2018

Understanding The Impact Of Land Use On Microbial Water Quality To Support Decisions For A Future Land Use Plan

Mohammed Tawfiq Aljassim
University of South Carolina

UNDERSTANDING THE IMPACT OF LAND USE ON MICROBIAL WATER QUALITY TO SUPPORT DECISIONS FOR A FUTURE LAND USE PLAN

By

Mohammed Tawfiq Aljassim

Bachelor of Science
King Faisal University, 2008

Master of Science
University of Findlay, 2012

Submitted in Partial Fulfillment of the Requirements

For the Degree of Doctor of Philosophy in

Environmental Health Sciences

The Norman J. Arnold School of Public Health

University of South Carolina

2018

Accepted by:

Dwayne E. Porter, Major Professor

Geoff Scott, Committee Member

Robin W. Kloot, Committee Member

Bo Cai, Committee Member

Cheryl L. Addy, Vice Provost and Dean of the Graduate School

DEDICATION

I dedicate this dissertation to my parents, my wife and my whole family, and to my country. I dedicate this work to Imam Abdulrahman Bin Faisal University.

ACKNOWLEDGEMENTS

As this dissertation is accomplished with supports from several people and institutions, I would love to thank Imam Abdulrahman Bin Faisal University for its full support to go overseas and accomplishing this work. I would also love to thank my parents for their prayers and their supportive inspirations. I will always thank my wife for standing with me overseas and for her infinite support. From the bottom of my heart, my thanks go to my committee members Dr. Dwayne E. Porter, Dr. Geoff Scott, Dr. Robin (Buz) Kloot, and Dr. Bo Cai. I should also thank Mr. Paul Conrads, a committee member who passed away in December of 2017. I would also like to thank the South Carolina Department of Health and Environmental Control and the government of the Town of Bluffton for providing me with my needs for this dissertation. My endless thanks go to my friends in the United States and Saudi Arabia for their supportive words and prayers.

From the bottom of my heart, I thank you all.

ABSTRACT

May River meandering through the Town of Bluffton, South Carolina has been known for its high level of water quality. The May River is one of the major sources of shellfish harvesting in South Carolina. Due to the non-point source pollution of fecal coliform, the water quality of the May River became degraded. As a result, shellfish harvesting classification in the May River was downgraded for the first time in its history. The rapid growth of the Town has led to a change in land use, which might be related to the increased levels of fecal coliform in the stream of the river. Furthermore, future land use developments are planned, and continues population growth is anticipated in the Town. Therefore, this research attempts to support the decision related to the implementation of a proposed land use plan by understanding the impact of the recent land use change on microbial water quality. Geographic Information System (GIS) technology was integrated with statistical analysis to assess the spatial relationships between different land use types and fecal coliform levels in the May River. This research found that residential areas, forestlands and golf courses are significantly correlated with fecal coliform. Geographically Weighted Regression (GWR) was used to examine the spatially varying relationships between specific land use types and fecal coliform concentrations among the sampling locations. Predictive models were developed to predict fecal coliform concentrations by including data of land use types and meteorological and environmental factors. In order to determine the optimal spatial scale for the land use variables, several circular buffer sizes were developed and examined for their appropriateness in supporting significant models for fecal coliform prediction. It was shown that land use percentages within 1800 meters radius were most significantly correlated with fecal coliform. Rainfall measurements, water temperature, air temperature, salinity, and tide stage in addition to land use classes (residential areas, forestlands, and open spaces) within the 1800 meters radius were able to provide the most significant models for fecal coliform prediction. In order to predict the impact of the future land use developments plan, two rainfall scenarios (average and maximum precipitation) were used in the predictive models. The findings of this research indicated that the future land use plan will not lead to higher fecal coliform loadings than the current land use.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AICc .... Akaike Information Criterion

ANN .... Artificial Neural Network

BOD .... Biological Oxygen Demand

CV .... Cross Validation

EPA .... Environmental Protection Agency

ERT .... Exploratory Regression Tool

ESRI .... Environmental Systems Research Institute

FC .... Fecal Coliform bacteria

GIS .... Geographic Information Systems

GWL .... Geographically Weighted Lasso

GWR .... Geographically Weighted Regression

GWRR .... Geographically Weighted Ridge Regression

HUC .... Hydrologic Unit Code

LC .... Land Cover

LMM .... Linear Mixed Model

LU .... Land Use

MLR .... Multiple Linear Regression

NEXRAD .... Next Generation Weather Radar

NOAA .... National Oceanic and Atmospheric Administration

NSSP .... National Shellfish Sanitation Program

OLS .... Ordinary Least Squares

ORW .... Outstanding Resource Water

PSD .... Public Service District

SA .... Spatial Autocorrelation

SCDHEC....................... South Carolina Department of Health and Environmental Control

SMLR.. Stepwise Multiple Linear Regression

SWAT .... Soil and Water Assessment Tool

TSS.. Total suspended solids

USFDA .... The United States Food and Drug Administration

VIF .... Variance Inflation Factor

CHAPTER 1

INTRODUCTION

1.1 BACKGROUND

1.1.1 What are land use and land cover?

Land use (LU) and land cover (LC) are the topographies of a specific region, country or even an area of a wider extent. The LU is a man-made activity, whereas LC is the nature's cover of the land. The LC can be forestlands, wetlands, and many other natural covers of land as well as bodies of water (US Department of Commerce, 2015). The LU tells us how the land is used by people; these uses can be human developments like residential, industrial, and commercial and transportation area (MRLC, 2016). Human activities have altered a large fraction of the earth's surface due to the transformations of natural landscapes for different human uses (Foley et al., 2005). These human activities can be deforestation, agriculture, farming productions, and urbanization, and these activities may impact the landscape's interaction with surface water (Foley et al., 2005).

Coastal and land managers look at the LU and LC maps to evaluate the past and the future developments in terms of their impacts on the environment. Furthermore, scholars and scientists use LU and LC change information to link their relationships to air and water quality as well as many other ecological impacts. Many federal, state and local agencies use LU and LC data for several proposes such as flood control, wastewater treatment planning, and water resources inventory (Anderson, 1976). The data of LU and LC are also needed by the federal agencies for environmental impact assessments and wildlife resources management to link the impact of human activities, like habitat fragmentation, to their impacts on wildlife resources and environmental quality (Anderson, 1976). LU and LC changes have been found to be related to the change of the global climate by altering the spatial and temporal patterns of thunderstorms (Pielke, 2005). Therefore, LU and LC are critical elements to many environmental studies, which became exceedingly valuable to many researchers and decision makers.

1.1.3 The relationship between LU/LC and water quality

Watersheds receive and transport waters that pass through their lands. Every watershed has different components of LU and LC which interact with the precipitation water that eventually drain into a waterbody's receiving stream. The interaction mechanisms between the different LU and LC and the water may degrade the quality of the drained runoff waters. Consequently, the waterbody will be loaded with certain types of pollutants that existed within the LU and LC. For example, wetland is a natural LC which may contribute to minimum pollutants in the runoff as it is not commonly known as a major source of pollutants. Conversely, agricultural LU will have different sources of pollutants in its runoff due to the presence of fertilizers and pesticides. Urban areas will have high percentages of developed lands and imperviousness. Impervious areas may negatively impact the quantity of the runoff water as they may reduce or prevent the water infiltration through the soil. Consequently, part of this runoff will be controlled by some storm water management practice while some other parts will be drained into the water bodies carrying different types of pollutants. Therefore, the relationship between land use and water quality must be considered to maintain water quality.

Ecologists environmentalists, and natural resources managers and policy makers acknowledge that the past LU activities can continue to influence the structure and the function of the ecosystem in the future (Foster et al., 2003). Consequently, historical LU changes and their impacts to the ecosystem have expanded the understandings of why the ecosystem has been influenced and how LU activities can be improved in the future to reduce the environmental impacts (Foster et al., 2003). A study found that changes in LU and LC in a global extent will result in a significant impact to our earth system functioning (Lambin et al., 2003). In addition, one of the main contributors to climate change is the changes in LU and LC (Searchinger et al., 2008). It was found that biodiversity is directly impacted by LU and LC changes globally (Hansen, DeFries, & Turner, 2012). Soil degradation is also one of the major consequences of LU and LC change activities (Tolba et al., 1992). Many other recent studies focused on the relationships between LU and LC and other impacts such as landslide (Persichillo, Bordoni, & Meisina, 2017), infectious diseases (Patz & Olson, 2017), and atmospheric mercury (Zhang, Holmes, & Wu, 2016).

It is very important to understand how water quality is sensitive to LU change. It is also important to link land management and land use planning to water resources management. With the existence of the advanced technologies such as remote sensing and Geographic Information Systems (GIS), it is possible to detect LU changes, model the relationship between LU changes and water quality, and predict the spatial and temporal impacts of LU change on the quality of the water. Therefore, many studies have assessed the relationships between LU change and different water quality parameters such as nutrients, Fecal Coliform bacteria (FC), specific conductivity (SC), water temperature, salinity, turbidity, and many other parameters depending on the study objectives and data availability.

1.2 STATEMENT OF THE PROBLEM

The United States Food and Drug Administration (USFDA) follows the guidelines of the National Shellfish Sanitation Program (NSSP) in evaluating the states' shellfish sanitation programs. These guidelines require that each shellfish harvesting area be surveyed for its human consumption conditions (classifications); they can be classified as Approved, Conditionally Approved, Restricted, or Prohibited. In 2009, the South Carolina Department of Health and Environmental Control (SCDHEC) reported that there are five shellfish harvesting stations that were classified as a "Restricted Area" for human consumption in the Town of Bluffton, SC due to elevated levels of FC at the five stations. A most recent report of the SCDHEC in 2016 stated the same problem within the five stations (Moody, 2016). The pollution source was found to be a non-point source due to storm water runoff (Moody, 2016). The SCDHEC is authorized to prohibit shellfish harvesting when there are unsafe conditions for human consumption. Therefore, it is important to know why these five stations often have elevated levels of FC, and to know how the FC levels are sensitive to LU activities in May River.

1.3 STUDY OBJECTIVE AND HYPOTHESIS

This dissertation aims to assess and model the relationship between LU and water quality by focusing on FC as a main indicator of microbial non-point source water pollutant. This research will assess and model the relationship between the LU types and the levels of FC bacteria at the Town of Bluffton, South Carolina. This research will integrate spatial and statistical analysis by using GIS and statistical methodologies. The following questions will be answered by this dissertation research: (1) What are the spatial relationships between the LU types and FC bacteria? (2) How do the relationships between LU types and FC bacteria vary locally between the shellfish monitoring stations? (3) What are the most applicable and suitable LU and environmental predictors for FC in the May River? (4) How are FC bacteria predictable for a future land use change scenario?

Only one hypothesis will be tested in this dissertation, and it is for the first question. The null hypothesis is that the type, amount, and location of land use within a watershed does not impact the concentrations of fecal coliform.

By answering these questions, testing the hypothesis, and accomplishing this work, this dissertation will contribute in adding valuable knowledge to the Town of Bluffton government and its water resource management efforts. This study will be one of very few studies that include the LU for FC prediction and modeling. Additionally, this research will add further findings to the spatial statistical analysis.

1.4 STUDY AREA

The study area is the May River watershed which is located at Beaufort County in South Carolina (Figure 1.1). The Town of Bluffton has the May River as its significant estuary that drains into the Atlantic Ocean in the eastern coast of South Carolina. The May River watershed is encompassed in a 12-digit hydrologic unit code (HUC) and its size is about 10,353 hectares (40 square miles) (Van Dolah, Sanger, & Filipowics, 2004). The LU types in the May River watersheds are as the followings: residential areas, non-residential areas, forestlands, forested wetlands, non-forested wetlands, civic, commercial, golf courses, open spaces, open water, transportations, and undeveloped areas. The population of the Town of Bluffton in the year 2015 was about 16,728 (U.S Census Bureau, 2015).

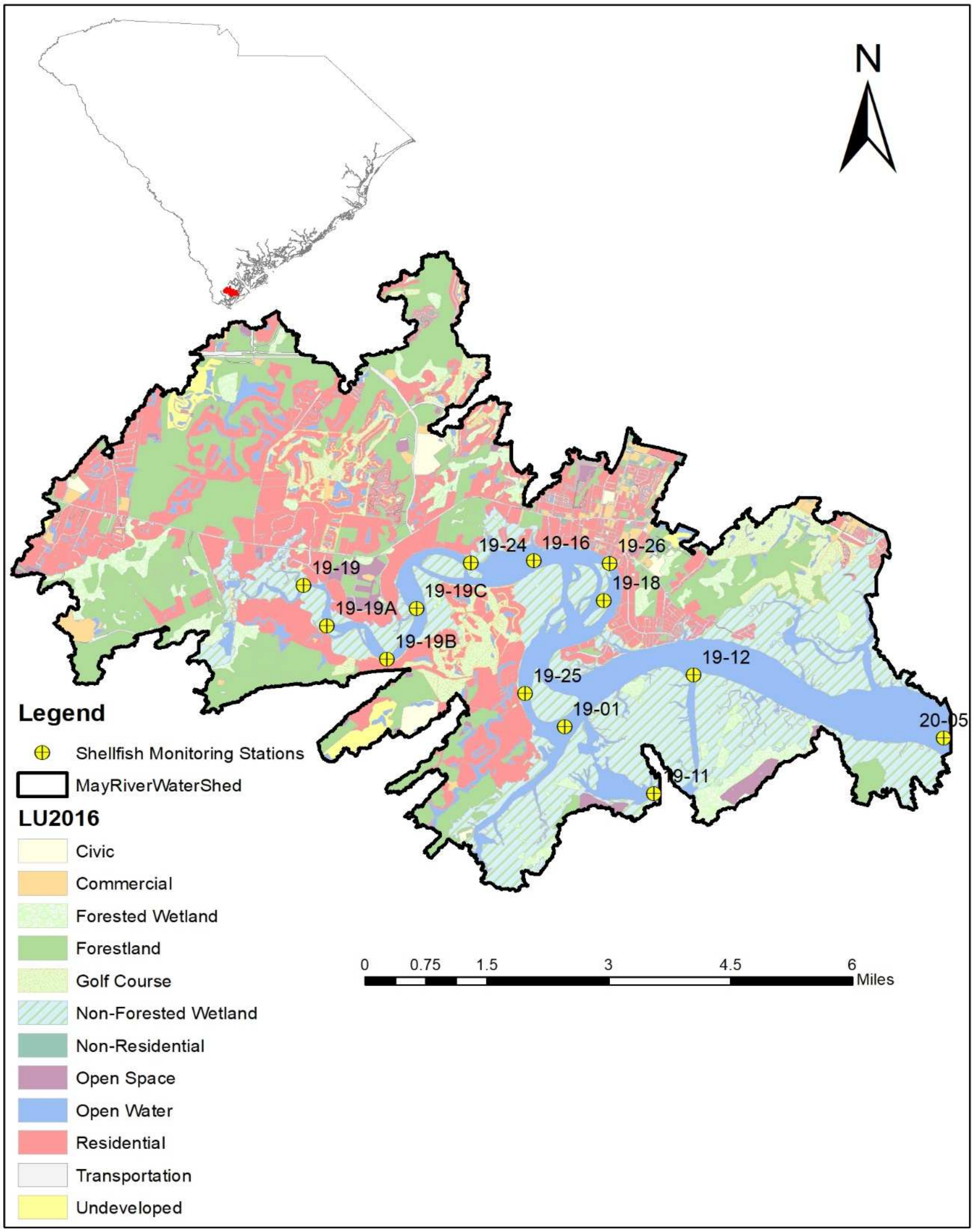

Figure 1.1: May River watershed with sample sites and LU classes in 2016

Due to the water quality of the May River, in 2001, it was considered as an Outstanding Resource Water (ORW) by the SCDHEC (Barber, 2008). The May River has been identified as a priority watershed by the Environmental Protection Agency (EPA) and SCDHEC. The May River is significantly valued by the residents since it has a great recreational and economical importance. The aesthetics and views, and the abundant natural resources have led to an increase in population and commercial growth in the area. As a result, major LU changes are planned for developments by the Town of Bluffton's government; thus, there are many concerns about the subsequent adverse impacts on water quality of May River. In addition, oyster production is the main economic activity in the May River and polluting this river possibly will impact the quality of the oyster beds and the harvestability of human consumption.

The Town of Bluffton has grown rapidly in the past few decades and this growth is expanding. In 1852, the area of the Town was about 640 acres (1 square mile), and in 1987, many annexations were approved by the Town's government until the area of the town increased to 34,560 acres (54 square miles) in 2015 to become on the top five largest municipality by land area in South Carolina (Town of Bluffton, 2015). The population size in 1990 was 713 people, which in 2005 increased to 4,885 people, which then increased to 12,893 people in 2010 to reach to 16,728 people in 2015 (Town of Bluffton, 2015). The estimated build-out projection by the Town of Bluffton's government is about 70,000 people (Town of Bluffton, 2015).

There were LU changes as the town was developing and growing in its population and commercial activities, which can impact the quality of the May River. The LU in the watershed is majorly residential, with some minor commercial uses, and there are no heavy industrial activities within the watershed. However, the May River for its first time in the history was downgraded for its shellfish harvesting classification in 2009. Microbial contamination of FC bacteria was and is still the main concerned pollutant in the river. The May River watershed is within the Shellfish Growing Area 19 (Figure 1-2). Five shellfish monitoring stations within the shellfish growing area 19 are classified as A “Restricted Area” in 2016 (Moody, 2016). These stations (Figure 1-2) are 19-19, 19-19A, 19-19B, 19-19C, and 19-24.

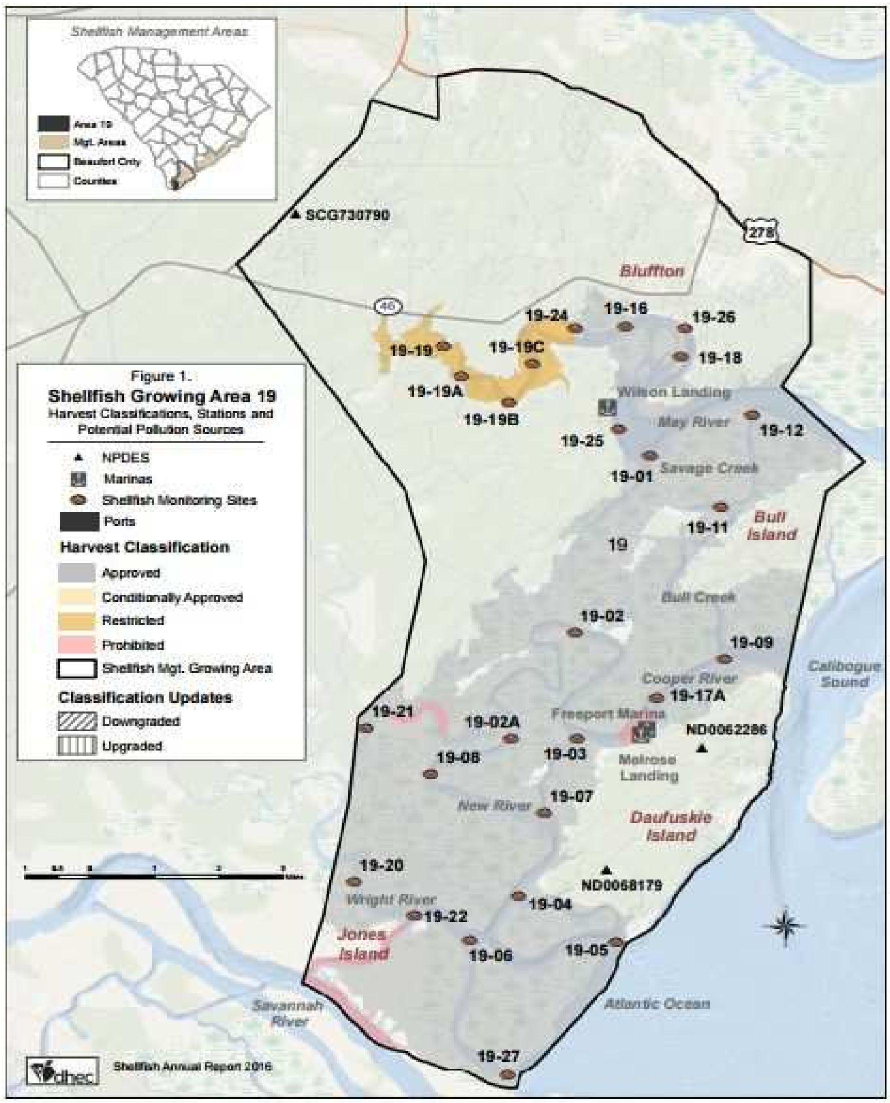

Figure 1.2: Shellfish Growing Area 19 with SCDHEC shellfish monitoring stations
Source: SCDHEC

The following chapters will go through different forms of analysis and assessments to answer the dissertation questions. Chapter 2 will assess the spatial relationships between

LC types and FC by using GIS and spatial statistical methods. Chapter 3 will model the relationships between LU types and FC by integrating environmental and meteorological factors. Chapter 4 will apply the models that are developed in Chapter 3 to predict the future impact of an alternative LU change scenario on FC at the specified study area.

REFERENCES

Anderson, J. R. (1976). *A land use and land cover classification system for use with remote sensor data* (Vol. 964). US Government Printing Office.

Barber, B. P. (2008). Town of Bluffton May River watershed monitoring program. *Prepared for Town of Bluffton, Bluffton, SC. 143pp*.

Foley, J. A., DeFries, R., Asner, G. P., Barford, C., Bonan, G., Carpenter, S. R., … others. (2005). Global consequences of land use. *Science*, *309*(5734), 570–574.

Foster, D., Swanson, F., Aber, J., Burke, I., Brokaw, N., Tilman, D., & Knapp, A. (2003). The importance of land-use legacies to ecology and conservation. *BioScience*, *53*(1), 77–88.

Hansen, A. J., DeFries, R. S., & Turner, W. (2012). Land use change and biodiversity. In *Land Change Science* (pp. 277–299). Springer. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4020-2562-4_16

Lambin, E. F., Turner, B. L., Geist, H. J., Agbola, S. J., Angelsen, A., Bruce, J. W., … Folke, C. (2003). The causes of land-use and land-cover change: moving beyond the myths. *Global Environmental Change*, *11*(4). Retrieved from http://www.cifor.org/online-library/browse/view-publication/publication/1329&pf=1.html

Moody, R. (2016). Shellfish Management Area 19-2016 Annual Update. Submitted by the Shellfish Sanitation Section, Water Monitoring, Assessment and Protection Division, Environmental Quality Control–Bureau of Water, South Carolina

Department of Health and Environmental Control, Columbia, SC. Retrieved from http://scdhec.gov/FoodSafety/Docs/sfrp_19.pdf

Multi-Resolution Land Characteristics Consortium (MRLC). (2016). National Land Cover Database (NLCD). Retrieved January 21, 2017, from https://www.mrlc.gov/nlcd92_leg.php

Patz, J. A., & Olson, S. H. (2017). Land Use/Land Change and Health. In S. R. Quah (Ed.), *International Encyclopedia of Public Health (Second Edition)* (pp. 358–363).Oxford: Academic Press. Retrieved from //www.sciencedirect.com/science/article/pii/B9780128036785002460

Persichillo, M. G., Bordoni, M., & Meisina, C. (2017). The role of land use changes in the distribution of shallow landslides. *Science of The Total Environment*, *574*, 924–937. https://doi.org/10.1016/j.scitotenv.2016.09.125

Pielke, R. A. (2005). Land use and climate change. *Science*, *310*(5754), 1625–1626.

Searchinger, T., Heimlich, R., Houghton, R. A., Dong, F., Elobeid, A., Fabiosa, J., … Yu, T.-H. (2008). Use of US croplands for biofuels increases greenhouse gases through emissions from land-use change. *Science*, *319*(5867), 1238–1240.

Tolba, M. K., El-Kholy, O. A., El-Hinnawi, E., Holdgate, M. W., McMichael, D. F., & Munn, R. E. (1992). *The World Environment, 1972–92*. Chapman and Hall, London.

Town of Bluffton. (2015). Growth Framework. Retrieved from https://www.townofbluffton.sc.gov/Documents/blufftoninnovation.pdf

U.S Census Bureau. (2015). 2015 Population Estimate for Town of Bluffton. Retrieved January 26, 2017, from //www.census.gov/quickfacts/table/PST045215/4507210

US Department of Commerce (2015). What is the difference between land cover and land use? Retrieved January20, 2017, from http://oceanservice.noaa.gov/facts/lclu.html

Van Dolah, R. F., Sanger, D. M., & Filipowics, A. B. (2004). A baseline assessment of environmental and biological conditions in the May River, Beaufort County, South Carolina. *Final Report to the Town of Bluffton, SC. Submitted by the Marine Resources Division, South Carolina Department of Natural Resources, Charleston, SC, and the Center of Coastal Environmental Health and Biomolecular Research, National Ocean Service, NOAA, Charleston, SC*.

Zhang, H., Holmes, C. D., & Wu, S. (2016). Impacts of changes in climate, land use and land cover on atmospheric mercury. *Atmospheric Environment*, *141*, 230–244

CHAPTER 2

ASSESSING LAND USE IMPACT ON MICROBIAL WATER QUALITY

ABSTRACT

Land use (LU) is widely known as one of the factors that has a direct impact on the hydrology of a watershed. Some land uses can negatively impact our water resources, which subsequently impact the health of the living organisms and human beings. This research assesses the spatial relationships between different types of LU and microbial water quality to see how the different types of LU influence the level of fecal coliform (FC) bacteria concentrations in the May River in the Town of Bluffton. Pearson product-moment correlation was used to explore the spatial relationships between the selected LU and FC. Geographically Weighted Regression (GWR) analysis was used to examine the variations of the spatial non-stationary relationships between LU and FC. An Exploratory Regression Tool (ERT) developed by the Environmental Systems Research Institute (ESRI) was used in determining the ideal combination of explanatory variables for the GWR models. The study employed a Geographic Information System (GIS) for spatial analysis and for mapping the spatial outcomes for aspects of the study. The data for this research included LU data from the government of the Town of Bluffton and FC data from the South Carolina Department of Health and Environmental Control (SCDHEC). Residential areas and Forestlands are the main land use components in the study area and were found to have significant positive relationships with FC. Even with the small size of the study area, obvious variabilities of spatial non-stationary relationships were explored

by the GWR. The ERT was found to be a substantial tool for saving time and effort and was very supportive in exploring the LU variables for GWR models. The findings of this study suggest that the GWR can be used for water quality assessments and watershed management research.

2.1 INTRODUCTION

There is nothing more important than ensuring access to clean and safe water resources for the community. The water cycle is the natural phenomenon that makes water continuously available in our environment. Through this cycle, storm water runs through different land covers (LC) and LU. The LC is the natural cover of the land, such as forestland or open water, and the LU is the man-made cover such as roads, building and lakes (Cihlar & Jansen, 2001). The LU can be defined as the manner in which LC is used by humans (Cihlar & Jansen, 2001). In fact, the chemical, physical, and biological properties of water can be changed due to the interaction of different environmental factors with LU or LC. Alteration to these properties may result in polluted water. For example, LU such as an urban-developed can be a main cause for surface water pollution, as it increases the impervious surface which will increase the volume of the storm water runoff (Wilson & Weng, 2010). Accordingly, different water pollutants may be related to the different LU and LC types.

