THESIS

CHARACTERIZATION OF WATER QUALITY POLLUTION IN MIXED LAND USE WATERSHEDS

Submitted by

Madeline Ludwig

Department of Civil and Environmental Engineering

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2020

Master's Committee:

Advisor: Mazdak Arabi

Susan De Long
Mike Wilkins

ABSTRACT

THE CHARACTERIZATION OF POLLUTION IN MIXED LAND USE WATERSHEDS

Anthropogenic sources of pollution often lead to degraded surface water quality in urban and agricultural streams. The Clean Water Act was developed to mitigate the negative effects of urbanization on water quality through the development of water quality targets and the Total Maximum Daily Load program. In this study, a probabilistic framework was developed to quantitatively assess how indicators of human influence impact vulnerability to *E. coli* impairment and nutrient concentrations in mixed land use watersheds across the state of Colorado. The models derived using this method can be used to predict instream pollutant concentrations and help regulatory agencies create sampling programs for at risk waterbodies.

Specifically, the first part of this study explores vulnerability to *E. coli* impairment under varying levels of upstream anthropogenic influences and develops a probabilistic method for assessing *E. coli* pollution based on the regulatory monitoring program. In this study, vulnerability is defined as the probability that ambient instream pollutant concentrations exceed numeric water quality standards. The study objective was examined for 28 sites along the Cache la Poudre River and its tributaries including: Boxelder Creek, Fossil Creek, and Spring Creek in northern Colorado. Indicators of urban influence include land use, wastewater treatment plant discharge capacity, combined animal feeding operation capacity, and population. Multiple linear regressions analysis between anthropogenic indicators, *E. coli* concentrations and vulnerability provide significant ($p < 0.05$) and strong ($R^2 > 0.7$) relationships. In general, land use predictor variables were able to accurately predict *E. coli* load, however the most important indicator of human influence differed between *E. coli* concentration response variables.

Additionally, the second part of this study expands upon the multiple linear regression framework to develop regression models that can predict base level nutrient concentrations for stream segments in three regions of Colorado. Regression models were developed using data from 89 sampling locations upstream of wastewater treatment plants and 84 sampling locations downstream of wastewater treatment plants. An initial analysis of gaged sampling locations showed that flow was a significantly influenced instream nutrient concentrations. Area and slope of the contributing sub watershed were then analyzed in a regression analysis and were found to be a surrogate for streamflow. Strong ($R^2 > 0.7$) and significant ($p < 0.05$) regression models for upstream and downstream locations were developed using area and slope, hydrologic, point, and non-point source predictor variables. The models showed that agricultural and urban activity significantly impacted instream baseline nutrient concentrations. The methodology developed in this study can be used to predict instream pollutant concentration and assist in the development of monitoring programs for at risk waterbodies.

ACKNOWLEDGEMENTS

I am extremely grateful for the help and support I have received to complete my degree and research. I would like to thank Dr. Mazdak Arabi for his guidance and encouragement that have not only benefitted this project, but my academic and professional career. This project would not have been possible without assistance from Tyler Wible, Tyler Dell, Caroline Van Horn, C.J. Schieman, Reed Featherstone, Will Rainey, and Canon Furth. I would additionally like to thank my committee members, Dr. Susan De Long and Dr. Mike Wilkins for their input.

This research was funded by the Colorado Department of Public Health and the Environment Water Quality Control Division and could not have been completed without the help and support of the Water Quality Control Division's team.

Finally, I cannot thank my family and friends enough for their love and support they have given me through this experience. The friends I have made at Colorado State have not only provided valuable advice but have been a great group to explore the outdoors with and made Colorado feel like home. I will be forever grateful for my wonderful family. Mom, Dad, Kate, Honey, Grandma, Grandpa, and the Wheeler family - I couldn't have done it without you. You have been the best support system I could ask for.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 Background

Nutrients and bacteria are both necessary in maintaining diversity in an aquatic ecosystem. However, disproportionate amounts can lead to degradation of surface waters and risk to human health. Increased fecal indicator bacteria from various urban activities are an indicator of recent fecal contamination and can indicate the co-presence of human pathogens that can lead to gastrointestinal disease in generally healthy humans (Feng et al., 2018). The presence of high NO_3 levels in drinking water has been associated with methemoglobinemia (or “blue baby syndrome”) and increased risk for cancer (Bryan & van Grinsven, 2013). Excessive concentrations of nitrogen and phosphorus in instream environments result in eutrophication, the decline of species diversity, and increased risk to human health due to harmful algal blooms (Klein, 1979; United States Environmental Protection Agency, 2001). The Clean Water Act (CWA) was established in 1972 to combat these issues and improve surface water quality, through the establishment of water quality standards and the development of the Total Maximum Daily Load (TMDL) program (Congress, 2002).

Regulatory approaches under the Clean Water Act, such as the Total Maximum Daily Load (TMDL) program and numeric water quality standards aim to improve the physical, biological, and chemical integrity of surface water (Congress, 2002). The state of Colorado, with guidance from the United States Environmental Protection Agency (USEPA), has established a classification system and basic standards for the state’s surface waters (CDPHE, 2013, 2018). The classification system assigns a numerical water quality standard based on the beneficial use of the water body. Waterbodies that consistently fail to meet standards assigned by the EPA and state are placed on the 303(d) list of impaired waterbodies as outlined by the CWA. Waterbodies with this classification are required to develop a TMDL and implement the program until the waterbody can meet the standards and remove it from the impairment list.

Understanding the interaction between anthropogenic activity and water quality is crucial to investigating water quality degradation and implementing water quality regulations. This is especially true in watersheds with a diverse combination of land uses and point source pollutants. Population growth, urbanization and the conversion of natural land cover to farmed and grazed lands are commonly identified some of the main drivers of stream impairment (Kang et al., 2010; Selvakumar & Borst, 2006; Young & Thackston, 1999). These anthropogenic activities are directly reflected in the landscape and allow morphological, land use, and point source attributes of the contributing watershed to be linked to biological and chemical water quality characteristics (Kang et al., 2010).

Multiple linear regression (MLR) models have been used to successfully relate anthropogenic activities with water quality degradation. This relatively simple approach allows for the characterization of sources of variability in water quality data over a large geographic region. Many studies have found strong and significant relationships between anthropogenic activities and instream pollutant load (Donnison et al., 2004; Francy et al., 2000; Harmel et al., 2010; Mallin et al., 2000, 2009; Pandey et al., 2012; Paule-Mercado et al., 2016; Selvakumar & Borst, 2006; Sylvestre et al., 2020; Tasdighi et al., 2017; Wickham et al., 2000; Williams et al., 2014; Young & Thackston, 1999). Several studies have concluded that percent urban land use was the main cause of pollution (Chang, 2008; Francy et al., 2000; Tasdighi et al., 2017). Other studies have determined that precipitation produced runoff is the leading cause of water quality degradation (Pandey et al., 2012; Paule-Mercado et al., 2016; Sylvestre et al., 2020). Harmel (2010) and Donnison (2004) both concluded that *Escherichia coli* (*E. coli*) concentrations were majorly impacted by the presence of grazing animals (Donnison et al., 2004; Harmel et al., 2010). Williams (2014) found that for nutrient concentrations in urban settings, waste discharge dominated low flows (Williams et al., 2014). However, older studies found that the degraded quality of a stream is not the result of any single factor, but the combined interaction of more than one anthropogenic influence (Klein, 1979; McMahon & Cuffney, 2000; Selvakumar & Borst, 2006; Young & Thackston, 1999). While these studies identify factors that lead to impairment of streams and rivers, they do not incorporate risk of impairment into the framework. Water quality impairment has already been widely identified across the United States and, with continued urbanization, it is critical to understand source of pollutants in order to meet goals set by the CWA and state agencies.

1.2 Regulatory approaches

The state of Colorado has implemented numeric bacteria limits for surface water based on their designated beneficial (CDPHE, 2018). The standards of concern for this study were for recreation class 1a and 1b. Stream segments classified as 1a are used for primary contact recreation, where primary contact recreation is defined as activities that include total body immersion with the potential for ingestion. Water bodies classified as 1b are used for potential primary contact recreation, at these locations primary contact is not currently occurring although such uses could potentially happen in the future (CDPHE, 2018). The geometric mean value for streams designated as recreation class 1a should not exceed 126 colony forming units per 100 milliliters (CFU/100mL), while the value for streams designated as recreation class 1b should not exceed 205 CFU/100mL (CDPHE, 2018). The geometric mean must be determined based on five water samples taken at least seven days apart during a two-month period (CDPHE, 2018).

In 2012, the Colorado Department of Public Health and Environment (CDPHE) implemented numeric nutrient limits for surface water in order to reduce the eutrophication (CDPHE, 2013). Like the numeric limit established for *E. coli*, the nutrient standards are determined based on the designated use of the water body. Surface waters are classified by warm or cold-water aquatic use. Warm water rivers and streams support biota that exist in waters with average summer temperatures that exceed 20 °C, while cold waters support biota that thrive in waters where this 20 °C threshold is typically not exceeded.

Warm water median nitrogen concentration is limited to 2.01 mg/L and the annual median phosphorous concentration is limited to 0.17 mg/L (CDPHE, 2013). The annual median concentrations for cold water are 1.25 mg/L and 0.11 mg/L for total nitrogen and total phosphorous respectively (CDPHE, 2013).

1.3 Research objectives

The thesis presented here is comprised of four chapters. The first chapter is a literature review that outlines the impact of anthropogenic influences on instream water quality and the methods that have previously been used to link these influences to water quality impairment. The objectives of the second chapter are to (i) investigate trends in water quality impairment and identify sources that improve and degrade surface water, (ii) identify the effect of precipitation on *E. coli* water quality exceedances, and (iii) explore the variability of vulnerability to *E. coli* impairment under varying levels of upstream anthropogenic influences and, (iv) develop a probabilistic method for assessing *E. coli* pollution based on the regulatory monitoring program in a mixed land use watershed. A multiple linear regression was performed to predict ambient *E. coli* concentrations and vulnerability as functions of anthropogenic influences.

The third chapter expands on the multiple linear regression framework developed in the second chapter to describe baseline nutrient concentrations across the state of Colorado. Specifically, the objectives of this chapter are to (i) examine trends in nutrient concentration in the different regions in Colorado, (ii) expand upon the multiple linear regression framework to develop regression models that can predict base-level nutrient concentrations for stream segments in three regions of Colorado.

The final chapter of this works provides a general synopsis of the work and summarizes the results of chapters two and three.

1.4 Developing models to predict vulnerability and concentrations of pollutants

In this study, a probabilistic framework was developed to characterize vulnerability and concentrations of pollutants to better understand the effects of anthropogenic influences on water quality degradation. Adequate water quality data are often lacking because of the cost associated with implementing extensive water quality monitoring and need for skilled labor to ensure equipment is adequately maintained. For this reason, accurately determining water quality impairment is difficult or even impossible to do in some areas. Utilizing an MLR model allows the user to rely on publicly available data to determine baseline conditions and vulnerability to impairment.

In chapter 2, observed *E. coli* concentration were fit to a lognormal distribution. The expected value was computed and compared with the water quality regulation target. The proportion of samples that exceeded the numeric target were analyzed to identify if a significant difference exists between rain and non-rain sampling events. Precipitation data were also used identify patterns of impairment. Vulnerability was defined as the likelihood that ambient water quality exceeds numeric targets and was calculated as the probability that the geometric mean of instream *E. coli* concentrations exceed Colorado Regulation 31 water quality standards. An MLR analysis was performed between vulnerability, *E. coli* concentrations and sources of pollution to relate water quality deterioration to anthropogenic influences.

The second part of this research, outlined in chapter 3, utilizes a similar MLR analysis method to analyze baseline nutrient conditions for sites in three regions in Colorado, the South Platte River basin, the Arkansas River basin, and the western slope. Stream flow has been found to be positively correlated to nutrient concentrations. However, because stream flow is not available at all sites the gauged sites were analyzed to find relationships between slope, area, precipitation, and stream flow that could be used as a surrogate in the regression models. The later data are more widely available. Geospatial factors were extracted from each watershed and used to create a series of linear regression models that could relate median nutrient concentrations to anthropogenic factors. The models developed can be applied to all stream segments across Colorado to predict instream pollutant concentration and assist in the development of monitoring programs for at risk waterbodies.

CHAPTER 2: CHARACTERIZATION OF VULNERABILITY TO *E. COLI* POLLUTION ALONG MIXED LAND USE STREAMS

2.1 Background

Urbanization and the conversion of natural land cover to farmed and grazed lands have led to stream impairment across the United States(Kang et al., 2010; Selvakumar & Borst, 2006; Young & Thackston, 1999). Increased fecal indicator bacteria from various urban activities are an indicator of recent fecal contamination and can indicate the co-presence of human pathogens that can lead to gastrointestinal disease in generally healthy humans (Feng et al., 2018). *Escherichia coli* (*E. coli*) is commonly used as a fecal indicator bacteria because *E. coli*'s only natural habitat is the gastrointestinal tract of humans and other warm-blooded animals (Feng et al., 2018; Odonkor & Ampofo, 2013; Standridge, 2008). *E. coli* is rarely found in other habitats due to its inability to survive long outside of the gastrointestinal tract of a host (Odonkor & Ampofo, 2013) or multiply appreciably in the environment(Edberg et al., 2000). E. coli has been found to survive longer in sediment than in the overlying water because the sediment protects the E. coli bacteria from U.V. exposure and predators, namely protozoa (Garzio-Hadzick et al., 2010). These limitations of the bacteria imply that the presence of *E. coli* in the environment is an indicator of recent fecal contamination.

The Clean Water Act (CWA) was established in 1972 in order to improve surface water quality in the United States through the development of numeric water quality standards and the development of the Total Maximum Daily Loads (TMDL) program. States were tasked with the creation, implementation, and monitoring of water quality standards. The state of Colorado has implemented numeric bacteria limits for surface waters based on their designated beneficial use (CDPHE, 2018). The standards of concern for this study were for recreation classes 1a and 1b. Stream segments classified as 1a are used for primary contact recreation, where primary contact recreation is defined as activities that include total body immersion with the potential for ingestion. Waters bodies classified as 1b have potential for primary contact recreation, at these locations primary contact is not currently occurring although such use could happen in the future (CDPHE, 2018). Geometric means of streams designated as recreation class 1a are limited to 126 CFU/100mL, while those designated as recreation class 1b should not exceed 205 CFU/100mL (CDPHE, 2018). The geometric mean must consist of 5 water samples taken at least 7 days apart during a two-month period (CDPHE, 2018).

The large number of water bodies in the United States makes monitoring the quality of all segments difficult. However, using models to predict concentration of pollutants can help identify waterbodies that are vulnerable to pollution. Understanding the effect of precipitation is the first step to properly characterize sources of impairment. Stormwater runoff transports pollutants that have accumulated during dry weather periods to the nearest waterbody. Paule-Mercado et al. (2016) found that precipitation can be a significant source of non-point source pollutants, especially in urban watersheds (Paule-Mercado et al., 2016). Similarly, Pandey et al. (2012) found that runoff from crop land can potentially increase *E. coli* concentrations in areas where manure is used as fertilizer (Pandey et al., 2012).

Multiple linear regression (MLR) models have been used in past studies to successfully relate anthropogenic activities with water quality degradation. This straightforward approach allows for the characterization of sources of variability in water quality data over a large geographic region. Many studies have found strong and significant relationships between anthropogenic activities and instream pollutant load (Donnison et al., 2004; Francy et al., 2000; Harmel et al., 2010; Mallin et al., 2009; Pandey et al., 2012; Paule-Mercado et al., 2016; Selvakumar & Borst, 2006; Sylvestre et al., 2020; Young & Thackston, 1999). Francy (2000) concluded that increased percent urban land use was the main cause of pollution (Francy et al., 2000). Other studies have determined that precipitation produced runoff is the leading cause of water quality degradation (Pandey et al., 2012; Paule-Mercado et al., 2016; Sylvestre et al., 2020). Harmel (2010) and Donnison (2004) both concluded that *E. coli* concentrations were majorly impacted by the presence of grazing animals (Donnison et al., 2004; Harmel et al., 2010). While newer studies have been able to identify specific sources of pollution, older studies have shown that the degraded quality of a stream is not the result of any single factor, but the combined interaction of several anthropogenic influences (Klein, 1979; Selvakumar & Borst, 2006; Young & Thackston, 1999).

While these studies identify factors that lead to impairment of streams and rivers, they do not incorporate risk of impairment into the framework. In a study conducted by Tasdighi et al. (2017), a statistical framework is developed to quantify vulnerability. Vulnerability was defined as the probability that ambient contaminant levels will exceed the numeric standard. Heiden et al. (2019) adopted this statistical framework and applied it in an MLR analysis to characterize vulnerability of urban streams to nutrient pollution across four distinct urban gradients. Strong and significant models were not found for all sites in the study, because of the rigorous nature of the numeric targets for nutrients. However, vulnerability was found to be strongly correlated to total phosphorous pollution and found to be strongly correlated to impervious surface cover and wastewater treatment plant discharge.

Although multiple regression analyses in combination with geographic information systems (GIS) have been widely used to identify exploratory geospatial variables effect on bacterial water quality, no previous studies have applied a probabilistic framework that integrates ambient *E. coli* concentrations and its numeric targets to describe the risk of water quality exceedances. More importantly, no study has incorporated this rigorous statistical procedure into a regression model to characterize how vulnerability to *E. coli* pollution changes with variations in anthropogenic influence. This study takes the vulnerability framework developed by Tasdighi and Heiden and adopts it to characterize vulnerability of a system to bacterial pollution (Heiden, 2019; Tasdighi et al., 2017).

This study investigates the relationship between anthropogenic indicators of urbanization and *E. coli* vulnerability and concentrations in the Cache la Poudre watershed. Where vulnerability to *E. coli* pollution is a function of ambient contaminant levels and the numeric standard. The specific objectives are to (i) investigate trends in water quality impairment and identify sources that improve and degrade surface water, (ii) identify the effect of precipitation on *E. coli* water quality exceedances, and (iii) explore the variability of vulnerability to *E. coli* impairment under varying levels of upstream anthropogenic influences, and (iv) develop a probabilistic method for assessing *E. coli* pollution based on the regulatory monitoring program in a mixed land use watershed

2.2 Methodology

A complete geospatial analysis was performed in order to quantify human sources of pollution that lead to water quality degradation. These point and non-point sources of pollution were assessed alongside vulnerability to *E. coli* exceedance in an exhaustive multiple linear regression (MLR) analysis to relate impairment to anthropogenic influences. A *z*-test analysis was also performed to determine if the proportion of samples that exceed numeric standards is significantly different during precipitation events.

2.2.1 Study area

The Cache la Poudre (CLP) Watershed drains approximately 4896 km^2 (1890.5 mi^2) of Northern Colorado and Southern Wyoming. The headwaters begin in Rocky Mountain National Park and flow approximately 205 km (127 mi) before the river's confluence with the South Platte River just east of Greeley (Figure 1). The CLP watershed is largely comprised of undeveloped forested land until it reaches the mouth of the canyon 89 km (55 mi) northwest of the confluence with the South Platte River. Here it enters area of mixed land use ranging from urban to mixed agriculture. The CLP watershed drains the Colorado cities of Fort Collins, Windsor, Greeley, and Wellington. The lower portion of the watershed is used extensively for irrigated agriculture and combined animal feeding operations (CAFOs).

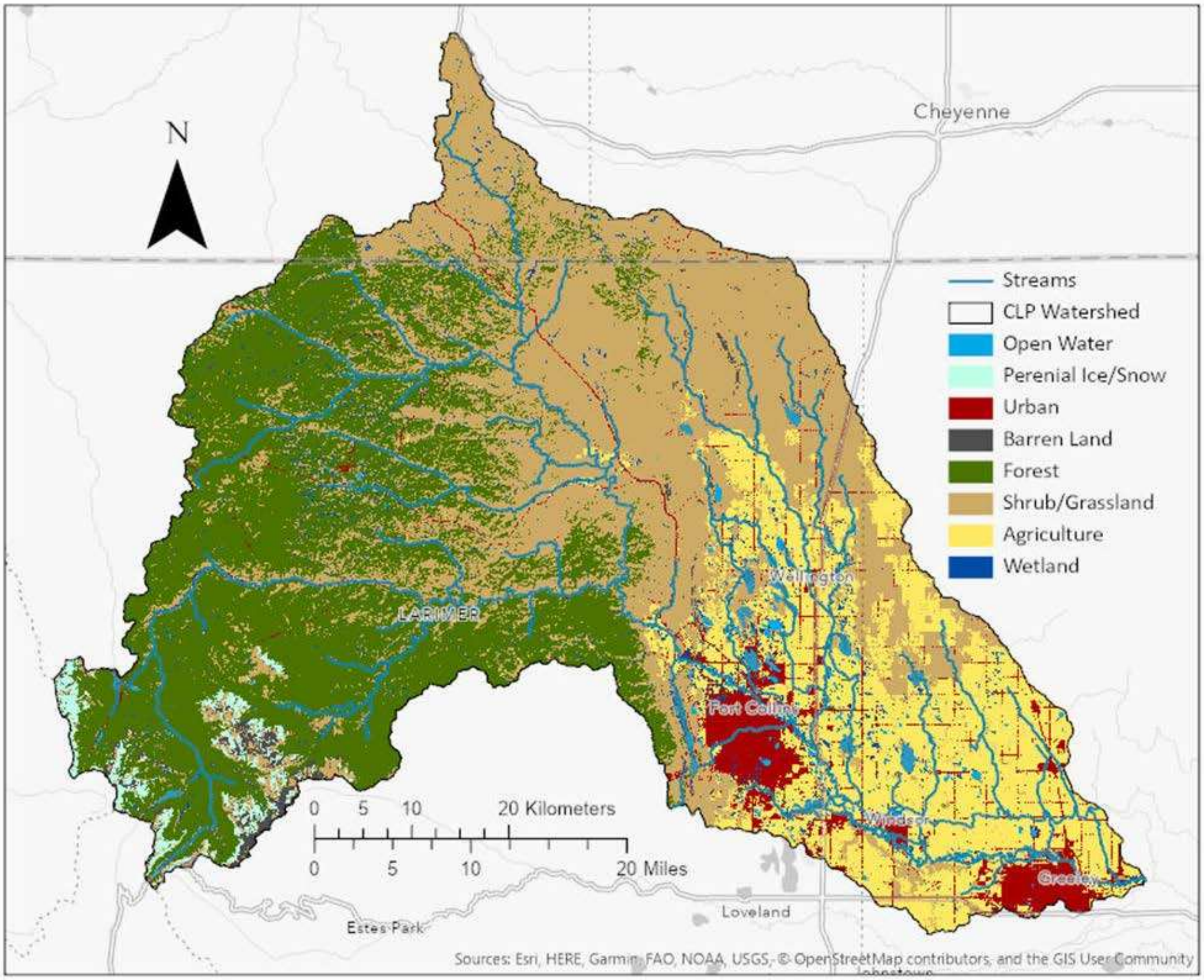


Figure 1 The study area located in Northern Colorado: Cache la Poudre (CLP) watershed land use

The lower portion of the CLP watershed was the focus of this study. 28 sampling locations were selected along the CLP River, Spring Creek, Fossil Creek, and Boxelder Creek (Figure 2). Each sampling location encompass a unique combinations of land uses, levels of human impact, and water quality standards (Appendix A).

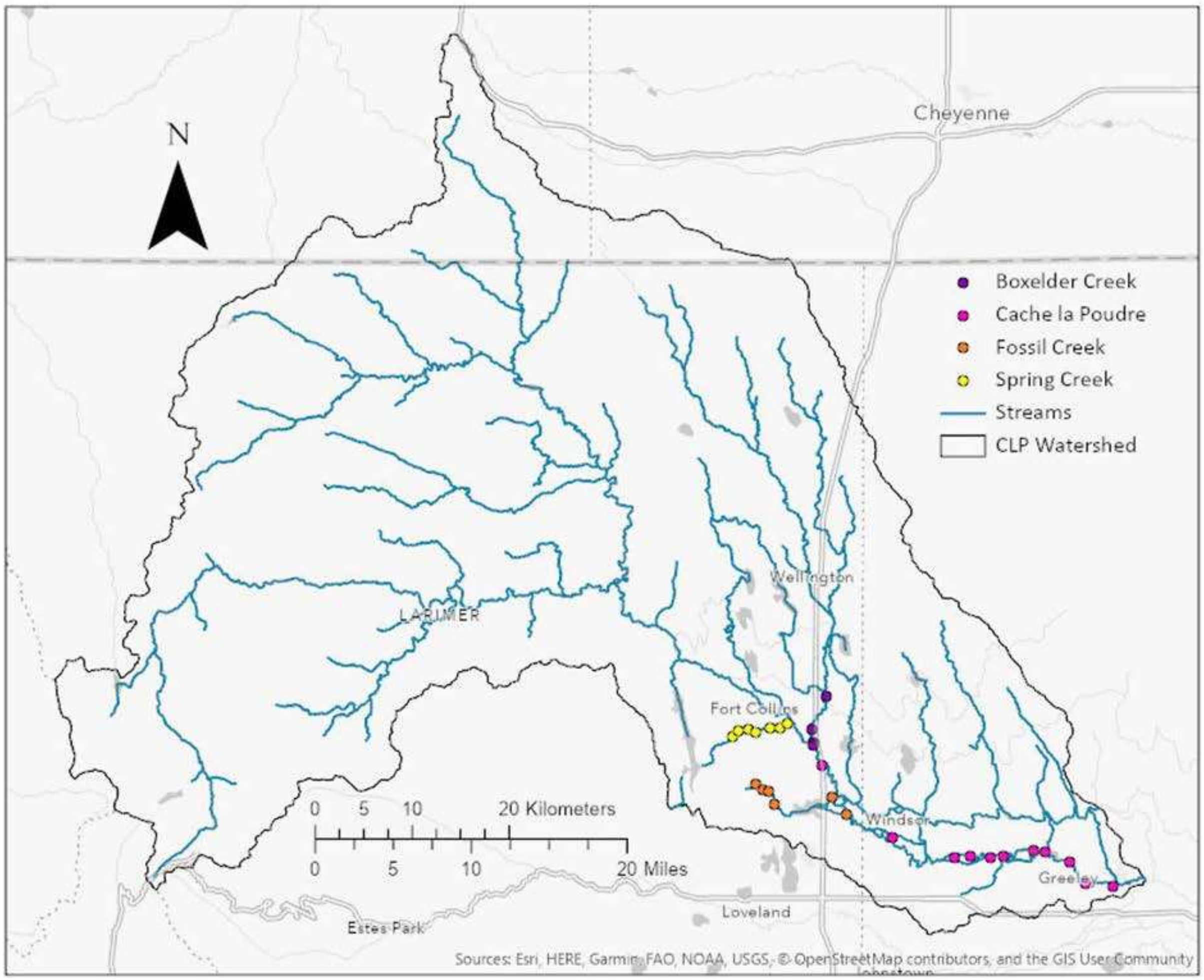


Figure 2 The CLP watershed sampling locations along the Boxelder Creek (purple), CLP River (pink), Fossil Creek (orange), and Spring Creek (yellow). The blue lines represent streams and rivers in the watershed and the black outline represents the CLP watershed.

2.2.1.1 Sample Collection Procedure

Sample collection occurred May through October during the 2018 and 2019 calendar year, at a frequency of 5 sampling events (for each site) every 61 days, with each sampling event occurring at least 7 days apart. Samples were taken over this time period because e. coli is most strictly regulated during these months due to many waters' recreational use. The samples were collected according to the standard operating procedures for the collection of water chemistry samples developed by the Water Quality Control Division of CDPHE (CDPHE, 2015). Laboratory analysis of the grab samples was performed according to EPA Method 1103.1 (USEPA, 2002). A regression of order statistics was performed for samples that had concentrations below the detection limit using a lognormal distribution in ProUCL 5.1 (Anita & Maichle, 2015).

Field measurements for temperature, specific conductivity, dissolved oxygen, and pH were conducted along with sample collection. This was accomplished using a multi-parameter sonde. Stream discharge measurements were measured at each site using a portable velocity meter. Sampling occurred in a downstream to upstream sequence to avoid the possibility of influencing water quality parameters with field work activity.

2.2.1.2 Collected Data

Over the course of the study 819 samples were collected. 41 samples resulted in non-detect readings. The samples had a mean of 493 CFU/100 mL and median of 1140 CFU/100 mL. The minimum *E. coli* level was 4.43 CFU/100mL and the maximum was 20,000 mL/100 mL. Flow measurements were taken on 641 of the 819 visits. Flow measurements were not taken during visits when the water was too deep or fast to enter safely or in instances when there was an equipment malfunction. Water quality (temperature, specific conductivity, dissolved oxygen, and pH) readings were taken at every visit, however during the first sampling events at the CLP locations and Fossil Creek locations DO measurements were not taken due toa malfunctioning sensor.

101 blanks, 102 duplicates, and 83 spike samples were also collected. All blank samples were reported back as non-detects, duplicate samples have a mean relative percent difference of -12.86%, and a mean spiked recovery of 279.30%.

2.2.2 Geospatial analysis

Vulnerability to microbiological pollution was related to various anthropogenic sources of pollution within each sub-basin. The contributing watershed for each sampling location was delineated using ArcGIS Pro version 2.4.0 (Esri, 2019). A comprehensive analysis of anthropogenic influence was performed in order to quantify urban, agricultural, and natural sources of pollution, as well as determine characteristics of the sub-basin that could be used to describe water quality impairment.

2.2.2.1 Watershed Delineation

The watershed associated with each sampling location was delineated using the ArcHydro toolbox in ArcGIS Pro. The drainage basins were found using a 1/3 arcsec digital elevation model from the National Elevation Dataset and the National Hydrography Dataset (NHD) from the USGS. The boundaries were carefully determined in the lower portion of the CLP watershed, to account for the irrigation ditches that change the natural hydrology of the area. Slope was found using the 1/3 arcsec digital elevation model and the slope and zonal statistics tool in ArcGIS Pro.

2.2.2.2 Characterizing anthropogenic indicators

Various anthropogenic sources lead to the degradation of streams (Table 1). Indicators of these sources include wastewater treatment plant capacity, animal feeding operation capacity, percent impervious land cover, population density, and watershed/hydrologic characteristics. These indicators were used to characterize sources of vulnerability. Using a variety of point and non-point source variables allows for a multiple linear regression with little to no collinearity.

Table 1 Summary of anthropogenic predictor variables employed in multiple linear regression models

Variable	Type	Unit
WWTP Capacity	Point Source/Facilities	MGD
CAFO Capacity	Point Source/Facilities	# animals
IDW WWTP Capacity	Point Source/Facilities	MGD/mi
IDW CAFO Capacity	Point Source/Facilities	# animals/mi
Population Density	Non-Point Source	People/mi^2
% Impervious Surface Cover	Land Use	%
% Urban land Use	Land Use	%
% Crop Land Use	Land Use	%
% Forest Land Use	Land Use	%
% Range Land Use	Land Use	%
Watershed Area	Watershed Characteristic	mi^2
Slope	Watershed Characteristic	ft/ft
Average Annual Precipitation	Hydrologic	in

The wastewater treatment plant (WWTP) and confined animal feeding operations (CAFO) locations and capacities were found using the Environmental Resource Assessment and Management System (One Water Solutions Institute, 2020). The wastewater treatment plant flow capacities were extracted from Colorado Department of Public Health and the Environment (CDPHE) Regulation 85 data (Water Quality Monitoring Council, 2020). The confined animal feeding operation capacity was calculated as the sum of animal units. The contribution of these variables was summarized by the cumulative contribution and by using an inverse distance weighting (IDW) method. The distance used for IDW was calculated as the sum of the overland distance to the nearest stream and the instream distance to the nearest sampling location.

Land use and impervious surface data was collected from the 2016 National Land Cover Database (NLCD) (Multi-Resoulution Land Characteristic Constortium, 2016). Urban land use was defined as a combination of low, medium, and high intensity developed land, and developed open spaces; cropland is a combination of pasture, hay, and cultivated crops; forest land cover includes deciduous, evergreen, and mixed forest; rangeland is the combination of grassland, shrub, and scrub. Tract level population data was collected from the U.S. Census and used to calculate population density (U.S. Census Bureau, 2019).

Average annual precipitation for each site was procured from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset (Northwest Alliance for Computational Science and Engineering, 2020).

2.2.3 The Relationship between *E. coli* and anthropogenic indicators

In this research, vulnerability is defined as the probability of the geometric mean of the ambient water quality exceeding Colorado Regulation 31 water quality standards. *E. coli* data was fit to a lognormal distribution and the likelihood of the geometric mean exceeding the numeric target was determined using a student *t* distribution. Vulnerability was then compared between reaches to assess the impact of anthropogenic factors on water quality. A *z*-test was used to determine if precipitation significantly impacted the proportion of the samples that exceeded the water quality standards. Finally, a multiple linear regression (MLR) analysis was performed between point and non-point source of pollutants and vulnerability, mean, and median concentrations of *E. coli* to relate impairment to anthropogenic influences.

