

Biologic and Hydrologic Controls of Water Quality in  
Urbanizing Semi-Arid Watersheds

Erin Fleming Jones

A dissertation submitted to the faculty of  
Brigham Young University  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy

Zachary T. Aanderud, Chair  
Benjamin W. Abbott  
Michelle A. Baker  
Gregory T. Carling  
Neil C. Hansen

Department of Plant and Wildlife Sciences

Brigham Young University

Copyright © 2019 Erin Fleming Jones

All Rights Reserved

## ABSTRACT

### Biologic and Hydrologic Controls of Water Quality in Urbanizing Semi-Arid watersheds

Erin Fleming Jones

Department of Plant and Wildlife Sciences, BYU

Doctor of Philosophy

This dissertation analyzed the effect of biologic and hydrologic processes on water quality in urban, semi-arid watersheds. In the first chapter, we analyzed bacterioplankton and water quality along elevation and urbanization gradients in three Wasatch Mountain watersheds across three seasons. We found that trace metals correlated with bacterioplankton composition and that the typical dispersal of bacteria from headwater sources (soil or groundwater) along the longitudinal pathway was drastically disrupted by the presence of large reservoirs. In the second chapter, we used high-frequency sensor data collected in streams above and below the urban center in the three watersheds to estimate the relative contribution of biologic, hydrologic, and anthropogenic processes to changes in nitrate concentration. In-stream metabolism correlated with less than 38% of diel fluctuations in nitrate, but diel nitrate concentration only represented 10% of the total nitrate variability, demonstrating how in-stream uptake can easily be overwhelmed by nutrient loading in even moderately modified watersheds. A majority of the nitrate was associated with hydrologic variables, specifically discharge and specific conductivity, with pulses of nitrate corresponding to anthropogenic activity that far exceeded the capability of the system to remove or process the nitrogen. In the third chapter, we used citizen science to collect synoptic solute data to analyze the catchment hydrology in one of the Wasatch watersheds (Provo River and Utah Lake). Unlike previous research from humid and temperate catchments, we did not observe a systematic decrease in spatial variability with watershed size in this semi-arid, endorheic basin. Our results demonstrate the value of combining participatory science with modern ecohydrological methods to determine catchment chemistry and hydrology. This dissertation shows how hydrology, and anthropogenic changes to watersheds that affect hydrology, are largely responsible for determining water quality in urbanizing, semi-arid watersheds.

Keywords: Water quality, catchment hydrology, microbial ecology, stream biogeochemistry, urban, semi-arid

## ACKNOWLEDGEMENTS

They say that it takes a village to raise a child. This dissertation is like a second baby to me, and it certainly took a village to produce. My committee have been an enormous help and support to me during my time at BYU. A huge thank you to many iUTAH technicians and research assistants (Dylan, Dave, Joe, Julie, Amber, Ka-voka, Chris, and Beth), including co-authors listed within each of the chapters. My fellow Aanderud lab mates (Natasha, Scott, Tayte, David, Tylan, Jason, Jordan, Eric, Amber, Jonathon, Madeline, Madeleine, and Gabbi) also deserve acknowledging for their help and support. This research also benefited from collaboration with other labs in the PWS department, including the Geary Lab (Madsen and Trevor) and Abbott Lab (Becca, Trevor, Rachel, Rhetta, Sam and Adam). This work could not have been done without funding from iUTAH NSF EPSCoR project, BYU Graduate Studies, and the Provo River Watershed Council.

Finally, I would like to thank my family: my mom and sister who encouraged me and helped watch/raise my son; my husband for his extreme dedication to supporting me; and for David. To David (and any of my future children): This is for you.

## TABLE OF CONTENTS

|  |      |
|--|------|
| TITLE PAGE .....   | i    |
| ABSTRACT .....   | ii   |
| ACKNOWLEDGEMENTS .....   | iii  |
| TABLE OF CONTENTS.....   | iv   |
| LIST OF FIGURES .....  | viii |
| LIST OF TABLES.....  | xi   |
| CHAPTER 1 .....  | 1    |
| ABSTRACT .....   | 1    |
| INTRODUCTION .....   | 2    |
| METHODS .....  | 7    |
| Study Sites.....   | 7    |
| Bacterioplankton Community Composition.....  | 8    |
| Environmental Factors .....  | 9    |
| Statistical Analyses .....   | 12   |
| RESULTS.....   | 15   |
| Environmental Drivers of Bacterioplankton Communities.....                             | 15   |
| Geographic and Dispersal-Based Influences on Stream Bacterioplankton Communities ..... | 16   |
| Human Influence on Bacterioplankton Communities .....                                  | 18   |
| DISCUSSION.....  | 20   |

|  |    |
|--|----|
| Environmental Drivers of Bacterioplankton Communities.....                             | 20 |
| Geographic and Dispersal-Based Influences on Stream Bacterioplankton Communities ..... | 21 |
| Human Influence on Bacterioplankton Communities .....                                  | 23 |
| <br>CONCLUSION .....   | 25 |
| <br>LITERATURE CITED.....  | 27 |
| <br>FIGURES.....   | 44 |
| <br>TABLES .....   | 52 |
| <br>SUPPLEMENTAL MATERIAL .....  | 56 |
| <br>CHAPTER 2 .....  | 63 |
| ABSTRACT .....   | 63 |
| INTRODUCTION .....   | 64 |
| METHODS .....  | 66 |
| Study Area.....  | 66 |
| Data Collection and Analysis.....  | 68 |
| RESULTS .....  | 70 |
| Time Series Data .....   | 70 |
| N v N .....  | 72 |
| Diel NO <sub>3</sub> <sup>-</sup> .....  | 73 |
| Metabolism and Nitrate.....  | 74 |
| Hydrology and Nitrate.....   | 75 |
| DISCUSSION.....  | 76 |

|   |     |
|---|-----|
| When and How Much are These Reaches Gaining or Losing Nitrate? .....    | 76  |
| How Much of the Change in Nitrate is Biological v Hydrologic?.....      | 77  |
| How Much Nitrate May be Anthropogenic?.....                             | 78  |
| CONCLUSION .....  | 79  |
| LITERATURE CITED.....   | 81  |
| FIGURES.....  | 88  |
| TABLES .....  | 95  |
| SUPPLEMENTAL MATERIAL .....   | 99  |
| CHAPTER 3 .....   | 110 |
| ABSTRACT .....  | 110 |
| INTRODUCTION .....  | 111 |
| METHODS .....   | 114 |
| Synoptic Sampling .....   | 114 |
| Site Description and Selection .....                                    | 114 |
| Citizen Science.....  | 116 |
| Laboratory and Statistical Analysis.....                                | 116 |
| RESULTS.....  | 117 |
| Solute Concentrations .....   | 118 |
| Leverage and Spatial Stability.....                                     | 119 |
| DISCUSSION.....   | 119 |
| Utah Lake Watershed has Novel Spatial and Temporal Hydrochemistry ..... | 119 |

|  |     |
|--|-----|
| Urban Sources Affect all the Solutes Except DOC..... | 120 |
| Citizen Science.....                                 | 121 |
| CONCLUSION .....                                     | 122 |
| LITERATURE CITED.....                                | 123 |
| FIGURES.....   | 128 |
| TABLES .....   | 133 |
| SUPPLEMENTAL MATERIAL .....                          | 134 |

## LIST OF FIGURES

- Figure 1-1. Conceptual diagram of the proposed relationships between hydrologic conditions and bacterioplankton communities. Residence time, a hydrologic condition related to flow velocity and volume, influences the extent that dispersal and stochastic processes alter community composition. Residence time also controls the extent to which environmental pressures act on communities to create species sorting, or selection. Bacterial metabolic activity affects water chemistry, creating a feedback loop within the model (hypothesis 1, purple). The community composition similarity along and between watersheds is related to longitudinal and lateral dispersal (hypothesis 2, orange). Anthropogenic changes to watersheds affect hydrologic conditions, and with repercussions on all aspects of the system (hypothesis 3, brown). ..... 44
- Figure 1-2. Map showing Logan (green), Red Butte (red) and Provo (blue) Watersheds in the Wasatch Range Metropolitan Area and locations within watersheds where bacterial communities and water quality data were collected. Location markers indicate position relative to man-made reservoirs and urban centers in all three watersheds: A= Above, D=below Dam, U= Urban. Site metadata is included in Table 1-1. Sources: ESRI, USGS, NOAA..... 45
- Figure 1-3. Vectors are significant environmental factors, indicating positive correlation. Location indicates position relative to man-made reservoirs and urban centers. Variables were selected from the combination of four RDA models of standard field parameters, nutrients, major ions, and trace elements (Table 1-3). ..... 46
- Figure 1-4. Longitudinal profiles of stream physiochemical variables selected by backwards stepwise RDA ordination of bacterial communities in streams in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Location indicates position relative to reservoirs and urban centers. Means and standard errors, where determinable, are shown (some data points were removed for QA/QC violations). ..... 47
- Figure 1-5. Reservoirs alter stream bacterial community composition from three watersheds in Utah's Wasatch Range Metropolitan Area (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Graph represents a Principal Coordinate Analysis (PCOA) ordination, with each point representing a community of bacterioplankton classified at the 97% similarity from 16s rRNA gene amplicon sequencing of OTUs. Location indicates position relative to man-made reservoirs and urban centers. Points closer together are more similar (based on Bray-Curtis dissimilarity index), while points farther apart are dissimilar..... 48
- Figure 1-6. Observed richness and Shannon diversity of bacterial communities for three watersheds over three seasons in Wasatch Range Metropolitan Area (WRMA). Boxes depict interquartile ranges, with the center line on the median and whiskers showing the extent of values. ..... 49
- Figure 1-7. Network co-occurrence models for bacterial communities collected along elevation and urbanization gradients in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Nodes indicate taxa (OTUs, taxa with 97% similar sequences) and edges connect where significant co-occurrence was detected in 75% of samples. Topologic statistics of the network are shown in Table 1-6..... 50
- Figure 1-8. Heatmap showing ANCOM results of differentially expressed family for sites in three watersheds across three seasons (Fall, Winter, and Spring). Taxa were selected using a

|  |     |
|--|-----|
| family relative abundance at least 0.5% for all samples. Location indicates position within watershed relative to man-made reservoirs, with Dam sites located immediately downstream of reservoir outlets (Table 1-1). Darker squares indicate higher relative abundance for bacterial families in each sample. Similarity between samples and taxa is indicated by dendrograms on each axis.....  | 51  |
| Figure 2-1. Theoretical diagram of possible nitrate concentration relationship between upstream and downstream sites. Along the 1:1 line represents net conservative transport, with deviations in some range of net N addition or N removal due to: biological activity, such as N <sub>2</sub> fixation or denitrification; natural hydrologic processes, such as groundwater or overland sources, or dilution; and anthropogenic influence. ....    | 88  |
| Figure 2-2. Map showing study reaches with Upstream (U) and Downstream (D) sensor locations in three Wasatch watersheds (Logan, Red Butte, and middle Provo). Source credits: Esri, USGS, and NOAA. ....   | 89  |
| Figure 2-3. Time series for one year of sensor data from upstream (grey) and downstream (black) sites in Logan (A), Red Butte (B), and middle Provo (C) watersheds. Parameters measured include temperature, dissolved oxygen (DO), discharge, NO <sub>3</sub> <sup>-</sup> , specific conductivity (Sp. Cond.), fluorescent dissolved organic matter (fDOM), and turbidity. Summary statistics of values by hydroperiod can be found in Table S1..... | 90  |
| Figure 2-4. Nitrate concentrations at upstream versus downstream sites for three Wasatch watersheds (Logan, A; Red Butte, B; Provo, C) demonstrate periods of net N-addition and removal across periods of the hydrograph (Spring Runoff, Summer growing period, and Winter low flow).....   | 91  |
| Figure 2-5. Time series of diel ΔNO <sub>3</sub> <sup>-</sup> values at upstream (Up) and downstream (Down) sites in Logan, Red Butte, and Provo watersheds. Diel NO <sub>3</sub> <sup>-</sup> values were calculated by subtracting the average concentration during the 1400 hour from the 0200 hour; positive values indicate net daytime NO <sub>3</sub> <sup>-</sup> assimilation.....  | 92  |
| Figure 2-6. One year of stream metabolism (i.e. gross primary productivity, GPP, and ecosystem respiration, ER) for three Wasatch stream reaches (Logan, Red Butte, and middle Provo). ..  | 93  |
| Figure 2-7. Normalized concentration-discharge (C-Q) relationships for NO <sub>3</sub> <sup>-</sup> in three Wasatch watersheds (Logan, A; Red Butte, B; and middle Provo, C) showing seasonal differences in hysteresis behavior.....   | 94  |
| Figure 3-1. Theoretical diagram showing existing concept of spatial variability collapse (left) and the pattern we observed in the Utah Lake watershed, with spatial variability expanse (right)..   | 128 |
| Figure 3-2. Map of Utah Lake watershed sites synoptically sampled, colored by land use and hydrologic modification category (red= Agricultural unregulated, green= Lake tributary, blue= Mixed dammed, purple= Mountain urban). Yellow triangles represent wastewater treatment plants.....  | 128 |
| Figure 3-3. Scaled concentration of solutes in surface water samples collected during three synoptic sampling events (Fall, Spring, Summer) of areas within the Utah Lake watershed with different characteristics (Agricultural unregulated, Lake tributary, Mixed dammed, and Mountain urban). Boxplots represent the 25 <sup>th</sup> , 50 <sup>th</sup> , and 70 <sup>th</sup> percentiles, points within 1.5 times the                            |     |

interquartile range, and points beyond. The notches represent the 95% confidence interval of the median (non-overlapping notches suggest differences between populations)..... 129

Figure 3-4. Maps showing average concentration of solutes from three synoptic sampling events across the Utah Lake watershed (outlined in black). Point color represents numeric water quality standards for N and P values, or 25<sup>th</sup> and 75<sup>th</sup> percentile for others. Point size is scaled to concentration..... 130

Figure 3-5. Scaled solute concentration (left) and leverage (right) by catchment area for solutes of interest (dissolved organic carbon, phosphorus, total nitrogen, dissolved inorganic nitrogen, chloride, and sulfate) for sites within four land use and hydrologic categories of the Utah Lake watershed. Samples were collected using participatory science volunteers on three synoptic sampling events conducted in March (Spring), July (Summer), and October (Fall) of 2018.... 131

Figure 3-6. Spatial stability for solutes (dissolved organic carbon, phosphate, total nitrogen, dissolved inorganic nitrogen, chloride, and sulfate) for different categories of land use within the Utah Lake watershed. Spatial stability is calculated as a Spearman rank correlation by comparing multiple synoptic samplings (Spring, Summer, and Fall 2018) pairwise by site. Points represent mean Spearman, and error bars represent the range of values..... 132

## LIST OF TABLES

|   |    |
|---|----|
| Table 1-1. Stream site characteristics in the Logan, Red Butte, and Provo River Watersheds (Figure 1-2). Location indicates site position relative to reservoir(s) in the watershed, with Dam referring to sites immediately downstream of reservoir outflow (Table 1-2). Percent developed was calculated for the entire upstream watershed using NLCD 2011 data, not including “developed open space” (green space within urban areas). Mean discharge (and standard deviation) was calculated based on the 2014 water year, which is when sampling occurred. ....  | 52 |
| Table 1-2. Physical specifications and discharge of the reservoirs and dams within the three study reaches. The Logan Watershed has a series of multiple impoundments with the lowest, First Dam, being the largest. Using maximum capacity and discharge at gages located immediately downstream, we calculated annual residence time. ....  | 53 |
| Table 1-3. Redundancy analysis (RDA) model results indicating the stream physiochemical variables structuring bacterial communities in streams across three watersheds (Logan, Red Butte, and Provo) and three seasons (Fall, Winter, and Spring). Environmental variables were separated into five categories and analyzed separately using backwards stepwise selection of AIC scores to identify significant components ( $p<0.1$ ). Variables from the best-fit models (excluding number five, which used fewer samples than other models) were combined and reported as an overall model. We report summary statistics, including adjusted $R^2$ , RDA axis percents, and proportion of community variation explained by constrained model (CP). Visualizations of the ordinations are included in supplemental figures 1-5. All models were significant ( $p$ -value=0.001). .... | 54 |
| Table 1-4. Pairwise PERMANOVA results comparing bacterioplankton community composition across three watersheds, at five locations, and three seasons. Higher R <sup>2</sup> values suggest more difference in community between two groups; smaller numbers are more similar in composition. * indicates $p$ -value<0.05. ....  | 55 |
| Table 1-5. Topological metrics calculated for network co-occurrence models of bacterial communities collected at five locations along an elevational gradient in three watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Nodes indicate taxa (OTUs) and edges connect where significant co-occurrence was detected in 75% of samples, edges represent significant co-occurrence connections that occur in at least 75% of samples in each watershed and have an MIC that is both > 0.7 and statistically significant. Mean path length: the average number of steps to connect each node; Mean degree: average number of edges for each node; Mean clustering coefficient: a measure of how completely neighboring nodes are connected to each other. ....  | 55 |
| Table 2-1. Site metadata for upstream (Up) and downstream (Down) locations of reaches within three Wasatch watersheds (Logan, Red Butte, and middle Provo). Stream length indicates distance from the upstream site. Distance integrated was calculated using velocity and K600 values estimated by streamMetabolizer or measured using SF6 injection. ....   | 95 |
| Table 2-2. Summary statistics for NO <sub>3</sub> <sup>-</sup> concentrations for upstream and downstream sites in three Wasatch watersheds (Logan, Red Butte, and Provo). ....   | 96 |
| Table 2-3. GLS models correlating diel change in NO <sub>3</sub> <sup>-</sup> concentration using daily gross primary productivity (GPP), daily ecosystem respiration (ER), and daily average upstream nitrate  |    |

|  |     |
|--|-----|
| (UpNitrate). Parameters listed show terms selected in the best model by AICc score, *denotes the term had a significant p-value. We list model R2 and phi1 (the autocorrelation term). .....   | 97  |
| Table 2-4. GLS models predicting NO <sub>3</sub> – concentration in three reaches for three hydroperiods using daily average water temperature (temp), sum of daily precipitation (precip), daily average chlorophyll α (chla), daily average specific conductivity (SpCond), upstream nitrate (UpNitrate), and daily average discharge (Q). Parameters listed show terms selected in the best model by AICc score, *denotes the term had a significant p-value. We list model R2 and phi1 (the autocorrelation term). ..... | 98  |
| Table 3-1. Watershed characteristics for contributing areas to Utah Lake, calculated using USGS StreamStats. Discharge data represents annual discharge from 1980-2003 (PSOMAS 2007). Land use was calculated using 2011 NLCD (% forest, % developed, % impervious surface) and 1992 NLCD (% herbaceous upland). Where categories represent multiple subwatersheds, statistics for the major contributors are given. ....  | 133 |
| Table 3-2. Linear regressions of solute concentration (dissolved organic carbon, phosphorus, total nitrogen, dissolved inorganic nitrogen, chloride, and sulfate) by land use (% impervious, % developed, % forest, % herbaceous upland) and Event (Spring, Summer, and Fall). Model selection was done based on AICc scores of REML models. ....  | 133 |

## CHAPTER 1

### Stream Microbial Community Structured by Trace Elements, Headwater Dispersal, and Large Reservoirs in Sub-Alpine and Urban Ecosystems

Erin Fleming Jones, Natasha A. Griffin, Julia E. Kelso, Gregory T. Carling, Michelle A. Baker,  
Zachary T. Aanderud

Department of Plant and Wildlife Sciences, Brigham Young University, Provo, UT  
Doctor of Philosophy

#### ABSTRACT

Stream bacterioplankton communities, a crucial component of aquatic ecosystems and surface water quality, are shaped by environmental selection (i.e., changes in taxa abundance associated with more or less favorable abiotic conditions) and passive dispersal (i.e., organisms' abundance and distribution is a function of the movement of the water). These processes are a function of hydrologic conditions such as residence time and water chemistry, which are mediated by human infrastructure. To quantify the role of environmental conditions, dispersal mechanisms, and human infrastructure (dams) on stream bacterioplankton, we measured bacterioplankton community composition in rivers from sub-alpine to urban environments in three watersheds (Utah, USA) across three seasons. Of the 43 environmental parameters measured (including physicochemical parameters, solute concentrations, and catchment characteristics), trace element concentrations explained the most variability in bacterioplankton community composition using Redundancy Analysis (RDA) ordination. Trace elements may correlate with bacterioplankton due to the commonality in source of water and microorganisms, and/or environmental selection creating more or less favorable conditions for bacteria. Patterns in community composition (i.e., consistent groupings within watersheds regardless of season), suggested bacteria entered the community predominantly through dispersal from headwater soils (lateral dispersal) and then were transported along the longitudinal gradient (longitudinal

dispersal), except in cases where large reservoirs drastically changed water residence time. Upslope anthropogenic development had no detectable effect on community composition, but large reservoirs (where water residence time increased orders of magnitude) had marked effects on diversity, community similarity (Bray-Curtis distance), and co-occurrence networks. Communities downstream of reservoirs were enriched with anaerobic Sporichthyaceae, methanotrophic Methylococcaceae, and iron-transforming Acidimicrobiales, suggesting alternative metabolic pathways became active in the hypolimnion of large reservoirs. Our results identify that human activity affects river microbial communities, with potential impacts to water quality through modified biogeochemical cycling.

## INTRODUCTION

Bacterioplankton, the portion of stream microbial communities suspended within the water column, are a crucial component of water quality, but are often treated as a black box in aquatic environments (Allison and Martiny, 2008). As land and water use alter biogeochemical fluxes and hydrological characteristics of aquatic ecosystems (Abbott et al., 2018; Blaszcak et al., 2019), understanding the drivers of microbial community composition is critical to protecting and restoring freshwater systems (Lindström and Östman, 2011). Stream microbial communities are shaped by two interacting processes: environmental selection (Fierer et al., 2007; Fierer and Lennon, 2011; Zwart et al., 2002), and dispersal (Albright and Martiny, 2018; Crump et al., 2007, 2012; Findlay, 2010; Savio et al., 2015). Environmental selection, or species sorting, is the process of the abundance of taxa gradually shifting based on advantage or disadvantages based on environmental conditions. Dispersal, the movement of organisms from one environment or

location into another, is particularly relevant for bacterioplankton, because the organisms are passively transported with the constant movement of water.

The most important environmental factors for structuring bacterial communities are pH, salinity, temperature, and dissolved oxygen, because of their role in controlling, or being controlled by, cellular activity (Doherty et al., 2017; Fierer et al., 2007; Fierer and Jackson, 2006). Water residence time moderates the duration that bacterial communities are influenced by the environment (Abbott et al., 2016; Ben Maamar et al., 2015; Niño-García et al., 2016), potentially leading to feedback loops where bacteria engineer new conditions that select for an altered set of taxa (Figure 1-1). For example, longer residence times lead to anoxic conditions at and below the sediment-water interface as microbial decomposition exceeds reaeration rates (Baker and Valett, 2000; Zarnetske et al., 2011). Once anoxic conditions are created, alternative terminal electron acceptor pathways become activated, with many subsequent changes in water chemistry; for example, anoxic ecosystems switch from net bacterial nitrification to denitrification (Briggs et al., 2013; Kolbe et al., 2019; Oldham et al., 2013). However, dozens of environmental parameters, including physicochemical conditions and solute concentrations, affect the abundance of bacterial taxa due to selective pressure (i.e. species sorting). Trace elements, such as molybdenum (a cofactor in the enzyme nitrogenase) and rare earth elements like lanthanum, may stimulate growth at low concentrations, but may also be toxic at high concentrations (Herrmann et al., 2016; Smedley and Kinniburgh, 2017). Macronutrients and ions generally considered limiting for autotrophic organisms, such as nitrogen (N), phosphorus (P), and potassium (K), may be less important for explaining bacterial community composition, because bacterial metabolic pathways are able to metabolize even recalcitrant substrates into more bioavailable species (Zeglin, 2015). Organic matter quality and quantity in soils and

streams also correlate with specific bacterial communities and ultimately bacterial activity across a broad range of climates and bedrock materials (Gabor et al., 2014; Miller et al., 2016; Ruiz-González et al., 2015). Explaining bacterioplankton community taxonomic composition with instantaneous stream chemistry conditions is complicated by temporal changes in the source and flowpath of water entering a channel (Dahlke et al., 2012; Leff and Lemke, 1998; Moatar et al., 2017). Seasonal changes in hydrology and environmental parameters affect the concentration, form, and downstream availability of organic and inorganic matter necessary for, or inhibitory to, microbial function (Duff and Triska, 2000; Hendricks and White, 2000).

The effect of selective pressure on a community is constrained by which taxa are present through dispersal. Stream-lake networks are unique because of their dendritic nature, which creates specific patterns in biodiversity (Widder et al., 2014). Bacterioplankton dispersal in streams is constrained to unidirectional flow paths and, like stream water chemistry, is ultimately the combination of longitudinal (i.e. upstream to downstream in channel), vertical (groundwater-surface-water exchange), and lateral (i.e. stream-bank exchange and tributaries) contributions (Covino, 2017). Headwater inputs and longitudinal connectivity are crucial for maintaining downstream community composition and potentially biogeochemical function; alpha and beta diversity are highest in low-order streams, and decrease as stream order increases. (Besemer et al., 2013; Crump et al., 2007; Savio et al., 2015). River systems are naturally punctuated by lakes and other abrupt changes in the natural topography (Ward and Stanford, 1983); global estimates of river water volume are around 2,000 km<sup>3</sup>, compared to 8,000 km<sup>3</sup> in artificial reservoirs and 200,000 km<sup>3</sup> in lakes (Abbott et al., 2019; Messager et al., 2016; Shiklomanov, 1993). Natural streams are less of a gradual longitudinal continuum and more often a series of highly distinct environments, particularly when considering microhabitats of bacteria, but lakes and reservoirs

are rarely incorporated into studies of riverine bacterial communities (Adams et al., 2014; Ben Maamar et al., 2015; Lindström et al., 2006).

Discontinuity in modern streams comes from both natural and human sources as humans have increasingly altered landscapes and introduced infrastructure along rivers that impede the flow of water and associated matter (Grill et al., 2019; Grimm et al., 2008; Hale et al., 2015). Links among bacterial communities, stream characteristics, and dispersal must consider the influence that urbanization exerts on natural processes due to increased infrastructure development to support growing populations in nearly every aspect of streams and rivers (Meybeck, 2003). Reservoirs and regulated lakes (natural lakes modified with infrastructure to provide managers with control over flow and lake elevation) behave differently than natural lakes (Döll et al., 2009), so understanding their unique impact on stream bacterial communities is crucial. Land use changes, such as agriculture, mining, forestry, and urbanization, may result in less direct changes to stream conditions by impacting hydrologic connections between upland, groundwater, and stream ecosystems (Meybeck, 2003; Rose and Peters, 2001). An estimated 77% of global land area is affected directly by land use change; indirect anthropogenic effects increase that number to 100% (Vitousek et al., 1997; Watson et al., 2018). Disturbance of the land surface affects water flowpath and chemistry, altering the microbial community in soils, aquifers, and surface waters (Covino, 2017; Moatar et al., 2017). Without more comprehensive understanding of the controls on stream bacterioplankton, it is almost impossible to account for the complex interactions between land use change, altered hydrology, and water chemistry, and alterations in bacterial communities (Fisher et al., 2015; Hall et al., 2018; Van Rossum et al., 2015).

Northern Utah streams provide an ideal setting to understand longitudinal, anthropogenic, and seasonal impacts on bacterioplankton dispersal and the interactions between water chemistry and bacterial communities. Land uses in the region include agriculture, forestry, mining and urban development, with built infrastructure encompassing reservoirs, irrigation systems, cross-basin diversions and flood control to support human industries (Hall et al., 2015). To understand the effect of environmental conditions, dispersal mechanisms, and human infrastructure on stream bacterioplankton community composition, we collected bacterioplankton and a suite of environmental parameters from five locations along montane to urban gradients in three northern Utah watersheds, over three seasons (fall, spring, and winter). We hypothesized that: 1) bacterioplankton communities are structured by standard water quality metrics, (pH, temperature, specific conductivity, and dissolved oxygen), more than other groups of variables (nutrients, major ions, trace elements and organic matter) because of their importance in regulating essential metabolic processes (Fierer and Jackson, 2006; Lindström et al., 2005; Niño-García et al., 2016); 2) the dominant bacterioplankton dispersal in these streams occurs by initial lateral and subsequent longitudinal transport, because external sources with high density of organisms will overwhelm contributions from local sources such as groundwater and stream biofilms (i.e. mass effects; Lindström and Östman, 2011), and 3) urbanization of watersheds decreases bacterial biodiversity and community connectivity because of the consequent alteration, specifically homogenization, of hydrologic conditions and flowpath (Figure 1-1).

## METHODS

### *Study Sites*

Our project was conducted in three northern Utah, USA watersheds, selected as part of the iUTAH (innovative Urban Transitions and Aridregion Hydro-sustainability) project, funded by the NSF Established Program to Stimulate Competitive Research (EPSCoR) (Jones et al., 2017). The watersheds include Red Butte Creek, Logan River, and Provo River (Fig. 2). Within these watersheds, long-term stream sampling sites were selected from subalpine (up to 2368 m.a.s.l) elevations to low-elevation (down to 1353 m.a.s.l.) urban and agricultural land uses to capture the effects of elevation and urbanization on water resources (Table 1-1). Snowmelt-dependent streams in the region flow from undeveloped mountains into densely populated valleys. In the three study watersheds, this transition is demarcated by dams of varying sizes built to meet urban and agricultural water demand in the semi-arid climate of the basins below. Logan River passes through a series of smaller impoundments compared to the other watersheds, with shallow reservoirs and much lower residence times (Table 1-2). We included sites above and immediately downstream of the reservoirs to measure the effect of the introduced infrastructure on the bacterial community composition.

The degree and type of watershed development is variable within and among the three watersheds. The headwaters of each watershed are federal land and have some degree of protection from urbanization. Red Butte Creek has the most stringent protections; its headwaters are designated as a natural research preserve. The most impacted headwaters are in the Logan River watershed, which is opened to livestock grazing each summer (Hall 2005). Each of the three watersheds' valleys has a unique urban development type, allowing a comparison of a range of human activity. The Red Butte Creek watershed experienced urbanization beginning

two centuries ago with Fort Douglas, one of the oldest permanently maintained developments in Utah; the subsequent construction of the University of Utah campus and residential areas exclude any present-day agricultural land use (Ehleringer et al., 1992). The Logan River watershed has valley sites that are a mixture of agriculture with some urban land uses, as the slow population growth in Cache County has resulted in a gradual shift from an agriculture-dominant to urbanized landscape. The middle region of the Provo River is also shifting from agricultural land use to newer urban development and was designated the fastest growing area in the country in 2016 (4.7% annual increase in population, US Census Bureau).

### *Bacterioplankton Community Composition*

We designed a sampling regime to capture longitudinal gradients and the effect of seasonal changes in hydrology and environmental conditions on bacterioplankton communities. We collected water column samples at fifteen sites, including five locations (named for position relative to reservoirs) in each of the three study watersheds (Figure 1-2). We collected suspended bacteria in an attempt to collect a more comprehensive sample and eliminate possible differences due to cross-site variable streambed material, which influences biofilm establishment based on size fraction and mineral type (Donlan, 2002; McCormick et al., 2014). Streams were sampled in November 2014 (Fall), which is dominated by low-flow conditions and subsequently high residence times, with leaf litter potentially providing organic matter subsidies to inorganic groundwater contributions; in February 2015 (Winter) to capture low-flow, snow-covered conditions when flows and inputs are likely most homogenous within and across watersheds; and in May 2015 (Spring) to capture peak runoff conditions, when residence times are lowest and lateral connectivity is high.

We used a target metagenomic approach for identifying bacterioplankton community composition. In the field, we filtered water onto 47-mm 0.2- $\mu$ m PES Supor filters (Pall) using autoclaved filter cups (Nalgene) and stored filters in cryovials in liquid nitrogen to immediately suspend microbial activity (adapted from Somerville et al., 1989). We stored samples at -80 °C until extracting with PowerSoil DNA extraction kits according to manufacturer instructions (MOBIO). We PCR-amplified the V4 region of the bacterial 16S rRNA gene with primer set 515F and 806R (Caporaso et al., 2011). After checking that amplification proceeded normally using gel electrophoresis, we purified and normalized samples (SequalPrep Normalization Plate Kit, Invitrogen). Samples were submitted to the Brigham Young University DNA Sequencing Center (<http://dnac.byu.edu/>) for 2x250 base pair paired-end sequencing on an Illumina HiSeq 2500 System. We processed sequences using a modified Mothur pipeline (Schloss et al., 2009). We calculated the relative abundances of bacterial taxa at the 97% operational taxonomic unit (OTU) similarity cutoff. All community inferences were based on 43 samples with 138,458 total sequences rarefied to 15,445 sequences, and 1,450 unique OTUs with samples possessing an average sequencing coverage of  $90.1\% \pm 0.62$  (mean and standard error). Four samples were removed during QA/QC due to low read counts including: Provo Above2 Fall, Logan Above1 Winter, Logan Dam Winter, and Logan Below1 Winter. All bacterial sequences are available at <http://www.hydroshare.org/resource/48fc6871c51d436b83000a8d29ddb702>, and code used in Mothur can be downloaded at <https://github.com/erinfjones>.

### *Environmental Factors*

We quantified a wide range of environmental conditions concurrent with bacterial sampling to identify which parameters correlated to community changes (see 2.4.1). We measured

standard water quality parameters (pH, water temperature, dissolved oxygen, conductivity) using either a YSI Quatro multiparameter probe or YSI EXO2 sonde (data downloaded from iUTAH web services using the R package WaterML; Jones et al., 2017; Kadlec et al., 2015). We analyzed nutrients potentially related to bacterial activity, including total nitrogen (TN, persulfate oxidation digestion and cadmium reduction method), total phosphorus (TP, persulfate oxidation digestion and ascorbic acid method), nitrate (EPA 353.2), ammonia (EPA 350.1), and dissolved orthophosphate (EPA 365.1) colorimetrically on an autoanalyzer (Astoria-Pacific). Total and volatile suspended solids were determined from combustion of pre-ashed Glass Fiber Filters (GF/F, Whatman) at 450° C for 2 hours, and chlorophyll  $\alpha$  was analyzed using ethanol extraction of filters followed by analysis on a handheld Turner Aquaflor fluorometer (Sartory and Grobbelaar, 1984; Steinman et al., 2017). Dissolved organic carbon (DOC), determined by acidification or sparging of inorganic carbon followed by combustion catalytic oxidation and NDIR detection, and total dissolved nitrogen (TDN), using the catalytic thermal decomposition/chemiluminescence method, were determined using the Shimadzu TOC analyzer. We measured major anion concentrations (F, Cl, and SO<sub>4</sub>) on a Dionex ICS-90 ion chromatograph. Major cations (Ca, K, Mg, Na) and trace elements (Ag, Al, As, B, Ba, Be, Cd, Ce, Co, Cr, Cs, Cu, Eu, Fe, La, Li, Lu, Mn, Mo, Nd, Ni, Pb, Rb, Sb, Se, Sm, Sr, Tb, Ti, Tl, U, V, Y, and Zn), which potentially shape bacterial community structure (Zeglin, 2015), were measured using an Agilent 7500ce quadrupole inductively coupled plasma mass spectrometer (ICP-MS, Goodsell et al., 2017). Stable isotopes in water ( $\delta^{18}\text{O}$  and  $\delta\text{D}$ ), which we included with trace elements as a surrogate of water source (Follstad Shah et al., 2019), were measured on unfiltered aliquots using a Los Gatos Research Liquid Water Isotope Analyzer (LWIA-24d, Carling et al., 2015).

We characterized spectrofluorometric properties of dissolved organic matter (hereafter “organic matter”) from excitation emission matrices (EEMs) using a Horiba Aqualog spectrofluorometer (Horiba Scientific). EEMs were collected over excitation wavelengths 248-830 nm at 6 nm increments and over emissions 249.4-827.7 nm at 4.7 nm (8 pixel) increments. All samples were collected in ratio mode (S/R), and samples that exceeded 0.3 absorbance units at excitation 254 nm were diluted with deionized water. All samples were corrected for inner filter effects, Rayleigh scatter, and blank subtracted in MATLAB<sup>TM</sup> (version 6.9; MathWorks) as described in Murphy et al. (2013). We calculated six indices from the EEMs, including: the beta:alpha index (BIX), where higher values represent more microbially derived DOM (Huguet et al., 2009; Parlanti et al., 2000); humification index (HIX) with higher values representing more humic-like material (Zsolnay et al., 1999); fluorescence index (FI), a ratio of fulvic- vs humic-like organic matter (McKnight et al., 2001); TC index, the ratio of maximum fluorescence in the peak T region (protein-like) versus peak C region (humic-like), with higher values representing more protein-like organic matter, including WWTP effluent (Baker et al., 2008); SUVA<sub>254</sub>, specific ultraviolet absorbance at 254 nm, an indicator of aromaticity (Weishaar et al., 2003); and Total EEM intensity, which correlates to the concentration of organic carbon in the sample.

EEMS were also used to resolve a 4 component PARAFAC model following protocols outlined in (Murphy et al., 2013). PARAFAC was used to identify humic and protein-like fluorescent components of DOM to elucidate differences in DOM that varied by watershed and location within a watershed. The drEEM toolbox was used to create a PARAFAC model in MATLAB<sup>TM</sup> following Murphy et al. (2013). Resolved PARAFAC components were then compared to previously found fluorophores in the open source library OpenFluor (Murphy et al.

2014). A total of 499 EEMs, collected as part of a previous synoptic sampling effort from July 2014 to December 2015, were used to create the PARAFAC model. Component 1 represented percent organic matter within component 1 (humic-like, developed), component 2 (humic-like, forested), component 3 (protein-like, tryptophan-like, developed), and component 4 (protein-like, tyrosine and tryptophan, forested); and %Protein, the percent protein-like organic matter as indicated by the sum of components 3 and 4 (Supplementary Figure 1). The model included EEMs from all 3 watersheds, each of which composed 11-36% of all EEMs used for the model. Of the EEMs from the model, 19 were collected concurrent with samples for this study, and were used for further analysis.

We also calculated two land use parameters that may explain changes in bacterioplankton communities: percent developed area (not including developed open space such as parks) and percent impervious surface, using 2011 National Land Cover Database (NLCD) and watersheds delineated from 10-m digital elevation models (DEM) from the Utah Automated Geographic Reference Center (AGRC) in ESRI ArcMap 10 (Chen et al., 2018; Van Rossum et al., 2015).

### *Statistical Analyses*

#### *Environmental Drivers of Bacterioplankton Communities*

To evaluate the effect of environmental factors on the bacterial communities, we grouped variables into five categories and performed multiple redundancy analyses (RDA) and selected significant variables ( $P < 0.1$ ) using backwards step-wise regression using the R package vegan (Oksanen et al., 2015; Zelený, 2011). The five categories were standard field parameters (pH, DO, temperature, specific conductivity, turbidity), nutrients ( $\text{NH}_4$ ,  $\text{NO}_3$ , TDN,  $\text{PO}_4$ , TN, DOC,  $\text{SO}_4$ ), major ions (Na, Mg, K, Ca, F, Cl), trace elements and isotopes (Li, B, Al, V, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Rb, Sr, Y, Mo, Sb, Ba, La, Ce, Eu, Pb, U,  $\delta^{18}\text{O}$ ,  $\delta\text{D}$ ), and dissolved organic

matter (BIX, HIX, FI, TC, SUVA<sub>254</sub>, Total intensity, %Protein, component 1, component 2, component 3, component 4). We used RDA and not canonical correspondence analysis (CCA) after determining that our environmental variables followed a linear and not unimodal distribution (DCA axis lengths < 3.0, Lepš and Šmilauer, 2003). For each RDA, we report adjusted R<sup>2</sup> (the percent of variability in community composition explained, calculated by a PERMANOVA test of the model), constrained proportion (cp; the amount of variability in community composition explained with environmental variables, i.e. the constrained model, compared to without, i.e. the unconstrained model), and axis values (the percent of community composition explained by the first and second ordination axes). Ordination plots of models 1-5 are included in Supplementary Figures 1-5. Variables selected by RDA models 1-4 were combined and used in a sixth backwards step-wise regression. Model 5, correlating dissolved organic matter characterization with communities, was generated using a subset of only 19 samples for which EEMs were available; to avoid reducing the number of samples included in the combined model or skewing results by attempting to interpolate missing values, we excluded organic matter variables in the combined model. The variables selected by the combined model were tested for differences between sites using ANOVA and TukeyHSD.

