UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

FRESHWATER STREAM MONITORING PROCESS IMPROVEMENTS FOR FECAL INDICATOR IMPAIRMENT DESIGNATION

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

GRANT MATTHEW GRAVES

Norman, Oklahoma

2023

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/337499

FRESHWATER STREAM MONITORING PROCESS IMPROVEMENTS FOR FECAL INDICATOR IMPAIRMENT DESIGNATION

A DISSERTATION APPROVED FOR THE

SCHOOL OF CIVIL ENGINEERING AND ENVIRONMENTAL SCIENCE

BY THE COMMITTEE CONSISTING OF

Dr. Jason Vogel, Chair

Dr. Kato Dee

Dr. Robert Nairn

Dr. Keith Strevett

Dr. Darrell Townsend

© Copyright by GRANT MATTHEW GRAVES 2023

All Rights Reserved.

To my former self who wasn't sure he could.

To my family who knew I could and

encouraged, loved and supported me

along my path.

Cheers!

Acknowledgments

I would first like to thank my incredible committee that I have had the opportunity to work with at OU. Dr. Nairn, Dr. Strevett, Dr. Dee and Dr. Townsend, I am very appreciative of your support in my research. Whether it was lunch chats, paper reviews, Zoom meetings, or coursework you have all made a positive impact on my foundation in research and becoming a role-model scientist and leader in my field. Thank you! Furthermore, I would like to acknowledge Dr. Tanner for his mentorship, guidance and support throughout many aspects of my research topics. You have inspired me to try new ideas, stretch the boundaries of research in the areas of microbiology and environmental science, and instill your knowledge and expertise into making my research superb. Thank you!

Dr. Vogel, first, I must say I'm glad you picked up the phone over 6 years ago to ask if I was interested in coming back for my PhD. As you would say, "it was a no-brainer", and I undoubtedly have enjoyed the journey. Thank you for all of the support, mentorship, friendship and challenging me to become a better scientist, mentor and leader. Cheers to you!

To all the students and staff, past and present, thank you! You have all made this work happen with your dedication to data collection including many long hours in the field and lab.

Lastly, I would like to acknowledge my family. Thank you Mom and Dad for always being there, loving and supporting me along the way and listening to me talk about bacteria and statistics for five years. To the love of my life, Alexandra, I'm not sure how I can express the love and shoulder to lean on that you have provided to me along the way, but I would not be at this point without you. Thank you and much love!! To my dogs, Cooper and Lyra, thanks for always being a distraction from the real world and playing tug-of-war or fetch with me.

Abstract

Recreational water quality standards for freshwater streams and rivers are important to understand the potential human health risks associated with primary body contact recreation. Indicator bacteria, *Enterococcus* and *Escherichia coli*, are used to routinely monitor and assess waterbodies for impairment. The 2020 Clean Water Act 303(d) Integrated Report indicates that approximately 7500 miles of streams and rivers are impaired for both *E. coli* and *Enterococcus* in Oklahoma. Fecal indicator bacteria (FIB) sources are often difficult to assess as they are from numerous anthropogenic, wildlife and environmental non-point sources and require consistent monitoring and assessment due to potential dynamic spatial and temporal factors within streams. The Oklahoma water quality standards provide threshold criteria and a general sampling frequency for FIB to make an impairment assessment, but do not provide guidelines for how, when, or where water samples should be collected in a waterbody. Furthermore, there is evidence from recent studies to suggest that *Enterococcus* may not be the strongest predictor for freshwater impairment criteria and may be of non-enteric origin and that fluorogenic substrate methods (i.e., Enterolert™ [ELT]) used to analyze *Enterococcus* samples may result in false positives. Resources are often limited for many agencies that routinely monitor these streams and new approaches and tools are needed to develop effective monitoring plans. Given the immense resource effort required to monitor and assess these streams, more research is needed to understand and improve the FIB monitoring process for primary body contact recreation assessment. Therefore, the objectives of this dissertation were to 1) evaluate spatial and temporal factors in Oklahoma streams that may influence FIB, 2) investigate stream sediment as a contributing factor to *Enterococcus* in streams and rivers, 3) evaluate the ELT enumeration method for applicability to analyze freshwater stream samples for *Enterococcus*, and 4) explore

existing geospatial and water quality data to develop correlation factors and regression equations to improve prediction of FIB for monitoring and assessment.

Studies that were conducted in this dissertation included a 1) field water quality spatiotemporal study at two cross sections in Spring Creek (Ch. 2), 2) spatiotemporal assessment of six streams and two laboratory microcosms for *Enterococcus* survivability in sediment and water and related environmental factors (Ch. 3), 3) investigation of the ELT method for *Enterococcus* false positives from stream water and sediment samples (Ch. 4), and 4) development of multiple linear regressions for FIB using water quality monitoring data collected from the Oklahoma Conservation Commission (Ch. 5). In brief, the results of these studies revealed that spatial and temporal factors (i.e., sampling location and time) and water quality and geographical characteristics (i.e., land use) can influence FIB in Oklahoma streams (Ch. 2 and Ch. 3). Furthermore, these spatiotemporal factors can be used to predict FIB concentrations in stream water and sediment (Ch. 3). *Enterococcus* showed extended survival and stability in stream sediments greater than 31-d under stable laboratory microcosms (Ch. 3). False positive bacteria were identified in 25% of all ELT samples analyzed with greater than 90% of those identified as *Paenibacillus* spp. from the microcosm and field studies (Ch. 4). Regression equations can be developed from water quality and geospatial variables to provide an initial reconnaissance of the expected FIB concentrations within a stream and/or region (Ch. 5). The outcomes of this work indicate that more emphasis should be placed on evaluation of the sampling process design and methodology for assessing Oklahoma streams for FIB impairment determination.

Table of Contents

Acknowledgments v

Abstract vi

Table of Contents viii

List of Tables xi

List of Figures xiii

Chapter 1 : Introduction 1

1.1 Research Questions 1

1.2 Hypotheses 1

1.3 Water Quality Status and History of Surface Waters in Oklahoma 2

1.4 Dissertation Objectives 3

1.5 References 7

Chapter 2 : Spatiotemporal Variability Comparisons of Water-Quality and *Escherichia coli* in an Oklahoma Stream 9

2.1 Introduction 10

2.2 Methods 13

2.3 Results and Discussion 18

2.4 Conclusions 24

2.5 References 25

Chapter 3 : Investigation of environmental factors on *Enterococcus* survival in Oklahoma streams 32

3.1 Introduction 33

3.2 Methods 35

3.3 Results and Discussion 42

3.4 Conclusions 59

3.5 References 60

Chapter 4 : The Effect of *Paenibacillus* on IDEXX Enterolert™ Results from Stream Environments 70

4.1 Introduction 71

4.2 Methods 73

4.3 Results and Discussion 78

4.4 Conclusions 94

4.5 References 95

Chapter 5 : Exploring Regional and Statewide Relationships between Fecal Indicator Bacteria, Water Quality and Geospatial Data in Oklahoma Streams 102

5.1 Introduction 103

5.2 Methods 105

5.3 Results and Discussion 113

5.4 Conclusions 134

5.5 References 136

Chapter 6 : Conclusions and Future Directions 144

6.1 Research Summary 145

6.2 Recommendations for Future Research 146

Appendices 149

Appendix 1. Correlogram of environmental factors analyzed in Chapter 3. 150

Appendix 2. Boxplots of water quality and hydrologic variable data used in the Chapter 5 analyses. 151

Appendix 3. Calculating preceding rainfall and antecedent dry periods from Mesonet data 162

List of Tables

Table 3-1. Descriptive geographical and soil data for each stream sampling site in the study.. .. 43

Table 3-2. Summary statistics of water quality, hydrologic and fecal indicator bacteria (FIB) concentrations for sampling locations during the field sampling events (n=10). 44

Table 4-1. Summary of isolates from Enterolert™ panels analyzed from the Walnut Creek microcosm stream sediment for a 31-d study duration. 81

Table 4-2. Isolate species identifications from Enterolert™ panels from two microcosms at two time periods, D_I and D_F, from two stream sediment sources 83

Table 4-3. Summary of isolates from Enterolert™ panels at four stream sampling locations at two time periods, T_I and T_F 85

Table 4-4. Isolate species identifications from Enterolert™ panels from four field stream sample locations at two time periods, T_I and T_F 87

Table 4-5. Antibiotic resistances of isolates recovered from Enterolert™ panels from six waterbodies. 93

Table 5-1. Data sources and types that are incorporated into FIB statistical analyses. 107

Table 5-2. Description of parameters and factors collected for the tabular datasets to be incorporated into FIB statistical analyses. 108

Table 5-3. Descriptive statistics of the contributing drainage areas (CONTDA) and sites within each region 114

Table 5-4. Significantly correlated ($p < 0.05$) variables from the Spearman correlation for all water all hydrologic, water quality and geospatial variables from 308 stream sampling locations and 2860 sampling points at Oklahoma streams. 122

Table 5-5. Spearman correlation strength association table for *Escherichia coli* and *Enterococcus* by region 125

Table 5-6. Significant ($p < 0.05$) linear regression model coefficients for *Enterococcus* (dependent variable) by region and by model predictors (independent variables)... 131

Table 5-7. Significant ($p < 0.05$) linear regression model coefficients for *Escherichia coli* (dependent variable) by region and by model predictors (independent variables)... 132

List of Figures

Figure 1-1. Process diagram of the regulatory decision process for impairment status of Oklahoma freshwater bodies. 5

Figure 1-2. Process diagram of Oklahoma water quality standards, the resources required for impairment of freshwater bodies, and the proposed research contributions to improve the scientific process of impairment determination..... 6

Figure 2-1. Site sampling locations at Spring Creek in central Oklahoma..... 15

Figure 2-2. Cross-section water-quality sampling at the "run" location at Spring Creek. 17

Figure 2-3. Standard box plots of a) T, b) DO, c) SC and d) Turbidity..... 20

Figure 2-4. Combination of plot of morning and afternoon TSS at the pool and run cross sections.. 22

Figure 2-5. *Escherichia coli* concentrations at two cross section locations (pool and run) at the Spring Creek study site for two time periods, morning (0800 and 0815) and afternoon (1500 and 1515). 23

Figure 3-1. Map showing sampling site locations, delineated watershed basins of sampling sites and associated river drainage basins..... 38

Figure 3-2. Time-series of *Enterococcus* sediment concentrations (log scale) for sampling sites during July-September for a total of 10 sampling visits at each of the six locations..... 47

Figure 3-3. Pearson correlation matrix of all parameters measured from field water samples and associated soil, geographical and hydrologic data..... 49

Figure 3-4. Scatter plot of log Enterococcus water versus log Enterococcus sediment concentrations from water and sediment samples collected during the field study over a period of ten sampling events. 53

Figure 3-5. Time-series plot of microcosm Enterococcus sediment concentrations over a period of 31 days from Cat Creek and Walnut Creek.. 57

Figure 5-1. Site map of Oklahoma with selected subregions and watersheds. 111

Figure 5-2. Site locations with delineated subbasins for each site location..... 115

Figure 5-3. Log-Boxplot of Enterococcus concentrations from stream sampling locations statewide by year for the period of 2001-2011..... 117

Figure 5-4. Log-Boxplot of E. coli concentrations from stream sampling locations statewide by year for the period of 2001-2011..... 118

Figure 5-5. Spearman correlation matrix of all hydrologic, water quality andgeospatial variables from 308 stream sampling locations and 2860 sampling points at Oklahoma streams.. 120

Chapter 1 : Introduction

The dissertation is formatted as a series of four separate publication-style research projects (Chapter 2-5) that are interconnected with the general research topic of fecal indicator bacteria in freshwater stream environments. An introduction and dissertation objectives are provided in Chapter 1 and conclusions and future research directions are discussed in Chapter 6.

1.1 Research Questions

The specific research questions for this dissertation are:

a) How can spatiotemporal environmental factors play a role in fecal indicator bacteria concentrations in Oklahoma freshwater streams?

b) Can *Enterococcus* survive and stabilize in freshwater stream sediments and water?

c) Is Enterolert™ (ELT) an appropriate method for evaluating *Enterococcus* concentrations in freshwater streams?

d) What can water quality and hydrologic data reveal about the State's efforts to evaluate, mitigate and enhance surface waters for beneficial uses?

1.2 Hypotheses

The hypotheses for this dissertation are developed based on the research questions in Section 2.2 and are denoted by chapter in parentheses at the end of each statement.

a) Time of day and stream location will influence fecal indicator bacteria concentrations and water quality parameters. (Chapter II)

b) *Enterococcus* can survive and stabilize over a 30-d period under controlled laboratory conditions. (Chapter III)

c) ELT will not inhibit growth of non-enterococci or streptococci species, resulting in false positives. (Chapter IV)
d) Spatial (regional) factors, watershed characteristics and water quality parameters will influence *Escherichia coli* and *Enterococcus* concentrations (Chapter III, Chapter V).

1.3 Water Quality Status and History of Surface Waters in Oklahoma

Since the passage of the Clean Water Act of 1972, scientists have worked to advance techniques to monitor, assess, remediate and preserve water resources (USEPA, 2019). With emerging technologies in modeling, engineering, and technologies, our work has become more efficient, yet more challenging, to solve critical issues in water resources. Regulations at both the federal and state levels created metrics for achievement and to "restore and maintain the chemical, physical and biological integrity of the Nation's waters" (USEPA, 2013). While we have reached many milestones in water resources protections and enhancements, new insights and advancements in technologies are needed as dynamic shifts are evident in water quality and quantity demands (Vliet et al., 2021). Furthermore, analyses of decades of data and re-evaluation of criteria and funding, such as Keiser & Shapiro (2019) indicated, are needed to improve our understanding of the next phases of a project and progress of our successes. Future advancements in best management practices and water treatment technologies are also needed as urban development and anthropogenic activities threaten hydrologic and ecosystem balances (Teurlincx et al., 2019).

The State of Oklahoma has a specific set of water quality standards (WQS) promulgated by the Oklahoma Water Resources Board (OWRB) that apply to specific beneficial use criteria such as recreational waters, fish and wildlife propagation and agriculture (OWRB, 2017).

Additionally, point-source, or end-of-pipe, regulations that set forth requirements for treated effluent into waterways are regulated by the Oklahoma Department of Environmental Quality (ODEQ, 2019). However, non-point sources polluting waterbodies are more challenging to address due to the complexity of potential sources (OCC, 2019). No direct authority regulates non-point sources in Oklahoma, but many are working towards monitoring, evaluating and developing innovative solutions to address water quality issues (OWRB, 2020b). Many gaps exist due to the enormity of resources required to monitor, assess, and develop solutions for impacted waters (OWRB, 2020b). Therefore, new, cost-effective solutions are needed using a combination of existing and applied techniques and technologies to assess and mitigate impacts to the State's waters.

1.4 Dissertation Objectives

The purpose of this dissertation is to explore a subset of water challenges within Oklahoma streams that pertain to fecal indicator bacteria and develop methodology, results, and scientific processes that can be used to improve the waters for the State and other regions. The topic area was chosen based on support and need from external entities at the federal and state levels, and include 1) investigation of *Enterococcus* as an indicator in freshwater streams and related environmental factors, 2) water quality analyses of historical stream data to understand and relate water quality conditions to fecal indicator bacteria, 3) spatial and temporal correlations between fecal indicator bacteria and abiotic and physiochemical parameters, and 4) evaluation of the fluorescent indicator method (ELT) for determining *Enterococcus* concentrations in freshwater.

Fecal indicator bacteria are important for evaluating stream reaches as a part of the Clean Water Act and Oklahoma beneficial use water quality standards to provide awareness of human

health risks. However, the water quality standards set forth by the State of Oklahoma and approved by the USEPA do not provide explicit details and scientific methodology of the how, where, when and why to sample freshwater streams. Furthermore, questions persist with using *Enterococcus* as a freshwater indicator and the methods that are used to quantify impairment status, and potentially, provide evidence for removal of streams from the 303(d) list that may not be impaired due to unexplored environmental factors, inaccurate methodologies, and/or varying sampling protocols. Process diagrams are presented in Fig. 1-1 and Fig. 1-2 which provide a conceptual view of a) the general process of assessing and designating impairment criteria to streams (Fig. 1-1), and b) the contributions that this dissertation will explore to provide steps to improve the scientific rigor of assessment criteria for fecal indicator bacteria in Oklahoma streams and rivers (Fig. 1-2). Within the Fig. 1-2 process diagram, the box in the far right indicates the dissertation research questions and how they interrelate with the regulatory decision process for fecal indicator bacteria and water quality standards.

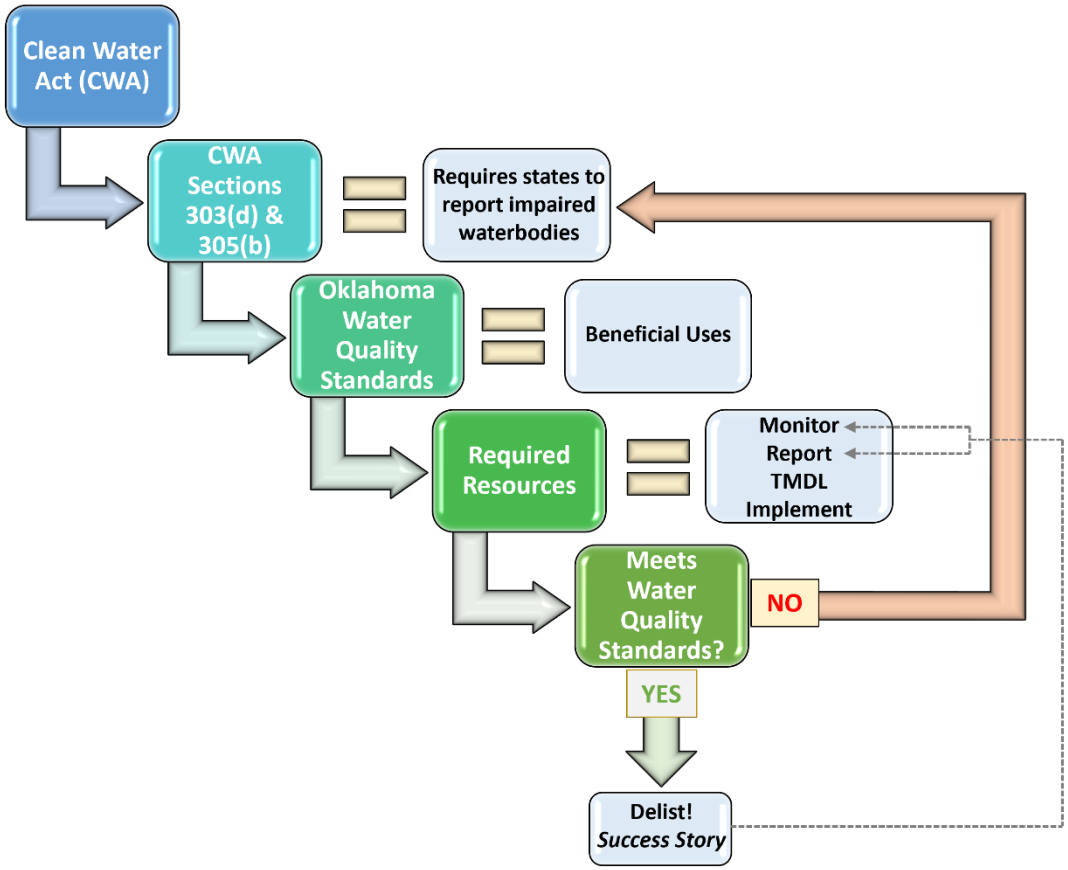


Figure 1-1. Process diagram of the regulatory decision process for impairment status of Oklahoma freshwater bodies.

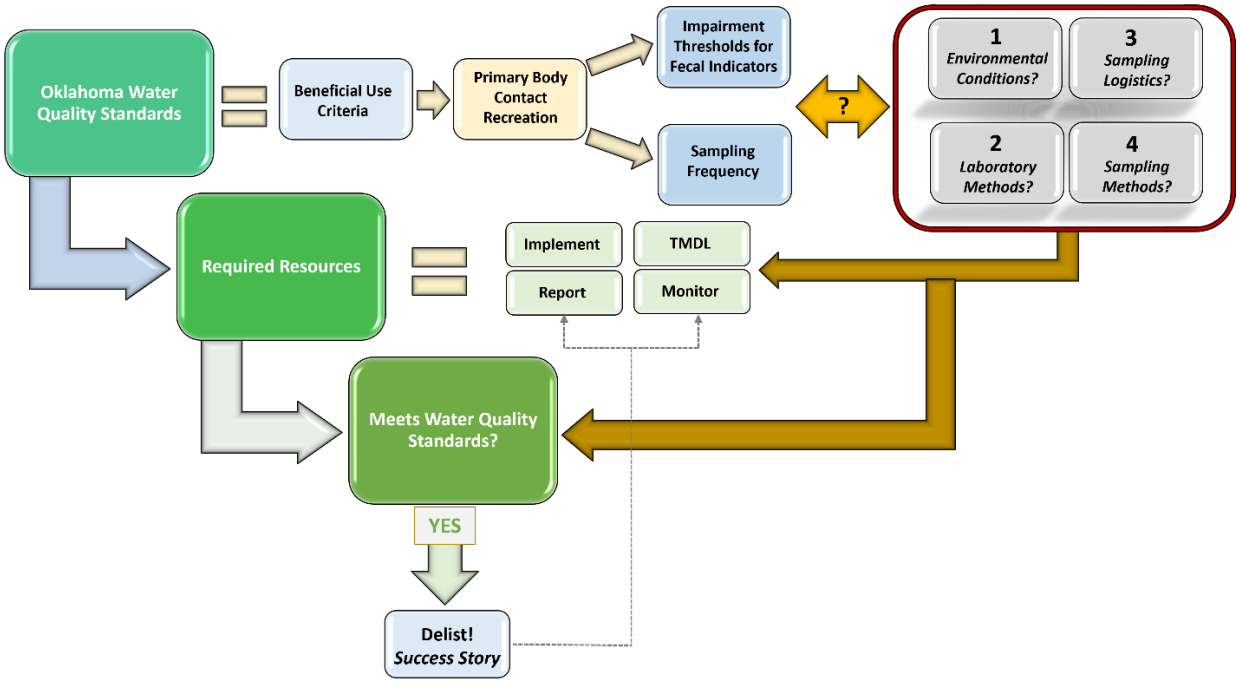


Figure 1-2. Process diagram of Oklahoma water quality standards, the resources required for impairment of freshwater bodies, and the proposed research contributions to improve the scientific process of impairment determination. Research questions related to the fecal indicator bacteria monitoring process for recreational water quality criteria are identified in the upper right box of the diagram and are 1) environmental conditions, 2) laboratory methods, 3) sampling logistics, and 4) sampling methods.

1.5 References

Keiser, D. A., & Shapiro, J. S. (2019). Consequences of the Clean Water Act and the Demand for Water Quality. The Quarterly Journal of Economics, 134(1), 349–396. https://doi.org/10.1093/qje/qjy019

ODEQ (Oklahoma Department of Environmental Quality) (2019). Title 252. Department of Environmental Quality Chapter 606. Oklahoma Pollutant Discharge Elimination System (OPDES) Standards.

Oklahoma Conservation Commission (OCC) (2019). Oklahoma's Nonpoint Source Management Program Plan 2019-2029.

OWRB (Oklahoma Water Resources Board) (2017). Oklahoma Water Quality Standards. Oklahoma Administrative Code 785:45. oar.state.ok.us. Accessed 21 Feb. 2021.

OWRB (Oklahoma Water Resources Board) (2020b). Status of Water Quality Monitoring in Oklahoma 2019-2020. https://www.owrb.ok.gov/quality/monitoring/StatusReport.pdf. Accessed 13 March 2021.

Teurlincx, S., Kuiper, J. J., Hoevenaar, E. C., Lurling, M., Brederveld, R. J., Veraart, A. J., Janssen, A. B., Mooij, W. M., & de Senerpont Domis, L. N. (2019). Towards restoring urban waters: Understanding the main pressures. Current Opinion in Environmental Sustainability, 36, 49–58. https://doi.org/10.1016/j.cosust.2018.10.011

USEPA (United States Environmental Protection Agency). (2013). Summary of the Clean Water Act [Overviews and Factsheets]. US EPA. https://www.epa.gov/laws-regulations/summary-clean-water-act

USEPA (United States Environmental Protection Agency). (2019). Milestones in EPA's Water Research History [Overviews and Factsheets]. US EPA. https://www.epa.gov/research/milestones-epas-water-research-history

Vliet, M. T. H. van, Jones, E. R., Flörke, M., Franssen, W. H. P., Hanasaki, N., Wada, Y., & Yearsley, J. R. (2021). Global water scarcity including surface water quality and expansions of clean water technologies. Environmental Research Letters, 16(2), 024020. https://doi.org/10.1088/1748-9326/abbfc3

Chapter 2 : Spatiotemporal Variability Comparisons of Water-Quality and *Escherichia coli* in an Oklahoma Stream

*This chapter is a **published** research note in the Journal of Contemporary Water Research and Education that has been formatted for the dissertation. The citation is as follows:*

Graves, G.M. and Vogel, J.R. (2023). Spatiotemporal Variability Comparisons of Water-Quality and *Escherichia coli* in an Oklahoma Stream. Journal of Contemporary Water Research & Education (JCWRE).

Abstract

Fecal indicator bacteria, *Escherichia coli*, for primary body contact recreation (PBCR) in Oklahoma waterbodies, is defined as the geometric mean of ten samples from the recreation season, May 1 to September 30, with an impairment threshold of 126 colony forming units (cfu) per 100 mL. However, the water quality standards provide limited guidance on spatiotemporal and environmental factors that could influence samples collected and analyzed. In this study, two stream cross sections under baseflow conditions in a central Oklahoma urban perennial stream, Spring Creek, were densely sampled to investigate temporal and spatial variability of *E. coli* concentrations and water quality parameters across the stream channel. Water quality parameters (specific conductivity, temperature, dissolved oxygen, pH, turbidity, and total suspended solids), stream discharge, and bacteria samples were collected simultaneously at equal intervals across the two cross sections in the morning and afternoon during one summer day with sunny, dry, and hot weather conditions. Results indicate a significant difference between time-of-day samples and water quality parameters and *E. coli* concentrations. Strong correlations between temperature, dissolved oxygen, and time versus *E. coli* concentrations were observed, while location, turbidity, and TSS were not significant or correlated to measured values. Furthermore,

E. coli concentrations were highly variable spatially across each stream cross section, regardless of time of day or location. Results from this study provide an initial indication that stream water quality, spatial cross section sample location, and diurnal variations may be influencing factors on bacteria concentrations.

Keywords

Escherichia coli, fecal indicator bacteria, sampling, freshwater, stream

2.1 Introduction

Fecal indicator bacteria (FIB) such as *Escherichia coli* in freshwater waterbodies are frequently monitored to assess potential human health risk from pathogen contact in recreational waters. The State of Oklahoma and U.S. Environmental Protection Agency water quality standard criteria for FIB for primary body contact recreation (PBCR) in waterbodies is defined as the geometric mean of 10 samples from the recreation season, May 1 to September 30, with an impairment threshold of 126 colony forming units (cfu) per 100 mL for *E. coli* (OWRB 2017). Thresholds were derived from epidemiology studies in freshwater and marine swimming beach areas in lakes and oceans where subjects contacted potential contaminated water and incidents of gastrointestinal illness occurred (USEPA 1986; 2012).

E. coli has been studied extensively for fecal source tracking, pathogenic strains, waterbody conditions, and other associated research questions related to human health risk and fecal water quality indicators for PBCR (Gitter et al. 2020). However, water quality standards provide limited guidance of how samples should be collected during the recreation season. State agencies and other entities that collect samples and make assessments often develop their own sampling metrics but are not standardized to sampling protocols (USEPA 2012). The U.S.

Environmental Protection Agency and others recognize that temporal and spatial factors could play significant roles in bacteria concentrations within a stream (USEPA 2010; Muirhead and Meenken 2018). Recent studies found that sampling location and frequency were significant factors when developing a monitoring plan to obtain representative samples for the evaluation of potential fecal contamination (Crosby et al. 2019; Stocker et al. 2019). In addition, previous research has indicated that sample type and technique when monitoring a stream should be considered to reduce uncertainty in analyses (Harmel et al. 2016). Gregory et al. (2019) determined there was a significant difference between streamflow thresholds (i.e., baseflow, floods) and *E. coli* concentrations, and indicated that specific hydrologic factors may provide stronger relationships to FIB stream concentrations and associated human health risk. Therefore, given the number of temporal and spatial factors within a stream sample reach, determining sample representativeness could be an important consideration for waterbody impairment designation.

Stream characteristics and environmental conditions have been shown to influence FIB and have been used to develop relationships between parameters and FIB concentrations (Dwivedi et al. 2013). Particularly, suspended solids, turbidity, water temperature, and habitat have previously been used as predictors for *E. coli* densities (Desai and Rifai 2010; Petersen and Hubbart 2020). Others have found significant relationships between nutrients, turbidity, and FIB in streams that can be used to predict bacteria concentrations (Christensen et al. 2002). Furthermore, discharge and precipitation, along with turbidity, have been found to strongly correlate with *E. coli* concentrations in streams (Hamilton and Luffman 2009). Comparison of stream reaches within similar land use segments has been explored with differentiating results for variable fecal indicator concentrations and environmental conditions (Stocker et al. 2016).

Results indicated that there were significant differences between stream sampling locations, and that more research is needed to understand stream dynamics that may affect FIB. Diurnal variation and sunlight are also important considerations for evaluating FIB in streams and rivers (Desai and Rifai 2013). Previous research has indicated that FIB concentrations in waterbodies is cyclical, with decay shown during high sunlight periods and increases in bacteria concentrations during low light periods (Whitman et al. 2004; Schultz-Fademrecht et al. 2008). Hydrologic extremes such as floods and droughts can increase variability within stream reaches due to external bacterial inputs from stormwater conveyance, wastewater overflows, and non-point sources (Vogel et al. 2009; McKergow and Davies-Colley 2010; Sanders et al. 2013; Verhougstraete et al. 2015; Rochelle-Newall et al. 2016; Stocker et al. 2018). Furthermore, Piorkowski et al. (2014) showed a variable spatial distribution of FIB in stream sediments under different flow conditions and sampling location. Sediment type and stream habitats have also shown to be *E. coli* reservoirs within streams (Brinkmeyer et al. 2015; Devane et al. 2020). Stream bed sediments have the potential to provide a consistent source of resuspended FIB in the stream water column due to dynamic hydrologic conditions and can create variable sampling conditions (Jamieson et al. 2005; Haller et al. 2009; Bradshaw et al. 2016).

While environmental and hydrologic conditions have been extensively studied to develop relationships between these factors and *E. coli* within streams and rivers, limited information exists to understand the variability of bacteria concentrations within the longitudinal and cross-section profiles of streams. The objectives of this study were to 1) investigate spatial and temporal variability in two stream cross sections, 2) evaluate physical and chemical factors for correlations between variables and evaluate statistical trends, and 3) provide preliminary information for future research targeting specific environmental and spatiotemporal factors that

may influence bacteria concentrations in streams and rivers, and ultimately, drive impairment criteria for water quality monitoring.

2.2 Methods

Two stream cross sections in a central Oklahoma urban perennial stream, Spring Creek, under baseflow conditions (less than 2.54 mm precipitation in previous seven days) were densely sampled during a seasonally average dry and hot, central Oklahoma summer day (Figure 1). Additionally, in-situ water quality parameters were collected across the stream channel sections at sampling points. Spring Creek is located in northwest Oklahoma City, OK at 35° 36' 18.7" N and -97° 36' 29.3" W, and the site location has an approximate drainage area of 30 km^2 as calculated in Stream Stats (Smith and Esralew 2010). The land use category of the watershed is highly urban (>90%) with silty clay to clay loam soil types (USDA NRCS 2023). Potential bacteria inputs are primarily from non-point sources from urban runoff, as no septic tanks, wastewater discharges, or agriculture are located in the watershed. Stream cross sections were evaluated at two daily time periods, morning (0800) and afternoon (1500), at two locations. The two measured cross section stream feature morphologies were a pool (upstream) and a run (downstream) and were separated by 200 m of a series of riffles, glides, pools, and runs. The upstream cross section had a width of 6.7 m and downstream location had a cross section width of 7.3 m.

Factors investigated included *E. coli* concentration, dissolved oxygen (DO), specific conductivity (SC), total suspended solids (TSS), turbidity, water temperature (T), stream velocity and flow, channel depth, stream location and cross section, and time. Water quality samples and parameters were collected across the cross section simultaneously by our sampling team for evaluation of spatial variability (Figure 2-1). Grab samples were collected at evenly spaced 1.2 m

cross section locations (minimum of six sampling locations) at mid-depth in sterile 1 L polypropylene bottles and split into respective subsamples for bacteria (*E. coli*), water quality parameters (turbidity, pH, conductivity), and sediment (TSS) analyses (Figure 2). Sampling protocols adhered to the U.S. Geological Survey sampling methods (USGS 2014). Discharge measurements were collected using a Sontek Flowtracker2® handheld-ADV (acoustic Doppler velocimeter) at each cross section, following collection of water quality samples. At each time period, samples were first collected at the downstream location to minimize disturbance of the water column from the upstream location. *E. coli* concentrations in water were analyzed using IDEXX Quantitray Colilert (SM9223-B) to determine most probable number (MPN) per 100 ml (Baird and Bridgewater 2017). TSS analyses were completed using SM 2540-D and turbidity was measured using a Hach® portable turbidity meter. Water temperature, pH, DO, and SC were measured using a ThermoFisher Scientific Orion Star A329 multiparameter meter.

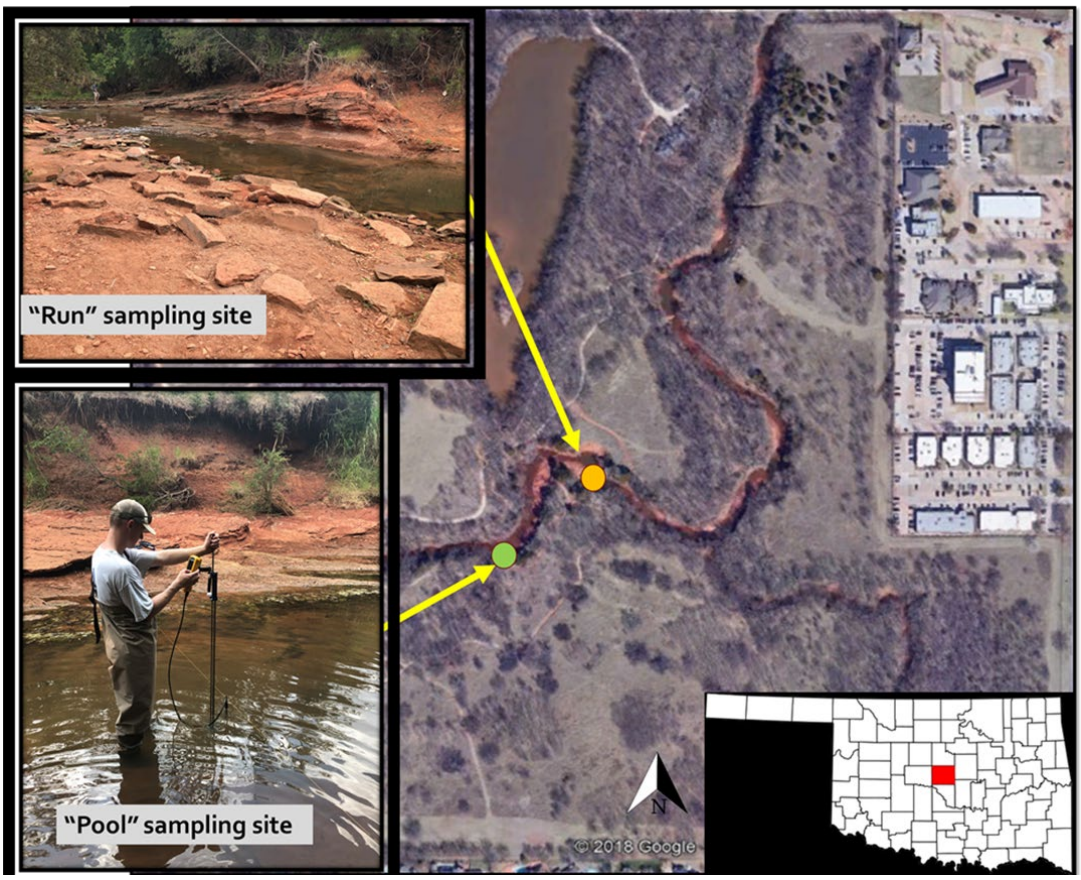


Figure 2-1. Site sampling locations at Spring Creek in central Oklahoma.

Factors investigated included *E. coli* concentration, dissolved oxygen (DO), specific conductivity (SC), total suspended solids (TSS), turbidity, water temperature (T), stream velocity and flow, channel depth, stream location and cross section, and time. Water quality samples and parameters were collected across the cross section simultaneously by our sampling team for evaluation of spatial variability (Figure 2-2). Grab samples were collected at evenly spaced 1.2-m cross section locations (minimum of six sampling locations) at mid-depth in sterile 1L-polypropylene bottles and split into respective subsamples for bacteria (*E. coli*), water quality parameters (turbidity, pH, conductivity), and sediment (TSS) analyses (Figure 2-2). Sampling protocols adhered to the U.S. Geological Survey sampling methods (USGS, 2014). Discharge measurements were collected using a Sontek Flowtracker2® handheld-ADV (acoustic Doppler velocimeter) at each cross section following collection of water quality samples. At each time period, samples were first collected at the downstream location to minimize disturbance of the water column from the upstream location. *E. coli* concentrations in water were analyzed using IDEXX Quantitray Colilert (SM9223-B) to determine most probable number (MPN) per 100-ml (Baird and Bridgewater 2017). TSS analyses were completed using SM 2540-D and turbidity was measured using a Hach® portable turbidity meter. Water temperature, pH, DO and SC were measured using a ThermoFisher Scientific Orion Star A329 multiparameter meter.



Figure 2-2. Cross-section water-quality sampling at the “run” location at Spring Creek.

Data Analysis

Data were analyzed using Microsoft Excel® and R statistical software. Differences in means were evaluated using a two-sample t-test with unequal variances. A Pearson correlation test with a two-sample t-test with unequal variances was performed to determine significant linear relationships between variables. An F-test was used to evaluate variance of water quality data collected from each stream section. All statistical figures were generated using R and Excel®.

2.3 Results and Discussion

Stream flow characteristics were measured at both the morning and afternoon sampling periods. Stream locations had mean column depths of 0.35 m at the pool and 0.15 m at the run. Discharge during the morning and afternoon periods (measurement was within ± 0.01 m^3s^{-1} at both the upstream and downstream locations) was 0.08 m^3s^{-1} and 0.04 m^3s^{-1}, respectively, which is within range of the estimated 50% flow-duration for Spring Creek in July (0.05 m^3s^{-1}) (Smith and Esralew 2010). The drainage area is characterized as highly urban, silty clay soils (hydrologic soil group D), which could increase the potential for anthropogenic influences and explain the higher discharge in the morning period when lawn irrigation is most common. No measurable precipitation (>2.54 mm) was recorded at the nearest Oklahoma City East Mesonet station for the preceding seven days (Brock et al. 1995; McPherson et al. 2007).

From a two-sample t-test with unequal variances, *E. coli* concentrations between the upstream (pool) and downstream (run) were not significantly different between the means for each location for all time periods ($p=0.23$). However, a significant difference ($p<0.001$) between time periods (morning and afternoon) was shown between each location for *E. coli* densities. The

geometric mean in the morning for *E. coli* was 664 MPN/100 ml (SD ± 116) and was 137 MPN/100 ml (SD ± 108) in the afternoon. Results from a t-test comparing Pearson correlation coefficients between factors indicate that time, DO, SC, and T were significant ($p < 0.05$) for *E. coli* concentrations. Furthermore, DO was significantly higher ($p < 0.01$) in the morning than afternoon and displayed a strong positive correlation of 0.69 to *E. coli* concentrations. Conversely, a very strong negative correlation (-0.93) of T was shown and a strong positive relationship with SC (0.78) was found versus *E. coli* concentrations ($p < 0.01$). The mean DO and T for both locations was 9.05 mg/L (SD ± 0.12) and 26.63°C in the morning, and 7.02 mg/L (SD ± 0.42) and 29.14°C in the afternoon. When comparing SC to *E. coli* concentrations, a significant difference was statistically determined, however, the means of SC for the morning and afternoon were 1167 (SD ± 1.72) and 1171 µS/cm (SD ± 4.59), respectively, which provides limited inference for interpretation given the minute difference between time points. However, the flow was a factor of two higher in the morning than in the afternoon and could suggest that more flow slightly altered the water chemistry through dilution. Significant differences ($p < 0.01$) were found from the Pearson correlation coefficient t-test when comparing *E. coli* concentrations from both sampling locations to water quality parameters (DO, T), water column depth, and time. However, no significant differences were shown ($p > 0.05$) for TSS, turbidity, and stream velocity. Boxplots of water quality parameters are shown in Figure 2-3.

