Evaluating the Impact of Onsite Wastewater Treatment Systems on Watershed Contamination, Choccolocco Creek Watershed, Alabama

by

Mallory Ann Jordan

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Master of Science in Geography

Auburn, Alabama
August 6, 2022

Keywords: Onsite wastewater treatment system (OWTS), multi-criteria decision analysis (MCDA), surface water, *E. coli*, geographic information system (GIS)

Approved by

Dr. Stephanie Rogers, Co-chair, Assistant Professor Department of Geosciences
Dr. Ann Ojeda, Co-chair, Assistant Professor Department of Geosciences
Dr. Ryan Thomson, Assistant Professor Department of Agricultural Economics and Rural Sociology

Abstract

There is increasing evidence that onsite wastewater treatment systems (OWTSs) can be a significant, and possibly underestimated, source of water contamination. OWTSs effectively treat wastewater when located in suitable environmental conditions and regularly maintained. However, these criteria are not always met, and the system can become susceptible to failure resulting in excess nutrients and harmful pathogens entering the surrounding environment. More information is needed to understand if and how OWTSs are correlated to surface water contamination and how the correlation changes with watershed scale. Thus, the objectives of this study are to 1) model the susceptibility of OWTS failure throughout the Choccolocco Creek watershed, and 2) determine if there is a relationship between *E. coli* concentrations in surface water with both modeled OWTS failure and land cover type and determine if watershed scale affects results in the Choccolocco Creek watershed. The Choccolocco Creek, a tributary to the Coosa River, is on the Alabama 303(d) List of Impaired Waterbodies for elevated concentrations of *E. coli,* a fecal indicator bacterium. The source(s) and relative contributions of *E. coli* are unclear, and typical of mixed-use watersheds, it is difficult to identify *E. coli* source(s) without advanced chemical analyses. However, geospatial methods can assist in identifying potential sources by exploring the geographic relationships between source areas and *E. coli* concentrations.

To address the first objective, GIS-based multi-criteria decision analyses (MCDA) were used to determine locations that have increased susceptibility to OWTS failure based on environmental variables (soil characteristics, proximity to surface water, and slope) and OWTS variables (age and density). With the model that included environmental and OWTS variables,

an area of 44.3 km^2 was identified as having high susceptibility to OWTS failure. Results indicate that OWTS age could be a driving factor of OWTS failure in the watershed.

For the second objective, results from OWTS failure models and distribution of land cover type were correlated to *E. coli* concentrations measured monthly from April to September 2021 at nine water sampling locations along the Choccolocco Creek. The water sampling locations were used to delineate nine sample point watersheds (SPWs) and 63 distance derived watersheds (DDWs) to summarize land cover distribution and OWTS failure models and assess how results vary with watershed scale. The analysis yielded several key results: 1) significant, positive correlations were found between *E. coli* concentrations and OWTS failure models for both SPWs and DDWs; 2) a significant, positive correlation between *E. coli* and OWTS count was found. Additionally, variation in the significance of correlations differed with watershed scale, demonstrating the importance of selecting an appropriate unit of analysis.

Our results suggest a relationship between OWTSs and elevated *E. coli* concentrations observed in Choccolocco Creek. Nonpoint source attribution challenges are not unique to the Choccolocco Creek watershed and methods outlined here could be applied to other watersheds to elucidate if OWTSs contribute to *E. coli* contamination of surface waters.

Acknowledgments

First, I want to thank my advisors, Dr. Stephanie Rogers and Dr. Ann Ojeda. Thank you both for your support and patience through this thesis research, challenging me with thoughtful questions and critiques, and teaching me the art of telling a (science) story. I would also like to thank Dr. Ryan Thomson, my committee member, for his input and guidance in this thesis. I am also very grateful for the support from my lab members in the GeoIDEA and Auburn Contaminants lab groups. Furthermore, I would like to thank my fellow graduate students, faculty, and staff in the Geosciences department for your help and support over the past two years. Finally, I am thankful for the love and support from my family.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ADEM	Alabama Department of Environmental Management
ADPH	Alabama Department of Public Health
cfu	Colony forming unit
DDW	Distance derived watershed
DEM	Digital elevation model
E. coli	*Escherichia coli*
EPA	(U.S.) Environmental Protection Agency
ESV	Environmental and system variables
EV	Environmental variables
GIS	Geographic information system
HUC	Hydrologic unit code
IDW	Inverse distance weighted
km	Kilometers
m	Meters
MAUP	Modifiable areal unit problem
MCDA	Multi-criteria decision analysis
mL	Milliliters
NHD	National Hydrography Dataset
NLCD	National Land Cover Dataset
OWTS	Onsite wastewater treatment system
R	Correlation coefficient
SPW	Sample point watershed

SSURGO	Soil Survey Geographic Database
TMDL	Total maximum daily loads
U.S.	United States
USGS	United States Geological Survey

1. Introduction

1.1. Onsite Wastewater Treatment System (OWTS) Overview

Onsite wastewater treatment systems (OWTSs) are an integral component of wastewater treatment infrastructure. There are two common ways to handle household wastewater in the United States (U.S.): connect to a municipal sewer system and send wastewater to a wastewater treatment plant or use an OWTS. An OWTS collects, transports, treats, and provides subsurface dispersal of sewage from establishments or dwellings (U.S. EPA 2002). In small-town, rural, and suburban areas, OWTSs provide an affordable means to treat wastewater (Lusk et al. 2017). As of 2019, approximately 23 percent of single-family and mobile homes use OWTSs (U.S. Census Bureau, American Housing Survey 2019), although prevalence varies by location. For example, OWTSs are utilized by only 10 percent of homes in California, 55 percent homes in Vermont, and more than 40 percent of homes in Alabama (U.S. Census Bureau 1990).

There are several types of OWTSs, but septic tank systems are the most conventional. Septic tank systems have two main components: the septic tank and a soil treatment unit or drain field. Figure 1 shows how wastewater is treated in a septic tank system. The wastewater first enters the septic tank by a pipe connected to the dwelling on site. Then, the wastewater is treated in the two components: (1) in the septic tank the primary treatment of solids occurs by anaerobic digestion and the wastewater leaves the septic tank as a septic tank effluent, and (2) in the drain field septic tank effluent percolates through the natural soils, contaminant concentrations are reduced, and the septic tank effluent moves to the groundwater (Lusk et al. 2017).

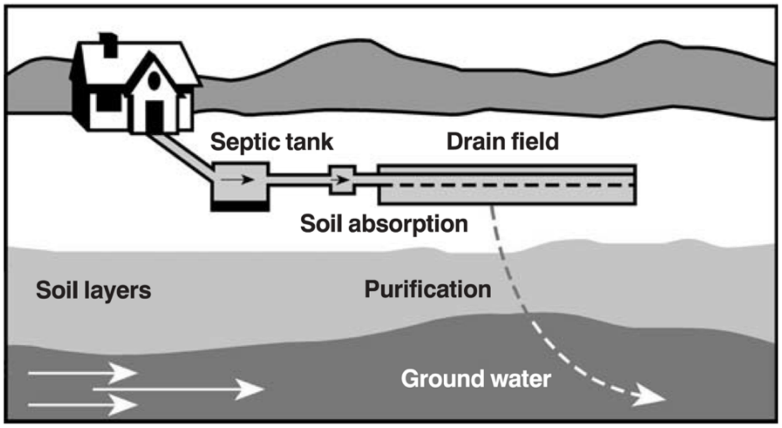

Figure 1. *A diagram outlining the path and processes of wastewater remediation in a septic tank system. Figure from U.S. EPA (2002).*

State public health departments are responsible for developing and enforcing OWTS codes. Historically, performance-based monitoring of OWTSs has not been required by regulatory agencies (U.S. EPA 2002). Regulatory agencies oversee the system siting, design, and construction for compliance with administrative codes and certification of site evaluators, designers, and other service providers. If all components meet the code requirements a system construction permit will be administered. At this point, the regulatory agency often relinquishes any further system oversight, and the owner becomes responsible for the OWTS. Regulatory agencies usually do not mandate or track maintenance and repairs but may become involved if a system failure is reported. As a result, there is minimal regulated performance oversight of the systems, potentially exacerbating the consequences of failed systems.

When OWTSs are planned, designed, installed, used, and maintained properly, pollutant attenuation rates are generally high, ranging from 70 percent to greater than 90 percent (Withers et al. 2013). OWTSs efficacy depends on multiple criteria, including soil type (Carroll et al. 2006), underlying lithology (Katz et al. 2010), septic tank and drain field design (Cooper et al. 2015; Cogger and Carlile 1984), and slope of the terrain (U.S. EPA 2002), among other factors. If these criteria are not suitable, OWTSs are susceptible to failure. OWTS failure is common across the U.S. with communities reporting failure rates between 10 and 70 percent, but the true extent is not known as no state has directly measured its failure rate (U.S. EPA 2000). Additionally, it is estimated that more than half of OWTSs are more than 30 years old and many outdated systems have known performance problems (U.S. EPA 2002). Although, OWTS failure is a pervasive problem, a comprehensive dataset of OWTS failure is not maintained by a single organization (U.S. EPA 2002).

Definitions of OWTS failure vary by regulatory agency with some defining failure as noncompliance with local regulations and others defining failure by inadequate system performance, with acceptable performance thresholds also varying by regulatory agency (U.S. EPA 2002). Examples of performance failure include hydraulic overloading of the system (wastewater pools on the ground surface or the system backs-up in the plumbing of the building) or pollutants, such as microbes (bacteria, viruses, or protozoa), nutrients (nitrogen and phosphorous), or trace organic compounds, entering the ground or surface waters (U.S. EPA 2002; Lusk et al. 2017). Therefore, failed OWTSs can serve as a pathway for pollutants, pathogens, and excess nutrients to enter the environment and adversely impact human and ecosystem health. Specifically, this leads to human exposure via well water, direct consumption, recreational contact with contaminated surface waters, sewage backups into homes, and pooling of septage on the ground. However, their overall contribution to water contamination is unknown (Withers et al. 2013), stemming largely from the difficulties associated with monitoring and modeling nonpoint source pollution (Patterson, Smith, and Bellamy 2013; Yuan, Sinshaw, and Forshay 2020). Without targeted data collection and monitoring of failure rates and effects, it is challenging to understand the extent and ramifications of OWTS failure.

1.2. Causes of OWTS Failure

There are multiple criteria that affect OWTS efficacy and susceptibility to failure. These criteria can be aggregated in two broad categories: system variables (e.g., system age) and environmental variables (e.g., soil characteristics). Criteria that affect OWTS failure cannot be considered independently; the system should be evaluated as a whole and relationships between criteria should be considered. The follow section provides a review of criteria that impact OWTS propensity of failure and description of how they impact systems.

1.2.1. System Variables

Age: There is a clear correlation between the age of an OWTS and the risk of failure. In general, the older an OWTS, the higher risk of failure. The average lifespan of an OWTS is 20 to 30 years; after 20 or more years of use, septic tanks and pipes can deteriorate and require repairs or replacement, and the soil surrounding drain fields may become inadequate for optimal system performance (U.S. EPA 2002). For example, spaces between soil particles can become filled with contaminants or the soil structure may change with time (U.S. EPA 2002). The useful lifespan on an OWTS will vary with system type, how well the system was maintained, and if it was used according to the intended usage.

Density: Water quality degradation attributed to high OWTS density was initially documented in an EPA report to congress, which detailed wastewater disposal practices and their effects on groundwater (U.S. EPA 1977). Yates (1985) connected groundwater contamination to high septic tank system density areas through an investigation of case studies and made three conclusions: septic tank systems are major contributors of wastewater to the subsurface; septic tank systems are the most frequently reported cause of groundwater contamination; and the most crucial factor influencing groundwater contamination by septic tank systems is the density. Updated guidelines for OWTS performance continue to cite high OWTS density as a contributing factor to ground or surface water contamination, but the threshold for "high density" is not well defined and is dependent on environmental factors, such as soil suitability (U.S. EPA 2002).

1.2.2. Environmental Variables

Soil characteristics: It is well understood that suitable soil characteristics are important for proper OWTS function, with soil properties that characterize water movement of utmost importance (U.S. EPA 2002). Soil variables of interest include the soil hydrologic group, which is an estimate of runoff potential, and soil drainage class, which describes the rate water is removed from the soil and is related to soil hydraulic conductivity. Soil hydraulic conductivity is a measurement of the ability of a soil to transmit water when subjected to a hydraulic gradient (NRCS 2004) and has consistently been cited as a critical factor of OWTS efficacy as it indicates the 'speed' at which septic tank system effluent moves through the soil (Collick et al. 2006; Beal et al. 2006; U.S. EPA 2002). Septic effluent should not move too quickly through the soil, which could lead to contaminants not being removed, or too slowly, where ponding may occur (U.S. EPA 2002).

Proximity to surface waters: The proximity of an OWTS to surface water can also contribute to performance failure (U.S. EPA 2002); if an OWTS is too close to a waterbody, there may not be enough travel distance for attenuation of pollutants from the septic system effluent. Setback distances are often established in OWTS regulations, however, they can be arbitrary and vary by regulatory agency (U.S. EPA 2002). The necessary setback distance for optimal system performance can differ based on other characteristics (e.g., soil characteristics).

Depth to groundwater: Similar to the proximity to surface water, the depth to the water table can contribute to OWTS failure if there is not enough travel distance to remove pollutants from the septic tank effluent; a minimum separation distance of 18 inches from the water table (U.S. EPA 2002). However, the effluent quality, hydraulic loading rate, soil characteristics, and wastewater effluent distribution methods can affect the soil depth necessary for to remove an acceptable concentration of pollutants (U.S. EPA 2002). Generally, 2-to-4-foot separation distances have allowed for adequate removal of fecal coliforms in septic tank effluent (Ayres Associates 1993).

Slope of terrain: The slope of terrain proximal to an OWTS should be relatively flat to allow for adequate effluent dispersion through the soil. The maximum slope of the terrain should be no greater than 10 to 20 percent (U.S. EPA 2002). However, the local topography should also be examined, as depressions, where ponding of water can occur, are not optimal of OWTSs.

1.3. GIS-Based Methods to Assess OWTS Susceptibility to Failure

Multiple, spatially dependent factors contribute to the efficacy of an OWTS, and a geographic information system (GIS) is a unique platform to study how these criteria interact spatially. In fact, several studies have already shown how a GIS-based, multi-criteria decision analysis (MCDA) approach, enable the creation of OWTS susceptibility to failure maps (Carroll et al. 2006; Oosting and Joy 2011; Capps et al. 2020). A MCDA is defined as a collection of techniques to analyze geographic phenomena where the analysis results are dependent upon the spatial relationships of the phenomena (Malczewski 1999). To conduct an MCDA, the user should: identify the problem or goal, determine input criteria, rank input criteria based on a standardized scale, weigh input criteria, aggregate input criteria, and validate the output (Eastman et al. 1995). Previous work in the field has shown that GIS-based MCDA have been useful in locating areas at risk of OWTS failure, helping to guide targeted management and oversight.

For example, Carroll et al. (2006) were one of the first to provide a novel framework to assess OWTS failure in a spatial context. They used three assessment categories: environmental, public health, and OWTS siting and design. Their model integrated phosphorus and nitrogen concentrations of surface and groundwaters, fecal bacteria and bacteria source tracking data, soil properties, lot size, setback distances, slope of the terrain, and floodplain boundaries. The risk classification and weighting schemes for each parameter were based on local OWTS regulations in Australia. Each criterion was represented by a geographic raster data layer and aggregated in a GIS. The resulting output identified areas 'at risk' and 'low risk' from OWTSs. This research has provided a conceptual framework detailing how to integrate environmental condition, OWTS data, and stakeholder knowledge to assess the risk of OWTS failure which has served as the foundation for subsequent studies.

Similarly, Oosting and Joy (2011) developed a GIS model to provide a risk rating of ground and surface water contamination from OWTSs in Ontario, Canada. The model calculated the cumulative risk of water pollution from OWTSs using nine risk parameters: soil type, slope, lot size, surface water proximity, floodplain, groundwater intrinsic susceptibility, recharge areas, and water supply proximity. The parameter weights in the risk model were determined through a survey of experts that included chief building officials, registered code agencies, public health inspectors, and researchers active in OWTS studies. The greatest contributors to the overall risk for the model were soil type, groundwater intrinsic susceptibility, and system age. This study also discussed the relationship between OWTS density and water quality but did not include it as a risk parameter like other studies (Carroll et al. 2006; Welhan and Moore 2012; Whittier and El-Kadi 2009). The model developed by Oosting and Joy (2011) successfully identified at-risk areas for

OWTS pollution and was useful in developing targeted management strategies for reinspection programs.