2.2.3.1 Descriptive statistics

Sample geometric mean, arithmetic mean, median, and standard deviation were calculated for the *E. coli* concentrations at each sampling site. The geometric mean was found as the average of the two-month geometric mean. The data analyzed is from the May through October time frame. For this study it was assumed that the limited range of the samples adequately describes the basic statistics of the *E. coli* concentrations.

2.2.3.2 Characterizing vulnerability

A proper statistical distribution to describe *E. coli* concentration was essential to determining the number of samples required for each site and characterizing vulnerability and its relationship to anthropogenic indicators. The Kolmogorov-Smirnov test for normality was used to analyze the fit of lognormal, normal, exponential, and gamma distributions computed using the maximum likelihood estimation. *E. coli* concentrations for the selected sites were found to fit a lognormal distribution and non-detect values were accounted for using a lognormal linear regression of order statistics.

Vulnerability to *E. coli* impairment was characterized as the probability that the geometric mean of *E. coli* concentrations exceeds the numeric target (*T*). Measured *E. coli* concentrations (*x*) were assumed to be lognormally distributed such that *y*=log*(x)* is normally distributed with sample mean (Θ) and sample standard deviation (ω).

$$x = \{x_1, x_2, \dots x_n\} \qquad [1]$$

$$y \sim N(\theta, \omega) \quad [2]$$

In order to determine if the geometric mean of ambient water quality exceeds the numeric target, the expected value (*Y*) and the standard deviation (σ) of the data for a normal distribution was calculated as

$$Y = mean[\ln(x)] \quad [3]$$

$$\sigma = \frac{1}{\sqrt{n}} std[\ln(x)] \quad [4]$$

where *n* is the sample size.

The standard normal variable (z_q) for a given quantile can be obtained from the standard normal table based on the quantile of interest (*q*).

$$z_q = \phi^{-1}(q) \quad [5]$$

For any set of n log-transformed water quality observations (*y*), the probability that the expected value (*Y*) exceeds the log-transformed numeric target concentration (*T*), can be computed as

$$P = 1 - F_Y[\log(T)] \quad [6]$$

where $F_Y[\log(T)]$ is the cumulative distribution of *Y* and *P* is the probability of exceedance. Applying the student t distribution, the vulnerability or probability of exceedance (*P*) was expressed as

$$Vulnerability = P = 1 - \phi\left[\frac{\log(T) - Y}{\sigma}\right] \quad [7]$$

where Φ is the non-exceedance probability, *T* is the target concentration, *Y* is the mean of the log transformed values, and σ is the log transformed standard deviation.

2.2.3.3 Determining the effect of precipitation on instream E. coli concentrations

Rain and non-rain events were analyzed to assess if there is a statistical difference between the proportion of *E. coli* samples that exceed the water quality standard at each site. Rain events were defined as events where rainfall exceeding 0.1 inch occurred in the 24 hours prior to sampling, while dry events had less than 0.1 in of rainfall in the 24-hour period prior to sampling. This analysis was performed using the two-sample *z*-test (Walpole et al., 1998). This analysis tests the null hypothesis that proportion samples that exceed the standard during rain and non-rain are not statistically different at α=0.5 significance level.

The two-sample *z*-test is a parametric test that compares proportions of two independent, normally distributed data sets. Measured *E. coli* concentrations (*x*) were assumed to be lognormally distributed such that *y*=log(*x*) is normally distributed. The proportion of log transformed samples that exceed the log transformed standard was computed as:

$$P_1 = \frac{count(y_1 > s)}{n_1} \tag{8}$$

$$P_2 = \frac{count(y_2 > s)}{n_2} \tag{9}$$

Where y_1 and y_2 are the rain and non-rain populations, *s* is the log transformed water quality standard, and n_1 and n_2 are the sample sizes. The overall proportion of samples that exceed the log transformed standard for each site was calculated as:

$$\hat{P} = \frac{count(y > s)}{n} \tag{10}$$

The test statistic z was calculated as a function of P_1, P_2, $\hat{P}$, and sample size.

$$z = \frac{P_1 - P_2}{\sqrt{\hat{P}\left(1-\hat{P}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \tag{11}$$

2.2.3.4 Characterizing anthropogenic intensity

Mean, geometric mean, and median *E. coli* concentration, and vulnerability to *E. coli* impairment were fit to multiple linear regression (MLR) models based on different combinations of watershed (area, slope), hydrologic (average annual precipitation), urban (population, WWTP capacity, IDW WWTP capacity, % impervious surface cover, % urban land cover), agricultural (% cropland cover, CAFO capacity, IDW CAFO capacity), and natural (% rangeland cover, % forest) characteristics.

The regression models were determined using the *regress* function in MATLAB v9.6 (R2019a) (MathWorks, 2019). Vulnerability, mean, and median concentrations were transformed with power functions using a box-cox transformation. The box-cox transformation identifies the transformation of the response variable (y) that is most appropriate for correcting skewness of the error terms, unequal error variance, and nonlinearity of the regression function (Kutner et al., 2005).

$$y(\lambda) = \begin{cases} \dfrac{c^{\lambda} - 1}{\lambda}; \lambda \neq 0 \\ \log(c); \lambda = 0 \end{cases} \qquad [12]$$

where c represents the observed *E. coli* concentrations in CFU/100mL and λ is the box-cox transformation constant. An iterative procedure was used to identify the λ value for each *E. coli* response variable that maximized the goodness of fit of the MLR model.

A comprehensive analysis was performed by pairing point and non-point anthropogenic predictor variables together to develop MLR models for *E. coli* vulnerability, mean, and median. The Schwarz' Bayesian criterion (SBC) and Akaike information criterion (AIC) were used to select the best model to describe *E. coli* vulnerability, mean, and median (Kutner et al., 2005).

Various statistical tests were used to assess the appropriateness of the chosen MLR models. The coefficient of multiple determination (R^2) and adjusted coefficient of multiple determination (Adj R^2) were used to determine how much of the variability the model accounted for. The overall significance of the regression model was determined using the lack of fit F-test. The T-test was used to evaluate the significance of model parameters. The normality of the residuals was analyzed using the Shapiro-Wilk test and the Lillie Test. Multicollinearity in the matrix of predictor variables was assessed using the variance inflation factor (VIF). Predictor variables should have a VIF value near 1 and collectively the sum of the VIF values of all model variables should be less than 10. Multicollinearity was avoided by only using one anthropogenic variable to describe each category in the model's predictor matrix.

2.3 Results

Vulnerability to *E. coli* impairment was fit to a lognormal distribution when non-detect values were accounted for. *E. coli* concentrations and vulnerability generally increase with distance downstream due to the cumulative impact of anthropogenic influence. Precipitation did not significantly impact the number of samples that exceeded the numeric water quality target; however, general trends were identified. Precipitation produced runoff generally lead to increased *E. coli* concentration in locations that are not greatly impacted by point source pollutants. While sites dominated by point source pollutants experienced a reduction in concentration caused by dilution during rain events. An MLR analysis was performed to fit models to predict *E. coli* vulnerability, geometric mean, mean, and median *E. coli* concentrations based on hydrologic and anthropogenic predictor variables. Significant and strong models were found. For sites in the Cache la Poudre watershed, land use was a valid surrogate for point source inputs.

2.3.1 *E. coli* concentration along a gradient of anthropogenic impact

Table 2 shows the various anthropogenic predictor variables were found as a result of the geospatial analysis. Sources of pollution increase with distance downstream. The capacity of WWTP, number of CAFOs, and population increase cumulatively as the CLP river nears the confluence with the South Platte River. The inverse distance weighted method used to describe CAFOs and WWTPs cause these parameters to not increase cumulatively with distance downstream. This method also reduced the effect of collinearity between the anthropogenic factors which exists due to the cumulative nature of the variables with distance downstream.

Table 2 Variables used in multiple linear regression compiled through geospatial analysis. Wastewater treatment plant [WWTP IDW] and confined animal feeding operations [CAFO IDW] are shown cumulatively for each site and using and inverse distance weighted method. Non-point source (population, impervious surface cover percent [IS], *urban land cover percent, crop land cover, forest land cover, and rangeland) are shown for each sub watershed as well as average annual precipitation.*

Location	River Mile mi	WS Area mi^2	Average Precip in/yr	Average WS Slope ft/ft	Urban	Crop	Forest	Range	IS	WWTP MGD	WWTP IDW MGD	CAFO #	CAFO IDW #	Total Pop #
					-----------	-----------	%	-----------	-----------					
CLP 01	2.796	1870.7	21.58	0.158	5.1	16.9	33.7	38.5	2	28.84	248.77	106190.3	5548	342528
CLP 02	4.918	1864.2	21.58	0.159	5.1	16.7	33.8	38.6	2	28.84	88.27	106190.3	5585	331215
CLP 03	7.007	1818.7	20.38	0.162	4.6	15.4	34.7	39.3	1.9	19.51	122.55	81083.3	5645	308043
CLP 04	9.076	1751.5	20.38	0.167	4.3	13.5	36	40.1	1.7	19.51	118.88	60766	5753	279143
CLP 05	9.902	1751	20.38	0.167	4.3	13.5	36	40.2	1.7	19.51	119.61	60748	4743	274568
CLP 06	13.559	1736	20.38	0.169	4.2	12.9	36.3	40.5	1.7	19.51	123.78	60748	4390	265692
CLP 07	14.786	1719.5	20.38	0.170	4	12.4	36.6	40.8	1.6	19.51	128.95	60748	4173	248086
CLP 08	17.167	1714.1	20.05	0.171	4	12.1	36.7	40.9	1.6	19.51	135.04	57903	5545	244654
CLP 09	18.512	1705.6	20.05	0.171	4	11.8	36.9	41.1	1.6	19.51	140.47	57903	5518	242655
CLP 10	25.494	1591.3	20.05	0.182	3.9	7.8	39.5	42.7	1.5	17.33	135.76	40252	6359	221740
CLP 11	46.567	1229.4	20.28	0.216	2.1	3	50.5	38.1	0.6	3.08	0.00	3000	3000	98761
BC01	38.304	285.8	18.00	0.074	2.8	20.6	2.4	69.7	0.8	2.79	4.27	30162	6356	36624
BC02	38.475	285.7	18.00	0.074	2.8	20.6	2.4	69.7	0.8	0.52	4.89	30162	6342	36624
BC03	39.371	285	17.9983	0.074	2.8	20.5	2.5	69.8	0.8	0.52	0.22	30162	6262	36624
BC04	43.337	259.8	18.26	0.079	2.3	16.7	2.7	75	0.6	0.52	0.96	27962	5735	19750
FC00	33.476	1.8	17.55	0.013	13.9	60.4	1.6	5.8	16.7	0	0	0	0	20251
FC01	31.61	33.2	18.40	0.031	33.1	42.2	0.7	15.5	16.1	2.77	5.14	7090	7090	73404
FC02	37.102	14.7	21.22	0.067	27.5	21.3	0.9	45.1	10.1	0	0	0	0	30773
FC03	38.342	13.9	21.22	0.069	25.1	21.6	1.4	47	10.1	0	0	0	0	30773
FC04	38.86	12.5	21.22	0.073	18.9	23.3	1	52.1	7.5	0	0	0	0	23579
FC05	40.475	11	21.22	0.080	14.1	22.2	1.2	57.8	5.1	0	0	0	0	19010
SC01	41.267	8.6	21.41	0.058	66.2	9.1	1.6	19.6	22.7	0	0	0	0	65270
SC02	41.844	8.5	21.41	0.568	66.1	8.8	1.6	19.9	22.4	0	0	0	0	57550
SC03	42.612	7.8	21.41	0.060	63.4	9.3	1.8	21.7	21.6	0	0	0	0	50451
SC04	43.805	6.6	21.41	0.069	56.6	11	2	25.7	18.7	0	0	0	0	39739
SC05	44.398	6.4	21.41	0.070	55.4	11.3	2	26.4	15.3	0	0	0	0	36253
SC06	45.171	4.5	21.4092	0.076692	47.7	12.5	2.9	30.7	15.3	0	0	0	0	28582
SC07	45.773	3.7	21.4092	0.0902621	35.4	15.2	3.7	37.9	12.1	0	0	0	0	23710

Figure 3 provides box plots of the *E. coli* data collect from 2018-2019 and characterizes *E. coli* concentration from upstream to downstream. Each reach is depicted separately to better identify trends between the Cache la Poudre River and its tributaries. As a general trend, the mean was much higher than the median and geometric mean showing that, while extreme events skewed the average towards higher concentration, the geometric mean more accurately represented the bulk of the data.

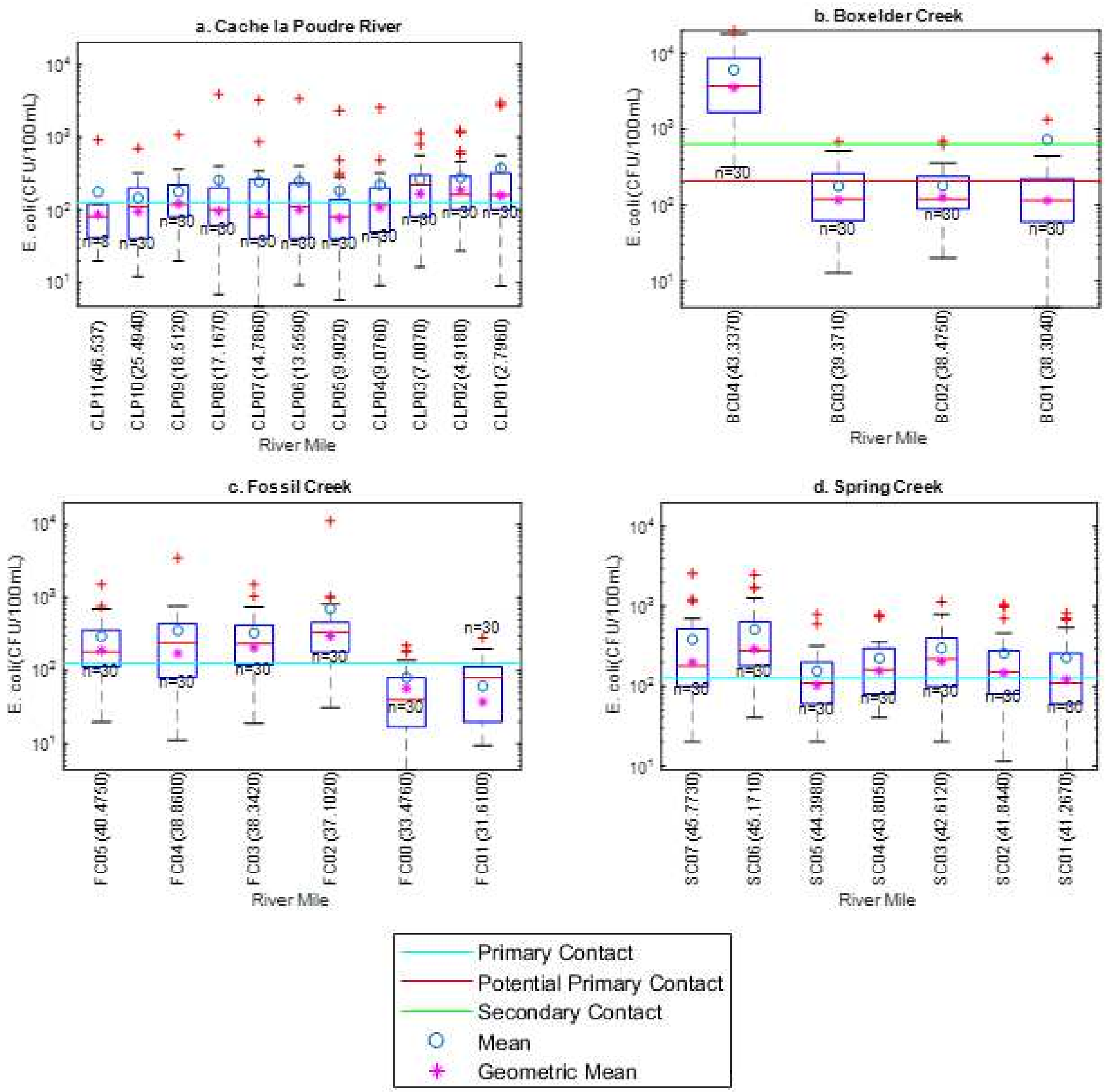


Figure 3 Concentrations of E. coli during 2018 and 2019 period along the Cache la Poudre River and its tributaries for (a) Cache la Poudre River, (b) Boxelder Creek, (c) Fossil Creek, and (d) Spring Creek. Sites are ordered by the river distance to the downstream confluence with the South Platte River. On each On each box, the red line in the center of the box is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most

extreme data points not considered outliers, and outliers are plotted individually. The circle represents the mean concentration. The blue, red, and green line represents the numeric standard.

The Cache la Poudre sites (Figure 3.a) show that, as a general trend, *E. coli* concentration tended to increase as the water moves downstream, likely due to increased human activity. The *E. coli* concentrations increased gradually until the concentrations were reduced at CLP05. This could be due to influent water from Greeley No. 3 Ditch diluting the concentration of pathogens. After CLP05 as the water moves downstream the concentrations of *E. coli* tended to increase. The geometric mean for sites is below the numeric standard until the water reaches CLP03, at this point bacterial concentration exceeded the water quality limit.

Sites on Boxelder creek (Figure 3.b) had an initially high concentration of *E. coli*, BC04 is located near a small farm with several cattle and other grazing animals that regularly drink and wade in the water. This was likely the cause of the extreme degradation in this segment of the stream. However, as the water moved downstream the *E. coli* attenuated and the concentrations reduced. BC02 and BC03 are upstream of the Boxelder Sanitation District effluent and show that the water is typically of an acceptable quality before the discharge location. Background *E. coli* levels at BCO1 were generally within the limits, however, there were several extreme events that had concentrations that were several magnitudes higher than the geometric mean or the standard.

The *E. coli* levels for Fossil Creek (Figure 3.c) experience gradual increased in concentration until the water reaches the Fossil Creek Reservoir. Water discharged out of the reservoir into Fossil Creek (FC01) and Fossil Creek Reservoir Outlet (FC00) had reduced bacterial loads likely due to the attenuation that occurred as result of retention in the reservoir. Spring Creek shows similar trends to Fossil Creek. The *E. coli* levels on Spring Creek (Figure 3.d) increased with distance downstream from SC07 to SC06. SC05 experienced a decreased *E. coli* concentration after being retained in Privy Pond. The *E. coli* concentration continued to increase until SCO3. As the water moves from SC03 to SC02 the water is retained in a small pond in Eldora Park before being released back into the main channel.

Vulnerability was calculated as a function of ambient *E. coli* concentrations and state level water quality standards. Table 3 shows the results of the vulnerability analysis for each site. The vulnerability for sites where V=1 does not indicate the magnitude of excursion.

Table 3 Results of vulnerability analysis for E. coli assuming Colorado Regulation 31 state level standards

Location	River Mile	Vulnerability	Geometric Mean	Mean	Median
			-----------CFU/100mL------------		
CLP 01	2.80	0.83	186.48	372.20	161.50
CLP 02	4.92	0.99	196.77	274.80	163.50
CLP 03	7.01	0.93	204.08	260.20	220.00
CLP 04	9.08	0.23	131.25	218.90	120.00
CLP 05	9.90	0.01	105.50	183.00	80.00
CLP 06	13.56	0.16	149.22	243.47	110.00
CLP 07	14.79	0.07	108.94	243.33	80.00
CLP 08	17.17	0.11	132.46	255.67	100.00
CLP 09	18.51	0.41	144.56	177.00	120.00
CLP 10	25.49	0.06	125.33	143.43	110.00
CLP 11	46.57	0.17	87.63	177.13	80.00
BC 01	38.30	0.03	144.18	728.54	115.00
BC 02	38.48	0	130.80	177.87	120.00
BC 03	39.37	0	249.80	176.17	120.00
BC 04	43.34	1.00	4752.90	6044.00	3750.00
FC 00	33.48	0	48.87	57.50	40.00
FC 01	31.61	0	65.11	79.33	80.00
FC 02	37.10	1.00	361.36	707.67	335.00
FC 03	38.34	0.99	230.03	327.67	235.00
FC 04	38.86	0.92	231.21	353.33	240.00
FC 05	40.48	0.99	218.44	295.47	180.00
SC 01	41.27	0.41	125.08	225.73	110.00
SC 02	41.84	0.76	149.45	258.00	150.00
SC 03	42.61	1.00	228.84	298.00	220.00
SC 04	43.81	0.90	178.87	224.33	160.00
SC 05	44.40	0.10	116.83	155.00	110.00
SC 06	45.17	1.00	406.50	513.43	280.00
SC 07	45.77	0.98	262.06	382.63	180.00

As a general trend, increased anthropogenic influence increased vulnerability to *E. coli* pollution. Vulnerability fluctuated due to factors that were not described in the geospatial analysis. Unlike nutrients, which persist in the system, *E. coli* does not multiply in water (Standridge, 2008) and will die off over time. This rate of die off is difficult to determine due to numerous influences, because of this, it is challenging to model accurately (Baudisova, 1997). Inflow to the Cache la Poudre River from various irrigation canals could be another source of variation in vulnerability. When inflows from these canals have better ambient quality than the receiving stream it acts to dilute the instream pathogen load.

Boxelder creek experienced improved water quality with distance downstream and distance away from farms and other small non-point sources of pollutants. As Boxelder Creek nears the confluence with the Cache la Poudre River, the water quality improved to the point where there is no vulnerability to exceedances. Vulnerability increased downstream of the Boxelder Creek Sanitation District's effluent outflow. The quality of the outflow decreased the ambient instream water quality and in turn increase the vulnerability at this location.

Fossil Creek experienced increased vulnerability until water reached Fossil Creek Reservoir, as identified previously retention appears to allow adequate time for bacterial die off. The sites downstream from the outflow were not vulnerable to impairment. Spring Creek experienced similar fluctuations in vulnerability that seem to be the result of retention in small ponds.

2.3.2 Variation in water quality exceedances due to precipitation events

There was a significant difference between the number of samples that exceed the water quality limit during rain (Figure 4) and non-rain (Figure 5) events for CLP11. This indicates that stormwater considerably impacts the concentration of instream *E. coli*. Stormwater runoff events can mobilized and transport non-point source microbials to the stream or, in some cases, increased flow due to precipitation can lead to re-suspension of microbes in sediment (Kistemann et al., 2002; Rossi et al., 2020; Signor et al., 2005; Sylvestre et al., 2020). E. coli has been found to survive longer in sediment than in the overlying waterbody, mainly due to the protecting the sediment provides from U.V. exposure and predators (Garzio-Hadzick et al., 2010). CLP11 is unique in that it is located the furthest upstream and the contributing watershed is largely undeveloped, natural land. This watershed has fewer point-source inputs than other locations in this study resulting in a lower baseline bacterial concentrations and increased impact of non-point sources during rain events.

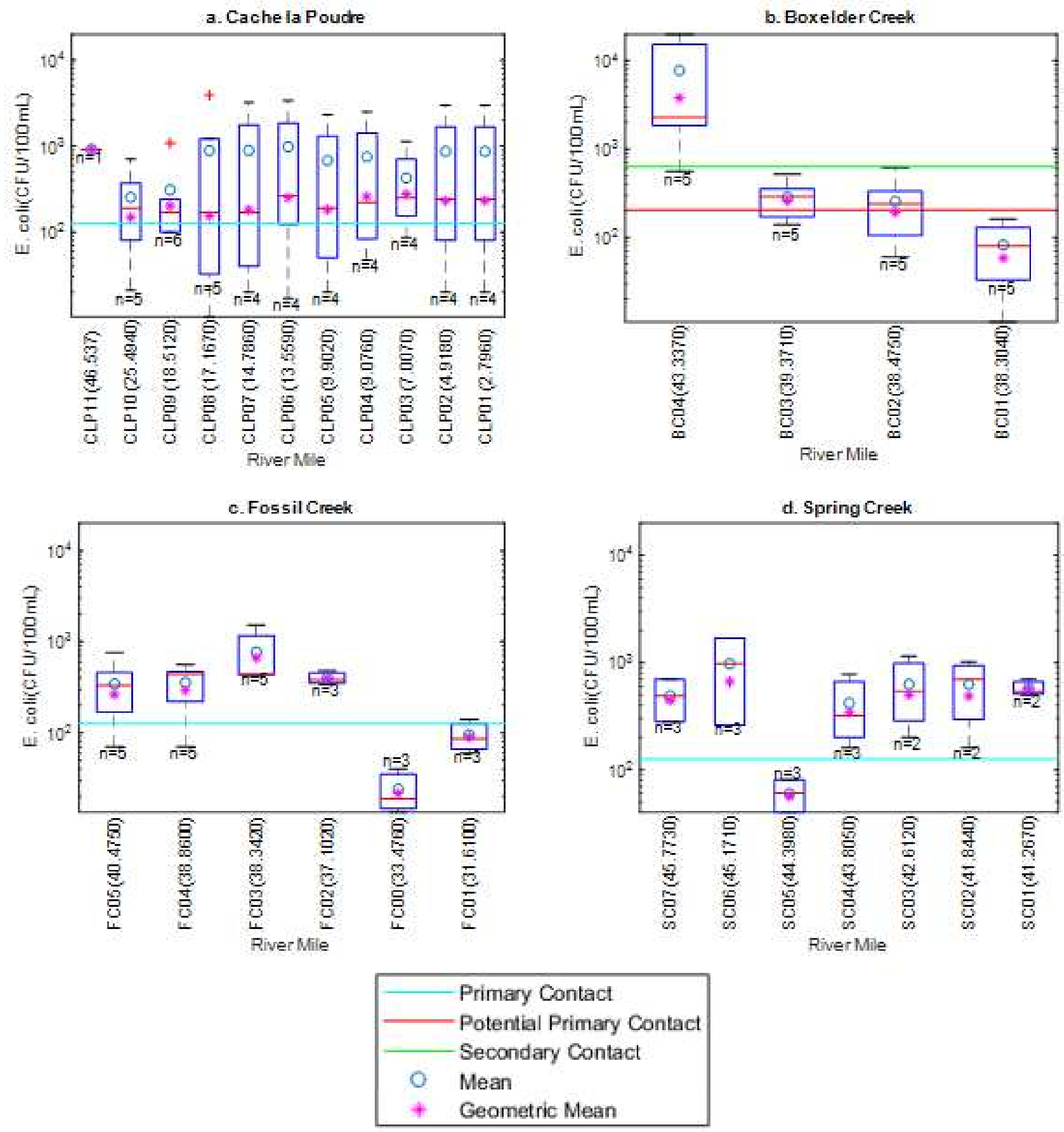


Figure 4 Concentrations of E. coli during rain events that occurred during the 2018 and 2019 sampling periods along the Cache la Poudre River and its tributaries for (a) Cache la Poudre River, (b) Boxelder Creek, (c) Fossil Creek, and (d) Spring Creek. Sites are ordered by the river distance to the downstream confluence with the South Platte River. On each On each box, the red line in the center of the box is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. The circle represents the mean concentration. The blue, red, and green line represents the numeric standard. N equals the sample size.

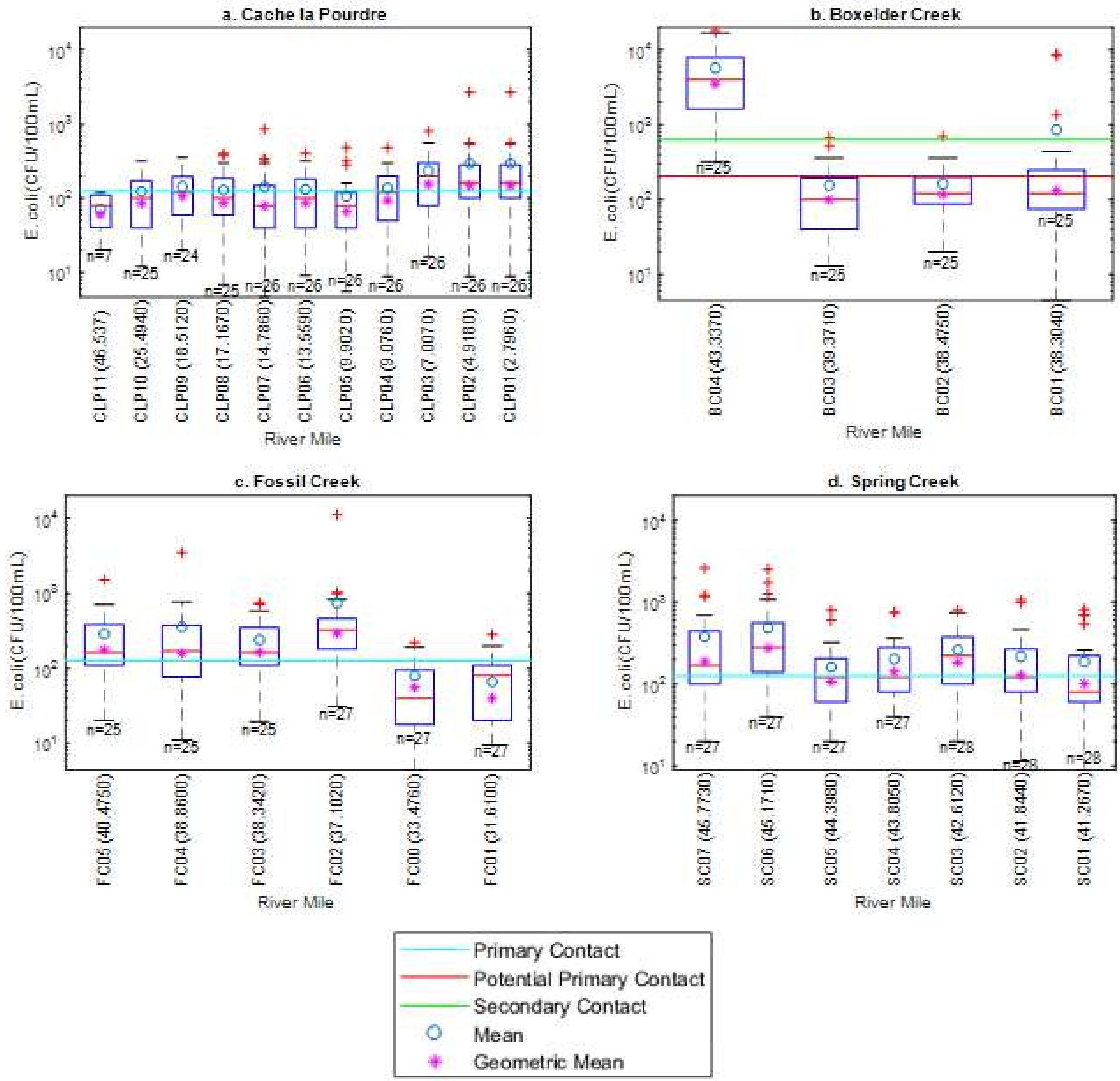


Figure 5 Concentrations of E. coli during non-rain events that occurred during the 2018 and 2019 sampling periods along the Cache la Poudre River and its tributaries for (a) Cache la Poudre River, (b) Boxelder Creek, (c) Fossil Creek, and (d) Spring Creek. Sites are ordered by the river distance to the downstream confluence with the South Platte River. On each On each box, the red line in the center of the box is the median, the edges of the box are the 25th and 75th percentiles, the whiskers extend to the most extreme data points not considered outliers, and outliers are plotted individually. The circle represents the mean concentration. The blue, red, and green line represents the numeric standard. N equals the sample size.

The remaining sites in the study did not exhibit a statistical difference between the number of samples that exceed the water quality standards for rain and non-rain events. However, these sites can still be used as a tool to help identify significant sources of pollution. Precipitation lead to increased *E. coli* concentrations at most of the sites.

There are several sites that do not respond to precipitation with increased *E. coli* concentrations. BC01 shows that precipitation events caused a decrease in *E. coli* concentration. This site is located below the Box Elder Sanitation District effluent outlet. Water quality at this site is dominated by the quality of the effluent. The contributing watershed between BC01 and BC02 is comprised of grasslands, forest, and wetland. Wetlands and emergent vegetation can effectively remove fecal microbes (Rossi et al., 2020) and improve the quality of runoff. The input of this runoff did little to introduce microbes into the system and instead worked to dilute the bacterial concentration of the effluent.

Retention appeared to be a considerable factor in reducing *E. coli* concentrations during precipitation events. FC00 is located downstream of Fossil Creek Reservoir and SC05 is located downstream of Privy Pond. FC00 is located on an irrigation ditch and most of the contributing watershed is used for agricultural production. This site consistently had the lowest *E. coli* concentrations, evidence that the quality of water exiting the reservoir is low in bacterial pollutants bacterial load does not increase due to non-point sources. Retention in Privy Pond upstream of SC05 also worked to reduce the concentration of *E. coli* in the effluent water. For both cases retention attenuated the *E. coli* load and precipitation produced runoff dilutes instream concentrations.

