EXAMINATION OF THE SPATIAL RELATIONSHIP BETWEEN DEVELOPMENT METRICS AND TOTAL PHOSPHORUS IN THE GALVESTON BAY ESTUARY

A Thesis

by

HELEN MARIE WALTERS

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF MARINE RESOURCES MANAGEMENT

Chair of Committee,	Samuel Brody
Co-Chair of Committee,	Wesley Highfield
Committee Member,	Antonietta Quigg
Head of Department,	Kyeong Park

August 2016

Major Subject: Marine Resources Management

ABSTRACT

Urban development can cause increased nutrient loads in nearby streams and rivers. Understanding how the pattern of urban development affects the level of nutrients, specifically phosphorus, within the Galveston Bay Estuary is particularly important for planners and policymakers working to maximize the water quality within the region. The problem of eutrophication that results from increased nutrients can be detrimental to the health of the ecosystem; further, the rapid population growth within the Galveston Bay Estuary is increasing development within the area. The ecosystem-based study described here examines 99 watersheds across the Galveston Bay Estuary, Texas. Multiple development metrics are evaluated for both high and low intensity development and these development patterns are related to total phosphorus as an indicator of water quality. Spatial lag models were used to determine the relationship between the high intensity and low intensity development and phosphorus levels. It was hypothesized and validated by the results that less fragmented and more connected urban development patches within the Galveston Bay Estuary relate to lower phosphorus levels. In addition, as the proportion of low intensity development increases within a watershed, phosphorus levels are also increased due to runoff from fertilizer. Phosphorus-based fertilizer runoff has increased in the region and is likely driven by the use of fertilizers on urban and rural homes. The results from this study can be implemented in planning and policy through a series of tools including development clustering, urban growth boundaries, transfer of development

rights, education and outreach, and implementation of laws. Each planning tool offers a way to aggregate the low intensity development in a manner that will reduce the phosphorus levels within the study area; this, in turn, will decrease the probability of eutrophication that can result in streams with nutrient loading problems. In addition, there are a large amount of phosphorus-based fertilizers used in the region, and reducing these levels will aid in decreasing the phosphorus levels within the rivers and streams.

DEDICATION

This thesis is dedicated to my parents, who are my role models and share my perpetual thirst for knowledge. The support and love of learning that I have originates from my parents and without them I would not be the person I am today. Through their encouragement and statistical-themed Christmas presents I have been able to accomplish and complete my Master's degree.

ACKNOWLEDGEMENTS

I would like to thank my committee co-chairs, Dr. Brody and Dr. Highfield for the constant help and support throughout the entire study. I would also like to thank Dr. Quigg for guidance and support. This Master's program has been a very formative period of my life and I am thankful for the direction and help that I have received throughout the whole process. I also want to extend my thanks to Texas A&M University for providing the 2-Year Competitive Graduate Merit Fellowship that funded this project and making this Master's thesis possible.

I would also like to extend a big thank you to friends and Center for Texas Beaches and Shores (CTBS) lab-mates for assisting me throughout the process. There are many people with whom this process would not have been as much fun and the product not as insightful and thorough. I also want to acknowledge my family for being a constant source of advice and support: including both of my wonderful parents, and siblings Kevin, Samuel, and Margaret. They gave me encouragement throughout the process of writing and defending my thesis.

NOMENCLATURE

AGNPS	Agricultural Nonpoint Source Pollution Model
AREA	Mean patch area metric
CONTIG	Contiguity metric
DEM	Digital Elevation Model
EPA	Environmental Protection Agency
ESRI	Environmental Systems Research Institute
GBE	Galveston Bay Estuary
HARC	Houston Advanced Research Center
H-GAC	Houston Galveston Area Council
HUC12	Hydrologic Unit Code 12
LPI	Largest Patch Index metric
NHDPlusV2	National Hydrographic Dataset Plus Version 2
NOAA-C-CAP	National Oceanic and Atmospheric Administration; Coastal Change Analysis Program
NRCS	Natural Resources Conservation Service
OSSF	On Site Septic Facility
PD	Patch Density metric
PLADJ	Percent of Like Adjacencies metric
PN	Patch number metric
PX_HID	Proximity measurement of high intensity development patch to streams

PX_LID	Proximity measurement of low intensity development patch to streams
SWQMIS Program	Surface Water Quality Monitoring Information System
TCEQ	Texas Commission on Environmental Quality
TN	Total Nitrogen
TNRIS	Texas Natural Resources Information System
TP	Total Phosphorus
USDA	United States Department of Agriculture
USGS	United States Geologic Survey
VIF	Variance Inflation Factor

TABLE OF CONTENTS

Page

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

1. INTRODUCTION

1.1. Problem statement

Urban development can have many effects on the natural ecosystem. One of these effects is hypoxia of waters due to nutrient loading from urban runoff. Diaz and Rosenburg (2008) show that the hypoxia problems across the globe are at least in part driven by urbanization. In addition, evidence shows that before the boom of human industrialization, hypoxia only occurred in natural environments (an example of which is the upwelling zone in the Pacific Ocean that causes seasonal natural hypoxia) (Chan et al., 2008; Diaz and Rosenberg, 2008). Now, hypoxia is a widespread and globally recognized problem that is related to increasing urbanization and anthropogenic impacts (Diaz and Rosenberg, 2008). While Diaz and Rosenburg (2008) discuss hypoxia problems on a global scale, narrowing down the scope to a region in order to understand the stream water quality problems associated with the specific region and how the problems can best be remedied through land use planning and individual behavior is the focus of this study.

The impact that urbanization has on nutrients has been looked at previously in many locations and across multiple time frames (Alberti, 2005; Alberti et al., 2007; Allan, 2004; Carpenter et al., 1998; Halstead et al., 2014; Hogan et al., 2014; Lenat and Crawford, 1994; Paul and Meyer, 2001; Zampella et al., 2007). One aspect of this research involves the study of the spatial patterns of urban development. Studies have examined the patterns of urban development and what spatial configuration of development best minimizes the negative effects on water quality (Alberti et al., 2007; Carle et al., 2005; Sun et al., 2014). In general, the more clustered the urban development, the less fragmentation exists in the surrounding ecosystem and therefore the impacts of the urban development are minimized (Alberti et al., 2007).

Up to this point there have been multiple studies across the nation evaluating development patch metrics effects on various aspects of water quality (Alberti et al., 2007; Carle et al., 2005; Sun et al., 2013; Sun et al., 2014; Tran et al., 2010). The patterns of development have been shown to be important in understanding how the spatial dynamics of development influence nutrient levels in the streams/rivers. While there have been several studies focusing on the importance of different development metrics on water quality (Alberti et al., 2007; Carle et al., 2005; Sun et al., 2013), little or no work has been done to examine the effects of development patterns on nutrient levels in the Galveston Bay Estuary (GBE), despite the fact that it is one of fastest growing areas in the country. For this reason, this study looks at the urban development landscape metrics and their effects on phosphorus (one component of nutrient loading) in the GBE.

Nutrient loading from nutrients including nitrogen and phosphorus are one cause of eutrophication (King et al., 2012). Hypoxic zones are considered areas that have lower than 2 mg of oxygen per liter of water (Diaz and Rosenberg, 2008; Dodds, 2006).These zones are caused by high levels of nutrients (such as nitrogen and phosphorus) that result in phytoplankton blooms and when these blooms die the bacteria that decompose them consume oxygen in the process, thus leaving the water low in oxygen (Dodds, 2006).

Many types of nutrients together (like nitrate, nitrite, ammonia, and total phosphorus) drive this hypoxia problem (Dodds, 2006).

Urban development is a driver of nutrient loading and polluted stormwater runoff from urban lands is the primary reason why 40% of the surveyed U.S. bodies of water did not meet EPA water quality standards in 2005 (Hogan et al., 2014; Paul and Meyer, 2001; USEPA, 2005) Stormwater runoff from urban development contains pollutants and nutrients that degrade the water quality of the streams it flows into. This type of pollution is considered non-point source pollution, and can contains nutrients such as nitrogen and phosphorous (Muscutt et al., 1993). Fertilizers that are applied to urban lawns are one component of the nutrient runoff entering the streams and rivers. Urban fertilizer can increase the nutrient levels in the water, which can be harmful to the stream ecosystem (King et al., 2012; Lehman et al., 2011).

Phosphorus is the limiting nutrient for algal growth in freshwaters (Soldat and Petrovic, 2008). Although, because estuarine and freshwater environments are interconnected, there will be effects with both nitrogen and phosphorus when alteration occurs to one of the ecosystems (Paerl et al., 2014). In the NOAA State of the Coast Report in 2001, 45% of the estuaries in the Gulf of Mexico were categorized as having a high level of eutrophication including the GBE (Clement et al., 2001). The Trinity River/Galveston Bay watershed also has some of the highest total phosphorus loadings compared to other watersheds in the region (Rebich et al., 2011). The USEPA 2010 Texas Water Quality Report states that urban-related runoff/stormwater is one of the top five largest contributors to the impairment of river and stream water quality in the state (U.S.

EPA, 2010). According to Rebich et al. (2011) urban development contributes about 80% of the total phosphorus export from the Trinity River/Galveston Bay watershed. For these reasons, it is important to understand further the relationship that humans have on the nutrient levels in the GBE.

The GBE is a very economically productive and urbanized part of the nation. It includes 60% of the large industrial facilities in the whole state (Örnólfsdóttir et al., 2004). High level of urbanization in the GBE can lead to urban sprawl. Urban sprawl has been shown to worsen water quality (Sun et al., 2014). In areas with larger lots this urban sprawl is low intensity development, which is classified as 21-49% impervious surface (NOAA-CCAP, 2010). The runoff and fertilizer use from sprawling urban development has been correlated with increased nutrient loading (like phosphorus) in the streams (Moore et al., 2003).

1.2. Overview of study

This quasi-experimental design evaluates the relationship between seven development metrics for too types of development: high intensity and low intensity development and nearby stream or river phosphorus levels. Additionally, it utilizes multiple control variables to isolate the influence of the development metrics on the dependent variable of total phosphorus. For the duration of this paper, streams and rivers within watersheds will be referred to simply as rivers.

Many nutrients have strong correlations with urban land use such as pH, temperature, nitrite, nitrate, ammonia, and biological health indicators (Halstead et al., 2014). Total phosphorus is a good indicator of nutrient loading because normally phosphorus enters into streams from surface water opposed to groundwater (Correll, 1998). Total phosphorus (TP) is used as one of the many different types of nutrient loading that influences the water quality of the rivers and in the GBE there is limited availability of consistently collected data other water quality indicators. TP is selected because in the study region only Harris County has sufficient data from other variables such as nitrate (NO_3^-) and nitrite (NO_2^-). TP is sufficiently and repeatedly sampled in the entire GBE and this is one of the reasons it will be used as an indicator for aquatic nutrient loading and water quality in this study.

This study was conducted at a watershed level. The benefit of using watersheds opposed to jurisdictional boundaries such as city, county, or state lines is that the ecosystem defines the watershed boundaries, thus allowing for an inclusive natural system to be studied. When a river is being studied, taking a watershed approach means not only studying the lake, river, or stream, but all of the land that drains into the water as well (Randolph, 2003). Often times this drainage has a significant influence on the water quality downstream. Utilizing a watershed approach allows a study to be driven by the natural components of the system such as hydrology and topography.

1.3. Research question and objectives

The purpose of this study is to examine the relationship between spatial development patterns and water quality (utilizing phosphorus as an indicator) in the Galveston Bay Estuary region (as defined by the National Estuary Program) (Figure 1). This study asks: ‘what is the spatial relationship between development metrics and phosphorus levels in the streams and rivers of the Galveston Bay Estuary evaluated at a watershed level?’

The primary objective of this study is to statistically identify the relationship between various development metrics and phosphorus levels measured in GBE. I will calculate and analyze the following seven different development metrics:

1. Average patch number
2. Average patch density
3. Average patch size
4. Average contiguity of patches
5. Largest patch index
6. Percent of like adjacencies
7. Average proximity of development patch to streams

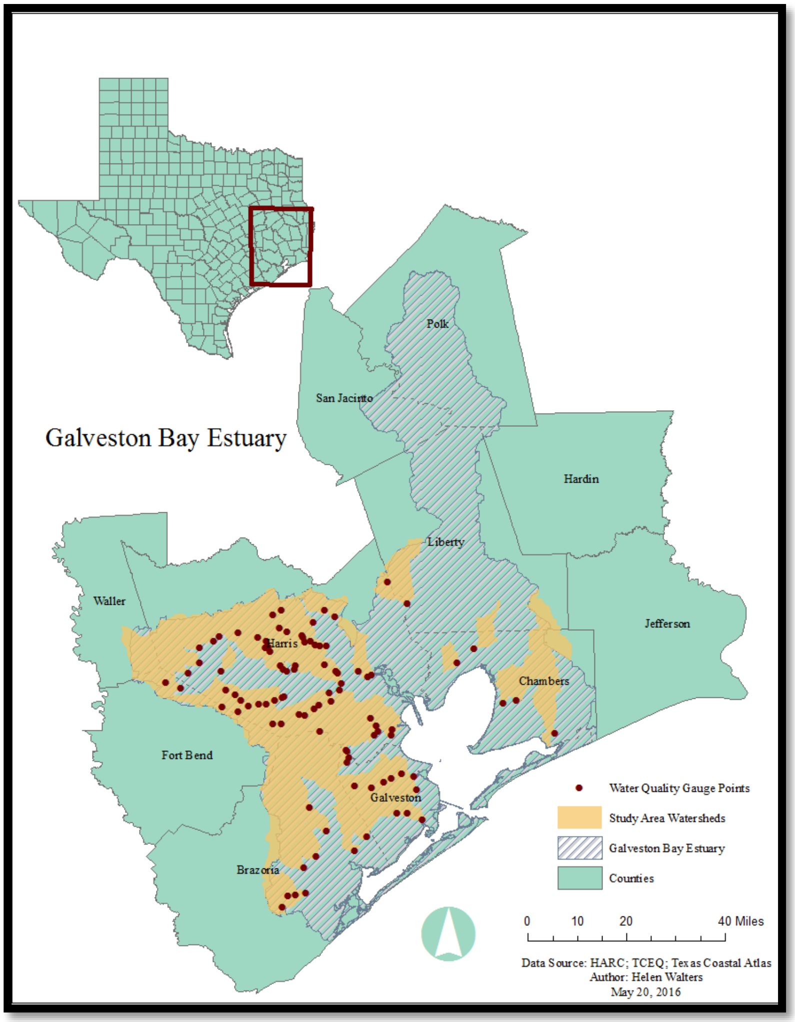


Figure 1: Galveston Bay Estuary and study area with surrounding counties. The water quality gauge points are the monitoring points from the Texas Commission on Environmental Quality. The Galveston Bay Estuary outline is the estuary as defined by the National Estuary program. The yellow study area watersheds show the area that is covered in this study and the counties are listed from the Texas Coastal Atlas.

When discussing urban development, the primary factor is not solely the percent of impervious surface in a watershed but the specific spatial distribution of impervious surfaces (Brody et al., 2013). All of these metrics can serve to help urban planners and policymakers understand the role the spatial patterns of development play in influencing phosphorus levels.

1.4. Contribution of research

The Houston/Galveston region is one of the fastest growing regions in the United States. According to the U.S. Census Bureau (2012), had the largest growth of any metro center in the United States between 2000 and 2010. Between 2010 and 2011 there was an increase of 140,000 people in Houston alone (U.S. Census Bureau, 2012). This rapid expansion of population (36% between 1997 and 2012) stimulates urban development and results in greater demand for industry, commercial, and housing development (Texas Land Trends, 2014). With this increase in urban development comes a decrease in water quality. Specifically, increased nutrient loading (nitrogen specifically in (Sun et al., 2013)) in rivers from urban development can lead to increased anthropogenic water pollution. River water quality is also a good indicator of ecological habitat quality and therefore ecologically sound rivers aid in cumulative ecosystem health (Rapport et al., 1998).