In addition to MCDAs, GIS-based statistical analyses have been used to investigate the relationship between OWTS failure and surface water quality. Flanagan et al. (2019) aimed to understand if and how OWTSs prevalence is spatially correlated to various environmental and surface water quality parameters. Flanagan et al. (2019) examined the spatial distribution of septic tank systems, coastal surface water contamination, and relevant environmental factors of coastal Florida watersheds. Spatially explicit explanatory data (e.g., land cover, population density, OWTS density, and soil drainage) were correlated to *in situ* nitrogen concentrations, Enterococci counts, and beach closures. Significant positive correlations ($\alpha = 0.05$) to Enterococci counts were found with percent agricultural cover, percent combined urban and agricultural cover, septic tank density, population density, and septic tank density in poorly drained soils. The authors also hypothesized that the combination of septic tanks in urban regions with high impervious cover, where there is elevated runoff, could also be a factor in surface water contamination leading to beach closures. This research provided a statistics-based framework to identify areas where septic systems may be causing environmental and human health issues.

1.4. *Study Area: Choccolocco Creek Watershed*

The study area for this research is the Choccolocco Creek watershed located in northeast Alabama (Figure 2A). This creek is on the Alabama 303(d) List of Impaired Waters, meaning the creek does not meet water quality standards set by the Alabama Department of Environmental Management (ADEM). Section 303(d) of the Clean Water Act requires that each state identify surface waters that do not currently support their designated use; Choccolocco Creek serves as a local recreational area and is designated as a fish and wildlife and public water use (ADEM 2022).

The Choccolocco Creek has been 303(d) listed due to high concentrations of polychlorinated biphenyls and mercury, since 1996 and 2010, respectively (ADEM 2020a). More recently, sections of the Choccolocco Creek have been added to the 303(d) list due to *Escherichia coli* (*E. coli)* contamination. *E. coli* is a fecal indicator bacterium and originates from wastewater collection system failure, OWTSs, agricultural runoff, urban runoff, and storm sewers. The average *E. coli* concentration in Choccolocco Creek during the summer of 2021 was 224 cfu per 100 mL; the EPA threshold for fresh, recreational water is 126 cfu per 100 mL (U.S. EPA 2012; Larson 2022). Sections of the Choccolocco Creek have continuously been 303(d) listed for *E. coli* contamination, with 33 miles listed in 2018, 44 miles listed in 2020, and two miles listed in 2022 (ADEM 2020b; 2022). The listed sources of *E. coli* are animal feeding operations, pasture grazing, collection system failure, and urban runoff and storm sewers (ADEM 2022). However, the relative contaminate loading from nonpoint sources, such as failed OWTSs and agriculture runoff, is unclear.

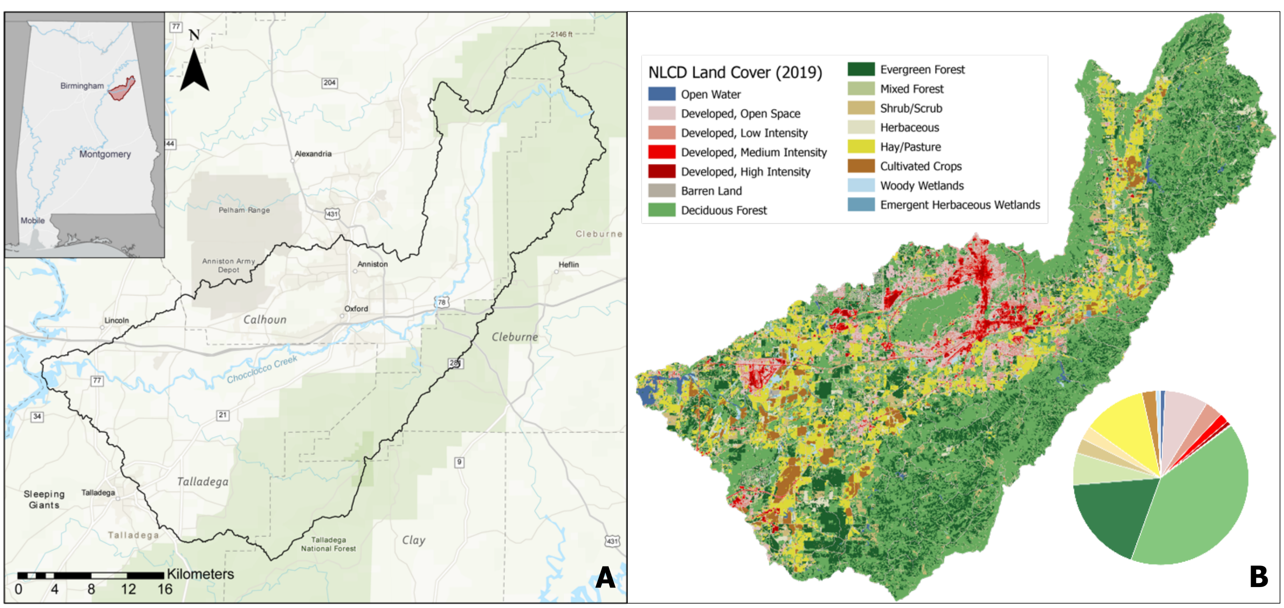

Figure 2. *The area of study, the Choccolocco Creek HUC-10 watershed boundary (A) and land cover (USGS 2019) within the Choccolocco Creek watershed (B).*

The study area is defined at the 10-digit hydrologic unit code (HUC) watershed and spans Calhoun, Clay, Cleburne, and Talladega counties. The Choccolocco Creek watershed is 1,323 km^2 (511 mi^2) in area and includes a mix of land cover types, making it difficult to identify *E. coli* source(s) in Choccolocco Creek watershed (Figure 2B). The central portion of the watershed is primarily urban and the location of the two prominent cities in the watershed: Oxford and Anniston. The north portion of the watershed is part of the Talladega National Forest and largely undeveloped. Additionally, there is crop and pasture agricultural land use throughout the watershed, accounting for approximately 14 percent of the watershed area. Choccolocco Creek, approximately 126.5 km in length (USGS 2021), is a major tributary to the Coosa River, which flows through Georgia and Alabama. The Choccolocco Creek flows southwest from headwaters in the Talladega National Forest to the confluence with Lake Logan Martin.

1.5. Objectives

There are multiple potential sources of *E. coli* in the Choccolocco Creek watershed that could contribute to water contamination. The varied land use and poor understanding of loading from nonpoint sources makes it unclear which source(s) are leading to the elevated levels of *E. coli*. Additionally, within the Choccolocco Creek watershed the propensity and spatial variability of OWTS failure and the relationship to water contamination in Choccolocco Creek are unknown. Furthermore, limited availability to OWTS data and a poor understanding of their spatial distribution makes it difficult to characterize the impact of failed OWTSs. Thus, the objective of this thesis is to improve our understanding of the contribution of OWTSs to *E. coli* contamination of surface waters by modeling OWTS susceptibility to failure in the Choccolocco Creek watershed. Based on the assumption that there are failed OWTSs within the Choccolocco Creek watershed, it was hypothesized that OWTSs are contributing to the high *E. coli* concentrations observed. The following research questions (RQs) will assess the hypothesis:

RQ 1. What is the spatial distribution of the susceptibility to OWTS failure in the Choccolocco Creek watershed?

RQ 2. What is the spatial correlation between OWTS variables, modeled susceptibility to OWTS failure, and land cover distribution with *E. coli* concentrations across the Choccolocco Creek watershed?

2. Methodology

2.1. Summary of Methods

First, the OWTSs were located within the Choccolocco Creek watershed using permit data from the Alabama Department of Public Health (ADPH). OWTSs were geocoded from table-based permits to geographic coordinates, represented as a point shapefile, for the watershed. Then, susceptibility of OWTSs failure, where failure is defined as the contamination of surface waters, was modeled using a MCDA (Figure 3); detailed methods on model development are provided in Section 2.2. The MCDA was completed in ArcGIS Pro 2.9. Two models were created to compare how environmental and OWTSs conditions affect model results: one model included only environmental variables (EV), and the second included the system and environmental variables (ESV). Next, MCDA results (that describe the susceptibility to OWTS failure), OWTS variables, and land cover percentages were correlated with *E. coli* concentrations in Choccolocco Creek, with specific methodology provided in Section 2.3. The steps taken for this analysis can be summarized in three general steps: delineation of sample point watersheds (SPWs), creation of distance derived watersheds (DDWs) based on surface water flow distance, and correlation of modeled OWTS failure, OWTS variables, and distribution of land cover types to *E. coli* concentrations.

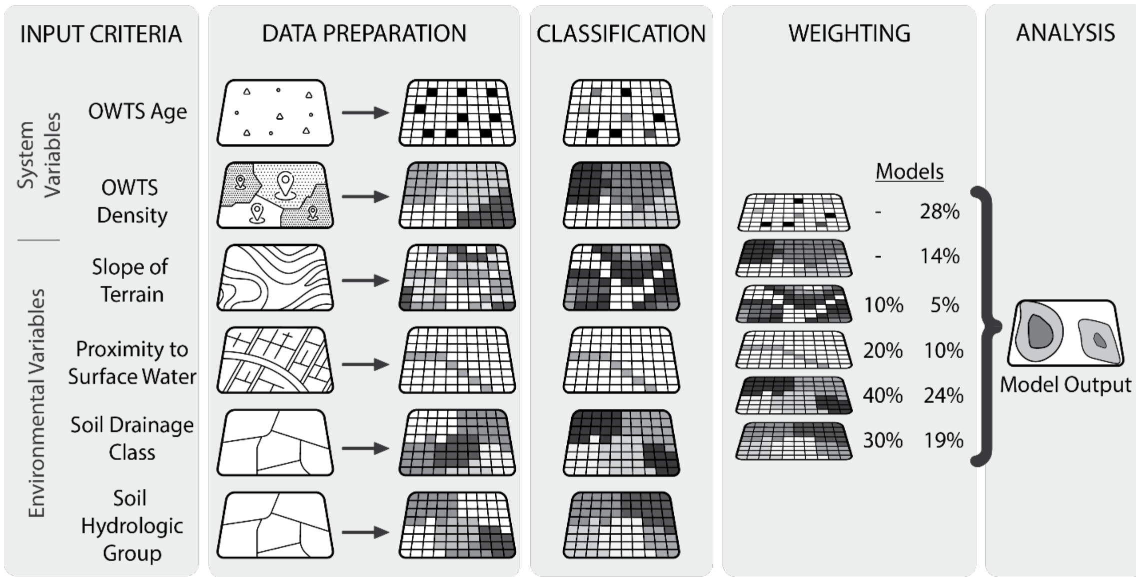

***Figure 3**. The framework for the GIS-based multi-criteria decision analysis (MCDA) to assess the susceptibility of OWTS pollution. Steps include selection of input criteria, data preparation, classification of data on common scale, development of a weighting scheme, and aggregation of data. Figure concept from Oosting and Joy (2011).*

2.2. *MCDA: Susceptibility to OWTS Failure (RQ 1)*

2.2.1. Data Preparation

Input criteria for the MCDA include OWTS age, OWTS density, slope of the terrain, proximity to surface water, soil drainage class, and soil hydrologic group. These criteria were chosen because they are shown to have an effect on OWTS efficacy and have been included in similar studies (Carroll et al. 2006; Oosting and Joy 2011; Capps et al. 2020). Data sources are listed in Table 1. All datasets were transformed to raster format with a cell size of 1/3 arc-second (approximately 10 m), as this was the coarsest resolution of all input data. Any location that did not have data for all criteria was excluded from the model results. A description of how all data were transformed to raster format is given in this section.

Table 1. Input data sources to assess onsite wastewater treatment system (OWTS) failure.

Data Source	Derived Input Criteria	Original Format	Resolution
National Hydrography Dataset (NHD)	Proximity to surface water	Vector	1:24,000*
U.S. Geological Survey Digital Elevation Model (USGS DEM)	Slope of terrain	Raster	1/3 arc-second (approximately 10 m)
Soil Survey Geographic database (SSURGO)	Drainage class and hydrologic group	Vector	1:12,000 to 1:63,360*
Alabama Department of Public Health (ADPH)	Septic system age and density	Tabular	

*The scale from which the data were digitized.

OWTS age and density: OWTS permit data were provided by the ADPH, in a table-based format. Permit data were provided for the four counties that intersect the Choccolocco Creek watershed: Calhoun, Clay, Cleburne, and Talladega. Permits were categorized into two types: new approvals and repair and included all systems permitted from January 1, 2000, to August 11, 2020. OWTSs were geocoded using the ArcGIS World Geocoder, a locator developed by Esri that converts a text address into the representative geographic coordinates (Esri 2022a). For quality control purposes, all geocoded addresses were required to have a minimum match score of 80. Geocoded addresses with lower scores were rematched, meaning the geocoded point was matched to a different address to improve the location accuracy. Rematched addresses were manually reviewed to check for incomplete or misspelled addresses, and the address was either manually corrected or a different address candidate was selected. Additionally, there were some addresses that contained multiple permits and in this case it was assumed that an address has only one permitted OWTS. If there was a new approval and repair permit(s) for the same address, the repair permit data was appended to the new approval point, so there was one point representing the system. Then, the OWTS age was calculated by the difference in the permit issuance date and the date of permit data collection (August 11, 2020). If there were multiple new approval permits for the same address, the most recent approval date was used to calculate the system age. If there was a repair permit for an address but no new approval permit, it was assumed the system was permitted prior to the start of recorded permit data, and the system age was calculated from January 1, 2000. OWTS age point data were transformed to a raster format using the *Inverse Distance Weighted (IDW)* tool. The *IDW* tool calculates cell values using a linearly weighted combination of sample points (i.e., OWTSs), and the weight is a function of inverse distance (Esri 2022b). Thus, this tool calculates an interpolated surface of a dependent variables (i.e., system age); this method assumes that the influence of the variable mapped decreases with distance. A search radius distance of 152.4 m was used, based on the largest required setback distance for septage set by the ADPH (ADPH 2017). The OWTS density was calculated with the *Point Density* tool. This calculates the density of a point feature within a specified neighborhood around each output raster cell; the output raster cells represent a point count per unit area (Esri 2022c). A neighborhood radius of 564.19 m was used to calculate the number of OWTSs per one km^2.

<u>Slope of terrain:</u> The slope of the terrain, in percent, was calculated using the *Slope* tool (Esri 2022d) with Digital Elevation Models (DEMs) as the input (Carroll et al. 2006; Oosting and Joy 2011). DEMs were sourced from the USGS (USGS 2020a; 2020b).

<u>Proximity to surface waters:</u> The proximity to surface waters was calculated using the distance from National Hydrography Dataset (NHD) Flowline and Waterbody layers (USGS 2021). The distance was calculated using the *Euclidean Distance* tool, and the two layers were combined with *Mosaic to New Raster* tool with the mosaic operator as minimum.

<u>Soil characteristics:</u> The soil characteristics were accounted for using the drainage class and hydrologic group from the Soil Survey Geographic Database (SSURGO), a description of the soil attributes are given by Soil Survey Staff (1993) and Soil Survey Staff (2006), respectively. These two variables were selected as data were available across the entire watershed and they have been

used in similar analysis (Oosting and Joy 2011). SSURGO is in vector format; therefore, the fields Hydrologic Group Dominant Condition (hydgrpdcd) and Drainage Class Dominant Condition (drdclassdcdm) were transformed from categorical data to unique values with the *Reclassify Field* tool and then, converted to a raster format with the *Feature to Raster* tool.