2.3.3 Multiple linear regression models for *E. coli* concentrations

The regression analyses for the Cache la Poudre River and the Cache la Poudre watershed were performed using different combinations of hydrologic variables, non-point source land use variables, and point source anthropogenic variables. The MLR analysis was performed to analyze four *E. coli* response variables: vulnerability to impairment, mean, geometric mean, and median *E. coli* concentrations. Correlations between the variables were identified using scatter plots that showed how mean, and median concentrations were transformed with power functions using a box-cox transformation, where a lambda value of 0 is a log transformation. The lambda values for the box-cox transformations are listed in Table 4 and Table 5.

In general, WWTP and IDW WWTP capacity were not as strong of indicators of urban influence as percent urban, percent impervious and population. This is likely because WWTPs are required to disinfect the water before it can be discharged into the receiving water body. WWTP capacity and CAFO capacity increase cumulatively with distance downstream leading to the significant collinearity in the models (represented by VIF), indicating that using both point sources to describe the response variables is likely not the best choice. Alternatively, non-point sources can be used to represent urban and agricultural influence. Percent cropland is highly correlated with CAFO capacity due to the application of manure as fertilizer. In some cases, percent cropland could be used as a surrogate for the impact of agriculture in areas without CAFO presence.

Strong ($R^2 > 0.7$) and significant (*p-value* < 0.05) correlations were found between the *E. coli* response variables and anthropogenic factors for sites on the Cache la Poudre River (Table 4). A complete table of all diagnostic test performed can be found in Appendix C. Mean *E. coli* concentration was highly correlated to percent rangeland and CAFO capacity. Geometric mean *E. coli* concentration was related to CAFO capacity. Median *E. coli* concentration and vulnerability to impairment were both correlated with percent rangeland and inverse distance weighted CAFO capacity. Increase in percent rangeland use was associated with a decrease in *E. coli* load and vulnerability. This could indicate that little grazing occurs and natural land inputs little *E. coli* into the system. Increases in CAFO capacity, both inverse distances weighted and cumulative, lead to increases in *E. coli* loads. The increased load could be associated with runoff from feeding operations, small scale hobby farms, or the application of manure as fertilizer.

Table 4 Cache la Poudre River multiple linear regression models for mean, geometric mean (geo mean), median, and vulnerability (vul) with precipitation (Precip), percent urban (Urban), percent rangeland (Range), and percent *forest (forest). With R^2 and adjusted R^2, P value for the appropriateness of the model, lambda, N, and degrees of freedom (DOF).*

Model	Linear Model	R^2	Adj. R^2	P	λ	N	DOF
Mean	$8.00 - 0.07(Range) + 6.18 * 10^{-6}(CAFO)$	0.72	0.65	6E-3	0	11	8
Geo Mean	$4.47 + 1.37 * 10^{-6}(CAFO)^{1.15}$	0.73	0.70	8E-4	0	11	9
Median	$9.69 - 0.16(Range) + 3.0 * 10^{-4}(CAFO\ IDW)$	0.73	0.66	6E-3	0	11	8
Vul	$7.51 - 0.25(Range) + 3.8 * 10^{-2}(CAFO\ IDW)^{0.5}$	0.78	0.72	2E-3	1	11	8

Figure 6 shows mean, geometric mean, median, and vulnerability as a function of the anthropogenic variables described in Table 4. The slope of the contour plots for vulnerability, mean, and median *E. coli* concentration are more sensitive to changes in CAFO capacity, than in percent rangeland. Indicating that animal waste and the effect of agriculture greatly effect water quality along the CLP river.

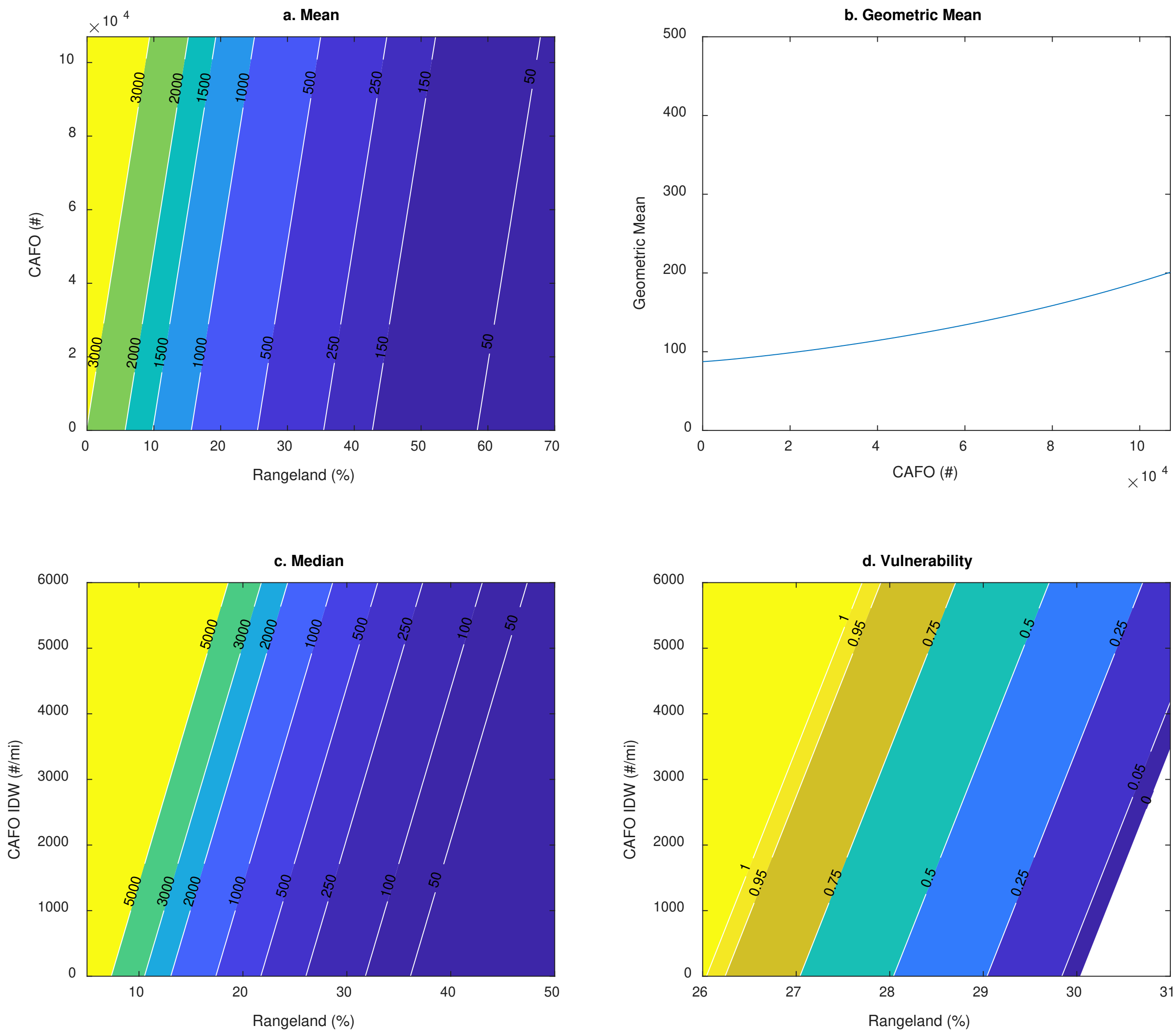


Figure 6 Contour plots for sites on the Cache la Poudre River for (a) mean, (b) geometric mean, (c) median, and (d) vulnerability as functions of different anthropogenic variables

When analyzing all the study sites in the Cache la Poudre watershed, there are several sites that are not influenced by CAFO or WWTP inputs. Because of this, land use may more accurately described *E. coli* variations. There is very little collinearity between land use factors used in this analysis, indicating that it is possible obtain significant models when the analysis preformed with several land use parameters.

The MLR analysis for the Cache la Poudre watershed was performed using different combinations of hydrologic and land use predictor variables (Table 5 and Appendix C). Precipitation was found to be an important indicator for estimating *E. coli* response. For each model increased precipitation correlated to increased E. coli load suggesting that bacteria are collected from urban and agricultural areas in runoff and transported to streams. Strong (R^2 > 0.7) and significant (*p-value* < 0.05) MLR models were found for the geometric mean and median *E. coli* concentrations. These models included average annual precipitation, percent forest, and percent rangeland. As percent forest increases the *E. coli* load decreases, signifying that very little *E. coli* enters the system through natural land. Increases in percent rangeland correlates with increased *E. coli* loads. Grazing activities occurring on rangeland likely increased *E. coli* concentrations. Mean *E. coli* concentration was related to average annual precipitation, percent urban land use, and percent rangeland. Like the models for geometric mean and median *E. coli* concentrations, increases *E. coli* mean concentrations can be associated with grazing activities occurring on rangeland. Increased urban land use was also found to lead to increased *E. coli* loads. Waste produced by domesticated animals was found a potential source of *E. coli* in urban areas (Selvakumar & Borst, 2006).

Table 5 Cache la Poudre Watershed multiple linear regression models for mean, geometric mean (geo mean), median, and vulnerability (vul) with precipitation (Precip), percent urban (Urban), percent rangeland (Range), and percent forest (forest). With R2 and adjusted R2, P value for the appropriateness of the model, lambda, N, and degrees of freedom (DOF).

Model	Linear Model	R^2	Adj. R^2	P	λ	N	DOF
Mean	$-0.68 + 0.23(Precip) + 0.26(Urban)^{0.25} + 0.03(Range)$	0.60	0.54	9E-5	0	27	23
Geo Mean	$0.9768 + 5.5 * 10^{-3}(Precip)^2 - 4.5 * 10^{-4}(Forest)^{1.85} + 0.66(Range)^{0.3}$	0.70	0.67	3E-6	0	27	23
Median	$-8.10 + 0.23(Precip) - 2.76 * 10^{-4}(forest)^2 + 7.14(Range)^{0.05}$	0.71	0.68	2E-6	0	27	23
Vul	$-9.83 + 0.34(Precip) + 2.44(Crop)^{0.1} - 6.1 * 10^{-5}(Forest)^2$	0.69	0.65	5E-6	1	27	23

Figure 7 shows mean, geometric mean, and median *E. coli* concentration, and vulnerability as a function of the anthropogenic variables described in Table 5. The slope and gradient of the contour plots for mean *E. coli* concentration shows that mean concentration is more sensitive to changes in percent rangeland than in percent urban. Figure 7.b and 7.c shows that geometric mean and median *E. coli* concentration are more responsive to changes in in percent forest than in percent rangeland. Vulnerability to impairment is more sensitive to changes in percent forest than in percent cropland. These results could signify that the conversion of natural land greatly increase instream *E. coli* loads.

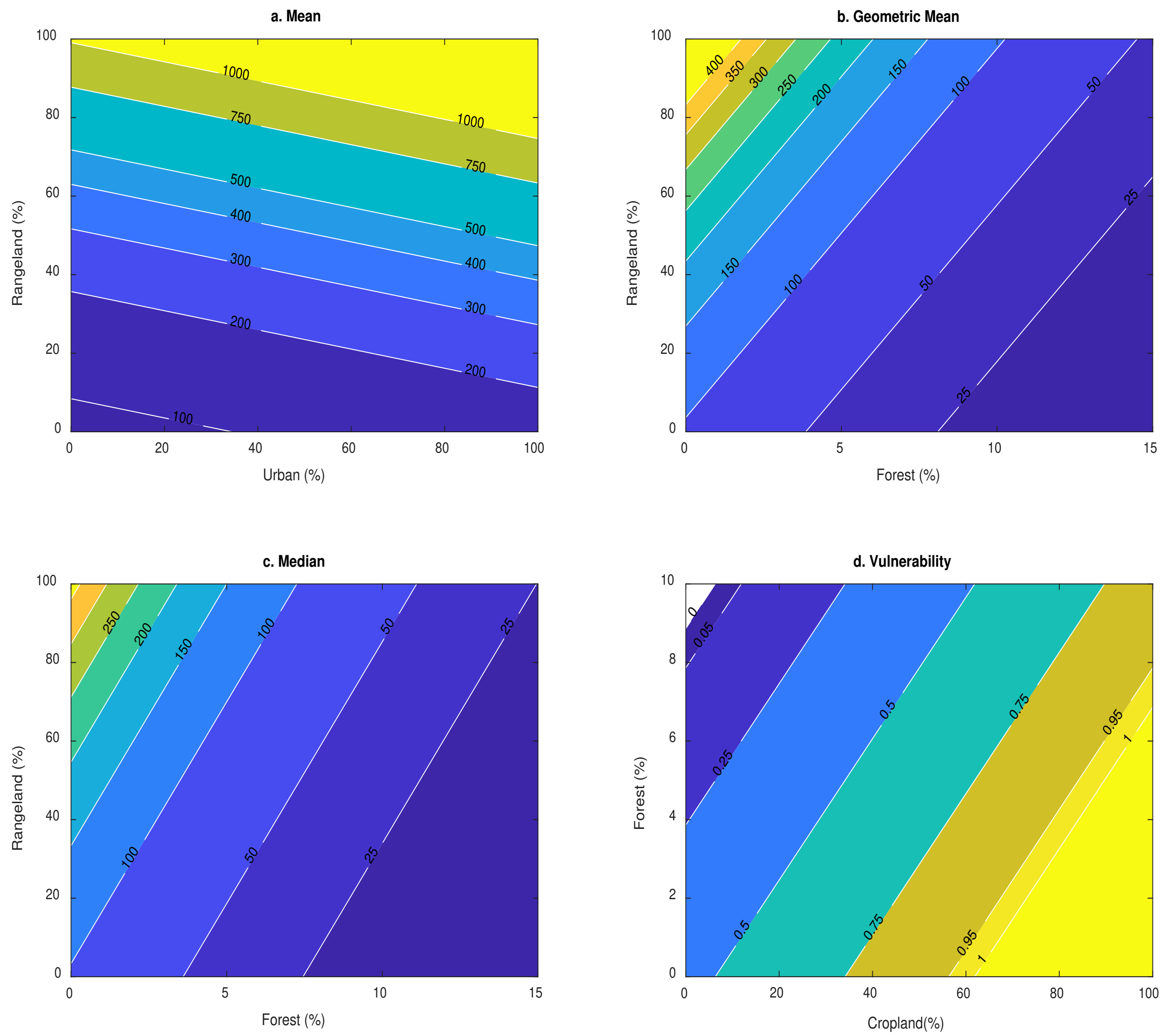


Figure 7 Contour plots for sites on the Cache la Poudre River Watershed for (a) mean, (b) geometric mean, (c) median, and (d) vulnerability as functions of different anthropogenic variables

2.4 Conclusion

An in-depth investigation of *E. coli* concentrations was performed in order to identify the impact of point and non-point sources that could be attributed to increased pollutant concentrations. Probabilistic methods for quantifying *E. coli* vulnerability, mean, and median concentrations were developed in this research. Precipitation analysis was also performed in order to identify the relative importance of point and non-point source pollutant sources. Regression models were adopted to analyze the effect of anthropogenic and natural influences on *E. coli*.

The Cache la Poudre Watershed generally experienced increased *E. coli* concentrations from upstream to downstream due to increased anthropogenic activity. However, several outside factors were identified that cause ambient bacterial concentrations to decrease. CLP05 experienced a decrease in concentration related to influent from Greeley No. 3 Ditch diluting the instream concentration of microbes. Sites downstream of reservoirs or retention ponds (FC00, FC01, SC05, and SC02) showed a decrease in *E. coli* concentration from the site directly upstream. Unlike nutrients which persist in the environment, *E. coli* is only able to survive a short time outside the host and is unable to multiply in the environment. Retention allows adequate time for microbial attenuation before sampling occurs at the downstream site.

Vulnerability was calculated based on Colorado Regulation 31 *E. coli* water quality limits and observed *E. coli* concentrations. Vulnerability tends to increase from upstream to downstream. One noticeable exception was Boxelder Creek, here vulnerability started out at V=1 and decreased with distance downstream. BC04 is located near a small farm with livestock that freely enter the stream. This leads to high bacterial concentrations in the waters. As the water moves downstream, the *E. coli* attenuated in the system until the vulnerability was nearly 0. Vulnerability increased again just downstream of the Boxelder Creek Sanitation District effluent, where concentrations sometimes exceeded instream limits. Like concentration, vulnerability decreased due to inflow from irrigation canals and retention.

There was a significant difference in the proportion of samples that exceeded the numeric water quality standards for *E. coli* during rain and non-rain events at CLP11. *E. coli* concentrations at this site during rain events were significantly higher than during sampling trips with no antecedent precipitation. The other sites in the study did not experience a significant difference in water quality exceedance, they did however provide insight on trends that occur. In general, locations that are dominated by non-point sources experience increases in *E. coli* concentrations during rain events. Water quality at BC01 was significantly influenced by point sources. Effluent from Boxelder Creek Sanitation District controlled the water quality downstream. During rain events *E. coli* concentrations were lowered as a result of dilution.

The relationship between anthropogenic influence and *E. coli* response in the Cache la Poudre watershed was explored with multiple linear regression models. The MLR approach was used to predict mean, geometric mean and median *E. coli* concentrations, as well as, vulnerability to *E. coli* impairment for sites along the Cache la Poudre River and in the Cache la Poudre watershed. Hydrologic and natural land cover descriptors, and urban and agricultural predictor variables were able to sufficiently describe the *E. coli* response.

When considering the regression models for sites along the Cache la Poudre River, CAFOs were a significant indicator of bacterial water quality for all the response variables. There was a significant amount of collinearity between point source predictor variables, because of this WWTP capacity and CAFO capacity were not be used in the same regression model. Land use was proven to be a good surrogate for point source predictor variables and because of the limited collinearity between the land use variables used, several of these variables could be used together to produce a valid MLR model.

Twelve of the 28 sites studied were not impacted by CAFOs or WWTPs, indicating that these predictor variables were not the best indicators for the regression model. When analyzing the Cache la Poudre watershed, average annual precipitation, percent urban land use, percent cropland, percent rangeland, and percent forest strongly correlated to the response variables. Adopting land use improved the overall performance of the individual regression models because it can more accurately describe changes in the unimpacted watersheds.

Land use data are much easier to obtain and process than WWTP and CAFO capacity data, a regression model that uses land use may be more practical and would allow this approach to be more easily adopted in future studies. Models developed for using this methodology can help researchers and regulators accurately predict instream E. coli concentrations and identify which streams are at risk of impairment. Regulators will be to save both time and money by focusing on developing monitoring programs for streams that are impaired or at risk of being impaired.

The identification of sources of pollution done in this work could also assist future studies in the development of best management practices for agricultural land use, urban land use, and confined animal feeding operations. Improvements to green infrastructure could also be made to increase the retention time before stormwater runoff is discharge back into streams and rivers.

Rapid urbanization and the degradation of surface water systems has given increased importance to understanding the relationships between anthropogenic influence and water quality degradation. This work demonstrates that watershed, hydrologic, urban, agricultural, and natural characteristics can be used to predict instream *E. coli* concentrations and vulnerability to impairment. However, future research is recommended to further explore this approach in other geographic areas. Additional sampling throughout the year would also provide additional data to help illustrate fluctuations in instream concentrations throughout the entire year.

CHAPTER 3: MODELING BASELINE NUTRIENT CONCENTRATIONS IN STREAMS ACROSS COLORADO

3.1 Background

While nutrients are critical to maintaining a healthy ecosystem, excess amounts can lead to eutrophication, the decline of species diversity, and increase in risk to human health due to harmful algal blooms(Klein, 1979; United States Environmental Protection Agency, 2001). Cultural eutrophication, (human-caused inputs of excess nutrients in waterbodies) is one of the primary factors resulting in impairment of surface waters (USEPA, 2000). Excess nutrient concentrations have led to the degradation of 40% of the U.S. rivers and streams (USEPA, 2000). The Clean Water Act (CWA) was established to combat these issues and improve surface water quality, through the establishment of water quality standards and the development of the Total Maximum Daily Load (TMDL) program.

Under the CWA, states were tasked with the creation, implementation, and monitoring of water quality standards. In 2012, the Colorado Department of Public Health and Environment (CDPHE) implemented numeric nutrient limits for surface water in order to reduce the eutrophication of surface waters of Colorado (CDPHE, 2013). The nutrient standards are determined based on the designated use of the water body. Surface waters are classified by warm or cold-water aquatic use. Warm water rivers and streams support biota that exist in waters with average summer temperatures that exceed 20 °C, while cold waters support biota that thrive in waters where this threshold is not normally exceeded. Warm water median nitrogen concentration is limited to 2.01 mg/l and the annual median phosphorous concentration is limited to 0.17 mg/l (CDPHE, 2013). While the annual median concentrations for cold water are 1.25 mg/L and 0.11 mg/L for total nitrogen and total phosphorous respectively (CDPHE, 2013).

Nutrient studies

Understanding the interaction between land use and water quality is crucial to investigating water quality degradation and assigning water quality regulations. This is especially true in watersheds with a diverse combination of land uses and point source pollutants. Population growth, urbanization and the conversion of natural land cover to farmed and grazed lands are commonly identified some of the main drivers stream impairment (Kang et al., 2010; Klein, 1979; Wickham et al., 2000). These anthropogenic activities are directly reflected in the landscape and allow morphological, land use, and point source attributes of the contributing watershed to be linked to biological and chemical water quality characteristics (Kang et al., 2010).

Multiple linear regression models have been used to successfully related anthropogenic influences to increased nutrient concentration in receiving rivers and streams. This relatively simple approach allows for the characterization of sources of variability in water quality data over a large geographic region. Many studies have found strong and significant relationships between anthropogenic activities and instream nutrient load (Francy et al., 2000; Sylvestre et al., 2020; Tasdighi et al., 2017; Wickham et al., 2000; Williams et al., 2014). Several studies have concluded that percent urban land use was the main cause of pollution (Chang, 2008; Francy et al., 2000; Tasdighi et al., 2017). Williams (2014) found that for nutrient concentrations in urban settings were dominated by waste discharge during periods of low flows (Williams et al., 2014). However, most studies agreed that the degraded quality of a stream is not the result of any single factor, but the combined interaction of different sources of anthropogenic influence (Klein, 1979; McMahon & Cuffney, 2000).

Modeling has been performed on a large scale to show that nutrients can be accurately predicted on the river basin scale. Several studies performed in the Han River basin have shown that nutrient concentration is positively correlated with urban and agricultural land cover (Chang, 2008; Li et al., 2009). In a study performed by Bouraoui, the model developed was able to predict a range of nitrate concentrations in surface water (Bouraoui et al., 2005). These studies show that nutrient concentrations can be accurately be predicted on the basin scale. This research takes that principle a step further to develop regression models for regions in Colorado. The models developed were applied to all stream segments in Colorado.

In this study, baseline ambient nutrients pollution in Colorado was characterized using regression models. The baseline concentration is the present nutrient level prior to any anticipated improvements in nutrient management due to new regulations. The specific objectives are to assist the CDPHE water quality control division in: (i) examining trends in nutrient concentration in the different regions in Colorado and (ii) expanding upon the multiple linear regression framework to develop regression models that can predict base level nutrient concentrations for stream segments in three regions of Colorado. The models developed will be used to assist in the identification of water bodies that are impaired or at risk for being impaired and allow for the development of water quality monitoring programs for these stream segments.

3.2 Methodology

A complete geospatial analysis was performed in order to collect quantify anthropogenic sources of water quality impairment. These point and non-point sources of pollution were assessed alongside nutrient concentrations in an exhaustive multiple linear regression (MLR) analysis to relate baseline nutrient concentrations to anthropogenic influences in three regions of Colorado. The regression models were used to predict the baseline nutrient concentrations for every stream segment in Colorado.

3.2.1 Study area

The state of Colorado is divided into 7 water divisions: The South Platte; Arkansas; Rio Grande; Gunnison; Colorado; Yampa, White, and North Platte; and San Juan and Dolores. A total of 89 sampling locations upstream of WWTPs were selected for this study, of these 30 had flow measurement. Because the spread of sampling locations was not equal across all divisions, the divisions were combined into three groups: Division 1 (South Platte River Basin), Division 2 (the Arkansas River Basin), and the

Western Slope. There are 41 upstream sites located in Division 1, 16 located Division 2, and 32 located on the western slope (Figure 8).

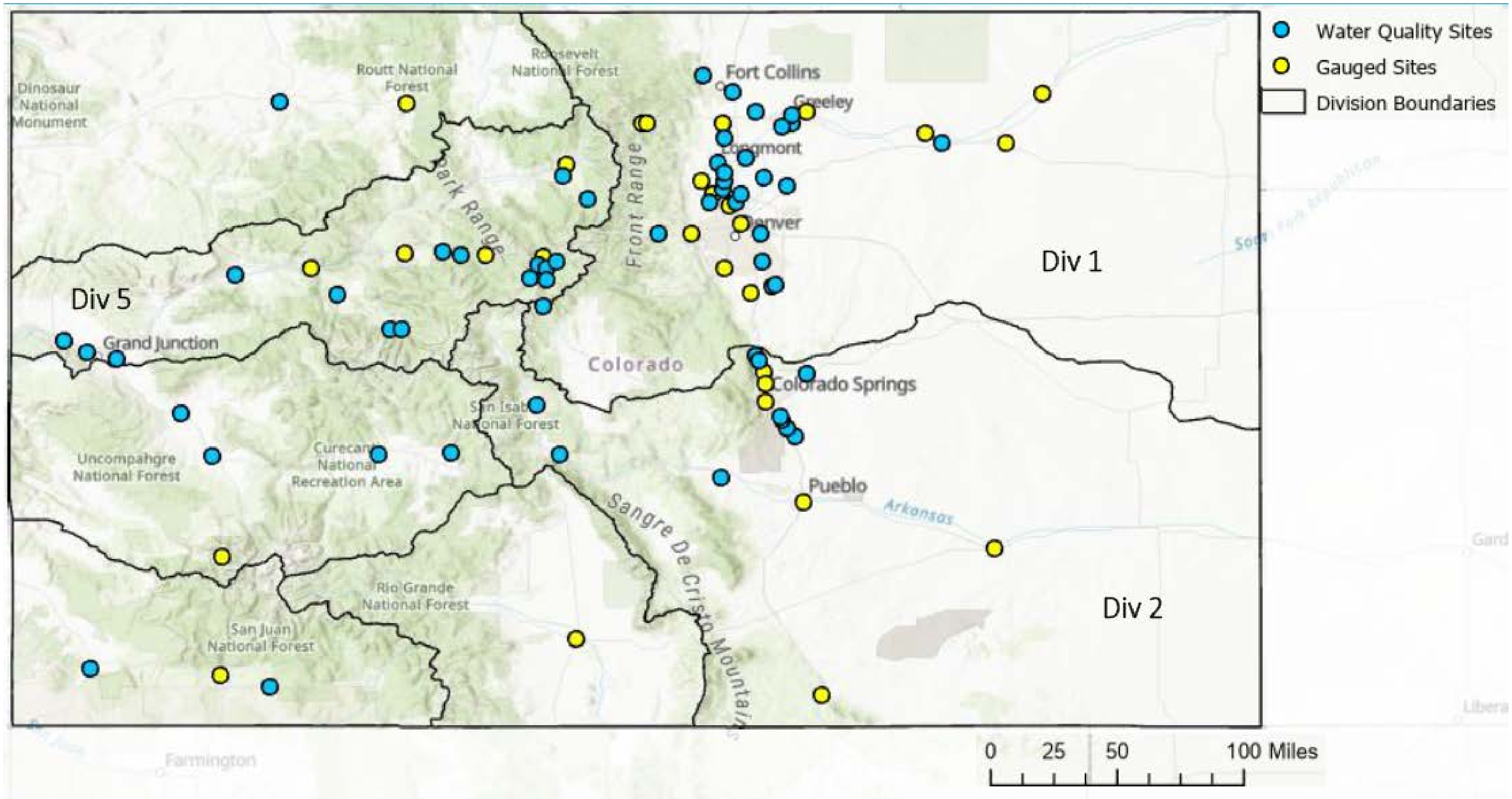


Figure 8 The nutrient sampling locations. The blue dots represent all water quality sites, while the yellow dots represent gauged locations. The black lines represent the different water divisions in Colorado

84 sites downstream of WWTPs were also analyzed in this study to ensure the accuracy of baseline nutrient predictions. 41 of these sites were in Division 1, 16 located Division 2, and 27 on the western slope.

3.2.2 Water quality data

Under the CWA, many states and cities around the United States have implemented extensive monitoring programs to assist in the development of numeric water quality standards and TMDLs. Publicly available water quality data was obtained for TN and TP from the Colorado Regulation 85 data set (Water Quality Monitoring Council, 2020). The data used for this study was collected from April 2013 to December 2018. This study includes data from locations both upstream and downstream of WWTPs.

3.2.3 Geospatial analysis

Nutrient concentrations were related to various indicators of anthropogenic influence. The entire contributing watershed was considered to account for the cumulative effect of both point and non-point sources of pollution. The contributing watershed for each sampling location was delineated using the Environmental Resource Assessment and Management System (eRAMs) (One Water Solutions Institute, 2020). The drainage basins and watershed slope were found using a 1/3 arcsec digital elevation model from the National Elevation Dataset and the National Hydrography Dataset (NHD) from the USGS. These factors along with others were used in an extensive analysis of geospatial factors was completed to quantify both natural and human causes eutrophication.

3.2.3.1 Characterizing anthropogenic indicators

Many point and non-point sources lead to eutrophication in streams, including wastewater treatment plants, confined animal feeding operations, land use, population density, and various watershed and hydrologic characteristics (Table 6). These sources of pollution where used to characterize instream concentrations of nutrients. Using a variety of factors to describe the impact of urban, agricultural, and natural sources of pollutants reduces the collinearity of the regression models.

Table 6 Summary of anthropogenic predictor variables employed in multiple linear regression models

Variable	Type	Unit
WWTP Capacity	Point Source/Facilities	MGD
WWTP Technology Type	Treatment Type	-
CAFO Capacity	Point Source/Facilities	# animals
Population	Non-Point Source	People/mi^2
% Impervious Surface Cover	Land Use	%
% Urban land Use	Land Use	%
% Crop Land Use	Land Use	%
% Forest Land Use	Land Use	%
% Range Land Use	Land Use	%
Watershed Area	Watershed Characteristic	mi^2
Slope	Watershed Characteristic	ft/ft
Average Annual Precipitation	Watershed Characteristic	in
Atmospheric Deposition	Watershed Characteristic	kg N/ha

The Watershed Rapid Assessment Program (WRAP) tool in eRAMs (One Water Solutions Institute, 2020) was used to obtain capacities of WWTPs and CAFOs, atmospheric deposition, precipitation, population, and land use data. Urban land use was defined as a combination of low, medium, and high intensity developed land and open spaces; cropland is a combination of pasture, hay, and cultivated crops; forest land cover includes deciduous, evergreen, and mixed forest; rangeland is the combination of grassland, shrub, and scrub (Multi-Resoulution Land Characteristic Constortium, 2016). Tract level population data was collected from the U.S. Census and used to calculate population density (U.S. Census Bureau, 2019).

The eRAMs platform was used to extract the geospatial data. The wastewater treatment plant flow capacity was extracted from Colorado Department of Public Health and the Environment (CDPHE) Regulation 85 data (Water Quality Monitoring Council, 2020). The WWTP technology type was determined for the WWTP directly upstream of the sampling location according to Regulation 85 data (Water Quality Monitoring Council, 2020). The treatment types were categorized based on treatment type (primary, secondary, tertiary). The confined animal feeding operation capacity was calculated as the sum of animal's units.

Average annual precipitation for each site was procured from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) (Northwest Alliance for Computational Science and Engineering, 2020). Atmospheric deposition was data was obtained from the National Atmospheric Deposition Program (NADP) for total nitrogen deposition (National Atmospheric Deposition Program, 2020).

3.2.3 Characterizing sources of nutrient pollution

Flow is a significant indicator of nutrient load (Williams et al., 2014) For that reason, the gauged water quality locations were analyzed to find a relationship between watershed characteristics that could be used as a surrogate for flow in the regression models. The median TN and TP concentrations were modeled using MLR models that utilized different combinations of point and non-point anthropogenic predictor variable and watershed characteristics. Data transformations, regression analysis, and regression diagnostic test were performed using MATLAB v9.6 (R2019a) (MathWorks, 2019). Nutrient concentrations were transformed with a box-cox transformation, which determines the best transformation of the response variable (y):

$$y(\lambda) = \begin{cases} \dfrac{c^{\lambda} - 1}{\lambda}; \lambda \neq 0 \\ \log(c); \lambda = 0 \end{cases} \qquad [13]$$

where c represents the observed nutrient concentrations in mg/L and λ is the box-cox transformation constant. An iterative procedure was used to identify the λ value for each nutrient response variable that maximized the goodness of fit of the MLR model.

A complete analysis was performed by pairing watershed, urban, and agricultural predictor variables together to develop MLR models for median nutrient concentrations. The Schwarz' Bayesian criterion (SBC) and Akaike information criterion (AIC) were used to select the best model to describe median TN and TP concentrations (Kutner et al., 2005).