#### *Geographic and Dispersal-Based Influences on Stream Bacterioplankton Communities*

Community similarity between watersheds, locations, and seasons was calculated using principal coordinate analysis (PCoA) of Bray-Curtis dissimilarities (phyloseq package; McMurdie and Holmes, 2013) and statistically tested using permutational multivariate analysis of variance (PERMANOVA) and visualized using ggplot2 (Wickham, 2016). To determine which of the watersheds, locations or seasons were different from the others, we used pairwise

PERMANOVA tests, using the Holm method to correct for multiple comparisons (pairwise.adonis R package; Martinez Arbizu, 2019). We tested for changes in bacterial diversity related to watershed, location, and season by calculating observed richness and Shannon diversity index. Values were graphed using box and whisker plots in R package ggplot2, with the box indicating interquartile range and whiskers showing the high and low extent of observations, with medians shown as a middle bar (Wickham, 2016).

### *Human Influence on Bacterioplankton Communities*

We calculated community co-occurrence network models to compare the core community topology by analyzing samples grouped by watershed. The models were based on the maximal information coefficient (MIC) analysis in R package minerva (Filosi et al., 2014; Reshef et al., 2011). The nodes in the network models represent OTUs and edges represent significant co-occurrence connections that occur in at least 75% of samples in each watershed and have an MIC that is both  $> 0.7$  and statistically significant ( $P$  value  $< 0.01$ ; Junker and Schreiber, 2011). We exported the graphs from R using igraph into Gephi (v. 0.8.2-beta; Bastian et al.), where we visualized networks and calculated network statistics (Campbell, 2015).

To identify which taxa had unique patterns in abundance, we used analysis of composition of microbiomes (ANCOM) to determine which taxa were significantly different in relative abundance between watersheds when controlled for location and season (Mandal et al., 2015). The results of the ANCOM were visualized using heatmap (R basic), where darker color indicates higher relative abundance. Samples and taxa are clustered by similarity in expression patterns, indicated by dendrograms on relevant axes (i.e. trees do not represent phylogeny of taxa). We calculated changes in percent relative abundance between different categories and

reported means and standard deviations. All R code used for analysis is available at <https://github.com/erinfjones/GAMUTdownload>.

## RESULTS

### *Environmental Drivers of Bacterioplankton Communities*

Of the six RDA models, encompassing 53 environmental variables, trace elements and isotopes explained the most community variation ( $R^2=0.22$ ,  $P$  value= 0.001,  $df=6$ , constrained proportion (cp)=0.38; Table 1-3), while nutrients and major ions explained the least ( $R^2=0.09$ ,  $P$  value=0.001,  $df=3$ , cp=0.08). Models incorporating standard field parameters ( $R^2 = 0.15$ ,  $P$  value= 0.001,  $df=5$ , cp=0.25) and organic matter ( $R^2 = 0.19$ ,  $P$  value= 0.001,  $df=5$ , cp=0.42) were intermediate in explaining bacterial community composition.

The combined RDA model, incorporating significant variables from models one through four, included B, Ba, F, La, Mo, Mg, U,  $\text{NO}_3^-$ , temperature, and  $\delta^{18}\text{O}$ , and explained more community variation than any other RDA (adj.  $R^2 = 0.25$ ,  $P$  value= 0.001,  $df=9$ , cp=0.44; Table 1-3). Positive associations (represented by arrows indicating direction and magnitude of correlation between environmental factors and communities) occurred between high elevation Provo River bacterial taxa and La, Red Butte Creek bacteria exiting the reservoir with  $\delta^{18}\text{O}$ , and Logan stream bacteria and  $\text{NO}_3^-$  concentrations and temperature (Figure 1-3). Uranium, Mg, Mo, Ba, and B corresponded most with bacterial communities in Dam and Urban bacterioplankton communities in Red Butte and Provo watersheds. The parameters from the combined model (excluding La, which decreased the model  $R^2$ ) explained 39% of the bacterial community variability. For comparison, the sampling design (Watershed + Location + Season) explained 43% of variation in the bacterial community.

Variables used in the combined model (model 6) highlight the biogeochemical differences among watersheds and in urban environments (Figure 1-4). All parameters from the combined model except  $\text{NO}_3^-$  and temperature differed by watershed (ANOVA,  $P<0.05$ ; Supplemental table 1). Red Butte had the highest concentrations of B, Mg, Mo, and U. Some solutes increased between montane and urban sites (B, Ba, F, Mg, Mo), while only La decreased from upstream to downstream (TukeyHSD,  $P\text{-adj.}<0.05$ ). Boron, Ba, F, La, Mg, and temperature differed by location (ANOVA,  $P<0.05$ ,  $df=4$ ) as well as the interaction of watershed and season (ANOVA,  $P<0.05$ ,  $df=4$ ). Lanthanum, Mg, and temperature were the only parameters in model 6 that varied with season (ANOVA,  $P<0.05$ ,  $df=2$ ).

#### *Geographic and Dispersal-Based Influences on Stream Bacterioplankton Communities*

Headwater bacterioplankton communities were similar across all watersheds and seasons, despite large geographic distances and unique environmental conditions. Communities from the Logan watershed and high-elevation locations (Above1 and 2) in the Provo and Red Butte watersheds clustered together on the PCoA, indicating similar community compositions, but Dam sites in Provo and Red Butte watersheds (where longitudinal transport was disrupted by dams that increased residence times by 2-4 orders of magnitude) were markedly unique (Figure 1-5). Low elevation locations in the Red Butte watershed (Dam, Urban1 and 2) had a clear pattern of both location and season. Provo communities shifted further along Axis1 than Red Butte communities (corresponding to a longer residence time), gradually moving back towards the Above and Logan watershed cluster. Dam locations in Provo and Red Butte watersheds were most similar to the upstream community in May, when flows were highest and residence time

lowest. The Logan watershed, despite covering the most stream miles (Table 1-1), was the most similar in community composition between sites.

The effect of watershed and location relative to season in driving bacterioplankton community composition was confirmed by PERMANOVA and pairwise post-hoc PERMANOVA tests. The differences in community composition observed between groups in the PCoA ordination were significant for all three main effects (Watershed, Season, and Location), as well as the interactions of watershed with season and location (PERMANOVA, Supplemental table 2,  $P=0.001$ ). The  $R^2$  values for watershed, location, and their interaction were each around 21%, while the  $R^2$  values for seasonal changes were much lower, around 8%. Logan watershed was the least similar to Provo and Red Butte (22% and 18%), according to pairwise PERMANOVA tests (Table 1-4). All watersheds were different from each other, regardless of season and/or location (i.e. p-values did not change when either were added as strata to the command model). The PERMANOVA test was unable to differentiate seasons, although controlling for location or watershed made the difference between Spring and Winter significant.

Dispersal driven by mass effects was also supported by richness and diversity patterns. Both gradually decreased in the Logan watershed, dropping by 18.7% (richness) and 8.7% (Shannon index) from the headwaters into the urban environment (Figure 1-6). Richness and diversity in Provo and Red Butte watersheds decreased slightly between Above1 and Above2 locations, then strikingly dropped at the Dam location. Diversity and richness then gradually increased downstream, despite moving into an urban environment, but remained slightly less diverse than the headwater community. Richness and diversity were 25% and 11% higher for samples collected in Spring than Fall or Winter (TukeyHSD, adj.  $P<0.05$ ).

### *Human Influence on Bacterioplankton Communities*

Large reservoirs had substantial effects on bacterioplankton richness, diversity, and community composition. Both species richness and diversity were highest in the Logan watershed, where reservoirs are shallow with much lower residence times (TukeyHSD, adj.  $P<0.05$ ). In the other two watersheds, where reservoir residence time was higher, bacterial diversity decreased by 25% and taxa richness by 67% from upstream to downstream of large reservoirs. Location was selected as a component of the RDA model (Table 1-3), and in the pairwise PERMANOVA tests the Above2 and Dam locations had the most dissimilarity of any comparison in bacterioplankton communities between main effects (27%, Table 1-4).

Development on the landscape was comparatively less influential on bacterioplankton richness, diversity, and community composition. Richness and diversity increased downstream of dams in Provo and Red Butte watersheds, despite moving into an urban environment, but remained slightly less diverse than the headwater community (Figure 1-6). Percent developed land use was not included by the step-wise RDA model selection (Table 1-3). Despite being the most developed and in different ways (densely urban versus mostly agricultural), Red Butte and Provo watersheds were the most similar (Table 1-4). Dam, Urban1 and Urban2 locations were the most similar, and differences between them were not significant regardless of whether comparisons were corrected for watershed and season (data not shown).

Longitudinal discontinuity from the large reservoirs also affected the community network structure for each of the watersheds. Logan watershed had the highest co-occurrence network complexity, with 3-4 times as many nodes as the other two watersheds (Figure 1-7, Table 1-5). Logan and Provo watersheds had similar modularities (0.856 and 0.767), mean path lengths (the average number of steps to connect each node, 6.739 and 6.029), and mean degrees (average

number of edges for each node, 5.533 and 6.662). Red Butte watershed had nearly as many edges as Logan, despite having much fewer nodes; as a result, Red Butte had a much more tightly clustered network, with over 5 times as high density (0.074), 3 times as high degree (19.83) and 0.5 times as high modularity (0.397) as the other two watersheds.

Large reservoirs, more than percent catchment development, influenced bacterial taxa relative abundance, with unique distribution patterns in each watershed. The ANCOM test (displayed using an abundance heatmap in Figure 1-8) returned taxa unique across watershed, location, and season. Similarities of taxa expression patterns (shown by dendrogram on x-axis) matched the clusters in the PCoA ordination. Sporichthyaceae (Actinobacteria) was an important component of communities at Dam and Urban sites in Red Butte and Provo watersheds ( $16.1 \pm 6.9\%$ ), while absent from Logan watershed and Above locations ( $1.1 \pm 3.1\%$ ). Cryomorphaceae (Bacteroidetes) were also enriched directly below reservoirs, with  $3.8 \pm 2.7\%$  in Red Butte and Provo Dam sites, and only  $0.3 \pm 0.1\%$  in Logan and Above sites. Red Butte watershed had higher relative abundance of Acidimicrobiales (Actinobacteria) in February ( $0.57 \pm 0.07\%$ ), while Provo communities in February were enriched with methanotrophic Methylococcaceae (Gammaproteobacteria,  $1.4 \pm 0.01\%$ ). Logan watershed and Above locations had higher densities of Cellvibrionales (Gammaproteobacteria), including over six times the relative abundance of Halieaceae ( $0.88 \pm 0.56\%$ ) and three times the relative abundance of Cellvibrionaceae ( $0.77 \pm 0.59\%$ ). The most abundant taxa, the aerobic, motile Comamonadaceae (Betaproteobacteria; Willems, 2014), was present in all samples at relative abundances between 5.5 and 53.4% (mean of 13.3%), but was removed by the ANCOM test because it was not differentially abundant between watershed, season, or location.

## DISCUSSION

### *Environmental Drivers of Bacterioplankton Communities*

Contrary to our hypothesis, trace element concentrations, and not standard field parameters like pH and dissolved oxygen, were best correlated with bacterioplankton communities. This finding is contrary to observations in soil bacterial communities and longitudinal riverine studies on the Amazon and Mississippi Rivers (Doherty et al., 2017; Henson et al., 2017). However, a recent meta-analysis of stream bacteria compositions found that trace metals, when included in ordinations, always correlated with community composition (Zeglin, 2015). Many studies that report the importance of standard chemical and physical parameters do not include trace elements in their analyses, possibly overestimating the bacterial variability correlated to the standard variables and failing to quantify the role of other unmeasured parameters. Trace elements may capture the variability among bacterial communities due to unique variations in geology, groundwater, and resultant stream chemistry better than the relatively narrow range of conditions in pH (7.24-8.5), dissolved oxygen (8.4-14.7 mg O<sub>2</sub>/L), and temperature (0-9.5°C) in these streams.

Statistical testing of the 53 variables returned many differences between the watersheds, locations and seasons (data not shown), but in some cases these differences failed to impact bacterial community composition, while variables with no statistical difference did affect composition. For example, nitrate concentrations were statistically indistinguishable between any factor of watershed, season or location, but correlated to 10% of bacterioplankton community variation, demonstrating that ecological significance may occur where statistical significance does not. Trace elements are no longer included in many water quality studies because their concentration is usually in a narrow range below toxicity levels for aquatic organisms at higher

trophic levels, but the slight differences in concentration might be high enough to differentiate the multitude of unique microbial taxa (Herrmann et al., 2016; Smedley and Kinniburgh, 2017).

Our highest RDA  $R^2$  values summed to around 50%, meaning we were unable to attribute half of community composition to any factor, which is similar to  $R^2$  values reported in a meta-analysis of 22 environmental microbiology studies (Hanson et al., 2012). Stochasticity, including genetic drift (e.g. die-off, bacterivory) and mutation (Evans et al., 2017; Hanson et al., 2012) may contribute to differences in bacterioplankton community proportional to residence time, but we were unable to distinguish their relative impacts with this study design. Bacterial dormancy and horizontal gene transfer are ecological processes that may also account for part of the 50% of community composition we were unable to associate with environmental conditions (Jones and Lennon, 2010; Trevors et al., 1987).

#### *Geographic and Dispersal-Based Influences on Stream Bacterioplankton Communities*

As hypothesized, our results reaffirm that the dominant dispersal pathway of bacterioplankton is longitudinal transport from headwater streams (where lateral dispersal from soil occurs), a process documented in arctic and temperate regions (Besemer et al., 2013; Crump et al., 2012; Febria et al., 2015). Upstream of dams, diversity decreased with increasing stream distance; however, the rapid decrease in diversity from above to below dams and then the gradual increase in diversity moving downstream demonstrated a deviation from typical patterns of bacterioplankton dispersal (Figure 1-6; Chen et al., 2018). The change in bacterioplankton community composition, presumably occurring within the reservoir, may be due to changes in dominant dispersal mechanisms. In reservoirs, longer residence times separate bacterioplankton from both the longitudinal dispersal from upstream terrestrial soils (Crump et al., 2012; Lennon

and Jones, 2011) and lateral dispersal from internal sources, such as lake sediments, that act as seed banks for microbial diversity (Comte et al., 2017). The mass effect responsible for colonizing headwaters from nearby soils may also be responsible for the pattern of community similarities seen below reservoirs; Red Butte Creek, which has a larger exposure ratio (surface area of bed:streamwater volume) than the Provo River, returned to the Above site cluster in much fewer stream kilometers than the Provo. Hydrologic variability, significantly decreased from static flows released from the reservoirs, may also play a role in structuring the bacterioplankton community indirectly through biofilm establishment (Widder et al., 2014).

Patterns in community similarity within and across watersheds also support the conclusion that lateral dispersal in headwaters followed by longitudinal dispersal is the dominant transport mechanism. Most of the community similarity was attributed to watershed, possibly due to longitudinal transport of a common community through each system from headwater to outlet. The lack of significant interaction between location and season (Supplemental table 1) also indicates headwater-driven mass effects play a role in determining community composition, because an interaction would have indicated important contributions of unique local taxa based on changes in lateral sources and hydrologic connectivity through time. We expected to see a higher variability between communities due to seasonal changes hydrologic conditions (Table 1-2), but our model determined only 7% of stream bacterial communities were associated with seasonal changes. Further, season was only significant in a pairwise PERMANOVA when controlled by watershed and/or location, suggesting that seasonal changes in residence time or lateral connectivity are less significant than expected (due to small and/or redundant contribution of bacterial taxa), unlike other stream bacterial communities that displayed strong seasonal

changes (Crump and Hobbie, 2005; Leff et al., 1998; Niño-García et al., 2017; Yannarell et al., 2003; Zeglin, 2015).

#### *Human Influence on Bacterioplankton Communities*

Developed land use was not associated with changes in bacterioplankton communities (Table 1-3), unlike previous studies identifying correlations between the two (Chen et al., 2018; Van Rossum et al., 2015). However, bacterial community composition was strongly affected by reservoirs, as demonstrated by large changes in communities where reservoir residence time was high (Table 1-2, Figure 1-5). Moving into the urban environment, communities became more similar to the Above locations, suggesting that either 1) lateral dispersal occurred from low elevation soil communities, which were not strongly affected by development in these watersheds, or 2) changes in environmental conditions promoted taxa being transported from above the reservoir to once again increase in abundance. The community network topology was also affected by reservoirs' disruption to longitudinal dispersal, with watershed-specific responses: Red Butte watershed maintained a consistent, highly connected core community while in Provo, taxa were less connected and may be comprised of more generalist taxa functioning independently (Figure 1-7). The network topology may be an artefact of the shorter stream distance and catchment area in the Red Butte watershed, which affected clustering of co-occurrence networks in an Austrian catchment (Widder et al., 2014). The negligible effect of changes in land use at higher order streams suggests instream bacterial communities may not always be dramatically affected by land use changes or stream channelization in urban areas.

Many environmental factors are affected by reservoirs that might be responsible for the change in bacterioplankton community below reservoirs (Comte et al., 2017; Lindström et al.,

2005). Drastic changes in bacterial communities are reported in artificial water systems (e.g. stormwater outfall, sewage, and 9-km long drinking water delivery pipe), but these changes were mostly attributed to biofilm development within pipes (Fisher et al., 2015; Van Rossum et al., 2015). Another potential factor is the removal of the suspended sediment load and particle-associated bacteria. As water slows, sediment drops from the water column and removes a subset of the overall taxa that might be dispersing attached to soil particles in swift currents (Atkinson et al., 1992; Doherty et al., 2017). However, turbidity was not a significant factor in the RDA selection, suggesting that this mechanism was not responsible for the changes observed (Table 1-3).

Large reservoirs may impact bacterioplankton communities by altering substrate availability (Ruiz-González et al., 2015). Carbon substrate quality changes depending on source, and large, deep lakes have very different carbon sources and cycles than shallow streams. As more labile carbon is metabolized in reservoirs, DOC levels increase. DOC concentrations correlated in an RDA ordination with the sites downstream of reservoirs (Supplemental Figure 2), suggesting that the carbon substrate availability was related to changes in the bacterial community. BIX values, indicating an increase in autochthonous production, also correlated with sites below reservoirs (Supplemental Figure 5). Cryomorphaceae, which increased in relative abundance below reservoirs, are chemoorganotrophs that are thought to metabolize these simple organic compounds (Bowman, 2014a). Oxygen is another metabolic component that likely controlled changes in taxa with increased residence time in reservoirs. Unfortunately, because outflow from reservoirs was quickly re-aerated, the oxygen levels measured were not representative of the environment where the community was formed, and our models did not detect correlations between community composition and oxygen concentrations (Figure 1-3). For example, samples

collected below the larger reservoirs had substantially higher concentrations of the facultative anaerobe Sporichthyaceae (Tamura, 2014), suggesting the anoxic hypolimnion played a role in driving community change. Sulfur concentrations also correlated with below reservoir samples, possibly due to oxidizing hydrogen sulfide created by sulfur-reducing bacteria in reservoir sediments. Increased Methylococcaceae relative abundance in Provo reservoir outflows indicates both methanogenesis and methanotrophy were also responsible for some of the changes in community (Bowman, 2014b). These changes in taxa may correspond to an increase in mobilized heavy metals from reservoir sediments; for example, Acidimicrobiales in Red Butte outflow indicates iron-reduction and iron-oxidation may be increased within the reservoir (Clark and Norris, 1996; Itoh et al., 2011). Changes in bacterial taxa may relate to changes in community function, especially those rare metabolic processes for which functional redundancy may be low (Adams et al., 2014; Comte and del Giorgio, 2010).

## CONCLUSION

We tested three predictions regarding bacterioplankton community composition along a longitudinal-elevational gradient, including multiple seasons and anthropogenic alterations to instream and landscape processes to determine the role of environmental conditions, dispersal mechanisms, and human infrastructure on stream bacterioplankton community composition. Trace element concentrations explained more variability in bacterioplankton community composition than other environmental parameters and should be included in more analyses of aquatic microbial ecology. Our findings support the hypothesis that dispersal from headwater soils and along the longitudinal gradient is the driving dispersal mechanism, except in cases where large reservoirs drastically changed water residence time. Reservoirs and lakes (anywhere

residence time changes) should be included in longitudinal stream bacteria studies; residence time may be a better metric than stream distance measurements in studies of bacterioplankton biogeography. Large reservoirs may have more impact on bacterioplankton communities than other aspects of watershed urbanization. More research is needed to quantify the magnitude of the effect of components of Figure 1-1, such as anthropogenic infrastructure and changes in residence time, on stream bacterial communities (e.g. when does longer residence time have positive or negative effects on alpha or beta diversity, and what controls the extent of that effect?). Understanding the drivers of bacteria and other microbial communities in streams will allow improved predictions of how watershed and stream development will affect biogeochemical processes and resultant water quality.

## LITERATURE CITED

- Abbott, B. W., Baranov, V., Mendoza-Lera, C., Nikolakopoulou, M., Harjung, A., Kolbe, T., et al. (2016). Using multi-tracer inference to move beyond single-catchment ecohydrology. *Earth-Science Reviews* 160, 19–42. doi:10.1016/j.earscirev.2016.06.014.
- Abbott, B. W., Bishop, K., Zarnetske, J. P., Minaudo, C., Chapin, F. S., Krause, S., et al. (2019). Human domination of the global water cycle absent from depictions and perceptions. *Nature Geoscience*, 1. doi:10.1038/s41561-019-0374-y.
- Abbott, B. W., Moatar, F., Gauthier, O., Fovet, O., Antoine, V., and Ragueneau, O. (2018). Trends and seasonality of river nutrients in agricultural catchments: 18 years of weekly citizen science in France. *Science of The Total Environment* 624, 845–858. doi:<https://doi.org/10.1016/j.scitotenv.2017.12.176>.
- Adams, H. E., Crump, B. C., and Kling, G. W. (2014). Metacommunity dynamics of bacteria in an arctic lake: the impact of species sorting and mass effects on bacterial production and biogeography. *Front. Microbiol.* 5. doi:10.3389/fmicb.2014.00082.
- Albright, M. B. N., and Martiny, J. B. H. (2018). Dispersal alters bacterial diversity and composition in a natural community. *The ISME Journal* 12, 296–299. doi:10.1038/ismej.2017.161.
- Allison, S. D., and Martiny, J. B. (2008). Resistance, resilience, and redundancy in microbial communities. *Proceedings of the National Academy of Sciences* 105, 11512–11519.
- Atkinson, C. F., Aumen, N. G., Miller, G. L., and Ward, G. M. (1992). Influence of Particle Shapes and Size Distributions on Fine Particulate Organic Matter Surface Area in

Streams. *Journal of the North American Benthological Society* 11, 261–268.

doi:10.2307/1467646.

Baker, A., Tipping, E., Thacker, S. A., and Gondar, D. (2008). Relating dissolved organic matter fluorescence and functional properties. *Chemosphere* 73, 1765–1772.  
doi:10.1016/j.chemosphere.2008.09.018.

Baker, M. A. ,. C. N. Dahm, and Valett, H. M. (2000). “Anoxia, anaerobic metabolism biogeochemistry of the stream water- ground water interface.,” in *Streams and Ground Waters*, eds. Jr. J.B. Jones and P. J. Mulholland (Academic Press, San Diego.), 259–284.

Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: An Open Source Software for Exploring and Manipulating Networks. 2.

Ben Maamar, S., Aquilina, L., Quaiser, A., Pauwels, H., Michon-Coudouel, S., Vergnaud-Ayraud, V., et al. (2015). Groundwater Isolation Governs Chemistry and Microbial Community Structure along Hydrologic Flowpaths. *Front. Microbiol.* 6.  
doi:10.3389/fmicb.2015.01457.

Besemer, K., Singer, G., Quince, C., Bertuzzo, E., Sloan, W., and Battin, T. J. (2013). Headwaters are critical reservoirs of microbial diversity for fluvial networks. *Proceedings of the Royal Society B: Biological Sciences* 280, 20131760–20131760.  
doi:10.1098/rspb.2013.1760.

Blaszcak, J. R., Delesantro, J. M., Urban, D. L., Doyle, M. W., and Bernhardt, E. S. (2019).

Scoured or suffocated: Urban stream ecosystems oscillate between hydrologic and dissolved oxygen extremes. *Limnology and Oceanography* 64, 877–894.  
doi:10.1002/lno.11081.

Bowman, J. P. (2014a). “The Family Cryomorphaceae,” in *The Prokaryotes: Other Major Lineages of Bacteria and The Archaea*, eds. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin, Heidelberg: Springer Berlin Heidelberg), 539–550. doi:10.1007/978-3-642-38954-2\_135.

Bowman, J. P. (2014b). “The Family Methylococcaceae,” in *The Prokaryotes: Gammaproteobacteria*, eds. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin, Heidelberg: Springer Berlin Heidelberg), 411–440.  
doi:10.1007/978-3-642-38922-1\_237.

Briggs, M. A., Lautz, L. K., Hare, D. K., and González-Pinzón, R. (2013). Relating hyporheic fluxes, residence times, and redox-sensitive biogeochemical processes upstream of beaver dams. *Freshwater Science* 32, 622–641. doi:10.1899/12-110.1.

Campbell, T. (2015). Tree Islands of Fertility Structure Bacterial Community Assembly and Functional Genes Contributing to Ecosystem Processes. *All Theses and Dissertations*. Available at: <https://scholarsarchive.byu.edu/etd/5241>.

Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Lozupone, C. A., Turnbaugh, P. J., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *PNAS* 108, 4516–4522. doi:10.1073/pnas.1000080107.

Carling, G. T., Tingey, D. G., Fernandez, D. P., Nelson, S. T., Aanderud, Z. T., Goodsell, T. H., et al. (2015). Evaluating natural and anthropogenic trace element inputs along an alpine to urban gradient in the Provo River, Utah, USA. *Applied Geochemistry* 63, 398–412. doi:10.1016/j.apgeochem.2015.10.005.

Chen, W., Wilkes, G., Khan, I. U. H., Pintar, K. D. M., Thomas, J. L., Lévesque, C. A., et al. (2018). Aquatic Bacterial Communities Associated With Land Use and Environmental Factors in Agricultural Landscapes Using a Metabarcoding Approach. *Front Microbiol* 9. doi:10.3389/fmicb.2018.02301.

Clark, D. A., and Norris, P. R. (1996). Acidimicrobium ferrooxidans gen. nov., sp. nov.: mixed-culture ferrous iron oxidation with Sulfovibrio species. *Microbiology* 142, 785–790. doi:10.1099/00221287-142-4-785.

Comte, J., Berga, M., Severin, I., Logue, J. B., and Lindström, E. S. (2017). Contribution of different bacterial dispersal sources to lakes: Population and community effects in different seasons: Bacterial dispersal in lakes. *Environmental Microbiology* 19, 2391–2404. doi:10.1111/1462-2920.13749.

Comte, J., and del Giorgio, P. A. (2010). Linking the patterns of change in composition and function in bacterioplankton successions along environmental gradients. *Ecology* 91, 1466–1476. doi:10.1890/09-0848.1.

Covino, T. (2017). Hydrologic connectivity as a framework for understanding biogeochemical flux through watersheds and along fluvial networks. *Geomorphology* 277, 133–144. doi:10.1016/j.geomorph.2016.09.030.

Crump, B. C., Adams, H. E., Hobbie, J. E., and Kling, G. W. (2007). Biogeography of Bacterioplankton in Lakes and Streams of an Arctic Tundra Catchment. *Ecology* 88, 1365–1378.

Crump, B. C., Amaral-Zettler, L. A., and Kling, G. W. (2012). Microbial diversity in arctic freshwaters is structured by inoculation of microbes from soils. *The Isme Journal* 6, 1629. doi:10.1038/ismej.2012.9  
<https://www.nature.com/articles/ismej20129#supplementary-information>.

Crump, B. C., and Hobbie, J. E. (2005). Synchrony and seasonality in bacterioplankton communities of two temperate rivers. *Limnology and Oceanography* 50, 1718–1729. doi:10.4319/lo.2005.50.6.1718.

Dahlke, H. E., Easton, Z. M., Lyon, S. W., Walter, M. T., Destouni, G., and Steenhuis, T. S. (2012). Dissecting the variable source area concept—subsurface flow pathways and water mixing processes in a hillslope. *Journal of hydrology* 420, 125–141.

Doherty, M., Yager, P. L., Moran, M. A., Coles, V. J., Fortunato, C. S., Krusche, A. V., et al. (2017). Bacterial Biogeography across the Amazon River-Ocean Continuum. *Frontiers in Microbiology* 8. doi:10.3389/fmicb.2017.00882.

Döll, P., Fiedler, K., and Zhang, J. (2009). Global-scale analysis of river flow alterations due to water withdrawals and reservoirs. *Hydrol. Earth Syst. Sci.*, 20.

Donlan, R. M. (2002). Biofilms: Microbial Life on Surfaces. *Emerging Infectious Diseases* 8, 881–890. doi:10.3201/eid0809.020063.

Duff, J. H., and Triska, F. J. (2000). Nitrogen biogeochemistry and surface-subsurface exchange in streams.

Ehleringer, J. R., Arnow, L. A., Arnow, T., McNulty, I. B., McNulty, I. R., and Negus, N. C. (1992). RED BUTTE CANYON RESEARCH NATURAL AREA: HISTORY, FLORA, GEOLOGY, CLIMATE, AND ECOLOGY. *The Great Basin Naturalist* 52, 95–121. doi:10.2307/41712704.

Evans, S., Martiny, J. B. H., and Allison, S. D. (2017). Effects of dispersal and selection on stochastic assembly in microbial communities. *ISME J* 11, 176–185. doi:10.1038/ismej.2016.96.

Febria, C. M., Hosen, J. D., Crump, B. C., Palmer, M. A., and Williams, D. D. (2015). Microbial responses to changes in flow status in temporary headwater streams: a cross-system comparison. *Frontiers in Microbiology* 6. doi:10.3389/fmicb.2015.00522.

Fierer, N., and Jackson, R. B. (2006). The diversity and biogeography of soil bacterial communities. *Proceedings of the National Academy of Sciences* 103, 626–631. doi:10.1073/pnas.0507535103.

Fierer, N., and Lennon, J. T. (2011). The generation and maintenance of diversity in microbial communities. *American Journal of Botany* 98, 439–448.

Fierer, N., Morse, J. L., Berthrong, S. T., Bernhardt, E. S., and Jackson, R. B. (2007). Environmental Controls on the Landscape-Scale Biogeography of Stream Bacterial Communities. *Ecology* 88, 2162–2173.

Filosi, M., Visintainer, R., Riccadonna, S., Jurman, G., and Furlanello, C. (2014). Stability Indicators in Network Reconstruction. *PLOS ONE* 9, e89815.  
doi:10.1371/journal.pone.0089815.

Findlay, S. (2010). Stream microbial ecology. *Journal of the North American Benthological Society* 29, 170–181. doi:10.1899/09-023.1.

Fisher, J. C., Newton, R. J., Dila, D. K., and McLellan, S. L. (2015). Urban microbial ecology of a freshwater estuary of Lake Michigan. *Elementa (Wash D C)* 3.  
doi:10.12952/journal.elementa.000064.

Follstad Shah, J. J., Jameel, Y., Smith, R. M., Gabor, R. S., Brooks, P. D., and Weintraub, S. R. (2019). Spatiotemporal Variability in Water Sources Controls Chemical and Physical Properties of a Semi-arid Urban River System. *JAWRA Journal of the American Water Resources Association* 55, 591–607. doi:10.1111/1752-1688.12734.

Gabor, R. S., Eilers, K., McKnight, D. M., Fierer, N., and Anderson, S. P. (2014). From the litter layer to the saprolite: Chemical changes in water-soluble soil organic matter and their correlation to microbial community composition. *Soil Biology and Biochemistry* 68, 166–176. doi:10.1016/j.soilbio.2013.09.029.

Goodsell, T. H., Carling, G. T., Aanderud, Z. T., Nelson, S. T., Fernandez, D. P., and Tingey, D. G. (2017). Thermal groundwater contributions of arsenic and other trace elements to the middle Provo River, Utah, USA. *Environ Earth Sci* 76, 268. doi:10.1007/s12665-017-6594-9.

Grill, G., Lehner, B., Thieme, M., Geenen, B., Tickner, D., Antonelli, F., et al. (2019). Mapping the world's free-flowing rivers. *Nature* 569, 215. doi:10.1038/s41586-019-1111-9.

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., et al. (2008). Global Change and the Ecology of Cities. *Science* 319, 756–760. doi:10.1126/science.1150195.

Hale, R. L., Turnbull, L., Earl, S. R., Childers, D. L., and Grimm, N. B. (2015). Stormwater Infrastructure Controls Runoff and Dissolved Material Export from Arid Urban Watersheds. *Ecosystems* 18, 62–75. doi:10.1007/s10021-014-9812-2.

Hall, E. K., Bernhardt, E. S., Bier, R. L., Bradford, M. A., Boot, C. M., Cotner, J. B., et al. (2018). Understanding how microbiomes influence the systems they inhabit. *Nature Microbiology* 3, 977. doi:10.1038/s41564-018-0201-z.

Hall, S. J., Hale, R. L., Baker, M. A., Bowling, D. R., and Ehleringer, J. R. (2015). Riparian plant isotopes reflect anthropogenic nitrogen perturbations: robust patterns across land use gradients. *Ecosphere* 6, art200. doi:10.1890/ES15-00319.1.

Hanson, C. A., Fuhrman, J. A., Horner-Devine, M. C., and Martiny, J. B. H. (2012). Beyond biogeographic patterns: processes shaping the microbial landscape. *Nature Reviews Microbiology* 10, 497–506. doi:10.1038/nrmicro2795.

Hendricks, S. P., and White, D. S. (2000). Stream and Groundwater Influences on Phosphorus Biogeochemistry-9.

Henson, M. W., Hanssen, J., Spooner, G., Flemming, P., Pukonen, M., Stahr, F., et al. (2017).

Nutrient dynamics and stream order influence microbial community patterns along a 2914 km transect of the Mississippi River. *bioRxiv*. doi:10.1101/091512.

Herrmann, H., Nolde, J., Berger, S., and Heise, S. (2016). Aquatic ecotoxicity of lanthanum – A review and an attempt to derive water and sediment quality criteria. *Ecotoxicology and Environmental Safety* 124, 213–238. doi:10.1016/j.ecoenv.2015.09.033.

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

Itoh, T., Yamanoi, K., Kudo, T., Ohkuma, M., and Takashina, T. (2011). Aciditerrimonas ferrireducens gen. nov., sp. nov., an iron-reducing thermoacidophilic actinobacterium isolated from a solfataric field. *International Journal of Systematic and Evolutionary Microbiology* 61, 1281–1285. doi:10.1099/ijsm.0.023044-0.

Jones, A. S., Aanderud, Z. T., Horsburgh, J. S., Eiriksson, D. P., Dastrup, D., Cox, C., et al. (2017). Designing and Implementing a Network for Sensing Water Quality and Hydrology across Mountain to Urban Transitions. *JAWRA Journal of the American Water Resources Association* 53, 1095–1120. doi:10.1111/1752-1688.12557.

Jones, S. E., and Lennon, J. T. (2010). Dormancy contributes to the maintenance of microbial diversity. *Proceedings of the National Academy of Sciences of the United States of America* 107, 5881–5886. doi:10.1073/pnas.0912765107.

Junker, B. H., and Schreiber, F. (2011). *Analysis of Biological Networks*. John Wiley & Sons.

Kadlec, J., StClair, B., Ames, D. P., and Gill, R. A. (2015). WaterML R package for managing ecological experiment data on a CUAHSI HydroServer. *Ecological Informatics* 28, 19–28. doi:10.1016/j.ecoinf.2015.05.002.

Kolbe, T., Dreuzey, J.-R. de, Abbott, B. W., Aquilina, L., Babey, T., Green, C. T., et al. (2019). Stratification of reactivity determines nitrate removal in groundwater. *PNAS* 116, 2494–2499. doi:10.1073/pnas.1816892116.

Leff, L. G., Leff, A. A., and Lemke, M. J. (1998). Seasonal changes in planktonic bacterial assemblages of two Ohio streams. *Freshwater Biology* 39, 129–134. doi:10.1046/j.1365-2427.1998.00269.x.

Leff, L. G., and Lemke, M. J. (1998). Ecology of Aquatic Bacterial Populations: Lessons from Applied Microbiology. *Journal of the North American Benthological Society* 17, 261–271. doi:10.2307/1467967.

Lennon, J. T., and Jones, S. E. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. *Nature Reviews Microbiology* 9, 119–130. doi:10.1038/nrmicro2504.

Lepš, J., and Šmilauer, P. (2003). *Multivariate Analysis of Ecological Data Using CANOCO*. Cambridge University Press.

Lindström, E. S., Agterveld, M. P. K.-V., and Zwart, G. (2005). Distribution of Typical Freshwater Bacterial Groups Is Associated with pH, Temperature, and Lake Water Retention Time. *Appl. Environ. Microbiol.* 71, 8201–8206. doi:10.1128/AEM.71.12.8201-8206.2005.

Lindström, E. S., Forslund, M., Algesten, G., and Bergström, A.-K. (2006). External Control of Bacterial Community Structure in Lakes. *Limnology and Oceanography* 51, 339–342.

Lindström, E. S., and Östman, Ö. (2011). The Importance of Dispersal for Bacterial Community Composition and Functioning. *PLOS ONE* 6, e25883. doi:10.1371/journal.pone.0025883.

Mandal, S., Van Treuren, W., White, R. A., Eggesbø, M., Knight, R., and Peddada, S. D. (2015). Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb. Ecol. Health Dis.* 26, 27663. doi:10.3402/mehd.v26.27663.

Martinez Arbizu, P. (2019). *pairwiseAdonis: Pairwise multilevel comparison using adonis*. Available at: <https://github.com/pmartinezarbizu/pairwiseAdonis> [Accessed March 6, 2019].

McCormick, A., Hoellein, T. J., Mason, S. A., Schluep, J., and Kelly, J. J. (2014). Microplastic is an Abundant and Distinct Microbial Habitat in an Urban River. *Environmental Science & Technology* 48, 11863–11871. doi:10.1021/es503610r.

McKnight, D. M., Boyer, E. W., Westerhoff, P. K., Doran, P. T., Kulbe, T., and Andersen, D. T. (2001). Spectrofluorometric characterization of dissolved organic matter for indication of precursor organic material and aromaticity. *Limnology and Oceanography* 46, 38–48. doi:10.4319/lo.2001.46.1.00038.

McMurdie, P. J., and Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* 8, e61217. doi:10.1371/journal.pone.0061217.

Messager, M. L., Lehner, B., Grill, G., Nedeva, I., and Schmitt, O. (2016). Estimating the volume and age of water stored in global lakes using a geo-statistical approach. *Nature Communications* 7, 13603. doi:10.1038/ncomms13603.

Meybeck, M. (2003). Global analysis of river systems: from Earth system controls to Anthropocene syndromes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 358, 1935–1955. doi:10.1098/rstb.2003.1379.

Miller, M. P., Boyer, E. W., McKnight, D. M., Brown, M. G., Gabor, R. S., Hunsaker, C. T., et al. (2016). Variation of organic matter quantity and quality in streams at Critical Zone Observatory watersheds. *Water Resources Research* 52, 8202–8216. doi:10.1002/2016WR018970.

Moatar, F., Abbott, B. W., Minaudo, C., Curie, F., and Pinay, G. (2017). Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. *Water Resources Research* 53, 1270–1287.

Murphy, K. R., A. Stedmon, C., Graeber, D., and Bro, R. (2013). Fluorescence spectroscopy and multi-way techniques. PARAFAC. *Analytical Methods* 5, 6557–6566. doi:10.1039/C3AY41160E.