Galveston Bay and the Houston Ship Channel contain a large petrochemical and oil presence while also providing many economically important resources to local communities. Additionally, Galveston Bay offers many aesthetic benefits, areas for recreation, and ecotourism opportunities. This region, home to a population of over two million people (in Houston alone) and growing, is vital to the national economy (U.S. Census Bureau, 2012). Water quality degradation will have a negative effect not only on the resources of this region but on the aesthetic value of the natural ecosystem as well.

The results of this study can be utilized when future land use planners and policy writers are working to reduce pollution in streams and rivers.

Research surrounding the effects of urban development on water quality is greatly needed in this region of Texas due to the fast growing population and unique ecological habitats found in this area. There has been a lack of previous research conducted on this topic within this region and it is critical to gain an understanding considering that Sun et al., (2013) claimed that the relationship of these factors is often “region specific.” In order to have a thorough understanding of the effects increased development have on water quality in the GBE and how to best shape management and development practices in the future, a site-specific study needs to be conducted. The unique relationship between the variables is essential to understanding the implications of increased development in close proximity to rivers and the effects of increased urbanization and sprawl (Sun et al., 2014).

Specifically regarding phosphorus, the trend of phosphorus in the tributaries of the GBE has seen an incline in the past 15 years. Figure 2 shows the trend since the late 1960s, where there is a lot of fluctuation in the data. However, when only the last 15 years is isolated (Figure 3) there is a steady increase in the average phosphorus levels (data obtained from the Texas Commission on Environmental Quality (TCEQ)).

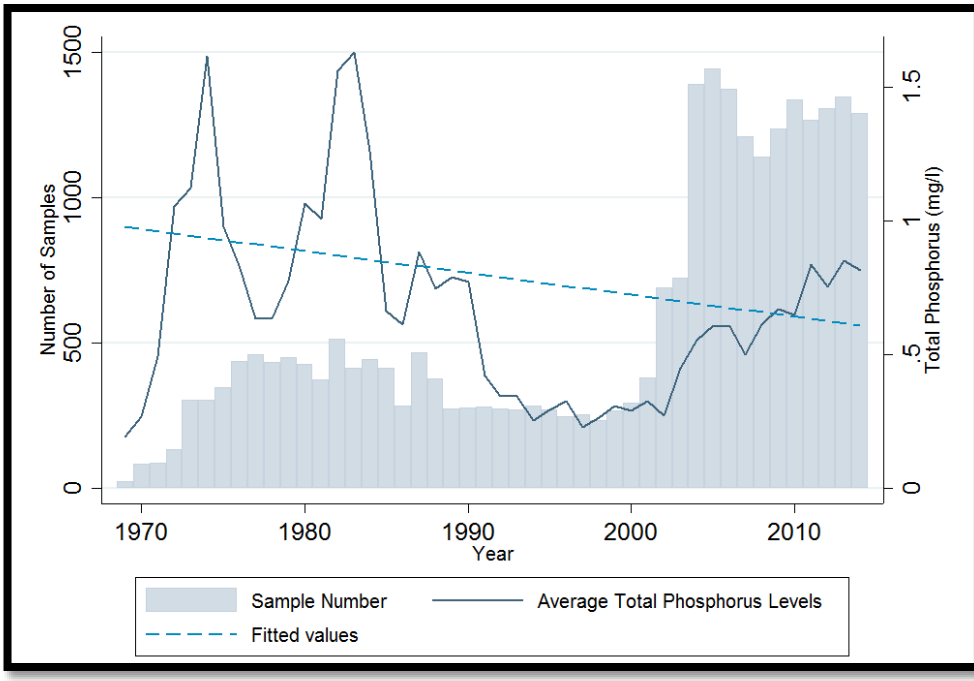


Figure 2: Trend of total phosphorus in the tributaries of the Galveston Bay Estuary. Samples are taken at 0.3 meters and above. The trend is provided for 1970-2010. The trend has an R^2 value of 0.14. Data is taken from the Texas Commission on Environmental Quality.

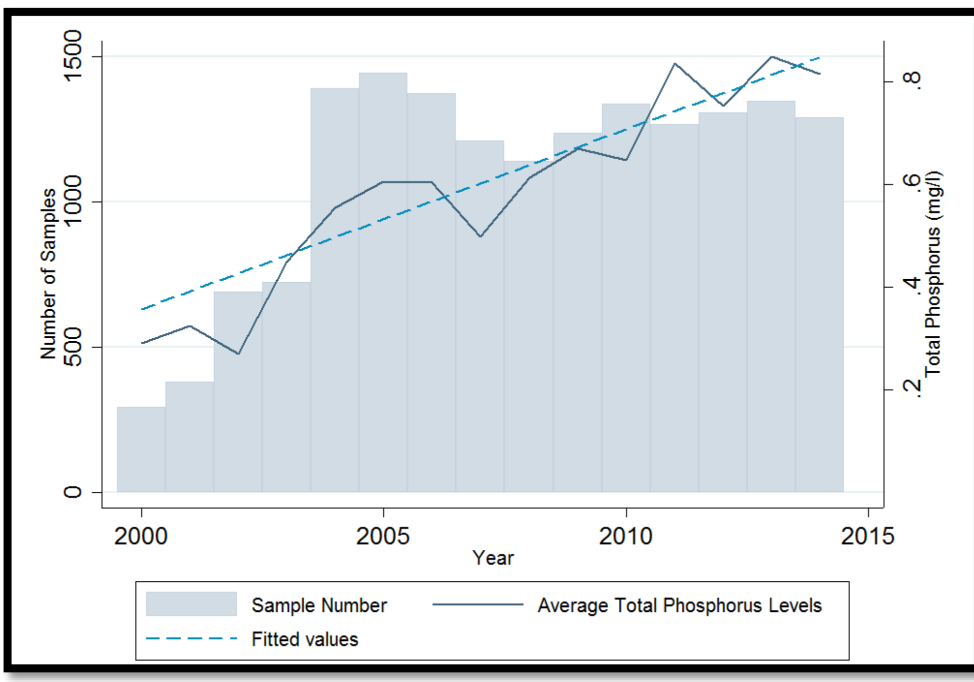


Figure 3: Total phosphorus levels within the Galveston Bay Estuary tributaries. Samples are taken at 0.3 meters and above. The time period of 2000-2014 is shown. The trend has an R^2 value of 0.8. Data is taken from the Texas Commission on Environmental Quality.

The increase shown in figure 3 shows a need for understanding how changing development influences the phosphorus levels of the estuary. Some background on the literature behind the research between spatial development patterns and water quality indicators follow in the literature review. The literature review discusses the relationship between land use and water quality (specifically the land use classes of urban development, wetlands, agricultural land, and forests), thresholds in the ecosystem, water quality variables, watershed delineations, developmental factors, scale, landscape metrics, and a case study of the Chesapeake Bay watershed.

2. LITERATURE REVIEW, CONCEPTUAL MODEL, AND HYPOTHESES

There has been much research conducted on the relationship between water quality and development in the past few decades. Certain studies focus on the relationship between land cover and biological conditions, which are indicators of stream health (Booth et al., 2004; Halstead et al., 2014). Utilizing multiple biological indicators such as taxa richness (Booth et al., 2004) is one part of the relationship. Looking at the health of the biological indicators in the stream is a good way to look at the overall health of the stream. Another general indicator of stream health is water quality. Some studies examine water quality indicators like nitrogen, phosphorus, total suspended solids, pH, and temperature (Coulter et al., 2004; Zampella et al., 2007). Other studies narrow down the focus to only the water quality parameters that contribute to the nutrient loading (i.e. nitrate, nitrite, total phosphorus, ammonia) (Dietz and Clausen, 2008). Each of these approaches focus on different aspects of the correlation between water quality and development. All of the studies show there is a relationship and it needs to be addressed particularly in areas with a high amount of urban development.

One of the main issues is referred to as urban stream syndrome (Halstead et al., 2014; Paul and Meyer, 2001; Walsh et al., 2005). The urban stream syndrome comes from nutrient or pollutant overloading in the streams from urban development. This urban stream syndrome can be detected on both large and small scales. When runoff from urban areas drains into streams this can be very detrimental and degrading to the stream environment. Understanding this relationship can help developers and planners make sure that the negative effects on streams are minimized.

The existing literature concludes that increasing urban land cover increases impervious surface areas and subsequently decreases nearby river water quality (Chang, 2008; Chang et al., 2014; Hogan et al., 2014; Paul and Meyer, 2001). Increasing impervious surface area has been shown to increase water pollution (Chang, 2008). In Dietz and Clausen (2008) there were significant relationships found between development of impervious surface area and increased runoff as well as nitrogen and phosphorus export. The runoff from impervious surface areas contains lots of pollutants which decrease the water quality of the stream.

2.1. Literature review

2.1.1. Relationship between land use and water quality

In studying the relationship between land use and water quality it is important to know what types of land covers are within the study watersheds. For watersheds that are not a single land cover type (mixed watersheds), knowing the proportion of each land cover types is necessary (Coulter et al., 2004). Based on previous literature, there are some main land cover types that stand out and influence on water quality levels. These land cover types are: agriculture (Lenat and Crawford, 1994), wetlands (Johnston, 1991; Liu and Cameron, 2001) and forests (Lee et al., 2009). All of the previously mentioned land cover types will be controlled for along with different types of urban development in this study. One important potential difference to note is that there can be different classification used for the land cover types. Each land cover type is derived from aerial imagery and can be different based on different classification software and operators. For the purposes of this study the National Oceanic and Atmospheric Administration (NOAA) Coastal Change Analysis Program (C-CAP) from the Digital Coast Program will be utilized and reclassified (Halstead et al., 2014). The following explains the relationship of each land cover type with water quality in more depth.

2.1.1.1. Urban development

There is a lot of discussion about the relationship between the urban development and the natural environment and the interactions and feedbacks that are present. One diagram that sums up very nicely the complex feedbacks that occur between the different facets of the human environment and natural environment is shown in figure 4.

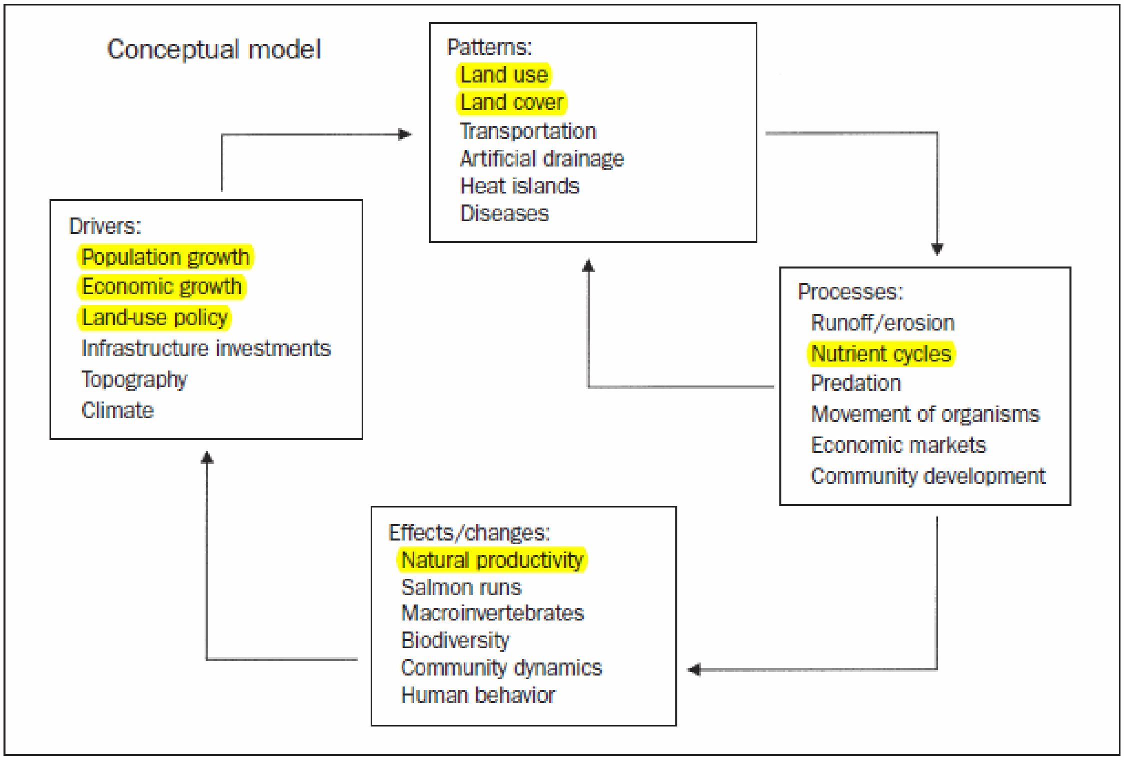


Figure 4: Feedbacks and interactions between human and biophysical variables. The drivers are the forces that are driving the specific patterns of urban development. The patterns of urban development influence the processes within the environment and then the effects/causes are what happens from the resulting processes. In addition these effects can cause more drivers and the process is cyclical and dynamic. This figure is from Alberti et al., 2003 with the highlighted items being the items that are important and relevant to this study.

There is an overall consensus of a negative relationship between urban land cover and water quality; studies suggest that this research may vary when examining different types of land cover or specific water quality indicators. Tong and Chen (2002) and Ahearn et al., (2005) both found a positive relationship between total nitrogen (TN) and total phosphorous (TP) levels and urban land cover. In Williams et al., (2005), however, no significant correlation was found between TN, TP, and urban land. Zampella et al., (2007) found conflicting results that validated portions of both Tong and Chen (2002) and

Williams et al. (2005) stating no relationship between TP and urban land use but a significant positive relationship between TN and urban land use.

Determining an appropriate spatial scale for these studies is critical, as different spatial scales may show varying relationships between water quality indicators and urban land cover (Dietz and Clausen, 2008). Specifically, small-scale studies may incorporate different geographic locations than large scale studies. These inconsistencies make comparisons between studies at different spatial scales extremely difficult and may result in suspect generalizations. Because of these inconsistencies it is important to conduct studies that vary in geography. It is not an accurate assumption that a study done in Saratoga County, NY will have the same results as one conducted in Waterford, CT. However, that being said, the general trends appear to be similar across landscapes based on different literature. Along with the negative relationship between area of urban land cover and water quality, the majority of research points to a positive relationship between nutrient levels and urban development (Halstead et al., 2014).

Stormwater is the primary driver behind the negative relationship between water quality and impervious surface cover (Walsh et al., 2005). Different land use types affect the input of stormwater into watersheds in various ways. The different proportions of agriculture, forested areas, urban development, as well as other land use categories have a large influence on the water quality within the watershed (Alberti et al., 2007; Coulter et al., 2004).

2.1.1.2. Wetlands

Wetlands are able to retain 4 to 80% of the input phosphorus (Johnston, 1991; Reddy et al., 1999). This means that the amount of phosphorus going into the wetlands are greater than the amount of phosphorus coming out of the wetland. The phosphorus filtration that wetlands have serves as a way to filter and clean the water. Wetlands help serve as a retentive ecosystem which serves as a benefit to the water downstream after it goes through the wetlands (Johnston, 1991). Wetlands are very important for retaining nutrients that could enter into streams and thus reducing the negative anthropogenic influence that sometimes can be see with additional pollutants and nutrients entering into streams (Liu and Cameron, 2001).

Wetlands do have a cap on how much phosphorus they can retain. For instance, if there is an overload of phosphorus onto a system, this can be detrimental to the wetland. Some of the things that cause the retention rate to vary in streams are flow velocity, discharge, biological fluxes, and chemical characteristics (Reddy et al., 1999). Wetland land cover often has a negative relationship with nutrients, and the more wetlands there are, the more the nutrients are filtered out and therefore the less nutrients there are in the watershed (Johnston, 1991).