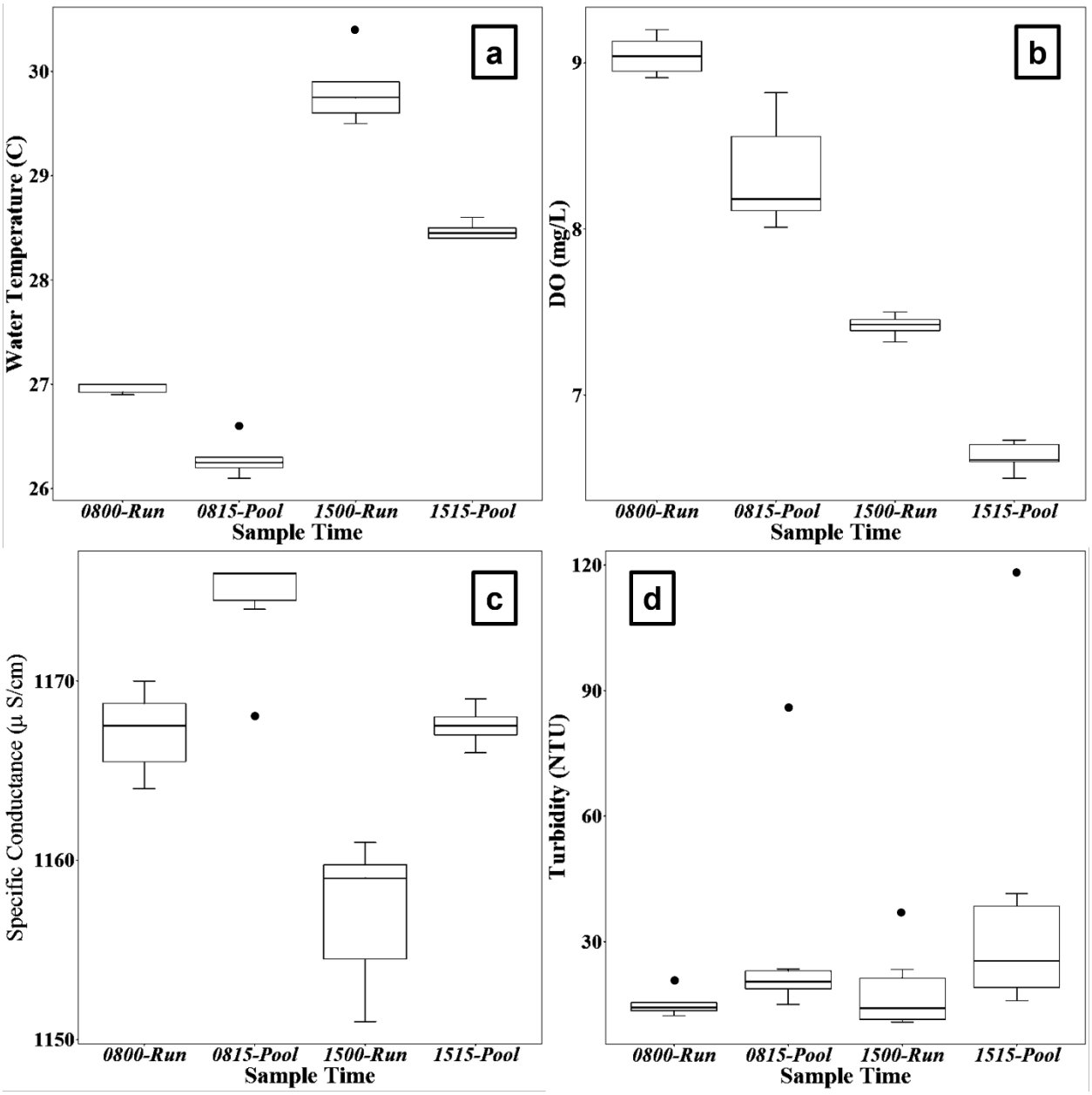


Figure 2-3. Standard box plots of a) T, b) DO, c) SC and d) Turbidity showing the median (line in box), lower (Q1) and upper (Q3) (T bars outside of box) and outlier values (points) grouped by sample time at each of the two Spring Creek sampling locations.

Previous research has indicated that sediment parameters are strong predictors for FIB sampling (Stocker et al. 2019). However, our results from the Pearson correlation indicated high variability and no significant relationship between turbidity, TSS, and *E. coli* for each cross section and location. Stream cross section versus TSS is presented in Figure 2-4, and visually demonstrates the variability of suspended sediments at time points and cross section location. Stream cross sections at both locations were evaluated using a two-sample F-test to determine if variability exists across the lateral profile of the stream for *E. coli* concentrations. Results show significant high variability between the pool and run locations ($p<0.01$) at both times, where the standard deviation was approximately a factor of three lower in the run location than the pool location. No significant difference in variability was found when comparing two time periods for the pool location ($p=0.44$), whereas a significant difference was indicated for the run location ($p=0.038$) when comparing two different time periods. *E. coli* stream cross section concentrations for two time periods and locations are displayed in Figure 2-5.

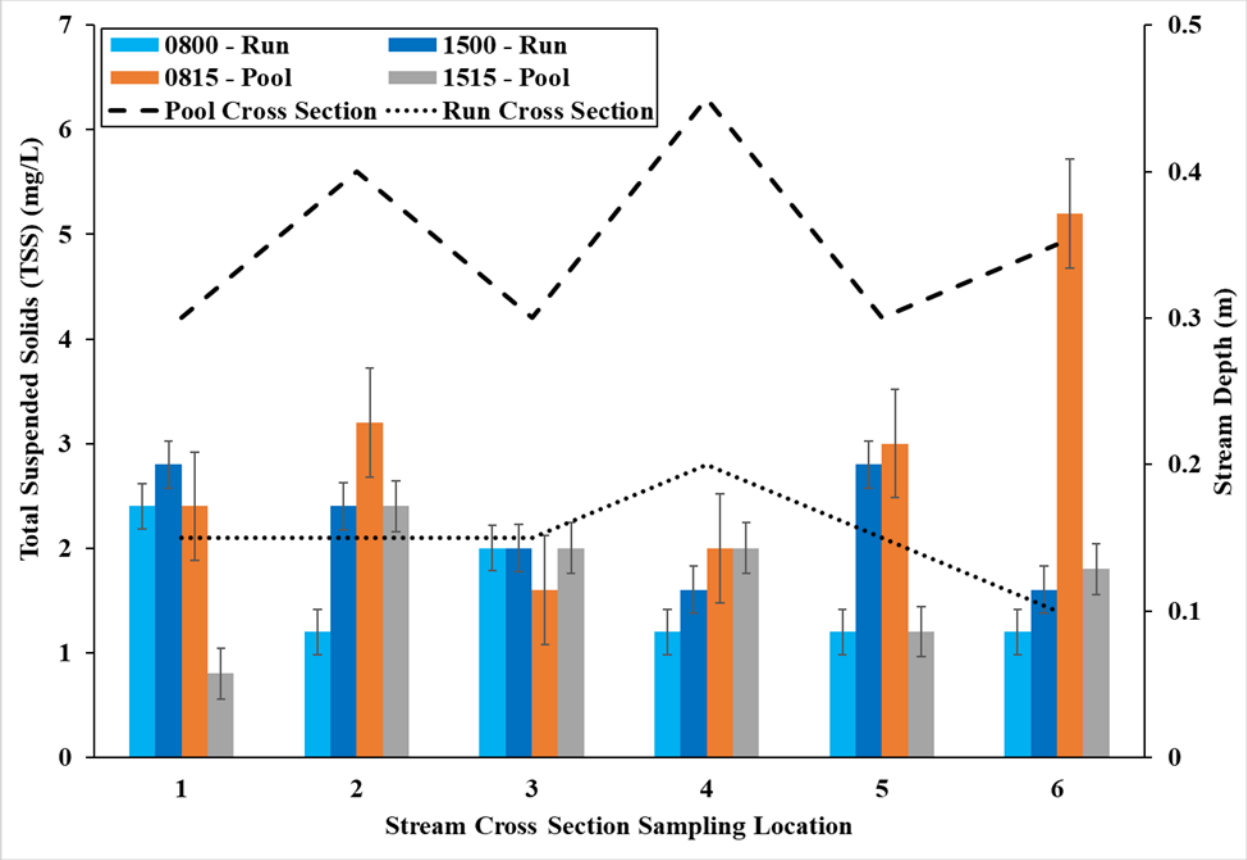


Figure 2-4. Combination of plot of morning and afternoon TSS at the pool and run cross sections. Cross section depth for each location is indicated by the dashed lines. Standard error is represented by the error bars.

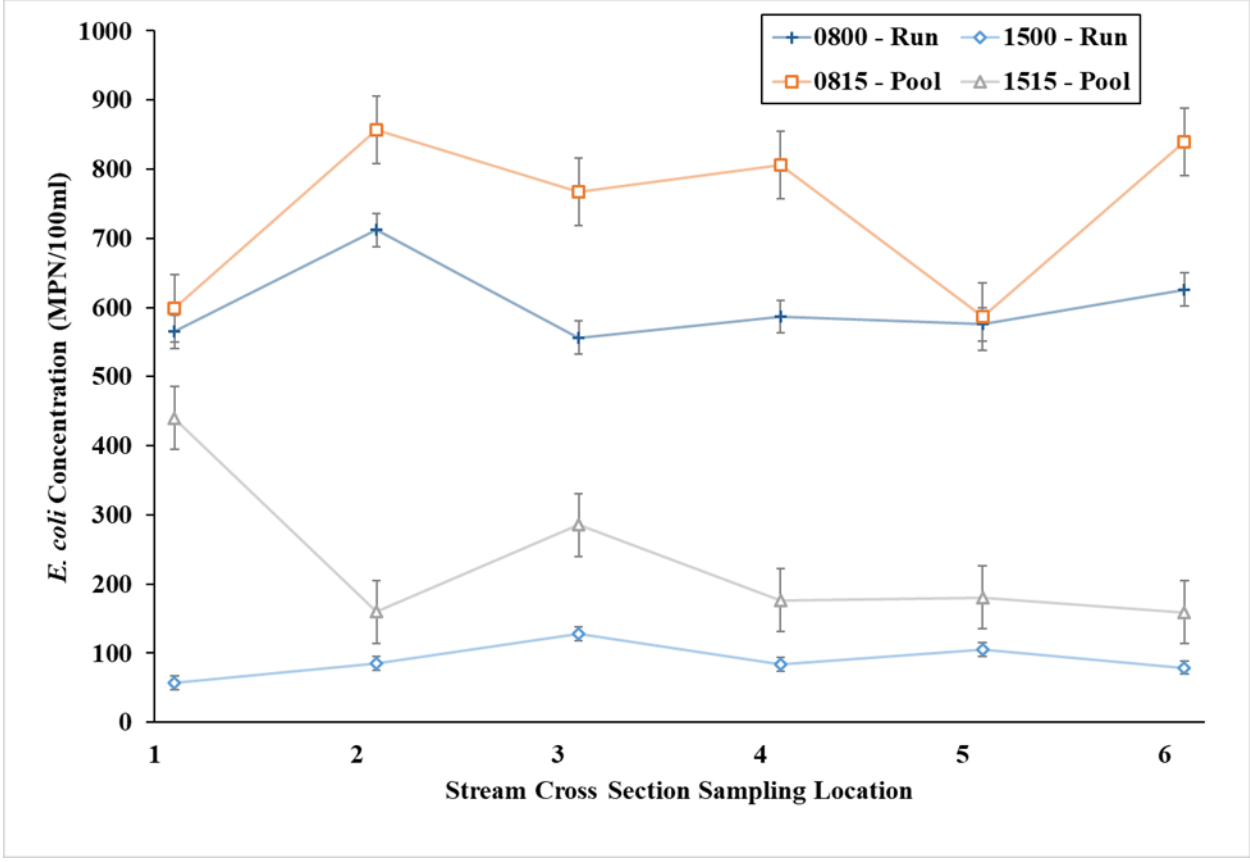


Figure 2-5. *Escherichia coli* concentrations at two cross section locations (pool and run) at the Spring Creek study site for two time periods, morning (0800 and 0815) and afternoon (1500 and 1515). Standard error is represented by the error bars.

While sediment is generally highly correlated to *E. coli* concentrations, variability between sample times has been shown to skew results while monitoring (Crosby et al. 2019). Results from our cross-section study comparing stream location indicates that variability of FIB concentrations, specifically *E. coli*, can be reduced if samples are collected in a well-mixed stream reach, such as from the stream run location, with consideration that variability can occur across the cross section even when hydrologic conditions and other factors are considered. Others have indicated that composite samples may be a better representation of stream water quality parameters when compared to other sample types (e.g., grab samples) (Harmel et al. 2016). In our preliminary research, water quality parameters (DO, T, and SC) were better predictors for *E. coli* than sediment, which may be related to time-of-day conditions within the stream since T can influence DO, SC, and *E. coli* concentrations. Diurnal variation and percent sunlight at each location were not measured for this study, but when comparing to previous research, this variable may be an important consideration of where and when to sample. More research is needed in various stream types, geographic locations, and spatial and temporal resolutions to validate the variability within stream cross sections and longitudinal segments.

2.4 Conclusions

Sampling FIB for water quality impairment determination is important to evaluate recreational waterbodies for potential pathogen presence that can affect human health. However, the water quality standards do not provide detailed guidance of the spatial and temporal distribution of water samples at a point of interest in a waterbody. Our research provides initial evidence that sampling methods should be investigated further to properly evaluate streams for water quality fecal indicators. We demonstrated that high spatial variability of bacteria concentrations across both stream reaches was shown regardless of time of day or other

waterbody conditions. Furthermore, basic water quality parameters (DO, T, and SC), time of day, and stream section locations may be useful predictors when selecting a representative location. This proof-of-concept study indicates that more emphasis should be placed on selecting site conditions that are representative (e.g., sampling reach) of the waterbody being sampled, with spatial and temporal considerations. Furthermore, other water quality and hydrologic factors could potentially be used to target stream reaches that are impaired and improve sampling protocols by understanding stream dynamics to obtain quality samples. Future work in this research area is needed to improve the water science community's approaches to enhance our understanding of streams and rivers and use our resources effectively.

2.5 References

Baird, R. and Bridgewater L. 2017. *Standard Methods for the Analysis of Water and Wastewater* (23rd edn.) Section 9223B and 2540D: Washington, D.C. American Public Health Association, American Public Health Association, American Water Works Association, and Water Pollution Control Federation.

Bradshaw, J. K., Snyder, B. J., Oladeinde, A., Spidle, D., Berrang, M. E., Meinersmann, R. J., Oakley, B., Sidle, R. C., Sullivan, K., & Molina, M. 2016. Characterizing relationships among fecal indicator bacteria, microbial source tracking markers, and associated waterborne pathogen occurrence in stream water and sediments in a mixed land use watershed. Water Research, 101, 498–509. https://doi.org/10.1016/j.watres.2016.05.014

Brinkmeyer, R., Amon, R. M. W., Schwarz, J. R., Saxton, T., Roberts, D., Harrison, S., Ellis, N., Fox, J., DiGuardi, K., Hochman, M., Duan, S., Stein, R., & Elliott, C. 2015. Distribution and persistence of Escherichia coli and Enterococci in stream bed and bank sediments from two

urban streams in Houston, TX. *Science of The Total Environment*, *502*, 650–658. https://doi.org/10.1016/j.scitotenv.2014.09.071

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet: A technical overview. J. Atmos. Oceanic Technol., 12, 5-19.

Christensen, V. G., Rasmussen, P. P., & Ziegler, A. C. 2002. Real-time water quality monitoring and regression analysis to estimate nutrient and bacteria concentrations in Kansas streams. Water Science and Technology, 45(9), 205–219. https://doi.org/10.2166/wst.2002.0240

Crosby, S. C., Spiller, N. C., Tietz, K. E., Cooper, J. R., & Fraboni, P. J. 2019. Temporal and spatial variability of instream indicator bacteria (Escherichia coli) and implications for water quality monitoring. *Environmental Monitoring and Assessment*, *191*(12), 745. https://doi.org/10.1007/s10661-019-7930-1

Desai, A. M., & Rifai, H. S. 2010. Variability of *Escherichia coli* Concentrations in an Urban Watershed in Texas. *Journal of Environmental Engineering*, *136*(12), 1347–1359. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000290

Desai, A. M., & Rifai, H. S. 2013. Escherichia coli Concentrations in Urban Watersheds Exhibit Diurnal Sag: Implications for Water-Quality Monitoring and Assessment. JAWRA Journal of the American Water Resources Association, 49(4), 766–779. https://doi.org/10.1111/jawr.12039

Devane, M. L., Moriarty, E., Weaver, L., Cookson, A., & Gilpin, B. 2020. Fecal indicator bacteria from environmental sources; strategies for identification to improve water quality monitoring. *Water Research*, *185*, 116204. https://doi.org/10.1016/j.watres.2020.116204

Dwivedi, D., Mohanty, B. P., & Lesikar, B. J. 2013. Estimating Escherichia coli loads in streams based on various physical, chemical, and biological factors. *Water Resources Research*, *49*(5), 2896–2906. https://doi.org/10.1002/wrcr.20265

Gitter, A., Mena, K. D., Wagner, K. L., Boellstorff, D. E., Borel, K. E., Gregory, L. F., Gentry, T. J., & Karthikeyan, R. 2020. Human Health Risks Associated with Recreational Waters: Preliminary Approach of Integrating Quantitative Microbial Risk Assessment with Microbial Source Tracking. *Water*, *12*(2), 327. https://doi.org/10.3390/w12020327

Gregory, L.F., Glitter A., Muela S., Wagner K.L. 2019. Should Contact Recreation Water Quality Standards be Consistent across Hydrological Extremes? Journal of Contemporary Water Research & Education. 166, 12-23, April 2019.

Haller, L., Amedegnato, E., Poté, J., & Wildi, W. (2009). Influence of Freshwater Sediment Characteristics on Persistence of Fecal Indicator Bacteria. Water, Air, and Soil Pollution, 203(1), 217–227. https://doi.org/10.1007/s11270-009-0005-0

Hamilton, J. L., & Luffman, I. 2009. Precipitation, Pathogens, and Turbidity Trends in the Little River, Tennessee. Physical Geography, 30(3), 236–248. https://doi.org/10.2747/0272-3646.30.3.236

Harmel, R. D., Hathaway, J. M., Wagner, K. L., Wolfe, J. E., Karthikeyan, R., Francesconi, W., & McCarthy, D. T. 2016. Uncertainty in monitoring E. coli concentrations in streams and stormwater runoff. *Journal of Hydrology*, *534*, 524–533. https://doi.org/10.1016/j.jhydrol.2016.01.040

Jamieson, R. C., Joy, D. M., Lee, H., Kostaschuk, R., & Gordon, R. J. 2005. Resuspension of Sediment-Associated Escherichia coli in a Natural Stream. Journal of Environmental Quality, 34(2), 581–589. https://doi.org/10.2134/jeq2005.0581

McKergow, L. A., & Davies-Colley, R. J. 2010. Stormflow dynamics and loads of Escherichia coli in a large mixed land use catchment. *Hydrological Processes*, *24*(3), 276–289. https://doi.org/10.1002/hyp.7480

McPherson, R. A., C. Fiebrich, K. C. Crawford, R. L. Elliott, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, K. A. Kloesel, S. J. Stadler, A. D. Melvin, A.J. Sutherland, and H. Shrivastava. 2007. Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. J. Atmos. Oceanic Technol., 24, 301–321.

Muirhead, R. W., & Meenken, E. D. 2018. Variability of Escherichia coli Concentrations in Rivers during Base-Flow Conditions in New Zealand. *Journal of Environmental Quality*, *47*(5), 967–973. https://doi.org/10.2134/jeq2017.11.0458

Oklahoma Water Resources Board (OWRB). 2017. Oklahoma Water Quality Standards. Oklahoma Administrative Code 785:45. Available at: https://www.owrb.ok.gov/rules/pdf/current/Ch45.pdf. Accessed February 21, 2023.

Petersen, F., & Hubbart, J. A. 2020. Physical Factors Impacting the Survival and Occurrence of Escherichia coli in Secondary Habitats. *Water*, *12*(6), 1796. https://doi.org/10.3390/w12061796

Rochelle-Newall, E. J., Ribolzi, O., Viguier, M., Thammahacksa, C., Silvera, N., Latsachack, K., Dinh, R. P., Naporn, P., Sy, H. T., Soulileuth, B., Hmaimum, N., Sisouvanh, P., Robain, H.,

Janeau, J.-L., Valentin, C., Boithias, L., & Pierret, A. 2016. Effect of land use and hydrological processes on Escherichia coli concentrations in streams of tropical, humid headwater catchments. *Scientific Reports*, *6*(1), 32974. https://doi.org/10.1038/srep32974

Sanders, E. C., Yuan, Y., & Pitchford, A. 2013. Fecal Coliform and E. coli Concentrations in Effluent-Dominated Streams of the Upper Santa Cruz Watershed. *Water*, *5*(1), 243–261. https://doi.org/10.3390/w5010243

Schultz-Fademrecht, C., Wichern, M., & Horn, H. 2008. The impact of sunlight on inactivation of indicator microorganisms both in river water and benthic biofilms. *Water Research*, *42*(19), 4771–4779. https://doi.org/10.1016/j.watres.2008.08.022

Smith, S. J., & Esralew, R. A. 2009. StreamStats in Oklahoma– Drainage-basin characteristics and peak-flow frequency statistics for ungaged streams (Report No. 2009–5255; Scientific Investigations Report). USGS Publications Warehouse. https://doi.org/10.3133/sir20095255

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. Available online at http://websoilsurvey.nrcs.usda.gov/ accessed [2023/02/01].

Stocker, M. D., Penrose, M., & Pachepsky, Y. A. 2018. Spatial Patterns of Escherichia coli Concentrations in Sediment before and after High-Flow Events in a First-Order Creek. *Journal of Environmental Quality*, *47*(5), 958–966. https://doi.org/10.2134/jeq2017.11.0451

Stocker, M. D., Rodriguez-Valentín, J. G., Pachepsky, Y. A., & Shelton, D. R. 2016. Spatial and temporal variation of fecal indicator organisms in two creeks in Beltsville, Maryland. *Water Quality Research Journal*, *51*(2), 167–179. https://doi.org/10.2166/wqrjc.2016.044

Stocker, M. D., Smith, J. E., Hernandez, C., Macarisin, D., & Pachepsky, Y. 2019. Seasonality of E. coli and Enterococci Concentrations in Creek Water, Sediment, and Periphyton. *Water, Air, & Soil Pollution*, *230*(9), 223. https://doi.org/10.1007/s11270-019-4263-1

USEPA (United States Environmental Protection Agency). 1986. Bacteriological Water Quality Criteria for Marine and Fresh Recreational Waters. EPA- 440/5-84-002. US Environmental Protection Agency, Office of Water Regulations and Standards, Cincinnati, OH, USA.

USEPA (United States Environmental Protection Agency). 2010. Sampling and Consideration of Variability (Temporal and Spatial) for Monitoring of Recreational Waters. EPA-823-R-10-005 US Environmental Protection Agency, Washington, DC, USA.

USEPA (United States Environmental Protection Agency). (2012) Recreational Water Quality Criteria. Office of Water 820-F-12-058. https://www.epa.gov/sites/production/files/2015-10/documents/rwqc2012.pdf

USGS (U.S. Geological Survey). 2014. National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1-A10, available online at http://pubs.water.usgs.gov/twri9A.

Verhougstraete, M. P., Martin, S. L., Kendall, A. D., Hyndman, D. W., & Rose, J. B. 2015. Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. *Proceedings of the National Academy of Sciences*, *112*(33), 10419–10424. https://doi.org/10.1073/pnas.1415836112

Vogel, J.R., Frankforter, J.D., Rus, D.L., Hobza, C.M. and Moser, M.T., 2009. *Water quality of combined sewer overflows, stormwater, and streams, Omaha, Nebraska, 2006-07*. U. S. Geological Survey.

Whitman, R. L., Nevers, M. B., Korinek, G. C., & Byappanahalli, M. N. 2004. Solar and Temporal Effects on Escherichia coli Concentration at a Lake Michigan Swimming Beach. Applied and Environmental Microbiology, 70(7), 4276–4285. https://doi.org/10.1128/AEM.70.7.4276-4285.200

Chapter 3 : Investigation of environmental factors on *Enterococcus* survival in Oklahoma streams

*This chapter is a **published** journal article in Aquatic Sciences that has been formatted for the dissertation and reproduced with permission from Springer Nature. The full citation is as follows:*

Graves, G. M., Vogel, J. R., & Tanner, R. S. (2023). Investigation of environmental factors on *Enterococcus* survival in Oklahoma streams. Aquatic Sciences, 85(2), 34. https://doi.org/10.1007/s00027-023-00933-3

Abstract

In this study, we assessed six Oklahoma streams for *Enterococcus* sediment and water concentrations along with water quality, sediment, hydrologic and geographical factors. We also conducted a microcosm experiment from two stream sediments to evaluate *Enterococcus* survivability under stable laboratory conditions. Stream sites exhibited common relationships between *Enterococcus* and other environmental factors, including significant correlations to antecedent dry period, *Escherichia coli*, impervious area, dissolved oxygen, and turbidity. These correlations were found for *Enterococcus* in both water and sediment. Specifically for *Enterococcus* in sediment, concentrations were also significantly correlated to turbidity and sediment percent organic matter, but not to hydrological conditions. Conversely, concentrations of *Enterococcus* in water exhibited significant moderate correlations to precipitation, antecedent dry period, drainage area, impervious area, and discharge, as well as streambed particle size. High variability between geographical attributes and stream conditions increased uncertainties and relationships between *Enterococcus* concentrations in the stream among most factors. However, when grouping sites by similar watershed and sediment characteristics, strong significant relationships for water quality parameters and *Enterococcus* concentrations in water

and sediment were observed. The microcosm study indicated that sediment *Enterococcus* concentrations for two streams with contrasting sediment properties were stable, except for a considerable increase between day 0 to day 1, with no decay shown for a 31-day period. Collectively, our field and laboratory results revealed that *Enterococcus* can survive for extended periods under both dynamic and stable sediment and water conditions, and that environmental factors can be used to characterize freshwater streams and rivers for Enterococcus concentrations in freshwater streams and rivers.

Keywords: freshwater; stream; *Enterococcus*; indicator; impairment; water quality

3.1 Introduction

Pathogens from environmental and anthropogenic sources have the potential to degrade water quality below that required for beneficial use of streams and rivers (Holcomb and Stewart 2020). Fecal indicator bacteria (FIB) — *Escherichia coli* and *Enterococcus*— are commonly used as a measure to determine potential fecal contamination in freshwater streams and rivers for the United States Environmental Protection Agency (USEPA) 303(d) impairment determination and beneficial uses (OWRB 2017; USEPA 1986; USEPA 2012). FIB concentrations used to determine human-health risk, PBCR, were established by the USEPA through a series of studies of marine and freshwater beaches in the 1980's (USEPA 2012). Water thresholds for impairment are related to number of gastrointestinal illnesses versus FIB concentrations (UESPA 1986, 2012). From these previous studies, *Enterococcus* has been assumed and established as a quality indicator of human-health risk for all recreational waters, including streams and rivers. *E. coli* is well documented in literature as a quality indicator bacterium in freshwater for predicting human health hazards and fecal contamination in both lentic and lotic freshwater bodies (Odonkor and

Ampofi 2013). Conversely, limited information is available to understand the dynamics of *Enterococcus* populations in freshwater lotic waterbodies.

There is evidence to suggest from related research, including recent USEPA work, that *Enterococcus* in freshwater bodies from environmental and animal sources may not be the best indicator for human-health risk (Cloutier and McLellan 2017; USEPA 2010). Previous studies have deducted that due to the number of non-human *Enterococcus* sources in the environment such as animal feces, soils, plants, and decaying matter that *Enterococcus* can replicate and survive outside of enteric environments (Boehm and Sassoubre 2014; Byappanahalli et al. 2012; Devane et al. 2020). *Enterococcus* is often used as a primary stable indicator in brackish or saline waters, whereas *E. coli* is often considered a more sensitive indicator in freshwater environments for fecal contamination (Jin et al. 2004). Recent research indicates that temporal and spatial factors due to climate change, seasonality and environmental conditions may impact how we currently assess waterbodies for fecal contamination, specifically *Enterococcus* as an indicator in freshwater (Petersen and Hubbart 2020; Teixeira et al. 2020). Furthermore, these factors may have an impact on FIB concentrations in streams and rivers as bacteria colonies have been shown as dynamic and in constant flux between the sediment and stream column (Litton et al. 2010; Stocker et al. 2016).

Bed sediments in streams are known to be stable reservoirs for the persistence and proliferation of *Enterococcus* in the environment which have the potential to be reintroduced into the water column through flow changes and bed disturbances within the stream (Bradshaw et al. 2016; Brinkmeyer et al. 2015; Haller et al. 2009). A study by Stocker et al. (2019) indicated that FIB can display persistence in periphyton and can contribute as a source of *Enterococcus* in sediment and the water column. Sediment and submerged aquatic vegetation may provide a

reservoir of *Enterococcus* populations that do not correspond with external contaminant sources (Badgley et al. 2010). Therefore, additional research is needed to quantify environmental factors that may play roles in *Enterococcus* survivability.

Our research team is not currently aware of similar research that has been completed in for *Enterococcus* in freshwater streams in Oklahoma or elsewhere. With more than 260,000 km of rivers and streams and 88,000 km of shoreline, Oklahoma is known for its water recreation and tourism opportunities (OWRB 2020). However, Oklahoma currently has over 12,000 kilometers of streams that are listed on the 2020 303(d) list for both *E. coli* and *Enterococcus*, which are used for primary body contact recreation indicators (PBCR) as defined in Chapter 45 of Oklahoma Water Quality Standards (ODEQ 2021; OWRB 2020). The results of this study are intended provide insight on how to approach fecal indicator bacteria analyses for beneficial use criteria and identify any factors that can be used to predict bacteria loads when developing monitoring strategies in freshwater streams and rivers. In this paper we describe results from a field and microcosm study in six Oklahoma streams to evaluate *Enterococcus* survivability in streams and potential environmental factors that have influence on their persistence in the environment.

3.2 Methods

3.2.1 Field study

Six perennial Oklahoma streams representing variable sediment types, flow conditions and ecoregions were monitored for FIB and water quality weekly for ten weeks from July-September 2021 in the stream water column and benthic substrate. Water quality parameters, sediment and water samples, and hydrologic measurements were collected to evaluate and

compare the stream reaches. Stream sampling points were selected to represent varied site conditions (e.g., urban, rural, ecoregion) and associated stream reaches were listed as impaired for both *Enterococcus* and *E. coli* in the most recent 2020 USEPA 303(d) list (ODEQ 2020). The sites monitored were located in the Upper Neosho-Grand (n=5) and Upper Canadian basins (n=1) to provide a geographic contrast to evaluate between varied stream types in Oklahoma (Fig. 3-1). Stream sediments were characterized during site reconnaissance and ranged from silty sand to medium gravel. A particle size distribution using method ASTM 6913 was performed in the laboratory to confirm the median particle size (D50) for each stream. Drainage area for each site was calculated from the pour point at the site location and delineated using U.S. Geological Survey (USGS) Stream Stats (Smith and Esralew 2010). Percent impervious area was calculated using the 2019 National Land Cover Dataset Imperviousness class from the USGS and using the Extract by Mask tool in ArcGIS 10.8 to clip the raster layer to the watershed area as delineated in USGS Stream Stats. Percent imperviousness for the purposes of this study was determined where percent imperviousness of the layer was greater than or equal to 10% of the raster grids, which is a typically used cutoff for when rivers and streams begin to erode, and sediment has the potential to be transported in the stream (Chithra et al. 2015). Sediment samples were sent to an external laboratory to determine percent organic matter for each sediment using Loss on Ignition (LOI) methods (Ball et al. 1964). Sampling visits were conducted during expected baseflow conditions to reduce potential variability of external influences (i.e., runoff), based on historical precipitation and stream flow for the sampling period of July through September in Oklahoma.

However, samples were collected weekly, regardless of precipitation or change in streamflow, which allowed for analysis of variable hydrologic conditions.

Site information collected on ten weekly occasions included hydrologic conditions, water samples, sediment samples, water quality parameters and other relevant watershed and stream information. Specifically, water quality parameters (pH, dissolved oxygen [DO], water temperature [T], specific conductance [SC], turbidity, total suspended solids [TSS]), and hydrologic parameters (stream discharge, and precipitation [nearest Oklahoma Mesonet station 24-h precipitation]) were collected along with bacteria water and sediment samples. Bacteria samples were collected based on USGS methods for collecting water samples (USGS 2014). A representative water sample (well-mixed, with adequate flow) was collected at the thalweg of the stream. Samples were stored on ice during field transport and lab storage and were processed within a 24-h hold time. Additional stream water was collected in 1-L polypropylene bottles and cooled to <6°C for total suspended solids and turbidity analysis. Antecedent dry period days were calculated from daily Oklahoma Mesonet rainfall (Brock et al. 1995; McPherson et al. 2007). Water quality parameters (pH, DO, SC, T) were measured in-situ using an Orion Star™ A325 portable multimeter (Thermo Fisher Scientific, Waltham, MA). Turbidity was measured with a Hach® 2100Q (Hach Company, Loveland, CO) portable turbidity meter. TSS was calculated by following ASTM 2540D. Stream discharge was measured using a SonTek FlowTracker2® Handheld-ADV® (SonTek YSI, San Diego, CA).

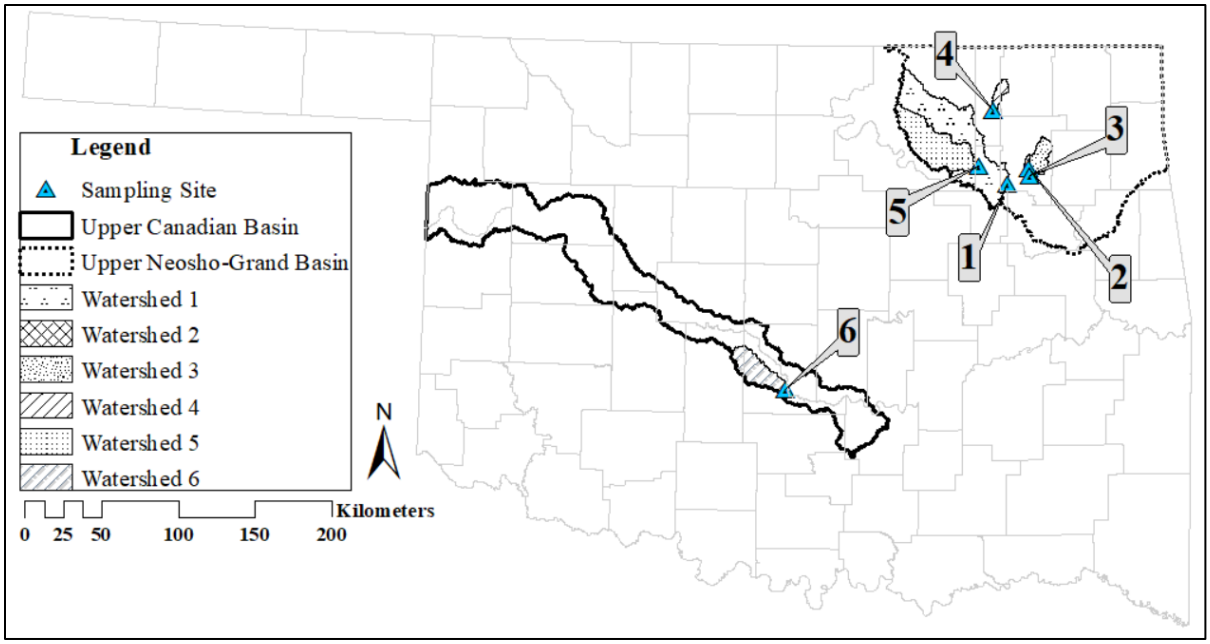


Figure 3-1. Map showing sampling site locations, delineated watershed basins of sampling sites and associated river drainage basins of Bird Creek (1), Cat Creek (2), Dog Creek (3), Hogshooter Creek (4), Hominy Creek (5) and Walnut Creek (6).

3.2.2 Microcosm study

A microcosm study was conducted from two streams, Cat and Walnut, to replicate stream conditions in a controlled environment to understand survivability of bacteria cells. Sample water and sediment from the upper 5 cm of the benthic stream substrate in the thalweg were collected in 290 ml sterile polystyrene bottles during April 2021. Stream temperatures in the water column were 17 °C and 16 °C, respectively, for Cat and Walnut during time of sediment collection. A total of 22 bottles for each site were collected with approximately 150 ml of stream water and 100 ml of sediment by volume. The microcosm trial was set up in an analogous manner as those routinely performed using soils (Schmidt and Scow 1997). The microcosm container tops (polystyrene collection bottles) were loosely covered with aluminum foil and the sides and bottoms were wrapped tight with aluminum foil to simulate a controlled, dark environment. The containers were held at a constant room temperature of 22 °C in a laboratory and placed on a horizontal orbital shaker at 100 revolutions per minute to encourage aerobic mixing as found in natural streambed conditions. For each day of analysis, two bottles from each site were removed randomly from the storage location on days 0, 1, 2, 5, 10, 17, 24 and 31 days and preserved and processed using methods described below in Section 2.3. Day 0, when the samples were collected, were processed upon arrival to the lab within 24 hours to determine initial concentrations of FIB from stream conditions. A decay rate was calculated for a 31-d period for each microcosm based on methods from Anderson et al. (2005) and Badgley et al. (2010) to understand and relate decay rates to similar studies using the following equation:

$$r = [\ln(N_t) - \ln(N_0)]/t \qquad \textit{Eq. 3-1}$$

where r = decay rate, N_t = *Enterococcus* Most Probable Number (MPN) 100 ml^{-1} at time t, N_0 = *Enterococcus* MPN 100 ml^{-1} at time zero, and t = time (31 days). The magnitude of the r value is relative where a positive value indicates positive growth, and a negative value indicates decay.

3.2.3 Microbiological analysis

Bacteria water samples were processed using the IDEXX most probable number (MPN) methods as defined under SM9223B and SM9230D for *E. coli* and *Enterococcus* concentrations, respectively (Baird et al. 2012). Sediment samples were stored in the dark at <6°C and processed within 24-h for *Enterococcus* using sodium pyrophosphate microbial detachment and soil dispersion methods modified from Ogram et al. (2007). A 2% sodium pyrophosphate solution was developed by mixing the tetrasodium pyrophosphate with sterile, reverse osmosis (RO) water, and then adjusting the pH to 7.0. Sediment samples were processed by carefully decanting the water from the top of the bottle using a sterile serological pipetter. Next, 200 ml of pyrophosphate solution was added to the bottle containing saturated sediment and the samples were dispersed by capping the bottles and manually shaking them vigorously for two minutes before they were placed on a horizontal orbital shaker at 200 revolutions per minute for 15 min. Serial dilutions were made (1:100, 1:500, 1:1000) from the sediment and pyrophosphate samples with buffered (7.0 pH) sterile RO water. Diluted samples were processed as water samples using standard methods (SM9230D) as previously described in this section.

3.2.4 Quality Control

Quality control samples, which included duplicates, field blanks and lab blanks, for field FIB enumeration were conducted at a rate of 5% for all water and sediment samples collected. Average duplicate results were ±22% (SD=15%) for all samples and no positive counts resulted

from field or laboratory blanks. All field and laboratory water quality meters were calibrated per specification standards once per week. Turbidity readings were taken five times for each sample and the median value was reported. TSS sample blanks and duplicates were performed at a rate of 5% of samples. TSS sample duplicates were within ±5% for each sample and lab blanks reported a <0.01% change in filter mass.