2.2.2. Data Classification and Weighting

The criteria presented above (Section 2.2.1) were then classified and weighted according to the schemes provided in Table 2. Data were classified with the *Reclassification* tool on a standardized scale of one to five, where one represents a low contribution to system failure and five represents a high contribution to system failure. Data were classified so each input criteria was represented by a common scale. Classification schemes (Table 2) were based on previous studies, EPA guidelines, and Alabama OWTS codes. Older OWTSs were given higher values, as there is an increased risk of failure (Oosting and Joy 2011). Higher densities of OWTSs were given higher values (Yates 1985). Higher slopes in the terrain were assigned higher values (Oosting and Joy 2011). The proximity to surface water classification breaks were based on Alabama OWTS regulations; systems within the required setback distance were classified as a high risk of failure and systems more than two times the required setback distance classified as low risk, similar to Oosting and Joy (2011). Soil characteristics (i.e., drainage class and hydrologic group) were classified based on Oosting and Joy (2011) and EPA guidelines (U.S. EPA 2002). Criteria weights were then assigned using the rank sum procedure and where higher weights were assigned to criteria that were more important to the susceptibility of system failure, using the following formula from Nyerges and Jankowski (2009):

$$w_j = \frac{n - r_j + 1}{\sum_{k=1}^{n}(n - r_k + 1)} \qquad \textit{Equation } 1$$

where w_j is the normalized weight (ranking in value from 0 to 1) for the criterion j, n is the number of criteria under consideration, and r_j is the rank position of the criterion. Ordinal ranks for all criteria were based on Oosting and Joy (2011). Classified data were then aggregated using the Weighted Sum tool, and the calculated criteria weights using the following formula:

$$Susceptibility = \sum (Susceptibility\ rating) \times (weight) \qquad Equation\ 2$$

where susceptibility is the susceptibility to OWTS failure, susceptibility rating is the classified value (from 1 to 5), and the weight calculated by Equation 1. High values in the model output represent higher susceptibility to system failure. The model outputs were classified into four classes based on the susceptibility to OWTS failure: minimal, low, moderate, and high. Category breaks were based on those developed by Oosting and Joy (2011).

Table 2. Input criteria, classification scheme, and weight for multi-criteria decision analyses to model onsite wastewater treatment system (OWTS) failure.

	Criteria	**Classification Value**					**Weight (%)**	
		Low		Moderate		High		
		1	2	3	4	5	EV	ESV
System Variables	OWTS Age (years)	< 5	5 – 10	10 – 15	15 – 20	> 20	-	28
	OWTS Density (OWTSs/km^2)	< 1	1 – 3.8	3.8 – 15.4	15.4 – 38.6	> 38.6	-	14
Environmental Variables	Slope of Terrain (%)	< 6	6 – 12	12 – 20	20 – 25	> 25	10	5
	Proximity to Surface Water (m)	> 30.48			15.24 – 30.48	< 15.24	20	10
	Drainage Class	Well drained	Moderately well drained	Somewhat poorly, somewhat excessively drained	Poorly drained	Very poorly drained, excessively drained	40	24
	Hydrologic Group	B and B/D		C and C/D		A and D	30	19

EV = Environmental variables
ESV = Environmental and system variables

2.3. Statistical Analysis: Correlations to E. coli Concentration (RQ 2)

2.3.1. E. coli Characteristics of the Choccolocco Creek

Water quality samples of Choccolocco Creek, which included *E. coli* concentration measurements, were taken monthly at nine locations over a six-month period, from April 2021 to September 2021, in the Choccolocco Creek watershed by Larson (2022) (Figure 4A). A total of sixteen *E. coli* samples were taken, as samples were taken in triplicate for all months except August where only one sample was taken (Table A.1). *E. coli* concentrations were averaged for each site for correlations. The average *E. coli* concentration for all sample sites was 224 cfu per 100 mL; site five had the highest concentration of 378 cfu per 100 mL and the lowest concentration at site two of 92 cfu per 100 mL (Figure 4B) (Table A.1). *E. coli* concentrations increased downstream from site nine to sites eight and seven. Then there was a decrease in concentration at site six. There was another increase in *E. coli* concentration at site five, downstream of Oxford and Anniston, with a general decrease in the average concentration to site one.

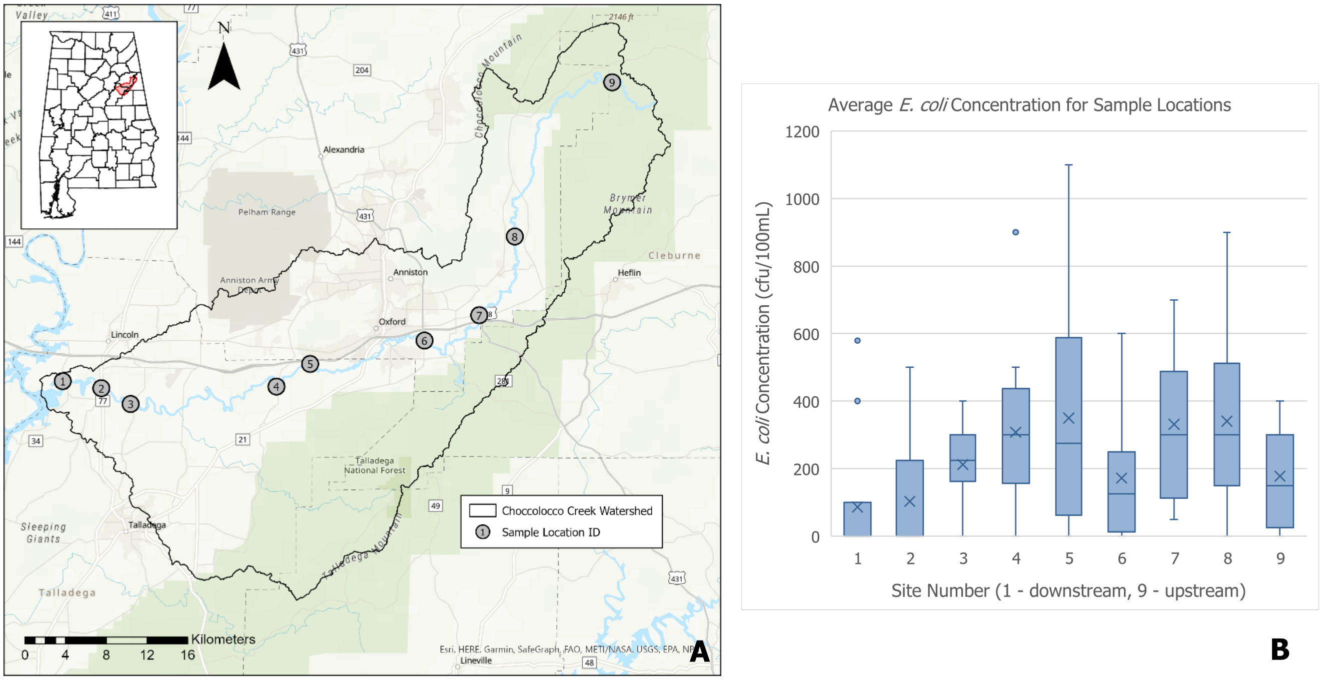

Figure 4. *Water sampling locations of the Choccolocco Creek, numbered 1, downstream, through 9, upstream (A). The average E. coli concentration (cfu/100 mL) for all sampling events is provided in the box plot (B). The lower and upper bounds of the boxes represent the values at the first and third quartiles, respectively. The line through the middle of the box denotes the median value and the 'x' denotes the mean. Outliers are represented as points.*

2.3.2. Delineation of Sample Point Watersheds (SPWs)

Watersheds were delineated using the water sampling locations (Figure 4A) and USGS DEMs (USGS 2020a; 2020b). For this study, it was important to understand which parts of the landscape were affecting potential *E. coli* concentrations at the sample points, thus, watersheds were created to obtain a representative SPW. The ArcGIS Pro Hydrology toolset was used to delineate watersheds. The imperfections from the DEM were removed, the flow direction was calculated, and flow accumulation was calculated, and watersheds were delineated from the nearest pour point within 30 m of the sample location (Esri 2016). A 30 m search distance was used because it was the smallest distance that produced watersheds that were large enough for analysis (Lindsay, Rothwell, and Davies 2008). Nine SPWs were delineated, one for each sample location.

2.3.3. Delineation of Distance Derived Watersheds (DDWs)

Using the sample point watersheds, DDWs were created to represent the upstream area that is within a given surface water flow distance of the pour point; distance intervals of 250 m, 500 m, 750 m, 1,000 m, 5,000 m, and 10,000 m were calculated with *Flow Length* tool and used to determine the effect of using watersheds of different scales as a unit of analysis. This was completed using ModelBuilder in ArcGIS Pro. First, the flow direction raster was clipped to each SPW. Then, the model iterated through each flow direction raster to calculate the surface water flow length from each sample location. Next, all flow length rasters were reclassified, using the *Reclassify* tool, based on the target flow length buffer distance; raster values within the desired distance were reclassified as 1, and all flow lengths greater than the desired distance were reclassified as 0. This was an iterative process for all flow distances. Then, all reclassified rasters were converted to polygons with the *Raster to Polygon* tool. Thus, the final model outputs were polygon DDWs delineating the upstream catchment area for all desired flow distances (250 m, 500 m, 750 m, 1,000 m, 5,000 m, and 10,000 m) (Figure 5).

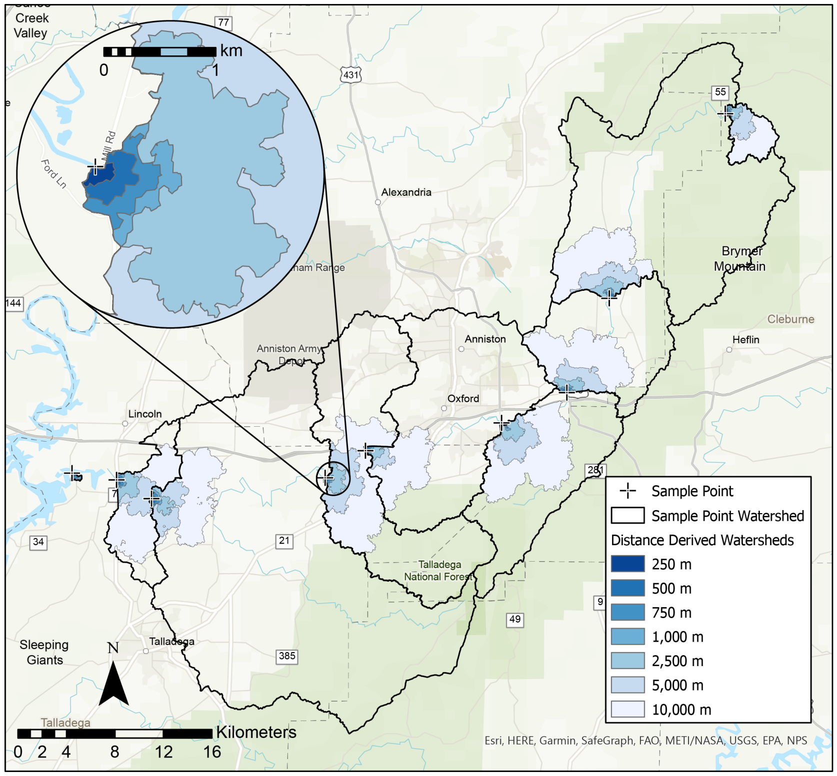

Figure 5. *The sample point watersheds (SPWs), denoted by the black line, and distance derived watersheds (DDWs), shaded in blue, for each sample point delineated from water-sampling locations along the Choccolocco Creek.*

2.3.4. Zonal Statistics

Zonal statistics, a raster-based calculation that provides statistics for a value raster within certain zones, were calculated using the *Zonal Statistics as Table* tool. In this case, zones were defined by the SPWs and DDWs and the pixel values summarized were the EV-MCDA and ESV-MCDA statistics (min, max, range, mean, standard deviation, median, and 90th percentile). Additionally, the count and density of OWTSs and land cover percentages were summarized for SPWs and DDWs. Land cover data were sourced from the 2019 national land cover dataset (NLCD) (USGS 2019). OWTSs and land cover data were summarized for all SPWs and DDWs as they represent possible *E. coli* sources.

2.3.5. *Correlations*

Correlation statistics were used to evaluate the relationship between the ES- and ESV-MCDA statistics (i.e., min, max, range, mean, standard deviation, median, and 90^{th} percentile), OWTS variables, distribution of land cover types with *E. coli* concentrations for all SPWs and DDWs. First all variables were checked for normality, using the Shapiro-Wilk test (Ghasemi and Zahediasl 2012). A Pearson correlation was used if data were normally distributed, and a Spearman correlation was used if data were not normally distributed (Schober, Boer, and Schwarte 2018). Analysis was completed using Prism (ver. 9.3.1).

3. Results

3.1. Susceptibility to OWTS Failure (RQ 1)

3.1.1. *OWTSs*

There were 3,844 OWTSs permits issued by the ADPH from January 1, 2000, to August 11, 2020, in the Choccolocco Creek watershed (Table 3). Of the permits issued, 2,530 were new approvals and 1,314 were for repairs. There were 3,717 OWTSs identified, and the average system age was 16.8 ± 3.7 years, where OWTS age was calculated from the date of permit data collection (August 11, 2020). The OWTS density for the watershed was 2.81 OWTSs per km^2 with class breaks based on categories established by Yates (1985) (Figure 6). OWTSs are prevalent throughout the developed portion of the Choccolocco Creek watershed, which encompasses the northwestern portion of the watershed; are sparsely located in the Talladega National Forest, which spans across the eastern boundary; and are not present in the Anniston Army Depot, located in the northwest portion (Figure 6).

Table 3. Onsite wastewater treatment system (OWTS) permit data summary for the Choccolocco Creek watershed.

Permit Type	Permit Count	Number of OWTSs Located
New Approval	2,530	2,499
Repair	1,314	1,218
Total	*3,844*	*3,717*

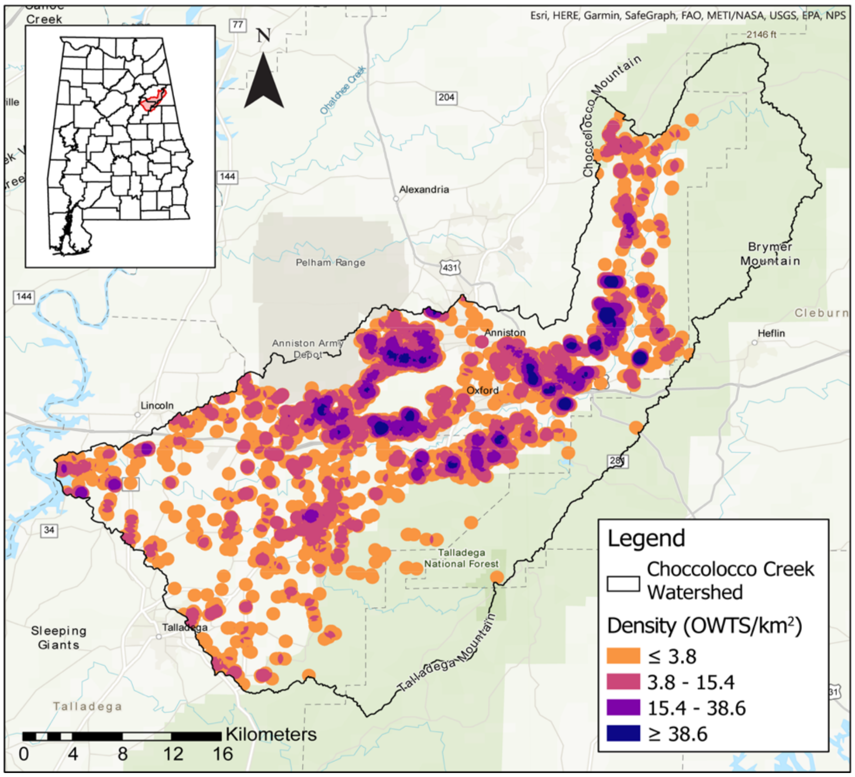

***Figure 6**. Density of onsite wastewater treatment systems (OWTSs) in the Choccolocco Creek watershed.*

3.1.2. MCDA: Susceptibility to OWTS Failure

All criteria were classified on a common scale, so that each criterion layer described the susceptibility to OWTS failure. Criteria were classified on a scale of 1 (beige) to 5 (dark red), with greater values denoting higher susceptibility to OWTS failure (Figure 7). Table 4 provides the minimum, maximum, mean, and standard deviation of classified criteria data. Classified values ranged from 1 to 5 for all criteria except soil drainage class that ranged from 1 to 4, as there were not very poorly drained or excessively drained soils. Slope of the terrain had the highest average classification value and OWTS age had the lowest average classification value for the watershed.

Then, aggregated data in the model outputs were classified into four classes based on the susceptibility to OWTS failure: minimal, low, moderate, and high. Minimal risk to susceptibility includes all cell values less than or equal to the mean raster value. Low, moderate, and high susceptibility to OWTS failure is defined by subsequent standard deviations above the mean. Therefore, low susceptibility to OWTS failure represents all cells between the mean and the first standard deviation above the mean, moderate susceptibility to OWTS failure represents all cells between the first standard deviation and second standard deviation above the mean, and high susceptibility to OWTS failure represents all cells between greater than the second deviation above the mean.

