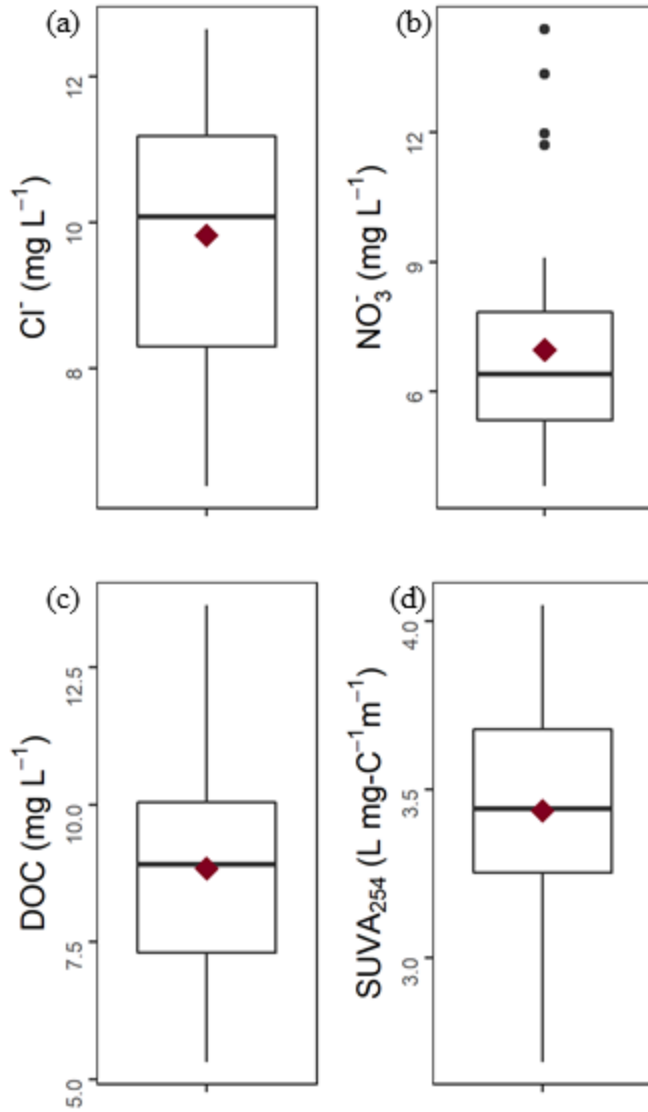


Figure 10 - The observed Cl⁻, NO₃⁻, DOC, and SUVA₂₅₄ conditions across the entire Augusta Creek stream network from the surface water synoptic sampling campaign (n=39 sample sites). Box and whiskers represent interquartile ranges, while red diamonds represent mean values and black circles represent outliers.

Watersheds and land use results

Surface watersheds ranged in size from 0.80 km² at the smallest first-order watershed to 97.38 km² at the largest third-order watershed, while groundwatersheds varied more from 0.05 to 112.81 km² (Figure 11). For each water sampling point, upstream groundwater and surface watershed areas were strongly and significantly correlated ($R^2 = 0.98$, $p < 0.05$). Below ~25 km²,

surface watersheds were typically larger than groundwatersheds, while, above $\sim 25 \text{ km}^2$, groundwatersheds were always larger than the corresponding surface watersheds. Surface watersheds had a range of percentage uniqueness ranging from 12.53 up to 93.89% with a median value of 43.90% (Figure 11). Groundwatersheds had a range of percentage uniqueness ranging from 0 to 96.90% with a median value of 35.05%. In terms of absolute size, 27 out of 39 groundwatersheds had larger areas than their surface watershed counterparts (Appendix A, Table 3).

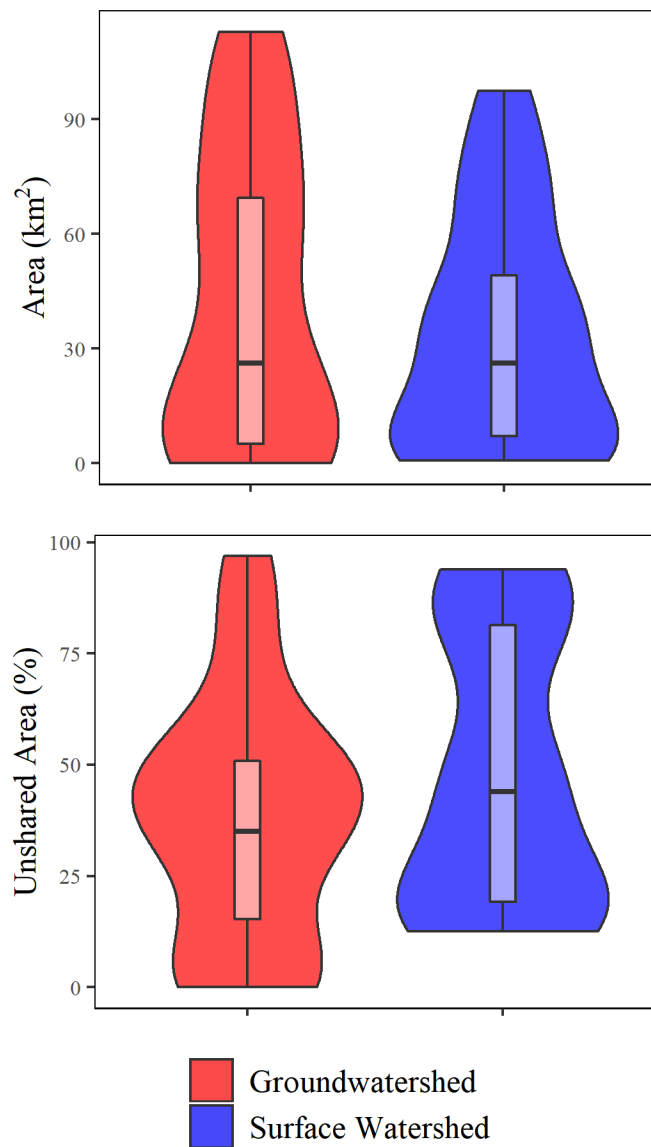


Figure 11 - Violin plots showing the distribution of (a) watershed areas (b) and unshared area percentage for groundwatersheds and surface watersheds (defined via Equations 1 and 2). The curvilinear outline and color fill of the violin plot shows a mirrored kernel density plot of values, while the internal boxplots show interquartile ranges. It is evident in (a) that the majority of areas for both surface watersheds and groundwatersheds are concentrated in smaller ($<30 \text{ km}^2$) areas. While the overall distributions of areas vary, including a greater maximum area in groundwatersheds, median values are similar between both. Note that the distributions of unshared areas differ substantially between groundwatersheds and surface watersheds.

In Augusta Creek, the land cover from the NLCD 2011 classification system (Homer et al. 2015) included Open Water, all levels of Developed (Low, Medium, High, and Open), Barren

Land, all types of Forest (Deciduous, Evergreen, and Mixed), Shrub/Scrub, Grassland/Herbaceous, Pasture/Hay, Cultivated Crops, Woody Wetlands, and Emergent Herbaceous Wetlands. The predominance of LULC type was different between the surface and groundwater watershed areas. For the largest surface watershed extent, Deciduous Forest is the predominant land cover, followed by Cultivated Crops and Pasture/Hay, while High Developed was the least common land cover, followed by Medium Developed and Barren. By contrast, for the largest groundwater watershed extent, Pasture/Hay was predominant, followed by Cultivated Crops and Deciduous forest, while the least common landcover types were the same as surface watersheds (See Figure 12).

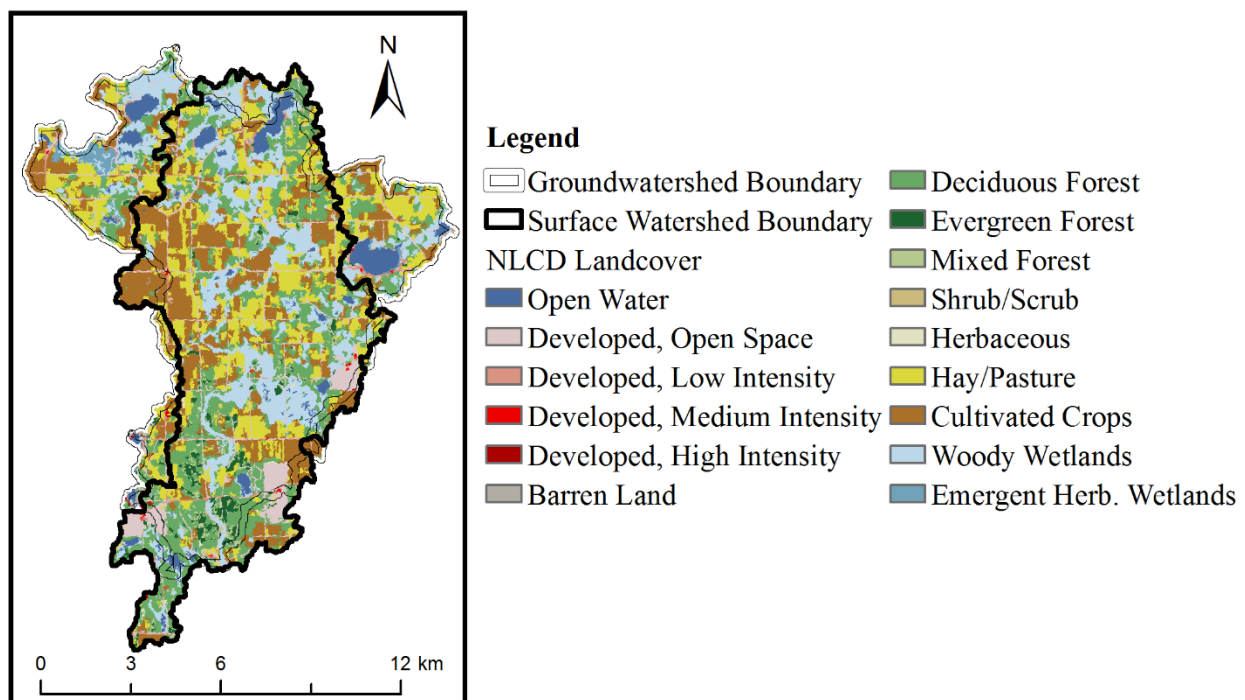


Figure 12 - LULC in the August Creek stream network for both surface watershed (black border) and groundwater (white border) extent. Note that the groundwater includes large open water body, wetlands, and additional agricultural lands that are not captured in the surface watershed.

Model results

Ultimately, following stepwise variable reduction, all selected models contained between 3 and 5 independent variables encompassing 12 land use categories (Appendix A, Table 4). The number of models that passed our criteria for any given biogeochemical measurement varied from 1 to 3 (Appendix A, Table 5)

For all biogeochemical measurements assessed, groundwater models had the lowest AIC (i.e., performed as the best-fit models; Figure 13). If we consider a threshold ΔAIC of 2 to indicate similar model performance (a typical assumption; Burnham and Anderson, 2002), surface watershed and groundwater models performed similarly for both DOC and $SUVA_{254}$. Conversely, Cl^- and NO_3^- surface watershed models both had poor performance as compared to their respective groundwater models.

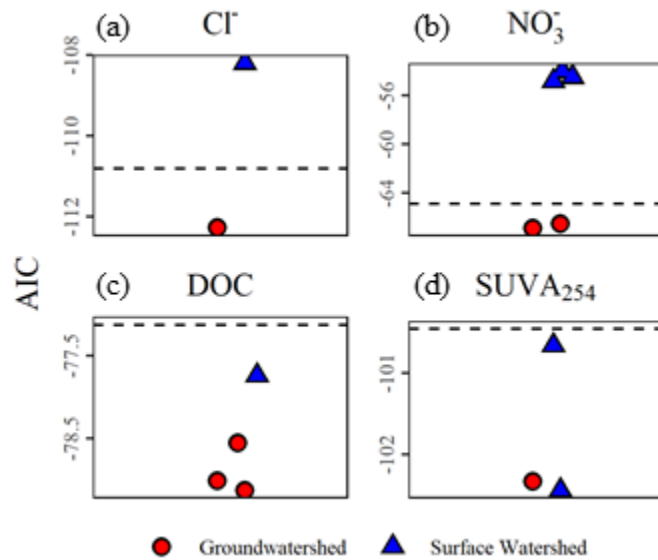


Figure 13 - An overall comparison of the groundwater versus surface watershed model performance is shown in the AIC values for all best-fit models across biogeochemical measurements Cl^- (a), NO_3^- (b), DOC (c), and $SUVA_{254}$ (d). Each model is represented by a point, red circles are groundwater models and blue triangles are surface watershed models. All points that fall below the dashed line are considered to be comparable to the best-fit model. Models above the dashed line are considered to be worse than the best-fit model.

For our second assessment of model performance, we found that for Cl^- and NO_3^- the groundwatershed models had greater Pseudo R^2 than surface watershed models (Figure 14). In contrast, we found that surface watershed models outperformed the groundwatershed models—according to Pseudo R^2 —for DOC and its optical property SUVA_{254} . Overall, we found that Cl^- models for both watershed types had the highest Pseudo R^2 values, followed by DOC and NO_3^- , and then by SUVA_{254} .

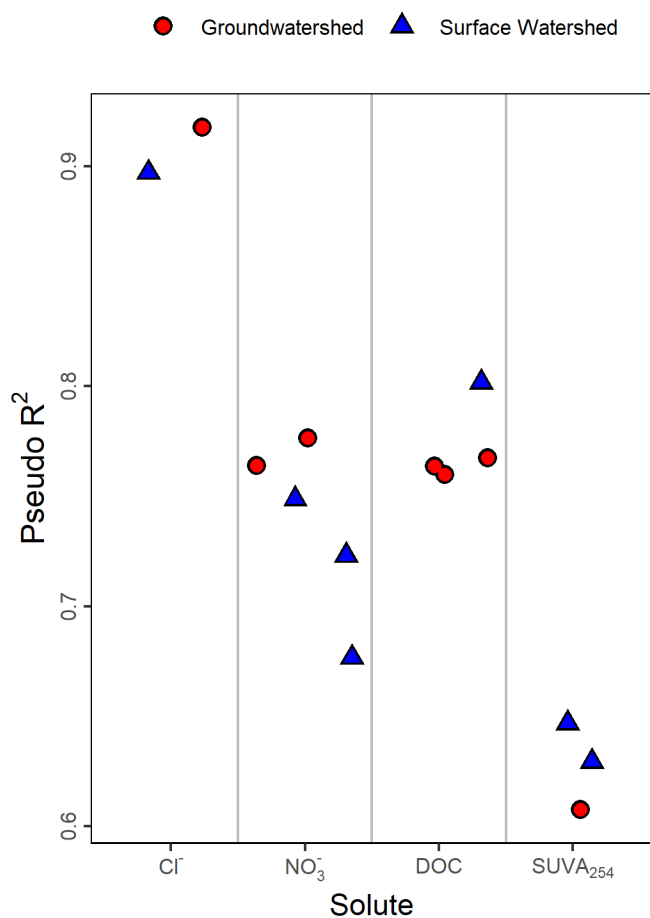


Figure 14 - A comparison of how groundwatershed and surface watershed models compare across solutes. Pseudo R^2 values for best-fit models based on either groundwatersheds or surface watersheds across all biogeochemical measurements. Red circles are groundwatersheds and blue triangles are surface watersheds. Values closer to 1 indicate the variance of predicted values is closer to the variance of observed values.

Discussion

Surface watershed and groundwatershed differences and potential implications

Only 4 groundwatershed-surface watershed pairs had comparable areas (i.e., < 10% difference in size) of the 39 unique surface subwatersheds delineated to our sampling points in Augusta Creek, , and of these four there was only a maximum of ~50% overlap between their surface and groundwatershed. While the relative overlap between surface watersheds and groundwatersheds varied depending upon location within the Augusta Creek stream network (Figure 15 and Appendix A, Table 3), it is clear that the selected contributing watershed delineation method will likely affect any subsequent analysis dependent upon the watershed area or spatial attributes of that watershed area. Therefore, researchers and watershed stakeholders should be mindful of which delineation is selected and how this decision affects their subsequent research objectives, analyses, and decisions. For instance, at a sampling point near to the headwaters of a second-order watershed in Augusta Creek, we find that a transition from a surface watershed to a groundwatershed shifts the land use type driving the modeled conditions from deciduous forest to woody wetland. This means that the initial watershed delineation step may influence any management decisions based on LULC made within this watershed.

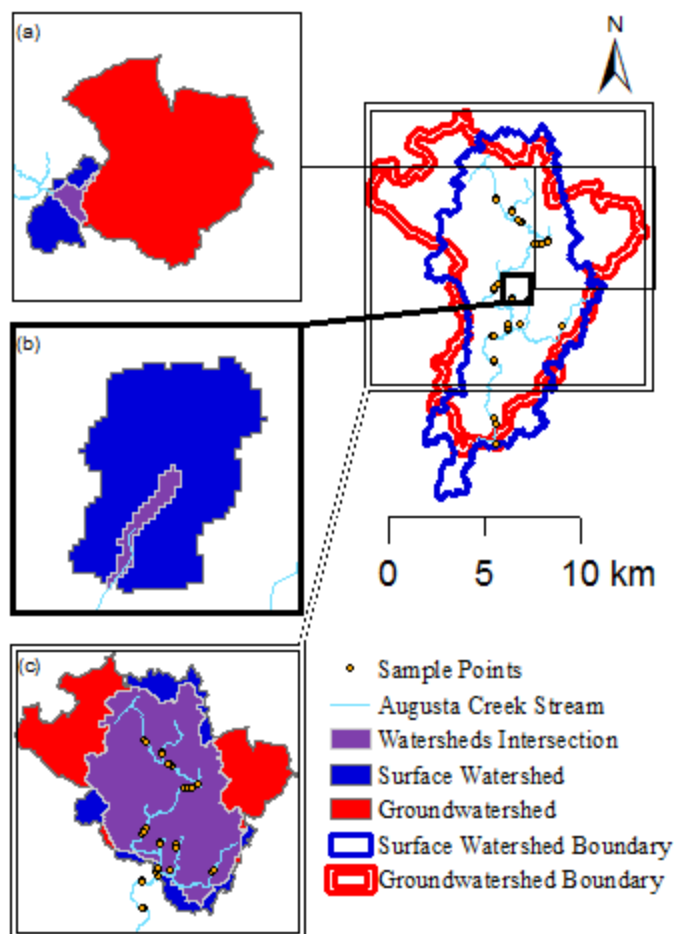


Figure 15 - Comparison of groundwatersheds (red) and surface watersheds (blue), as well as their overlap (purple) in Augusta Creek. The three insets present examples of (a) a low degree of overlap for surface and watersheds delineated to a point, (b) a high degree of overlap for the groundwatershed, and (c) high degree of overlap for both surface watersheds and groundwatersheds.

We think the clear differences documented in Augusta Creek between surface and groundwatersheds delineations supports the need to, at a minimum, consider groundwatersheds in lowland freshwater ecosystem research and management. Further, there is evidence here suggesting that groundwatersheds and surface watersheds may differ considerably for a stream network. This should highlight the need for hydrologists, biogeochemists, ecologists, and managers to work together to understand these freshwater systems. This supports past calls (see e.g., Woessner, 2000) for greater interdisciplinary collaboration in efforts to understand the

interactions between different components of a watershed (e.g., groundwater, riparian zone, hyporheic zone, etc.). Here we demonstrate that the typical assumption of the surface watershed approach—that the surface water is derived across the topographically defined surface watershed—could be in error if, in fact, the water is primarily sourced from the less commonly used groundwatershed. Reevaluating this fundamental assumption could impact, for example, how land managers evaluate stream water quality as it relates to the surrounding LULC.

It should be noted that there are many other empirical and process-based modeling methods of groundwater contributing area delineations that could provide different interpretations than those presented in this study. Again, the primary function of watershed delineation and linear regression models used in this study was to compare their performance when using either surface versus groundwatershed information, rather than trying to create the best absolute model for predicting Augusta Creek stream biogeochemical conditions. More complex, process-based approaches for groundwater flow and solute transport at the watershed scale are rapidly developing and are likely to outperform the approach used for this study. Nevertheless, we have demonstrated that groundwatershed models created from a straightforward, widely transferrable method can outperform surface watershed models within the August Creek Watershed for multiple aspects of water chemistry fundamental to water quality.

Comparisons of different groundwater and mass transport modeling methods is an active area of hydrological research but is beyond the scope of this initial assessment focused on comparisons of comparable surface and groundwatershed delineation methods (i.e., comparable in terms of data needs and complexity of implementation). Most groundwatershed delineation methods require additional information and effort to implement than that used in this study. For

instance, Chow et al. (2016) created groundwatersheds using several different finite element and finite difference flow and transport models. Their models include hydraulic gradient, as we did in this study, as well as flow velocity and multiple hydrogeologic subsurface properties. While Chow et al. (2016) and the many other methods capable of producing groundwatersheds exist and will likely produce unique results, there is no consensus across these models for which is the ‘correct’ watershed and often there is a large data and computational burden for the researcher to meet before implementing these models. Overall, we recommend that future studies trying to predict stream chemistry via statistical and process-based models should be pursued in common locations where the data burden is met for comparing multiple groundwater delineation methods.