Various diagnostic tests were used to assess the appropriateness of the chosen MLR models. The coefficient of multiple determination (R^2) and adjusted coefficient of multiple determination (Adj R^2) were used to determine how much of the variability the model accounted for. The overall significance of the regression model was determined using the lack of fit F-test. The T-test was used to evaluate the significance of model parameters. The normality of the residuals was analyzed using the Shapiro-Wilk test and the Lillie Test. Multicollinearity in the matrix of predictor variables was assessed using the variance inflation factor (VIF). Multicollinearity was avoided by only using one anthropogenic variable to describe each urban, agricultural, natural, and watershed impact in the model's predictor matrix.

3.3 Results

An MLR analysis was performed to fit models to predict mean TN and TP concentration for locations upstream and downstream of point source inputs for 3 regions in Colorado: the South Platte River Basin (Division 1), the Arkansas River Basin (Division 2), and the Western Slope. Area and Slope were found to be strong indicators of instream flow and were used as surrogates for flow in the models. The selected model for each area and nutrient consists of a combination of urban, agricultural, and natural predictor variables.

3.3.1 Multiple linear regression models for nutrient concentrations

3.3.1.1 Upstream Sampling Locations

Through an initial analysis of gaged sampling locations, flow was found to be a significant indicator of instream nutrient concentrations. Watershed characteristics were then analyzed in a regression analysis to find a surrogate for streamflow. Watershed area and Slope were analyzed, and the following relationship was found:

$$Q = 2.3A^{0.89}S^{1.2} \quad [14]$$

where Q is flow, A is area, and S is slope. Area and Slope were strongly correlated to instream flow and the model developed had as R^2 value of 0.90 (Figure 9).

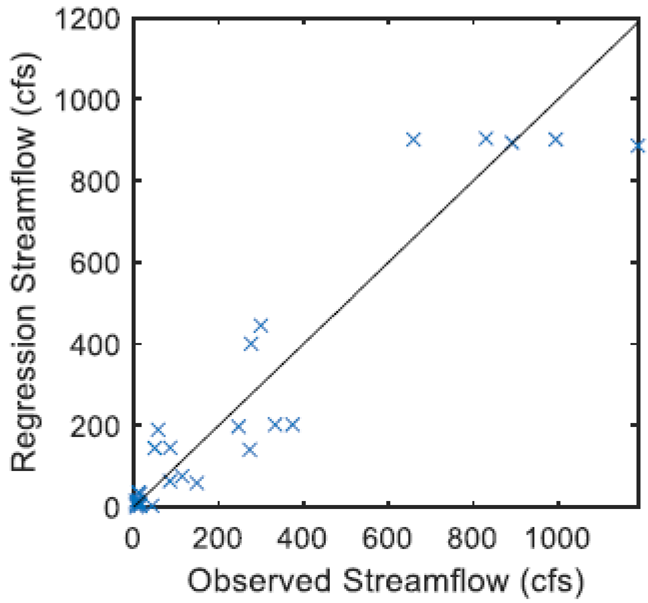


Figure 9 Observed vs. Calculated stream flow for gauged sites in Division 1

WWTP capacity and CAFO capacity increased cumulatively with distance downstream leading to the significant collinearity in the models (represented by VIF) and indicated that using both point sources to describe the response variables is not the best choice for most models. Alternatively, non-point sources can be used to represent urban and agricultural influence. Percent cropland is highly correlated with CAFO capacity and in some cases could be used as a surrogate for the impact of agriculture in areas without CAFO presence. Percent urban land use, percent impervious surface cover, and population were highly correlated with WWTP capacity and could be used to better predict nutrient concentrations.

Strong ($R^2 > 0.7$) correlations were found for all the upstream locations (Table 7). A complete table with all diagnostic tests can be found in Appendix E. Division 1 TN concentrations were strongly correlated to area, slope, population, WWTP capacity, and CAFO capacity. Division 1 TP concentrations were related to area, slope, WWTP capacity, and percent cropland. Increased watershed area leads to a decrease in TP concentrations, this could be associated with dilution that occurs in larger watersheds. Division 2 TN concentration was correlated to area, slope, rangeland, and WWTP capacity. TN concentration increased as drainage area increased, indicating that TN compounded with distance downstream. Median TP concentrations in Division 2 were associated with area, WWTP capacity, and CAFO capacity. TN concentrations for locations on the western slope were correlated to area, slope, and percent cropland. Median TP concentrations on the western slope were correlated with area, average annual precipitation, percent impervious surface cover, percent urban land use, and percent rangeland.

Table 7 Multiple linear regression models for upstream locations. With R^2 and adjusted R^2, P value for the appropriateness of the model, lambda, N, and degrees of freedom (DOF).

	Division	Model	Linear Model	R^2	Adj. R^2	P	λ	N	DoF
Upstream	1	TN	$-0.43 - 3.32 * 10^{-4} Area^{0.89*1.2} - 5.38\, Slope^{1.2*1.2} + 0.06\, Pop + 0.59\, WWTP^{0.25} + 8.67 * 10^{-6}\, CAFO^{0.25}$	0.77	0.74	7E-9	0	36	30
	1	TP	$-3.26 - 3.0 * 10^{-4}(Area)^{0.89} - 4.90(Slope)^{1.2} + 1.40(WWTP)^{0.15} + 0.12(Crop)^{0.6}$	0.71	0.68	3E-8	0	37	32
	2	TN	$5.29 + 3 * 10^{-4} Area^{0.89} - 14.22\, Slope^{1.2} - 0.83 Range^{0.5} + 1.67\, WWTP^{0.1}$	0.87	0.82	2E-4	0	15	10
	2	TP	$-3.07 - 1.4 * 10^{-3}\, Area^{0.89} + 1.9 WWTP^{0.1} + 3.00 * 10^{-5} CAFO$	0.79	0.72	2E-3	0	13	9
	West	TN	$0.97 + 1.2 * 10^{-5} Area^{0.89} - 0.75 Slope^{1.2} + 0.01\, Crop^2$	0.77	0.74	9E-9	1	31	27
	West	TP	$-3.37 + 2.33 * 10^{-9}(Area)^2 - 7.0 * 10^{-4} Precip^2 + 0.27 Urban + 0.07 Range^{0.75} - 0.64 Imperv$	0.71	0.61	1E-3	0	21	15

Negatively correlated area, like that seen in Division 1 TN, Division 1 TP, and Division 2 TN, signify that nutrient concentrations decreased with increased watershed area, it is likely that the instream nutrient concentrations were diluted as it moves downstream. Positive correlations between nutrient concentrations and area indicated that nutrient concentrations increased with increased watershed area. For Division 2 TN, TN and TP on the western slope nutrient concentrations compound due to the cumulative impact of human activity.

Slope was negatively correlated for TN sites in Division 1, Division 2, and on the western slope as well as for TP sites in Division 1. Steeper areas have lower nutrient concentrations, this could be associated with the greater percent of natural land cover on steeper slopes. Precipitation and impervious area were negatively correlated for TP concentrations on the westerns slope. This could signify that runoff is not a significant source of nutrients.

Negatively correlated rangeland, like what is seen for Division 2 TN, indicated that nutrient concentrations did not increase with rangeland. This indicated that natural area did not input a significant amount of nutrients. Conversely, median TP concentrations on the western slope, increased with percent rangeland; rangeland is likely used for grazing and nutrient input is from animal waste.

Population, WWTP capacity, and percent urban were all surrogates for urban activity. Positive correlations between these variables exist in of the regression equations for Division 1, the western slope, and Division 2 TP. This indicates that urban activity was a significant source of nutrient pollution. Different urban sources could potentially be causing the increase in nutrient load, including, but not limiting, WWTP discharge, fertilizer, and domesticated animal waste.

CAFO capacity and percent cropland were used as a proxy for urban activity. Agricultural activity increases instream nutrient concentrations. Nutrient inputs could be from farm animals, manure application, and fertilizer application.

3.3.1.2 Downstream Sampling Locations

Strong models were found a few of the downstream response variables (Appendix E). Strong and significant models were found for Division 2 TN, and TN and TP on the Western slope. Division 2 TN was correlated with area, average annual precipitation, percent impervious surface, and WWTP technology (Table 8). Western slope TN was associated with percent urban and percent cropland. Median TP concentrations for the western slope were correlated to watershed area, percent urban and percent cropland. Models were also developed for the other downstream locations. Division 1 TN was associated with watershed area, average annual precipitation, population, and percent cropland.

Median TP concentration was related to watershed area, population, percent cropland, and rangeland. Division 2 TP concentration was modeled using watershed area, average annual precipitation, percent forest, percent cropland, and percent urban.

Table 8 Multiple linear regression models for downstream locations. With R2 and adjusted R2, P value for the appropriateness of the model, lambda, N, and degrees of freedom (DOF).

	Division	Model	Linear Model	R2	Adj. R^2	P	λ	N	DoF
Downstream	1	TN	$-3.23 - 3.8 * 10^{-4}\ Area^{0.89} - 5.0 * 10^{-3} Precip^{1.5} + 1.46\ Pop^{0.1} + 0.23 Crop^{0.25}$	0.67	0.63	4E-7	**0**	36	31
	1	TP	$1.48 - 3.9 * 10^{-4}\ Area^{0.89} + 1.89 Pop^{0.1} + 0.07 Crop^{0.5} - 6.41 Range^{0.1}$	0.62	0.58	5E-7	0	40	34
	2	TN	$43.97 - 1.9 * 10^{-3}\ Area^{0.89} - 15.26\ Precip^{0.25} + 0.03 Imperv^{2} - 7.33 WWTP\ Tech^{0.1}$	0.86	0.81	3E-4	1	15	10
	2	TP	$9.72 - 1.0 * 10^{-3} Area^{0.89} - 0.43\ Precip + 1.21 Forest^{0.1} - 1.95\ Crop^{0.1} - 3.7 * 10^{-3} Urban^{2}$	0.67	0.50	3E-2	0	16	10
	West	TN	$-033 + 0.32 Urban + 0.20 Crop$	0.79	0.78	1E-8	1	26	22
	West	TP	$1.11 - 0.69 Area^{0.1} + 0.28 Urban + 0.20 Crop$	0.84	0.81	9E-9	1	26	22

Negatively correlated area indicated that as area increased nutrient concentration decreased. This along with the negatively correlate precipitation that is seen in several of the models indicated that runoff tends to dilute nutrient concentrations as water moves downstream.

Rangeland was found to be negatively correlated to nutrient concentrations for Division 1 TP. This indicated natural land use does not input nutrients into the system. Inversely, forest land was positively correlated to nutrient concentration for TP sites in Division 2. In this case nutrients enter the system from natural land cover.

Population, impervious surface cover, and percent urban were used in this study as surrogates for urban activity. Positive correlations between these variables exist in of the regression equations for Division 1, the western slope, and Division 2 TN. This indicated that urban activity is a significant source of nutrient pollution. Different urban sources could potentially be causing the increase in nutrient load, including, but not limited to, WWTP discharge, fertilizer, and domesticated animal waste. Median TP concentration in Division 2 showed a negative trend for urban influence, indicating that urban activity had little influence on nutrient loads.

Cropland was used in these models to represent the influence of agricultural activity. Cropland was positively correlated to nutrient concentrations for Division 1 and the western slope. Agricultural activity increased instream nutrient concentrations. Nutrient inputs could be from farm animals, manure application, and fertilizer application. Division 2 median TP concentrations increase with decreased cropland use. This indicated that agricultural activity had little influence for sites in this location.

WWTP technology was used to predict TN concentrations in Division 2. Improved WWTP technology was associated with improved nutrient concentrations. Indicating that WWTPs were able to successfully remove nitrogen from wastewater.

3.4 Conclusion

The influence of anthropogenic activity on baseline nutrient concentrations was analyzed using a multiple linear regression. The regression analysis was performed for TN and TP concentrations for Division 1, Division 2, and the Western Slope. The regression model was performed using combinations hydrologic variables, non-point source land use variables, and point source anthropogenic variables. Through an initial analysis of gauged sampling locations, flow along with geospatial factors were found to be a significant indicator of instream nutrient concentrations. Watershed characteristics were then analyzed in a regression analysis to find a surrogate for streamflow.

The regression models for upstream and downstream locations were found to be functions of watershed, hydrologic, point, and non-point source predictor variables. The models all showed that agricultural and urban activity significantly impacted instream ambient nutrient concentrations. These models save time and money by assisting the CDPHE Water Quality Control Division in identify stream segments that are already impaired or at risk of being impaired. Rigorous water quality monitoring plans can be developed for these segments to help assist in the development of TMDLs or new regulation.

CHAPTER 4: CONCLUSION

4.1 *E. coli*

An in-depth investigation of *E. coli* concentrations was performed in order to identify the impact of point and non-point sources that could be attributed to increased pollutant concentrations. Probabilistic methods for quantifying *E. coli* vulnerability, mean, and median *E. coli* concentrations were developed in this research. Precipitation analysis was also performed in order to identify the relative importance of point and non-point source pollutant sources. Regression models were adopted to analyze the effect of anthropogenic and natural influences on *E. coli*.

The Cache la Poudre Watershed generally experienced increased *E. coli* concentrations from upstream to downstream due to increased anthropogenic activity. However, inflows from irrigation canals and retention ponds were outside factors that were identified as mechanisms that decrease ambient bacterial concentrations. Vulnerability was developed based on Colorado Regulation 31 *E. coli* water quality limits and observed *E. coli* concentrations. Vulnerability tended to increase from upstream to downstream. Boxelder Creek showed a notable exception. Vulnerability had a value of V=1 upstream and attenuated in the system as it moves downstream. Like concentration, vulnerability decreases due to inflow from irrigation canals and retention.

There was a significant difference in the proportion of samples that exceeded the numeric water quality standards for *E. coli* during rain and non-rain events at CLP11. *E. coli* concentrations at this site during rain events were significantly higher than during sampling trips with no antecedent precipitation. The other sites in the study did not experience a significant difference in water quality exceedance, they did, however, provide insight on trends that occur. In general, locations that are dominated by non-point sources experience increases in *E. coli* concentrations during rain events. Water quality at BC01 was significantly influenced by point sources. Effluent from Boxelder Creek Sanitation District controlled the water quality downstream. During rain events bacterial concentrations decreased as a result of dilution.

The relationship between anthropogenic influence and *E. coli* response in the Cache la Poudre watershed was explored with multiple linear regression models. The MLR approach was used to predict mean, geometric mean and median *E. coli* concentrations, as well as, vulnerability to *E. coli* impairment for sites along the Cache la Poudre River and in the Cache la Poudre watershed. Hydrologic and natural land cover descriptors, and urban and agricultural predictor variables were able to sufficiently describe the *E. coli* response.

When considering the regression models for sites along the Cache la Poudre River, CAFOs were a significant indicator of bacterial water quality for all the response variables. There was a significant amount of collinearity between point source predictor variables, because of this WWTP capacity and CAFO capacity were not be used in the same regression model. Land use was proven to be a good surrogate for point source predictor variables and because of the limited collinearity between the land use variables used, several of these variables could be used together to produce a valid MLR model. Twelve of the 28 sites studied were not impacted by CAFOs or WWTPs, indicating that these predictor variables were not the best indicators for the regression model. When analyzing the Cache la Poudre watershed, average annual precipitation, percent urban land use, percent cropland, percent rangeland, and percent forest strongly correlated to the response variables.

Adopting land use improved the overall performance of the individual regression models because it can more accurately describe changes in the unimpacted watersheds. Models created using the methodology developed in this study, with land use explanatory variables, can help researchers and regulators accurately predict instream E. coli concentrations and identify which streams are at risk of impairment. Regulators will be to save both time and money by focusing on developing monitoring programs for streams that are impaired or at risk of being impaired.

The identification of sources of pollution done in this work could also assist future studies in the development of best management practices for agricultural land use and confined animal feeding operations. Improvements to green infrastructure in urban areas could also be made to increase the retention time before stormwater runoff is discharge back into streams and rivers.

4.2 Nutrients

The impact of anthropogenic influence on baseline nutrient concentrations was analyzed using a multiple linear regression. The regression analysis was performed for TN and TP concentrations for Division 1, Division 2, and the Western Slope. The regression model was performed using combinations hydrologic variables, non-point source land use variables, and point source anthropogenic variables. Through an initial analysis of gaged sampling locations, flow was found to be a significant indicator of instream nutrient concentrations. Area and slope were then analyzed in a regression analysis and were found to be a surrogate for streamflow.

The regression models for upstream and downstream locations were found to be functions of area and slope, hydrologic, point, and non-point source predictor variables. The models all showed that agricultural and urban activity significantly impacted instream baseline nutrient concentrations. These models can help identify waterbodies that are at vulnerable to impairment and will be used to assist regulatory agencies in developing water quality monitoring programs to set TMDLs.

REFERENCES

Anita, R. S., & Maichle. (2015). *ProUCL Version 5.1 User Guide Statistical Software for Environmental Applications for Data Sets with and without Nondetect Observations*.

Baudisova, D. (1997). Evaluation of Escherichia Coli as the main indicator of Faecal Pollution. *Elsevier Science Ltd.*, *35*(11–12), 333–336.

Bouraoui, F., Benabdallah, S., Jrad, A., & Bidoglio, G. (2005). Application of the SWAT model on the Medjerda river basin (Tunisia). *Physics and Chemistry of the Earth*, *30*(8–10), 497–507. https://doi.org/10.1016/j.pce.2005.07.004

Bryan, N. S., & van Grinsven, H. (2013). *The Role of Nitrate in Human Health*. *Advances in Agronomy* (Vol. 119). Elsevier. https://doi.org/10.1016/B978-0-12-407247-3.00003-2

CDPHE. (2013). Regulation # 85: Nutrient Management Control Regulation, (May), 5 CCR 1002-43.

CDPHE. (2015). *Standard Operating Procedures for the Collection of Water Quality and Biological Samples*. *April 2015*. Retrieved from https://www.swfwmd.state.fl.us/files/database/site_file_sets/2684/Chemistry_Laboratory_Manual.pdf

CDPHE. (2018). *Regulation no. 31-The basic standards and methodologies for surface water 5 CCR 1002-31*. Demver.

Chang, H. (2008). Spatial analysis of water quality trends in the Han River basin, South Korea. *Water Research*, *42*(13), 3285–3304. https://doi.org/10.1016/j.watres.2008.04.006

Congress, U. S. Federal water pollution control act, 33 U.S.C. 1251 et seq. § (2002).

Donnison, A., Ross, C., & Thorrold, B. (2004). Impact of land use on the faecal microbial quality of hill-country streams. *New Zealand Journal of Marine and Freshwater Research*, *38*(5), 845–855. https://doi.org/10.1080/00288330.2004.9517284

Edberg, S. C., Rice, E. W., Karlin, R. J., & Allen, M. J. (2000). Escherichia coli: The best biological drinking water indicator for public health protection. *Journal of Applied Microbiology Symposium*

Supplement, *88*(29). https://doi.org/10.1111/j.1365-2672.2000.tb05338.x

Esri. (2019). ArcGIS Pro 2.4.

Feng, P., Weagant, S. D., Grant, M. A., & Burkhardt, W. (2018). Enumeration of Escherichia coli and the Coliform Bacteria. In *Bacteriological Analytical Manual* (pp. 1–14). Retrieved from https://www.fda.gov/Food/FoodScienceResearch/LaboratoryMethods/ucm064948.htm

Francy, D. S., Helsel, D. R., & Nally, R. A. (2000). Occurrence and Distribution of Microbiological Indicators in Groundwater and Stream Water. *Water Environment Research*, *72*(2), 152–161. https://doi.org/10.2175/106143000x137220

Garzio-Hadzick, A., Shelton, D. R., Hill, R. L., Pachepsky, Y. A., Guber, A. K., & 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

Harmel, R. D., Karthikeyan, R., Gentry, T., & Srinivasan, R. (2010). Effects of agricultural management, land use, and watershed scale on E. coli concentrations in runoff and streamflow. *Transactions of the ASABE*, *53*(6), 1833–1841. https://doi.org/10.13031/2013.35809

Heiden, C. (2019). *Submitted by*.

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

Kistemann, T., Claßen, T., Koch, C., Dangendorf, F., Fischeder, R., Gebel, J., et al. (2002). Microbial load of drinking water reservoir tributaries during extreme rainfall and runoff. *Applied and Environmental Microbiology*, *68*(5), 2188–2197. https://doi.org/10.1128/AEM.68.5.2188-2197.2002

Klein, R. D. (1979). Urbanization and Stream Quality Impairment. *Water Resources Bulletin*, *15*(4). https://doi.org/10.1111/j.1752-1688.1979.tb01074.x

Kutner, M. H., Nachtsheim, C., Neter, J., & Li, W. (2005). *Applied linear statistical models*.

Li, S., Liu, W., Gu, S., Cheng, X., Xu, Z., & Zhang, Q. (2009). Spatio-temporal dynamics of nutrients in the upper Han River basin, China. *Journal of Hazardous Materials*, *162*(2–3), 1340–1346. https://doi.org/10.1016/j.jhazmat.2008.06.059

Mallin, M. A., Williams, K. E., Esham, E. C., & Lowe, R. P. (2000). Effect of human development on bacteriological water quality in coastal watersheds. *Ecological Applications*, *10*(4), 1047–1056. https://doi.org/10.1890/1051-0761(2000)010[1047:EOHDOB]2.0.CO;2

Mallin, M. A., Johnson, V. L., & Ensign, S. H. (2009). Comparative impacts of stormwater runoff on water quality of an urban, a suburban, and a rural stream. *Environmental Monitoring and Assessment*, *159*(1–4), 475–491. https://doi.org/10.1007/s10661-008-0644-4

MathWorks. (2019). MATLAB v9.6 (R2019a).

McMahon, G., & Cuffney, T. F. (2000). Quantifying urban intensity in drainage basins for assessing stream ecological conditions. *Journal of the American Water Resources Association*, *36*(6), 1247–1261. https://doi.org/10.1111/j.1752-1688.2000.tb05724.x

Multi-Resoulution Land Characteristic Constortium. (2016). Land Cover Data. Retrieved from https://www.mrlc.gov/data

National Atmospheric Deposition Program. (2020). Deposition Data. Retrieved from http://nadp.slh.wisc.edu/data/

Northwest Alliance for Computational Science and Engineering. (2020). PRISM Data. Retrieved from https://prism.oregonstate.edu/

Odonkor, S. T., & Ampofo, J. K. (2013). Escherichia coli as an indicator of bacteriological quality of water: an overview. *Microbiology Research*, *4*(1), 2. https://doi.org/10.4081/mr.2013.e2

One Water Solutions Institute. (2020). Environmental Resource Assessment and Management System (eRAMS).

Pandey, P. K., Soupir, M. L., Haddad, M., & Rothwell, J. J. (2012). Assessing the impacts of watershed indexes and precipitation on spatial in-stream E. coli concentrations. *Ecological Indicators*, *23*, 641–652. https://doi.org/10.1016/j.ecolind.2012.05.023

Paule-Mercado, M. A., Ventura, J. S., Memon, S. A., Jahng, D., Kang, J. H., & Lee, C. H. (2016). Monitoring and predicting the fecal indicator bacteria concentrations from agricultural, mixed land use and urban stormwater runoff. *Science of the Total Environment*, *550*, 1171–1181. https://doi.org/10.1016/j.scitotenv.2016.01.026

Rossi, A., Wolde, B. T., Lee, L. H., & Wu, M. (2020). Prediction of recreational water safety using Escherichia coli as an indicator: case study of the Passaic and Pompton rivers, New Jersey. *Science of the Total Environment*, *714*. https://doi.org/10.1016/j.scitotenv.2020.136814

Selvakumar, A., & Borst, M. (2006). Variation of microorganism concentration in urban stormwater runoff with land use and seasons. *Journal of Water and Health*, *4*(1), 109–124. https://doi.org/10.2166/wh.2005.063

Signor, R. S., Roser, D. J., Ashbolt, N. J., & Ball, J. E. (2005). Quantifying the impact of runoff events on microbiological contaminant concentrations entering surface drinking source waters. *Journal of Water and Health*, *3*(4), 453–468. https://doi.org/10.2166/wh.2005.052

Standridge, J. (2008). E. coli as a public health indicator of drinking water quality. *Journal / American Water Works Association*, *100*(2). https://doi.org/10.1002/j.1551-8833.2008.tb08143.x

Sylvestre, É., Burnet, J. B., Smeets, P., Medema, G., Prévost, M., & Dorner, S. (2020). Can routine monitoring of E. coli fully account for peak event concentrations at drinking water intakes in agricultural and urban rivers? *Water Research*, *170*. https://doi.org/10.1016/j.watres.2019.115369

Tasdighi, A., Arabi, M., & Osmond, D. L. (2017). The Relationship between Land Use and Vulnerability to Nitrogen and Phosphorus Pollution in an Urban Watershed. *Journal of Environmental Quality*, *46*(1), 113–122. https://doi.org/10.2134/jeq2016.06.0239

U.S. Census Bureau. (2019). Tract Level Populaiton Data. Retrieved from https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html

United States Environmental Protection Agency, (Usepa). (2001). Ambient water quality criteria recommendations lakes and reservoirs in nutrient ecoregion XIV EPA 822-B-01-011, (December).

USEPA. (2000). Nutrient criteria technical guidance manual: Lakes and reservoirs, EPA-822-B00-001, (July), 1–231.

USEPA. (2002). *Method 1103.1: Escherichia coli (E. coli) in Water by Membrane Filtration Using membrane-Thermotolerant Escherichia coli Agar (mTEC).*

Walpole, R. E., Myers, R. H., & Myers, S. L. (1998). *Probability and Statistics for Engineers and Scientists* (6th ed.). Prentice Hall.

Water Quality Monitoring Council. (2020). Water Quality Data. Retrieved from https://www.waterqualitydata.us/portal/#organization=21COL001_WQX&project=Reg_85_Monitoring&mimeType=csv

Wickham, J. D., Riitters, K. H., O'Neill, R. V., Reckhow, K. H., Wade, T. G., & Jones, K. B. (2000). Land cover as a framework for assessing risk of water pollution. *Journal of the American Water Resources Association*, *36*(6), 1417–1422. https://doi.org/10.1111/j.1752-1688.2000.tb05736.x

Williams, R. E., Arabi, M., Loftis, J., & Elmund, G. K. (2014). Monitoring Design for Assessing Compliance with Numeric Nutrient Standards for Rivers and Streams Using Geospatial Variables. *Journal of Environmental Quality*, *43*(5), 1713–1724. https://doi.org/10.2134/jeq2013.12.0528

Young, K. D., & Thackston, E. L. (1999). Housing Density and Bacterial Loading in Urban Streams. *Journal of Environmental Engineering*, *600*(December), 1–5.

Appendix A

Cache la Poudre

Eleven sampling locations were selected on the CLP River (Figure 10). TMDL-CLP-01, TMDL-CLP-02, TMDL-CLP-03 are in industrial or commercial areas in Greeley, CO. TMDL-CLP-04, TMDL-CLP-05, TMDL-CLP-06, TMDL-CLP-07, TMDL-CLP-08, and TMDL-CLP-09 are located along the Poudre Trail and near various natural areas in Greeley, CO. TMDL-CLP-10 is located in a park and near a neighborhood in Windsor, CO and TMDL-CLP-11 is located near several natural areas in Fort Collins, CO. Over the course of the study, 308 samples were collected and tested for *E. coli*, complete water quality parameters were measured on 298 visits, and flow was measured 190 times (Table 9). 17 of the analyzed *E. coli* samples resulted in non-detect readings. According to Regulation 31, a geometric mean of five samples take at least 7 days apart during a two-month period cannot exceed 126 CF/100 mL.

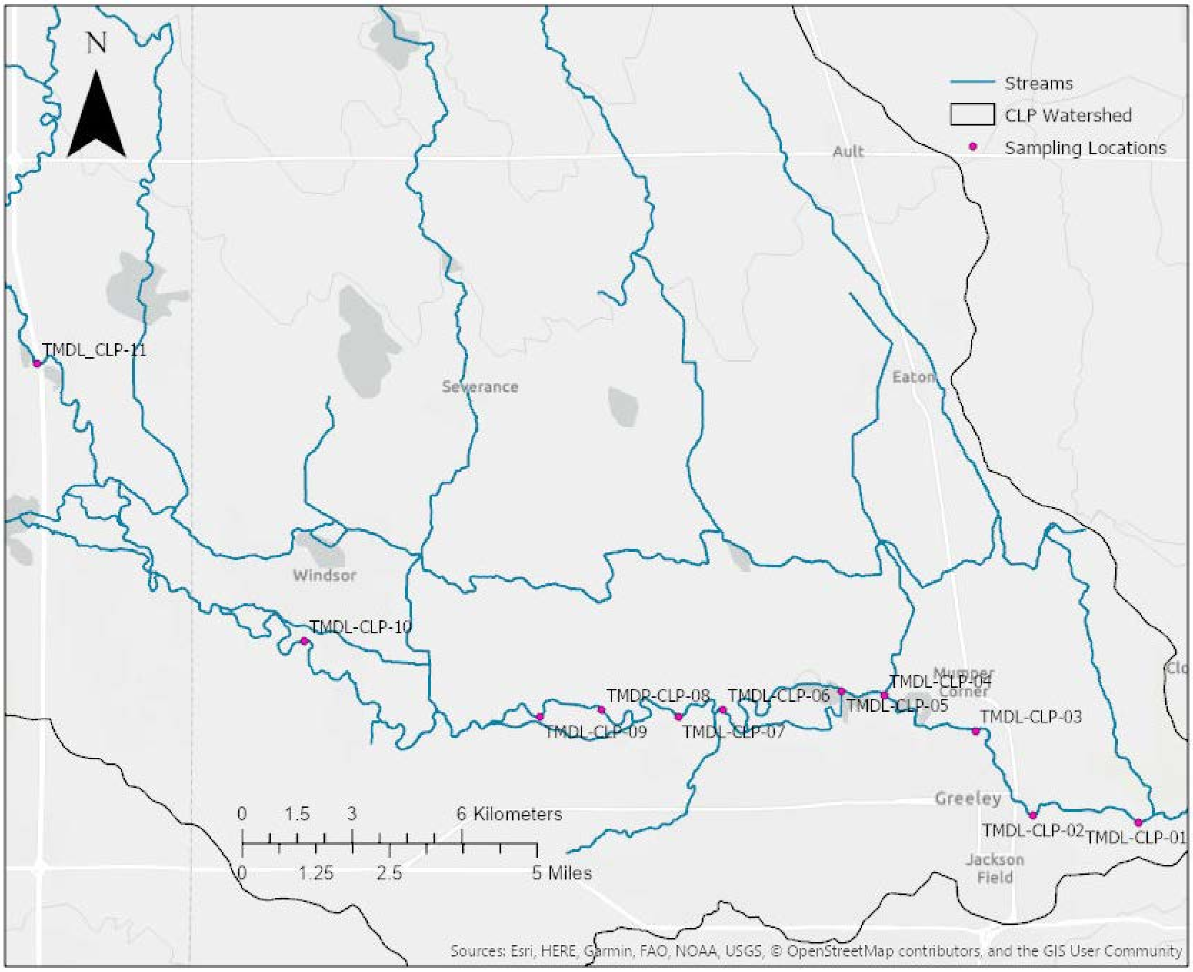


Figure 10 Cache la Poudre E. coli sampling locations

Table 9 Cache la Poudre (CLP) sampling site description and summary of collected data during the 2018 and 2019 sampling periods. Watershed area is the area that drains to the sampling site, River Mile is the distance from the site to the CLP Rivers confluence with the South Platte River. Number of E. coli samples and flow are the number of readings collected at each site. Water Quality is the number of complete water quality profiles collected at each location (pH, specific conductivity, DO concentration, and temperature). Non-detect readings are the number of samples that had E. coli concentrations below the detection limits (20 CFU/100mL).

Site ID	Watershed Area (mi^2)	River Mile	Number of *E. coli* samples	Flow	Water Quality	Non-Detect
CLP01	1837.60	2.796	30	19	29	1
CLP02	1831.99	4.918	30	13	29	1
CLP03	1824.86	7.007	30	17	29	1
CLP04	1743.36	9.076	30	18	29	1
CLP05	1742.83	9.902	30	18	29	2
CLP06	1732.08	13.559	30	19	29	4
CLP07	1714.04	14.786	30	20	29	1
CLP08	1705.46	17.167	30	19	29	2
CLP09	1702.95	18.512	30	20	29	0
CLP10	1584.83	25.494	30	20	29	4
CLP11	1237.18	46.567	8	7	8	0

TMDL-CLP-01

TMDL-CLP-01 is located on Fern Avenue near Weld County Airport, 2.8 miles from the CLP River's confluence with the South Platte River. 1837.60 mi^2 drain to this location. The Ogilvy Ditch takes water out for agricultural purposes upstream of this site. Greeley Number 3 Ditch, upstream of TMDL-CLP-01, introduces water that is being return to the system upstream. 30 *E. coli* samples were taken at this sampling location over the course of the study period, along with 19 flow and 29 water quality measurements. One *E. coli* sample collected resulted in a non-detect reading. The mean *E. coli* concentration at this location was found to be 372.2 CFU/mL, the median concentration was 161.5 CFU/100mL. The minimum was 8.932 CFU/100mL and the maximum was 2980 CFU/100mL.