3.2.5 Data Analysis

Data were analyzed using Microsoft Excel® software and R Studio statistical software to evaluate statistical inferences between bacteria, water quality and hydrologic metrics. Correlations between water quality, sediment, hydrology, and bacteria parameters were analyzed using the 'stats' package within R using the correlation and Pearson functions. A Pearson correlation matrix was used to evaluate trends, specifically for parameters related to *Enterococcus* sediment and water concentrations, and a Welch's t-test was used to determine significance. Prior to the correlation analysis, data were evaluated for skewness and log-transformations were performed (Helsel et al. 2020). A Kruskal-Wallis test in R was performed to evaluate differences among means of watershed characteristics from Table 1 (n = 6) (R Core Team 2013). For bacteria concentrations, a geometric mean was used to normalize right skewed data and is often used for regulatory limit reports for primary body contact recreation for *Enterococcus* and *E. coli* (OWRB 2017). A log-linear regression (α=0.05) was used to evaluate the field sampling time-series *Enterococcus* sediment concentration data (R Core Team 2013). An antecedent dry period was calculated using a custom script in R by calculating consecutive run length of days less than 2.54-mm precipitation for the sampling days and site locations (Appendix 3). The closest Oklahoma Mesonet stations were spatially matched to the water quality sampling locations in ArcGIS.

3.3 Results and Discussion

Site information collected for watershed and soil characteristics are presented in Table 1. Drainage area (DA), impervious drainage percentage (IA), particle size (D50) and percentage organic matter (OM) were analyzed by ranking parameters. A Kruskal-Wallis test was performed, and a significant difference ($p < 0.001$) was shown between all sites for all parameters. From these results, two groups were identified for further exploration of factors based on significant differences ($p < 0.001$) in DA, D50 and particle texture class from a Kruskal-Wallis test between the groups. No significant difference ($p > 0.05$) between the groups for all sites was identified for impervious drainage and percent organic matter. However, two of the smaller sites (Location [ID] 2 and 5) were subdrainages of the larger drainages of ID 2 and 3, respectively (Table 3-1). Hogshooter and Walnut were hydrologically disconnected from the other watersheds. The two groups identified for further analysis based on similar stream characteristics were Group 1: site ID 1 (Bird), 5 (Hominy) and 6 (Walnut), and Group 2: site ID 2 (Cat), 3 (Dog) and 4 (Hogshooter).

Table 3-1. Descriptive geographical and soil data for each stream sampling site in the study. ID corresponds to the map Site ID in Fig. 3-1.

ID	Group	Stream Name	Drainage Area (km^2)	Percent Impervious	Particle Size D50 (mm)	Particle Texture Class	Percent Organic Matter
1	1	Bird	2940	9	1.78	Med-coarse sand	2.4
5	1	Hominy	1060	2	0.26	Fine silty sand	1.0
6	1	Walnut	523	5	0.38	Fine-med sand	0.4
2	2	Cat	25	5	5.08	Fine-med gravel	1.2
3	2	Dog	270	11	8.89	Fine-med gravel	0.6
4	2	Hogshooter	109	2	10.7	Medium gravel	0.1

3.3.1 Field study

Summary statistics from sediment and water samples from six creeks were monitored for ten weekly sampling events for FIB and water-quality parameters and results are presented in Table 3-2.

Table 3-2. Summary statistics of water quality, hydrologic and fecal indicator bacteria (FIB) concentrations for sampling locations during the field sampling events (n=10).

24-hr Precipitation (mm)					**Antecedent Dry Period (days)**			
Location	Mean	Min	Max	SD	Mean	Min	Max	SD
Bird	0.41	0	3.30	1.04	4	1	8	3
Cat	1.12	0	9.14	2.89	3	0	8	3
Dog	0.99	0	9.14	2.87	3	0	8	3
Hogshooter	1.83	0	18.29	5.78	5	0	23	7
Hominy	0.41	0	3.30	1.04	4	1	8	3
Walnut	0.79	0	5.84	1.89	6	0	14	5
Discharge (m^3s^{-1})					**Dissolved Oxygen ($mg\ L^{-1}$)**			
Location	Mean	Min	Max	SD	Mean	Min	Max	SD
Bird	20.95	4.45	82.97	29.28	7.13	6.32	8.34	0.63
Cat	0.07	0.01	0.39	0.12	5.3	2.34	7.14	1.38
Dog	0.68	0.01	2.45	0.98	4.77	2.51	6.87	1.70
Hogshooter	0.72	0.01	4.00	1.22	6.45	5.85	7.13	0.40
Hominy *	5.13	NA	NA	NA	7.98	7.44	8.97	0.56
Walnut	1.58	0.62	4.46	1.12	7.46	6.98	8.47	0.43
***Enterococcus* water (MPN g^{-1})**					***Enterococcus* sediment (MPN g^{-1})**			
Location	Geo-mean	Min	Max	SD	Geo-mean	Min	Max	SD
Bird	265	86	3873	1162	305	5	3951	1313
Cat	2939	448	9678	2731	472	36	2278	701
Dog	1636	241	9678	2724	266	38	1674	619
Hogshooter	1399	373	9678	2725	57	10	480	166
Hominy	589	168	1102	337	59	36	256	83
Walnut	324	98	7945	2476	18	1	980	322

Table 3-2 (Cont.)

***E. coli* water (MPN 100 ml^{-1})**					**pH**			
Location	Geo-mean	Min	Max	SD	Mean	Min	Max	SD
Bird	65	4	1935	593	7.96	7.70	8.75	0.30
Cat	367	48	2595	942	7.88	7.48	8.80	0.39
Dog	113	4	1741	529	7.62	7.15	8.76	0.50
Hogshooter	166	30	6212	1921	8.11	7.73	8.90	0.40
Hominy	70	48	100	21	8.06	7.57	9.05	0.58
Walnut	129	16	1670	670	8.36	8.19	8.51	0.10
Specific Conductance (µS cm^{-1})					**Turbidity (NTU)**			
Location	Mean	Min	Max	SD	Mean	Min	Max	SD
Bird	342.9	240.8	389.7	57.4	55.9	20.7	215.0	58.7
Cat	556.9	435.6	689.3	109.1	24.4	5.8	73.2	21.8
Dog	242.4	28.6	350.6	91.2	23.8	8.0	47.4	14.6
Hogshooter	403.0	172.0	510.5	115.3	33.4	4.5	185.0	54.1
Hominy	252.6	232.5	270.9	11.1	25.1	19.1	38.4	6.4
Walnut	740.5	532.7	807.7	86.1	47.1	10.4	227.5	65.3
Water Temperature (°C)					**Total Suspended Solids (mg L^{-1})**			
Location	Mean	Min	Max	SD	Mean	Min	Max	SD
Bird	26.2	22.0	28.9	2.5	8.2	1.0	36.0	13
Cat	25.5	22.6	26.9	1.4	0.8	0.3	2.0	0.6
Dog	26.3	23.5	28.5	1.4	1.7	0.4	5.0	1.6
Hogshooter	24.4	21.2	25.9	1.5	3.3	0.4	13.0	4.1
Hominy	21.5	18.9	22.5	1.4	6.1	1.0	15.0	5.7
Walnut	30.7	25.0	33.8	3.0	5.5	1.0	26.0	8.4

3.3.2 Enterococci in Sediment

Samples were collected over ten consecutive weeks to identify trends between weeks and understand background levels of *Enterococcus* in streambed sediments during a primary body contact recreation sampling period (Fig. 3-2). Overall, sediment *Enterococcus* concentrations were variable between sampling weeks for all locations. Only one significant relationship was determined from a log-linear regression for concentration versus time at Hominy ($p<0.001$), where a decreasing trend was shown over the sampling period. No other significant relationships ($p>0.05$) were determined for the other locations. The maximum recorded concentration for all sites was Bird on July 26 at 3951 MPN/g wet sediment, and the minimum concentration for all sites was 1.3 MPN/g wet sediment at Walnut on August 31. The geometric mean for all sites was 122 MPN/g wet sediment and geometric means ranged from 18 MPN/g wet sediment at Walnut to 472 MPN/g wet sediment at Bird.

Notably, sediment samples from the week of July 27 and August 17 showed a mean increase of 79% (SD ± 16%) in *Enterococcus* sediment concentrations. Average ADP at the nearest Oklahoma Mesonet station for those days at all sampling locations was 1.6 days for precipitation less than 2.5 mm and an average precipitation of 3.6 mm whereas on all the other sampling dates the ADP was 3.3 days and an average precipitation of 1.3 mm in the previous 24-h period. The relationship between ADP and FIB concentrations is evaluated further in Section 3.3.3. The time-series results indicate that *Enterococcus* concentrations in sediment, regardless of location or time, are consistent with previous studies showing viable streambed populations that have the potential to interact with the water column (Brinkmeyer et al. 2015).

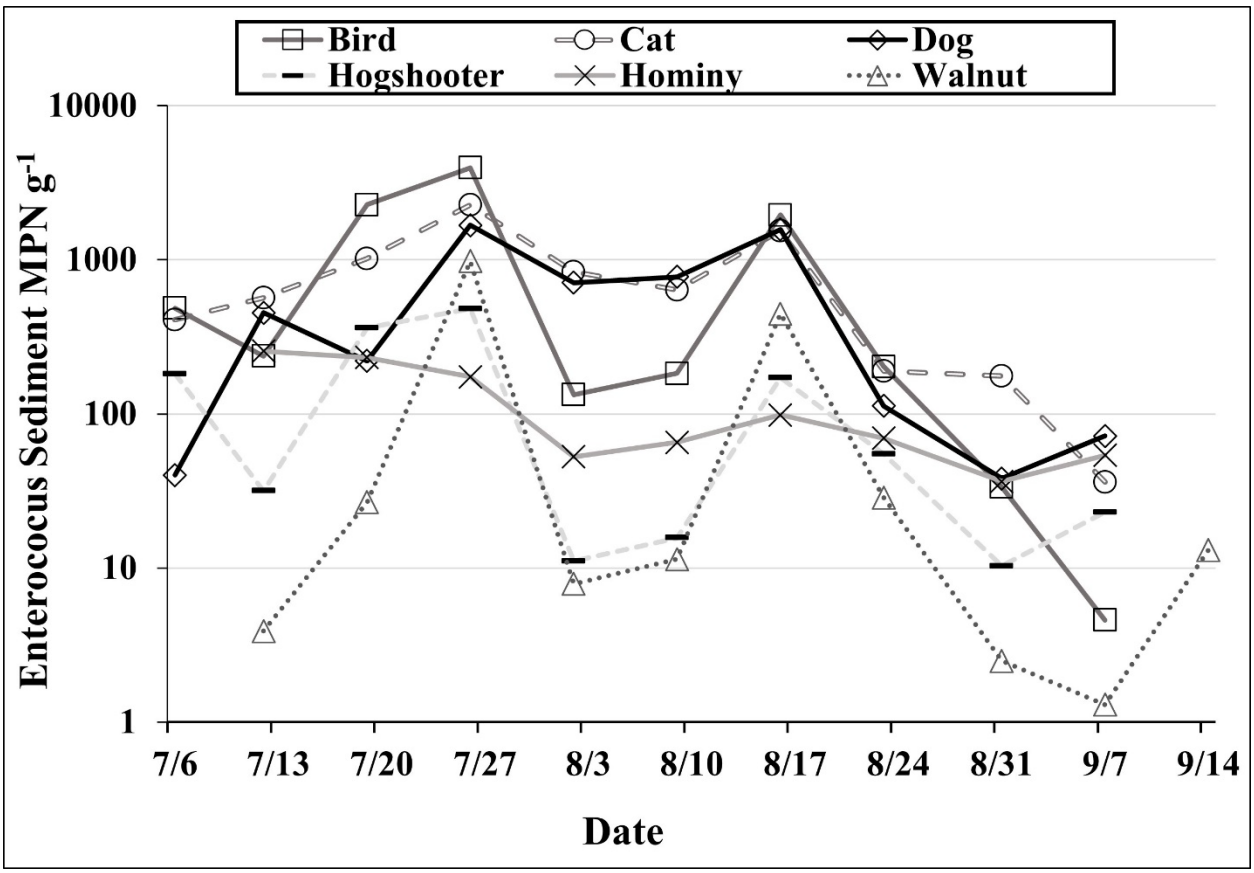


Figure 3-2. Time-series of *Enterococcus* sediment concentrations (log scale) for sampling sites during July-September for a total of 10 sampling visits at each of the six locations. MPN = Most Probable Number.

3.3.3 Environmental Factor Correlation

Summary statistics (mean, minimum, maximum and standard deviation) were calculated for the twelve stream metrics collected over a period of ten sampling visits (Table 3-2). Log transformations were performed for discharge, turbidity, total suspended solids, *E. coli* water, *Enterococcus* water, and *Enterococcus* sediment based on methods from Helsel et al. (2020) to reduce skewness. Results from the Pearson correlation matrix values ranged from a strong linear positive correlation of 1 to a strong linear negative correlation of -1 and significance was determined as $p < 0.05$. Values between ± 0.3 and 0.7 are moderately correlated, values less than ± 0.3 are weakly correlated and values greater than ± 0.7 are strongly correlated, in respect to positive or negative values. Factors were explored by each stream for all factors and by the two groups previously identified by site characteristics. A correlation matrix with significant parameters is displayed in Fig. 3-3.

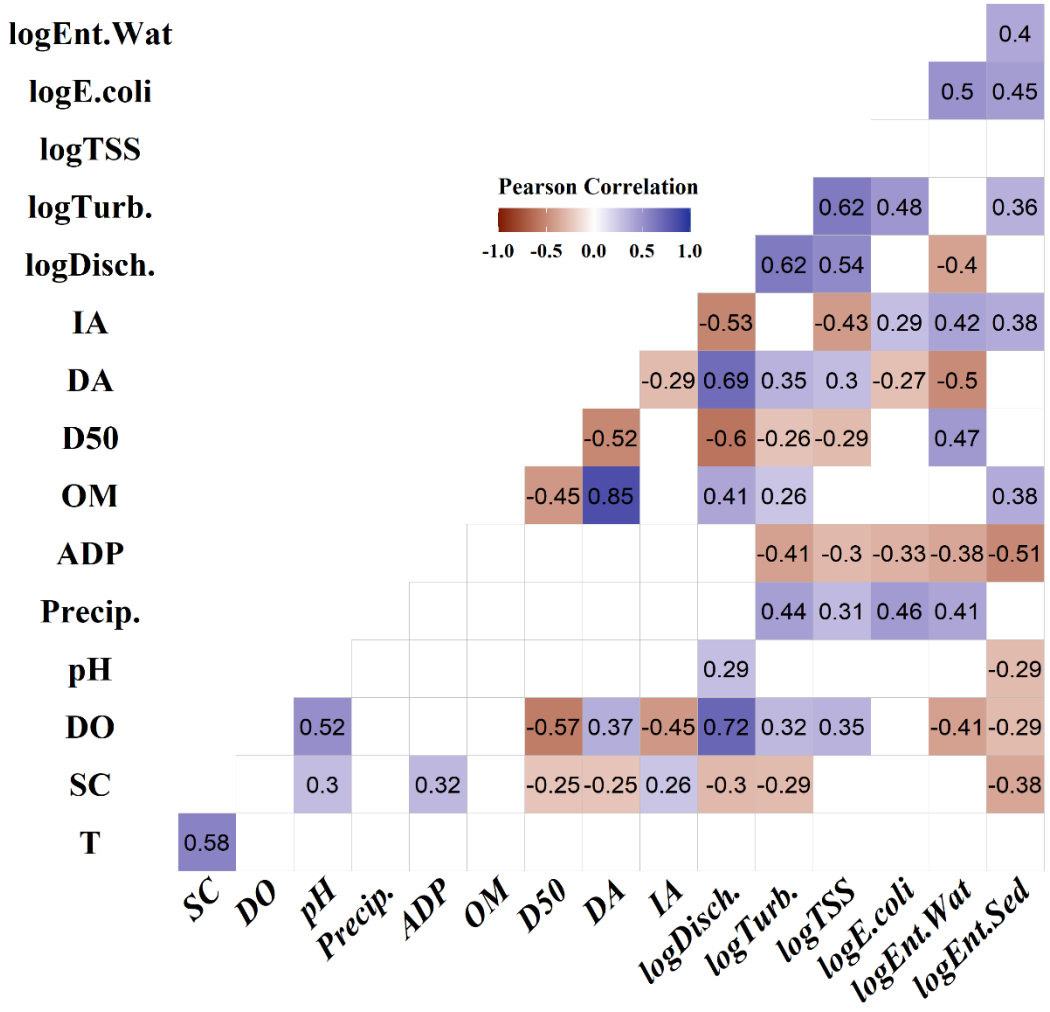


Figure 3-3. Pearson correlation matrix of all parameters measured from field water samples and associated soil, geographical and hydrologic data. The correlation matrix values range from a strong linear positive correlation of 1 indicating a very strong positive correlation to a negative correlation of -1 indicating a very strong negative linear relationship. Values and blocks shown are those that showed a significant ($p<0.05$) relationship from a t-test. Acronym explanation: T= water temperature, SC = specific conductance, DO = dissolved oxygen, ADP = antecedent dry period, OM = organic matter, D50 = median particle size, DA = drainage area, IA= impervious area. Parameters that were log-transformed are indicated by "log" before the description.

Significant parameters ($p < 0.05$) related to *Enterococcus* water concentrations included moderate positive correlations to *E. coli* water concentrations (0.5), median particle size (D50) (0.47), percent impervious area (IA) (0.42), and 24-h precipitation (0.41). Weak negative correlations were observed for *Enterococcus* water concentrations to percent drainage area (DA) (-0.5), dissolved oxygen (DO) (-0.41), discharge (-0.4) and antecedent dry period (ADP) (-0.38) . *Enterococcus* sediment concentrations displayed moderate positive correlations to *E. coli* (0.45), *Enterococcus* water (0.4), and weak positive correlations to IA (0.38), percent organic matter (OM%) (0.38), and turbidity (0.36). Negative weak correlations were shown for *Enterococcus* sediment to conductivity (-0.38) and dissolved oxygen (DO) (-0.38). Based on correlation comparisons, *Enterococcus* in the sediment does not appear to be significantly influenced by hydrology but does appear to correlate to sediment differences such as OM% and turbidity. However, *Enterococcus* water concentrations and precipitation and discharge were moderately correlated, which corresponds to previous studies showing precipitation influence on FIB water concentrations in freshwater streams from external influences (Ibekwe et al. 2011). ADP displayed a moderate negative correlation for both water and sediment *Enterococcus* concentrations, indicating that potential dry periods allow for concentrations to decrease in the water column due to reduced fluctuation in hydrology that could disrupt sediment from increased runoff. Previous research similarly found that ADP has the potential to influence FIB by creating a flushing effect in the stream sediments and can potentially be used as a predicting factor for bacteria concentrations (Christian et al. 2020; Phillips et al., 2011).

Relationships between *E. coli* and *Enterococcus* in both the sediment and water are key considerations for evaluating fecal indicators as *E. coli* is used as a primary indicator of recreational criteria and is often related to *Enterococcus* concentrations (Stocker et al. 2019). We

found that *E. coli* was not significantly related to sediment (TSS, OM, D50) or other water quality parameters except for turbidity (0.48). However, *E. coli* was moderately correlated to sediment and water *Enterococcus* concentrations, which corresponds to results in previous studies of freshwater streams where sediment was found to be a significant contributor to *E. coli* and *Enterococcus* in the water column (Brinkmeyer et al. 2015). Similarly, others have indicated relationships between watershed characteristics (i.e., percent imperviousness and watershed area) and *E. coli.* which corresponds with our correlation results for both fecal indicator bacteria (Chen and Chang 2014). Correlation of these two fecal indicators is important because there is evidence that suggests water quality monitoring for human health can be impacted by naturalized bacteria that are potential reservoirs and sources of contamination in freshwaters (Devane et al. 2020).

Hydrologic characteristics within watersheds and stream reaches are often used to evaluate water quality trends for abiotic and biotic factors (Bojarczuk et al. 2018; Economy et al. 2019). Discharge showed a significant moderate negative correlation (-0.4) for *Enterococcus* water concentrations and no correlation with *Enterococcus* concentrations, potentially indicating that small fluctuations in flow are not as representative for evaluating *Enterococcus* concentrations without additional water quality parameters measured such as turbidity and TSS. Discharges for this study were magnitudes smaller than what would typically occur during the spring or fall precipitation events and discharge could play a more important role for correlating FIB concentrations during high flow conditions (Garbrecht et al. 2004). Additionally, smaller order streams and mixed land use, as most of the streams in this study represent, potentially have higher hydrologic variability and influence from precipitation events than higher order streams on FIB concentrations (Dila et al. 2018; Zhang et al. 2020). Furthermore, drainage area showed a moderate negative correlation to *Enterococcus* water correlations, which could be due to

potential dilution from precipitation and other watershed inputs (i.e., mixed land use) that could influence *Enterococcus* concentrations in the stream column (Islam et al. 2017). No significant relationship for *E. coli*, *Enterococcus* sediment or water and TSS was determined for this study, which may be indicative of the distribution of streambed particle sizes and external influences of suspended particles. Other studies have shown that turbidity is often a stronger predictor of fecal indicators, and that the particle-bound *Enterococcus* relationship is not well-understood (Suter et al. 2011).

Based on Group 1 and Group 2 identified earlier in Section 3.3, factors between each group were explored to determine if watershed and water quality characteristics potentially influenced *Enterococcus* concentrations in streams. A log-log regression was performed to evaluate prediction between *Enterococcus* sediment and water. Results showed that Group 1 had an R^2 of 0.003 and Group 2 had an R^2 of 0.51 and a significant difference ($p<0.05$) was shown between groups for *Enterococcus* water and sediment from a Welch's t-test (Fig. 3-4).

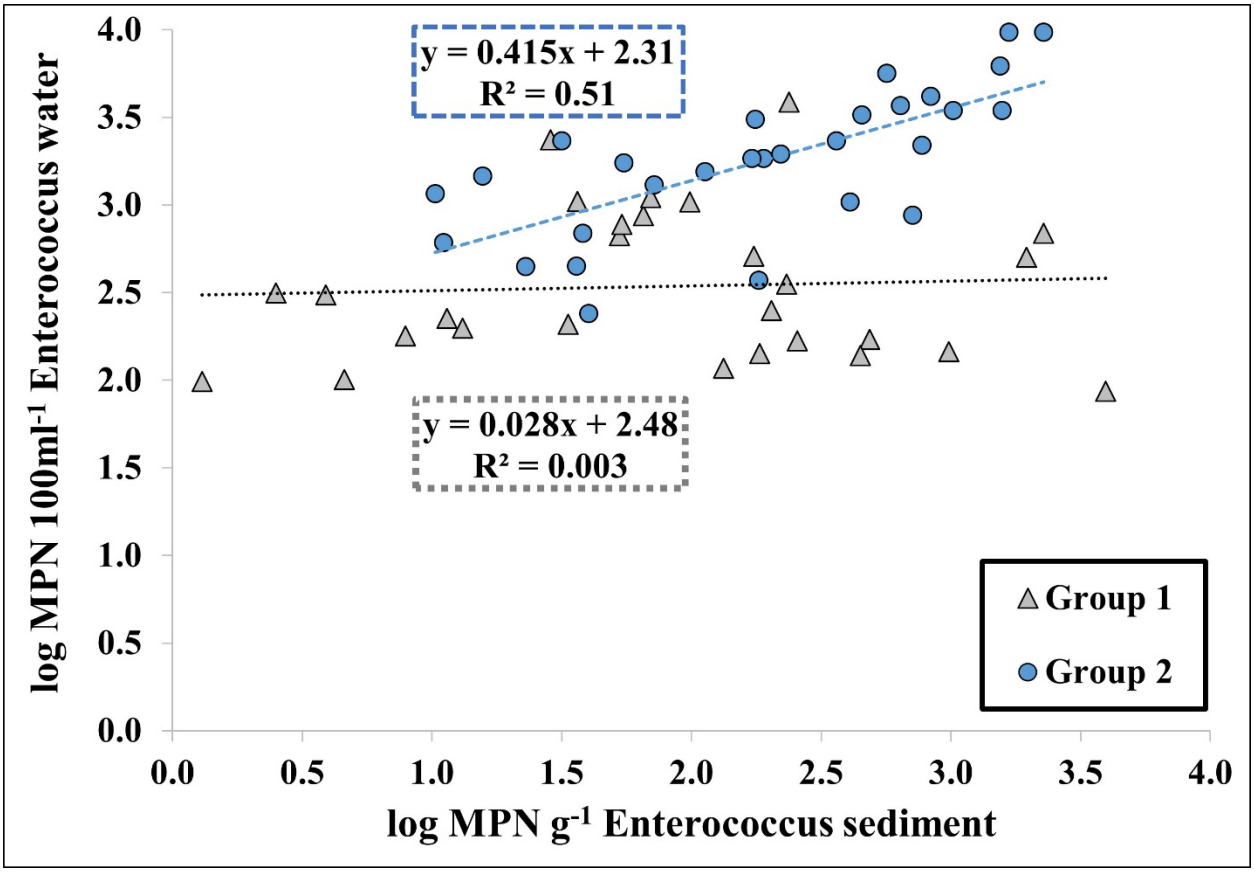


Figure 3-4. Scatter plot of log *Enterococcus* water versus log *Enterococcus* sediment concentrations from water and sediment samples collected during the field study over a period of ten sampling events for Group 1: site ID 1 (Bird), 5 (Hominy), and 6 (Walnut), and Group 2: site ID 2 (Cat), 3 (Dog), and 4 (Hogshooter). Linear trendlines are displayed and next to each line are associated equations and R^2 values. MPN = Most Probable Number.

Stream characteristics between Group 1 and Group 2 were evaluated to determine if relationships existed between *Enterococcus* water and sediment concentrations. Significant differences between means ($p<0.05$) resulted for DO, pH, discharge, *Enterococcus* water, OM, TSS, D50, DA, IA, turbidity, and *E. coli* water. Conversely, no statistical differences ($p>0.05$) were found for *Enterococcus* sediment, temperature, conductivity, precipitation, and ADP.

Relationships within each group were then compared to determine where relationships exist, if any, and if any parameters correlated to *Enterococcus* sediment and water concentrations. Results indicate that within Group 1, discharge (0.56), turbidity (0.56), OM (0.54), DA (0.53), and D50 (0.45) showed significant moderate correlations within *Enterococcus* sediment. *Escherichia coli* (0.43) and turbidity (0.4) resulted in a moderate positive correlation with *Enterococcus* water concentration, and conductivity (SC) (-0.52) and ADP (-0.54) were moderately negative correlated with *Enterococcus* sediment. Within Group 2, however, *Enterococcus* sediment indicated significant positive moderate correlations for OM (0.54), *E. coli* (0.48), IA (0.46), water temperature (0.46), and a negative moderate correlation for D50 (-0.5) and ADP (-0.48), while *Enterococcus* water concentrations showed positive moderate correlations between precipitation (0.55), *E. coli* water (0.48), turbidity, (0.5), TSS (0.48)), water temperature (0.41) and a negative moderate correlation for ADP (-0.52). The relationship within Group 2 for *Enterococcus* water and *Enterococcus* sediment had a significant strong positive correlation (0.71), whereas no significant relationship was found in Group 1. The similarities between groups for both *Enterococcus* sediment and water concentrations were turbidity and *E. coli*, and ADP and OM for *Enterococcus* sediment.

Previous research has indicated that sediment and particles are related to persistence of *Enterococcus* in the water column, and our results similarly demonstrate that smaller drainages

may be easier to predict the concentrations of *Enterococcus* from hydrological (precipitation and discharge) and sediment characteristics (turbidity, TSS, D50) (Myers and Juhl 2020). From the sediment characteristics, TSS, turbidity, OM and D50 were significantly different between each group. Brinkmeyer et al. (2015) found that most *Enterococcus* in the water column were correlated with suspended sediment from silt to fine sand grains, and Haller et al. (2009) showed that smaller particles have the potential to resuspend FIB and have higher interaction with the water column. Between the two groups, Group 2 had a larger mean size substrate (gravel size particles), lower OM, lower discharge, and lower mean turbidity and TSS, which resulted in reduced variability between *Enterococcus* concentrations in the sediment and water. In larger streams and river drainages, where suspended sediment and higher OM from higher turbulent discharge is possible, such as found in Group 2 in this study, *Enterococcus* concentrations in sediment and water may be highly variable due to the continuous interaction between the streambed and water column (Grant et al. 2011).

Differences between sites were highly variable for all site characteristics, hydrology, and water-quality parameters versus *Enterococcus* concentrations in sediment and water. Conversely, similarities existed between groupings of sites (i.e., drainage area) and could provide insight when selecting sites for monitoring and evaluation of water quality and impairment for *Enterococcus*. Similar research has revealed that grouping watershed and stream characteristics can be important when developing spatial and temporal monitoring studies (Piorkowski et al. 2014; Stocker et al. 2016). We found that when evaluating streams for *Enterococcus* concentrations, hydrologic and geologic factors such as discharge, sediment (OM, turbidity, TSS, D50), antecedent dry period, and drainage and impervious area may be influential on where to monitor streams and expected relationships of environmental factors. Furthermore, water-quality

parameters (T, DO and pH) were shown in this study to be significantly different depending on the watershed and could be important considerations when evaluating *Enterococcus* levels within in stream. The variability of these characteristics was shown to increase uncertainty of predictors for determining *Enterococcus* concentrations, regardless of stream conditions (i.e., water quality parameters).

3.3.4 Microcosm study

A microcosm study for two streams, Cat, and Walnut, was performed for a period of 31-d to investigate the sediment *Enterococcus* concentrations. The geometric means for *Enterococcus* concentrations were 729 MPN/g and 7 MPN/g wet sediment (n=8), for Cat and Walnut, respectively. Maximum and minimum values were 2,055 MPN/g and 439 MPN/g wet sediment for Cat, and 17 MPN/g and 4 MPN/g wet sediment for Walnut. A time series plot for both microcosms is presented in Fig. 3-5. A Pearson correlation with a paired t-test was performed to determine correlation between Cat and Walnut *Enterococcus* sediment concentrations and results indicated a significant ($p<0.01$) strong positive correlation (0.76). For both microcosms, the *Enterococcus* concentrations increased between Day 0 and Day 1 before exponentially declining to stabilization around Day 10 to Day 17. The decay rates for Cat and Walnut were calculated for Day 0 and Day 31 from Eq. 1 and resulted in r=-0.032 and r=0.001, respectively. From these values, no discernable difference in concentrations was shown for the study period with a slight decay in Cat and neutral growth for Walnut.

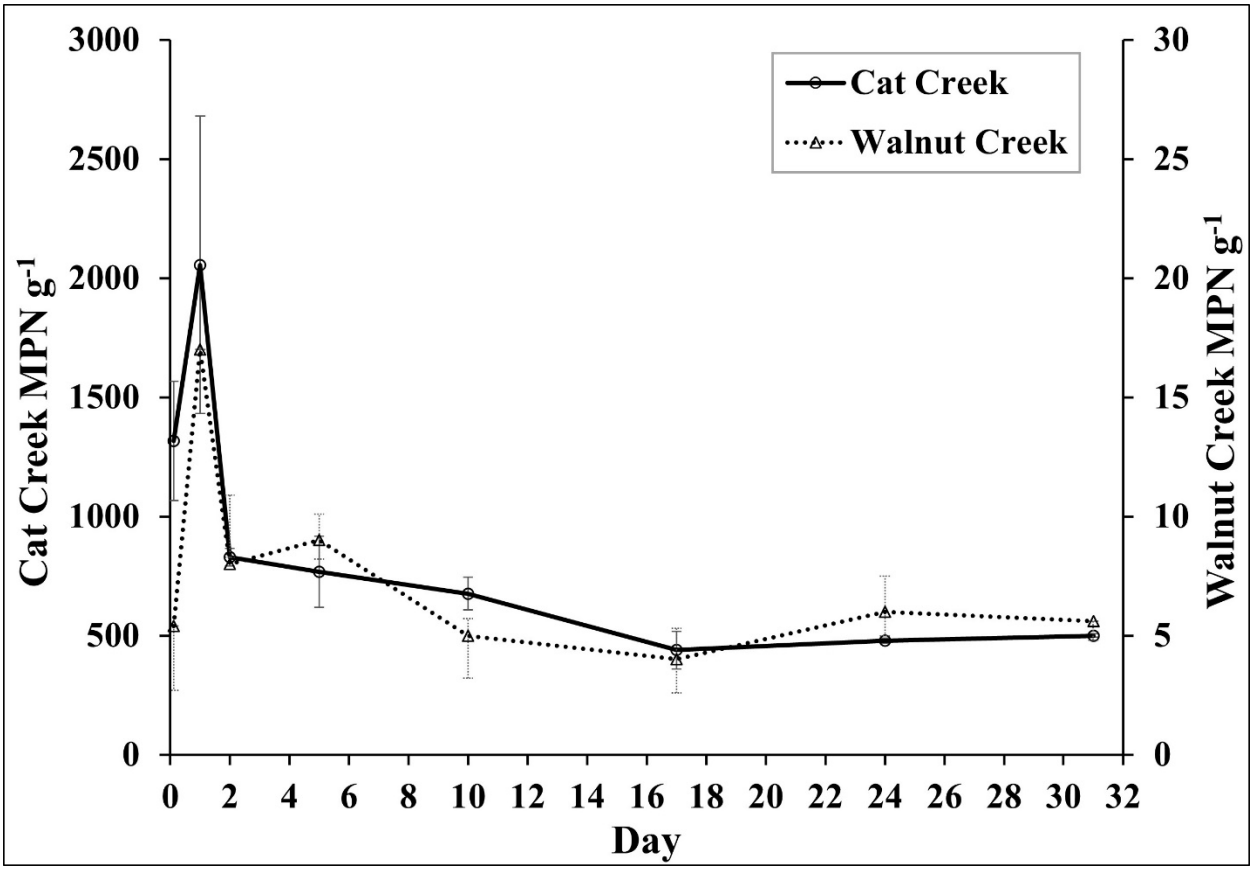


Figure 3-5. Time-series plot of microcosm *Enterococcus* sediment concentrations over a period of 31 days from Cat Creek and Walnut Creek. Standard error is represented for each time-series by error bars. MPN = Most Probable Number.

Throughout the 31-d study period, both microcosms with contrasting substrate types, organic matter and *Enterococcus* concentrations showed persistence in *Enterococcus* viability under stable conditions. A comparable study to our results by Kim and Wuertz (2015) indicated a rapid ten-fold increase in *Enterococcus* counts for the initial two to three days followed by a gradual decay and stabilization in numbers over a 40-d period. Similarly, decay rates and survival of *Enterococcus* in a microcosm study by Haller et al. (2009) were shown for a period of 50-d, whereas *E. coli* and total coliforms appeared to decrease to non-detectable concentrations. Other related research conducted in mesocosm studies have shown that *Enterococcus* and FIB decay was significantly reduced in sediment and organic matter (Tiwari 2019). Furthermore, aquatic vegetation such as periphyton has been shown to play a key role in *Enterococcus* survivability and growth (Stocker et al. 2019). Sediment for our study was collected from the upper benthic substrate and had the potential for inclusion of periphyton and biofilm that accumulated from the natural stream conditions.

Since our microcosm experiments were under no-light conditions, the *Enterococcus* colonies may have experienced the rapid growth shown early in the time period due to abundance of organic and plant matter before metabolizing the available nutrients resulting in a decay and stabilization in concentrations and is analogous to conclusions from studies by Kim and Wuertz (2015) and Zimmer-Faust et al. (2017). *Enterococcus* has been found to persist in many different environments (e.g., soil and plant matter) regardless of external inputs such as fecal contamination (Staley et al. 2014). Additionally, organic carbon and nutrient inputs have been found to stimulate growth of FIB in stream sediments and enhance population stability (Korajkic et al. 2019). Many of the streams in the study region have dense canopy cover, high

potential for nutrient and organic carbon inputs and mobile substrates, which could enhance *Enterococcus* survivability.

Our microcosm experiment along with previous experiments provide convincing evidence that benthic streambed sediments under stable conditions have the potential to be reservoirs and sources of *Enterococcus*. Given the dynamic nature of mobile streambeds, resuspension of these sediment-laden *Enterococcus* could increase the potential for these fecal indicators to persist within the stream water column for extended periods without external inputs (e.g., stormwater runoff) as evident in studies relating *E. coli* and streambed sediments (Garzio-Hadzick et al. 2010; Stephenson et al. 1982). Therefore, sediment sources could create interferences with accurately assessing human-health risk and stream impairment criteria. More research is needed to understand the in-situ relationship of streambed sediment influences on *Enterococcus* concentrations in freshwater streams.

3.4 Conclusions

Enterococcus is often used to determine recreational water-quality for the purposes of limiting or preventing potential gastrointestinal illness. However, questions remain on the validity of using *Enterococcus* to make regulatory decisions given the potential for persistence in the environment without external inputs of fecal sources. Additionally, limited information exists on the relationships between water quality, geography, stream substrate properties, and hydrologic conditions that have the potential to influence *Enterococcus* concentrations in the stream water column. Our study aimed to understand stream dynamics in the field and laboratory to assess potential persistence in the environment and relate stream factors to *Enterococcus* concentrations. Results indicate, in general, that hydrologic conditions, watershed area, sediment properties and multiple water quality parameters are correlated to *Enterococcus* concentrations in

the water column and sediments. Furthermore, relationships between sediment and water *Enterococcus* sediment existed when grouping sites by geographical and sediment characteristics. The microcosm *Enterococcus* sediment study corresponded with the field study in that concentrations remained stable throughout the study period except for during the first day after the start of the trial. Conditions as in the first days of the microcosm study could also occur in the streambed where variations in sediment *Enterococcus* concentrations may exist due to external inputs (e.g., rainfall runoff) and mobile beds. Implications from this work emphasize that more research is needed to evaluate *Enterococcus* as a regulatory indicator, given counts have the potential to remain viable in recreational freshwater streams and are often ubiquitous in concentrations above the regulatory thresholds for a majority of the recreational season. This study indicates that monitoring plans should consider environmental factors as influencers on *Enterococcus* concentrations within freshwater streams.

3.5 References

Anderson, K. L., Whitlock, J. E., and Harwood, V. J. (2005) Persistence and Differential Survival of Fecal Indicator Bacteria in Subtropical Waters and Sediments. *Applied and Environmental Microbiology*, *71*(6), 3041–3048. https://doi.org/10.1128/AEM.71.6.3041-3048.2005

Badgley, B. D., Nayak, B. S., and Harwood, V. J. (2010) The importance of sediment and submerged aquatic vegetation as potential habitats for persistent strains of enterococci in a subtropical watershed. *Water Research*, *44*(20), 5857–5866. https://doi.org/10.1016/j.watres.2010.07.005

Baird, Rodger, et al. (2012) *Standard Methods for the Analysis of Water and Wastewater* (22nd edn.) Section 9223B and 2540D: Washington, D.C. American Public Health Association, American Public Health Association, American Water Works Association, and Water Pollution Control Federation.

Ball, D. F., (1964) Loss-on-ignition as an estimate of organic matter and organic carbon in non-calcareous soils. J. Soil Sci. 15: 84–92.

Boehm, A. B., and Sassoubre, L. M. (2014) Enterococci as Indicators of Environmental Fecal Contamination. In *Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet]*. Massachusetts Eye and Ear Infirmary. https://www.ncbi.nlm.nih.gov/sites/books/NBK190421/

Bojarczuk, A., Jelonkiewicz, Ł., and Lenart-Boroń, A. (2018) The effect of anthropogenic and natural factors on the prevalence of physicochemical parameters of water and bacterial water quality indicators along the river Białka, southern Poland. *Environmental Science and Pollution Research*, *25*(10), 10102–10114. https://doi.org/10.1007/s11356-018-1212-2

Bradshaw, J. K., Snyder, B. J., Oladeinde, A., Spidle, D., Berrang, M. E., Meinersmann, R. J., Oakley, B., Sidle, R. C., Sullivan, K., and Molina, M. (2016) Characterizing relationships among fecal indicator bacteria, microbial source tracking markers, and associated waterborne pathogen occurrence in stream water and sediments in a mixed land use watershed. *Water Research*, *101*, 498–509. https://doi.org/10.1016/j.watres.2016.05.014

Brinkmeyer, R., Amon, R. M. W., Schwarz, J. R., Saxton, T., Roberts, D., Harrison, S., Ellis, N., Fox, J., DiGuardi, K., Hochman, M., Duan, S., Stein, R., and Elliott, C. (2015) Distribution and persistence of *Escherichia coli* and *Enterococci* in stream bed and bank sediments from

two urban streams in Houston, TX. *Science of The Total Environment*, *502*, 650–658. https://doi.org/10.1016/j.scitotenv.2014.09.071

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and Eilts, M. D. (1995) The Oklahoma Mesonet: A technical overview. J. Atmos. Oceanic Technol., 12, 5-19.