Many studies were conducted to investigate the relationships between LU and water pollutants (BenDor, Jordanova, & Miles, 2017; Cha, Park, Lee, Kim, & Cho, 2016; Chen et al., 2016; Deason, Seekamp, & Barbieri, 2014; Dheenan et al., 2016; Kelsey, Porter, Scott, Neet, & White, 2004; Li, Zhao, Miaomiao, & Wang, 2010; Mallin, Williams, Esham, & Lowe, 2000; Paule-Mercado et al., 2016; Pettus, Foster, & Pan, 2015; Reano, Haver, Oki, & Yates, 2015; Schoonover & Lockaby, 2006; Su, Xiao, & Zhang, 2012; Sun, Guo, Liu, & Wang, 2014; Tu & Xia, 2008; Wilson & Weng, 2010). These studies found that there is a relationship between LU types and water quality parameters. The common statistical methods used in these studies were correlation analysis, and Ordinary Least Squares (OLS). Furthermore, most of these studies used simple statistical methods since they are simpler and more robust than developing complex hydrological models (Wan et al., 2014). Practically, LU percentages are used as independent variables, and water quality parameters from each sampling station are used as dependent variables when a regression or correlation analysis is used to study the relationship between LU and water quality.

Although correlation analysis and OLS are simple and robust statistical methods and are widely used in evaluating the relationship between dependent and independent variables in different disciplines, they cannot explain the complex relationships between LU and water quality parameters by using these regression methods (Giri & Qiu, 2016). However, GWR can illuminate the complex relationships, because it is integrated with GIS, which implicates the coordinate system into the regression model to make it capable of examining the spatial variations between the dependent and independent variables (Giri & Qiu, 2016)**.** The GWR is a spatial statistical analysis method that is extended form the OLS which estimates local regression coefficients instead of the global ones (Fotheringham, Charlton, & Brunsdon, 2001). The GWR is therefore a local statistical method because it explains the variations of a relationship over the space by allowing the local regression coefficients and local coefficient of determination (R^2) to change over space which make it possible to explain the causes of spatial patterns (Fotheringham, Brunsdon & Charlton, 2003). Therefore, unlike the OLS, which is a stationary spatial method, the GWR is a non-stationary spatial method and considers the positions of the dependent variables in the analysis. Furthermore, the GWR can detect the autocorrelation in the model (Fotheringham et al., 2003), and this gives it an additional advantage over the OLS method. The spatial autocorrelation can happen when values of a certain variable in more than one location are neighboring and have correlated values (Zhang, Bi, Cheng, & Davis, 2004).

Tu and Xia (2008) were the first to apply the GWR to study the relationship between LU and water quality. They used the GWR to examine the spatially varying relationships between LU and selected water quality parameters. They found that the GWR gives better predictions than the OLS because the GWR explains the spatial variations between the different LU percentages and water quality parameters. Some other studies also found that GWR was more robust than OLS and accurately interpreted the relationship between urban landscape patterns and water quality (Sun et al., 2014), and the relationship between LU and water quality (Yu, Shi, Liu, & Xun, 2013). The GWR was also used in other research purposes such as urbanization impacts on agricultural landscape (Su et al., 2012), the impact of environmental factors on land surface temperature (Li et al., 2010), air temperature and its influence factors (Ivajnšič, Kaligarič, & Žiberna, 2014), social life and HIV analysis (Wabiri, Shisana, Zuma, & Freeman, 2016), and air pollutants and human health (You et al., 2015).

Using the GWR, previous studies often use a single LU indicator as a dependent variable when examining the relationship with water quality parameters (Brown et al., 2012; (Huang, Huang, Pontius, & Zhang, 2015; Tu, 2013; Tu & Xia, 2008). This technique has been applied to avoid multicollinearity among LU variables. The LU variables are usually correlated, and consequently, collinearity will occur between them. With the presence of multicollinearity, the GWR model will not be able to predict the relationship between the studied dependent and independent variables. However, using this univariate method may hide some important relationships by missing important explanatory variables in the model. This is because that more than one LU type can exist in a sub-watershed. Actually, the variations of a water quality parameters' levels will not be completely explained if we, for example, only observed agriculture LU percentage in a sub-watershed while the main cause of water pollution is the runoff from high impervious surface dominated by urban developed LU (Chen et al., 2016).

The issue of multicollinearity in multivariate GWR modelling has been investigated by Wheeler and Tiefelsdorf (2005). They found that the GWR lacks well-established diagnostic tools that are used in standard global regression analysis like the OLS. In 2007, Wheeler developed Geographically Weighted Ridge Regression (GWRR) to overcome the multicollinearity problem when multivariate variables are used in the GWR model (Wheeler, 2007). Furthermore, (Wheeler, 2009) introduced a penalized form of GWR, Geographically Weighted Lasso (GWL), which limits the correlation effects of the independent variables by restraining the magnitude of the regression coefficients. Using the spatially multivariate technique will help to account for non-point source pollution by adding more related variables in the analysis (Su et al., 2011). However, very few studies used the multivariate GWR model to examine the effects of LU as well as other factors on surface water quality. A study used Stepwise Multiple Linear Regression (SMLR) to find the explanatory variables, including LU variables, that are strongly correlated with water quality variables to be then used in the GWR model (Pratt & Chang, 2012). The most recent study (Chen et al., 2016) used a manual variable excluding-selecting method and compared it with the SMLR that Pratt and Chang (2012) used, and found that the manual variable excluding-selecting method was better in finding the most predictive and significant models.

However, there is a data mining tool called Exploratory Regression (ERT) that was developed by the ESRI team (Rosenshein, Scott, & Pratt, 2011). This tool evaluates all possible combinations of independent variables and chooses only models that successfully pass the required OLS assumptions. The OLS assumptions are: (a) Linear relationship between the dependent and independent variables, (b) the mean of the error term should be zero, (c) observations should be randomly sampled (d) no multicollinearity between the independent variables, (e) the errors should be uncorrelated, (f) the errors should be normally distributed (Montgomery, Peck, & Vining, 2012). Furthermore, the concept of the ERT is similar to SMLR, but it also has a function that automatically selects the variables that provide models that meet all the assumptions required by the OLS method. This is an advantage that will help in finding the best GWR models. On this basis, the ERT is used in this study in order to see if it is possible to develop multivariate GWR models to assess the relationships. The objective of this study is to assess both the spatial relationships and the spatially varying relationships between selected LU types and FC levels in the adjacent May River in the Town of Bluffton, South Carolina, and to see if it is possible to apply multivariate GWR models.

2.2 STUDY AREA

The May River watershed is located in Beaufort County in the southeastern part of South Carolina (Figure 2.1). The Town of Bluffton has the May River as its significant estuary, which drains into the Atlantic Ocean on the eastern coast of South Carolina. The population of the Town of Bluffton in the year 2015 was about 16,728 (U.S Census Bureau, 2015). Due to the high water quality of the May River in 2001, it was classified Outstanding Resource Water (ORW) by the SCDHEC (Barber, 2008). The Environmental Protection Agency (EPA) and SCDHEC have acknowledged the May River as a priority watershed. The May River is highly valued by the residents of the Town of Bluffton due to its great

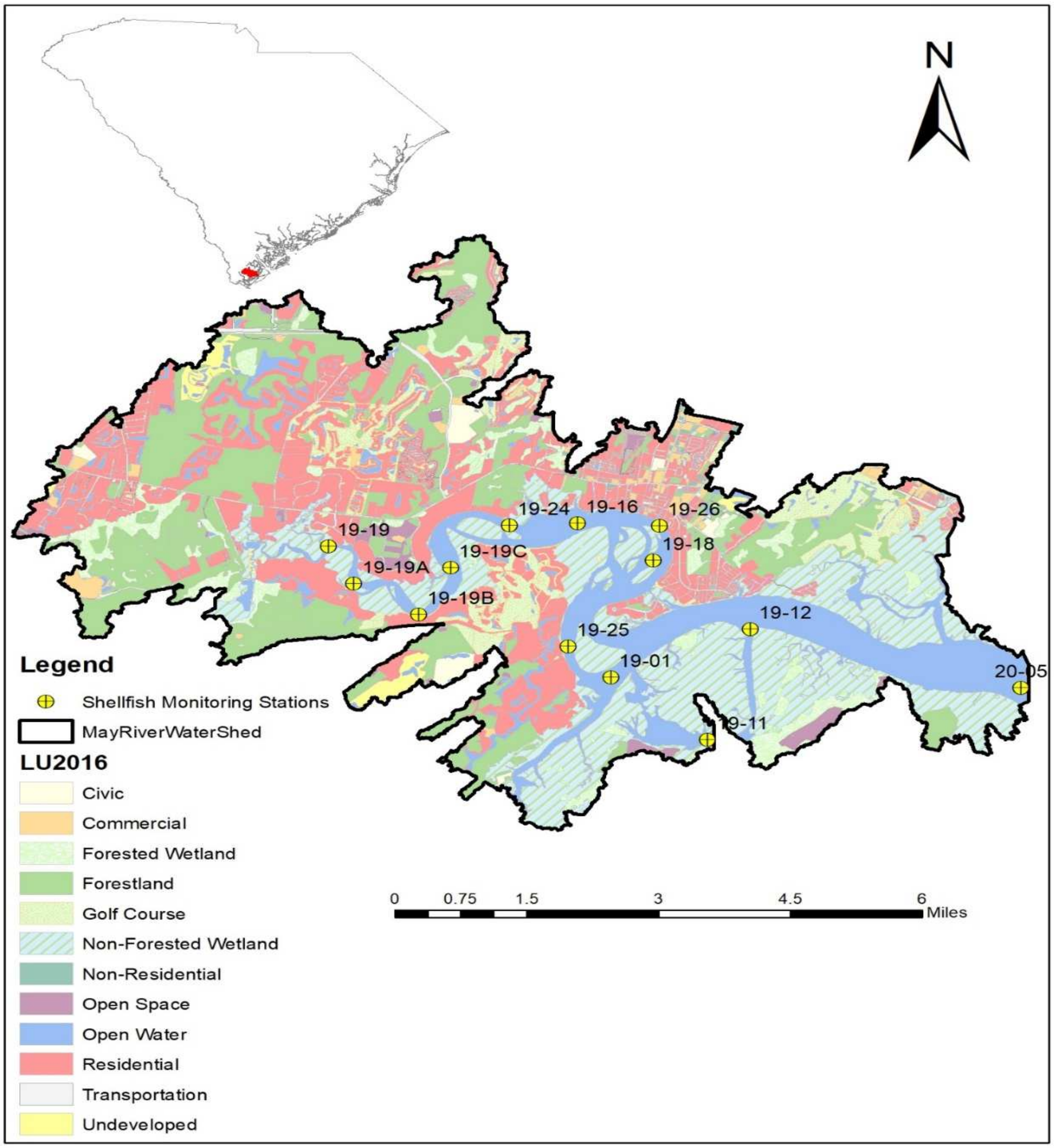

Figure 2.1: May River watershed with shellfish monitoring station sites and LU classes in 2016. Source: The government of Town of Bluffton

recreational and economical importance. The aesthetics and views, and the abundant natural resources in the Town of Bluffton have led to an increase in population and commercial growth in the area. As a result, major LU changes are planned for developments by the government of the Town of Bluffton. However, there are many concerns about the subsequent adverse impacts on water quality of the river. In addition, oyster harvesting is the main economic activity in the May River and polluting this river may impact the quality of the oyster beds.

During the past few decades, the Town of Bluffton has grown rapidly, and this growth is continuing. There were LU changes as the town was developing and growing in population and commercial activities which eventually have changed the quality of the May River. The LU in the watershed is majorly residential, with some minor commercial uses, and there are no heavy industrial activities within the watershed. However, the May River, for the first time in its history was downgraded in its shellfish harvesting classification in 2009. FC bacteria were and are still the main pollutants in the river. The May River watershed is within Shellfish Growing Area 19 (Figure 2.2). Five shellfish monitoring stations within shellfish growing area 19 were classified as Restricted in 2016 (Moody, 2016). These stations are 19-19, 19-19A, 19-19B, 19-19C, and 19-24 (Figure 2.2). The other eight stations were classified as Approved in 2016 (Moody, 2016).

2.3 DATA SOURCES AND METHODS

2.3.1 Fecal coliform data

Data for FC concentrations were retrieved from SCDHEC's shellfish-monitoring program. The program was established primarily to maintain the health and quality standards of the shellfish and their harvesting areas by following federal guidelines and state regulations (SCDHEC, 2017a). Also, this program was established to enhance water quality for the shellfish harvesting areas. Each shellfish growing area in South Carolina is comprehensively evaluated under this program. Annual evaluation is conducted for the

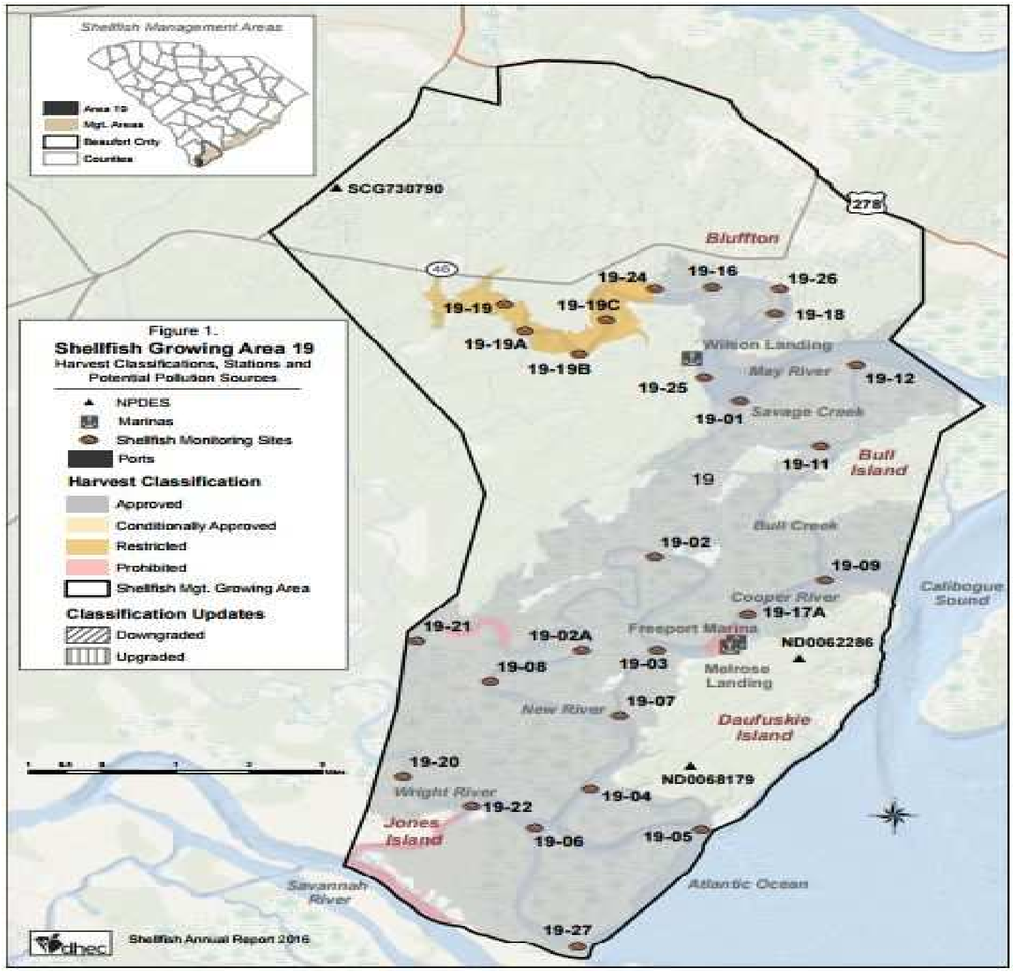

Figure 2.2: Shellfish Growing Area 19 map with SCDHEC shellfish monitoring stations Source: SCDHEC

shellfish growing areas that meet the requirements of the National Shellfish Sanitation Program (NSSP). Monthly routine sampling and laboratory analysis are conducted for bacteriological water quality monitoring at the SCDHEC's designated sampling sites. The SCDHEC's shellfish-monitoring program complies with the NSSP for standards, sampling and monitoring methods, and laboratory analysis. Most Probable Number (MPN) per 100 milliliters (ml) is the measurement unit used for the FC concentrations. As the shellfish monitoring stations are within Approved and Restricted areas, the guidelines for these two shellfish areas are as the followings: 1- For the Approved area, the median or the geometric mean of FC should not exceed 14 MPN per 100 milliliters, 2- A shellfish area should be classified Restricted when it is proven by the survey data that there are a reasonable level of pollutants exist or if there are harmful substances (deleterious or poisonous) which can unpredictably change the water quality or when it is not possible to classify the area as Conditionally Approved (SCDHEC, 2017b). There are 13 stations located in the study area which are included in this research. The data for FC concentrations from 1999 to 2015 are only available for 10 stations, while the data from three stations are available from 2009 to 2015. This is because these three stations were installed in 2009.

2.3.2 Land use and watersheds data

LU data for the years 2007 (Figure 2.3) and 2016 (Figure 2.4) were retrieved from the governments of the Town of Bluffton and Beaufort County. The data are shapefile feature classes that were developed by polygons that reflect the LU classes that were captured by the satellite imagery. The projected coordinate system for the data is NAD 1983 State Plane South Carolina FIPS 3900 Feet Intl, and the projection is Lambert Conformal Conic. The geographic coordinate system for the data is GCS North American 1983. Using ArcMap 10.4 software, the LU data were edited to prepare it for the analysis of the study. Twelve major LU types were identified: residential areas, forestlands, forested wetlands, non-forested wetlands, open water, commercial, transportation, non-residential, civic, golf courses, open space, and undeveloped areas. ArcMap 10.4 was used to calculate the percentages of each land use classes from 2007 and 2016 in the May River watershed. The

May River watershed layer used in the study is a 12-Digits Hydrologic Unit Code (HUC) that were retrieved from the government of the Town of Bluffton. Sub-watersheds data were also retrieved from the governments of the Town of Bluffton and Beaufort County (Figure 2.5). For the purpose of the study, the percentages of each land use class in each sub- watershed are required to assess the impact of LU classes on the concentration of FC

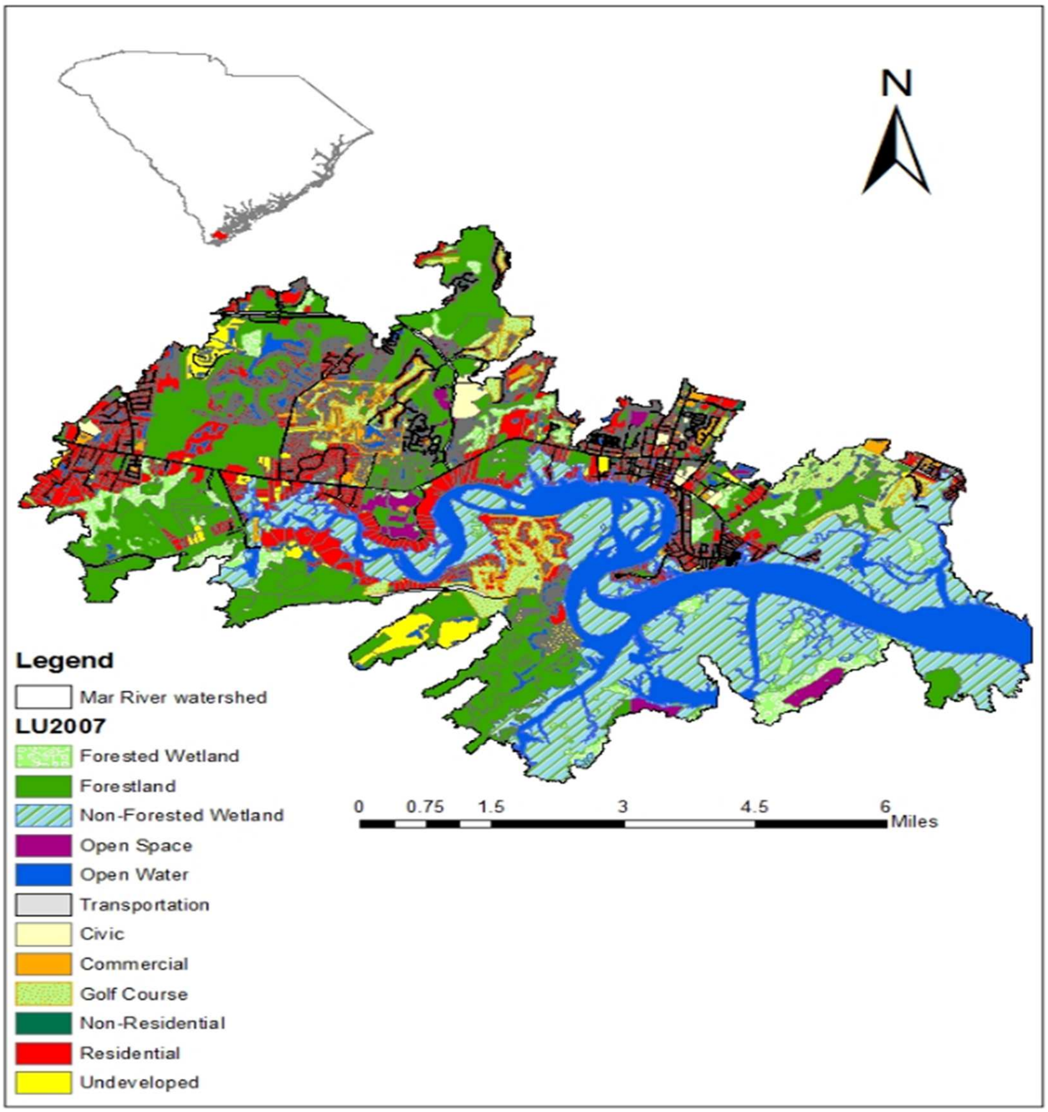

Figure 2.3: Land use map in 2007
Source: The government of Town of Bluffton

at each sub-watershed. In other words, the LU for every shellfish monitoring station was calculated based on the sub-watershed in which the monitoring station is located.

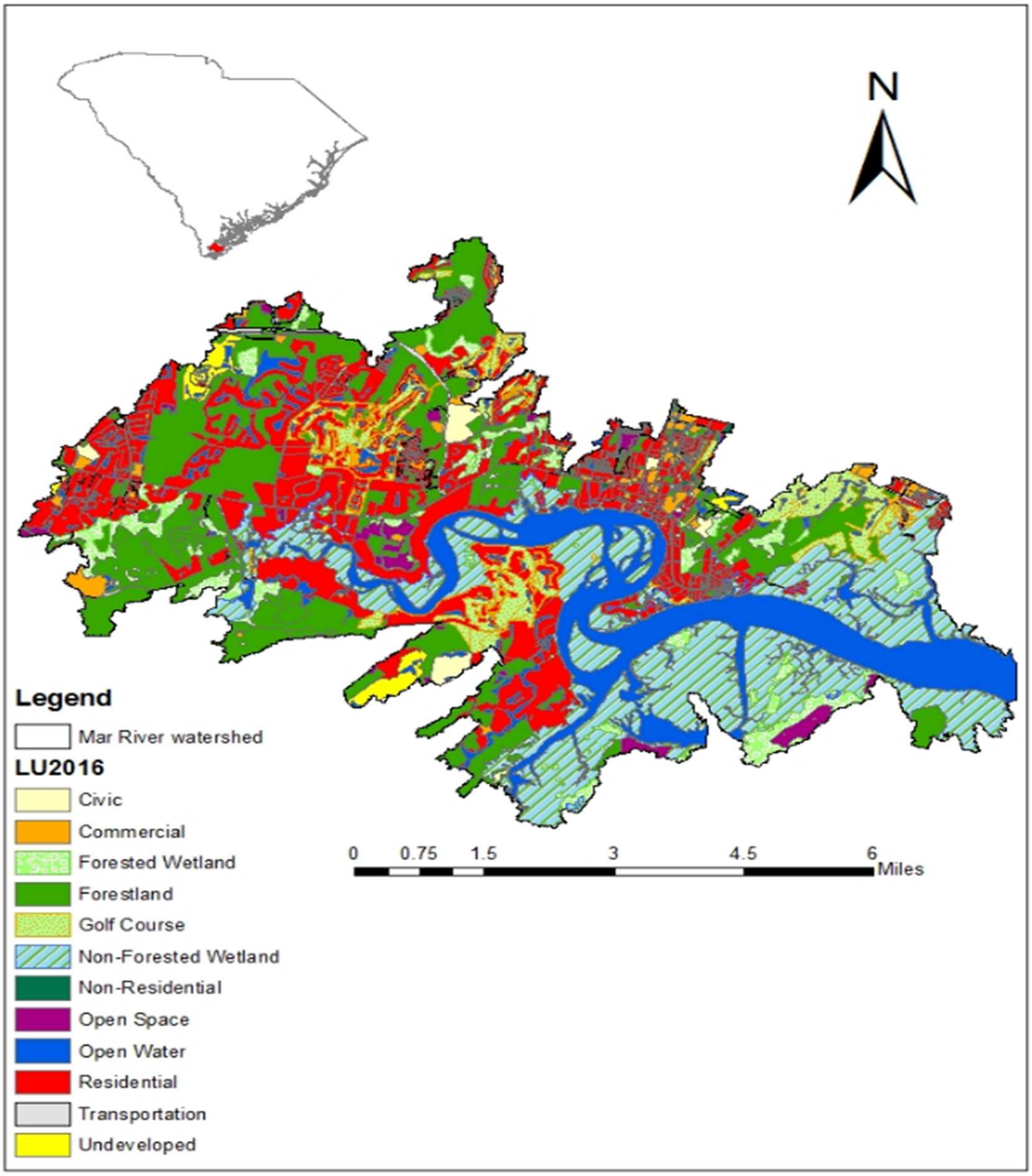

Figure 2.4: Land use map in 2016
Source: The government of Town of Bluffton

2.3.3 Statistical analysis for the relationship between LU and FC

Pearson product-moment correlation was used to assess the spatial relationships between land use variables for each shellfish monitoring station and their corresponding

FC concentrations to find the correlation coefficients (r) and the correlation tests. The resulted r-values showed the strength and the direction of the spatial relationships between

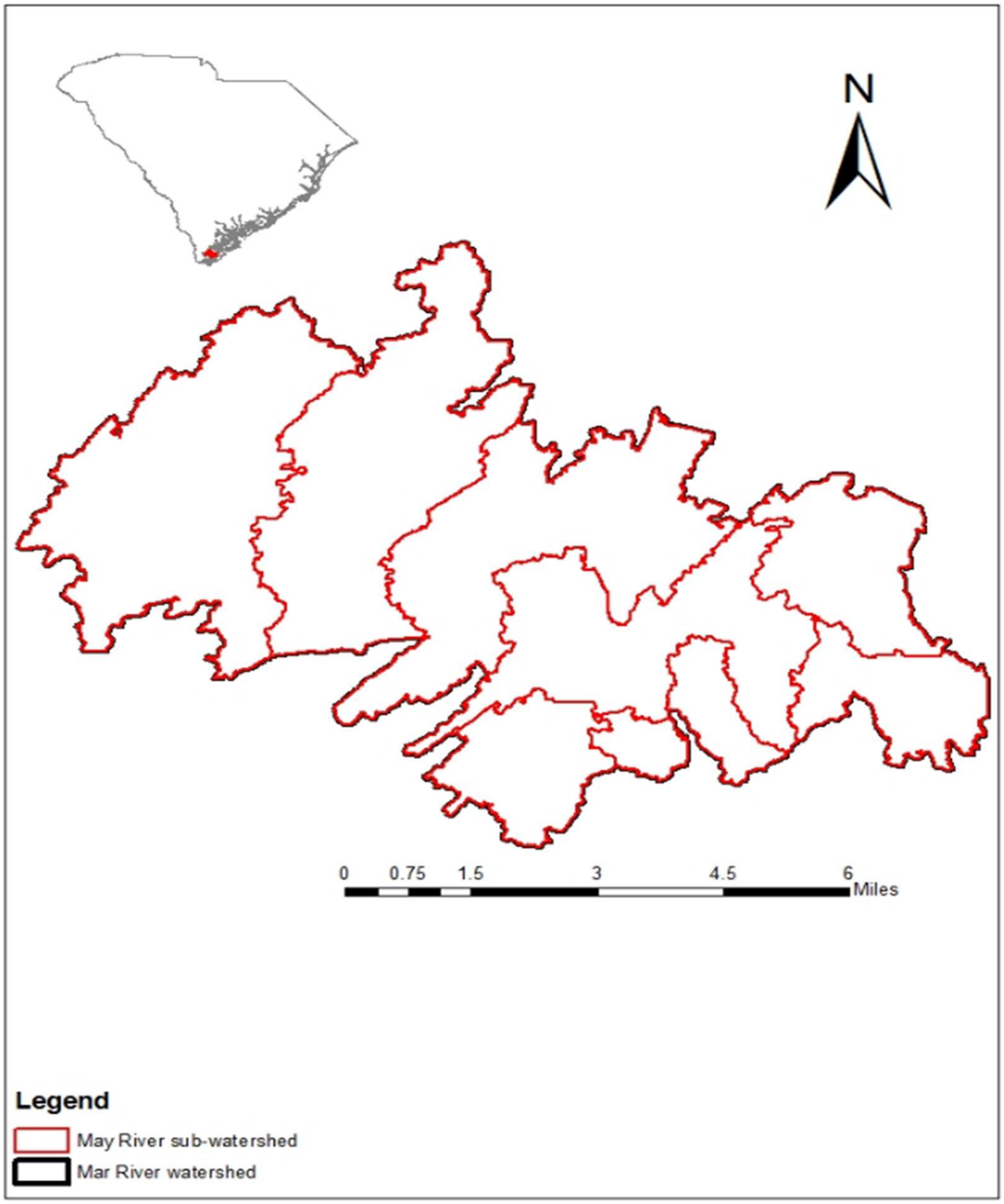

Figure 2.5: May River's sub-watersheds
Source: The government of Town of Bluffton

the selected variables. In other words, these r values indicated if there were significant or non-significant positive or negative correlations between the LU percentages and FC concentrations. This analysis was implemented for two periods. For land use data from

2007, the corresponding FC concentrations for 2009 were used; this is because there are three stations that started to collect FC concentration data in 2009. For 2016 LU data, the corresponding FC concentrations for the year 2015 were used, since the FC data for 2016 was not available. Scatterplot diagrams for the correlations for the two periods were also produced to help in explaining the relationships. The land use variables that showed significant positive or negative relationships with FC concentrations were used in the next analysis to examine the variations of the spatial relationships between the stations.