2.1.1.3. Agricultural land

Another important land cover type that has influences on nutrient levels in streams is agriculture land. There has been much literature stating that increasing the amount of agriculture land increases the levels of phosphorus (and other nutrient loading) to nearby rivers. According to Lenat and Crawford (1994) agricultural lands produce the highest amount of nutrient loading (including total phosphorus) in nearby rivers. In a discussion of just crop lands (a portion of the total agricultural land), Peterjohn and Correll (1984) stated that 16% of phosphorus exports from crop land were in surface runoff. In addition, Peterjohn and Correll (1984) looked at the interaction between cropland and riparian forest and showed that 94% of the phosphorus inputs into the riparian forest came from surface runoff. This shows that there can be an interaction between cropland and nutrient levels.

Phosphorus that is accumulated in soil through use of phosphorus based fertilizers can be released through surface runoff (Hart et al., 2004). When there is a large rain event, or the general watering process causes some of the fertilizer to runoff of the crop land it often ends up in the nearby rivers or other water bodies. This over-fertilization of the soil can contribute towards the eutrophication problem seen in some water bodies (Hart et al., 2004). One of the largest issues is that in many cases, a small amount of phosphorus can have a large impact on the waters and can potentially cause these eutrophication problems (Hart et al., 2004). Agricultural land is positively correlated with TP (Liu et al., 2009; Tong and Chen, 2002). Agricultural land nearby streams and rivers can make for poorer river water quality because the high levels of nutrients and pollutants, originating from fertilizer, run off the land and into these bodies of water (Tong and Chen, 2002).

2.1.1.4. Forests

In many cases, forest land is negatively correlated with TN, nitrate, and TP (Halstead et al., 2014; Tong and Chen, 2002). Larger percentages of forest land surrounding streams and rivers often results in better water quality. Forest is an important variable when determining water quality after the landscape has been altered by urban development (Sliva and Williams, 2001). It has been shown that often a higher percentage of forest within the landscape results in improved water quality (Tu, 2011). Therefore, TN, TP, nitrate, and fecal coliform levels are positively related to the amount of commercial, residential, and agricultural lands; in a similar fashion, they are negatively related with the amount of forest land (Tong and Chen, 2002).

While there is overall a negative relationship between water quality and forest, it has been shown that the landscape of these forested areas are very important (Lee et al., 2009). In addition, there is variability between different seasons and the relationship between forests and certain water quality variables. For instance, in the spring there is no significant relationship seen between TP and forest but in the fall there is shown a significant negative relationship in a study conducted in South Korea (Lee et al., 2009).

In addition, there are different forestry patch metrics that are shown to have a positive relationship with total phosphorus. For instance, Lee et al., (2009) showed that there was a positive relationship of TP with patch density and the study resulted in saying that the less fragmented but more complex the forest area is seems to preserve the water quality the best. This study went on to state that the degradation of water quality can come not only from increasing urban lands but also decreasing the quality of the remaining forests that have not been converted into urban lands (Lee et al., 2009). Other studies have also shown that the as the forested land cover becomes more fragmented its ability to intercept runoff becomes reduced (Alberti et al., 2007). The summation of this is to say that highly fragmented forests do not function as a filter to provide the generally understood negative relationship with total phosphorus as is shown in numerous other studies.

2.1.2. Thresholds

It is important to note the definitions of impervious land cover and urban development. Impervious surfaces are impenetrable materials that cover land. Urban development is the classification of land cover that includes impervious cover. However, the term urbanization is a very broad term that includes all different types of development including industrial as well as all types of residential (Booth et al., 2004). Percent impervious cover is a way of measuring urban development and therefore the meanings can be synonymous and can be used interchangeably (Zampella et al., 2007). It should be noted that while there are many studies that use impervious surface as an indicator for human impact it cannot necessary be considered a replicate for human impact on the natural ecosystem (Booth et al., 2004).

The threshold where urban development causes a negative effect on stream water quality is readily debated in literature. There are two types of thresholds when referring to impervious land: percent of impervious land cover in the watershed and size of the buffer from impervious surface cover to streams and rivers nearby. Zampella et al., (2007) states that watersheds with 10% or more altered land represents the threshold where water quality decreases due to impervious surface effects. A third argument from Walsh et al., (2005) states that rivers within a watershed are in good condition until the watershed is covered with 12% impervious surfaces. Then, after the 12% threshold has been crossed, the river conditions become consistently poorer.

Buffers around streams can be used to prevent pollution. There are differences in the literature about how large a buffer needs to be in order to prevent the majority of pollutants from impervious surfaces from entering the stream. Certain studies like Ou and Wang (2011) state that a 300 meter buffer is the threshold where impervious surface cover has less of an impact on the nearby stream and river water quality. Another study (Chang, 2008) uses a 500 m buffer for the relationship between urban land cover and water quality. And other studies still show that the threshold where impervious surface cover has less of an impact on the nearby river water quality is over 200 meters away from the stream (Halstead et al., 2014; Sun et al., 2013; Tran et al., 2010). May et al. (1999) states that there are no thresholds for urban development in relation to stream characteristics and instead there is a continuous scale of urbanization alteration and stream water change.

2.1.3. Water quality

There are many factors that influence the quality of water within a waterbody. One hydrological factor is the amount of water in a stream. Large rain events increase the total amount of water that is within a stream and these large amounts of precipitation can increase stormwater runoff. This increased stormwater runoff can result in a higher volume of water in the rivers, which can lead to flooding of the riverbanks. When there is more stormwater input, this can increased pollutants and nutrients from the stormwater runoff. It is documented in literature that increased impervious surface area results in an increase in runoff volume (Dietz and Clausen, 2008; Sun et al., 2013). This five and a half fold increase can lead to a greater volume of water in the stream than is normal in the natural system.

There are many variables that can be used to assess water quality such as total suspended solids (Ahearn et al., 2005; Coulter et al., 2004), fecal coliform (Nagy et al., 2012; Tong and Chen, 2002), total nitrogen and total phosphorus (Chang, 2008; Halstead et al., 2014; Tong and Chen, 2002), pH (Chang, 2008; Coulter et al., 2004; Tong and Chen, 2002; Zampella et al., 2007), taxa richness (Booth et al., 2004), temperature (Chang, 2008; Coulter et al., 2004), and sodium, cadmium, lead, and conductivity (Tong and Chen, 2002). All of these are acceptable methods of analyzing water quality. In addition to these water quality indicators, there are other factors that should be are accounted for when evaluating watershed systems such the number of streams within a watershed, size of watershed, point and nonpoint sources of pollution, soil type, wetland density, drainage density, and location of watershed with respect to the coast (located along the coast or inland) (Booth et al., 2004; Carle et al., 2005; Coulter et al., 2004; Halstead et al., 2014; Hogan et al., 2014; Nagy et al., 2012; Nelson and Booth, 2002; Ou and Wang, 2011; Paul and Meyer, 2001).

Seasonal variability is another factor that influences water quality. The wet season, where there is more precipitation, has a greater level of runoff, which includes pollutants and nutrients, ending up in rivers (Huang and Klemas, 2012). For this reason, when comparing within a year it is important to understand the nutrient fluctuations occurring with rising water levels and precipitation. Some studies use temporal averages to remove the inter-annual hydro-climatic variability (Sliva and Williams, 2001). Another seasonal variation that occurs in rivers is temperature. The warm season means that there are higher surface water temperatures, which can stimulate algal growth (Maranon et al., 2014). However, if the temperatures reach such an extreme level, algae may be unable to survive. Therefore, understanding the seasonal fluctuation specific to a region and hydrologic system is important to understanding its water quality.

Another key component of water quality is nutrient loading. One way of assessing water quality is by looking at the nutrient loading in the water. When nutrient levels are higher than normal in a river, there can be detrimental effects on the organisms within the stream. Some of these effects include algal blooms, loss of sensitive invertebrates and nekton, and decreases in fish populations (Walsh et al., 2005). There are many different nutrients that can be used to assess water quality including nitrate, nitrate, ammonia, and phosphorous. All of these water quality indicators have been used in literature, TP among them (Bedan and Clausen, 2009; Simeonov et al., 2003; Turner and Rabalais, 1991). In addition, one of the benefits of TP is that it is an important nutrient for both fresh and salt water phytoplankton growth (Withers et al., 2009). The hypoxia area or “dead zone” that occurs from increased nutrient loading in the Gulf of Mexico is due to high amounts of nutrients introduced from stormwater runoff that in turn lead to large algal blooms followed by algal die offs. Ultimately this leads to an increase in the “duration and magnitude” of hypoxia, causing harm to the organisms living within the Gulf of Mexico (Carpenter et al., 1998; Varlamoff et al., 2001; Williams et al., 2005).

2.1.4. Delineation of watersheds

In all studies that utilize watersheds (whether delineated by the scientist based on water quality monitoring gauge points or based off of pre-delineated databases such as National Hydrography Dataset) calculations are performed based on elevation, flow direction, and flow accumulation to create the watershed boundary. Some studies use the National Hydrography Dataset (NHD), which contains different sized watersheds (hydrologic unit code (HUC) 2-14 with each increasing number defining a smaller watershed area) (USDA-NRCS, and USGS, and EPA). Studies that use NHD HUCs allow for a larger sample size because of the reduced time in using pre-delineated watersheds. For instance, Liu et al., (2009) had over 1,000 sample points within 299 watersheds in the study. Other studies delineate watersheds based on water quality monitoring points and Digital Elevation Models (DEMs), which incorporate elevation, flow direction, and flow accumulation (Alberti et al., 2007; Coulter et al., 2004; Ou and Wang, 2011; Tran et al., 2010).

Delineation of watersheds allows the area of study to be controlled by the hydrologic environment. Watersheds are the “the upslope area that drains to a specific point on a river" (Kroll et al., 2004). Water quality monitoring gauge points are used as a starting point to trace the flow lines upstream of the watershed. When delineating a watershed the three variables previously stated are used to calculate the hydrology of the watershed and define the boundary based on this hydrology. When a watershed is delineated based on a water quality monitoring gauge point water conditions from upstream are factored in to the quality measurement which is important because there can be many influencers originating upstream of the gauge points that impact the water quality downstream (Hogan et al., 2014; Sun et al., 2013). More details of watershed delineations are included in Appendix D.

Understanding the assumptions of a study that uses delineated watersheds is important. When a watershed is delineated from a stream water quality monitoring gauge point it is assumed that the poorest the water quality can be in the stream is at the gauge point (the base hydrologically). Therefore, any pollutant that comes down the stream is going to pass through this gauge point and therefore the water quality can't be any more degraded than it is at the gauge used in the delineation.

2.1.5. Developmental factors

Developmental variables are correlated with nutrient loading and water quality within rivers. Population density is a variable that is strongly correlated with nitrogen export (Halstead et al., 2014). There are other developmental indicators, including contiguous impervious surface area, spatial patterning of impervious surfaces, portion of various land cover in watershed, stormwater connectivity, density sewer system connections, septic tank density, percent watershed within city limits, percent impervious area, household density, and house age (Booth et al., 2004; Carle et al., 2005).

Urban development is one of the major factors affecting fragmentation of landscapes (Alberti, 2005; Alberti et al., 2003; Matte et al., 2015). When there is high ecosystem fragmentation due to urban development (roads and urban centers, for example), the natural ecosystem becomes cut up into patches. This fragmentation that is seen across the globe occurs on many different types of land cover: including farmland, forest, wetlands, grasslands and other land use categories (Matte et al., 2015; Su et al., 2014). In order to reduce this rampant fragmentation, consolidating (clustering) of urban development has been used as a strategy to ensure that the greatest portion of the natural ecosystem is preserved.

Another aspect to take into consideration is the specific measure of land cover. A commonly used land cover source is the National Oceanic and Atmospheric Administration's Coastal Change Analysis Program: Digital Coast Land Cover (NOAA C-CAP) (Brody et al., 2013; Halstead et al., 2014). This source has land cover for multiple years and is processed in the same resolution and processed into a land cover format. In order to be of use, the land cover has to be reclassified into fewer groups in order to minimize the number of land cover types. In some studies, urban development is referred to as a single variable (Coulter et al., 2004; Halstead et al., 2014; Sun et al., 2013) or total impervious surface area (Dietz and Clausen, 2008) is used. However, the different types of urban development (high, medium, and low intensity development) are very different:

High intensity development is 80-100% impervious surface; medium intensity development is 50-79% impervious surface; and low intensity development is 21-49% (NOAA-CCAP, 2010). By this classification standard high intensity development refers to industrial infrastructure, medium intensity refers to dense suburban landscapes, and low intensity development refers to low density suburban landscapes. Due to how different these landscapes are it is important to split up each type of development when measuring (Brody et al., 2011). Understanding that each type of development refers to a different landscape type and therefore can potentially have different effects on the water quality is important. The way the landscape metrics are can have potentially large implications for developers and policy makers.

2.1.6. Spatial scale

There are many different approaches that can be taken to evaluate the relationship between urban development land use and water quality. There is debate about what scale the interaction between urban development and water quality occurs on. In addition, there is discussion about how interactions and alterations on one scale will affect the relationships and feedbacks on other scales (Alberti et al., 2003). The urban-ecosystem interaction is dynamic and occurs over time and space. Some of these approaches include watershed based and buffer zone approaches as well as geographically weighted regression models that help remove some of the potential collinearity between watersheds

(Halstead et al., 2014). The fundamental difference between the two types of approaches is the spatial extent of the study area that is examined.

There is minimal consensus regarding which spatial scale is optimal for this sort of study (Chang et al., 2014). For instance, Gburek and Folmar (1999) state that the most effective way of examining correlations between land use and water quality is through looking at a "*very local scale*". Conversely, Gove (2001) and Bruno et al., (2014) emphasize that the strength of the relationship between the two variables increases as the study area increases meaning a basin-wide approach is preferred.

The watershed approach evaluates the entire catchment as a whole and compares the percent of development within the watershed to the water quality. In some circumstances The difference between this approach and the buffer zone approach is that it deals with larger scale studies and looks proportionally instead of at a certain buffer distance. In addition, using watersheds to look at the water quality as a whole is an effective approach when the sample size of watersheds is large as shown in Tong and Chen (2002).

The buffer zone approach involves measuring how far development patches are from streams and having a riparian buffer of a certain size between the river and urban development. This study has positive results in certain studies (Alberti et al., 2007; Tran et al., 2010). Tran et al., (2010) states that while there were no significant correlations between land use and water quality when looking at the watershed level some were found when looking at a 200 meter buffer. One potential explanation for this is that the study scale is smaller and therefore the level of detail is greater which means that there are less assumptions made and a potentially more valid result. Other studies such as Halstead et al (2014) show that there is no difference in how far away the stream is from the urban development. Both scales have benefits and detriments. As Chang (2008) states, it is important to consider evaluation at multiple scales because the interaction of water quality and land use is very complex.

One methodological difference in the literature is the number of samples used within the study. Some studies use as small as six study watersheds to infer relationships between these variables (Carle et al., 2005). Other studies incorporate between 28-299 watersheds (Alberti et al., 2007; Liu et al., 2009; Ou and Wang, 2011; Tran et al., 2010; Williams et al., 2005). While fewer watersheds can be used to study small details within a watershed, a greater number of samples are needed for statistical validity when looking at large-scale correlations.

This study will utilize class metrics to measure development patterns. Class metrics (the spatial arrangement of certain land use patches) can have a great impact on the ecological integrity of a system. This method of analyzing developmental patterns has been used for years (Gustafson, 1998). Using class metrics is a way to take specific spatial characteristics of development patches and quantify them (Alberti et al., 2007; Brody et al., 2013; Leitao et al., 2006). There are many different class metrics that can be utilized to assess development at a watershed level. The section below discusses the specific metrics that will be measured in this study.