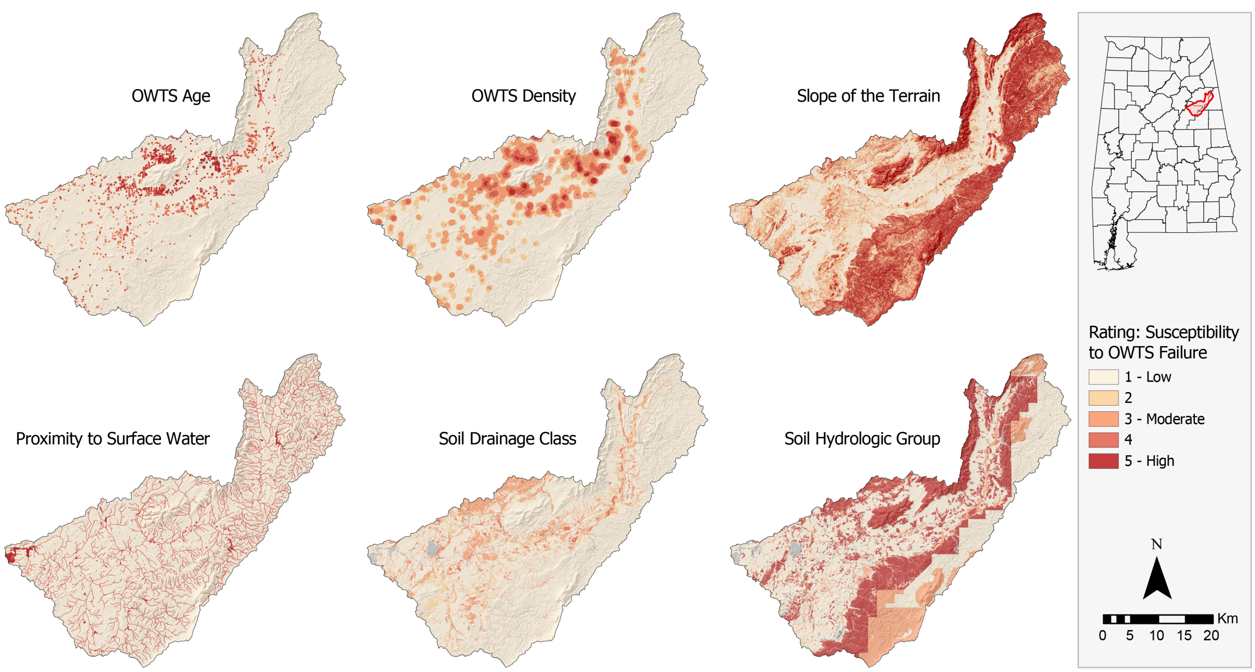

***Figure** 7. Classified susceptibility to OWTS failure for all input criteria and are classified on a scale of 1 to 5, with greater values denoting greater susceptibility to failure.*

Table 4. Summary statistics for classified input criteria.

Criterion	Min.	Max.	Mean	Standard Deviation
OWTS Age	1	5	1.26	0.87
OWTS Density	1	5	1.54	0.94
Slope of the Terrain	1	5	2.76	1.59
Proximity to Surface Water	1	5	1.41	1.14
Soil Drainage Class	1	4	1.33	0.74
Soil Hydrologic Group	1	5	2.54	1.80

OWTS = Onsite wastewater treatment system

The result of the susceptibility to OWTS failure EV-MCDA (slope of the terrain, proximity to surface water, soil drainage class, and soil hydrologic group) is provided in Figure 8. Darker, purple areas represent areas of higher susceptibility to OWTS failure. The model output ranged from 1 to 4.60 with an average value of 1.85 ± 0.71 (Table 5). Within the watershed, 29.8 km^2 were classified as high risk (> 3.27) to OWTS failure, 2.3% of the model area, and 669.7 km^2 were classified as minimal risk (Table 5). Areas of increased susceptibility to OWTS failure were observed in the western and southern sections of the watershed, and areas proximal to Choccolocco creek in the northern area of the watershed. The susceptibility to failure is low in the Talladega National Forest.

The result of the susceptibility to OWTS failure ESV-MCDA (slope of the terrain, proximity to surface water, soil drainage class, soil hydrologic group, OWTS age, and density) is provided in Figure 9. Darker, blue areas represent areas of higher susceptibility to OWTS failure. The model output ranged from 1 to 4.42 with an average value of 1.65 ± 0.51 (Table 5). There were 44.3 km^2 were classified as high risk (> 3.27) to OWTS failure, 3.4% of the model area, and 716.6 km^2 were classified as minimal risk (Table 5). The susceptibility to failure is lowest throughout the Talladega National Forest, located along the eastern boundary of the watershed, where there is a low density of OWTSs. The susceptibility to failure is highest outside of Oxford and Anniston where there is a high density of OWTSs.

Table 5. Summary statistics for susceptibility to onsite wastewater treatment system (OWTS) models.

Model	**Min.**	**Max.**	**Mean**	**Standard Deviation**	**Susceptibility to OWTS Failure Area (km^2)**			
					Minimal	***Low***	***Moderate***	***High***
Environmental	1	4.60	1.85	0.71	669.7 (51.7%)	322.3 (24.9%)	273.8 (21.1%)	29.8 (2.3%)
Environmental + System	1	4.42	1.65	0.51	716.6 (55.3%)	385.8 (29.8%)	148.9 (11.5%)	44.3 (3.4%)

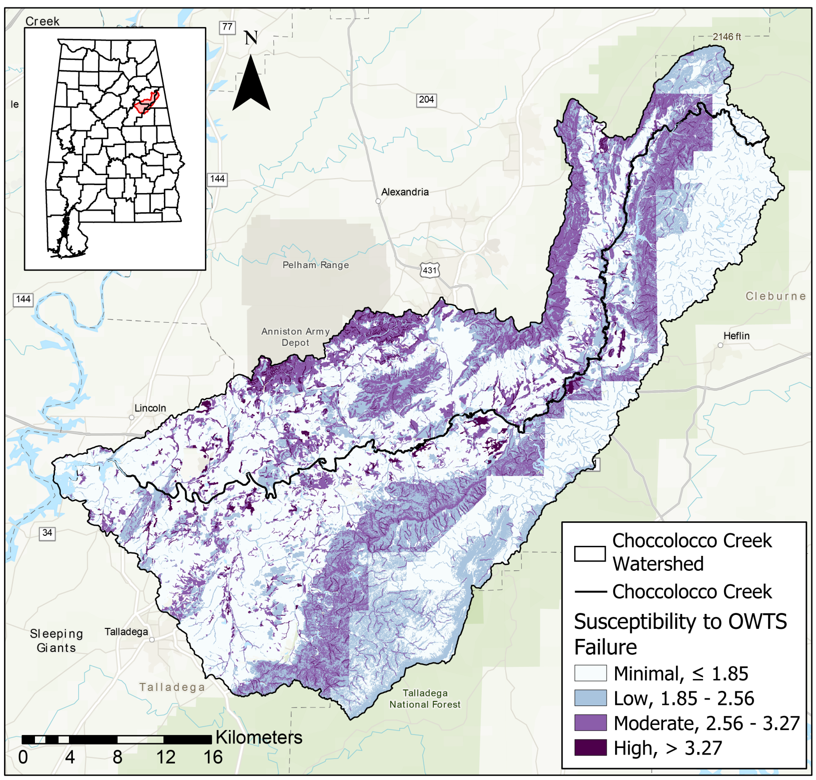

Figure 8. *The susceptibility to OWTS failure in the Choccolocco Creek watershed based on environmental OWTS variables. The Choccolocco creek is shown by the black line. Darker, purple areas represent areas of higher susceptibility to OWTS failure.*

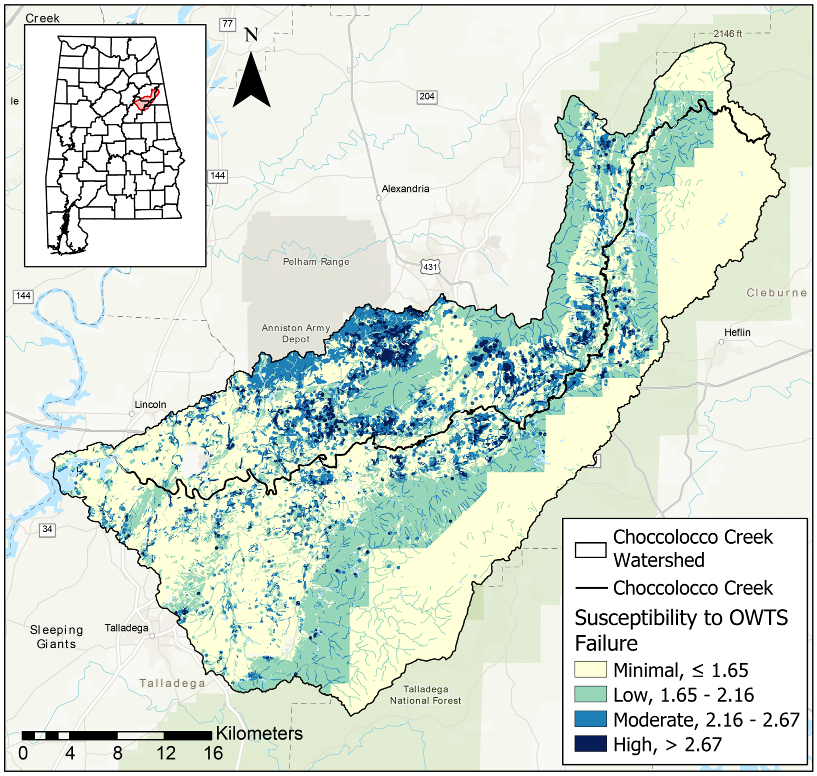

Figure 9*. The susceptibility to OWTS failure in the Choccolocco Creek watershed based on environmental and OWTS variables. The Choccolocco creek is shown by the black line. Darker, blue areas represent areas of higher susceptibility to OWTS failure.*

3.2. Statistical Analysis: Correlations to E. coli Concentrations (RQ 2)

The correlations of *E. coli* concentrations to MCDA values, OWTS variables, and land cover percentages for SPWs and DDWs are provided in Table A.2. The correlation coefficient (R) and R-squared values for significant relationships ($p < 0.05$) are provided in Table 6. There were 12 significant relationships between *E. coli* concentrations and OWTSs, MCDA summary statistics at various buffer intervals, and land cover percentages in the watershed. There were five significant correlations found using SPWs and seven significant correlations using DDWs, with significant correlations found at the 500 m, 750 m, and 1,000 m DDWs. There were significant positive correlations between the EV-MCDA mean, median, and 90th percentile and ESV-MCDA mean and median. There was a significant positive correlation between OWTS count and *E. coli* concentration, and between wetland land cover and *E. coli* concentration.

Table 6. Correlation coefficients (R), R-squared, and p-values for significant correlations ($p < 0.05$) to average *E. coli* concentrations (n = 9) in the Choccolocco Creek.

Variable	Shapiro-Wilks Test		Correlations			
	p-value	*Normal? ($\alpha < 0.05$)*	*Type*	*R*	*R-squared*	*P-value*
SPW ESV-MCDA mean	0.3603	Yes	Pearson	0.6732	0.4532	0.0469
SPW ESV-MCDA median	0.2800	Yes	Pearson	0.8729	0.7620	0.0021
SPW EV-MCDA mean	0.0880	Yes	Pearson	0.7109	0.5054	0.0318
SPW EV-MCDA median	0.4372	Yes	Pearson	0.8444	0.7131	0.0042
500 m DDW EV-MCDA 90th percentile	0.0547	Yes	Pearson	0.8064	0.6503	0.0086
750 m DDW EV-MCDA 90th percentile	0.1597	Yes	Pearson	0.8241	0.6792	0.0063
1000 m DDW EV-MCDA mean	0.6033	Yes	Pearson	0.7221	0.5214	0.0280
1000 m DDW EV-MCDA 90th percentile	0.1915	Yes	Pearson	0.8214	0.6748	0.0066
SPW OWTS count	0.0859	Yes	Pearson	0.7630	0.5821	0.0168
500 m DDW open water	0.0266	No	Spearman	-0.7120	0.5069	0.0402
500 m DDW woody wetlands	0.0541	Yes	Pearson	0.7076	0.5007	0.0330
1000 m DDW emergent herbaceous wetlands	<0.0001	No	Spearman	0.7303	0.5333	0.0278

SPW = Sample point watershed
DDW = Distance derived watershed
ESV = Environmental and system variables
EV = Environmental variables
MCDA = multi-criteria decision analysis
OWTS = Onsite wastewater treatments systems

4. Discussion

4.1. MCDA: Susceptibility to OWTS Failure (RQ 1)

4.1.1. Model Outputs

The first objective of this thesis was to model the susceptibility to OWTS failure in the Choccolocco Creek watershed using MCDA. The EV-MCDA could provide guidance on areas that may not be suitable for OWTSs or require advanced treatment systems. Additionally, OWTS data are difficult to obtain, and the EV-MCDA model provides a basis for locating areas where system failure is more probable without OWTS data. The ESV-MCDA could provide insight as to where OWTS may fail or have failed, leading to contamination of surface waters. Besides model outputs, which provide generalized information on areas susceptible to OWTS failure, the classified input criteria (Figure 7) render how individual criteria are suited for OWTSs across the watershed. These data provide more explicit information as to why a given location may be more susceptible to OWTS failure. For example, locations with poorly suited soil characteristics may require an advanced treatment system.

Comparison of model outputs provides information on how the inclusion of different variables affect the model results and allows for discerning driving criteria for OWTS failure. As expected, there is a shift in areas of high susceptibility to OWTS failure between the models. In the EV-MCDA, the areas of high susceptibility to OWTS failure were scattered in the eastern portion of the watershed, generally proximal to Choccolocco Creek. However, for the ESV-MCDA areas of high susceptibility to OWTS failure are mostly located in the central portion of the watershed and coincide with areas of higher OWTS density. The ESV-MCDA also has more area in the high susceptibility category (29.8 km^2 for the EV vs 44.3 km^2 for ESV) (Table 5). The shift in high susceptibility locations and increase in area with the inclusion of system variables suggests that system variables may be the greatest contributors to system failure in the watershed. Moreover, system variables account for 42 percent of the ESV-MCDA, meaning system variables have less weight than environmental variables. Furthermore, the majority of OWTSs had a classified susceptibility rating of 4 or 5 for system age (Figure 7), and the average age was 16.7 years old. An OWTS lifespan is between 20 and 30 years, so many of the OWTSs in watershed may be nearing the end of their intended lifespan (U.S. EPA 2002). This suggests that outdated OWTS could be a driving force of OWTS failure in the watershed. Similarly, Capps et al. (2020) found that only 8 percent of the registered OWTS presented potential environmental risk due to stream proximity, soil type, and slope; however, almost 70 percent of the OWTS presented potential environmental risk due to their age. This further highlights the importance of system variables – particularly age – on potential OWTS failure, and thus the importance of having access to robust OWTS data.

4.1.2. Model Limitations

The input criteria data preparation methods, classification schemes, and criteria weights vary in similar OWTS risk analyses, but in all analyses, it is apparent that a holistic view is needed. A discussion of uncertainties and limitations of the models and comparison to previous analyses is provided in this section.

<u>OWTS age and density:</u> OWTS permit data used to calculate system age and density introduced uncertainty and limitations to the model. First, the 3,717 OWTS located from the permit data do not represent all OWTSs in the Choccolocco Creek watershed. There are likely systems in use prior to start of the permit data collection (January 1, 2000), and systems that are no longer in use, where a homeowner connected to the sewer system. It is hypothesized that this study has underestimated the number of OWTSs in the watershed, considering it is estimated that more than half of OWTSs are more than 30 years old (U.S. EPA 2002). Second, the location of OWTSs within the parcel is not known, and the OWTS points were geocoded to the routing location, meaning the location next to the street segment associated with the address. Recently the parcel centroid has been used to estimate OWTS location (Capps et al. 2020) and the address location, represented by the rooftop centroid for the parcel, could be a more accurate approximation of the OWTS location. Third, the date of permit issuance was used to calculate the age of the system, but OWTS permits are valid for five years in Alabama, meaning an OWTS could be installed up to five years after a permit is issued (ADPH 2017). A lack of OWTS data or quality control methods,

such as those previously mentioned, are often cited as a major limitation in similar analyses (Capps et al. 2020). This highlights the need for publicly available OWTS permit databases, possibly georeferenced to parcels to eliminate the uncertainty introduced in geocoding.