2.3.4 MODELING METHODS

To assess the variations of the spatial relationships between the LU percentages and FC concentrations, the mean values of FC concentrations were calculated for each station from 2009 to 2015. Land use percentages for 2016 were used in the analysis to examine the spatially varying relationships between these land use percentages and the calculated mean values of FC concentrations. To avoid spatial auto-correlation and to maintain the assumption of statistical independence of the predicted values in the analysis, only one station was included for each sub-watershed when more than one station was located in a single sub-watershed. Doing this entailed taking the mean FC concentrations of all the stations located in a single sub-watershed and to considering them one station. Therefore, out of the 13 stations included in the study only six stations were examined at six sub-watersheds.

The GWR method was used for the LU variables that were significantly correlated with FC concentrations. The GWR analysis was performed by using ArcMap10.4. Before developing the GWR models, it was necessary to develop OLS models for the same selected variables, and to ensure that these OLS models are not violating the linear regression assumptions. The ERT was used in ArcMap10.4 to help in selecting the LU variables that can be used without violating the assumptions of the linear regression method. This is to get the best models that can explain the spatial variations of the selected LU variables with their corresponding FC concentrations. The ERT tests all the given selected variables at one time by testing all the possible combinations of variables that are not violating the assumptions of the OLS, and that give the best fitting models. A successful model that is not violating the OLS assumptions is mandatory before running the GWR; this is because the GWR is very sensitive to multicollinearity and autocorrelation in performing the analysis.

After the selected variables were tested and elected by the Explanatory Regression Tool (ERT), the GWR models were developed to evaluate the local coefficients (parameter estimates), and local R-squared to exhibit how the relationships between LU percentages and FC concentrations can be spatially variable between the shellfish monitoring stations. The GWR can let the regression coefficients and the R-squared to change over the space. This means that every shellfish monitoring station will have its local regression coefficient and the local R-squared for its variables. The local coefficients were used to show how the relationships between the selected LU variables and their related FC variables varies between the monitoring stations, and the local R^2 was used to explain how the FC concentration values in the stations are variably influenced by the LU variables. The concept of the GWR models is as the following:

$$y_j = \beta_0\,(uj, vj) + \sum_{i=1}^{\rho} \beta_i\,(uj, vj)x_{ij} + \varepsilon_j$$

Where β_0 is the intercept for the coordinates $(uj, \nu j)$ for location j; β_i is the local regression coefficient for x_i which is the independent variable at location j, and ε_j is the local error term at location j.

In the GWR, the locations of the shellfish monitoring stations are geographically weighted in the model by using a distance decay function. This function basically assumes that the local coefficient of an observation has higher impact from the other observations that are closer to it than the ones that are further. A spatial kernel bandwidth is a main component in the weighting process. The distance decay function in the model is calculated as follows:

$$w_{ij} = \exp\left(-d^2_{ij}/b^2\right)$$

Were w_{ij} is the distance decay function's weight for the observations i and j, d_{ij} is the distance between the observations i and j, and b is the spatial kernel bandwidth. In ArcMap10.4, there are two types of kernel bandwidths which are FIXED or ADAPTIVE, from which the user can choose. The FIXED kernel bandwidth is fixed for the distance between the observations and thus it is appropriate when the observations are not randomly distributed. The ADAPTIVE kernel bandwidth is adaptable for the number of the neighboring observations and thus it is applicable for the observations that are randomly distributed. Three methods can be used to find the optimum bandwidth that can give the best prediction. These three methods are Akaike Information Criterion (AICc), Cross Validation (CV), and Bandwidth Parameter. The AICc method finds the optimum bandwith by minimizing the AICc value. The CV method finds the optimum bandwidth by minimizing the CV score. Both the AICc and CV work automatically in finding the optimum bandwidth. The Bandwidth parameter method is used when there is a need to manually specify the bandwidth. In this research, ADAPTIVE kernel bandwidth and AICc method were used in ArcMap10.4 to run the GWR models.

2.4 RESULTS AND DISSCUSSION

2.4.1 Characterization and percentage change of LU types for the years 2007 and 2016

May River watershed LU types and their percentages for the years 2007 and 2016 are shown in Table 2.1. The May River watershed was mostly forestlands in 2007 with a percentage about 24%. Following forestlands, non-forested wetland occupied about 20.5% of May River watershed. Residential area was about 20% in 2007. Open water

Table 2.1: land use types percentages for May River watershed

Land use type	2007	2016
Residential	19.96%	24.21%
Commercial	1.50%	1.60%
Civic	0.85%	1.00%
Open Space	1.86%	1.53%
Forestlands	23.95%	21.26%
Undeveloped	2.77%	1.17%
Transportation	2.48%	2.36%
Non-Residential	0.28%	0.28%
Non-Forested Wetland	20.50%	20.50%
Open Water	15.28%	15.60%
Forested Wetland	5%	5%
Golf Course	5.57%	5.50%

area was about 15.28% in 2007 and 15.6% in 2016. Golf courses cover about 5% of the watershed. In fact, the forestlands, residential areas, open water, and golf courses make up most of the area of May River watershed while the other land uses cover a very small portion of the watershed with percentages not more than 5%. There was about a 4% increase in the residential areas from 2007 to 2016 which made it the dominant land use in

2016 with 24.21% of the watershed. Forestlands decreased about 3% in 2016, and this is because the residential areas expanded in this period. For the same reason, open space and undeveloped lands decreased about 0.2% and 1.6%, respectively. The other land use types were either the same for the two periods or were changed very slightly.

2.4.2 The spatial relationships between LU types and FC

Pearson's correlation results for the correlation between LU types and FC in 2009 and 2015 are shown in Table 2.2. To get better understandings and better interpretations for the spatial relationships between each LU percentages and their relative FC concentrations in MPN/100 ml, scatterplots were used for the two periods (Figure 2.6 and

Table: 2.2 Pearson's correlation results for land use and FC for 2009 and 2015

Land use type	Pearson's correlation coefficient (r) 2009	P- value < 0.05	Pearson's correlation coefficient (r) 2015	P- value < 0.05
Residential	**0.81**	0.04	**0.83**	0.03
Commercial	**0.89**	0.01	0.40	0.42
Civic	**0.91**	0.009	0.68	0.13
Open Space	0.53	0.27	-0.03	0.94
Forestlands	**0.83**	0.03	**0.87**	0.02
Undeveloped	0.25	0.62	0.15	0.76
Transportation	0.77	0.06	0.77	0.06
Non-Residential	0.32	0.52	0.32	0.53
Non-Forested Wetland	-0.71	0.11	-0.75	0.08
Open Water	-0.72	0.10	-0.65	0.16
Forested Wetland	-0.33	0.51	-0.44	0.37
Golf Course	**0.87**	0.02	**0.89**	0.01

Bold**:** Significant correlation

Figure 2.7). Residential land use has significant positive relationships with FC in the two periods, 2009 and 2015, with correlation coefficient (r)-value ranged from 0.81 to 0.83 and p-values ranged from 0.04 to 0.03, respectively. This result was expected since more residential areas can discharge more point and non-point source pollutants. The non-point

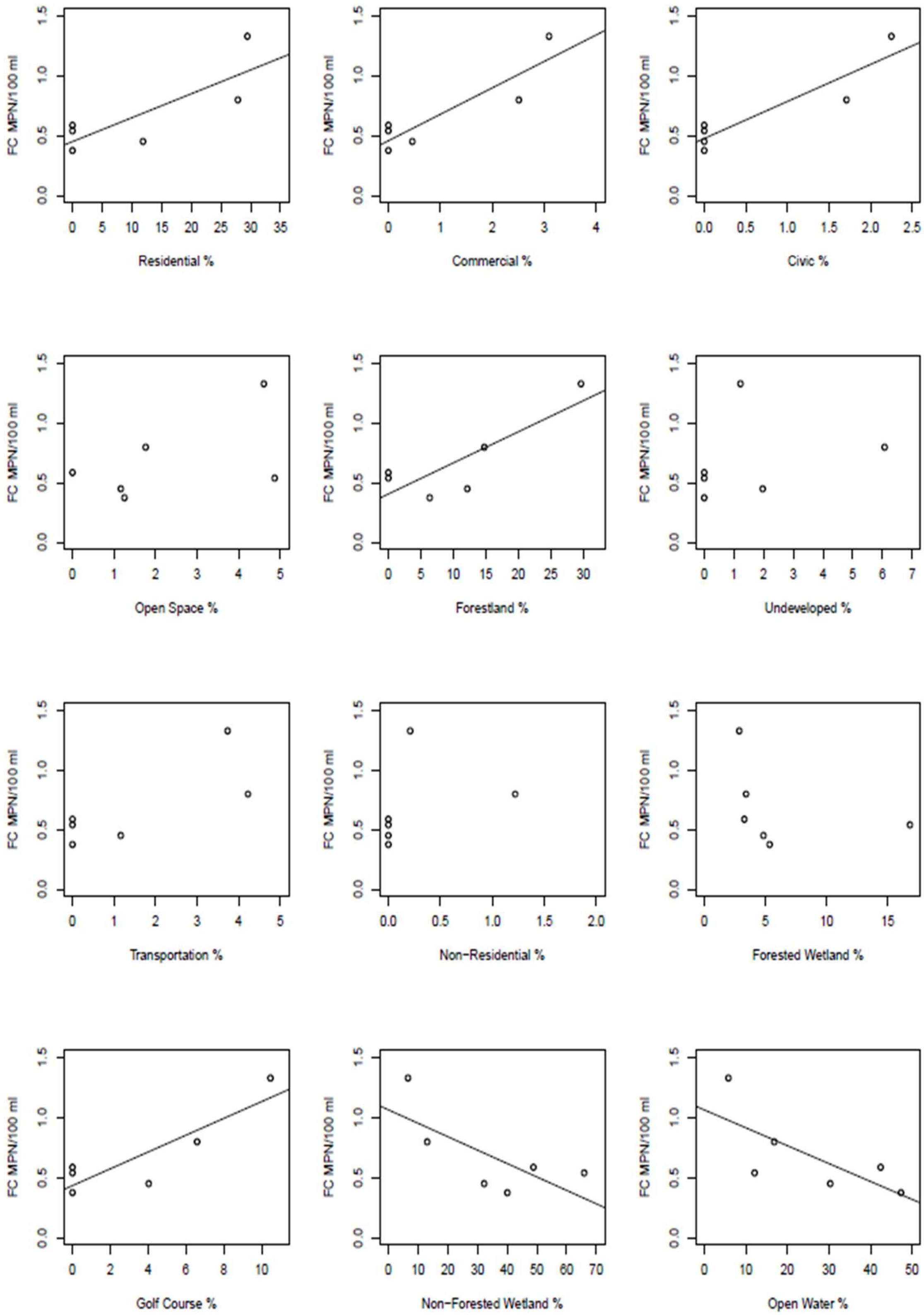

Figure 2.6: Scatter plots for land use percentages of 2007 and FC concentrations in 2009

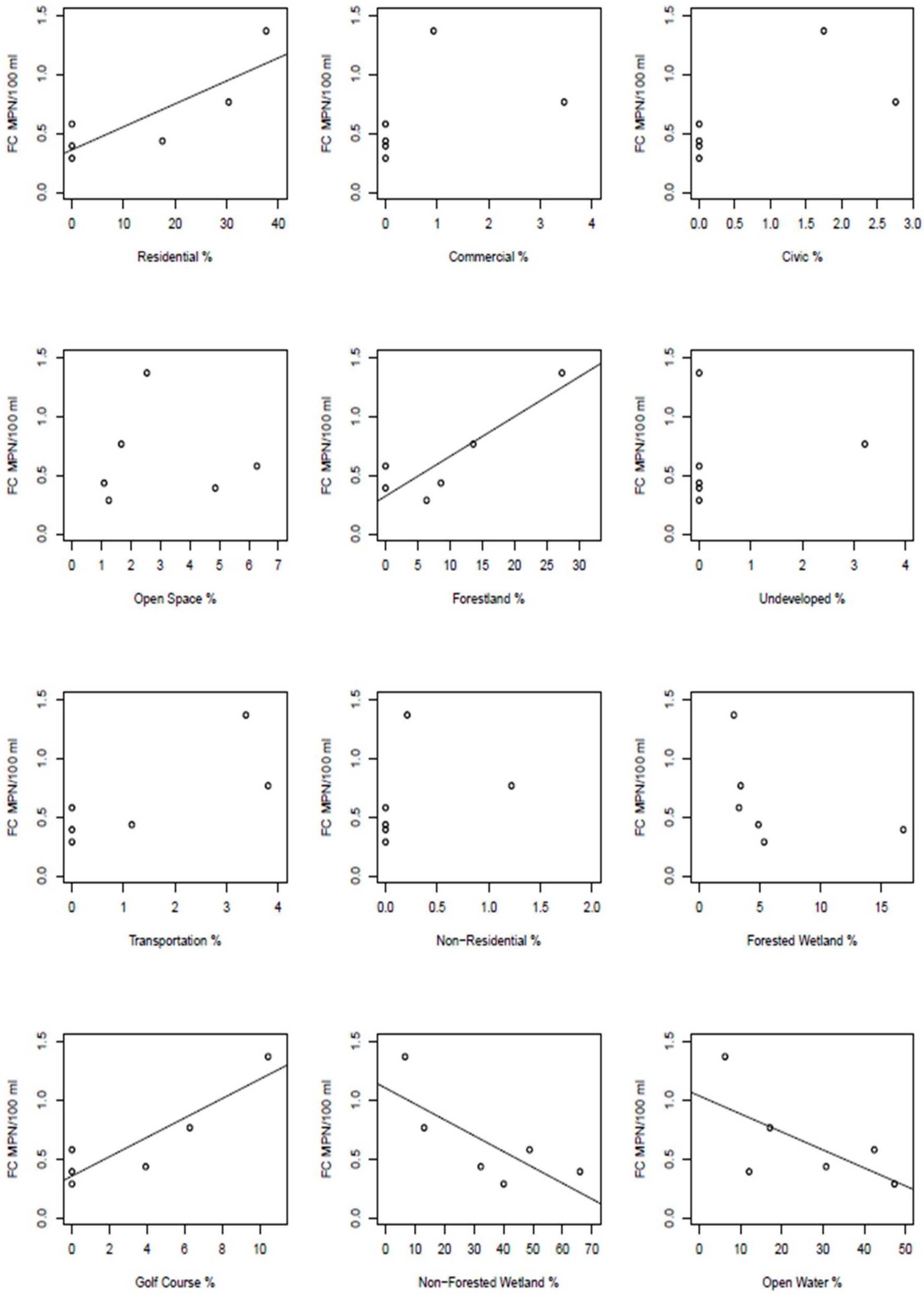

Figure 2.7: Scatterplots for land use percentages of 2016 and FC concentrations in 2015

source of FC from the residential areas is mainly caused by the runoff from these areas. It is well known that imperviousness of a specific land is an important factor in increasing its runoff. In fact, residential areas were found to have higher imperviousness compared to the other land uses. Furthermore, the runoff from these residential areas is commonly polluted with different biological and chemical wastes including fertilizers, pesticides, wastewater treatment plants, on-site septic systems, and pets and wild animals' wastes. Most of these sources can pollute the river stream with FC which will be carried by the precipitation runoff. This result is consistent with the outcome of many previous studies that assessed the the relationship between residential areas and FC (Kelsy et al., 2004, Reano et al., 2015, Cha et al., 2016, Paule-Mercado et al., 2016).

Same as residential areas, forestlands and golf courses were also significantly positively correlated with FC in the two periods with r-value ranged from 0.83 to 0.87, and from 0.87 to 0.89 and p-values ranged from 0.02 to 0.03 and from 0.01 to 0.02, respectively. The forestlands can be sources of FC due to wildlife animals' feces and due to fertilizers from agricultural areas as they are categorized under forestlands in this study. Therefore, runoffs from these areas are often loaded with fecal wastes in which FC bacteria exist. This result is similar to many previous studies (Cha et al., 2016; Paule-Mercado et al., 2016; Pettus et al., 2015; Schoonover & Lockaby, 2006). Golf courses can also be sources of FC due to fertilizers which accumulate in the river stream. Moreover, all the golf courses are located nearby the May River. This result is not different than the findings of some previous studies (Deason et al., 2014; Kelsy et al., 2008). The only negatively correlated LU with FC were open water, non-forested wetland, and forested Wetland. Open water was found to have a strong but not significant negative correlation in 2009 and 2015 with r-value ranged from

-0.72 to -0.65 and p-values from 0.10 to .16, respectively. This negative correlation of FC with open water is due to dilution processes. Higher dilution occurs in bigger open water areas, and consequently lower FC concentrations were found in sub-watersheds that are within bigger portion of the river stream. This finding is considered new in terms of assessing the relationships between LU and FC since most of the studies have not focused in open water impact on FC. Non-forested wetlands were also found to have strong negative correlation with FC in 2009 and 2015 with r-value ranged from -0.75 to -0.71 and p-values ranged from 0.11 to 0.08. This can be explained by the same reason for the open water as more dilution to FC concentrations can occur in bigger non-forested wetland areas. Additionally, the areas of non-forested wetlands are directly proportional to open water areas.

Transportation was strongly positively correlated with FC in 2009 and 2015 with matched r-values of 0.77 and matched p-values of 0.06. This correlation is not reliable since transportation percentages are very small at each sub-watershed; this strong positive correlation is because transportation percentage is greater in sub-watersheds that have greater percentages of residential areas which are significantly positively correlated with FC. Commercial and civic lands were found to have significant positive correlation with FC only in 2009. In fact, they represent a small land use portion in very few sub-watersheds, and most of the sub-watersheds have 0% of both commercial and civic land uses. Thus, the significant positive correlations in 2015 for the commercial and civic lands are not reliable. However, undeveloped, open space, non-residential, and forested wetland were found to have weak, positive and negative, correlations with FC in 2009 and 2015.

2.4.1 Examining the variations of the spatial non-stationary relationships between FC and LU types

After using the ERT, the only land uses that passed the assumptions for inclusion in a linear regression model are golf courses, forestlands and residential areas. One of the main purposes of using this tool was to know if it is possible to include more than one LU variable in the regression model. Based on the data of this study, this ERT indicated that there is no chance to have a model that passes the assumptions when more than one explanatory variable is included in the model. Furthermore, this tool indicated that it is possible to include up to two LU variables, but it would violate the assumptions of the multiple linear regression. The violations can be due to multicollinearity, model performance, spatial autocorrelation of the residuals, or due to no normal distribution of the residuals. The ERT checked on six parameters in selecting the passing models (Table 2.3). A passing model can consist of one or more LU variable.

The ERT selected three models. Each model has only one LU variable. The three selected models had golf courses, forestlands and residential areas for FC prediction; these models had an adjusted coefficient of determination (R^2) ranged from 0.70 to 0.83. The AICc compared the performance for the models; the lower AICc is the better model. Jarque-Bera tested the normal distribution of the residuals, and it tested not significant for the three selected variables. This means that there are normal distributions for the residuals of the selected models. Koenker (bp) or K(bp) in Table 2.3 is a test used by the same tool to see if there are spatial non-stationary variations that can be detected in the selected models if GWR is going to be used. The k(bp) tested significant for the three selected models. Variance Inflation Factor (VIF) was also used in this tool to test the multicollinearity among the explanatory variables. The VIF should be lower than 7.5 to avoid multicollinearity. All the selected variables had a value of 1 for the VIF. Spatial Autocorrelation (SA) for the residuals was also tested in the three selected models which tested not significant for each of them.

Table 2.3: Explanatory regression tool passing models

AdjR2	AICc	JB	K(BP)	VIF	SA	Model
0.831	10.01	0.75	0.60	1.0	0.11	+GOLF***
0.78	11.5	0.71	0.34	1.0	0.32	+FORESTLANDS***
0.70	13.4	0.78	0.04	1.0	0.14	+RESIDENTIAL**

***: Model variable significance at p <0.01
**: Model variable significance at p <0.05
AdjR2: Adjusted R-Squared
AICc: Akaike's Information Criterion
JB: Jarque-Bera p-value
K(BP):Koenker (BP) Statistic p-value
VIF: Max Variance Inflation Factor
SA: Global Moran's I p-value

As it is shown in Figure 2.8, there are local coefficients and local R^2 values generated by the GWR for the relationship between residential areas and FC. There were also clear spatial variations in the local coefficients and local R^2 between the stations. Higher coefficients were found at the stations located in sub-watersheds that have higher residential percentage and a higher FC concentration. This positive relationship result was expected since significant positive correlations between residential areas and FC were found in Table 2.2. The local R^2 for all the stations ranged from 0.51 to 0.72. From Figure 2.8, it was obvious that lower R^2 resulted at stations located in sub-watersheds with lower percentages of residential Areas. The higher local coefficients showed the higher R^2, and the higher coefficient was found at sub-watersheds with higher percentages of residential areas. This result shows that residential area percentage is an important factor that has to be considered when FC concentration is to be assessed. This result is consistent with many

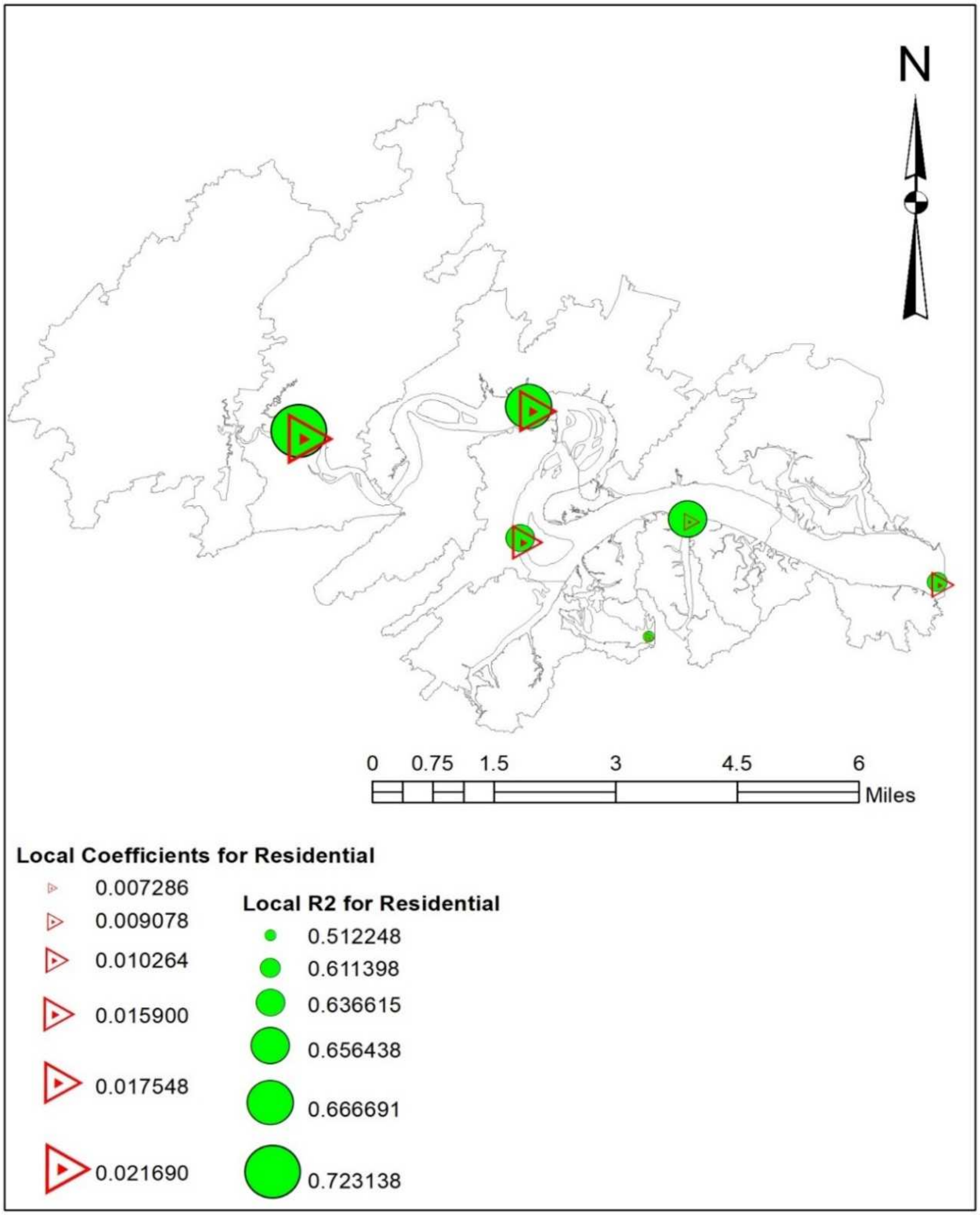

Figure 2.8: Results of GWR model for the spatial variations of the relationships between residential percentages and FC concentrations.

previous studies that examined the spatial variations of the relationship between residential areas and FC using the GWR (Dheenan et al., 2016; Kelsy et al.,2004; Paule-Mercado et al., 2016; Yu et al., 2013).