TMDL-CLP-02

TMDL-CLP-02 is located on Ash Ave in Greeley, CO 4.918 river miles from the confluence with the South Platte River and drains 1831.99 mi^2 of land. This sampling location is downstream of Kaylor of Colorado Production, Leprino Foods' Manufacturing Facility and the JBS meat processing plant. The Mead lateral returns water to the CLP river upstream of the sampling location. 30 *E. coli* samples from this site were analyzed for *E. coli* concentration. One of these samples had concentrations too low to be detected. 13 flow measurements were taken from this location and 29 complete water quality profiles were measured. The average *E. coli* concentration was 274.8 CFU/100mL and the median concentration of the samples was 163.5 CFU/100mL. The minimum concentration was 27.2 CFU/100mL and the maximum concentration was 1240 CFU/100mL.

TMDL-CLP-03

TMDL-CLP-03 is located on 8th Avenue, upstream of the JBS meat processing plant. The Graham Seep and two small tributaries enter the CLP River upstream of this site. This site is located 7.01 miles from the confluence with the South Platte River and Drains 1824.9 mi^2. 30 *E. coli* samples were collected from this site, flow was measured at 17 of the visits, and complete water quality measurements were recorded at 29 of the visits. One of the samples has *E. coli* concentrations below the detection limits. The mean *E. coli* concentration was 260.2 CFU/100mL, the median concentration was 220 CFU/100mL. The minimum concentration was 16. 3 CFU/100mL and the maximum concentration was 1120 CFU/100mL.

TMDL-CLP-04

Located 9.08 miles upstream of the South Platte River, TMDL-CLP-04 drains approximately 1743.4 mi^2 of land. This site is located off 25th Avenue. near the Poudre Ponds. 30 samples were collected throughout the study period, one sample was analyzed and found to have levels below the detection limit. A complete water quality profiles were collected during 29 of the visits and on 18 visits conditions allowed for flow measurements to be collected. The average *E. coli* concentration of the samples was 218.9 CFU/100mL, while the median concentration was 120 CFU/100mL. The minimum *E. coli* concentration was 9 CFU/100mL and the maximum was 2520 CFU/100mL.

TMDL-CLP-05

Located along the Poudre Trail east of North 35th Avenue, TMDL-CLP-05 is 9.9 miles upstream of the CLP Rivers confluence with the South Platte River and drains approximately 1742.8 mi^2. There is one diversion upstream. 30 *E. coli* samples were collected and analyzed for bacterial concentration, two of the collected samples had non-detect readings. The mean concentration was 183 CFU/100mL and the median concentration was 80 CFU/100mL. The minimum concentration was 5.7 CFU/100mL and the maximum was 2320 CFU/100mL. Water quality parameters were measured during 29 visits and flow was measured during 18 visits.

TMDL-CLP-06

TMDL-CLP-06 is located off 59th Avenue between Greeley's Cottonwood and Sheep Draw Natural Area. Sheep Draw discharges into the CLP river upstream of the sampling location. This site is located 13.6 miles from the CLP River's confluence with the South Platte River and drains 1732 mi^2. 30 samples were collected at this site, four of the samples were analyzed and found to have *E. coli* concentrations below the test's detection limits. Flow was measured during 19 of the visits and complete water quality profiles were collected during 29 visits. The average *E. coli* concentration was 243.5 CFU/100mL and the median concentration was 110 CFU/100mL. The minimum concentration was 9.2 CFU/100mL and the maximum concentration was 3380 CFU/100mL.

TMDL-CLP-07

Downstream of the Signature Bluffs Natural Area, TMDL-CLP-07 is 14.8 mile from the CLP River's confluence with the South Platte River and drains 1714 mi^2. This site is near a livestock fence that extends into the water and there are regularly cows in or near the river. 30 samples were collected and one of the samples had a non-detect reading. Flow was measured during 20 of the visits. 29 of the visits had complete water quality readings. The mean *E. coli* concentration was 243 CFU/100mL and the median concentration was 80 CFU/100mL. The minimum concentration was 4.7 CFU/100mL and the maximum was 3220 CFU/100mL.

TMDL-CLP-08

TMDL-CLP-08 is located on 83rd Avenue near the Poudre Learning Center. A small tributary enters the stream upstream of this location. This site is located 17.2 miles from the CLP River's confluence with the South Platte River and drains 1705.5 mi^2 of land. 30 samples were collected at this site and 2 of these samples had non-detect readings. Flow was measured during 19 trips and water quality readings were measured at 29 of the 30 visits. The mean *E. coli* concentration was 255.7 CFU/100mL and the median concentration was 100 CFU/100mL. The minimum concentration was 6.8 CFU/100mL and the maximum was 3900 CFU/100mL.

TMDL-CLP-09

Located off 95th Avenue, TMDL-CLP-09 is 18.5 miles upstream of the CLP River's confluence with the South Platte River. 1702 mi^2 drain to this location. The Jones Ditch takes water out of the river upstream of the sampling location. 30 *E. coli* samples were collected over the course of the study period, along with 20 flow measurements. Complete water quality profiles were measured during 29 of the visits. The samples had a mean *E. coli* concentration of 177 CFU/100mL and a median concentration of 120 CFU/100mL. The samples had a minimum concentration of 20 CFU/100mL and a maximum concentration of 1080 CFU/100mL.

TMDL-CLP-10

TMDL-CLP-10 is located near in Eastman Park in Windsor, CO near the community garden. There are three irrigation ditches that are taking water out of the CLP River directly upstream of the sampling site. These include Greeley Number 2 Canal, Eaton Ditch, and Whitney Ditch. Fossil Creek Reservoir also discharges into the CLP River upstream of this site. TMDL-CLP-10 is located 25.5 miles from the confluence with the South Platte River and it drains 1584.8 mi^2 of land. 30 samples were collected at this site, including four with non-detect readings. Flow measurements were taken on 20 of the visits and complete water quality panels were found at 28 of the sites. The mean *E. coli* concentration was 143.4 CFU/100mL and the median was 110 CFU/100mL. The minimum concentration was 12 CFU/100mL and the maximum was 700 CFU/100mL.

TMDL-CLP-11

46.6 miles from the confluence with the South Platte River, TMDL-CLP-11 is located off Timberline in Fort Collins. This site is near the Kind Fisher Point Natural Area to the south and a commercial area to the north. This site drains 1237.18 mi^2 of land. Sampling began at this site July 31, 2019, eight samples were collected during this time. At each of the visits water quality parameters were collected. Flow measurements were taken during 7 of the visits. The mean *E. coli* concentration was 177 CFU/100mL and the median concentration was 80 CFU/100mL. The minimum concentration was 20 CFU/100mL and the maximum was 914 CFU/100mL.

Boxelder Creek

Four sampling locations were selected along Boxelder Creek in Fort Collins, CO (Figure 11). TMDL-BC-01 and TMDL-BC-02 are located near the Boxelder Sanitation District. TMDL-BC-03 and TMDL-BC-04 are located in areas highly influenced by agriculture. 119 samples were collect at the Boxelder Creek locations, 5 of the analyzed samples had non-detect readings (Table 10). 120 water quality measurements and 113 flow measurements were also collected. Colorado Regulation 31 requires a geometric mean of five *E. coli* concentrations taken at least 7 days apart over 61 days have a concentration of 205 CFU/100mL May 15th through September 15th and a concentration of 630 September 16th through May 14th.

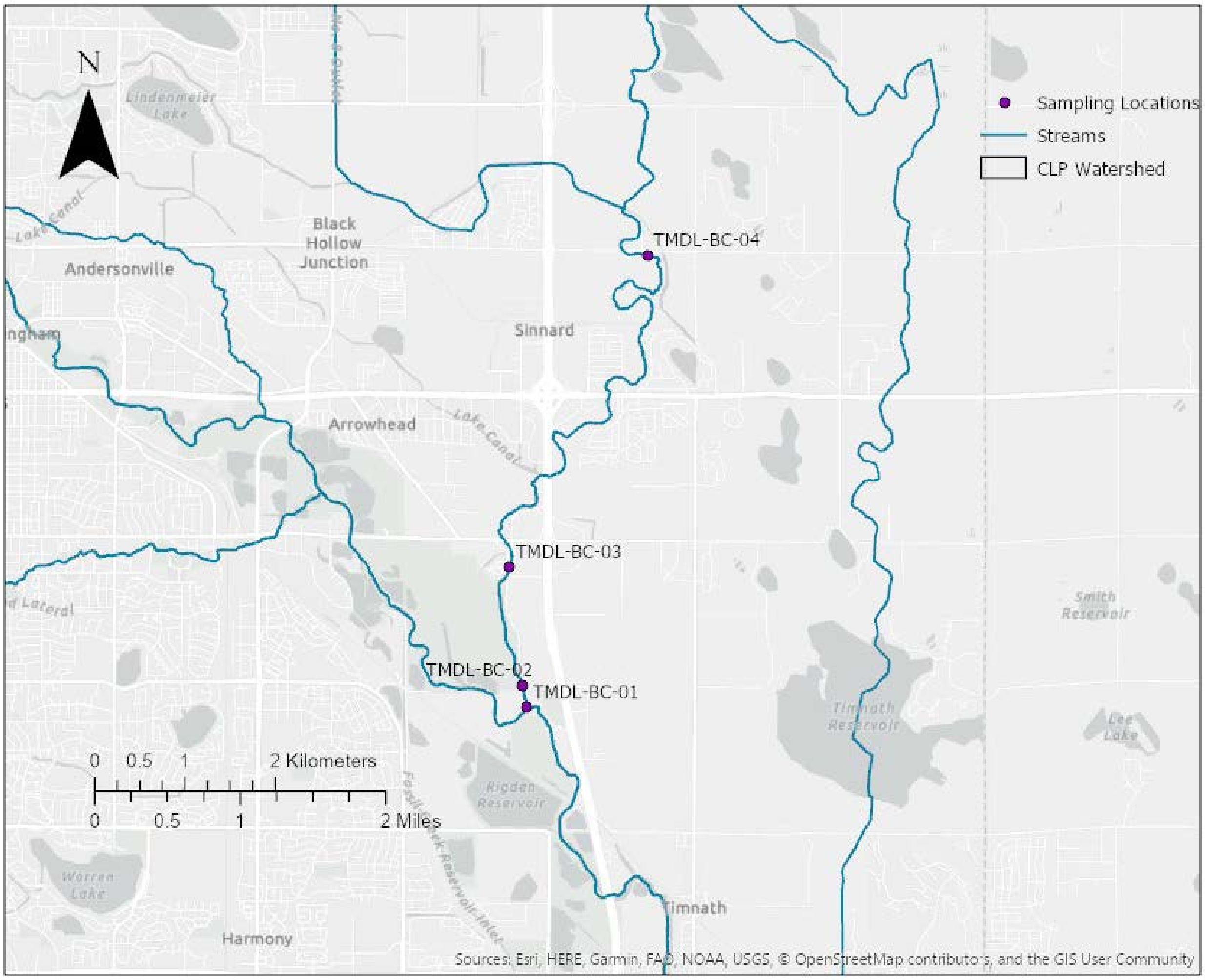


Figure 11 Boxelder Creek E. coli sampling locations

Table 10 Boxelder Creek (BC) sampling site description and summary of collected data during the 2018 and 2019 sampling periods. Watershed area is the area that drains to the sampling site, river mile is the distance from the site to the CLP Rivers confluence with the South Platte River. Number of E. coli samples and flow are the number of readings collected at each site. Water Quality is the number of complete water quality profiles collected at each location (pH, specific conductivity, DO concentration, and temperature). Non-detect readings are the number of samples that had E. coli concentrations below the detection limits (20 CFU/100mL).

Site ID	Watershed Area (mi^2)	River Mile	Number of *E. coli* samples	Flow	Water Quality	Non-Detect
BC01	318.40	38.304	29	26	30	3
BC02	318.37	38.475	30	30	30	0
BC03	318.10	39.371	30	28	30	2
BC04	313.97	43.337	30	29	30	0

TMDL-BC-01

Located 38.3 miles from the CLP River's confluence with the South Platte River, TMDL-BC-01 drains 318.4 mi^2. This site is directly downstream of the Boxelder Sanitation District effluent and upstream of Boxelder Creek's confluence with the CLP River. Because of the proximity to the CLP River, this site is regularly inundated due to high flows in the CLP River. 30 samples were collected over the study period; however, one sample results are missing due to a laboratory error. Three of the samples were found to have non-detect levels of *E. coli*. The mean *E. coli* concentration was 728.5 CFU/100mL and the median was 115 CFU/100mL. The minimum *E. coli* concentration was 4.5 CFU/100mL and the maximum concentration was 8800 CFU/100mL. Complete water quality profiles were taken at every visit along with 26 flow readings.

TMDL-BC-02

TMDL-BC-02 is located just upstream of TMDL-BC-01 at river mile 38.48 and it drains 318.37 mi^2. This site is in the Running Deer Natural Area near a stream gauge just upstream of the Boxelder Sanitation District. 30 *E. coli* samples, flow measurements, and water quality profiles were gathered at this site. The samples collected had an average *E. coli* concentration of 177.9 CFU/100mL and a median concentration of 120 CFU/100mL. The site had a minimum concentration of 20 CFU/200mL and a maximum of 697 CFU/100mL.

TMDL-BC-03

TMDL-BC-03 is located off Prospect Road near the Colorado Department of Transportation Poudre Rest Area. This site is 39.4 miles from the CLP River's confluence with the South Platte River and drains 318.1 mi^2. 30 *E. coli* samples were collected from this site, with two having non-detect readings. Water quality data was collected during every visit and flow measurements taking place during 28 visits. The mean *E. coli* concentration was 176 CFU/100mL and a median concentration of 120 CFU/100mL. The minimum concentration was 12.9 CFU/100mL and the maximum concentration was 680 CFU/100mL

TMDL-BC-04

43.3 miles upstream of the CLP River's confluence with the South Platte River, TMDL-BC-04 is located off East Vine Drive in a heavily farmed area. This site drains 314 mi^2 and has several diversions upstream into small reservoirs. 30 *E. coli* samples were collected from this location over the course of the study, along with 30 water quality readings and 28 flow readings. The average *E. coli* concentration at this site was 6044 CFU/100mL and the median concentration was 3750 CFU/100mL. The minimum level was 320 CFU/100mL and the maximum was 20,000 CFU/100mL.

Fossil Creek

Five sampling locations were selected on Fossil Creek, along with one location on the Fossil Creek Reservoir Outlet (Figure 12). These location are primarily located in residential areas and near developed open spaces. Throughout the study 180 samples were collected from the Fossil Creek sites and analyzed for *E. coli*, 16 of these samples had non-detect readings (Table 11). Water quality parameters were measured during 177 of the visits and flow was collected during 141 visits. Colorado Regulation 31 states that location on this tributary have a geometric mean *E. coli* concentration of 126 CFU/100mL. The geometric mean should consider at least five measurements taken 7 or more days apart over a two month period.

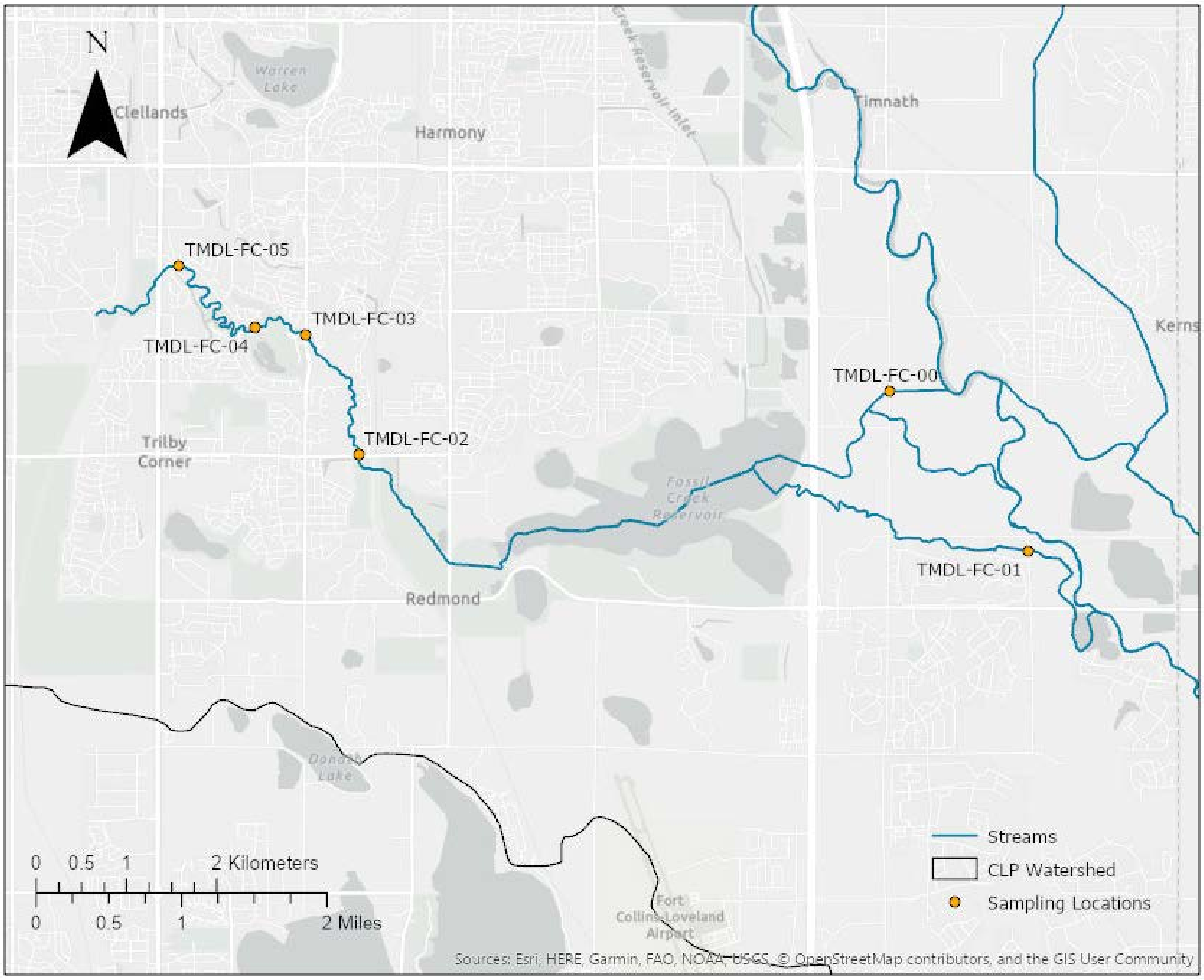


Figure 12 Fossil Creek E. coli sampling locations

Table 11 Fossil Creek (FC) sampling site description and summary of collected data during the 2018 and 2019 sampling periods. Watershed area is the area that drains to the sampling site, river mile is the distance from the site to the CLP Rivers confluence with the South Platte River. Number of E. coli samples and flow are the number of readings collected at each site. Water Quality is the number of complete water quality profiles collected at each location (pH, specific conductivity, DO concentration, and temperature). Non-detect readings are the number of samples that had E. coli concentrations below the detection limits (20 CFU/100mL).

Site ID	Watershed Area (mi^2)	River Mile	Number of *E. coli* samples	Flow	Water Quality	Non-Detect
FC00	3.69	33.476	30	0	29	10
FC01	35.40	31.61	30	29	29	3
FC02	14.50	37.102	30	27	29	1
FC03	12.71	38.342	30	29	30	1
FC04	12.14	38.86	30	28	30	1
FC05	11.67	40.475	30	28	30	0

TMDL-FC-00

TMDL-FC-00 is located 33.5 miles upstream of the CLP River's confluence with the South Platte River and drains 3.69 mi^2. This site is located on the Fossil Creek Reservoir Outlet. This outlet is used as an irrigation ditch and because of this no flow measurements were collected at this site to avoid potential contamination. 30 samples were collected and tested for *E. coli*, 10 of these samples came back as non-detect. The mean concentration was 33.5 CFU/100mL and the median concentration was 40 CFU/100mL. The minimum *E. coli* concentration of the samples was 4.4 CFU/100mL and the maximum concentration was 220 CFU/100mL. Complete water quality profiles were collected during 29 of the visits.

TMDL-FC-01

Downstream of the Ptarmigan Country Club and Golf Course, TMDL-FC-01 is 31.6 miles from the CLP River's confluence with the South Platte River. This site drains 35.4 mi^2 and is feed by the Fossil Creek Reservoir. 30 *E. coli* samples were collected at this site, including three with a non-detect readings. Flow was measured and water quality readings were taken during 29 of the visits. The mean *E. coli* level was 79.3 CFU/100mL and the median concentration was 80 CFU/100mL. The maximum concentration was 280 CFU/100mL and the minimum was 9.4 CFU/100mL.

TMDL-FC-02

TMDL-FC-02 is located off East Trilby in a residential area near the Power Trail and the Fossil Creek Trail 37.1 miles upstream of the CLP River's confluence. This site has several small tributaries that feed into Fossil Creek upstream. 14.5mi^2 drain into TMDL-FC-02. Over the course of the study 30 samples were collected and tested for *E. coli*, one was found non-detect. Flow was measured 27 times and water quality characteristics were analyzed during 29 visits. The mean *E. coli* concentration at this site was 707.7 CFU/100mL and the median concentration was 335 CFU/100mL. The minimum concentration was 30.9 CFU/100mL and the maximum concentration was 11,100 CFU/100mL.

TMDL-FC-03

Located off South Lemay Avenue, TMDL-FC-03 is 38.3 miles upstream of the CLP River's confluence with the South Platte River and drains 12.71 mi^2. This site is nestled between the Southridge Golf Course, Fossil Creek Dog Park, and Fossil Creek Dog Park. Mail Creek, a tributary to Fossil Creek enters upstream of the sampling location. 30 *E. coli* samples were collected from this site, one of which had *E. coli* concentrations below the detection limit. Water quality parameters were analyzed during every visit and flow was measured during 29 of the visits. The average *E. coli* concentration was 327.7 CFU/100mL and the median concentration was 235 CFU/100mL. The minimum concentration was 19.3 CFU/100mL and the maximum concentration was 3460 CFU/100mL.

TMDL-FC-04

TMDL-FC-04 is in Fossil Creek Park, 38.9 miles upstream of the CLP River's confluence with the South Platte River and drains 12.1 mi^2. This site is adjacent to several neighborhoods and the Fossil Creek Trail. 30 *E. coli* samples were collected over the course of the study, including one non-detect. The mean concentration of the samples was 353.3 CFU/100mL and the median concentration was 240 CFU/100mL. The minimum concentration was 11 CFU/100mL and the maximum was 3460 CFU/100mL. Flow was measured during 28 of the visits and water quality parameters were measured at every visit.

TMDL-FC-05

TMDL-FC-05 is located downstream of the Redtail Grove Natural Area, Commercial businesses, and several residential neighborhoods. This site is 40.5 miles upstream of the CLP River's confluence with the South Platter River and it drains 11.7 mi^2. 30 samples were collected and measured for *E. coli*. Resulting in an average concentration of 295.5 CFU/100mL, a median concentration of 180 CFU/100mL, a minimum concentration of 20 CFU/100mL and a maximum concentration of 1520 CFU/100mL. Flow was measured at this site during 28 visits and water quality parameters were analyzed during all the visits.

Spring Creek

Seven sites were chosen on Spring Creek in Fort Collins, CO (Figure 13). These sites are in areas that are predominantly urban or developed open spaces. 210 samples were collected and analyzed for *E. coli* concentrations on this reach, 4 of the collected samples had non-detect readings (Table12). Water quality parameters were measured during every visit and flow was measured during 199 visits. According to Regulation 31, a geometric mean of five samples take at least 7 days apart during a two-month period cannot exceed 126 CF/100 mL for sites on this reach.

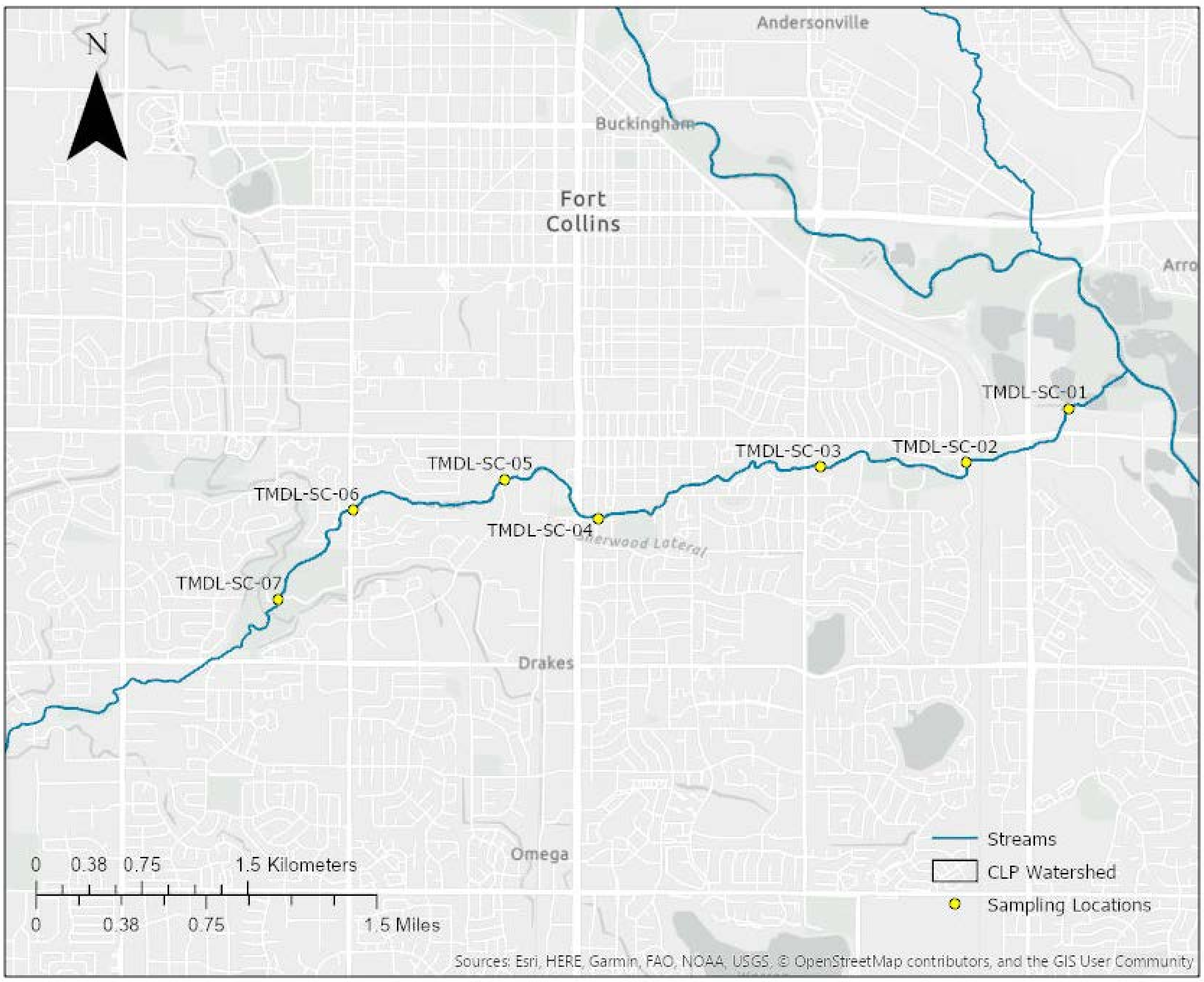


Figure 13 Spring Creek E. coli sampling site locations

Table 12 Spring Creek (SC) sampling site description and summary of collected data during the 2018 and 2019 sampling periods. Watershed area is the area that drains to the sampling site, river mile is the distance from the site to the CLP Rivers confluence with the South Platte River. Number of E. coli samples and flow are the number of *readings collected at each site. Water Quality is the number of complete water quality profiles collected at each location (pH, specific conductivity, DO concentration, and temperature). Non-detect readings are the number of samples that had E. coli concentrations below the detection limits (20 CFU/100mL).*

Site ID	Area (mi^2)	River Mile	Number of *E. coli* samples	Flow	Water Quality	Non-Detect
SC01	9.83	41.267	30	29	30	2
SC02	9.64	41.844	30	29	30	1
SC03	8.94	42.612	30	29	30	1
SC04	7.80	43.805	30	29	30	0
SC05	6.68	44.398	30	30	30	0
SC06	6.10	45.171	30	28	30	0
SC07	3.84	45.773	30	25	30	0

TMDL-SC-01

Located 41.3 miles from the CLP River's confluence with the South Platte River, TMDL-SC-01 drains 9.83 mi^2. This site is located off East Prospect Road near a school and a commercial area. 30 samples were collected from this site and tested for *E. coli*. In two of these samples *E. coli* levels were found to be below the detection limit. The mean *E. coli* level was found to be 225.7 CFU/100mL and the median concentration was 110 CFU/100mL. The minimum concentration was 9 CFU/100mL and the maximum concentration was 820 CFU/100mL. Water quality measurements were recorded during all visits and flow was measured during 29 visits.

TMDL-SC-02

TMDL-SC-02 is located 41.8 miles upstream the CLP River's confluence with the South Platte River and drains 9.64 mi^2. Located in Edora Park this site underwent construction during sampling in 2019. 30 samples were taken from this site, along with 30 water quality measurements, and 29 flow measurements. One *E. coli* sample was reported as non-detect. The mean concentration was 258 CFU/100mL and the median concentration was 150 CFU/100mL. The minimum concentration was 11.5 CFU/100mL and the maximum was 1060 CFU/100mL.

TMDL-SC-03

Off South Lemay Avenue, TMDL-SC-03 is located near an apartment building and commercial areas. This site is 42.6 miles upstream of the CLP River's Confluence with the South Platte River and drains 8.9 mi^2. 30 samples were collected and analyzed for *E. coli* concentrations, one of the samples had levels below the detection limit. The mean concentration was 298 CFU/100mL and the median concentration was 220 CFU/100mL. The minimum concentration was 20 CFU/100mL and the maximum concentration was 1140 CFU/100mL. Water quality parameters were recorded at every visit and flow was measured during 29 visits.

TMDL-SC-04

Located in a residential area off Remington Street, TMDL-SC-04 is 43.8 miles upstream of the CLP River's confluence with the South Platte River. This site drains 7.8 mi^2 of Fort Collins, CO. 30 samples were collected and analyzed for *E. coli* concentrations, along with 30 water quality measurements and 29 flow measurements. The mean *E. coli* concentration was 224.3 CFU/100mL and the median concentration was 160 CFU/100mL. The minimum concentration was 40 CFU/100mL and the maximum concentration was 780 CFU/100mL.

TMDL-SC-05

TMDL-SC-05 is located off Centre Avenue near the Hilton Fort Collins. This site is 44.4 miles upstream from the CLP River's confluence with the South Platte River and drains 6.68 mi^2. One small tributary enters spring creek upstream of this site before the water is stored in Privy Pond and discharged back into Spring Creek. 30 *E. coli* samples, flow measurements, and water quality measurements were collected from this location. The mean *E. coli* concentration was 155 CFU/100mL and the median concentration was 110 CFU/100mL. The minimum value was 20 CFU/100mL and the maximum value was 800 CFU/100mL.

TMDL-SC-06

Located 45.2 miles from the CLP River's confluence with the South Platte River. This site is located near some commercial and residential areas and drains 6.1 mi^2 of Fort Collins. 30 samples were taken and analyzed for *E. coli*. The mean concentration was 513.4 CFU/100mL and a median concentration of 280 CFU/100mL. The minimum concentration was 40 CFU/100mL and the maximum concentration was 2500 CFU/100mL. Water quality was recorded during all the visits and flow was recorded during 28 of the visits.

TMDL-SC-07

Located in Rolland Moore Park downstream of the Ross Natural Area, TMDL-SC-07 drains 3.84 mi^2 and is located 45.8 miles from the CLP River's confluence with the South Platte River. 30 samples were collected at this site along with 30 water quality readings. Flow was taken during 25 visits, because on several visits the flow was too low to measure or there was equipment malfunction. The mean *E. coli* level for this site was 513.4 CFU/100mL and the median was 280 CFU/100mL. The minimum concentration 20 CFU/100mL and the maximum concentration was 2580 CFU/100mL.