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

Niño-García, J. P., Ruiz-González, C., and del Giorgio, P. A. (2016). Interactions between hydrology and water chemistry shape bacterioplankton biogeography across boreal freshwater networks. *The ISME Journal* 10, 1755–1766. doi:10.1038/ismej.2015.226.

Niño-García, J. P., Ruiz-González, C., and del Giorgio, P. A. (2017). Exploring the Ecological Coherence between the Spatial and Temporal Patterns of Bacterioplankton in Boreal Lakes. *Front. Microbiol.* 8. doi:10.3389/fmicb.2017.00636.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., Simpson, G. L., et al. (2015). Package ‘vegan.’

Oldham, C. E., Farrow, D. E., and Peiffer, S. (2013). A generalized Damköhler number for classifying material processing in hydrological systems. *Hydrology and Earth System Sciences* 17, 1133–1148. doi:<https://doi.org/10.5194/hess-17-1133-2013>.

Parlanti, E., Wörz, K., Geoffroy, L., and Lamotte, M. (2000). Dissolved organic matter fluorescence spectroscopy as a tool to estimate biological activity in a coastal zone submitted to anthropogenic inputs. *Organic Geochemistry* 31, 1765–1781. doi:10.1016/S0146-6380(00)00124-8.

Reshef, D. N., Reshef, Y. A., Finucane, H. K., Grossman, S. R., McVean, G., Turnbaugh, P. J., et al. (2011). Detecting Novel Associations in Large Data Sets. *Science* 334, 1518–1524. doi:10.1126/science.1205438.

Rose, S., and Peters, N. E. (2001). Effects of urbanization on streamflow in the Atlanta area (Georgia, USA): a comparative hydrological approach. *Hydrological Processes* 15, 1441–1457.

Ruiz-González, C., Niño-García, J. P., Lapierre, J.-F., and Giorgio, P. A. del (2015). The quality of organic matter shapes the functional biogeography of bacterioplankton across boreal freshwater ecosystems. *Global Ecology and Biogeography* 24, 1487–1498.  
doi:10.1111/geb.12356.

Sartory, D. P., and Grobbelaar, J. U. (1984). Extraction of chlorophyll a from freshwater phytoplankton for spectrophotometric analysis. *Hydrobiologia* 114, 177–187.  
doi:10.1007/BF00031869.

Savio, D., Sinclair, L., Ijaz, U. Z., Parajka, J., Reischer, G. H., Stadler, P., et al. (2015). Bacterial diversity along a 2600 km river continuum: River bacterioplankton diversity. *Environmental Microbiology* 17, 4994–5007. doi:10.1111/1462-2920.12886.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Applied and Environmental Microbiology* 75, 7537–7541. doi:10.1128/AEM.01541-09.

Shiklomanov, I. A. (1993). “World fresh water resources,” in *Water in Crisis: A Guide to the World’s Fresh Water Resources*, ed. Peter H. Gleick (New York: Oxford University Press), 13–24.

Smedley, P. L., and Kinniburgh, D. G. (2017). Molybdenum in natural waters: A review of occurrence, distributions and controls. *Applied Geochemistry* 84, 387–432.  
doi:10.1016/j.apgeochem.2017.05.008.

Somerville, C. C., Knight, I. T., Straube, W. L., and Colwell, R. R. (1989). SIMPLE, RAPID METHOD FOR DIRECT ISOLATION OF NUCLEIC-ACIDS FROM AQUATIC ENVIRONMENTS. *Applied and Environmental Microbiology* 55, 548–554.

Steinman, A. D., Lamberti, G. A., Leavitt, P. R., and Uzarski, D. G. (2017). “Chapter 12 – Biomass and Pigments of Benthic Algae,” in *Methods in Stream Ecology, Volume 1 (Third Edition)*, eds. F. R. Hauer and G. A. Lamberti (Boston: Academic Press), 223–241. doi:10.1016/B978-0-12-416558-8.00012-3.

Tamura, T. (2014). “The Family Sporichthyaceae,” in *The Prokaryotes*, eds. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin, Heidelberg: Springer Berlin Heidelberg), 883–888. doi:10.1007/978-3-642-30138-4\_182.

Trevors, J. T., Barkay, T., and Bourquin, A. W. (1987). Gene transfer among bacteria in soil and aquatic environments: a review. *Canadian Journal of Microbiology* 33, 191–198. doi:10.1139/m87-033.

US Census Bureau Maricopa County Added Over 222 People Per Day in 2016. *The United States Census Bureau*. Available at: <https://www.census.gov/newsroom/releases/2017/cb17-44.html> [Accessed February 23, 2019].

Van Rossum, T., Peabody, M. A., Uyaguari-Diaz, M. I., Cronin, K. I., Chan, M., Slobodan, J. R., et al. (2015). Year-Long Metagenomic Study of River Microbiomes Across Land Use and Water Quality. *Front. Microbiol.* 6. doi:10.3389/fmicb.2015.01405.

Vitousek, P. M., Mooney, H. A., Lubchenco, J., and Melillo, J. M. (1997). Human Domination of Earth’s Ecosystems. *Science* 277, 494–499. doi:10.1126/science.277.5325.494.

Ward, J. V., and Stanford, J. (1983). The serial discontinuity concept of lotic ecosystems.

*Dynamics of lotic ecosystems* 10, 29–42.

Watson, J. E. M., Venter, O., Lee, J., Jones, K. R., Robinson, J. G., Possingham, H. P., et al.

(2018). Protect the last of the wild. *Nature* 563, 27. doi:10.1038/d41586-018-07183-6.

Weishaar, J. L., Aiken, G. R., Bergamaschi, B. A., Fram, M. S., Fujii, R., and Mopper, K.

(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.

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer.

Widder, S., Besemer, K., Singer, G. A., Ceola, S., Bertuzzo, E., Quince, C., et al. (2014). Fluvial network organization imprints on microbial co-occurrence networks. *PNAS* 111, 12799–12804. doi:10.1073/pnas.1411723111.

Willems, A. (2014). “The Family Comamonadaceae,” in *The Prokaryotes*, eds. E. Rosenberg, E. F. DeLong, S. Lory, E. Stackebrandt, and F. Thompson (Berlin, Heidelberg: Springer Berlin Heidelberg), 777–851. doi:10.1007/978-3-642-30197-1\_238.

Yannarell, A. C., Kent, A. D., Lauster, G. H., Kratz, T. K., and Triplett, E. W. (2003). Temporal Patterns in Bacterial Communities in Three Temperate Lakes of Different Trophic Status. *Microb Ecol* 46, 391–405. doi:10.1007/s00248-003-1008-9.

Zarnetske, J. P., Haggerty, R., Wondzell, S. M., and Baker, M. A. (2011). Dynamics of nitrate production and removal as a function of residence time in the hyporheic zone. *Journal of Geophysical Research: Biogeosciences* 116. doi:10.1029/2010JG001356.

Zeglin, L. H. (2015). Stream microbial diversity in response to environmental changes: review and synthesis of existing research. *Frontiers in Microbiology* 6. doi:10.3389/fmicb.2015.00454.

Zelený, D. (2011). Analysis of community ecology data in R. *Analysis of community ecology data in R*. Available at: <https://www.davidzeleny.net/anadat-r/doku.php/en:start> [Accessed March 7, 2019].

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

Zwart, G., Crump, B., Kamst-van Agterveld, M., Hagen, F., and Han, S. (2002). Typical freshwater bacteria: an analysis of available 16S rRNA gene sequences from plankton of lakes and rivers. *Aquatic Microbial Ecology* 28, 141–155. doi:10.3354/ame028141.

## FIGURES

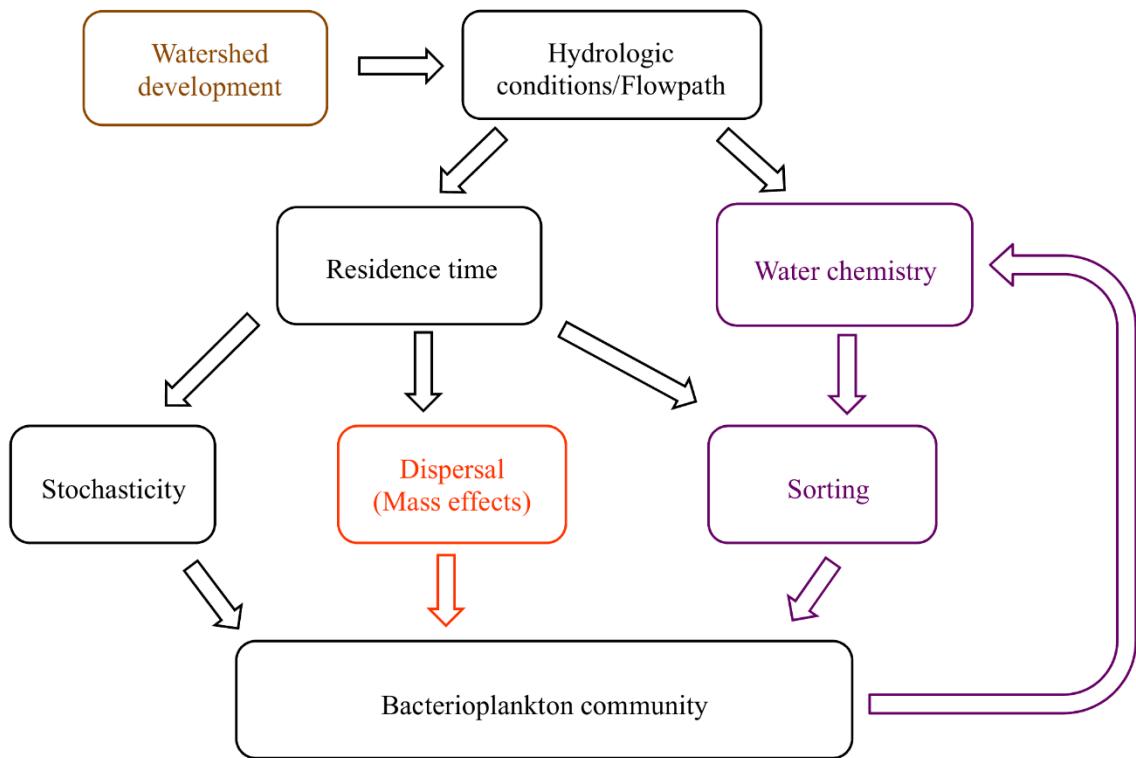


Figure 1-1. Conceptual diagram of the proposed relationships between hydrologic conditions and bacterioplankton communities. Residence time, a hydrologic condition related to flow velocity and volume, influences the extent that dispersal and stochastic processes alter community composition. Residence time also controls the extent to which environmental pressures act on communities to create species sorting, or selection. Bacterial metabolic activity affects water chemistry, creating a feedback loop within the model (hypothesis 1, purple). The community composition similarity along and between watersheds is related to longitudinal and lateral dispersal (hypothesis 2, orange). Anthropogenic changes to watersheds affect hydrologic conditions, and with repercussions on all aspects of the system (hypothesis 3, brown).

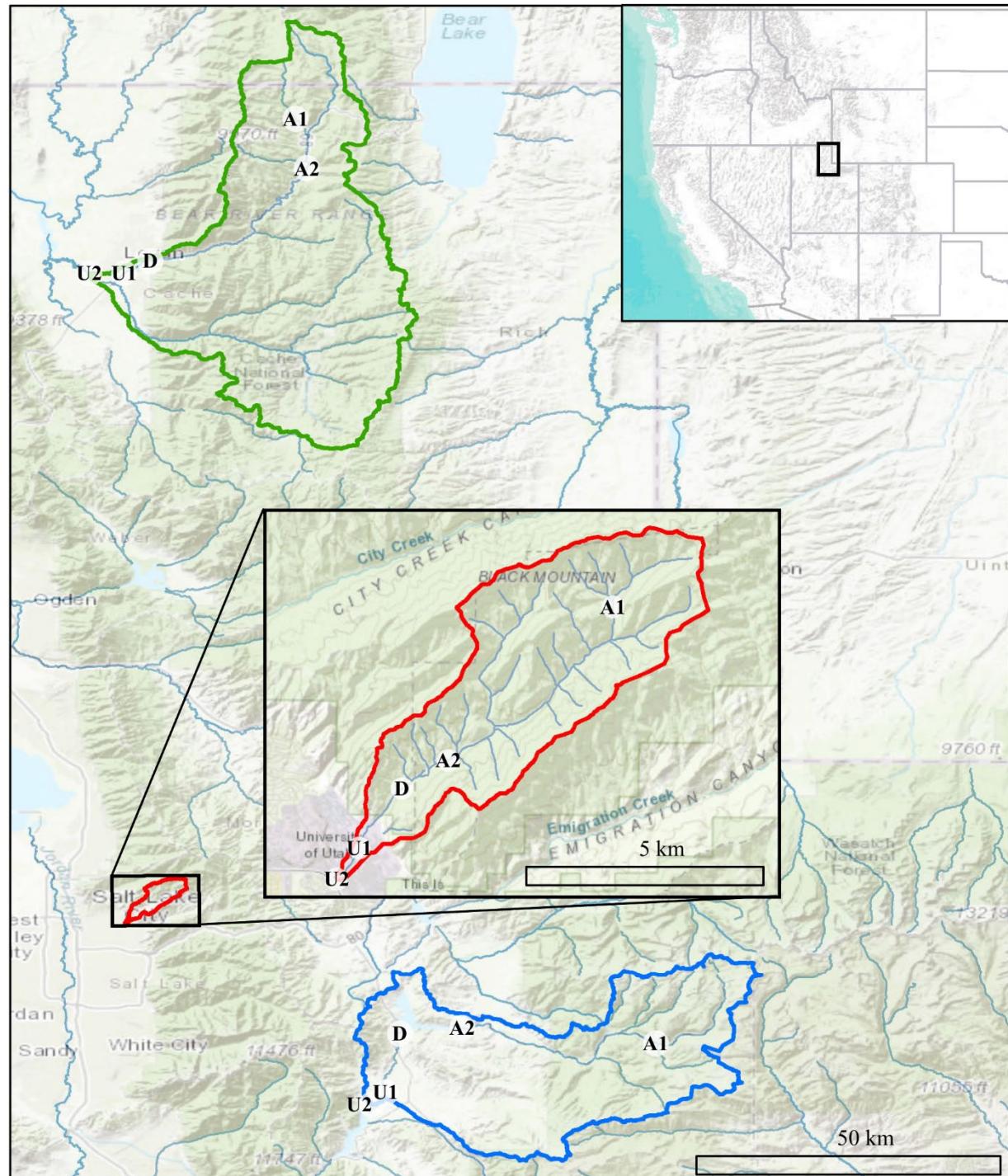


Figure 1-2. Map showing Logan (green), Red Butte (red) and Provo (blue) Watersheds in the Wasatch Range Metropolitan Area and locations within watersheds where bacterial communities and water quality data were collected. Location markers indicate position relative to man-made reservoirs and urban centers in all three watersheds: A= Above, D=below Dam, U= Urban. Site metadata is included in Table 1-1. Sources: ESRI, USGS, NOAA.

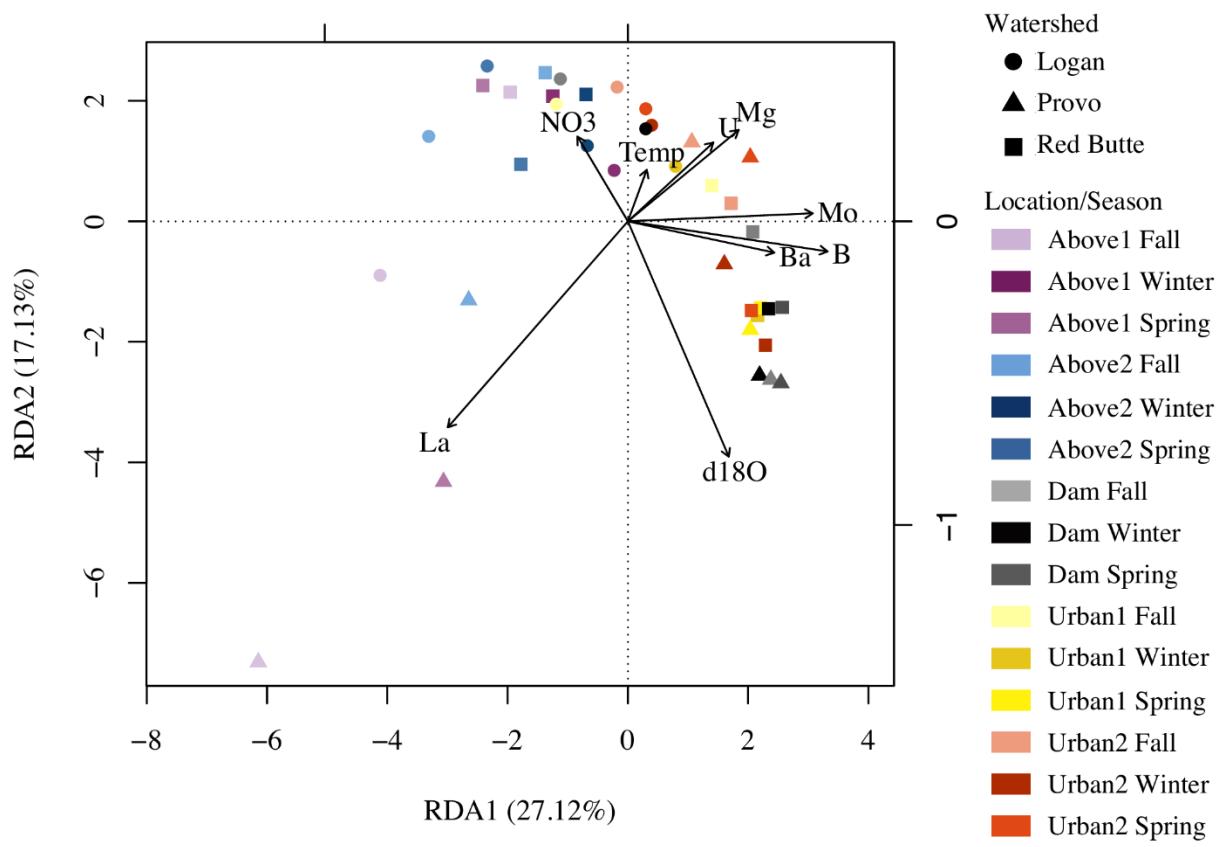


Figure 1-3. Vectors are significant environmental factors, indicating positive correlation. Location indicates position relative to man-made reservoirs and urban centers. Variables were selected from the combination of four RDA models of standard field parameters, nutrients, major ions, and trace elements (Table 1-3).

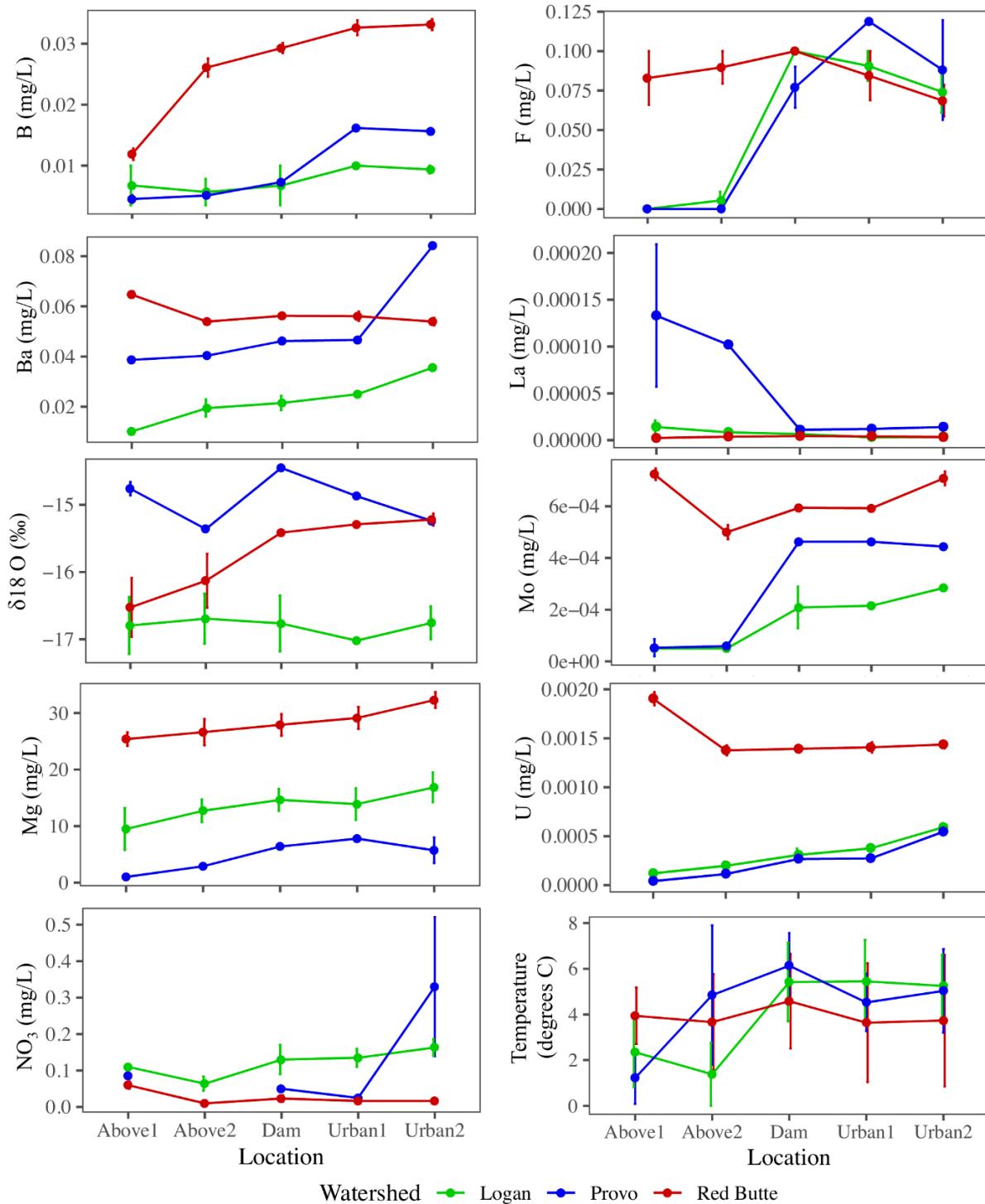


Figure 1-4. Longitudinal profiles of stream physiochemical variables selected by backwards stepwise RDA ordination of bacterial communities in streams in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Location indicates position relative to reservoirs and urban centers. Means and standard errors, where determinable, are shown (some data points were removed for QA/QC violations).

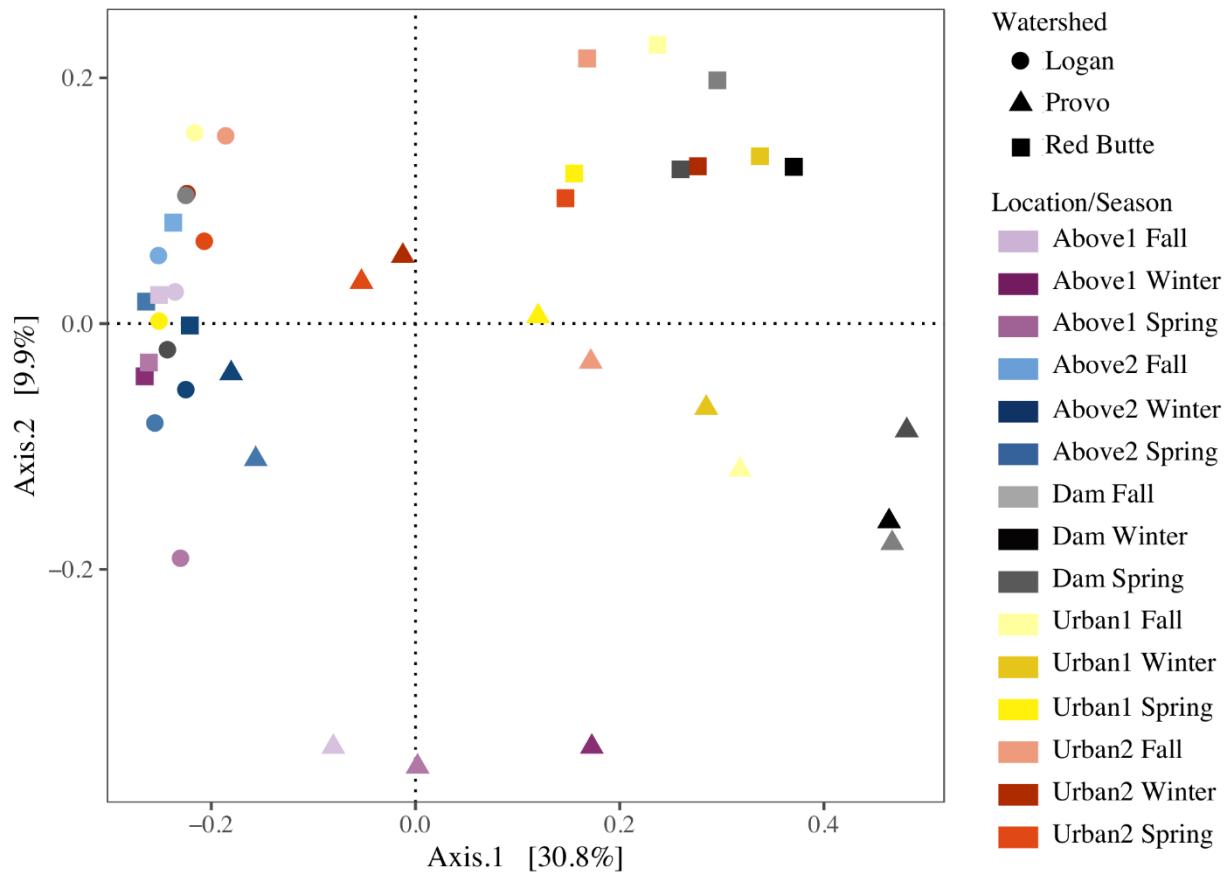


Figure 1-5. Reservoirs alter stream bacterial community composition from three watersheds in Utah's Wasatch Range Metropolitan Area (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Graph represents a Principal Coordinate Analysis (PCOA) ordination, with each point representing a community of bacterioplankton classified at the 97% similarity from 16s rRNA gene amplicon sequencing of OTUs. Location indicates position relative to man-made reservoirs and urban centers. Points closer together are more similar (based on Bray-Curtis dissimilarity index), while points farther apart are dissimilar.

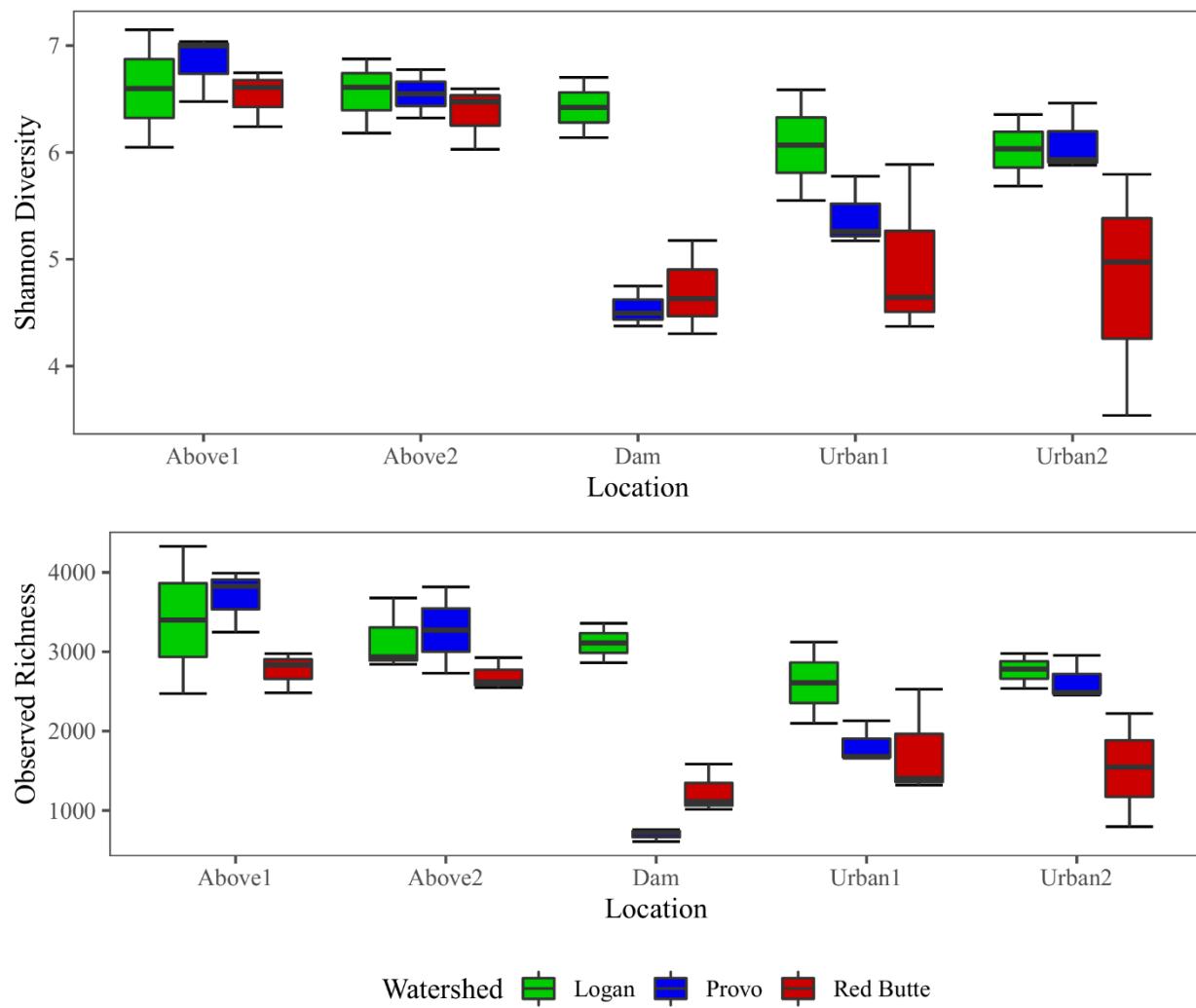


Figure 1-6. Observed richness and Shannon diversity of bacterial communities for three watersheds over three seasons in Wasatch Range Metropolitan Area (WRMA). Boxes depict interquartile ranges, with the center line on the median and whiskers showing the extent of values.

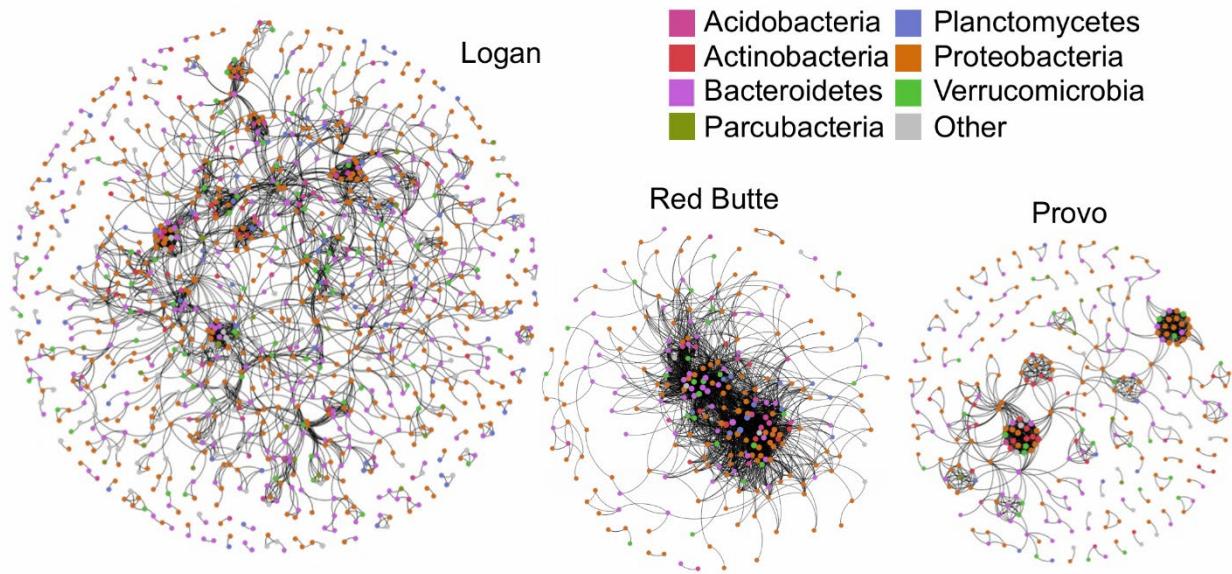


Figure 1-7. Network co-occurrence models for bacterial communities collected along elevation and urbanization gradients in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Nodes indicate taxa (OTUs, taxa with 97% similar sequences) and edges connect where significant co-occurrence was detected in 75% of samples. Topologic statistics of the network are shown in Table 1-6.

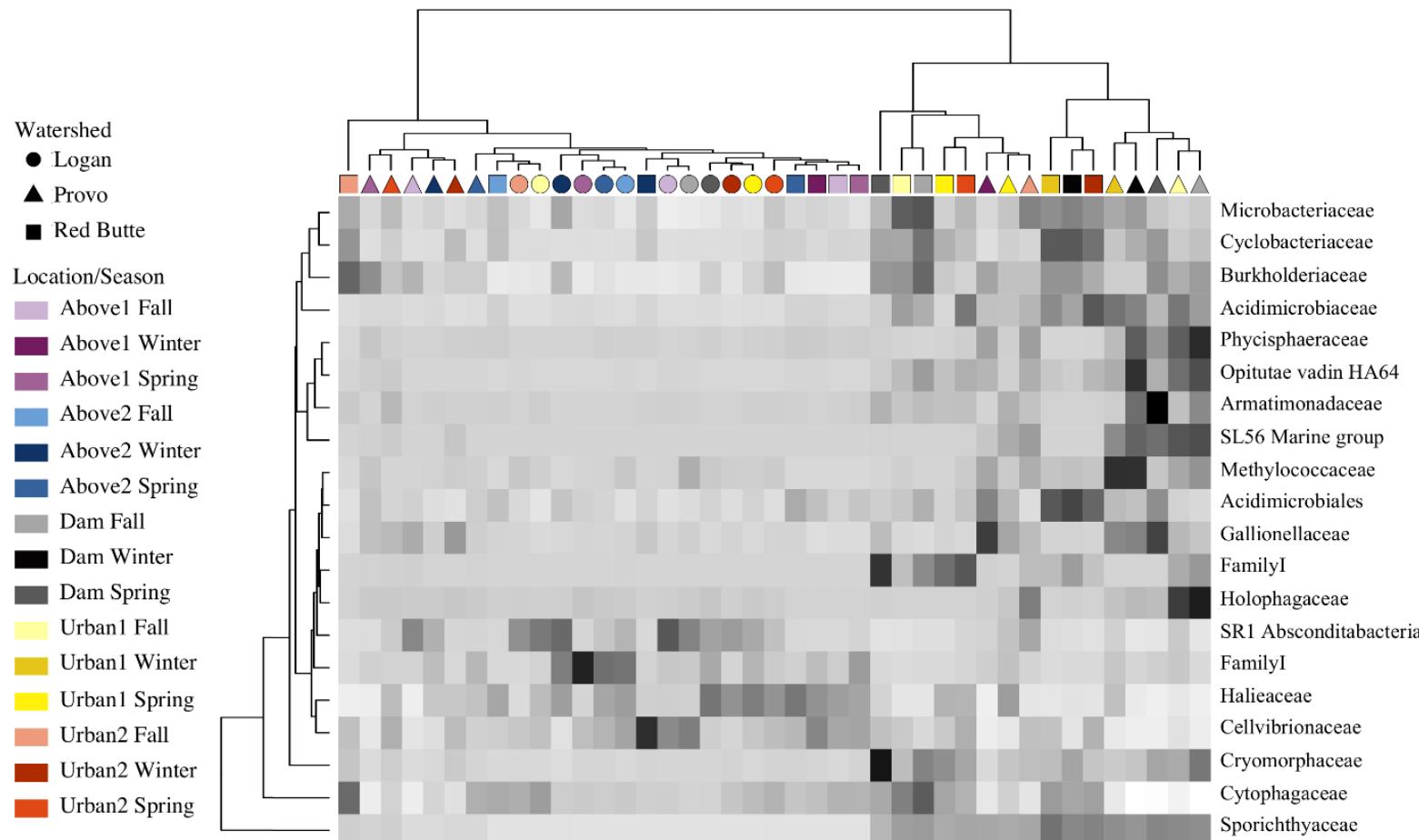


Figure 1-8. Heatmap showing ANCOM results of differentially expressed family for sites in three watersheds across three seasons (Fall, Winter, and Spring). Taxa were selected using a family relative abundance at least 0.5% for all samples. Location indicates position within watershed relative to man-made reservoirs, with Dam sites located immediately downstream of reservoir outlets (Table 1-1). Darker squares indicate higher relative abundance for bacterial families in each sample. Similarity between samples and taxa is indicated by dendrograms on each axis.