From Figure 2.9, local R^2 for forestlands ranged from 0.06 to 0.99 for all the stations. This means that forestlands explanation to FC varied from 6% up to 99% at the

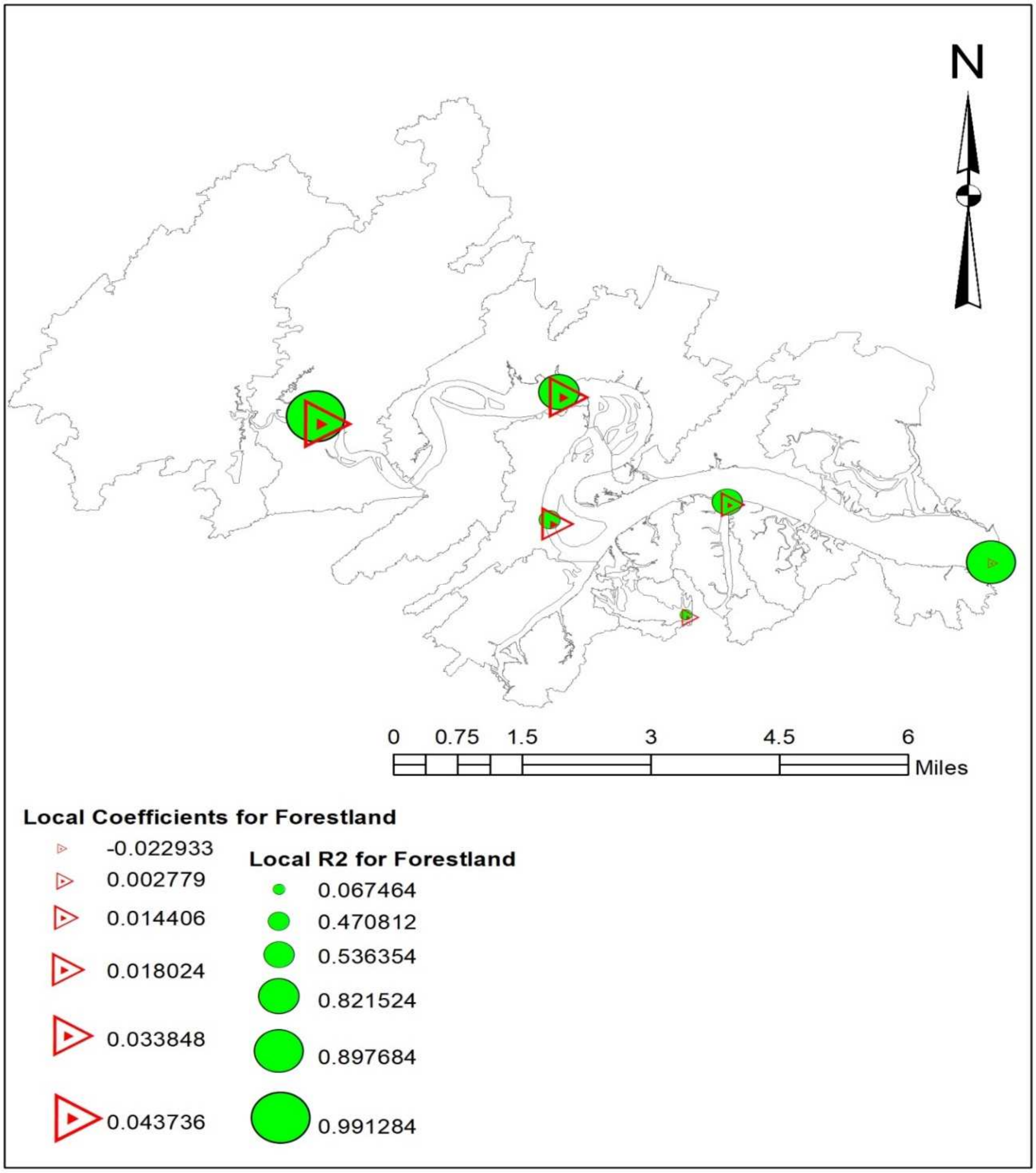

Figure 2.9: Results of GWR model for the spatial variations of the relationships between forestlands percentages and FC concentrations

sampling stations located in the sub-watersheds. This showed that there are great spatial variations of forestlands percentages in the sub-watersheds. Furthermore, a value of 0.99

for R^2 in one of the stations suggests that forestlands percentage is a very important predictor for FC. Additionally, greater R^2 were always found when there were high forestlands percentages in the sub-watersheds. A negative local coefficient was only found in one station, and it is the station that had the value of R^2 =0.06. All other stations had positive local coefficients. This negative correlation might be due to the presence of this station in a sub-watershed that has a high percentage of open water which dilutes FC densities. Also, there were only 6% of forestlands available in that sub-watershed, which is a very small amount. Overall, this GWR model result is compatible with the same indication of the significant positive correlation between forestlands and FC in Table 2.2. This result is consistent with several previous studies that examined the relationship between Forestlands and FC concentration (Cha et al., 2016; Paule-Mercado et al., 2016; Pettus et al., 2015; Schoonover & Lockaby, 2006), while some other studies found negative relationships between Forestlands and FC (BenDor et. al., 2017; Pettus et al., 2015).

Figure 2.10 shows the GWR model for the golf courses. The variations of the golf courses relationships with FC were very similar to the residential areas. In other words, the spatial variations of local R^2 and local coefficients for golf course and residential areas are almost the same. This indicated that the variation of the proportions of the percentages of the golf course and the residential areas in the sub-watersheds is similar. The local R^2 for Golf course ranged from 0.47 to 0.86. This significant positive relationship was also indicated by the correlation analysis in Table 2.2. This result suggests that the golf courses percentage is an important variable in predicting FC concentrations. Runoff from golf courses is well known for its impaired water quality; therefore, these variations of its relationships with FC were expected.

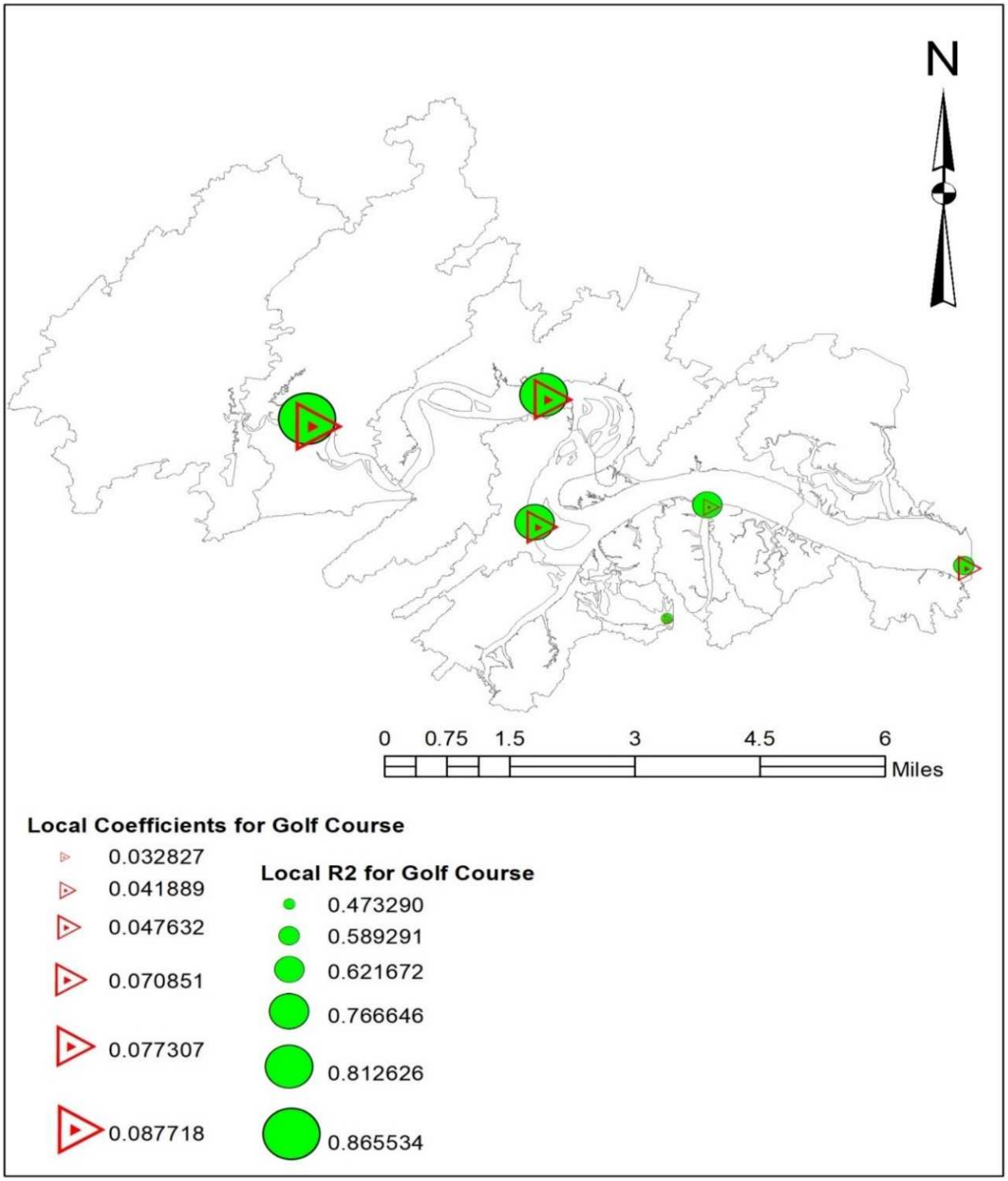

Figure 2.10: Results of GWR model for the spatial variations of the relationships between golf courses percentages and FC concentrations

The result for the GWR model for the open water (Figure 2.11) was different than the results for the models of the other land uses. While residential areas, forestlands, and golf courses showed positive relationships with FC concentration, open water showed negative correlation at most of the stations. This is similar to the findings of a previous

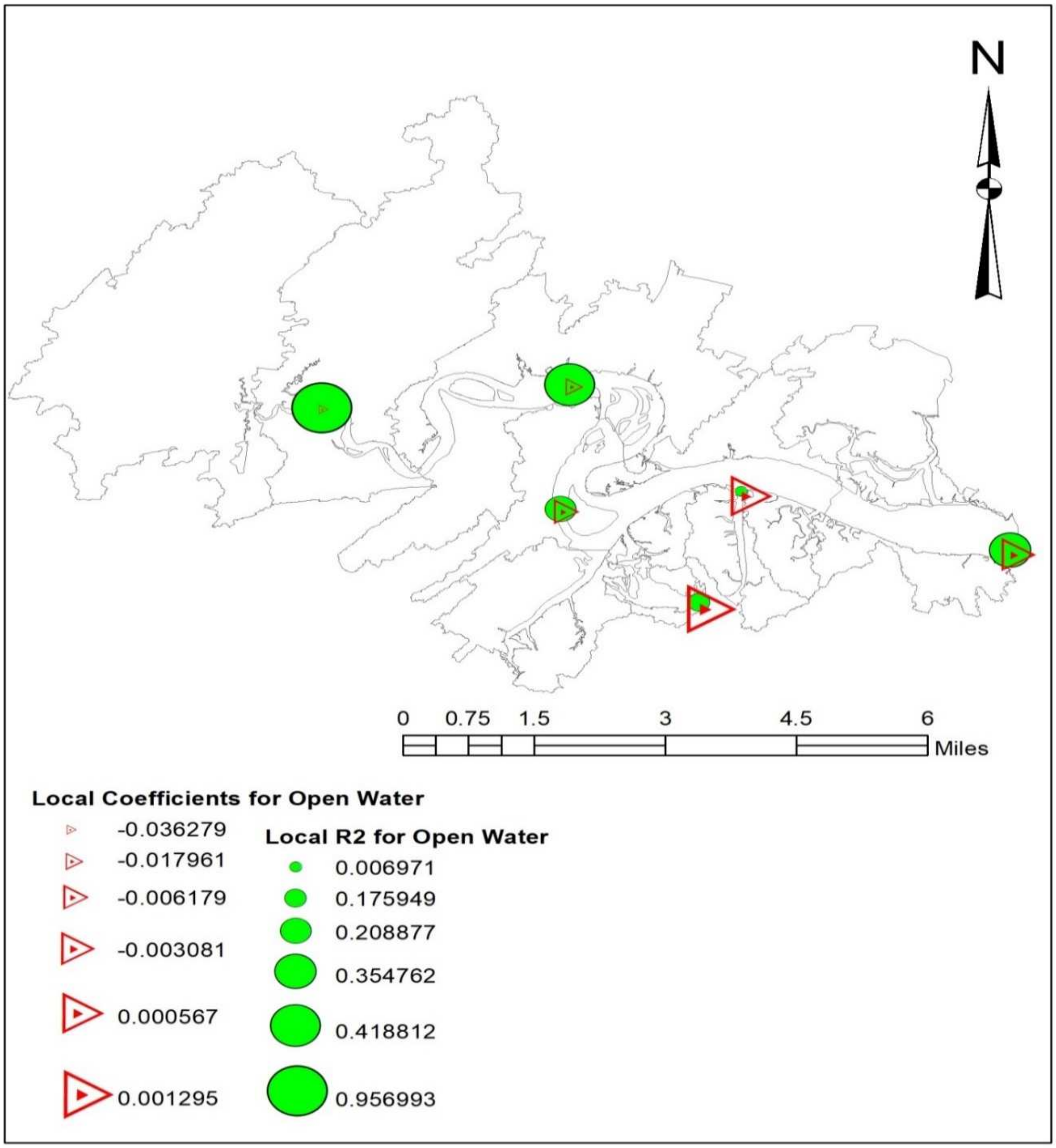

Figure 2.11: Results of GWR model for the spatial variations of the relationships between open water percentages and FC concentrations.

study (Dheenan et al., 2016), except for two stations. These two stations showed positive relationship with FC because of the two different situations for both stations. One station

(19-12) is located in the headwater of a creek that is connected to May River. The sub-watershed for this station included only the creek water which is a small percentage of open water that was considered to have low FC concentration. However, this station is located within greater open water area where dilution process occurs; thus, lower FC concentration was found, and a positive relationship was shown by the local coefficient of the GWR

model. The FC concentration was not found low in station (19-11) even though it had a high percentage of open because this station is not located in the stream of the May River. Instead, it is located at a creek far away from the River where no strong dilution is occurred; as a result, a positive relationship was shown by the GWR model. However, due to the locations of the stations in the sub-watersheds, these variations of negative and positive relationships for open water and FC concentrations were expected.

For better results, a larger study area and more sample numbers are required to explain more details about the variations of the impact of LU on FC. In addition, sub-watersheds with mixed land uses may reduce the precision of the results. The locations of the sampling station at each sub-watershed should be considered when choosing the sub-watersheds in the analysis. However, this was not possible due to data limitation.

2.5 CONCLUSION

This research attempted to assess the impact of different LU types on the microbial water quality of the shellfish harvest area in the May River at the Town of Bluffton, South Carolina. This research was conducted to examine the spatial (stationary and non-stationary) variations of the relationships between selected LU types and FC. Scatterplots and Pearson's correlation results of the stationary (global) spatial relationships showed numerus significant and non-significant positive and negative relationships between LU percentages and FC concentrations. The Scatterplots and Pearson's correlation were useful methods in explaining these relationships. The ERT was used to automatically select the explanatory LU variables that can be included in the GWR modelling. The ERT saved much time and effort in selecting the variables for GWR model's development. Although the purpose of using the ERT was to use multivariate variables in the GWR models, only a single variable was used for each GWR model. The GWR was able to explain detailed non-stationary (local) spatial variations of the relationships between the selected LU variables and FC. Residential areas, forestlands, and golf courses were found to have significant positive correlations with FC. On other hand, spatial variabilities of negative correlation with FC concentrations can be explained by open water. The GWR method is robust and sensitive in examining the spatial non-stationary relationships between the tested LU and FC. The visualization of the GWR makes it easier to know which shellfish monitoring station has either negative or positive local coefficient since the model is displayed in a map. Although it is preferable to use the GWR for larger study areas and lager sample numbers, it was capable to examine the spatial non-stationary variations in this study.

REFERENCES

Barber, B. P. (2008). Town of Bluffton May River watershed monitoring program. *Prepared for Town of Bluffton, Bluffton, SC. 143pp*.

BenDor, T. K., Jordanova, T. V., & Miles, B. (2017). A geospatial analysis of land use and stormwater management on fecal coliform contamination in North Carolina streams. *Science of the Total Environment*, *603*, 709–727.

Brown, S., Versace, V. L., Laurenson, L., Ierodiaconou, D., Fawcett, J., & Salzman, S. (2012). Assessment of spatiotemporal varying relationships between rainfall, land cover and surface water area using geographically weighted regression. *Environmental Modeling & Assessment*, *17*(3), 241–254.

Cha, Y., Park, M.-H., Lee, S.-H., Kim, J. H., & Cho, K. H. (2016). Modeling spatiotemporal bacterial variability with meteorological and watershed land-use characteristics. *Water Research*, *100*, 306–315.

Chen, Q., Mei, K., Dahlgren, R. A., Wang, T., Gong, J., & Zhang, M. (2016). Impacts of land use and population density on seasonal surface water quality using a modified geographically weighted regression. *Science of the Total Environment*, *572*, 450–466.

Cihlar, J., & Jansen, L. J. (2001). From land cover to land use: a methodology for efficient land use mapping over large areas. *The Professional Geographer*, *53*(2), 275–289.

Deason, G., Seekamp, E., & Barbieri, C. (2014). Perceived impacts of climate change, coastal development and policy on oyster harvesting in the Southeastern United States. *Marine Policy*, *50*, 142–150.

Dheenan, P. S., Jha, D. K., Das, A. K., Vinithkumar, N. V., Devi, M. P., & Kirubagaran, R. (2016). Geographic information systems and multivariate analysis to evaluate fecal bacterial pollution in coastal waters of Andaman, India. *Environmental Pollution*, *214*, 45–53.

Fotheringham, A. S., Brunsdon, C., & Charlton, M. (2003). *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons.

Fotheringham, A. S., Charlton, M. E., & Brunsdon, C. (2001). Spatial variations in school performance: a local analysis using geographically weighted regression. *Geographical and Environmental Modelling*, *5*(1), 43–66.

Giri, S., & Qiu, Z. (2016). Understanding the relationship of land uses and water quality in Twenty First Century: A review. *Journal of Environmental Management*, *173*, 41–48.

Huang, J., Huang, Y., Pontius, R. G., & Zhang, Z. (2015). Geographically weighted regression to measure spatial variations in correlations between water pollution versus land use in a coastal watershed. *Ocean & Coastal Management*, *103*, 14–24.

Ivajnšič, D., Kaligarič, M., & Žiberna, I. (2014). Geographically weighted regression of the urban heat island of a small city. *Applied Geography*, *53*, 341–353.

Kelsey, H., Porter, D. E., Scott, G., Neet, M., & White, D. (2004). Using geographic information systems and regression analysis to evaluate relationships between land use and fecal coliform bacterial pollution. *Journal of Experimental Marine Biology and Ecology*, *298*(2), 197–209.

Kelsey, R. H., Webster, L. F., Kenny, D. J., Stewart, J. R., & Scott, G. I. (2008). Spatial and temporal variability of ribotyping results at a small watershed in South Carolina. *Water Research*, *42*(8), 2220–2228.

Li, S., Zhao, Z., Miaomiao, X., & Wang, Y. (2010). Investigating spatial non-stationary and scale-dependent relationships between urban surface temperature and environmental factors using geographically weighted regression. *Environmental Modelling & Software*, *25*(12), 1789–1800.

Mallin, M. A., Williams, K. E., Esham, E. C., & Lowe, R. P. (2000). Effect of human development on bacteriological water quality in coastal watersheds. *Ecological Applications*, *10*(4), 1047–1056.

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). *Introduction to linear regression analysis* (Vol. 821). John Wiley & Sons.

Moody, R. (2016). Shellfish Management Area 19-2016 Annual Update. Submitted by the Shellfish Sanitation Section, Water Monitoring, Assessment and Protection Division, Environmental Quality Control–Bureau of Water, South Carolina Department of Health and Environmental Control, Columbia, SC. Retrieved from http://scdhec.gov/FoodSafety/Docs/sfrp_19.pdf

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

Pettus, P., Foster, E., & Pan, Y. (2015). Predicting fecal indicator organism contamination in Oregon coastal streams. *Environmental Pollution*, *207*, 68–78.

Pratt, B., & Chang, H. (2012). Effects of land cover, topography, and built structure on seasonal water quality at multiple spatial scales. *Journal of Hazardous Materials*, *209*, 48–58.

Reano, D. C., Haver, D. L., Oki, L. R., & Yates, M. V. (2015). Long-term characterization of residential runoff and assessing potential surrogates of fecal indicator organisms. *Water Research*, *74*, 67–76.

Rosenshein, L., Scott, L., & Pratt, M. (2011). Exploratory Regression: A tool for modeling complex phenomena. *ESRI Redlands*, *1*.

S.C. Dept. of Health and Environmental Control (SCDHEC). (2017a). DHEC: Food Safety - Shellfish Monitoring. Retrieved January 31, 2018, from http://www.scdhec.gov/FoodSafety/ShellfishMonitoring/

S.C. Dept. of Health and Environmental Control (SCDHEC). (2017b). Regulation 61-47, shellfish. Retrieved from http://www.scdhec.gov/Agency/docs/water-regs/r61-47.pdf

Schoonover, J. E., & Lockaby, B. G. (2006). Land cover impacts on stream nutrients and fecal coliform in the lower Piedmont of West Georgia. *Journal of Hydrology*, *331*(3), 371–382.

Su, S., Li, D., Zhang, Q., Xiao, R., Huang, F., & Wu, J. (2011). Temporal trend and source apportionment of water pollution in different functional zones of Qiantang River, China. *Water Research*, *45*(4), 1781–1795.

Su, S., Xiao, R., & Zhang, Y. (2012). Multi-scale analysis of spatially varying relationships between agricultural landscape patterns and urbanization using geographically weighted regression. *Applied Geography*, *32*(2), 360–375.

Su, Z. H., Lin, C., Ma, R. H., Luo, J. H., & Liang, Q. O. (2015). Effect of land use change on lake water quality in different buffer zones. *Applied Ecology and Environmental Research*, *13*(2), 489–503.

Sun, Y., Guo, Q., Liu, J., & Wang, R. (2014). Scale effects on spatially varying relationships between urban landscape patterns and water quality. *Environmental Management*, *54*(2), 272–287.

Tu, J. (2013). Spatial variations in the relationships between land use and water quality across an urbanization gradient in the watersheds of northern Georgia, USA. *Environmental Management*, *51*(1), 1–17.

Tu, J., & Xia, Z.-G. (2008). Examining spatially varying relationships between land use and water quality using geographically weighted regression I: Model design and evaluation. *Science of the Total Environment*, *407*(1), 358–378.

U.S Census Bureau. (2015). 2015 Population Estimate for Town of Bluffton. Retrieved January 26, 2017, from //www.census.gov/quickfacts/table/PST045215/4507210

Wabiri, N., Shisana, O., Zuma, K., & Freeman, J. (2016). Assessing the spatial nonstationarity in relationship between local patterns of HIV infections and the covariates in South Africa: a geographically weighted regression analysis. *Spatial and Spatio-Temporal Epidemiology*, *16*, 88–99.

Wan, R., Cai, S., Li, H., Yang, G., Li, Z., & Nie, X. (2014). Inferring land use and land cover impact on stream water quality using a Bayesian hierarchical modeling approach in the Xitiaoxi River Watershed, China. *Journal of Environmental Management*, *133*, 1–11.

Wang, R., Xu, T., Yu, L., Zhu, J., & Li, X. (2013). Effects of land use types on surface water quality across an anthropogenic disturbance gradient in the upper reach of the Hun River, Northeast China. *Environmental Monitoring and Assessment*, *185*(5), 4141–4151.

Wheeler, D. C. (2007). Diagnostic tools and a remedial method for collinearity in geographically weighted regression. *Environment and Planning A*, *39*(10), 2464–2481.

Wheeler, D. C. (2009). Simultaneous coefficient penalization and model selection in geographically weighted regression: the geographically weighted lasso. *Environment and Planning A*, *41*(3), 722–742.

Wheeler, D., & Tiefelsdorf, M. (2005). Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *Journal of Geographical Systems*, *7*(2), 161–187.

Wilson, C., & Weng, Q. (2010). Assessing surface water quality and its relation with urban land cover changes in the Lake Calumet Area, Greater Chicago. *Environmental Management*, *45*(5), 1096–1111.

You, W., Zang, Z., Zhang, L., Li, Z., Chen, D., & Zhang, G. (2015). Estimating ground-level PM10 concentration in northwestern China using geographically weighted regression based on satellite AOD combined with CALIPSO and MODIS fire count. *Remote Sensing of Environment*, *168*, 276–285.

Yu, D., Shi, P., Liu, Y., & Xun, B. (2013). Detecting land use-water quality relationships from the viewpoint of ecological restoration in an urban area. *Ecological Engineering*, *53*, 205–216.

Zhang, L., Bi, H., Cheng, P., & Davis, C. J. (2004). Modeling spatial variation in tree diameter–height relationships. *Forest Ecology and Management*, *189*(1), 317–329

CHAPTER 3

DEVELOPMENT OF PREDICTIVE MODELS TO MONITOR THE IMPACT OF LAND USE ON SHELLFISH HARVESTING WATER QUALITY

ABSTRACT

Determining the proper spatial scale for the land use (LU) classes is crucial to model their impacts on water quality and has been widely studied over the past few decades. In this research, multiple circular buffers that are different in size were created and used in developing predictive models that can estimate land use impact on water quality for a shellfish growing area in the May River at the Town of Bluffton, SC. Geographic Information System (GIS) was used in developing eight buffer sizes ranged from 500 to 2500 meters. Linear Mixed Model (LMM) was used for the statistical analysis in order to determine the most appropriate buffer size that its LU variables as well as environmental explanatory variables are significantly associated with FC concentration. Combined and seasonal models were developed and compared. The comparison between the models was based on the significance of the models' parameters. The data for this study included: (a) LU data from the government of the Town of Bluffton and (b) FC data from the South Carolina Department of Health and Environmental Control (SCDHEC). The LU percentages within the circular buffers of 1800 meters along with the environmental

explanatory variables were better associated with FC, and they provided significant predictive models. The results of the study appear to encourage the use of the LMM when the special scale selection for LU variables is required in developing water quality models.

3.1 INTRODUCTION

To protect the human health form the potential risk accompanied with shellfish consumption, the National Shellfish Sanitation Program (NSSP) has developed guidelines that have to be followed by all costal states to identify, survey, and classify their shellfish growing waters (USFDA, 1993). The shellfish growing areas may be classified as (1) Approved, (2) Conditionally approved, (3) Restricted, (4) Conditionally restricted, or (5) Prohibited. The main criterion for these classifications is the concentration of Fecal Coliform (FC) bacteria concentration. The FC bacteria can be derived from human waste, such as sewage, or animal wastes as well as some other point and non-point sources. Based on the epidemiological studies, the United States Environmental Protection Agency (USEPA) found that the concentration of Escherichia Coli (E. coli), a main species of fecal coliform indicator bacteria, is correlated with gastrointestinal illnesses for swimmers (Dufour, 1984; USEPA, 1986) as well as infections to ears, nose, eyes and throats in less cases (Hunter, 1997). Thus, the presence of FC bacteria in water streams is an indicator of a pathogenic contamination, which may pose a health risk due to shellfish consumption.

To prevent the occurrence of health problems related to shellfish consumption form water bodies that have exceeded fecal coliform indicator concentration, it is common to assess the water quality and issue shellfish harvest advisories. The main concern with issuing an advisory for the shellfish harvest areas is that the decision is based on the preceding day measurement for FC bacteria, because the FC bacteria take 18 to 24 hours for their incubation period (Whitman, Nevers, & Gerovac, 1999). The method that involves using models to estimate FC concentration based on preceding day measurements concentration is called the "persistence method" (USEPA, 2007).