Stream Biogeochemistry: Surface watershed linear models vs. groundwatershed linear models

The groundwatershed linear models generally performed better than surface watershed linear models in predicting biogeochemistry in the stream as it relates to the land use within the watershed area, especially for important water quality drivers of Cl^- and NO_3^- . When groundwatershed models did not perform better it was for DOC and SUVA_{254} , but the models were nearly always comparable to surface watershed model performance. This overall enhanced performance of groundwatershed based models in the lowland Augusta Creek watershed suggests that groundwatersheds may be preferable for research or water management goals, especially those associated with Cl^- and NO_3^- in the region of this study. Further, it suggests that groundwatershed spatial models may be comparable to surface watershed modeling efforts for the more complex organic matter stream properties of DOC, and SUVA_{254} . Overall, we consider the linear model results to be clear initial evidence that groundwatershed-based landscape models

are worth including in future biogeochemical or water contamination investigations of Augusta Creek and similar lowland groundwater stream networks of the study region.

The similarity of SUVA₂₅₄ and DOC model performance were most comparable between surface watershed models and groundwatershed models is interesting. It is unclear from our study approach why these DOC conditions were similar. However, it suggests that the DOC conditions measured across the stream network at the time of sampling were not strongly dominated by either surface watershed or groundwatershed attributes. This is likely due to the complex flowpaths and processing of DOC and SUVA₂₅₄ signals in the watershed (Ruhala et al., 2017). For SUVA₂₅₄, in the groundwater dominated Augusta Creek (Poff et al. 1997; Hamilton et al. 2018), the groundwater is very likely acting as a constant supply of poorly aromatic (i.e., low SUVA₂₅₄) DOC during the summer months when our stream measurements occurred, whereas the highly aromatic (i.e., high SUVA₂₅₄) representing the landscape signature will be more pronounced in the stream during the wet season when hillslope soil through flow is larger and more readily connecting local landscape features to the stream (see, e.g., Lambert et al., 2013). This could mean that the relationship between LULC and SUVA₂₅₄ is not easily observed in the stream during the summer months, where localized processing may have a greater impact on the conveyance of the groundwater SUVA₂₅₄ signal to the surface waters. Hence, neither surface nor groundwatershed approaches may be able to accurately predict SUVA₂₅₄ without a more complex approach that includes these temporal/seasonal drivers of SUVA₂₅₄.

Additionally, it was previously demonstrated that Augusta Creek stream water has a dynamic equilibrium with respect to DOC, and only has small fluctuations both within and between sites on a diurnal basis (Manny and Wetzel 1973). This dynamic equilibrium of DOC is attributed to the overall abundance of DOC and the complex influent sources of DOC within the

stream ranging from lakes to wetlands to groundwater flowpaths and their respective ecosystem processing of DOC. While these LULC categories appear in our models, the categorical nature does not allow for distinction in the specific function of a type of wetland at one location as compared to another location. While we know that SUVA₂₅₄ values can increase through a wetland (see e.g., Pinney et al., 2000), it would be inappropriate to assume that all wetlands are equally connected via surface and subsurface flow paths at all times in Augusta Creek. In some cases, wetland connectivity, and thus DOC inputs to the stream from wetlands, is seasonal or related to precipitation inputs, whereas in other cases wetlands are continuously connected (personal communication: S. Hamilton, Michigan State University).

We do not suggest that groundwatersheds must be considered by all watershed researchers and stakeholders. Together, our findings for concentrations and properties of DOC offer an initial highlight that the surface watershed versus groundwatershed approach may be less relevant when dealing with biogeochemical variables that are subject to very localized and rapid processing and transformation within the river network, such as what can happen to many solutes experiencing hyporheic, in channel, lake, or wetland transformations. These localized and rapid transformation locations are known as watershed “hot spots” and “control points” (McClain et al. 2003; Brazner et al. 2007; Bernhardt et al. 2017) and capturing them in watershed scale models is a challenge (Abbott et al., 2016). Or there may be broader, landscape scale zones adjacent to the stream network that are still smaller than the full watershed scale that exert greater influence on the biogeochemistry (e.g., Floyd et al., 2009). Neither surface nor groundwatershed scale statistical approaches will capture these more local scale drivers of instream chemistry and all watershed scale statistical water quality models are likely be

ineffective until these local scale hot spots and control points can be accounted for in the models—an area of rapidly growing research (e.g., Pinay et al. 2015; Abbott et al. 2016).

In order to produce the best possible predictive models of river network water chemistry, it might be useful to create combined surface watershed and groundwatershed models to represent all possible flowpaths to the stream. This more complex approach that utilizes information from both surface and groundwater topography observations so we would hypothesize that it would increase overall predictive power. That assessment of determining the best overall statistical model that uses all possible predictors for modeling Augusta Creek chemistry is not the goal of this initial assessment of surface versus groundwatershed. Instead we wanted to take a first important step toward examining whether surface watersheds and groundwatersheds produce similar or different outcomes when using a common modeling approach. This first step is part of putting forth the idea that groundwatersheds may be important to conceptualize in future ecological and water quality studies, and that not accounting for possible differences between surface and groundwatershed delineations may ultimately affect the way researchers interpret results based on areal LULC delineations. Still, to start to explore the potential for combining information from both surface and groundwatershed approaches and test our hypothesis that they would outperform the more parsimonious single surface or groundwatershed approach models, we conducted a combined analysis using the same linear regression modeling approach as described for Augusta Creek. Interestingly, in most cases, the combined surface watershed and groundwatershed models were comparable to the best-fit models from the presented models (Appendix A, Figure 26 and Figure 27). In the case of Cl^- , the combined models performed better than their surface and groundwatershed counterparts, while in the case of DOC they performed worse. In several instances the combined models produced were

exactly the same as either the surface watershed or groundwatershed models. These are unexpected results and, with exception of Cl^- , are in opposition to our hypothesis of better prediction with a combined watershed approach. Again, this is more evidence that the more biogeochemically reactive solutes of NO_3^- , DOC, and SUVA_{254} observed in the stream network are also controlled by local scale hot spots and control points that our watershed scale models do not capture.

Linear model performance for subsets of LULC and stream-order

We further explore if the different watershed delineations approaches provide further insight into how stream chemistry at a particular location in the stream network is or is not related to its upstream LULC or stream order. To do this we analyzed and modeled subsets of the stream network data that were partitioned by the predominant LULC and Strahler stream orders. While the subsets include a limited number of sites and therefore offer only some preliminary insights, it is seen that areas dominated by Hay/Pasture and Deciduous Forest areas yielded the best predictions of all measured biogeochemical properties in Augusta Creek. Conversely, both surface and groundwatersheds did poorly when predicting biogeochemical measurements in areas with a predominant LULC category of cultivated crops. While there is often a relationship between stream chemistry and cultivated crops in many watersheds (e.g., NO_3^- Schilling and Libra 2000), this relationship may not be so simple that a mixed regressive-spatial autoregressive model using watersheds areas can be used to make robust predictions in Augusta Creek. This inability of the model could be due to additional factors not captured in the LULC predictive framework, such highly variable practices of fertilizer application and tillage across cultivated lands of different farmers (Birgand et al. 2007; Thomas and Abbott 2018) or transformation (i.e., generation or removal) of solutes along groundwater flowpaths (Hill 1996; Ocampo et al. 2006).

Still, it is valuable for land and water managers to consider in Augusta Creek that there may be a lack of relationship between cultivated crops and the stream chemistry, and therefore a different approach beyond using spatial LULC predictors may be needed.

Across stream orders there was more variability in predictive power of the second- and third-order reaches as compared to the first-order reaches, this is consistent with the findings of Hedin et al. (1998) that narrow regions of the interface between terrestrial-aquatic environments can act as focal points for nutrient transformation. First-order reaches were generally well predicted for both surface watersheds and groundwatersheds, while second- and third-order watersheds did relatively poorly for all biogeochemical measurements except for Cl^- . This difference across stream orders is partially explained by the fact that first-order streams have substantially less area draining to them, which limits the variability of hydrologic flowpath lengths and biogeochemical source-sink processes contributing to the observed stream chemistry (Abbott et al. 2018). On the other hand, at larger watershed scales, such as the second- and third-order sites, the streams not only drain a larger area but also function as integration points of the biogeochemical composition from upstream subwatersheds, which can be a further complication in trying to understand reactive biogeochemistry source and sink processes as a function of LULC (Abbott et al. 2018). Overall, the Cl^- measurements are more easily predicted in second- and third-order streams, likely due to its more conservative nature along hydrologic flowpaths and through aquatic ecosystems when compared to the other more reactive N and C solutes in these larger order stream locations.

When we break up the data into subsets of particular land cover or stream order, it becomes less clear if using either using groundwatersheds or surface watersheds is advantageous. Again, as discussed above, this could be due to scale processes not captured in the LULC

approach are controlling the stream chemistry. Alternately, it could be due to the reduced statistical power of these smaller subsets. For example, it does appear that Cl^- is better predicted by groundwatersheds in second- and third-order sampling points. This may be attributed to groundwater acting as a long-term reservoir for human produced Cl^- (i.e., legacy effects; Kelly 2008, Kincaid and Findlay 2009). Also, given that Cl^- is a conservative (i.e., non-reactive) solute, it is more probable that any Cl^- signature derived from the landscape will be well preserved along the groundwater flow paths and through reactive control points, such groundwater-surface water interfaces (Abbott et al. 2016; Bernhardt et al., 2017), as opposed to the other biogeochemical measurements investigated here (NO_3^- , DOC, and DOC optical parameters) which are likely to undergo transformation in the groundwater and as they are transported through groundwater-surface water interfaces.

Limitations of this study

There are two primary limitations to be noted in this study. First, while the results are promising, in terms of using groundwatersheds to aid in future surface water biogeochemistry modeling and management, this study was only conducted in a single watershed during a single time of year. Future work should span several stream networks over multiple years and seasons to further assess the utility of using groundwatersheds versus surface watersheds. Second, it would be useful to have independent verification of the accuracy of groundwatershed delineations. Surface watersheds are relatively easy to delineate using high resolution digital elevation models, and the surface topography is relatively static through time scales of seasons and years. On the other hand, groundwatersheds vary with seasonal and annual hydrologic conditions and data collected for their delineation usually relies on aggregating multiple months and years of watershed measurement levels, including those used this study. Consequently, we

are not necessarily using the precise groundwatershed that was present at the time of biogeochemical sampling in Augusta Creek. However, the relatively uniform topography and similar soil characteristics across the study region make it reasonable to assume that even as climatic and seasonal conditions shift, groundwater divides are unlikely to substantially shift. These assumptions will not hold in regions with more topographic and subsurface hydrogeologic heterogeneity. Furthermore, the importance of vertical component of groundwater flowpaths varies across and within watersheds, which can affect groundwatershed area as well. Again, this is not a major factor for Augusta Creek, as the vertical flowpath component is nominal compared to the lateral components because the thickness of the aquifer is typically less than 50 m, compared to a typical interstream distance (and thus flow path length) of 1 km or more. These assumptions about surface and groundwater routing will not hold in regions with more topographic and subsurface hydrogeologic heterogeneity. Overall, a deeper investigation into the uncertainty and variability of groundwatershed delineations is important work for the future of surface water chemistry studies and models. For example, to fully verify the groundwatershed delineations from Augusta Creek used in this study would require substantial long-term work far beyond the scope of this initial investigation as both groundwater and biogeochemical synoptic sampling are limited in this watershed, as in most watersheds (McGuire et al. 2014; Ruhala et al. 2018; Abbott et al. 2018). Fundamentally, until more groundwater data is collected at greater spatiotemporal resolution, the limitation of available groundwater data is likely to remain a major impediment to advancing the use of groundwatershed approaches to predict and manage river water quality.

Model Performance vs. Stream Order and Dominant Land Use

After fitting the models on the entire dataset, we subset the sampling data and model predictions into two separate groupings: stream order and dominant land use to assess if the surface versus groundwater comparison is dependent on these spatial attributes. The Pseudo R^2 for each of these groupings are shown in Figure 7. Note that because these are subsets of the sites in the dataset, Pseudo R^2 values can exceed unity and offer limited ability to draw robust conclusions. Within each of the three stream orders present in Augusta Creek (first-, second-, and third-order) we typically found that the surface watershed models had greater Pseudo R^2 values in the first-order sampling points than for the second-order and third-order sampling points (Figure 16). Similarly, the groundwater models had Pseudo R^2 values closest to a value of 1 within the first-order sampling points. In general, surface watershed models outperformed groundwater models at first- and third-order scales, but groundwater models performed better for the intermediate spatial scale of the second-order streams.

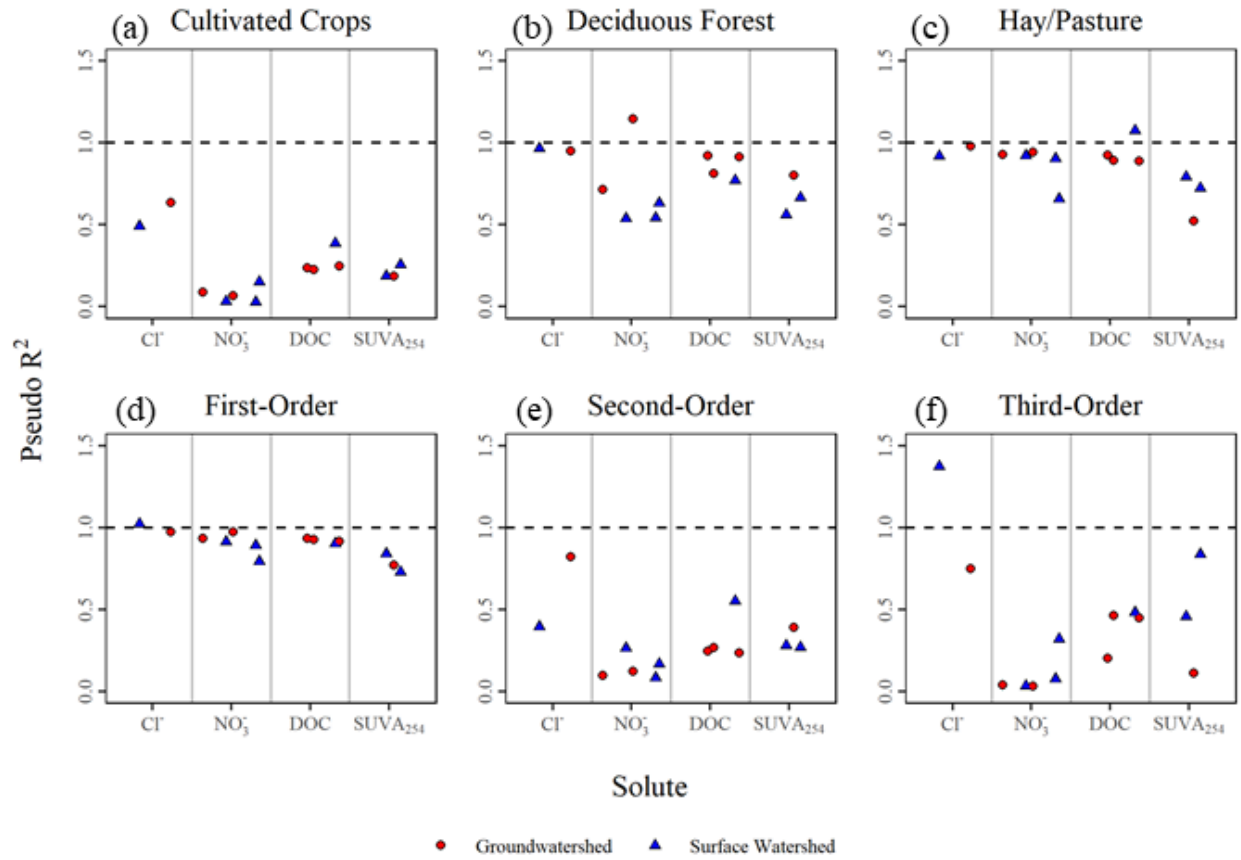


Figure 16 - A comparison of groundwater and surface watershed model performance as it relates to dominant land cover class and placement in the stream network is shown in the Pseudo R^2 values for sampling subsets related to predominant land use within an area, cultivated crops (a), deciduous forest (b), and hay/pasture (c); and stream order, first-order (d), second-order (e), and third order (f).

Dominant land use categories varied slightly for surface watersheds and groundwatersheds, but generally included cultivated crops, hay/pasture, and deciduous forest (Figure 16). We found that sites that predominantly drained hay/pasture or deciduous forest typically had Pseudo R^2 values closer to a value of 1 than sites predominantly draining cultivated crops. This pattern was true for both surface watershed and groundwater models.

Conclusion

In this study we investigated the relative predictive power of using LULC within groundwatersheds as compared to LULC within surface watersheds to predict the important water quality parameters of Cl^- , NO_3^- , DOC, and SUVA_{254} in the lowland Augusta Creek watershed in southwestern Michigan, United States. We provide clear evidence that groundwatersheds, created using data on groundwater levels from well logs collected over several decades, may be a useful tool in advancing the understanding and prediction of stream biogeochemistry. Specifically, in lowland watersheds, groundwatershed delineation may provide a better framework for relating land use to stream biogeochemistry than comparable surface watershed modeling approaches. In all instances of measured stream biogeochemistry at the network scale, groundwatersheds performed nearly as well or better than surface watersheds predictive models of stream biogeochemistry. Still, despite this study leveraging rarely available stream synoptic chemistry and having high spatial resolution of the groundwater topography, it is still clear that the availability of groundwater data limits rigorous assessments of the strengths and weaknesses of a groundwatershed versus surface watershed approach to modeling chemistry in Augusta Creek. Therefore, in streams where even less high spatial resolution water quality data is available, it will be difficult to assess the utility of surface versus groundwatershed approaches to managing water quality. Looking forward, this study suggests that hydrologists, biogeochemists, ecologists, and watershed managers should work closely together to explore collecting more groundwater data and to use groundwatersheds as a framework for predicting and protecting stream water quality conditions, much in the same way that these collaborations have advanced watershed sciences in the past with research on riparian buffer zones (Floyd et al. 2009), hyporheic zones (Boano et al. 2014), and stream restoration (Brazner et al. 2007).

Acknowledgments

This work was partially supported by the National Science Foundation [EAR-1446328, EAR-1846855, and DEB-1637653], Michigan State University Department of Earth and Environmental Sciences, Michigan State University Environmental Science and Policy Program, and Michigan State University Kellogg Biological Station Long Term Ecological Research site. Thank you to S. Hamilton for aiding with sampling site selection. Thank you to A. Shogren for reviewing an early draft. Thank you to S. Ruhala, S. Plont, T. Bigg, E. Wiewiora, and T. Hampton for assistance collecting samples.

CHAPTER 3: EVIDENCE FOR STREAM WATER-GROUNDWATER INTERFACE MODIFICATION OF LANDSCAPE BIOGEOCHEMICAL SIGNATURES

Abstract

There is a need to understand the stream water-groundwater interface at stream network scales relevant to the needs of land managers and modelers. Currently, research within this interface and at this scale is typically conceptual or strictly a numerical modeling effort, with most field studies focused on a local or feature scale. Here, we focus on assessing the reactivity of the stream water-groundwater interface at the network-scale across the Augusta Creek watershed in southwestern Michigan, USA. We use spatially-lagged linear models to assess first, if there is a predictable pattern of biogeochemistry in stream water that is related to the land cover, and second, whether this predictable pattern, or signal, propagates into and is preserved within the stream water-groundwater interface. We test this for several solutes including chloride, nitrate, and dissolved organic carbon. Our results indicate that the landscape biogeochemical signature found in surface waters was not preserved as it traveled through the shallow subsurface, where stream water and groundwater mix. We provide evidence that at this aggregated scale, surface water biogeochemical signals were changed as a result of reactions at the stream water-groundwater interface.

Introduction

The stream-groundwater interface (SGI) is an important ecotone with high rates of biogeochemical transformations (e.g., Findlay and Sobczak 1996; Lautz and Fanelli 2008), and often acts as a biogeochemical processing control point for water passing through stream networks (e.g., McClain et al. 2003; Bernhardt et al. 2017). Given the importance of the SGI for watershed biogeochemical transport and transformations, there is a need for both land managers

and researchers to understand this interface at scales exceeding what we have historically measured (Krause et al. 2011). Our fundamental understanding of the SGI is primarily based on studies of channel lengths ranging from less than 1 m to as much as approximately 1000 m, with little consideration of how processes within this interface affect biogeochemical and ecological patterns and processes across a river network (Ward 2016). Those studies that have investigated the SGI at the river network-scale are generally limited to models of reduced complexity focused on the river corridors and not the entire watershed (e.g., Kiel and Cardenas 2014; Gomez-Velez and Harvey 2014). Though the modeling community has made tremendous progress in assessing the physical processes occurring within the SGI at network scales, they have been limited by an absence of empirical biogeochemical data to inform the function with respect to biogeochemical processing.

The landscape within a catchment influences the signature of the surface water biogeochemistry (e.g., Herlihy et al. 1998; Inwood et al. 2005). Observed patterns of stream biogeochemistry through a river network cannot be explained only by the surface water perspectives of connectivity of streams in the network and the unidirectional movement of water. In fact, many stream ecosystems are controlled by varying degrees of groundwater upwelling, though there exists a knowledge gap in understanding exactly what type of responses freshwater ecosystems have to groundwater inputs (Boulton and Hancock 2006). To understand river network biogeochemistry we need to understand the role of the SGI is a control point across stream networks (McClain et al. 2003; Bernhardt et al. 2017) that affects processes such as retention, biodegradation, or sorption of materials in surface water or in groundwater emerging into streams (Edmonds and Grimm 2011). While statistical models have had some success in predicting stream surface water chemistry based on land use and cover in the surrounding

landscape (e.g., Mattsson et al. 2005; Dodds and Oakes 2006), the role of reactive control points like the SGI are generally not accounted for in the models.