Byappanahalli, M. N., Nevers, M. B., Korajkic, A., Staley, Z. R., and Harwood, V. J. (2012) Enterococci in the Environment. *Microbiology and Molecular Biology Reviews*, *76*(4), 685–706. https://doi.org/10.1128/MMBR.00023-12

Chen, H. J., and Chang, H. (2014). Response of discharge, TSS, and *E. coli* to rainfall events in urban, suburban, and rural watersheds. Environmental Science: Processes and Impacts, 16(10), 2313–2324. https://doi.org/10.1039/C4EM00327F

Cloutier, D. D., and McLellan, S. L. (2017) Distribution and Differential Survival of Traditional and Alternative Indicators of Fecal Pollution at Freshwater Beaches. *Applied and Environmental Microbiology*, *83*(4). https://doi.org/10.1128/AEM.02881-16

Chithra, S. V., Nair, M. H., Amarnath, A., and Anjana, N. S. (2015) Impacts of impervious surfaces on the environment. *International Journal of Engineering Science Invention*, *4*(5), 27-31.

Christian, L., Epps, T., Diab, G., and Hathaway, J. (2020) Pollutant Concentration Patterns of In-Stream Urban Stormwater Runoff. *Water*, *12*(9), 2534. https://doi.org/10.3390/w12092534

Devane, M. L., Moriarty, E., Weaver, L., Cookson, A., and Gilpin, B. (2020) Fecal indicator bacteria from environmental sources; strategies for identification to improve water quality monitoring. *Water Research*, *185*, 116204. https://doi.org/10.1016/j.watres.2020.116204

Dila, D. K., Corsi, S. R., Lenaker, P. L., Baldwin, A. K., Bootsma, M. J., and McLellan, S. L. (2018) Patterns of Host-Associated Fecal Indicators Driven by Hydrology, Precipitation, and Land Use Attributes in Great Lakes Watersheds. *Environmental Science & Technology*, *52*(20), 11500–11509. https://doi.org/10.1021/acs.est.8b01945

Economy, Louise M., Tracy N. Wiegner, Ayron M. Strauch, Jonathan D. Awaya, and Tyler Gerken. (2019) Rainfall and Streamflow Effects on Estuarine *Staphylococcus aureus* and Fecal Indicator Bacteria Concentrations. *Journal of Environmental Quality* 48, no. 6: 1711–21. https://doi.org/10.2134/jeq2019.05.0196

Garbrecht, J., Van Liew, M., and Brown, G. O. (2004) Trends in precipitation, streamflow, and evapotranspiration in the Great Plains of the United States. *Publications from USDA-ARS/UNL Faculty*, 462.

Garzio-Hadzick, A., Shelton, D. R., Hill, R. L., Pachepsky, Y. A., Guber, A. K., and Rowland, R. (2010) Survival of manure-borne *E. coli* in streambed sediment: Effects of temperature and sediment properties. Water Research, 44(9), 2753–2762. https://doi.org/10.1016/j.watres.2010.02.011

Grant, S. B., Litton-Mueller, R. M., and Ahn, J. H. (2011) Measuring and modeling the flux of fecal bacteria across the sediment-water interface in a turbulent stream. *Water Resources Research*, *47*(5). https://doi.org/10.1029/2010WR009460

Haller, L., Amedegnato, E., Poté, J., and Wildi, W. (2009) Influence of Freshwater Sediment Characteristics on Persistence of Fecal Indicator Bacteria. *Water, Air, and Soil Pollution*, *203*(1), 217–227. https://doi.org/10.1007/s11270-009-0005-0

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., (2020) Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chap. A3, 458 p., https://doi.org/10.3133/tm4a3

Holcomb, D. A., and Stewart, J. R. (2020) Microbial Indicators of Fecal Pollution: Recent Progress and Challenges in Assessing Water Quality. Current Environmental Health Reports, 7(3), 311–324. https://doi.org/10.1007/s40572-020-00278-1

Ibekwe, A. M., Lesch, S. M., Bold, R. M., Leddy, M. B., and Graves, A. K. (2011) Variations of indicator bacteria in a large urban watershed. *Transactions of the ASABE*, *54*, 10.

Islam, M. M. M., Hofstra, N., and Islam, Md. A. (2017) The Impact of Environmental Variables on Faecal Indicator Bacteria in the Betna River Basin, Bangladesh. Environmental Processes, 4(2), 319–332. https://doi.org/10.1007/s40710-017-0239-6

Jin, G., Englande, A. j., Bradford, H., and Jeng, H.-W. (2004) Comparison of *E. coli*, Enterococci, and Fecal Coliform as Indicators for Brackish Water Quality Assessment. Water Environment Research, 76(3), 245–255. https://doi.org/10.2175/106143004X141807

Kim, M., and Wuertz, S. (2015) Survival and persistence of host-associated *Bacteroidales* cells and DNA in comparison with *Escherichia coli* and *Enterococcus* in freshwater sediments as quantified by PMA-qPCR and qPCR. Water Research, 87, 182–192. https://doi.org/10.1016/j.watres.2015.09.014

Korajkic, A., Wanjugi, P., Brooks, L., Cao, Y., and Harwood, V. J. (2019) Persistence and Decay of Fecal Microbiota in Aquatic Habitats. *Microbiology and Molecular Biology Reviews*, *83*(4), e00005-19. https://doi.org/10.1128/MMBR.00005-19

Litton, R. M., Ahn, J. H., Sercu, B., Holden, P. A., Sedlak, D. L., and Grant, S. B. (2010) *Evaluation of Chemical, Molecular, and Traditional Markers of Fecal Contamination in an Effluent Dominated Urban Stream* (world). ACS Publications; American Chemical Society. https://doi.org/10.1021/es101092g

McPherson, R. A., C. Fiebrich, K. C. Crawford, R. L. Elliott, J. R. Kilby, D. L. Grimsley, J. E. Martinez, J. B. Basara, B. G. Illston, D. A. Morris, K. A. Kloesel, S. J. Stadler, A. D. Melvin, A.J. Sutherland, and H. Shrivastava. (2007) Statewide monitoring of the mesoscale environment: A technical update on the Oklahoma Mesonet. J. Atmos. Oceanic Technol., 24, 301–321.

Myers, E. M., and Juhl, A. R. (2020) Particle association of *Enterococcus* sp. Increases growth rates and simulated persistence in water columns of varying light attenuation and turbulent diffusivity. *Water Research*, *186*, 116140. https://doi.org/10.1016/j.watres.2020.116140

ODEQ (Oklahoma Department of Environmental Quality). (2020) Water Quality in Oklahoma 2020 Integrated Report. https://www.deq.ok.gov/water-quality-division/watershed-planning/integrated-report/. Accessed 25 June 2022.

Odonkor, S.T., Ampofi, J.K. (2013) *Escherichia coli* as an indicator of bacteriological quality of water: An overview. Microbiol. Res. 2013, 4, 1–11.

Ogram, A., Castro, H. and Chauhan, A. (2007) Methods of Soil Microbial Community Analysis, p 652-662. In Hurst, Christon J. and Crawford, Ronald L. and Garland, Jay L. and Lipson,

David A. and Mills, Aaron L. and Stetzenbach, Linda D.(ed), Manual of Environmental Microbiology, 3rd edn. doi:10.1128/9781555815882.ch53

OWRB (Oklahoma Water Resources Board). (2017) Oklahoma Water Quality Standards. Oklahoma Administrative Code 785:45. oar.state.ok.us. Accessed 25 June 2022.

OWRB (Oklahoma Water Resources Board). (2020) Status of Water Quality Monitoring in Oklahoma 2019-2020. https://www.owrb.ok.gov/quality/monitoring/StatusReport.pdf. Accessed 25 June 2022.

Petersen, F., and Hubbart, J. A. (2020) Advancing Understanding of Land Use and Physicochemical Impacts on Fecal Contamination in Mixed-Land-Use Watersheds. *Water*, *12*(4), 1094. https://doi.org/10.3390/w12041094

Piorkowski, G. S., Jamieson, R. C., Hansen, L. T., Bezanson, G. S., and Yost, C. K. (2014) Characterizing spatial structure of sediment *E. coli* populations to inform sampling design. *Environmental Monitoring and Assessment*, *186*(1), 277–291. https://doi.org/10.1007/s10661-013-3373-2

Phillips, M. C., Solo-Gabriele, H. M., Reniers, A. J. H. M., Wang, J. D., Kiger, R. T., and Abdel-Mottaleb, N. (2011) Pore Water Transport of Enterococci out of Beach Sediments. *Marine Pollution Bulletin*, *62*(11), 2293–2298. https://doi.org/10.1016/j.marpolbul.2011.08.049

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Schmidt SK, Scow KM (1997) Use of soil bioreactors and microcosms in bioremediation research, p. 822-829. In Hurst CJ, Knudsen GR, McInerney MJ, Stetzenbach LD, Walter MV (ed.), Manual of environmental microbiology. Published by ASM Press, Washington, D.C.

Smith, S. J., and Esralew, R. A. (2010) StreamStats in Oklahoma– Drainage-basin characteristics and peak-flow frequency statistics for ungaged streams (Report No. 2009–5255; Scientific Investigations Report). USGS Publications Warehouse. https://doi.org/10.3133/sir20095255

Staley, C., Dunny, G. M., and Sadowsky, M. J. (2014) Chapter Four—Environmental and Animal-Associated Enterococci. In S. Sariaslani and G. M. Gadd (Eds.), *Advances in Applied Microbiology* (Vol. 87, pp. 147–186). Academic Press. https://doi.org/10.1016/B978-0-12-800261-2.00004-9

Stephenson, G. R., and Rychert, R. C. (1982) Bottom sediment: a reservoir of *Escherichia coli* in rangeland streams. Rangeland Ecology and Management/Journal of Range Management Archives, 35(1), 119-123.

Stocker, M. D., Rodriguez-Valentín, J. G., Pachepsky, Y. A., and Shelton, D. R. (2016) Spatial and temporal variation of fecal indicator organisms in two creeks in Beltsville, Maryland. *Water Quality Research Journal*, *51*(2), 167–179. https://doi.org/10.2166/wqrjc.2016.044

Stocker, M. D., Smith, J. E., Hernandez, C., Macarisin, D., and Pachepsky, Y. (2019) Seasonality of *E. coli* and Enterococci Concentrations in Creek Water, Sediment, and Periphyton. *Water, Air, & Soil Pollution*, *230*(9), 223. https://doi.org/10.1007/s11270-019-4263-1

Suter, E., Juhl, A., and O'Mullan, G. (2011) Particle Association of *Enterococcus* and Total Bacteria in the Lower Hudson River Estuary, USA. Journal of Water Resource and Proection, 3(10). https://doi.org/10.4236/jwarp.2011.310082

Teixeira, P., Salvador, D., Brandão, J., Ahmed, W., Sadowsky, M. J., and Valério, E. (2020) Environmental and Adaptive Changes Necessitate a Paradigm Shift for Indicators of Fecal Contamination. *Microbiology Spectrum*, *8*(2). https://doi.org/10.1128/microbiolspec.ERV-0001-2019

Tiwari, A., Kauppinen, A. and Pitkänen, T. (2019) Decay of *Enterococcus faecalis, Vibrio cholerae* and MS2 coliphage in a laboratory mesocosm under brackish beach conditions. *Frontiers in public health*, *7*, p.269.

USEPA (United States Environmental Protection Agency). (1986) Bacteriological Water Quality Criteria for Marine and Fresh Recreational Waters. EPA- 440/5-84-002. US Environmental Protection Agency, Office of Water Regulations and Standards, Cincinnati, OH, USA.

USEPA (United States Environmental Protection Agency). (2010) Sampling and Consideration of Variability (Temporal and Spatial) for Monitoring of Recreational Waters. EPA-823-R-10-005 US Environmental Protection Agency, Washington, DC, USA.

USEPA (United States Environmental Protection Agency). (2012) Recreational Water Quality Criteria. Office of Water 820-F-12-058. https://www.epa.gov/sites/production/files/2015-10/documents/rwqc2012.pdf

USGS (U.S. Geological Survey). (2014) National field manual for the collection of water-quality data: U.S. Geological Survey Techniques of Water-Resources Investigations, book 9, chaps. A1-A10. http://pubs.water.usgs.gov/twri9A.

Zhang, X., Zhi, X., Chen, L., and Shen, Z. (2020) Spatiotemporal variability and key influencing factors of river fecal coliform within a typical complex watershed. *Water Research*, *178*, 115835. https://doi.org/10.1016/j.watres.2020.115835

Zimmer-Faust, A. G., Thulsiraj, V., Marambio-Jones, C., Cao, Y., Griffith, J. F., Holden, P. A., and Jay, J. A. (2017) Effect of freshwater sediment characteristics on the persistence of fecal indicator bacteria and genetic markers within a Southern California watershed. Water Research, 119, 1–11. https://doi.org/10.1016/j.watres.2017.04.028

Chapter 4 : The Effect of *Paenibacillus* on IDEXX Enterolert™ Results from Stream Environments

*This chapter is in **peer review** as of February 16, 2023, in Environmental Science and Pollution Research with a tentative citation:*

Graves G.M., Tanner R.S., Vogel J.R., De León K.B., and Walls A.W. (2023). The effect of *Paenibacillus* on IDEXX Enterolert™ results from freshwater stream environments. Environmental Science and Pollution Research.

Abstract

Enterolert™ (ELT), a fluorogenic substrate test, is used as a quantitative method for determining freshwater concentrations of *Enterococcus* for water quality indicators. However, there is some evidence from recent studies suggesting that ELT may not suppress false-positives due to pollution sources in waterbodies. In this study, we evaluated this method by analyzing field water and sediment samples from four freshwater streams. We also performed a laboratory microcosm study from two of the stream sediments. The ELT method was investigated by phenotypic and genomic analyses for accuracy of isolating and quantifying *Enterococcus* and/or *Streptococcus*. Additionally, we tested isolates from ELT panels for antibiotic resistance. Results from the field and microcosm studies from initial to final time points indicated that false positives were predominantly *Paenibacillu*s spp. and other non-fecal indicator bacteria. Furthermore, the microcosm study indicated shifts from lactic acid to non-lactic acid bacteria between initial to final time points, but *Enterococcus* concentrations from ELT panels remained stable for the duration of the study for both stream sediments. Antibiotic resistance indicated no distinct pattern of resistance or susceptibility to a suite of antibiotics. However, all isolates tested were resistant to bacitracin and nalidixic acid. In conclusion, we found that ELT was not exclusively selective for *Enterococcus* from freshwater environments, and that sediment and

polluted waterbodies have the potential to skew the presumed concentrations. More research is needed to evaluate the effectiveness and selectivity of the medium used for the fluorogenic substrate test for *Enterococcus* enumeration.

Keywords*:* Enterolert™; *Enterococcus;* freshwater; stream; sediment*; Paenibacillus*

4.1 Introduction

Fecal indicator bacteria (FIB) are important determinants of recreational water quality in freshwater streams and rivers. *Enterococcus* and *E. coli* are the most used FIB as they are generally found in enteric environments and can indicate the presence of fecal contamination of waterbodies. While *E. coli* is a primary indicator species for freshwater, *Enterococcus* or enterococci have been used as a primary indicator for marine and freshwater environments to predict gastrointestinal diseases in humans from direct contact with contaminated water at certain density thresholds (USEPA 1986; USEPA 2012). Conventional methods used to determine *Enterococcus* concentrations in water were membrane filtration or multiple-tube fermentation prior to 1996 (Koide et al. 2007). As the importance of water quality testing for FIB increased, a new technique was developed to provide a standard, easy-to-use method for investigators to evaluate recreational waters (Budnick et al. 1996). A fluorogenic enzyme test, Enterolert™ (ELT), was developed to enumerate FIB in water samples using a most probable number method based on fluorescence of a matrix, or defined substrate, in a multi-well tray that inhibits non-enterococci species from fluorescing and promotes esculinase production by enterococci (Baird and Bridgewater 2017; Chen et al. 1996).

Results from initial trials by Chen et al. (1996) showed ELT effective for selecting *Enterococcus* spp., and Budnick et al. (1996) and Noble et al. (2010) indicated that ELT

performed similarly to membrane filter and multiple tube fermentation techniques. Following initial studies and subsequent successful use by various entities, the method was formally adopted for use by the United States Environmental Protection Agency (USEPA) in 2003 (USEPA 2003). While regulatory entities have widely accepted this USEPA-approved method, early evidence from Kinzelman et al. (2003) found a low correlation between membrane filtration and the fluorogenic methods, and others expressed that more research was needed to determine its appropriateness for freshwater regulatory testing (Ferguson et al. 2013). Recent evidence from Peperzak and van Bleijswijk (2021) suggested that the fluorogenic method (ELT) has a potential for false positives in marine environments, caused predominantly by *Bacillus licheniformis*.

Limited research has been conducted to determine if the fluorogenic substrate method accurately quantifies *Enterococcus* indicators, and suppresses non-*Enterococcus* species, in freshwater samples. Research for method development was from a series of urban runoff and beach samples. Ferguson et al. (2013) showed that species selectivity for the two fecal-associated *Enterococcus* species, *E. faecium* and *E. faecalis*, was lower than that of the marine environment water samples and had a higher percentage of plant associated species. Furthermore, recent studies indicate that *Enterococcus* sources may be non-enteric and in natural aquatic environments such as sediment and decayed organic matter, plants, soils, and animals (Boehm and Sassoubre 2014; Byappanahalli et al. 2012). Badgley et al. (2010) and Devane et al. (2020) revealed that reservoirs of *Enterococcus* and other pathogens of interest exist within the stream water column, sediment, and aquatic vegetation. Additionally, *Enterococcus* populations in the environment have been shown to vary in species distributions and antimicrobial resistance

depending on the specific freshwater conditions and seasons (Alm et al. 2014; Cho et al. 2020; Lupo et al. 2012).

More evidence is needed to understand how *Enterococcus* persists in freshwater environments, and how population shifts and method interferences could potentially alter our understanding of monitoring streams and rivers for recreational water quality. In this paper, we describe results from a study to investigate freshwater streams and lakes in Oklahoma for *Enterococcus* population trends in water and sediment, assess antimicrobial resistance of isolates from ELT panels, and identify isolates from positive ELT panel wells. The objectives of the study were to 1) evaluate *Enterococcus* species distributions in freshwater environments under stable (lab) and field (variable) conditions, and 2) assess the fluorogenic substrate test method (IDEXX Enterolert™) for selectivity of *Enterococcus* in freshwater samples.

4.2 Methods

Representative impaired waterbodies for *Enterococcus* on the 2020 USEPA 303(d) list that included varying stream size, geomorphology, and geographic locations were chosen for this study to limit bias of environmental conditions (ODEQ 2020). Field water and sediment samples were collected at four stream sites (Cat, Bird, Hogshooter, and Walnut Creek) at two time points: Time 1 (Week 1 [T_I]) in July to Time Final (Week 10 [T_F]) in September. Samples were collected during a minimum of 5-d antecedent rainfall (< 0.254 mm) period. Water samples from field sites were collected in 120 ml sterile polystyrene bottles from the stream thalweg and waist-depth from lake shores using United States Geological Survey sampling procedures (USGS 2014). Positive control wastewater samples were collected from the City of Norman, OK wastewater reclamation influent water. Water samples were cooled on ice to <6°C and processed within 24-h hold time using the fluorogenic substrate most probable number (MPN) method, as

described in SM9230D (Baird and Bridgewater 2017). Sterile 290 ml polystyrene bottles were used to collect sediment and water from the upper 5 cm of the benthic stream substrate in the thalweg. The volume of the resulting sediment sample bottles was comprised of roughly 100 ml of sediment and 150 ml of stream water. Immediately following collection, the samples were cooled and stored at <6°C. To establish a controlled and dark environment, the 290 ml polystyrene collection bottles used in the microcosm study had the lids loosely covered and the sides and bottoms tightly wrapped with aluminum foil. The containers were placed on a horizontal orbital shaker at 100 revolutions per minute at constant room temperature (22 °C) to promote aerobic interaction with the sediment as found in natural streambeds. Sample bottles were randomly selected in duplicate at Day 1 (D_I) and Day 31 (D_F) for analysis.

Sediment samples were processed using adapted soil microbial detachment and dispersal methods from Ogram et al. (2007). A 2% sodium pyrophosphate solution was made by adding tetrasodium pyrophosphate to sterile, reverse osmosis (RO) water and buffered to pH 7.0. A sterile serological pipetter was used to decant water from the undisturbed sample bottle. The saturated sediment was dispersed by adding 200 ml of pyrophosphate solution, closing the bottle, and shaking manually for two minutes before placement on a horizontal orbital shaker at 200 revolutions per minute for 15 min. The sediment and pyrophosphate sample mixtures were serially diluted (1:500, 1:1000, 1:10000) with buffered (pH 7.0) sterile RO water and processed and analyzed as aqueous samples using SM9230D for *Enterococcus* (Baird and Bridgewater 2017). Similarly, we evaluated water samples using Enterolert™ (ELT) from the same field locations to verify and compare them to the diluted sediment samples. Following completion of bacteria enumeration, IDEXX processed sample well trays were preserved at < 6°C for short term preservation (<5 days) for Biolog processing and/or 16s sequencing. Representative

streambed sediment was collected and analyzed to determine percent organic matter and particle size distribution using Loss on Ignition (LOI) methods and ASTM 6913 (Ball et al. 1964).

4.2.1 Isolate analysis

Isolates were selected from a subset of the field and laboratory samples at initial and final times for identifying isolates in water and sediment through phenotypic and genomic analyses. Furthermore, isolates from all of the ELT panels were screened to determine phenotypic traits to characterize the enumerated bacteria densities. Isolates from ELT panels at T_I and T_F for the field study and at D_I and D_F for the microcosm study were used to determine the predominant species for each time point and location. Four stream sediments from were analyzed from the field study and both streams from the microcosm were analyzed. To validate methodologies between sediment and water samples, a subset of water samples was analyzed from the field study Week 10 samples. Additionally, isolates of *Enterococcus faecalis* were obtained from the raw wastewater samples and *Paenibacillus licheniformis* was isolated from selective screening methods for positive controls using the lactic minimal medium (LMM) agar method as described below.

Biolog phenotypic methods were used to identify isolates from ELT panels. ELT panels were sprayed with sterile 70% isopropyl alcohol (CiDehol 70, Decon Labs) and allowed to air dry at room temperature. Ten positive wells from each panel were sampled using BD allergy syringes (# 305541) and 5 µl were streaked onto plates of LMM agar. Plates were incubated at 37º C. Dominant colony types were plated further until pure cultures were obtained. Most ELT panels sampled and cultured had a single colony type. Cultures were screened for Gram reaction (Hardy Diagnostics kit GK400A), cell morphology (phase microscopy), motility, and catalase

reaction. If ELT wells contained more than one colony type, they were scored positive as a fecal indicator bacteria (FIB) if at least one was identified as an FIB.

LMM agar was adapted from LMM medium described by Ralph S. Wolfe, which was used in microbial diversity laboratories at the universities of Illinois, Massachusetts, Oklahoma, and elsewhere since the 1960s (Tanner 2007). The recipe for the LMM agar is (/L):10 ml mineral solution; 10 ml vitamin solution; 1 ml trace metal solution; 0.5 g K_2HPO_4; 8 g yeast extract (BD 212750 = Gibco 212750); 10 g glucose; 15 g purified agar (Oxoid LP0028). Precipitated chalk (2 g/L; $CaCO_3$) was added for the detection of acid-producing colonies. Sodium azide (0.1 g/L) may be added to inhibit non-LAB (lactic acid bacteria) and cycloheximide (0.15 g/L) to inhibit fungi but was not used in this study. LMM agar could support the growth of isolates that may not grow on TSA or BUG-blood agar (data not shown).

Isolates at initial and final times of the studies were identified using Biolog's GEN III system (Sandle et al. 2013) and microplates were scored by eye. The identity of approximately 20% of isolates was confirmed by 16S rDNA sequence analysis. Genomic DNA were extracted with the Promega Wizard® Genomic DNA purification kit according to the manufacturer's instructors for Gram-positive bacteria (Promega Corp. Madison, WI) and quantified with the Qubit dsDNA broad range assay (ThermoFisher Scientific, Waltham, MA, USA). The 16S rRNA gene was amplified by PCR with primers FD1 and 1492R and with Taq DNA polymerase and ThermoPol® buffer (New England BioLabs, Inc., Ipswich, MA) (Elnahas et al. 2017; Turner et al. 1999, Weisburg et al. 1991).

4.2.2 Antibiotic resistance

Water samples were collected from two reservoirs (Mountain Lake and Lake Thunderbird) and four stream locations (Bluff, Crooked Oak, Washington, and West Elm Creek) for antibiotic resistance and susceptibility screening. Samples were collected from these additional locations to further evaluate and understand potential differences between water types and various expected levels of pollution. Expected levels of pollution were based on the 303(d) list of impaired waterbodies. For example, Mountain Lake is an isolated rural lake with no known impairments whereas Lake Thunderbird is an urban lake and is impaired for turbidity, dissolved oxygen and nutrients and is expected to have additional inputs of FIB based on nearby urban runoff. Similarly, West Elm Creek and Bluff Creeks are urban streams whereas Crooked Oak and Washington are rural streams. However, all stream sites were listed on the 303(d) list for both *E. coli* and *Enterococcus* (DEQ 2020). Isolates from field water sample ELT panel wells (as described earlier in Methods) were evaluated for antibiotic susceptibility using the method as in Bauer et al. (1966), often referred as the Kirby-Bauer antibiotic disc assay. A total of 35 isolates were screened for antibiotic resistance or sensitivity. The antibiotics tested were: ampicillin (AM), 10 µg; bacitracin (B), 0.04 units; carbenicillin (CB), 100 µg; cefoxitin (FOX), 30 µg; doxycycline (D), 30 µg; erythromycin (E); 15 µg; gentamicin (GM), 10 µg; nalidixic acid (NA), 30 µg; sulfathiazole (ST), 250 µg; tetracycline (TE); 30 µg; trimethoprim (TMP); 5 µg; vancomycin (A), 30 µg.

4.2.3 Data analysis

Results from the Biolog analysis were analyzed by descriptive statistics to determine metrics such as percent false positives and percent FIB. Percent positive fecal indicator bacteria (FIB_P) is determined by the total number of isolates from ELT wells that are identified within the

19 spp. of *Enterococcus* and *Streptococcus* as described in Table 9230:1 of *9230 Fecal Enterococcus/Streptococcus Groups* in Baird and Bridgewater (2017) and divided by the total number of all isolates identified for each location and time point. Total isolates were all isolates recovered from ten positive (fluorescing) ELT panel windows from each individual sample. A Fisher's Exact test conducted in R was used to evaluate significance of FIB percentages and any population shifts between initial and final time points of the study (Fisher, 1934; R Core Team 2013). Species identified were compared and evaluated as a fecal or non-fecal *Enterococcus* source and screened by morphological characteristics for ELT accuracy where false positives are indicated. All data analyses were performed in R and Excel.

4.3 Results and Discussion

Enterococcus water concentrations from the ten-week field study from Cat, Bird, Hogshooter, and Walnut Creek water samples ranged from 171 to 1040 MPN (Most Probable Number) 100 ml^{-1} with an average of 457 MPN 100 ml^{-1} and 101 to 1300 MPN 100 ml^{-1} with an average of 499 MPN 100 ml^{-1} for July (T_I) and September (T_F), respectively. The sediment *Enterococcus* concentrations from the field ranged from 3.9 to 490 MPN per gram and an average of 224 MPN per gram for T_I and ranged from 4.6 to 36 MPN per gram with an average of 19 MPN per gram for T_F samples. The sediment *Enterococcus* concentrations from the laboratory microcosm experiment at Day 1 (D_I) were 1300 MPN per gram and 5.4 MPN per gram, and at Day 31 (D_F) concentrations were 500 MPN per gram and 5.6 MPN per gram for Cat and Walnut, respectively. The pH and water temperature were recorded for all field samples between time points with an average pH of 8.0 and 7.9 and average water temperature of 24.0 and 24.9 °C at T_I and T_F, respectively. Furthermore, pH was measured in the microcosm samples and was 7.7 and 7.4 for Cat and 8.3 and 8.5 for Walnut at T_I and T_F, respectively.

4.3.1 Enterolert™ panel isolates

Sediment samples were analyzed by phenotypic methods to identify isolates from Enterolert™ (ELT) panels on D_1 and D_F of the microcosm sampling period. The Walnut and Cat Creek microcosm results showed that while sediment *Enterococcus* counts remained stable for the 31-d period, the FIB_P percentage decreased from 82% to 0% for Walnut. Cat Creek decreased from 9% to 0% FIB_P for Cat Creek between D_1 and D_F with ELT sediment concentrations of 1320 and 500 MPN per gram, respectively. A Fisher's test was performed on both microcosms and significant differences between time points were shown for Walnut ($p=0.01$) and no significant difference ($p>0.05$) for Cat. Results show that while the panels indicated positive for a FIB, false positives were indicated for 100% of the samples after 31-d with the majority of the isolates (65%) identified by Biolog as *Paenibacillus* spp. Furthermore, D_1 isolates from the Cat samples resulted in 63% of isolates identified as *Paenibacillus* spp.

Even though the *Enterococcus* numbers, as determined by ELT panels, in Walnut and Cat Creek were stable over the 31-d microcosm period, the ELT positive population quickly shifted away from lactic acid bacteria (LAB) to the non-lactic rods, false positives. Specifically for Walnut Creek, the FIB_P isolates as indicated by Gram positive, catalase negative morphologies went from 82% of total isolates on D_I to 45% on Day 2 then averaged 14% (SD ± 5%) for days 5 through D_F (Table 4-1). Most species identifications were found using Biolog and some were confirmed by 16S partial sequence analysis on D_I and D_F. All other LAB isolates from this study were catalase negative, Gram positive, nonmotile cocci in pairs and short chains, and the unidentified non-LAB were catalase positive, Gram positive, motile rods. These were similar to *Paenibacillus* spp. isolates, especially as they all gave unusual reactions on Biolog plates: all false positives, including the negative control well, using protocol A; a yellow color instead of

the usual purple color, due to further reduction of the indicator, in the positive control well using protocol B.

Table 4-1. Summary of isolates from Enterolert™ panels analyzed from the Walnut Creek microcosm stream sediment for a 31-d study duration. FIB = fecal indicator bacteria. FIB percent (%) total is calculated as the number of FIB isolates divided by the total isolates.

Day	Total Isolates	FIB Isolates	FIB % Total	*Enterococcus* sediment (MPN g^{-1})
1	11	9	82%[1]	5
2	8	5	63%	17
5	9	2	22%	9
10	10	2	20%	6
17	12	1	8%	5
24	11	1	9%	4
31	11	0	0%[1]	6

[1]FIB isolates were confirmed by Biolog and/or 16s rDNA sequencing from Day 1 and Day 31.

The LAB from the D_I samples from Walnut microcosm sediment were presumptive enterococcal FIB. Predominant isolates from Walnut were identified as *Enterococcus mundtii*, usually associated with plants but also isolated from a wide variety of other sources (Švec and Franz 2014). Conversely, Cat isolates on D_I were primarily *Paenibacillus* species that are not commonly associated with a human health risk (Grady et al. 2016). D_F isolates for both Cat and Walnut resulted in a 100 percent decrease in FIB_P isolates with the predominate isolates being *Paenibacillus* and/or unidentified non-LAB, which were Gram positive, catalase positive, motile rods (Table 4-2).

Table 4-2. Isolate species identifications from Enterolert™ panels from two microcosms at two time periods, D_I and D_F, from two stream sediment sources. Ten windows from each Enterolert™ panel for each site and time period were sampled for identification tests. The number of multiple identifications (if any) from each sample location and time are denoted by parentheses next to the binomial nomenclature. Fecal indicator bacteria species that are found on the list of enterococci or streptococci fecal indicators in Baird and Bridgewater (2017) are highlighted in bold text.

Location	Sample	Day 1 (D_I) Identification	Day 31 (D_F) Identification
Cat Microcosm	Sediment	***Enterococcus faecium* (1)**	*Paenibacillus sp.* (2)
		Paenibacillus sp. (2)	*P. dendritiformis* (5)
		P. dendritiformis (3)	Unidentified non-LAB (5)
		P. thiaminolyticus (2)	
		Gemella sp. (1)	
		G. palanticanis (1)	
		Unidentified non-LAB (1)	
Walnut Microcosm	Sediment	***E. faecalis* (1)**	*Carnobacterium divergens* (1)
		***E. faecium* (1)**	*Paenibacillus* sp. (2)
		***E. mundtii* (6)**	*P. dendritiformis* (3)
		***E. gallinarium* (1)**	*P. thiaminolyticus* (1)
		Paenibacillus borealis (1)	*Staphylococcus epidermis* (1)
		Unidentified non-LAB (1)	Unidentified non-LAB (4)

Furthermore, field sediment samples from four streams (creeks) were analyzed by phenotypic methods using an identical approach from the microcosm study to identify isolates from ELT panels during a ten-week period. No discernable trend was observed between the July (T_I) and September (T_F) time periods for ELT panels from sediment concentrations in field samples (Table 4-3). A higher average percentage of FIB_P isolates were found in the sediment samples at T_F (50%) compared to T_I (27%). Water samples analyzed from T_F sampling time points resulted in an average of 71% FIB_P isolates. Results from the Fisher's exact test indicated that no significant difference ($p=0.13$) was shown between T_I and T_F for all locations and isolates. Individual sites were evaluated, similarly, and no significant differences ($p > 0.05$) were shown indicating that no change in population was found by chance.

However, when evaluating percent differences between T_I and T_F, the average FIB_P was 27% and 50%, respectively, indicating a quantitative shift increase of fecal indicators between time points. FIB_P isolates for all locations averaged 32% (SD±39%) for T_I and 27% (SD ± 8%) T_F and an average of 69% (SD ± 29%) for all time points. The results indicate that species in the streams are dynamic and that shifts in FIB occurred frequently during the ten-week sampling period for all locations. For example, at Bird Creek, isolates from ELT panels processed from stream sediment samples were identified as 10%, 30%, 90%, 90%, 50%, 60%, 100%, 80%, 10% and 30% FIB_P for consecutive weeks 1 through 10, respectively. *Paenibacillus* spp. were the predominant false positive species identified in 25% of all isolates identified from field sediment and water sample ELT panels.

Table 4-3. Summary of isolates from Enterolert™ panels at four stream sampling locations at two time periods, T_I and T_F. FIB = fecal indicator bacteria. FIB percent (%) total is calculated as the number of FIB isolates divided by the total isolates.

Site	Total Isolates	FIB Isolates	FIB % Total	*Enterococcus* concentration[1,2]
July (T_I) - Sediment				
Bird	14	0	0%	490
Cat	10	2	20%	410
Hogshooter	9	1	11%	40
Walnut	9	7	78%	3.9
September (T_F) - Sediment				
Bird	10	2	20%	4.6
Cat	11	5	45%	36
Hogshooter	9	6	67%	23
Walnut	9	6	67%	13
September (T_F) - Water				
Bird	10	9	90%	100
Cat	11	9	82%	450
Hogshooter	10	6	60%	440
Walnut	10	5	50%	140

[1] *Enterococcus* sediment units: MPN g^{-1}
[2] *Enterococcus* water units: MPN 100 ml^{-1}

A high percentage (>90%) of the non-FIB from sediment windows from the microcosm and field samples (113 of 531) were isolates of *Paenibacillus*, including *P. apiarus*, *P. borealis*, *P. dendritiformis*, *P. graminis*, *P. thiaminolyticus,* and *P. woosongensis*. Other species recovered from sediment windows were *Aeromonas veronii*, *Burkholderia multivorans*, *Carnobacterium divergens* (a LAB), *Gemella palaticanis,* and *Streptococcus acidominimus* (a LAB). Thirteen of 119 windows from aqueous stream sample ELT panel windows were non-FIB, but these were a diverse group of bacteria including *Carnobacterium gallinarium* (a LAB), *Cellulomonas hominis*, *Chryseobacterium humi*, *Enterococcus canintestini* (a LAB), *Lactococcus garviae* (a LAB), *Jonesia denitrificans*, *Ochrobactrum intermedium*, *Paenibacillis sanguinis*, *P. thinaminolyticus,* and *Proteus mirabilis* (Tables 4-2 and 4-4). The resulting identifications indicate that ELT panels are not exclusive for selection of the targeted FIB for freshwater samples and these isolates are potentially not considered a human health concern as many are of environmental origin (Devane et al. 2020). Furthermore, these isolates could create interferences and variability in most probable number calculations due to the high percentage of false positives and non-target selection characteristics of the medium.

Table 4-4 (next page). Isolate species identifications from Enterolert™ panels from four field stream sample locations at two time periods, T_I and T_F. Ten windows from each Enterolert™ panel for each site and time period were sampled for identification tests. The number of multiple identifications (if any) from each sample location and time are denoted by parentheses next to the binomial nomenclature. Fecal indicator bacteria species that are found on the list of enterococci or streptococci fecal indicators in Baird and Bridgewater (2017) are highlighted in bold text.

Location	Sample	July (T_I) Identification	September (T_F) Identification
Bird	Sediment	*Paenibacillus apiarius* (10) *Aeromonas veronii* (1) *Aeromonas* sp. (1) *Bacillus* sp. (1) *Paenibacillus* sp. (1)	***Enterococcus mundtii* (3)** *Ochrobactrum* sp. (5) *Micrococcus* sp. (1) *Paenibacillus* sp. (1)
Cat	Sediment	***E. casseliflavus* (2)** *Ochrobactrum* sp. (4) *Paenibacillus* sp. (1) *P. thiaminolyticus* (2) *P. apiarius* (1)	***E. faecalis* (2)** ***E. mundtii* (3)** *Carnobacterium divergens* (1) *Micrococcus* sp. (1) *P. thiaminolyticus* (2) Unidentified LAB (2)
Hogshooter	Sediment	***E. faecalis* (1)** *Burkholderia multivorans* (1) *Paenibacillus* sp. (1) *P. apiarius* (1) *P. thiaminolyticus* (5)	***E. faecalis* (6)** Unidentified LAB (3)
Walnut	Sediment	***E. casseliflavus* (4)** ***E. mundtii* (3)** *Enterococcus* sp. (1) *P. thiaminolyticus* (1)	***E.faecalis* (3)** ***E. mundtii* (3)** *Enterococcus* sp. (1) *P. sanguinis* (1) *P. thiaminolyticus* (1)
Bird	Water	-[1]	***Enterococcus mundtii* (9)** *Paenibacillus* sp. (1)
Cat	Water	-	***E. canintestini* (1)** ***E. faecalis* (4)** ***E. gallinarum* (3)** ***E. mundtii* (3)**
Hogshooter	Water	-	***E. faecalis* (6)** Unidentified LAB (4)
Walnut	Water	-	***E. casseliflavus* (2)** ***E. dispar* (1)** ***E. faecium* (1)** ***E. mundtii* (1)** *Cellulomonas hominis* (2) *Chryseobacterium humi* (1) *Kocuria* sp. (1) *Proteus mirabilis* (1)

[1]One time period was used (T_F) to compare water versus sediment samples.

Isolate identification showed that many isolates were not FIB and potentially were from other non-enteric sources. Many of the more recently added FIB from Baird and Bridgewater (2017) would be considered of animal origin (*E. hirae*, *E. columbae, E. cecorum, E. saccharolyticus,* and *E. asini*) or of environmental origin (*E. casseliflavus* and *E. mundtii*) rather than of human origin (Švec and Franz 2014). Additionally, the list of fecal enterococci/streptococci increased from four species in 1985 (Greenberg et al. 1985) to 18 species in 2017 (Baird and Bridgewater 2017); n.b., *Streptococcus bovis* is a synonym of *Streptococcus equinus* as shown in the List of Prokaryotic names with Standing in Nomenclature (Parte et al. 2020). However, limited information is available for the increase of fecal indicator species on this list. Many of these species that are not of animal origin are often associated in the same environmental conditions as fecal-origin bacteria except these LAB spp. identified may have been isolated from contaminated water samples (Korajkic et al. 2018). More research is needed to assess whether the fecal indicators added between 1985 to 2017 are of concern to human health in primary body contact recreation freshwaters and how to apply methodology for accurately identifying targeted species.