2.1.7. Landscape metrics

The spatial configuration of landscape patches plays an important role in the health of the surrounding ecosystems (Alberti et al., 2007; Carle et al., 2005; Sun et al., 2014). The more clustered the urban development, the less fragmentation exists in the surrounding ecosystem and therefore impacts of the development are minimized (Alberti et al., 2007). Vice versa, the less clustered the developmental patches are, the more edges present in the watershed. Large amounts of urban development edges are counteractive to preserving the biodiversity of the ecosystem because these types of edges leave the organisms within the natural ecosystem susceptible to outside pollutants (Perlman and Miller, 2005).

There are many different metrics that can be used in studies to indicate different parts of the landscape. Metrics such as connectivity, patch size, aggregation index, percentage of like adjacency, and contagion are examples (Alberti et al., 2007). In order to wade through the large amount of metrics used, it was important to cypher out the ones that were clear indicators of some aspect of the environment. In addition, it is important to look at metrics that can help focus future development. If there are metrics that have no bearing on creating a better development scheme in the future, it means they aren't valuable to current and future policy (Brody et al., 2013).

This study will only look at class level landscape metrics. This means that the metrics of low intensity development and high intensity development as a land cover class within a watershed will be examined. All of the class metrics are calculated from patches of the same type (in this case low intensity and high intensity development). A patch is a homogenous area that of a single land cover class (Leitao et al., 2006). Looking at class metrics is a way to measure fragmentation indices because it measures the different configurations and patterns of a particular land cover type (Leitao et al., 2006). Because the land cover classes of interest are low intensity and high intensity development, these are the only ones that will have class metrics looked at.

One metric that is studied in this quasi-experiment is proximity of development patch to nearest river or stream. This is a debated metric because in some circumstances there is a set distance from urban development patch to stream or river that can help increase the water quality of the stream. Ou and Wang (2011) show that when the proximity of urban development is less than 200 meters, the impact of runoff from impervious surface will have a greater negative impact on the water quality of the streams. However, in other cases there may be no relationship between water quality and distance of stream to patch (Halstead et al., 2014).

2.1.8. Chesapeake Bay case study

One case study that drives this study is the Chesapeake Bay watershed which is an example of an estuary that has become degraded throughout the past years and management practices have been developed and enforced to help maintain and restore certain ecosystem functions. Urban development in the watershed has drastically increased in the last 200 years and this population growth has resulted in an increase in nutrient loading (Kemp et al., 2005; Smith et al., 2003). This population growth has been dramatic as well as the increase in fertilizer (mostly nitrogen based in this location) (Kemp et al., 2005). In addition, sedimentary records indicate that the hypoxic events being seen in the Chesapeake Bay are a recent change. While there were some hypoxic events prior to the past 50 years, the past 50 years has shown a spike in intensity as well as seasonal regularity in these events.

What can be done to mitigate these effects of urban development and the consequences it is having on nutrient loading? Reducing nitrogen and phosphorus inputs in the tributaries was tried the result was significant ecological benefits that occur (Kemp et al., 2005). One case study in the Potomac River showed that when phosphorus inputs (and some nitrogen inputs) were reduced, there was a reduction in algal blooms, increased oxygen levels, and increased water clarity (Kemp et al., 2005). It is important to discover and target the source of the nutrient loading. The source might not always be next to the location where hypoxia is seen as a study in the Chesapeake Bay showed (Dauer et al., 2000). There can be a spatial lag between when the nutrients enter into the stream (from urban sources) to where the hypoxia or oxygen waters are seen (Dauer et al., 2000). Another study showed that there is a lag between the increase in fertilizer use globally and the increase of hypoxia (Diaz and Rosenberg, 2008). This means that there could be residential development upstream that is causing hypoxic zones downstream or in another location hydrologically connected in the watershed.

2.2. Conceptual model

Based on the literature review above, I have constructed a conceptual model (figure 5) that illustrates the variables and their relationships used to explain water quality in the Galveston Bay study area. The two types of control variables used in this study are environmental controls and land cover controls. The control variables are included in the model to better isolate the independent variables. The environmental controls consist of precipitation, contributing drainage, watershed area, and septic system count within the watershed. It is hypothesized that precipitation has a negative relationship with phosphorus (the more precipitation the less phosphorus because more precipitation will cause a dilution in the streams of the phosphorus levels.) Contributing drainage has a positive relationship because the more drainage area there is that can contribute to a stream, the more phosphorus there will be. It is also hypothesized that watershed area will have a positive relationship with phosphorus. When there is more watershed area the more phosphorus can contribute to the stream within the watershed. Septic system permits are hypothesized to have a positive relationship because the more septic systems there are the more nutrient discharge and therefore the higher phosphorus levels.

The land cover controls include forest, wetland, high intensity development, low intensity development, and cultivated crops proportion within each watershed. Forest is expected to have a negative relationship with phosphorus meaning that the greater the proportion of forests in the watershed the lower the phosphorus levels. Wetland land cover is also hypothesized to have a negative relationship with phosphorus levels in this study.

Low intensity development, high intensity development, and cultivated crops are also hypothesized to have a positive relationship with phosphorus. For all three of these land covers when there are more proportions in the watershed, the phosphorus levels are hypothesized to be higher.

The independent variables of interest are the low intensity and high intensity class metrics (proximity to river, average area, average contiguity, patch number, patch density, largest patch index, and percent of like adjacencies). When proximity to river, average area, average contiguity, largest patch index and percent of like adjacencies are greater it is hypothesized that there will be a lower level of phosphorus (negative relationship). It is also hypothesized that the greater the patch number and higher the patch density the higher the phosphorus (positive relationship). The dependent variable utilized in this study is average total phosphorus levels. Figure 5 is the conceptual model of the study and breaks down the different variables utilized into categories and shows how the independent variables as well as the control variables affect the total phosphorus levels. In addition, figure 5 shows the setup of the models that are utilized in this study and the different control, independent, and dependent variables that have been previously talked about.

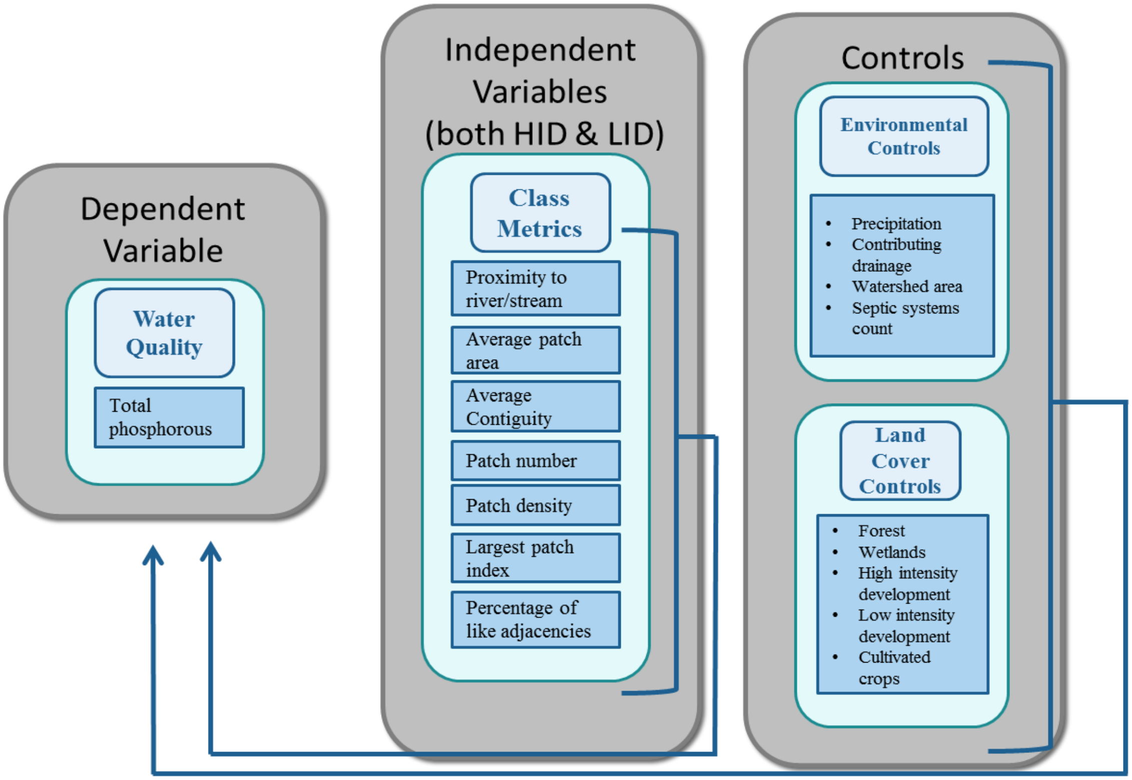


Figure 5: Conceptual model of study. The dependent variable of study is total phosphorus, the independent variables are patterns of landscape development metrics calculated for both high intensity development and low intensity development. The controls (both environmental and land cover) are included to account for ecosystem processes.

2.3. Hypotheses

The hypotheses are grouped into four categories. The hypothesis dealing with fragmentation of the landscape is based on the concept that urbanization reduces the connectedness of the natural ecosystem which can be harmful to the biodiversity of the species in the ecosystem (Alberti, 2005).

It is hypothesized that:

2.3.1 *The larger the fragmentation, the higher the phosphorus levels*

The larger the fragmentation is in the watershed, the higher the phosphorus levels. A less fragmented landscape means a larger average patch size, a smaller number of patches, and a smaller amount of patches within the landscape (patch density). It is hypothesized that when (for both low intensity and high intensity development metrics) the average patch sizes are larger, lower number of patches, and smaller patch density, the phosphorus levels will be lower in the rivers.

2.3.2 *The more connected the patches, the lower the phosphorus levels*

The more connected the patches are in the watershed, the lower the phosphorus levels. A more connected landscape is characteristic of a larger percent of like adjacencies, a greater largest patch index, and higher contiguity. This means that for both low and high intensity development when the percent of like adjacencies is larger, the largest patch index is higher and the contiguity is higher than the phosphorus levels will be lower within the rivers.

2.3.3 *Greater proximity to stream means less phosphorus*

When the proximity of urban development is less than 200 meters, the impact of impervious surface runoff will have crossed the threshold and have a greater negative impact on river and stream water quality. When the urban development patches (both high and low intensity development) are closer to a stream the higher the phosphorus levels will be within the stream. This is because the distance for the phosphorus to run off and reach the stream is less and therefore there are fewer ways that the nutrient could be filtered out before it reaches the rivers.

2.3.4 Greater proportion of low intensity development the higher the phosphorus

The greater the proportion of low intensity development is within the watershed, the higher the phosphorus levels. This hypothesis involves the control variable: percent of low intensity development, and states that when the percent of low intensity development within the landscape is higher the phosphorus levels within the watershed will also be higher. The use of fertilizers on lawns creates opportunities for phosphorus to runoff into the rivers and increase the levels.

3. RESEARCH METHODS AND DATA ANALYSIS

3.1. Study area

The study area consists of 99 watersheds within the Galveston Bay Estuary figure 6. The watersheds are delineated from water quality monitoring gauge points within the GBE that have enough samples (both ammonia and phosphorus although phosphorous is the only one focused on in this study- adequate number of samples was defined as having at least 20 samples in the four year period) and delineated watershed areas within 10 – 100 square miles to provide for valid statistical testing. The total phosphorus is taken from 2010-2013 and an average across these four years is taken in order to ensure removal of inter-annual as well as seasonal variation (Carle et al., 2005; Sliva and Williams, 2001). The GBE encompasses the spatial area defined by the National Estuary Program (figure 1). All of the watersheds used in the study are delineated from the TCEQ water quality monitoring gauge points (using the ESRI ArcMap10.2 model builder process shown in Appendix A) and utilize the NHDPlusV2 DEM, flow direction grid, and flow accumulation grid in Appendix A. Additionally, all watershed gauge points are located along streams or rivers.

The Galveston Bay Estuary is a diverse ecosystem with a large amount of environmental and economic value. The Houston region is one of the largest producers of oil in the region, and Galveston Bay is home to a wide variety of species that are sensitive to anthropogenic pollution. The combined Houston/Galveston region encompasses portions of Galveston, Harris, Brazoria, Chambers, Liberty, Polk and San Jacinto counties. The Houston/Galveston region has seen a large amount of growth over the past few decades, with a current population of 6.5 million people in the entire region that is only projected to grow in the coming years (Houston Galveston Area Council, 2014). This rapid growth has caused an increase in residential and commercial development. An increased Houston/Galveston population results in a greater demand for low intensity development, such as suburban housing.

The reason that the Galveston Bay Estuary is a prime location to study the relationship between water quality and development is based on a combination of qualities. As previously mentioned, the increasing population will drive increased housing and development which in turn could drive decreased water quality. Currently, 29% of stream miles within the Houston/Galveston Bay region exceed the screening levels for nutrients as set by TCEQ (Council, 2014). Since there is already an issue of having these streams be in exceedence, this location is prime to analyze what are some of the factors driving this problem and how can change be implemented. In addition, it is important to assess streams and areas where there is impairment of streams of occurring and develop management and policy to increase the quality of these streams and eventually remove them from the impaired stream list.

3.2. Concept measurement

The measurement of all variables used in the study are defined below as well as source of data and methodology. All of the summary statistics of the variables are listed in Table 1.

3.2.1. Average total phosphorus

The dependent variable in this study is average total phosphorus levels between 2010 and 2013. These total phosphorus levels are taken at TCEQ water quality monitoring gauge points and delineated to create watersheds. Taking a statistic (mean, median, maximum, etc.) over a multi-year period has been done to get a robust understanding of the water quality conditions within the system (Carle et al., 2005; Sliva and Williams, 2001; Zampella et al., 2007). While there are multiple different methods to obtain water quality data, such as remote sensing (Olmanson et al., 2013) and stream biological indicators (Booth et al., 2004), this study utilized total phosphorus levels for watersheds that were delineated from water quality gauge points. Total phosphorus data are obtained from TCEQ. However, some of the TCEQ water quality data from the SWQMIS Program (Surface Water Quality Monitoring Information System) are missing and some samples are not uniform. Therefore, we also used a smaller and cleaner set of data analyzed by the Houston Advanced Research Center (HARC). There are 235 gauge points that have adequate total phosphorus data in the study area between 2010 and 2013. Each of the monitoring gauge points were delineated based on a flow accumulation raster from NHDPlusV2 (description below) and then watersheds with an area between 10 and 110 mi^2 were selected. The total phosphorous measurement used for this study was the average phosphorous for these 99 water samples taken between the years 2010 and 2013. Having a four year average for the dependent variable will make sure that the inter-annual and intra-annual variation are accounted for in the analysis. This method is utilized in other studies where it is ideal to remove inner-annual and seasonal variation (Carle et al., 2005; Sliva and Williams, 2001). The distribution of phosphorus levels in the study area is shown in figure 6.

The trend that was shown previously in figure 2 details the trend of phosphorus over the past 40 years in the GBE tributaries and figure 3 shows the trend from 2000-2014 in the same location. This increasing trend from 2000-2014 shows the need for an examination of what is causing this increase. This study looks at one snapshot in time (2010-2013 dependent variable average) and how phosphorus is influenced by the development in the area.