Several studies have shown that the persistence method often post incorrect advisories (Christen, 2002; Frick, Ge, & Zepp, 2008; Whitman & Nevers, 2008; Whitman, Nevers, Korinek, & Byappanahalli, 2004). Therefore, this persistence method may provide either false positive or false negative advisories for a shellfish harvest area (Frick et al., 2008). False advisories, in some cases, could lead to either a closure for a safe condition or a non-closure for an unsafe condition. Consequently, this will result in exposing the public to health risks as well as economic losses. Therefore, many agencies and researchers have started to develop empirical predictive models to provide near real-time estimation for their surface water conditions (Francy, Darner, & Bertke, 2006; Francy & Darner, 2007). To some extent, these predictive models are useful in determining waterbodies that need stricter monitoring, and they are also helpful in estimating the conditions of the waterbodies between sampling times in order to avoid unsafe water conditions (USEPA, 2010).

To promote the efficiency of water quality monitoring programs, the use of the predictive models was highly recommended by the USEPA in its 2012 draft for recreational water quality criteria (USEPA, 2012). Decision making for a shellfish harvest area advisories can be improved by employing predictive models combined with their necessary data that can be attained from different accessible sources (Kelsey, Scott, Porter, Siewicki, & Edwards, 2010). Measurable physicochemical parameters of water can provide decision makers with feasible FC prediction based on a real-time forecasting approach (David & Haggard, 2011; Gonzalez, Conn, Crosswell, & Noble, 2012). Using a regression model incorporated with real-time measurements demonstrated a great promise for guiding decision makers towards water quality conditions (Olyphant & Whitman, 2004). Predictive models will still be necessary to guide decision makers towards water quality monitoring even with the possibility of measuring FC concentration directly (Olyphant & Whitman, 2004). Several studies have found that FC concentration can be better estimated by combining precipitation, climate and physiochemical variables in the models. (Brooks, Corsi, Fienen, & Carvin, 2016; David & Haggard, 2011; Galfi, Österlund, Marsalek, & Viklander, 2016; Gonzalez et al., 2012; Kelsey et al., 2010; Olyphant, Thomas, Whitman, & Harper, 2003; Olyphant & Whitman, 2004; Thoe et al., 2014; Thoe, Wong, Choi, & Lee, 2012). Therefore, the variations in FC densities are affected by the interaction of several environmental factors including water and climate parameters.

Furthermore, a study found that precipitation and temperature are crucial factors in determining the FC densities due to their impacts on FC survival (Leight, Hood, Wood, & Brohawn, 2016). Other studies also found that temperature, Biological Oxygen Demand (BOD), Turbidity, Total suspended solids (TSS) and antecedent dry days are significant parameters to provide detailed association with FC (Paule-Mercado et al., 2016). A study used Soil and Water Assessment Tool (SWAT) and concluded that temperature was the main significant variable in determining FC concentration (Cho et al., 2016). Moreover, a study found that FC densities can be better predicted using only water and weather parameters (Nevers & Whitman, 2005).

Several studies have suggested the integration of LU classes in predicting FC concentrations (Cha, Park, Lee, Kim, & Cho, 2016; Crowther, Wyer, Bradford, Kay, & Francis, 2003; Galfi et al., 2016; Paule-Mercado et al., 2016; Schoonover & Lockaby, 2006). Most importantly, some of these studies have shown that physiochemical parameters, LU classes, and imperviousness give a greater insight to model the FC bacteria. (Galfi et al., 2016; Schoonover & Lockaby, 2006). However, some other studies have revealed that the strength of the relationships between the modeled variables including FC are varied by the variation of the seasons and the geographical locations, and they suggested that the models should be site-specific for the predicted bacteria. (David & Haggard, 2011; Leight, et al., 2016).

Accordingly, precipitation is the most commonly used explanatory variable in the literature, and the other climate and physiochemical variables were used based on data availability and best fitted models resulted. Furthermore, other studies have recommended separating the models into base and high flow conditions to get more accurate FC prediction (Crowther et al., 2011; David & Haggard, 2011; Galfi et al., 2016; Paule-Mercado et al., 2016; Schoonover & Lockaby, 2006; Thoe & Lee, 2013). These studies used such approach in order to identify the essential variables that must be considered for both base and high flow since storm water runoff is the main nonpoint source of FC during the high flow condition.

Different statistical methods were used in different studies to predict FC such as Artificial Neural Network (ANN), Multiple Linear Regression (MLR), and quantitative PCR (qPCR). Quite a number of studies have found that the ANN is a promising method for developing predictive models for FC concentrations (He & He, 2008; Zhang, Deng, & Rusch, 2012). Several other studies used MLR method to predict the FC densities; these studies suggested that MLR is a promising tool and can provide a greater insight to understand the factors that influence the FC densities (Brooks et al., 2016; Frick et al., 2008; Gonzalez & Noble, 2014; He & He, 2008; Herrig, Böer, Brennholt, & Manz, 2015; Olyphant et al., 2003; Paule-Mercado et al., 2016; Zhang et al., 2012). Therefore, MLR as well as real-time explanatory variables can be used to develop empirical models, which can be used as the basis to issue health advisories (Frick et al., 2008).

As LU variables will be included in developing the FC models, the spatial scale for defining the LU variables is required. For this study, circular buffers method is the only option that can be used to define the LU variables because the shellfish monitoring stations are positioned in the stream of the May River. Several studies have used circular buffers to model the impact of LU on water quality (Azyana & Norulaini, 2012; Chang, 2008; Daugomah, Siewicki, Porter, & Scott, 2007; Gyawali, Techato, & Monprapusson, 2015; Sanger, Holland, & Hernandez, 2004; Zhai, Xia, & Zhang, 2014). Few studies have concluded that circular buffers were not better than other spatial scales such as watersheds in explaining the LU impact (Azyana & Norulaini, 2012; Gyawali et al., 2015; Zhai et al., 2014). However, several other studies were able to evaluate the influence of LU classes on water quality by developing circular buffers for the LU variables (Chang, 2008; Daugomah et al., 2007; Sanger et al., 2004). These studies have not come to an agreement on a consistent circular buffer size although it ranged from 400 to 500 meters.

Due to data collection approach in this study, there will be repeated measurements (observations) for each shellfish monitoring station. In this case, the LU variables will be matched (unchanged) for the observations of each station, and the MLR is not the appropriate statistical method to deal with the repeated measurements. There is a Linear Mixed Model (LMM), which can deal with the repeated measurements (Pinheiro and Bates, 2000). Several studies have used the LMM to predict FC by including LU as well as other factors in the modelling (Delpla & Rodriguez, 2014; Hurley & Mazumder, 2013; Ragosta et al., 2010). In this study, the LMM method will be used to find the circular buffer size for LU variables that can provide the most predictive models for FC bacteria.

3.2 STUDY AREA

The May River watershed is located in Beaufort County in south east of South Carolina (Figure 3.1). The Town of Bluffton has the May River as its significant estuary, which drains into the Atlantic Ocean on the eastern coast of South Carolina. The population of the Town of Bluffton in 2015 was about 16,728 according to the U.S Census Bureau, 2015. Due to the high water quality of the May River in 2001, it was classified Outstanding Resource Water (ORW) by the SCDHEC (Barber, 2008). The Environmental Protection Agency (EPA) and SCDHEC have acknowledged the May River as a priority watershed. The May River is highly valued by the residents of the Town of Bluffton due to its great recreational and economical importance. The aesthetics and views, and the abundant natural resources in the Town of Bluffton have led to an increase in population and commercial growth in the area. As a result, major LU changes are planned for developments by the government of the Town of Bluffton. However, there are many concerns about the subsequent adverse impacts on water quality of the river. In addition, oyster harvesting is the main economic activity in the May River and polluting this river may impact the quality of the oyster beds. During the past few decades, the Town of Bluffton has grown rapidly, and this growth is continuing. There were LU changes as the town was developing and growing in population and commercial activities have eventually changed the quality of the May River. The LU in the watershed is majorly residential, with some minor commercial uses, and there are no heavy industrial activities within the watershed. However, the May River, for the first time in its history, was downgraded in its

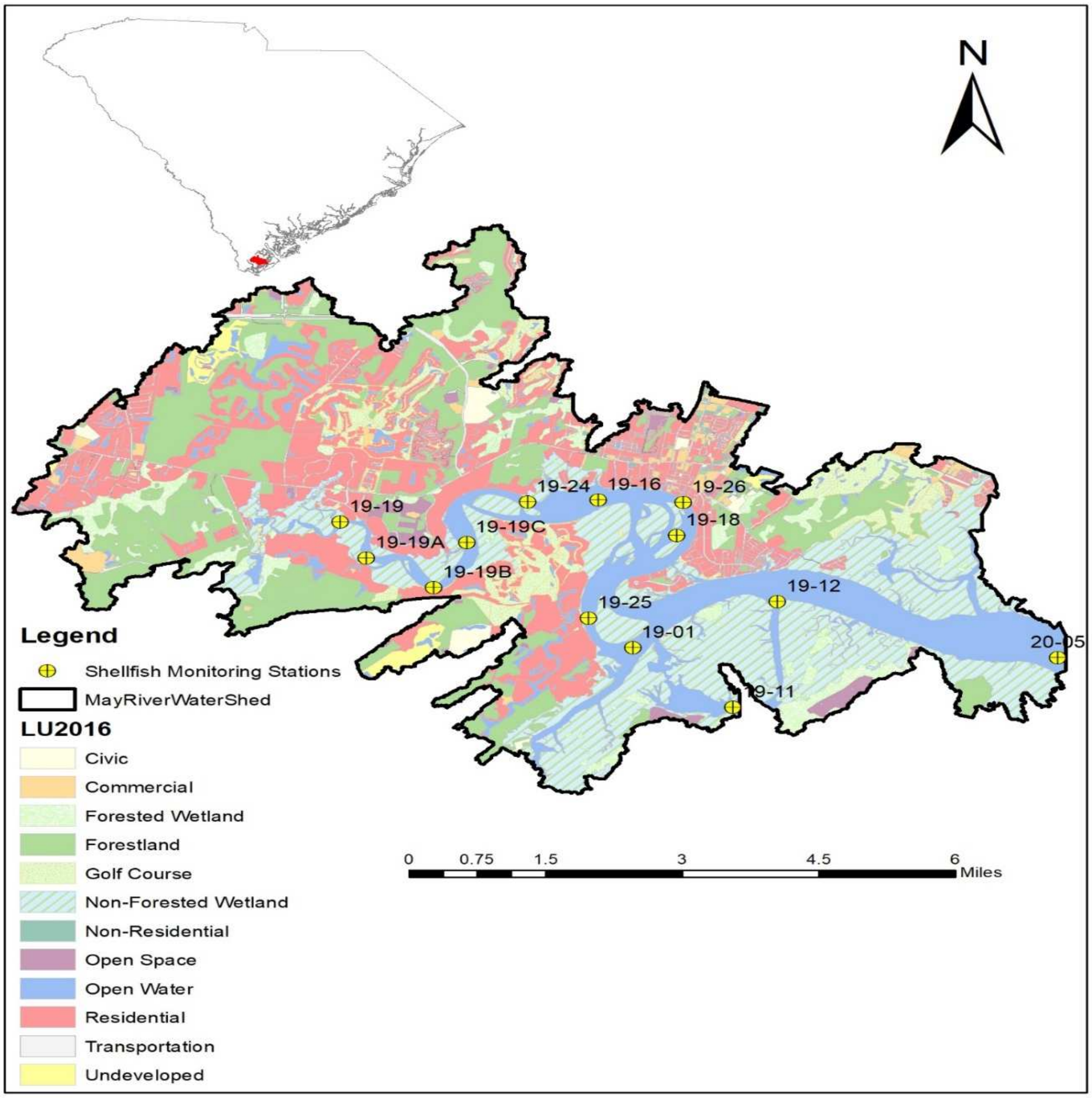

Figure 3.1: May River watershed with shellfish monitoring station sites and LU classes in 2016
Data source: The government of Town of Bluffton

shellfish harvesting classification in 2009. FC bacteria were and are still the main pollutants in the river. The May River watershed is within Shellfish Growing Area 19 (Figure 3.2). Five shellfish monitoring stations within shellfish growing area 19 were classified as Restricted in 2016 (Moody, 2016). These stations are 19-19, 19-19A, 19-19B, 19-19C, and

19-24 (Figure 2.2). The other eight stations were classified as Approved in 2016 (Moody, 2016).

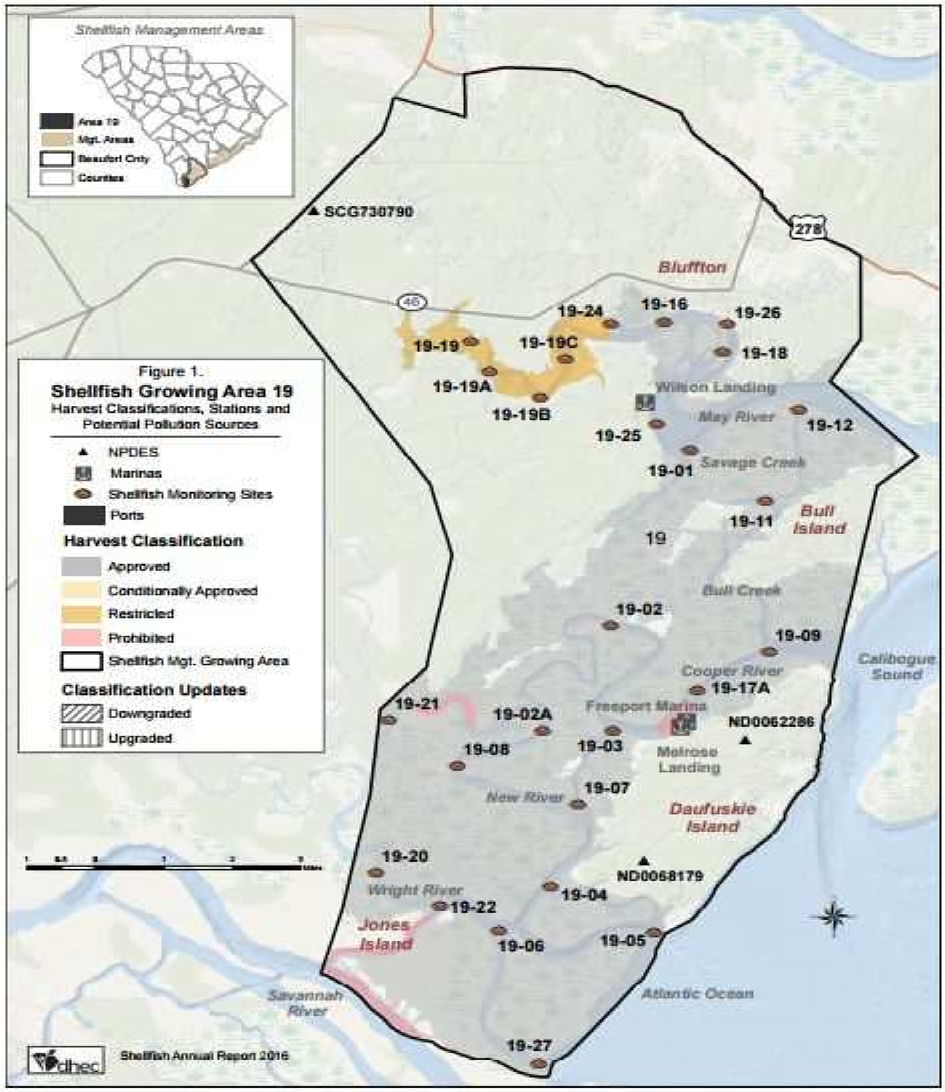

Figure 3.2: Shellfish Growing Area 19 map with SCDHEC shellfish monitoring stations
Source: SCDHEC

3.3 DATA SOURCES AND METHODS

3.3.1 Fecal Coliform Data

Data for FC concentrations were retrieved from SCDHEC's shellfish-monitoring program. The program was established primarily to maintain the health and quality standards of the shellfish and their harvesting areas by following federal guidelines and state regulations (SCDHEC, 2017a). Also, this program was established to enhance water quality for the shellfish harvesting areas. Each shellfish growing area in South Carolina is comprehensively evaluated under this program. Annual evaluation is conducted for the shellfish growing areas that meet the requirements of the National Shellfish Sanitation Program (NSSP). Monthly routine sampling and laboratory analysis are conducted for bacteriological water quality monitoring at the SCDHEC's designated sampling sites. The SCDHEC's shellfish-monitoring program complies with the NSSP standards, sampling and monitoring methods, and laboratory analysis. Most Probable Number (MPN) per 100 milliliters (ml) is the measurement unit used for the FC concentrations.

As the shellfish monitoring stations are within Approved and Restricted areas, the guidelines for these two shellfish areas are as the following: 1- For the Approved area, the median or the geometric mean of FC should not exceed 14 MPN per 100 milliliters, 2- A shellfish area should be classified as Restricted when it is proven by the survey data that there are a reasonable level of pollutants exist or if there are substances deleterious or poisonous substances which can unpredictably change the water quality or when it is not possible to classify the area as Conditionally Approved (SCDHEC, 2017b). There are 13 stations located in the study area which will be included in this research. The data for FC concentrations from 1999 to 2015 are only available for 10 stations, while the data from three stations are available from 2009 to 2015. This is because these three stations were installed in 2009.

3.3.2 Land use data

LU data for the year 2016 were retrieved from the governments of the Town of Bluffton. The data is a shapefile feature class that was developed by polygons that reflect the LU classes that were captured by the satellite imagery. The projected coordinate system for the data is NAD 1983 State Plane South Carolina FIPS 3900 Feet Intl, and the projection is Lambert Conformal Conic. The geographic coordinate system for the data is GCS North American 1983. Using ArcMap 10.4, the LU data were edited in order to be prepared for the analysis of the study. Twelve major LU types were identified: residential areas, forestlands, forested wetlands, non-forested wetlands, open water, commercial, transportation, non-residential, civic, golf courses, open space, and undeveloped areas.

3.3.3 Circular Buffers for land use data

ArcMap 10.4 was used to calculate the percentages of each land use class for the year 2016 in circular buffers. A shapefile feature class layer for the May River watershed was required for the buffers developments to clip the edges of the buffers within the watershed borders. The May River watershed's layer used in the study is a 12-Digits Hydrologic Unit Code (HUC) that was retrieved from the government of the Town of Bluffton. For the purpose of this study, the percentages of the LU classes for each shellfish monitoring station is required to get the LU variables for the stations. Therefore, a circular buffer was created for each shellfish monitoring station to calculate the LU percentages inside it. Since there are 13 shellfish monitoring stations in the study area, 13 circular buffers have been developed. As one of the study objectives is to identify the optimum buffer size for FC modelling, it was necessary to develop multiple buffers with different sizes for all the stations. For that reason, eight circular buffer sizes in meters were created: 500m, 800m, 1000m, 1200m, 1500m, 1800m, 2000m, and 2500m (Figure 3.3).

3.3.4 Precipitation data

Two rainfall datasets were used for fecal coliform modeling. One dataset is from rainfall monitoring gauge for Broad Creek Public Service District (PSD). This is the rainfall data source that is used by the SCDHEC for shellfish monitoring assessments in the May River. The other rainfall dataset was retrieved from the Next Generation Weather Radar (NEXRAD) system from the National Atmospheric and Oceanic Administration (NOAA). Historical rainfall data from 2009 to 2015 were used in this study. The data were prepared in two ways for the purposes of the study. Rainfall measurements for the day of sampling, and rainfall measurements for 24 hours, 48 hours and 72 hours preceding the day of sampling.

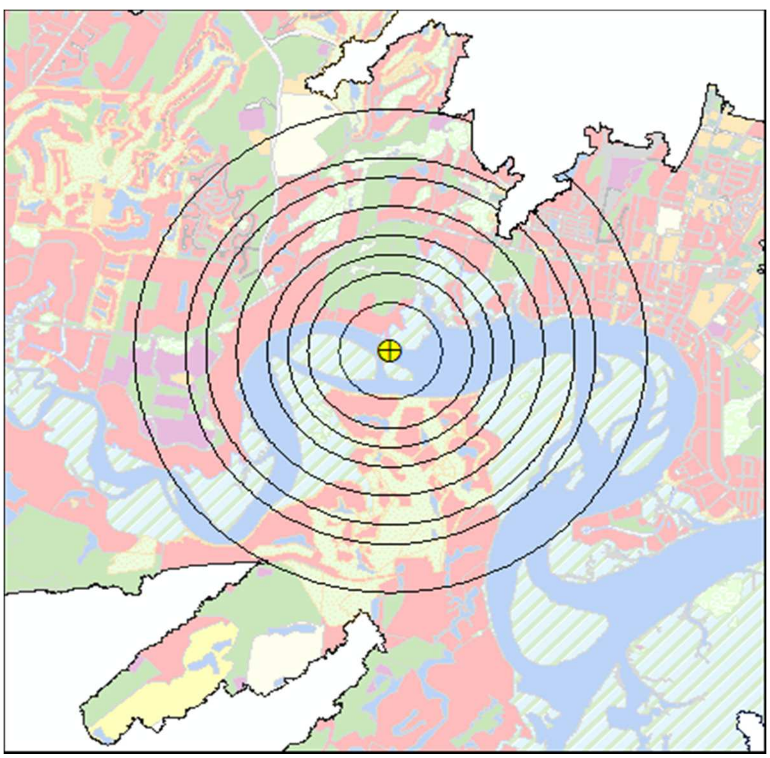

Figure 3.3: Example of the circular buffer sizes for one of the shellfish monitoring stations

3.3.5 Environmental factors data

The SCDHEC's shellfish monitoring stations measure the following parameters: salinity, water temperature, air temperature, salinity, wind direction and tide stage. These parameters were retrieved form the SCDHEC and were used in FC modeling. Air and water temperature are measured in Celsius, salinity is measured in part per thousand (ppt), wind direction is a categorical variable (eight categories) with eight directions, and tide stage is a categorical variable (eight categories) with eight levels.

3.3.5 Statistical analysis of the relationship between salinity and rainfall

Pearson product-moment correlation and scatterplots were used to assess the impact of rainfall on salinity. Precipitation measurements from rain gauge and NEXRAD for the day of sampling were used in the analysis to see if they influence the salinity concentration in the May River. The precipitation and salinity measurements from 2009 to 2015 were included in the analysis. The correlation coefficients (r) and the p-values for the correlation test explained the strength, the significance and the direction of the relationship between salinity and rainfall. The results of Pearson product-moment correlation can determine if there are significant or non-significant positive or negative correlations between salinity and rainfall.

3.3.6 Statistical analysis of the relationship between salinity and FC bacteria

Pearson product-moment correlation and a scatterplot between FC and salinity were used to explore how FC concentration is linked to salinity. Data from 2009 to 2015 for salinity and FC measurements for the day of sampling were included in the analysis. The r-value and the p-value for the correlation test illustrated the strength and significance of the correlation of the relationship between salinity and FC. The degree of significance and direction of the relationship is determined by using this test. The scatterplot was used to display the correlation and its direction between salinity and FC.

3.3.7 MODELING METHODS

3.3.7.1 Combined models

LMM was used to develop 16 predictive models for the entire sampling period (2009 to 2015). The LMM was used because it is an appropriate statistical model in determining associations between a dependent variable and a set of predictors when the data are repeatedly collected from the sampling stations. This is because the data is structured by sampling dates for each monitoring station, so in this case, the LU data are repeated for each sampling station, and because the LU data do not vary by date, they will be matched for the repeated measurements for each station.

Rian gauge and NEXRAD datasets were used in modelling for the eight buffer sizes to get eight models with rain gauge and eight models with NEXRAD. In each model, the following variables were included: LU classes (residential areas, forestlands, open spaces, golf courses, non-forested wetlands, open water), water temperature, air temperature, salinity, wind direction, tide stage, precipitation measurements (24 hours, 48 hours and 72 hours) preceding the day of sampling. R-studio software via “lme” function was used to perform the LMM to get the 16 models.

For each buffer size, a model selection function called “dredge” in R-studio was used to automatically select the variables that provide the best fitted model. This method was used to make sure that each buffer size has the best selected model. This is necessary before the models for the buffers are compared. The comparison objective is to find the buffer size that provides the most predictive model. However, this most predictive model must be significant. Therefore, the parameters of the models' outcomes were compared. This was done by identifying the model that has the most significant coefficients at a p-value less than 0.05. Since the LMM does not provide R^2 values, a function in R-studio was used to get marginal R^2 values for the models.

3.3.7.2 Seasonal models

After finding the optimum buffer size for providing the most significant combined model, the data were separated by season to get the results for the seasonal models. Therefore, there were four datasets used in the analysis for the seasonal model. The seasonal models were developed twice with the two precipitation datasets, rain gauge and NEXRAD. This is to compare the seasonal models' performance from with the two different precipitation datasets. Similar to the combined models, model selection was performed for the seasonal models to find the most significant models.

3.3 RESULTS AND DISSCUSSION

3.3.1 The relationship between salinity and rainfall

Pearson's correlation results (Table 3.1) and scatterplots (Figure 3.4) showed significant but weak positive correlations between rainfall and salinity. The results of the two precipitation variables (rain gauge and NEXRAD) are very similar. What was expected is that salinity is negatively correlated with rainfall. This is because salinity concentration should decrease as it is diluted by precipitation. In fact, most of the sampling days were dry days, and even though the correlation is positive, it is not strong enough to provide a clear explanation for this relationship. However, the reason for this positive relationship might be due to the soil particles that are carried by storm water runoffs, which may have increased the salinity in the river stream. This result is based on a monthly collected data for both of salinity and rainfall, and a daily data will help to explain more details for this relationship.

Table 3.1: Correlation test results for rainfall variables

Rainfall Variable	Pearson's correlation coefficient (r)	P- value < 0.05
Rain gauge	0.13	0.00002
NEXRAD	0.08	0.006962

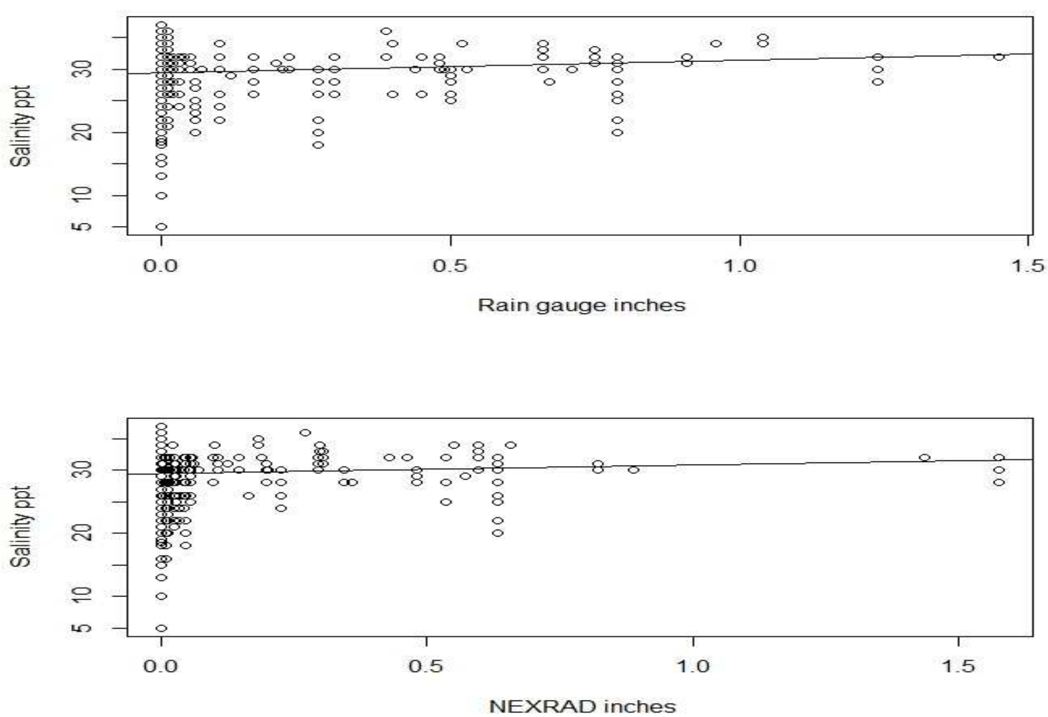

Figure 3.4: The relationship between salinity and rainfall

3.3.2 The relationship between salinity and FC

Salinity is one of the explanatory variables that will be used in predicting FC concentration. Pearson's correlation results (Table 3.2) and scatterplots (Figure 3.5) showed significant negative correlation with r-value of -0.33 and p-value of 2.2e-16. This significant linear negative relationship was expected because salinity influences the survival of FC. With higher salinity, less FC can survive in a fresh water body. This result is consistent with some previous studies that assessed the impact of salinity on FC concentration (Gonzalez et al.,2012; Gonzalez & Noble, 2014; Kelsey et al., 2010). Thus, it is important to include salinity concentration from each shellfish monitoring station in predicting FC concentration.