Additionally, this model does not take into account unpermitted or straight pipe systems, where untreated sewage is discharged through pipes that run from the residence to a surface trench or wooded area (Maxcy-Brown et al. 2021). The use of straight pipe systems is documented across the U.S., including rural, central Alabama (Maxcy-Brown et al. 2021). These systems are most common in poor, rural areas with unfavorable geologic settings for OWTSs where a more expensive advanced engineered system would be necessary (Maxcy-Brown et al. 2021). However, the prevalence and location are not known, and they could not be accounted for in this study. With available OWTS, sewer system, and building data a 'process of elimination' could be used to locate parcels potentially containing an unpermitted wastewater treatment system (i.e., straight pipe); similar to methods used by Capps et al. (2020) to locate potential septic tank systems. Untreated sewage from straight pipe systems is another source of onsite wastewater contamination, so their estimated locations could be incorporated into an ESV-MCDA or used to create a straight pipe system specific MCDA to better characterize sources of contaminants.

Beyond uncertainties inherent in the raw OWTS permit data, the methods used to rasterize and classify OWTS age and density impact model results. When OWTS age was incorporated in similar models it was linked to parcel polygons (Oosting and Joy 2011; Capps et al. 2020), but without available parcel boundaries an IDW surface was created to incorporate the system age in the model. This tool requires a search distance, which was based on Alabama septage regulations, but the boundary imposed is somewhat arbitrary in defining the spatial extent of OWTS age. A similar issue arises when selecting a search distance for calculating the OWTS density. The search distance used for this analysis was selected so that the number of OWTSs were calculated per 1 km^2 to better facilitate classification breaks based on previously established categories. Additionally, OWTS density classifications are generalized, and specific areas may be able to accommodate a high or lower of OWTS, based on environmental variables. For example, areas with more suitable soil conditions for OWTSs may be better equipped to handle more OWTSs than areas with less suitable soil conditions. Thus, generalized restrictions on OWTSs based on density may be inappropriate for planning purposes, as a holistic view of all criteria is required to understand the susceptibility to OWTS failure (U.S. EPA 2002). Furthermore, there are not OWTS density restrictions in the Alabama OWTS regulations.

Slope of the terrain: The topography of the land is important in designing an OWTS as flatter terrain does not strictly correlate to better performance. Collick et al. (2006) assessed OWTS performance with changing slope of the terrain in the Catskills, NY and found that flatter terrains correlated to an increased risk of OWTS failure. These findings were attributed to the undulating landscape of the region where topographic lows with a high water table and low hydraulic conductivity accumulate water and potential contaminants. However, OWTSs on steeper slopes required greater setback distances from waters. Collick et al. (2006) recommended that a prescriptive upper slope boundary may not appropriately describe the risk of failure associated with slope and all landscape factors should be considered for the region. The findings highlight that the most suitable slopes for system performance may not follow general guidelines and a comprehensive analysis of landscape patterns should be considered.

Proximity to surface water: The minimum setback distance to surface waters varies by regulatory agency. For example, Alabama requires a minimum setback distance of 50 ft (15.24 m) from surface waters to the drainfield, and the minimum distance can be reduced with proper documentations from an engineer or geologist (ADPH 2017). However, in Georgia the minimum setback is 25 ft (7.62 m) to surface waters (GDPH 2019). Thus, the minimum setback distance to surface waters is dependent upon the regulatory codes for the OWTS location. The use of prescribed setback distances can be arbitrary and are often based on standards used by others and not scientific evaluations of the site (U.S. EPA 2002).

Soil characteristics: The optimal soil characteristics are important for proper function of OWTS drainfield. Soil characteristics (drainage class and hydrologic group) were sourced from SSURGO, and these data are collected by county, leading to possible discrepancies in the soil characterization based on the county. As a result, the county boundaries may propagate through the dataset, and this is evident in the soil hydrologic group data, where there are abrupt changes in the hydrologic group are seen at the county boundaries (Figure 10). For example, the county boundary is apparent between Talladega and Clay counties; the soil hydrologic group is classified as moderate in Clay county while it is classified as high in Talladega county even though there is no geologic underpinning for this discrepancy. This is an artefact of the data aggregation methods. Higher-resolution soil data could provide more accurate model results.

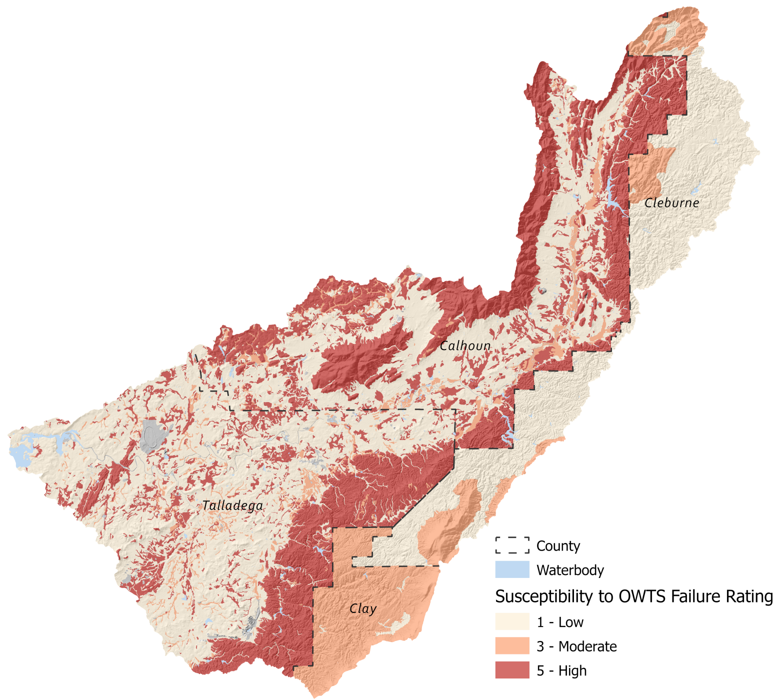

Figure 10. *The classified soil hydrologic groups overlaid with the county boundaries for the Choccolocco Creek watershed.*

Confounding criteria: There are additional variables that affect OWTS performance not included in this analysis, primarily due to a lack of available data, including OWTS design and system maintenance history. OWTSs vary in size and design depending upon the estimated wastewater flow, normally estimated on a per person or per bedroom basis (U.S. EPA 2002). The volume of wastewater entering the system may change over time as the number of residents in the building changes or plumbing is added, potentially leading to hydraulic failure from overloading the system. Additionally, OWTSs require routine maintenance after installation. Septic tanks should be pumped when the sludge accumulation exceeds 30 percent of the tank volume or when it is encroaching on the inlet and outlet entrances, with pumping normally needed every 3 to 5 years (U.S. EPA 2002). Unpumped septic tanks can result in clogged adsorption fields and hydraulic failure from overloading the system. Despite the importance of these criteria in system performance, there are minimal data on system design and maintenance history available in the state of Alabama, stemming from the fact that individual homeowners are responsible for the maintenance of their system, and regulatory agencies do not oversee system repairs.

4.1.3. Future Work

There are several ways the models developed in this study could be improved. First, more robust data, particularly OWTS data, would improve the model accuracy. Incorporating all known OWTSs in the watershed would improve OWTS density and age accuracy. Additionally, incorporating parcel boundaries in the analysis could provide a better unit of analysis for OWTS age. The model could be further improved by completing a sensitivity analysis, where the classification schemes and weights are altered to assess how the model changes. Similarly, the thresholds used to define minimal, low, moderate, and high susceptibility to OWTS failure could be altered to assess how the model results change. Review and input by local experts could also guide further improvements. For example, Oosting and Joy (2011) collected input from local experts (e.g., chief building officials, registered code agencies, public health inspectors and researchers actively studying OWTSs) on their weighting scheme, and after review experts suggested OWTS age receive a higher weight due to the overall importance of the age in assessing system failure and a high degree of confidence in the OWTS age data for the area of study. Moreover, the model could be validated by comparing the location of known failed systems and areas of high susceptibility to OWTS failure. We would hypothesize that there is a higher propensity of OWTS failure in areas of higher susceptibility; if that was not validated by field data an inverse approach could be applied, where explanatory variables are classified and weighted based on the attributes at known failed OWTSs.

There are several potential uses for this model beyond locating areas that may have surface water contamination from failed OWTSs. For example, results could be used to characterize

OWTS failure for *E. coli* loading estimations (e.g., Sowah et al. 2020) and source identification (e.g., Teaugue et al. 2012) research. The models could also support total maximum daily loads (TMDLs) development, which are used to establish allowable loads of pollutants for waterbodies based on the relationship between pollutant sources and in-stream concentrations for 303(d) listed waters. Thus, integration into TMDL protocol could better characterize the relationship between failed OWTSs and water contaminants, which is not explicitly accounted for as failed OWTSs are a nonpoint source of pollution. Additionally, model results could guide targeted local education or regulations for maintaining OWTSs. Furthermore, model results could aid in determining areas to expand sewer services; this could be particularly useful for areas that have a high susceptibility to OWTS failure due to environmental criteria, as these areas may require expensive, advanced OWTSs.

The model and results leave several outstanding questions: (1) Criteria were classified and weighted largely based on Alabama OWTS codes, how would the model change if codes from different states were used to classify criteria? (2) Are there variables not included in the model that better explain areas of potential OWTS failure? (3) If and how are the models related to failed OWTS in the Choccolocco Creek watershed?

4.2. Statistical Analysis (RQ 2)

4.2.1. Correlations

The second objective of this thesis was to determine if and how *E. coli* concentrations along the Choccolocco Creek were correlated to modelled susceptibility to OWTS failure, OWTS variables, and the distribution of land cover categories. At the SPW scale, *E. coli* was positively correlated to both the EV- and ESV-MCDA median and mean values (Table 6). Thus, as *E. coli* concentration increased, the mean and median model values OWTS failure models increased. At the 500 m, 750 m, and 1,000 m DDWs there was a positive correlation between *E. coli* and the 90th percentile EV-MCDA values (Table 6). A positive correlation was also observed between *E. coli* and the mean EV-MCDA values at the 1,000 m DDW (Table 6). The positive correlations observed with EV-MCDA may be reflective of poorly suited conditions for natural attenuation of contaminants. There was also a positive correlation between *E. coli* and OWTS count (Table 6); although, there was not a significant correlation to OWTS density (Table A.2). These results may suggest a relationship between OWTSs and elevated *E. coli* concentrations observed in Choccolocco Creek.

Several significant correlations were found between *E. coli* and land cover distribution. At the 500 m DDWs woody wetlands had a positive correlation to *E. coli*; emergent herbaceous wetlands also had a positive correlation with *E. coli* at the 1,000 m DDW (Table 6). However, wetland areas have been found to reduce coliform (Ibekwe, Grieve, and Lyon 2003), acting as a natural buffer of surface waters. At the same time, these natural wetland areas could attract wildlife and wildlife present in riparian zones could increase *E. coli* in proximal surface waters (Cox et al. 2005; Pandey et al. 2012). There was also a significant negative correlation between *E. coli* and open water at the 500 m DDW (Table 6), and this correlation is likely a result of an increased catchment area withing the watershed boundary. Considering the land cover distribution differs with SPWs (Figure 11), it is also important to note that no significant relationships were observed to between agricultural or urban land cover types that both serve as potential sources of *E. coli* (Table A.2). This is not to say these potential sources are not contributing to *E. coli* loading, but perhaps other sources – i.e., OWTSs – are driving elevated *E. coli* in the Choccolocco Creek. Based on the land cover distribution we expect the highest *E. coli* concentrations at sites one and five, since they have the highest proportion of agricultural and developed land cover, respectively. However, site one has a relatively low *E. coli* concentration even though most of the land cover in the SPW is either hay/pasture or developed. This may be a result of the site one SPW being relatively small and not fully capturing the land cover types in the catchment area. Site five has the highest average *E. coli* concentration and the highest proportion of developed land cover types for all SPWs, possibly driving the increase spike in *E. coli* observed at the site. These trends suggest that a combination of sources with variable relative contribution across the watershed are driving elevated *E. coli* concentrations in the Choccolocco Creek.

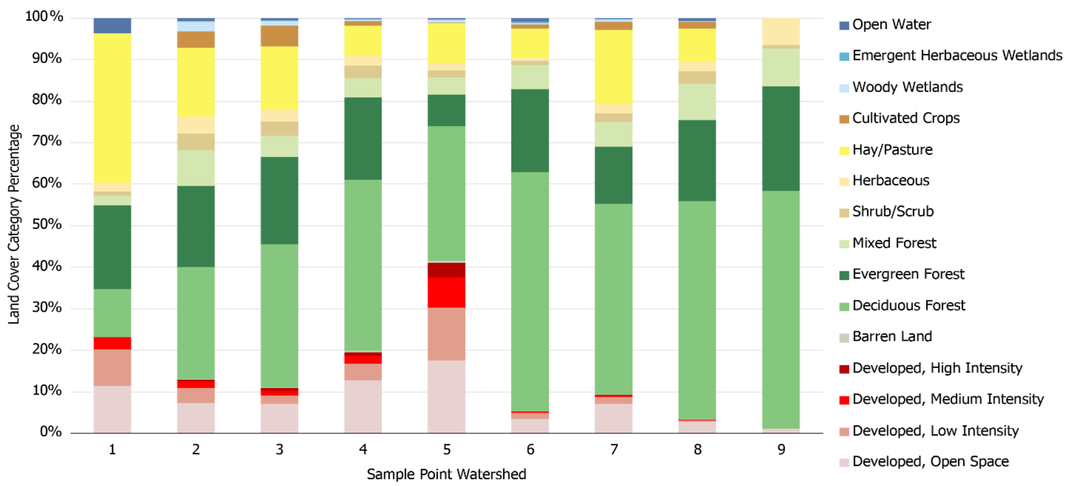

Figure 11. *The distribution of land cover types within the sample point watersheds delineated from water sampling locations along the Choccolocco Creek.*

Another objective was to elucidate how altering the unit of analysis, (i.e., watershed boundaries) affects the existence and significance of relationships. Thus, the SPWs and DDWs were delineated. As expected, presence and type of correlation (i.e., positive or negative) varied with watershed definition. Out of SPWs and all DDWs watersheds, most correlations were found with the SPWs (five out of 12). This may be because SPWs encapsulate the entire catchment area flowing into the pour point. For the DDWs, significant correlations were observed at the smaller watershed scales – 500 m, 750 m, and 1,000 m. No significant correlations were observed at the

250 m scale (Table A.2), but this could be a result of the 250 m watershed mostly encompassing water, where the MCDAs did not output a result. Furthermore, the variance in significant correlations with watershed boundaries is not unexpected, especially when viewed from a geospatial perspective; this is an example of the modifiable areal unit problem (MAUP). The MAUP is a statistical bias that can occur when aggregating data into a unit of analysis (Openshaw and Taylor 1979), and has been found in previous studies that examined water-quality indicators and watershed characteristics at various hydrologic units of analysis (Sun et al. 2014; Xiao et al. 2016; Zhang, Liu, and Zhou 2018; Sliva and Dudley Williams 2001). Moreover, studies that assessed the relationship between land cover with surface water quality indicators at multiple scales generally found correlations at smaller watershed scales, often between 100 m and 500 m (Xiao et al. 2016; Sun et al. 2014; Peng and Li 2021). Although, fecal coliforms were not used as a water quality indicator in these studies, possibly attributed to the variable lifespan of *E. coli* (Field and Samadpour 2007). In this analysis the DDWs did not provide significant supporting evidence to identify *E. coli* sources. However, the methods developed to delineate DDWs provide a more hydrologically correct catchment area compared to Euclidean distance buffers, which do not take into account surface water flow patterns, used in previous analyses (Sun et al. 2014; Xiao et al. 2016; Zhang, Liu, and Zhou 2018). Furthermore, the DDWs facilitate a more localized unit of analysis that could capture contaminant sources proximal to the sample location that may be muted in SPW-level analysis. Results from this analysis are not intended to prescribe a unit of analysis for future research, but to highlight the importance of selecting an appropriate unit of analysis, as there was variance in the presence and type of correlation (positive or negative) with watershed scales.