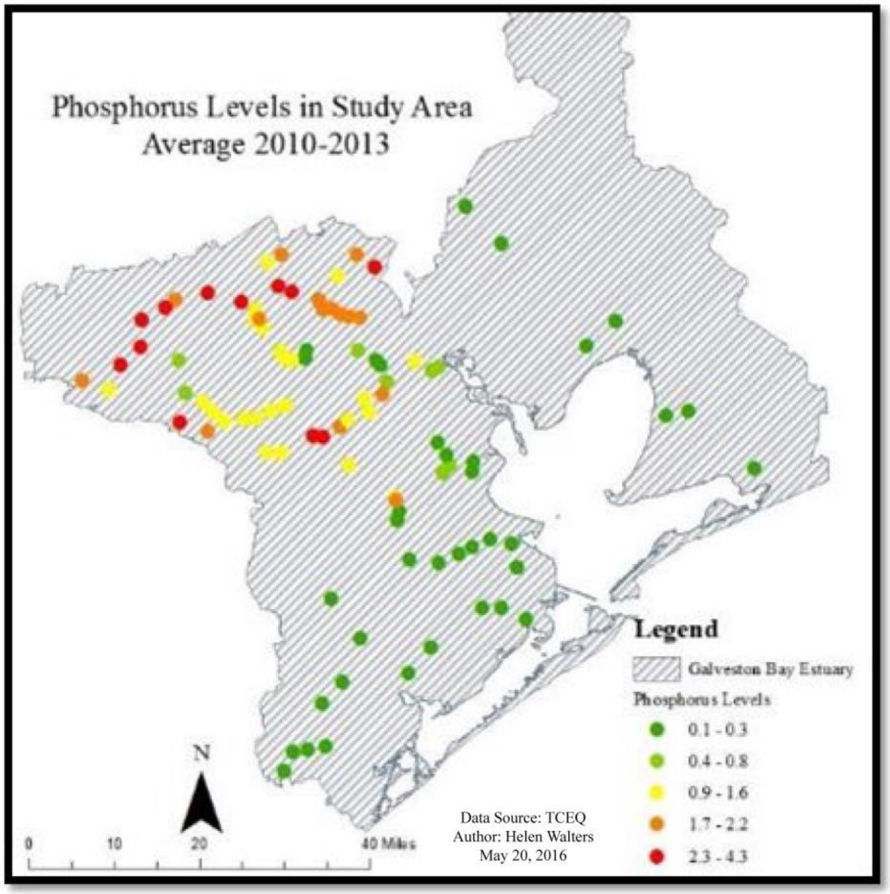


Figure 6: Total phosphorus levels in study area: average 2010-2013. This data is taken from the Texas Commission on Environmental Quality and the sample points are the points used in this study to delineate watersheds.

3.2.2. Land cover

The specific development types that are used in this study are low intensity and high intensity. The classification of low intensity development is from NOAA's C-CAP remote sensing data for 2010. The satellite imagery used for this program is from the Landsat Thematic Mapper data and is at 30 meter grid cell resolution. Low intensity development is classified as land that is 21-49% impervious surface (NOAA-CCAP, 2010). The type of development that is within this classification is large-lot suburban and sprawling rural development (Brody et al., 2013). High intensity development includes area that has constructed materials covering no less than 80% of the area. This land classification most often pertains to “heavily built-up urban centers” (NOAA, C-CAP).

In some studies, urban development is lumped into one large category (Halstead et al., 2014); however, other studies have shown that breaking development up into the three different intensities of development is more effective (Brody et al., 2013). As stated above, the different types of development refer to different percentages of impervious land cover. These different types of land cover (industrial versus suburban versus rural neighborhoods) can have very different effects on the environment and surrounding ecosystem. The land cover of the Galveston Bay Estuary is shown in figure 7. The land cover is 2010 from the National Oceanic and Atmospheric Administration Digital Coast Change Analysis Program. The area shown is the GBE and the land cover categories have been reclassified into the specific categories that are shown in figure 7 (the reclassification table is in Appendix D). The gray outline on the map is the area that the study watersheds cover and the proportion of the land cover within the study is: low intensity development- 15%; high intensity development -8.9%; wetlands – 7.1%; forest- 6.7%; cultivated crops – 6%. This study is driven by low intensity development (as it has the highest proportion of any land cover within the study area).

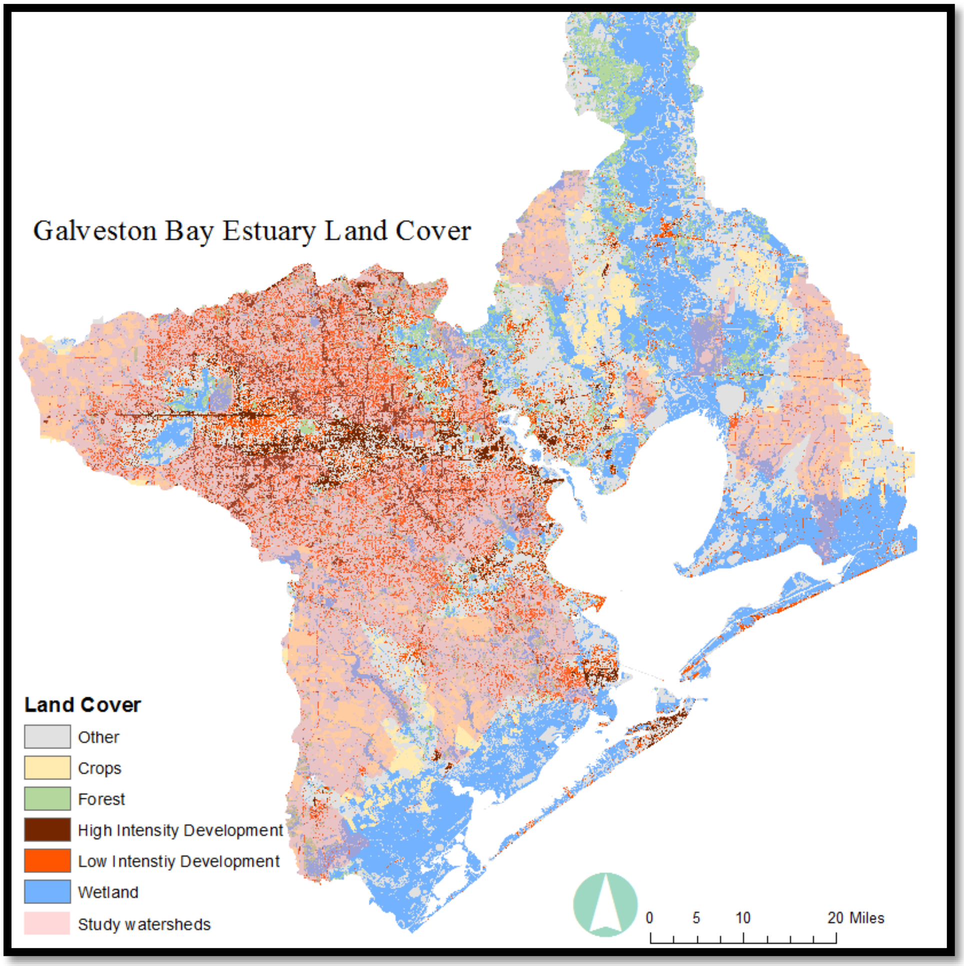


Figure 7: Land cover 2010 from the National Oceanic and Atmospheric Administration Digital Coast Change Analysis Program (NOAA-C-CAP). The land cover is reclassified from the NOAA-C-CAP land cover to highlight the land covers of interest. The area shown is the Galveston Bay Estuary and has been reclassified into the specific categories that are shown in the figure. The light pink outline on the map is the area that the watersheds cover for this study and the proportion of land that is within the study is low intensity development: 15.0%; high intensity development: 8.9%; wetlands: 7.1%; forest: 6.7%; and cultivated crops: 6.0%.

There are seven different landscape metrics that are utilized in this study (one calculated in ArcMap10.2 and six calculated in FRAGSTATS). The landscape metrics are discussed in the literature review and are chosen because they each explain a slightly different aspect of the spatial patterns of development in the landscape. The majority of them have been used in other studies when looking at the spatial configuration of landscapes (Alberti et al., 2007). A description and visual display of each of the metrics used in the study is elaborated on in the flowing sections.

3.2.3. ArcMap10.2 metrics

3.2.3.1. Proximity of development patch to stream (PX_HID and PX_LID)

The only metric that is calculated in ArcMap10.2 is the proximity of each development patch to the nearest stream or river. This metric is calculated using a flow chart that is in Appendix F. The distance is calculated from the centroid of each patch to the nearest river using the stream/river data from the National Hydrography Dataset (NHD).

The greater the development patch number is within a watershed, the more potential for urban runoff to enter the nearby rivers. Having larger development patches and preserving more of the natural ecosystem instead of having multiple smaller patches can be beneficial to ensuring the survival of the natural ecosystem. Having a few numbers of patches, even if they are large, in a watershed is beneficial because it limits the amount of edges (decreases fragmentation) that breakup the natural flow of the ecosystems still existing within the development (Irwin and Bockstael, 2007). A visual display of the difference between a “close” and “far” proximity to stream is shown in figure 8.

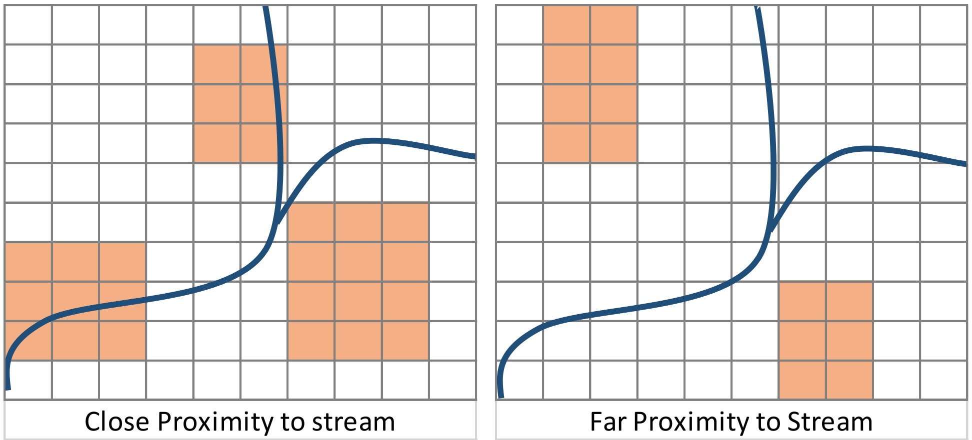


Figure 8: Proximity to stream development metric (calculated in ArcMap10.2). The left diagram shows a stream that has a close proximity to the development patches and the right diagram shows a stream with a farther away proximity to the stream.

3.2.4. FRAGSTATS metrics

3.2.4.1. Patch number (NP) and patch density (PD)

Patch number (NP) is a simple way to measure fragmentation of development. The limiting factor with this metric is that there is no information provided about the area, distribution, or density of the patches. If the areas of the watersheds are held constant, then the number of patches delivers the same message that the patch density metric does. However, for this study, the area is not held constant because the area of the watersheds are different. Because of this, NP only gives part of the story. Therefore, patch density (PD) must be investigated further. PD is the number of development patches divided by the total watershed area and multiplied by a conversion factor to get the units in number per 100 hectares. For both the NP and PD the 8 neighbor rule for delineating patches will be used (Carle et al., 2005; McGarigal, 2015). The visual display of PN and PD is shown in figure 9.

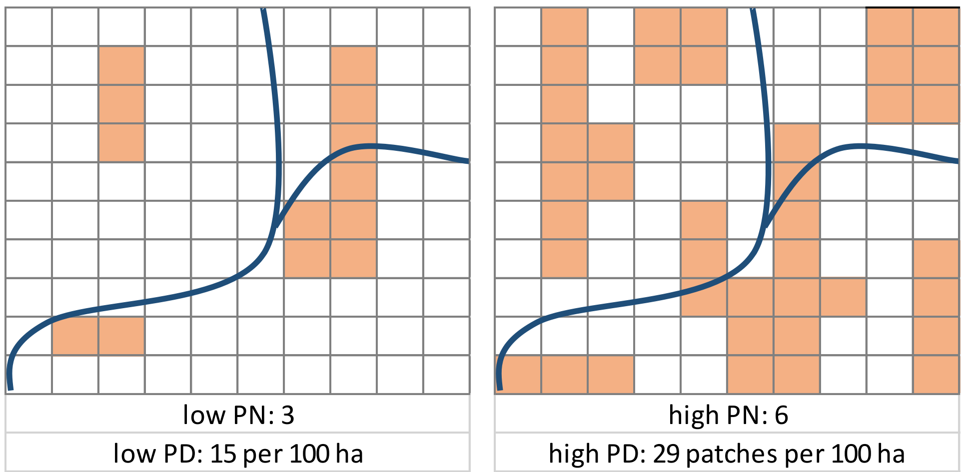


Figure 9 Patch density and patch number (calculated in FRAGSTATS). The left display shows a low patch number and patch density and the right display shows a higher patch number and patch density.

3.2.4.2. Mean patch area (AREA)

Mean patch area (AREA) is a measure of fragmentation. Assuming that the number of patches is held constant, smaller average patch sizes indicate a more fragmented the landscape (Leitao et al., 2006). Mean patch size does not does not address the spatial distribution of the patches, which means that there is no way to tell with this metric how far apart patches are or what their spatial arrangement is over the landscape. Additionally, AREA does not give any indication as to what the range of the patch sizes are within the watershed (McGarigal, 2015). This landscape development metric, calculated in FRAGSTATS is shown in figure 10.

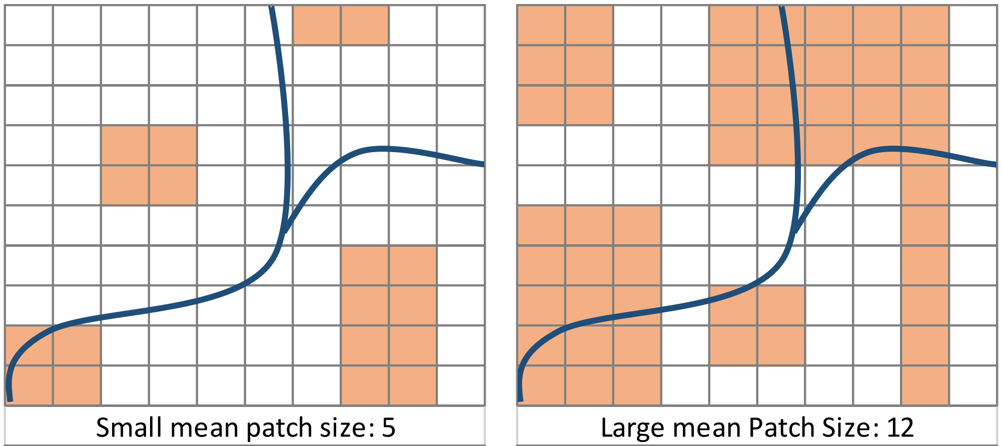


Figure 10: Average area patch size metric (calculated in FRAGSTATS). The left display shows a small average patch size and the right display shows a larger average patch size.

3.2.4.3. Contiguity (CONTIG)

The contiguity (CONTIG) metric measures the sum of all the development (either high intensity or low intensity) patches divided by the number of patches of the same class type within the watershed. CONTIG is used for analyzing the spatial connectedness of the cells within the patch. It gives an indication of the patch shape and how contiguous it is. This metric is quantified by using a 3x3 pixel grid that has integers assigned to each pixel based on how close it is to the centroid of the patch. The value is a function of the number and location of pixels of the same class. This means that the higher the contiguity index, the larger the contiguous patches. The contiguity value will equal zero if it is a one pixel patch and will increase to a value of one with increased connectedness (McGarigal, 2015). A visual display of the CONTIG metric is shown in figure 11.

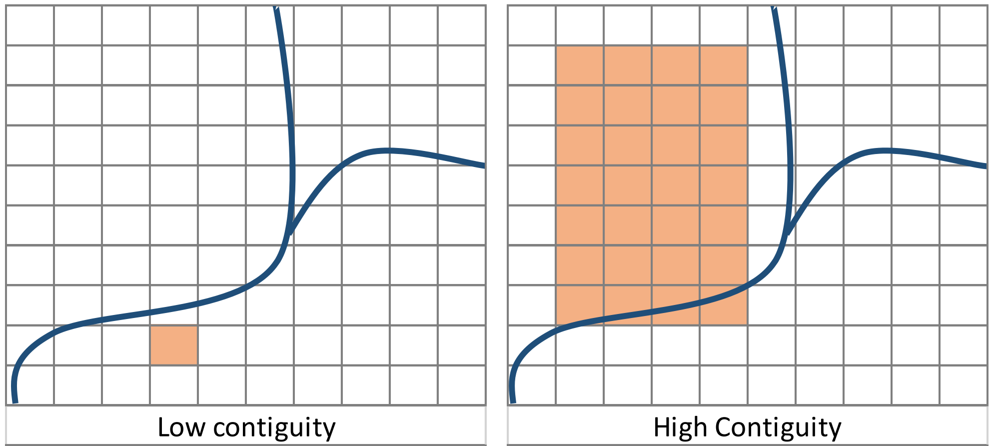


Figure 11 Average contiguity (calculated in FRAGSTATS). The left diagram shows a low contiguity and the right diagram shows a higher contiguity index.