Table 3.2: Correlation test results for salinity and FC

Tested variables	Pearson's correlation coefficient (r)	P- value < 0.05
Salinity and FC	-0.33	2.2e-16

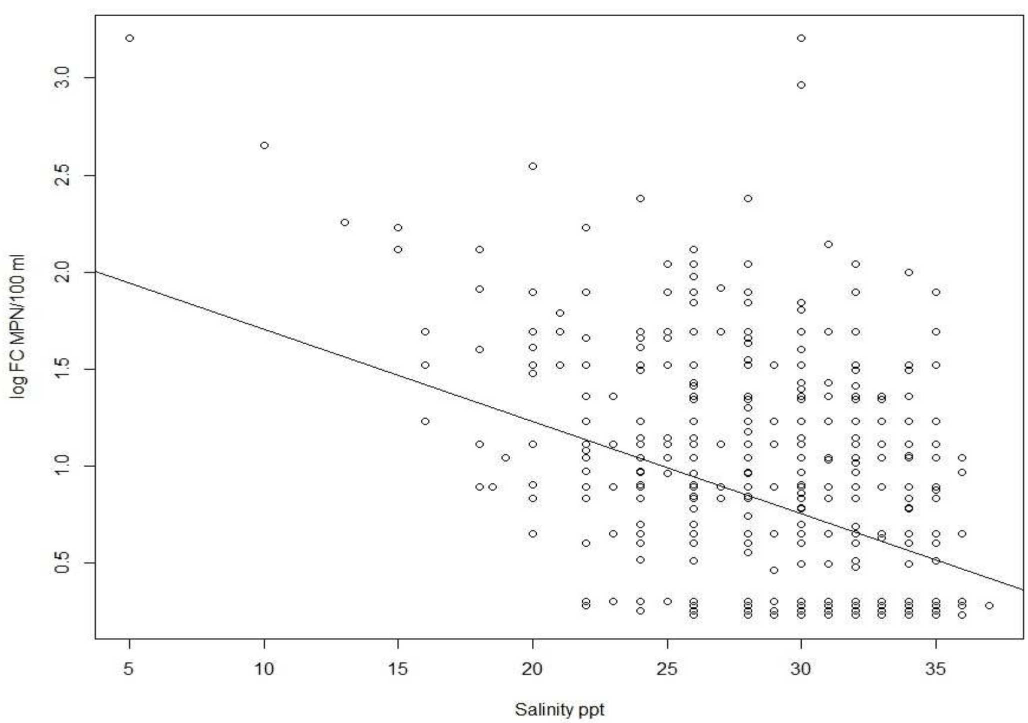

Figure 3.5: The relationship between log FC and salinity

3.3.3 Modeling LU and FC

3.3.3.1 Combined models

Linear Mixed Model (LMM) was used in developing 16 models, these are eight models for eight buffer sizes for two precipitation datasets (rain gauge and NEXRAD). Marginal R^2 values were used in comparisons between the models to evaluate models' performance for eight models for rain gauge and for eight models for NEXRAD (Table 3.3 and Table 3.4). The marginal R^2 value was the highest with the 2500m-buffer followed by the 2000m and then by the 1800m-buffer. The marginal R^2 for the models with rain gauge for all buffer sizes ranged from 0.395 to 0.567, and from 0.387 to 0.561 with NEXRAD models.

The p-values for the models' parameter estimates (coefficients) were also compared to evaluate the models' significance ($p < 0.05$) with rain gauge (Table 3.5) and NEXRAD (Table 3.6). It was found that the models (rain gauge and NEXRAD) for the 1800m-buffer were the only two models that had significant p-values for their coefficients. This means that FC concentration was significantly predicted by the LU variables as well as other variables that were retained in the models of the 1800m-buffer. Furthermore, it was shown that the p-value for the intercept was decreasing as the buffer size increases until it reached to less than 0.05 with 1800m-buffer models, which then returned to increase with the 2000m-buffer. The coefficient for the non-forested wetlands variable was never significant at any model in which it was included. It is interesting to find that the 1800m-buffer models (with rain gauge and NEXRAD) did not include the non-forested wetlands variable. However, the non-forested wetlands variable was included in the 2000m- buffer's model. In fact, the non-forested wetlands variable appeared to have negative correlation with FC in chapter 2 while it was always positive with all the models in which it was retained. This opposite correlation of this variable in the models might be the reason why it was not significant at any model, which may explain why it was not included in the 1800m-buffer's model. As for the 2500m-buffer's models (with rain gauge and NEXRAD), the signs of the coefficients of the golf course variable were changed from positive to negative; this negative correlation was not expected with FC concentration because golf course was

Table 3.3: Marginal R^2 values for the models with different buffer sizes (gauge)

Buffer size of the model	Marginal R^2
500m Gauge	0.395
800 m Gauge	0.453
1000 m Gauge	0.494
1200 m Gauge	0.518
1500 m Gauge	0.536
1800 m Gauge	0.546
2000 m Gauge	0.555
2500 m Gauge	0.567

Table 3.4: Marginal R^2 values for the models with different buffer sizes (NEXRAD)

Buffer size of the model	Marginal R^2
500 m NEXRAD	0.387
800 m NEXRAD	0.445
1000 m NEXRAD	0.486
1200 m NEXRAD	0.510
1500 m NEXRAD	0.528
1800 m NEXRAD	0.540
2000 m NEXRAD	0.549
2500 m NEXRAD	0.561

Table 3.5: linear mixed model parameters and their p-values for each buffer size (rain gauge)

Model parameters	P-values (< 0.05) for buffer sizes (Rain gauge)							
	500 m	800 m	1000 m	1200 m	1500 m	1800 m	2000 m	2500 m
Intercept	0.5476	0.7287	0.6596	0.1453	0.0918	**0.000**	0.1141	**0.000**
open space	0.6144	0.6505	0.2018	**0.0343**	**0.0146**	**0.0023**	**0.0194**	**0.0003**
open water	0.6360	0.9859	0.8079	-	-	-	-	-
residential	0.8126	0.2909	0.1430	**0.0415**	**0.0219**	**0.0031**	**0.0072**	**0.0012**
forestlands	0.8037	0.5623	0.1369	0.0702	**0.0438**	**0.0083**	**0.0038**	**0.0003**
non-forested wetlands	0.8142	0.3302	0.3254	0.2666	0.2976	-	0.1733	-
golf course	0.6830	0.8392	0.8692	0.6704	-	-	-	0.0542
water temp.	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
air temp.	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
salinity	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
gauge 24hrs	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**

Bold: significant ($p < 0.05$)

Table 3.6: linear mixed model parameters and their p-values for each buffer size (NEXRAD)

Model parameters	P-values (< 0.05) for buffer sizes (NEXRAD)							
	500 m	800 m	1000 m	1200 m	1500 m	1800 m	2000 m	2500 m
Intercept	0.5626	0.8817	0.8345	0.2554	0.1952	**0.0000**	0.3094	**0.0000**
open space	0.6080	0.5728	0.1744	**0.0318**	**0.0146**	**0.0025**	**0.0239**	**0.0004**
open water	0.6361	0.9017	0.7385	-	-	-	-	-
residential	0.8067	0.2734	0.1461	0.0552	**0.0312**	**0.0055**	**0.0103**	**0.0026**
forestlands	0.7675	0.5461	0.1329	0.0809	0.0501	**0.0078**	**0.0032**	**0.0003**
non-forested wetlands	0.8076	0.3280	0.3319	0.3041	0.3317	-	0.1700	-
golf course	0.6498	0.8270	0.8744	0.7060	-	-	-	0.0613
salinity	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**
NEXRAD 24hrs	**0.0004**	**0.0004**	**0.0004**	**0.0004**	**0.0004**	**0.0004**	**0.0004**	**0.0004**
NEXRAD 72hrs	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**

Bold: significant ($p < 0.05$)

shown to have positive correlation with FC in Chapter 2 as well as some other studies. Therefore, these results suggest that the circular 1800m-buffer is the optimum buffer size for using the LU variables in predicting FC concentration.

While some of the studies that used the circular buffers found that 400 to 500m buffer size were the most appropriate (Chang, 2008; Daugomah et al., 2007; Sanger et al., 2004), it is not surprising that the 1800m was found the most appropriate in this study for some reasons. The first reason is that the shellfish monitoring stations are located in the river's stream, so at 400m or 500m-buffer, most of the LU percentage in the buffer is going to be water. The second reason is that as the buffer size increases and because it is circular, more LU classes will be included form all the directions, which may be advantageous in explaining more detailed relationships between FC concentration, LU percentages and the environmental factors.

After model selection was completed for the models of all buffer sizes with rain gauge dataset, the following variables retained in all models: open space, residential, forestlands, water temperature, air temperature, salinity, tide stage, gauge-48hrs. The models for the buffers: 500m,800m and 1000m retained all the variables except for gauge-24hrs and gauge-72hrs because they were never shown in any model. Golf course disappeared in the model for the 1500m-buffer and it was shown again with negative sign at 2500m-buffer.

With NEXRAD dataset and after model selection was completed for the models of all buffer sizes, the following variables retained in all model: open space, residential, forestlands, salinity, tide stage, NEXRAD-24hrs, and NEXRAD-72hrs. Buffers 500m,800m and 1000m retained all the variables except for NEXRAD-24hrs, air temperature and water temperature because they were never shown in any model. However, air temperature and water temperature and gauge24hrs were involved in all rain gauge models. This might have occurred because all NEXRAD models included

NEXRAD-24hrs and NEXRAD-72hrs measurements in their models, which may would cause multicollinearity if the air temperature and water temperature measurements were included in the model. Like what happened with rain gauge models, with NEXRAD models, golf course disappeared at 1500m-buffer and it was shown again with negative sign at 2500m-buffer. These results suggest that open space, residential, forestlands, salinity, tide stage and rainfall measurements are important predictors for FC concentration in May River. These results are consistent with several studies that modeled FC concentration with LU and environmental factors (Cha et al., 2016; Crowther et al., 2003; Galfi et al., 2016; Paule-Mercado et al., 2016; Schoonover & Lockaby, 2006).

For the 1800m-buffer and rain gauge data, the following variables were involved in the model: open space, residential, forestlands, water temperature, air temperature, salinity, rain gauge-24hrs, and tide stage (Table 3.7). The model for 1800m-buffer with NEXRAD data included the following variables: open space, residential, forestlands, salinity, NEXRAD-24hrs, and NEXRAD-72hrs and tide stage (Table 3.7). Both models (rain gauge and NEXRAD) retained open space, residential, forestlands, salinity, and precipitation measurements as significant predictors in the models. This indicates that FC bacteria loading is controlled by the storm-water runoff from LU classes, and their survival is affected by salinity.

Salinity is negatively correlated with FC bacteria and this might because E. coli bacteria cannot survive in high-salinity waters. The marginal R^2 values for the 1800m-buffer models (rain gauge and NEXRAD) indicate that about 0.54 to 0.55 % of FC variations was explained by the variables in the selected models. These values seem to suggest that predicting FC bacteria with rain gauge and NEXRAD are very similar.

Table 3.7: FC prediction models using 1800m-buffer for LU variables

Model	Equation
Rain gauge	0.93+5.4 (open space) +1.18 (residential) + 2.5 (forestlands)-0.02 (water temp) +0.014 (air temp) - 0.019 (salinity) +0.389 (gauge-48hrs) + f (tide stage)
NEXRAD	0.78+5.5 (open space) +1.1(residential) +2.6 (forestlands)-0.02 (salinity) + 0.21 (NEXRAD-24hrs) + 0.11 (NEXRAD-72hrs) + f(tide stage)

f: Factor for tide's categorical variables

3.3.3.2 Seasonal models

There was an intent to develop two seasonal models based on wet and dry seasons. However, clustered column charts (Figure 3.6 and 3.7) for the rain gauge and NEXRAD datasets indicated that the climate in the May River has no wet and dry seasons. Therefore, four seasonal models were developed for each rainfall datasets using the 1800m-buffer for LU variables. The marginal R^2 values (Table 3.7) and the p-values (Table 3.8 and Table 3.9) for the models' parameters were compared to determine the significance levels and the prediction capabilities for the seasonal models.

The winter models for rain gauge and NEXRADS are the most predictive models and are the only two models that have significant coefficients for the parameters of the models (Table 3.8 and Table 3.9). The marginal R^2 values for all seasonal models ranged from 0.52 to 0.67 (Table 3.7). For winter models, the marginal R^2 value with rain gauge is 0.673, and with NEXRAD, the marginal R^2 is 0.665 (Table 3.7). The seasonal models with rain gauge had open space, residential and forestlands included in all the models. Unlike the combined models, the retained rainfall measurement variables were varied with the

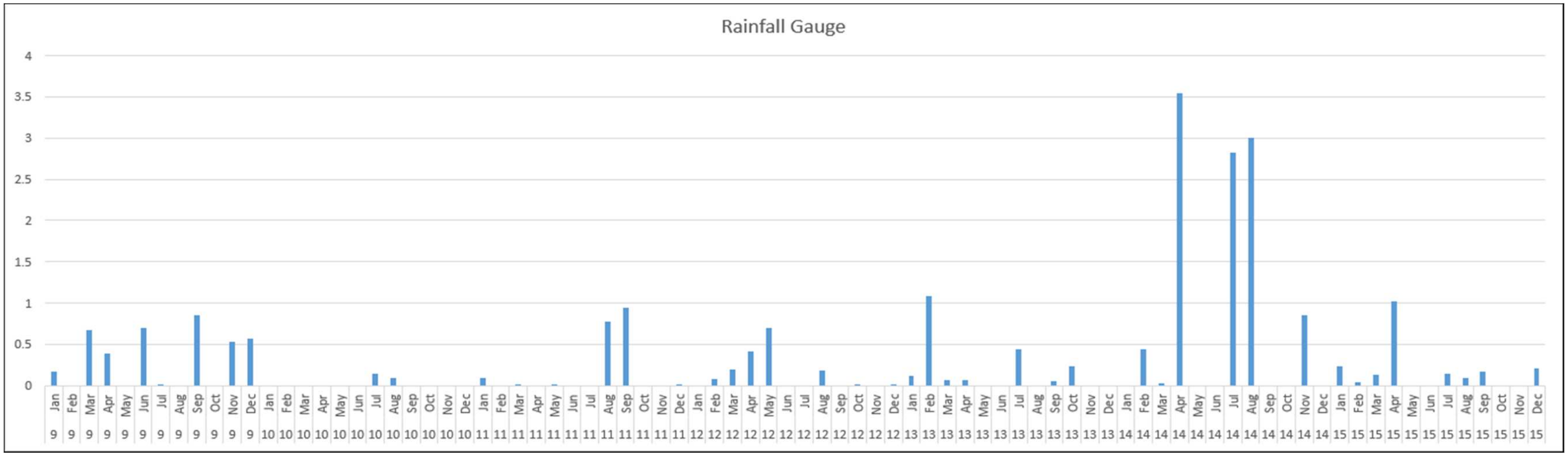

Figure 3.6: Monthly precipitation inches from 2009 to 2015 measured by rain gauge

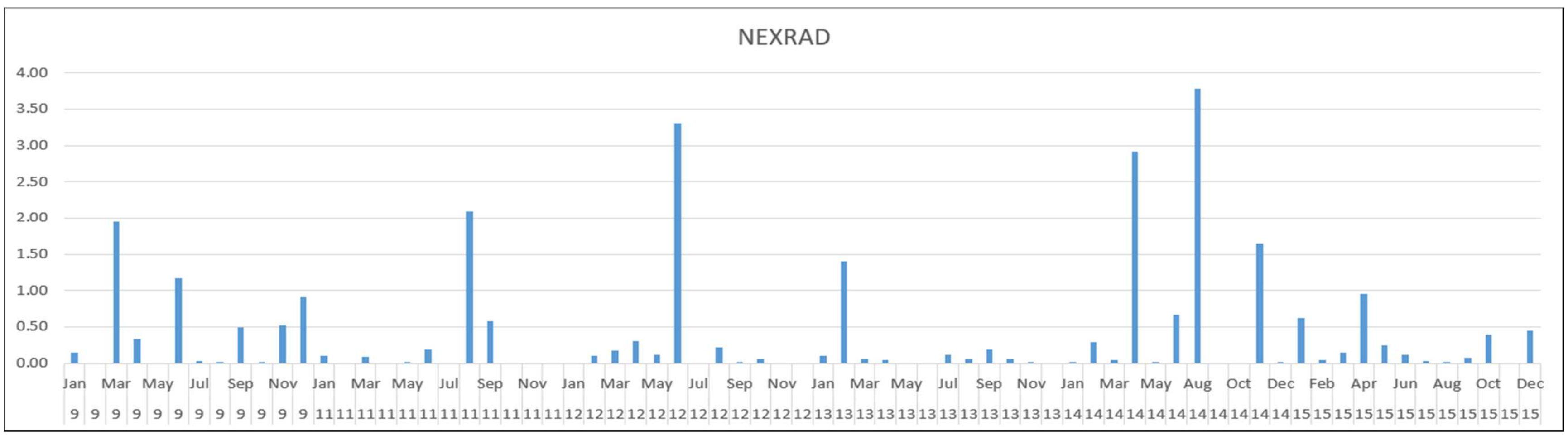

Figure 3.7: Monthly precipitation inches from 2009 to 2015 measured by NEXRAD

Table 3.8: Marginal R^2 values for seasonal models for 1800m-buffer (rain gauge and NEXRAD)

Seasonal models for buffer 1800 m	Marginal R^2
Fall Gauge	0.578
Fall NEXRAD	0.618
Spring Gauge	0.569
Spring NEXRAD	0.550
Summer Gauge	0.520
Summer NEXRAD	0.516
Winter Gauge	0.673
Winter NEXRAD	0.665

Table 3.9: linear mixed model parameters and their p-values for seasonal models (rain gauge)

Model parameters	P-values (< 0.05) for seasonal models (gauge)			
	Fall	Spring	Summer	Winter
Intercept	0.5950	0.4205	0.3247	**0.0000**
open space	**0.0057**	**0.0006**	**0.0145**	**0.0128**
open water	-	-	-	-
residential	**0.0114**	**0.0071**	**0.0043**	**0.0078**
forestlands	**0.0074**	**0.0012**	**0.0081**	**0.0464**
non-forested wetlands	-	0.1705	**0.0961**	-
golf course	-	-	-	-
salinity	-	**0.0008**	-	**0.0000**
air temp	-	-	**0.0012**	**0.0000**
gauge 24hrs	**0.0000**	-	-	-
gauge 48hrs	-	-	**0.0000**	**0.0000**
gauge 72hrs	-	**0.0000**	-	-

Bold: significant ($p < 0.05$)

seasonal models. For example, the fall model had gauge-24hrs, and the spring model had gauge-72hrs, while summer and winter models had gauge-72hrs. With NEXRAD, all the seasonal models retained the following variables: open space, residential, and forestlands. Salinity was only shown in winter model. Like the seasonal models with rain gauge, every seasonal model with NEXRAD retained only one rainfall measurement, which varied with the seasons. However, open water and golf course were not retained in any season with neither rain gauge nor NEXRAD models. This seems to emphasize that open water and golf course are not important predictors for FC when buffer 1800m is used.

Table 3.10: linear mixed model parameters and their p-values for seasonal models (NEXRAD)

Model parameters	P-values (< 0.05) for seasonal models (NEXRAD)			
	Fall	Spring	Summer	Winter
Intercept	**0.0090**	0.2468	0.1889	**0.0000**
open space	**0.0056**	**0.0005**	**0.0186**	**0.0202**
open water	-	-	-	-
residential	**0.0086**	**0.0091**	**0.0073**	**0.0147**
forestlands	**0.0087**	**0.0016**	**0.0091**	**0.0422**
non-forested wetlands	-	0.1952	0.1231	-
golf course	-	-	-	-
salinity	-	-	-	**0.0000**
air temp	-	-	**0.0003**	**0.0000**
NEXRAD 24hrs	**0.0000**	-	**-**	**0.0000**
NEXRAD 48hrs	-	-	**0.0000**	-
NEXRAD 72hrs	-	**0.0000**	-	-

Bold: significant ($p < 0.05$)

The equations for FC predictions for the seasonal models are shown in Table 3.10 and Table 3.11. With rain gauge models, tide stage was only retained in winter model, whereas with NEXRAD models, only fall and winter models retained the tide stage variable. Non-forested wetlands variable was not significant at any seasonal model, and it was not involved in both winter models (rain gauge and NEXRAD), which were found to be the most predictive models among the other seasonal models.

Data for precipitation and environmental factors for a longer period is needed to get more precise combined and seasonal models. In this study, FC concentrations were mostly sampled in dry days. Thus, precipitation data for a longer period can provide more information about the impact of rainfall on FC levels. The prediction power with rain gauge and NEXRAD were found to be very similar, thus, the NEXRAD is advantageous over the rain gauge since it can be accessed with no cost.

Table 3.11: Seasonal FC prediction models for 1800 meters circular buffer and rain gauge

Model	Equation
Fall	0.041+5.5 (open space) +1.12 (residential)+3.05 (forestlands)+0.42 (gauge-24hrs)
Spring	0.26+4.6 (open space) +1.6 (residential)+3.7 (forestlands) + 0.74 (non-forested wetlands)- 0.01 (salinity) + 0.14 (gauge-72hrs)
Summer	0.40+4.62 (open space) + 2.08 (residential) + 3.13 (forestlands) +1.06 (non-forested wetlands) - 0.03 (air temp) +0.27 (gauge-48hrs)
Winter	1.12+4.36 (open space) +1.1 (residential) +1.9 (forestlands) +0.03 (air temp)-0.04 (salinity) + 0.60 (gauge48hrs) + f (tide stage)

f: Factor for tide's categorical variables

Table 3.12: Seasonal FC prediction models for 1800 meters circular buffer and NEXRAD

Model	Equation
Fall	0.23+5.7 (open space) +1.21 (residential) +3.02 (forestlands) + 0.36 (NEXRAD-24hrs) + f (Tide stage)
Spring	-0.34+5.14 (open space) + 1.71 (residential) +3.85 (forestlands) +0.77 (non-forested wetlands) + 0.16 (NEXRAD-72hrs)
Summer	0.55 + 4.6 (open space) +2.0 (residential) + 3.2 (forestlands) +1.01(non-forested wetlands) - 0.03 (air temp) +0.15 (NEXRAD-48)
Winter	1.08+4.14 (open space) + 1.02 (residential) +2.02 (forestlands) + 0.03 (air temp) - 0.04 (salinity) + 0.78 (NEXRAD-24hrs) + f (tide stage)

f: Factor for tide's categorical variables

The knowledge for best management practices for LU impact on FC in the May River can be improved by further researches on other environmental factors such as the volume of the stream inside the buffer, dissolved oxygen (DO), PH, turbidity, and alkalinity. Including these parameters in further researches can be crucial in validating the size of the buffer for future studies on FC levels on the May River.

3.4 CONCLUSION

This study attempted to develop combined and seasonal models to predict FC concentration by developing several circular buffers that vary in size. The percentages of LU classes in the buffers and environmental factors were included in FC modellings. The models' outcomes for all the buffers were compared to identify which buffer size can provide the best fitted and the most significant model. The results of this study indicated that LU variables for 1800m-buffer scale were better correlated with FC when compared with the other buffer sizes. The LU variables for 1800m-buffer incorporated with environmental factors were able to provide significant models for FC prediction. Thus, it is important to consider LU impact in water quality modeling and in management practices. The seasonal models demonstrated the impact of seasonal variations on FC modelling as each season has a different model. This study suggests that LMM can be used to model the associations between water quality and LU as well as other environmental parameters when there are repeated samplings for the sampling locations.

REFERENCES

Azyana, Y., & Norulaini, N. N. (2012). The entire catchment and site buffer radii landscape variables, urban land use as predictors of water quality variation. *International Journal of Environmental Science and Development*, *3*(2), 141.

Barber, B. P. (2008). Town of Bluffton May River watershed monitoring program. *Prepared for Town of Bluffton, Bluffton, SC. 143pp*.

Brooks, W., Corsi, S., Fienen, M., & Carvin, R. (2016). Predicting recreational water quality advisories: A comparison of statistical methods. *Environmental Modelling & Software*, *76*, 81–94. https://doi.org/10.1016/j.envsoft.2015.10.012

Cha, Y., Park, M.-H., Lee, S.-H., Kim, J. H., & Cho, K. H. (2016). Modeling spatiotemporal bacterial variability with meteorological and watershed land-use characteristics. *Water Research*, *100*, 306–315. https://doi.org/10.1016/j.watres.2016.05.024

Chang, H. (2008). Spatial analysis of water quality trends in the Han River basin, South Korea. *Water Research*, *42*(13), 3285–3304.

Cho, K. H., Pachepsky, Y. A., Kim, M., Pyo, J., Park, M.-H., Kim, Y. M., … Kim, J. H. (2016). Modeling seasonal variability of fecal coliform in natural surface waters using the modified SWAT. *Journal of Hydrology*, *535*, 377–385.

Christen, K. (2002). Making accurate beach water-quality determinations. *Environmental Science & Technology*, *36*(19), 368A–369A.

Crowther, JMDW, Wyer, M. D., Bradford, M., Kay, D., & Francis, C. A. (2003). Modelling faecal indicator concentrations in large rural catchments using land use and topographic data. *Journal of Applied Microbiology*, *94*(6), 962–973.

Crowther, John, Hampson, D. I., Bateman, I. J., Kay, D., Posen, P. E., Stapleton, C. M., & Wyer, M. D. (2011). Generic Modelling of Faecal Indicator Organism Concentrations in the UK. *Water*, *3*(2), 682–701. https://doi.org/10.3390/w3020682

Daugomah, J. W., Siewicki, T. C., Porter, D. E., & Scott, G. I. (2007). The effects of urbanization on populations of grass shrimp Palaemonetes spp. in small, high salinity estuaries.

David, M. M., & Haggard, B. E. (2011). Development of regression-based models to predict fecal bacteria numbers at select sites within the Illinois River Watershed, Arkansas and Oklahoma, USA. *Water, Air, & Soil Pollution*, *215*(1–4), 525–547.

Delpla, I., & Rodriguez, M. J. (2014). Effects of future climate and land use scenarios on riverine source water quality. *Science of the Total Environment*, *493*, 1014–1024.