4.2.2. Limitations and Future Work

There are confounding environmental variables that can control in-stream *E. coli* concentrations. Environmental parameters such as water temperature (Blaustein et al. 2013), pH (Blackburn et al. 1997), solar insolation (Whitman et al. 2004; 2008), suspended and settled solids concentrations (Petersen and Hubbart 2020a; 2020b), and hydrologic conditions (Whitman et al. 2008; Wu et al. 2016) impact the survival of *E. coli*. For example, a positive correlation between solar insolation and *E. coli* inactivation rates have been observed, meaning there was a decrease in *E. coli* concentration with increased solar insolation (Whitman et al. 2004; 2008). As a result, the weather on sampling days (e.g., cloudy or sunny conditions), water clarity, and stream geometry (e.g., depth and width) could affect the amount of solar insolation and, thus, *E. coli* concentrations (Whitman et al. 2004). In the Choccolocco Creek watershed, significant correlations were found between precipitation (mm/day), pH, and water temperature with *E. coli* concentration (Larson 2022). Consequently, these environmental parameters could also drive spatial and temporal variance in the *E. coli* concentrations observed in the Choccolocco Creek.

There are several limitations to these methods of analysis. First, the water quality samples were only taken at nine sample locations and averaged over a six-month period. As a result, the correlations were calculated using nine data points. The small size could contribute to an increased Type I error rate, meaning a 'false positive' significant correlation (Bishara and Hittner 2012). Higher spatial and temporal resolution may yield different results because limited data points for *in situ* water quality parameters may introduce data bias, as noted in similar studies (Flanagan et al. 2019). Specifically, it would be useful to have water quality data from Choccolocco Creek tributaries, as this could provide insight to *E. coli* transport through the watershed and further narrow down specific sources. Greater temporal resolution could decrease the variability in *E. coli* concentrations and allow for consideration of seasonal changes in *E. coli* concentrations. Second, aggregating explanatory data in the SPWs and DDWs does not account for a potential downstream additive effect of variables on *E. coli*; each watershed was treated as an isolated unit. This is a simplistic view of the watershed dynamics used for analysis. Third, there is possible error in delineating watersheds using the snap pour point methodology (Lindsay, Rothwell, and Davies 2008). The *Snap to Pour Point* tool in ArcGIS Pro moves the outlet location to the cell with the highest accumulation flow within a given search distance. Watersheds delineated from these pour points can be highly sensitive to the search distance, and common errors that can occur are shown in Figure 12 (Lindsay, Rothwell, and Davies 2008). The SPW for sample one (Figure 5) may be smaller than the other watersheds due to an off-stream error, where the repositioned outlet point is not in the stream cells, as it was located near the river confluence (Figure 12A). A search distance up 90 m was tested with no major changes to the watershed delineated at sample site one. Therefore, a 30 m search radius was used as it was the lowest search distance that did not produce conspicuously small watersheds for all other sample locations. Thus, future research could explore different watershed delineation methods to assess accuracy. These methods of analysis are also largely dependent on the quality and resolution of input data and does not take into consider the influence of point sources (i.e., wastewater treatment plants). Additional improvements could include incorporating a more accurate estimation of possible *E. coli* sources, such as wildlife densities and

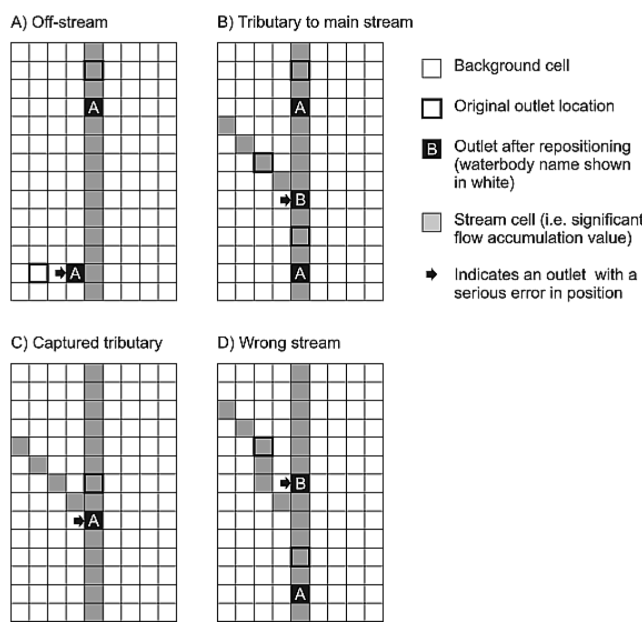

Figure 12. *Common errors that occur when repositioning an outlet point (Lindsay, Rothwell, and Davies 2008).*

agricultural (e.g., cows and chickens) densities, similar to methods used by Teague et al. (2012) and Thilakarathne, Sridhar, and Karthikeyan (2018).

5. Conclusions

The central objective of this thesis was to improve the understanding of how OWTSs potentially contribute to water contamination in the Choccolocco Creek using a GIS-based approach to (1) model the susceptibility to OWTS failure and (2) evaluate the relationship between the modeled OWTS failure, OWTS variables, and distribution of land cover types to in-stream *E. coli* concentrations. The first objective was addressed by developing EV- and ESV- MCDAs to model the susceptibility to OWTS failure. The results of the models changed with the inclusion of different input criteria and investigation of input criteria suggested that high OWTS age could drive OWTS failure in the Choccolocco Creek watershed. The second objective was met by correlating between *E. coli* concentration along the Choccolocco Creek to modeled OWTS failure, OWTS variables, and distribution of land cover types. SPWs and DDWs were delineated from water sampling locations and used to aggregate data, and significant correlations differed with the unit of analysis. Most notably, a positive correlation was found between both the EV- and ESV-MCDAs with *E. coli*. Results suggest a relationship between OWTSs and elevated *E. coli* concentrations in Choccolocco Creek. Through these analyses and previous studies, it is apparent that modeling OWTS failure and their relationship to water contamination requires a holistic approach in both analyzing the criteria that impact system efficacy and watershed variables that impact *E. coli* concentrations. Nonpoint source identification challenges are not unique to the Choccolocco Creek watershed and methods outlined here could be applied to other watersheds to assist in understanding how OWTSs may contribute to surface water *E. coli* contamination.

References

ADEM. 2020a. “2020 Alabama §303(d) List.” http://adem.alabama.gov/programs/water/wquality/2020AL303dList.pdf.

———. 2020b. “Alabama’s Water Quality Assessment and Listing Methodology.” http://adem.alabama.gov/programs/water/wquality/2020WAM.pdf.

———. 2022. “2022 Alabama Draft §303(d) List.” http://adem.alabama.gov/newsEvents/notices/jan22/pdfs/12022AL303dlist.pdf.

ADPH. 2017. “Onsite Sewage Treatment and Disposal, Chapter 420-3-1.”

Ayres Associates. 1993. “Onsite Sewage Disposal System Research in Florida.” https://www.floridahealth.gov/environmental-health/onsite-sewage/research/_documents/research-reports/_documents/1993-capability-of-fine-sandy-soil-for-septi-tank-effluent-treatment1.pdf.

Beal, C. D., E. A. Gardner, G. Kirchhof, and N. W. Menzies. 2006. “Long-Term Flow Rates and Biomat Zone Hydrology in Soil Columns Receiving Septic Tank Effluent.” *Water Research* 40 (12): 2327–38. https://doi.org/10.1016/j.watres.2006.04.018.

Bishara, Anthony J., and James B. Hittner. 2012. “Testing the Significance of a Correlation with Nonnormal Data: Comparison of Pearson, Spearman, Transformation, and Resampling Approaches.” *Psychological Methods* 17 (3): 399–417. https://doi.org/10.1037/a0028087.

Blackburn, C. de W, L. M Curtis, L Humpheson, C Billon, and P. J McClure. 1997. “Development of Thermal Inactivation Models for Salmonella Enteritidis and Escherichia Coli O157:H7 with Temperature, PH and NaCl as Controlling Factors.”

International Journal of Food Microbiology 38 (1): 31–44. https://doi.org/10.1016/S0168-1605(97)00085-8.

Blaustein, R. A., Y. Pachepsky, R. L. Hill, D. R. Shelton, and G. Whelan. 2013. "Escherichia Coli Survival in Waters: Temperature Dependence." *Water Research* 47 (2): 569–78. https://doi.org/10.1016/j.watres.2012.10.027.

Capps, Krista A., Jacob M. Bateman McDonald, Nandita Gaur, and Rebecca Parsons. 2020. "Assessing the Socio-Environmental Risk of Onsite Wastewater Treatment Systems to Inform Management Decisions." *Environmental Science & Technology* 54 (23): 14843–53. https://doi.org/10.1021/acs.est.0c03909.

Carroll, Steven, Ashantha Goonetilleke, Evan Thomas, Megan Hargreaves, Ray Frost, and Les Dawes. 2006. "Integrated Risk Framework for Onsite Wastewater Treatment Systems." *Environmental Management* 38 (2): 286–303. https://doi.org/10.1007/s00267-005-0280-5.

Cogger, C. G., and B. L. Carlile. 1984. "Field Performance of Conventional and Alternative Septic Systems in Wet Soils." *Journal of Environmental Quality* 13 (1): 137–42. https://doi.org/10.2134/jeq1984.00472425001300010025x.

Collick, Amy S., Zachary M. Easton, Franco A. Montalto, Bin Gao, Young-Jin Kim, Laurence Day, and Tammo S. Steenhuis. 2006. "Hydrological Evaluation of Septic Disposal Field Design in Sloping Terrains." *Journal of Environmental Engineering* 132 (10): 1289–97. https://doi.org/10.1061/(ASCE)0733-9372(2006)132:10(1289).

Cooper, Jennifer A., George W. Loomis, David V. Kalen, and Jose A. Amador. 2015. "Evaluation of Water Quality Functions of Conventional and Advanced Soil-Based

Onsite Wastewater Treatment Systems.” *Journal of Environmental Quality* 44 (3): 953–62. https://doi.org/10.2134/jeq2014.06.0277.

Cox, Peter, Merran Griffith, Mark Angles, Daniel Deere, and Christobel Ferguson. 2005. “Concentrations of Pathogens and Indicators in Animal Feces in the Sydney Watershed.” *Applied and Environmental Microbiology* 71 (10): 5929–34. https://doi.org/10.1128/AEM.71.10.5929-5934.2005.

Eastman, J. R., Jin Weigen, Peter Kyem, and James Toledano. 1995. “Raster Procedures for Multi-Criteria/Multi-Obiective Decisions.” *Photogrammetric Engineering & Remote Sensing* 61 (5): 539–47.

Esri. 2016. “How To: Create a Watershed Model Using the Hydrology Toolset.” 2016. https://support.esri.com/en/technical-article/000012346.

———. 2022a. “Geocode Addresses (Geocoding)—ArcGIS Pro | Documentation.” 2022. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/geocoding/geocode-addresses.htm.

———. 2022b. “IDW (Spatial Analyst)—ArcGIS Pro | Documentation.” 2022. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/idw.htm.

———. 2022c. “Point Density (Spatial Analyst)—ArcGIS Pro | Documentation.” 2022. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/point-density.htm.

———. 2022d. “Slope (Spatial Analyst)—ArcGIS Pro | Documentation.” 2022. https://pro.arcgis.com/en/pro-app/2.8/tool-reference/spatial-analyst/slope.htm.

Field, Katharine G., and Mansour Samadpour. 2007. “Fecal Source Tracking, the Indicator Paradigm, and Managing Water Quality.” *Water Research*, Identifying Sources of Fecal Pollution, 41 (16): 3517–38. https://doi.org/10.1016/j.watres.2007.06.056.

Flanagan, Kyle, Barnali Dixon, Tess Rivenbark, and Dale Griffin. 2019. "An Integrative GIS Approach to Analyzing the Impacts of Septic Systems on the Coast of Florida, USA." *Physical Geography*, September, 1–26. https://doi.org/10.1080/02723646.2019.1671297.

GDPH. 2019. "Manual for On-Site Sewage Management Systems." https://dph.georgia.gov/document/document/manual-site-sewage-management-systems-rules/download.

Ghasemi, Asghar, and Saleh Zahediasl. 2012. "Normality Tests for Statistical Analysis: A Guide for Non-Statisticians." *International Journal of Endocrinology and Metabolism* 10 (2): 486–89. https://doi.org/10.5812/ijem.3505.

Ibekwe, A. Mark, Catherine M. Grieve, and Stephen R. Lyon. 2003. "Characterization of Microbial Communities and Composition in Constructed Dairy Wetland Wastewater Effluent." *Applied and Environmental Microbiology* 69 (9): 5060–69. https://doi.org/10.1128/AEM.69.9.5060-5069.2003.

Katz, Brian, Dale Griffin, Peter Mcmahon, Harmon Harden, Edgar Wade, Richard Hicks, and Jeffrey Chanton. 2010. "Fate of Effluent-Borne Contaminants beneath Septic Tank Drainfields Overlying a Karst Aquifer." *Journal of Environmental Quality* 39 (July): 1181–95. https://doi.org/10.2134/jeq2009.0244.

Larson, Eleanore. 2022. "Determining Sources of E. Coli Using Human Proxies in Choccolocco Creek, AL." Auburn University.

Lindsay, John B., James J. Rothwell, and Helen Davies. 2008. "Mapping Outlet Points Used for Watershed Delineation onto DEM-Derived Stream Networks." *Water Resources Research* 44 (8). https://doi.org/10.1029/2007WR006507.

Lusk, Mary G., Gurpal S. Toor, Yun-Ya Yang, Sara Mechtensimer, Mriganka De, and Thomas A. Obreza. 2017. “A Review of the Fate and Transport of Nitrogen, Phosphorus, Pathogens, and Trace Organic Chemicals in Septic Systems.” *Critical Reviews in Environmental Science and Technology* 47 (7): 455–541. https://doi.org/10.1080/10643389.2017.1327787.

Malczewski, Jacek. 1999. *GIS and Multicriteria Decision Analysis*. John Wiley & Sons.

Maxcy-Brown, Jillian, Mark A. Elliott, Leigh Anne Krometis, Joe Brown, Kevin D. White, and Upmanu Lall. 2021. “Making Waves: Right in Our Backyard- Surface Discharge of Untreated Wastewater from Homes in the United States.” *Water Research* 190 (February): 116647. https://doi.org/10.1016/j.watres.2020.116647.

Nyerges, Timothy L., and Piotr Jankowski. 2009. *Regional and Urban GIS: A Decision Support Approach*. Guilford Press.

Oosting, Andrew, and Doug Joy. 2011. “A GIS-Based Model to Assess the Risk of On-Site Wastewater Systems Impacting Groundwater and Surface Water Resources.” *Canadian Water Resources Journal* 36 (September): 229–46. https://doi.org/10.4296/cwrj3603882.

Openshaw, S, and P. J. Taylor. 1979. “A Million or so Correlation Coefficients, Three Experiments on the Modifiable Areal Unit Problem.” *Statistical Applications in the Spatial Science*, 127–44.

Pandey, Pramod K., Michelle L. Soupir, Monica Haddad, and James J. Rothwell. 2012. “Assessing the Impacts of Watershed Indexes and Precipitation on Spatial In-Stream E. Coli Concentrations.” *Ecological Indicators* 23 (December): 641–52. https://doi.org/10.1016/j.ecolind.2012.05.023.

Patterson, James J., Carl Smith, and Jennifer Bellamy. 2013. "Understanding Enabling Capacities for Managing the 'Wicked Problem' of Nonpoint Source Water Pollution in Catchments: A Conceptual Framework." *Journal of Environmental Management* 128 (October): 441–52. https://doi.org/10.1016/j.jenvman.2013.05.033.

Peng, Shuangyun, and Shihua Li. 2021. "Scale Relationship between Landscape Pattern and Water Quality in Different Pollution Source Areas: A Case Study of the Fuxian Lake Watershed, China." *Ecological Indicators* 121 (February): 107136. https://doi.org/10.1016/j.ecolind.2020.107136.

Petersen, Fritz, and Jason A. Hubbart. 2020a. "Quantifying Escherichia Coli and Suspended Particulate Matter Concentrations in a Mixed-Land Use Appalachian Watershed." *Water* 12 (2): 532. https://doi.org/10.3390/w12020532.

———. 2020b. "Physical Factors Impacting the Survival and Occurrence of Escherichia Coli in Secondary Habitats." *Water* 12 (6): 1796. https://doi.org/10.3390/w12061796.

Schober, Patrick, Christa Boer, and Lothar A. Schwarte. 2018. "Correlation Coefficients: Appropriate Use and Interpretation." *Anesthesia & Analgesia* 126 (5): 1763–68. https://doi.org/10.1213/ANE.0000000000002864.