3.2.4.4. Largest patch index (LPI)

Largest patch index (LPI) is the area of the largest development patch within the watershed divided by the total landscape area and measured as a percentage. LPI is simply a measure of dominance, or how big the largest patch of interest is, (McGarigal, 2015). A visual display of LPI (a landscape metric calculated in FRAGSTATS) is shown in figure 12.

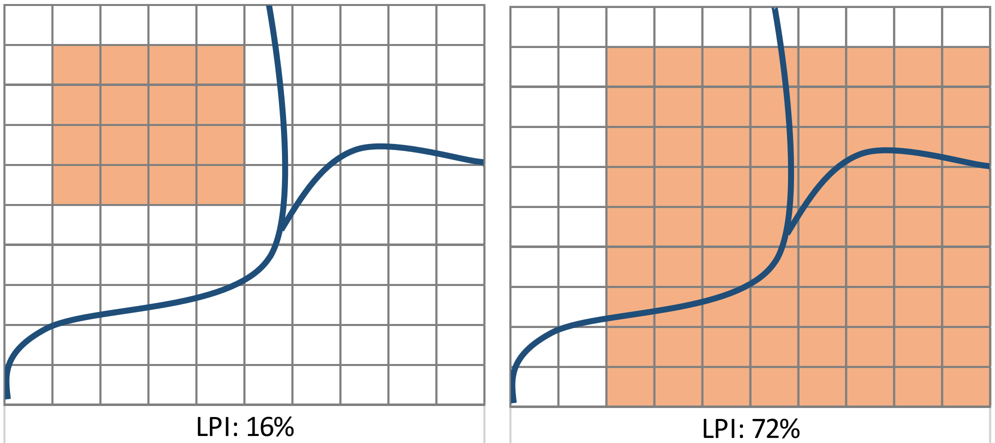


Figure 12: Largest patch index development metric (calculated in FRAGSTATS). The left diagram shows a low largest patch index and the right diagram has a higher largest patch index.

3.2.4.5. Percent of like adjacencies (PLADJ)

Percent of like adjacencies (PLADJ) is another measure of the connectedness within the landscape. This metric takes the number of like (same development type; i.e. high intensity or low intensity development) adjacencies divided by the total number of cells that are adjacent of the same development, expressed as a percentage. Cell adjacencies are calculated using the double-count method, meaning that each pixel order is preserved. This metric is a way to measure how aggregated the development patches are within the watershed. When PLADJ is closer to 0, the patch is highly disaggregated and there are no like adjacencies. When PLADJ is 100, it means that the landscape is entirely one patch (McGarigal, 2015). This metric has been used in multiple other studies to determine how aggregated a landscape is (Alberti et al., 2007). A visual display of the PLADJ metric is shown in figure 13.

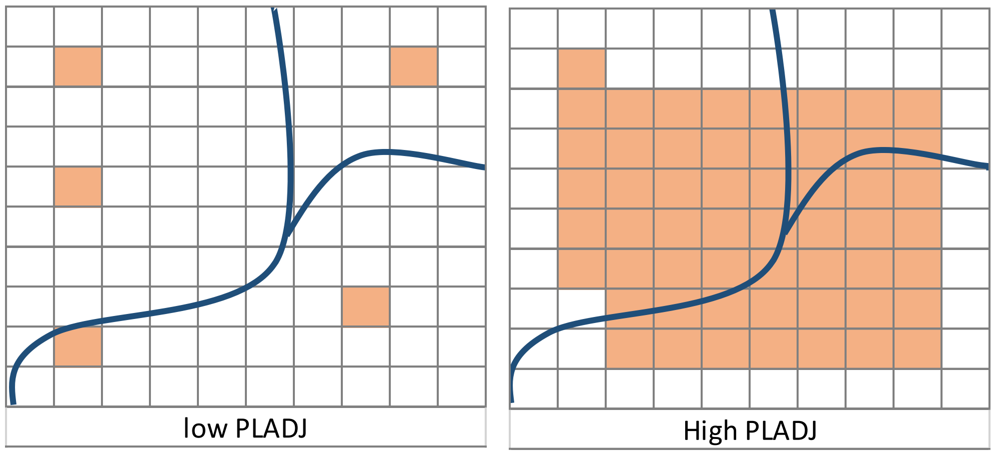


Figure 13: Percent of like adjacencies (calculated in FRAGSTATS). The left diagram shows a low percent of like adjacencies and the right diagram shows a higher percent of like adjacencies.

Each development landscape metric that is used is based on literature as well as how applicable they are to being implemented in planning and policy. Each of the development landscape metrics that are calculated in FRAGSTATS have what studies they were utilized in as well as the equations that FRAGSTATS utilizes to calculate each (table 1).

*Table 1: FRAGSTATS landscape metrics equations and the studies these metrics have previously been utilized in. Patch density: ni = number of patches in the landscape of class I; A = total landscape area (m). Largest patch index: a_{ij} = area (m^2) of patch ij; A = total landscape area (m^2). Patch size: a_{ij} = area of patch$_{ij}$; n_i = number of patches in landscape of patch type (class)$_{i.}$ Number of patches: n_i = number of patches in the landscape of class type i. Percent of like adjacencies: g_{ij} = number of like adjacencies (joins) between pixels of patch type (class) I based on the double-count method; g_{ik} = number of adjacencies (joins) between pixels of patch type (classes) I and k based on the double count method. Contiguity: C_{ijr}= contiguity value of r pixel r in patch ij; V = sum of values in a 3x3 cell template; A_{ij} * = area of patch ij in terms of number of cell.*

Patch Metrics	Study	Equation
Patch Density	Sun et al. (2014)	$PD = \frac{n_i}{A}(10{,}000)(100)$
	Huang et al. (2012)	
	Brody et al. (2013)	
	Lee et al. (2009)	
Largest patch index	Sun et al. (2014)	$LPI = \frac{\max_{j=1}^{n}(a_{ij})}{A}(100)$
	Huang et al. (2012)	
	Lee et al. (2009)	
Patch Size	Sun et al. (2013)	$Mean\ Patch\ Size = \sum_{j=1}^{n} \text{aij}$
	Alberti et al. (2007)	
	Carle et al. (2005)	
	Aguilera et al. (2011)	
Number of patches	Sun et al. (2013)	$NP = n_i$
	Brody et al. (2013)	
	Aguilera et al. (2011)	
Percent of like adjacencies	Alberti et al. (2007)	$PLADJ = \left(\frac{g_{ii}}{\sum_{k=1}^{m} g_{ik})}\right)(100)$
Contiguity	Connectance used in Brody et al. (2013)	$CONTIG = \frac{\left[\frac{\sum_{r=1}^{z} c_{ijr}}{a_{ij}^{*}}\right] - 1}{v-1}$
	Multiple connectivty metrics in Alberti et al. (2007)	
	cohesion index used in Lee et al. (2009)	

The descriptive statistics for patch metrics that are used for low and high intensity development are in tables 2 and 3. The descriptive statistics in both tables 2 and 3 include the independent variables of interest, detail about the measure, mean, standard deviation, minimum, and maximum of each variable.

Table 2: Low intensity development metrics descriptive statistics. Descriptions taken from (McGarigal, 2015).

Variable	Measure	Mean	Std. Dev.	Min	Max
PX_LID	Average proximity of development patch to stream/river.	1.05	0.30	0.53	2.20
AREA	Sum of all patches of corresponding landscape type divided by the number of patches.	0.70	0.22	0.26	1.46
CONTIG	Average contiguity value for the cells in a patch 1/ sum of the template values -1.	0.16	0.02	0.10	0.20
NP	Total number of patches	2169.83	1964.21	79	8594
PD	NP/Watershed Area	21.60	9.90	1.10	36.16
LPI	% Largest Patch	1.41	1.45	0.03	7.98
PLADJ	Number of like adjacencies between pixels of same class based on double count method/ number of adjacencies between pixels of classes summed then multiplied by 100 to obtain percent.	45.48	4.85	32.22	61.15

Table 3: High intensity development metrics descriptive statistics. Descriptions taken from (McGarigal, 2015).

Variable	**Measure**	**Mean**	**Std. Dev.**	**Min**	**Max**
PX_HID	Average proximity of development patch to stream/river.	1.07	0.34	0.07	2.29
AREA	Sum of all patches of corresponding landscape type divided by the number of patches.	0.95	0.56	0.15	2.66
CONTIG	Average contiguity value for the cells in a patch 1/ sum of the template values -1.	0.19	0.044	0.07	0.45
NP	Total number of patches	817.05	846.30	2	3674
PD	NP/Watershed Area	8.26	5.07	0.03	19.76
LPI	% Largest Patch	2.19	3.21	0.003	17.47
PLADJ	Number of like adjacencies between pixels of same class based on double count method/ number of adjacencies between pixels of classes summed then multiplied by 100 to obtain percent.	59.03	10.40	16.67	76.16

3.3. Control variables

There are eight control variables used in this study. All of the control variables are based off literature as well as what other aspects of the ecosystem affect phosphorus levels. Precipitation is one aspect of the ecosystem that can affect phosphorus because the more precipitation the more runoff there may be into the rivers but also the higher the dilution of phosphorus within the rivers. Septic systems also have the potential to influence phosphorus levels. Drainfields in septic systems are supposed to filter out the nutrients before the reach the ground water but sometime this is not effective and the nutrient levels will be greater around septic system. In addition, the more watershed area contributing drainage to a stream, (the more runoff and impacts from the development) the higher the nutrient runoff into the stream. These along with multiple land cover types can affect the phosphorus levels and are included in the study as control variables. Each control variable is discussed in more depth following.

3.3.1. Precipitation

The precipitation data used in this study are from the National Oceanic and Atmospheric Administration (NOAA) National Climatic Database and is average annual precipitation during 1981-2010 (data accessed from the Texas Natural Resources Information Systems (TNRIS) Climatological Data). The precipitation was calculated for each watershed and reported in centimeters. The National Climatic Data center provides monthly precipitation data for the United States, which is obtained through interpolations of daily precipitation records (figure 14).

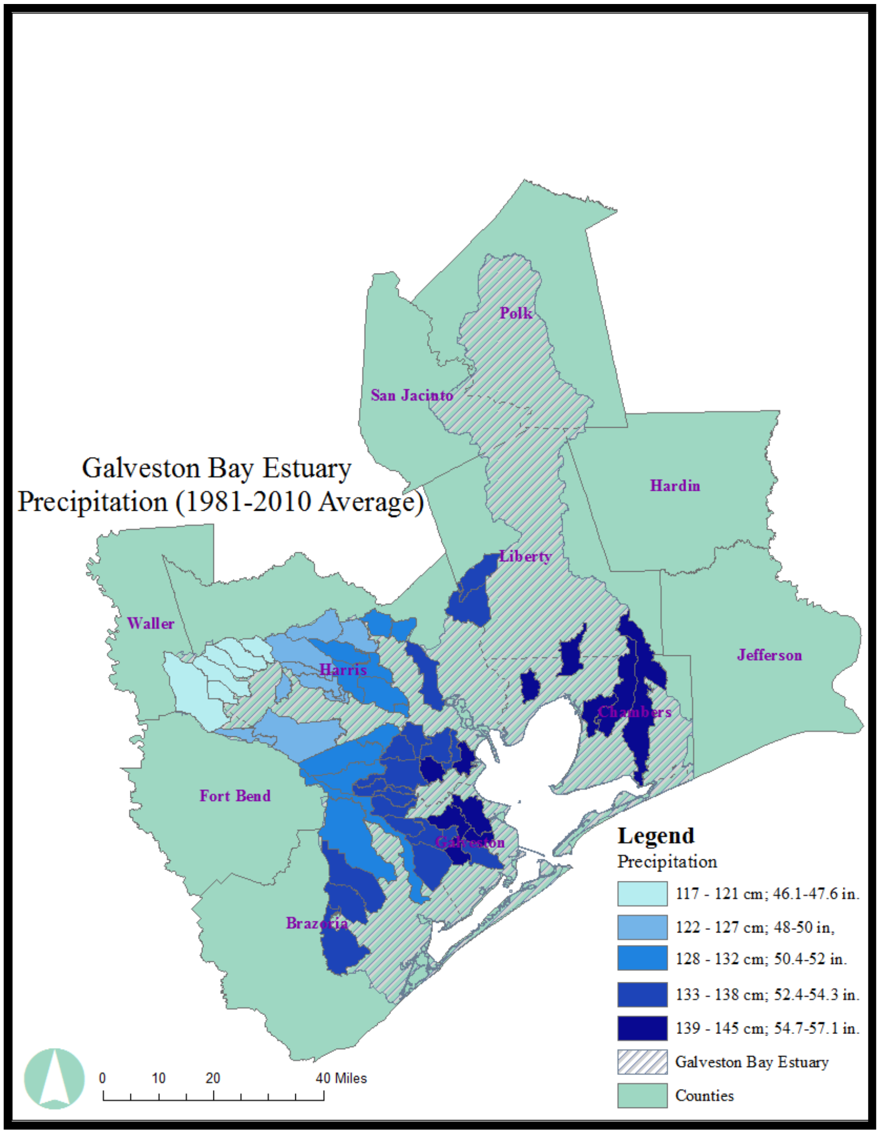


Figure 14: Galveston Bay Estuary annual precipitation based on NOAA National Climatic Database 30 year average. The National Climatic Data center provides monthly precipitation data for the United States, which is obtained through interpolations of daily precipitation records.

3.3.2. On site septic facilities (OSSF)

The OSSF permits were obtained from the Houston-Galveston Area Council (H-GAC). Phosphorous concentrations can be affected by septic systems because the drainfield has the potential to reduce the phosphorus levels (MPCA, 1999). A drainfield is the portion of the septic tank that serves as an area to remove contaminants from the fluid before it is released into the ground water. However, often times when there are lots of septic systems there is an increase in nutrients, which could potentially affect the phosphorus levels in streams. Dudley and May (2007) show that the proportion of phosphorus loading into rivers can be attributed to septic systems is on average of 12% (based on the catchments reviewed, with a range of 3-58%, and utilizing the middle value when a range is given for a catchment). A visual display of the septic systems in the study area is shown in figure 15. In other studies it has been shown that septic systems are linked to higher levels of phosphorus (Moore et al., 2003). It should be noted however that sometimes septic systems can also cause nitrogen loading within streams (Bernhardt et al., 2008). The time period of the septic systems is 1984-2013.