## TABLES

Table 1-1. Stream site characteristics in the Logan, Red Butte, and Provo River Watersheds (Figure 1-2). Location indicates site position relative to reservoir(s) in the watershed, with Dam referring to sites immediately downstream of reservoir outflow (Table 1-2). Percent developed was calculated for the entire upstream watershed using NLCD 2011 data, not including “developed open space” (green space within urban areas). Mean discharge (and standard deviation) was calculated based on the 2014 water year, which is when sampling occurred.

| Watershed | Location | Site name          | Elevation<br>(m.a.s.l.) | Distance from<br>outlet (km) | Watershed<br>area (km <sup>2</sup> ) | % Developed | Discharge<br>(m <sup>3</sup> s <sup>-1</sup> ) |
|-----------|----------|--------------------|-------------------------|------------------------------|--------------------------------------|-------------|--|
| Logan     | Above1   | Franklin Basin     | 2110                    | 75.67                        | 63.4                                 | 0.0         | 1.05   |
| Logan     | Above2   | Tony Grove         | 1886                    | 62.86                        | 277.6                                | 0.1         | 1.75   |
| Logan     | Dam      | Water Lab          | 1414                    | 17.34                        | 556.4                                | 0.2         | 4.14   |
| Logan     | Below1   | Main Street        | 1377                    | 11.49                        | 560.3                                | 0.6         | 3.14   |
| Logan     | Below2   | Mendon Road        | 1353                    | 0                            | 1924.6                               | 0.9         | 4.64   |
| Red Butte | Above1   | Knowlton Fork      | 1986                    | 9.00                         | 3.7                                  | 0.0         | 0.020  |
| Red Butte | Above2   | Above RB Reservoir | 1649                    | 3.72                         | 18.7                                 | 0.0         | 0.047  |
| Red Butte | Dam      | Red Butte Gate     | 1582                    | 2.45                         | 20.6                                 | 0.0         | 0.042  |
| Red Butte | Below1   | Cottam's Grove     | 1502                    | 0.85                         | 22.4                                 | 0.3         | 0.038  |
| Red Butte | Below2   | Foothill Drive     | 1449                    | 0                            | 22.8                                 | 1.3         | 0.032  |
| Provo     | Above1   | Soapstone          | 2368                    | 75.73                        | 154.7                                | 0.2         | 3.25   |
| Provo     | Above2   | Hailstone          | 1880                    | 31.07                        | 589.6                                | 0.2         | 7.25   |
| Provo     | Dam      | Below Jordanelle   | 1790                    | 20.42                        | 672.5                                | 0.2         | 7.31   |
| Provo     | Below1   | Lower Midway       | 1676                    | 4.65                         | 718.5                                | 0.3         | 5.44   |
| Provo     | Below2   | Charleston         | 1658                    | 0                            | 779.6                                | 1.3         | 6.27   |

Table 1-2. Physical specifications and discharge of the reservoirs and dams within the three study reaches. The Logan Watershed has a series of multiple impoundments with the lowest, First Dam, being the largest. Using maximum capacity and discharge at gages located immediately downstream, we calculated annual residence time.

| <b>Watershed</b>                                  | <b>Logan</b> | <b>Red Butte</b> | <b>Provo</b> |
|---|--------------|------------------|--------------|
| Reservoir   | First Dam    | Red Butte        | Jordanelle   |
| Volume at capacity ( $\text{m}^3$ )               | 172687       | 474890           | 395083644    |
| Residence time (day)                              | 0.5          | 43               | 1067         |
| Dam height (meter)                                | 9.14         | 39               | 105.16       |
| Average discharge ( $\text{m}^3 \text{ s}^{-1}$ ) | 4.14         | 0.042            | 7.32         |
| Average discharge November 2014 (Fall)            | 2.87         | 0.02             | 4.21         |
| Average discharge February 2015 (Winter)          | 2.92         | 0.05             | 4.24         |
| Average discharge May 2015 (Spring)               | 10.01        | 0.13             | 8.59         |

Table 1-3. Redundancy analysis (RDA) model results indicating the stream physiochemical variables structuring bacterial communities in streams across three watersheds (Logan, Red Butte, and Provo) and three seasons (Fall, Winter, and Spring). Environmental variables were separated into five categories and analyzed separately using backwards stepwise selection of AIC scores to identify significant components ( $p<0.1$ ). Variables from the best-fit models (excluding number five, which used fewer samples than other models) were combined and reported as an overall model. We report summary statistics, including adjusted  $R^2$ , RDA axis percents, and proportion of community variation explained by constrained model (CP). Visualizations of the ordinations are included in supplemental figures 1-5. All models were significant ( $p\text{-value}=0.001$ ).

| # | Model name      | Variables   | Significant variables                                    | Adj. $R^2$ | Axis 1, Axis 2 (%) | CP     |
|---|-----------------|---|--|------------|--------------------|--------|
| 1 | Basic           | Watershed, Season, Elevation, % Imp., Temp, pH, DO, Sp. Cond, Turbidity NH <sub>4</sub> , NO <sub>3</sub> , TDN, PO <sub>4</sub> , TN, DOC, SO <sub>4</sub> | Watershed, Temp, pH, Sp. Cond                            | 0.147      | 39.46, 21.08       | 0.2534 |
| 2 | Nutrients       |   | NO <sub>3</sub> , DOC, SO <sub>4</sub> , NH <sub>4</sub> | 0.086      | 40.09, 35.99       | 0.1909 |
| 3 | Major ions      | Na, Mg, K, Ca, F, Cl<br>Li, B, Al, V, Mn, Fe, Co, Ni, Cu, Zn, As, Se, Rb, Sr, Y, Mo, Sb, Ba, La, Ce, Eu, Pb, U, δ18O, δD                                    | Mg, K, F   | 0.077      | 50.55, 30.41       | 0.1565 |
| 4 | Trace elements  | BIX, HIX, FI, TC, TotInten, pFmax1, pFmax2, pFmax3, pFmax4, pProtein, SUVA  | B, Mo, La, U, Ba, δ18O                                   | 0.220      | 33.03, 19.99       | 0.376  |
| 5 | Organic matter  | Significant factors from models 1-4   | BIX, FI, pFmax1, pFmax4, TotInten                        | 0.190      | 34.98, 24.47       | 0.415  |
| 6 | Combined        | Watershed + Season + Location   | Temp, Mg, NO <sub>3</sub> , Ba, B, Mo, La, U, δ18O       | 0.255      | 27.12, 17.13       | 0.4415 |
| 7 | Sampling design |   | Watershed, Season, Location                              | 0.183      | 33.7, 16.64        | 0.3464 |

Table 1-4. Pairwise PERMANOVA results comparing bacterioplankton community composition across three watersheds, at five locations, and three seasons. Higher R<sup>2</sup> values suggest more difference in community between two groups; smaller numbers are more similar in composition.  
\* indicates p-value<0.05.

| <b>Watershed</b> | Logan  | Red Butte | Provo |        |        |
|------------------|--------|-----------|-------|--------|--------|
| Logan            | -      | -         | -     |        |        |
| Red Butte        | 0.18*  | -         | -     |        |        |
| Provo            | 0.22*  | 0.12*     | -     |        |        |
| <b>Location</b>  | Above1 | Above2    | Dam   | Urban1 | Urban2 |
| Above1           | -      | -         | -     | -      | -      |
| Above2           | 0.09   | -         | -     | -      | -      |
| Dam              | 0.22*  | 0.27*     | -     | -      | -      |
| Urban1           | 0.19*  | 0.21*     | 0.05  | -      | -      |
| Urban2           | 0.16*  | 0.16*     | 0.11  | 0.05*  | -      |
| <b>Season</b>    | Fall   | Winter    |       | Spring |        |
| Fall             | -      | -         |       | -      |        |
| Winter           | 0.05   | -         |       | -      |        |
| Spring           | 0.07   | 0.05      |       | -      |        |

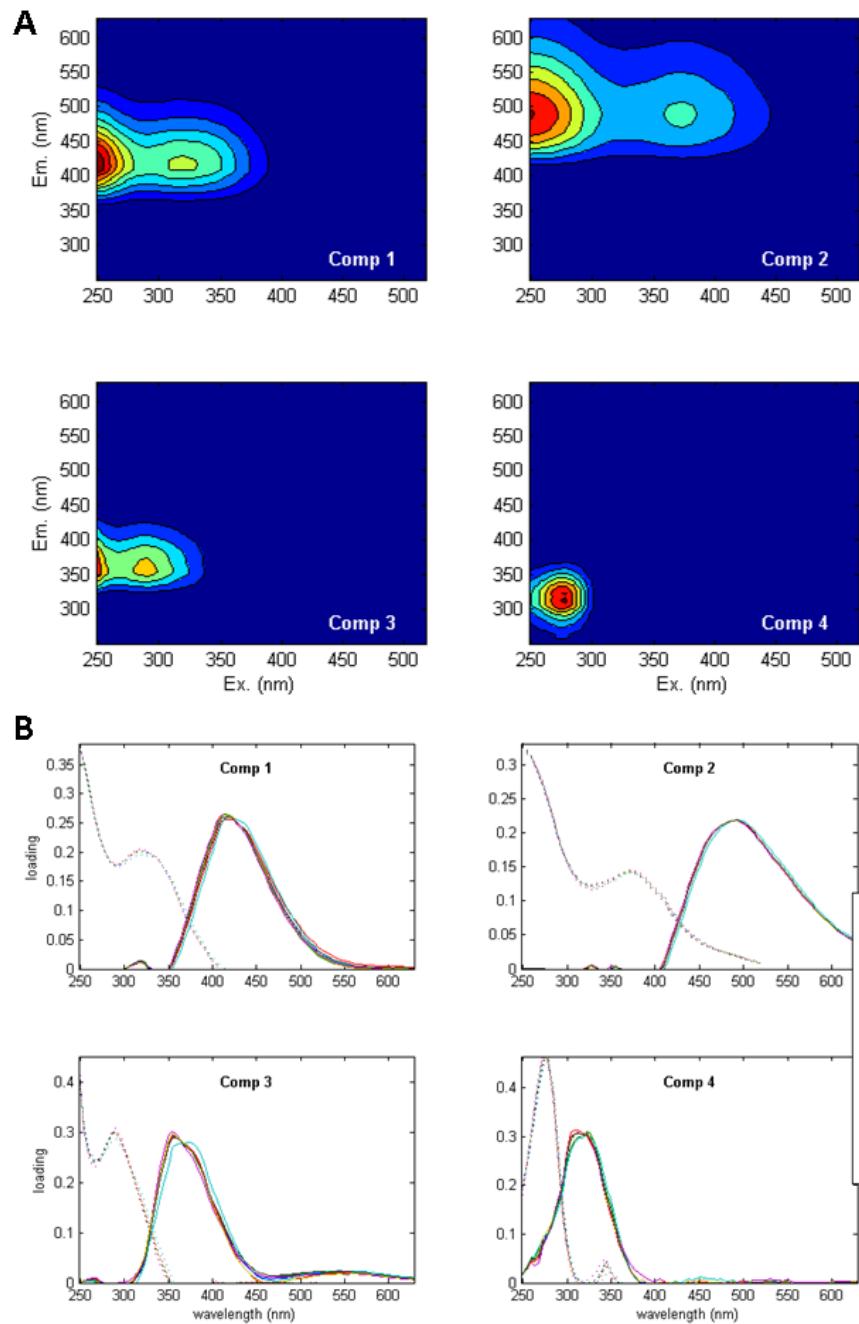
Table 1-5. Topological metrics calculated for network co-occurrence models of bacterial communities collected at five locations along an elevational gradient in three watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Nodes indicate taxa (OTUs) and edges connect where significant co-occurrence was detected in 75% of samples, edges represent significant co-occurrence connections that occur in at least 75% of samples in each watershed and have an MIC that is both > 0.7 and statistically significant. Mean path length: the average number of steps to connect each node; Mean degree: average number of edges for each node; Mean clustering coefficient: a measure of how completely neighboring nodes are connected to each other.

|                             | <b>Logan</b> | <b>Red Butte</b> | <b>Provo</b> |
|-----------------------------|--------------|------------------|--------------|
| Nodes                       | 1071         | 270              | 305          |
| Edges                       | 2963         | 2677             | 1016         |
| Mean Path length            | 6.739        | 3.452            | 6.029        |
| Mean Degree                 | 5.533        | 19.83            | 6.662        |
| Mean Clustering coefficient | 0.57         | 0.547            | 0.703        |
| Density                     | 0.005        | 0.074            | 0.022        |
| Modularity                  | 0.856        | 0.397            | 0.767        |

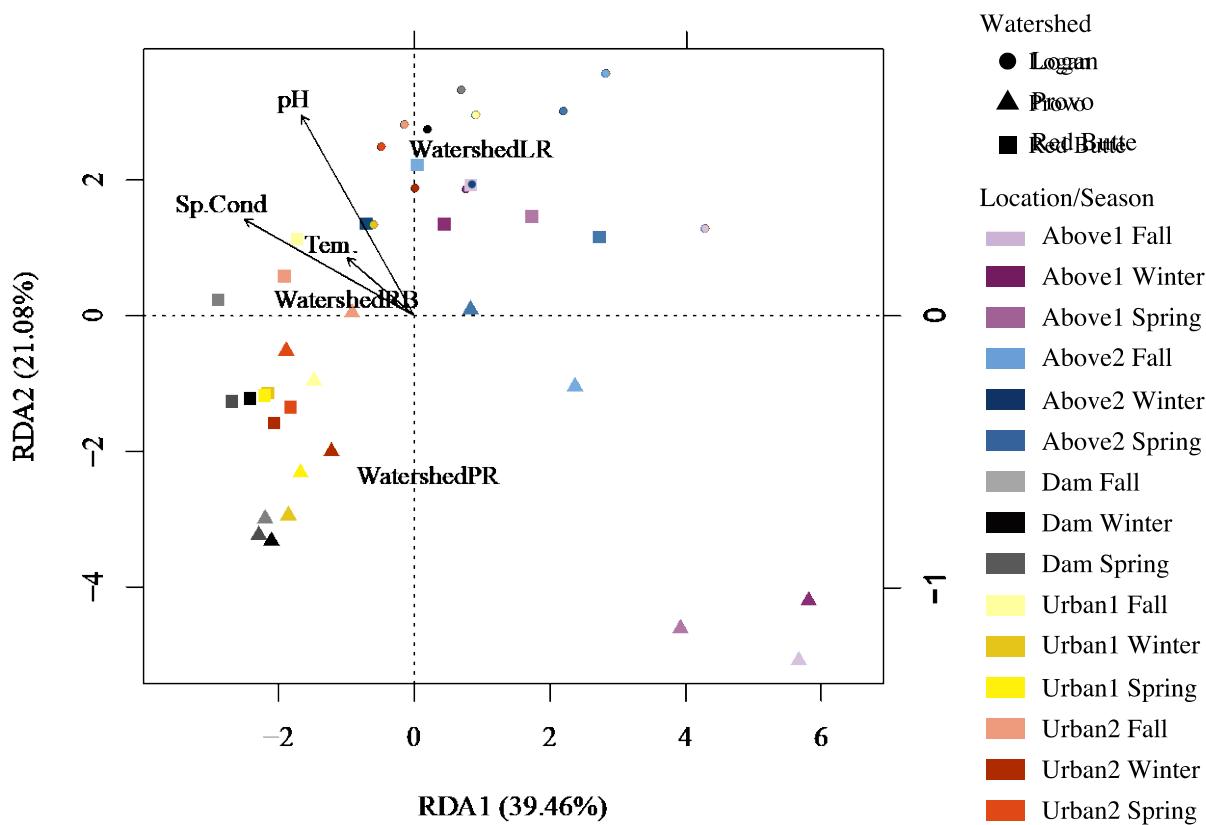
## SUPPLEMENTAL MATERIAL

Supplementary Table 1. PERMANOVA of bacterial communities in streams from three watersheds in the WRMA.

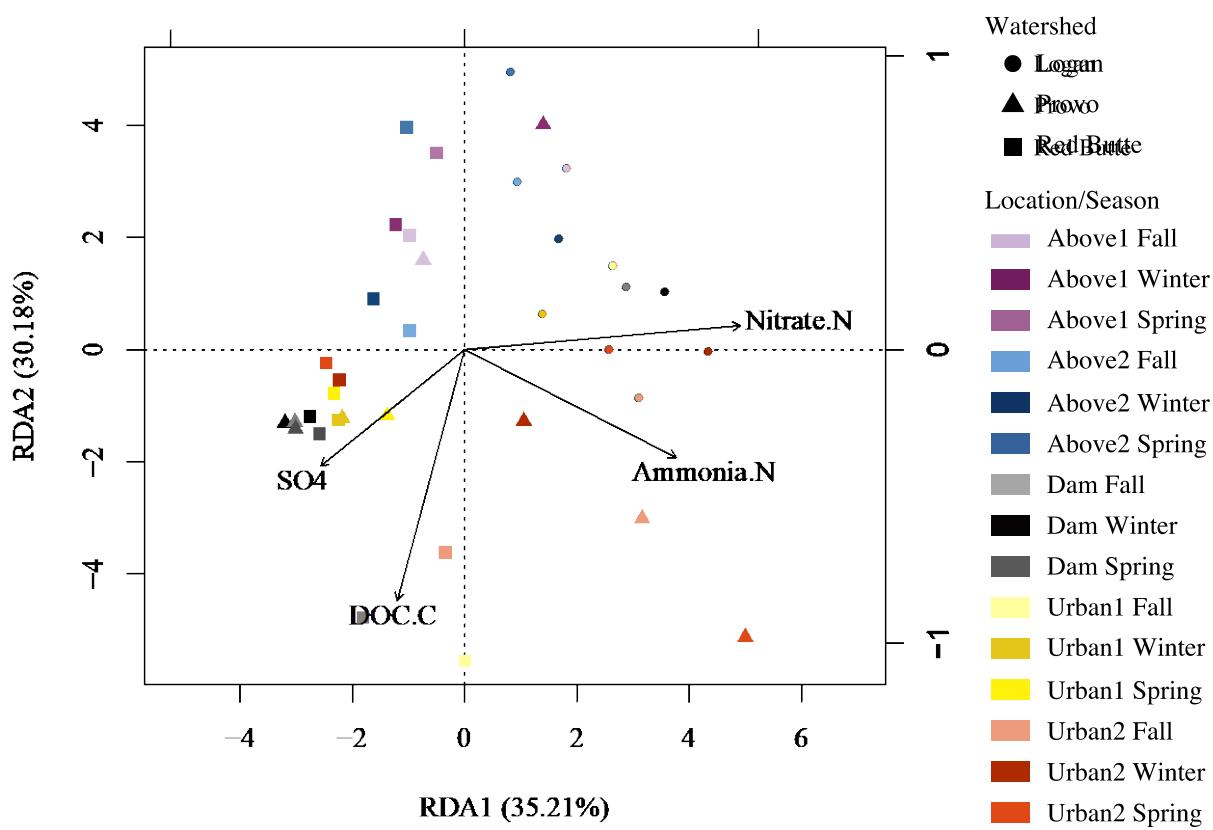
|  | Df | SumsOfSqs | MeanSqs | F.Model | R2      | Pr (>F) |     |
|--|----|-----------|---------|---------|---------|---------|-----|
| Watershed  | 2  | 1.8179    | 0.90894 | 10.6895 | 0.21505 | 0.001   | *** |
| Location   | 4  | 1.7840    | 0.44600 | 5.2451  | 0.21104 | 0.001   | *** |
| Season   | 2  | 0.5934    | 0.29668 | 3.4891  | 0.07019 | 0.001   | *** |
| Watershed:Location   | 8  | 1.8570    | 0.23212 | 2.7299  | 0.21968 | 0.001   | *** |
| Watershed:Season   | 4  | 0.6441    | 0.16103 | 1.8938  | 0.07620 | 0.007   | **  |
| Location:Season  | 8  | 0.7366    | 0.09207 | 1.0828  | 0.08713 | 0.340   |     |
| Residuals  | 12 | 1.0204    | 0.08503 |         | 0.12071 |         |     |
| Total  | 40 | 8.4533    |         |         | 1.00000 |         |     |
| ---  |    |           |         |         |         |         |     |
| Signif. codes: 0 '****' 0.001 '***' 0.01 '**' 0.05 '*' 0.1 '.' 1 |    |           |         |         |         |         |     |



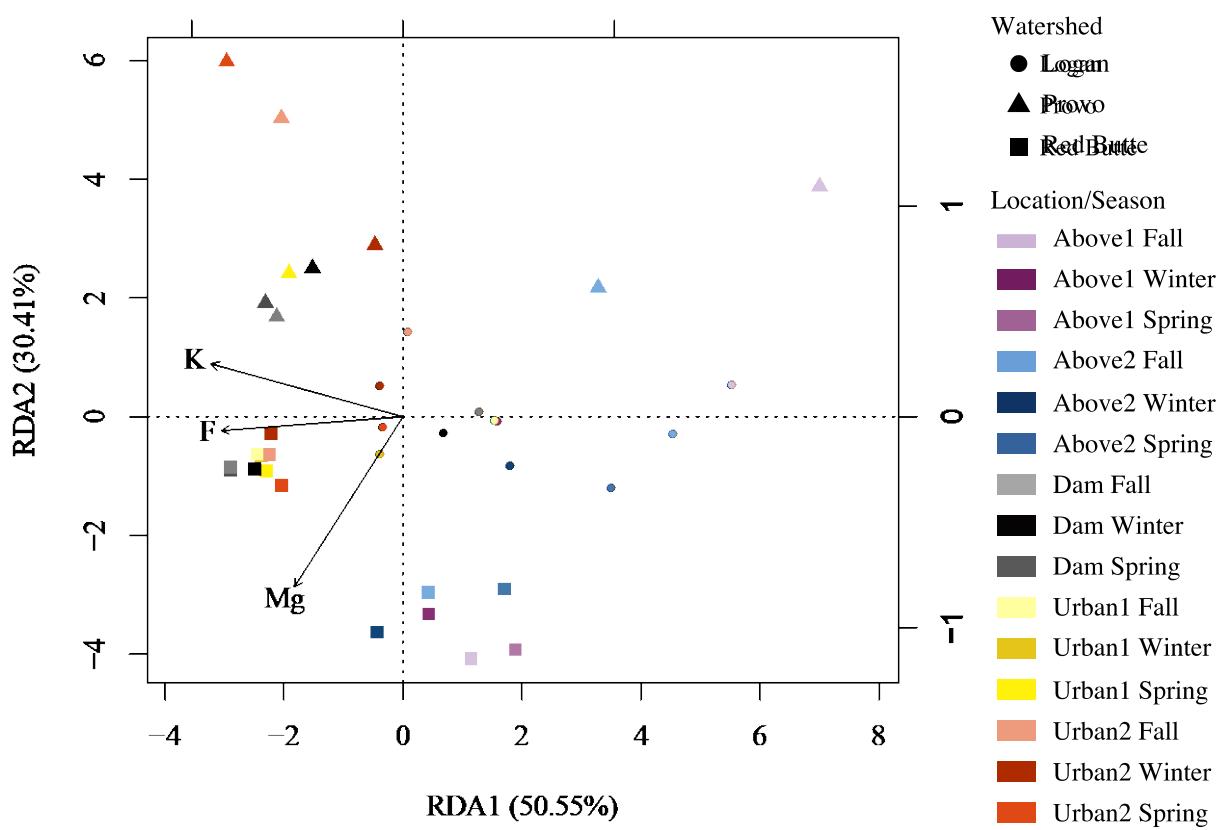
Supplementary Figure 1. A four-component PARAFAC model was resolved (A) and validated with split-half analysis where split models found a match with Tucker correlation coefficient  $> .95$  (B). Components 1 and 2 (C1, C2) were humic-like, and components 3 and 4 (C3, C4) were protein-like.



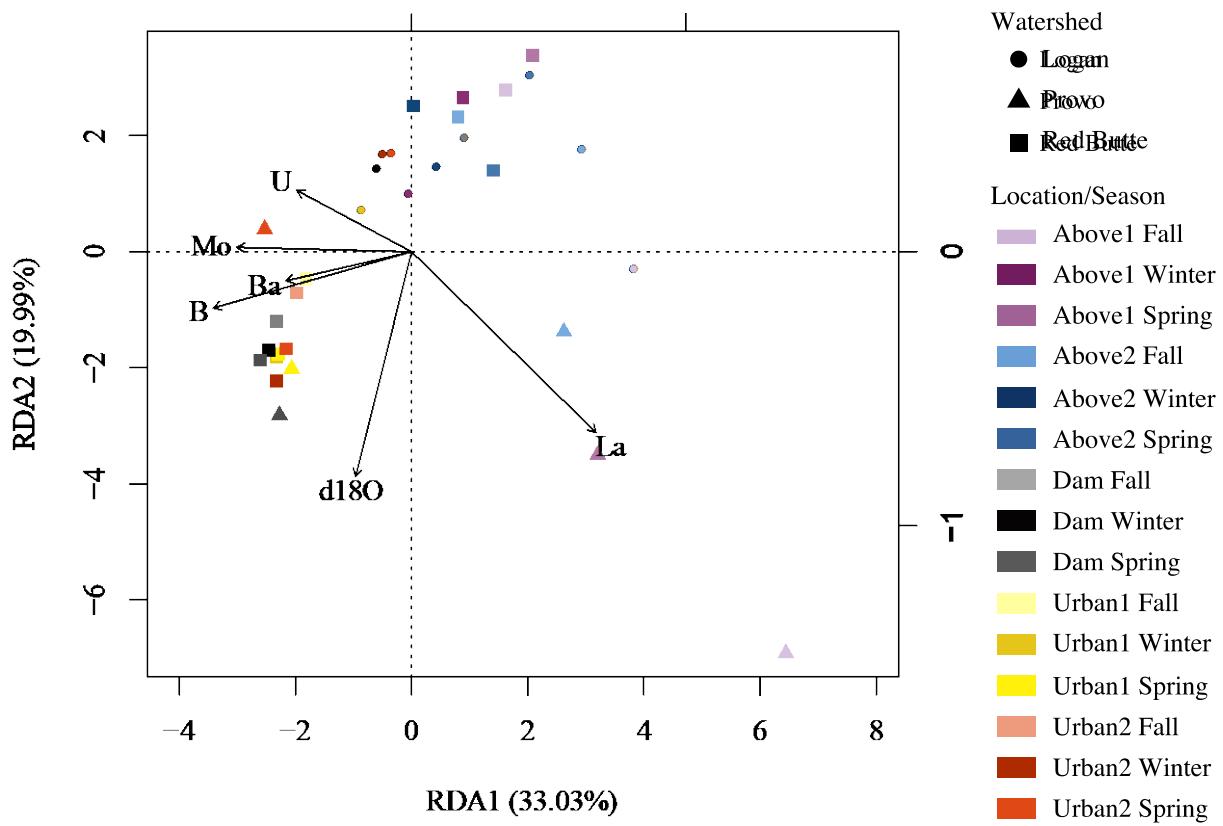
Supplementary Figure 2. Redundancy analysis (RDA) plot relating basic stream chemistry and physical characteristics with bacterial communities in streams in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Vectors represent positive correlations between environmental factors and a sample community composition. Location indicates position relative to man-made reservoirs and urban centers.



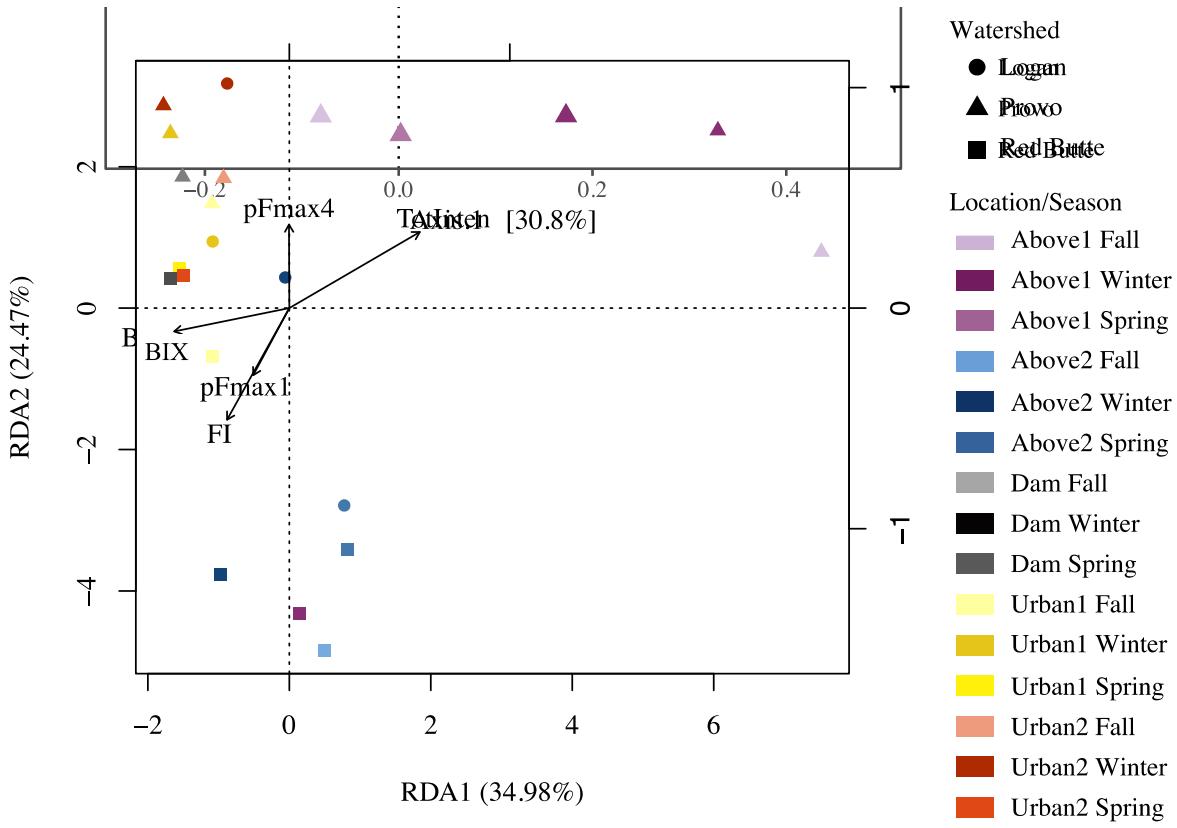
Supplementary Figure 3. Redundancy analysis (RDA) plot relating stream nutrient chemistry with bacterial communities in streams in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Vectors represent positive correlations between environmental factors and a sample community composition. Location indicates position relative to man-made reservoirs and urban centers.



Supplementary Figure 4. Redundancy analysis (RDA) plot relating stream major ion concentrations with bacterial communities in streams in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Vectors represent positive correlations between environmental factors and a sample community composition. Location indicates position relative to man-made reservoirs and urban centers.



Supplementary Figure 5. Redundancy analysis (RDA) plot relating stream minor ions, trace elements, and water isotopes with bacterial communities in streams in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Vectors represent positive correlations between environmental factors and a sample community composition. Location indicates position relative to man-made reservoirs and urban centers.



Supplementary Figure 6. Redundancy analysis (RDA) plot relating stream environmental variables and bacterial communities in streams in three Utah watersheds (Red Butte Creek, Provo River, and Logan River) across three seasons (Fall, Winter, and Spring). Vectors represent positive correlations between environmental factors and a sample community composition. Location indicates position relative to man-made reservoirs and urban centers.

## CHAPTER 2

### High-Frequency Nitrate and Stream Metabolism Reveal Sources and Sinks of Nutrients in Urbanizing Semi-Arid Rivers

Erin Fleming Jones, Dylan Dastrup, David P. Eiriksson, Benjamin W. Abbott,  
Michelle A. Baker, Zachary T. Aanderud

Department of Plant and Wildlife Sciences, Brigham Young University, Provo, UT  
Doctor of Philosophy

#### ABSTRACT

Human activity has increased nutrient loading to aquatic ecosystems, triggering eutrophication that causes damage to ecosystems, human health, and the economy at a global scale. Understanding the processes that regulate nutrient delivery and removal across the terrestrial-aquatic gradient is critical for effective protection and restoration of freshwater ecosystems. Here, we used a network of high-frequency water chemistry sensors to assess loading and in-stream loss of nitrate in three streams in northern Utah, USA. We calculated stream metabolism with diel O<sub>2</sub> data, which we compared with nitrate concentration and flux at two locations in each stream. By comparing upstream and downstream nitrate loads with other sensor metrics (e.g. changes in discharge, specific conductivity, pH, and fluorescent dissolved organic matter), we partitioned nitrate removal between biotic uptake and hydrological loss (i.e. dilution). Diel fluctuations in nitrate corresponded with stream metabolism in the Logan reach, but this only represented 10% of the total nitrate flux, demonstrating how in-stream uptake can easily be overwhelmed by nutrient loading in even moderately modified watersheds. The Red Butte reach had some diel fluctuations, but these did not correspond to instream metabolism, and likely represented upslope or riparian processes. This small, urban, and highly impervious watershed had large increases related to precipitation inputs, potentially from storm water.

## INTRODUCTION

Humans are now the primary source of reactive nitrogen (N) in the Earth system.

Agriculture, land conversion, fossil fuel use, and wastewater discharge have more than doubled pre-industrial inputs of N (Abbott et al., 2018; Foley et al., 2011). At the same time, aquatic ecosystems such as wetlands, floodplains, and river networks have been destroyed or altered at a global scale, reducing aquatic biomass by ~80% and decreasing ecosystem capacity to attenuate excess nutrients and recovery from disturbance (Abbott et al., 2019; Dupas et al., 2019).

Solutions to anthropogenic N loading have been difficult because of the complexity of N cycling. Non-point inputs are difficult to trace back to the original source, and even point sources may be highly temporally variable. Nitrogen has many phases and redox states, including mineral and gaseous phases, leading to a relatively complex cycling and making it difficult to distinguish between anthropogenic and natural N inputs. Variable flow paths and legacy N inputs can create lags between the pollution event and the downstream evidence of degradation (Basu et al., 2010; Zimmer et al., 2019). Natural N sources may be atmospheric, through deposition or N<sub>2</sub> fixation. Soils and groundwater accumulate ammonium or nitrate (NO<sub>3</sub><sup>-</sup>), which is leached into streams. Riparian zones may act as a source or a sink of nitrate, with temporal and seasonal variability (Hill, 2000). Microbes are responsible for a large majority of the aquatic biological cycling of N; benthic and hyporheic zone biofilms can act as sources or sinks of N as water passes between the stream channel and groundwater or adjacent soils (Duff & Triska, 2000). These processes all have natural variation, with some systems more resilient to external loading than others (Burns et al., 2019). The degree and source of N pollution may vary at a local scale depending on natural landscape and type/extent of alterations (Abbott et al., 2019; Moatar et al., 2017).

High frequency sensor data may capture unique signatures that indicate specific sources and pathways (Burns et al., 2019; Kirchner et al., 2004). High frequency sensors are probes installed in water bodies that measure different water quality variables at regular intervals, usually 15 minutes to an hour. The result is a temporally-dense, but because of prohibitive equipment and installation costs, spatially-sparse dataset. Using nitrate sensors and other water quality metrics (e.g. specific conductivity, oxygen, and discharge) between an upstream and downstream sensor station, patterns suggest specific processes (Burns et al., 2019; Jarvie et al., 2018; C. S. Jones et al., 2018; Wollheim et al., 2017; Figure 2-1). High frequency sensors allow scientists to infer differences in hydrologic flow paths based on patterns in storm events (Duncan et al., 2017). These concentration-discharge (C-Q) relationships (Evans & Davies, 1998) may be used to determine where and when solutes are being generated and transported within a watershed, through flushing and hysteresis indices (Baker & Showers, 2019; Blaen et al., 2016; Butturini et al., 2008; Vaughan et al., 2017). High frequency sensors are particularly important for identifying patterns in solutes in small, mixed land-use watersheds, where responses may be erratic and flashy to capture through less frequent sampling regimes (Bende-Michl et al., 2013; Schwientek et al., 2013). For example, diel patterns in NO<sub>3</sub> data give additional insight into both hydrologic and biologic processes (Heffernan & Cohen, 2010; Pellerin et al., 2009). Biological processes may also be inferred from high frequency sensor data because they have a strong diel signal, which is difficult to assess using any other sampling regime (Blaen et al., 2017).

In Utah, USA, eutrophication is an issue leading to harmful algal blooms and degraded ecosystems across the semi-arid northern Wasatch Front, including: the Logan River (Cutler Reservoir), Red Butte Creek (Jordan River), and Provo River (Deer Creek Reservoir and Utah Lake). We installed sensors in three Wasatch watersheds, at sites above and below urban

infrastructure to classify and quantify N processes within urban reaches. We interpreted processes from patterns of diel changes in  $\text{NO}_3^-$  and concentration-discharge (C-Q) hysteresis to answer the following questions: 1) when and why are these reaches gaining or losing  $\text{NO}_3^-$ ?, 2) what percent of  $\text{NO}_3^-$  is biologically transformed in-stream compared to hydrologic processes?, and 3) how much  $\text{NO}_3^-$  input can be attributed to anthropogenic pollution?

## METHODS

### *Study Area*

The Wasatch Range metropolitan area is located at the western edge of the Rocky Mountains and the Great Basin in Utah, USA. The region includes cold desert ecosystems to the west and semi-arid mountains to the east. In a narrow band between the two ecoregions, 2 million people (expected to double by 2050) rely on modest precipitation (254 mm in desert regions, 10164 mm in mountains, a majority of which is as snow) for irrigation (72%), mining (6%), other industrial activities (3%), and public use (15%; Dieter et al., 2018). Many of the watersheds along the Wasatch Range are experiencing eutrophication, due to a combination of natural factors and anthropogenic activities (Randall et al., 2019). We selected three watersheds within this region that originate at high elevations (2400-3000 masl) and capture a range of different types of urbanization at lower elevations (1400-1700 masl): Logan River (agricultural), Red Butte Creek (urbanized), and Provo River (transitioning from agricultural to urban). Watershed characteristics are described in Table 2-1. All three watersheds terminate in water bodies that have been listed as impaired for excess phosphorus (P), and experience seasonal harmful algal blooms: Cutler Reservoir, Jordan River, and Deer Creek Reservoir, respectively. This study uses stream sensor

stations installed above the populated zone (marked by a dam in all three watersheds) and above the impaired receiving waters (Figure 2-1).

Logan River originates in high-elevation limestone karst and ends in Lake Bonneville sediments. Livestock grazing occurs from the headwaters to the outlet, and over the entire range of elevations. The upstream site receives outflow from a small, mesotrophic impoundment with a low residence time (0.5 days). A large tributary (Blacksmith Fork) enters within the Logan reach, increasing the watershed area from 550 km<sup>2</sup> to 1400 km<sup>2</sup>. There are no waste water treatment plants in the reach, but septic systems and large feedlots in the watershed may be losing efficacy with age, and groundwater at the end of the reach has high nitrate levels.

Red Butte Creek is the smallest watershed and stream in our study (22.8 km<sup>2</sup>; Table 2-1). It originates in a protected watershed before entering a highly urban area (Salt Lake City, University of Utah campus; Ehleringer et al., 1992). No major tributaries enter within the study reach, but Red Butte Reservoir at the border between the protected watershed and the urban environment is large relative to the stream discharge, resulting in a high residence time (up to 120 days, based on an average flow of 0.046 m<sup>3</sup>/s and maximum reservoir capacity of 385 acre-feet). No agricultural activity takes place in the watershed, but extensive impervious surfaces drain through dozens of storm drains into the reach. Groundwater inputs have been experimentally determined to be responsible for driving solute concentrations in Red Butte Creek (Gabor et al., 2017; S. J. Hall et al., 2016).

The middle Provo reach represents a region that is rapidly transitioning from agricultural to urban and suburban land use. The reach is bounded by two large reservoirs; the upstream Jordanelle Reservoir (395 million m<sup>3</sup> at capacity) has an average residence time of over 1000 days. Dam managers select the release depth from one or more of 6 gates to maintain

temperatures in the reach, which is a blue-ribbon cold-water fishery. The increased discharge provides water users with agricultural irrigation during dry summer months when evapotranspiration is much higher than precipitation inputs. There are no major tributaries within the reach, but multiple canals divert to agriculture and provide return flow. The reach has one waste water treatment plant adjacent to the stream. Storm water is transported in canals or enters groundwater, which is very shallow for most of the year.

### *Data Collection and Analysis*

Sensor data is described in Jones et al., 2017. Briefly, each site included sensors measuring water temperature, dissolved oxygen, fluorescent dissolved organic matter (fDOM), and specific conductivity (EXO2, Yellow Springs Instruments);  $\text{NO}_3^-$  (SUNA V2, Satlantic); turbidity (DTS, Forest Technology Systems); and discharge. Data were quality controlled following guidelines published in S1 of Jones et al., 2017. We downloaded one year of sensor data from the iUTAH database using the waterml R package (Kadlec et al., 2015). Due to constraints in station operation, we used data from February 1<sup>st</sup>, 2015 to March 31<sup>st</sup>, 2016 for Red Butte and Provo watershed sites, and 2016 to 2017 for Logan. Biweekly grab samples were taken at sensor stations during the same period and analyzed for  $\text{NO}_3^-$ . These samples showed agreement between sensor values and those measured with traditional lab methods (Figure S1).

Time series data were separated into three hydrograph periods: April through June (snow melt and peak runoff; 91 days), July through September (descending limb and growing season; 92 days), and October through March (groundwater dominated low flow; 182 days). We filled gaps in sensor data that were less than two hours using linear interpolation but excluded longer gaps and tested for correlations using spearman's rank correlation coefficient ( $r$ ).

We used the streamMetabolizer R package (Appling et al., 2017) to estimate daily gross primary productivity (GPP) and ecosystem respiration (ER) from climate sensor data collected within each reach, including photosynthetically active radiation (PAR). We also used diel change in dissolved oxygen (as percent of saturation, %DO) to estimate instream metabolic activity at the downstream site in Red Butte, where streamMetabolizer was unable to generate realistic metabolism values, despite fixing K600 values using measurements from SF<sub>6</sub> injection at a nearby site for NEON observations (downloaded from <https://data.neonscience.org>, Wanninkhof et al., 1990). We calculated diel change in %DO as the average value of %DO during the 1600 hour minus the average of %DO over the 0500 hour. We used K600 values from the streamMetabolizer outputs and mean water velocity measurements collected with Flowtracker handhelds (Sontek) to calculate the upstream distance incorporated in oxygen sensor measurements (Grace & Imberger, 2006).

We attempted to quantify diel signals in N using EEMD analysis, but irregularity in seasonal NO<sub>3</sub><sup>-</sup> prevented a meaningful extraction of any pattern. Instead we used a low-tech option of taking the difference of the diel minimum from the maximum to approximate biological uptake/production of NO<sub>3</sub><sup>-</sup> (Figure 2-5, Table 2-2). Diel change in NO<sub>3</sub><sup>-</sup> was calculated by subtracting the average NO<sub>3</sub><sup>-</sup> concentration over the 1400 hour from the average 0200 hour NO<sub>3</sub><sup>-</sup> concentration. These times were selected to represent nighttime low (0200) and midday peak (1400) in light, with a slight lag. We were unable to calculate storm event C-Qs, because discharges in these reaches are highly regulated, meaning that increases in flow may not be associated with precipitation inputs, and storms may not result in a measurable difference in flow volume, despite altering flow paths. Also, the vast difference in watershed sizes and timing of storm responses made comparisons between the watersheds unrealistic, even when normalized.