4.3.2 Enterolert™ false positives

The most widely used application of ELT is for routine bacterial enumeration of *Enterococcus* in non-potable and potable water sources. However, we also evaluated this method to evaluate the efficacy of ELT with introduction of sediment by using sediment dispersion methods, diluting the sediment sample, and processing them as water samples as described in Methods section. The fluorogenic substrate in ELT is 4-methyl-umbelliferyl-β-D-glucoside, an analog of esculin (James et al. 1997). The same β-D-glucosidase that cleaves esculin also cleaves this fluorogenic substrate, releasing the chromophore. The medium used for ELT may

have an inducer of this esculin hydrolase and any esculin hydrolysis positive species could give a positive fluorescent result in an ELT window. The published false positive species for ELT, *Bacillus licheniformis* (Peperzak and van Bleijswijk 2021) is esculin positive (Logan and De Vos 2009). Most spp. of *Paenibacillus* are esculin positive (Priest 2009).

The stream substrates of the two sample locations were analyzed for sediment properties with D_{50} particle sizes of 5.08 and 0.38 mm and percent organic matter of 1.2% and 0.4% for Cat and Walnut, respectively. With the difference in stream substrate types and organic matter, potential interferences in the Cat samples may be due to the organic matter percentage. Zimmer-Faust et al. (2017) and others have indicated that organic material may have a positive effect on the viability and diversity of bacterial communities within stream environments. Additionally, sediments are known to increase the persistence of *Enterococcus* in the water column and could potentially create interferences with the medium and isolation of targeted species in ELT panels (Graves et al. 2023; Haller et al. 2009). Typical freshwater stream environments often carry a sediment load such as the sediment samples in this study that we analyzed as water samples (Stocker et al. 2019). The water samples collected and analyzed were typically less turbid than the diluted sediment samples, which had a known sediment load introduced into each water sample and may have been a contributing factor that resulted in fewer false positives as shown by similar enumeration methods (USEPA 2000). The variability of FIB_P with no distinct pattern identified between sediment and water samples indicates that the addition of sediment seems to increase irregularities in positively selecting targeted indicator species.

Results from the microcosm study correspond with the field species identification where high percentages of false positive indicators were present in positive ELT wells. We found that regardless of the of the FIB_P counts in all samples, false positives were consistent throughout all

panels and had an abundance of non-FIB_P bacterium (both LAB and non-LAB) between each individual panel. Furthermore, we used *P. licheniformis* and *E. faecalis* as presumed positive controls and both indicators resulted in positive fluorescence in all 96 ELT panel wells for both species with an approximate cell count of 500 cells per mL.

As the ELT method is a USEPA-approved method for recreational freshwater quality criteria, true numbers of FIB as indicated in Baird and Bridgewater (2017) could be misrepresented for concentration requirements for impaired waterbodies. Furthermore, selectivity for FIB_P in freshwater may be skewed by false positive indicators such as *Paenibacillus* spp. and not representative of the waters evaluated, especially if the waters have contributing sediment or pollution from sources such as wastewater or urban runoff (Ferguson et al. 2013; Peperzak and van Bleijswijk 2021; Suzuki et al. 2012). Overall, the results indicate that the number of species identified from the microcosm and field studies are of varying origin in the environment and suggest that more emphasis in selectivity of target fecal indicators should be considered when using ELT for analyzing freshwater samples.

4.3.3 Antibiotic resistance

Isolates from field water and sediment samples were evaluated for antibiotic resistance and susceptibility to understand potential predictors and characterization of *Enterococcus* spp. from freshwater samples processed using ELT methods. Isolates were cultured and recovered from direct plating of water samples and isolation from positive ELT panel windows. Results indicated that all 35 isolates were resistant to bacitracin and nalidixic acid. Bacitracin, in general, targets Gram positive bacteria (Dubos 1939), but our set of isolates contradicted this conclusion. However, nalidixic acid is generally more effective against Gram negative bacteria, so the resistance observed may not be unusual (Cook et al. 1966). All but one isolate was

resistant to sulfathiazole, which was a *Lactococcus garvieae* recovered from an ELT window inoculated with water from Crooked Oak. Additionally, sulfathiazole is a broad-spectrum antibiotic, therefore the resistance observed here was not expected but has been documented from fecal samples (Middleton and Ambrose 2005). We found limited differences between the overall antibiotic resistance of isolates between direct isolation (plating) or from an ELT window. The only isolates (n = 4) that were resistant to vancomycin were all isolated directly from the Norman Wastewater Reclamation Influent: *Enterococcus faecium*, *Leuconostoc lactis*, *Paenibacillus anaericanus,* and *Weissella halotolerans*. *Enterococcus* or Gram-positive species displayed an overall higher susceptibility to most of the other antibiotics selected (Table 4-5). Given the potential for antibiotic-resistant bacterium in the environment, more information is needed to understand which antibiotics have the potential to suppress non-indicator bacterium for inclusion into selective medium for bacteria quantification testing such as the fluorogenic substrate tests.

Table 4-5. Antibiotic resistances of isolates recovered from Enterolert™ panels from six waterbodies.[a] The antibiotics tested were ampicillin (AM), bacitracin (B)[d], carbenicillin (CB), cefoxitin (FOX), doxycycline (D), erythromycin (E), gentamicin (GM), nalidixic acid (NA)[d], sulfathiazole (ST)[d], tetracycline (TE), trimethoprim (TMP), and vancomycin (VA)[d]. The site locations were Bird (B), Crooked Oak (CO), Lake Thunderbird (LT), Mountain Lake (ML), Washington Creek (WA), West Elm (WE), and Norman Wastewater Reclamation Influent (WW).

Site	n[b]	AM[c]	CB	FOX	E	GM	D	TE	TMP
B	6	1	3	5	6	1	-	-	-
CO	1	-	-	1	-	-	-	-	-
LT	1	-	-	1	-	-	1	-	-
ML	1	-	-	-	-	-	-	-	1
WA	7	-	3	4	4	1	5	4	-
WE	1	-	-	1	1	-	-	-	-
WW	6	-	1	6	5	5	1	1	1

[a]Number of resistant isolates.
[b]Number of isolates from the site.
[c]Antibiotics and concentrations in Methods section.
[d]Bacitracin (B) and nalidixic acid (NA) are not shown in the table as all 35 isolates tested were resistant to both antibiotics.

4.4 Conclusions

Freshwater fecal indicators are important for designating impairment status and related potential for human health impacts. However, considerations for applicable methods used in freshwater waterbodies should be evaluated further for regulatory decisions to properly target and quantify fecal indicators. We investigated the ELT method using a combination of field and laboratory freshwater stream samples and determined potential interferences in *Enterococcus* enumeration. This work provides evidence that false positives can be dominant in freshwater stream samples using the fluorogenic substrate method for *Enterococcus* and that population shifts away from FIB can occur in both field and laboratory environments over short time periods. In addition, many of our identified LAB isolates from ELT panels could be considered of environmental or animal origin. *Paenibacillus* spp. (non-LAB) was the predominant false positive indicator among all isolates identified in sediment and water samples regardless of sample location, type, or time period. Antibiotic resistance from various waterbodies provided initial evidence of resistance to nalidixic acid and bacitracin that are not typical of Gram-positive species.

Our work coincides with and adds to the evidence of Peperzak and van Bleijswijk (2021) where they found false positives were in seawater samples with the genus *Bacillus* as the primary cause of interference. However, more studies are required to understand the impacts of sediment, organics, and other potential interactions and influences on the isolation of targeted Gram-positive FIB. Furthermore, research is needed to improve the fluorogenic substrate technology, such as ELT, to suppress non-FIB in freshwater samples.

4.5 References

Alm EW, Zimbler D, Callahan E, and Plomaritis E (2014) Patterns and persistence of antibiotic resistance in faecal indicator bacteria from freshwater recreational beaches. J of Appl Microbiol, 117:1, 273–285. https://doi.org/10.1111/jam.12512

Badgley BD, Nayak BS, and Harwood, VJ (2010) The importance of sediment and submerged aquatic vegetation as potential habitats for persistent strains of enterococci in a subtropical watershed. Water Res, 44:20 5857–5866. https://doi.org/10.1016/j.watres.2010.07.005

Baird R and Bridgewater L (2017) Standard methods for the examination of water and wastewater. 23rd edition. Washington, D.C., American Public Health Association.

Ball DF (1964) Loss-on-ignition as an estimate of organic matter and organic carbon in non-calcareous soils. J Soil Sci 15:84–92.

Boehm AB and Sassoubre LM (2014) Enterococci as indicators of environmental fecal contamination. In: Enterococci: From Commensals to Leading Causes of Drug Resistant Infection. Massachusetts Eye and Ear Infirmary. PMID: 24649503.

Budnick GE, Howard RT, and Mayo DR (1996) Evaluation of Enterolert for enumeration of enterococci in recreational waters. Appl and Environ Microbiol, 62:10, 3881–3884. https://doi.org/10.1128/aem.62.10.3881-3884.1996

Byappanahalli, MN, Nevers MB, Korajkic A, Staley ZR, and Harwood VJ (2012) Enterococci in the Environment. Microbiology and Molecular Biology Reviews, 76:4, 685–706. https://doi.org/10.1128/MMBR.00023-12

Chen CM, Doherty K, Gu H, Dichter G, and Naqui A (1996) Enterolert—A rapid method for the detection of *Enterococcus* spp. In: ASM abstracts, no. Q448. 96th American Society for Microbiology general meeting. New Orleans. p. 464.

Cho S, Jackson CR, and Frye JG (2020) The prevalence and antimicrobial resistance phenotypes of *Salmonella*, *Escherichia coli* and *Enterococcus* sp. In surface water. Letters in Applied Microbiology, 71:1, 3–25. https://doi.org/10.1111/lam.13301

Cook, TM, Brown KG, Boyle JV, and Goss WA (1966) Bactericidal action of nalidixic acid on *Bacillus subtilis*. Journal of Bacteriology 92:1510-1514.

Devane ML, Moriarty E, Weaver L, Cookson A, and Gilpin B (2020) Fecal indicator bacteria from environmental sources; strategies for identification to improve water quality monitoring. Water Res, 185, 116204. https://doi.org/10.1016/j.watres.2020.116204

Dubos RJ (1939) Studies on a bactericidal agent extracted from a soil bacillus: I. Preparation of the agent. Its activity in vitro. The Journal of Experimental Medicine 70:1-10.

Elnahas M, Amin M, Hussein M, Shanbhag V, Ali A, Wall J. (2017) Isolation, characterization and bioactivities of an extracellular polysaccharide produced from *Streptomyces* sp. MOE6. Molecules 22:1396.

Ferguson DM, Griffith JF, McGee CD, Weisberg SB, and Hagedorn C (2013) Comparison of *Enterococcus* species diversity in marine water and wastewater using Enterolert and EPA Method 1600. Journal of Environmental and Public Health, 2013, e848049. https://doi.org/10.1155/2013/848049

Fisher RA (1934) Statistical methods for research workers, 5th edn. Oliver and Boyd, Edinburgh.

Grady EN, MacDonald J, Liu L, Richman A, and Yuan ZC (2016) Current knowledge and perspectives of *Paenibacillus*: A review. Microbial Cell Factories, 15(1), 203. https://doi.org/10.1186/s12934-016-0603-7

Graves GM, Vogel JR, and Tanner RS (2023) Investigation of environmental factors on *Enterococcus* survival in Oklahoma streams. Aquatic Sciences, 85(2), 34.

Greenberg AE, Trussell RR, Clesceri LS (1985) Standard methods for the examination of water and wastewater, 16th edn. Published by American Public Health Association, Washington DC.

Kinzelman J, Ng C, Jackson E, Gradus S, and Bagley R (2003) Enterococci as indicators of lake michigan recreational water quality: comparison of two methodologies and their impacts on public health regulatory events. Applied and Environmental Microbiology, 69:1, 92–96. https://doi.org/10.1128/AEM.69.1.92-96.2003

Koide T, Sugita E, Iyo T, Miyata N, and Iwahori K (2007) Evaluation and Comparison of Various Enumeration Methods for Enterococci in Surface Water. *Japanese Journal of Water Treatment Biology*, *43*(3), 129–142. https://doi.org/10.2521/jswtb.43.129

Korajkic A, McMinn BR, and Harwood VJ (2018) Relationships between microbial indicators and pathogens in recreational water settings. International Journal of environmental research and public health, 15:12, Article 12. https://doi.org/10.3390/ijerph15122842

James AL, Perry JD, Ford M, Armstrong L, and Gould FK (1997) Cyclohexenoesculetin-β-D-glucoside: a new substrate for the detection of bacterial β-D-glucosidase. Journal of Applied Microbiology, 82:4, 532-536.

Logan NA and De Vos PG (2009) Genus I. *Bacillus*, p. 84. In: De Vos PG et al. (eds), Bergey's Manual of Systematic Bacteriology, 2nd edn., vol. 3. Springer, New York City.

Lupo A, Coyne S, and Berendonk T (2012) Origin and evolution of antibiotic resistance: The common mechanisms of emergence and spread in water bodies. frontiers in microbiology, 3. https://www.frontiersin.org/articles/10.3389/fmicb.2012.00018

Noble RT, Blackwood AD, Griffith JF, McGee CD, and Weisberg SB (2010) Comparison of rapid quantitative pcr-based and conventional culture-based methods for enumeration of *Enterococcus* spp. and *Escherichia coli* in recreational waters. Applied and Environmental Microbiology. https://doi.org/10.1128/AEM.00651-10

Middleton JH and Ambrose A (2005) Enumeration and antibiotic resistance patterns of fecal indicator organisms isolated from migratory Canada geese (Branta canadensis). Journal of Wildlife Diseases, 41:2, 334-341.

ODEQ (Oklahoma Department of Environmental Quality). (2020) Water Quality in Oklahoma 2020 Integrated Report. https://www.deq.ok.gov/water-quality-division/watershed-planning/integrated-report/. Accessed 25 February 2023.

Ogram A, Castro H and Chauhan A (2007) Methods of Soil Microbial Community Analysis, p 652-662. In Hurst et al. (eds), Manual of Environmental Microbiology, 3rd edn. doi:10.1128/9781555815882.ch53

Parte, AC, Sardia Carbasse J, Meier-Kolthoff JP, Reimer LC, Goker M. (2020) List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. Int J Syst Evol Microbiol 70:5607-5612.

Peperzak L and van Bleijswijk J (2021) False-positive enterococci counts in seawater with the IDEXX Enterolert-E most probable number technique caused by *Bacillus licheniformis*. Environmental Science and Pollution Research. https://doi.org/10.1007/s11356-020-11342-6

Priest FG (2009) Genus I. *Paenibacillus*, pp. 269-295. *In* De Vos, P., G. M. Garrity, D. Jones, N. R. Krieg, W. Ludwig, F. A. Rainey. K.-H. Schliefer and W. B. Whitman (ed.), Bergey's Manual of Systematic Bacteriology, 2nd edn., vol. 3. Springer, New York City.

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Sandle T, Skinner K, Sandle J, Gebala B, and Kothandaraman, P (2013) Evaluation of the GEN III OmniLog® ID System microbial identification system for the profiling of cleanroom bacteria. European Journal of Parenteral and Pharmaceutical Sciences, 18(2), 44-50.

Stocker MD, Smith JE, Hernandez C, Macarisin D, and Pachepsky Y (2019) Seasonality of *E. coli* and Enterococci Concentrations in Creek Water, Sediment, and Periphyton. Water, Air, and Soil Pollution, 230(9), 223. https://doi.org/10.1007/s11270-019-4263-1

Suzuki Y, Kanda N, and Furukawa T (2012) Abundance of *Enterococcus* species, *Enterococcus faecalis* and *Enterococcus faecium*, essential indicators of fecal pollution, in river water. Journal of Environmental Science and Health, Part A, 47:11, 1500–1505. https://doi.org/10.1080/10934529.2012.680315

Švec P, and Franz C (2014) The genus *Enterococcus*. In *Lactic Acid Bacteria* (pp. 175–211). John Wiley and Sons, Ltd. https://doi.org/10.1002/9781118655252.ch15

Tanner, RS (2007) Cultivation of Bacteria and Fungi. In Hurst CJ, Crawford RL, Mills AL, Garland JL, Stetzenbach LD, Lipson DA (eds) Manual of Environmental Microbiology 3rd edn. ASM Press, Washington DC, pp 69-78. https://doi.org/10.1128/9781555815882.ch6

Turner S, Pryer KM, Miao VPW, Palmer JD (1999) Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. J Eukaryotic Microbiology 46:327–338.

USEPA (United States Environmental Protection Agency) (2012) Recreational Water Quality Criteria. Office of Water 820-F-12-058. https://www.epa.gov/sites/production/files/2015-10/documents/rwqc2012.pdf

USEPA (U.S. Environmental Protection Agency) (2003) Guidelines Establishing Test Procedures for the Analysis of Pollutants; Analytical Methods for Biological Pollutants in Ambient Water; Final Rule. U.S. Federal Register - 40 CFR Part 136 Vol. 68, No. 139.

USEPA (U.S. Environmental Protection Agency) (2000) Improved enumeration methods for the recreational water quality indicators: enterococci and *Escherichia coli*. EPA/821/R-97/004. Office of Science and Technology, U.S. Environmental Protection Agency, Washington, D.C.

USEPA (United States Environmental Protection Agency) (1986) Bacteriological Water Quality Criteria for Marine and Fresh Recreational Waters. EPA- 440/5-84-002. US Environmental Protection Agency, Office of Water Regulations and Standards, Cincinnati, OH, USA.

Weisburg WG, Barns SM, Pelletier DA, Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173:697–703.

Zimmer-Faust AG, Thulsiraj V., Marambio-Jones C, Cao Y, Griffith JF, Holden PA, and Jay JA (2017) Effect of freshwater sediment characteristics on the persistence of fecal indicator bacteria and genetic markers within a Southern California watershed. Water Res, 119, 1–11. https://doi.org/10.1016/j.watres.2017.04.028

Chapter 5 : Exploring Regional and Statewide Relationships between Fecal Indicator Bacteria, Water Quality and Geospatial Data in Oklahoma Streams

This chapter is formatted for submission to the Journal of the American Water Resources Association.

Abstract

Understanding fecal indicator bacteria (FIB) in freshwater streams is important for protecting human health and meeting water quality standards. Current approaches for assessment of stream water quality includes routine sampling of point locations at priority probabilistic stream locations. While this approach provides an immense amount of important information for decision-making, it does not regularly account for spatial or temporal factors that may influence FIB concentrations in streams. For this study, we used a decade (2001-2011) of water quality and hydrologic data collected by the Oklahoma Conservation Commission in concert with geospatial data (i.e., watershed physical, hydrologic and geographical characteristics) from readily public-available sources to develop correlation and regression outputs that can be used to provide an initial remote assessment of streams by region in Oklahoma. Sites were grouped into five unique regions based on watershed drainages (HUC-4) and related hydrologic characteristics. A Spearman correlation was used to provide initial linear relationships among all 26 factors between *Enterococcus* and *Escherichia coli*. Results indicate that significant weak to moderate correlations between geospatial, water quality and hydrologic factors were shown for all regions. To further investigate relationships between variables, stepwise ordinary multiple linear regression models were developed from these data to identify significant variables that can be used to predict FIB concentrations in each region and statewide. Results from the model indicate that contributing drainage area (CONTDA) increased the number of covariates in the models between all regions. Furthermore, specific region-trends were evident and were explained by precipitation gradients,

CONTDA or known regional general water quality characterization. Model adjusted-R^2 was highest between *Enterococcus* and *E. coli,* but in almost all cases, introduction of additional independent variables reduced variability of the model. Turbidity was the only identified common predictor in all but one region for both FIB. The average unexplained model variability was approximately 40%. Results from the model indicate that significant model predictors are often region specific and have unaccounted variability that may be due to spatial and temporal factors. This study demonstrates that large geospatial and hydrologic datasets can be used to develop initial assessments of expected FIB in freshwater streams by region and watershed.

Keywords: *water quality, fecal indicator bacteria, geospatial, regression, correlation, streams*

5.1 Introduction

Non-point source pollution concerns in Oklahoma streams have led to a committed effort by state agencies, specifically the Oklahoma Conservation Commission (OCC), since 2000 to routinely monitor streams and rivers for water quality impairments. Routine monitoring includes selecting a sub-group of stream segments within basins to synoptically measure physical, chemical and biological parameters on a rotating basis every five years. From a period of 2000 to 2019, 727 priority streams and rivers were sampled across Oklahoma for routine water quality parameters (dissolved oxygen, pH, temperature, specific conductance, hardness and alkalinity), nutrients and ions (nitrogen, phosphorus, sulfate, and chlorides), fecal indicator bacteria (*Escherichia coli* and *Enterococcus*) and stream discharge. The data are used by state agencies to identify streams that are potentially degrading due to non-point or other unidentified sources (OCC, 2020).

The extensive water quality dataset from the OCC also provides an opportunity to explore potential patterns and relationships between water quality parameters and other spatiotemporal datasets. For example, water quality monitoring data has been used to predict and develop relationships between fecal indicator bacteria (*E. coli* and *Enterococcus*) and water quality conditions (Devane et al. 2020). Furthermore, inclusion of watershed characteristics and land use features may provide additional insight to interconnectedness with FIB and water quality (Vitro et al., 2017). Hydrologic characteristics including precipitation, antecedent dry periods and stream discharge have all been indicated as drivers for FIB in streams (Vidon et al. 2008). Geospatial data has been shown to be a valuable tool in predicting FIB levels on local and watershed scales (Fisher et al., 2000; Petersen and Hubbart, 2020). Spatial studies have also been used to develop relationships between fecal indicator bacteria (FIB) and geospatial and environmental variables (Kang et al. 2010; Piorkowski et al. 2014).

FIB are extensively used to evaluate and characterize streams for impairment status related to recreational waters and waterbody standards (OWRB 2017; USEPA 2012). However, temporal and spatial variations of FIB in recreational waters are not well-understood. The United States Environmental Protection Agency (USEPA) and others have indicated that variability in sampling time and location, environmental conditions, and water quality factors may contribute to higher variability among samples (Devane et al. 2020; Stocker et al. 2019; Stocker et al. 2018; USEPA 2010). Therefore, an assessment of a large monitoring dataset using novel statistical methods and approaches could provide insights to understanding water quality dynamics in relation to surface water beneficial use criteria. Prediction models have been used to evaluate and develop relationships between FIB, spatial water quality, land use and related geographical datasets to prioritize best management practices for certain watersheds (Kang et al. 2010). Implications for assessing existing water quality and related geographical datasets can provide

new tools for water resources managers and planners to understand aquatic ecosystem health (Liao et al. 2018). Given the substantial investment required to routinely monitor each stream within a state or watershed, as reported in waterbody success stories by the USEPA (2018), it is imperative that alternative, complementary solutions are developed to maintain and improve our waters.

Water quality data from the OCC from a period of 2000 to 2019 were collected as part of the non-point source (NPS) pollution USEPA 319 grant program for the State of Oklahoma to monitor and assess priority wadeable (or lower order) stream reaches. Compiled data for each rotating basin, or small watershed region, is used for determining and reporting the extent of non-point source pollution impacts to the State's waters (OCC 2020). However, the data collected also have the potential to provide a detailed analysis of any stream parameters and conditions that may play roles in predicting fecal indicator bacteria (FIB) concentrations in Oklahoma streams and rivers. While others have focused on smaller watershed approaches to develop geospatial FIB relationships such as from agricultural watersheds, limited studies exist to evaluate these metrics on a regional or statewide scale. The objectives of this study were to 1) explore relationships between hydrological, chemical, physical and geographical stream characteristics through a correlation analysis and 2) develop potential predictors for FIB in freshwater streams for statewide and multi-watershed level scales through multiple linear regression models.

5.2 Methods

Water quality data was compiled from various sources to develop a comprehensive dataset to effectively evaluate FIB and water quality among related geographical, hydrological, chemical and biological variables. Data included water quality and hydrologic data from the Oklahoma Conservation Commission (OCC) from 2001 to 2011, geographical coverages from the U.S.

Geological Survey (USGS) and Natural Resources Conservation Service (NRCS), and meteorological data from the Oklahoma Mesonet. A summarized table of data sources, types, and year periods is provided in Table 5-1 and a summary of stream water chemistry and parameters are shown in Table 5-2.

Table 5-1. Data sources and types that are incorporated into FIB statistical analyses.

Data Sources			
Data Source	**Type**	**Data Description**	**Year/Period**
OCC	Tabular Data	Stream Water Chemistry	2001-2011
NRCS, USGS	Raster	National Land Cover Dataset	2011
OK Mesonet	Tabular Data	Daily Precipitation	2001-2011
USGS	Shapefile	Watershed Boundary Dataset	2022
USGS	Shapefile	StreamStats Basins/Characteristics	2022

Table 5-2. Description of parameters and factors collected for the tabular datasets to be incorporated into FIB statistical analyses.

Study Data Summary

Data Source	Data Type	Parameters
OCC	Water Chemistry	NO3-, NO2-, Ortho-P, Total P, TKN NH3, Cl-, Sulfate, TSS, TDS
OCC	Field Parameters	Water temperature, DO, pH, Conductivity, Alkalinity, Hardness, Turbidity
OCC	Microbiology	*E. coli*, Enterococcus
USGS	Coverage/Landscape	Contributing Drainage Area, Canopy Coverage, Impervious Area, Soil Permeability, Annual Precipitation
OCC, Mesonet	Hydrology	Daily precipitation, Stream discharge
Calculated	Hydrology	24-,48- and 72-hr Precipitation totals, Antecedent dry period

5.2.1 Spatial Extent of Sites

OCC data include over 22,000 observations from 727 sites over a 19-year period. However, *Enterococcus* sampling occurred for a period of 2001-2011 and this data subset was used in the analyses to ensure pairwise comparisons of all water quality parameters and *E. coli*. Additionally, sampling points were filtered to those that are only during the primary body contact recreation period for fecal indicator bacteria impairment assessment criteria from April through September (OWRB 2017). Water quality data was cleaned for missing observations, anomalies, and other extraneous errors in R. Following, observations and sites were organized by region to further investigate any regional differences. Sites were first separated by Hydrologic Unit Code (HUC) 4 basins, and where sites were close to boundaries HUC 8 basins were used to further separate regions. The basin site separations are similar to those used for the OCC rotating basin monitoring program's regions (OCC 2020). The five final resulting regions were Central-East (CentE), Northeast (NE), Northwest (NE), Southeast (SE) and Southwest (SW) as identified in Figure 5-1.

Oklahoma Mesonet daily rainfall data was used to develop four additional datasets from a custom R-code: antecedent dry period, one-day preceding rainfall, two-day preceding rainfall and three-day preceding rainfall. Sites were spatially related in ArcGIS with the nearest straight-line distance Oklahoma Mesonet station to match precipitation data to the sites. Watersheds (subbasins) for each sampled stream location were delineated from U.S. Geological Survey Stream Stats using the batch processing tool, where contributing drainage area and average annual precipitation were calculated (USGS 2019). Land use data (soil permeability, impervious area, canopy coverage) from the 2011 National Land Cover Dataset from the USGS and NRCS was then clipped to each subbasin in ArcGIS (Homer et al. 2015). Finally, calculated spatial data were

matched with site identifiers from the OCC water quality datasets for analysis with the correlation and multiple regression analyses.

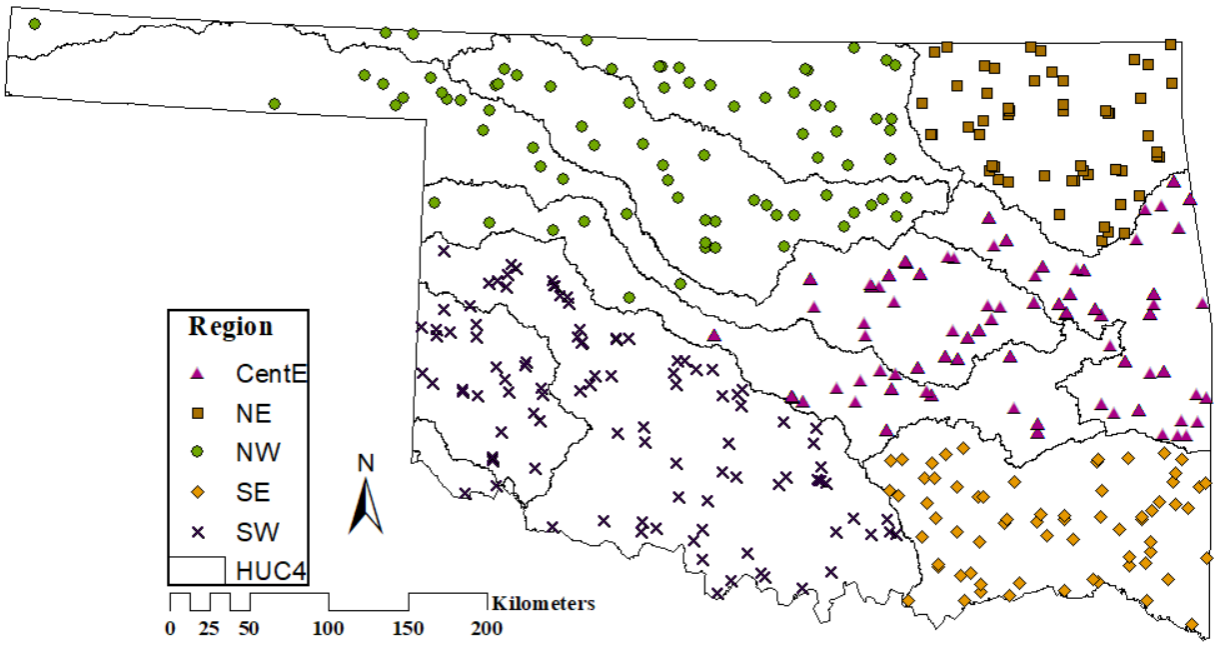


Figure 5-6. Site map of Oklahoma with selected subregions and HUC (Hydrologic Unit Code) -4 watersheds. Regions are defined as CentE = Central-East, NE = Northeast, NW = Northwest, SE = Southeast and SW = Southwest.

5.2.2 Statistical Analysis

Descriptive statistics for water quality parameters were calculated in R and Microsoft Excel®. Initial correlations for the water quality dataset were developed using a Spearman correlation matrix in R using the 'gcorrplot' package. A two-sided t-test ($\alpha = 0.05$) was used to evaluate where significant linear relationships exist (R Core Team 2013). Furthermore, a multiple linear stepwise regression analysis was performed using R and JASP software to analyze all water quality and hydrological values against fecal indicator bacteria concentrations to evaluate any potential significant predictor variables at $\alpha = 0.05$ (JASP Team 2023). A stepwise regression was chosen due to the a) the large number of interaction terms and b) has been previously used for similar water quality and environmental studies (Yang et al. 2017). Data was screened and identified as skewed from residuals and skewness calculations, therefore, values were log-transformed to symmetrize the residuals and improve the potential linear variance relationships of the continuous variables (Helsel et al. 2020).

A natural log transformation in the form of $y = \ln(x+1)$ was applied on all variables. The constant (+1) was used in this analysis given that many of the continuous variables were skewed with true zeros such as in the case of precipitation amounts or percent impervious area. Constants are commonly used in many different analyses of continuous and count data, with +1 being the most frequently chosen value when handling zeros (Bellégo et al. 2022). Therefore, the following equation form for this model was:

$$\ln(FIB) = b_0 + b_1 * \ln(x_1 + 1) + b_2 * \ln(x_2 + 1) + \cdots + b_n * \ln(x_n + 1), \qquad \textit{Eq. 5-1}$$

and results in the subsequent equation after taking exponentials,

$$FIB = e^{b_0} + e^{b_1 * \ln(x_1+1)} - 1 + e^{+ b_2 * \ln(x_2+1)} - 1 + \ldots + e^{b_n * \ln(x_n+1)} - 1, \qquad \textit{Eq. 5-2}$$

where FIB = fecal indicator bacteria concentration (*Enterococcus* or *E. coli*), b_0 to b_n = regression coefficients, and (x_1-1) to (x_n-1) = independent covariates (water quality and watershed variables) and accounting for the constant introduced during the initial log transformation.

5.3 Results and Discussion

Before correlation or regression analyses, 3402 sampling site visits from 306 locations were split into five subregions and ranged from minimum of 318 sampling points in the Southeast (SE) to a maximum of 957 sites in the Northwest (NW). Contributing drainage area (CONTDA) for individual sites ranged from a median area of 44.4 km^2 in the SE to 145 km^2 in the NW with a statewide total drainage area for all locations of 739,000 km^2. The minimum CONTDA was located in the NE, NW and SE regions of 0.049 km^2 and a maximum of 10600 km^2 in the Central-East (CentE) region. A site layout with delineated drainages is shown in Figure 5-2 and associated contributing drainage areas and number of site observations per region is presented in Table 5-3.

Hydrologic conditions were additional variables that were used to characterize regions and understand and predict FIB concentrations in streams and have been used extensively in related FIB studies (Gregory et al. 2019; Rochelle-Newall et al. 2016; Verhougstraete et al. 2015). The defined regions resulted in variable differences in precipitation amounts, preceding rainfall periods and stream discharge. Annual precipitation in subbasins ranged from an average of 762 inches in the NW region to 1270 mm in the SE with an average of 990 mm (SD ± 115 mm) across all regions. ADP ranged from an average of 5 days in the SE to 8 days in the NW, with a maximum of 49 days in the NE, and an average of 6 days (SD ± 6 d) statewide. Stream discharge ranged from an average of 0.27 m^3s^{-1} in the SE to 3.85 m^3s^{-1} in the NE and an average of 0.91 m^3s^{-1} (SD = 13.5 m^3s^{-1}).

Table 5-3. Descriptive statistics of the contributing drainage areas (CONTDA) and sites within each region. Regions are defined as CentE = Central-East, NE = Northeast, NW = Northwest, SE = Southeast and SW = Southwest. Site observations are defined as the total number of samples for all sites in each region.

	Contributing Drainage Area (CONTDA)				
Statistic	**CentE**	**NE**	**NW**	**SE**	**SW**
Site Observations (n)	478	325	957	318	782
Median (km^2)	129	65.9	145	44.4	95.2
Minimum (km^2)	0.086	0.049	0.049	0.049	0.065
Maximum (km^2)	10600	5930	9690	652	6450

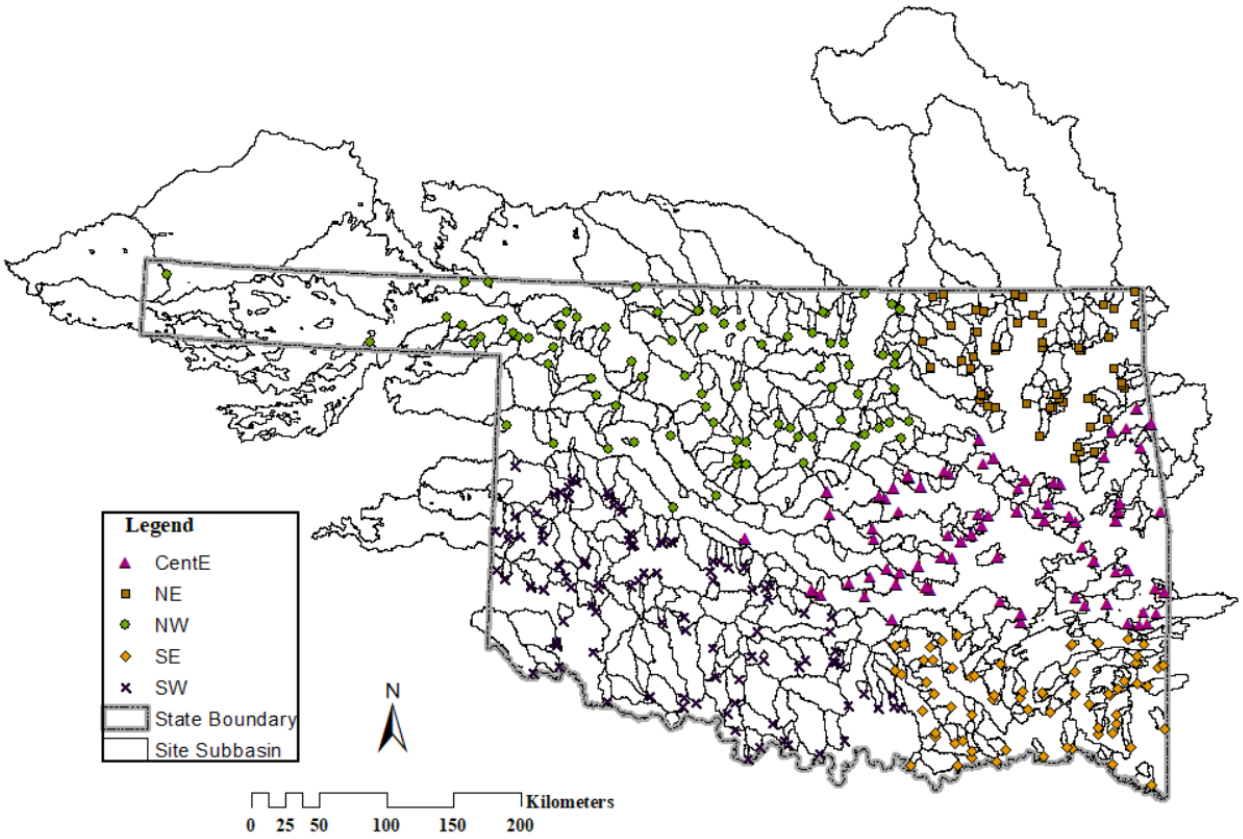


Figure 5-7. Site locations with delineated subbasins for each site location. Regions are defined as CentE = Central-East, NE = Northeast, NW = Northwest, SE = Southeast and SW = Southwest.

Descriptive statistics and general trends for fecal indicator bacteria (FIB), *Enterococcus* and *Escherichia coli*, were calculated for the period of record and all regions. *Enterococcus* concentrations ranged from a minimum of 5 to a maximum reported of >10,000 CFU 100 ml^{-1} with a geometric mean of 127 CFU 100 ml^{-1} (SD ± 4.4 CFU 100 ml^{-1}) for all sites. *E. coli* ranged from a minimum of 5 to a maximum of 10,000 CFU 100 ml^{-1} with a geometric mean of 104 CFU 100 ml^{-1} (SD ± 4.8 CFU 100 ml^{-1}) for all sites. The range of all locations was 9,995 CFU 100 ml^{-1}, except for the SE region (1,555 CFU 100 ml^{-1}). Based on the range of values, it is assumed that the minimum and maximum values from the dataset were calculated as the minimum and maximum detection levels reported either due to dilution factors out of range or another quality assurance procedure that the lab the Oklahoma Conservation Commission used for their analyses. However, log transformations, as described in the stepwise linear regression section address the skewness of the data and how outliers were addressed. Boxplots in Figures 5-3 and 5-4 show the general trends of *Enterococcus* and *E. coli* for the study period from 2001 to 2011.

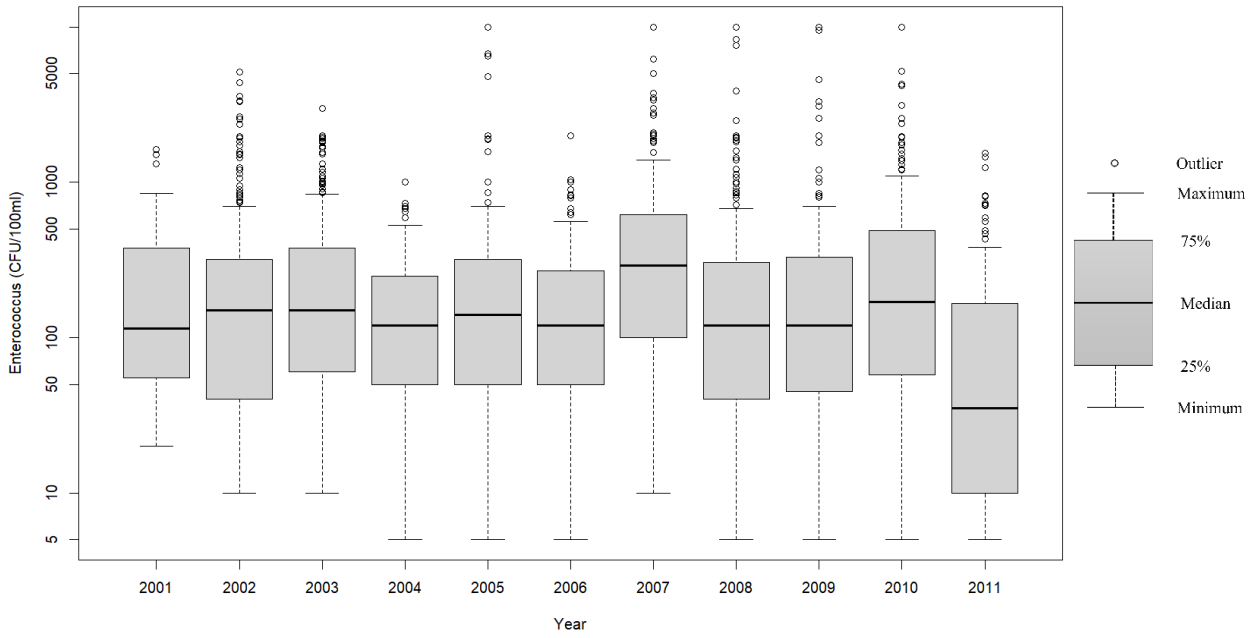


Figure 5-8. Log-Boxplot of Enterococcus concentrations from stream sampling locations statewide by year for the period of 2001-2011.