To better understand how the SGI affects water quality at the watershed scale, this study utilized a recently generated data set of SGI biogeochemical data collected at the network scale (Ruhala et al. 2018) to assess whether the SGI has a consistent function as a control point across the entire stream network as typically applied in models, or if their function is inconsistent. We hypothesize that the biogeochemical carbon and nutrient signature of the SGI will be independent of the overlying surface water and thus not predictable from watershed land use and cover upstream of the sampling point. Further, we examine this hypothesis at the relatively understudied network-scale with a focus on understanding how the landscape signal propagates into and through the SGI. The Ruhala et al. (2018) study showed that the upper 20cm of the SGI sediments is the approximate region where most of the surface and groundwater exchange occurs across the lowland Augusta Creek stream network, and therefore for our study we define the SGI as the upper 20cm of sediments on the stream bed. Our biogeochemical species of interest are dissolved organic carbon (DOC) along with its spectral properties—specific ultraviolet absorbance at 254 nm ($SUVA_{254}$, Weishaar et al. 2003) and spectral slope ratio (S_R , Helms et al. 2008), as well as nitrate (NO_3^-), and chloride (Cl^-). We selected DOC because it is a fundamental control on water quality due to its role as an electron donor in nutrient cycling and control of photochemistry, among other properties (Aiken 2014). Both $SUVA_{254}$ and S_R are commonly used to investigate stream ecosystems and their relationship to land use and land cover (e.g., Yates et al. 2016). We included Cl^- and NO_3^- as conservative and reactive solutes, respectively, commonly studied in freshwater systems (e.g., Triska et al. 1993; Barber et al. 2005; Zarnetske et al. 2011), and Cl^- can help assess solute transport direction given its nonreactive properties. We

test our hypothesis by using mixed spatial lag models to assess landscape-biogeochemistry correlations in surface water, then test these models at different depths beneath the SGI to determine if they maintain their predictive power at depth.

Materials and Methods

Site Description

The water chemistry data used in this study were synoptically collected across the stream network of the Augusta Creek watershed (Figure 17), a third-order, low-gradient watershed covering 98 km² in southwestern Michigan, United States. Augusta Creek drains a mixed-use landscape composed of wetlands, lakes, row crops, pasture, and upland forests. The stream network channel is primarily composed of eroded glacial till sediments, and surface water flows are predominantly groundwater-fed with a dampened discharge response to storm events due to its abundant wetlands and minimal overland flow (Poff et al. 1997; Hamilton et al. 2018). This study includes sampling sites from across all three orders originating from a variety of hydrologic settings including lake outflows, upland forested headwaters, and wetland outflows. Augusta Creek flows through the Kellogg Biological Station Long Term Ecological Research program and has an active USGS gaging station (04105700).

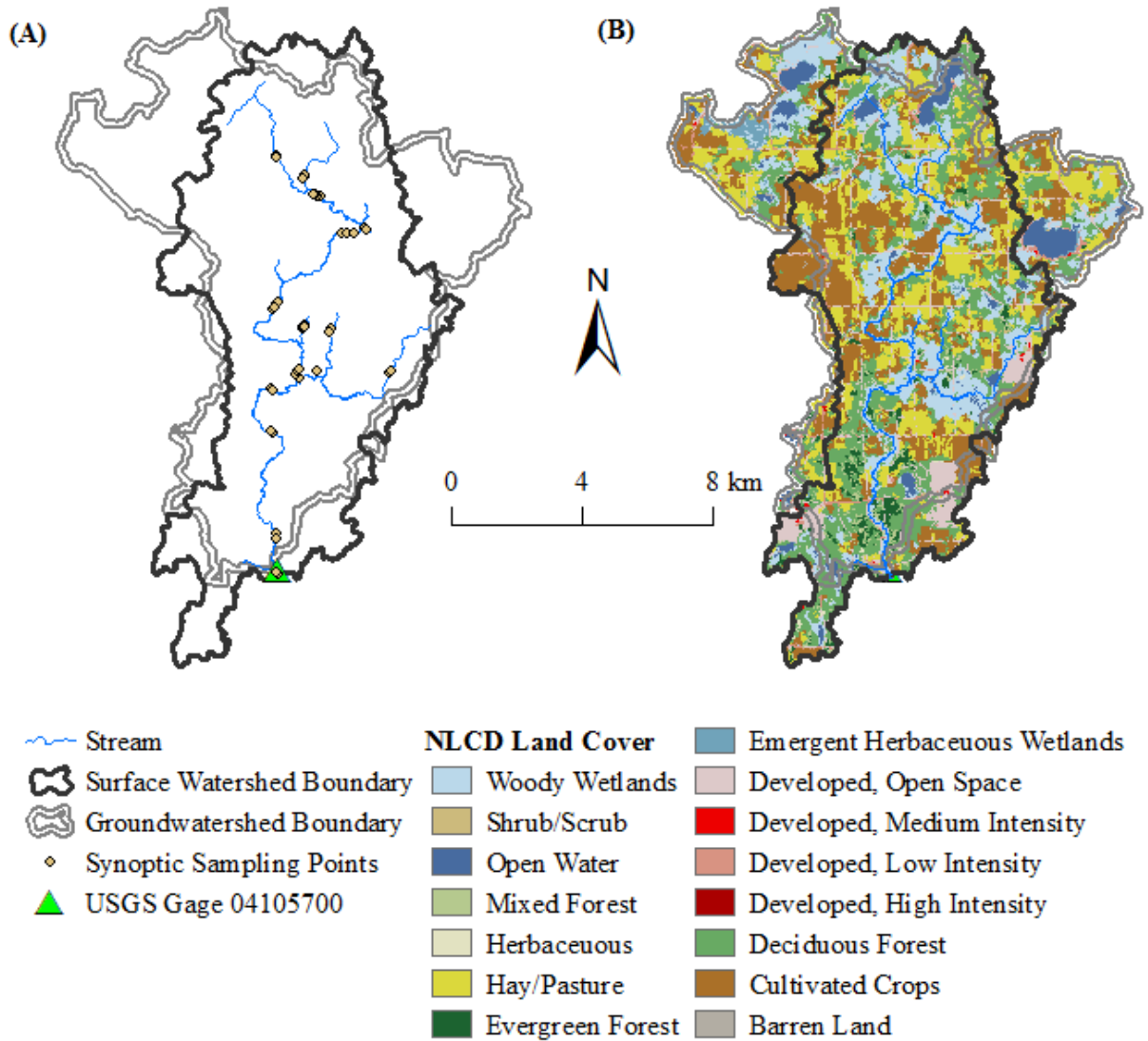


Figure 17 - Map of Augusta Creek including (a) with the watershed boundaries in which the double gray line is the groundwatershed and the solid black line is the surface (topographic) watershed, the blue line is the stream network, yellow points represent the sampling points, and the green triangle represents the USGS Gaging Station (04105700); (b) displays the National Land Cover Database 2011 with categories found in the Augusta Creek watershed indicated in the legend.

Synoptic Sampling Campaign

The collection of stream water and SGI samples at 39 sites distributed across first-, second-, and third-order stream reaches was detailed in Ruhala et al. (2018) and is only briefly reviewed here. The collection locations were stratified to reflect the relative stream length of

each stream order with 16 first-order sites, 14 second-order sites, and 9 third-order sites. At each site sample collection optimally included a surface water sample, 6 SGI samples at discrete depths beneath the sediment surface (2.5, 5, 7.5, 10.0, 15.0, and 20 cm) for a total of 307 samples (38 surface water and 269 SGI). All samples were collected in the shortest period of time under similar weather conditions to the extent possible. In total, synoptic sampling took 7 days to complete between 16 and 22 August 2016.

Samples at the SGI were collected using the relatively non-invasive MINIPPOINT porewater piezometers (Duff et al. 1998; Harvey and Fuller 1998), ideal for collecting SGI samples in unconsolidated sandy and gravelly sediments. Each individual MINIPPOINT porewater piezometer was attached to a Masterflex peristaltic pump (Cole-Parmer) using L/S Tygon tubing, with water drawn from the sediments at a rate of 2.5 mL min^{-1} to ensure minimal disruption of subsurface flowpaths (Duff et al. 1998). For each sample, 80 mL of water was collected. From this 80 mL, 20 mL of water was filtered through a series of filters as a rinse, first through a coarser filter to remove particulates and larger microbes (Whatman GF/F 0.7 μm nominal pore size) and then through a finer filter to remove remaining microbes, leaving only the dissolved components (Sartorius Stedim cellulose acetate 0.2 μm nominal pore size). The remaining 60 mL was filtered into acid-rinsed HDPE amber bottles and stored on ice until the end of the sampling day; thereafter it was stored in the dark at 4°C and analyzed within 28 d.

Each filtered sample was analyzed for DOC as non-purgeable organic carbon using a TOC-L total organic carbon analyzer (Shimadzu Scientific Instruments, Kyoto, Japan) with Pt catalyzed oxidation at 680°C . Concentrations of Cl^{-} and NO_3^{-} were measured using a Dionex ICS-2100 Ion Chromatography System (Thermo Fisher Scientific, Massachusetts, USA). Optically-derived DOC properties were determined through absorbance measurements on a

Shimadzu dual-beam ultra-violet 1800 spectrophotometer (Shimadzu Scientific Instruments, Kyoto, Japan). The optical properties used in this study are SUVA₂₅₄ and the S_R. SUVA₂₅₄ is defined as the absorbance at $\lambda = 254$ nm divided by the DOC concentration, and S_R is the ratio of the absorbance slope from 275 to 295 nm to the absorbance slope from 350 to 400 nm.

Watershed Delineation & Land Use

To determine the area from which drainage to the surface water occurred in this lowland stream network, we delineated a surface watershed area (i.e., the area of surface water drainage based on surface topography) and a groundwatershed (i.e., the area of surface water drainage based on groundwater levels) to each sampling site. Previous studies in this watershed and similar lowland watersheds have shown that the groundwater and surface watershed areas can be different (Lee-Cullin et al. *Chapter 2*, Boutt et al. 2001). We created surface watersheds and groundwatersheds for all 39 sites as detailed in Lee-Cullin (*Chapter 2*; see Figure 18 for an example of the delineation of a surface watershed as compared to a groundwatershed). Briefly summarized, surface watersheds were created using the traditional approach based upon surface topography while groundwatersheds were created using freely available databases on static water levels in mostly residential water-supply wells to create groundwater flow and capture zones draining to a sample site. For the groundwatershed, we assumed two-dimensional flow and small vertical gradients in the hydraulic head. We then interpolated water table elevations using static water levels. With this we used the potentiometric surface of the water table (i.e., the water table topography) to delineate groundwatersheds.

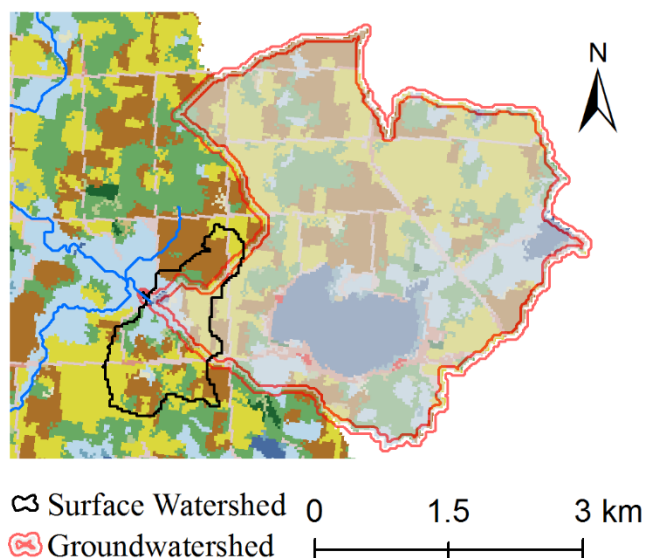


Figure 18 - Comparison of surface watershed (black outline) and groundwatershed (red double outline) for a particular sampling point on a tributary of Augusta Creek, showing a case of particularly large disagreement between the two watershed areas. The image displays the different land use categories (various colors defined in Figure 17) that are included depending on which watershed was selected.

The Augusta Creek land use data were obtained using the National Land Cover Database 2011 (NLCD 2011; U.S. Geological Survey 2014; Homer et al. 2015). The NLCD 2011 is divided into 8 categories each with sub-categories: Water, Developed, Barren, Forest, Shrubland, Herbaceous, Planted/Cultivated, and Wetlands. For the purpose of this study, given the predominantly rural and undeveloped characteristics of the Augusta Creek watershed, we combined all Developed sub-categories (i.e., Open Space, Low Intensity, Medium Intensity, High Intensity) into a singular Developed category. In total, this developed area represents 2.5% of the largest, farthest downstream groundwatershed areal extent and 7.1% of the largest downstream surface watershed areal extent. All other land cover categories present were left in their original classification categories. We determined the relative proportion of all land use categories for each of the subwatersheds described above by taking the area of a land use type within a watershed divided by the total area of that watershed.

Linear Models to Assess Landscape Controls on Stream and SGI Chemistry

We use linear models parameterized for surface water chemistry based on land use type from the contributing surface and groundwatershed to see if the SGI chemistry can be predicted by propagating surface water models to discrete SGI depths. If the linear model performance degrades in the SGI for a solute, it indicates additional processes unique to the SGI are altering the solute chemistry and hence is evidence in support of our hypothesis. If the linear model performance remains comparable to the surface water model for a solute, it indicates that the SGI is not altering the solute chemistry and refutes our hypothesis. And we predicted that linear models for reactive DOC and NO_3^- characteristic will perform poorly in the SGI, while linear models for the more conservative Cl^- will perform comparable in the SGI.

An analysis of the ability of the groundwatershed versus surface watershed modeling approach for predicting surface water chemistry in Augusta Creek using linear models is detailed in Lee-Cullin (*Chapter 2*). We briefly review this to provide context for assessing how water delivered to the SGI via the surface watershed as compared to the groundwatershed controls the chemistry seen in the SGI. A linear modeling approach was used to assess how watershed inputs to the stream network may be modified by SGI interactions. This method was selected because it is relatively parsimonious and compatible with the available data.

The linear modeling approach starts with constructing spatial lag models to account for spatial autocorrelation while attempting to model stream biogeochemistry based on landscape attributes. All modeling was carried out using R (R Core Team 2017). We first constructed multiple linear regression models using the generalized ordinary least squares (OLS) equation:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \quad (1)$$

where y is the biogeochemical parameter we wish to predict (i.e., DOC, SUVA₂₅₄, S_R, Cl⁻, or NO₃⁻), x is the landscape unit proportion (note that we created both surface watershed models and groundwatershed models, each with distinct landscape proportions), β is the slope as it relates to the landscape unit proportion, and ε is the random error. Here, we used surface water data to calibrate the linear models. We started with all landscape variables in the model and removed independent variables in a stepwise fashion using the Variance Inflation Factor (VIF) to reduce multicollinearity in the models. We set a VIF threshold of 3 (Zuur et al. 2010) for all independent variables and each model that did not meet this criterion had the two variables with the highest VIF removed to create two new models. This created a suite of models. Once all independent variables had a VIF of less than 3 we used stepwise removal of variables using a significance threshold of 0.05 to reduce model complexity (after de Koning et al. 1998).

Due to spatial autocorrelation in the model residuals, as indicated using Moran's I test on the OLS models, we converted all OLS models into mixed regressive-spatial autoregressive models, thereby accounting for spatial autocorrelation (see Anselin 1988; Overmars et al. 2003). The mixed spatial models were created using the following formulation:

$$Y = \rho WY + X\beta + WX\gamma + e \quad (2)$$

where ρ is fit first and is the autoregression coefficient, W is a spatial weights matrix developed using nearest neighbors, X is the predictor matrix from Equation(1) β is a vector representing slopes of the independent variables from the predictor matrix, γ is a fit parameter representing the autoregression coefficient of the spatially lagged explanatory values, and e represents the spatially independent errors.

In our final step of modeling, we used Akaike's Information Criteria (AIC) to choose the most parsimonious models within each biogeochemical solute category and for each type of

watershed (i.e., surface watershed or groundwater watershed). The AIC determined the most parsimonious model as the one with the lowest AIC value. We also selected models within an AIC value of 2 relative to the models with the lowest AIC (i.e., $\Delta AIC < 2$), a threshold commonly considered to indicate an equally parsimonious model (Burnham and Anderson 2002). Therefore, the lowest AIC model for each watershed-biogeochemical pair and all models within a ΔAIC value of 2 were selected as “best-fit” models. Based on AIC, groundwatersheds and surface watersheds performed comparably for DOC and $SUVA_{254}$, groundwatersheds were considered better for predicting Cl^- and NO_3^- , and surface watersheds performed better predicting S_R across the Augusta Creek watershed (*Chapter 2*; Figure 19).

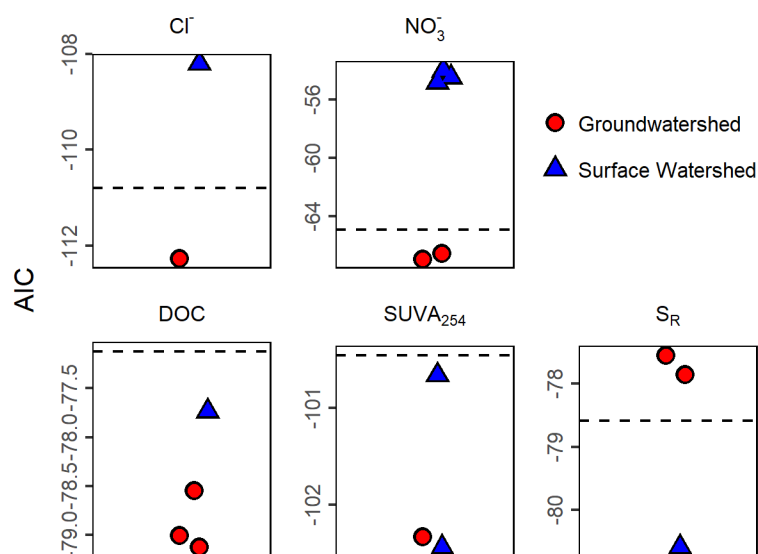


Figure 19 - Comparison of AIC values for surface water spatial lag models, modified from Lee-Cullin et al. (*Chapter 2*). Red circles indicate groundwatersheds were used for drainage area delineation while blue triangles indicate that surface watersheds were used for drainage area delineation. The dashed line indicates the threshold at which models below the line are considered ‘best fit’ models. In AIC, lower values mean a more parsimonious model.

The identified “best-fit” models for prediction of surface water chemistry from land cover were then used to parameterize models for each subsequent discrete SGI depth from 2.5 to 20 cm and the GW at 60 cm. Specifically, we used each best-fit model’s (for each surface watershed and groundwater delineations) chosen landscape characteristics and created new linear

models at every given depth. In creating these linear models, we simply kept the landscape proportions and allowed the slopes and random error to change to develop the best model possible. This led to the creation of a suite of models, all based off the “best-fit” surface water models, at each depth across the sediment water interface (2.5, 5.0, 7.5, 10.0, 15.0, 20.0, and 60.0 cm). Each model’s performance was then assessed by its pseudo R^2 value. We used a pseudo R^2 because the traditional coefficient of determination R^2 is incompatible with spatial lag models. Pseudo R^2 accounts for the fact that one cannot give each measure equal weight in a measured fit in spatial lag models. The pseudo R^2 is given as the variance of predicted values divided by variance of the observed values (i.e., $\frac{\text{Variance}_{\text{predicted}}}{\text{Variance}_{\text{observed}}}$; Anselin 1988; Overmars et al. 2003). Note that pseudo R^2 used here matches R^2 when used for OLS models.

Results

Stream Network Biogeochemistry

Results of biogeochemistry of the surface water and the SGI across the entire watershed are summarized in Figure 20. In general, DOC decreased moving from the surface water down through 20 cm in the SGI. The optical properties of DOC, SUVA_{254} and S_R , both had greater variability at depth as compared to the surface water. The Cl^- mean and median concentrations decreased slightly in the SGI but more notably increased in variability as compared to surface water. Similarly, NO_3^- had greater variability at depth than in the surface water and had a generally increasing trend through depth. Both NO_3^- and Cl^- showed occasional very high concentrations in the SGI.

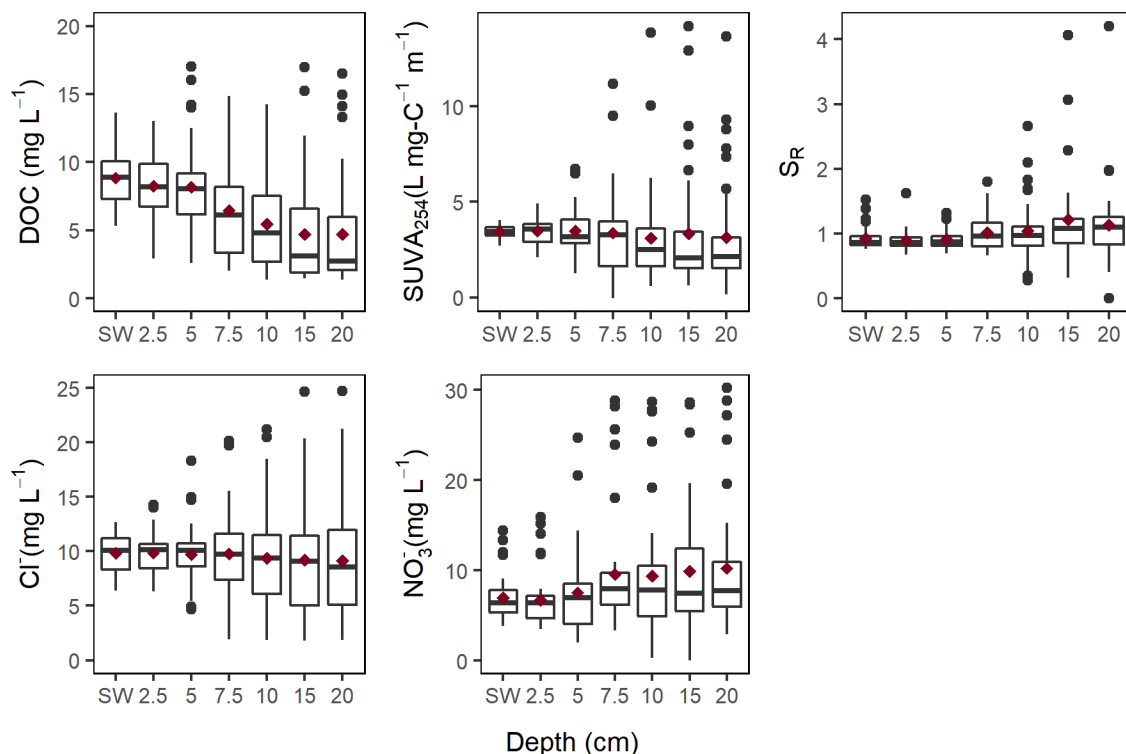


Figure 20 - Biogeochemical results of surface water and at various depths beneath the sediment surface in the SGI across Augusta Creek. Measurements include DOC, SUVA₂₅₄, S_R, Cl⁻, and NO₃⁻. Box and whiskers represent interquartile range, red diamonds represent mean values, and black circles represent outliers.