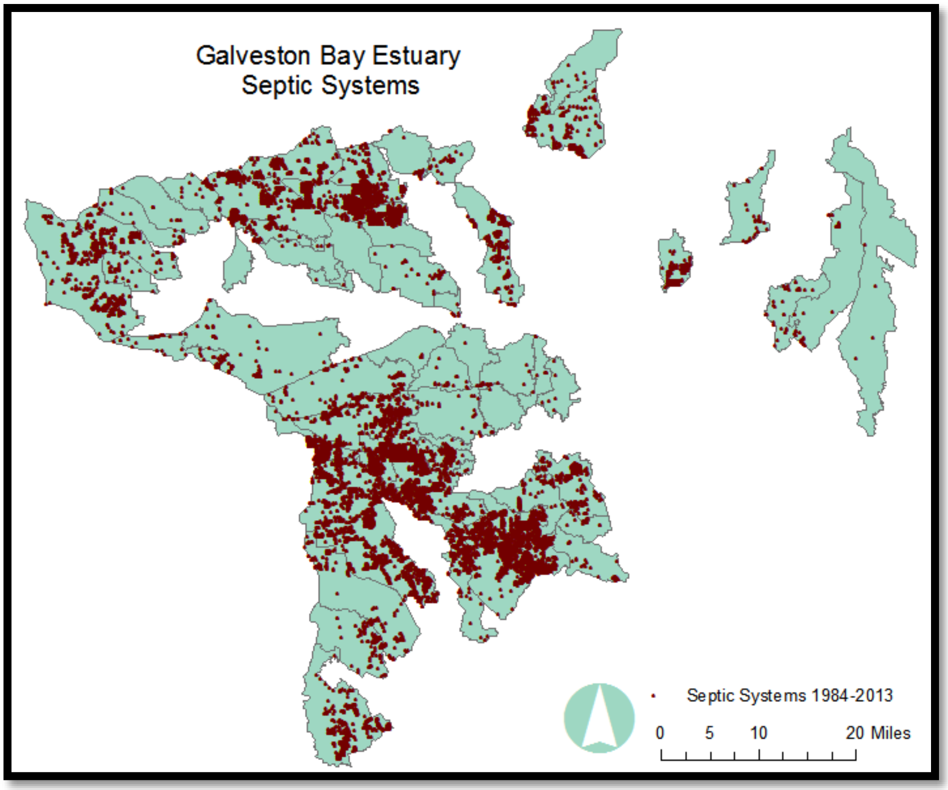


Figure 15: On-site septic facilities in the study area. Data obtained from the Houston-Galveston Area Council. The septic system data has septic systems from 1984-2013 shows each of the permits that were approved to be put in.

3.3.3. Area and contributing drainage

Two other control variables used in the study are derived from the watershed characteristics: watershed area and contributing drainage. When delineations were performed in ArcMap10.2, the area for each watershed was also calculated. Then, a "nesting order" was assigned to each watershed to show which ones fall spatially on top of other ones. The area of the largest watershed in the nesting order would be assigned the contributing drainage area and all of the smaller watersheds within would be assigned the same area for the contributing drainage variable.

For example, figure 16 shows one set of watersheds that are spatially nested within each other. The black watershed has the largest area; therefore, the area of the largest watershed within the nesting structure was assigned as the contributing drainage for all of the other watersheds that lie on top of it (figure 16).

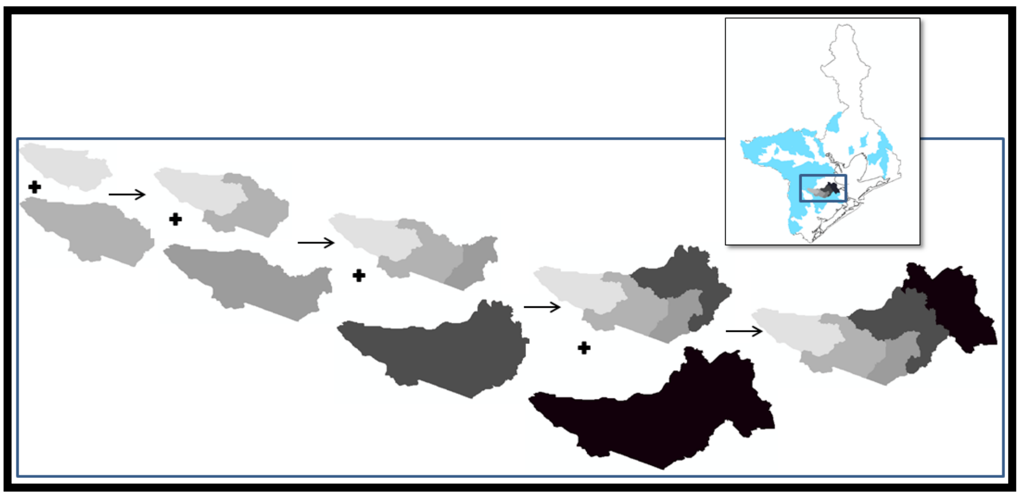


Figure 16: An example of the nesting watershed patterns that occur when watersheds are delineated from TCEQ water quality monitoring gauge points. The black watershed is the bottom watershed that has the most area and each of the grey watersheds (all varying shades of grey) fit spatially within the black watershed. This nesting pattern is due to the flat topography as well as how close together the water quality monitoring gauges are located.

This nesting occurs due to the relatively flat topography of the Galveston Bay Estuary and the relatively small distance between which the samples were taken. Each watershed was delineated from a water quality monitoring gauge point, and some gauge points delineated a short distance while others delineated a larger distance that directly overlapped the small distances. The process by which the watersheds nest is shown for a subset of the study area (figure 16). Utilizing watersheds that are spatially nested has been done in other studies (Carle et al., 2005). In Carle et al. (2005) it is noted that there may be some spatial autocorrelation that is due to this nesting structure.

3.3.4. Land cover controls

The other four control variables that are used in the study are different types of land cover. The details about how these land cover types affect phosphorus levels is discussed in the Literature Review. Forests are one land cover type that can affect phosphorus. Forests can act as a filtration system for nutrients and reduce the amount of nutrients that run off into the stream. However, the landscape metrics of forests can influence the level at which they act as a filtration system.

Wetlands are another land cover that can affect phosphorus. Wetlands can serve as filtration systems for nutrients (Reddy et al., 1999). This means that the more wetlands there are in a given area, the fewer nutrients one could expect to see in surrounding rivers and streams. Agricultural land can be a large contributor of nutrient loading into rivers (Lenat and Crawford, 1994). High levels of fertilizer use and runoff in crops or residential land can result in increased levels of phosphorus in nearby streams and rivers due to surface runoff (Hart et al., 2004).

The two types of development that are controlled for in this study are high and low intensity development. Even though the independent variables of interest in the study are high and low intensity metrics, it is important to control for the proportion of both of these types of development within the study area. The different types of development can have large effects on the nutrients that enter into nearby streams. Medium intensity development is not included because it is highly collinear with the percent of high intensity development and the percent of low intensity development (0.80 correlation with high intensity development and 0.79 with low intensity development opposed to the correlation between low intensity development and high intensity development of 0.62). These land cover types are shown in figure 7 on page 43. All of the descriptive statistics for the control variables used in the study, as well as the dependent variable (average total phosphorus), are in table 4.

Table 4: Descriptive statistics for control variables

Variable	**Mean**	**Std. Dev.**	**Min**	**Max**
Phosphorus (log)	-0.41	1.10	-2.80	1.47
Phosphorus (untransformed) (mg/l)	1.08	0.93	0.06	4.34
Area (km^2)	103.54	68.73	25.98	272.9
Precipitation (in)	51.47	2.74	45.88	57
Contributing Drainage (km)	159.62	82.91	26.46	272.90
OSSF Count	388.899	459.27	0	2149
Wetland Land Cover (%)	7.08	9.12	0.134	60.80
Low Intensity Development (%)	14.96	7.46	0.2827	29.40
High Intensity Development (%)	8.862	7.62437	0.0176	34.5358
Cultivated Crops Land Cover (%)	6.5	9.69	0.0073	43.37
Forest Land Cover (%)	6.743	4.657848	0.2633	23.4466

3.4. Analytical process

The analytical process is a detailed flow chart of the conceptual model (figure 5 on page 34). The process by which each variable was analyzed and utilized in the statistical models is shown in the analytical process (figure 17). Patch metrics (independent variable) were analyzed for the study area along with both environmental and land cover controls. The model's output shows the relationship between the dependent variable (total phosphorus) and the independent development patch metrics examined in this study.

There are many different data sources used in this study. The watersheds were delineated using a Houston/Galveston DEM from NHDPlusV2 and the water quality points were obtained from TCEQ SWQMS data that was previously cleaned by HARC. The land cover that was used for the FRAGSTATS development metrics as well as the land cover controls is from NOAA C-CAP Digital Coast. Precipitation is from NOAA Climatological Database (obtained from the TNRIS database) and septic systems are from H-GAC.

Each data source underwent a manipulation process to quantify it in terms of the watersheds used in the study. The analytical process (figure 17) shows the relationships between the raw data, processes conducted to manipulate the data, programs used to conduct the manipulation, and the statistical analysis that was conducted at the conclusion of the study.

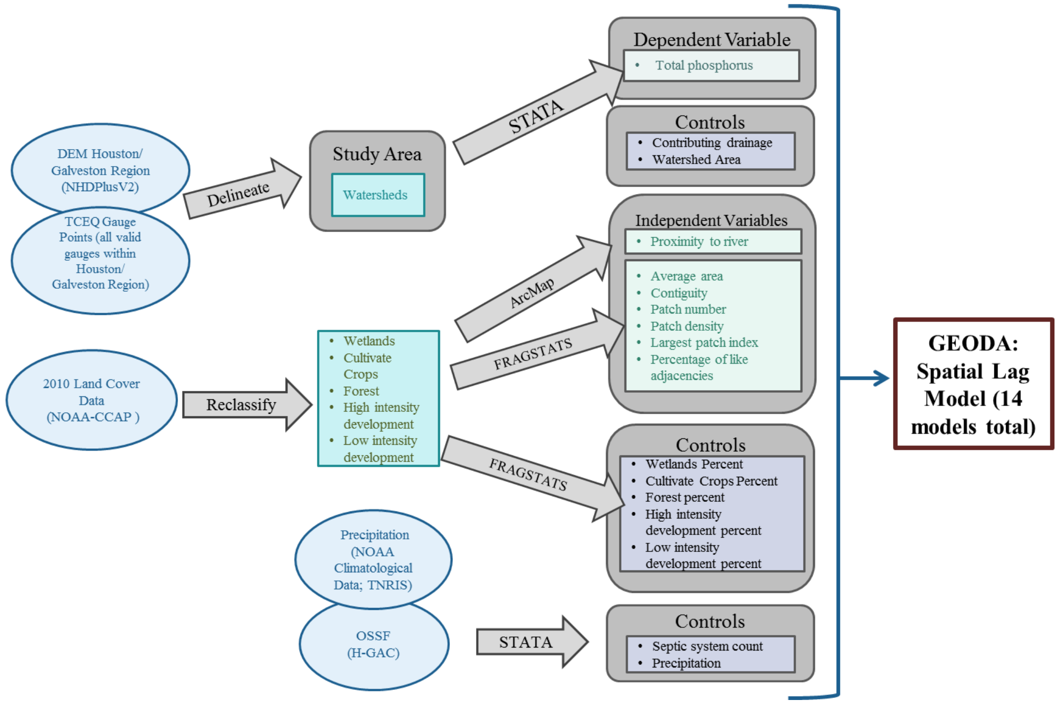


Figure 17: Analytical process: All of the data sources utilized in this study along with the processes that were carried out to create the variables used in the 14 different spatial lag models.

4. STATISTICAL METHODOLOGY

Initially, simple ordinary least squared (OLS) regression was attempted but there was high collinearity and spatial autocorrelation among the resulting variables. One of the largest problems was derived from the nesting pattern of the watersheds within the study area. Occasionally, when observations are spatially clustered, using a robust cluster can help account for the multicollinearity and spatial autocorrelation between the observations. Using a robust cluster can also help remove the independent assumption between the groups of clustered variables (STATA, 2013). However, this method does not always correct for all of the spatial autocorrelation issues and additionally, there can still be problems of inter-cluster correlation (Williams, 2000). This problem arose in this study when the robust OLS models were run; there was still spatial autocorrelation in the residuals for the majority of the models. For this reason, robust clustered OLS regression no longer was a viable option.

There are two statistical options present for controlling for spatial autocorrelation in a model: spatial lag and spatial error models. These options are both specifically for spatial regressions and help account for spatial autocorrelation. In choosing which model (either spatial lag or spatial error) to utilize the Lagrange multiplier is looked at. In this study, the maximum likelihood spatial lag model had Lagrange multiplier that was the best fit for the data.

The program utilized for this spatial statistical analysis is GEODA. The spatial lag model utilizes a spatial lag variable that is derived from a spatial weights matrix. The weights matrix for all of the development metrics models is a Euclidean distance based matrix with lowest distance threshold of 4.28 miles, determined by GEODA. We then checked to ensure that there were no observations without neighbors in the weights matrix. This is verified in the histogram shown in figure 18. This histogram shows that there are no observations without any neighbors based on the Euclidean distance based matrix utilized for this study. We used a geographic coordinate system North American Datum 1983 with an Albers projection to create the latitude and longitude for the weights matrix.

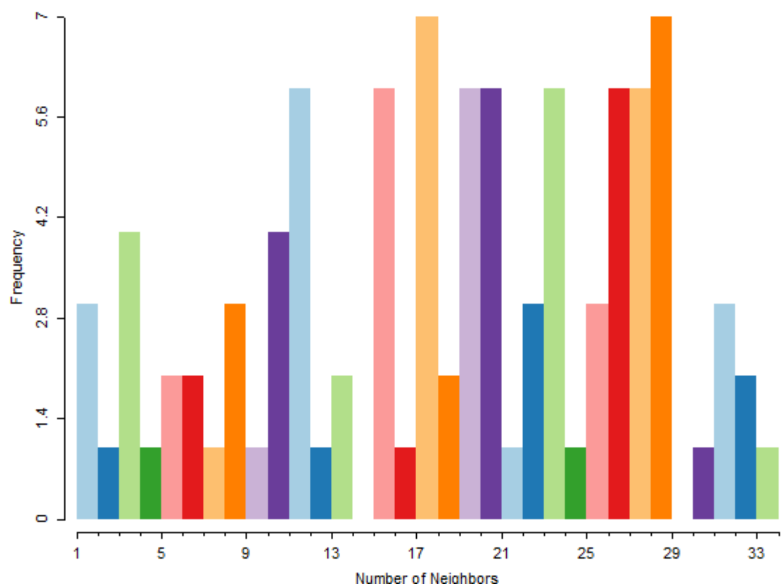


Figure 18: Histogram showing the number of neighbors for each sample in the weights matrix utilized for the maximum likelihood spatial lag models.

A maximum likelihood spatial lag model was used to explain the relationship between average total phosphorus levels and development metrics within the study area. Each development metric model was analyzed with all of the control variables previously mentioned (unless there was high collinearity between some of the variables and then this was accounted for by removing the highly correlated control variable). Each of the independent low intensity development and high intensity development class metrics were analyzed in separate spatial lag models to avoid multicollinearity between development metrics. In addition, the control variables that were used in all of the models were only selected if they were not highly correlated with other control variables (based on the correlation matrix in Appendix A and the variance inflation factor (VIF) obtained from a simple OLS regression shown in Appendix B). Halstead et al. (2014) also used this method of removing multicollinearity by removing control variables if necessary. Each of the development metrics were tested for multicollinearity using a variance inflation factor (VIF) test on a simple OLS model. If there was higher than acceptable VIF (greater than 10) then the control variable creating the high VIF was removed. If the variable that has the high VIF is the independent variable then this variable is left in and it is noted in Appendix B that this issue has occurred. The reason that even when there is high collinearity with the independent variable the variable could not be dropped is because all of the variables that are included in the model have an influence on the phosphorus levels. In this case (which is shown in Appendix B if it occurs) there should be caution used when interpreting coefficients. This is the same process that was utilized in Carle et al. (2005) if the situation occurred.

In total, there were fourteen models run in this study. There were seven different development metrics utilized for both high intensity and low intensity development. Each model contained all of the control variables (except when previously noted) and the dependent variable for all models was total phosphorus. Specifically, for the low intensity development metrics, the only one that had issues with the VIF was patch density. In the patch density model the variable that had a higher than accepted VIF (greater than 10) was the independent variable of interest: low intensity development patch density. Because this is the variable that is of interest in the model it cannot be removed from the model and instead it should be noted that the variance inflation is slightly higher than accepted and the patch density variable could not be removed.