All analyses were done in RStudio, using zoo, base and vegan packages. All code and data used is available at <https://github.com/erinfjones/sensormetab>. We created linear models of data using gls and an autocorrelation term based on the previous time step.

## RESULTS

### *Time Series Data*

Discharge increased from around 3 m<sup>3</sup>/s in April to over 20 m<sup>3</sup>/s in July as a result of spring runoff (Figure 2-3a). Peak runoff was followed by lowest flows in the growing period at the downstream site, though upstream discharge was not diminished, suggesting agricultural withdrawals occurred. Nitrate at the upstream site was highest during spring runoff. Specific conductivity decreased from 425 µS/cm to 300 µS/cm during peak runoff at the upstream site, suggesting dilution. Nitrate (0.5 to 2.5 mg N/L), specific conductivity (400 to 700 µS/cm), and fDOM (10 to 50 QSU) peaked during summer low flows at the downstream site. Water temperature in the Logan reach also peaked during summer months from winter lows of 3°C to 15-20°C, with dissolved oxygen following an inverse relationship (Figure S2 and S3). Summer diel swings in oxygen (4 mg O<sub>2</sub>/L) were much larger than those in spring runoff and winter (1-2 mg O<sub>2</sub>/L).

Red Butte, the smallest and most urbanized watershed, had much flashier (i.e. variable) time series, particularly at the downstream site. Spring runoff in this watershed began and ended earlier than the other two watersheds, because it has a lower mean elevation (Table 2-1) and therefore snow-free before the others. The upstream site had a clear spring runoff peak (an increase from 0.05 to 0.8 m<sup>3</sup>/s) which, due to the reservoir, was muted at the downstream site. However, high amounts of impervious surface and storm drains within the reach created peaks in

discharge (between 0.4 and 0.6 m<sup>3</sup>/s) at the downstream site that correspond to drastic changes in the other parameters, including NO<sub>3</sub><sup>-</sup>, specific conductivity, fDOM, and turbidity. Nitrate concentration at the upstream site increased from 0.05 to 0.2 mg N/L during spring runoff and remained around 0.1 mg N/L throughout the growing season. Specific conductivity was inverted with discharge during spring runoff ( $r=0.72$  upstream and 0.69 downstream) but correlated in an overall “L” shape opposed to a straight line (Figure S4 and S5). Dissolved oxygen was highly correlated with water temperature ( $r= -0.99$ ).

The middle Provo reach was strongly affected by Jordanelle reservoir, with dam releases delaying “peak runoff” into late July, stepping from 4 to 17 m<sup>3</sup>/s in discrete intervals. Discharge decreased from the top to the bottom of the reach by about 10% (2 m<sup>3</sup>/s) in June through September, likely due to agricultural withdrawals. Increased variability in discharge, as well as specific conductivity, fDOM, and turbidity at the downstream site despite no major tributaries indicate possible groundwater exchange and agricultural withdrawals/returns within the reach. Larger diel swings in temperature (from 1°C to 5°C) and dissolved oxygen (from 1 to 3 mg O<sub>2</sub>/L) also highlighted the role of in-stream processes in introducing variability downstream of the reservoir. Significant dilution of specific conductivity from 210 µS/cm to 160 µS/cm, as well as a 2°C increase in temperature and 20 QSU in fDOM occurred at the upstream site mid-June corresponding to a subtle change in discharge, suggesting a change in reservoir release depth. The upstream site of the Provo reach had lowest NO<sub>3</sub><sup>-</sup> concentrations during spring runoff (Figure 2-3c), suggesting dilution occurred. The effect of homogenous dam release on initial water chemistry is highlighted by C-Q relationships shown in Figure S6 and S7.

*N v N*

Nitrate concentration and variability increased longitudinally in every watershed (Table 2-2, Figure 2-4). The Logan sites had very nearly 1:1 concentration relationships between upstream and downstream during the spring runoff and winter low flow portions of the hydrograph, indicating conservative transport of  $\text{NO}_3^-$  (Figure 2-4a). Average nitrate concentration increased from upstream to the downstream site by 42% in spring runoff, 83% in summer, and 51% in winter low flow (Table 2-2). When combined with discharge data, the load of N to the reach was 30000 kg N over the one-year period.

Nitrate concentration in the Red Butte reach also generally increased from upstream to downstream (Figure 2-4b). The downstream site was 61% and 32% higher in  $\text{NO}_3^-$  concentration than the upstream site during the summer growing season, and winter low flow, respectively. Spring runoff  $\text{NO}_3^-$  concentration decreased by 14%. Some exceptions to the pattern of decreasing  $\text{NO}_3^-$  occurred in each season, some of which may have been artifacts of sensor malfunction or inaccuracy at low concentrations (Figure 2-4b). Winter low flow nitrate was especially poorly quantified, partially because values were near or below the minimum detection limit for the sensor.

Nitrate concentration increased from upstream to downstream in the middle Provo reach (Figure 2-4c), except for during a small period of summer during the high flow dam release. The downstream site had 33%, 42%, and 52% increase in  $\text{NO}_3^-$  for spring runoff, growing season, and winter low flow, respectively. This increase represented almost an addition of 6000 kg N within the reach over the year of study. The  $\text{NO}_3^-$  concentrations for the upstream site were quantized (visible as horizontal lines in the scatter) because the probe was not precise enough to detect the subtle changes in  $\text{NO}_3^-$  in the homogenous outflow. The Provo reach had relatively

little change in  $\text{NO}_3^-$  through time, evidenced by the tight clustering of observations with no obvious hysteresis events (Figure 2-4c). Nitrate concentrations were highest in the winter hydroperiod, except for some dilution that began in March before discharge increased (orange points in Figure 2-4c).

#### *Diel $\text{NO}_3^-$*

Diel change in  $\text{NO}_3^-$  concentration at the upstream end of the Logan reach had little variability (between 0.01 and -0.04 mg N/L; Figure 2-5). However,  $\text{NO}_3^-$  generally peaked midday and decreased overnight, regardless of season (Table 2-2). Midday  $\text{NO}_3^-$  maxima despite high photoautotrophic potential may be due to instream storage lags, especially in larger watersheds (Burns et al., 2019; Cohen et al., 2012). At the downstream site,  $\text{NO}_3^-$  concentration tended to peak midday in spring runoff and overnight during the other two hydroperiods, resulting in positive diel  $\text{NO}_3^-$  values (Table 2-2). A period of high variability in diel  $\text{NO}_3^-$  change in August and September corresponded to erratic changes in  $\text{NO}_3^-$  that were unlikely actual diel signals (Figure 2-3, Figure S10).

The upstream Red Butte reach had very consistent diel  $\text{NO}_3^-$  patterns, with peaks overnight and decreases midday. Spring runoff changes were slightly less than summer changes, even though the stream is narrow and solar radiation to the stream is decreased by overhead canopy. The magnitude of diel swings decreased to essentially zero during winter low flow, in addition to overall lower concentrations (0.07 mg N/L to 0.035 mg N/L, Table 2-2). At the downstream site, the timing of  $\text{NO}_3^-$  peaks was similar, but the magnitude and variability increased. The summer growing season had nighttime peaks of 0.051 mg N/L above midday lows.

Nitrate concentrations at the upstream middle Provo site had little diel variability (as well as the smallest range of observed  $\text{NO}_3^-$  concentrations, Table 2-2), but a small midday decrease was detected during spring runoff (Figure 2-5). Downstream, no consistent pattern of diel  $\text{NO}_3^-$  occurred, although small periods of nighttime peaks can be seen during the lowest concentrations of winter low flow, and slight midday peaks in periods of the summer growing period. Diel changes in  $\text{NO}_3^-$  concentration (0.5 mg N/L) were lower at the downstream Provo site than at downstream site of the other two reaches.

### *Metabolism and Nitrate*

The three reaches were largely heterotrophic, which resulted in low estimates of GPP. Logan River had peak respiration during spring runoff, with  $3 \text{ g m}^{-2} \text{ d}^{-1}$ . The reach had slight increases in GPP and ER during winter compared to summer (Figure 2-6a). The metabolism estimates for Red Butte Creek (Figure 2-6b) had large uncertainty and negative rates because the small stream has low productivity and high gas flux, making it difficult for the program to accurately model. We updated the metabolism estimates using a binned model. Gas flux rates (K600) were determined experimentally upstream using SF6 methods and gave K600 values around  $9\text{-}17 \text{ d}^{-1}$  compared to values up to  $180 \text{ d}^{-1}$  using the streamMetabolizer package. However, correlations can still be calculated from the rough estimates that were generated, with some caution. GPP was very low (less than  $0.1 \text{ g m}^{-2} \text{ d}^{-1}$ ) with no seasonal trend. ER was usually between 0 and  $2 \text{ g m}^{-2} \text{ d}^{-1}$ , except for a few peaks during spring runoff and in winter that exceeded  $4 \text{ g m}^{-2} \text{ d}^{-1}$ . The middle Provo reach had the highest metabolic rates, which peaked in summer and late fall at  $4 \text{ g m}^{-2} \text{ d}^{-1}$  for GPP and  $10 \text{ g m}^{-2} \text{ d}^{-1}$  for ER (Figure 2-6c). Winter rates were around  $0.8$  and  $5 \text{ g m}^{-2} \text{ d}^{-1}$  for GPP and ER, respectively.

In Logan, GPP was a significant component of models across all hydroperiods (Table 2-3). In winter and summer, upstream Nitrate was included, although it was only significant in summer. Logan winter  $\text{NO}_3^-$  and metabolism had the highest  $R^2$  (0.38), while summer had the lowest (0.11). Because of the uncertainty in metabolic rates calculated with the streamMetabolizer, we calculated diel change in percent dissolved oxygen to include as a parameter in Red Butte diel change in  $\text{NO}_3^-$  (Figure S9), but %DO was not selected as a significant model parameter for any of the hydroperiods. Summer was the only Red Butte model with significant terms (GPP) and had an  $R^2$  of 0.05. None of the Provo models had significant terms, and  $R^2$  values remained below 0.02.

### *Hydrology and Nitrate*

In Logan, Spring runoff and summer hydroperiods had L-shaped C-Q curves (Figure 2-7a), suggesting initial dilution occurred, but a chemostatic condition was eventually reached. Winter low flows had more of a 1:1 relationship, and large, counterclockwise hysteresis loops can be seen at multiple points in the time series. Counterclockwise hysteresis patterns can reflect either dilution or concentration during the rising hydrograph (Burns et al., 2019).

Red Butte CQ relationships had little pattern, likely a factor of the flashy responses of the small watershed and stream (Figure 2-7b). Much of the C-Q relationship is peaks in concentration without any change in discharge. The spring runoff period had some hysteresis loops, but they were inconsistent in size, shape, and direction. The occasional hysteresis events in summer were also irregular, but in winter were generally counterclockwise.

The Provo C-Q graphs have multiple clusters joined by single lines transitioning between them (Figure 2-7c). This unusual behavior is artifacts of the managed dam releases and represent

the steps that are clearly seen in the hydrograph (Figure S1). Homogeneity of the dam outflows prevented the large hysteresis loops observed in the Logan reach, but some evidence of dilution can be seen. In Winter, snowmelt began at the end of March and decreased NO<sub>3</sub> concentrations even before an increase in discharge occurred.

In Logan, nitrate concentration was best predicted by sensor data in winter, specifically water temperature, precipitation, and discharge ( $R^2=0.55$ , Table 2-4). In summer, upstream nitrate had an  $R^2=0.28$ , but the p-value for nitrate was  $>0.05$ . No significant model was generated for Logan spring. In Red Butte, the top model for winter included chlorophyll  $\alpha$ , precipitation, upstream nitrate, and discharge ( $R^2=0.32$ ). In spring, only precipitation was included but the predictive power increased ( $R^2=0.45$ ). In summer, precipitation and upstream nitrate best explained nitrate concentrations ( $R^2=0.14$ ). In the middle Provo reach, winter nitrate concentration correlated with specific conductivity ( $R^2=0.29$ ), spring nitrate correlated slightly with chlorophyll  $\alpha$  ( $R^2=0.03$ ), and in summer nitrate correlated with precipitation and specific conductivity ( $R^2=0.79$ )

## DISCUSSION

### *When and How Much are These Reaches Gaining or Losing Nitrate?*

Most of the reaches increased in NO<sub>3</sub><sup>-</sup> concentration regardless of hydroperiod, with the exception of spring runoff in the Red Butte reach (Table 2-2). Because of decrease in discharge from the stream channel to groundwater, Red Butte during winter also slightly decreased in N load over the winter. Discharge also affected the N load in Logan, where spring runoff concentrations increased only slightly, but when compounded with increase in water volume,

represented an added 127 kg N/day. In Provo, the reach increase in  $\text{NO}_3^-$  concentration was highest in winter, unlike the other two watersheds which increased most in summer.

#### *How Much of the Change in Nitrate is Biological v Hydrologic?*

In the Logan reach, diel changes in  $\text{NO}_3^-$  were 3.2 to 8.7% of the total  $\text{NO}_3^-$  concentration, and only 11 to 38% of the diel change correlated with in-stream metabolic rates (specifically GPP; Table 2-2 and 3). The Logan reach was the only reach that had consistently negative  $\text{NO}_3^-$  changes, or a decrease in  $\text{NO}_3^-$  during the day. Diel patterns of daytime troughs of  $\text{NO}_3^-$  have been linked to assimilation driven by photosynthetic N uptake (Burns et al., 2019; Cohen et al., 2012; Rode, Halbedel née Angelstein, et al., 2016). Other processes that may have been responsible for diel patterns in  $\text{NO}_3^-$  in the Logan reach include photooxidation of organic nitrogen to  $\text{NO}_3^-$  (Sandford et al., 2007) and snowmelt dilution (Burns et al., 2019). The importance of snowmelt and mixing of sources (e.g. groundwater versus upstream) is suggested by the model of winter  $\text{NO}_3^-$  concentrations (Table 2-4), where 55% of  $\text{NO}_3^-$  correlated to hydrologic variables, including temperature, daily precipitation, and discharge. The importance of hydrologic processes in determining  $\text{NO}_3^-$  within the Logan reach are also evident from the C-Q relationships (Figure 2-7).

In Red Butte, precipitation and discharge were significant components of the model selected to explain  $\text{NO}_3^-$  concentrations, highlighting the role of hydrologic processes in this groundwater dominant system (Gabor et al., 2017; S. J. Hall et al., 2016). The increase in diel change from spring to summer, despite overhead canopy closure, suggests riparian and upland evapotranspiration (ET) may be responsible for the observed pattern (Pellerin et al., 2009; Rode, Wade, et al., 2016). Indirect biological processes may also manifest as a diel pattern; riparian

plant evapotranspiration may lower groundwater levels and thereby alter  $\text{NO}_3^-$  concentrations in nearby streams (Flewelling et al., 2014). An additional line of evidence supporting upslope processes over instream production is the tight correlation between water temperature and DO (Figure S2 and S3), suggesting that DO concentration was controlled by physical processes (dissolution from atmosphere to saturation) and instream biological activity was minimal.

We predicted that  $\text{NO}_3^-$  concentration would be negatively correlated with GPP and ER, to the extent that in-stream metabolism was responsible for N cycling. Whether due an actual ecological phenomenon or inaccuracies in metabolism predictions, models of diel change in  $\text{NO}_3^-$  with metabolic rates yielded little correlation. In Provo, the size of the watershed may result in lags in  $\text{NO}_3^-$  transport and obscure relationships between metabolism and  $\text{NO}_3^-$  concentrations.

Diel  $\Delta\text{NO}_3$  was 1.3 to 26.4% of seasonal  $\Delta\text{NO}_3^-$  across all three reaches (Table 2-2), compared to 75% reported in Grand Teton National Park (R. O. Hall & Tank, 2003). These values compare to 47% and 75% of daily  $\text{NO}_3^-$  load in agricultural and forested central Europe streams (Rode, Halbedel née Angelstein, et al., 2016), less than 20% N retention in a spring-fed Florida river (Heffernan & Cohen, 2010), and less than 10% of mean daily concentration (Burns et al., 2016), to >50% of the peak value (Moraetis et al., 2010).

### *How Much Nitrate May be Anthropogenic?*

Nitrate loading to the Logan reach during summer, in conjunction with an increase in specific conductivity and fDOM levels (Figure 2-3), strongly suggests an agricultural return flow from fields or farms within the watershed. While a relatively short time period, the elevated  $\text{NO}_3^-$  concentration represent an input of 2000 kg N (Table 2-2). The magnitude of average diel and

seasonal variability in  $\text{NO}_3^-$  at the upstream site (0.02 mg N/L and 0.15 mg N/L, respectively) suggests that the added  $\text{NO}_3^-$  (up to 1.56 mg N/L) far exceeds the capability of the stream to absorb or transform the  $\text{NO}_3^-$ . The inclusion of upstream  $\text{NO}_3^-$  concentration suggests an interaction between the amount of  $\text{NO}_3^-$  entering the reach and the capacity of instream productivity.

The Red Butte stream had large increases in  $\text{NO}_3^-$  concentration which correlated to precipitation inputs (Table 2-4), potentially indicating nutrient loading from storm water. Nitrate increases in this small, urban, and highly impervious watershed exceeded the removal capacity of the stream ecosystem.

In the middle Provo, a lack of disturbance from hydrologic modifications may have lessened the reach's ability to respond to process  $\text{NO}_3^-$ , chronic addition from WWTP or high groundwater  $\text{NO}_3^-$  concentrations with little removal within the riparian corridor may be responsible for winter increases in  $\text{NO}_3^-$ .

## CONCLUSION

Diel fluctuations in nitrate corresponded with stream metabolism in the Logan reach, but this only represented 10% of the total nitrate flux, demonstrating how in-stream uptake can easily be overwhelmed by nutrient loading in even moderately impacted watersheds. The amount of diel  $\text{NO}_3^-$  concentrations correlated with metabolism increased in winter. Instream metabolic processes may have more biogeochemical importance relative to terrestrial activity during winter in high latitudes, where soils may become dormant while frozen, but stream environments maintain low levels of activity. In the other reaches, net  $\text{NO}_3^-$  concentrations do not appear to be influenced by instream metabolic processes.

Nitrate concentrations within the Red Butte reach were driven by hydrologic processes. The Red Butte reach had some diel fluctuations, but these did not correspond to instream metabolism, and likely represented groundwater affected by upslope or riparian ET. This small, urban, and highly impervious watershed had large increases related to precipitation inputs, potentially from storm water. All three reaches generally increased in  $\text{NO}_3^-$  concentration (32% to 83%), except for Red Butte during spring runoff (lost 14%), possibly due to replacement of recent stream water with  $\text{NO}_3^-$  depleted groundwater.

The Provo reach had little diel fluctuations, and little of the  $\text{NO}_3^-$  associated with any biologic metrics (i.e., GPP, ER, and chlorophyll). Most of the  $\text{NO}_3^-$  was associated with hydrologic factors, including groundwater  $\text{NO}_3^-$  concentrations (indicated by increased  $\text{NO}_3^-$  reach loads during winter low flows, Figure 2-3c, Table 2-2).

Our results emphasize the need for improved management to prevent nitrogen pollution entering aquatic ecosystems. Pulses of nitrate corresponding to anthropogenic activity far exceed the capability of the system to remove or process the nitrogen. This dataset has generated hypotheses that can be further tested with additional measurements, experiments, tracers, and other targeted approaches (Burns et al., 2019).

## LITERATURE CITED

- Abbott, B. W., Moatar, F., Gauthier, O., Fovet, O., Antoine, V., & Ragueneau, O. (2018). Trends and seasonality of river nutrients in agricultural catchments: 18 years of weekly citizen science in France. *Science of The Total Environment*, 624, 845–858.  
<https://doi.org/10.1016/j.scitotenv.2017.12.176>
- Abbott, B. W., Bishop, K., Zarnetske, J. P., Hannah, D. M., Frei, R. J., Minaudo, C., et al. (2019). A water cycle for the Anthropocene. *Hydrological Processes*.  
<https://doi.org/10.1002/hyp.13544>
- Appling, A. P., Hall, R. O. J., Arroita, M., & Yackulic, C. B. (2017). *streamMetabolizer: Models for estimating aquatic photosynthesis and respiration*.
- Baker, E. B., & Showers, W. J. (2019). Hysteresis analysis of nitrate dynamics in the Neuse River, NC. *Science of The Total Environment*, 652, 889–899.  
<https://doi.org/10.1016/j.scitotenv.2018.10.254>
- Basu, N. B., Destouni, G., Jawitz, J. W., Thompson, S. E., Loukinova, N. V., Darracq, A., et al. (2010). Nutrient loads exported from managed catchments reveal emergent biogeochemical stationarity. *Geophysical Research Letters*, 37(23).  
<https://doi.org/10.1029/2010GL045168>
- Bende-Michl, U., Verburg, K., & Cresswell, H. P. (2013). High-frequency nutrient monitoring to infer seasonal patterns in catchment source availability, mobilisation and delivery. *Environmental Monitoring and Assessment*, 185(11), 9191–9219.  
<https://doi.org/10.1007/s10661-013-3246-8>
- Blaen, P. J., Khamis, K., Lloyd, C. E. M., Bradley, C., Hannah, D., & Krause, S. (2016). Real-time monitoring of nutrients and dissolved organic matter in rivers: Capturing event

dynamics, technological opportunities and future directions. *Science of The Total Environment*, 569–570, 647–660. <https://doi.org/10.1016/j.scitotenv.2016.06.116>

Blaen, P. J., Khamis, K., Lloyd, C., Comer-Warner, S., Ciocca, F., Thomas, R. M., et al. (2017). High-frequency monitoring of catchment nutrient exports reveals highly variable storm event responses and dynamic source zone activation: High-Frequency Storm Event Monitoring. *Journal of Geophysical Research: Biogeosciences*, 122(9), 2265–2281. <https://doi.org/10.1002/2017JG003904>

Burns, D. A., Miller, M. P., Pellerin, B. A., & Capel, P. D. (2016). Patterns of diel variation in nitrate concentrations in the Potomac River. *Freshwater Science*, 35(4), 1117–1132. <https://doi.org/10.1086/688777>

Burns, D. A., Pellerin, B. A., Miller, M. P., Capel, P. D., Tesoriero, A. J., & Duncan, J. M. (2019). Monitoring the riverine pulse: Applying high-frequency nitrate data to advance integrative understanding of biogeochemical and hydrological processes. *Wiley Interdisciplinary Reviews: Water*, 0(0), e1348. <https://doi.org/10.1002/wat2.1348>

Butturini, A., Alvarez, M., Bernal, S., Vazquez, E., & Sabater, F. (2008). Diversity and temporal sequences of forms of DOC and NO<sub>3</sub>-discharge responses in an intermittent stream: Predictable or random succession? *Journal of Geophysical Research: Biogeosciences*, 113(G3). <https://doi.org/10.1029/2008JG000721>

Cohen, M. J., Heffernan, J. B., Albertin, A., & Martin, J. B. (2012). Inference of riverine nitrogen processing from longitudinal and diel variation in dual nitrate isotopes: ISOTOPES TO INFER RIVER NITROGEN PROCESSING. *Journal of Geophysical Research: Biogeosciences*, 117(G1). <https://doi.org/10.1029/2011JG001715>

Dieter, C. A., Maupin, M. A., Caldwell, R. R., Harris, M. A., Ivahnenko, T. I., Lovelace, J. K., et al. (2018). *Estimated use of water in the United States in 2015* (USGS Numbered Series No. 1441) (p. 76). Reston, VA: U.S. Geological Survey. Retrieved from <http://pubs.er.usgs.gov/publication/cir1441>

Duff, J. H., & Triska, F. J. (2000). Nitrogen biogeochemistry and surface-subsurface exchange in streams.

Duncan, J. M., Welty, C., Kemper, J. T., Groffman, P. M., & Band, L. E. (2017). Dynamics of nitrate concentration-discharge patterns in an urban watershed. *Water Resources Research*, 53(8), 7349–7365. <https://doi.org/10.1002/2017WR020500>

Dupas, R., Minaudo, C., & Abbott, B. W. (2019). Stability of spatial patterns in water chemistry across temperate ecoregions. *Environmental Research Letters*, 14(7), 074015. <https://doi.org/10.1088/1748-9326/ab24f4>

Ehleringer, J. R., Arnow, L. A., Arnow, T., McNulty, I. B., McNulty, I. R., & Negus, N. C. (1992). RED BUTTE CANYON RESEARCH NATURAL AREA: HISTORY, FLORA, GEOLOGY, CLIMATE, AND ECOLOGY. *The Great Basin Naturalist*, 52(2), 95–121. <https://doi.org/10.2307/41712704>

Evans, C., & Davies, T. D. (1998). Causes of concentration/discharge hysteresis and its potential as a tool for analysis of episode hydrochemistry. *Water Resources Research*, 34(1), 129–137. <https://doi.org/10.1029/97WR01881>

Flewelling, S. A., Hornberger, G. M., Herman, J. S., Mills, A. L., & Robertson, W. M. (2014). Diel patterns in coastal-stream nitrate concentrations linked to evapotranspiration in the riparian zone of a low-relief, agricultural catchment. *Hydrological Processes*, 28(4), 2150–2158. <https://doi.org/10.1002/hyp.9763>

Foley, J. A., Ramankutty, N., Brauman, K. A., Cassidy, E. S., Gerber, J. S., Johnston, M., et al. (2011). Solutions for a cultivated planet. *Nature*, 478(7369), 337–342.

<https://doi.org/10.1038/nature10452>

Gabor, R. S., Hall, S. J., Eiriksson, D. P., Jameel, Y., Millington, M., Stout, T., et al. (2017).

Persistent Urban Influence on Surface Water Quality via Impacted Groundwater.

*Environmental Science & Technology*, 51(17), 9477–9487.

<https://doi.org/10.1021/acs.est.7b00271>

Grace, M., & Imberger, S. (2006). Stream Metabolism: Performing & Interpreting Measurements. *Water Studies Centre Monash University, Murray Darling Basin Commission and New South Wales Department of Environment and Climate Change*.

Hall, R. O., & Tank, J. L. (2003). Ecosystem metabolism controls nitrogen uptake in streams in Grand Teton National Park, Wyoming. *Limnology and Oceanography*, 48(3), 1120–1128.

<https://doi.org/10.4319/lo.2003.48.3.1120>

Hall, S. J., Weintraub, S. R., Eiriksson, D., Brooks, P. D., Baker, M. A., Bowen, G. J., & Bowling, D. R. (2016). Stream Nitrogen Inputs Reflect Groundwater Across a Snowmelt-Dominated Montane to Urban Watershed. *Environmental Science & Technology*, 50(3), 1137–1146. <https://doi.org/10.1021/acs.est.5b04805>

Heffernan, J. B., & Cohen, M. J. (2010). Direct and indirect coupling of primary production and diel nitrate dynamics in a subtropical spring-fed river. *Limnology and Oceanography*, 55(2), 677–688. <https://doi.org/10.4319/lo.2010.55.2.0677>

Hill, A. R. (2000). Stream chemistry and riparian zones. *Streams and Ground Waters*, 83–110.

Jarvie, H. P., Sharpley, A. N., Kresse, T., Hays, P. D., Williams, R. J., King, S. M., & Berry, L.

G. (2018). Coupling High-Frequency Stream Metabolism and Nutrient Monitoring to Explore Biogeochemical Controls on Downstream Nitrate Delivery. *Environmental Science & Technology*, 52(23), 13708–13717. <https://doi.org/10.1021/acs.est.8b03074>

Jones, A. S., Aanderud, Z. T., Horsburgh, J. S., Eiriksson, D. P., Dastrup, D., Cox, C., et al.

(2017). Designing and Implementing a Network for Sensing Water Quality and Hydrology across Mountain to Urban Transitions. *JAWRA Journal of the American Water Resources Association*, 53(5), 1095–1120. <https://doi.org/10.1111/1752-1688.12557>

Jones, C. S., Kim, S., Wilton, T. F., Schilling, K. E., & Davis, C. A. (2018). Nitrate uptake in an agricultural stream estimated from high-frequency, in-situ sensors. *Environmental Monitoring and Assessment*, 190(4), 226. <https://doi.org/10.1007/s10661-018-6599-1>

Kadlec, J., StClair, B., Ames, D. P., & Gill, R. A. (2015). WaterML R package for managing ecological experiment data on a CUAHSI HydroServer. *Ecological Informatics*, 28, 19–28. <https://doi.org/10.1016/j.ecoinf.2015.05.002>

Kirchner, J. W., Feng, X., Neal, C., & Robson, A. J. (2004). The fine structure of water-quality dynamics: the (high-frequency) wave of the future. *Hydrological Processes*, 18(7), 1353–1359. <https://doi.org/10.1002/hyp.5537>

Moatar, F., Abbott, B. W., Minaudo, C., Curie, F., & Pinay, G. (2017). Elemental properties, hydrology, and biology interact to shape concentration-discharge curves for carbon, nutrients, sediment, and major ions. *Water Resources Research*, 53(2), 1270–1287.

- Moraetis, D., Efstathiou, D., Stamati, F., Tzoraki, O., Nikolaidis, N. P., Schnoor, J. L., & Vozinakis, K. (2010). High-frequency monitoring for the identification of hydrological and bio-geochemical processes in a Mediterranean river basin. *Journal of Hydrology*, 389(1), 127–136. <https://doi.org/10.1016/j.jhydrol.2010.05.037>
- Pellerin, B. A., Downing, B. D., Kendall, C., Dahlgren, R. A., Kraus, T. E. C., Saraceno, J., et al. (2009). Assessing the sources and magnitude of diurnal nitrate variability in the San Joaquin River (California) with an in situ optical nitrate sensor and dual nitrate isotopes. *Freshwater Biology*, 54(2), 376–387. <https://doi.org/10.1111/j.1365-2427.2008.02111.x>
- Randall, M. C., Carling, G. T., Dastrup, D. B., Miller, T., Nelson, S. T., Rey, K. A., et al. (2019). Sediment potentially controls in-lake phosphorus cycling and harmful cyanobacteria in shallow, eutrophic Utah Lake. *PLOS ONE*, 14(2), e0212238. <https://doi.org/10.1371/journal.pone.0212238>
- Rode, M., Halbedel née Angelstein, S., Anis, M. R., Borchardt, D., & Weitere, M. (2016). Continuous In-Stream Assimilatory Nitrate Uptake from High-Frequency Sensor Measurements. *Environmental Science & Technology*, 50(11), 5685–5694. <https://doi.org/10.1021/acs.est.6b00943>
- Rode, M., Wade, A. J., Cohen, M. J., Hensley, R. T., Bowes, M. J., Kirchner, J. W., et al. (2016). Sensors in the Stream: The High-Frequency Wave of the Present. *Environmental Science & Technology*, 50(19), 10297–10307. <https://doi.org/10.1021/acs.est.6b02155>
- Sandford, R. C., Exenberger, A., & Worsfold, P. J. (2007). Nitrogen Cycling in Natural Waters using In Situ, Reagentless UV Spectrophotometry with Simultaneous Determination of Nitrate and Nitrite. *Environmental Science & Technology*, 41(24), 8420–8425. <https://doi.org/10.1021/es071447b>

Schwientek, M., Osenbrück, K., & Fleischer, M. (2013). Investigating hydrological drivers of nitrate export dynamics in two agricultural catchments in Germany using high-frequency data series. *Environmental Earth Sciences*, 69(2), 381–393.

<https://doi.org/10.1007/s12665-013-2322-2>

Vaughan, M. C. H., Bowden, W. B., Shanley, J. B., Vermilyea, A., Sleeper, R., Gold, A. J., et al. (2017). High-frequency dissolved organic carbon and nitrate measurements reveal differences in storm hysteresis and loading in relation to land cover and seasonality.

*Water Resources Research*, 53(7), 5345–5363. <https://doi.org/10.1002/2017WR020491>

Wanninkhof, R., Mulholland, P. J., & Elwood, J. W. (1990). Gas exchange rates for a first-order stream determined with deliberate and natural tracers. *Water Resources Research*, 26(7), 1621–1630. <https://doi.org/10.1029/WR026i007p01621>

Wollheim, W. M., Mulukutla, G. K., Cook, C., & Carey, R. O. (2017). Aquatic Nitrate Retention at River Network Scales Across Flow Conditions Determined Using Nested In Situ Sensors. *Water Resources Research*, 53(11), 9740–9756.

<https://doi.org/10.1002/2017WR020644>

Zimmer, M. A., Pellerin, B., Burns, D. A., & Petrochenkov, G. (2019). Temporal Variability in Nitrate-Discharge Relationships in Large Rivers as Revealed by High-Frequency Data.

*Water Resources Research*, 55(2), 973–989. <https://doi.org/10.1029/2018WR023478>

## FIGURES

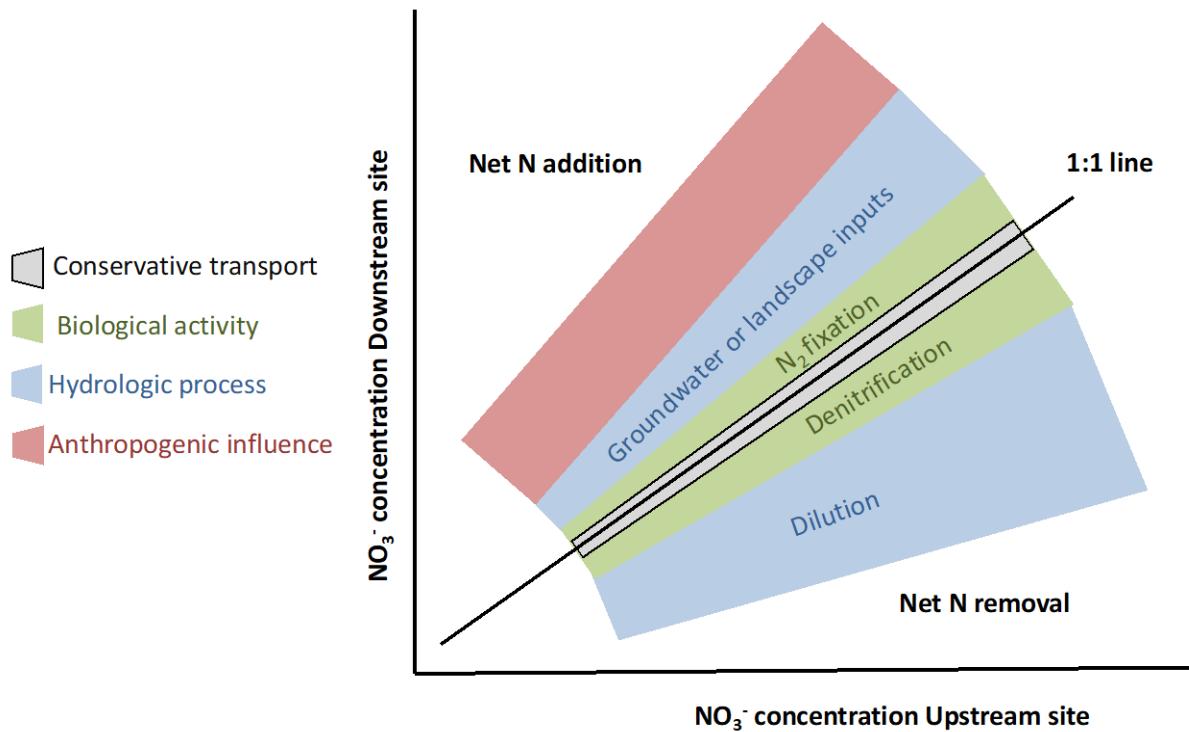


Figure 2-1. Theoretical diagram of possible nitrate concentration relationship between upstream and downstream sites. Along the 1:1 line represents net conservative transport, with deviations in some range of net N addition or N removal due to: biological activity, such as N<sub>2</sub> fixation or denitrification; natural hydrologic processes, such as groundwater or overland sources, or dilution; and anthropogenic influence.

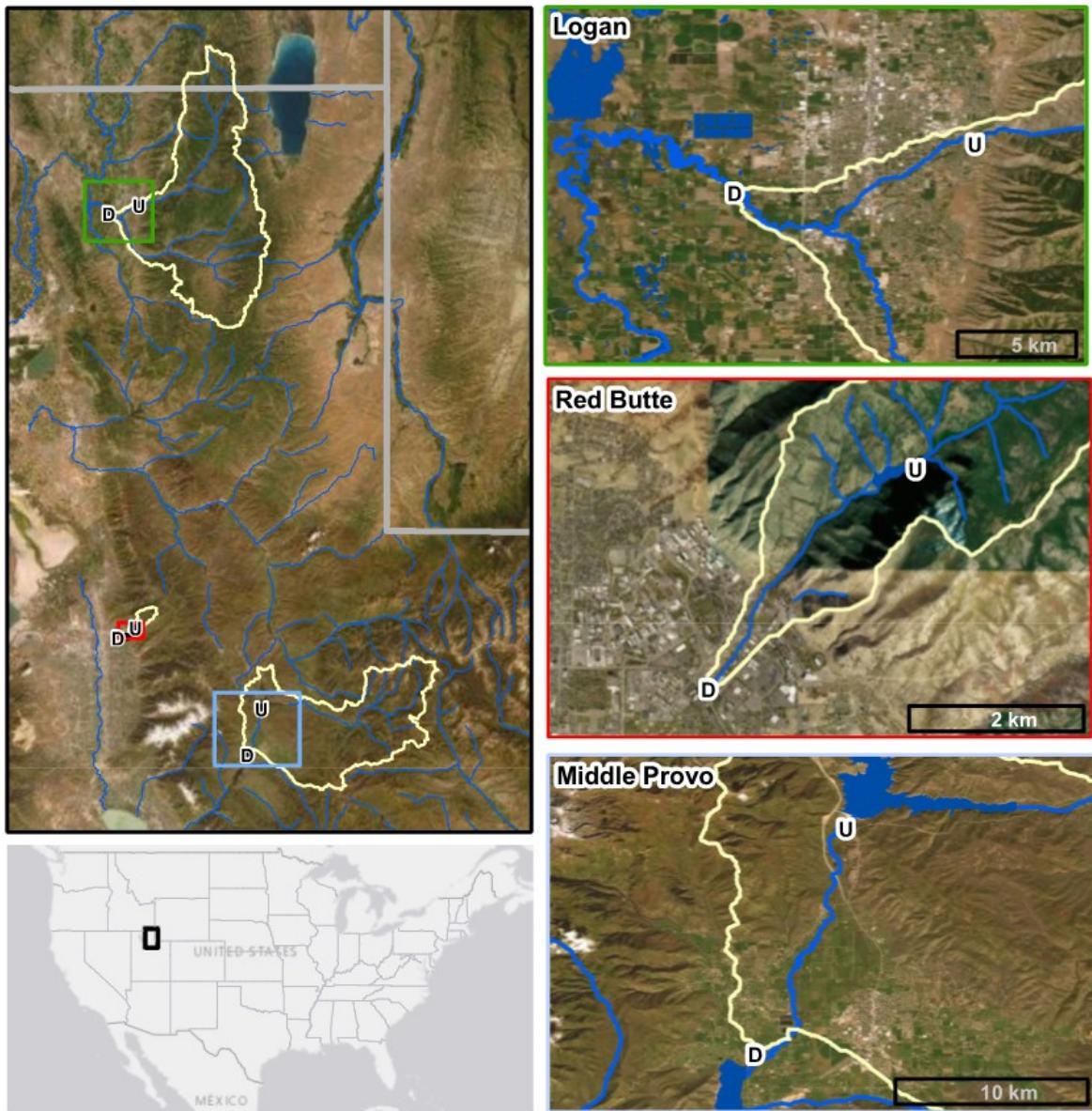


Figure 2-2. Map showing study reaches with Upstream (U) and Downstream (D) sensor locations in three Wasatch watersheds (Logan, Red Butte, and middle Provo). Source credits: Esri, USGS, and NOAA.