Sliva, Lucie, and D Dudley Williams. 2001. "Buffer Zone versus Whole Catchment Approaches to Studying Land Use Impact on River Water Quality." *Water Research* 35 (14): 3462–72. https://doi.org/10.1016/S0043-1354(01)00062-8.

Sun, Yanwei, Qinghai Guo, Jian Liu, and Run Wang. 2014. "Scale Effects on Spatially Varying Relationships Between Urban Landscape Patterns and Water Quality." *Environmental Management* 54 (2): 272–87. https://doi.org/10.1007/s00267-014-0287-x.

Teague, Aarin, R. Karthikeyan, Meghna Babbar-Sebens, R. Srinivasan, and R. Persyn. 2009. “Spatially Explicit Load Enrichment Calculation Tool to Identify Potential E. Coli Sources in Watersheds.” *Transactions of the ASABE* 52 (July). https://doi.org/10.13031/2013.27788.

Thilakarathne, Madusanka, Venkataramana Sridhar, and Raghupathy Karthikeyan. 2018. “Spatially Explicit Pollutant Load-Integrated in-Stream E. Coli Concentration Modeling in a Mixed Land-Use Catchment.” *Water Research* 144 (November): 87–103. https://doi.org/10.1016/j.watres.2018.07.021.

U.S. Census Bureau. 1990. “We Asked... You Told Us: Source of Water and Sewage Disposal,” 2.

U.S. Census Bureau, American Housing Survey. 2019. “American Housing Survey (AHS) - AHS Table Creator.” 2019. https://www.census.gov/programs-surveys/ahs/data/interactive/ahstablecreator.html?s_areas=00000&s_year=2019&s_tablename=TABLE4&s_bygroup1=3&s_bygroup2=1&s_filtergroup1=1&s_filtergroup2=1&s_show=S.

U.S. EPA. 1977. “Waste Disposal Practices and Their Effects on Groundwater.”

———. 2000. “EPA Guildelines for Management of Onsite/Decentralized Wastewater Systems.” https://www.epa.gov/sites/default/files/2015-06/documents/2004_07_07_septics_septic_guidelines_factsheet.pdf.

———. 2002. “Onsite Wastewater Treatment System Manual.” https://www.epa.gov/sites/default/files/2015-06/documents/2004_07_07_septics_septic_2002_osdm_all.pdf.

USGS. 2019. “National Land Cover Database (NLCD) 2019 Products.” doi:10.5066/P9KZCM54.

———. 2020a. “USGS 1/3 Arc-Second N34w086 1 x 1 Degree: U.S. Geological Survey.” https://apps.nationalmap.gov/downloader/#/.

———. 2020b. “USGS 1/3 Arc-Second N34w087 1 x 1 Degree: U.S. Geological Survey.” https://apps.nationalmap.gov/downloader/#/.

———. 2021. “National Hydrography Dataset.” https://www.usgs.gov/core-science-systems/ngp/national-hydrography/access-national-hydrography-products.

Welhan, John, and Carol Moore. 2012. “An Objective GIS Screening Tool for Rating the Suitability of Land for Septic-Based Development.” Idaho Geologic Survey. https://www.idahogeology.org/product/s-12-1.

Whitman, Richard L., Meredith B. Nevers, Ginger C. Korinek, and Muruleedhara N. Byappanahalli. 2004. “Solar and Temporal Effects on Escherichia Coli Concentration at a Lake Michigan Swimming Beach.” *Applied and Environmental Microbiology* 70 (7): 4276–85. https://doi.org/10.1128/AEM.70.7.4276-4285.2004.

Whitman, Richard L., Katarzyna Przybyla-Kelly, Dawn A. Shively, Meredith B. Nevers, and Muruleedhara N. Byappanahalli. 2008. “Sunlight, Season, Snowmelt, Storm, and Source Affect E. Coli Populations in an Artificially Ponded Stream.” *Science of The Total Environment* 390 (2): 448–55. https://doi.org/10.1016/j.scitotenv.2007.10.014.

Whittier, Robert B, and Aly I El-Kadi. 2009. “Human and Environmental Risk Ranking of Onsite Sewage Disposal Systems,” December, 142.

Withers, Paul JA, Philip Jordan, Linda May, Helen P. Jarvie, and Nancy E. Deal. 2013. "Do Septic Tank Systems Pose a Hidden Threat to Water Quality?" *Frontiers in Ecology and the Environment* 12 (2): 123–30. https://doi.org/10.1890/130131.

Wu, Jianyong, Mohammad Yunus, Md. Sirajul Islam, and Michael Emch. 2016. "Influence of Climate Extremes and Land Use on Fecal Contamination of Shallow Tubewells in Bangladesh." *Environmental Science & Technology* 50 (5): 2669–76. https://doi.org/10.1021/acs.est.5b05193.

Xiao, Rui, Guofeng Wang, Qianwen Zhang, and Zhonghao Zhang. 2016. "Multi-Scale Analysis of Relationship between Landscape Pattern and Urban River Water Quality in Different Seasons." *Scientific Reports* 6 (1): 25250. https://doi.org/10.1038/srep25250.

Yuan, Lifeng, Tadesse Sinshaw, and Kenneth J. Forshay. 2020. "Review of Watershed-Scale Water Quality and Nonpoint Source Pollution Models." *Geosciences* 10 (1): 25. https://doi.org/10.3390/geosciences10010025.

Zhang, Xin, Yuqi Liu, and Lin Zhou. 2018. "Correlation Analysis between Landscape Metrics and Water Quality under Multiple Scales." *International Journal of Environmental Research and Public Health* 15 (8): 1606. https://doi.org/10.3390/ijerph15081606.

Appendix

Table A.1. Descriptive statistics for *E. coli* concentrations measured along the Choccolocco Creek by Larson (2022).

Site Number	Site Name	Sample Count	Minimum (cfu/100 mL)	Maximum (cfu/100mL)	Mean (cfu/100mL)	Standard Deviation (cfu/100 mL)
1	Stemley Rd	16*	0	579.4	141.0	227.4
2	Hwy 77	16	0	316.7	91.7	126.8
3	Jackson Trace	16	0	316.7	188.9	116.7
4	Phillip Watts Property	16*	116.7	500	293.6	153.3
5	Silver Run	16	50	833.3	377.8	302.5
6	Boiling springs	16*	0	450	152.8	162.1
7	Hwy 78	16	50	633.3	300.0	215.5
8	Chosea Springs	16	50	666.7	313.9	229.4
9	Talladega 540	16	0	350	158.3	125.9

* Sample on August 26, 2021, provided by Coosa Riverkeepers

Table A.2. Correlation coefficients (R), R-squared, and p-values to average *E. coli* concentrations in the Choccolocco Creek.

Variable	Shapiro-Wilks Test		Correlations			
	p-value	*Normal? (α < 0.05)*	*Type*	*R*	*R-squared*	*P-value*
SPW ESV-MCDA min	<0.0001	No	Spearman	-0.4108	0.1688	0.4444
SPW ESV-MCDA max	0.001	No	Spearman	0.4118	0.1696	0.2707
SPW ESV-MCDA range	0.0013	No	Spearman	0.4118	0.1696	0.2707
SPW ESV-MCDA mean	**0.3603**	**Yes**	**Pearson**	**0.6732**	**0.4532**	**0.0469***
SPW ESV-MCDA standard deviation	0.0202	No	Spearman	0.2594	0.0673	0.4980
SPW ESV-MCDA median	**0.2800**	**Yes**	**Pearson**	**0.8729**	**0.7620**	**0.0021***
SPW ESV-MCDA 90th percentile	0.2312	Yes	Pearson	0.4067	0.1654	0.2774
250 m DDW ESV-MCDA min	0.0289	No	Spearman	0.2623	0.0688	0.4926
250 m DDW ESV-MCDA max	0.9196	Yes	Pearson	0.2013	0.0405	0.6036
250 m DDW ESV-MCDA range	0.8783	Yes	Pearson	0.0832	0.0069	0.8315
250 m DDW ESV-MCDA mean	0.6289	Yes	Pearson	-0.0057	0.0000	0.9884
250 m DDW ESV-MCDA standard deviation	0.7579	Yes	Pearson	-0.0881	0.0078	0.8216
250 m DDW ESV-MCDA median	0.5515	Yes	Pearson	-0.0951	0.0090	0.8077
250 m DDW ESV-MCDA 90th percentile	0.7528	Yes	Pearson	-0.0217	0.0005	0.9558
500 m DDW ESV-MCDA min	0.0005	No	Spearman	-0.0996	0.0099	0.8016
500 m DDW ESV-MCDA max	0.3263	Yes	Pearson	0.2165	0.0469	0.5757
500 m DDW ESV-MCDA range	0.4821	Yes	Pearson	0.2423	0.0587	0.5299
500 m DDW ESV-MCDA mean	0.6894	Yes	Pearson	-0.0457	0.0021	0.9070

500 m DDW ESV-MCDA standard deviation	0.2169	Yes	Pearson	0.0387	0.0015	0.9213
500 m DDW ESV-MCDA median	0.4855	Yes	Pearson	-0.1276	0.0163	0.7436
500 m DDW ESV-MCDA 90th percentile	0.6997	Yes	Pearson	-0.0244	0.0006	0.9502
750 m DDW ESV-MCDA min	0.0005	No	Spearman	-0.0996	0.0099	0.8016
750 m DDW ESV-MCDA max	0.0393	No	Spearman	0.0833	0.0069	0.8432
750 m DDW ESV-MCDA range	0.2397	Yes	Pearson	0.2701	0.0729	0.4821
750 m DDW ESV-MCDA mean	0.7093	Yes	Pearson	0.0570	0.0032	0.8842
750 m DDW ESV-MCDA standard deviation	0.9749	Yes	Pearson	0.1168	0.0136	0.7648
750 m DDW ESV-MCDA median	0.5966	Yes	Pearson	0.0853	0.0073	0.8273
750 m DDW ESV-MCDA 90th percentile	0.6984	Yes	Pearson	0.0160	0.0003	0.9673
1000 m DDW ESV-MCDA min	0.0009	No	Spearman	-0.0199	0.0004	0.9683
1000 m DDW ESV-MCDA max	0.0293	No	Spearman	0.3000	0.0900	0.4366
1000 m DDW ESV-MCDA range	0.0141	No	Spearman	0.4833	0.2336	0.1938
1000 m DDW ESV-MCDA mean	0.7151	Yes	Pearson	0.2998	0.0899	0.4332
1000 m DDW ESV-MCDA standard deviation	0.8214	Yes	Pearson	0.1514	0.0229	0.6975
1000 m DDW ESV-MCDA median	0.7896	Yes	Pearson	0.2668	0.0712	0.4876
1000 m DDW ESV-MCDA 90th percentile	0.7314	Yes	Pearson	0.1863	0.0347	0.6313
2500 m DDW ESV-MCDA min	<0.0001	No	Spearman	-0.0456	0.0021	0.9444
2500 m DDW ESV-MCDA max	0.3443	Yes	Pearson	0.3208	0.1029	0.3999
2500 m DDW ESV-MCDA range	0.3330	Yes	Pearson	0.3154	0.0995	0.4084
2500 m DDW ESV-MCDA mean	0.6801	Yes	Pearson	0.4018	0.1615	0.2837
2500 m DDW ESV-MCDA standard deviation	0.1879	Yes	Pearson	0.2184	0.0477	0.5723
2500 m DDW ESV-MCDA median	0.3606	Yes	Pearson	0.4267	0.1821	0.2521
2500 m DDW ESV-MCDA 90th percentile	0.3615	Yes	Pearson	0.2579	0.0665	0.5028
5000 m DDW ESV-MCDA min	<0.0001	No	Spearman	-0.4108	0.1688	0.4444
5000 m DDW ESV-MCDA max	0.2411	Yes	Pearson	0.3497	0.1223	0.3563
5000 m DDW ESV-MCDA range	0.2812	Yes	Pearson	0.3670	0.1347	0.3313
5000 m DDW ESV-MCDA mean	0.5092	Yes	Pearson	0.2355	0.0555	0.5419
5000 m DDW ESV-MCDA standard deviation	0.0399	No	Spearman	0.1187	0.0141	0.7721
5000 m DDW ESV-MCDA median	0.1323	Yes	Pearson	0.3474	0.1207	0.3596
5000 m DDW ESV-MCDA 90th percentile	0.5402	Yes	Pearson	0.2054	0.0422	0.5959
10000 m DDW ESV-MCDA min	<0.0001	No	Spearman	-0.4108	0.1688	0.4444
10000 m DDW ESV-MCDA max	0.0064	No	Spearman	0.2259	0.0510	0.5582
10000 m DDW ESV-MCDA range	0.0104	No	Spearman	0.2259	0.0510	0.5582
10000 m DDW ESV-MCDA mean	0.6224	Yes	Pearson	0.5837	0.3407	0.0989
10000 m DDW ESV-MCDA standard deviation	0.0004	No	Spearman	0.4622	0.2136	0.2129
10000 m DDW ESV-MCDA median	0.1507	Yes	Pearson	0.6538	0.4275	0.0561
10000 m DDW ESV-MCDA 90th percentile	0.0268	No	Spearman	0.3431	0.1177	0.3640
SPW EV-MCDA min			No variance			
SPW EV-MCDA max	0.0013	No	Spearman	0.2469	0.0610	0.5185
SPW EV-MCDA range	0.0013	No	Spearman	0.2469	0.0610	0.5185
SPW EV-MCDA mean	**0.0880**	**Yes**	**Pearson**	**0.7109**	**0.5054**	**0.0318***
SPW EV-MCDA standard deviation	0.0114	No	Spearman	0.6276	0.3939	0.0776

SPW EV-MCDA median	**0.4372**	**Yes**	**Pearson**	**0.8444**	**0.7131**	**0.0042***
SPW EV-MCDA 90th percentile	0.0284	No	Spearman	0.6061	0.3674	0.0947
250 m DDW EV-MCDA min	<0.0001	No	Spearman	0.0685	0.0047	0.8889
250 m DDW EV-MCDA max	0.0551	Yes	Pearson	0.5865	0.3439	0.0969
250 m DDW EV-MCDA range	0.1841	Yes	Pearson	0.3737	0.1397	0.3218
250 m DDW EV-MCDA mean	0.6963	Yes	Pearson	0.1455	0.0212	0.7088
250 m DDW EV-MCDA standard deviation	0.5661	Yes	Pearson	0.6019	0.3623	0.0864
250 m DDW EV-MCDA median	0.5478	Yes	Pearson	-0.1798	0.0323	0.6434
250 m DDW EV-MCDA 90th percentile	0.1095	Yes	Pearson	0.6390	0.4083	0.0639
500 m DDW EV-MCDA min			No variance			
500 m DDW EV-MCDA max	0.5616	Yes	Pearson	0.5145	0.2647	0.1564
500 m DDW EV-MCDA range	0.5616	Yes	Pearson	0.5145	0.2647	0.1564
500 m DDW EV-MCDA mean	0.2795	Yes	Pearson	0.4777	0.2282	0.1935
500 m DDW EV-MCDA standard deviation	0.3215	Yes	Pearson	0.4777	0.2282	0.1934
500 m DDW EV-MCDA median	0.3120	Yes	Pearson	0.4049	0.1639	0.2797
500 m DDW EV-MCDA 90th percentile	**0.0547**	**Yes**	**Pearson**	**0.8064**	**0.6503**	**0.0086***
750 m DDW EV-MCDA min			No variance			
750 m DDW EV-MCDA max	0.5518	Yes	Pearson	0.6563	0.4307	0.0549
750 m DDW EV-MCDA range	0.5518	Yes	Pearson	0.6563	0.4307	0.0549
750 m DDW EV-MCDA mean	0.5487	Yes	Pearson	0.6113	0.3737	0.0803
750 m DDW EV-MCDA standard deviation	0.5561	Yes	Pearson	0.5249	0.2755	0.1468
750 m DDW EV-MCDA median	0.1044	Yes	Pearson	0.3286	0.1080	0.3879
750 m DDW EV-MCDA 90th percentile	**0.1597**	**Yes**	**Pearson**	**0.8241**	**0.6792**	**0.0063***
1000 m DDW EV-MCDA min			No variance			
1000 m DDW EV-MCDA max	0.9285	Yes	Pearson	0.6215	0.3862	0.0740
1000 m DDW EV-MCDA range	0.9285	Yes	Pearson	0.6215	0.3862	0.0740
1000 m DDW EV-MCDA mean	**0.6033**	**Yes**	**Pearson**	**0.7221**	**0.5214**	**0.0280***
1000 m DDW EV-MCDA standard deviation	0.7768	Yes	Pearson	0.5184	0.2688	0.1528
1000 m DDW EV-MCDA median	0.0656	Yes	Pearson	0.5576	0.3109	0.1188
1000 m DDW EV-MCDA 90th percentile	**0.1915**	**Yes**	**Pearson**	**0.8214**	**0.6748**	**0.0066***
2500 m DDW EV-MCDA min			No variance			
2500 m DDW EV-MCDA max	0.1798	Yes	Pearson	0.4053	0.1643	0.2792
2500 m DDW EV-MCDA range	0.1798	Yes	Pearson	0.4053	0.1643	0.2792
2500 m DDW EV-MCDA mean	0.1026	Yes	Pearson	0.6358	0.4043	0.0657
2500 m DDW EV-MCDA standard deviation	0.4070	Yes	Pearson	0.5341	0.2853	0.1385
2500 m DDW EV-MCDA median	0.0629	Yes	Pearson	0.5696	0.3244	0.1094
2500 m DDW EV-MCDA 90th percentile	0.5107	Yes	Pearson	0.6049	0.3659	0.0844
5000 m DDW EV-MCDA min			No variance			
5000 m DDW EV-MCDA max	0.0755	Yes	Pearson	0.4115	0.1693	0.2712
5000 m DDW EV-MCDA range	0.0755	Yes	Pearson	0.4115	0.1693	0.2712
5000 m DDW EV-MCDA mean	0.3985	Yes	Pearson	0.3079	0.0948	0.4203
5000 m DDW EV-MCDA standard deviation	0.1841	Yes	Pearson	0.4077	0.1662	0.2760
5000 m DDW EV-MCDA median	0.0396	No	Spearman	0.2992	0.0895	0.4276