Surface versus Groundwater Watersheds

Both surface watersheds and groundwatersheds varied in area across Augusta Creek, depending on the locations of the water sampling points within the stream network. For both types of watersheds, first-order watersheds had areas as small as < 1km² while the third-order watersheds reached areas big as approximately the same size as the entire surface watershed area (~100 km²). The surface watersheds were typically larger than the groundwatersheds for the first-order sampling sites, while the opposite was true in third-order watersheds because groundwatersheds extended beyond topographic boundaries especially in the lowlands of Augusta Creek. Although different across the stream network, surface watershed and groundwatersheds areas for each of the 39 sampling points were strongly correlated ($R^2 = 0.98$, $p < 0.05$). Each watershed type had different predominant landcovers at the largest areal extent—

surface watersheds were predominantly Deciduous Forest, groundwatersheds were predominantly Pasture/Hay.

Linear Models

For all biogeochemical measurements except $SUVA_{254}$, at each depth across the SGI, our best-fit linear models performed worse than they did for surface water (Figure 21). The exception, $SUVA_{254}$, is where the surface watershed models performed better from 7.5 to 20 cm depth than in surface water. For DOC and NO_3^- , both types of models performed relatively poorly at intermediate depths before improved performance at the greatest depth of 20 cm. Both S_R and Cl^- performed poorly relative to surface water at depth, with S_R exhibiting an increase in model performance from 5 to 7.5 cm.

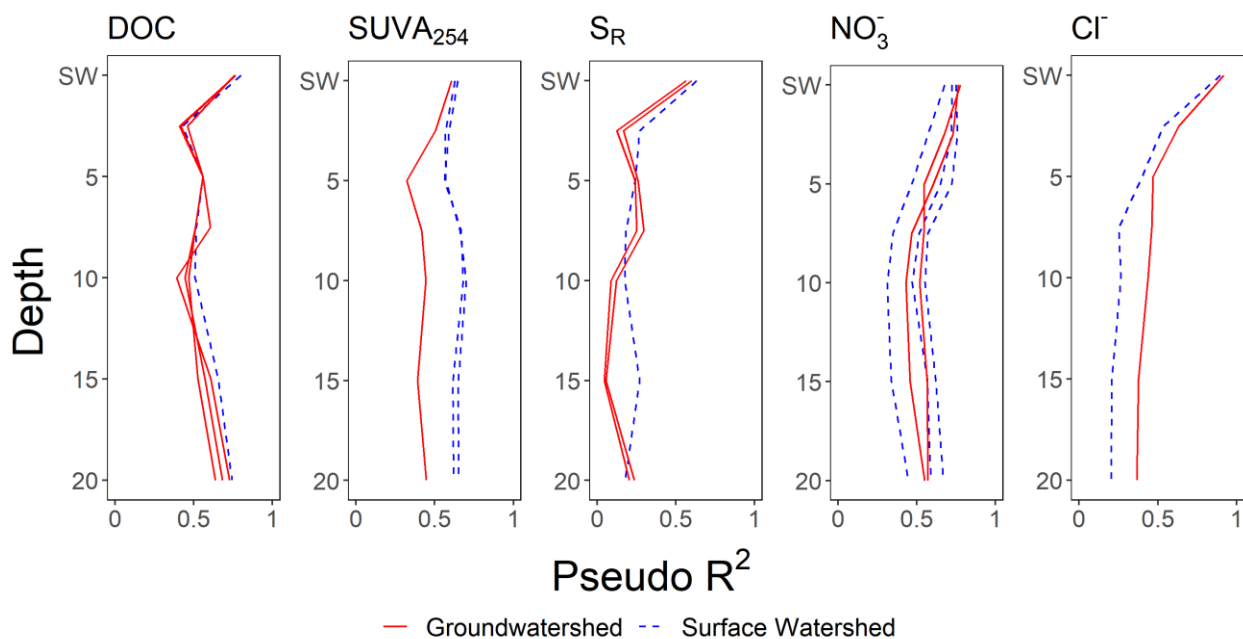


Figure 21 - Pseudo R^2 results for spatial lag models across all depths for each biogeochemical measurement across all depths within the Augusta Creek stream network. Solid red lines represent groundwatershed models while dashed blue lines represent surface watershed models. Pseudo R^2 values close to one indicate a higher proportion of the overall variance of modeled values representing the overall variance of observed values.

Discussion

The decreasing performance of spatial lag models from the surface water into the SGI for DOC, S_R , NO_3^- , and Cl^- across the stream network indicates that the landscape signal evident in the surface water chemistry across the stream network is not fully preserved in the SGI. This supports our hypothesis that the SGI generally functions as an ecotone that modifies the landscape biogeochemical signals across the entire stream network. This study also adds further support that the SGI functions as a biogeochemical control point across the landscape (Bernhardt et al. 2017) and builds upon the large number of studies finding the SGI as a highly reactive ecotone.

We find that both DOC and NO_3^- in the surface water samples are well predicted by either groundwatershed or surface watershed models. However, the ability of landscape attributes to predict DOC and NO_3^- generally decreased when trying to model conditions in the SGI. These DOC and NO_3^- model results suggest that both the surface water and deeper SGI samples are by the upstream landscape, but as these solutes move along flowpaths through the 20 cm of observed SGI, the signals are attenuated or lost. While understudied for lowland watershed SGIs, a number of studies from upland SGIs have found that concentrations of DOC and NO_3^- , which are typically abundant solutes in surface and groundwaters, decrease through the SGI as compared to the surface water (e.g., Pinay et al. 1994; Schindler and Krabbenhoft 1998; Sobczak and Findlay 2002). Further, Zarnetske et al. (2011a; b) found that labile fraction of the DOC, critical to NO_3^- processing within the SGI (e.g., fueling denitrification), decreased more rapidly along SGI flowpaths than did total DOC concentration. Hence, it is probable that the similar patterns of predictability in our models between DOC and NO_3^- are due to their coupled biogeochemical relationship in SGIs.

It is notable that, despite the similarities in performance between the DOC and NO_3^- predictive models, the NO_3^- model does not have as large of an increase in its pseudo R^2 value at depth in the SGI. This distinction between models at depth can be attributed to the process by which NO_3^- is utilized in SGIs. For example, removal of NO_3^- via denitrification occurs primarily in SGI zones absent of oxygen. Conversely, DOC processing in the SGI occurs both in the presence of oxygen (e.g., via aerobic respiration) as well as in the presence of NO_3^- (e.g., via anaerobic respiration). Therefore, the DOC is likely to be processed throughout the entire SGI, while NO_3^- may still be at an intermediate stage of its processing at depth because its removal is dependent upon DOC and redox conditions.

The S_R models performed poorly overall, even in surface water, with a marked decrease in performance when moving into the SGI. The S_R is inversely correlated to molecular weight of DOC (Helms et al. 2008). In some cases, higher S_R values have been linked to protein-like, generally fresher DOC as compared to terrestrially derived, humic-like DOC, often associated with baseflow (Fasching et al. 2015b) or upstream locations (Helms et al. 2008). Further, S_R is associated with clearer waters (Helms et al. 2008; O'Donnell et al. 2012), consistent with degraded DOC in freshwater (Mann et al. 2012). Across Augusta Creek, S_R generally illustrated that there was a decrease in molecular weight and greater variability with depth in the SGI (Figure 21). Our inability to accurately predict this optical DOC property at depth indicates that the landscape fingerprint is modified in the SGI. This modification diminishes our ability to predict DOC composition in the SGI and, together with the greater variance and generally decreasing S_R in the interface across the stream network, illustrates the reactivity of the SGI, especially with respect to the biotic (microbial degradation) and abiotic (sorption) processes. It is well established that DOC may undergo several biotic (e.g., Amon and Benner 1996) and abiotic

(e.g., Nelson et al. 1993; Cory et al. 2014) processes during its transportation through stream networks. It appears that these transformation processes are consistently occurring across the entire Augusta Creek network within the SGI.

Surface watershed and groundwatershed models for SUVA₂₅₄ had distinct results from each other, with the surface watershed models substantially outperforming the groundwatershed models. This is troubling as SUVA₂₅₄ is one of the most frequently and readily measured attributes of DOC in aquatic ecosystem studies. SUVA₂₅₄ has been useful in that it is well-correlated to the aromaticity of DOC (Weishaar et al. 2003). Both surface and groundwatershed models performed similarly for the surface water models. However, the groundwatershed performance decreased moving down to 5 cm depth and the pseudo-R² remained at lower values through the SGI, while the surface watershed stayed relatively consistent moving from surface water through the SGI. This unique result has implications for selection of a surface watershed or a groundwatershed delineation. Selection of one watershed delineation over the other produces different results across the SGI watershed and our interpretation thereof—selection of the surface watershed would lead to an interpretation that the landscape signature observed in surface water can be seen through the entire SGI, while the groundwatershed would lead to an interpretation that this signature is removed through the SGI.

Our vertical profiles through the SGI showed decreasing median values of SUVA₂₅₄ moving from the surface water through the SGI, with outlier points that are substantially higher at each depth from 7.5 down to 20 cm. This general pattern of a decrease through the SGI can be interpreted as decreasing aromaticity in the DOC pool at depth. This stands in contrast to other studies comparing surface DOC to comparable (15-35 cm) and deeper (1.5 m) sediment porewater DOC (e.g., Lambert et al. 2013; Fasching et al. 2015a). There is evidence of DOC

processing in the SGI for this aromatic pool of DOC. However, given that we cannot determine whether surface watershed or groundwatershed models are superior, it is difficult to draw broader conclusions about the propagation of the landscape fingerprint of SUVA_{254} into the SGI.

The conservative solute Cl^- had the highest pseudo R^2 values in surface water, in contrast to low pseudo R^2 values deeper in the SGI. It is unlikely that this disparity can be attributed to reactivity, given the conservative nature of Cl^- . More likely in this instance is that the surface water Cl^- expression is closely linked to the landscape while groundwater flowpaths and mixing in the SGI decouple this relationship locally. The difference between the diminished pseudo R^2 for Cl^- as compared to the reactive solutes can be attributed to legacy Cl^- concentrations building up along the groundwater flow paths due to de-icing agents used in watersheds (Kaushal et al. 2018). We assert that poorer performance of SGI models can be attributed to factors we did not include in the models. This includes factors such as hydrodynamic mixing between groundwater and surface water Cl^- and local heterogeneities. The effective result from the processes listed above is a much greater variability of Cl^- concentrations in the SGI (Figure 20).

In this study, we focused on the biogeochemistry of the SGI across an entire stream network, and as a result less emphasis went into characterizing the heterogeneity at each individual sampling site. This is in line with the recommendations of (Lee-Cullin et al. 2018) for how best to characterize longitudinal heterogeneity of the SGI across the the Augusta Creek stream network. However, it must be noted that local heterogeneity in space and time can be highly variable in the SGI. For example, intrameander flow paths may cause a shift from upwelling to downwelling in a given bend of a river (Nelson et al. 2019), and variability in stream stage can change exchange rates through the SGI (Boano et al. 2014). However, small changes in stage are observed in Augusta Creek across variable flow conditions (Poff et al. 1997;

Ruhala et al. 2018) and in many cases the sediment heterogeneity is relatively small across the watershed as it consists of mostly well sorted and high permeability sands and gravel with sandy matrix (Ruhala et al. 2018). Tonina et al. (2016) showed that when there are low levels of sediment heterogeneity in the SGI, the hydraulics of the SGI can be assumed to be homogeneous when working at scales larger than a single streambed morphologic feature.

Finally, the timescale of groundwater inputs into the stream may be well beyond the scope and means of the data and modeling approach of this study. Land changes over the timeframe of years and this is not congruent with the timescales of groundwater flow into recently glaciated landscapes which may be on the order of many decades (e.g., Stewart et al. 2010; Hamilton 2012). This long residence time of solutes in the groundwater allows for 1) many biogeochemical processes such as sorption and uptake to occur and 2) the overlaying landscape to be modified. The latter means that the modeling approaches used in this study for relating the stream and SGI biogeochemistry to LULC may be temporally decoupled from landscape changes and therefore models that are better able to account for these temporal legacies in groundwater is an opportunity for future research.

Conclusion

The reactivity of the SGI is widely documented and often included in conceptual biogeochemical models (e.g., McClain et al. 2003; Bernhardt et al. 2017), but this study is one of the first to provide a framework that quantifies patterns across a river network using SGI measurement. It provides clear evidence that the SGI is unique from surface and ground water conditions and that landscape signals are modified as they transit through the SGI. This study of the propagation of the landscape biogeochemical signals imparted on stream biogeochemistry into the SGI indicates that 1) stream surface water biogeochemistry can be predicted using both

surface watersheds and groundwatersheds and their landscape cover in a lowland stream, 2) this predictable network-wide signal associated with landscape cover diminishes as stream water and groundwater mix in the SGI, and 3) selection of areal delineation of watersheds draining to a point in the stream network should be done carefully according to this study, because, for some solutes, the groundwatershed delineations approach outperforms the more traditionally used surface watershed delineation approach. This research builds upon and provides quantitative evidence in support of the SGIs operating as biogeochemical control points altering landscape biogeochemical conditions (McClain et al. 2003; Bernhardt et al. 2017) and provides one of the few efforts to date to expands the focus of SGI field assessments from point and reach scales to network scales as called for by Ward (2016). Consequently, this study informs water quality and ecological efforts attempting larger-scale watershed modeling and watershed management practices.

Acknowledgments

This work was partially supported by the National Science Foundation [EAR-1446328, EAR-1846855, and DEB-1637653], Michigan State University Department of Earth and Environmental Sciences, Michigan State University Environmental Science and Policy Program, and Michigan State University Kellogg Biological Station's Long Term Ecological Research site [DEB-1637653]. Thank you to S. Hamilton for helpful comments in aiding with sampling site selection. Thank you to Anthony Kendall for providing the script to help delineate groundwatersheds. Thank you to S. Ruhala, S. Plont, T. Bigg, E. Wiewiora, and T. Hampton for assistance collecting samples.

CHAPTER 4: EVALUATING THE SURFACE WATER-GROUNDWATER INTERFACE FUNCTION IN TRANSFORMING AND DEGRADING OF DISSOLVED ORGANIC CARBON ACROSS SIX MIDWESTERN STREAMS.

Abstract

Complex dissolved organic carbon (DOC) compounds are a key control of the ecology and water quality of freshwater systems. There is much uncertainty about the fate and transport of dissolved organic carbon, particularly through highly reactive stream-groundwater interfaces (SGIs). This study assesses the molecular composition and transformation of specific molecular DOC properties in surface waters and the underlying sediments of SGIs. Specifically, we attempt to understand what, if any, consistent DOC processing occurs within the SGI across a variety of stream systems. To guide this assessment, we hypothesize that distinct DOC sources, with unique molecular properties and composition, will decay within the SGI and the remaining DOC will approach similar composition signatures, regardless of DOC origin or the specific SW-GWSGI the DOC passes through. We assess this hypothesis by using controlled laboratory bioassay and DOC molecular characterization methods to document changes in DOC molecular properties when different natural DOC sources are exposed to the surface waters and SGI sediments collected from 6 unique stream sites. Using a mixed-effects linear model, our results generally support the hypothesis but indicate that there is very little consistency in transformations of most components of the DOC due to interactions with stream sediments. Most dissolved organic carbon components decompose based on more complicated interactions between sources of carbon and the specific surface water components.

Introduction

Organic carbon is a key control on water quality and the ecology of stream networks, controlling nutrient and contaminant cycles, food webs, and drinking water quality and treatment (Aiken 2014). In streams, organic carbon is predominantly in the form of dissolved organic

carbon (DOC, Cole et al. 2007; Battin et al. 2008). The composition of DOC is complex and highly variable because it depends, in part, on the origin of the carbon in the landscape (e.g., high humic content from soils as compared to wetland vegetation; Hansen et al. 2016). Ultimately, DOC composition affects its fate and reactivity in the environment (e.g., DOC participation in biogeochemical reactions such as denitrification). Much DOC research has relied on measurements of the total concentration without information of its composition or properties. Emerging techniques and technologies to infer the molecular properties of DOC now allow investigation of the complexity of DOC, enabling more advanced understanding of DOC reactivity and function.

The fate of instream DOC is largely driven by metabolic activity at the interface of streams and groundwaters (Grimm and Fisher 1984). The stream-groundwater interface (SGI) is an ecotone where metabolic rates are often orders of magnitude greater than surface waters due to favorable physical and chemical conditions (Findlay and Sobczak 1996). Therefore, this zone is considered a “hotspot” and “control point” of nutrient and reactive solute transformation in watersheds (McClain et al. 2003; Bernhardt et al. 2017). Despite this interface being a known biogeochemical hotspot, there is still uncertainty in how DOC properties are altered in subsurface zones. There are few studies to date assessing DOC alteration in SGIs, and those that exist reveal different fates of DOC properties in the SGI. For example, Sobczak and Findlay (2002) did detailed characterizations at a one stream site and found that lower concentrations of DOC in the SGI were due to labile DOC removal and accumulation of recalcitrant DOC at depth. On the other hand, Helton et al. (2015), working at a single stream site, observed an increase in DOC lability as it moved along subsurface flowpaths. Because study observations at from different site SGIs can be contradictory, it is important to identify whether there are any

consistent patterns of DOC transformation and degradation across multiple stream sites, and to see what factors drive changes in DOC composition. Identifying consistent DOC transformations in SGIs would enable possible simplified representations SGI effects on the fate and transport of DOC in watershed process-based models. However, if there are not consistent DOC transformations across SGIs then models may need to account for local, site-specific conditions (e.g., water and sediment geochemistry) controlling the fate and transport of DOC through SGIs.

There are several factors that can complicate the interpretation of DOC transformations within the stream and SGI, and some of these key factors are briefly reviewed here. First, hydrologic conditions are highly dynamic in most stream networks and these dynamics influence how streams and groundwaters interact in the SGI. The SGI sometimes manifests as an ecotone called the hyporheic zone, where surface water enters the subsurface, mixes with groundwater, and returns to the surface (Boano et al. 2014). The relative presence and exchange rates of the hyporheic zone vary depending on factors such as the sinuosity, longitudinal elevation profile, and upwelling and downwelling of the stream water (Cardenas et al. 2004; Boano et al. 2014). Because the hyporheic zone has a substantial effect on solute transport processes (Stream Solute Workshop 1990; Ward et al. 2010; Aubeneau et al. 2016), variation in hyporheic zone hydrology will further control solute transport and transformation. Second, the source of stream DOC that can exchange with the SGI is highly variable. Much of the freshwater DOC may be derived from soil processes (e.g., Lambert et al. 2011) or from more direct inputs, such as litterfall, which are often exposed to highly reactive biological processes before ever reaching the stream (e.g. Amon & Benner, 1996). Also, DOC can be sourced from other aquatic ecosystems connected to streams, such as wetlands. In some cases, wetland complexes can contribute a disproportionate amount of DOC to streams, where DOC is tightly linked to wetland coverage within the

watershed (Bouchard 2007; Hansen et al. 2018). Third, DOC is highly reactive and undergoes both biotic and abiotic transformation processes such as photolytic degradation (e.g., Cory et al. 2014; Wagner and Jaffé 2015) or sorption to mineral surfaces (e.g., Nelson et al. 1993). All these DOC fate and transport factors interact in the SGI to create a complex system that is difficult to isolate specific controlling factors. Here, we isolate one aspect of the complexity—the role of the SGI physical transport by conducting bioassay studies—to help determine how similarly or differently unique sources of DOC will be processed.

In this study, we attempt to understand what, if any, consistent DOC processing occurs within the SGI across a variety of stream systems. To guide this assessment, we hypothesize that distinct DOC sources, with unique molecular properties and composition, will decay within the SGI and the remaining DOC will approach similar composition signatures, regardless of DOC origin or the specific SGI the DOC passes through. We propose this hypothesis because there several biotic and abiotic factors driving the fate of DOC in SGIs discussed above may act in concert to homogenize DOC properties (Figure 22). The processes we expect to occur in the SGI include physical processes such as mineral sorption of humic substances (e.g., Murphy et al. 1990) and biological degradation (Hur et al. 2011; Sleighter et al. 2014) of “fresher” or more labile DOC components. With these two processes acting as the primary modes of transformation of DOC, there will be a diverse array of DOC constituents remaining in solution, but we expect to see a predictable pattern emerging from exposure to stream water alone as compared to treatment with SGI sediments. We expect that the labile and sorption properties of DOC will be rapidly diminished, leaving relatively inert and recalcitrant DOC within the SGI. Over time, we expected to observe relatively inert DOC within the SGI, which ultimately would be available for downstream transport to other aquatic ecosystems (Kerner et al. 2003).