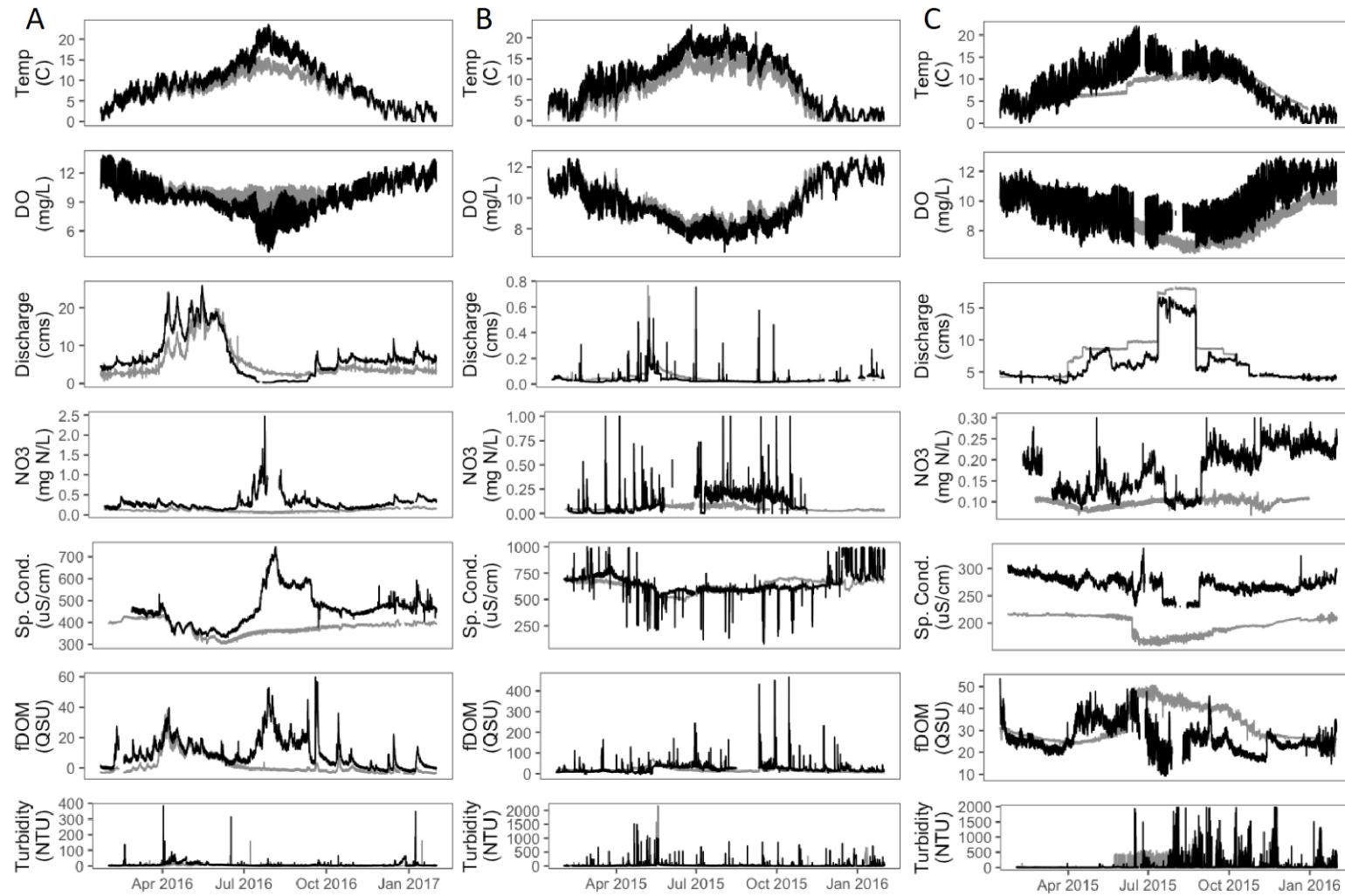


Figure 2-3. Time series for one year of sensor data from upstream (grey) and downstream (black) sites in Logan (A), Red Butte (B), and middle Provo (C) watersheds. Parameters measured include temperature, dissolved oxygen (DO), discharge, NO<sub>3</sub><sup>-</sup>, specific conductivity (Sp. Cond.), fluorescent dissolved organic matter (fDOM), and turbidity. Summary statistics of values by hydroperiod can be found in Table S1.

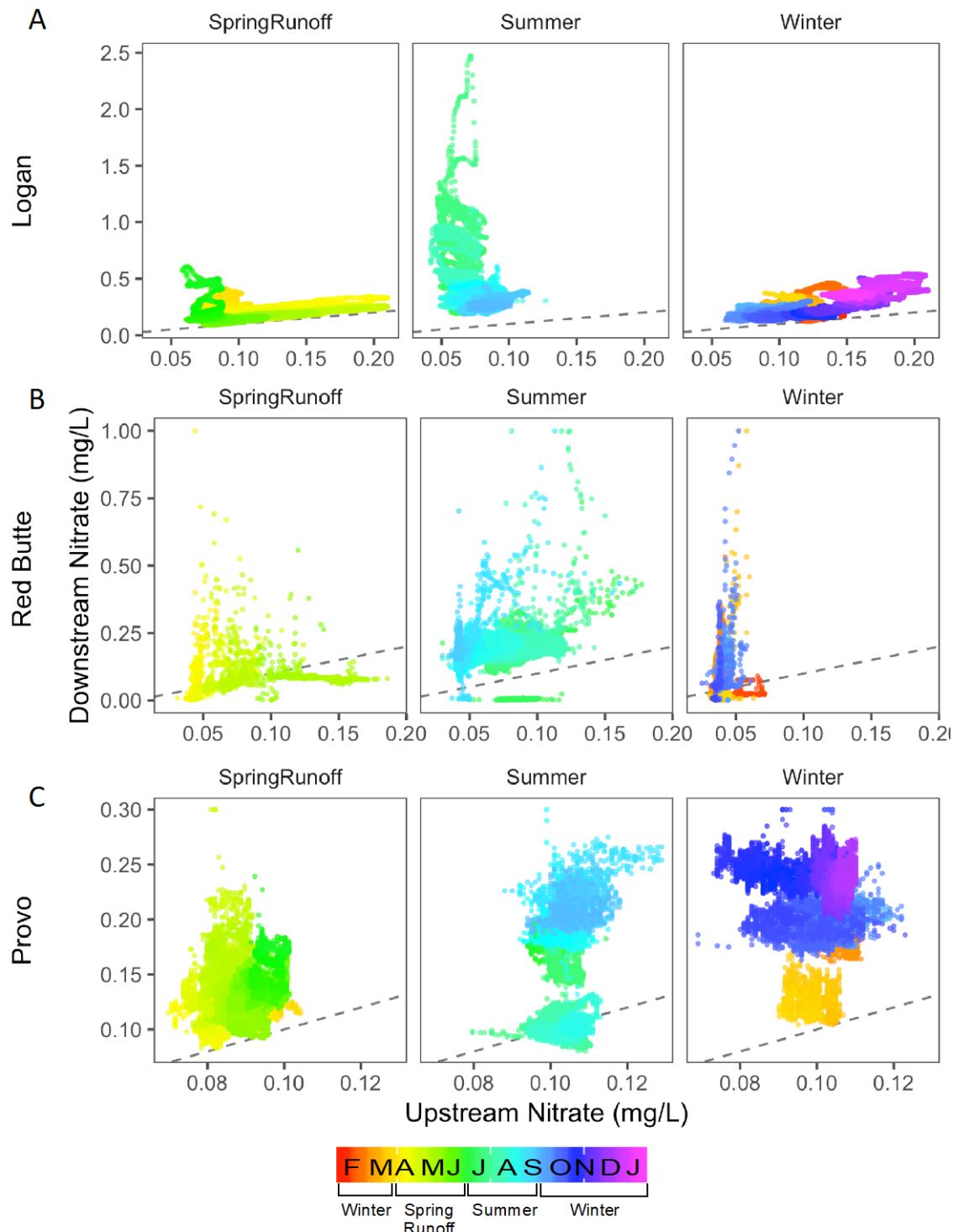


Figure 2-4. Nitrate concentrations at upstream versus downstream sites for three Wasatch watersheds (Logan, A; Red Butte, B; Provo, C) demonstrate periods of net N-addition and removal across periods of the hydrograph (Spring Runoff, Summer growing period, and Winter low flow).

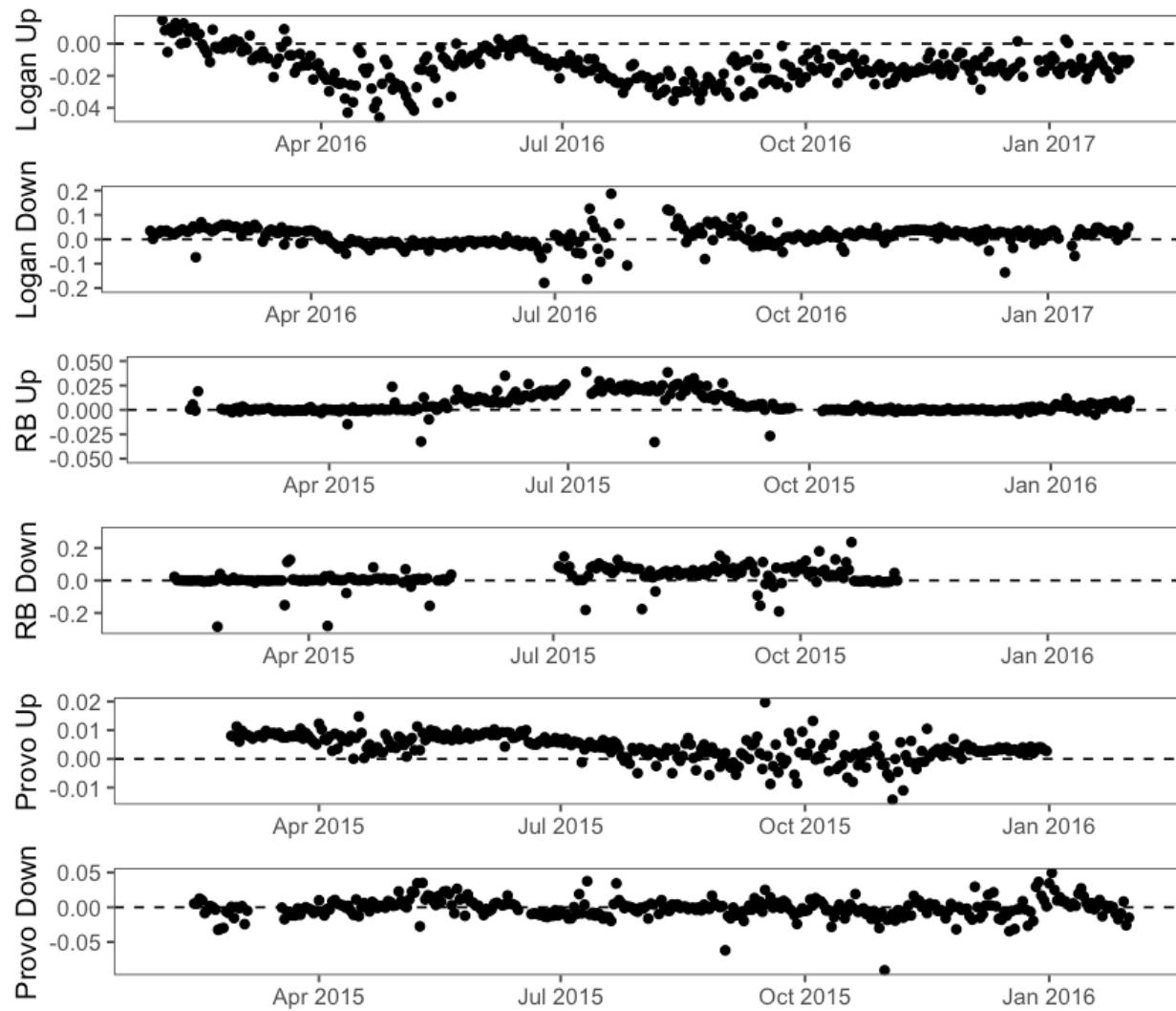


Figure 2-5. Time series of diel  $\Delta\text{NO}_3^-$  values at upstream (Up) and downstream (Down) sites in Logan, Red Butte, and Provo watersheds. Diel  $\text{NO}_3^-$  values were calculated by subtracting the average concentration during the 1400 hour from the 0200 hour; positive values indicate net daytime  $\text{NO}_3^-$  assimilation.

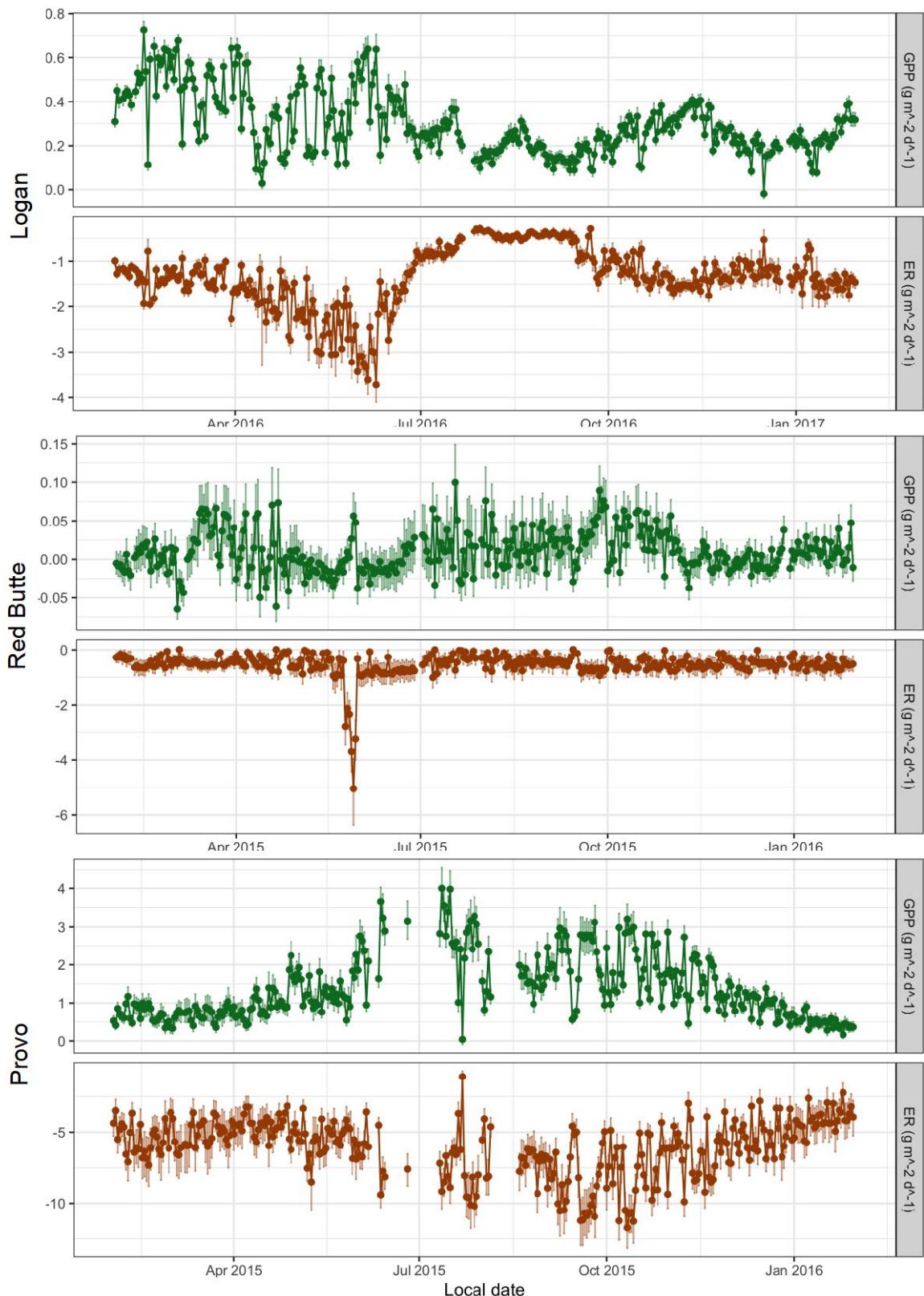


Figure 2-6. One year of stream metabolism (i.e. gross primary productivity, GPP, and ecosystem respiration, ER) for three Wasatch stream reaches (Logan, Red Butte, and middle Provo).

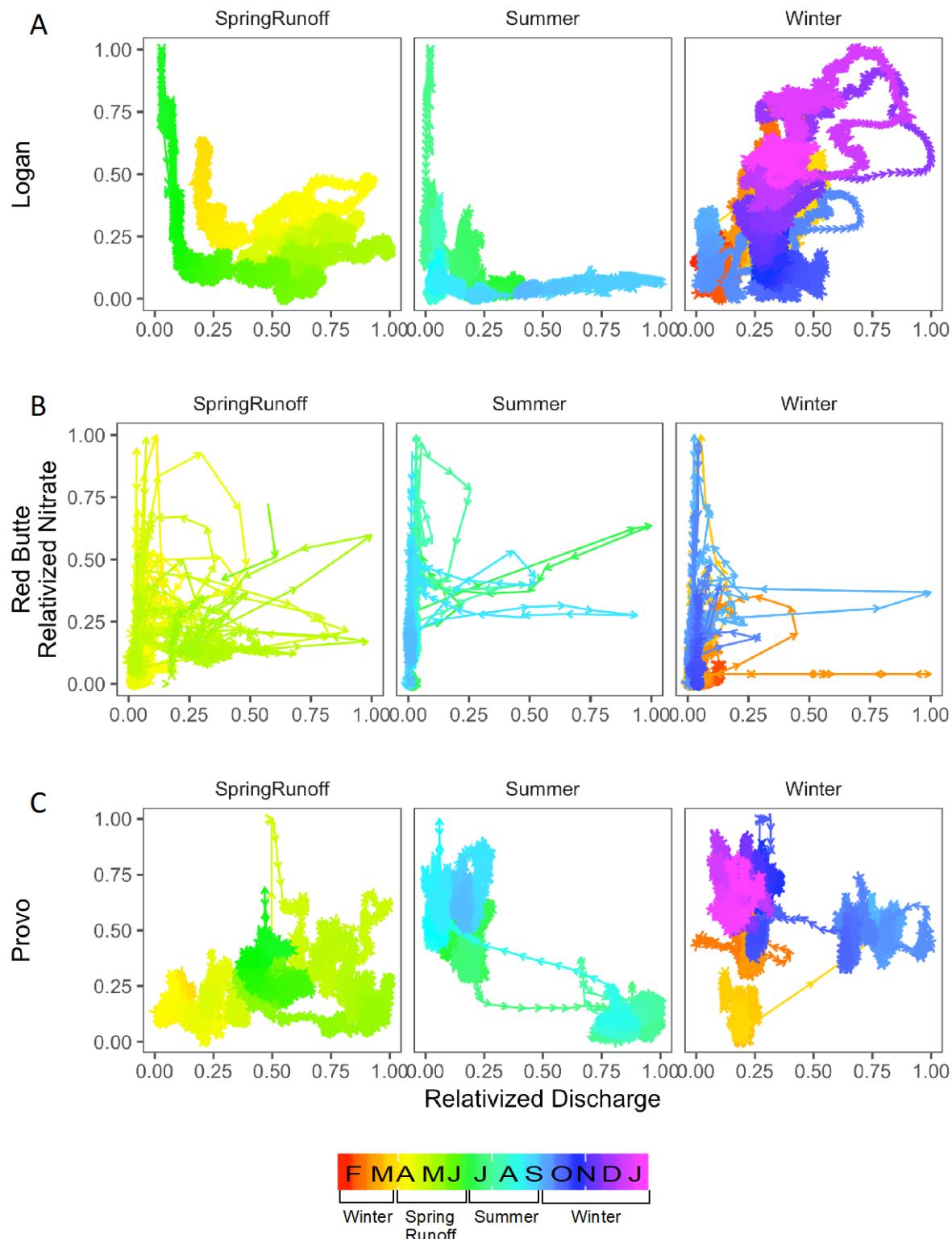


Figure 2-7. Normalized concentration-discharge (C-Q) relationships for  $\text{NO}_3^-$  in three Wasatch watersheds (Logan, A; Red Butte, B; and middle Provo, C) showing seasonal differences in hysteresis behavior.

## TABLES

Table 2-1. Site metadata for upstream (Up) and downstream (Down) locations of reaches within three Wasatch watersheds (Logan, Red Butte, and middle Provo). Stream length indicates distance from the upstream site. Distance integrated was calculated using velocity and K600 values estimated by streamMetabolizer or measured using SF<sub>6</sub> injection.

| <b>Watershed/<br/>Location</b> | <b>Elev. (masl)</b> | <b>Watershed<br/>area (km<sup>2</sup>)</b> | <b>Stream length<br/>(km)</b> | <b>Distance integrated<br/>(km)</b> |
|--------------------------------|---------------------|--|-------------------------------|-------------------------------------|
| Logan Up                       | 1414                | 557  | 0                             | -                                   |
| Logan Down                     | 1353                | 1367                                       | 17.3                          | 1.6 – 14.5                          |
| Red Butte Up                   | 1649                | 18.6                                       | 0                             | -                                   |
| Red Butte Down                 | 1449                | 22.8                                       | 3.7                           | 0.8 – 8.7                           |
| Provo Up                       | 1790                | 678  | 0                             | -                                   |
| Provo Down                     | 1658                | 976  | 20.4                          | 6.0 – 8.5                           |

Table 2-2. Summary statistics for NO<sub>3</sub><sup>-</sup> concentrations for upstream and downstream sites in three Wasatch watersheds (Logan, Red Butte, and Provo).

|  | Logan up  | Logan down  | Red Butte up  | Red Butte down   | Provo up   | Provo down   |
|--|---|---|---|--|--|--|
| <b>Min.-max. NO<sub>3</sub><sup>-</sup> (mg N/L)</b>                                   | 0.051-0.201   | 0.105-1.67  | 0.023-0.192   | 0.003-0.401  | 0.077-0.112  | 0.087-0.260  |
| <b>Range NO<sub>3</sub><sup>-</sup> (mg N/L)</b>                                       | 0.15  | 1.56  | 0.169   | 0.398  | 0.035  | 0.173  |
| <b>Average NO<sub>3</sub><sup>-</sup> conc.<br/>± s.d. (mg N/L)</b>                    | <b>Spring Runoff</b><br><i>Summer</i> 0.071 ± 0.013<br><i>Winter</i> 0.128 ± 0.033  | <b>0.202 ± 0.074</b><br><i>Summer</i> 0.44 ± 0.28<br><i>Winter</i> 0.264 ± 0.087    | <b>0.07 ± 0.025</b><br><i>Summer</i> 0.074 ± 0.023<br><i>Winter</i> 0.035 ± 0.007 | <b>0.055 ± 0.059</b><br><i>Summer</i> 0.193 ± 0.081<br><i>Winter</i> 0.063 ± 0.085 | <b>0.089 ± 0.007</b><br><i>Summer</i> 0.104 ± 0.004<br><i>Winter</i> 0.101 ± 0.008 | <b>0.131 ± 0.025</b><br><i>Summer</i> 0.153 ± 0.051<br><i>Winter</i> 0.212 ± 0.036   |
| <b>Diel NO<sub>3</sub><sup>-</sup> conc.<br/>± s.d. (mg N/L)</b>                       | <b>Spring Runoff</b><br><i>Summer</i> -0.017 ± 0.012<br><i>Winter</i> -0.01 ± 0.009 | <b>-0.017 ± 0.025</b><br><i>Summer</i> 0.014 ± 0.153<br><i>Winter</i> 0.023 ± 0.025 | <b>0.007 ± 0.01</b><br><i>Summer</i> 0.015 ± 0.018<br><i>Winter</i> 0.001 ± 0.002 | <b>0.005 ± 0.073</b><br><i>Summer</i> 0.051 ± 0.059<br><i>Winter</i> 0.015 ± 0.059 | <b>0.007 ± 0.002</b><br><i>Summer</i> 0.002 ± 0.004<br><i>Winter</i> 0.003 ± 0.004 | <b>0.003 ± 0.011</b><br><i>Summer</i> -0.002 ± 0.012<br><i>Winter</i> -0.003 ± 0.015 |
| <b>% of overall NO<sub>3</sub><sup>-</sup><br/>conc. change in diel</b>                | <b>Spring runoff</b><br><i>Summer</i> -14.7<br><i>Winter</i> -29.6                  | <b>-8.4</b><br><i>Summer</i> 3.2<br><i>Winter</i> 8.7                               | <b>10.0</b><br><i>Summer</i> 20.3<br><i>Winter</i> 2.9                            | <b>9.1</b><br><i>Summer</i> 26.4<br><i>Winter</i> 23.8                             | <b>7.9</b><br><i>Summer</i> 1.9<br><i>Winter</i> 3.0                               | <b>2.3</b><br><i>Summer</i> -1.3<br><i>Winter</i> -1.4                               |
|  |   | Logan   | Red Butte   | Provo  |  |  |
| <b>Average daily reach<br/>conc. change (mg<br/>N/L)</b>                               | <b>Spring Runoff</b><br><i>Summer</i><br><i>Winter</i>                              | 0.084<br>0.363<br>0.135   | -0.007<br>0.118<br>0.018  |  | 0.043<br>0.049<br>0.109  |  |
| <b>Average daily reach<br/>load (kg N/d)</b>   | <b>Spring Runoff</b><br><i>Summer</i><br><i>Winter</i>                              | 127<br>27.1<br>101  | -0.241<br>0.183<br>-0.002   |  | 7.39<br>4.87<br>44.6   |  |
| <b>Annual reach load<br/>(kg N)</b>  | <b>Spring Runoff</b><br><i>Summer</i><br><i>Winter</i>                              | 11401<br>2058<br>17458  | -12.6<br>14.3<br>-0.149   |  | 635<br>443<br>4866   |  |
| <b>Change in NO<sub>3</sub><sup>-</sup> conc.<br/>from up to down as<br/>% of down</b> | <b>Spring Runoff</b><br><i>Summer</i><br><i>Winter</i>                              | 42.1<br>83.2<br>51.5  | -13.7<br>61.6<br>32.0   |  | 32.6<br>32.2<br>51.9   |  |

Table 2-3. GLS models correlating diel change in  $\text{NO}_3^-$  concentration using daily gross primary productivity (GPP), daily ecosystem respiration (ER), and daily average upstream nitrate (UpNitrate). Parameters listed show terms selected in the best model by AICc score, \*denotes the term had a significant p-value. We list model  $R^2$  and phi1 (the autocorrelation term).

| <b>Watershed/<br/>hydroperiod</b> | <b>Phi1</b> | <b>Selected parameters *p-value&lt;0.05</b> | <b>R<sup>2</sup></b> |
|-----------------------------------|-------------|---|----------------------|
| Logan                             |             |   |                      |
| Winter                            | 0.11        | (int)*, GPP*, UpNitrate                     | 0.38                 |
| Spring runoff                     | 0.53        | int*, GPP*                                  | 0.17                 |
| Summer                            | -0.08       | int*, GPP*, UpNitrate*                      | 0.11                 |
| Red Butte                         |             |   |                      |
| Winter                            | 0.04        | int, GPP, UpNitrate                         | 0.09                 |
| Spring runoff                     | -0.04       | int, UpNitrate                              | 0                    |
| Summer                            | 0.17        | int*, GPP*                                  | 0.05                 |
| Provo                             |             |   |                      |
| Winter                            | 0.28        | int, UpNitrate                              | 0                    |
| Spring runoff                     | 0.07        | int, UpNitrate                              | 0.02                 |
| Summer                            | 0.28        | int, UpNitrate                              | 0.02                 |

Table 2-4. GLS models predicting  $\text{NO}_3^-$  concentration in three reaches for three hydroperiods using daily average water temperature (temp), sum of daily precipitation (precip), daily average chlorophyll  $\alpha$  (chla), daily average specific conductivity (SpCond), upstream nitrate (UpNitrate), and daily average discharge (Q). Parameters listed show terms selected in the best model by AICc score, \*denotes the term had a significant p-value. We list model  $R^2$  and phi1 (the autocorrelation term).

| <b>Watershed/<br/>hydroperiod</b> | <b>Phi1</b>   | <b>Selected parameters *p-value&lt;0.05</b> | <b>R<sup>2</sup></b> |
|-----------------------------------|---------------|---|----------------------|
| Logan                             |               |   |                      |
|                                   | Winter        | 0.97 int*, temp*, precip*, Q*               | 0.55                 |
|                                   | Spring runoff | 0.99 int, precip                            | 0                    |
| Red Butte                         | Summer        | 0.88 int*, UpNitrate                        | 0.28                 |
|                                   | Winter        | 0.94 int, chla*, precip, UpNitrate, Q*      | 0.32                 |
|                                   | Spring runoff | 0.75 int*, precip*                          | 0.45                 |
| Middle Provo                      | Summer        | 0.84 int, precip*, UpNitrate*               | 0.14                 |
|                                   | Winter        | 0.99 int, SpCond*                           | 0.29                 |
|                                   | Spring runoff | 0.9 int*, chla*                             | 0.03                 |
|                                   | Summer        | 0.98 int*, precip*, SpCond*                 | 0.79                 |

## SUPPLEMENTAL MATERIAL

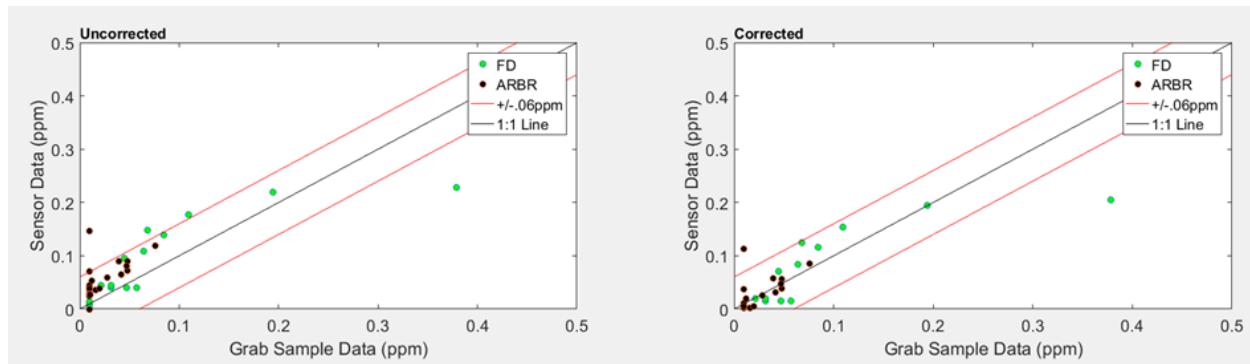


Figure S1. N QA/QC for Red Butte.

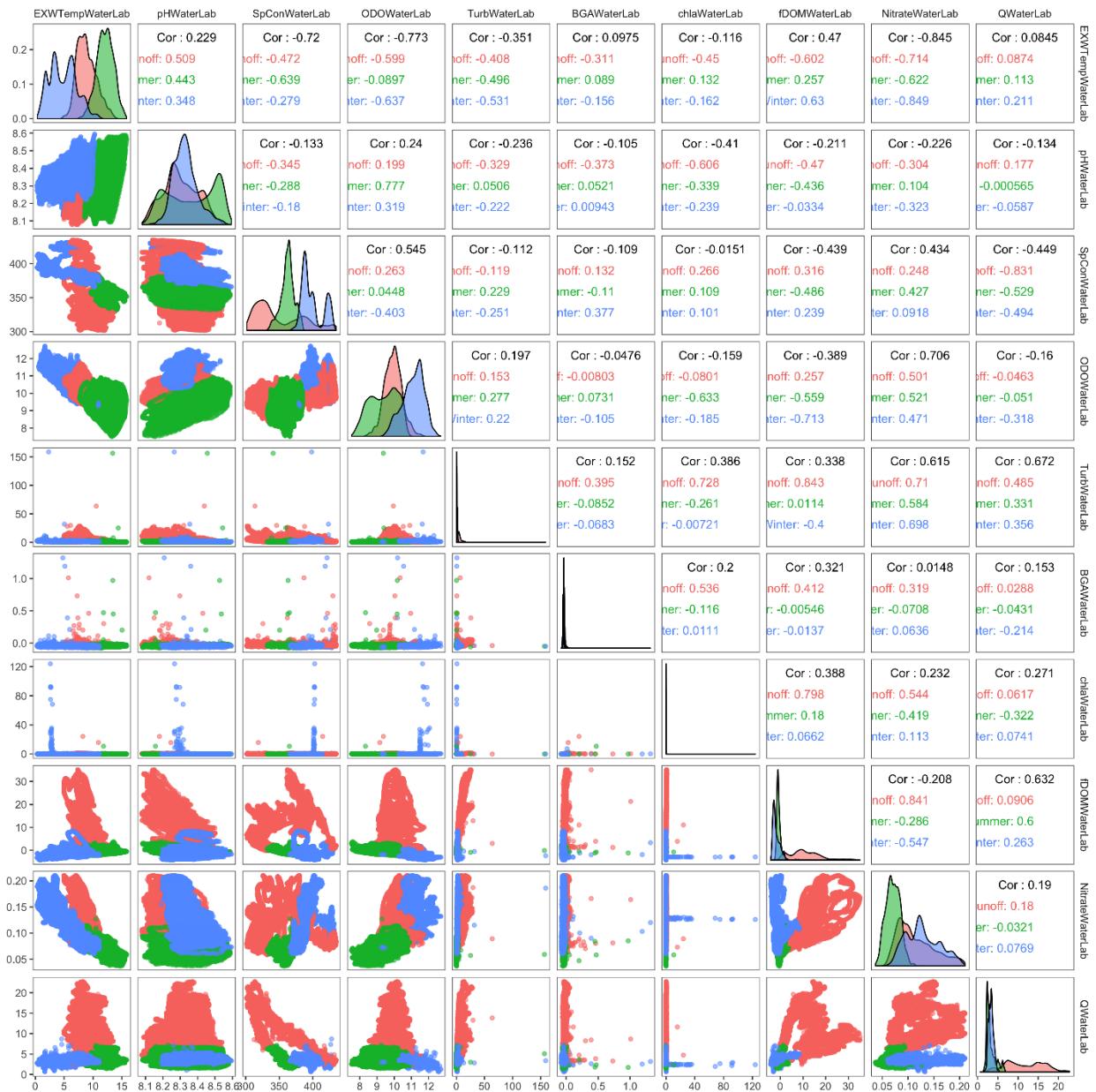


Figure S2. Correlation of sensor data at upstream Logan site, using Pearson's correlation.

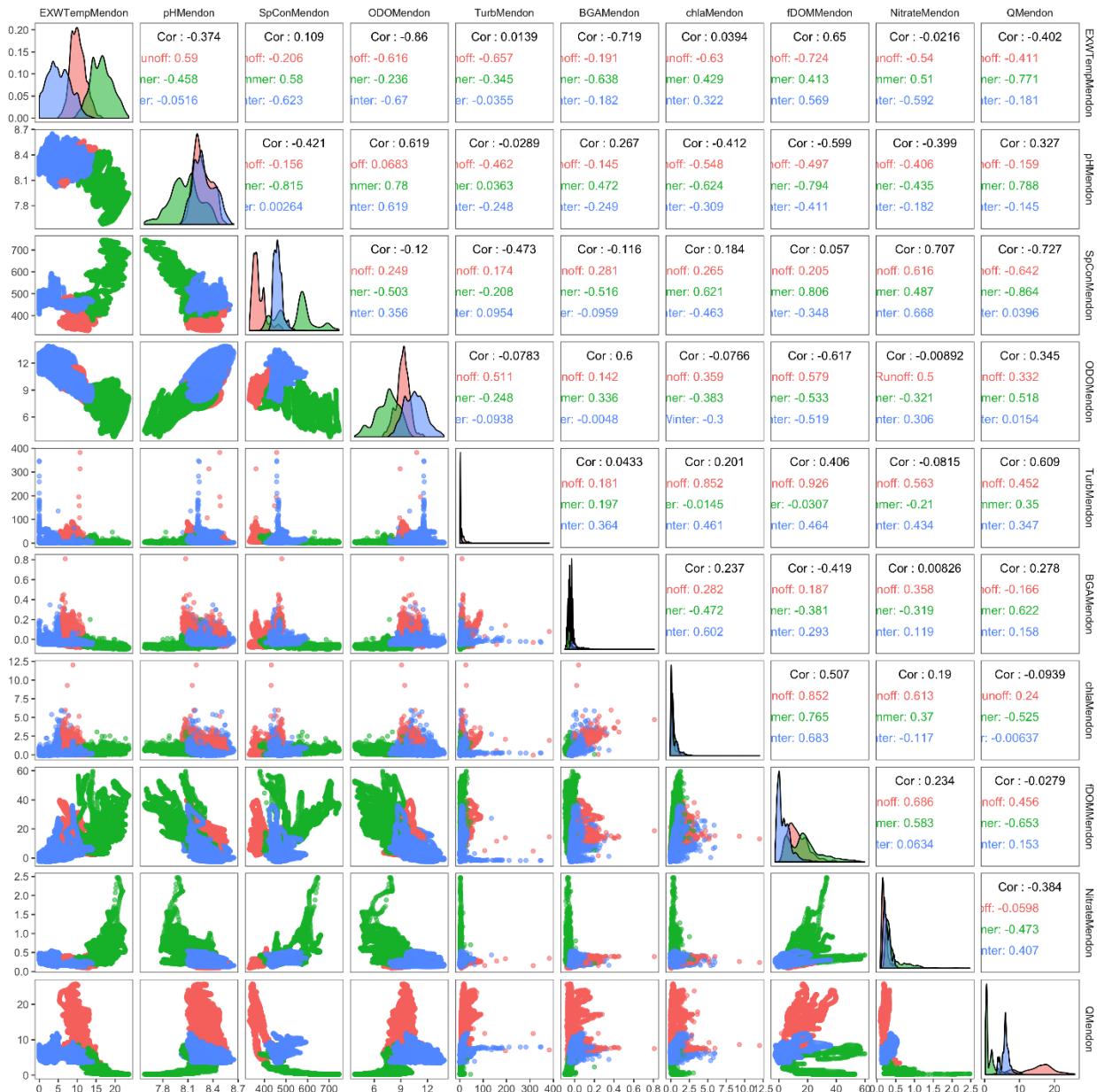


Figure S3. Correlation of sensor data at downstream Logan site, using Pearson's correlation.