Appendix B

Variable	Data Sources
WWTP Capacity	Colorado Department of Public Health and Environment (CDPHE) Regulation 85
WWTP Technology Type	Colorado Department of Public Health and Environment (CDPHE) Regulation 85
CAFO Capacity	A summary of the total number of animal units, by type of animal, is included for watersheds larger than 20 square miles.
Population Density	U.S. 2010 Census
% Impervious Surface Cover	The National Land Cover Database (NLCD) of Percent Developed Imperviousness
% Urban land Use	The National Land Cover Database (NLCD)
% Crop Land Use	The National Land Cover Database (NLCD)
% Forest Land Use	The National Land Cover Database (NLCD)
% Range Land Use	The National Land Cover Database (NLCD)
Average Annual Deposition	National Atmospheric Deposition Program (NADP
Average Annual Precipitation	Parameter-elevation Regressions on Independent Slopes Model (PRISM)

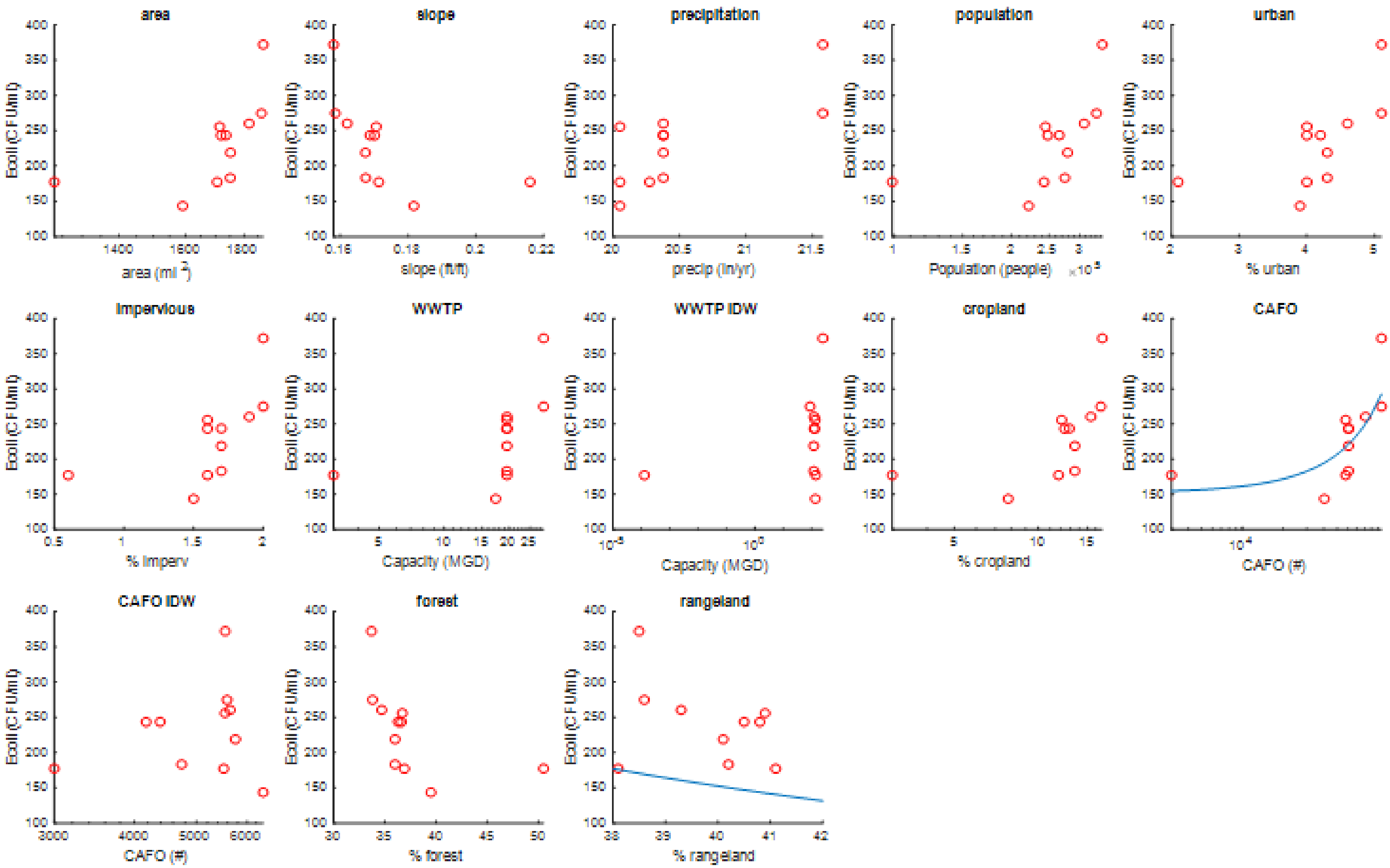
Appendix C

Cache la Poudre River

Mean

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Area	$4.48 + 3.2 * 10^{-7}(Area)^2$	0.43	0.37	3E-2	-	0.2918	0.0140	-32.46	-31.66	0	11	9
Precipitation	$-66.80 + 53.38(Precip)^{0.1}$	0.52	0.47	1E-2	-	0.2524	0.9144	-34.31	-33.52	0	11	9
Slope	$24.50 - 22.77(Slope)^{0.1}$	0.39	0.32	4E-2	-	0.1722	0.7561	-31.66	-30.87	0	11	9
Urban	$4.87 + 3.1 * 10^{-2}(Urban)^2$	0.46	0.40	2E-2	-	0.5000	0.7835	-32.98	-32.19	0	11	9
Cropland	$4.99 + 2.5 * 10^{-3}(Crop)^2$	0.64	0.60	3E-3	-	0.5000	0.7166	-37.52	-36.72	0	11	9
Forest	$19.25 - 9.64(Forest)^{0.1}$	0.34	0.27	6E-2	-	0.1073	0.8527	-30.89	-30.09	0	11	9
Rangeland	$7.58 - 1.3 * 10^{-3}(Range)^2$	0.31	0.23	8E-2	-	0.5000	0.1240	-30.31	-29.52	0	11	9
Impervious	$4.92 + 0.18(Imperv)^2$	0.45	0.39	2E-2	-	0.5000	0.2615	-32.90	-32.10	0	11	9
WWTP	$5.06 + 8.3 * 10^{-4}(WWTP)^2$	0.52	0.47	6E-3	-	0.0305	0.1239	-34.41	-33.61	0	11	9
WWTP IDW	$5.23 + 9.9 * 10^{-6}(WWTP\ IDW)^2$	0.34	0.26	6E-2	-	0.4156	0.9077	-30.77	-29.97	0	11	9

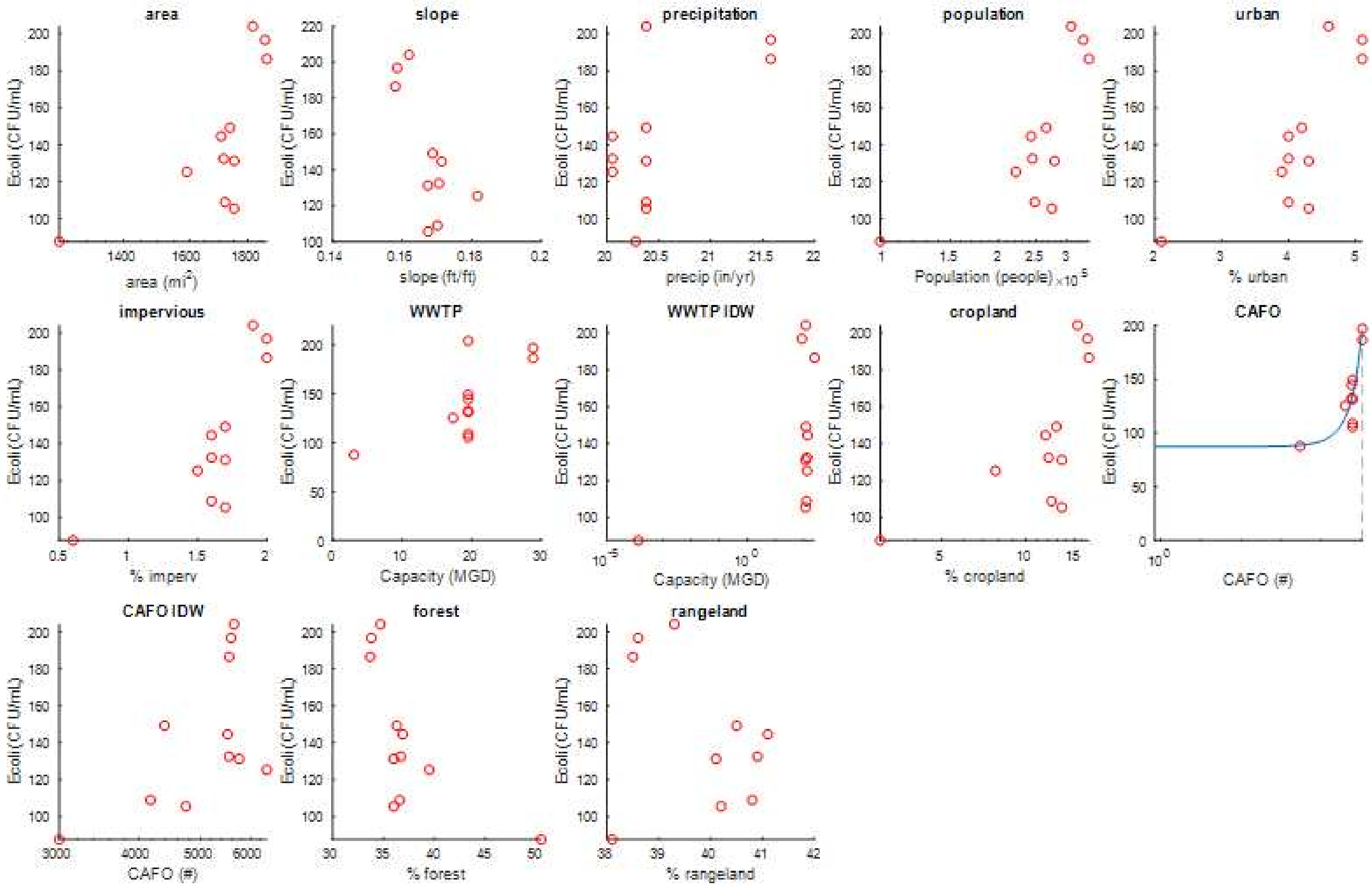
CAFO	$5.13 + 2.9 * 10^{-11}(CAFO)^2$	0.64	0.60	3E-3	-	0.1041	0.5272	-37.51	-36.72	0	11	9
CAFO IDW	$3.70 + 0.73(CAFO\ IDW)^{0.1}$	0.02	-0.09	7E-1	-	0.2492	0.4764	-26.44	-25.64	0	11	9
Population	$4.94 + 6.6 * 10^{-12}(Pop)^2$	0.55	0.50	9E-3	-	0.5000	0.7539	-35.04	-34.24	0	11	9
River Mile	$8.08 - 2.08(RM)^{0.1}$	0.61	0.57	2E-3	-	0.2406	0.6732	-36.63	-35.84	0	11	9
Stepwise	$6.13 - 0.47(Urban) + 2.0 * 10^{-5}(CAFO)$	0.75	0.69	4E-3	24.6	0.5000	0.8166	-39.58	-38.39	0	11	8
Include ag	$8.00 - 0.07(Range) + 6.18 * 10^{-6}(CAFO)$	0.72	0.65	6E-3	2.2	0.4324	0.7138	-38.31	-37.11	0	11	8
Include urban	$4.97 - 0.12(Range) + 1.48(Urban)^{0.1}$	0.69	0.62	9E-3	2	0.4198	0.9903	-37.24	-36.04	0	11	8


area
slope
precipitation
population
urban
impervious
WWTP
WWTP IDW
cropland
CAFO
CAFO IDW
forest
rangeland
Ecoli (CFU/ml)
area (mi 2)
slope (ft/ft)
precip (in/yr)
Population (people)
% urban
% imperv
Capacity (MGD)
% cropland
CAFO (#)
% forest
% rangeland

Geometric Mean

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Area	$3.80 + 3.9 * 10^{-7}(Area)^2$	0.61	0.57	4E-3	-	0.0682	0.7244	-36.33	-35.53	0	11	9
Precipitation	$2.08 + 6.8 * 10^{-3}(Precip)^2$	0.34	0.26	6E-2	-	0.2786	0.0308	-30.41	-29.61	0	11	9
Slope	$28.53 - 28.16(Slope)^{0.1}$	0.58	0.53	7E-3	-	0.1257	0.8904	-35.39	-34.60	0	11	9
Urban	$4.24 + 3.9 * 10^{-2}(Urban)^2$	0.70	0.67	1E-3	-	0.4289	0.9906	-39.23	-38.43	0	11	9
Cropland	$4.49 + 2.6 * 10^{-3}(Crop)^2$	0.66	0.62	2E-3	-	0.5000	0.5330	-37.79	-37.00	0	11	9
Forest	$22.66 - 12.35(Forest)^{0.1}$	0.55	0.50	9E-3	-	0.5000	0.4400	-34.67	-33.88	0	11	9
Rangeland	$5.82 - 5.5 * 10^{-4}(Range)^2$	0.05	-0.06	5E-1	-	0.5000	0.5257	-26.48	-25.68	0	11	9
Impervious	$4.29 + 0.23(Imperv)^2$	0.73	0.70	8E-4	-	0.0572	0.3445	-40.30	-39.51	0	11	9
WWTP	$4.44 + 7.4 * 10^{-3}(WWTP)^{1.4}$	0.59	0.55	6E-3	-	0.0891	0.0714	-35.83	-35.04	0	11	9
WWTP IDW	$4.30 + 0.41(WWTP\ IDW)^{0.1}$	0.32	0.25	7E-2	-	0.5000	0.9214	-30.20	-29.40	0	11	9
CAFO	$4.42 + 8.0 * 10^{-6}(CAFO)$	0.73	0.69	9E-4	-	0.2847	0.8787	-40.12	-39.33	0	11	9

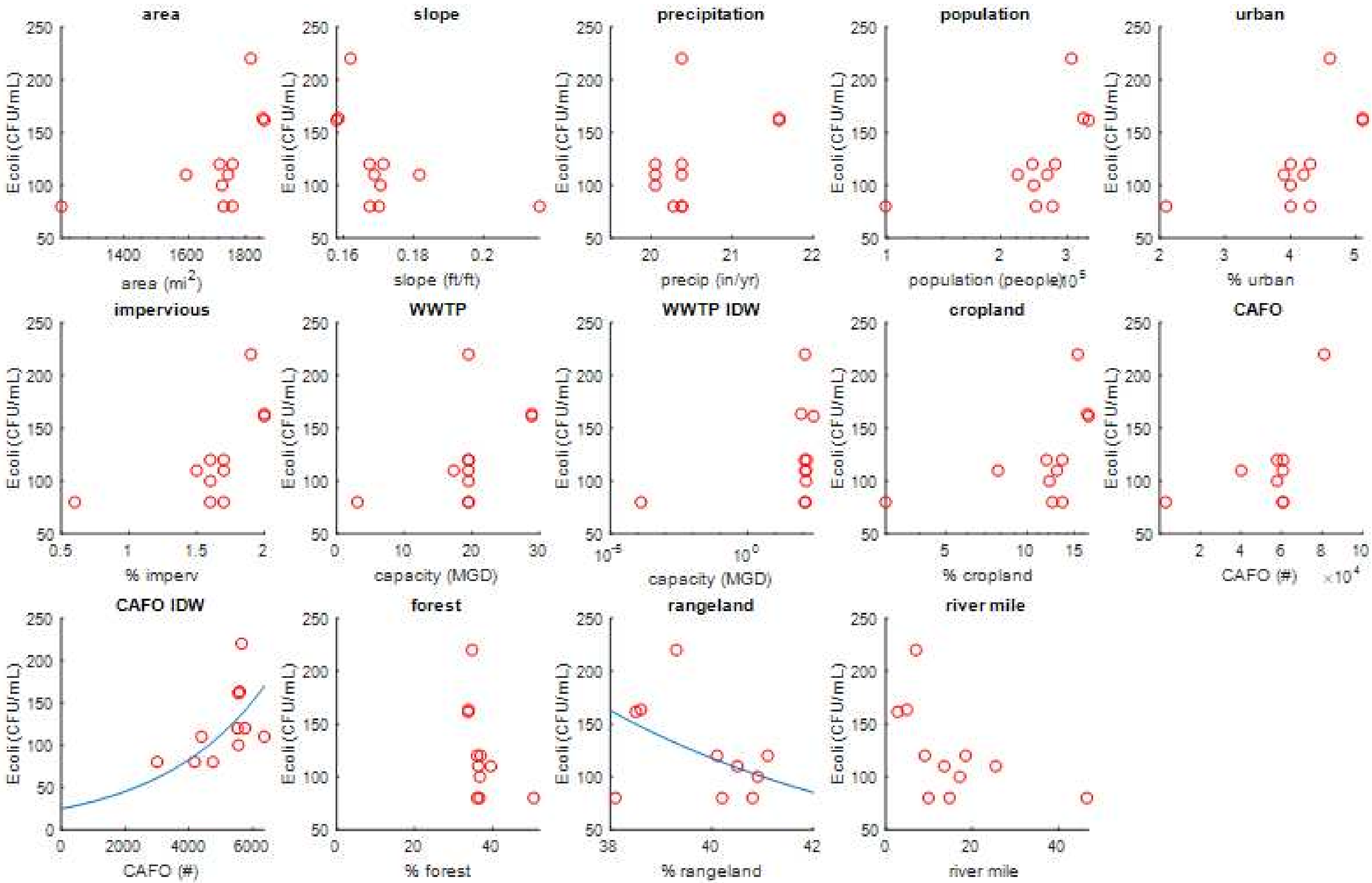
CAFO IDW	$3.70 + 0.73(CAFO\ IDW)^{0.1}$	0.41	0.35	3E-2	-	0.2563	0.5595	-31.72	-30.93	0	11	9
Population	$4.39 + 7.6 * 10^{-12}(Pop)^{2}$	0.71	0.68	1E-3	-	0.4842	0.6290	-39.46	-38.66	0	11	9
River Mile	$7.61 - 2.08(RM)^{0.1}$	0.60	0.55	5E-3	-	0.3599	0.5777	-35.92	-35.12	0	11	9
Stepwise	$4.42 + 8.0 * 10^{-6}(CAFO)$	0.73	0.69	9E-4	-	0.2847	0.8787	-40.12	-39.33	0	11	9


area
slope
precipitation
population
urban
impervious
WWTP
WWTP IDW
cropland
CAFO
CAFO IDW
forest
rangeland
Ecoli (CFU/mL)
area (mi^2)
slope (ft/ft)
precip (in/yr)
Population (people) $\times 10^5$
% urban
% imperv
Capacity (MGD)
% cropland
CAFO (#)
% forest
% rangeland

Median

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Area	$3.67 + 3.7 * 10^{-7}(Area)^2$	0.38	0.31	4E-2	-	0.3497	0.8547	-26.74	-25.94	0	11	9
Precipitation	$1.77 + 7.1 * 10^{-3}(Precip)^2$	0.25	0.16	1E-1	-	0.1720	0.0601	-24.63	-23.84	0	11	9
Slope	$26.99 - 26.53(Slope)^{0.1}$	0.34	0.27	6E-2	-	0.4668	0.6928	-26.16	-25.37	0	11	9
Urban	$4.06 + 3.9 * 10^{-2}(Urban)^2$	0.49	0.43	2E-2	-	0.3552	0.3182	-28.88	-28.09	0	11	9
Cropland	$4.30 + 2.7 * 10^{-3}(Crop)^2$	0.47	0.42	2E-2	-	0.5000	0.6010	-28.59	-27.79	0	11	9
Forest	$21.16 - 11.43(Forest)^{0.1}$	0.32	0.24	7E-2	-	0.5000	0.5324	-25.69	-24.90	0	11	9
Rangeland	$7.69 - 7.3 * 10^{-2}(Range)$	0.09	-0.01	4E-1	-	0.3654	0.1827	-22.58	-21.78	0	11	9
Impervious	$4.10 + 0.23(Imperv)^2$	0.51	0.45	1E-2	-	0.4315	0.7505	-29.35	-28.56	0	11	9
WWTP	$4.40 + 8.4 * 10^{-4}(WWTP)^2$	0.35	0.27	6E-2	-	0.0044	0.0163	-26.17	-25.38	0	11	9
WWTP IDW	$4.24 + 0.34(WWTP\ IDW)^{0.1}$	0.15	0.05	2E-1	-	0.1415	0.9366	-23.27	-22.48	0	11	9
CAFO	$4.37 + 1.3 * 10^{-8}(CAFO)^{1.5}$	0.52	0.47	1E-2	-	0.3601	0.3619	-29.63	-28.84	0	11	9

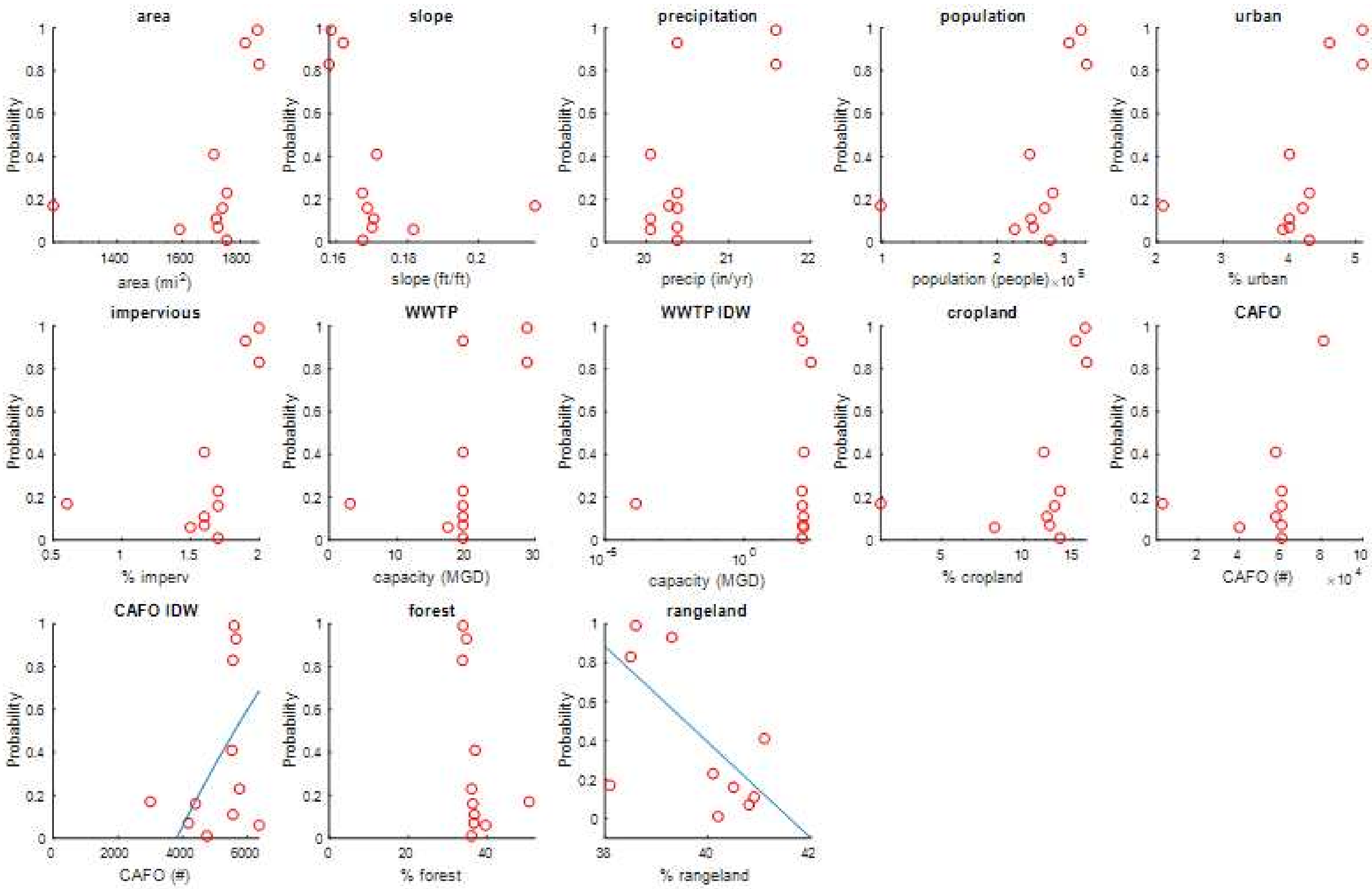
CAFO IDW	$3.30 + 3.7 * 10^{-3}(CAFO\ IDW)^{0.7}$	0.36	0.28	5E-2	-	0.5000	0.8826	-26.36	-25.57	0	11	9
Population	$4.18 + 8.0 * 10^{-12}(Pop)^{2}$	0.53	0.48	1E-2	-	0.1495	0.5138	-29.79	-29.00	0	11	9
River Mile	$7.60 - 2.21(RM)^{0.1}$	0.45	0.39	2E-2	-	0.1365	0.6562	-28.13	-27.34	0	11	9
Stepwise	$4.24 + 8.1 * 10^{-6}(CAFO)$	0.50	0.44	2E-2	-	0.2734	0.3086	-29.05	-28.26	0	11	9
Median	$9.69 - 0.16(Range) + 3.0 * 10^{-4}(CAFO\ IDW)$	0.73	0.66	6E-3	2.4	0.5000	0.9570	-33.73	-32.54	0	11	8


area
slope
precipitation
population
urban
impervious
WWTP
WWTP IDW
cropland
CAFO
CAFO IDW
forest
rangeland
river mile
Ecoli (CFU/mL)
area (mi^2)
slope (ft/ft)
precip (in/yr)
population (people) $\times 10^5$
% urban
% imperv
capacity (MGD)
% cropland
CAFO (#) $\times 10^4$
% forest
% rangeland
river mile

Vulnerability

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Area	$-0.74 + 3.7 * 10^{-7}(Area)^2$	0.30	0.22	8E-2	-	0.1057	0.2063	-22.66	-21.86	1	11	9
Precipitation	$-4.61 + 1.2 * 10^{-2}(Precip)^2$	0.53	0.48	1E-2	-	0.0164	0.0420	-27.00	-26.20	1	11	9
Slope	$22.08 - 25.90(Slope)^{0.1}$	0.25	0.17	1E-1	-	0.0507	0.1780	-21.92	-21.12	1	11	9
Urban	$-0.41 + 0.04(Urban)^2$	0.46	0.40	2E-2	-	0.0271	0.3104	-25.48	-24.69	1	11	9
Cropland	$-0.19 + 3.3 * 10^{-3}(Crop)^2$	0.54	0.49	1E-2	-	0.1669	0.7271	-27.17	-26.37	1	11	9
Forest	$15.71 - 10.69(Forest)^{0.1}$	0.21	0.13	2E-1	-	0.0525	0.1622	-21.34	-20.55	1	11	9
Rangeland	$207.65 - 144.74(Range)^{0.1}$	0.35	0.28	5E-2	-	0.2984	0.5474	-23.51	-22.71	1	11	9
Impervious	$-0.33 + 0.25(Imperv)^2$	0.44	0.37	3E-2	-	0.0381	0.2506	-25.00	-24.20	1	11	9
WWTP	$-2.91 + 2.9 * 10^{-3}(WWTP)^2$	0.47	0.41	2E-2	-	0.4881	0.3319	-25.65	-24.85	1	11	9
WWTP IDW	$0.22 + 7.7 * 10^{-6}(WWTP\ IDW)^2$	0.10	0.00	3E-1	-	0.3025	0.0401	-19.88	-19.09	1	11	9
CAFO	$-0.06 + 8.8 * 10^{-11}(CAFO)^2$	0.72	0.68	1E-3	-	0.3323	0.9823	-32.53	-31.74	1	11	9

CAFO IDW	$-0.93 + 1.8 * 10^{-2}(CAFO\ IDW)^{0.5}$	0.12	0.2	3E-1	-	0.4975	0.4900	-20.07	-19.27	1	11	9
Population	$-0.29 + 9.1 * 10^{-12}(Pop)^{2}$	0.52	0.47	1E-2	-	0.2806	0.5260	-26.85	-26.05	1	11	9
River Mile	$3.73 - 2.62(RM)^{0.1}$	0.49	0.43	2E-2	-	0.5000	0.9559	-26.10	-25.30	1	11	9
Stepwise	$6.18 - 0.81(Precip) + 0.20(Forest) + 5.0 * 10^{-5}(CAFO)$	0.92	0.88	4E-4	71.8	0.4431	0.3438	-41.90	-40.31	1	11	7
Reduced Collinearity	$37.23 - 8.69(Precip)^{0.5} + 8.0 * 10^{-4}(Forest)^{2} + 2.8 * 10^{-10}(CAFO)^{2}$	0.92	0.88	4E-4	35.9	0.5000	0.8061	-41.87	-40.28	1	11	7
Signs make sense	$7.51 - 0.25(Range) + 3.8 * 10^{-2}(CAFO\ IDW)^{0.5}$	0.78	0.72	2E-3	2.4	0.5000	0.9132	-33.25	-32.06	1	11	8

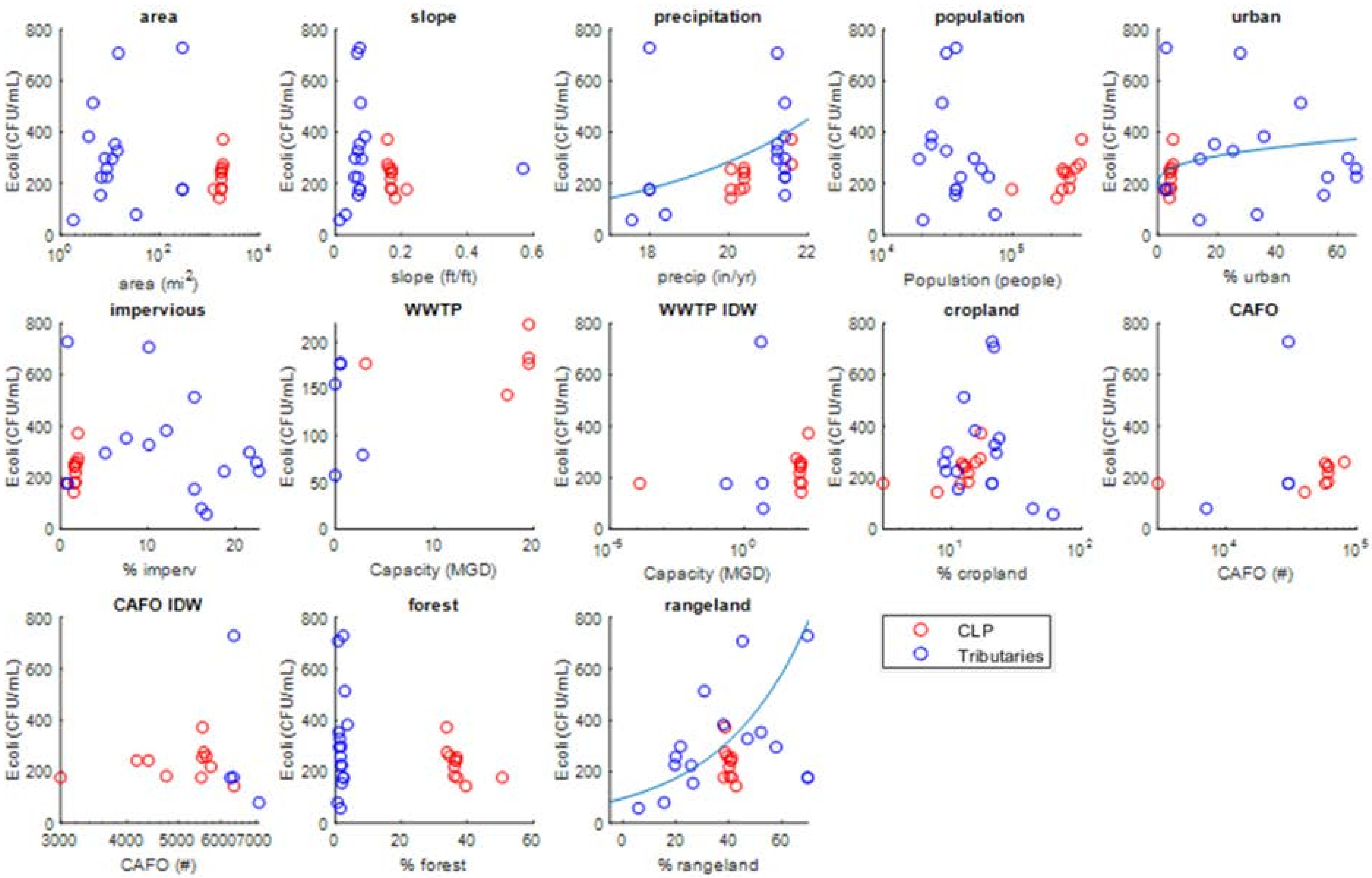

area
slope
precipitation
population
urban
impervious
WWTP
WWTP IDW
cropland
CAFO
CAFO IDW
forest
rangeland
Probability
area (mi^2)
slope (ft/ft)
precip (in/yr)
population (people) $\times 10^5$
% urban
% imperv
capacity (MGD)
% cropland
CAFO (#) $\times 10^4$
CAFO (#)
% forest
% rangeland