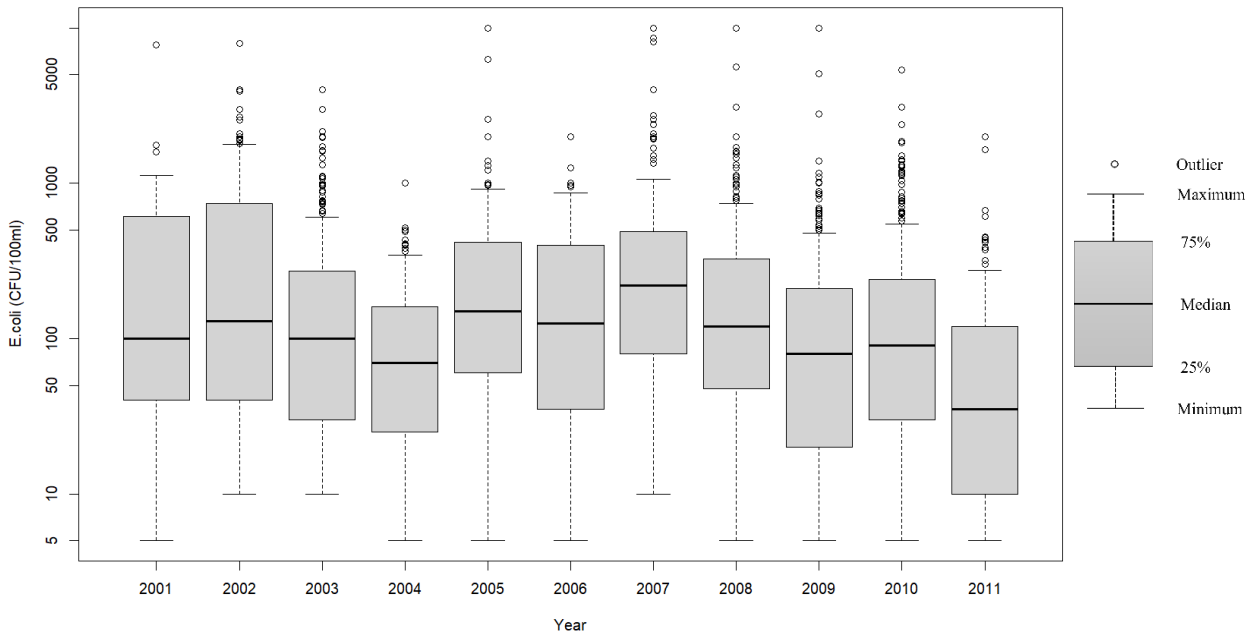


Figure 5-9. Log-Boxplot of *E. coli* concentrations from stream sampling locations statewide by year for the period of 2001-2011.

5.3.1 Correlation analysis

A Spearman correlation analysis was used to understand what correlations, if any, were present between site variables. Spearman rank-order correlation was chosen as the appropriate method as all data were right-skewed and non-linear and has been used in similar studies to relate geospatial and water-quality data (Brendel and Soupir 2017). Correlation analysis is often used as an important procedure to detect trends before developing a regression model to understand the linear relationships between variables (Gauther 2001). Therefore, Spearman correlations were performed on 28 water quality and geospatial variables in each of the five HUC-based regions and statewide. Results were categorized by a) strength of association, b) direction of correlation (+ or -) and c) significant parameters. Spearman correlations range from a strong linear positive correlation of 1 to a strong linear negative correlation of -1. In general, values less than ± 0.3 indicate weak correlations, ± 0.3 to 0.7 are moderately correlated, values greater than ± 0.7 are strongly correlated (Akoglu 2018) (Figure 5-5).

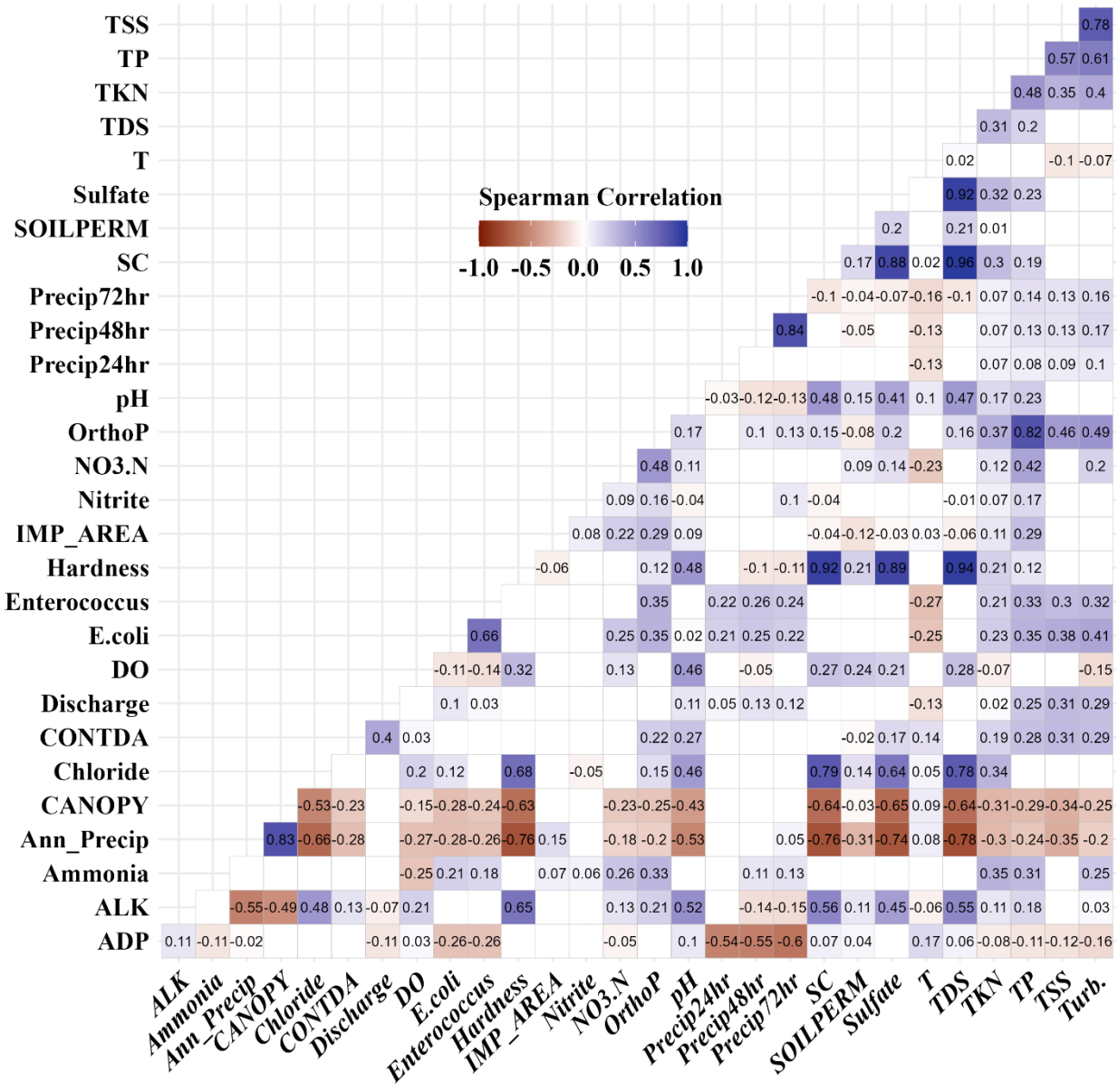


Figure 5-10. Spearman correlation matrix of all hydrologic, water quality and geospatial variables from 308 stream sampling locations and 2860 sampling points at Oklahoma streams. The Spearman values range from a strong positive correlation of 1 to a negative correlation of -1 indicating a very strong positive/negative monotonic relationship. Values and blocks shown are significant from a t-test analysis ($p<0.05$). Acronym explanation: ADP = antecedent dry period, ALK = Alkalinity, CONTDA = contributing drainage area, DO = dissolved oxygen, IMP_AREA= impervious area, NO3.N = Nitrate-Nitrogen, SC = specific conductance, SOILPERM = soil permeability, T= water temperature, TDS = total dissolved solids, TKN = Total-Kjeldahl Nitrogen, TP = total phosphorus, TSS = total suspended solids.

Spearman correlation analysis for all regions and data was conducted to understand initial relationships between factors and FIB and for comparison to the region analysis. Overall, the statewide correlation indicated that no significant strong correlations were shown between *E. coli* or *Enterococcus* and the variables. However, weak to moderate relationships were prominent between many hydrological and water quality variables (Table 5-4). From the results of the statewide Spearman correlation showing that most variables were not strongly correlated to FIB, a cutoff for strength of relationship was used for these variables and to easily compare regions. Significant ($p < 0.05$) Spearman correlations that were $\geq \pm 0.3$ were used as the threshold for determining strong correlation trends between regions, however, all significant correlations were considered and included in the results. Related studies have used the ±0.3 as a metric indicating the minimum level for a moderate correlation (Akoglu 2018). Results indicate that *E. coli* and *Enterococcus* were strongly positively correlated for all regions. All regions, except for Central-East (CentE) for *Enterococcus* and Northwest (NW) for *E. coli*, exhibited significant positive correlations for turbidity. All regions showed significant weak positive correlations for Total-Kjeldahl nitrogen (TKN). Significant positive correlations were also indicated for most regions (min. of 8 regions between both FIB) for ADP, AnnPrecip, discharge, ortho-phosphate (OrthoP), Precip24hr, Precip48hr, Precip72hr, T, TP, and TSS (Table 5-4).

Table 5-4. Significantly correlated ($p < 0.05$) variables from the Spearman correlation for all water all hydrologic, water quality and geospatial variables from 308 stream sampling locations and 2860 sampling points at Oklahoma streams.

Significant factor ($p < 0.05$)	***E. coli***	***Enterococcus***
Antecedent dry period (ADP)	-0.26	-0.26
Annual Average Precip (AnnPrecip)	-0.28	-0.27
Ammonia	-0.21	-0.18
Canopy Percentage (CANOPY)	-0.28	0.24
Dissolved Oxygen (DO)	-0.14	-0.18
Orthophosphate (OrthoP)	0.35	0.35
24-h precipitation (Precip24hr)	0.21	0.22
48-h precipitation (Precip48hr)	0.25	0.26
72-h precipitation (Precip72hr)	0.22	0.24
Water Temperature (T)	-0.25	-0.27
Total Kjeldahl Nitrogen (TKN)	0.23	0.21
Total Phosphorus (TP)	0.35	0.33
Total Suspended Solids (TSS)	0.38	0.3
Turbidity (Turb.)	0.41	0.32

Hydrologic factors such as discharge and precipitation were found in many cases to correlate to FIB, however, no significant trend ($p > 0.05$) was found from a one-way ANOVA when comparing AnnPrecip, Discharge, Precip24hr, Precip48hr, Precip72hr. Similarly, no significant trends were determined when grouping other factors (water quality parameters, or watershed characteristics). While no significant trends between regions are determined from the correlation analysis, unique correlations appeared when comparing correlations between environmental variables and *Enterococcus* and *E. coli*. For example, contributing drainage area showed significant weak correlations in the NW and SW regions, however, no correlations were found for *E. coli* for any region. Additionally, contributing drainage area (CONTDA) showed a weak negative correlation with *Enterococcus* for the west regions (NW, SW), however no correlation was found for the east regions (CentE, NE, SE). Differences in significant correlations were also shown for precipitation (Precip24hr, -48hr, and -72hr). The NW and SW showed no correlation with *Enterococcus* and precipitation whereas significant correlations were found for the CentE, NE, and SE regions. *E. coli*, however, did not show any obvious correlation trends with precipitation.

Many of the significant factors from our results such as hydrologic and sediment variables have been shown to be strong predictors of FIB (Haller et al. 2009). Additionally, nutrients in the forms of OrthoP, TKN, and TP were moderately correlated to FIB in at least 80% of the regions. Other studies have indicated that due to the linear relationship between nutrients and sediments in freshwater streams that FIB are also inherently correlated with nutrients due to similar sorption characteristics (Christensen et al. 2002). The SE region was unique in that the region had a larger number of correlated factors, 19 for *Enterococcus* and 23 for *E. coli*, compared to the average of 13 for all regions. The SE region was shown to have a larger number of factors that correlated with FIB and some of these variables were unique to the region compared to all others: chloride,

hardness, impervious area (IMPAREA), nitrate, nitrite, soil permeability (SOILPERM) and TDS. Considerations for the SE region may include CONTDA is approximately 50% smaller on average than the remaining regions, and as IMPAREA increases and SOILPERM decreases, the linear relationship shows that *Enterococcus* and *E. coli* both increase in concentration. Furthermore, as chloride, TDS, hardness and nitrate are highly mobile in water, the flashiness of the smaller watersheds could be contributing factors for why these variables are positively correlated with FIB. Conversely, the SE region was the only region that hydrologic (discharge, ADP, Precip24, -48, -72hr) and water temperature (T) did not show a correlation with FIB. Again, this could be due to hydrologic differences compared to the rest of the state as the SE region receives considerably more annual rainfall (38%) than the rest of the subbasins in this study. Wetter regions have been shown to exhibit consistently higher FIB concentrations and are tied to precipitation (Dila et al. 2018). However, in our study, the variation of precipitation in the other regions that have longer ADPs may result in higher FIB concentrations due to the hydrologic response that is magnitudes higher than typical baseflow conditions in the streams and the CONTDA size.

Table 5-5. Spearman correlation strength association table for *Escherichia coli* and *Enterococcus* by region. A positive (+) value indicates a significant ($p<0.05$) positive correlation and a negative (-) value indicates a significant negative correlation. Values that are highlighted are those that have ≥ ±0.3 Spearman rank correlation coefficient.

	Enterococcus Region						*E. coli* Region					
Variable	State	CentE	NE	NW	SE	SW	State	CentE	NE	NW	SE	SW
ADP	-	-	-	-		-	-	-	-	-		-
ALK			-	-	+				-	-	+	
AnnPrecip	-	-	-		-		-	-	-		-	
CANOPY	-	-			-		-		-		-	
Chloride					+						+	
CONTDA				-		-						
Discharge	+	+	+	+		+		+	+			+
DO	-			-	-		-			-	-	
E. coli	+	+	+	+	+	+						
Enterococcus							+	+	+	+	+	+
Hardness					+						+	
IMP_AREA					+						+	
Nitrite					+						+	
NO3-N					+		+				+	
OrthoP	+			+	+	+	+	+	+	+	+	+
pH			-						-		+	
Precip24hr	+	+		+		+	+	+	+	+	+	+
Precip48hr	+	+		+		+	+	+	+	+	+	+
Precip72hr	+	+		+		+	+	+	+		+	+
SC				-	+					-	+	
SOILPERM					-						-	
Sulfate											+	
T	-	-		-		-	-	-	-			-
TDS	+				+						+	
TKN	+	+	+	+	+	+	+	+	+	+	+	+
TP	+	+		+	+	+	+	+	+	+		+
TSS	+	+		+	+	+	+		+	+	+	+
Turbidity	+	+	+	+	+	+	+	+	+		+	+

5.3.2 Stepwise linear regression

Ordinary stepwise linear regression methods were used to reject statistically insignificant independent variables ($p > 0.05$) from the FIB regression model. The independent variables included all 26 water quality and geospatial variables and the dependent variables were *E. coli* and *Enterococcus* (FIB) concentrations. The stepwise model involves iteration and replacement of independent (predictor) variables to include in the final regression model that can explain the variation of the dependent variables. In this study, a forward stepwise regression was chosen as Spearman correlations (Section 5.3.1) between the FIB and independent variables were generally moderate or low, except between *E. coli* and *Enterococcus*. This exploratory approach allows for unbiased variable inclusion by adding each one step by step; it keeps only the statistically significant variables when compared to the group (Henderson & Denison, 1989).

Regression model coefficients were developed to determine the significant predictors for FIB concentrations and were evaluated by regions identified in Section 5.2.1. In all cases, *Enterococcus* and *E. coli* were significant model predictors and had a strong influence explaining the variability of the response. The average adjusted-R^2 for the linear regression between *E. coli* and *Enterococcus* for all watershed regions and statewide was 0.44 (SD ± 0.06) and the average RMSE was 1.07 (SD ± 0.06). However, adding water quality and geospatial models improved all other regional model outcomes, with an average adjusted-R^2 of 0.53 (SD ± 0.06) for *Enterococcus* and 0.55 (SD ± 0.06) for *E. coli*, with an average number of covariates (nCOV) of 9 and 7, respectively. The model region with the highest adjusted-R^2 and lowest Root Mean Square Error (RMSE) was the NE region for both FIB with 0.63 (RMSE = 0.84, nCOV = 4) and 0.64 (RMSE = 0.92, nCOV = 6) for *Enterococcus* and *E. coli*, respectively. Conversely, the CentE model region resulted in minimum adjusted-R^2 0.47 (RMSE = 1.02, nCOV = 7) and 0.45 (RMSE

= 1.08, nCOV = 5) for *Enterococcus* and *E. coli*, respectively. Summary tables of significant ($p < 0.05$) covariates of the final models are provided in Tables 5-5 and 5-6 for *Enterococcus* and *E. coli*, respectively.

The NW region exhibited both a larger nCOV and contrasted with many of the other regions. Additionally, it provided similarities with the statewide regression coefficients. For *Enterococcus*, specifically, the nCOV was 12 whereas all other locations averaged 6 nCOV. An explanation of this occurrence could be due to hydrologic and landscape characteristics that influence how specific predictors respond to physical or chemical changes. For example, contributing drainage area (CONTDA) for the NW is larger (48% of total) than the other regions (10% average) and also potentially skews the statewide average given the percentage of watershed coverage of all sites and CONTDA is indicated as a significant predictor for the NW region for both FIB. Furthermore, differences are shown between east and west regions of the state for CONTDA where the NW, SW and CentE regions are significant predictors and the NE and SE were excluded from the model ($p > 0.05$). Discharge was the only hydrologic significant predictor included for *Enterococcus* in the NW region, however, the *E. coli* regression model included AnnPrecip, Precip24hr, and Precip48hr as significant predictors. Conversely, the SE region had a lower nCOV than other regions with a smaller CONTDA, on average, and no commonalities were shown between SE predictors for *E. coli* and *Enterococcus*. However, when looking at an overall analysis of CONTDA versus nCOV a logarithmic trend relationship was found with a resulting R^2 of 0.55 (n=10). Similar studies from agricultural watersheds have indicated that drainage area size influences water quality parameters and interrelatedness (Brenner and Brenner 1995). Therefore, the overall conclusion is that CONTDA seems to influence the nCOV and subsequently the variance and complexity of the model due to the number of interactive predictors with FIB.

Multiple studies have indicated that precipitation is strongly correlated to antecedent dry periods and hydrological conditions (Chen & Chang, 2014; Hamilton & Luffman, 2009; Zubizarreta, 2018). However, in this regression analysis, no discernable trends were identified when comparing hydrologic variables and FIB between the regions. The results indicate that precipitation is a significant predictor in the statewide regression for both FIB and in some regions annual average rainfall, daily preceding rainfall, and ADP were significant variables included in the models. The NW and SE regions were found to be significant predictors between ADP and *Enterococcus*, which is unique in that the gradient extreme of precipitation in Oklahoma is from the NW (lowest) to the SE (highest) (Oklahoma Climatological Survey, 2023). However, precipitation was not included in the final model for either region. Only one inclusion of discharge from the NW region for *Enterococcus* was included in the final models. Overall, hydrologic influences were shown to improve the model in some regions but was not a prominent predictor compared to all other factors.

The only predictor that was found in the majority of regions (5 of 6) was turbidity for *E. coli*. In all other cases, no trends were apparent between or among each FIB. However, the uniqueness of each region may be related to the uniqueness of the regions and watersheds. Other studies have indicated that water quality differences can occur by watershed regions and should be considered when understanding relationships between water quality at statewide or larger scales (Soranno et al. 2011). For example, total suspended solids (TSS) were included as significant predictors in the statewide, NW and CentE regions for *Enterococcus* and only in the SE for *E. coli.* Whereas water temperature (T) and dissolved oxygen (DO) (which are often correlated) were included as significant predictors with the three largest average watershed drainage areas, Statewide, CentE and NW for *Enterococcus*, but were only found in the SE for *E. coli.* Furthermore, in the NW region, both FIB models included Nitrate as a predictor. The OCC

water quality data indicates that the NW region has a 54% higher mean nitrate concentration compared to all other regions. From other studies nitrate is known to be in the highest concentrations in NW Oklahoma, specifically for alluvial aquifers which contribute to a majority of stream baseflow in the region (Masoner and Mashburn 2004). Therefore, it is not surprising that this relationship occurs, but also provides evidence that some parameters are region-specific and can potentially be used as prediction variables for FIB in streams.

Remotely sensed and readily available geospatial and hydrologic data that are easily collected may provide useful information for reconnaissance efforts for FIB prediction (Sokolova et al. 2022; Verhougstraete et al. 2015). In this regression analysis, precipitation (24-, 48- and 72-hr preceding rainfall, ADP, and AnnPrecip), contributing drainage area, soil permeability, canopy percentage and impervious area percentage were all found as significant predictors to include in some or all region models. When only considering these nine variables in the stepwise regression model, the resulting adjusted R^2 averaged 0.23 for *Enterococcus* and 0.24 for *E. coli*. Results indicate that while remotely sensed data can improve the models, field data collection is important to improve the models. For example, in the lowest performing remotely sensed region (SW), including basic water quality parameters (T, SC, DO, Turb., and pH) improved the adjusted R^2 from 0.13 to 0.34 and 0.13 to 0.27 for *E. coli* and *Enterococcus*, respectively. The results from these explorative combinations of remotely sensed and in-situ water quality data indicate that adding more variables can reduce uncertainty and are important to explain the variability in the model.

Given the dynamic nature of freshwater stream systems, the statewide probabilistic sampling locations, number of samples collected in a stream reach, and time that samples were collected are only a snapshot of the potentially spatial and temporal variability that could occur in

the watershed (USEPA 2010). Additionally, geographic factors were aggregated either from closest weather station (precipitation) to aggregation of an average across the watershed (i.e., percent canopy, impervious area or soil permeability) and could increase spatial uncertainty based on sampling points (Desai and Rifai 2010). Even with these unknowns, the model was able to show significant interactions and interconnectedness of parameters that explained at least half variability in the model and compares to a similar spatiotemporal model analysis by Guo et al. (2020). As demonstrated by Brendel and Soupir (2017) on similar analysis of smaller agricultural watersheds, it is expected that a more densely sampled data set, both temporally and spatially, would improve the model based on similar studies from smaller agricultural watersheds (Brendel and Soupir 2017). More research is needed to understand the uncertainties of dynamic stream systems and how we can best model and predict FIB concentrations in streams with field and geospatial data for remote water quality assessments.

Table 5-6. Significant ($p < 0.05$) linear regression model coefficients for *Enterococcus* (dependent variable) by region and by model predictors (independent variables). Regions in the table correspond to regions in the map previously introduced in Methods (Figure 5-1) and variable descriptions are listed in Figure 5-5.

	Enterococcus Coefficients by Region					
Model Predictor[a]	State	CentE	NE	NW	SE	SW
Intercept	10.1	18.2	1.65	7.97	4.35	5.03
ADP				-0.12	-0.20	
ALK	0.07			0.18		
AnnPrecip	-1.27	-2.70				
CANOPY	0.13				-0.39	-0.14
Chloride	-0.07					-0.15
CONTDA	-0.09	-0.24		-0.05		-0.10
Discharge				-0.08		
DO	-0.56			-0.47	-0.48	-0.70
E. coli	0.50	0.38	0.49	0.57	0.58	0.50
Hardness						-0.41
Nitrate	0.31			0.36		
pH				-1.85		
Precip24hr	0.47		0.73			0.43
Precip48hr	0.47		0.80			
Precip72hr		0.62				0.45
SC				-0.46		
SOILPERM	-0.26					
Sulfate		-0.19				0.13
TDS	0.07			0.48		0.31
Temperature	-0.81	-1.14		-0.51		
TKN					0.68	
TSS	0.12	0.29		0.14		
Turbidity			0.21			
Model Adj. R^2	0.52	0.47	0.63	0.52	0.51	0.52
RMSE	0.98	1.02	0.84	0.98	1.00	0.86
n Covariates	14	7	4	12	5	10

[a]Insignificant variables that were not included in the final model were: Ammonia, OrthoP, TP, Nitrite

Table 5-7. Significant ($p < 0.05$) linear regression model coefficients for *Escherichia coli* (dependent variable) by region and by model predictors (independent variables). Regions in the table correspond to regions in the map previously introduced in Methods (Figure 5-1) and variable descriptions are listed in Figure 5-5.

	Escherichia coli Coefficients by Region					
Model Predictor[a]	State	CentE	NE	NW	SE	SW
Intercept	6.71	1.78	1.52	8.14	0.04	10.4
ADP	-0.07		-0.21			
ALK				-0.22	0.22	
AnnPrecip	-0.45			-0.66		
CANOPY	-0.07	-0.31	-0.17			
Chloride			-0.22			
CONTDA				-0.10		
DO						0.49
Enterococcus	0.55	0.40	0.63	0.54	0.50	0.59
Hardness	0.18					
IMP_AREA				-0.53		
Nitrate				0.34		
pH	-1.18					-3.02
Precip24hr				0.45		
Precip48hr	0.36	0.96				
Precip72hr				0.42		
Sulfate		0.15	0.27			
TDS	-0.13					
Temperature	-0.42			-0.83		-1.23
TKN				0.31		
TSS			0.26			
Turbidity	0.27	0.31		0.20	0.40	0.35
Model Adj. R^2	0.54	0.45	0.64	0.55	0.56	0.53
RMSE	1.02	1.08	0.92	0.96	0.91	1.01
n Covariates	10	5	6	11	3	5

[a]Insignificant variables that were not included in the final model were: Ammonia, Nitrite, OrthoP, SOILPERM, TP, and SC.

5.3.3 Model Limitations and Assumptions

Multiple linear regressions are used for developing statistical models that can provide estimates of dependent variables from regression estimates and computed statistics. However, the accuracy of the model can depend on the unexplained error of the dependent variables that includes both the model and sample error. The unaccounted variability in the model may be due to the temporal and spatial availability of water quality data and FIB concentrations (Guo et al. 2020). In this multiple linear stepwise regression, variables were added and removed based on a p-value threshold of 0.1 for a conservative approach of including interactions between the independent (water quality, hydrology, geospatial) and the dependent variables (FIB). This conservative approach is often used to ensure that non-significant variables are included that may have underlying relationships with other variables that increase the strength of the model (Berger et al. 2018). The model also may overpredict based on training data due to introduction of many predictor variables, which can lead to false positives and overfitting of the model data (Steyerberg 2019). To account for this false positive potential, we used bootstrapping methods with 200 replications to ensure that model errors were not significantly different ($p<0.05$) between the original stepwise and bootstrapped samples. Estimates of predictors in this regression model followed assumptions of normality, homoscedasticity, independence and linearity based on the initial and final analysis of the residuals of the coefficients as linear regression models require for validity of the results (Ghani and Ahmad 2010).

Our models suggest that in many cases approximately half of the variability is unaccounted for in the models. This could potentially be due to 1) the limitation of the model to predict FIB due to unknown variations in FIB concentrations in the river or stream, 2) low correlation between predictors that leads to multicollinearity and 3) limitation of the dataset due

to skewed or complex non-linear relationships that are not well-related or represented in a linear model and 4) in the cases where zeroes are present (i.e., precipitation) data may have higher variance even with log transformation. Other uncertainty considerations could include sampling methods, laboratory methods, and variable hydrologic conditions (Harmel et. al 2016). Therefore, improvement of spatiotemporal water quality datasets is needed to refine and reduce the variability of the model. The predictive outputs from this study should be used with caution for understanding relationships between predictors and response variables as it ultimately may not explain the best variables to use in each particular watershed, stream, study reach or season. For example, understanding ambient conditions, potential point sources or other bacteria influences in the studied watersheds may provide additional insight into the variability of each model (Zhang et al. 2020). While our model could most likely be improved with more data, refined model inputs and other advanced statistical tools, these may come with a tradeoff of increased resources required (Gholizadeh et al. 2016). However, overall, our study does provide evidence that regression models can be used to preliminarily understand the predictors which may have influence on FIB concentrations in freshwater streams.

5.4 Conclusions

Water quality data is often collected for regulatory or beneficial use criteria assessments and geospatial data is collected for survey and other related purposes. For this study, a decade of data from water quality samples collected and associated watershed characteristics were explored through correlation and regression analyses. The purpose of this study was to provide an initial exploratory approach of how environmental factors such as water quality parameters and watershed characteristics are correlated and used to predict FIB concentrations in freshwater streams and rivers. Spearman correlations indicated that weak to moderate significant linear

relationships were evident between both FIB and many hydrologic, water quality and geospatial data. However, this approach only identified individual relationships between FIB and did not account for potential interaction terms as found in a regression analysis. Therefore, a stepwise multiple linear regression was performed on all 26 variables from 5 regions and statewide to develop regression models for FIB. The results from the model indicate that *E. coli* and *Enterococcus* showed the strongest relationship between each associated concentration and explained a majority of the variability in the model for all sites and regions. The addition of the included significant independent variables were found to improve the models in all cases and moderately reduce variability. Regional differences were shown and contributing drainage area was a significant contributor to increased number of variables in each model. Variability in the regression model outputs was further reduced in the regions with smaller watersheds. However, in almost all regions at least 40% of the variability was undetermined. Overall, our models indicated that variability of the model was reduced depending on the region, contributing drainage area, and hydrologic and water quality differences between regions.

More research is needed to understand how we can use environmental parameters to predict and characterize FIB in freshwater streams. The results from our regression models indicate that predictors and uncertainty may be improved by increasing the density of data both spatially and temporally and accounting for specific physical, chemical, and geographic complexities in each region and watershed. As monitoring efforts are often resource-limited, effective monitoring strategies are needed to improve the reconnaissance efforts and target potential impaired waterbodies through collection of basic water quality parameters and geospatial data. This research demonstrates that regression equations can be used to develop initial predictors between FIB and watershed and water quality parameters on a region-based scale using extensive geospatial, hydrologic and water quality datasets.

5.5 References

Akoglu H. (2018). User's guide to correlation coefficients. Turkish journal of emergency medicine, 18(3), 91–93. https://doi.org/10.1016/j.tjem.2018.08.001

Bellégo, C., Benatia, D., & Pape, L.-D. (2022). *Dealing with Logs and Zeros in Regression Models* (SSRN Scholarly Paper No. 3444996). https://doi.org/10.2139/ssrn.3444996

Berger, P. D., Maurer, R. E., & Celli, G. B. (2018). Multiple Linear Regression. In P. D. Berger, R. E. Maurer, & G. B. Celli (Eds.), Experimental Design: With Application in Management, Engineering, and the Sciences. (pp. 505–532). Springer International Publishing. https://doi.org/10.1007/978-3-319-64583-4_15

Brendel, C. and Soupir, M. L. (2017). Relating Watershed Characteristics to Elevated Stream Escherichia coli Levels in Agriculturally Dominated Landscapes: An Iowa Case Study. *Water*, *9*(3), Article 3. https://doi.org/10.3390/w9030154

Brenner, E. K. and Brenner, F. J. (1995). Impact of Precipitation and Drainage Area Size on Water Quality in Stream Systems Supplying a Rural Residential Lake. *Journal of the Pennsylvania Academy of Science*, *69*(3), 88–92.

Chen, H. J. and Chang, H. (2014). Response of discharge, TSS, and E. coli to rainfall events in urban, suburban, and rural watersheds. Environmental Science: Processes & Impacts, 16(10), 2313–2324. https://doi.org/10.1039/C4EM00327F

Christensen, V. G., Rasmussen, P. P., & Ziegler, A. C. (2002). Real-time water quality monitoring and regression analysis to estimate nutrient and bacteria concentrations in Kansas streams. Water Science and Technology, 45(9), 205–219. https://doi.org/10.2166/wst.2002.0240

Devane, M. L., Moriarty, E., Weaver, L., Cookson, A., & Gilpin, B. (2020). Fecal indicator bacteria from environmental sources; strategies for identification to improve water quality monitoring. *Water Research*, *185*, 116204. https://doi.org/10.1016/j.watres.2020.116204

Desai, A. M., & Rifai, H. S. (2010). Variability of *Escherichia coli* Concentrations in an Urban Watershed in Texas. Journal of Environmental Engineering, 136(12), 1347–1359. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000290

Dila, D. K., Corsi, S. R., Lenaker, P. L., Baldwin, A. K., Bootsma, M. J., & McLellan, S. L. (2018). Patterns of Host-Associated Fecal Indicators Driven by Hydrology, Precipitation, and Land Use Attributes in Great Lakes Watersheds. Environmental Science & Technology, 52(20), 11500–11509. https://doi.org/10.1021/acs.est.8b01945

Fisher, D. S., Steiner, J. L., Endale, D. M., Stuedemann, J. A., Schomberg, H. H., Franzluebbers, A. J., & Wilkinson, S. R. (2000). The relationship of land use practices to surface water quality in the Upper Oconee Watershed of Georgia. *Forest Ecology and Management*, *128*(1), 39–48. https://doi.org/10.1016/S0378-1127(99)00270-4

Gauthier, T.D. (2001). Detecting trends using Spearman's rank correlation coefficient. Environmental Forensics, Vol.2, pp. 359-362, ISSN 1527-5930

Ghani, I. M. M., & Ahmad, S. (2010). Stepwise multiple regression method to forecast fish landing. Procedia-Social and Behavioral Sciences, 8, 549-554.

Gholizadeh, M. H., Melesse, A. M., & Reddi, L. (2016). A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. Sensors, 16(8), Article 8. https://doi.org/10.3390/s16081298

Gregory, L. F., Gitter, A., Muela, S., & Wagner, K. L. (2019). Should Contact Recreation Water Quality Standards be Consistent across Hydrological Extremes? Journal of Contemporary Water Research & Education, 166(1), 12–23. https://doi.org/10.1111/j.1936-704X.2019.03298.x

Guo, D., Lintern, A., Webb, J. A., Ryu, D., Bende-Michl, U., Liu, S., & Western, A. W. (2020). A data-based predictive model for spatiotemporal variability in stream water quality. Hydrology and Earth System Sciences, 24(2), 827–847. https://doi.org/10.5194/hess-24-827-2020

Haller, L., Amedegnato, E., Poté, J., & Wildi, W. (2009). Influence of Freshwater Sediment Characteristics on Persistence of Fecal Indicator Bacteria. Water, Air, and Soil Pollution, 203(1), 217–227. https://doi.org/10.1007/s11270-009-0005-0

Hamilton, J. L., and Luffman, I. (2009). Precipitation, Pathogens, and Turbidity Trends in the Little River, Tennessee. Physical Geography, 30(3), 236–248. https://doi.org/10.2747/0272-3646.30.3.236

Harmel, R. D., Hathaway, J. M., Wagner, K. L., Wolfe, J. E., Karthikeyan, R., Francesconi, W., & McCarthy, D. T. (2016). Uncertainty in monitoring *E. coli* concentrations in streams and stormwater runoff. Journal of Hydrology, 534, 524–533. https://doi.org/10.1016/j.jhydrol.2016.01.040

Helsel, D.R., Hirsch, R.M., Ryberg, K.R., Archfield, S.A., and Gilroy, E.J., (2020) Statistical methods in water resources: U.S. Geological Survey Techniques and Methods, book 4, chap. A3, 458 p., https://doi.org/10.3133/tm4a3

Henderson, D. A., & Denison, D. R. (1989). Stepwise Regression in Social and Psychological Research. *Psychological Reports*, *64*(1), 251–257. https://doi.org/10.2466/pr0.1989.64.1.251

Homer, C. G., Dewitz, J. A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N. D., Wickham, J. D., & Megown, K. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States - Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing* 81(5), 345-354. https://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=309950

JASP Team (2023). JASP (Version 0.17) [Computer software].

Kang, J.-H., Lee, S. W., Cho, K. H., Ki, S. J., Cha, S. M., & Kim, J. H. (2010). Linking land-use type and stream water quality using spatial data of fecal indicator bacteria and heavy metals in the Yeongsan river basin. *Water Research*, *44*(14), 4143–4157. https://doi.org/10.1016/j.watres.2010.05.009

Liao, H., Sarver, E., & Krometis, L.-A. H. (2018). Interactive effects of water quality, physical habitat, and watershed anthropogenic activities on stream ecosystem health. *Water Research*, *130*, 69–78. https://doi.org/10.1016/j.watres.2017.11.065

Masoner, J. R., and Mashburn, S. L. (2004). Water quality and possible sources of nitrate in the Cimarron Terrace Aquifer, Oklahoma, 2003 (No. 4). US Department of the Interior, US Geological Survey.

Oklahoma Conservation Commission (OCC). (2020). WQ Assessment Rotating Basin Monitoring Program. Accessed 13 December 2022 https://www.ok.gov/conservation/Agency_Divisions/Water_Quality_Division/WQ_Monitoring/WQ_Assessment_Rotating_Basin_Monitoring_Program.html

Oklahoma Climatological Survey. (2023). Retrieved February 21, 2023, from https://climate.ok.gov/index.php/site/page/climate_of_oklahoma

OWRB (Oklahoma Water Resources Board) (2017). Oklahoma Water Quality Standards. Oklahoma Administrative Code 785:45.

Petersen, F., & Hubbart, J. A. (2020). Advancing Understanding of Land Use and Physicochemical Impacts on Fecal Contamination in Mixed-Land-Use Watersheds. Water, 12(4), 1094. https://doi.org/10.3390/w12041094

Piorkowski, G. S., Jamieson, R. C., Hansen, L. T., Bezanson, G. S., & Yost, C. K. (2014). Characterizing spatial structure of sediment E. coli populations to inform sampling design. *Environmental Monitoring and Assessment*, *186*(1), 277–291. https://doi.org/10.1007/s10661-013-3373-2

R Core Team (2013) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rochelle-Newall, E. J., Ribolzi, O., Viguier, M., Thammahacksa, C., Silvera, N., Latsachack, K., Dinh, R. P., Naporn, P., Sy, H. T., Soulileuth, B., Hmaimum, N., Sisouvanh, P., Robain, H., Janeau, J.-L., Valentin, C., Boithias, L., & Pierret, A. (2016). Effect of land use and hydrological processes on Escherichia coli concentrations in streams of tropical, humid headwater catchments. Scientific Reports, 6(1), Article 1. https://doi.org/10.1038/srep32974

Steyerberg, E. W. (2019). Overfitting and Optimism in Prediction Models. In E. W. Steyerberg (Ed.), Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating (pp. 95–112). Springer International Publishing. https://doi.org/10.1007/978-3-030-16399-0_5

Sokolova, E., Ivarsson, O., Lilliestöm, A., Speicher, N. K., Rydberg, H., & Bondelind, M. (2022). Data-driven models for predicting microbial water quality in the drinking water source using *E. coli* monitoring and hydrometeorological data. Science of The Total Environment, 802, 149798. https://doi.org/10.1016/j.scitotenv.2021.149798

Soranno, P. A., Wagner, T., Martin, S. L., McLean, C., Novitski, L. N., Provence, C. D., & Rober, A. R. (2011). Quantifying regional reference conditions for freshwater ecosystem management: A comparison of approaches and future research needs. Lake and Reservoir Management, 27(2), 138–148. https://doi.org/10.1080/07438141.2011.573614

Stocker, M. D., Penrose, M., & Pachepsky, Y. A. (2018). Spatial Patterns of Escherichia coli Concentrations in Sediment before and after High-Flow Events in a First-Order Creek. *Journal of Environmental Quality*, *47*(5), 958–966. https://doi.org/10.2134/jeq2017.11.0451

Stocker, M. D., Smith, J. E., Hernandez, C., Macarisin, D., & Pachepsky, Y. (2019). Seasonality of E. coli and Enterococci Concentrations in Creek Water, Sediment, and Periphyton. *Water, Air, & Soil Pollution*, *230*(9), 223. https://doi.org/10.1007/s11270-019-4263-1

USEPA (United States Environmental Protection Agency). (2010). Sampling and Consideration of Variability (Temporal and Spatial) for Monitoring of Recreational Waters. EPA-823-R-10-005 US Environmental Protection Agency, Washington, DC, USA.