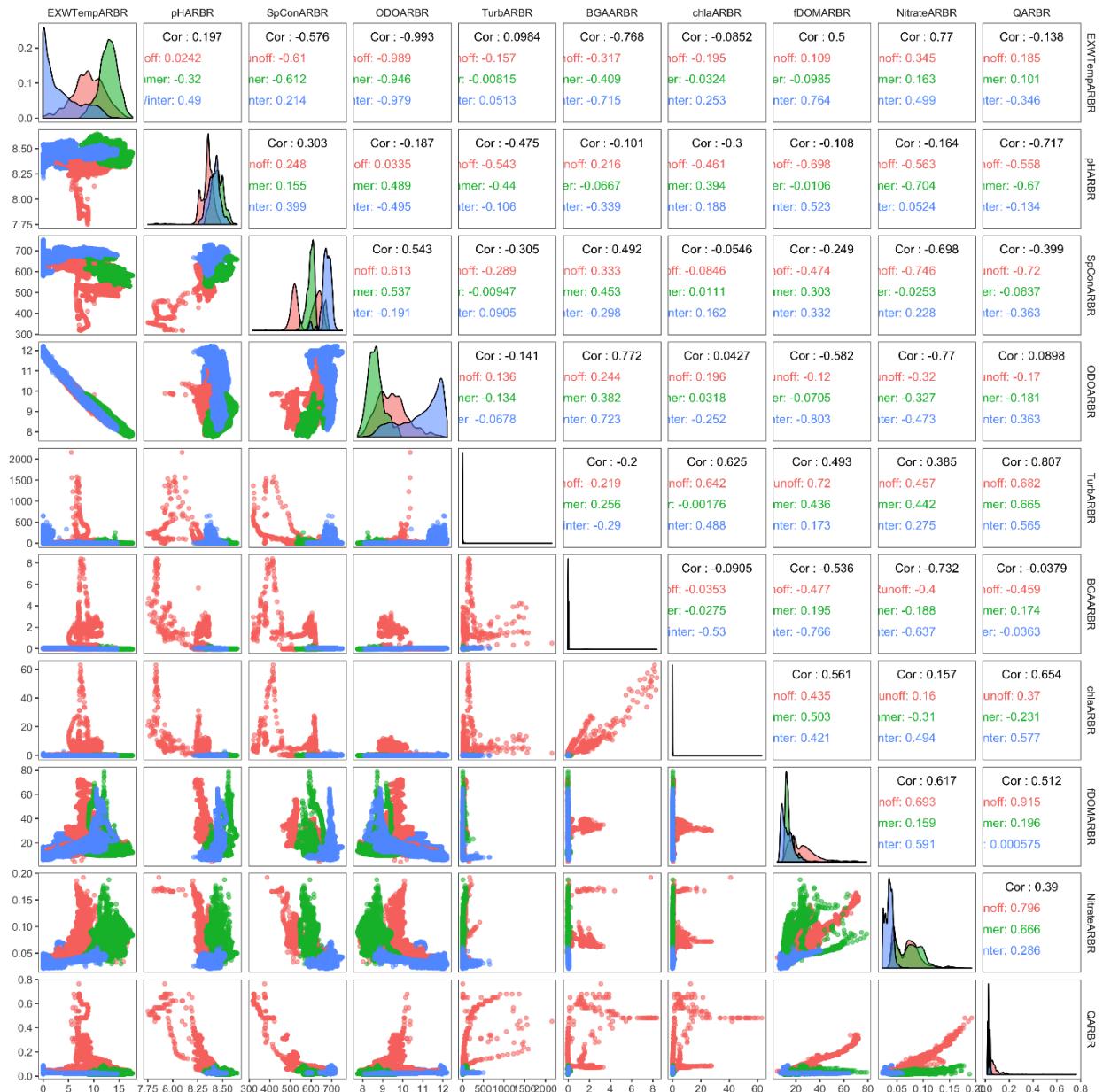


Figure S4. Correlation of sensor data at upstream Red Butte site, using Pearson's correlation.

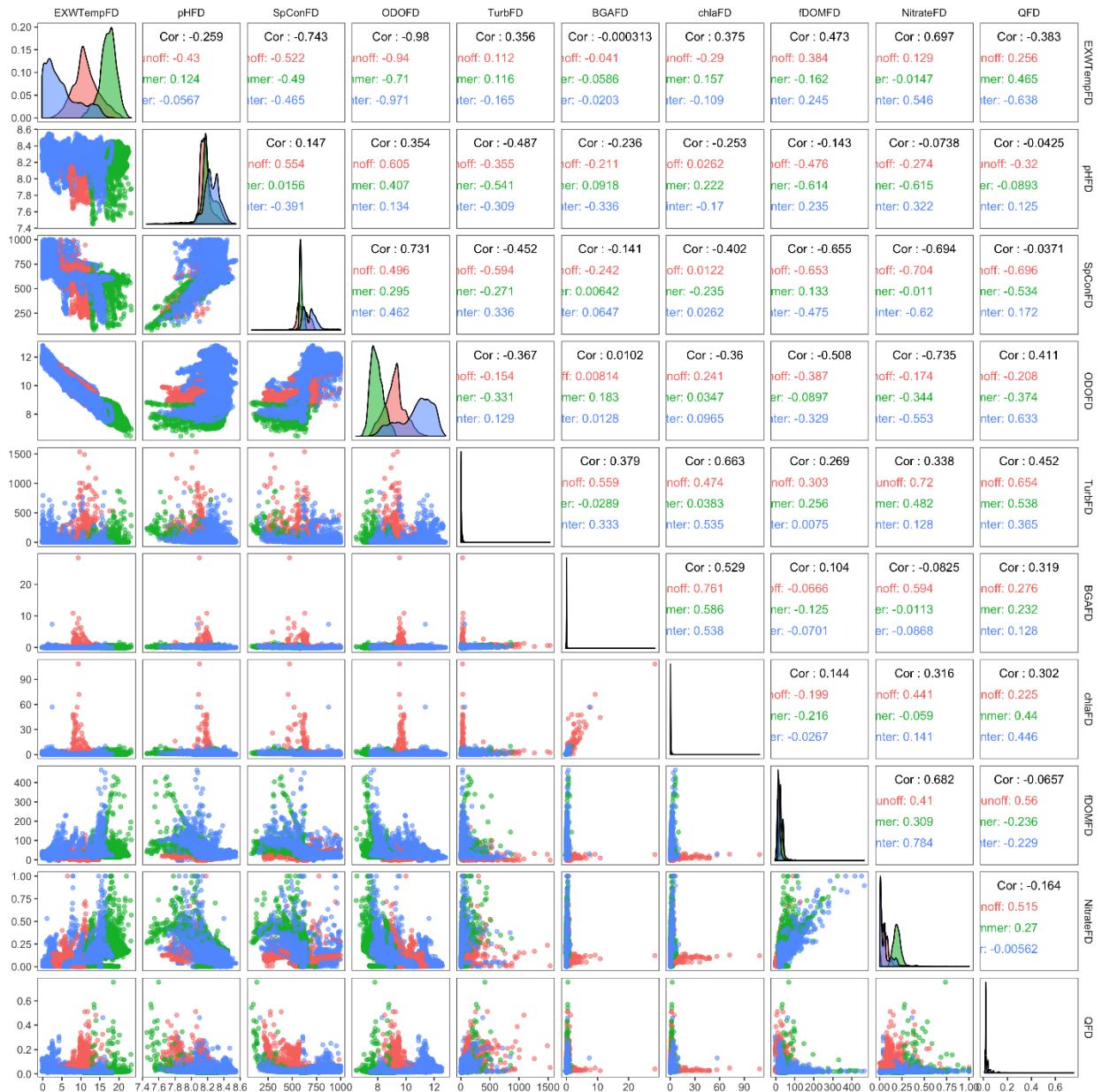


Figure S5. Correlation of sensor data at downstream Red Butte site, using Pearson's correlation.

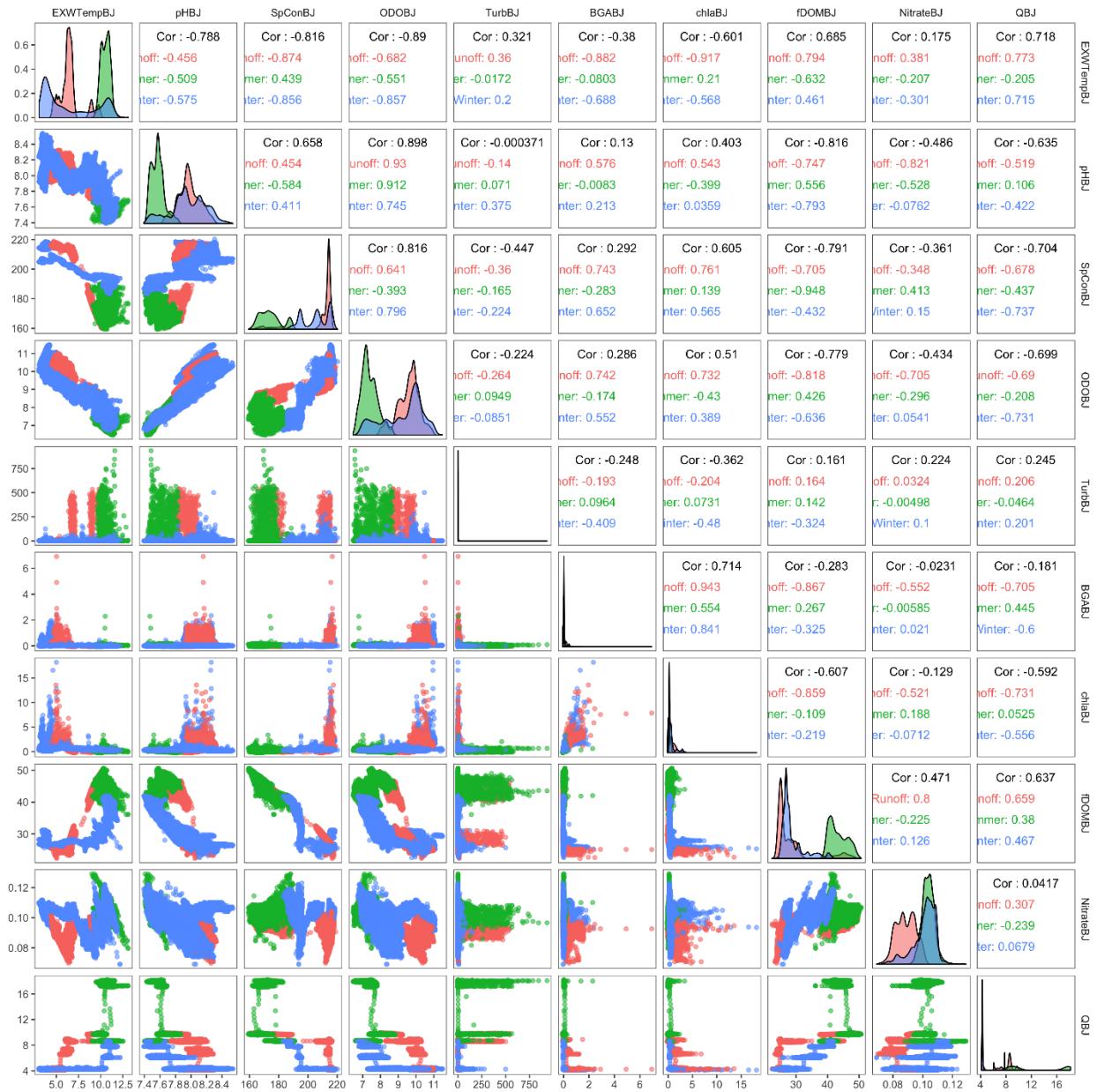


Figure S6. Correlation of sensor data at upstream Provo site, using Pearson's correlation.

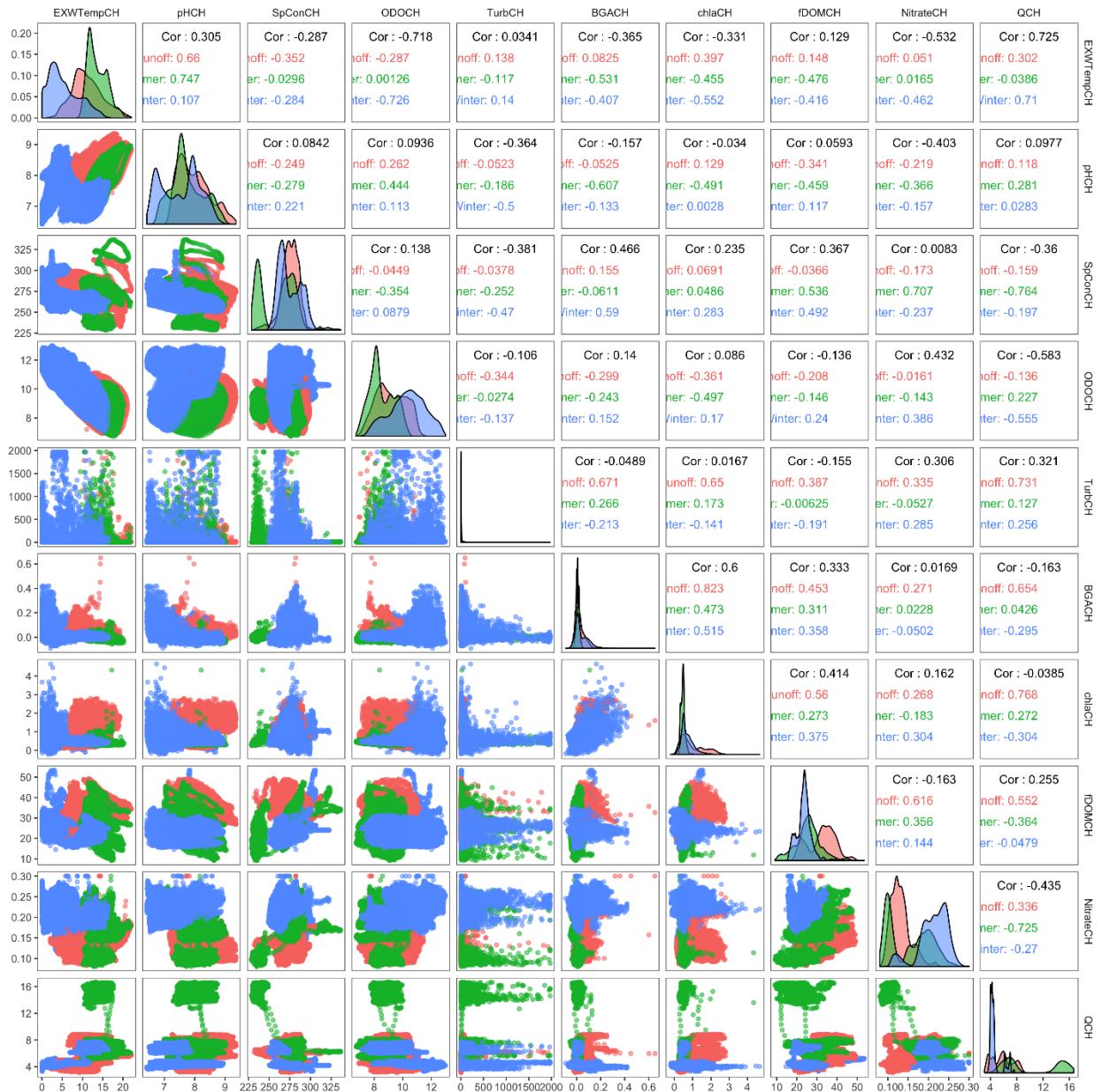


Figure S7. Correlation of sensor data at downstream Provo site, using Pearson's correlation.

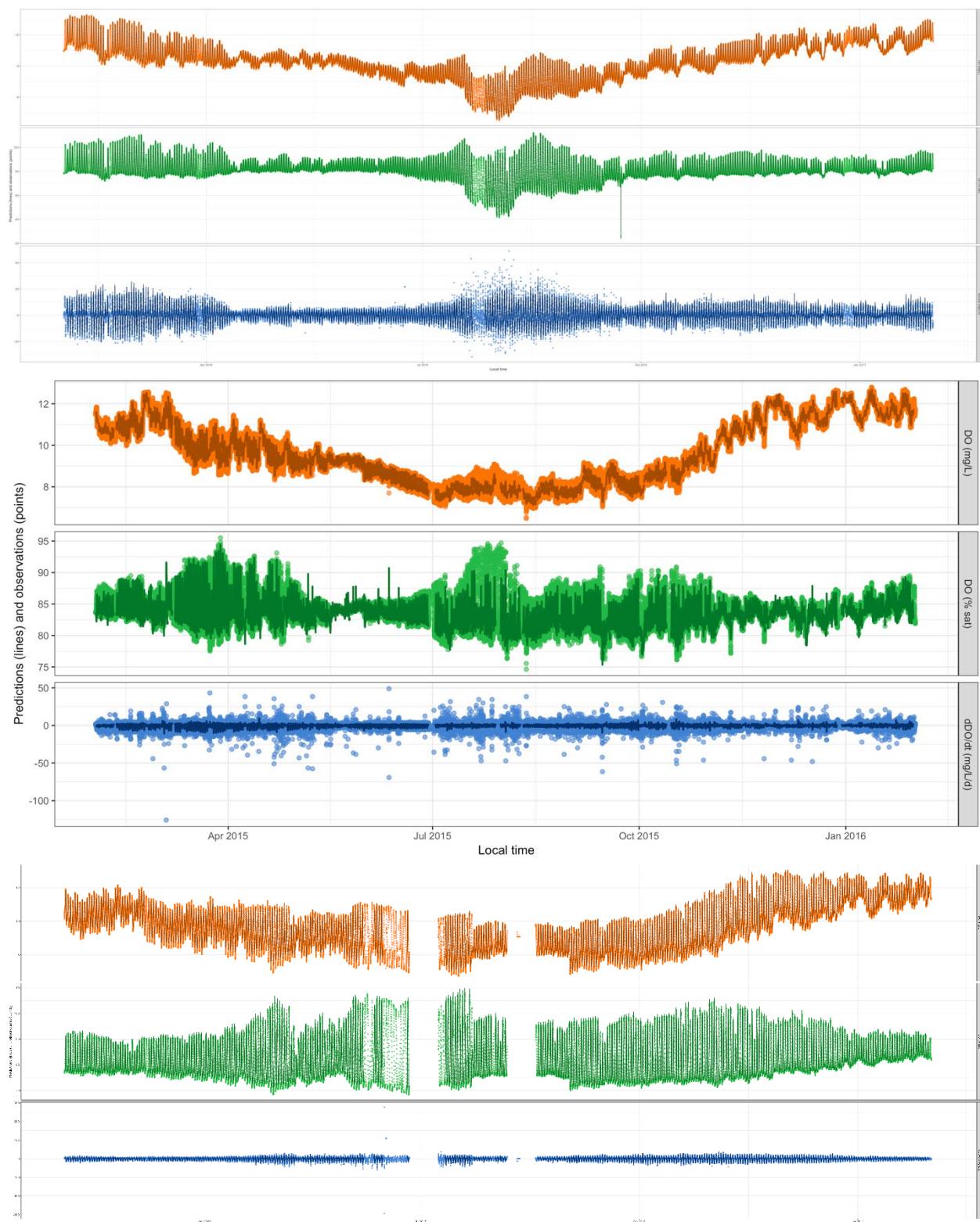


Figure S8. Metabolism model assessment

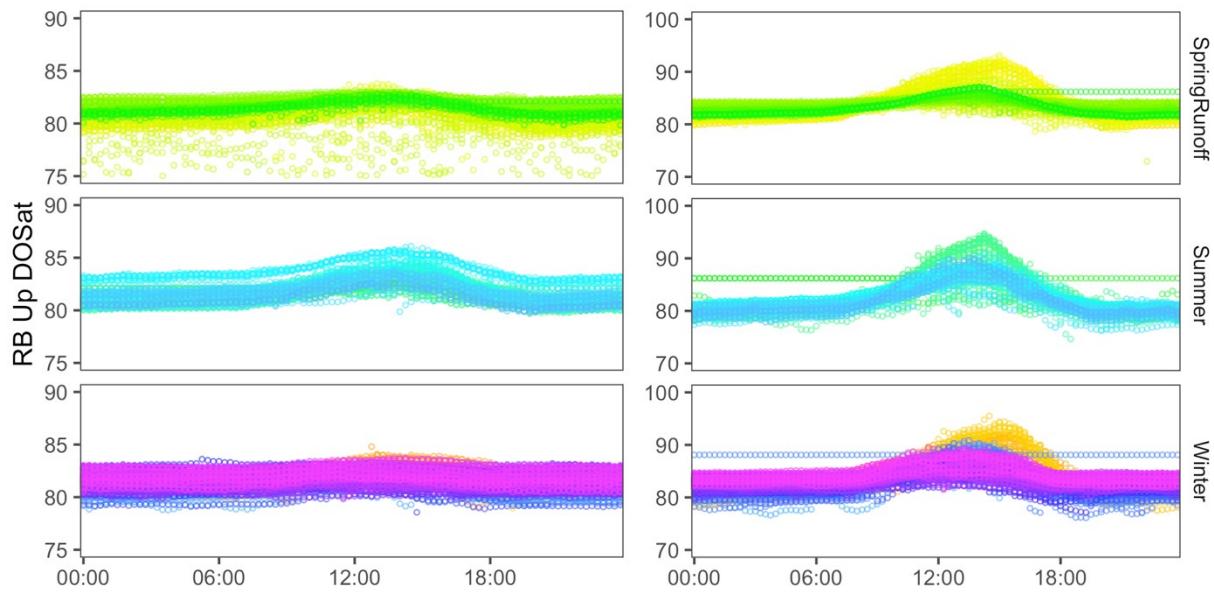


Figure S9. Diel dissolved oxygen as percent of saturation at Red Butte sites.

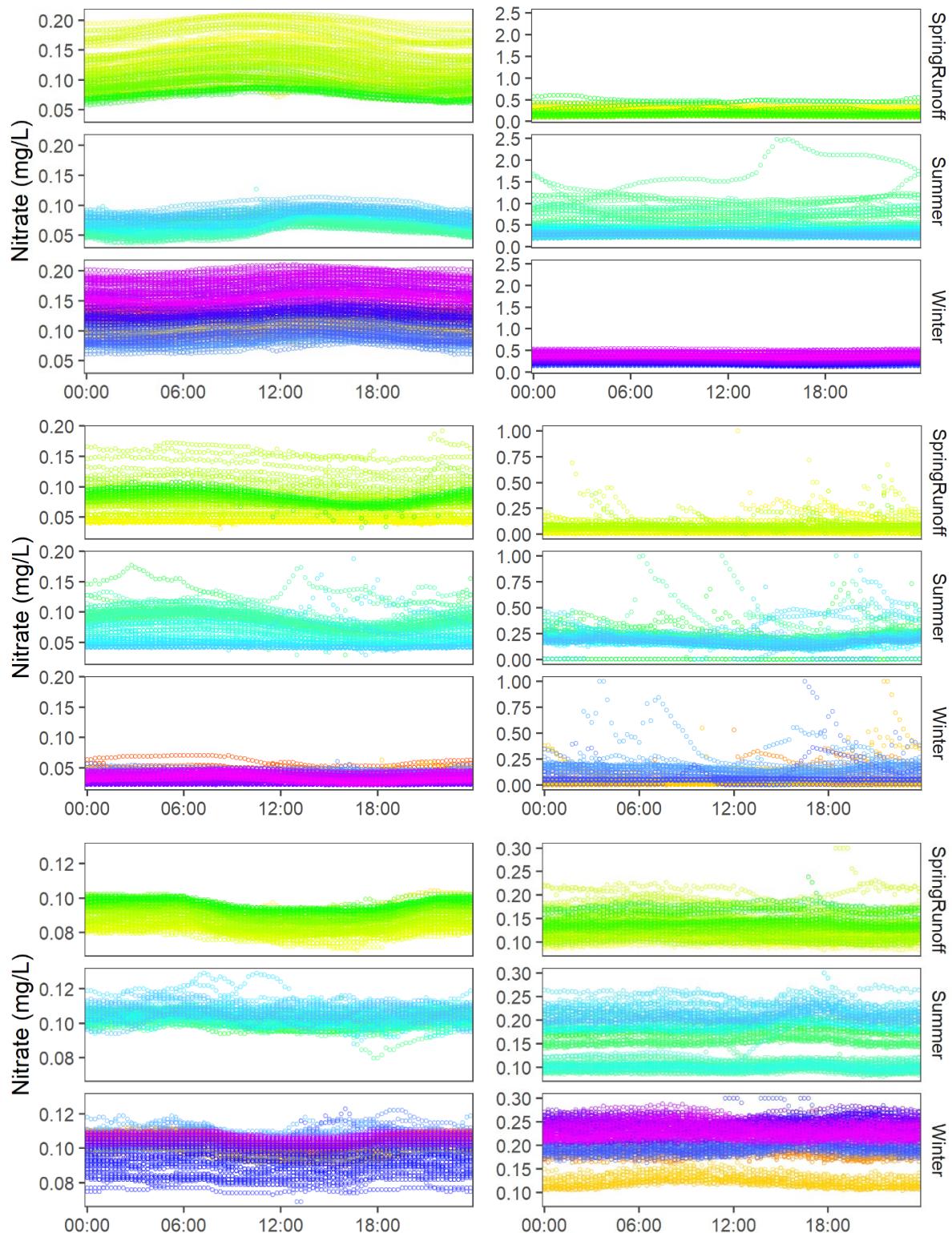


Figure S10. Diel  $\text{NO}_3^-$  concentrations for upstream (left) and downstream (right) sites in three Wasatch watersheds (Logan, A; Red Butte, B; and middle Provo, C) show seasonal differences in concentration, hydrology, anthropogenic additions, and in-stream processes.

Table S1. Summary statistics of sensor data at upstream and downstream sites in three Wasatch watersheds (Logan, Red Butte, and middle Provo).

|                |               | Temp (°C)   | DO (mg/L)   | Q (cms)     | NO <sub>3</sub> <sup>-</sup> (mg N/L) | SpCon (µS/cm) | fDOM (QSU)  | Turb (NTU)  |
|----------------|---------------|-------------|-------------|-------------|---------------------------------------|---------------|-------------|-------------|
| Logan Up       | Spring Runoff | 8.9 (1.65)  | 9.94 (0.51) | 11.6 (4.64) | 0.12 (0.04)                           | 349 (35)      | 10.7 (6.8)  | 5.69 (4.1)  |
|                | Summer        | 12.1 (1.43) | 9.36 (0.78) | 3.2 (1.15)  | 0.07 (0.01)                           | 362 (9)       | -0.7 (0.9)  | 0.6 (1.75)  |
|                | Winter        | 4.82 (2.27) | 11.1 (0.62) | 3.3 (0.61)  | 0.13 (0.03)                           | 398 (15)      | -1.8 (1.7)  | 1.45 (1.4)  |
| Logan Down     | Spring Runoff | 10.2 (1.91) | 9.21 (0.74) | 14.5 (5)    | 0.2 (0.07)                            | 373 (30)      | 13.1 (7.1)  | 15.6 (12.9) |
|                | Summer        | 16.4 (2.79) | 7.48 (1.3)  | 1.4 (1.5)   | 0.44 (0.28)                           | 542 (82)      | 18.0 (11.4) | 3.82 (2.53) |
|                | Winter        | 5.25 (3.01) | 10.58 (1.2) | 5.9 (1.17)  | 0.26 (0.09)                           | 463 (20)      | 3.5 (5)     | 5.51 (10.0) |
| Red Butte Up   | Spring Runoff | 8.97 (3.09) | 9.45 (0.74) | 0.09 (0.07) | 0.07 (0.02)                           | 576 (63)      | 25.3 (11.5) | 20.9 (84.0) |
|                | Summer        | 12.89 (1.8) | 8.66 (0.41) | 0.03 (0.01) | 0.07 (0.02)                           | 608 (30)      | 15.35 (6.6) | 2.45 (6.5)  |
|                | Winter        | 3.4 (3.57)  | 10.93 (1.0) | 0.03 (0.01) | 0.04 (0.01)                           | 675 (23)      | 13.0 (6.1)  | 9.36 (38.1) |
| Red Butte Down | Spring Runoff | 11.3 (3.29) | 9.19 (0.71) | 0.05 (0.04) | 0.06 (0.06)                           | 604 (74)      | 23.3 (15.3) | 35.1 (64.6) |
|                | Summer        | 17.1 (2.11) | 7.94 (0.45) | 0.02 (0.02) | 0.19 (0.08)                           | 585 (47)      | 31.6 (23.8) | 13.9 (29.6) |
|                | Winter        | 4.9 (4.34)  | 10.7 (1.1)  | 0.03 (0.02) | 0.06 (0.09)                           | 683 (83)      | 19.3 (17.9) | 13.6 (50.2) |
| Provo Up       | Spring Runoff | 6.55 (1.2)  | 9.6 (0.6)   | 7.7 (1.72)  | 0.09 (0.01)                           | 208 (13)      | 28.6 (6.2)  | 12.3 (59.8) |
|                | Summer        | 10.5 (0.51) | 7.4 (0.45)  | 12.7 (4.31) | 0.1 (0)                               | 174 (7)       | 43.3 (2.9)  | 18.7 (73.4) |
|                | Winter        | 6.4 (2.91)  | 9.26 (1.13) | 4.7 (1.07)  | 0.1 (0.01)                            | 205 (8)       | 28.8 (3.7)  | 3.22 (9.3)  |
| Provo Down     | Spring Runoff | 11.1 (3.72) | 9.2 (1.01)  | 5.7 (1.57)  | 0.13 (0.02)                           | 275 (10)      | 32.4 (6.7)  | 14.2 (73.4) |
|                | Summer        | 13.6 (2.38) | 8.5 (0.85)  | 10.2 (4.26) | 0.15 (0.05)                           | 260 (21)      | 25.5 (6.3)  | 57.8 (155)  |
|                | Winter        | 5.34 (3.75) | 10.3 (1.24) | 4.6 (0.92)  | 0.21 (0.04)                           | 274 (13)      | 23.7 (4.0)  | 51.8 (187)  |

## CHAPTER 3

### Network-Wide Water Sampling Reveals Nutrient Sources in a Eutrophic, Semi-Arid, Urban Watershed

Erin Fleming Jones, Rebecca Frei, Andrew Follett, Madeline Buhman, Madeleine Malmfeldt,  
Gabriella Lawson, Jansen Howe, Rachel Watts, Rhetta Shoemaker, Trevor Crandall, Jordan  
Maxwell, Adam Norris, Zachary T. Aanderud, Benjamin W. Abbott

Department of Plant and Wildlife Sciences, Brigham Young University, Provo, UT  
Doctor of Philosophy

#### ABSTRACT

Anthropogenic inputs of reactive nitrogen and phosphorus have led to eutrophication of water bodies worldwide, intensifying water resource scarcity. While treating point sources has reduced nutrient loading in many areas, non-point sources continue to fuel a second wave of eutrophication. To make improvements in a specific watershed, the relative contribution of point and non-point sources of nutrients needs to be quantified. Synoptic sampling of many points throughout the surface-water network can provide a spatially dense dataset of water chemistry, which recent research suggests can efficiently identify nutrient sources. Here, we used synoptic sampling to characterize nutrient sources to Utah Lake, a large, shallow, eutrophic lake that has been listed as impaired for phosphorus and total dissolved solids. To address policy disagreement about the source of solutes entering the lake and widespread public apathy, we collaborated with community members to collect ~175 samples in March, July, and October of 2018 across the 7640 km<sup>2</sup> Utah Lake watershed. For the four major tributaries to Utah Lake, we calculated ecohydrological metrics of concentration and flux for phosphate, dissolved inorganic nitrogen (DIN), total dissolved nitrogen (TN), dissolved organic carbon (DOC), sulfate, and chloride. Solute concentrations and leverage (influence on flux) were highest for all solutes except DOC in low-lying reaches draining directly to the lake, indicating the dominance of urban point and non-

point sources. Phosphate and DOC showed spatial instability (inconstant spatial pattern through time) relative to other solutes and only the smallest mountain watershed had a persistent spatial pattern. Unlike previous research from humid and temperate catchments, we did not observe a systematic decrease in spatial variability with watershed size in this semi-arid, endorheic basin. Instead of a funnel shape, there was an hourglass shape: high variability in the relatively pristine low-order reaches, low variability in the mid-order mainstems, and high variability in the high-order mainstems. This was attributable to semi-natural solute sourcing in the headwaters and return flows, diversions, and overall human footprint in the tailwaters. Our results demonstrate the value of combining participatory science with modern ecohydrological methods to determine catchment chemistry and hydrology. In addition to the scientific value, collaborative science can rehabilitate individual and community relationships with local ecosystems, representing an opportunity to improve understanding and stewardship of threatened and modified socio-ecological systems.

## INTRODUCTION

Eutrophication, a water quality impairment which affects two-thirds of freshwater ecosystems, is primarily caused by anthropogenic phosphorus (P) and nitrogen (N) additions (Galloway *et al.* 2004; Haygarth *et al.* 2005; Foley *et al.* 2011). However, removing point sources of nutrients has, in many instances, not alleviated algal blooms, suggesting that non-point sources play an important role in nutrient loading (Le Moal *et al.* 2019). Non-point nutrient sources in headwaters have been overlooked because of legal and practical constraints. For example, headwater chemistry is considered extremely variable in space and time due to changes in water flow, biological activity, and human disturbance (Heathwaite 2010; Abbott *et al.*

2018a). Headwater chemistry is a major driver of stream chemistry; headwaters make up the majority of global stream length (Downing 2012), and contribute solutes to watershed outlets (Alexander *et al.* 2000). Long-term synoptic sampling, however, is beginning to show spatial and temporal patterns that suggest that at certain spatial scales and for some solutes, large-scale patterns are more stable than previously assumed (Abbott *et al.* 2018a).

Locating point sources can be fairly straightforward, but non-point pollution is difficult to measure and determine the source. However, understanding the hydrology of a catchment can give clues about the nature of non-point sources (Temnerud and Bishop 2005; Abbott *et al.* 2016). One hydrologic metric that assists in the determination of pollution sources is spatial stability. The concept of spatial stability, as described by Abbott *et al.* (2018), is a quantification of how consistent solute concentrations are through time. Stated differently, in a catchment with high spatial stability, stream sites (or subcatchments) with relatively high concentrations of a specific solute are consistently higher and locations with relatively low concentrations are consistently lower. In practice, high spatial stability allows for a single sampling to capture nutrient dynamics across a watershed. Spatial stability has been calculated for many solutes in multiple watersheds, and many watersheds exhibit a high degree of stability (Abbott *et al.* 2018a; Dupas *et al.* 2019b). Some solutes are less stable than others, with low stability frequently noted for P. Stability can vary between solutes because N and P inputs involve different processes, including weathering products, groundwater sources and flow paths (Ayraud *et al.* 2008; Kolbe *et al.* 2019), and anthropogenic sources. Solute stability may be influenced by state factors such as watershed age and climate, but many of the studies have been done in temperate regions. Semi-arid regions might have higher hydrologic variability, leading to less stability.

Another hydrologic metric that may help managers understand the source and nature of stream solutes is spatial variability threshold, also called representative elemental area (Wood *et al.* 1988; Blöschl *et al.* 1995, Figure 3-1). This threshold represents multiple watershed characteristics, including the patch size of solutes, and the watershed area which has the highest leverage on the overall stream concentration for a catchment. One way to calculate threshold area is by using pruned exact linear time to identify where concentration and leverage, or influence on the overall catchment median, rapidly approach the median watershed value (Asano *et al.* 2009; Abbott *et al.* 2018a). In large, remote watersheds, synoptic sampling to obtain the data necessary to perform these calculations can be prohibitive. Citizen science is the perfect tool to address this issue, and because public incentive to address non-point pollution is often low, engaging the public in data collection can generate grassroots interest in improving local water quality (Church *et al.* 2018).

Utah Lake is listed as impaired for P and has regular summer algal blooms, with P coming from natural and anthropogenic non-point and point sources (PSOMAS 2007). Utah's semi-arid climate, age and parent material of watersheds, and differences in human development present a unique environment to calculate spatial stability. To generate a spatially dense dataset, we organized synoptic citizen science sampling events, which has also been shown to improve attitudes and awareness of environmental issues (Bonney *et al.* 2009; Crall *et al.* 2012; Church *et al.* 2018). Our study addresses two questions about stream nutrient sources and variability using participatory science in the Utah Lake Watershed. First, where (and what) are the point and non-point sources of N and P in this watershed? Second, which solutes (if any) are spatially stable in this watershed?

## METHODS

### *Synoptic Sampling*

We used a systems approach informed by landscape ecology and catchment hydrology for analyzing spatial and temporal variance of water chemistry in stream networks (Dupas *et al.* 2019b). Streams experience temporal variability in chemistry because of hydrologic pulses and fluctuations in biogeochemical activity (Rinaldo *et al.* 1998; Erlandsson *et al.* 2008; Raymond *et al.* 2016). As pulses move through stream networks, their downstream attenuation or preservation depends on the synchrony of pulse generation in subcatchments (Abbott *et al.* 2018a). To identify nutrient sources, quantify the spatial stability of those sources through time, and assess what stream and catchment characteristics most affect resilience to nutrient loading, we designed a spatially dense synoptic sampling of solute and particulate concentrations in the river's main stem and tributaries. This allowed us to identify sources of pollution in urban and non-urban areas of the watershed, which will help land managers to determine optimal areas for future remediation. This study includes three synoptic sampling events, conducted in March (Spring), July (Summer), and October 2018 (Fall).

### *Site Description and Selection*

This study was executed in the Utah Lake watershed, including five major tributaries of the lake: Provo River, Spanish Fork River, American Fork River, Hobble Creek, and Benjamin Slough (Table 3-1). Utah Lake is the second largest remnant of Lake Bonneville, an inland sea that covered 52,000 km<sup>2</sup> (Gilbert 1890; Hunt *et al.* 1953). The lake drains into the largest remnant, Great Salt Lake, to the north. The watershed is characterized by relatively pristine high-elevation headwaters (although mining, intensive livestock grazing, rural subdivisions and ski

resorts are present) with denser development in low-elevation valleys. As Utah is one of the fastest growing states in the US (US Census Bureau), areas of agricultural operations in the watershed's valleys have increasingly been converted to urbanized landscapes. All of the watersheds have some degree of hydrological modification to meet water demands, but one of the watersheds (Provo River) has two large reservoirs that drastically alter the hydrologic condition of the watershed. We classified the watersheds into one of four categories based on land use and hydrologic condition (Table 3-1): Agriculture unregulated (Spanish Fork River), Mixed dammed (Provo River), Mountain urban (American Fork River and Hobble Creek), and Lake tributaries (Benjamin Slough, Currant Creek, and other valley tributaries).

Utah Lake watershed's hydrology consists mainly of gaining reaches (i.e., net movement of soil and ground water into the channel) at high elevations, but at the base of the mountains (where rivers flow into Lake Bonneville sediment) reaches quickly become losing (i.e., net movement from the channel into soil and groundwater) and transport largely occurs through shallow groundwater (Cederberg *et al.* 2009). Near the lake, streams once again become gaining reaches and many groundwater-fed springs generate new streams that flow into the lake (e.g. Lake tributary watersheds). Seven wastewater treatment plants, serving approximately 600,000 people in the valley region, also discharge effluent into small tributaries to the lake.

We initially selected 500 sites from the Department of Water Quality Ambient Water Quality Monitoring System (AWQMS) database and Utah State University Water Quality Extension citizen science program Utah Water Watch (UWW). We used a clustering technique effort by including sites just above and below a confluence to maximize watershed coverage and minimize travel distance. We consolidated the initial 500 sites to 270 (Figure 3-2) by removing sites that were inaccessible, channels that were no longer active, and merging redundant sites. We

delineated catchments and reported watershed area and percent land use (% developed, % impervious, % forested, and % herbaceous upland) for each site using USGS StreamStats.

### *Citizen Science*

This study included two levels of public participation. University students in a Watershed Ecology course helped design the study, recruit volunteers, and perform sample and data analysis. To cultivate a sense of stewardship and responsibility, we invited community members to participate in sample collection (Church *et al.* 2018). Volunteers were recruited using social media, fliers, and by coordinating with other organizations, such as the Provo River Watershed Council, UWW and Utah Division of Water Quality. We presented a model watershed demonstration (Enviroscapes) at local community events to help recruit participants. Students sponsored two seminar events about Utah Lake watershed and its history of water quality, attended by community members. We provided training and copies of sampling instructions for participants when they picked up sampling materials. With this instruction, we are satisfied that the sampling techniques were acceptable and that the data collected is reliable and accurate. All sample analyses were done by trained lab technicians at Brigham Young University.

### *Laboratory and Statistical Analysis*

Samples were filtered in the field with pre-rinsed 0.45  $\mu\text{m}$  cellulose acetate filters (Millipore Millex-GV) and immediately frozen or analyzed within 2 weeks. Quantified analytes included many common water quality parameters including ions, soluble reactive phosphorus (SRP), dissolved organic carbon (DOC), and total dissolved nitrogen (TDN, hereafter TN). Anions ( $\text{NO}_3^-$ ,  $\text{NO}_2^-$ ,  $\text{SO}_4^{2-}$ ,  $\text{Cl}^-$ ,  $\text{F}^-$ ,  $\text{PO}_4^{3-}$ , and  $\text{Br}^-$ ) and cations ( $\text{Li}^+$ ,  $\text{Na}^+$ ,  $\text{NH}_4^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{+2}$ ,  $\text{Ca}^{+2}$ , and  $\text{Sr}^{+2}$ )

were quantified by ion chromatography (Dionex Thermo Fisher HPIC). SRP was quantified colorimetrically using the ascorbic acid method. DOC and TDN were quantified using an Elementar auto-analyzer. We used ArcGIS Pro (ESRI) to map solute concentrations.