Dufour, A. P. (1984). Health effects criteria for fresh recreational waters. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=US201300708833

Francy, D. S., Darner, R. A., & Bertke, E. E. (2006). Models for predicting recreational water quality at Lake Erie beaches. Scientific Investigations Report 2006–5192. *US Geological Survey, Columbus, OH*.

Francy, D. S., & Darner, R. A. (2007). *Nowcasting beach advisories at Ohio Lake Erie beaches*. Geological Survey (US). Retrieved from https://pubs.er.usgs.gov/publication/ofr20071427

Frick, W. E., Ge, Z., & Zepp, R. G. (2008). Nowcasting and forecasting concentrations of biological contaminants at beaches: A feasibility and case study. *Environmental Science & Technology*, *42*(13), 4818–4824.

Galfi, H., Österlund, H., Marsalek, J., & Viklander, M. (2016). Indicator bacteria and associated water quality constituents in stormwater and snowmelt from four urban catchments. *Journal of Hydrology*, *539*, 125–140. https://doi.org/10.1016/j.jhydrol.2016.05.006

Gonzalez, R. A., Conn, K. E., Crosswell, J. R., & Noble, R. T. (2012). Application of empirical predictive modeling using conventional and alternative fecal indicator bacteria in eastern North Carolina waters. *Water Research*, *46*(18), 5871–5882. https://doi.org/10.1016/j.watres.2012.07.050

Gonzalez, R. A., & Noble, R. T. (2014). Comparisons of statistical models to predict fecal indicator bacteria concentrations enumerated by qPCR-and culture-based methods. *Water Research*, *48*, 296–305.

Gyawali, S., Techato, K., & Monprapusson, S. (2015). Assessing the Impact of Land Use on Water Quality across Multiple Spatial Scales in U-tapao River Basin, Thailand. *Journal of Institute of Science and Technology*, *20*(2), 54–60.

He, L.-M. L., & He, Z.-L. (2008). Water quality prediction of marine recreational beaches receiving watershed baseflow and stormwater runoff in southern California, USA. *Water Research*, *42*(10), 2563–2573.

Herrig, I. M., Böer, S. I., Brennholt, N., & Manz, W. (2015). Development of multiple linear regression models as predictive tools for fecal indicator concentrations in a stretch of the lower Lahn River, Germany. *Water Research*, *85*, 148–157. https://doi.org/10.1016/j.watres.2015.08.006

Hunter, P. R. (1997). Drinking water and waterborne diseases. *Waterborne Disease: Epidemiology and Ecology*, 27–41.

Hurley, T., & Mazumder, A. (2013). Spatial scale of land-use impacts on riverine drinking source water quality. *Water Resources Research*, *49*(3), 1591–1601.

Kelsey, R. H., Scott, G. I., Porter, D. E., Siewicki, T. C., & Edwards, D. G. (2010). Improvements to shellfish harvest area closure decision making using GIS, remote sensing, and predictive models. *Estuaries and Coasts*, *33*(3), 712–722.

Leight, A. K., Hood, R., Wood, R., & Brohawn, K. (2016). Climate relationships to fecal bacterial densities in Maryland shellfish harvest waters. *Water Research*, *89*, 270–281. https://doi.org/10.1016/j.watres.2015.11.055

Moody, R. (2016). Shellfish Management Area 19-2016 Annual Update. Submitted by the Shellfish Sanitation Section, Water Monitoring, Assessment and Protection Division, Environmental Quality Control–Bureau of Water, South Carolina Department of Health and Environmental Control, Columbia, SC. Retrieved from http://scdhec.gov/FoodSafety/Docs/sfrp_19.pdf

Nevers, M. B., & Whitman, R. L. (2005). Nowcast modeling of Escherichia coli concentrations at multiple urban beaches of southern Lake Michigan. *Water Research*, *39*(20), 5250–5260.

Olyphant, G. A., Thomas, J., Whitman, R. L., & Harper, D. (2003). Characterization and statistical modeling of bacterial (Escherichia coli) outflows from watersheds that discharge into southern Lake Michigan. In *Coastal Monitoring through Partnerships* (pp. 289–300). Springer.

Olyphant, G. A., & Whitman, R. L. (2004). Elements of a predictive model for determining beach closures on a real time basis: the case of 63rd Street Beach Chicago. *Environmental Monitoring and Assessment*, *98*(1–3), 175–190.

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

Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: basic concepts and examples. *Mixed-Effects Models in S and S-Plus*, 3–56.

Ragosta, G., Evensen, C., Atwill, E. R., Walker, M., Ticktin, T., Asquith, A., & Tate, K. W. (2010). Causal connections between water quality and land use in a rural tropical island watershed. *EcoHealth*, *7*(1), 105–113.

Sanger, D. M., Holland, A. F., & Hernandez, D. L. (2004). Evaluation of the impacts of dock structures and land use on tidal creek ecosystems in South Carolina estuarine environments. *Environmental Management*, *33*(3), 385–400.

S.C. Dept. of Health and Environmental Control (SCDHEC). (2017a). DHEC: Food Safety Shellfish Monitoring. Retrieved January 31, 2018, from http://www.scdhec.gov/FoodSafety/ShellfishMonitoring/

S.C. Dept. of Health and Environmental Control (SCDHEC). (2017b). Regulation 61-47, shellfish. Retrieved from http://www.scdhec.gov/Agency/docs/water-regs/r61-47.pdf

Schoonover, J. E., & Lockaby, B. G. (2006). Land cover impacts on stream nutrients and fecal coliform in the lower Piedmont of West Georgia. *Journal of Hydrology*, *331*(3), 371–382.

Thoe, W., Gold, M., Griesbach, A., Grimmer, M., Taggart, M. L., & Boehm, A. B. (2014). Predicting water quality at Santa Monica Beach: evaluation of five different models for public notification of unsafe swimming conditions. *Water Research*, *67*, 105–117.

Thoe, Wai, & Lee, J. H. (2013). Daily forecasting of Hong Kong beach water quality by multiple linear regression models. *Journal of Environmental Engineering*, *140*(2), 04013007.

Thoe, Wai, Wong, S. H., Choi, K. W., & Lee, J. H. W. (2012). Daily prediction of marine beach water quality in Hong Kong. *Journal of Hydro-Environment Research*, *6*(3), 164–180.

United States Environmental Protection Agency (USEPA). (1986). Ambient Water Quality Criteria for Bacteria. Retrieved from http://www.ugra.org/pdfs/EPA1986.pdf

United States Environmental Protection Agency (USEPA). (2007). Critical Path Science Plan for the Development of New or Revised Recreational Water Quality Criteria. Retrieved from https://nepis.epa.gov/Exe/ZyNET.exe

United States Environmental Protection Agency (USEPA). (2010). Predictive Tools for Beach Notification Volume I: Review and Technical Protocol. Retrieved from https://www.epa.gov/sites/production/files/2015-11/documents/predictive-tools-volume1.pdf

United States Environmental Protection Agency (USEPA). (2012). Recreational Water Quality Criteria. Retrieved from https://www.epa.gov/sites/production/files/2015-10/documents/rwqc2012.pdf

U.S Census Bureau. (2015). 2015 Population Estimate for Town of Bluffton. Retrieved January 26, 2017, from //www.census.gov/quickfacts/table/PST045215/4507210

U.S. Food and Drug Administration (FDA). (1993). National Shellfish Sanitation Program Manual of Operations: Part I, Sanitation of Shellfish Growing Areas. Retrieved from http://oceanservice.noaa.gov/websites/retiredsites/sotc_pdf/SGW.PDF

Whitman, R. L., & Nevers, M. B. (2008). Summer E. coli patterns and responses along 23 Chicago beaches. *Environmental Science & Technology*, *42*(24), 9217–9224.

Whitman, R. L., Nevers, M. B., & Gerovac, P. J. (1999). Interaction of ambient conditions and fecal coliform bacteria in Southern Lake Michigan beach waters: Monitoring program implications. *Natural Areas Journal*, *19*(2), 166–171.

Whitman, R. L., Nevers, M. B., Korinek, G. C., & Byappanahalli, M. N. (2004). Solar and temporal effects on Escherichia coli concentration at a Lake Michigan swimming beach. *Applied and Environmental Microbiology*, *70*(7), 4276–4285.

Zhai, X., Xia, J., & Zhang, Y. (2014). Water quality variation in the highly disturbed Huai River Basin, China from 1994 to 2005 by multi-statistical analyses. *Science of the Total Environment*, *496*, 594–606.

Zhang, Z., Deng, Z., & Rusch, K. A. (2012). Development of predictive models for determining enterococci levels at Gulf Coast beaches. *Water Research*, *46*(2), 465–474.

CHAPTER 4

PREDICTING THE IMPACT OF A FUTURE LAND USE PLAN ON SHELLFISH HARVESTING WATER QUALITY

ABSTRACT

This study assesses the impact of a future land use (LU) plan fecal coliform (FC) concentrations in the May River at the Town of Bluffton, South Carolina. Predictive models were used for two different rainfall scenarios. Average and maximum rainfall scenarios in addition to LU percentages as well as tide stage were used in FC prediction. Circular buffers that are 1800 meters in size were used as the special scale for the proposed LU percentages that were included in the models. The results of FC prediction for the future LU scenarios were compared with the predicted FC in 2015 for the same scenarios. The comparison indicated that FC loading from the future LU plan will not to be higher than FC level in 2015. The data for this research included LU data from the Town of Bluffton and FC concentrations from the South Carolina Department of Health and Environmental Control (SCDHEC). The results of this research recommend the use of the Linear Mixed Model (LMM) in developing predictive models in order to serve as decision support tool for water quality and land management.

4.1 INTRODUCTION

To maintain the quality of surface waters, it is important to understand the links between the conditions of waterbodies and land management practices. Land management can help in understanding the spatial and temporal variabilities of water pollutants. One of the life-threatening water pollutants for humans and aquatic creatures is fecal coliform. As prevention is always better than treatment, land management can provide the prevention required to limit the fecal coliform levels. Land use (LU) activities are among the most concerned issues in land management practices. A comprehensive understanding of how the hydrologic cycle is impacted by LU change is needed to achieve the optimum natural resources management (Scanlon, Reedy, Stonestrom, Prudic, & Dennehy, 2005). The LU became an important factor in studying their impacts on the environment and the ecosystem.

The variabilities of fecal coliform densities of the future can be predicted by analyzing the past LU changes, and many studies have been conducted to assess the impact of future LU on water quality (De Girolamo & Porto, 2012; Delpla & Rodriguez, 2014; Garmendia, Mariel, Tamayo, Aizpuru, & Zabaleta, 2012; Isik, Kalin, Schoonover, Srivastava, & Lockaby, 2013; Karlsson et al., 2016; Mango, Melesse, McClain, Gann, & Setegen, 2010; Neupane & Kumar, 2015; Tong, Liu, & Goodrich, 2009; Tu, 2009; Vaché, Eilers, & Santelmann, 2002; Wijesekara et al., 2012; Yira, Diekkrüger, Steup, & Bossa, 2016). These studies assessed the impacts of future (proposed) LU plan on water quality by using a scenario-based approach. Furthermore, these studies have used different methods and models for the assessments. Some of these studies have used physically-based scale models like SWAT and some others have used statistical models like Multiple Linear

Regression (MLR) and Artificial Neural Network (ANN). The scenario-based analysis can be utilized as a tool to examine the potential impacts for a future LU change (Verburg et al., 2008).

To manage the quality of surface waters, it is important to have robust modeling tools capable of assessing LU and management scenarios (Elliott et al., 2016). Developing LU change scenarios is an effective method in projecting the impact of future developments and providing tools to support policies and decisions that maintain sustainable strategies for decision makers (Nakicenovic et al., 2000). As a mean of mitigating the negative environmental impact of future LU change, the scenario-based modeling approach is necessary to provide the support in the analysis of the future LU impact (Sohl et al., 2012). Decisions in LU management can be supported if enough information concerning the impacts of future LU and climate change scenarios are available (Lin, Hong, Wu, & Lin, 2007).

Most of the studies that used scenario-based approach generated arbitrary future LU scenarios and evaluated their impacts. There are many physically-based watershed simulation models that have been developed to estimate the impact of future LU change scenarios. Some of these simulation models include: AGNPS (Young & Shepherd, 1995), ANSWERS (Beasley and Huggins 1982), HSPF (Bicknell, Imhoff, Kittle Jr, Donigian Jr, & Johanson, 1997), LTHIA (Bhaduri, Grove, Lowry, & Harbor, 1997) and SWAT (Arnold, Williams, Srinivasan, King, & Griggs, 1994). Some studies have compared the efficiencies between two different simulation models and found that it is hard to determine which model can perform better in terms of the statistical results (Niraula, Kalin, Srivastava, & Anderson, 2013; Sharifi et al., 2017). Several studies have used SWAT to assess the impact of future LU scenarios on the ecosystem including the impact on surface water quality (De Girolamo & Porto, 2012; Mancosu et al., 2015; Mango et. al, 2010; Mehdi, Ludwig, & Lehner, 2015; Neupane & Kumar, 2015; Tong, Sun, Ranatunga, He, & Yang, 2012; Vaché et. al, 2002). However, a study used Multi Linear Regression (MLR) to predict the impact of future LU scenarios on water quality (Garmendia et. al, 2012). Delpla & Rodriguez, (2014) used Linear Mixed Model (LMM) to predict FC and turbidity from multiple LU and climate scenarios.

Technically, to assess the impact of LU change on surface water quality, the current LU of a watershed is modeled, and then the models are used for future LU scenarios to compare the differences between the two modeled conditions (Elfert & Bormann, 2010).

Using the scenario-based approach with the integration of Geographical Information System (GIS) can serve as a powerful decision support tool for LU policy and management (Vaché et al., 2002). Numerous previous studies that assessed the impact of LU change on water quality, which used scenario-based approach, have proposed future LU scenarios based on "what if" scenario analysis (Bussi et al., 2016; Garmendia et al., 2012;Karlsson et al., 2016; Mango et al., 2010; Neupane & Kumar, 2015; Rajib, Ahiablame, & Paul, 2016; Sharifi et al., 2017; Wilson & Weng, 2011; Yira et. al, 2016). This study will assess the impact of a proposed LU plan on fecal coliform densities by using previously developed models.

4.2 STUDY AREA

The May River watershed is located in Beaufort County in south east of South Carolina (Figure 4.1). The Town of Bluffton has the May River as its significant estuary, which drains into the Atlantic Ocean on the eastern coast of South Carolina. The population of the Town of Bluffton in 2015 was about 16,728 according to the U.S Census Bureau, 2015. Due to the high water quality of the May River in 2001, it was classified Outstanding Resource Water (ORW) by the SCDHEC (Barber, 2008). The Environmental Protection

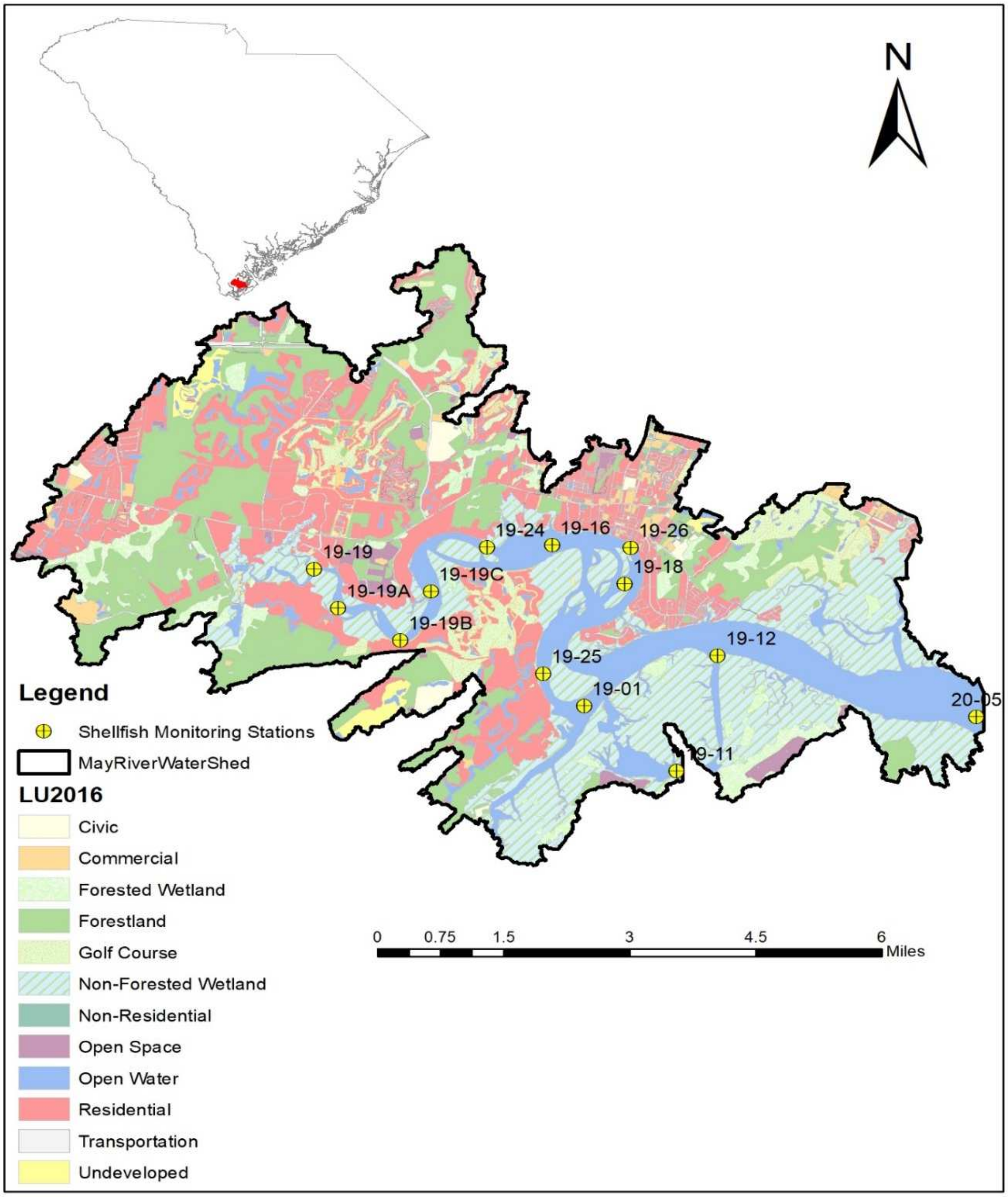

Figure 4.1: May River watershed with shellfish monitoring station sites and LU classes in 2016
Data source: The government of Town of Bluffton

Agency (EPA) and SCDHEC have acknowledged the May River as a priority watershed. The May River is highly valued by the residents of the Town of Bluffton due to its great recreational and economical importance. The aesthetics and views, and the abundant natural resources in the Town of Bluffton have led to an increase in population and commercial growth in the area. As a result, major LU changes are planned for developments by the government of the Town of Bluffton. The May River watershed is within Shellfish Growing Area 19 (Figure 4.2).

4.3 DATA SOURCES AND METHODS

4.3.1 Fecal Coliform Data

Data for FC concentrations were retrieved from SCDHEC's shellfish-monitoring program. The program was established primarily to maintain the health and quality standards of the shellfish and their harvesting areas by following federal guidelines and state regulations (SCDHEC, 2017a). Also, this program was established to enhance water quality for the shellfish harvesting areas.

Each shellfish growing area in South Carolina is comprehensively evaluated under this program. Annual evaluation is conducted for the shellfish growing areas that meet the requirements of the National Shellfish Sanitation Program (NSSP). Monthly routine sampling and laboratory analysis are conducted for bacteriological water quality monitoring at the SCDHEC's designated sampling sites. The SCDHEC's shellfish-monitoring program complies with the NSSP standards, sampling and monitoring methods, and laboratory analysis. Most Probable Number (MPN) per 100 milliliters (ml) is the measurement unit used for the FC concentrations. As the shellfish monitoring stations are within Approved and Restricted areas, the guidelines for these two shellfish areas are as

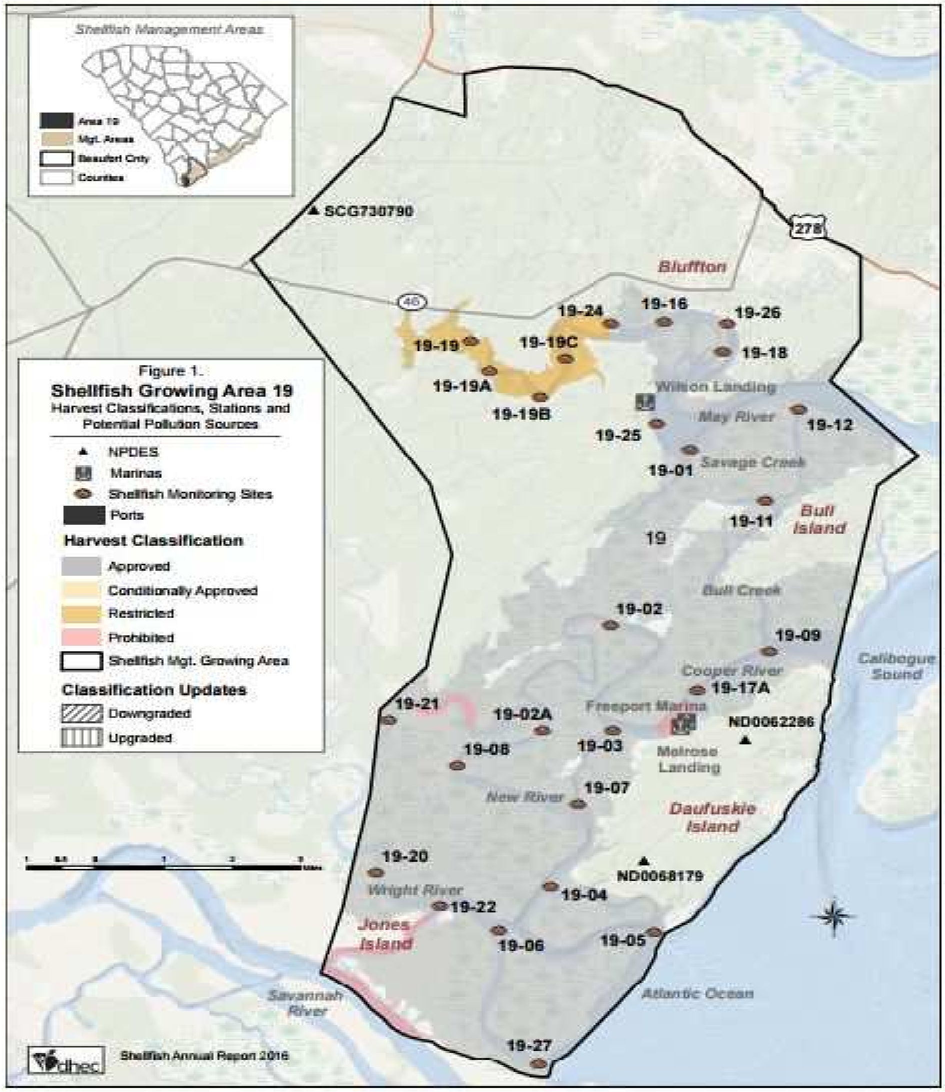

Figure 4.2: Shellfish Growing Area 19 map with SCDHEC shellfish monitoring stations
Source: SCDHEC

the following: 1- For the Approved area, the median or the geometric mean of FC should not exceed 14 MPN per 100 milliliters, 2- A shellfish area should be classified as Restricted when it is proven by the survey data that there are a reasonable level of pollutants exist or if there are substances deleterious or poisonous substances which can unpredictably change the water quality or when it is not possible to classify the area as Conditionally Approved (SCDHEC, 2017b). There are 13 stations located in the study area which will be included in this research. The data for FC concentrations from 1999 to 2015 are only available for

10 stations, while the data from three stations are available from 2009 to 2015. This is due to the fact that these three stations were installed in 2009.

4.3.2 Land use data

LU data for the for the proposed LU plan were retrieved from the governments of the Town of Bluffton (Figure 4.4). The data is a shapefile feature class that was developed by polygons that reflect the LU classes that were captured by the satellite imagery. The projected coordinate system for the data is NAD 1983 State Plane South Carolina FIPS 3900 Feet Intl, and the projection is Lambert Conformal Conic. The geographic coordinate system for the data is GCS North American 1983. Using ArcMap 10.4, the LU data were edited in order to be prepared for the analysis of the study. Three major LU types were used in the analysis: (1) residential areas, (2) forestlands, and (3) open space

4.3.3 Circular Buffers for land use data

In Chapter 3, it was found that 1800 meter is the optimum size for LU buffers. ArcMap 10.4 was used develop 1800m-circular buffers for each shellfish monitoring stations to calculate the percentages of each land use class for the proposed LU plan within the circular buffers (Figure 4.3). A shapefile feature class layer for the May River watershed was required for the buffers developments to clip the edges of the buffers within the watershed borders. The May River watershed’s layer used in the study is a 12-Digits Hydrologic Unit Code (HUC) that was retrieved from the government of the Town of Bluffton. For the purpose of this study, the percentages of the LU classes for each shellfish monitoring station are required to get the LU variables for the stations.