Cache la Poudre Watershed

Mean

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	**N**	**DOF**
Area	$53.53 + 1.7 * 10^{-4}(Area)^{0.85}$	0.01	-0.03	7E-1	-	0.3459	-18.02	-29.21	-26.61	0	11	9
Precipitation	$3.32 + 5.2 * 10^{-3}(Precip)^{2}$	0.22	0.19	1E-2	-	0.0084	0.0131	-35.88	-33.29	0	11	9
Slope	$3.66 - 2.29(Slope)^{0.1}$	0.05	0.02	2E-1	-	0.2698	0.8922	-30.55	-27.96	0	11	9
Urban	$5.08 + 0.32(Urban)^{0.1}$	0.01	-0.03	6E-1	-	0.0563	0.1541	-29.26	-26.66	0	11	9
Cropland	$5.67 + 4.3 * 10^{-4}(Crop)^{2}$	0.32	0.29	2E-3	-	0.1518	0.1439	-39.25	-36.66	0	11	9
Forest	$5.55 - 1.2 * 10^{-4}(Forest)^{2}$	0.02	-0.01	4E-1	-	0.3690	0.0940	-29.69	-27.10	0	11	9
Rangeland	$-0.57 + 4.24(Range)^{0.1}$	0.31	0.28	3E-3	-	0.5000	0.6330	-38.92	-36.32	0	11	9
Impervious	$5.54 - 4.1 * 10^{-4}(Imperv)^{2}$	0.02	-0.02	5E-1	-	0.3446	0.4873	-29.47	-26.88	0	11	9
WWTP	$5.59 - 0.15(WWTP)^{0.1}$	0.03	-0.01	4E-1	-	0.0901	0.1061	-29.81	-27.22	0	11	9
WWTP IDW	$5.56 - 0.10(WWTP\ IDW)^{0.1}$	0.02	-0.02	5E-1	-	0.4036	0.1582	-29.58	-26.99	0	11	9
CAFO	$5.58 - 0.06(CAFO)^{0.1}$	0.02	-0.01	4E-1	-	0.4313	0.2156	-29.71	-27.12	0	11	9

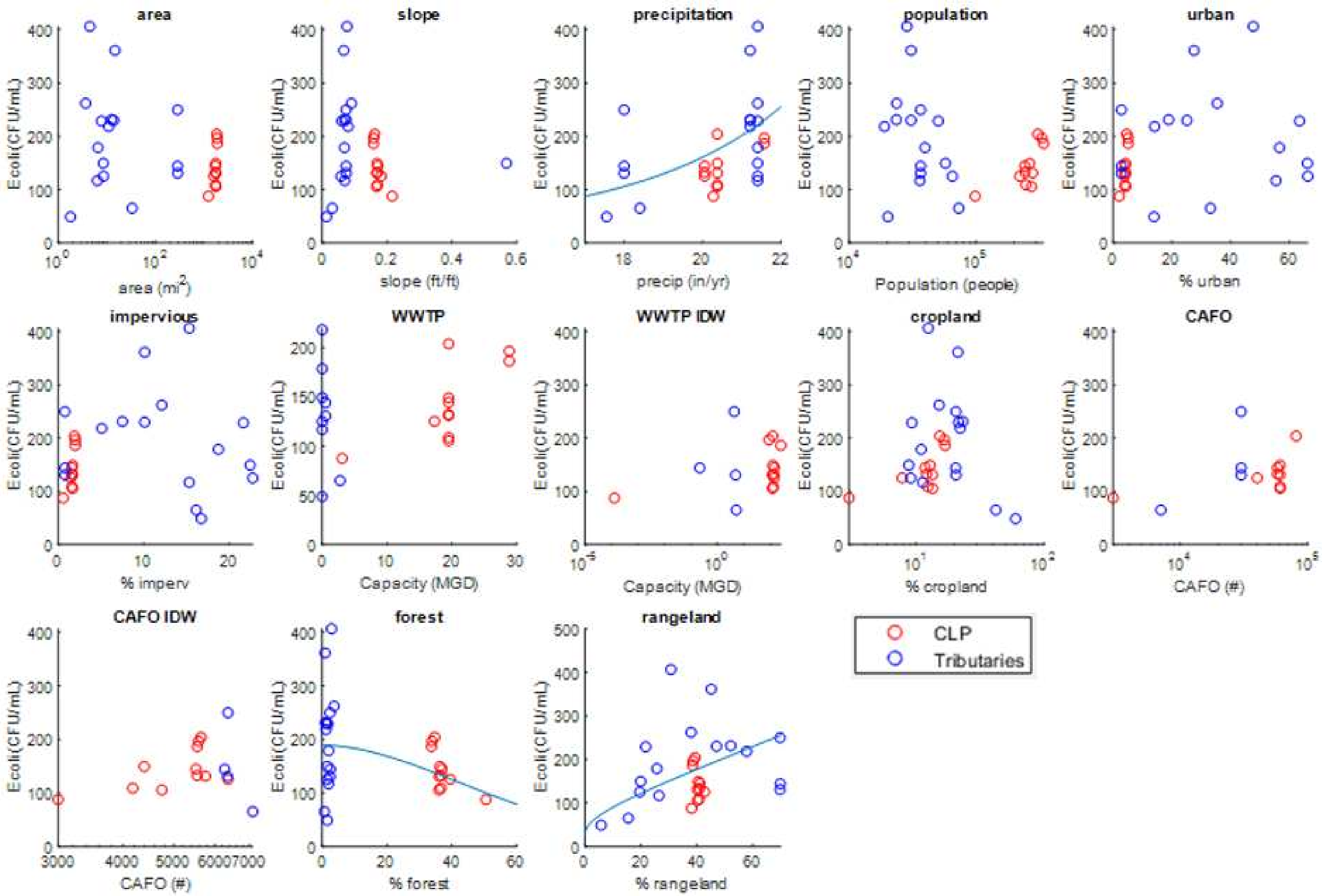
CAFO IDW	$5.61 - 7.1 * 10^{-9}(CAFO\ IDW)^2$	0.05	0.01	3E-1	-	0.163 6	0.083 5	-30.44	-27.85	0	11	9
Population	$6.10 - 0.20(Pop)^{0.1}$	0.01	-0.03	6E-1	-	0.500 0	0.159 4	-29.40	-26.81	0	11	9
River Mile	$5.38 + 1.0 * 10^{-4}(RM)^2$	0.02	-0.02	5E-1	-	0.500 0	0.325 1	-29.56	-26.97	0	11	9
Stepwise	$-0.87 + 0.27(Precip) + 0.02(Range)$	0.56	0.52	1E-4	2.2	0.107 2	0.017 2	-49.03	-45.15	0	27	23
Best	$-0.68 + 0.23(Precip) + 0.26(Urban)^{0.25} + 0.03(Range)$	0.60	0.54	9E-5	4.8	0.319 5	0.128 6	-49.57	-44.38	0	27	23


area
slope
precipitation
population
urban
impervious
WWTP
WWTP IDW
cropland
CAFO
CAFO IDW
forest
rangeland
Ecoli (CFU/mL)
area (mi^2)
slope (ft/ft)
precip (in/yr)
Population (people)
% urban
% imperv
Capacity (MGD)
% cropland
CAFO (#)
% forest
% rangeland
CLP
Tributaries

Geometric Mean

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Area	$5.20 + 1.9 * 10^{-2}(Area)^{0.35}$	0.05	0.01	3E-1	-	0.5000	0.2414	-38.48	-35.89	0	27	25
Precipitation	$2.96 + 5.0 * 10^{-3}(Precip)^{2}$	0.28	0.26	4E-3	-	0.5000	0.8653	-46.24	-43.65	0	27	25
Slope	$4.29 - 0.96(Slope)^{0.1}$	0.01	-0.03	6E-1	-	0.5000	0.2557	-37.59	-35.00	0	27	25
Urban	$4.14 + 0.72(Urban)^{0.1}$	0.06	0.02	2E-1	-	0.5000	0.4735	-38.81	-36.22	0	27	25
Cropland	$5.2 + 3.4 * 10^{-4}(Crop)^{2}$	0.27	0.24	6E-3	-	0.2804	0.7205	-45.69	-43.10	0	27	25
Forest	$5.17 - 1.9 * 10^{-4}(Forest)^{2}$	0.09	0.05	1E-1	-	0.4409	0.0867	-39.75	-37.16	0	27	25
Rangeland	$0.35 + 3.29(Range)^{0.1}$	0.25	0.22	8E-3	-	0.5000	0.5152	45.04	-42.44	0	27	25
Impervious	$4.61 + 0.39(Imperv)^{0.1}$	0.01	-0.03	6E-1	-	0.5000	0.5396	-37.62	-35.02	0	27	25
WWTP	$5.22 - 0.23(WWTP)^{0.1}$	0.10	0.06	1E-1	-	0.5000	0.2892	-40.02	-37.43	0	27	25
WWTP IDW	$5.18 - 0.16(WWTP\ IDW)^{0.1}$	0.07	0.03	1E-1	-	0.5000	0.4160	-39.05	-36.46	0	27	25
CAFO	$5.21 - 0.09(CAFO)^{0.1}$	0.09	0.05	1E-1	-	0.5000	0.3429	-39.65	-37.09	0	27	25
CAFO IDW	$5.23 - 0.12(CAFO\ IDW)^{0.1}$	0.11	0.08	9E-2	-	0.2767	0.2131	-40.47	-37.88	0	27	25
Population	$6.09 - 0.33(Pop)^{0.1}$	0.05	0.01	3E-1	-	0.2544	0.2162	-38.69	-36.10	0	27	25
River Mile	$4.91 + 1.4 * 10^{-4}(RM)^{2}$	0.05	0.01	3E-2	-	0.4273	0.4423	-38.60	-36.00	0	27	25
Stepwise	$-0.80 + 0.26(Precip) - 8.9 * 10^{-3}(Forest) + 0.02(Range)$	0.68	0.64	7E-6	3.2	0.5000	0.5384	-64.07	-58.89	0	27	23

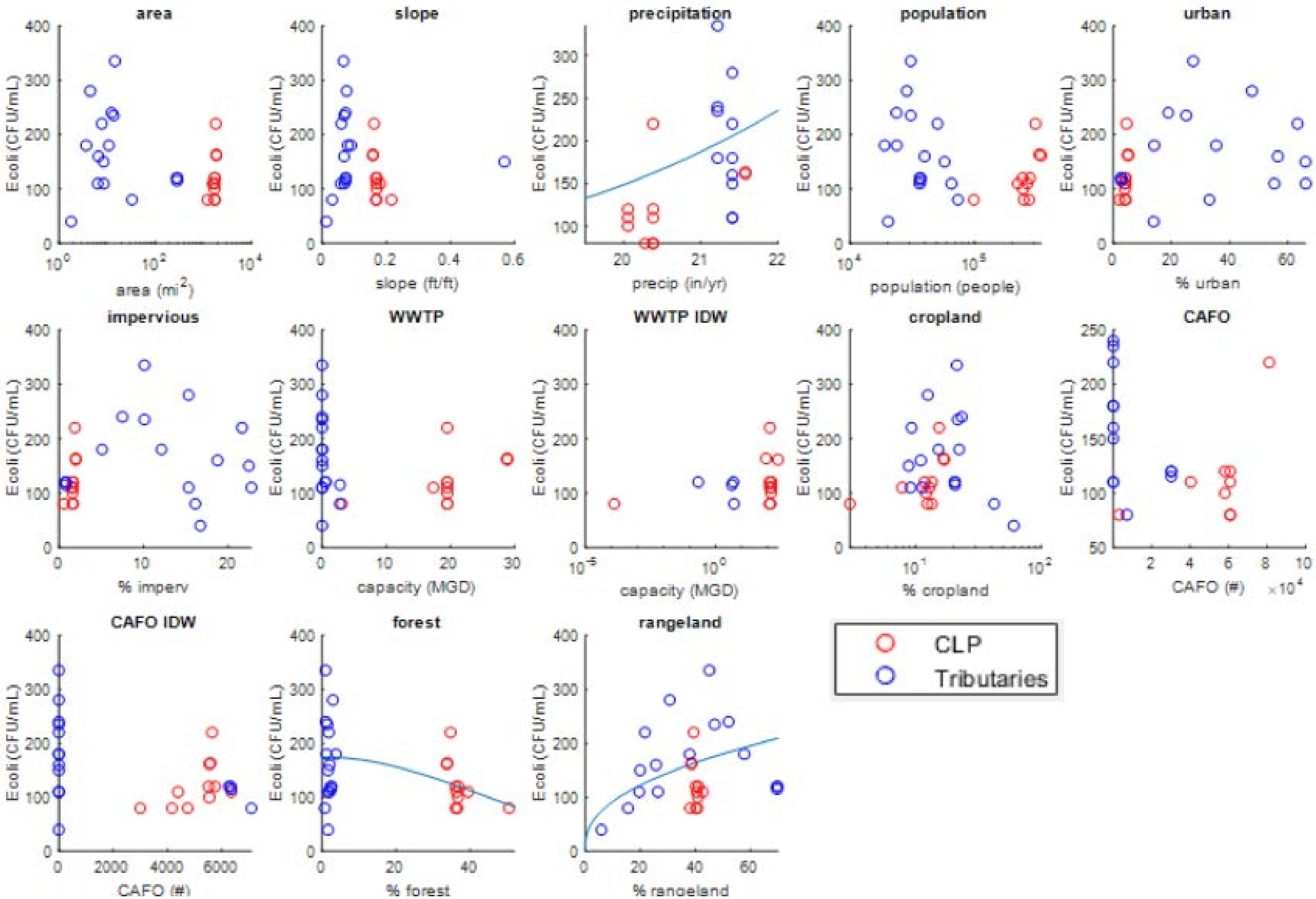
Best	$1.60 + 7.0 * 10^{-3}(Precip)^2 - 1.4 * 10^{-5}(CAFO\ IDW) - 2.5 * 10^{-4}(Forest)^2 + 0.02(Range)$	0.70	0.64	2E-5	8.6	0.3252	0.4791	-63.38	-56.90	0	27	23


area
slope
precipitation
population
urban
impervious
WWTP
WWTP IDW
cropland
CAFO
CAFO IDW
forest
rangeland
Ecoli(CFU/mL)
area (mi2)
slope (ft/ft)
precip (in/yr)
Population (people)
% urban
% imperv
Capacity (MGD)
% cropland
CAFO (#)
% forest
% rangeland
CLP
Tributaries

Median

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Area	$5.40 - 0.30(Area)^{0.1}$	0.07	0.04	2E-1	-	0.5000	0.1351	-40.24	-37.64	0	27	25
Precipitation	$2.52 + 5.7 * 10^{-3}(Precip)^2$	0.38	0.35	6E-4	-	0.2057	0.3335	-51.06	-48.47	0	27	25
Slope	$4.24 - 0.84(Slope)^{0.1}$	0.01	-0.03	6E-1	-	0.2314	0.3636	-38.46	-35.86	0	27	25
Urban	$3.51 + 1.08(Urban)^{0.1}$	0.13	0.10	6E-2	-	0.5000	0.5913	-42.08	-39.49	0	27	25
Cropland	$5.03 + 3.1 * 10^{-4}(Crop)^2$	0.23	0.20	1E-2	-	0.1172	0.7741	-45.35	-42.76	0	27	25
Forest	$5.02 - 2.0 * 10^{-4}(Forest)^2$	0.11	0.7	1E-1	-	0.3336	0.2266	-41.20	-38.66	0	27	25
Rangeland	$0.35 + 3.29(Range)^{0.1}$	0.15	0.12	4E-2	-	0.3283	0.3644	-42.71	-40.12	0	27	25
Impervious	$3.95 + 0.82(Imperv)^{0.1}$	0.06	0.03	2E-1	-	0.5000	0.2766	-39.99	-37.40	0	27	25
WWTP	$5.09 - 0.28(WWTP)^{0.1}$	0.15	0.11	5E-2	-	0.2131	0.0238	-42.48	-39.89	0	27	25
WWTP IDW	$5.06 - 0.20(WWTP\ IDW)^{0.1}$	0.11	0.07	1E-1	-	0.2976	0.1073	-41.21	-38.62	0	27	25
CAFO	$5.09 - 0.12(CAFO)^{0.1}$	0.14	0.10	6E-2	-	0.2674	0.0332	-42.19	-39.59	0	27	25

CAFO IDW	$5.11 - 0.16(CAFO\ IDW)^{0.1}$	0.16	0.13	4E-2	-	0.2561	0.0168	-43.03	-40.44	0	27	25
Population	$5.92 - 0.33(Pop)^{0.1}$	0.05	0.02	2E-1	-	0.5000	0.2752	-39.66	-37.07	0	27	25
River Mile	$4.91 + 1.4 * 10^{-4}(RM)^2$	0.05	0.02	2E-1	-	0.5000	0.9454	-39.70	-37.11	0	27	25
Stepwise	$-1.11 + 0.28(Precip) - 9.7 * 10^{-3}(Forest) + 0.01(Range)$	0.69	0.65	5E-6	3.2	0.0495	0.2210	-65.83	-60.65	0	27	23
Best	$-44.64 + 33.58(Precip)^{0.1} - 2.0 * 10^{-4}(Forest)^2 + 3.02(Range)^{0.1}$	0.71	0.68	2E-6	3	0.0311	0.1661	-67.87	-62.69	0	27	23

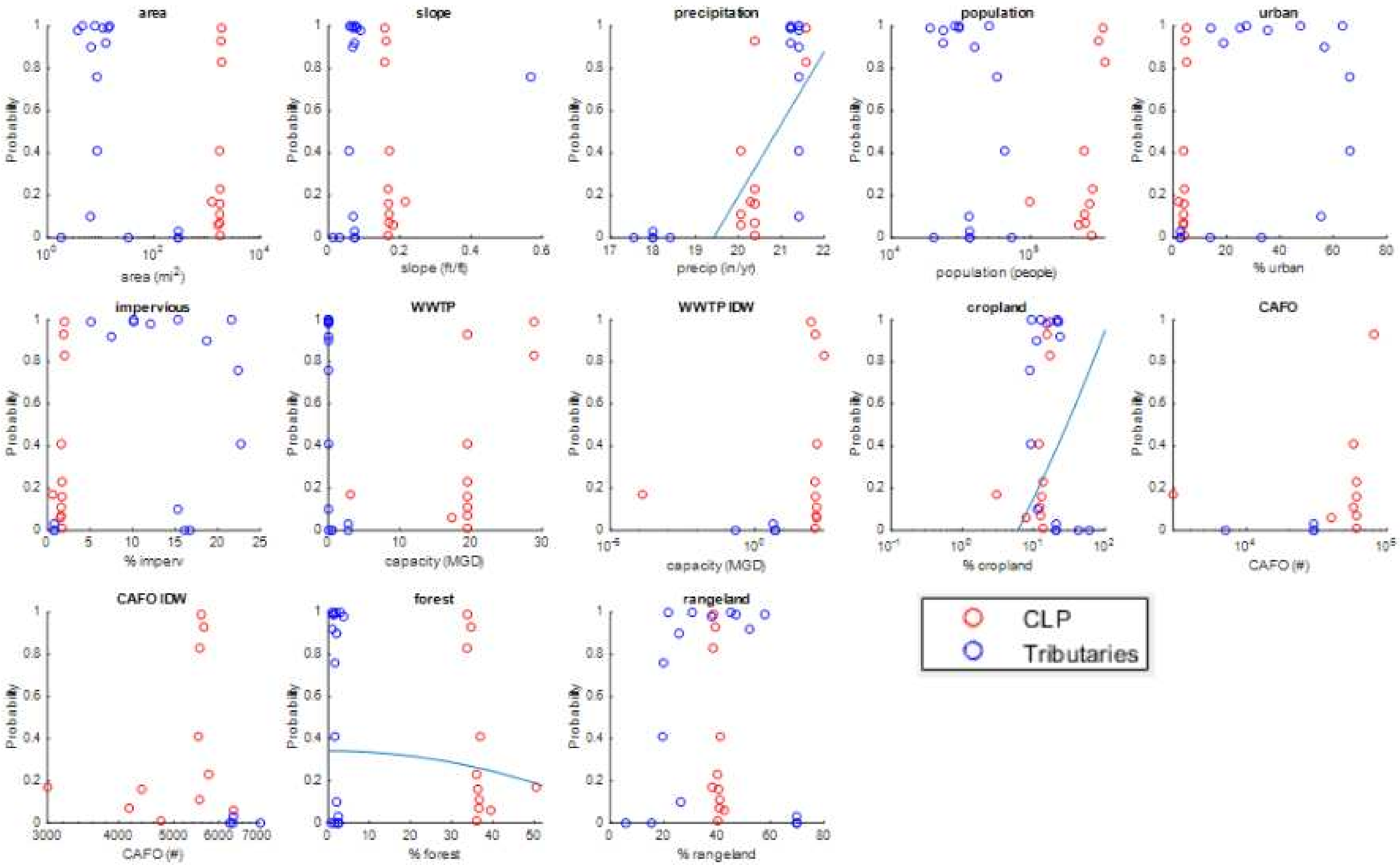


area
slope
precipitation
population
urban
impervious
WWTP
WWTP IDW
cropland
CAFO
CAFO IDW
forest
rangeland
Ecoli (CFU/mL)
area (mi^2)
slope (ft/ft)
precip (in/yr)
population (people)
% urban
% imperv
capacity (MGD)
capacity (MGD)
% cropland
CAFO (#)
CAFO (#)
% forest
% rangeland
CLP
Tributaries

Vulnerability

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Area	$1.14 - 0.40(Area)^{0.1}$	0.15	0.12	4E-2	-	0.1892	0.5873	-46.80	-44.21	1	27	25
Precipitation	$-2.17 + 6.4 * 10^{-3}(Precip)^2$	0.55	0.52	1E-5	-	0.0201	0.0698	-63.84	-61.24	1	27	25
Slope	$0.47 + 0.52(Slope)^2$	0.01	-0.03	7E-1	-	0.0019	0.0002	-42.47	-39.88	1	27	25
Urban	$-1.27 + 1.37(Urban)^{0.1}$	0.25	0.22	8E-3	-	0.5000	0.8751	-50.12	-47.52	1	27	25
Cropland	$0.55 - 1.5 * 10^{-4}(Crop)^2$	0.06	0.02	2E-1	-	0.0092	0.0005	-44.01	-41.42	1	27	25
Forest	$0.58 - 1.7 * 10^{-4}(Forest)^2$	0.09	0.05	1E-1	-	0.0154	0.0144	-44.76	-42.17	1	27	25
Rangeland	$0.59 - 6.1 * 10^{-5}(Range)^2$	0.04	-0.01	3E-1	-	0.0027	0.0028	-43.31	-40.72	1	27	25
Impervious	$-1.10 + 1.36(Imperv)^{0.1}$	0.21	0.18	2E-2	-	0.3462	0.4661	-48.66	-46.07	1	27	25
WWTP	$0.72 - 0.34(WWTP)^{0.1}$	0.25	0.22	8E-3	-	0.4789	0.8972	-50.15	-47.55	1	27	25
WWTP IDW	$0.68 - 0.26(WWTP\ IDW)^{0.1}$	0.20	0.17	9E-3	-	0.3759	0.5917	-48.49	-45.90	1	27	25
CAFO	$0.73 - 0.15(CAFO)^{0.1}$	0.28	0.25	5E-3	-	0.3778	0.7425	-51.01	-48.42	1	27	25

CAFO IDW	$0.75 - 0.21(CAFO\ IDW)^{0.1}$	0.33	0.30	2E-1	-	0.2470	0.6469	-52.97	-50.38	1	27	25
Population	$1.42 - 0.30(Pop)^{0.1}$	0.05	0.01	3E-1	-	0.0142	0.0174	-43.73	-41.17	1	27	25
River Mile	$0.34 + 1.3 * 10^{-4}(RM)^2$	0.05	0.01	3E-1	-	0.0216	0.0059	-43.74	-41.15	1	27	25
Stepwise	$-5.92 + 0.34(Precip) - 8.8 * 10^{-3}(Urban) - 0.02(Forest)$	0.71	0.68	2E-6	7	0.3981	0.2775	-71.99	-66.80	1	27	23
Reduced Collinearity	$-2.60 + 8.1 * 10^{-3}(Precip)^2 - 5.0 * 10^{-4}(Urban)^{1.65} - 2.0 * 10^{-3}(Forest)^{1.5}$	0.72	0.68	1E-6	5.4	0.3718	0.1810	-72.70	-67.52	1	27	23
Signs make sense	$-9.83 + 0.34(Precip) + 2.44(Crop)^{0.1} - 6.1 * 10^{-5}(Forest)^2$	0.69	0.65	5E-6	3.1	0.3124	0.6162	-70.05	-64.87	1	27	23


area
slope
precipitation
population
urban
impervious
WWTP
WWTP IDW
cropland
CAFO
CAFO IDW
forest
rangeland
Probability
area (mi^2)
slope (ft/ft)
precip (in/yr)
population (people)
% urban
% imperv
capacity (MGD)
% cropland
CAFO (#)
% forest
% rangeland
CLP
Tributaries

Appendix D

Cache la Poudre River

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Mean	$8.00 - 0.07(Range) + 6.18 * 10^{-6}(CAFO)$	0.72	0.65	6E-3	2.2	0.4324	0.7138	-38.31	-37.11	0	11	8
Geo Mean	$4.47 + 1.37 * 10^{-6}(CAFO)^{1.15}$	0.73	0.70	8E-4	-	0.5000	0.8064	-40.23	-39.43	0	11	9
Median	$9.69 - 0.16(Range) + 3.0 * 10^{-4}(CAFO\ IDW)$	0.73	0.66	6E-3	2.4	0.5000	0.9570	-33.73	-32.54	0	11	8
Vul	$7.51 - 0.25(Range) + 3.8 * 10^{-2}(CAFO\ IDW)^{0.5}$	0.78	0.72	2E-3	2.4	0.5000	0.9132	-33.25	-32.06	1	11	8

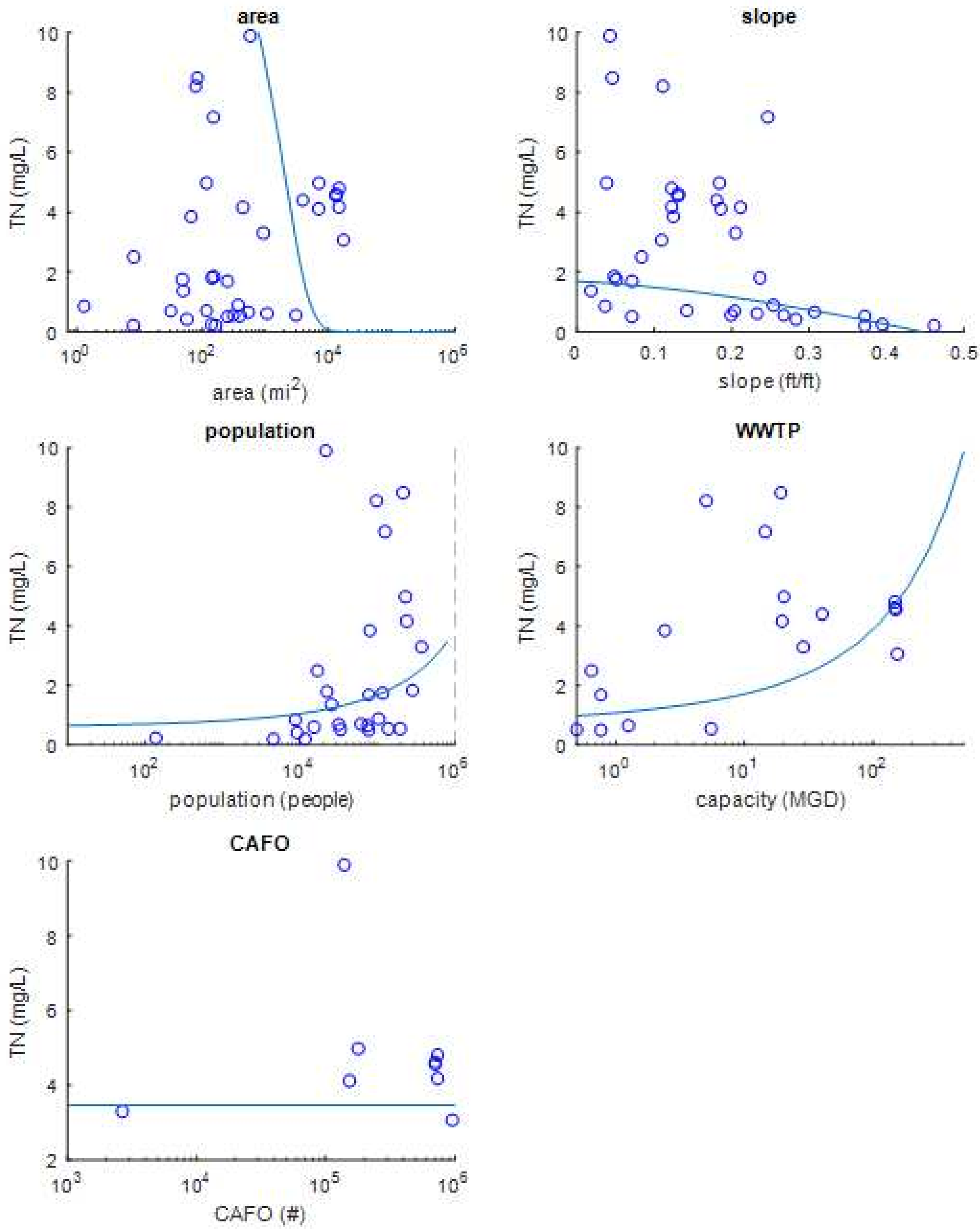
Cache la Poudre Watershed

Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DOF
Mean	$-0.68 + 0.23(Precip) + 0.26(Urban)^{0.25} + 0.03(Range)$	0.60	0.54	9E-5	4.8	0.3195	0.1286	-49.57	-44.38	0	27	23
Geo Mean	$0.98 + 5.5 * 10^{-3}(Precip)^2 - 4.5 * 10^{-4}(Forest)^{1.85} + 0.66(Range)^{0.3}$	0.70	0.67	3E-6	3	0.3294	0.5080	-66.11	-60.93	0	27	23
Median	$-8.10 + 0.23(Precip) - 2.76 * 10^{-4}(forest)^2 + 7.14(Range)^{0.05}$	0.71	0.68	2E-6	3	0.0306	0.1713	-67.87	-62.68	0	27	23
Vul	$-9.83 + 0.34(Precip) + 2.44(Crop)^{0.1} - 6.1 * 10^{-5}(Forest)^2$	0.69	0.65	5E-6	3.1	0.3124	0.6162	-70.05	-64.87	1	27	23