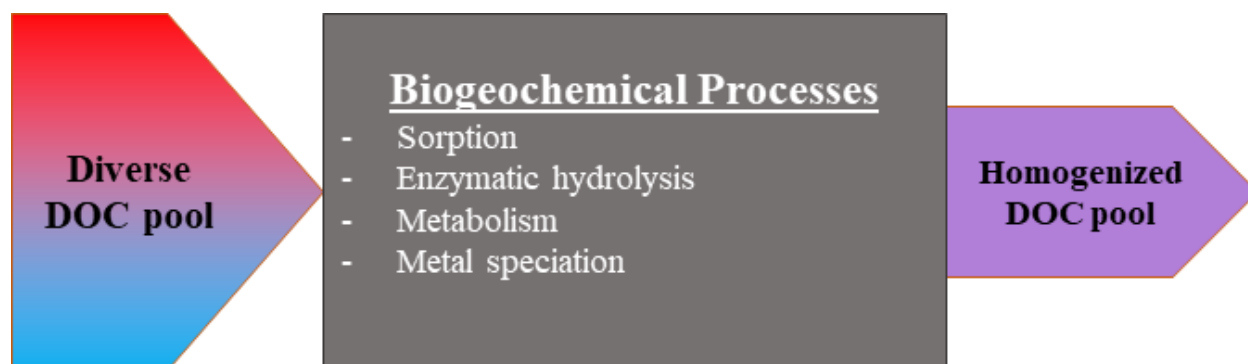


Figure 22 - Conceptual figure representing the proposed mechanisms by which DOC is consistently processed in the SGI. Here DOC enters the SGI as a unique, diverse pool of constituents. Once in the SGI processes such as sorption (e.g., Nelson et al. 1993), enzymatic hydrolysis (e.g., Volk et al. 1997), biological metabolism (e.g., Cammack et al. 2004), and redox transformation (e.g., Lovley et al. 1996) all occur, leaving only the less reactive components of the DOC pool to exit the SGI back into the surface stream water.

Materials and Methods

Field Sites and Sample collection/preparation

This study encompasses six sites across Michigan and Oklahoma (Table 1) including samples from a headwaters of the Kalamazoo River in southwest Michigan (Augusta Creek watershed), the Manistee and Au Sable watersheds in central Michigan, two tributaries to Lake Michigan coastal watersheds, and a tributary to the Red River in Oklahoma (Kiamichi River watershed).

Table 1 – Outline of site locations and the types of DOC sources used across each site.

Site	Location	DOC sources
Augusta Creek	Michigan	floc, 2 sources fresh leaves
Manistee	Michigan	aquatic vegetation, shrub foliage
Au Sable	Michigan	ground vegetation, shrub foliage
Lake Michigan Coastal 1	Michigan	fresh leaves, shrub foliage
Lake Michigan Coastal 2	Michigan	aquatic vegetation, shrub foliage
Kiamichi Watershed	Oklahoma	senesced leaves

Sources for DOC leaching were collected within the watershed for each stream studied and, in total, included 12 different sources of DOC likely to interact with the site SGIs. These sources encompassing aquatic vegetation, ground vegetation, shrub foliage, tree leaves, and flocculent organic sediment (floc). Vegetation sources included freshly clipped vegetation and previously senesced leaves recovered from the ground during water and sediment sampling. The primary purpose of a wide range of DOC sources was to ensure a broad range of unique DOC compositional properties. Each DOC sourced was collected and prepared as follows. Floc was collected by visually identifying floc patches in the streambed and collecting organic sediment from the mineral sediment surface. This floc sediment was then filtered through a 1 mm sieve and stored at 4°C until use in experiments. The other fresh DOC sources were dried in an oven at 65°C for 48 hours and stored in vacuum sealed plastic containers until use. Leachates of DOC were created using dried DOC sources. First, DOC sources were broken into coarse pieces. Then, approximately 6 g of dried mass was added to 1 L of ultrapure deionized water and allowed to steep for 24 hours. After 24 hours, the leachates were filtered through 0.45 µm glass microfiber filters (Whatman GF/F, 25 mm diameter). Leachates were stored in amber bottles at 4°C and used within 24 hours.

For each study site, stream water was collected from the stream channel in I-Chem LDPE Cubitainers (Thermo-Scientific) that were triple rinsed with ultrapure water. Water was stored in the dark and kept at 4°C until use. Stream sediment at each site was collected as a core within the top 10 cm of the streambed and kept saturated with water and at least 3 cm of water above it. Sediment was also stored in the dark and kept at 4°C until use. Before use, sediment was stirred to create a well-mixed sediment sample.

Laboratory bioassays

Our bioassay methods followed Dahm (1981) because it is transferable across the stream sites and DOC sources and only required minor modifications for our particular experiment as detailed below. We created a series of treatments that included ultrapure water with leachate, stream water with leachate, stream water and sediment with leachate, stream water only, and stream water with sediment only. Separate leachate treatments were created for each DOC source (Table 1) and all treatments were conducted in triplicate. For each replicate we added 450 mL of water (either stream or ultrapure) to acid-washed glass amber bottles. For replicates containing sediment, we weighted out 5 g of sediment and added it to the 450 mL of stream water. At the beginning of the experiment we added 16 mL of leachate to each replicate containing leachates and swirled bottles to mix well. Throughout the bioassay experiment we periodically (every 6-8 hours) mixed open bottles for 30 seconds to oxygenate the water via removing the top and swirling the bottle. The room temperature was monitored and maintained at $21\pm 2^{\circ}\text{C}$ to ensure there were minimal temperature effects on the experiment.

All treatments were sampled at three intervals: initial samples were collected from each bottle at 0 hours (i.e., immediately after all leachates were added and mixed); at 24 hours; and finally, at 48 hours. For each interval, a 60 mL sample of the solution was withdrawn. Of these 60 mL, first 10 mL were pushed through a $0.7\ \mu\text{m}$ glass microfiber filter (Whatman GF/F) and $0.2\ \mu\text{m}$ cellulose acetate filters (Sartorius) in series as a filter rinse. The final 50 mL was pushed through the filters into a 60 mL acid-washed plastic bottle. Filtered samples were stored at 4°C until sample analysis for DOC molecular properties as described below.

DOC analysis

As this study is focused on how DOC properties shift when interacting with SGI sediments, quantifying and analyzing multiple DOC molecular properties is the focus of the DOC analysis. These DOC molecular properties were characterized via optical characterization metrics described below. The sample optical measurements were determined using an Aqualog (Horiba Scientific) with 1 cm quartz cuvettes. We measured the spectral slope ratio (S_R), which serves as a proxy for molecular weight, as the slope from 275 to 295 nm divided by the slope from 350 to 400 nm (Helms et al. 2008). Fluorescence spectra were measured from excitation-emission matrices (EEMs) following Cory et al. (2010) to generate multiple DOC optical metrics. EEMs were all corrected for inner-filter effects as well as for instrument-specific corrections in Matlab (Cory et al. 2010). Fluorescence measures included the Fluorescence Index (FI), Freshness Index ($\beta:\alpha$), Humification Index (HIX), Peak A, Peak C, and Peak T. The FI was calculated as the ratio of emission intensity of 470 nm to the intensity of 520 nm at an excitation wavelength of 370 nm. The ratio $\beta:\alpha$, originally used in marine environments, represents the relative freshness of DOC and was calculated as the ratio at an emission of 380 nm over the maximum emission between 420 and 435 nm both at an excitation of 310 nm (Wilson and Xenopoulos 2009). The HIX represents the relative level of humification and is calculated as the area under the emission of 435 to 480 nm divided by the peak area from 310 to 345 nm all using an excitation wavelength of 254 nm (Ohno 2002). Peak A and Peak C both represent humic-like fluorophores and are calculated as the maximum emission from 400 to 460 nm at an excitation of 260 nm and the emission from 420 to 460 nm at an excitation emission of 320 to 360 nm, respectively (Coble et al. 2014). Finally, Peak T represents Tryptophan-like fluorophores (protein-like) and was calculated using the emission intensity at 340 nm from an excitation of 275 nm (Coble et al. 2014).

We used principal components analysis (PCA) to visualize the optical properties of the various sources of DOC used in the bioassays. This multivariate analysis allows us to look at the multidimensional DOC optical space in a simpler two-dimensional space. We included the fluorescence and absorbance metrics mentioned above. This enables us to explore how DOC sources varied relative to each other based on their optical properties.

Mixed effects model

To understand the relationship between the transformation of different sources of DOC in relation to their exposure to the SGI we developed mixed-effects linear models (Baayen et al. 2008b) using the lme4 package (Bates et al. 2015) in the open source computing software R (R Core Team 2017). We selected a mixed effects model to assess our hypothesis because it is a straightforward method that is robust for our sample size and has fewer assumptions that could potentially be violated as compared to some other statistical methods. In this modeling technique we specify the fixed and random effects (i.e., those parameters that do not vary and those that vary randomly). In our formulation we set time, the source of DOC, and the medium (e.g., stream water alone or stream water and sediment together) as fixed effects and the individual bioassay bottle as a random effect. The bottle was the only parameter set as a random effect as it would also encompass any heterogeneities in inter- and intra-site heterogeneity. Each DOC optical property, the dependent or response variable, was given its own mixed effect linear model. These response variables were transformed individually using Box-Cox transformations (Box and Cox 1964) to meet assumptions of normality.

Models were fitted using restricted maximum likelihood methods, which are less likely to commit Type 1 Errors (false positive) than the alternative, maximum likelihood (Luke 2017). Restricted maximum likelihood deals with small sample sizes better than maximum likelihood

methods (Pinheiro and Bates). Models were checked for normality of residuals at both the first (individual observations) and second (random) levels. Finally, models were evaluated in two ways. First, the significance of random effects was evaluated using a likelihood ratio test (LRT; Baayen et al. 2008) from the lmerTest package (Kuznetsova et al. 2017) in R (R Core Team 2017) where the model with a random effect included is compared to the model with the random effect removed to see if the random effect produces a significantly different model ($\alpha = 0.05$). Second, we corrected p-values of the fixed effects using the Kenward-Rogers adjustment (Littell et al. 2002) from the package stats (R Core Team 2017), which is used to adjust p-values due to the high dimensionality of our mixed effects linear model (Bühlmann 2013). The Kenward-Rogers adjustment is an approximation for degrees of freedom for the F distribution, which is appropriate for restricted maximum likelihood and relatively robust among other types of p-value adjustments (Luke 2017).

Results

DOC optical properties and Principal Components Analysis

The optical properties of DOC varied within a narrow range across all DOC sources (Table 1) and media, with the exception of one DOC source. The DOC from *Thuja occidentalis* from a Lake Michigan coastal watershed had Peak A and Peak C values that were consistently several factors larger than across other DOC sources. Average optical property values for each DOC source/medium combination can be found in Appendix B, Table 6.

The initial DOC properties and their changes during the bioassays were analyzed together in the PCA. The first principal component from the PCA explains 87.8% of the variance of the data, while the second principal component explains 5.9%. The first two principal components together explain 93.7% of the variance of the optical properties for this study (Figure 23). From the PCA we can see distinct clusters for DOC sources, with some sites clustering more closely

than others. Each DOC source has a trend of change through time of incubation, with values appearing to decrease along the first two principal components in several cases (Figure 24).

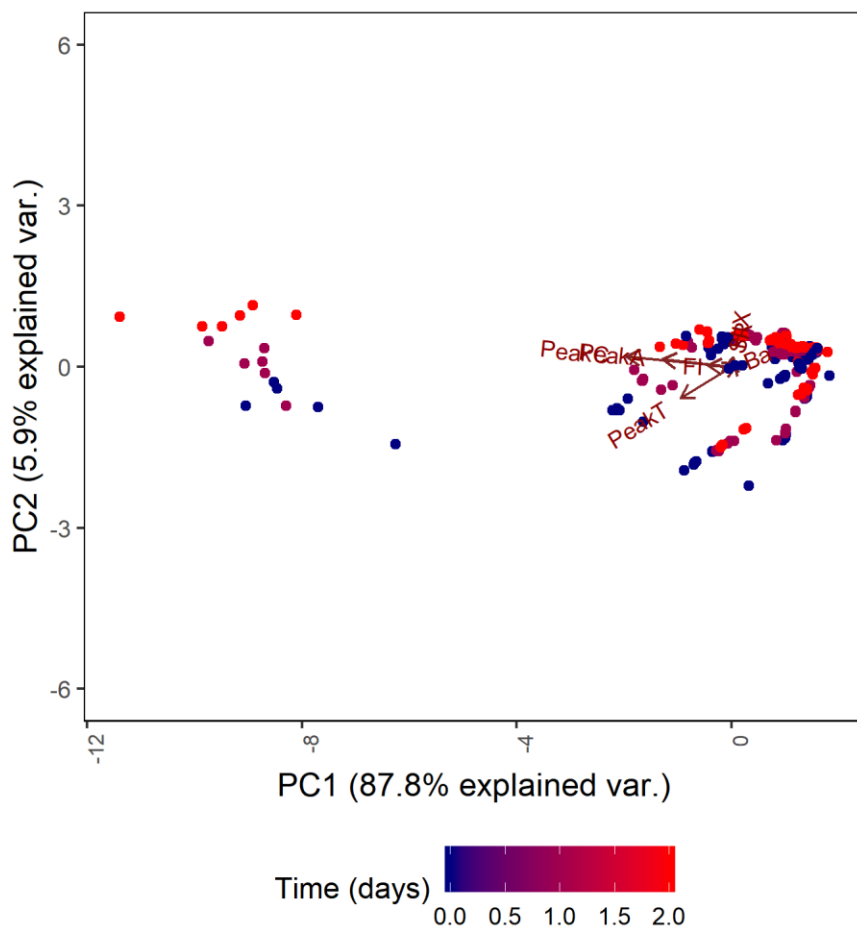


Figure 24 - PCA results of all optical measurements for all DOC sources included in this study, plotted the same as Figure 23 but with symbol colors indicating time of incubation in the bioassays, with initial samples appearing as navy blue and end point samples appearing as red. *Mixed effects models*

Across the mixed effects models, all fixed effect parameters and interaction parameters were considered significant ($p < 0.05$), but were dependent on each optical property (Table 2). For the Peak A model, the time parameter was not significant. For the FI model, one DOC source parameter was not significant. Finally, for the Peak T model, none of the interaction terms were significant. The random effects of bioassay bottle were significant for every model ($p < 0.05$). For all models, there was no overdispersion (i.e., observed variance is higher than expected) nor were any of the produced models singular (i.e., perfect correlation between columns of data in the model matrix).

Table 2 - Results of each mixed effects model for individual optical properties. An 'x' denotes significance ($p < 0.05$) of model parameters.

Independent Variable		Fixed Effects			Random Effect
Optical Property	Medium	DOC Treatment	Medium Treatment Interaction	Time	Bottle
S _R	x	x	x	x	x
Peak A	x	x	x		x
Peak C	x	x	x	x	x
Peak T	x	x		x	x
B:α	x	x	x	x	x
FI	x	x	x	x	x
HIX	x	x	x	x	x

Discussion

DOC Optical Properties

Our PCA approach revealed discrete clusters of the various DOC sources (Figure 23) demonstrating that our DOC leachates were distinct in their molecular composition. While we did not have an *a priori* concept of the magnitude of difference between leachates, we did expect that each leachate would cluster uniquely in a PCA, expecting that the DOC leached from various sources will produce unique fluorescent signatures as seen in prior studies (e.g., Cuss and Guéguen 2015). We note that most DOC sources fall between values of -2 and 2 on the PC1 axis, though one DOC source is clustered very far to the left on the PC1 axis. The large separation between *Thuja occidentalis* (commonly known as northern white-cedar, the fresh leaf DOC source at Lake Michigan Coastal 1 site) and the other DOC sources appears to be due to a high fluorescent value of the Peak T value, representing tryptophan-like protein fluorescence. For these leaf leachates, Peak T was orders of magnitude higher than other DOC sources, including other kinds of tree leaves used in this study. Studies of EEMs in leaf leachates have found that several components of the EEMs, observed through Parallel Factor Analysis, are strongly

associated with proteins commonly found in conifers (see e.g., Wickland et al. 2007; Cuss and Guéguen 2012), and *Thuja occidentalis* is broadly categorized as a conifer.

The bioassay method was used to specifically isolate the SGI effects on processing DOC. However, because of the controlled nature of this method there are several differences between the bioassay and in-stream conditions including, but not limited to, the flow of water, the oxic/anoxic gradient, and diurnal temperature fluctuations. There may be specific aspects of DOC processing that are overlooked due to maintenance of the bioassays in oxic conditions, such as the role of humic substances in microbial respiration under anoxic conditions (e.g., Scott et al. 1998). We also carried these studies out over approximately 48 hours, which is sufficient time to assess decomposition of the most labile DOC, but there is evidence that longer-term incubations can reveal more about DOC changes, particularly for the recalcitrant components (i.e., Peak A, Peak C) of the DOC pool (Kothawala et al. 2012).

Mixed Effects Models

In the mixed effects models for each optical property, most if not all fixed effects and interaction terms representing DOC sources were significant. When the interaction terms were significant, this suggests that DOC sources and the medium of interest (either stream water alone or stream water with sediment) have interdependencies that we cannot separate from each other. This can indicate that the DOC signal we see after DOC transformations occur in the stream depends both on its origin in the landscape and how that specific DOC source interacts with either sediment or stream water alone. In other words, we cannot state with confidence that the SGI sediment by itself is significant to driving DOC molecular properties without also considering the origin of the DOC. This result does not indicate a clear homogenization of DOC

signal from the SGI as hypothesized in this study; rather, it suggests that unique DOC signals may propagate into and out the SGI.

The one exception to this interdependence of DOC source and SGI effects and that was observed Peak T properties. The Peak T model had no significant interaction terms. This means that although the DOC source terms in the model were significant, so too was the medium term. Therefore, there was a predictable, attributable role of the medium across all treatment/medium combinations at all sites for Peak T. This further implies that there was a difference in how the Peak T portion of the DOC was transformed when exposed to sediments compared to exposure to stream water without sediments. For Peak T, there is no significant interaction between DOC sources and medium and therefore no interdependencies between Peak T and the medium. Together, this indicates that there is some sort of SGI homogenization effect on the Peak T signal across all unique sites and sources. Hence, there is the distinct possibility of a consistent function of the SGI across Peak T.

It should be noted that for several of the optical measurements of the bioassay water during experiments are likely impacted by the background DOC that could not be controlled for when adding natural SGI sediment and stream water to the bioassays. There is evidence in most of the bioassays that there was a background DOC source and that it is likely more highly decomposed DOC across systems than any of the fresher leachate derived DOC. The constant effect on Peak T by interaction with sediments as compared to stream water can likely be attributed to the specific characteristics of Peak T. Typically, Peak T is considered to be a proxy for labile, easily degraded DOC (Yamashita and Tanoue 2003; Nieto-Cid et al. 2006). In contrast, the optical properties Peak A, Peak C, and HIX all represent the humic, highly degraded components of the DOC pool. It is possible that these measurements could have been obscured

by background DOC already present in the stream water, as surface water DOC concentrations are strongly correlated to humic fluorescence (Wu et al. 2007). Humic substances make up a majority of freshwater DOC (McKnight and Aiken 1998; McKenna 2004) and undergo substantial decomposition upon leaving freshwater and entering into estuarine or marine systems (Kerner et al. 2003; Kisand et al. 2008), indicating a that there is additional processing potential of DOC that is not present in stream systems. FI may have similarly been obscured by the background measurement in the stream water. FI represents the DOC origin, binned categorically as allochthonous or autochthonous (McKnight et al. 2001), but is also an indicator of the aromaticity from fulvic acids (Battin 1998). It is highly probable that fulvic acids were already present in both the stream water and sediment used in this study (see e.g., Miller et al. 2006). The S_R , indicative of the molecular weight of DOC, may have been similarly hidden by the presence of highly decomposed DOC in stream water samples. Larger molecules associated with fresher materials make up a very small portion of the entire DOC pool (Wu et al. 2003) and given the highly labile nature of the fresh materials, could have been degraded down to smaller molecules in the stream water-only treatments just as easily as the sediment treatments.

Overall, our results indicate that there is not necessarily a ubiquitous function of the SGI as we predicted. Notably, the lack of a discernable stream-groundwater effect, as evidenced by the interdependence of sediment/no sediment with DOC sources in the mixed effects models, strongly implies that DOC transformations from location to location may vary dramatically. These results may help to clarify the previously introduced contradictory results such as between Helton et al. (2015) and Findlay and Sobczak (2002), where one group found an increase in biodegradable DOC and the other found a decrease within the SGI. Broadly speaking, our results indicate that future studies should be careful about generalizing and broadly categorizing these

SGI zones when it comes to how they process the very complex suite of DOC found in the environment.