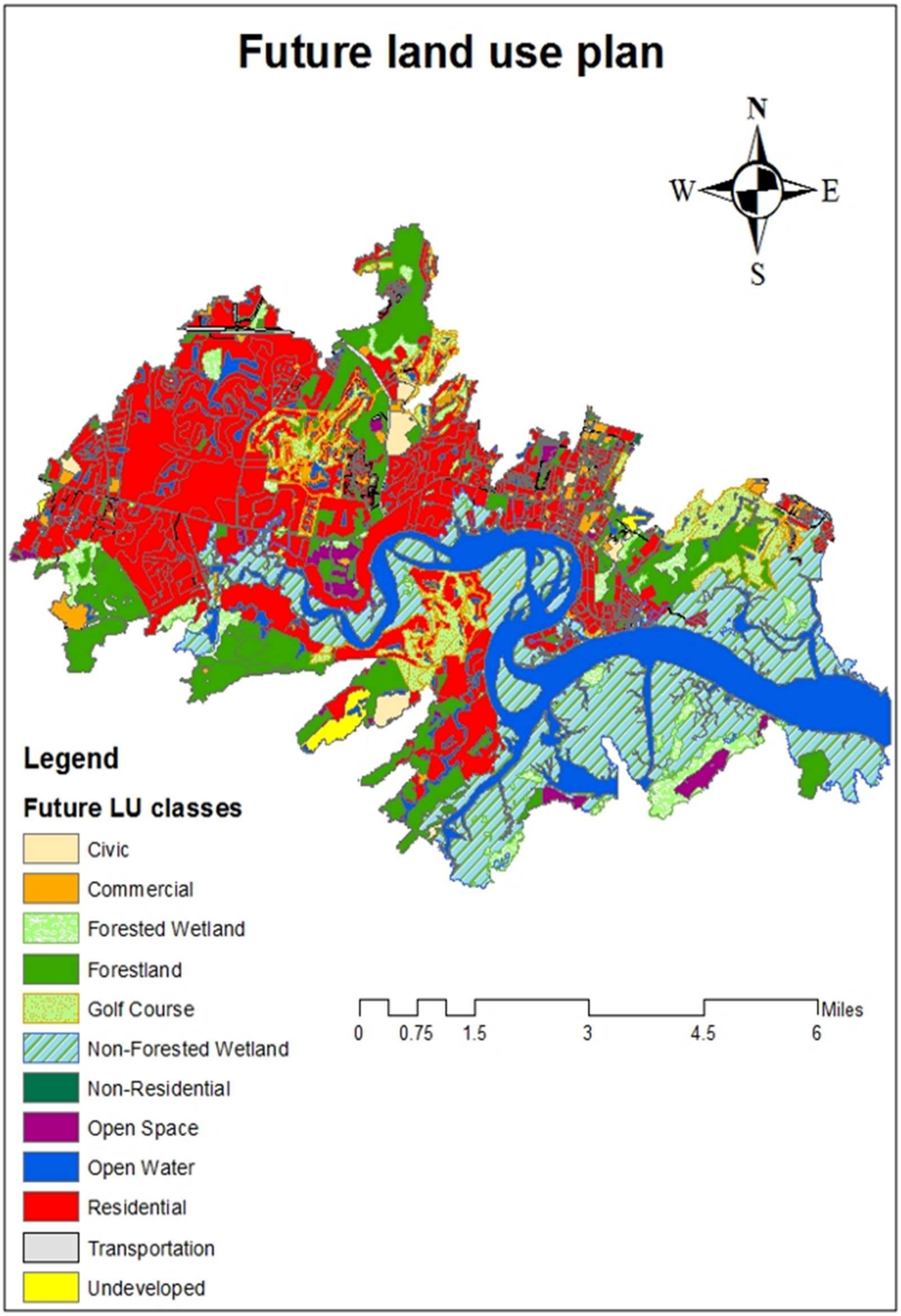

Figure 4.3: Future LU plan of the Town of Bluffton

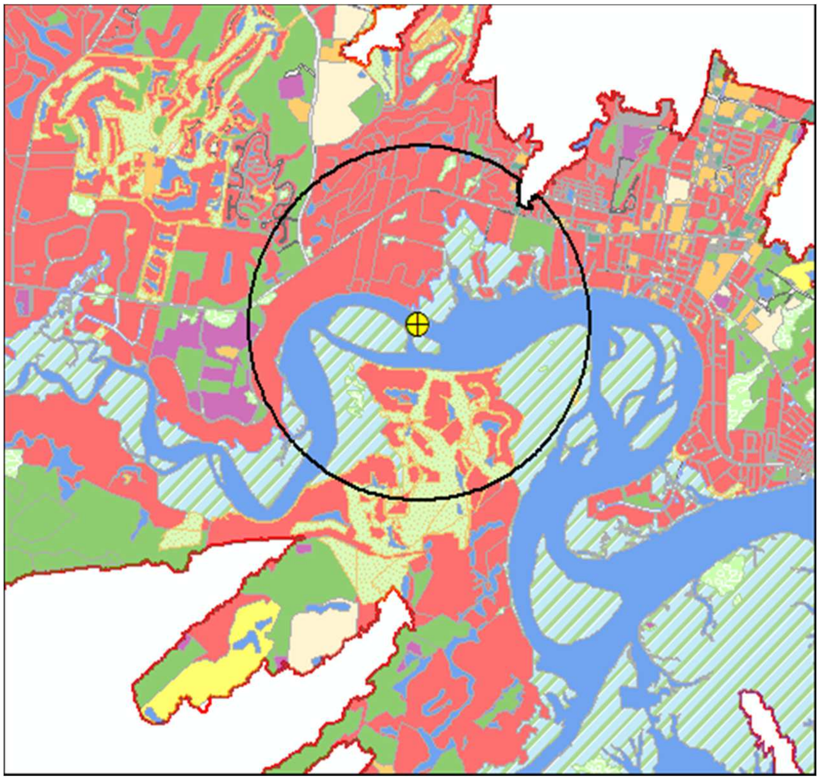

Figure 4.4: Circular buffer of 1800 meters for one of the shellfish monitoring stations

4.3.4 Precipitation and tides data

Two rainfall datasets were used for fecal coliform modeling. One dataset is from rainfall monitoring gauge for Broad Creek Public Service District (PSD). This is the rainfall data source that is used by the SCDHEC for shellfish monitoring assessments in the May River. The other rainfall dataset was retrieved from the Next Generation Weather Radar (NEXRAD) system from the National Atmospheric and Oceanic Administration (NOAA). Historical rainfall data from 2009 to 2015 were used in this study. Average and maximum rainfall events for Fall 2015 were calculated and used in the modeling. Tide stages data were retrieved from the SCDHEC's shellfish monitoring stations. The tide stages are categorical variables (eight categories) with eight levels.

4.3.5 Modeling methods

Two equations for FC from Chapter 3 were used in the modeling (Table 4.1). These two equations were developed for the fall season. One equation is for FC modeling using rain gauge data and the other equation is for FC modeling using NEXRAD precipitation data. The fall season model was selected for two reasons. The first reason is that the fall is the season for shellfish harvesting. The second reason is that there was a flood event in South Carolina in Fall 2015.

In order to compare FC levels of the future LU plan with their levels in the current LU, two future scenarios (with rain gauge and NEXRAD) were developed and compared with two modeled scenarios for FC in 2015. One scenario included average rainfall in fall 2015, and the other scenario included the maximum rainfall event in 2015. Therefore, FC levels of the average and maximum rainfall scenarios (with rain gauge and NEXRAD) for the future LU were compared with FC levels of the average and maximum rainfall scenarios for LU in 2015. For each scenario, the equation was used to calculate FC concentration for the 13 sampling locations in and then the average was calculated. Column charts were used to demonstrate the comparisons and to determine whether the future LU could cause water quality degradation

Table 4.2 Rain Gauge and NEXRAD models for the Fall season

Model	Equation
Rain Gauge	0.041+5.5 (open space) +1.12 (residential) +3.05 (forestlands) + 0.42 (Gauge-24hrs)
NEXRAD	0.23+5.7 (open space) +1.21 (residential) + 3.02 (forestlands) + 0.36 (NEX-24hrs) + f (Tide stage)

f: Factor for tide’s categorical variables

4.4 RESULTS AND DISCUSSION

For each shellfish monitoring station (Table 4.1), the modeled (predicted) log FC concentration in 2015 and the modeled future log FC concentration and their averages were calculated based on the average rainfall (Rain Gauge and NEXRAD). The average FC concentration for the modeled scenarios with average rainfall are shown in Figure 4.1. What can be clearly seen from the average rainfall models in Figure 4.1 is that FC levels from the modeled future LU, with Rain Gauge and NEXRAD, is slightly lower than FC levels from the 2015 modeled LU. By using maximum rainfall from Rain Gauge and NEXRAD, modeled FC concentration in 2015 and modeled future FC for each station and their averages were calculated (Table 4.2). Like the average rainfall models, the results of the maximum rainfall models show that FC levels of the modeled future LU are lower than FC levels of the modeled 2015 LU (Figure 4.2). Interestingly, similar results were found from the average and the maximum rainfall models.

An initial objective of the study was to assess the impact of the future LU plan and climate scenarios on fecal coliform levels. The results suggest that the future LU plan should not lead to a higher water quality degradation which may occur as a result of higher FC loading in the May River. This study suggests concerning LU, rainfall, and tide stage in predicting FC of a future LU plan. Contrary to expectations, the future LU plan will not lead to higher FC loadings although the plan will have higher residential areas, which were found positively correlated with FC in chapters 2 and 3. These results may be explained by the fact that most the proposed residential areas are far from the shellfish monitoring stations; and thus, their impact will not reach the river.

Table 4.2: Modeled FC concentration in 2015 and in the future with Average Rain (Gauge and NEXRAD)

Station ID	Modeled 2015 FC with Average Rain (Gauge)	Modeled Future FC Average Rain (Gauge)	Modeled 2015 FC Average Rain (NEXRAD)	Modeled Future FC Average Rain (NEXRAD)
19-19	1.37	1.35	1.61	1.59
19-19A	1.45	1.45	1.68	1.68
19-19B	1.10	1.10	1.31	1.31
19-19C	1.03	1.00	1.26	1.23
19-24	0.78	0.66	0.99	0.89
19-16	0.61	0.53	0.83	0.75
19-26	0.80	0.80	1.02	1.02
19-18	0.61	0.61	0.82	0.82
19-25	0.58	0.58	0.80	0.80
19-01	0.48	0.48	0.69	0.69
19-11	0.45	0.45	0.65	0.65
19-12	0.38	0.38	0.58	0.58
20-05	0.29	0.29	0.48	0.48
Average FC from all the stations	0.76	0.74	0.98	0.96

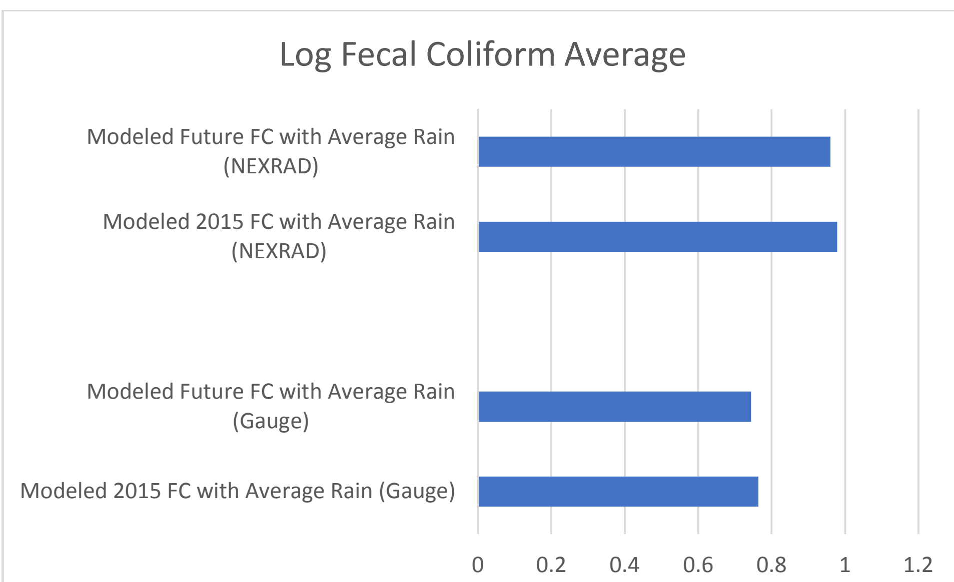

Figure 4.5: Average FC levels for Modeled 2015 and future FC with Average Rain

Table 4.3: Modeled 2015 and future FC with Maximum Rain (Gauge and NEXRAD)

Station ID	Modeled 2015 FC Maximum Rain (Gauge)	Modeled Future FC Maximum Rain (Gauge)	Modeled 2015 FC Maximum Rain (NEXRAD)	Modeled future FC Maximum Rain (NEXRAD)
19-19	1.89	1.86	2.23	2.20
19-19A	1.97	1.97	2.29	2.29
19-19B	1.61	1.61	1.93	1.93
19-19C	1.55	1.52	1.88	1.85
19-24	1.29	1.18	1.61	1.50
19-16	1.13	1.04	1.44	1.37
19-26	1.32	1.32	1.63	1.63
19-18	1.13	1.13	1.44	1.44
19-25	1.10	1.10	1.42	1.42
19-01	0.99	0.99	1.30	1.30
19-11	0.96	0.96	1.26	1.26
19-12	0.90	0.90	1.20	1.20
20-05	0.80	0.80	1.09	1.09
Average FC from all the stations	1.28	1.26	1.59	1.58

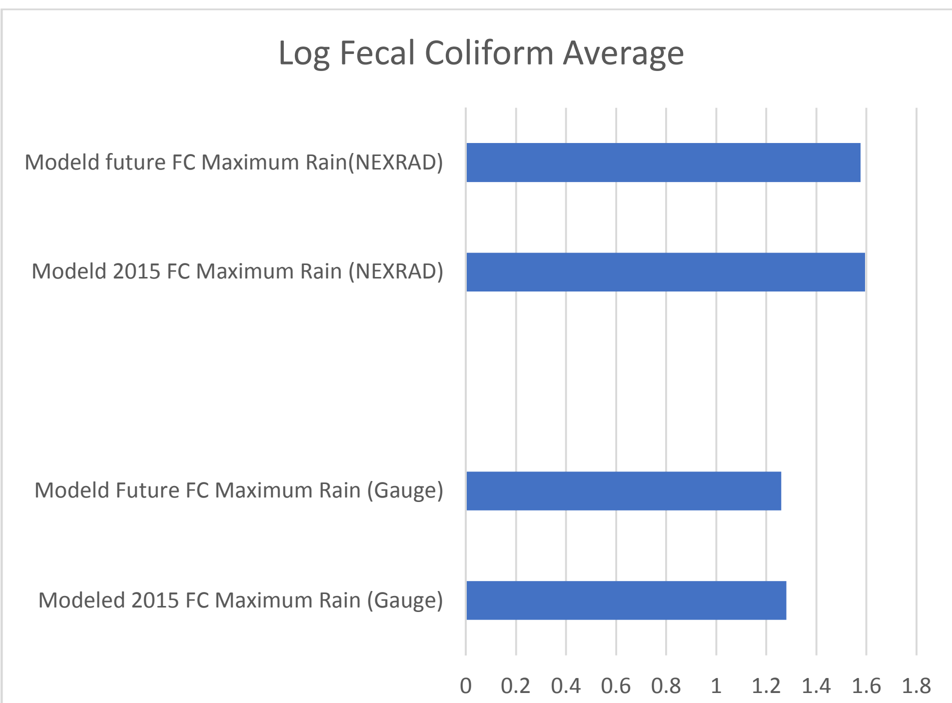

Figure 4.6: Average FC levels for Modeled 2015 and future FC with Maximum Rain

A prior study has noted the importance of integrating a scenario-based modeling with the GIS in order to have a powerful decision support tool for LU management (Vaché et al., 2002). Similar to the study conducted by Delpla & Rodriguez, (2014), LMM was used in modeling the impact of future LU and climate scenarios in FC levels. This study seems to suggest using the LMM and the GIS in developing the basis of a decision support tool for forecasting LU impact in water quality. The findings of the study have significant implications for developing decision support tools for May River's water quality.

The results of this study may rather be limited by the locations of the shellfish monitoring stations because the basis of the modes is based on the relationships between LU percentages in the sampling locations and FC, which is an important issue for future research. Further research should be undertaken to develop predictive models for water quality and apply it to multiple land use scenarios in addition to climate scenarios.

4.5 CONCLUSION

This study attempted to use formerly developed models to predict the impact of a proposed LU plan on FC levels in the May River. Two rainfall scenarios (average and maximum precipitation) were used in the models. The models that predicted the future FC levels and the models that simulated FC levels in 2015 were compared. The results indicated that the proposed LU plan will not lead to a higher water quality degradation which may occur as a result of higher FC loading in the May River. By integrating LU percentages with rainfall and tide stage the models successfully predict FC concentration for different LU scenarios. Such models can participate in land management practices through identifying the spatial and temporal variabilities of water pollutants. This study suggests using circular buffer in defining the spatial scale for the LU variables that are used in the models. Moreover, this study suggests using the LMM for model developments.

REFERENCES

Arnold, J. G., Williams, J. R., Srinivasan, R., King, K. W., & Griggs, R. H. (1994). SWAT: soil and water assessment tool. *US Department of Agriculture, Agricultural Research Service, Grassland, Soil and Water Research Laboratory, Temple, TX*.

Barber, B. P. (2008). Town of Bluffton May River watershed monitoring program. *Prepared for Town of Bluffton, Bluffton, SC. 143pp*.

Beasley, D. B., & Huggins, L. F. (1982). *ANSWERS-Users Manual. US Environment Protection Agency*. Report EPA-905/9-82-001.

Bhaduri, B., Grove, M., Lowry, C., & Harbor, J. (1997). Assessing long-term hydrologic effects of land use change. *American Water Works Association. Journal*, *89*(11), 94.

Bicknell, B. R., Imhoff, J. C., Kittle Jr, J. L., Donigian Jr, A. S., & Johanson, R. C. (1997). Hydrological simulation program—Fortran: User's manual for version 11. *Environmental Protection Agency Report No. EPA/600/R-97/080. US Environmental Protection Agency, Athens, Ga*. Retrieved from http://infohouse.p2ric.org/ref/07/06417.pdf

Bussi, G., Whitehead, P. G., Bowes, M. J., Read, D. S., Prudhomme, C., & Dadson, S. J. (2016). Impacts of climate change, land-use change and phosphorus reduction on phytoplankton in the River Thames (UK). *Science of the Total Environment*.

Retrieved from http://www.sciencedirect.com/science/article/pii/S0048969716303199

De Girolamo, A. M., & Porto, A. L. (2012). Land use scenario development as a tool for watershed management within the Rio Mannu Basin. *Land Use Policy*, *29*(3), 691–701.

Delpla, I., & Rodriguez, M. J. (2014). Effects of future climate and land use scenarios on riverine source water quality. *Science of the Total Environment*, *493*, 1014–1024.

Elfert, S., & Bormann, H. (2010). Simulated impact of past and possible future land use changes on the hydrological response of the Northern German lowland 'Hunte'catchment. *Journal of Hydrology*, *383*(3), 245–255.

Elliott, A. H., Semadeni-Davies, A. F., Shankar, U., Zeldis, J. R., Wheeler, D. M., Plew, D. R., … Harris, S. R. (2016). A national-scale GIS-based system for modelling impacts of land use on water quality. *Environmental Modelling & Software*, *86*, 131–144.

Garmendia, E., Mariel, P., Tamayo, I., Aizpuru, I., & Zabaleta, A. (2012). Assessing the effect of alternative land uses in the provision of water resources: evidence and policy implications from southern Europe. *Land Use Policy*, *29*(4), 761–770.

Isik, S., Kalin, L., Schoonover, J. E., Srivastava, P., & Lockaby, B. G. (2013). Modeling effects of changing land use/cover on daily streamflow: an artificial neural network and curve number based hybrid approach. *Journal of Hydrology*, *485*, 103–112.

Karlsson, I. B., Sonnenborg, T. O., Refsgaard, J. C., Trolle, D., Børgesen, C. D., Olesen, J. E., … Jensen, K. H. (2016). Combined effects of climate models, hydrological model structures and land use scenarios on hydrological impacts of climate change. *Journal of Hydrology*, *535*, 301–317.

Lin, Y.-P., Hong, N.-M., Wu, P.-J., & Lin, C.-J. (2007). Modeling and assessing land-use and hydrological processes to future land-use and climate change scenarios in watershed land-use planning. *Environmental Geology*, *53*(3), 623–634.

Mancosu, E., Gago-Silva, A., Barbosa, A., De Bono, A., Ivanov, E., Lehmann, A., & Fons, J. (2015). Future land-use change scenarios for the Black Sea catchment. *Environmental Science & Policy*, *46*, 26–36.

Mango, L. M., Melesse, A. M., McClain, M. E., Gann, D., & Setegen, S. G. (2010). Land use and climate change impacts on the hydrology of the upper Mara River Basin, Kenya: results of a modeling study to support better resource management. Retrieved from https://works.bepress.com/daniel_gann/17/

Mehdi, B., Ludwig, R., & Lehner, B. (2015). Evaluating the impacts of climate change and crop land use change on streamflow, nitrates and phosphorus: A modeling study in Bavaria. *Journal of Hydrology: Regional Studies*, *4*, *Part B*, 60–90. https://doi.org/10.1016/j.ejrh.2015.04.009

Nakicenovic, N., Alcamo, J., Davis, G., de Vries, B., Fenhann, J., Gaffin, S., … others. (2000). *Special report on emissions scenarios: a special report of Working Group III of the Intergovernmental Panel on Climate Change*. Pacific Northwest National Laboratory, Richland, WA (US), Environmental Molecular Sciences Laboratory (US). Retrieved from http://www.osti.gov/scitech/biblio/15009867

Neupane, R. P., & Kumar, S. (2015). Estimating the effects of potential climate and land use changes on hydrologic processes of a large agriculture dominated watershed. *Journal of Hydrology*, *529*, 418–429.

Niraula, R., Kalin, L., Srivastava, P., & Anderson, C. J. (2013). Identifying critical source areas of nonpoint source pollution with SWAT and GWLF. *Ecological Modelling*, *268*, 123–133.

Rajib, M. A., Ahiablame, L., & Paul, M. (2016). Modeling the effects of future land use change on water quality under multiple scenarios: A case study of low-input agriculture with hay/pasture production. *Sustainability of Water Quality and Ecology*, *8*, 50–66.

S.C. Dept. of Health and Environmental Control (SCDHEC). (2017a). DHEC: Food Safety - Shellfish Monitoring. Retrieved January 31, 2018, from http://www.scdhec.gov/FoodSafety/ShellfishMonitoring/

S.C. Dept. of Health and Environmental Control (SCDHEC). (2017b). Regulation 61-47, shellfish. Retrieved from http://www.scdhec.gov/Agency/docs/water-regs/r61-47.pdf

Scanlon, B. R., Reedy, R. C., Stonestrom, D. A., Prudic, D. E., & Dennehy, K. F. (2005). Impact of land use and land cover change on groundwater recharge and quality in the southwestern US. *Global Change Biology*, *11*(10), 1577–1593.

Sharifi, A., Yen, H., Boomer, K. M., Kalin, L., Li, X., & Weller, D. E. (2017). Using multiple watershed models to assess the water quality impacts of alternate land development scenarios for a small community. *Catena*, *150*, 87–99.

Sohl, T. L., Sleeter, B. M., Sayler, K. L., Bouchard, M. A., Reker, R. R., Bennett, S. L., … Zhu, Z. (2012). Spatially explicit land-use and land-cover scenarios for the Great Plains of the United States. *Agriculture, Ecosystems & Environment*, *153*, 1–15.

Tong, S. T., Liu, A. J., & Goodrich, J. A. (2009). Assessing the water quality impacts of future land-use changes in an urbanizing watershed. *Civil Engineering and Environmental Systems*, *26*(1), 3–18.

Tong, S. T., Sun, Y., Ranatunga, T., He, J., & Yang, Y. J. (2012). Predicting plausible impacts of sets of climate and land use change scenarios on water resources. *Applied Geography*, *32*(2), 477–489.

Tu, J. (2009). Combined impact of climate and land use changes on streamflow and water quality in eastern Massachusetts, USA. *Journal of Hydrology*, *379*(3), 268–283.

U.S Census Bureau. (2015). 2015 Population Estimate for Town of Bluffton. Retrieved January 26, 2017, from //www.census.gov/quickfacts/table/PST045215/4507210

Vaché, K. B., Eilers, J. M., & Santelmann, M. V. (2002). *Water quality modeling of alternative agricultural scenarios in the US Corn belt1*. Wiley Online Library. Retrieved from http://onlinelibrary.wiley.com/doi/10.1111/j.1752-1688.2002.tb00996.x/abstract

Verburg, P. H., Eickhout, B., & van Meijl, H. (2008). A multi-scale, multi-model approach for analyzing the future dynamics of European land use. *The Annals of Regional Science*, *42*(1), 57–77.

Wijesekara, G. N., Gupta, A., Valeo, C., Hasbani, J.-G., Qiao, Y., Delaney, P., & Marceau, D. J. (2012). Assessing the impact of future land-use changes on hydrological processes in the Elbow River watershed in southern Alberta, Canada. *Journal of Hydrology*, *412*, 22 232.

Wilson, C. O., & Weng, Q. (2011). Simulating the impacts of future land use and climate changes on surface water quality in the Des Plaines River watershed, Chicago Metropolitan Statistical Area, Illinois. *Science of the Total Environment*, *409*(20), 4387–4405.

Yira, Y., Diekkrüger, B., Steup, G., & Bossa, A. Y. (2016). Modeling land use change impacts on water resources in a tropical West African catchment (Dano, Burkina Faso). *Journal of Hydrology*, *537*, 187–199.

Young, R. A., & Shepherd, R. G. (1995). AGNPS-AGRICULTURAL NONPOINT SOURCE MODEL. In *WORKSHOP ON COMPUTER APPLICATIONS IN WATER MANAGEMENT* (p. 33). Retrieved from http://www.cwi.colostate.edu/publications/IS/79.pdf#page=43

CHAPTER 5

CONCLUSION

The initial objectives of this research were to assess the impact of LU types on microbial water quality by integrating GIS as well as spatial and non-spatial statistical analysis, and to predict the impact of a proposed LU plan in a shellfish harvesting area in the May River at the Town of Bluffton, SC. This dissertation aimed to achieve the following:

- Assess the impact of different LU types on the FC loadings in the river.
- Examine the spatial (stationary and non-stationary) variations of the relationships between selected LU types and FC.
- Develop combined and seasonal models to predict FC concentration by developing several circular buffers sizes to determine the one that can provide the most significant models.
- Use the developed models to predict the impact of a proposed LU plan on FC levels in the May River by examining two rainfall scenarios (average and maximum precipitation).

The results of this dissertation indicated that:

- The stationary (global) spatial relationships showed numerus significant and non-significant positive and negative correlations between LU percentages and FC concentrations.

- The GWR was able to explain detailed non-stationary (local) spatial variations of the relationships between the selected LU variables and FC.
- Residential areas, forestlands, and golf courses were found to have significant positive correlations with FC.
- The ERT is a suitable tool to automatically select the explanatory LU variables that can be included in the GWR modeling.
- The ERT saves much time and effort in selecting the variables for GWR model's.
- Although it is preferable to use the GWR for larger study areas and lager sample numbers, it was capable to examine the spatial non-stationary variations in this study.
- The LU variables for 1800m-buffer scale were better correlated with FC when compared with the other buffer sizes.
- The LU variables for 1800m-buffer incorporated with environmental factors were able to provide significant models for FC prediction.
- It is important to consider LU impact in water quality modeling and in management practices.
- FC bacteria are affected by seasonal variations, and therefore, each season should have its own model.
- The proposed LU plan will not lead to a higher water quality degradation level than currently exists.
- By integrating LU percentages with rainfall and tide stage the models successfully predict FC concentration for different LU and climate scenarios.

The significance of the results of this study can be summarized as the following:

- The results suggest that GWR can be used for small study areas
- The statistical bias was avoided by using the ERT
- The ability to spatially link different LU classes and FC
- Found that for this study area an 1800m buffer size supports the most significant model for FC.
- Reached a suitable statistical method for including LU in the modeling.
- Provided seasonal models to predict FC in a specific season.

The limitations of this dissertation are the following:

- Three sampling stations were installed in 2009, therefore, the data for FC and all other variables other than the LU data in this study were collected in 2009 without including older historical data.
- The size of the study area is small and therefore small sampling stations conducted in the study.
- The sampling stations are located within the stream of the May River, therefore, none of these stations were in any sub-watersheds of the study area.
- FC were mostly sampled in dry days, so precipitation data for a longer period can provide more information about the impact of rainfall on FC levels.
- Only one proposed LU plan was assessed for FC prediction.

Investigations for future research could include the following:

- Assessing the relationships between LU and FC in a larger study area that has more sampling stations and historical data.

- Comparing the relationships between LU and FC between several study areas in South Carolina to explain more details about the variations of the impact of LU on FC.
- Assessing the relationships between LU and other water quality parameters such as DO, heavy metals and nutrients.
- Developing wet and dry models, if applicable, to predict FC.
- Examining the models with different spatial scales such as comparing models' performances of circular buffers with sub-watersheds.
- Examining the relationships between LU and water quality within sub-watersheds can explore more facts about the relative impacts.
- Developing multiple land use scenarios in addition to climate scenarios to provide more information about water quality prediction.

The findings of this research have implications for LU and water resources managers:

- The models that were developed in the study can participate in land management practices through identifying the spatial and temporal variabilities of water pollutants.
- Water resources managers can implement the modeling concept of this research to predict the impact of LU on other water quality parameter or pollutant, or to predict the impact of other explanatory factors that have potential impact on water quality.
- Land use managers can use similar models to predict the impact of land use and other explanatory variables on air quality.
- The LU and water resources mangers in the Town of Bluffton can apply the developed models in this study to predict the impact of further future land use

scenarios. In addition, the combined and seasonal models can be used to predict the current FC concentrations from the current LU and climate data.

- The technique of the developed models can be geographically transferred to other regions.