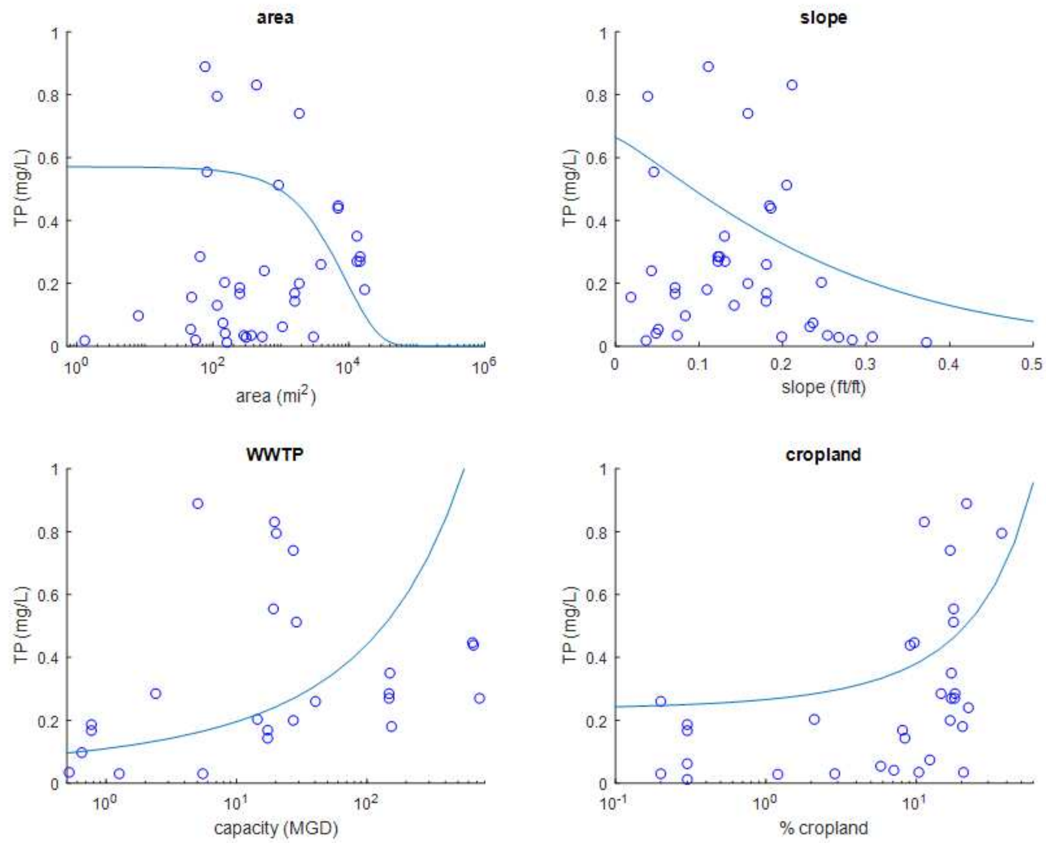
Appendix E

Upstream

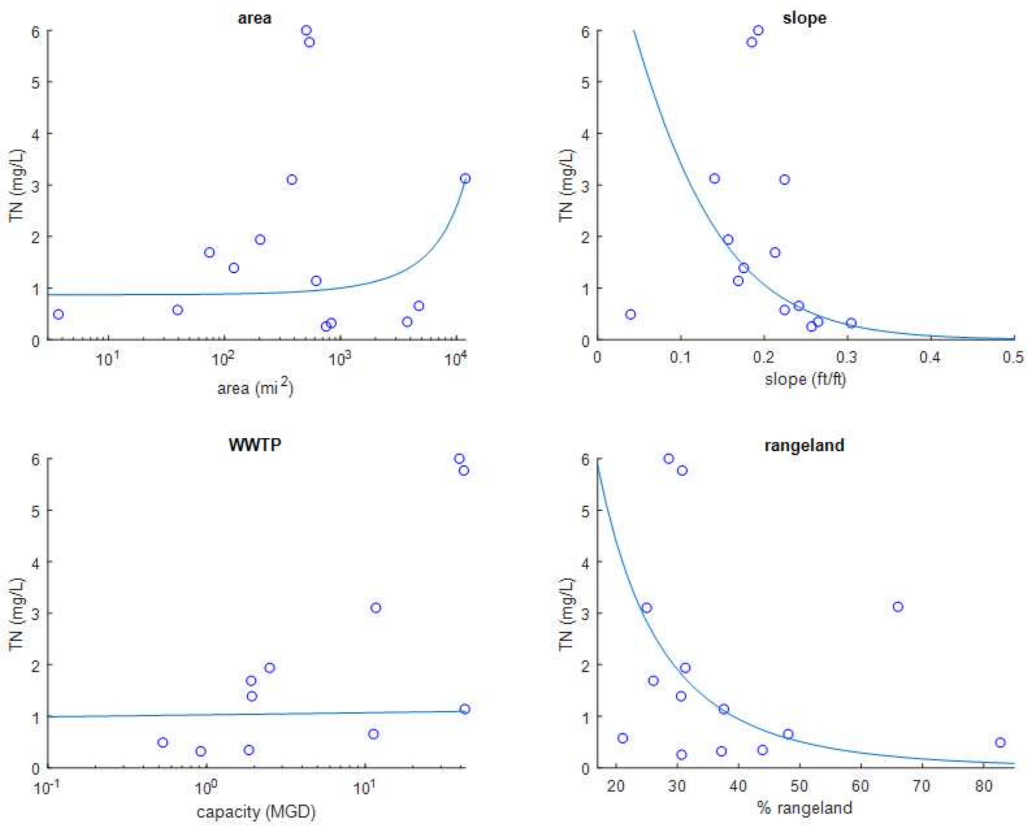
Division 1 TN



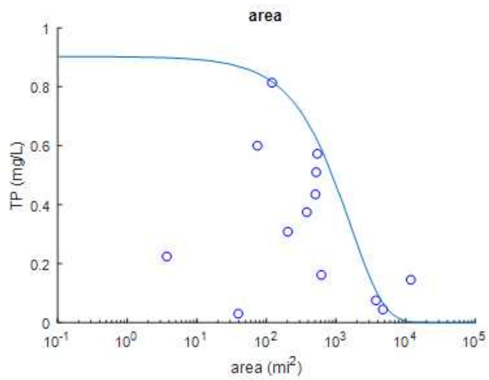
Division 1 TP

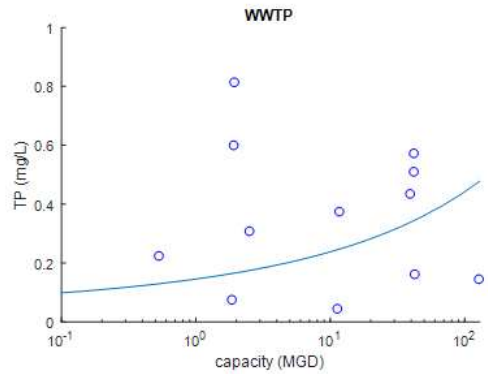


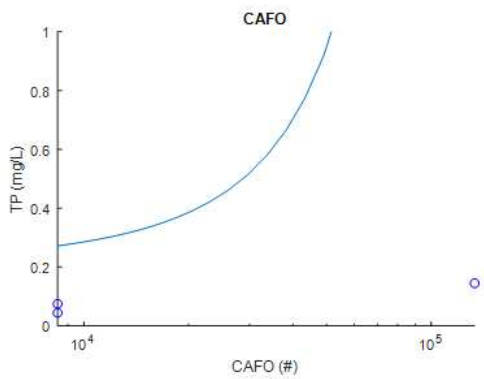
Division 2 TN



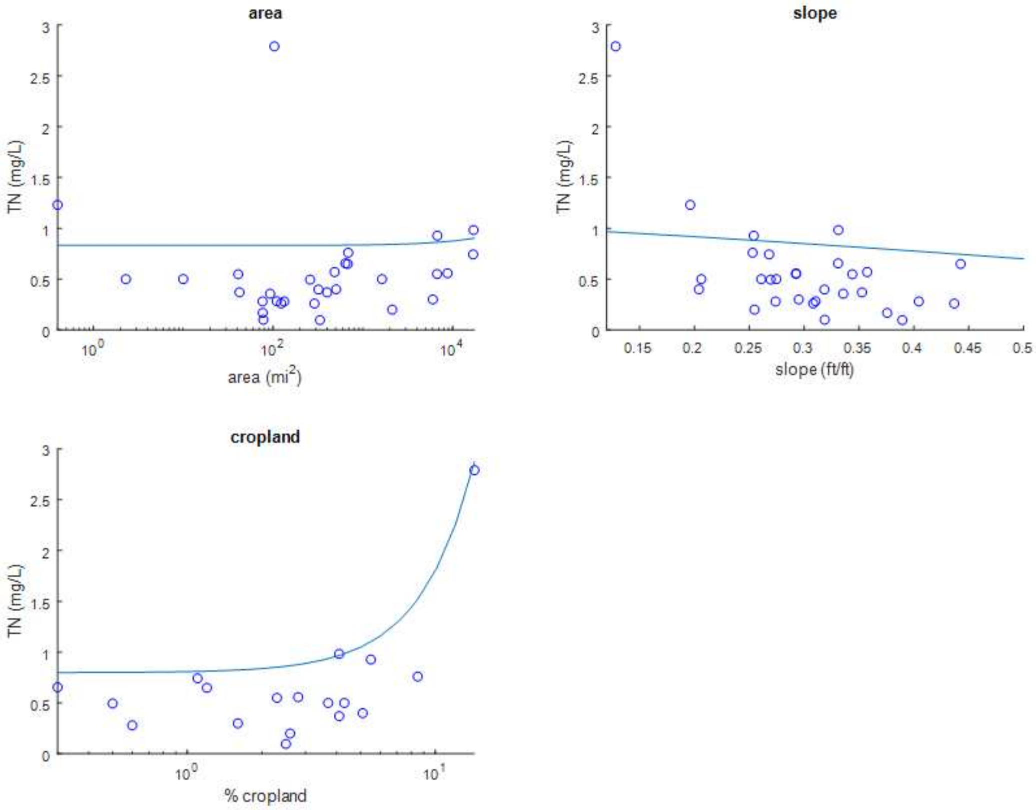
Division 2 TP



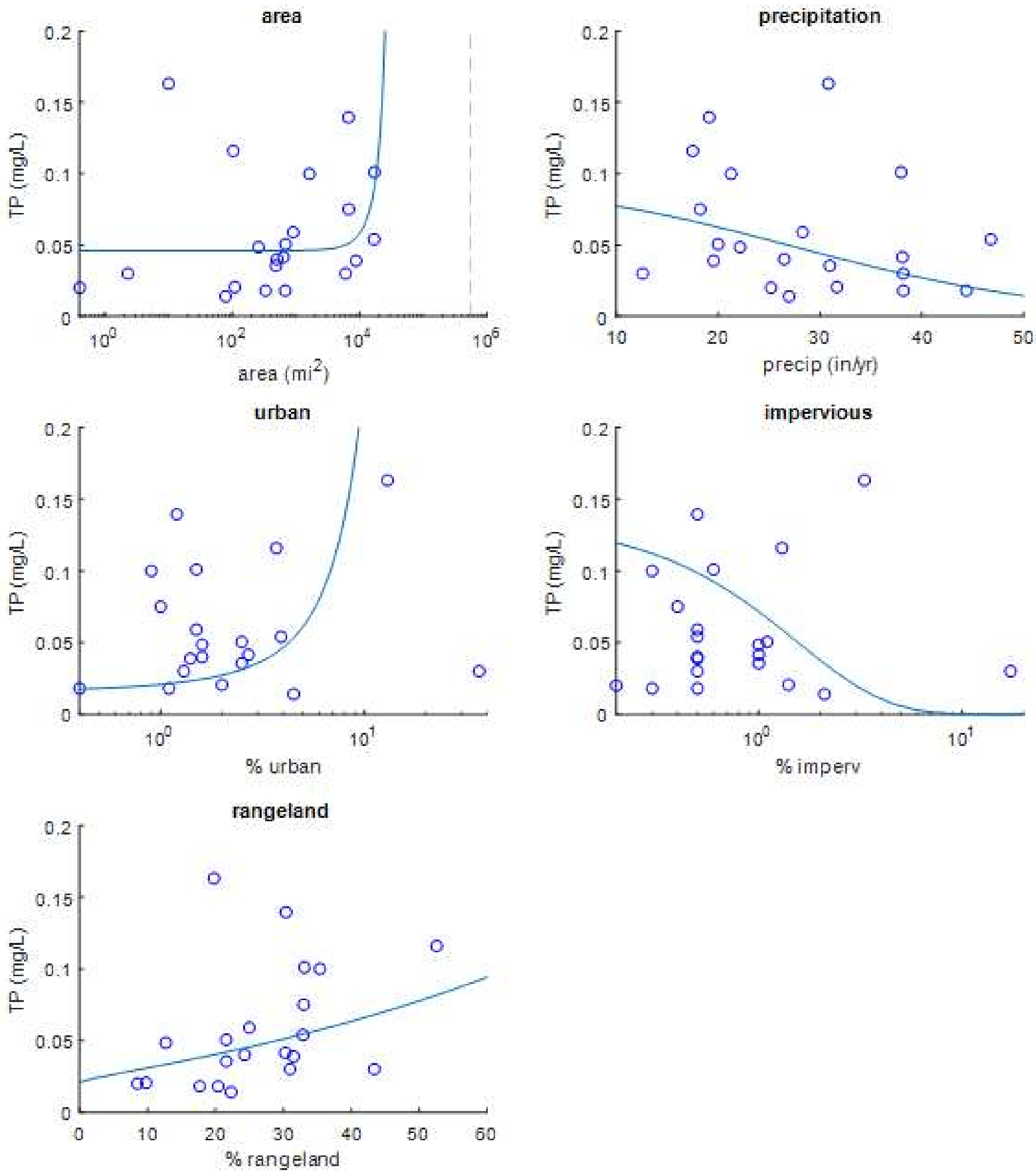




Western Slope TN

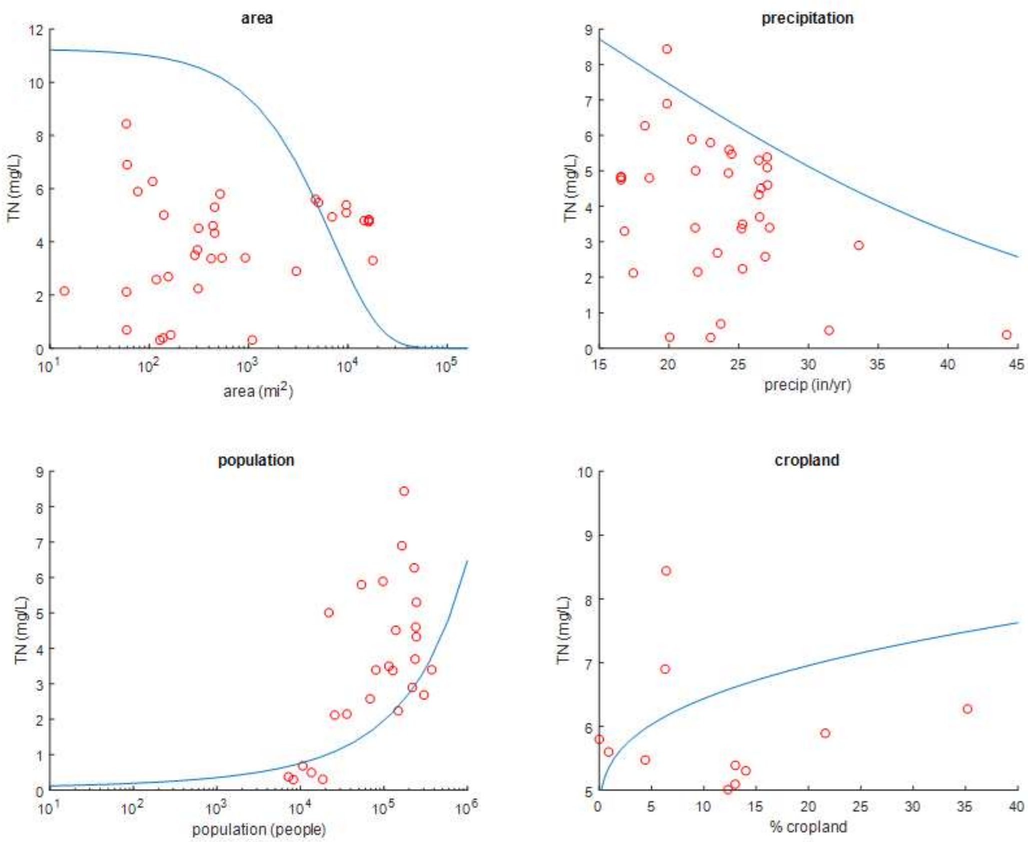


Western Slope TP

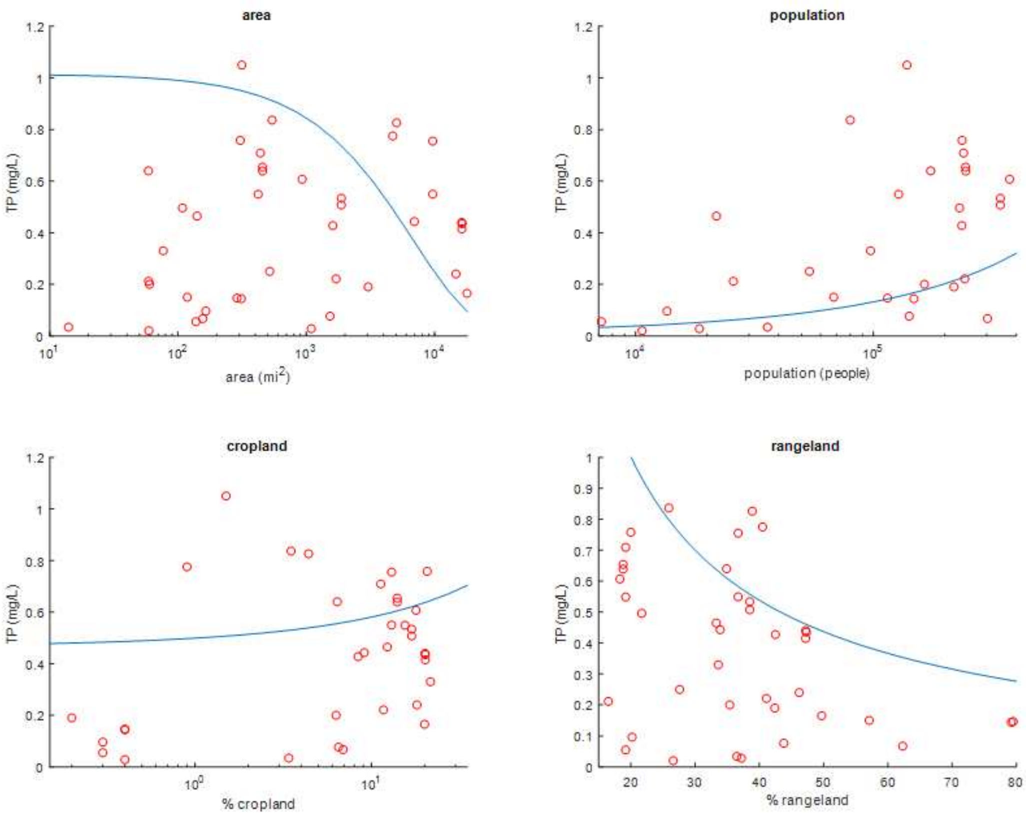


Downstream

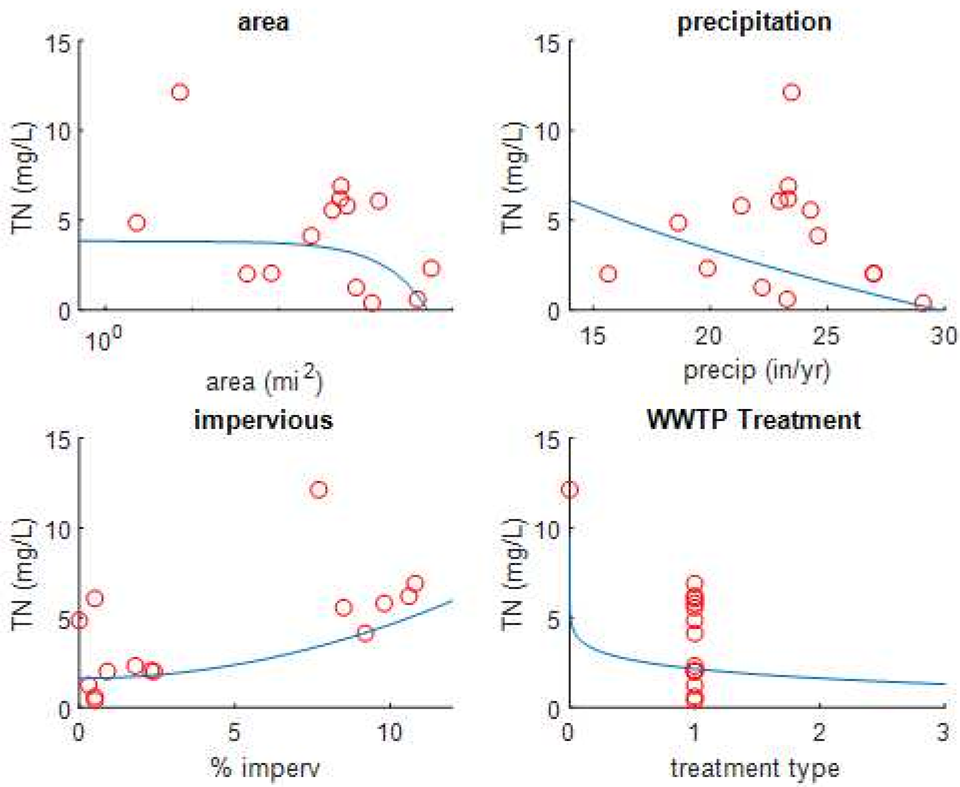
Division 1 TN



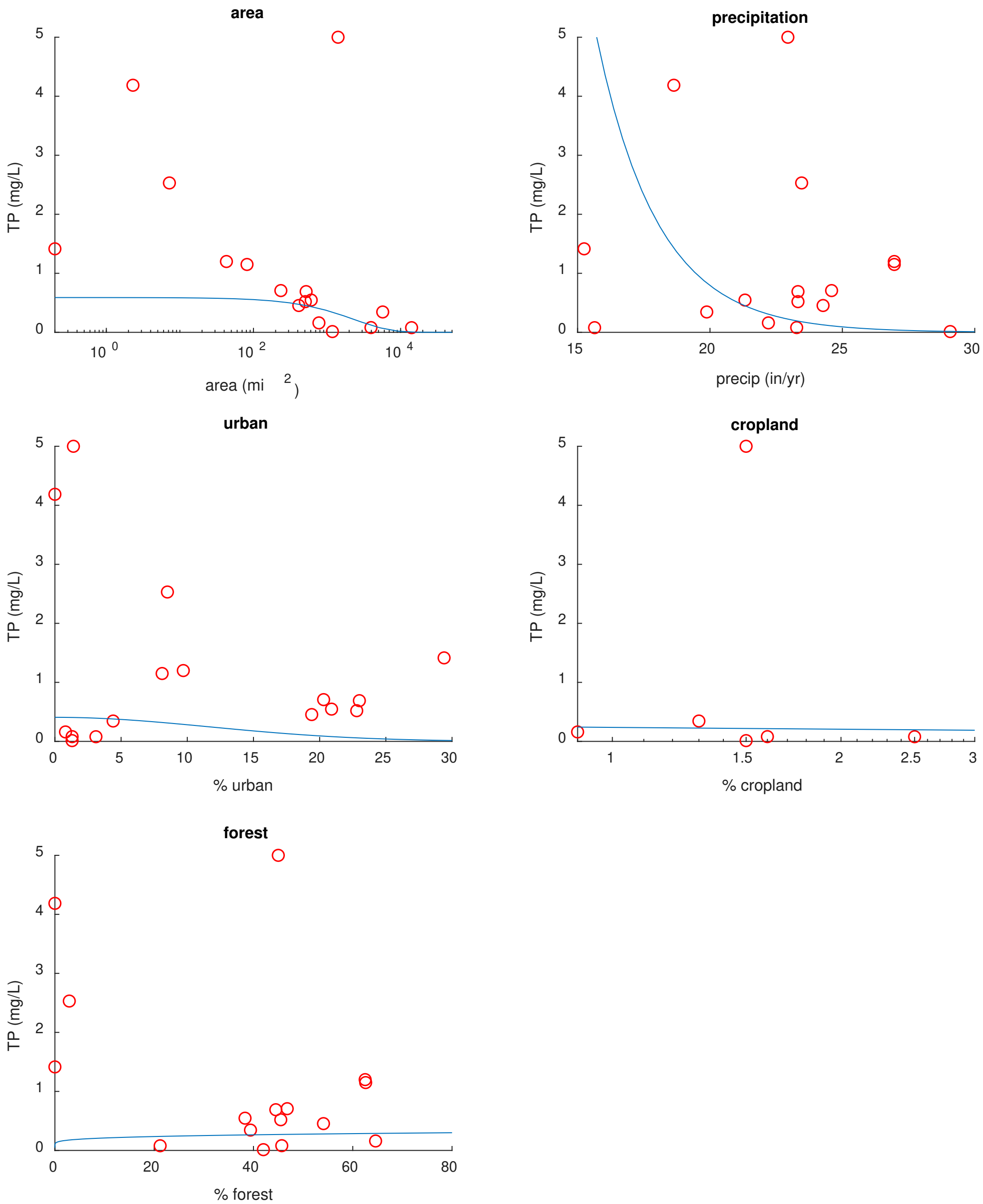
Division 1 TP



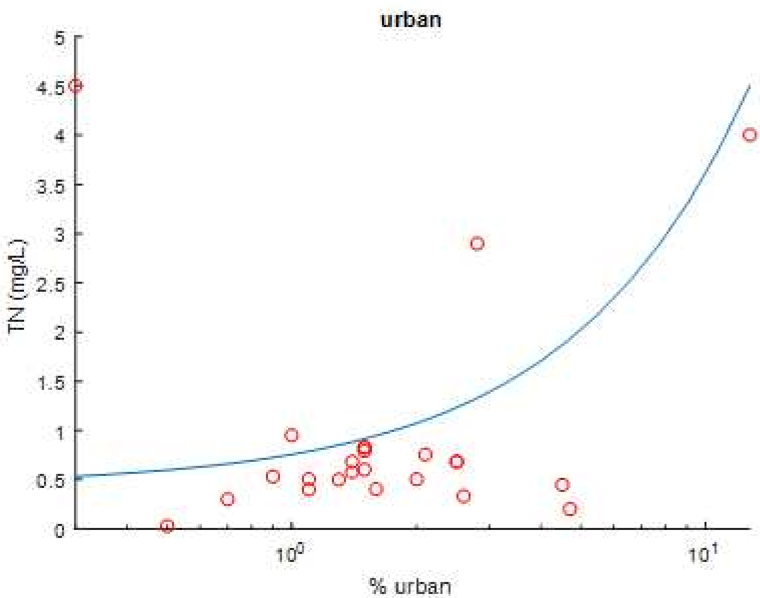
Division 2 TN

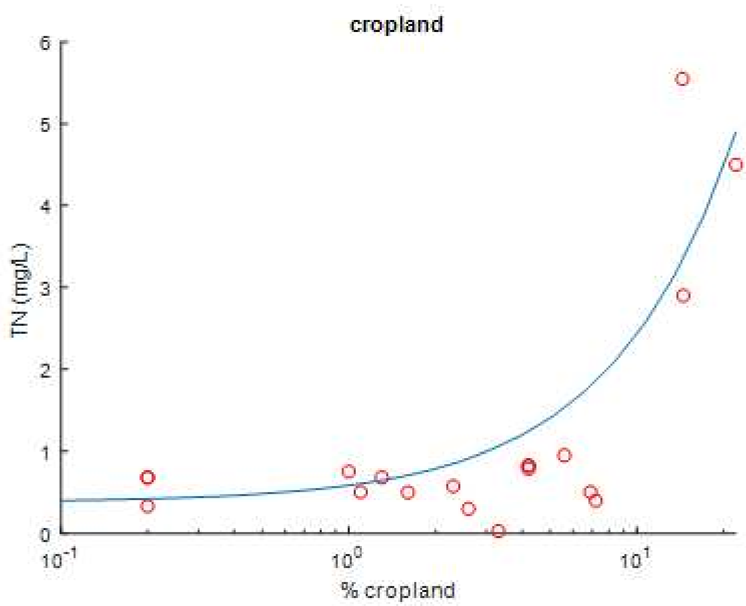


Division 2 TP

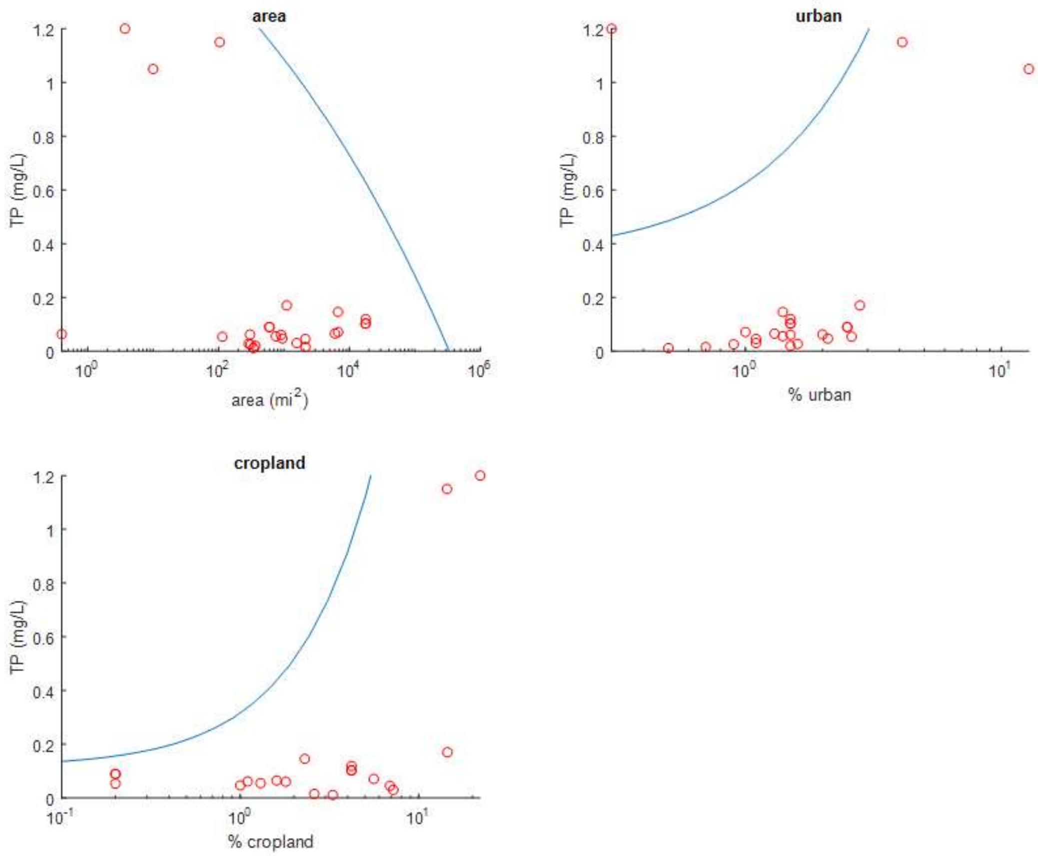


Western Slope TN





Western Slope TP



Appendix F

	Division	Model	Linear Model	R^2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DoF
Upstream	1	TN	$-0.43 - 3.32 * 10^{-4} Area^{0.89*1.2} - 5.38\, Slope^{1.2*1.2} + 0.06\, Pop + 0.59\, WWTP^{0.25} + 8.67 * 10^{-6}\, CAFO^{0.25}$	0.77	0.74	7E-9	131.7	0.1791	0.1252	-32.86	-23.36	0	36	30
	1	TP	$-3.26 - 3.0 * 10^{-4} (Area)^{0.89} - 4.90 (Slope)^{1.2} + 1.40 (WWTP)^{0.15} + 0.12 (Crop)^{0.6}$	0.71	0.68	3E-8	6.7	0.4315	0.5182	-24.09	-16.03	0	37	32
	2	TN	$5.29 + 3 * 10^{-4} Area^{0.89} - 14.22\, Slope^{1.2} - 0.83 Range^{0.5} + 1.67\, WWTP^{0.1}$	0.87	0.82	2E-4	7.1	0.5000	0.3809	-18.97	-15.43	0	15	10
	2	TP	$-3.07 - 1.4 * 10^{-3}\, Area^{0.89} + 1.9 WWTP^{0.1} + 3.00 * 10^{-5} CAFO$	0.79	0.72	2E-3	13.4	0.2552	0.2678	-12.01	-9.75	0	13	9
	West	TN	$0.97 + 1.2 * 10^{-5} Area^{0.89} - 0.75 Slope^{1.2} + 0.01\, Crop^2$	0.77	0.74	9E-9	3.6	0.2864	0.0465	-82.91	-77.18	1	31	27
	West	TP	$-3.37 + 2.33 * 10^{-9} (Area)^2 - 7.0 * 10^{-4} Precip^2 + 0.27 Urban + 0.07 Range^{0.75} - 0.64 Imperv$	0.71	0.61	1E-3	71.1	0.5000	0.5870	-28.74	-22.47	0	21	15

	Division	Model	Linear Model	R2	Adj. R^2	P	VIF	Lillie P-value	SW p-value	AIC	SCB	λ	N	DoF
Downstream	1	TN	$-3.23 - 3.8 * 10^{-4}\, Area^{0.89} - 5.0 * 10^{-3} Precip^{1.5} + 1.46\, Pop^{0.1} + 0.23 Crop^{0.25}$	0.67	0.63	4E-7	10.3	0.5000	0.4899	-39.70	-31.78	**0**	36	31
	1	TP	$1.48 - 3.9 * 10^{-4}\, Area^{0.89} + 1.89 Pop^{0.1} + 0.07 Crop^{0.5} - 6.41 Range^{0.1}$	0.62	0.58	5E-7	10	0.1746	0.7407	-28.75	-20.31	0	40	34
	2	TN	$43.97 - 1.9 * 10^{-3}\, Area^{0.89} - 15.26\, Precip^{0.25} + 0.03 Imperv^{2} - 7.33 WWTP\ Tech^{0.1}$	0.86	0.81	3E-4	6.5	0.0180	0.0114	13.47	17.01	1	15	10
	2	TP	$9.72 - 1.0 * 10^{-3} Area^{0.89} - 0.43\, Precip + 1.21 Forest^{0.1} - 1.95\, Crop^{0.1} - 3.7 * 10^{-3} Urban^{2}$	0.67	0.50	3E-2	17.4	0.0254	0.0211	7.76	12.40	0	16	10
	West	TN	$-033 + 0.32 Urban + 0.20 Crop$	0.79	0.78	1E-8	2	0.0552	0.446	-17.77	-13.99	1	26	22
	West	TP	$1.11 - 0.69 Area^{0.1} + 0.28 Urban + 0.20 Crop$	0.84	0.81	9E-9	3.2	0.0413	0.894	-21.65	-16.62	1	26	22