USEPA (United States Environmental Protection Agency). (2012). Recreational Water Quality Criteria. Office of Water 820-F-12-058. https://www.epa.gov/sites/production/files/2015-10/documents/rwqc2012.pdf

USEPA (United States Environmental Protection Agency). (2018). Success Stories about Restoring Water Bodies Impaired by Nonpoint Source Pollution. Website: https://www.epa.gov/nps/success-stories-about-restoring-water-bodies-impaired-nonpoint-source-pollution

U.S. Geological Survey (USGS). (2019). The StreamStats program. StreamStats v5.3.05 Batch Processing Tool. Online at https://streamstatsags.cr.usgs.gov/ss_bp/, accessed on December 15, 2022.

Verhougstraete, M. P., Martin, S. L., Kendall, A. D., Hyndman, D. W., & Rose, J. B. (2015). Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. Proceedings of the National Academy of Sciences, 112(33), 10419–10424. https://doi.org/10.1073/pnas.1415836112

Vidon, P., Tedesco, L. P., Wilson, J., Campbell, M. A., Casey, L. R., & Gray, M. (2008). Direct and Indirect Hydrological Controls on *E. coli* Concentration and Loading in Midwestern Streams. Journal of Environmental Quality, 37(5), 1761–1768. https://doi.org/10.2134/jeq2007.0311

Vitro, K. A., BenDor, T. K., Jordanova, T. V., & Miles, B. (2017). A geospatial analysis of land use and stormwater management on fecal coliform contamination in North Carolina streams. Science of The Total Environment, 603–604, 709–727. https://doi.org/10.1016/j.scitotenv.2017.02.093

Yang, X., Liu, Q., Luo, X., & Zheng, Z. (2017). Spatial Regression and Prediction of Water Quality in a Watershed with Complex Pollution Sources. *Scientific Reports*, *7*, 8318. https://doi.org/10.1038/s41598-017-08254-w

Zhang, X., Zhi, X., Chen, L., & Shen, Z. (2020). Spatiotemporal variability and key influencing factors of river fecal coliform within a typical complex watershed. Water Research, 178, 115835. https://doi.org/10.1016/j.watres.2020.115835

Zubizarreta, M. E. (2018). Environmental Impacts on Enterococcus in Shem Creek, South Carolina, and Characterization of Changing Land Use. Theses and Dissertations. https://scholarcommons.sc.edu/etd/4501

Chapter 6 : Conclusions and Future Directions

Recreational water quality is vital to not only human health but also for high-quality tourism experiences at Oklahoma streams, rivers and lakes. Therefore, a proper assessment of the water quality, specifically for fecal indicator bacteria (FIB), of these waters is important to understand potential pathogenic contact with humans. Since the conception of the Clean Water Act, many different sectors have worked to monitor, assess and improve our surface water quality for beneficial uses. Point source, or end-of-pipe discharges, monitoring and regulatory requirements have been successful in improving water quality for many waterbodies across the United States. However, non-point sources remain a challenge for water resources stakeholders who must invest in high-dollar resources such as large spatial and temporal routine monitoring networks to begin answering water quality assessment questions.

While conventional approaches for FIB have been successful for improving our understanding of potential non-point and point source pathogens, more research is needed to evaluate the processes that we are using to gauge human health risk. Resources not only are needed for data collection and analysis, but also for remediating and improving waterbodies. Therefore, strategic plans from reconnaissance, research and regulation of beneficial use categorical requirements must be developed. The goal of these efforts is to ensure that our resources are allocated and invested properly to enhance our understanding and stretch our resource boundaries to improve impaired waterbodies. The purpose of this dissertation was to investigate new research methods in the laboratory, field and spatial statistics with the overall goal of improving our approach to assess, mitigate and enhance the quality of our freshwater streams and rivers.

A dissertation research outcome summary is provided below in Chapter 6.1. Research points where my hypotheses (a, b, c, and d) from Chapter 1.2 were supported are denoted by letters in parentheses at the end of each point in the numerical list.

6.1 Research Summary

1) Sample location and time can influence *E. coli* concentrations in streams and rivers and variability in FIB concentrations can occur both longitudinally and laterally in a stream channel. **(a)**
2) Monitoring approaches should consider sampling location, time and other environmental conditions when sampling for fecal indicator bacteria. **(a)**
3) Stream sediment can influence *Enterococcus* concentrations in streams and rivers by providing a consistent and stable source in both laboratory (stable) and field (dynamic) conditions. **(b)**
4) Contributing drainage area, stream substrate properties, and watershed characteristics can be used to understand and predict potential FIB concentrations in stream sediments. **(d)**
5) Relationships between *Enterococcus* in sediment and water samples were indicated when grouping streams by streambed characteristics (particle size, organic matter), drainage area and impervious area. **(d)**
6) Enterolert™ (ELT), which is used for enumeration of *Enterococcus*, was found to exhibit false positives, predominantly *Paenibacillus* spp., from stream sediment and water samples. **(c)**
7) ELT concentrations remained stable throughout field and microcosm studies, however in almost all cases, the isolates identified rapidly shifted away from FIB to non-lactic acid bacteria (non-FIB) during the study periods. **(b) (c)**

8) ELT may require additional research to evaluate the false positives and how to improve selectivity of the media for FIB in freshwater streams. **(c)**
9) Hydrologic, geographic, and water quality variables from large monitoring datasets can be used to predict fecal indicator bacteria concentrations by region to make initial assessments of important factors to consider when monitoring FIB in streams. **(d)**
10) Begin the conversation for re-thinking how we assess and monitor freshwater streams. Specifically, by 1) addressing spatial and temporal sampling factors, 2) appropriateness of *Enterococcus* as a freshwater indicator and accuracy of enumeration methods, 3) understanding the influences of sediment, water quality, hydrologic and geographic factors on FIB and 4) using available datasets and geospatial information to understand and predict potential site factors that can influence FIB in Oklahoma streams.

6.2 Recommendations for Future Research

The impetus for this research stems from the research questions posed by multiple entities that indicate historical methods and approaches of monitoring and assessing streams in general do not account for the how, when, where and why we are sampling a waterbody, specifically for microbiological indicators. Monitoring streams for regulatory purposes for FIB should require that samples are 1) representative of time and space, 2) accurately quantified and collected and 3) provide meaningful results to properly assessment. In our water quality standards, we are often given ambiguous terminology for how to best approach monitoring and assessment with the only requirement is a minimum number of samples over a certain time period. However, this criteria does not go into detail of standardizing or understanding the potential environmental factors, sampling protocols, and/or enumeration methods for freshwater streams that could influence bacteria concentrations, and often applies criteria that was developed from a series of beach

studies on reservoirs and oceans by the USEPA in the 1980s. Therefore, we need additional research to understand if we are hitting our targets for properly characterizing and assessing freshwater streams for associated human health risks with FIB.

For example, when I have discussed *Enterococcus* sampling with invested sampling entities in Oklahoma, they indicate that almost any waterbody that is sampled will most likely be impaired based on the primary body contact recreation water quality standards. However, while many have alluded to this being an issue, limited research is available to understand why this presumed phenomenon is occurring. As a first step as a follow-up from my work, an important research question would be to continue understanding the relationship between *Enterococcus* in the sediment and water and associated interactions. My research gave an initial glimpse of what may be occurring within freshwater streams and further research could validate the hypothesis of stream sediments as a stable source of *Enterococcus* that is often of environmental or animal origin. Second, further investigating the ELT method for suppression of non-FIB is an important consideration for future research. If we continually use this method to make assessments of our freshwater streams, lakes, and rivers, we may be biasing our results with a high percentage of false positives or other interferences. This research may need to come in the form of contacting the manufacturer of the product for potential consideration of our results and how we can improve our accuracy of results for these types of samples. Lastly, additional research is needed to correlate and relate *Enterococcus* concentrations in streams with environmental factors as this type of analysis could prove vital given the limited resources available to routinely monitor and enumerate bacteria samples.

As provided in Chapter 5, we can also use readily available geospatial, hydrologic and water quality datasets to begin developing predictors and relationships between FIB to improve

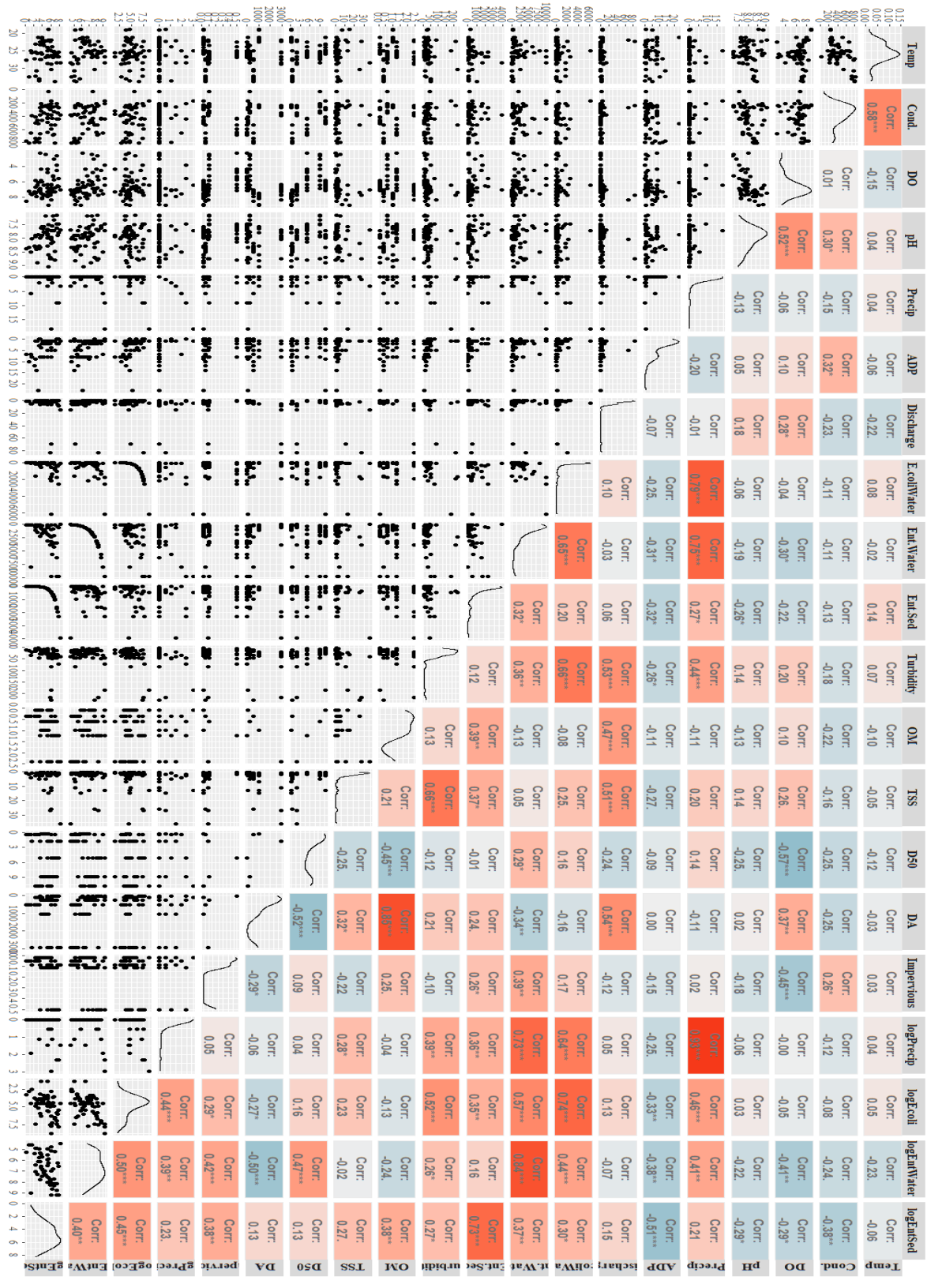
our assessment protocols. While there are many underlying assumptions with these approaches, they also provide us an easier look at what may be occurring in stream reaches where we are not able to collect samples due to budget restraints, access, or other factors. Therefore, continuing to routinely collect water quality and hydrology data along with FIB is critical to improving the models. Additionally, a multiple watershed analysis with dense sampling and monitoring would be an ideal project for refining the models and attempting to answer the unknown variability of the initial models that were developed. I think it would be interesting to compare a 10-week study such as from Chapter 3 where we had a continual 10 weeks of data versus Chapter 5 where over 10 years of data from many different points and time but complete this type of analysis on a paired set of locations and/or watersheds with a focus on temporal and spatial metrics along with hydrologic factors. The outputs from this type of analysis might lead us towards explaining some of the variability of the models and allow us to hone our monitoring and assessment approaches.

When I began this research, my initial goal was to develop a path forward for how we can develop new strategies and approaches for assessing freshwater streams for fecal indicator bacteria to ensure that we are accurately and efficiently making the optimal decisions for protecting and enhancing water quality of our rivers and streams. Referring back to Figure 1-2, I think that my research has just scratched the surface and started the conversation for research questions 1-4 in the upper right corner of the process diagram. The outcome of this work created more important questions to lead us to the ultimate goal of sampling process improvements. I hope the next phase of this research includes further evaluation of many of these decision-making tools that we use to provide regulatory water quality designations. In conclusion, the need for future research in this topic area is critical as we are faced with many economic, environmental and climate challenges to ensure that we are effectively monitoring and protecting our invaluable water resources.

Appendices

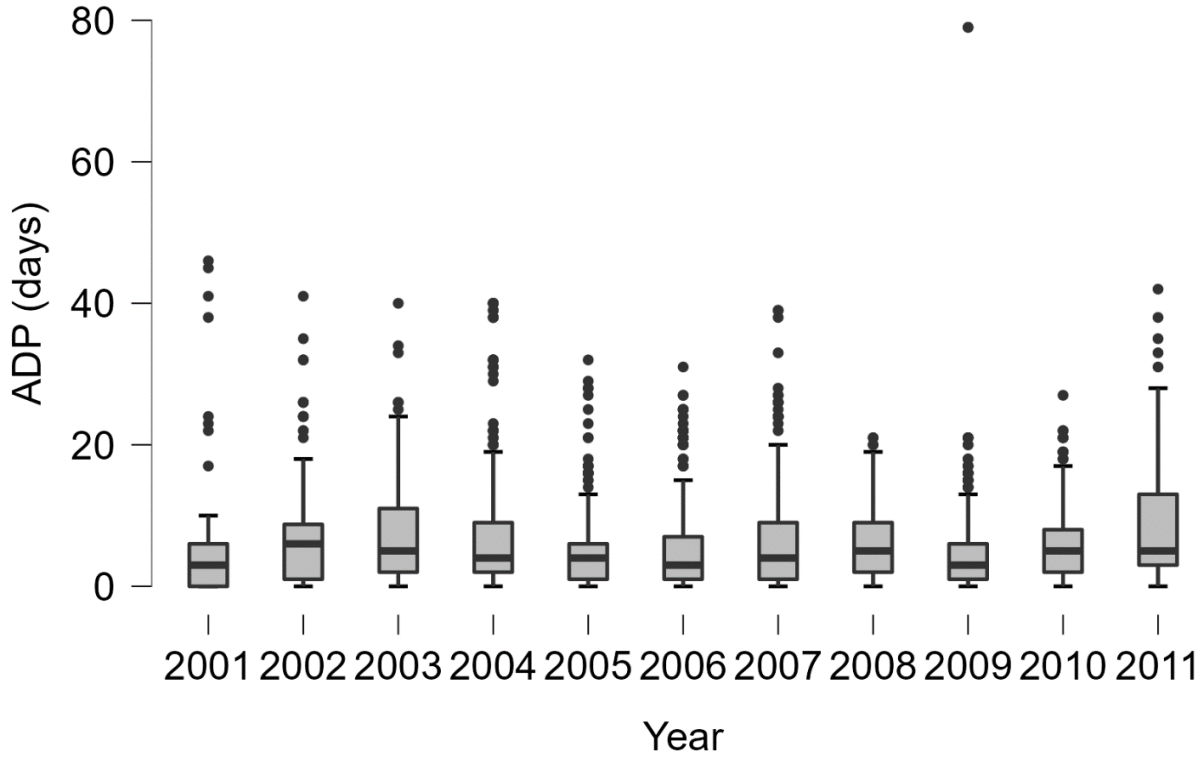
The appendices are organized by items as they appear in descending chapters throughout the dissertation. The figures, tables and code here were not directly included in the publications but were important for developing the final results and conclusions of chapters and provides additional context for data interpretations.

Appendix 1. Correlogram of environmental factors analyzed in Chapter 3.

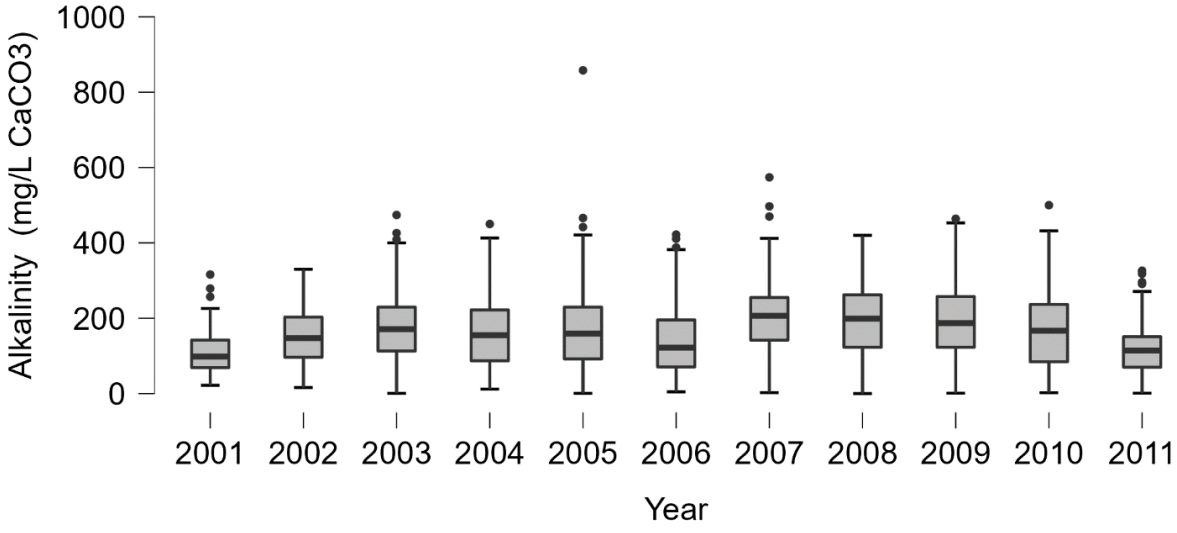


Appendix 2. Boxplots of water quality and hydrologic variable data used in the Chapter 5 analyses.