There is plentiful evidence that DOC composition depends upon its origin within the landscape (e.g., Stedmon and Markager 2005; Cuss and Guéguen 2015; Hansen et al. 2016) and that stream chemistry (including DOC) is often a reflection of its surrounding landscape (e.g., McGuire et al. 2014). However, the interaction between the stream and its specific DOC sources may be an important consideration for field researchers and modelers trying to understand how DOC is processed moving from headwaters to downstream locations or trying to upscale SGI function to the watershed scale (Krause et al. 2011). As mentioned above, for all variables except Peak T we found statistically significant interdependency between the DOC sources and the bioassay medium. Our inability to separate which variable has a more profound effect further emphasizes a need to understand a watershed's particular placement in the landscape and the types of DOC inputs into the stream network. To some degree, this is already evident among microbial ecologists who have presented evidence that DOC source is linked to the composition of microbial communities (e.g., Docherty et al. 2006), but these details about quantity and qualities of DOC are rarely included in process-based watershed DOC models.

We selected a mixed effects model to assess our hypothesis because it is a is robust for our sample size and contains fewer assumptions to be violated compared to other statistical methods that could be used for our dataset. However, given the complex results of this initial investigation of DOC transformations with SGI sediment interactions across a range of stream sites, the scope of this research needs to expand. In this initial investigation, bioassays were logistically limited in the number of total treatments and bottles by the amount of time necessary to collect and store samples, the amount of space needed to create bioassays, and the time and

cost necessary for the analyses listed below. Clearly, there is a need to expand the bioassay experiments to better assess the background DOC in the SGI sediments and stream waters and to do more sites. Hence, we recommend that future studies with the resources and labor to carry out a study with more replicates from a greater number of samples across more sites, which also has the added benefit of being able to use more robust statistics, such as a repeated measures MANOVA or Bayesian methods.

Alternatively, there are many experiments that can be conducted within SGIs in the field that can also start to assess how DOC properties are altered in SGIs. These in situ studies of SGIs would provide a means to independently assess paired bioassay results. One of the common methods of testing in situ biogeochemical processing of SGI and other shallow groundwater environments is the push-pull method. However, this method will not work in all sites. In fact, this type of in situ push-pull testing was attempted to complement the bioassays of some of our study sites Augusta Creek. However, for these Augusta Creek SGI push-pull sites, the local water exchange rate through the SGI was large enough to prevent the assumptions from being met for of the push-pull experimental design and analysis (i.e., we observed that the tracer plume was being diluted by new, unlabeled groundwater). Beyond trying in situ investigations to complement the bioassays, there may be alternative lab experiments that can be used, such as column-scale experiments. Overall, bioassay efforts should be increased, and complementary field or laboratory investigations are likely needed to clearly flesh out a specific site, watershed, or cross watershed assessment of SGI functioning on DOC properties.

Conclusion

In this study we found that unique DOC sources from different watersheds were not consistently processed in the SGI of their respective watersheds according to most DOC optical

properties investigated. Because the interactions between sediment and DOC across six distinct watersheds were significant, the relative contribution to the final, decomposed DOC product can neither be attributed solely to the origin of the DOC nor to the sediment within which the DOC is processed. The one exception was the humic-like Peak T component, for which the SGI sediments significantly modified its signal. The results of this study indicate that researchers should take care when assessing DOC transformations in SGIs, which are commonly known to function as biogeochemical control points and reaction hot spots across watersheds. Contrary to our expectations, this initial study of multiple stream SGI sites indicates that there is no common function across watersheds for DOC processing in SGIs, despite the high level of reactivity. Clearly, more work using different bioassay methods and larger spatiotemporal representations of DOC processing from many different sites is needed before the similarities or heterogeneities of SGI biogeochemical functions will be revealed. We can suggest from this study that future work investigating DOC or DOC coupled reactions in SGIs should consider *in situ* or detailed bioassay tests of DOC processing in their study watersheds because there is little evidence in our findings outside of Peak T that a study can assume a priori that and SGI consistently alters DOC properties..

Acknowledgements

This work was partially supported by the National Science Foundation [EAR-1446328, EAR-1846855, and DEB-1637653], Michigan State University Department of Earth and Environmental Sciences, Michigan State University Environmental Science and Policy Program, and Michigan State University Kellogg Biological Station's Long Term Ecological Research site [DEB-1637653]. Thank you to R. Cory, A. Trusiak, J. Bowen, and L. Treibergs for assistance collecting EEMs data. Thank you to S. Hamilton and D. Kincaid for assistance with Augusta

Creek access and floc collection. Thank you to A. Shogren for reviewing an early draft. Thank you to R. Geiger, T. Parr, and T. Hampton for assistance collecting samples and conducting the bioassays. Thank you to B. Hannah, J. Michienzi, Q. Hamlin for obtaining samples.

APPENDICES

APPENDIX A: Chapter 2 Supplemental Tables and Figures

Table 3 - Information on groundwatersheds (GWshed) and surface watersheds (SWshed) across all 39 sampling points in the Augusta Creek watershed

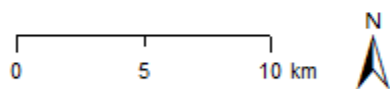
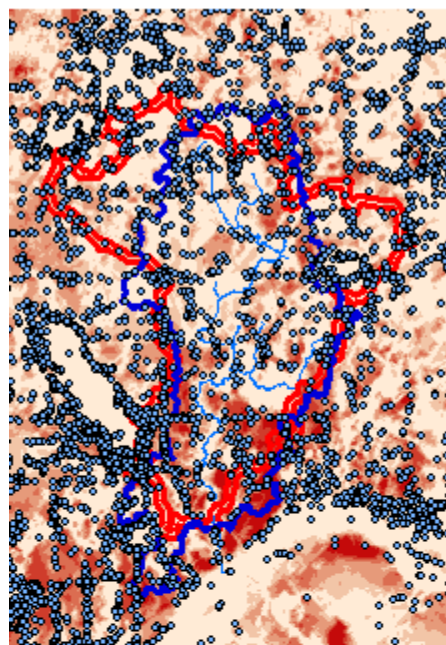
Watershed Designation	GWshed Area (km²)	SWshed Area (km²)	Shared Area (km²)	GWshed % Uniqueness	SWshed % Uniqueness
1	89.13	68.74	57.89	35.05	15.79
2	0.15	2.02	0.14	5.56	93.19
3	0.39	3.12	0.36	7.36	88.44
4	89.13	68.74	57.89	35.05	15.79
5	63.19	41.12	33.34	47.24	18.91
6	62.77	40.87	33.05	47.35	19.13
7	86.08	67.18	56.44	34.43	15.99
8	69.45	49.14	39.41	43.25	19.79
9	95.58	71.81	62.81	34.28	12.53
10	69.43	49.11	39.39	43.26	19.79
11	69.42	49.02	39.37	43.28	19.68
12	5.11	9.21	5.09	0.4	44.69
13	5.11	9.07	5.09	0.4	43.9
14	5.12	9.23	5.1	0.4	44.77
15	12.89	1.76	0.4	96.89	77.28
16	19.97	7.25	5.66	71.64	21.86
17	44.13	32.41	5.66	87.17	82.53
18	44.05	32.4	18.37	58.3	43.3
19	112.81	97.38	18.37	83.72	81.14
20	19.97	7.25	5.66	71.64	21.89
21	109.64	86.15	74.86	31.72	13.1
22	26.23	26.12	12.89	50.84	50.64
23	26.22	25.46	12.88	50.86	49.39
24	26.18	25.45	12.85	50.93	49.51
25	26.18	25.43	12.84	50.95	49.51
26	1.55	7.1	1.29	16.87	81.83
27	110.1	86.48	75.29	31.62	12.94
28	0.06	0.81	0.06	0	93.01
29	0.05	0.8	0.05	0	93.89
30	0.06	0.8	0.06	0	93.05
31	0.05	0.8	0.05	0	93.85
32	92.52	71.72	60.19	34.94	16.07
33	62.51	40.77	32.8	47.53	19.56
34	62.88	40.99	33.16	47.27	19.11
35	0.18	2.05	0.15	13.69	92.53
36	12.89	1.76	0.4	96.9	77.31
37	112.81	97.38	79.36	29.65	18.5
38	1.55	7.05	1.29	16.87	81.7
39	43.88	32.33	18.2	58.52	43.71

Table 4 - Best fit models for all measurements of surface watersheds (S) and groundwatersheds (G). A value in a column indicates that a the landcover category—Barren (B), Cultivated Crops (C), Emergent Herbaceous Wetland (EHW), Evergreen Forest (EF), Hay/Pasture (HP), Deciduous Forest (DF), Herbaceous (H), Open Water (OW), Shrub Scrubs (SS), Woody Wetland (WW), and Developed (D)—indicates coefficient value

	B	CC	EHW	EF	HP	DF	H	MF	OW	SS	WW	D
DOC S1		-0.66			-0.72		22.84		0.55	159.91		
DOC G1			-5.75	-3.29	-0.47					90.07		
DOC G2			-4.96		-0.52				1.28	39.16		
DOC G3			-4.32		-0.74			-7.30	2.05			
Cl ⁻ S1	13.35		18.32				15.64		1.53			7.66
Cl ⁻ G1	-109.66		4.04	7.92	0.17		23.50					
NO ₃ ⁻ S1	-16.29		-30.89				-22.42			-72.03		-0.60
NO ₃ ⁻ S2	-55.23					-1.72	-17.00					2.25
NO ₃ ⁻ S3								-21.42	-1.46		-1.67	
NO ₃ ⁻ G1		-0.64			1.55					-47.68		
NO ₃ ⁻ G2				14.76	1.65		-29.25	-17.00				
SUVA ₂₅₄ S1	-40.08		-18.77				-16.99			59.20		-2.71
SUVA ₂₅₄ S2	-53.412					-0.71	-13.13			84.68		-2.43
SUVA ₂₅₄ G1	44.666		-3.355				-60.125			53.023		
S _R S1									1.801	-109.903	1.479	
S _R G1	-84.801		5.256		-0.03							
S _R G2	-73.734		5.084									

Table 5 - Number of models passing our criteria for each biogeochemical measurement.

<i>Biogeochemical Measurement</i>	Groundwatershed Models (n)	Surface Watershed Models (n)
<i>DOC</i>	3	1
<i>NO₃⁻</i>	1	1
<i>Cl⁻</i>	3	2
<i>SUVA₂₅₄</i>	1	2
<i>S_R</i>	3	1



Depth to Water Table (ft)

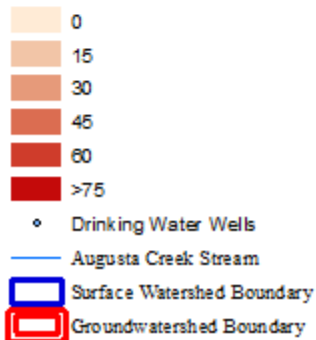


Figure 25 - Location of groundwater wells in the Augusta Creek watershed from which groundwater levels were determined. The pink to red gradient represents depth to water table ranging from 0 feet down to 75 feet or greater. Well data available through the State of Michigan Environment, Great Lakes & Energy Wellogic.

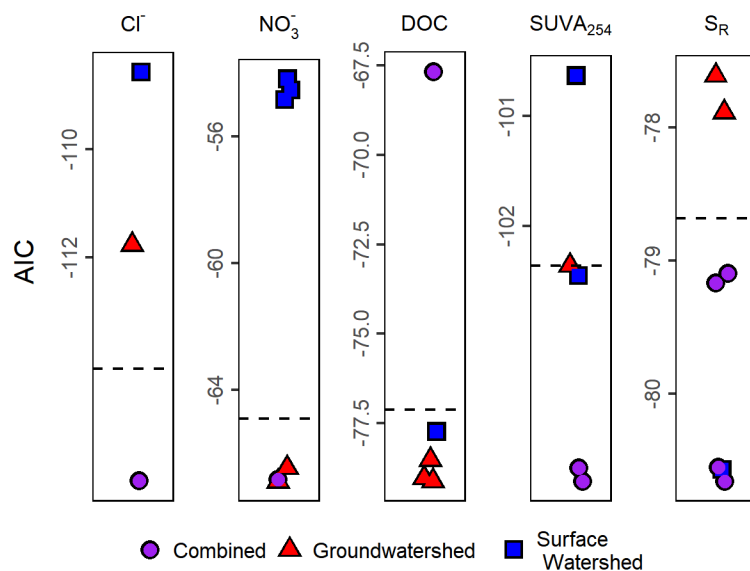


Figure 26 - AIC values for all best-fit models across each biogeochemical measurement. Each model is represented by a point, red circles are groundwater watershed models, blue triangles are surface watershed models, and purple circles are combined watersheds. All points that fall below the dashed line are considered to be comparable to the best-fit model. Models above the dashed line are considered to be worse than the best-fit model.

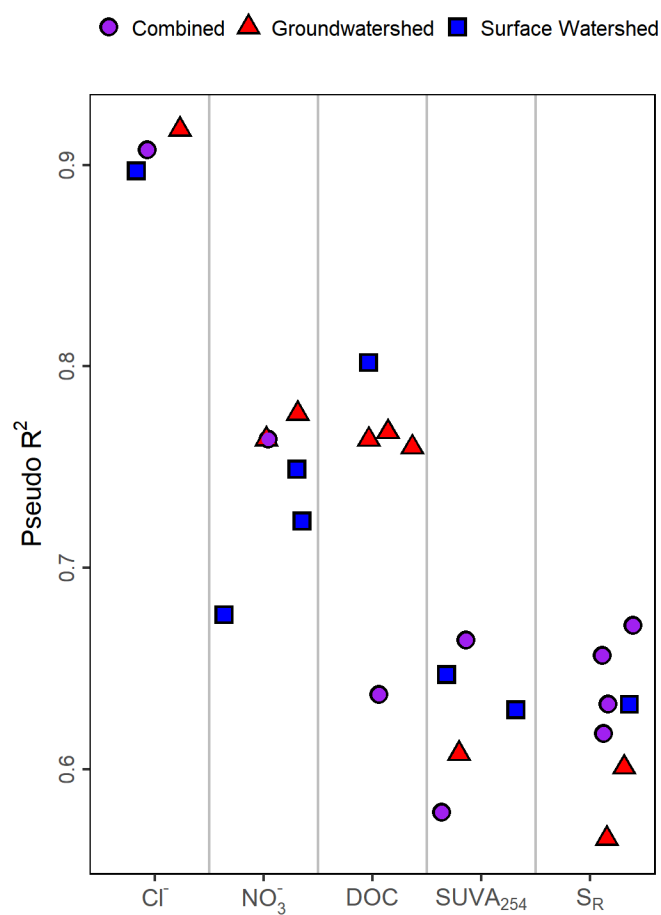


Figure 27 - Pseudo R² for all biogeochemical measurements for surface watersheds (blue square), groundwatersheds (red triangle), and combined watersheds (purple circles).

APPENDIX B: Chapter 4 Supplemental Table

Table 6 - Average optical properties for each medium/DOC source combination

Site	Medium	DOC source	S _R	FI	HIX	Peak A	Peak C	Peak T
Augusta Creek	Stream Water	<i>Ulmus americana</i>	1.34	1.58	0.83	1.50	0.56	0.46
Augusta Creek	Stream Water and Sediment	<i>Ulmus americana</i>	1.24	1.55	0.84	1.48	0.55	0.44
Augusta Creek	Stream Water	<i>Larix laricina</i>	1.09	1.51	0.81	1.44	0.53	0.55
Augusta Creek	Stream Water and Sediment	<i>Larix laricina</i>	1.15	1.53	0.82	1.44	0.53	0.52
Augusta Creek	Stream Water	Organic flocculent sediment	0.83	1.57	0.86	1.19	0.43	0.25
Augusta Creek	Stream Water and Sediment	Organic flocculent sediment	1.02	1.54	0.87	1.24	0.45	0.25
Manistee	Stream Water	<i>Potamogeton richardsonii</i>	1.22	1.48	0.83	1.52	0.65	0.57
Manistee	Stream Water and Sediment	<i>Potamogeton richardsonii</i>	1.21	1.47	0.83	1.54	0.66	0.60
Manistee	Stream Water	<i>Rosa palustris</i>	1.04	1.42	0.18	1.16	0.49	1.52
Manistee	Stream Water and Sediment	<i>Rosa palustris</i>	1.06	1.45	0.23	1.20	0.49	1.50
Au Sable	Stream Water	ground vegetation	1.43	1.38	0.59	1.01	0.45	1.13
Au Sable	Stream Water and Sediment	ground vegetation	1.64	1.38	0.63	0.99	0.43	0.95
Au Sable	Stream Water	<i>Aronia prunifolia</i>	0.78	2.03	0.41	1.88	1.19	2.58

Table 6 (cont'd)

Au Sable	Stream Water and Sediment	<i>Aronia prunifolia</i>	0.73	2.11	0.42	1.82	1.12	2.44
Lake Michigan Coastal 1	Stream Water	<i>Thuja occidentalis</i>	1.30	3.18	0.61	5.64	7.78	3.98
Lake Michigan Coastal 1	Stream Water and Sediment	<i>Thuja occidentalis</i>	1.24	3.27	0.63	6.20	8.13	4.24
Lake Michigan Coastal 1	Stream Water	<i>Asarum canadense</i>	0.91	1.78	0.74	4.06	1.75	1.91
Lake Michigan Coastal 1	Stream Water and Sediment	<i>Asarum canadense</i>	0.81	1.87	0.74	3.87	1.72	1.80
Lake Michigan Coastal 2	Stream Water	<i>Phragmites australis</i>	2.22	1.47	0.83	1.94	0.70	0.60
Lake Michigan Coastal 2	Stream Water and Sediment	<i>Phragmites australis</i>	2.69	1.49	0.85	1.95	0.71	0.56
Lake Michigan Coastal 2	Stream Water	<i>Cornus foemina</i>	0.99	1.66	0.84	3.03	1.69	0.88
Lake Michigan Coastal 2	Stream Water and Sediment	<i>Cornus foemina</i>	0.99	1.68	0.84	2.57	1.25	0.75
Kiamichi	Stream Water	<i>Platanus occidentalis</i>	1.11	1.99	0.86	2.01	1.29	0.61
Kiamichi	Stream Water and Sediment	<i>Platanus occidentalis</i>	1.13	2.02	0.86	1.98	1.30	0.50

REFERENCES

REFERENCES

- Abbott, B. W., V. Baranov, C. Mendoza-Lera, and others. 2016. Using multi-tracer inference to move beyond single-catchment ecohydrology. *Earth-Science Rev.* **160**: 19–42. doi:10.1016/J.EARSCIREV.2016.06.014
- Abbott, B. W., G. Gruau, J. P. Zarnetske, and others. 2018. Unexpected spatial stability of water chemistry in headwater stream networks J. Grover [ed.]. *Ecol. Lett.* **21**: 296–308. doi:10.1111/ele.12897
- Aiken, G. R. 2014. Dissolved organic matter in aquatic systems, p. 205–220. *In* S. Ahuja [ed.], *Comprehensive Water Quality and Purification*. Elsevier Inc.
- Alexander, R. B., E. W. Boyer, R. A. Smith, G. E. Schwarz, and R. B. Moore. 2007. The Role of Headwater Streams in Downstream Water Quality. *J. Am. Water Resour. Assoc.* **43**: 41–59. doi:10.1111/j.1752-1688.2007.00005.x
- Allan, D., D. Erickson, and J. Fay. 1997. The influence of catchment land use on stream integrity across multiple spatial scales. *Freshw. Biol.* **37**: 149–161. doi:10.1046/j.1365-2427.1997.d01-546.x
- Allen, T. F. H., and T. B. Starr. 1982. *Hierarchy : perspectives for ecological complexity*, T. Allen and T.B. Starr [eds.]. University of Chicago Press Chicago, IL.
- Amon, R. M. W., and R. Benner. 1996. Photochemical and microbial consumption of dissolved organic carbon and dissolved oxygen in the Amazon River system. *Geochim. Cosmochim. Acta* **60**: 1783–1792. doi:10.1016/0016-7037(96)00055-5
- Anselin, L. 1988. Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity. *Geogr. Anal.* **20**: 1–17. doi:10.1111/j.1538-4632.1988.tb00159.x
- Anselin, L. 1992. *SpaceStat TUTORIAL A Workbook for Using SpaceStat in the Analysis of Spatial Data*.
- Argerich, A., R. Haggerty, E. Martí, F. Sabater, and J. Zarnetske. 2011. Quantification of metabolically active transient storage (MATS) in two reaches with contrasting transient storage and ecosystem respiration. *J. Geophys. Res.* **116**: G03034. doi:10.1029/2010JG001379
- Aubeneau, A. F., B. Hanrahan, D. Bolster, and J. Tank. 2016. Biofilm growth in gravel bed streams controls solute residence time distributions. *J. Geophys. Res. Biogeosciences* **121**: 1840–1850. doi:10.1002/2016JG003333
- Baayen, R. H., D. J. Davidson, and D. M. Bates. 2008a. Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* **59**: 390–412. doi:10.1016/J.JML.2007.12.005