We calculated and graphed leverage, Spearman's rank correlation (Abbott *et al.* 2018a) and spatial stability (Dupas 2019) in RStudio, using the base, vegan, ggplot2, and cowplot packages. We used phosphorus concentrations (PO<sub>4</sub>.P), which are the average of phosphorus values determined by ion chromatography and ascorbic acid. Total inorganic nitrogen (TIN) was calculated as the sum of N species (i.e. NO<sub>3</sub><sup>-</sup>, NO<sub>2</sub><sup>-</sup>, and NH<sub>4</sub><sup>+</sup>) from the ion chromatography analysis. We calculated spatial stability by calculating spearman's rank correlation coefficient for each watershed category and solute across each pair of samplings and calculated the average of the three coefficients. We scaled solute concentrations (i.e., report each value as a standard deviation from the mean) for each solute to create notched boxplots and test using Analysis of Variance (ANOVA). We scaled solute concentrations by solute and land use category to create spatial variability graphs. Leverage was calculated by scaling each concentration by the concentration of the largest catchment area for each combination of watershed category and solute. All data and code used is available at <https://github.com/erinfjones/citizenscience>.

## RESULTS

As part of our recruiting efforts, we made 3,700 contacts with the watershed model presentation, and had 85 attendees at a seminar event featuring local researchers. We had over 150 participants at sampling events. When asked about their experience, responses have been overwhelmingly positive, including statements like, "I had no idea there were so many beautiful

streams in Utah Valley,” and “This opened my eyes to how much we depend on this water. It actually comes from somewhere before my sink!”

### *Solute Concentrations*

Solute concentrations were different across Events (ANOVA, F-stat=5.896, p-value=0.0028) and categories (ANOVA, F-stat=96.055, p-value<0.0001) regardless of solute (Figure 3-3, Table S1). Pairwise analysis determined that solute concentrations in Spring were higher than Summer or Fall (Tukey, p-adj.=0.002). Concentrations at Lake tributary sites were higher than the other three categories regardless of solute (p-adj.<0.001) and Mountain urban were higher than Mixed Dammed (p-adj=0.005). The change in certain solute concentrations depended on Event (ANOVA, F-stat=8.057, p-value<0.0001) and land use category (ANOVA, F-stat=4.34, p-value<0.001). A pairwise analysis determined that Summer DOC concentrations were higher than Fall or Spring values (Tukey, p-adj<0.001), and Fall TN concentrations were higher than Summer or Spring (Tukey, p-value<0.05).

The spatial distribution of solute concentrations across the watershed can be seen in Figure 3-4. Point color represents numeric water quality standards for N and P values, or 25<sup>th</sup> and 75<sup>th</sup> percentile for others. High and low DOC and PO4-P concentrations were distributed across the watershed, while DIN-N, TN, sulfate and chloride concentrations were highest at sites near or on Utah Lake (Tukey, adj. p-value<0.001).

Regressions of solute concentration by event and land use found that % impervious surface was correlated with higher concentrations for all solutes (Table 3-2). TN, DOC, and sulfate models included Event; sulfate and chloride included % herbaceous upland. Correlation coefficients ( $R^2$ ) for the models were between 0.11 and 0.25.

### *Leverage and Spatial Stability*

Scaled concentration and leverage by watershed area for the different solutes did not show a typical funnel shape (Figure 3-5, Asano *et al.* 2009). Instead, many had hourglass shapes, with occasional higher leverage and scaled concentration at the largest watershed size. Occasional outliers (single points at mid-range and large watershed size) may have had an oversized effect on the overall pattern. DOC concentrations showed no variance collapse. Lake tributary sites within smaller watersheds had particularly high leverage on DIN-N and TN (Figure 3-5).

Spatial stability was solute and land-use specific (Figure 3-6, Table S2). DOC and PO4-P had lower spatial stability than the other solutes (0.25-0.5). Chloride and sulfate had the highest spatial stability (>0.7). DIN-N and TN were intermediate (0.5-0.7). Stability was highest overall at Mountain urban and lowest in Lake tributary and Agricultural unregulated reaches, although the order of most to least stable was dependent on solute (ANOVA, F-statistic=3.514, df=30, p-value<0.001).

## DISCUSSION

### *Utah Lake Watershed has Novel Spatial and Temporal Hydrochemistry*

Our results emphasize the unique hydrochemistry of semi-arid regions. PO4-P was less spatially stable, which was also the case in watersheds in France (Abbott *et al.* 2018a; Dupas *et al.* 2019b) and Arctic tundra (Shogren *et al.* 2019). The low spatial stability of P when compared with other solutes like chloride suggests two non-exclusive facts: one, sources other than natural geologic weathering introduce variability and two, biological processes are variable. For example, P may switch between excess and limiting across space and/or time (Hoellein *et al.* 2011).

Lake tributary reaches stood out as having particularly low spatial stability, across all solutes (Figure 3-6). We hypothesize that the low stability of solute concentrations at Lake tributary sites is at least partially a function of the difference in concentration between snow-melt stream water and evaporate-rich groundwater, combined with a change in source based on hydrologic conditions (Cederberg *et al.* 2009). When discharge is high, water diluted with snow-melt extends further into the valleys before the change from losing to gaining occurs. During low flow, the switching point of losing to gaining moves upstream, creating large variability in concentrations observed at valley sites between the mouth of the canyons and the lake.

Our analysis did not return a clear spatial threshold for concentration collapse, unlike other studies (Likens and Buso 2006; Tiwari *et al.* 2017; Coble *et al.* 2019; Hale and Godsey 2019). For example, Northern Boreal catchments tended to have variance collapse in DOC at 15 km<sup>2</sup> (Temnerud and Bishop 2005). Arctic tundra watersheds had thresholds of 10-20 km<sup>2</sup> for DOC and nitrate, with slightly larger (25 km<sup>2</sup>) for phosphorus (Shogren *et al.* 2019). Mined Kentucky headwaters had variance collapse in major anion and cation concentrations between 15 and 75 km<sup>2</sup> (Johnson *et al.* 2019). The lack of variance collapse, like the low spatial stability, may have also been due to increased solute concentrations at groundwater-influenced sites near the lake (Cederberg *et al.* 2009).

#### *Urban Sources Affect all the Solutes Except DOC*

This study identified hot spots of solutes of concern (Figure 3-4), and determined that position within watershed is important in determining concentration dynamics (Dupas *et al.* 2019a). Phosphorus, TN, DIN, chloride, and sulfate were highest in Lake tributary reaches, and Solute removal or dilution occurs at mid-elevation reaches, indicated by the map of

concentrations (Figure 3-4) and narrowing of solute concentration (Figure 3-5). Impervious surfaces contributed to overall higher solute concentrations (Table 3-2), but solute concentrations were more variable in reaches with agricultural activity (Figure 3-6). Decreases in concentration at sites in the valley could be due to losses to groundwater (Cederberg *et al.* 2009) or sorption of P to Lake Bonneville sediments (Randall *et al.* 2019).

Lake tributary reaches were significant sources of N, a majority of which is DIN (Figure 3-3, Figure 3-5). Linear models of TN and DIN-N had higher correlation with % impervious surface than any of the other solutes (Table 3-2), and both solutes had lower stability across Lake tributary reaches than the other land use categories. Lake tributary reaches, which had the highest percent developed land use (Table 3-1), had the highest variability in concentrations, unlike urban watersheds in New York which decreased in variability in the urban environment (Hoellein *et al.* 2011).

### *Citizen Science*

We provided the opportunity for thousands of local citizens to learn more about point and non-point sources of water pollution, in a deep and meaningful way (Bonney *et al.* 2009). This engagement has the possibility of creating public support for efforts to address water quality in the Utah Lake watershed (Dickinson *et al.* 2012). Future directions of this research include using educational research tools to quantify the impact of participation on knowledge, attitude, and behavior. We conclude that citizen science can extend scientific observation and fundamentally change public awareness and mentality, which subsequently influences how water resources are managed (Church *et al.* 2018). In this sense, participatory water quality monitoring is not only a means of generating understanding of how water and nutrients propagate through catchments; it

is a mechanism to improve water quality itself and encourage sustainable stewardship (Abbott *et al.* 2018b).

## CONCLUSION

Utah Lake watershed is fundamentally different in network-scale hydrochemistry than previously described regions. Part of this could be because this is a hydrologically-losing area with extreme hydrological modification, non-random distribution of human footprint creates a strong influence on water chemistry (urban and agricultural activity is concentrated in the lowlands). We encourage including semi-arid regions, especially endorheic basins, in hydrologic studies because understanding their distinctive hydrologic characteristics are critical to preserving these unique and threatened ecosystems (Wurtsbaugh *et al.* 2017).

Our results demonstrate the high spatial and temporal variability of phosphorus within this watershed. However, at intermediate watershed sizes, phosphate removal or dilution occurred (Figure 3-5). Point sources and groundwater around the lake contributed N in the form of DIN, and like phosphate, decreased in concentration in mid-range watershed size. In addition to high concentrations of solutes, Lake tributary reaches had low spatial stability.

## LITERATURE CITED

- Abbott BW, Baranov V, Mendoza-Lera C, *et al.* 2016. Using multi-tracer inference to move beyond single-catchment ecohydrology. *Earth-Science Reviews* **160**: 19–42.
- Abbott BW, Gruau G, Zarnetske JP, *et al.* 2018a. Unexpected spatial stability of water chemistry in headwater stream networks. *Ecology Letters* **21**: 296–308.
- Abbott BW, Moatar F, Gauthier O, *et al.* 2018b. Trends and seasonality of river nutrients in agricultural catchments: 18 years of weekly citizen science in France. *Science of The Total Environment* **624**: 845–58.
- Alexander RB, Smith RA, and Schwarz GE. 2000. Effect of stream channel size on the delivery of nitrogen to the Gulf of Mexico. *Nature* **403**: 758.
- Asano Y, Uchida T, Mimasu Y, and Ohte N. 2009. Spatial patterns of stream solute concentrations in a steep mountainous catchment with a homogeneous landscape. *Water Resources Research* **45**.
- Ayraud V, Aquilina L, Labasque T, *et al.* 2008. Compartmentalization of physical and chemical properties in hard-rock aquifers deduced from chemical and groundwater age analyses. *Applied Geochemistry* **23**: 2686–707.
- Blöschl G, Grayson RB, and Sivapalan M. 1995. On the representative elementary area (REA) concept and its utility for distributed rainfall-runoff modelling. *Hydrological Processes* **9**: 313–30.
- Bonney R, Cooper CB, Dickinson J, *et al.* 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience* **59**: 977–84.
- Cederberg JR, Gardner PM, and Thiros SA. 2009. Hydrology of Northern Utah Valley, Utah County, Utah, 1975–2005. Reston, VA: U.S. Geological Survey.

- Church SP, Payne LB, Peel S, and Prokopy LS. 2018. Beyond water data: benefits to volunteers and to local water from a citizen science program. *Journal of Environmental Planning and Management*: 1–21.
- Coble AA, Koenig LE, Potter JD, et al. 2019. Homogenization of dissolved organic matter within a river network occurs in the smallest headwaters. *Biogeochemistry* **143**: 85–104.
- Crall AW, Jordan R, Holfelder K, et al. 2012. The impacts of an invasive species citizen science training program on participant attitudes, behavior, and science literacy. *Public Understanding of Science* **22**: 745–64.
- Dickinson JL, Shirk J, Bonter D, et al. 2012. The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment* **10**: 291–7.
- Downing J. 2012. Global abundance and size distribution of streams and rivers. *Inland Waters* **2**: 229–36.
- Dupas R, Abbott B, Minaudo C, and Fovet O. 2019a. Distribution of Landscape Units Within Catchments Influences Nutrient Export Dynamics. *Frontiers of Environmental Science & Engineering* **7**: 43.
- Dupas R, Minaudo C, and Abbott BW. 2019b. Stability of spatial patterns in water chemistry across temperate ecoregions. *Environmental Research Letters* **14**: 074015.
- Erlandsson M, Buffam I, Fölster J, et al. 2008. Thirty-five years of synchrony in the organic matter concentrations of Swedish rivers explained by variation in flow and sulphate. *Global Change Biology* **14**: 1191–8.
- Foley JA, Ramankutty N, Brauman KA, et al. 2011. Solutions for a cultivated planet. *Nature* **478**: 337–42.

- Galloway JN, Dentener FJ, Capone DG, *et al.* 2004. Nitrogen Cycles: Past, Present, and Future. *Biogeochemistry* **70**: 153–226.
- Gilbert GK. 1890. Lake Bonneville. U.S. Government Printing Office, Washington D.C.: U.S. Geological Survey.
- Hale RL and Godsey SE. 2019. Dynamic stream network intermittence explains emergent dissolved organic carbon chemostasis in headwaters. *Hydrological Processes* **0**.
- Haygarth PM, Condron LM, Heathwaite AL, *et al.* 2005. The phosphorus transfer continuum: Linking source to impact with an interdisciplinary and multi-scaled approach. *Science of The Total Environment* **344**: 5–14.
- Heathwaite AL. 2010. Multiple stressors on water availability at global to catchment scales: understanding human impact on nutrient cycles to protect water quality and water availability in the long term. *Freshwater Biology* **55**: 241–57.
- Hoellein TJ, Arango CP, and Zak Y. 2011. Spatial variability in nutrient concentration and biofilm nutrient limitation in an urban watershed. *Biogeochemistry* **106**: 265–80.
- Hunt CB, Varnes HD, and Thomas HE. 1953. Lake Bonneville: Geology of northern Utah Valley, Utah. Washington, D.C.: U.S. Government Printing Office.
- Johnson B, Smith E, Ackerman JW, *et al.* 2019. Spatial Convergence in Major Dissolved Ion Concentrations and Implications of Headwater Mining for Downstream Water Quality. *JAWRA Journal of the American Water Resources Association* **55**: 247–58.
- Kolbe T, Dreuzé J-R de, Abbott BW, *et al.* 2019. Stratification of reactivity determines nitrate removal in groundwater. *Proceedings of the National Academy of Sciences* **116**: 2494–9.
- Le Moal M, Gascuel-Odoux C, Ménesguen A, *et al.* 2019. Eutrophication: A new wine in an old bottle? *Science of The Total Environment* **651**: 1–11.

- Likens GE and Buso DC. 2006. Variation in Streamwater Chemistry Throughout the Hubbard Brook Valley. *Biogeochemistry* **78**: 1–30.
- PSOMAS. 2007. Utah Lake TMDL: Pollutant Loading Assessment and Designated Beneficial Use Impairment Assessment. State of Utah Division of Water Quality.
- Randall MC, Carling GT, Dastrup DB, *et al.* 2019. Sediment potentially controls in-lake phosphorus cycling and harmful cyanobacteria in shallow, eutrophic Utah Lake. *PLOS ONE* **14**: e0212238.
- Raymond PA, Saiers JE, and Sobczak WV. 2016. Hydrological and biogeochemical controls on watershed dissolved organic matter transport: pulse-shunt concept. *Ecology* **97**: 5–16.
- Rinaldo A, Rodriguez-Iturbe I, and Rigon R. 1998. Channel Networks. *Annual Review of Earth and Planetary Sciences* **26**: 289–327.
- Shogren AJ, Zarnetske JP, Abbott BW, *et al.* 2019. Revealing biogeochemical signatures of Arctic landscapes with river chemistry. *Sci Rep* **9**: 1–11.
- Temnerud J and Bishop K. 2005. Spatial Variation of Streamwater Chemistry in Two Swedish Boreal Catchments: Implications for Environmental Assessment. *Environ Sci Technol* **39**: 1463–9.
- Tiwari T, Buffam I, Sponseller RA, and Laudon H. 2017. Inferring scale-dependent processes influencing stream water biogeochemistry from headwater to sea. *Limnology and Oceanography* **62**: S58–70.
- US Census Bureau. Utah is Nation's Fastest-Growing State, Census Bureau Reports <https://www.census.gov/newsroom/press-releases/2016/cb16-214.html>. Viewed 23 Feb 2019.

Wood EF, Sivapalan M, Beven K, and Band L. 1988. Effects of spatial variability and scale with implications to hydrologic modeling. *Journal of Hydrology* **102**: 29–47.

Wurtsbaugh WA, Miller C, Null SE, *et al.* 2017. Decline of the world's saline lakes. *Nature Geoscience* **10**: 816.

## FIGURES

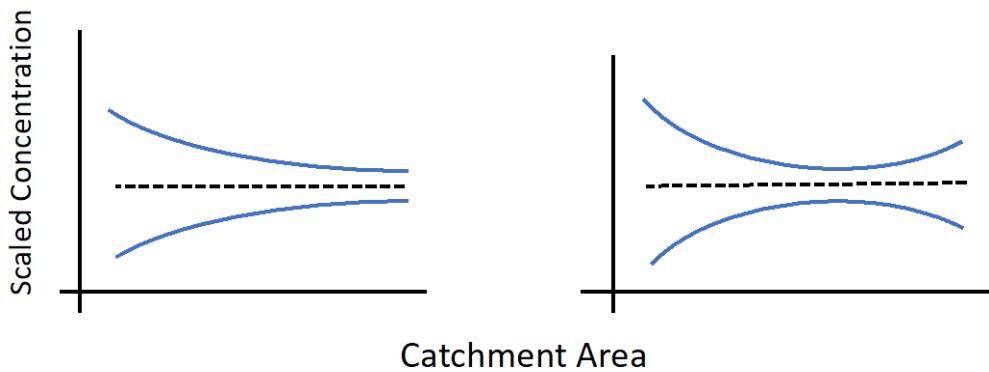


Figure 3-1. Theoretical diagram showing existing concept of spatial variability collapse (left) and the pattern we observed in the Utah Lake watershed, with spatial variability expanse (right).

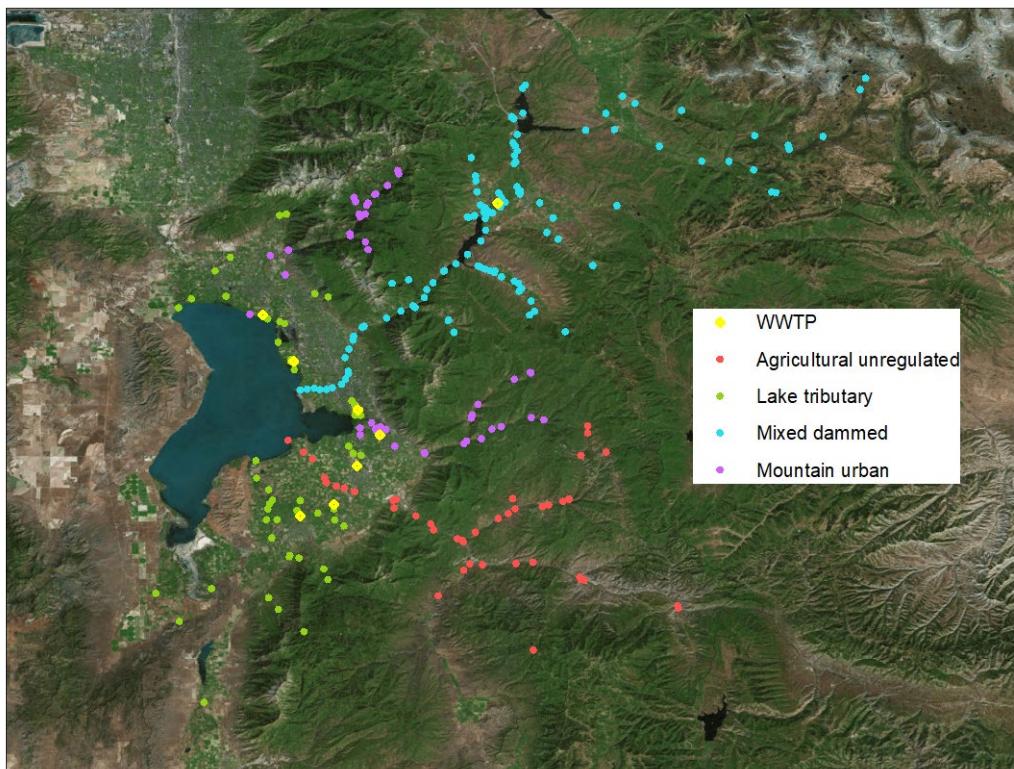


Figure 3-2. Map of Utah Lake watershed sites synoptically sampled, colored by land use and hydrologic modification category (red= Agricultural unregulated, green= Lake tributary, blue= Mixed dammed, purple= Mountain urban). Yellow triangles represent wastewater treatment plants.

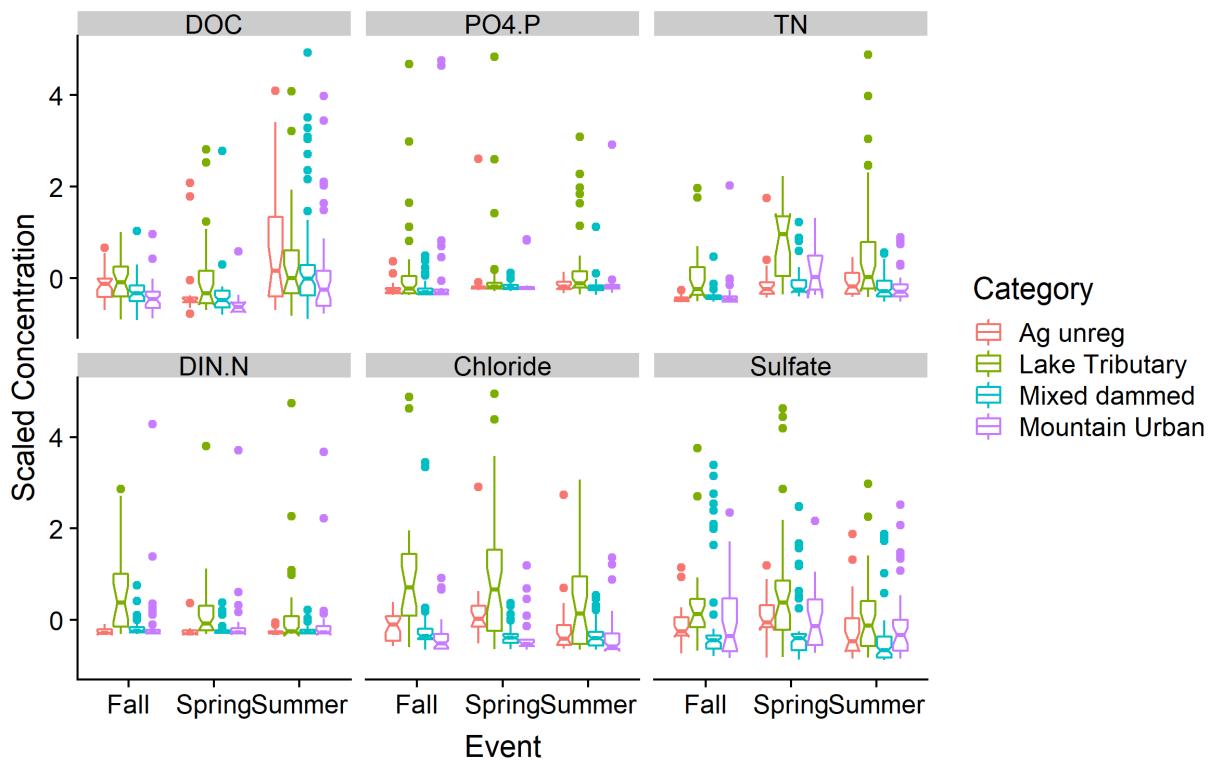


Figure 3-3. Scaled concentration of solutes in surface water samples collected during three synoptic sampling events (Fall, Spring, Summer) of areas within the Utah Lake watershed with different characteristics (Agricultural unregulated, Lake tributary, Mixed dammed, and Mountain urban). Boxplots represent the 25<sup>th</sup>, 50<sup>th</sup>, and 70<sup>th</sup> percentiles, points within 1.5 times the interquartile range, and points beyond. The notches represent the 95% confidence interval of the median (non-overlapping notches suggest differences between populations).

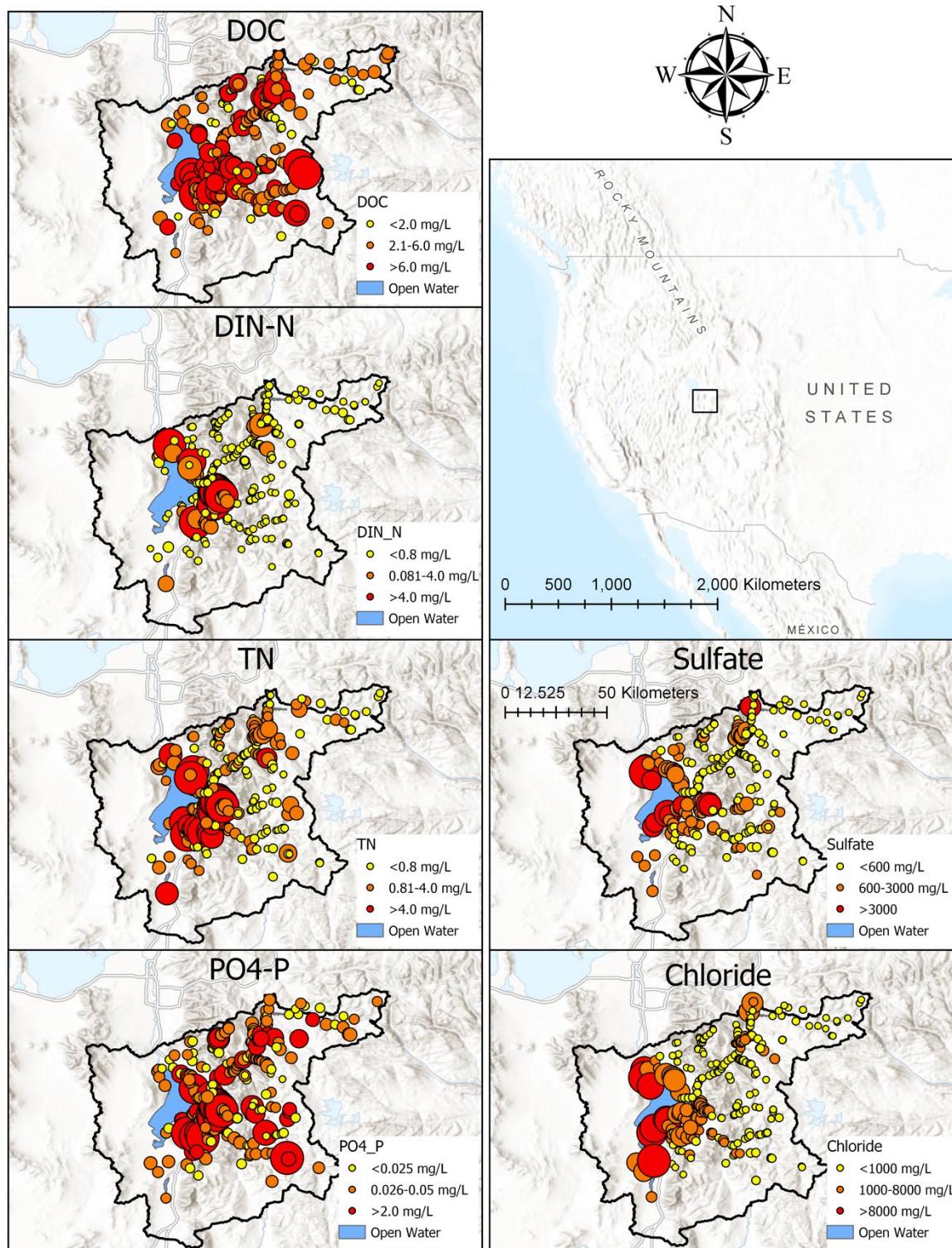


Figure 3-4. Maps showing average concentration of solutes from three synoptic sampling events across the Utah Lake watershed (outlined in black). Point color represents numeric water quality standards for N and P values, or 25<sup>th</sup> and 75<sup>th</sup> percentile for others. Point size is scaled to concentration.

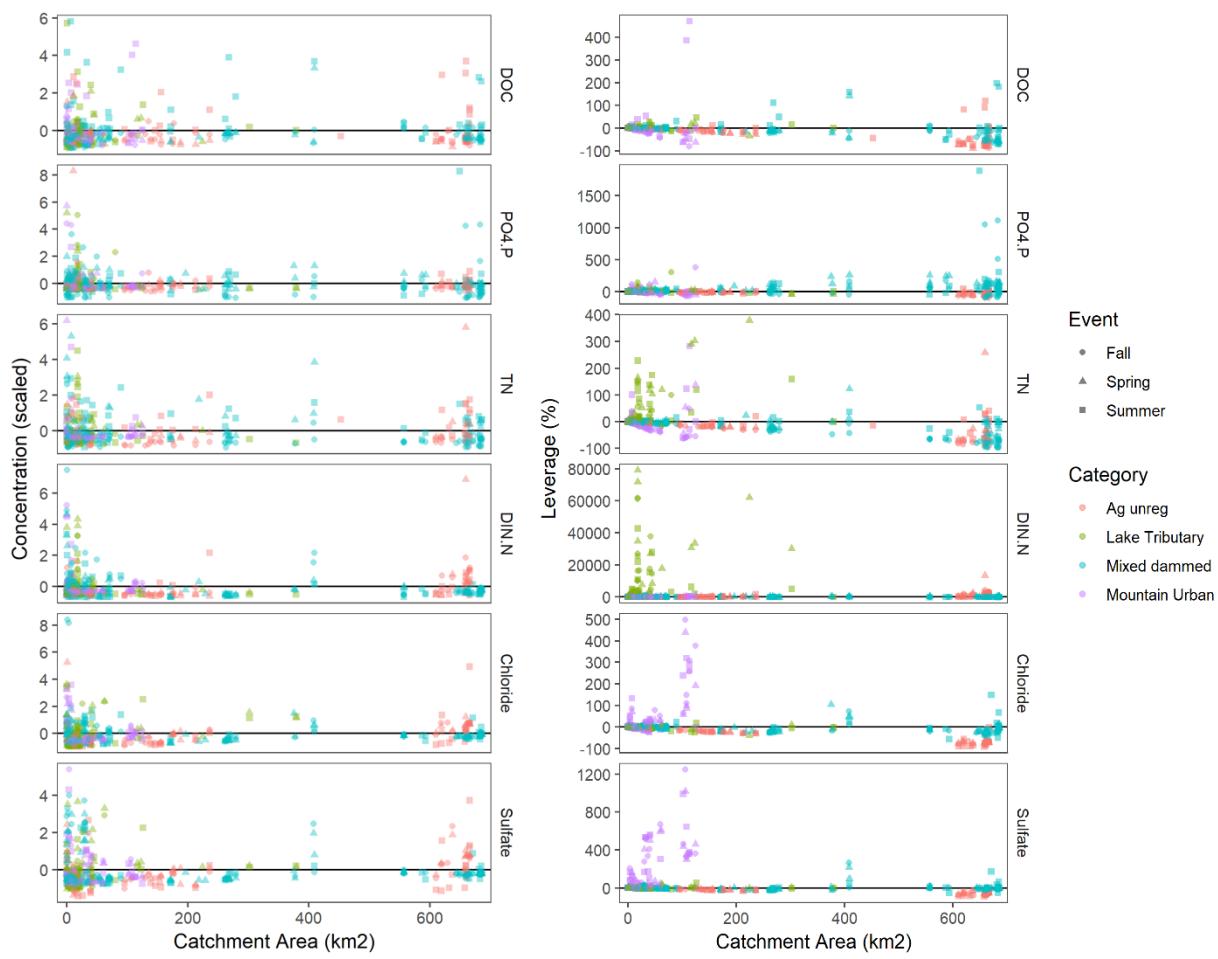


Figure 3-5. Scaled solute concentration (left) and leverage (right) by catchment area for solutes of interest (dissolved organic carbon, phosphorus, total nitrogen, dissolved inorganic nitrogen, chloride, and sulfate) for sites within four land use and hydrologic categories of the Utah Lake watershed. Samples were collected using participatory science volunteers on three synoptic sampling events conducted in March (Spring), July (Summer), and October (Fall) of 2018.

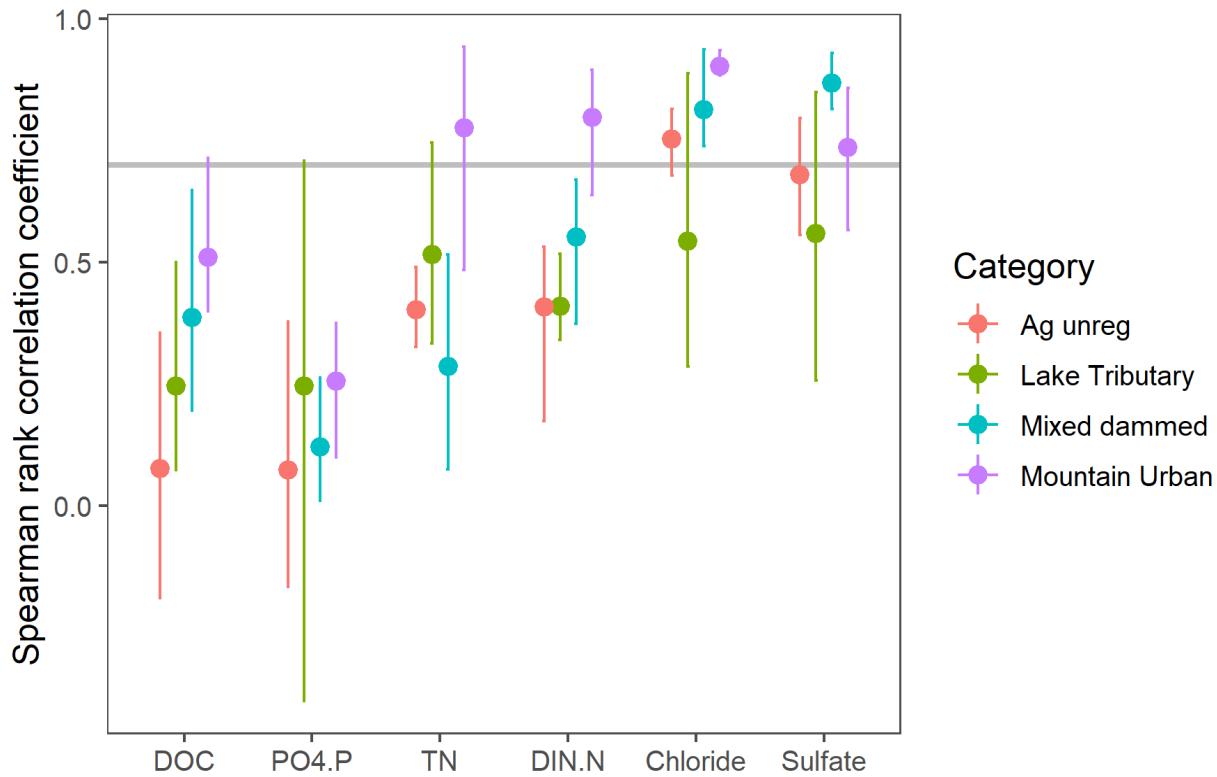


Figure 3-6. Spatial stability for solutes (dissolved organic carbon, phosphate, total nitrogen, dissolved inorganic nitrogen, chloride, and sulfate) for different categories of land use within the Utah Lake watershed. Spatial stability is calculated as a Spearman rank correlation by comparing multiple synoptic samplings (Spring, Summer, and Fall 2018) pairwise by site. Points represent mean Spearman, and error bars represent the range of values.

## TABLES

Table 3-1. Watershed characteristics for contributing areas to Utah Lake, calculated using USGS StreamStats. Discharge data represents annual discharge from 1980-2003 (PSOMAS 2007). Land use was calculated using 2011 NLCD (% forest, % developed, % impervious surface) and 1992 NLCD (% herbaceous upland). Where categories represent multiple subwatersheds, statistics for the major contributors are given.

| <b>Watershed</b>  | <b>Area<br/>(km<sup>2</sup>)</b> | <b>Mean<br/>Elevation<br/>(MASL)</b> | <b>Annual<br/>discharge<br/>(m<sup>3</sup> x 10<sup>6</sup>)</b> | <b>%<br/>Forest</b>          | <b>%<br/>Dev</b>            | <b>%<br/>Imp</b>      | <b>%<br/>Herb</b>       |
|---|----------------------------------|--------------------------------------|--|------------------------------|-----------------------------|-----------------------|-------------------------|
| <b>Mixed dammed</b><br>Provo River  | 1774                             | 2320                                 | 186.3  | 64.9                         | 5.5                         | 1.1                   | 3.5                     |
| <b>Agricultural<br/>unregulated</b><br>Spanish Fork River                             | 1725                             | 2137                                 | 123.0  | 55.9                         | 1.9                         | 0.5                   | 5.52                    |
| <b>Mountain Urban</b><br>American Fork<br>River<br>Hobble Creek                       | 160<br>298                       | 2493<br>2158                         | 7.28<br>24.4   | 69.7<br>58.4                 | 3.92<br>2.1                 | 1.3<br>0.6            | 3.5<br>7.54             |
| <b>Lake tributaries</b><br>Mill Race<br>Dry Creek<br>Currant Creek<br>Benjamin Slough | 46.6<br>111<br>1046<br>326       | 1899<br>2048<br>1896<br>1771         | 41.3<br>1.11<br>-<br>45.3  | 47.1<br>40.5<br>36.1<br>37.6 | 40.6<br>20.4<br>3.5<br>10.3 | 23<br>7<br>0.7<br>3.6 | 2.89<br>7.7<br>7.1<br>7 |
| <b>Utah Lake (total)</b>  | 7640                             | 1990                                 | 520.0  | 43.8                         | 10.7                        | 2.1                   | 6.2                     |

Table 3-2. Linear regressions of solute concentration (dissolved organic carbon, phosphorus, total nitrogen, dissolved inorganic nitrogen, chloride, and sulfate) by land use (% impervious, % developed, % forest, % herbaceous upland) and Event (Spring, Summer, and Fall). Model selection was done based on AICc scores of REML models.

| <b>Solute</b> | <b>Significant variables</b>      | <b>R<sup>2</sup></b> |
|---------------|-----------------------------------|----------------------|
| PO4-P         | % impervious                      | 0.146                |
| DIN-N         | % impervious                      | 0.245                |
| TN            | % impervious, Event               | 0.231                |
| DOC           | % impervious, Event               | 0.192                |
| Sulfate       | % impervious, % herbaceous, Event | 0.115                |
| Chloride      | % impervious, % herbaceous        | 0.123                |

## SUPPLEMENTAL MATERIAL

Table S1. ANOVA test comparing solute concentrations in streams from different land use categories over three synoptic sampling events in the Utah Lake watershed.

|                         | Df | Sum Sq | Mean Sq | F value | Pr (>F)  |     |
|-------------------------|----|--------|---------|---------|----------|-----|
| Event                   | 2  | 10.2   | 5.09    | 5.896   | 0.00279  | **  |
| Solute                  | 5  | 0.1    | 0.02    | 0.023   | 0.99978  |     |
| Category                | 3  | 248.8  | 82.94   | 96.055  | < 2e-16  | *** |
| Event:Solute            | 10 | 69.6   | 6.96    | 8.057   | 6.37e-13 | *** |
| Event:Category          | 6  | 10.6   | 1.76    | 2.042   | 0.05699  | .   |
| Solute:Category         | 15 | 56.3   | 3.76    | 4.349   | 4.03e-08 | *** |
| Event:variable:Category | 30 | 27.0   | 0.90    | 1.043   | 0.40212  |     |

Table S2. ANOVA test comparing spatial stability values for solutes measured in different land use categories over three synoptic sampling events in the Utah Lake watershed.

|                 | Df | Sum Sq | Mean Sq | F value | Pr (>F)  |     |
|-----------------|----|--------|---------|---------|----------|-----|
| Solute          | 10 | 5.245  | 0.5245  | 21.542  | < 2e-16  | *** |
| Category        | 3  | 1.750  | 0.5834  | 23.960  | 1.74e-11 | *** |
| Event           | 3  | 0.365  | 0.1216  | 4.993   | 0.0030   | **  |
| Solute:Category | 30 | 2.567  | 0.0856  | 3.514   | 2.21e-06 | *** |
| Category:Event  | 9  | 1.137  | 0.1264  | 5.190   | 1.14e-05 | *** |
| Solute:Event    | 30 | 1.278  | 0.0426  | 1.750   | 0.0229   | *   |
| Residuals       | 90 | 2.191  | 0.0243  |         |          |     |