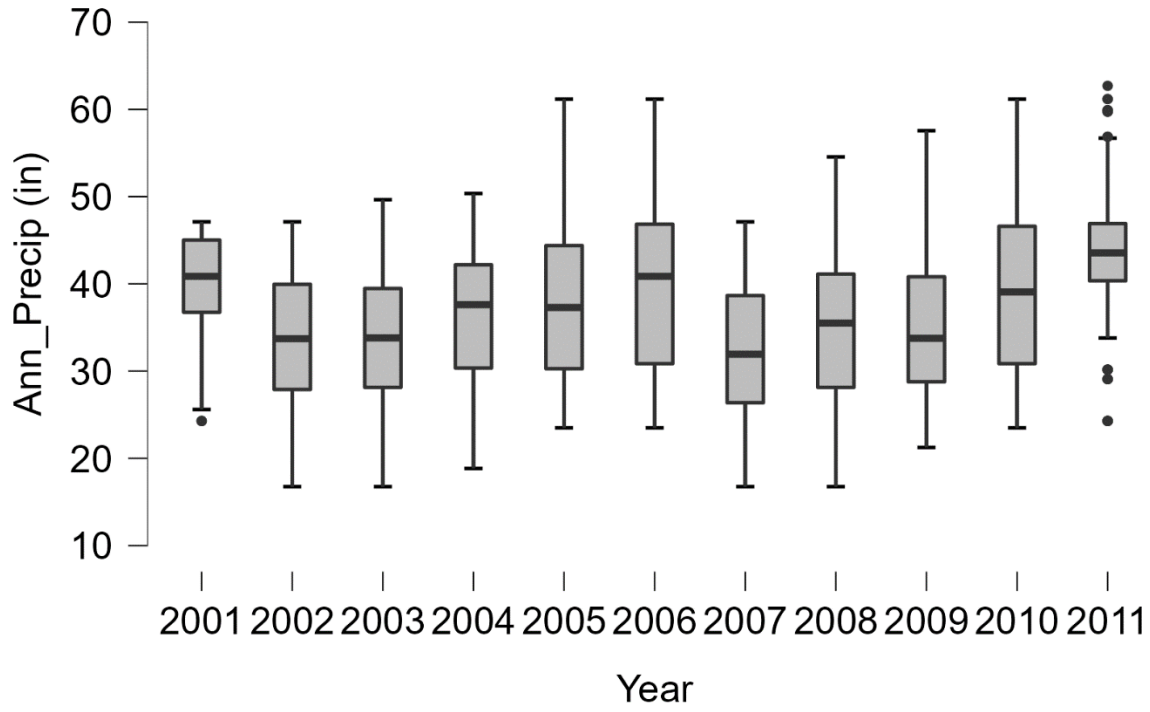
ADP



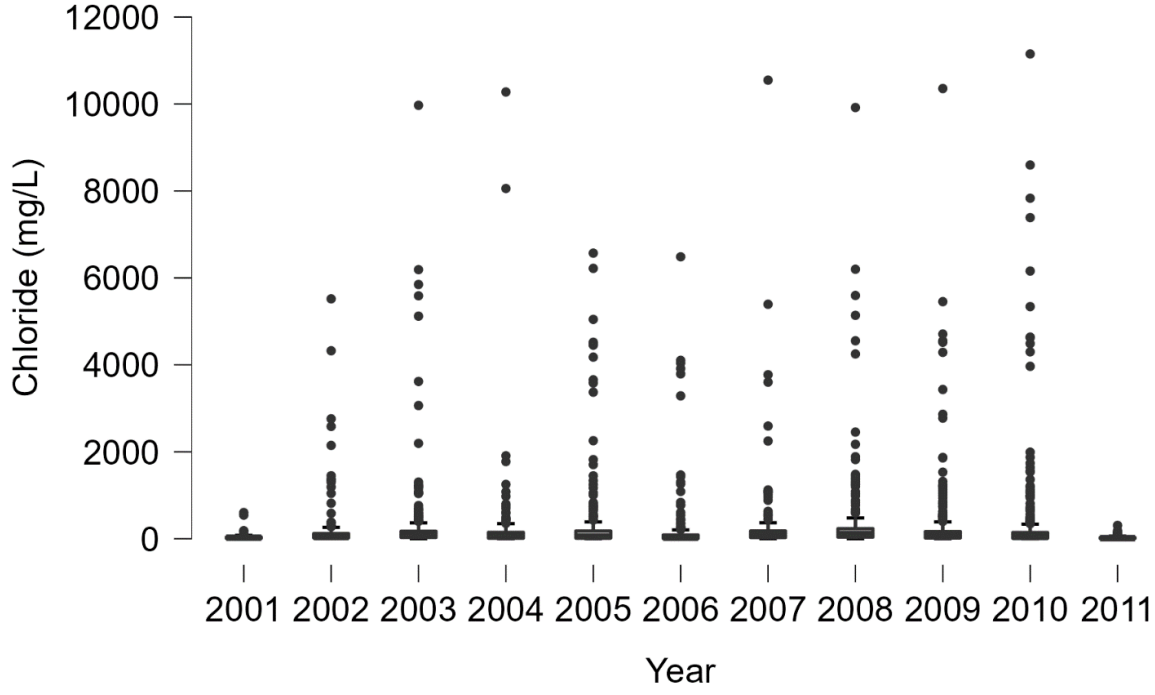
ALK



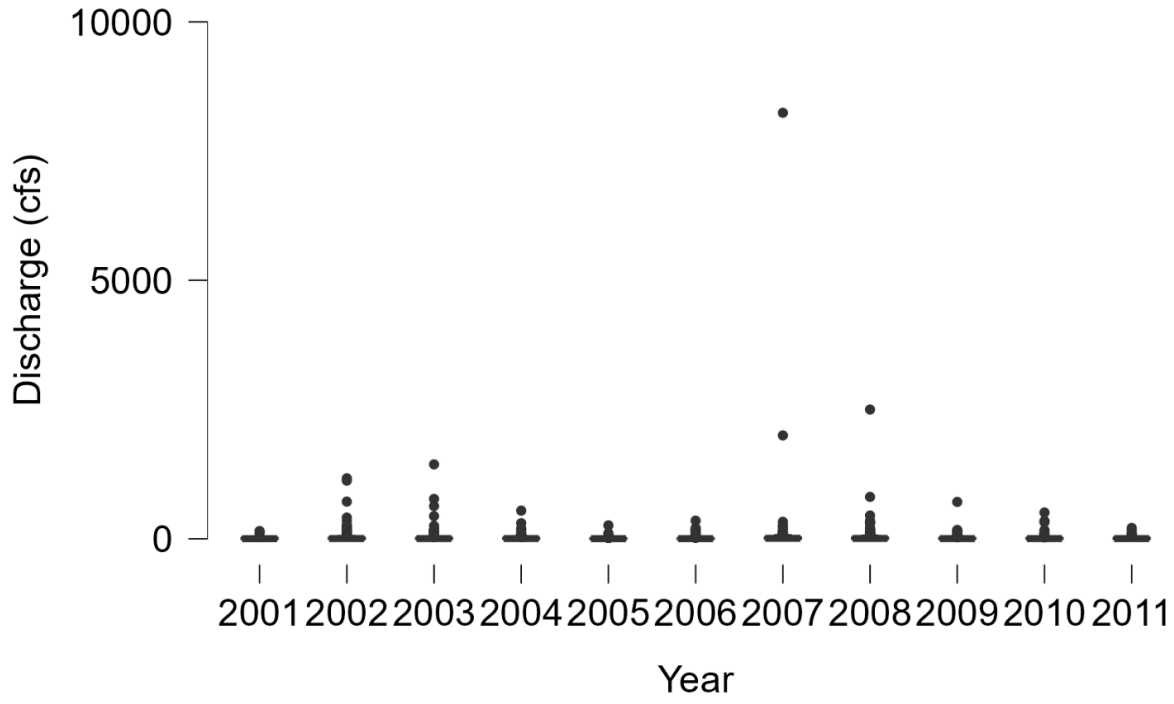
Ann_Precip



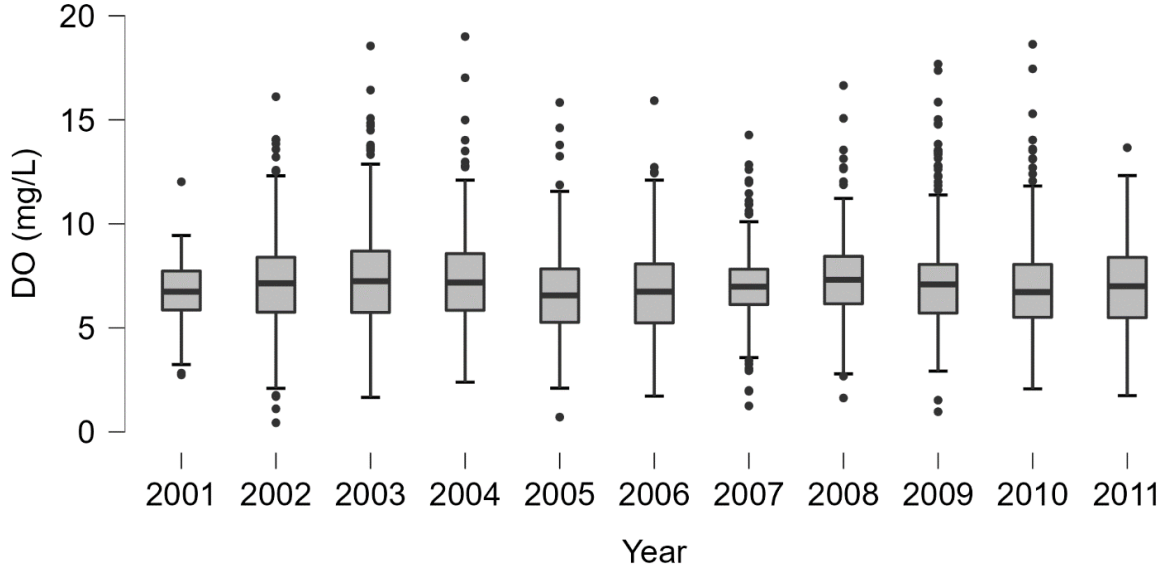
Chloride



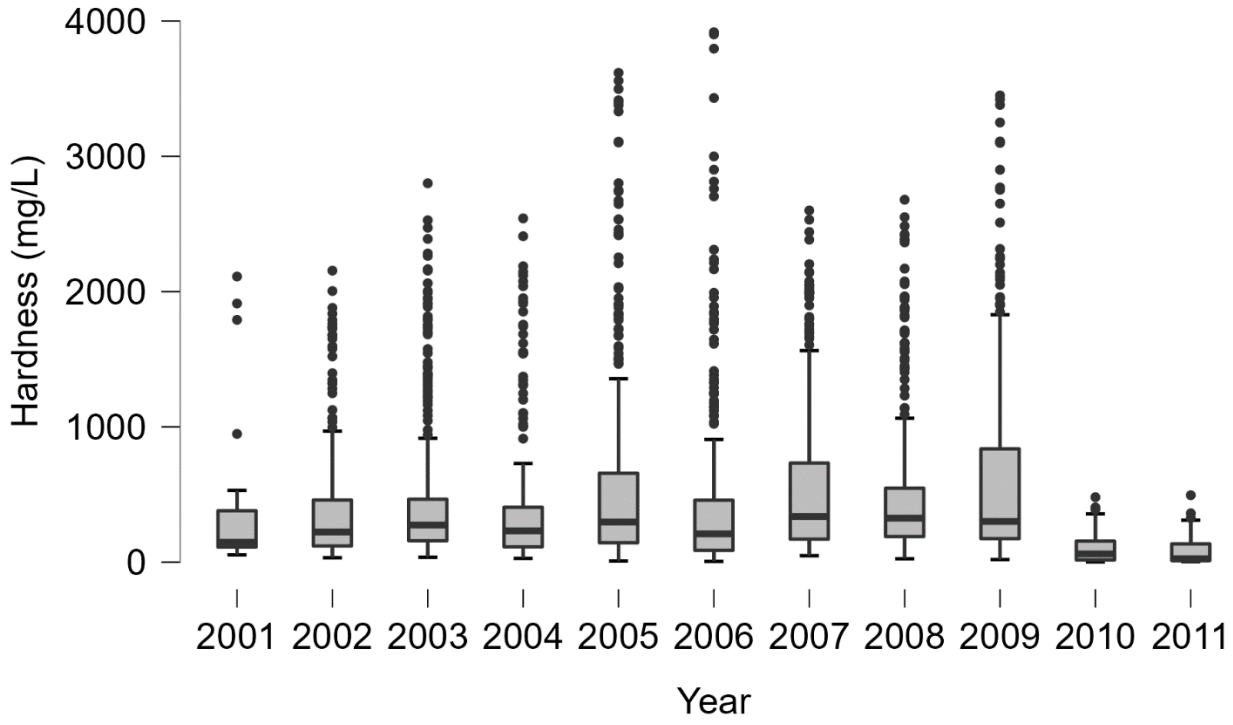
Discharge



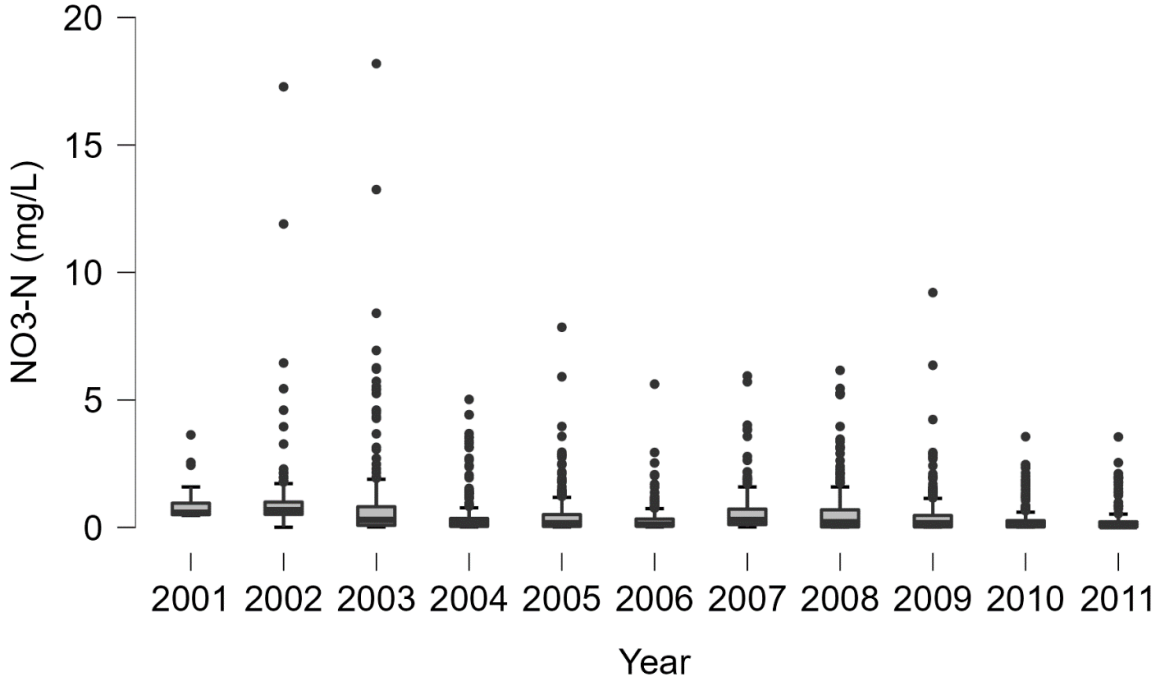
DO



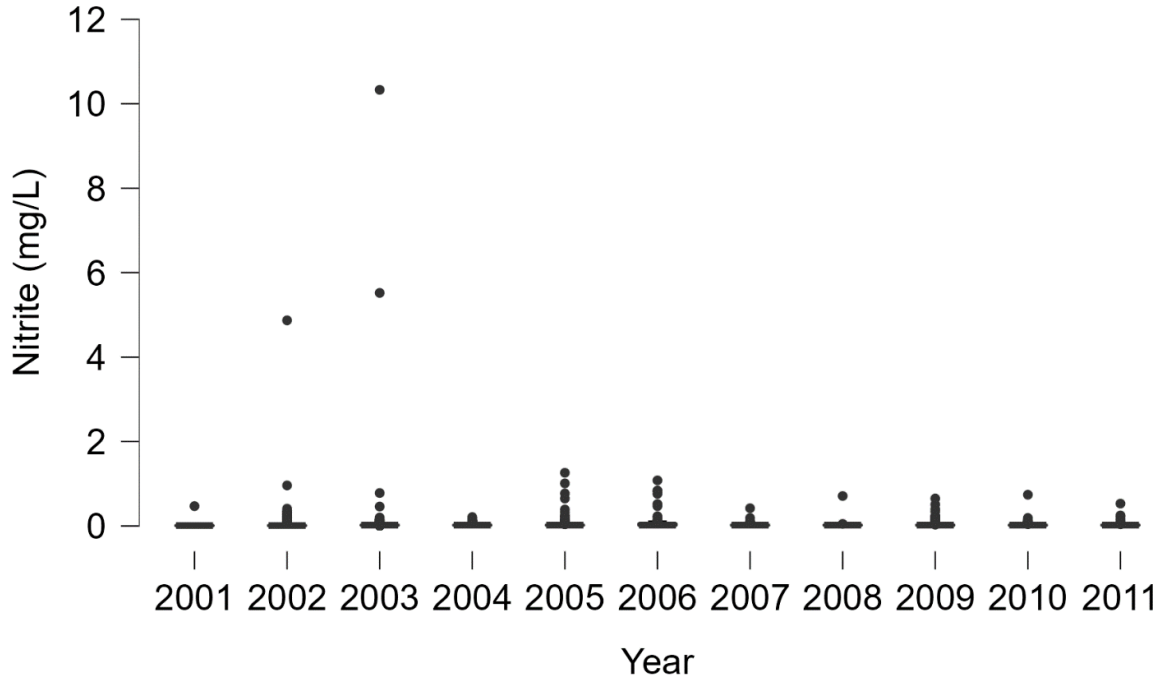
Hardness



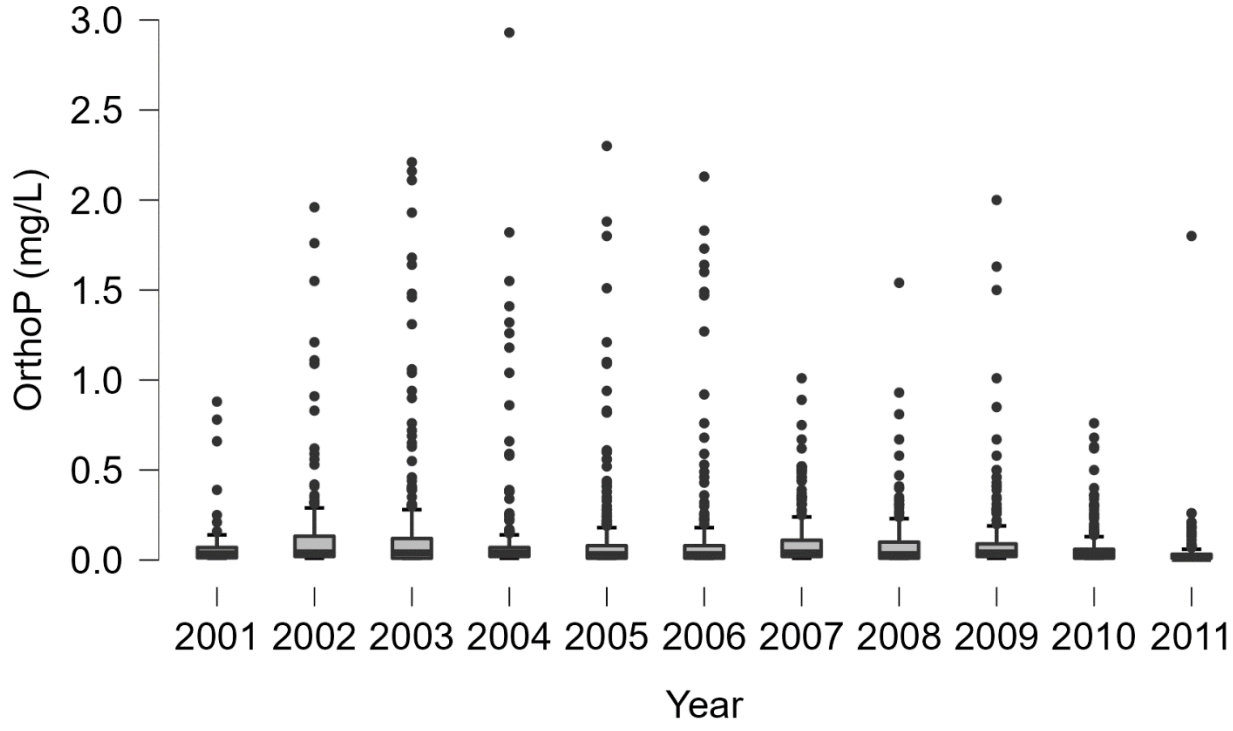
NO3.N



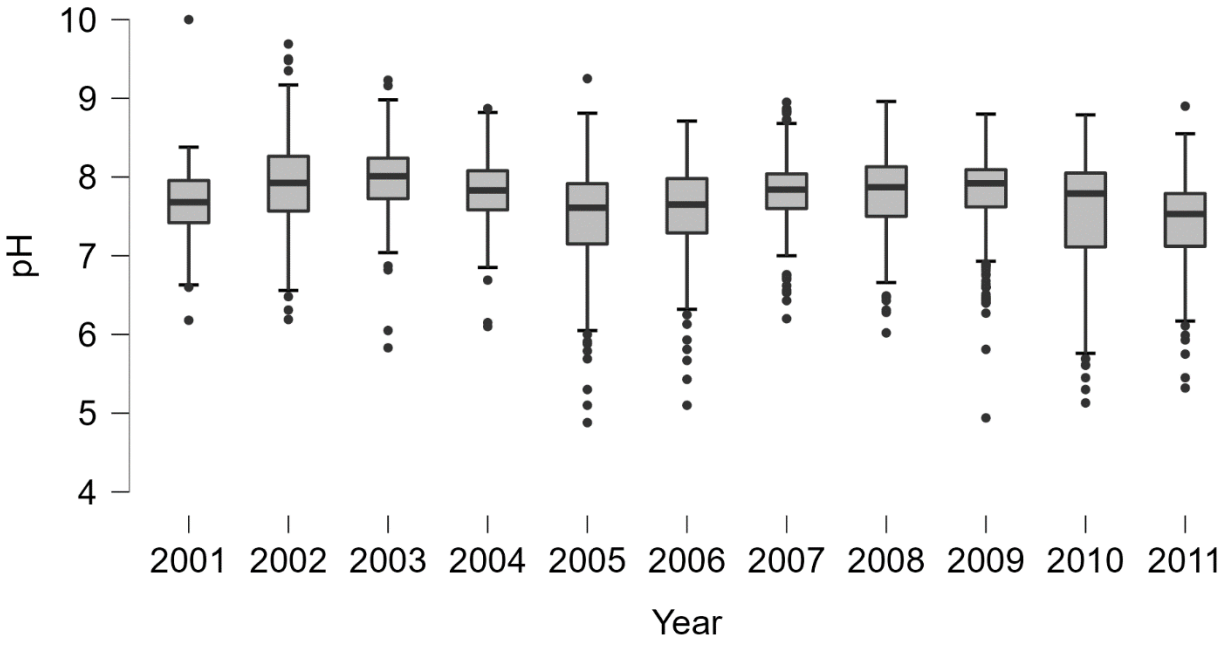
Nitrite



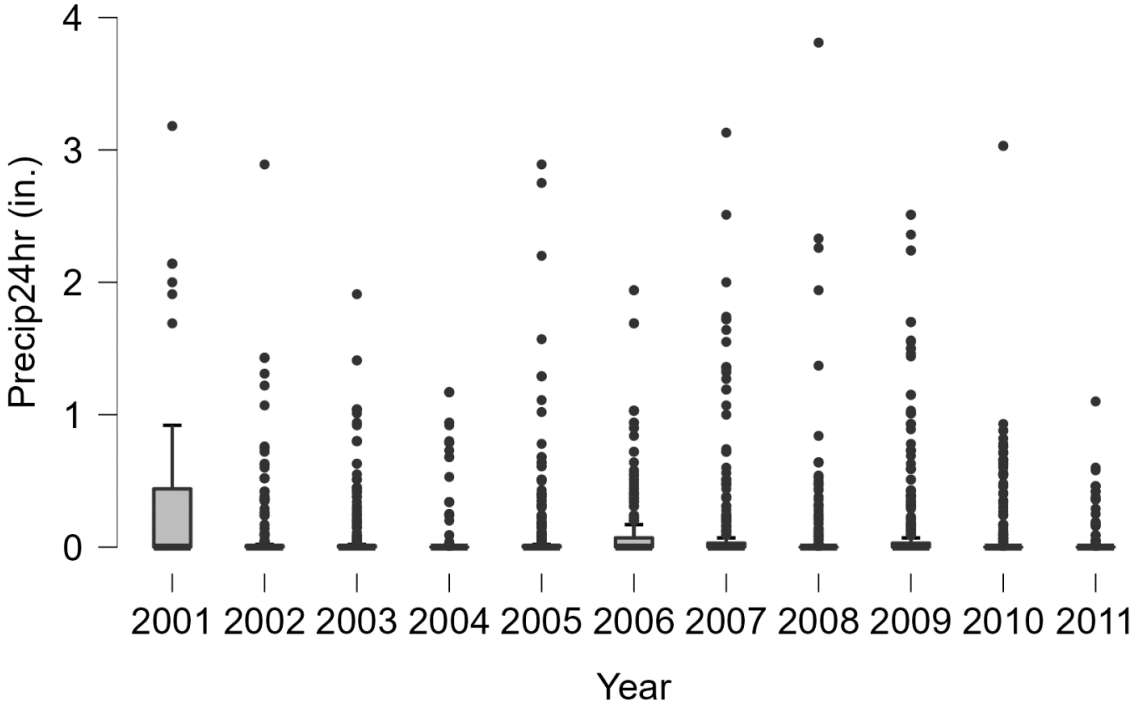
OrthoP



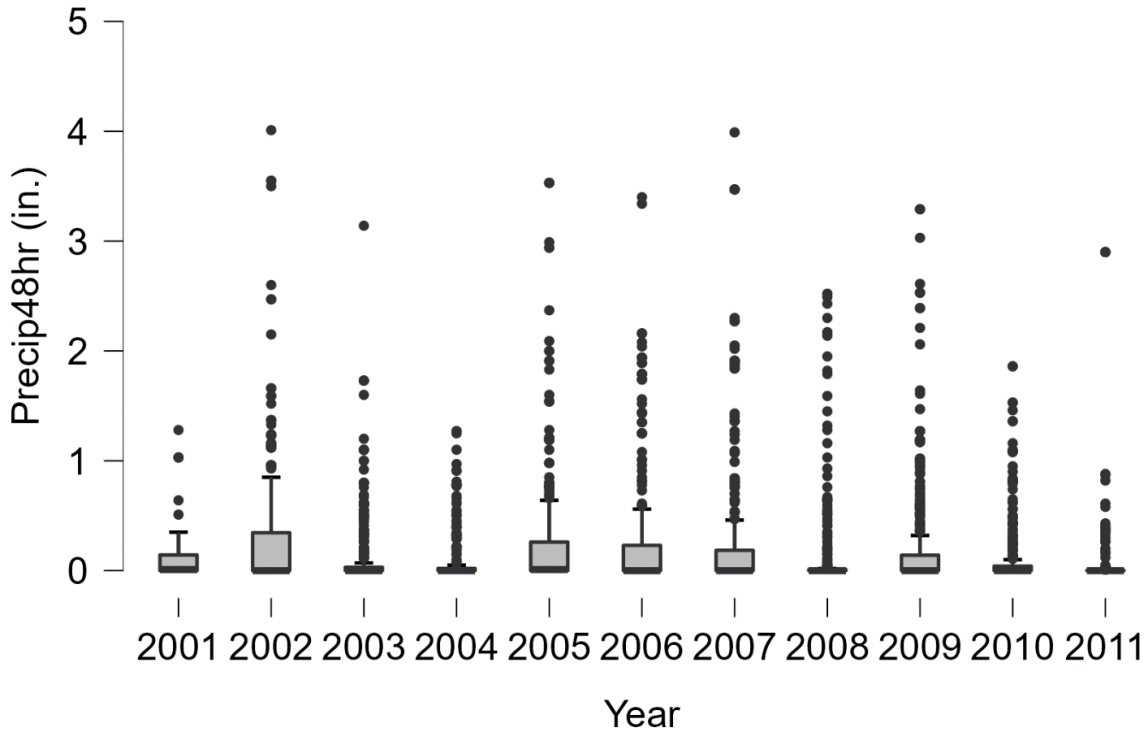
pH



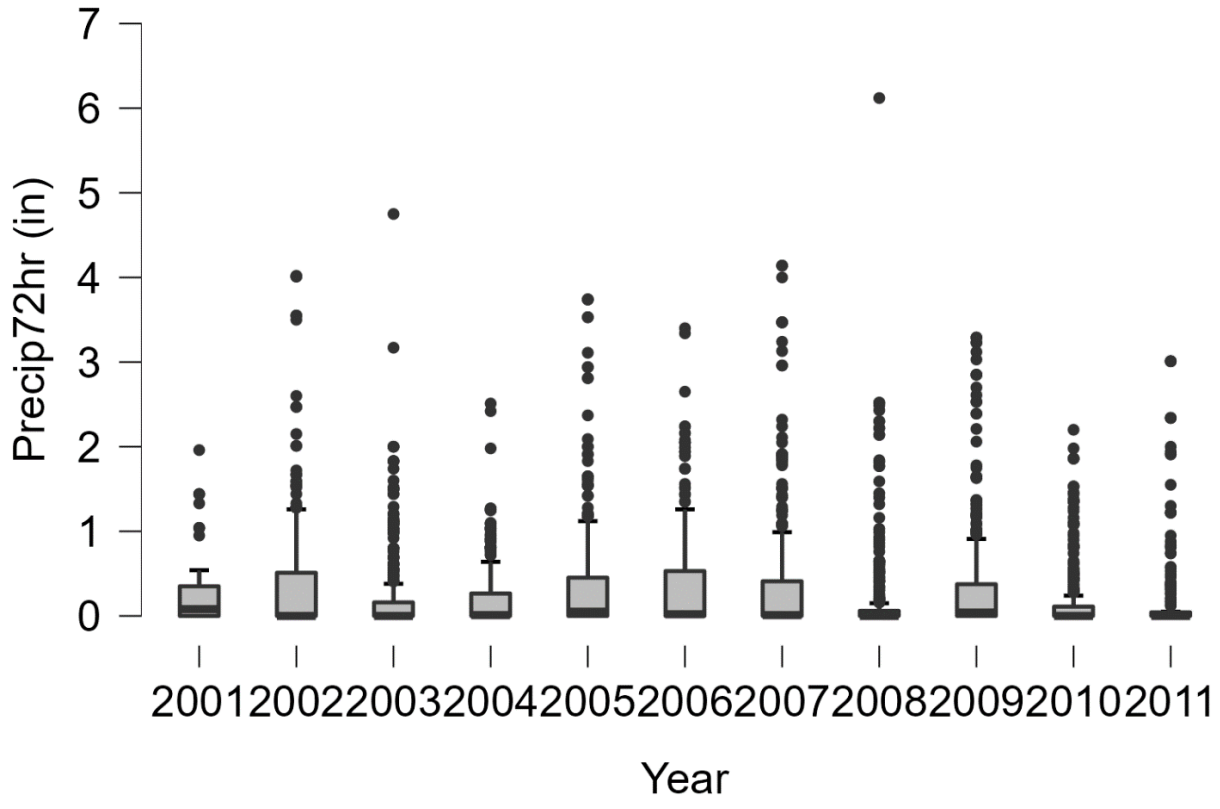
Precip24hr



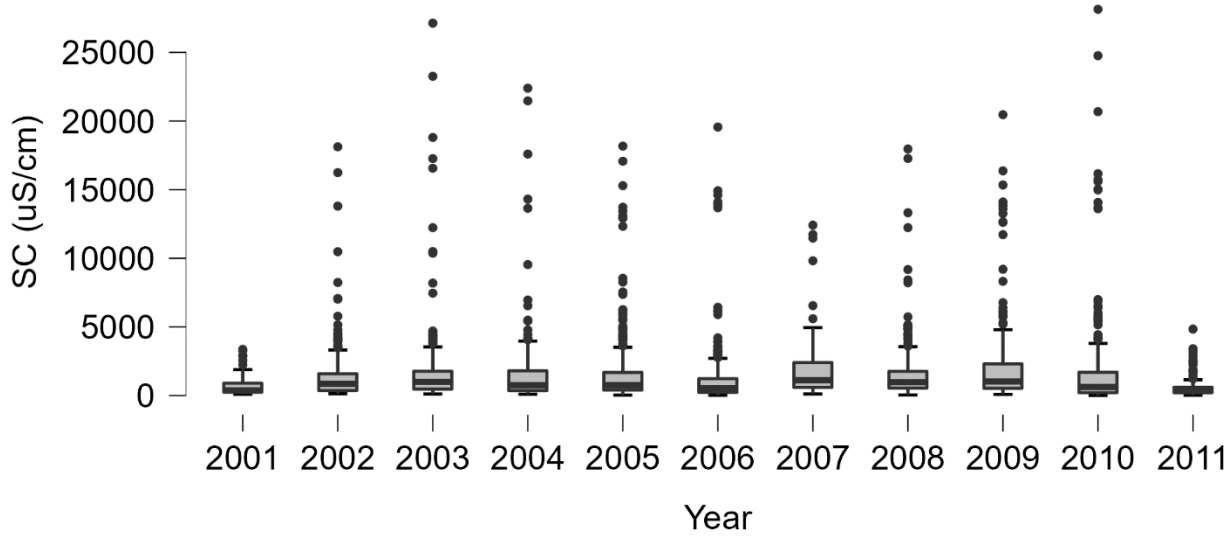
Precip48hr



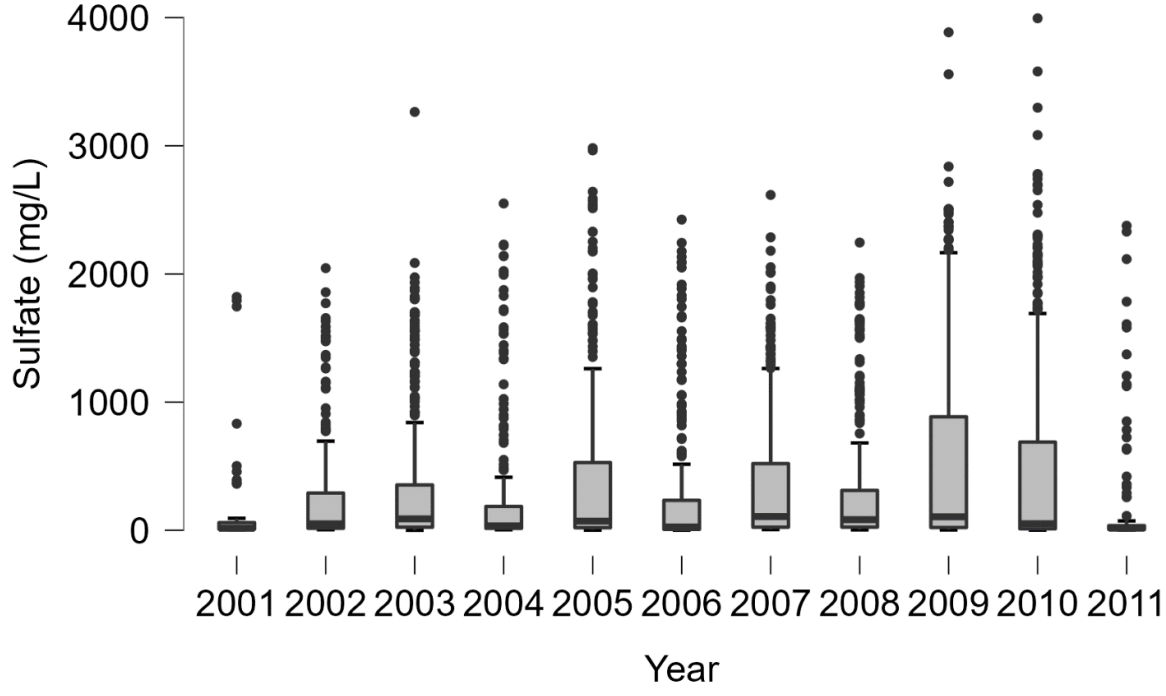
Precip72hr



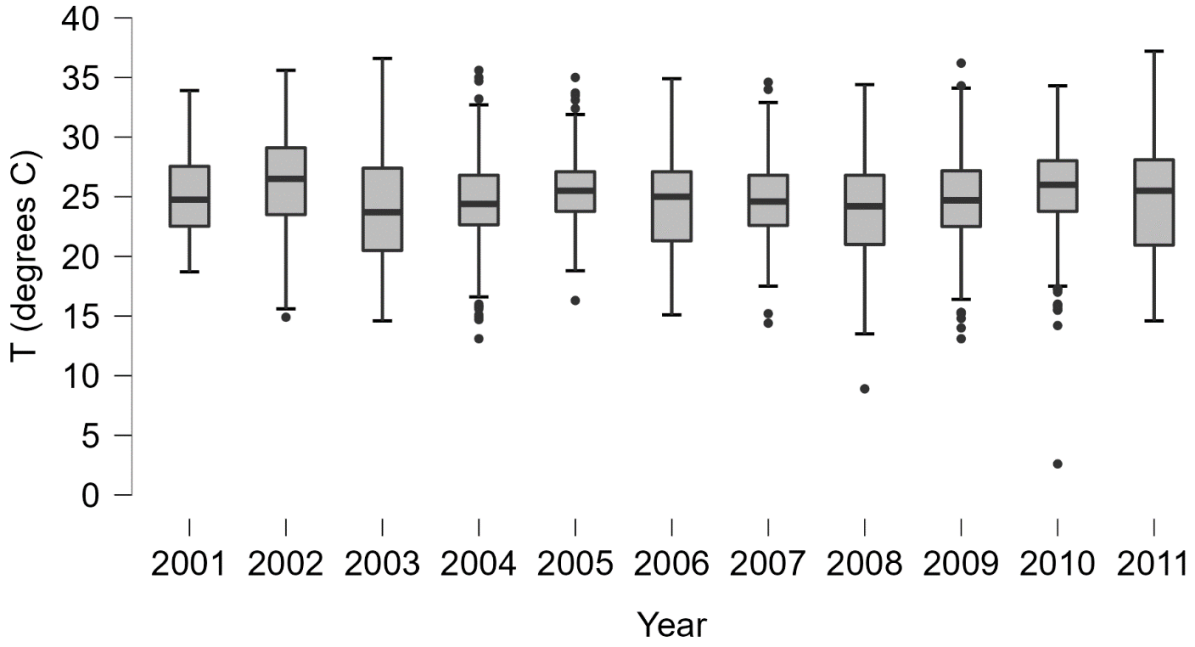
SC


SC (uS/cm)
25000
20000
15000
10000
5000
0
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
Year

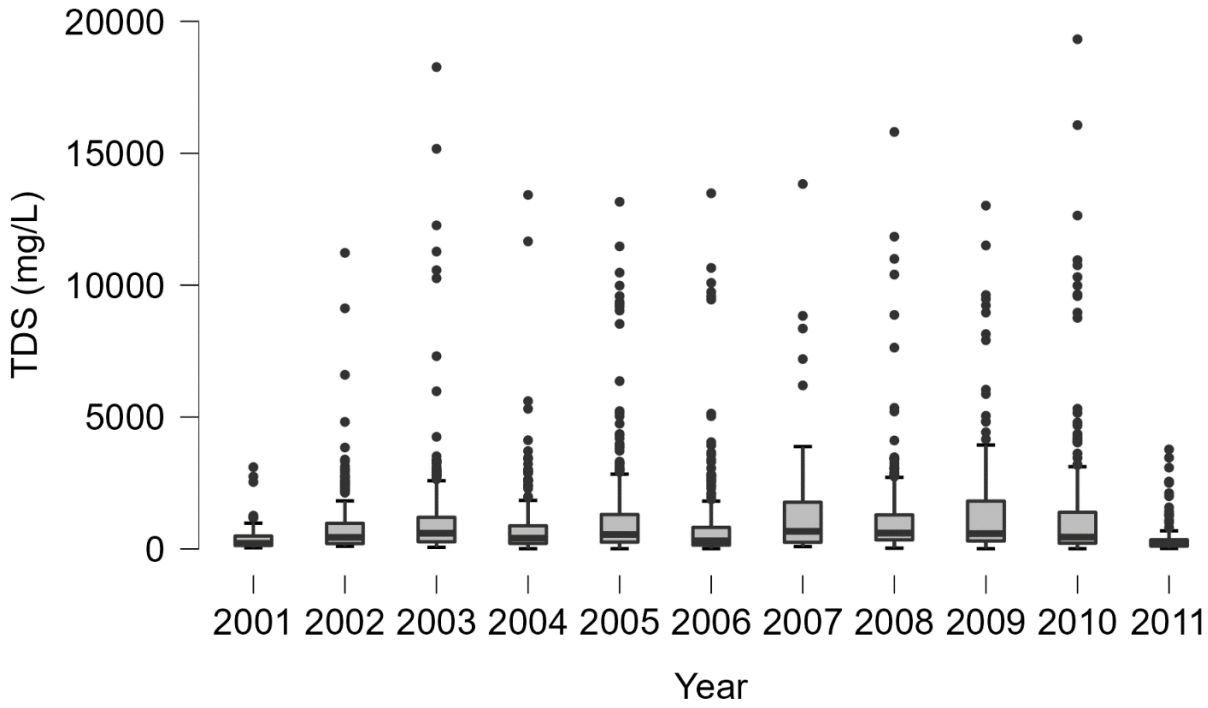
Sulfate


Sulfate (mg/L)
4000
3000
2000
1000
0
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
Year

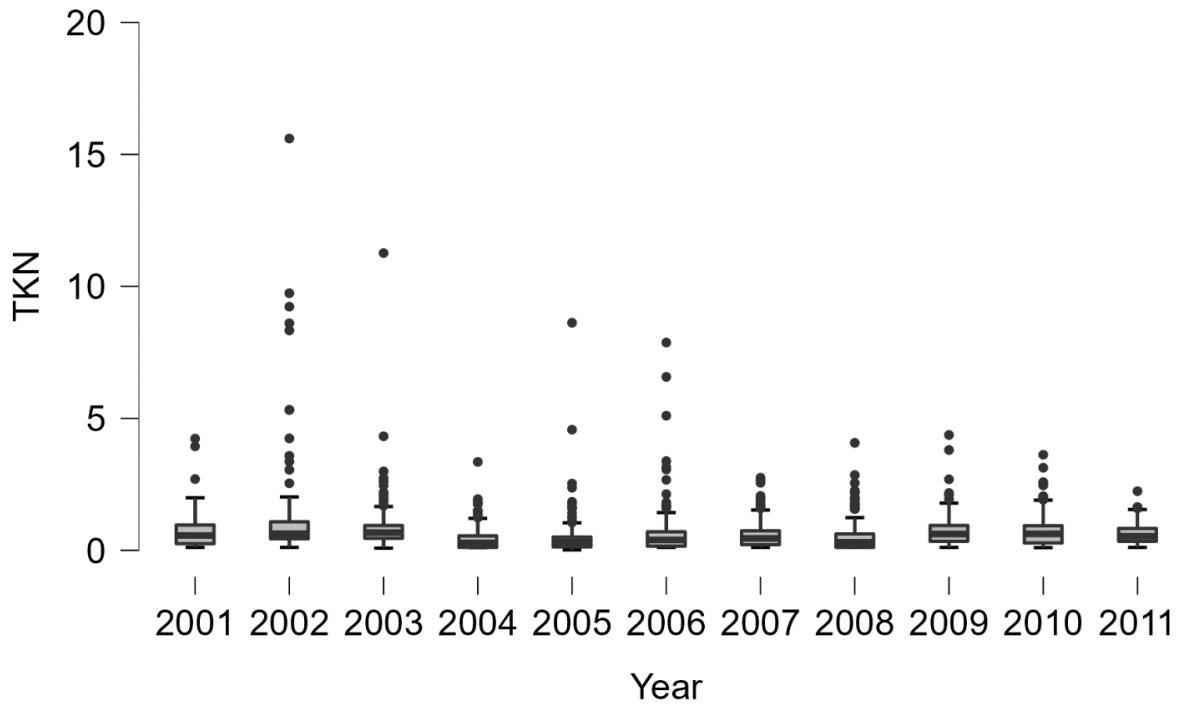
Water Temperature (T)



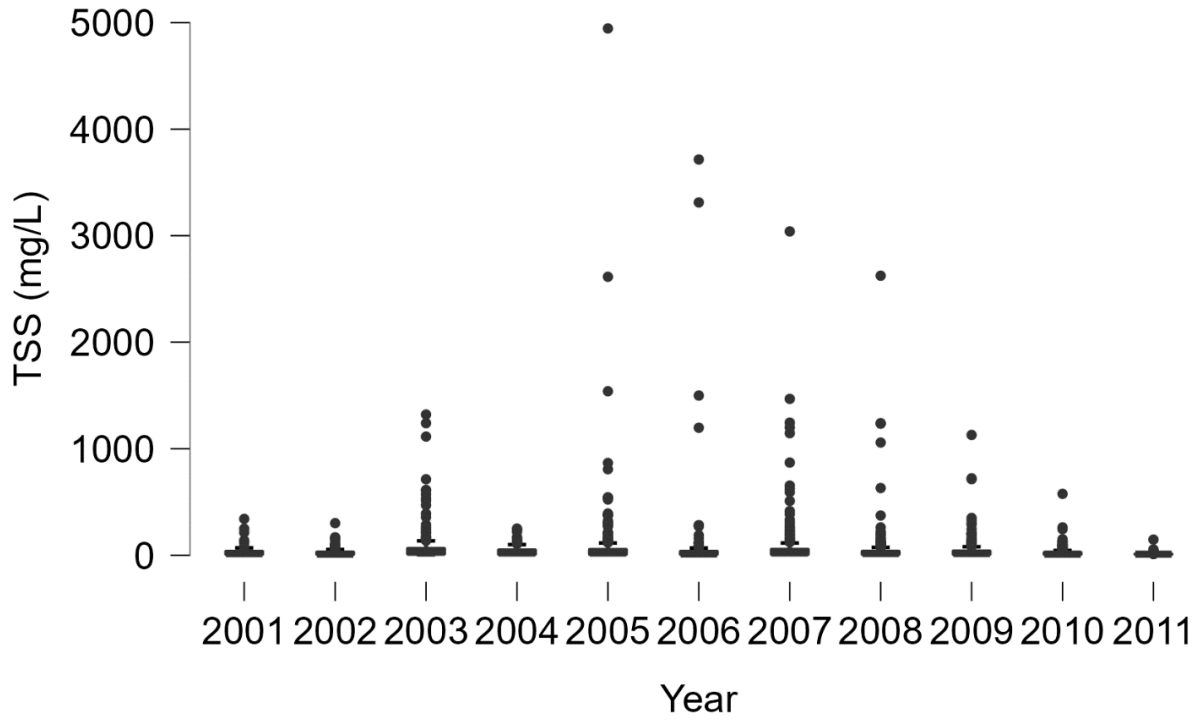
TDS



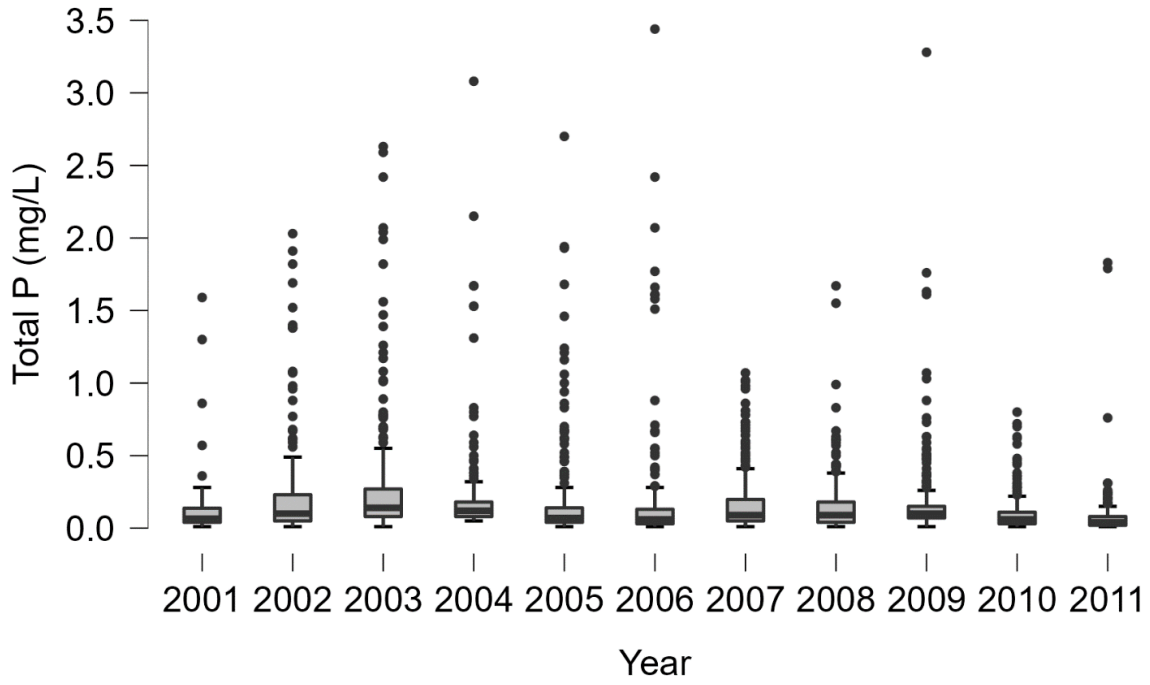
TKN


20
15
10
5
0
TKN
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
Year

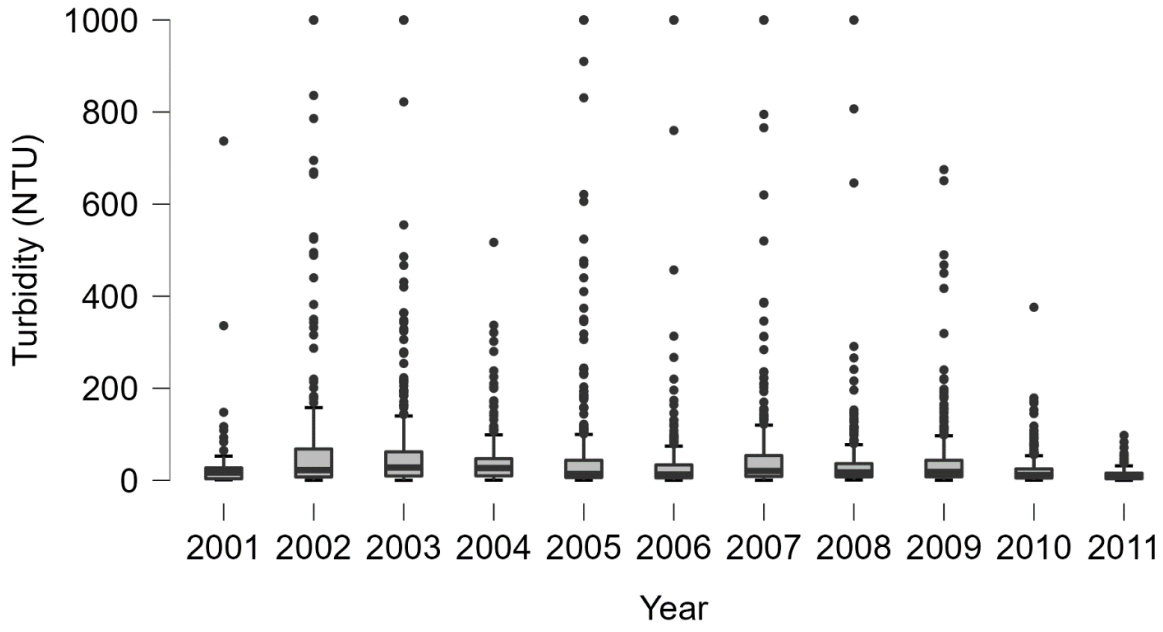
TSS


5000
4000
3000
2000
1000
0
TSS (mg/L)
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
Year

TP


Total P (mg/L)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
Year

Turb.


Turbidity (NTU)
1000
800
600
400
200
0
2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
Year

Appendix 3. Calculating preceding rainfall and antecedent dry periods from Mesonet data

Antecedent dry periods and preceding rainfall days are often used as hydrologic indicators for many different types of analyses in environmental sciences and engineering. In this dissertation they are used as variables in Chapter 3 and 5, which in both analyses, they were important factors for relating fecal indicator bacteria concentrations in streams. While these data are very useful, it is often difficult to obtain or calculate this data as most calculations need to be performed manually. This can be an arduous task, especially when working with many years of data and multiple locations. Therefore, I created a custom R-script that incorporates a previously developed R package 'okmesonet' that retrieves five-minute Mesonet data from the data file server (https://rdrr.io/cran/okmesonet/man/okmts.html). From there, the user can input start and end times, variables to retrieve, and site locations. The script will then select the daily rainfall data for each site from the "0000-UTC" location where the daily total is stored, remove missing values and flagging any error codes (https://www.mesonet.org/index.php/site/about/mdf_mts_files). Additionally, when data is retrieved from the MTS files and data is missing from the beginning or end, it will often include calculate the antecedent dry period, 48-hr preceding rainfall and 72-hr preceding rainfall and save it to a CSV file where the user specifies.

The goal of this R-script is to provide Oklahoma Mesonet data users with a simple tool that can assist with performing advanced data calculations of large precipitation (or other parameters) datasets. However, there are limitations and user inputs that require manual interpretation when evaluating large datasets that may have erroneous values due to a number of factors such as instrument malfunction, frozen precipitation, freezing weather, and extreme events. I suggest that users sort the data by the flagged errors presented and interpret these

intervals before proceeding with calculating antecedent dry days or preceding rainfall. Otherwise, the data has the potential to be highly skewed and not accurate for the data represented.

When I began to research the best way to calculate these precipitation metrics, I quickly discovered there was not a direct solution that I could currently use. I hope that this tool is useful for others in the hydrologic sciences and that the code here can be used for other research projects. Additionally, the next steps would be to develop and integrate this tool into a web-based interface that the typical web user could use to access these calculations.

The R code is as follows:

```
####The purpose of this tool/code is to retrieve Oklahoma Mesonet
###data using the 'okmesonet' package
##and calculating the daily antecedent dry period
##and preceding rainfall (48 and 72 hour)

###Written by Grant Graves, grant.graves@outlook.com
###Last modified on Feb 6, 2023.


#install packages
if (!require('lubridate'))
  install.packages('lubridate')
library('lubridate')
if
(!require('tcltk'))
  install.packages('tcltk')
library('tcltk')
if (!require('dplyr'))
  install.packages('dplyr')
library('dplyr')
install.packages(
  "https://cran.r-project.org/src/contrib/Archive/okmesonet/
okmesonet_0.1.5.tar.gz",
  repos = NULL,
  method = "libcurl"
)
library('okmesonet')

okstations <- updatestn()
```

```
##create function for user inputs to retrieve data
runMesonet <- function() {
  Work_Dir <<-
    setwd(tk_choose.dir(getwd(), "Select where to put files..."))
  ## opens a dialog window
  print(Work_Dir)
  savename <<-
    readline(prompt = "Name of output file saved to directory: ")

  start.date <<- readline(prompt = "Start Date (YYYY-MM-DD): ")

  print(start.date)
  end.date <<- readline(prompt = "End Date (YYYY-MM-DD): ")

  print(end.date)

  ###define stations and variables
  ##to get 4 letter ID use: View(okstations)
  stations <<- unlist(strsplit(readline(prompt = "Stations: "), ", "))

  print(stations)
  variables <<-
    unlist(strsplit(readline(prompt = "Variables: "), ", "))

  print(variables)

}

##run function to begin user inputs
runMesonet()

##convert dates to satisify okmesonet package
startdate <- as.POSIXct(paste(start.date, "00:00:00"), tz = "")
enddate <- as.POSIXct(paste(end.date, "00:00:00"), tz = "")

###for loop to pull data from multiple stations

fivemin <- data.frame() ##create empty data frame

for (station in stations) {
  dat <- okmts(
    startdate,
    enddate,
    station,
    lat =
    NULL, lon
    = NULL,
    variables,
    localtime = TRUE,
    missingNA = TRUE,
    mcores = FALSE
  )
  fivemin = rbind(fivemin, dat)
```

```
}

fivemin$UTC <-
  with_tz(fivemin$TIME, tzone = "GMT")
##convert to Greenwich Mean Time (UTC)

Meso_sub <-
  subset(fivemin, hour(fivemin$UTC) == 00 &
           minute(fivemin$UTC) == 00)
##extract daily precip values which are held at 0000UTC

Meso_sub$Date <-
  as.Date(Meso_sub$UTC, format = "%Y-%m-%d", tz = "America/Chicago")

Meso <- subset(Meso_sub, select = -c(UTC, TIME))

meso_sort <-
  Meso[order(Meso[, 1], Meso[, 4]), ]
##sort by site ID then date descending

###now flag any error codes for future QA
meso_sort$flag <- case_when(
  meso_sort$RAIN ==  -999 ~ 1,
  meso_sort$RAIN ==  -998 ~ 2,
  meso_sort$RAIN ==  -997 ~ 3,
  meso_sort$RAIN ==  -996 ~ 4,
  meso_sort$RAIN ==  -995 ~ 5,
  meso_sort$RAIN ==  -994 ~ 6
)

meso_sort$RAIN2 <-
  as.numeric(replace(meso_sort$RAIN, which(meso_sort$RAIN < 0), NA))
  ##create field where anything <0 = NA
meso_sort$TF                        <-
  ifelse(meso_sort$RAIN2 < 0.1, 1, 0)
  ##boolean operation to determine if rainfall is <0.1 inch rainfall
meso_sort$TF1 <-
  ifelse(is.na(meso_sort$TF), 0, meso_sort$TF)
  ##create field where anything NA in TF = 0
#determine start date of dataset
startdate <- min(meso_sort$Date)
##make sure it's formatted
startdate <- as.Date(startdate, "%m/%d/%Y")

##define site name
sitename <-
meso_sort$STID y <-
meso_sort

###ADP########
#########################
adpsum <- function(x) {
  sequence(rle(x$TF)$lengths) * x$TF ##run length function>boolean column
```

```
"TF"
}
ADP <- rep(NA, length(y$TF))
for (site in sitename) {
  for (i in 1:length(y$TF)) {
    ADP[i] = adpsum(y)[i]
  }
  break
}
y$ADP <- ADP


##extract y, m, d for ease of sorting etc later
m1 <- y
m1 <- arrange(m1, STID, Date) ##sort by site ID then date descending
m1$Year <- year(m1$Date)
m1$Month <- month(m1$Date)
m1$Day <- day(m1$Date)

###calculate sums of preceding precip days
###including current day (minimum of 2 days)
###example: k =2 will sum the current and previous day)

startdate <- min(m1$Date)
startdate <- as.Date(startdate, "%m/%d/%Y")
sitename <- m1$STID

k <- 2 ## number of days precedingfor (site in sitename) {
  for (i in 1:length(m1$Date)) {
    day = m1$Date[i]
    if (day <= startdate) {
      ###this filters out the start date
      m1$precip48hr[i] <- NA
    }
    if (day >= startdate + (k - 1)) {
      ###time to calculate rolling sum
      m1$precip48hr[i] = (sum(m1$RAIN[i:(i - (k))], na.rm = TRUE)) -
m1$RAIN[i]
      #doesnt include current day (example: k = 2 for preceding day(s) sum)
    }
  }
  break
}

###including current day (minimum of 2 days)
###example: k =2 will sum the current and previous
day) j <- 3

for (site in sitename) {
  for (i in 1:length(m1$Date)) {
    day = m1$Date[i]
    if (day <= startdate) {
      ###this filters out the start date
      m1$precip72hr[i] <- NA
```

```
    }
    if (day >= startdate + (j - 1)) {
      ###time to calculate rolling sum
      m1$precip72hr[i] = (sum(m1$RAIN[i:(i - (j))], na.rm = TRUE)) -
m1$RAIN[i]
      #doesnt include current day (example: k = 2 for preceding day(s) sum)
    }
  }
  break
}
col_order <-
  c(
    'Date',
    'STNM',
    'STID',
    'RAIN',
    'flag',
    'RAIN2',
    'TF',
    'TF1',
    'ADP',
    'precip48hr',
    'precip72hr'
  ) ##reorder columns
m1 <- m1[, col_order] ##assign columns

write.csv(
  subset(m1, select = -c(TF, TF1, RAIN2)),
  file = paste0(Work_Dir, "\\", savename, ".csv"),
  row.names = F
) ##Write to CSV
```