- Baayen, R. H., D. J. Davidson, and D. M. Bates. 2008b. Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang.* **59**: 390–412. doi:10.1016/J.JML.2007.12.005
- Barber, L. B., S. F. Murphy, P. L. Verplanck, M. W. Sandstrom, H. E. Taylor, and E. T. Furlong. 2005. Chemical Loading into Surface Water along a Hydrological, Biogeochemical, and Land Use Gradient: A Holistic Watershed Approach. doi:10.1021/ES051270Q
- Bates, D., M. Mächler, B. Bolker, and S. Walker. 2015. Fitting Linear Mixed-Effects Models Using **lme4**. *J. Stat. Softw.* **67**: 1–48. doi:10.18637/jss.v067.i01
- Battin, T. J. 1998. Dissolved organic matter and its optical properties in a blackwater tributary of the upper Orinoco river, Venezuela. *Org. Geochem.* **28**: 561–569. doi:10.1016/S0146-6380(98)00028-X
- Battin, T. J., L. a. Kaplan, S. Findlay, C. S. Hopkinson, E. Marti, A. I. Packman, J. D. Newbold, and F. Sabater. 2008. Biophysical controls on organic carbon fluxes in fluvial networks. *Nat. Geosci.* **1**: 95–100. doi:10.1038/ngeo101
- Bencala, K. E., R. E. Rathbun, A. Jackman, V. C. Kennedy, G. W. Zellweger, and R. J. Avanzino. 1983. Rhodamine WT dye losses in a mountain stream environment. *J. Am. Water Resour. Assoc.* **19**: 943–950. doi:10.1111/j.1752-1688.1983.tb05944.x
- Bernhardt, E. S., J. R. Blaszczak, C. D. Ficken, M. L. Fork, K. E. Kaiser, and E. C. Seybold. 2017. Control Points in Ecosystems: Moving Beyond the Hot Spot Hot Moment Concept. *Ecosystems* **20**: 665–682. doi:10.1007/s10021-016-0103-y
- Birgand, F., R. W. Skaggs, G. M. Chescheir, and J. W. Gilliam. 2007. Nitrogen Removal in Streams of Agricultural Catchments—A Literature Review. *Crit. Rev. Environ. Sci. Technol.* **37**: 381–487. doi:10.1080/10643380600966426
- Boano, F., a. Demaria, R. Revelli, and L. Ridolfi. 2010. Biogeochemical zonation due to intrameander hyporheic flow. *Water Resour. Res.* **46**: n/a-n/a. doi:10.1029/2008WR007583
- Boano, F., J. W. Harvey, A. Marion, A. I. Packman, R. Revelli, L. Ridolfi, and A. Wörmann. 2014. Hyporheic flow and transport processes: Mechanisms, models, and biogeochemical implications. *Rev. Geophys.* **52**: 603–679. doi:10.1002/2012RG000417
- Bouchard, V. 2007. Export of organic matter from a coastal freshwater wetland to Lake Erie: An extension of the outwelling hypothesis. *Aquat. Ecol.* **41**: 1–7. doi:10.1007/s10452-006-9044-4
- Boulton, A. J., S. Findlay, P. Marmonier, E. H. Stanley, and H. M. Valett. 1998. The functional significance of the hyporheic zone in streams and rivers. *Annu. Rev. Ecol. Syst.* **29**: 59–81. doi:10.1146/annurev.ecolsys.29.1.59
- Boulton, A. J., and P. J. Hancock. 2006. Rivers as groundwater-dependent ecosystems: a review of degrees of dependency, riverine processes and management implications. *Aust. J. Bot.*

54: 133. doi:10.1071/BT05074

- Boutt, D. F., D. W. Hyndman, B. C. Pijanowski, and D. T. Long. 2001. Identifying potential land use - Derived solute sources to stream baseflow using ground water models and GIS. *Ground Water* **39**: 24–34. doi:10.1111/j.1745-6584.2001.tb00348.x
- Box, G. E. P., and D. R. Cox. 1964. An Analysis of Transformations. *J. R. Stat. Soc. Ser. B* **26**: 211–243. doi:10.1111/j.2517-6161.1964.tb00553.x
- Brazner, J. C., N. P. Danz, A. S. Trebitz, and others. 2007. Responsiveness of Great Lakes Wetland Indicators to Human Disturbances at Multiple Spatial Scales: A Multi-Assemblage Assessment. *J. Great Lakes Res.* **33**: 42–66. doi:10.3394/0380-1330(2007)33[42:ROGLWI]2.0.CO;2
- Briggs, M. A., M. N. Gooseff, C. D. Arp, and M. A. Baker. 2009. A method for estimating surface transient storage parameters for streams with concurrent hyporheic storage. *Water Resour. Res.* **45**: n/a-n/a. doi:10.1029/2008WR006959
- Briggs, M. A., L. K. Lautz, and D. K. Hare. 2014. Residence time control on hot moments of net nitrate production and uptake in the hyporheic zone. *Hydrol. Process.* **28**: 3741–3751. doi:10.1002/hyp.9921
- Bühlmann, P. 2013. Statistical significance in high-dimensional linear models. *Bernoulli* **19**: 1212–1242. doi:10.3150/12-BEJSP11
- Burnham, K., and D. Anderson. 2002. *Model Selection and Multimodel Inference* 2nd ed.,.
- Byrne, P., a. Binley, a. L. Heathwaite, and others. 2014. Control of river stage on the reactive chemistry of the hyporheic zone. *Hydrol. Process.* **28**: 4766–4779. doi:10.1002/hyp.9981
- Cammack, W. K. L., J. Kalff, Y. T. Prairie, and E. M. Smith. 2004. Fluorescent dissolved organic matter in lakes: Relationships with heterotrophic metabolism. *Limnol. Oceanogr.* **49**: 2034–2045. doi:10.4319/lo.2004.49.6.2034
- Cardenas, M. B., J. L. Wilson, and V. A. Zlotnik. 2004. Impact of heterogeneity, bed forms, and stream curvature on subchannel hyporheic exchange. *Water Resour. Res.* **40**. doi:10.1029/2004WR003008
- Chow, R., M. E. Frind, E. O. Frind, J. P. Jones, M. R. Sousa, D. L. Rudolph, J. W. Molson, and W. Nowak. 2016. Delineating baseflow contribution areas for streams – A model and methods comparison. *J. Contam. Hydrol.* **195**: 11–22. doi:10.1016/J.JCONHYD.2016.11.001
- Cole, J. J., Y. T. Prairie, N. F. Caraco, and others. 2007. Plumbing the global carbon cycle: Integrating inland waters into the terrestrial carbon budget. *Ecosystems* **10**: 171–184. doi:10.1007/s10021-006-9013-8
- Cory, R. M., M. P. Miller, D. M. McKnight, J. J. Guerard, and P. L. Miller. 2010. Effect of

- instrument-specific response on the analysis of fulvic acid fluorescence spectra. *Limnol. Oceanogr. Methods* **8**: 67–78. doi:10.4319/lom.2010.8.67
- Cory, R. M., C. P. Ward, B. C. Crump, and G. W. Kling. 2014. Sunlight controls water column processing of carbon in arctic fresh waters. *Science* (80-.). **345**: 925–928. doi:10.1126/science.1253119
- Creed, I. F., D. M. McKnight, B. A. Pellerin, and others. 2015. The river as a chemostat: fresh perspectives on dissolved organic matter flowing down the river continuum. *Can. J. Fish. Aquat. Sci.* 150416143426006. doi:10.1139/cjfas-2014-0400
- Cuss, C. W., and C. Guéguen. 2012. Determination of relative molecular weights of fluorescent components in dissolved organic matter using asymmetrical flow field-flow fractionation and parallel factor analysis. *Anal. Chim. Acta* **733**: 98–102. doi:10.1016/J.ACA.2012.05.003
- Cuss, C. W., and C. Guéguen. 2015. Relationships between molecular weight and fluorescence properties for size-fractionated dissolved organic matter from fresh and aged sources. *Water Res.* **68**: 487–97. doi:10.1016/j.watres.2014.10.013
- Dahl, M., B. Nilsson, J. H. Langhoff, and J. C. Refsgaard. 2007. Review of classification systems and new multi-scale typology of groundwater–surface water interaction. *J. Hydrol.* **344**: 1–16. doi:10.1016/J.JHYDROL.2007.06.027
- Dahm, C. N. 1981. Pathways and Mechanisms for Removal of Dissolved Organic Carbon from Leaf Leachate in Streams. *Can. J. Fish. Aquat. Sci.* **38**: 68–76.
- Delcourt, H. R., P. A. Delcourt, and T. Webb. 1982. Dynamic plant ecology: the spectrum of vegetational change in space and time. *Quat. Sci. Rev.* **1**: 153–175. doi:10.1016/0277-3791(82)90008-7
- Dent, C. L., and N. B. Grimm. 1999. Spatial heterogeneity of stream water nutrient concentrations over successional time. *Ecology* **80**: 2283–2298. doi:10.1890/0012-9658(1999)080[2283:SHOSWN]2.0.CO;2
- Dillon, P. J., and L. A. Molot. 1997. Effect of landscape form on export of dissolved organic carbon, iron, and phosphorus from forested stream catchments. *Water Resour. Res.* **33**: 2591–2600. doi:10.1029/97WR01921
- Docherty, K. M., K. C. Young, P. A. Maurice, and S. D. Bridgham. 2006. Dissolved Organic Matter Concentration and Quality Influences upon Structure and Function of Freshwater Microbial Communities. *Microb. Ecol.* **52**: 378–388. doi:10.1007/s00248-006-9089-x
- Dodds, W. K., and R. M. Oakes. 2006. Controls on Nutrients Across a Prairie Stream Watershed: Land Use and Riparian Cover Effects. *Environ. Manage.* **37**: 634–646. doi:10.1007/s00267-004-0072-3
- Downing, J. A., J. J. Cole, C. M. Duarte, and others. 2012. Global abundance and size

- distribution of streams and rivers. *Int. Waters* **2**: 229–236. doi:10.5268/IW-2.4.502
- Duff, J. H., F. Murphy, C. C. Fuller, F. J. Triska, J. W. Harvey, and A. P. Jackman. 1998. A mini drivepoint sampler for measuring pore-water solute concentrations in the hyporheic zone of sand-bottom streams. *Limnol. Oceanogr.* **43**: 1378–1383. doi:10.4319/lo.1998.43.6.1378
- Dunn, O. J. 1964. Multiple Comparisons Using Rank Sums. *Technometrics* **6**: 241–252. doi:10.1080/00401706.1964.10490181
- Edmonds, J. W., and N. B. Grimm. 2011. Abiotic and biotic controls of organic matter cycling in a managed stream. *J. Geophys. Res.* **116**: G02015. doi:10.1029/2010JG001429
- Fasching, C., B. Behounek, G. A. Singer, and T. J. Battin. 2015a. Microbial degradation of terrigenous dissolved organic matter and potential consequences for carbon cycling in brown-water streams. *Sci. Rep.* **4**: 4981. doi:10.1038/srep04981
- Fasching, C., A. J. Ulseth, J. Schelker, G. Steniczka, and T. J. Battin. 2015b. Hydrology controls dissolved organic matter export and composition in an Alpine stream and its hyporheic zone. *Limnol. Oceanogr.* n/a-n/a. doi:10.1002/lno.10232
- Findlay, S., and W. V. Sobczak. 1996. Variability in removal of dissolved organic carbon in hyporheic sediments. *J. North Am. Benthol. Soc.* **15**: 35–41.
- Fisher, S. G., R. A. Sponseller, and J. B. Heffernan. 2004. Horizons in stream biogeochemistry: flowpaths to progress. *Ecology* **85**: 2369–2379. doi:10.1890/03-0244
- Floyd, W. C., S. H. Schoenholtz, S. M. Griffith, P. J. Wigington, and J. J. Steiner. 2009. Nitrate-Nitrogen, Land Use/Land Cover, and Soil Drainage Associations at Multiple Spatial Scales. *J. Environ. Qual.* **38**: 1473. doi:10.2134/jeq2008.0099
- Frost, P. C., J. H. Larson, C. A. Johnston, K. C. Young, P. A. Maurice, G. A. Lamberti, and S. D. Bridgham. 2006. Landscape predictors of stream dissolved organic matter concentration and physicochemistry in a Lake Superior river watershed. *Aquat. Sci.* **68**: 40–51. doi:10.1007/s00027-005-0802-5
- Gomez-Velez, J. D., J. W. Harvey, M. B. Cardenas, and B. Kiel. 2015. Denitrification in the Mississippi River network controlled by flow through river bedforms. *Nat. Geosci.* **8**: 941–945. doi:10.1038/ngeo2567
- Gomez, J. D., J. L. Wilson, and M. B. Cardenas. 2012. Residence time distributions in sinuosity-driven hyporheic zones and their biogeochemical effects. *Water Resour. Res.* **48**: n/a-n/a. doi:10.1029/2012WR012180
- Grimm, N. B., and S. G. Fisher. 1984. Exchange between interstitial and surface water: Implications for stream metabolism and nutrient cycling. *Hydrobiologia* **111**: 219–228. doi:10.1007/BF00007202
- Haacker, E. M. K., A. D. Kendall, and D. W. Hyndman. 2016. Water Level Declines in the High

- Plains Aquifer: Predevelopment to Resource Senescence. *Groundwater* **54**: 231–242. doi:10.1111/gwat.12350
- Haggerty, R. 2009. STAMMT-L 3.0. A solute transport code for multirate mass transfer and reaction along flowlines.
- Hamilton, S. K. 2012. Biogeochemical time lags may delay responses of streams to ecological restoration. *Freshw. Biol.* **57**: 43–57. doi:10.1111/j.1365-2427.2011.02685.x
- Hamilton, S. K., M. Z. Hussain, C. Lowrie, B. Basso, and G. P. Robertson. 2018. Evapotranspiration is resilient in the face of land cover and climate change in a humid temperate catchment. *Hydrol. Process.* **32**: 655–663. doi:10.1002/hyp.11447
- Hansen, A. M., T. E. C. Kraus, B. A. Pellerin, J. A. Fleck, B. D. Downing, and B. A. Bergamaschi. 2016. Optical properties of dissolved organic matter (DOM): Effects of biological and photolytic degradation. *Limnol. Oceanogr.* **61**: 1015–1032. doi:10.1002/lno.10270
- Hansen, A. T., C. L. Dolph, E. Foufoula-Georgiou, and J. C. Finlay. 2018. Contribution of wetlands to nitrate removal at the watershed scale. *Nat. Geosci.* **11**: 127–132. doi:10.1038/s41561-017-0056-6
- Harvey, J., and M. Gooseff. 2015. River corridor science: Hydrologic exchange and ecological consequences from bedforms to basins. *Water Resour. Res.* **51**: 6893–6922. doi:10.1002/2015WR017617
- Harvey, J. W., and C. C. Fuller. 1998. Effect of enhanced manganese oxidation in the hyporheic zone on basin-scale geochemical mass balance. *Water Resour. Res.* **34**: 623–636. doi:10.1029/97WR03606
- Hedin, L. O., J. C. von Fischer, N. E. Ostrom, B. P. Kennedy, M. G. Brown, and G. P. Robertson. 1998. Thermodynamic constraints on nitrogen transformations and other biogeochemical processes at soil-stream interfaces. *Ecology* **79**: 684–703. doi:10.1890/0012-9658(1998)079[0684:TCNAO]2.0.CO;2
- Helms, J. R., A. Stubbins, J. D. Ritchie, E. C. Minor, D. J. Kieber, and K. Mopper. 2008. Absorption spectral slopes and slope ratios as indicators of molecular weight, source, and photobleaching of chromophoric dissolved organic matter. *Limnol. Oceanogr.* **53**: 955–969. doi:10.4319/lno.2008.53.3.0955
- Helton, A. M., M. S. Wright, E. S. Bernhardt, G. C. Poole, R. M. Cory, and J. A. Stanford. 2015. Dissolved organic carbon lability increases with water residence time in the alluvial aquifer of a river floodplain ecosystem. *J. Geophys. Res. Biogeosciences* n/a-n/a. doi:10.1002/2014JG002832
- Herlihy, A. T., J. L. Stoddard, and C. B. Johnson. 1998. The Relationship Between Stream Chemistry and Watershed Land Cover Data in the Mid-Atlantic Region, U.S. *Water, Air, Soil Pollut.* **105**: 377–386. doi:10.1023/A:1005028803682

- Hester, E. T., and M. N. Gooseff. 2010. Moving Beyond the Banks: Hyporheic Restoration Is Fundamental to Restoring Ecological Services and Functions of Streams. *Environ. Sci. Technol.* **44**: 1521–1525. doi:10.1021/es902988n
- Hill, A. R. 1990. Ground water flow paths in relation to nitrogen chemistry in the near-stream zone. *Hydrobiologia* **206**: 39–52. doi:10.1007/BF00018968
- Hill, A. R. 1996. Nitrate Removal in Stream Riparian Zones. *J. Environ. Qual.* **25**: 743. doi:10.2134/jeq1996.00472425002500040014x
- Hinton, M. J., S. L. Schiff, and M. C. English. 1993. Physical properties governing groundwater flow in a glacial till catchment. *J. Hydrol.* **142**: 229–249. doi:10.1016/0022-1694(93)90012-X
- Hollister, J. W., P. V. August, and J. F. Paul. 2008. Effects of spatial extent on landscape structure and sediment metal concentration relationships in small estuarine systems of the United States' Mid-Atlantic Coast. *Landsc. Ecol.* **23**: 91–106. doi:10.1007/s10980-007-9143-1
- Homer, C., J. Dewitz, L. Yang, and others. 2015. Completion of the 2011 National Land Cover Database for the Conterminous United States-Representing a Decade of Land Cover Change Information. *Photogramm. Eng. Remote Sens.* **81**: 345–354. doi:10.14358/PERS.81.5.345
- Hotchkiss, E. R., S. Sadro, and P. C. Hanson. 2018. Toward a more integrative perspective on carbon metabolism across lentic and lotic inland waters. *Limnol. Oceanogr. Lett.* **3**: 57–63. doi:10.1002/lol2.10081
- Hur, J., B.-M. Lee, and H.-S. Shin. 2011. Microbial degradation of dissolved organic matter (DOM) and its influence on phenanthrene–DOM interactions. *Chemosphere* **85**: 1360–1367. doi:10.1016/J.CHEMOSPHERE.2011.08.001
- Hynes, H. B. N. 1983. Groundwater and stream ecology. *Hydrobiologia* **100**: 93–99. doi:10.1007/BF00027424
- Inwood, S. E., J. L. Tank, and M. J. Bernot. 2005. Patterns of denitrification associated with land use in 9 midwestern headwater streams. *J. North Am. Benthol. Soc.* **24**: 227–245. doi:10.1899/04-032.1
- Jackson, T. R., R. Haggerty, S. V. Apte, A. Coleman, and K. J. Drost. 2012. Defining and measuring the mean residence time of lateral surface transient storage zones in small streams. *Water Resour. Res.* **48**: n/a-n/a. doi:10.1029/2012WR012096
- Kaplan, L. A., and J. D. Newbold. 2000. Surface and subsurface dissolved organic carbon, p. 237–258. *In* J.B. Jones and P.J. Mulholland [eds.], *Streams and Groundwaters*. Academic Press.
- Kelleher, C., T. Wagener, B. McGlynn, A. S. Ward, M. N. Gooseff, and R. A. Payn. 2013.