5000 m DDW EV-MCDA 90th percentile	0.1755	Yes	Pearson	0.4286	0.1837	0.2498
10000 m DDW EV-MCDA min			No variance			
10000 m DDW EV-MCDA max	0.0029	No	Spearman	0.3405	0.1159	0.3681
10000 m DDW EV-MCDA range	0.0029	No	Spearman	0.3405	0.1159	0.3681
10000 m DDW EV-MCDA mean	0.3848	Yes	Pearson	0.4702	0.2211	0.2015
10000 m DDW EV-MCDA standard deviation	0.0031	No	Spearman	0.5126	0.2628	0.1614
10000 m DDW EV-MCDA median	0.1098	Yes	Pearson	0.5519	0.3046	0.1234
10000 m DDW EV-MCDA 90th percentile	0.0401	No	Spearman	0.3405	0.1159	0.3681
SPW OWTS count	**0.0859**	**Yes**	**Pearson**	**0.7630**	**0.5821**	**0.0168***
250 m DDW OWTS count			No variance			
500 m DDW OWTS count	<0.0001	No	Spearman	-0.5477	0.3000	0.2222
750 m DDW OWTS count	0.0077	No	Spearman	-0.6336	0.4014	0.0754
1000 m DDW OWTS count	0.0648	Yes	Pearson	-0.3693	0.1364	0.3280
2500 m DDW OWTS count	0.0002	No	Spearman	-0.0251	0.0006	0.9574
5000 m DDW OWTS count	0.0148	No	Spearman	0.3431	0.1177	0.3640
10000 m DDW OWTS count	0.0326	No	Spearman	0.4000	0.1600	0.2912
SPW OWTS density	0.1047	Yes	Pearson	0.5304	0.2813	0.1418
250 m DDW OWTS density			No variance			
500 m DDW OWTS density	<0.0001	No	Spearman	-0.5477	0.3000	0.2222
750 m DDW OWTS density	0.0009	No	Spearman	-0.5660	0.3204	0.1210
1000 m DDW OWTS density	0.0297	No	Spearman	-0.5509	0.3035	0.1272
2500 m DDW OWTS density	0.0008	No	Spearman	-0.1757	0.0309	0.6517
5000 m DDW OWTS density	0.0875	Yes	Pearson	0.2863	0.0820	0.4551
10000 m DDW OWTS density	0.1687	Yes	Pearson	0.4755	0.2261	0.1958
SPW open water	0.0001	No	Spearman	-0.5000	0.2500	0.1777
SPW developed open space	0.6374	Yes	Pearson	0.4510	0.2034	0.2230
SPW developed low intensity	0.0372	No	Spearman	0.0333	0.0011	0.9484
SPW developed medium intensity	0.0060	No	Spearman	0.0333	0.0011	0.9484
SPW developed high intensity	0.0001	No	Spearman	0.2678	0.0717	0.4828
SPW total developed	0.1815	Yes	Pearson	0.4270	0.1824	0.2516
SPW barren land	0.2852	Yes	Pearson	0.4707	0.2216	0.2009
SPW deciduous forest	0.6101	Yes	Pearson	0.2019	0.0408	0.6024
SPW evergreen forest	0.0442	No	Spearman	-0.4833	0.2336	0.1938
SPW mixed forest	0.3452	Yes	Pearson	-0.1891	0.0358	0.6260
SPW shrub scrub	0.3903	Yes	Pearson	0.0604	0.0037	0.8773
SPW herbaceous	0.0624	Yes	Pearson	-0.3636	0.1322	0.3361
SPW hay pasture	0.1490	Yes	Pearson	-0.2715	0.0737	0.4797
SPW cultivated crops	0.0872	Yes	Pearson	-0.2712	0.0736	0.4802
SPW woody wetlands	0.0277	No	Spearman	0.0167	0.0003	0.9733
SPW emergent herbaceous wetlands	0.1541	Yes	Pearson	-0.2131	0.0454	0.5819
250 m DDW open water	0.0002	No	Spearman	-0.5919	0.3503	0.0993
250 m DDW developed open space	0.0517	Yes	Pearson	0.3624	0.1313	0.3378
250 m DDW developed low intensity	0.0012	No	Spearman	0.3051	0.0931	0.4270

250 m DDW developed medium intensity			No variance			
250 m DDW developed high intensity	<0.0001	No	Spearman	-0.2510	0.0630	0.5556
250 m DDW total developed	0.0586	Yes	Pearson	0.0140	0.0002	0.9715
250 m DDW barren land			No variance			
250 m DDW deciduous forest	0.7586	Yes	Pearson	0.3747	0.1404	0.3204
250 m DDW evergreen forest	0.0086	No	Spearman	0.0174	0.0003	0.9717
250 m DDW mixed forest	0.0001	No	Spearman	0.0198	0.0004	0.9762
250 m DDW shrub scrub			No variance			
250 m DDW herbaceous	<0.0001	No	Spearman	-0.3651	0.1333	0.3611
250 m DDW hay pasture	0.4985	Yes	Pearson	0.0594	0.0035	0.8794
250 m DDW cultivated crops			No variance			
250 m DDW woody wetlands	0.0436	No	Spearman	0.2959	0.0876	0.4365
250 m DDW emergent herbaceous wetlands			No variance			
500 m DDW open water	**0.0266**	**No**	**Spearman**	**-0.7120**	**0.5069**	**0.0402***
500 m DDW developed open space	0.7085	Yes	Pearson	0.0027	0.0000	0.9945
500 m DDW developed low intensity	0.0271	No	Spearman	0.0333	0.0011	0.9484
500 m DDW developed medium intensity	0.0035	No	Spearman	0.2698	0.0728	0.4772
500 m DDW developed high intensity	0.0138	No	Spearman	0.2373	0.0563	0.5470
500 m DDW total developed	0.0773	Yes	Pearson	0.0103	0.0001	0.9791
500 m DDW barren land			No variance			
500 m DDW deciduous forest	0.8265	Yes	Pearson	0.1291	0.0167	0.7406
500 m DDW evergreen forest	0.0784	Yes	Pearson	-0.1934	0.0374	0.6180
500 m DDW mixed forest	0.2593	Yes	Pearson	0.0031	0.0000	0.9937
500 m DDW shrub scrub	0.0040	No	Spearman	0.1826	0.0333	0.6481
500 m DDW herbaceous	<0.0001	No	Spearman	-0.1741	0.0303	0.6521
500 m DDW hay pasture	0.1088	Yes	Pearson	-0.2041	0.0417	0.5984
500 m DDW cultivated crops			No variance			
500 m DDW woody wetlands	**0.0541**	**Yes**	**Pearson**	**0.7076**	**0.5007**	**0.0330***
500 m DDW emergent herbaceous wetlands			No variance			
750 m DDW open water	0.0219	No	Spearman	-0.6611	0.4371	0.0632
750 m DDW developed open space	0.8374	Yes	Pearson	-0.3110	0.0967	0.4152
750 m DDW developed low intensity	0.1797	Yes	Pearson	0.0180	0.0003	0.9634
750 m DDW developed medium intensity	0.0138	No	Spearman	0.0170	0.0003	0.9815
750 m DDW developed high intensity			No variance			
750 m DDW total developed	0.5501	Yes	Pearson	-0.1227	0.0151	0.7531
750 m DDW barren land			No variance			
750 m DDW deciduous forest	0.5162	Yes	Pearson	-0.1729	0.0299	0.6564
750 m DDW evergreen forest	0.0018	No	Spearman	-0.2259	0.0510	0.5570
750 m DDW mixed forest	0.4133	Yes	Pearson	-0.0378	0.0014	0.9232
750 m DDW shrub scrub	0.0154	No	Spearman	-0.2350	0.0552	0.5409
750 m DDW herbaceous	<0.0001	No	Spearman	-0.0848	0.0072	0.8410
750 m DDW hay pasture	0.7586	Yes	Pearson	0.1140	0.0130	0.7703
750 m DDW cultivated crops	<0.0001	No	Spearman	0.2739	0.0750	0.6667

750 m DDW woody wetlands	0.0005	No	Spearman	0.6441	0.4149	0.0700
750 m DDW emergent herbaceous wetlands	<0.0001	No	Spearman	0.5477	0.3000	0.2222
1000 m DDW open water	0.0146	No	Spearman	-0.5085	0.2586	0.1688
1000 m DDW developed open space	0.3877	Yes	Pearson	-0.2373	0.0563	0.5387
1000 m DDW developed low intensity	0.2560	Yes	Pearson	0.0198	0.0004	0.9598
1000 m DDW developed medium intensity			No variance			
1000 m DDW developed high intensity			No variance			
1000 m DDW total developed	0.3911	Yes	Pearson	-0.0601	0.0036	0.8780
1000 m DDW barren land			No variance			
1000 m DDW deciduous forest	0.4615	Yes	Pearson	-0.0342	0.0012	0.9305
1000 m DDW evergreen forest	0.0648	Yes	Pearson	-0.5105	0.2606	0.1602
1000 m DDW mixed forest	0.4688	Yes	Pearson	-0.2230	0.0497	0.5642
1000 m DDW shrub scrub	0.0418	No	Spearman	-0.2712	0.0735	0.4811
1000 m DDW herbaceous	<0.0001	No	Spearman	0.0753	0.0057	0.8525
1000 m DDW hay pasture	0.1570	Yes	Pearson	0.1388	0.0193	0.7217
1000 m DDW cultivated crops	<0.0001	No	Spearman	-0.1369	0.0187	0.7500
1000 m DDW woody wetlands	<0.0001	No	Spearman	0.6272	0.3934	0.0808
1000 m DDW emergent herbaceous wetlands	**<0.0001**	**No**	**Spearman**	**0.7303**	**0.5333**	**0.0278***
2500 m DDW open water	0.1151	Yes	Pearson	-0.5515	0.3041	0.1238
2500 m DDW developed open space	0.0691	Yes	Pearson	0.0637	0.0041	0.8707
2500 m DDW developed low intensity	0.1878	Yes	Pearson	-0.0908	0.0082	0.8163
2500 m DDW developed medium intensity	0.0018	No	Spearman	0.1667	0.0278	0.6777
2500 m DDW developed high intensity	0.0004	No	Spearman	-0.0527	0.0028	0.9190
2500 m DDW total developed	0.0389	No	Spearman	-0.0500	0.0025	0.9116
2500 m DDW barren land	0.0550	Yes	Pearson	0.5603	0.3140	0.1166
2500 m DDW deciduous forest	0.3206	Yes	Pearson	-0.2878	0.0828	0.4527
2500 m DDW evergreen forest	0.6462	Yes	Pearson	-0.2576	0.0664	0.5034
2500 m DDW mixed forest	0.6988	Yes	Pearson	0.4062	0.1650	0.2779
2500 m DDW shrub scrub	0.0082	No	Spearman	-0.2259	0.0510	0.5570
2500 m DDW herbaceous	0.3172	Yes	Pearson	0.6532	0.4267	0.0564
2500 m DDW hay pasture	0.3849	Yes	Pearson	0.0464	0.0022	0.9056
2500 m DDW cultivated crops	0.0010	No	Spearman	0.0183	0.0003	0.9854
2500 m DDW woody wetlands	0.0033	No	Spearman	0.3390	0.1149	0.3733
2500 m DDW emergent herbaceous wetlands	0.0123	No	Spearman	0.4564	0.2083	0.2275
5000 m DDW open water	0.1027	Yes	Pearson	-0.5157	0.2660	0.1553
5000 m DDW developed open space	0.5858	Yes	Pearson	0.3716	0.1381	0.3248
5000 m DDW developed low intensity	0.0648	Yes	Pearson	-0.1567	0.0246	0.6873
5000 m DDW developed medium intensity	0.0661	Yes	Pearson	0.1756	0.0308	0.6514
5000 m DDW developed high intensity	<0.0001	No	Spearman	0.1261	0.0159	0.7482
5000 m DDW total developed	0.7531	Yes	Pearson	0.2045	0.0418	0.5977
5000 m DDW barren land	0.0723	Yes	Pearson	0.2107	0.0444	0.5864
5000 m DDW deciduous forest	0.2975	Yes	Pearson	-0.1070	0.0115	0.7840
5000 m DDW evergreen forest	0.3312	Yes	Pearson	-0.3849	0.1482	0.3063

5000 m DDW mixed forest	0.9267	Yes	Pearson	0.4686	0.2196	0.2033
5000 m DDW shrub scrub	0.0215	No	Spearman	0.0667	0.0044	0.8801
5000 m DDW herbaceous	0.3821	Yes	Pearson	0.4334	0.1878	0.2439
5000 m DDW hay pasture	0.1996	Yes	Pearson	-0.0450	0.0020	0.9084
5000 m DDW cultivated crops	0.0022	No	Spearman	0.1044	0.0109	0.7893
5000 m DDW woody wetlands	0.0025	No	Spearman	0.3347	0.1120	0.3738
5000 m DDW emergent herbaceous wetlands	0.0140	No	Spearman	0.4178	0.1746	0.2618
10000 m DDW open water	0.0156	No	Spearman	-0.6103	0.3725	0.0890
10000 m DDW developed open space	0.6661	Yes	Pearson	0.4719	0.2226	0.1997
10000 m DDW developed low intensity	0.4968	Yes	Pearson	0.0827	0.0068	0.8324
10000 m DDW developed medium intensity	0.1866	Yes	Pearson	0.1010	0.0102	0.7960
10000 m DDW developed high intensity	0.0066	No	Spearman	0.0335	0.0011	0.9370
10000 m DDW total developed	0.9413	Yes	Pearson	0.2993	0.0896	0.4340
10000 m DDW barren land	0.0001	No	Spearman	0.5714	0.3265	0.1141
10000 m DDW deciduous forest	0.6399	Yes	Pearson	-0.0031	0.0000	0.9936
10000 m DDW evergreen forest	0.4335	Yes	Pearson	-0.6519	0.4249	0.0571
10000 m DDW mixed forest	0.5678	Yes	Pearson	-0.0335	0.0011	0.9319
10000 m DDW shrub scrub	0.0118	No	Spearman	0.3667	0.1345	0.3363
10000 m DDW herbaceous	0.6783	Yes	Pearson	-0.0336	0.0011	0.9316
10000 m DDW hay pasture	0.5714	Yes	Pearson	0.0961	0.0092	0.8058
10000 m DDW cultivated crops	0.0550	Yes	Pearson	-0.0454	0.0021	0.9077
10000 m DDW woody wetlands	0.0543	Yes	Pearson	0.0928	0.0086	0.8123
10000 m DDW emergent herbaceous wetlands	0.0181	No	Spearman	0.4854	0.2356	0.1871

***** Significant correlation *($\alpha < 0.05$)*
SPW – Sample point watershed
DDW – Distance derived watershed
ESV – Environmental and system variables
EV – Environmental variables
MCDA – multi-criteria decision analysis
OWTS – Onsite wastewater treatments systems