- Identifiability of transient storage model parameters along a mountain stream. *Water Resour. Res.* **49**: 5290–5306. doi:10.1002/wrcr.20413
- Kelly, W. R. 2008. Long-Term Trends in Chloride Concentrations in Shallow Aquifers near Chicago. *Ground Water* **46**: ???-??? doi:10.1111/j.1745-6584.2008.00466.x
- Kerner, M., H. Hohenberg, S. Ertl, M. Reckermann, and A. Spitzzy. 2003. Self-organization of dissolved organic matter to micelle-like microparticles in river water. *Nature* **422**: 150–154. doi:10.1038/nature01469
- Kiel, B. A., and M. B. Cardenas. 2014. Lateral hyporheic exchange throughout the Mississippi River network. *Nat. Geosci.* **7**: 413–417. doi:10.1038/ngeo2157
- Kincaid, D. W., and S. E. G. Findlay. 2009. Sources of Elevated Chloride in Local Streams: Groundwater and Soils as Potential Reservoirs. *Water, Air, Soil Pollut.* **203**: 335–342. doi:10.1007/s11270-009-0016-x
- Kisand, V., D. Røcker, and M. Simon. 2008. Significant decomposition of riverine humic-rich DOC by marine but not estuarine bacteria assessed in sequential chemostat experiments. *Aquat. Microb. Ecol.* **53**: 151–160. doi:10.3354/ame01240
- de Koning, G. H. ., A. Veldkamp, and L. . Fresco. 1998. Land use in Ecuador: a statistical analysis at different aggregation levels. *Agric. Ecosyst. Environ.* **70**: 231–247. doi:10.1016/S0167-8809(98)00151-0
- Kothawala, D. N., E. von Wachenfeldt, B. Koehler, and L. J. Tranvik. 2012. Selective loss and preservation of lake water dissolved organic matter fluorescence during long-term dark incubations. *Sci. Total Environ.* **433**: 238–246. doi:10.1016/J.SCITOTENV.2012.06.029
- Kratz, T. K., B. J. Benson, E. R. Blood, G. L. Cunningham, and R. A. Dahlgren. 1991. The Influence of Landscape Position on Temporal Variability in Four North American Ecosystems. *Am. Nat.* **138**: 355–378. doi:10.1086/285222
- Krause, S., D. M. Hannah, J. H. Fleckenstein, and others. 2011. Inter-disciplinary perspectives on processes in the hyporheic zone. *Ecohydrology* **4**: 481–499. doi:10.1002/eco.176
- Krivoruchko, K. 2012. Empirical Bayesian Kriging Implemented in ArcGIS Geostatistical Analyst.
- Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen. 2017. lmerTest Package: Tests in Linear Mixed Effects Models. *J. Stat. Softw.* **82**. doi:10.18637/jss.v082.i13
- Lambert, T., A.-C. Pierson-Wickmann, G. Gruau, A. Jaffrezic, P. Petitjean, J.-N. Thibault, and L. Jeanneau. 2013. Hydrologically driven seasonal changes in the sources and production mechanisms of dissolved organic carbon in a small lowland catchment. *Water Resour. Res.* **49**: 5792–5803. doi:10.1002/wrcr.20466
- Lambert, T., A. C. Pierson-Wickmann, G. Gruau, J. N. Thibault, and A. Jaffrezic. 2011. Carbon

- isotopes as tracers of dissolved organic carbon sources and water pathways in headwater catchments. *J. Hydrol.* **402**: 228–238. doi:10.1016/j.jhydrol.2011.03.014
- Lautz, L. K., and R. M. Fanelli. 2008. Seasonal biogeochemical hotspots in the streambed around restoration structures. *Biogeochemistry* **91**: 85–104. doi:10.1007/s10533-008-9235-2
- Lee-Cullin, J. A., J. P. Zarnetske, S. S. Ruhala, and S. Plont. 2018. Toward measuring biogeochemistry within the stream-groundwater interface at the network scale: An initial assessment of two spatial sampling strategies. *Limnol. Oceanogr. Methods* **16**. doi:10.1002/lom3.10277
- Littell, R. C., W. W. Stroup, R. J. Freund, and R. C. Littell. 2002. SAS for linear models,.
- Lovley, D. R., J. D. Coates, E. L. Blunt-Harris, E. J. P. Phillips, and J. C. Woodward. 1996. Humic substances as electron acceptors for microbial respiration. *Nature* **382**: 445–448. doi:10.1038/382445a0
- Lowe, W. H., G. E. Likens, and M. E. Power. 2006. Linking Scales in Stream Ecology. *Bioscience* **56**: 591. doi:10.1641/0006-3568(2006)56[591:LSISE]2.0.CO;2
- Luke, S. G. 2017. Evaluating significance in linear mixed-effects models in R. *Behav. Res. Methods* **49**: 1494–1502. doi:10.3758/s13428-016-0809-y
- Mann, P. J., A. Davydova, N. Zimov, R. G. M. Spencer, S. Davydov, E. Bulygina, S. Zimov, and R. M. Holmes. 2012. Controls on the composition and lability of dissolved organic matter in Siberia's Kolyma River basin. *J. Geophys. Res. Biogeosciences* **117**. doi:10.1029/2011JG001798
- Manny, B. A., and R. G. Wetzel. 1973. Diurnal changes in dissolved organic and inorganic carbon and nitrogen in a hardwater stream. *Freshw. Biol.* **3**: 31–43. doi:10.1111/j.1365-2427.1973.tb00060.x
- Martin, S. L., D. B. Hayes, A. D. Kendall, and D. W. Hyndman. 2017. The land-use legacy effect: Towards a mechanistic understanding of time-lagged water quality responses to land use/cover. *Sci. Total Environ.* **579**: 1794–1803. doi:10.1016/J.SCITOTENV.2016.11.158
- Marzadri, A., M. M. Dee, D. Tonina, A. Bellin, and J. L. Tank. 2017. Role of surface and subsurface processes in scaling N₂O emissions along riverine networks. *Proc. Natl. Acad. Sci. U. S. A.* **114**: 4330–4335. doi:10.1073/pnas.1617454114
- Mattsson, T., P. Kortelainen, and A. Räike. 2005. Export of DOM from Boreal Catchments: Impacts of Land Use Cover and Climate. *Biogeochemistry* **76**: 373–394. doi:10.1007/s10533-005-6897-x
- McClain, M. E., E. W. Boyer, C. L. Dent, and others. 2003. Biogeochemical Hot Spots and Hot Moments at the Interface of Terrestrial and Aquatic Ecosystems. *Ecosystems* **6**: 301–312. doi:10.1007/s10021-003-0161-9

- McGuire, K. J., C. E. Torgersen, G. E. Likens, D. C. Buso, W. H. Lowe, and S. W. Bailey. 2014. Network analysis reveals multiscale controls on streamwater chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **111**: 7030–5. doi:10.1073/pnas.1404820111
- McKenna, J. H. 2004. DOC dynamics in a small temperate estuary: Simultaneous addition and removal processes and implications on observed nonconservative behavior. *Estuaries* **27**: 604–616. doi:10.1007/BF02907648
- McKnight, D. M., and G. R. Aiken. 1998. Sources and Age of Aquatic Humus, p. 9–39. *In* Springer, Berlin, Heidelberg.
- McKnight, D. M., E. W. Boyer, P. K. Westerhoff, P. T. Doran, T. Kulbe, and D. T. Andersen. 2001. Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnol. Oceanogr.* **46**: 38–48. doi:10.4319/lo.2001.46.1.0038
- Miller, M. P., D. McKnight, R. M. Cory, M. W. Williams, and R. L. Runkel. 2006. Hyporheic Exchange and Fulvic Acid Redox Reactions in an Alpine Stream/Wetland Ecosystem, Colorado Front Range. *Environ. Sci. Technol.* **40**: 5943–5949. doi:10.1021/ES060635J
- Mulholland, P. J. 1997. Dissolved Organic Matter Concentration and Flux in Streams. *J. North Am. Benthol. Soc.* **16**: 131–141. doi:10.2307/1468246
- Murphy, E. M., J. M. Zachara, and S. C. Smith. 1990. Influence of mineral-bound humic substances on the sorption of hydrophobic organic compounds. *Environ. Sci. Technol.* **24**: 1507–1516. doi:10.1021/es00080a009
- Nelson, A. R., A. H. Sawyer, R. S. Gabor, and others. 2019. Heterogeneity in Hyporheic Flow, Pore Water Chemistry, and Microbial Community Composition in an Alpine Streambed. *J. Geophys. Res. Biogeosciences* **124**: 3465–3478. doi:10.1029/2019JG005226
- Nelson, P. N., J. A. Baldock, and J. M. Oades. 1993. Concentration and composition of dissolved organic carbon in streams in relation to catchment soil properties. *Biogeochemistry* **19**: 27–50. doi:10.1007/BF00000573
- Nieto-Cid, M., X. A. Álvarez-Salgado, and F. F. Pérez. 2006. Microbial and photochemical reactivity of fluorescent dissolved organic matter in a coastal upwelling system. *Limnol. Oceanogr.* **51**: 1391–1400. doi:10.4319/lo.2006.51.3.1391
- O'Donnell, J. A., G. R. Aiken, M. A. Walvoord, and K. D. Butler. 2012. Dissolved organic matter composition of winter flow in the Yukon River basin: Implications of permafrost thaw and increased groundwater discharge. *Global Biogeochem. Cycles* **26**: n/a-n/a. doi:10.1029/2012GB004341
- Ocampo, C. J., C. E. Oldham, and M. Sivapalan. 2006. Nitrate attenuation in agricultural catchments: Shifting balances between transport and reaction. *Water Resour. Res.* **42**. doi:10.1029/2004WR003773

- Ohno, T. 2002. Fluorescence Inner-Filtering Correction for Determining the Humification Index of Dissolved Organic Matter. doi:10.1021/ES0155276
- Overmars, K. P., G. H. J. de Koning, and A. Veldkamp. 2003. Spatial autocorrelation in multi-scale land use models. *Ecol. Modell.* **164**: 257–270. doi:10.1016/S0304-3800(03)00070-X
- Pinay, G., N. E. Haycock, C. Ruffinoni, and R. M. Holmes. 1994. The role of denitrification in nitrogen removal in river corridors., p. 107–116. *In* W.J. Mitsch [ed.], *Global Wetlands: Old World and New*. Elsevier.
- Pinay, G., T. C. O’Keefe, R. T. Edwards, and R. J. Naiman. 2009. Nitrate removal in the hyporheic zone of a salmon river in Alaska. *River Res. Appl.* **25**: 367–375. doi:10.1002/rra.1164
- Pinay, G., S. Peiffer, J. R. De Dreuz, and others. 2015. Upscaling Nitrogen Removal Capacity from Local Hotspots to Low Stream Orders’ Drainage Basins. *Ecosystems* **18**: 1101–1120. doi:10.1007/s10021-015-9878-5
- Pinheiro, J. C., and D. Bates. Fitting Linear Mixed-Effects Models, p. 133–199. *In* *Mixed-Effects Models in S and S-PLUS*. Springer-Verlag.
- Pinney, M. L., P. K. W. M., and L. B. M. 2000. Transformation in dissolved organic carbon through constructed wetlands. *Water Research* **34**: 1897–1911. doi: 10.1016/S0043-1354(99)00330-9
- Poff, N. L., J. D. Allan, M. B. Bain, J. R. Karr, K. L. Prestegard, B. D. Richter, R. E. Sparks, and J. C. Stromberg. 1997. The Natural Flow Regime. *Bioscience* **47**: 769–784. doi:10.2307/1313099
- Poole, G. C. 2002. Fluvial landscape ecology: addressing uniqueness within the river discontinuum. *Freshw. Biol.* **47**: 641–660. doi:10.1046/j.1365-2427.2002.00922.x
- R Core Team. 2017. R: A language and environment for statistical computing.
- Rahimi, M., H. I. Essaid, and J. T. Wilson. 2015. The role of dynamic surface water-groundwater exchange on streambed denitrification in a first-order, low-relief agricultural watershed. *Water Resour. Res.* **51**: n/a-n/a. doi:10.1002/2014WR016739
- Rheaume, S. J. 1991. Hydrologic provinces of Michigan.
- Robertson, G. P., and S. K. Hamilton. 2015. Long-Term Ecological Research at the Kellogg Biological Station LTER Site Conceptual and Experimental Framework.
- Rosi-Marshall, E. J., K. L. Vallis, C. V. Baxter, and J. M. Davis. 2016. Retesting a prediction of the River Continuum Concept: autochthonous versus allochthonous resources in the diets of invertebrates. *Freshw. Sci.* **35**: 534–543. doi:10.1086/686302
- Ruhala, S. S., J. P. Zarnetske, D. T. Long, J. A. Lee-Cullin, S. Plont, and E. R. Wiewiora. 2018.

- Exploring dissolved organic carbon cycling at the stream–groundwater interface across a third-order, lowland stream network. *Biogeochemistry* **137**. doi:10.1007/s10533-017-0404-z
- Runkel, R. L., D. M. McKnight, and E. D. Andrews. 1998. Analysis of transient storage subject to unsteady flow : diel flow variation in an Antarctic stream. *J. Nor* **17**: 143–154. doi:10.2307/1467958
- Schilling, K. E., and R. D. Libra. 2000. The Relationship of Nitrate Concentrations in Streams to Row Crop Land Use in Iowa. *J. Environ. Qual.* **29**: 1846. doi:10.2134/jeq2000.00472425002900060016x
- Schindler, J. E., and D. P. Krabbenhoft. 1998. The hyporheic zone as a source of dissolved organic carbon and carbon gases to a temperate forested stream. 157–174.
- Scott, D. T., D. M. McKnight, E. L. Blunt-Harris, S. E. Kolesar, and D. R. Lovley. 1998. Quinone Moieties Act as Electron Acceptors in the Reduction of Humic Substances by Humics-Reducing Microorganisms. doi:10.1021/ES980272Q
- Sear, D. A., P. D. Armitage, and F. H. Dawson. 1999. Groundwater dominated rivers. *Hydrol. Process.* **13**: 255–276. doi:10.1002/(SICI)1099-1085(19990228)13:3<255::AID-HYP737>3.0.CO;2-Y
- Sleighter, R. L., R. M. Cory, L. A. Kaplan, H. A. N. Abdulla, and P. G. Hatcher. 2014. A coupled geochemical and biogeochemical approach to characterize the bioreactivity of dissolved organic matter from a headwater stream. *J. Geophys. Res. Biogeosciences* **119**: 1520–1537. doi:10.1002/2013JG002600
- Sobczak, W. V., and S. Findlay. 2002. Variation in Bioavailability of Dissolved Organic Carbon among Stream Hyporheic Flowpaths. *Ecology* **83**: 3194–3209.
- Soranno, P. A., S. L. Hubler, S. R. Carpenter, and R. C. Lathrop. 1996. Phosphorus Loads to Surface Waters: A Simple Model to Account for Spatial Pattern of Land Use. *Ecol. Appl.* **6**: 865–878. doi:10.2307/2269490
- Stanford, J. A. 1998. Rivers in the landscape: introduction to the special issue on riparian and groundwater ecology. *Freshw. Biol.* **40**: 402–406. doi:10.1046/j.1365-2427.1998.00398.x
- Stedmon, C. A., and S. Markager. 2005. Resolving the variability in dissolved organic matter fluorescence in a temperate estuary and its catchment using PARAFAC analysis. *Limnol. Oceanogr.* **50**: 686–697. doi:10.4319/lo.2005.50.2.0686
- Stewart, M. K., U. Morgenstern, and J. J. McDonnell. 2010. Truncation of stream residence time: How the use of stable isotopes has skewed our concept of streamwater age and origin. *Hydrol. Process.* **24**: 1646–1659. doi:10.1002/hyp.7576
- Strayer, D. L., R. E. Beighley, L. C. Thompson, S. Brooks, C. Nilsson, G. Pinay, and R. J. Naiman. 2003. Effects of Land Cover on Stream Ecosystems: Roles of Empirical Models and Scaling Issues. *Ecosystems* **6**: 407–423. doi:10.1007/PL00021506

- Stream Solute Workshop. 1990. Concepts and methods for assessing solute dynamics in stream ecosystems. *J. North Am. Benthol. Soc.* **13**: 15–33. doi:10.2307/1467445
- Temnerud, J., and K. Bishop. 2005. Spatial Variation of Streamwater Chemistry in Two Swedish Boreal Catchments: Implications for Environmental Assessment. doi:10.1021/ES040045Q
- Thomas, Z., and B. W. Abbott. 2018. Hedgerows reduce nitrate flux at hillslope and catchment scales via root uptake and secondary effects. *J. Contam. Hydrol.* **215**: 51–61. doi:10.1016/J.JCONHYD.2018.07.002
- Thorp, J. H., and R. E. Bowes. 2017. Carbon Sources in Riverine Food Webs: New Evidence from Amino Acid Isotope Techniques. *Ecosystems* **20**: 1029–1041. doi:10.1007/s10021-016-0091-y
- Tonina, D., F. P. J. de Barros, A. Marzadri, and A. Bellin. 2016. Does streambed heterogeneity matter for hyporheic residence time distribution in sand-bedded streams? *Adv. Water Resour.* **96**: 120–126. doi:10.1016/j.advwatres.2016.07.009
- Triska, F. J., J. H. Duff, and R. J. Avanzino. 1993. The role of water exchange between a stream channel and its hyporheic zone in nitrogen cycling at the terrestrial-aquatic interface. *Hydrobiologia* **251**: 167–184. doi:10.1007/BF00007177
- Turner, M. G., and R. H. Gardner. 2015. Introduction to Landscape Ecology and Scale, p. 1–32. *In* *Landscape Ecology in Theory and Practice*. Springer New York.
- U.S. Geological Survey. 2014. NLCD 2011 Land Cover (2011 Edition, amended 2014) - National Geospatial Data Asset (NGDA) Land Use Land Cover. Downloadable data.
- Valett, H. M., J. a. Morrice, C. N. Dahm, and M. E. Campana. 1996. Parent lithology, surface-groundwater exchange, and nitrate retention in headwater streams. *Limnol. Oceanogr.* **41**: 333–345. doi:10.4319/lo.1996.41.2.0333
- Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The River Continuum Concept. *Can. J. Fish. Aquat. Sci.* **37**: 130–137. doi:10.1139/f80-017
- Vinson, M. R., and C. P. Hawkins. 1998. Biodiversity of Stream Insects: Variation at Local, Basin, and Regional Scales. *Annu. Rev. Entomol.* **43**: 271–293. doi:10.1146/annurev.ento.43.1.271
- Volk, C. J., C. B. Volk, and L. A. Kaplan. 1997. Chemical composition of biodegradable dissolved organic matter in streamwater. *Limnol. Oceanogr.* **42**: 39–44. doi:10.4319/lo.1997.42.1.0039
- Wagner, S., and R. Jaffé. 2015. Effect of photodegradation on molecular size distribution and quality of dissolved black carbon. *Org. Geochem.* **86**: 1–4. doi:10.1016/J.ORGGEOCHEM.2015.05.005
- Walker, C. M., R. S. King, D. F. Whigham, and S. J. Baird. 2012. Landscape and Wetland

- Influences on Headwater Stream Chemistry in the Kenai Lowlands, Alaska. *Wetlands* **32**: 301–310. doi:10.1007/s13157-011-0260-x
- Wang, S.Y., E. B. Sudduth, M. D. Wallenstein, J. P. Wright, and E. S. Bernhardt. 2011. Watershed Urbanization Alters the Composition and Function of Stream Bacterial Communities J.A. Gilbert [ed.]. *PLoS One* **6**: e22972. doi:10.1371/journal.pone.0022972
- Ward, A. S. 2016. The evolution and state of interdisciplinary hyporheic research. *Wiley Interdiscip. Rev. Water* **3**: 83–103. doi:10.1002/wat2.1120
- Ward, A. S., M. N. Gooseff, and K. Singha. 2010. Imaging hyporheic zone solute transport using electrical resistivity. *Hydrol. Process.* **24**: 948–953. doi:10.1002/hyp.7672
- Ward, A. S., C. A. Kelleher, S. J. K. Mason, T. Wagener, N. McIntyre, B. McGlynn, R. L. Runkel, and R. A. Payn. 2017. A software tool to assess uncertainty in transient-storage model parameters using Monte Carlo simulations. *Freshw. Sci.* **36**: 195–217. doi:10.1086/690444
- Weishaar, J. L., G. R. Aiken, B. A. Bergamaschi, M. S. Fram, R. Fujii, and K. Mopper. 2003. Evaluation of Specific Ultraviolet Absorbance as an Indicator of the Chemical Composition and Reactivity of Dissolved Organic Carbon. *Environ. Sci. Technol.* **37**: 4702–4708. doi:10.1021/es030360x
- Wickland, K. P., J. C. Neff, and G. R. Aiken. 2007. Dissolved Organic Carbon in Alaskan Boreal Forest: Sources, Chemical Characteristics, and Biodegradability. *Ecosystems* **10**: 1323–1340. doi:10.1007/s10021-007-9101-4
- Wilcoxon, F. 1945. Individual Comparisons by Ranking Methods. *Biometrics Bull.* **1**: 80–83.
- Wilson, H. F., and M. A. Xenopoulos. 2009. Effects of agricultural land use on the composition of fluvial dissolved organic matter. *Nat. Geosci.* **2**: 37–41. doi:10.1038/ngeo391
- Woessner, W. W. 2000. Stream and Fluvial Plain Ground Water Interactions: Rescaling Hydrogeologic Thought. *Ground Water* **38**: 423–429. doi:10.1111/j.1745-6584.2000.tb00228.x
- Wolock, D. M., J. Fan, and G. B. Lawrence. 1997. Effects of basin size on low-flow stream chemistry and subsurface contact time in the Neversink River Watershed, New York. *Hydrol. Process* **11**: 1273–1286.
- Wörman, A., and P. Wachniew. 2007. Reach scale and evaluation methods as limitations for transient storage properties in streams and rivers. *Water Resour. Res.* **43**: n/a-n/a. doi:10.1029/2006WR005808
- Wu, F. C., R. D. Evans, and P. J. Dillon. 2003. Separation and Characterization of NOM by High-Performance Liquid Chromatography and On-Line Three-Dimensional Excitation Emission Matrix Fluorescence Detection. doi:10.1021/ES020244E

- Wu, F. C., D. N. Kothawala, R. D. Evans, P. J. Dillon, and Y. R. Cai. 2007. Relationships between DOC concentration, molecular size and fluorescence properties of DOM in a stream. *Appl. Geochemistry* **22**: 1659–1667. doi:10.1016/J.APGEOCHEM.2007.03.024
- Xiao, H., and W. Ji. 2007. Relating landscape characteristics to non-point source pollution in mine waste-located watersheds using geospatial techniques. *J. Environ. Manage.* **82**: 111–119. doi:10.1016/J.JENVMAN.2005.12.009
- Yamashita, Y., and E. Tanoue. 2003. Chemical characterization of protein-like fluorophores in DOM in relation to aromatic amino acids. *Mar. Chem.* **82**: 255–271. doi:10.1016/S0304-4203(03)00073-2
- Yates, C. A., P. J. Johnes, and R. G. M. Spencer. 2016. Assessing the drivers of dissolved organic matter export from two contrasting lowland catchments, U.K. *Sci. Total Environ.* **569–570**: 1330–1340. doi:10.1016/J.SCITOTENV.2016.06.211
- Zarnetske, J. P., R. Haggerty, S. M. Wondzell, and M. A. Baker. 2011a. Dynamics of nitrate production and removal as a function of residence time in the hyporheic zone. *J. Geophys. Res.* **116**: G01025. doi:10.1029/2010JG001356
- Zarnetske, J. P., R. Haggerty, S. M. Wondzell, and M. A. Baker. 2011b. Labile dissolved organic carbon supply limits hyporheic denitrification. *J. Geophys. Res.* **116**: G04036. doi:10.1029/2011JG001730
- Zarnetske, J. P., R. Haggerty, S. M. Wondzell, V. A. Bokil, and R. González-Pinzón. 2012. Coupled transport and reaction kinetics control the nitrate source-sink function of hyporheic zones. *Water Resour. Res.* **48**: n/a-n/a. doi:10.1029/2012WR011894
- Zimmer, M. A., S. W. Bailey, K. J. McGuire, and T. D. Bullen. 2013. Fine scale variations of surface water chemistry in an ephemeral to perennial drainage network. *Hydrol. Process.* **27**: 3438–3451. doi:10.1002/hyp.9449
- Zuur, A. F., E. N. Ieno, and C. S. Elphick. 2010. A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* **1**: 3–14. doi:10.1111/j.2041-210X.2009.00001.x