When looking at the high intensity development models, only one metric (average patch area) violated the variance inflation factors. The VIF for the average high intensity development patch area is higher than the accepted value and therefore for the high intensity development average patch area model the proportion of high intensity development control variable was removed from the model.

In addition, some of the models have high heteroscedasticity. Using a spatial regression can remove some of the heteroscedasticity but in some circumstances it does not remove all of it. Therefore, it should be noted that there are some models with higher than commonly accepted homoscedasticity.

As previously mentioned, the study area's watersheds exhibited a frequent nesting pattern. The spatial autocorrelation problem that arose from this issue was combated using a spatial lag model. Even though the nesting watersheds pose a statistical challenging, it is important to still include them in the model because they fully capture the area of study. Nested watersheds have been used in other literature and collinearity has been evaluated between the watersheds to limit the effect of spatial dependency (Carle et al., 2005). All of the regression results and VIF tables are included in appendix B.

5. RESULTS

Fourteen spatial lag models were used to examine the relationship between total phosphorus levels, low intensity development, and high intensity development metrics while controlling for multiple environmental and land cover variables. The results for the different models are presented in the following section. Both low intensity development and high intensity development model controls are also discussed.

5.1. Low intensity development models

Table 5 shows the effects of the low intensity development metrics for the seven development metrics. The models explain between 77 and 79 percent of the variance in the average total phosphorus levels.

*Table 5: Maximum likelihood spatial lag regression models, showing low intensity development metrics, coefficients and z-values. Significance levels are shown by * designations. P< 0.05: *; P< 0.01: **.*

Low Intensity Development Metric	**Coefficient**	**Z-value**
Patch Number (#)	0.00008	1.00
Patch Density (#/100 ha)	0.066	3.61**
Average Patch Size (ha)	-0.96	-2.31*
Average Contiguity (index)	-8.38	-1.88 (p=0.06)
Percent of like adjacencies (%)	-0.036	-2.16*
Largest Patch Index (%)	-0.081	-1.39
Average proximity of patch to streams (km)	0.27	1.37

P<0.05: *; P<0.01: **

There are three low intensity development metrics that were significant in the models: patch density ($p<0.01$), average patch size ($p<0.05$), and percent of like adjacencies ($p<0.05$). In addition, average contiguity was significant at $p = 0.06$. Patch density was significant positively whereas average patch size, percent of like adjacencies, and average contiguity were significant negatively.

As the patch density increases by 1 unit (one patch per 100 hectares) the phosphorus levels increase by 6.6%. This result backs up the hypotheses that states when the fragmentation of the landscape (more patches over a certain area) increases, the phosphorus levels increase within the streams. Average patch area has a negative relationship with phosphorus which shows that as the average patch size increases by 1 hectare the average total phosphorus levels decrease by 96%. This should be noted that the average low intensity development patch size within the study is 0.721 hectares, so an increase of 1 hectare is more than doubling the average patch size. As the percent of like adjacencies increases by 1 percent in this model, the average total phosphorus levels decrease by 3.6% across the study. This aligns well with the hypothesis that the less fragmented (more aggregated) the landscape is, the lower the phosphorus levels are within the rivers. All three of the low intensity development metrics that are significant in these models show that having a less fragmented landscape is beneficial to reducing the phosphorus levels in the study area.

5.2. High intensity development models

Table 6 shows the effects of the high intensity development metrics for the seven different high intensity development metrics. The models explain between 76 and 78 percent of the variance in the average total phosphorus levels.

*Table 6: Maximum likelihood spatial regression models high intensity development metrics coefficients and z-values. Significance levels are shown by * designations. P< 0.05: *; P< 0.01: **.*

High Intensity Development Metric	**Coefficient**	**Z-value**
Patch Number (#)	0.00008	0.53
Patch Density (#/100 ha)	0.031	1.12
Average Patch Size (ha)	-0.35	-2.89**
Average Contiguity (index)	0.71	0.51
Percent of like adjacencies (%)	0.014	1.46
Largest Patch Index (%)	-0.069	-1.84
Average proximity of patch to streams (km)	0.009	0.057

P<0.01: **

The only metric that is significant for the high intensity development metrics models is average patch area. It should be noted that for the high intensity development average patch area model the proportion of high intensity development control variable was removed from the model. The reason for this was because the VIF for the proportion of high intensity development was higher than acceptable. The high intensity average patch area shows that the greater the average patch area, the lower the phosphorus levels in the streams. This aligns with the low intensity development models showing that larger patch area means less edges and less area for runoff.

5.3. Control variables

Of the eight control variables used in the study, four were on average significant: precipitation, percent of low intensity development, percent of high intensity development, and percent of forest land cover. Every spatial lag model used the same set of control variables (except when specified: see Appendix B for regressions and more details). No variables that had different correlation signs (positive versus negative) for the various metrics were significant within the models. These variables are indicated with a "~". Tables 7 and 8 show the control variables for the low intensity development and high intensity development spatial lag models, respectively, the correlation signs, and the significance levels for each control variable.

*Table 7: Control variables entered into low intensity development regression models and their effects on average phosphorous levels (sign and significance shown). If p < 0.1 then there is *, if p <0.05 then there is **. The signs shown for the variables that are on average significant (at both 90 and 95% confidence levels) do not change signs from the different models. The ones with a ~ have at 6 out of the 7 metrics with the sign reported.*

Variable	***Sign and significance***
Precipitation (in)	-**
Contributing Drainage (km)	+
OSSF (Count)	-
Watershed Area (km)	+
Percent of wetland land cover	-
Percent of low intensity development land cover	+**
Percent of high intensity development land cover	-**
Percent of cultivated crops land cover	~
Percent of forested land cover	+**

P< 0.1: * P<0.05: **

*Table 8: Control variables entered into high intensity development regression models and their effects on average phosphorous levels (sign and significance shown). If p < 0.1 then there is *, if p <0.05 then there is **. The signs shown for the variables that are on average significant (at both 90 and 95% confidence levels) do not change signs from the different models. The ones with a ~ have at 6 out of the 7 metrics with the sign reported.*

Variable	***Sign and significance***
Precipitation (in)	-**
Contributing Drainage (km)	+
OSSF (Count)	-
Watershed Area (km)	+
Percent of wetland land cover	-*
Percent of low intensity development land cover	+*
Percent of high intensity development land cover	-**
Percent of cultivated crops land cover	~
Percent of forested land cover	+**

P< 0.1: * P<0.05: **

Precipitation is on average significant (P<0.05) negatively within the high intensity and low intensity development metrics models. This means that as the precipitation increases the phosphorus levels decrease. One of the other control variables that is on average significant is percent of low intensity development land cover (P<0.05) in the low intensity development metrics and P<0.1 in the high intensity development metrics models) positively. As the percent of low intensity development increases so do the phosphorus levels in the stream on average across the study area. Controversially, percent of high intensity development is negatively significant (P<0.05 for both the low intensity and high intensity development metrics models) which shows that the more high intensity development there is the lower the phosphorus levels across the study area. This could be an indication that phosphorus runoff is mostly from low intensity development (i.e. suburban development).

The final control variable that was on average significant is percent of forested cover. Percent of forestland cover is significant positively ($P<0.05$ for both the low intensity and high intensity development metrics models) which shows that the more forest there is the greater the phosphorus levels. This is not necessarily intuitive but could be because of a high level of forest fragmentation which can potentially be detrimental to water quality.

Percent of wetland land cover is on average significant in the high intensity development metrics model at a 90% confidence level ($P<0.1$). This land cover control shows that the greater the percent of wetland land cover within the watershed the lower the average phosphorus level which aligns with studies that show wetlands act as a filtration for nutrients (Johnston, 1991; Liu and Cameron, 2001; Reddy et al., 1999).

The other control variables that are included in the study are not significant on average are contributing drainage, septic system count, watershed area, and percent of cultivated crops land cover. Contributing drainage has a positive relationship showing that the greater the contributing drainage to a watershed the higher the phosphorus levels. Septic system count has a negative relationship meaning that the more septic systems within the watershed the lower the phosphorus levels. Watershed area shows a positive relationship, which means the greater the watershed area the greater the phosphorus levels. Cultivated crops land cover has differing relationship for multiple of the high intensity and low intensity development metrics models.

6. DISCUSSION

This study examines the relationship between phosphorus levels and low and high intensity development metrics in the Galveston Bay Estuary. Phosphorus is utilized as an indicator for nutrient loading (although it should be noted that there are other nutrients such as nitrogen that are also indicators of nutrient loading and the resulting eutrophication that can potentially occur). This discussion assesses the relationships between low and high intensity development with phosphorus levels within the Galveston Bay Estuary.

6.1. Low intensity development models

Average patch size has a significantly negative effect ($p<0.05$) on average total phosphorus levels in the low intensity development metrics models. This metric shows that the greater the average patch size of low intensity development within the watershed, the lower the phosphorus levels. Therefore, if there are on average very large patches within a watershed, the landscape is less fragmented compared to a landscape that has small broken up patches of low intensity development. This aligns with the hypothesis that having a less fragmented landscape of low intensity development metrics will result in lowering phosphorus levels. Based on the results, when the average patch size increases by 1 hectare (equal to 10,000 square meters) the average total phosphorus levels decrease by 95.7%. It should be noted that the average low intensity development patch size within the study is 0.70 hectares, so an increase of 1 hectare is more than doubling the average patch size.

Patch density (PD) has a positively significant ($P<0.01$) on phosphorous levels, meaning the greater the patch density the greater the phosphorus levels in surrounding rivers and streams. The patch density metric normalizes the patch number by dividing by the total area of the watershed. These results support previous findings that show the more patches there are per area the greater the anthropogenic impacts on the natural environment (Carle et al., 2005). Therefore, the more densely developed a watershed (in relation to the low intensity development that is categorized as 20-49% impervious surface), the greater the TP loading. It should be noted that patch number is sensitive to watershed size. This is one reason why watershed area is controlled for in the model. When a large watershed has the same number of patches as a smaller watershed, the development metric of patch number can be misleading. However, when patch density is used and the variation in area of a watershed is controlled for, a more complete picture of the fragmentation within the watershed is established. When the patch density decreases by 1 unit (one patch per kilometer which is equal to 100 hectares) then the average phosphorus levels decreases by 6.56%. This result backs up the hypothesis that states when the fragmentation of the landscape (more patches over a certain area) increases, the phosphorus levels increase within the rivers.

Percent of like adjacencies (PLADJ) is maximized if the watershed consists solely of one land cover type. PLADJ is minimized if the patch type is completely disaggregated. This study shows that percentage of like adjacencies ($P<0.05$) is greater when the phosphorus levels are lower, implying a negative relationship. The greater the PLADJ, the more continuous the low intensity development patches are in the landscape. This finding indicates that the higher the PLADJ (and therefore the higher the connectedness of the patch), the lower the phosphorus levels will be in the surrounding rivers. The reason for this relationship is potentially because a higher percent of like adjacencies indicates a more connected landscape of low intensity development and with a more connected and aggregated landscape of low intensity development, there is less fragmentation of the environment and a higher level of ecosystem preservation. In this case, when the average percent of like adjacencies are increased by 1 percent, the average phosphorus levels decrease by 3.6%. This aligns with the hypothesis that having a more connected landscape (higher percent of like adjacencies) will reduce the phosphorus levels within the watershed.

The final low intensity development metric of note is contiguity (CONTIG) which is significant at p= 0.06. This is not significant at the 95% confidence level, but significant at the 10% confidence level, which is worth discussing. This variable is negatively correlated with phosphorus levels in the watershed. The more contiguous the patches are, the lower the phosphorous levels are in the surrounding rivers. The reason for this is that (as shown by other metrics) increased fragmentation and decreased connectedness in areas of low intensity development increases the level of phosphorous in the watershed. As stated by the previous three low intensity development metrics that are significant in the low intensity development metrics models, the hypothesis of decreasing fragmentation is validated in the relationship seen with low intensity development contiguity.

6.2. High intensity development models

There is only one high intensity development metric that is significant out of the seven different high intensity development metrics that are analyzed. This study shows that the greater the high intensity development patch size, the lower the phosphorus levels. According to this study, if the average patch size increases by 1 hectare, the average total phosphorus levels decrease by 34.5% (the average patch size of high intensity development is higher that of low intensity development: 0.95 hectares). This aligns with the hypothesis that less fragmented development patterns results in less phosphorus levels within the streams. The significance of this high intensity development metric lines up with the low intensity development metric of average patch size which also has a negative relationship with phosphorus levels within the study.

One of the reasons that there may be no other significant metrics in the high intensity model is because the study area is driven by low intensity development as shown by the land cover in figure 7 on page 43. As mentioned previously, the proportion of low intensity development in the study area is 15%, while the high intensity development is only 8.9%. The dominance of low intensity development compared to high intensity development in this area means there is less high intensity development metrics to look at in the study which may influence the results of this study.

6.3. Control variables

There are four control variables that are on average significant within both the low intensity and high intensity development metrics models. The forested land cover is positively significant, meaning that the more forest in the watershed the higher the phosphorus levels. The reason forest cover and phosphorus levels have a positive relationship is potentially due to fragmentation. While many studies state that forests can filter out nutrients before they enter streams, there is research showing how forestry patch metrics play an important role in the positive relationship between forest and phosphorus levels. Lee et al. (2009) suggests that fragmentation of forest land cover might be detrimental to surrounding water quality, and sometimes stop the natural nutrient filtration system.

Precipitation and phosphorus levels have a negative relationship, which means as the precipitation increases the phosphorus levels decrease. The most likely cause of this relationship is that the extra precipitation dilutes the phosphorus levels within the stream. This area tends to get a lot of rain (mean annual precipitation = 51.47 inches or 131 centimeters), particularly in short but heavy downfalls. This large amount of rain can dilute the nutrient levels that are in the rivers.

Percent of low intensity development within a watershed has a negative relationship with phosphorus levels. This shows that the more low intensity development there is within a watershed the higher the phosphorus levels. The relationship could be driven by the use of fertilizers on urban lawns. This fertilizer would not be seen in the high intensity development parts of the watershed because high intensity development is 80-100% impervious surface cover. This explanation aligns with the relationship seen between high intensity development and phosphorus which is positive in this study. The greater the percent of high intensity development within a watershed the lower the phosphorus levels. In regards to home use fertilizers applied to urban lawns, they are not as frequently used in areas of high intensity development. There is a lot of literature supporting the correlation between fertilizers and increased phosphorus in streams (Qiu et al., 2014). According to Qiu et al. (2014), reducing phosphorus application rates lowers the total phosphorus loads in the streams significantly. In addition, the less phosphorus present in fertilizers, the less phosphorus ends up in rivers. When the phosphorus levels in fertilizers went from normal application levels to none at all, the phosphorous levels in the water phosphorus were reduced by 26% (Qiu et al., 2014). As stated in Paerl et al. (2014) in areas that are highly urbanized and have lot of agriculture, the phosphorus levels in the waters are high which is in part due to phosphorus-based fertilizer use. This problem is compounded by the large amounts of phosphorus-based fertilizer that are currently being added to the lawns of Texas. In 2011, the amount of phosphorus-based fertilizer (P_2O_5 which is 44% phosphorus) purchased in Texas was 144,209 (in 1000 kg of P_2O_5), up 6.014% up from 2009 (EPA, 2015). This level of phosphorus use makes Texas the 11th largest state consumer of P_2O_5 in the nation (EPA, 2015).

One of the biggest issues that arise with increased phosphorus loading in streams from fertilizer runoff is eutrophication (Carpenter et al., 1998). Nonpoint sources of pollution cause the largest input of phosphorus and nitrogen into streams and rivers in the

United States. While phosphorus does not have any adverse health effects, it does drastically increase the chances of eutrophication. This in turn can cause potentially harmful algal blooms (Carpenter et al., 1998; Correll, 1998; Varlamoff et al., 2001). This process from phosphorus loading to eutrophication is depicted in figure 19.

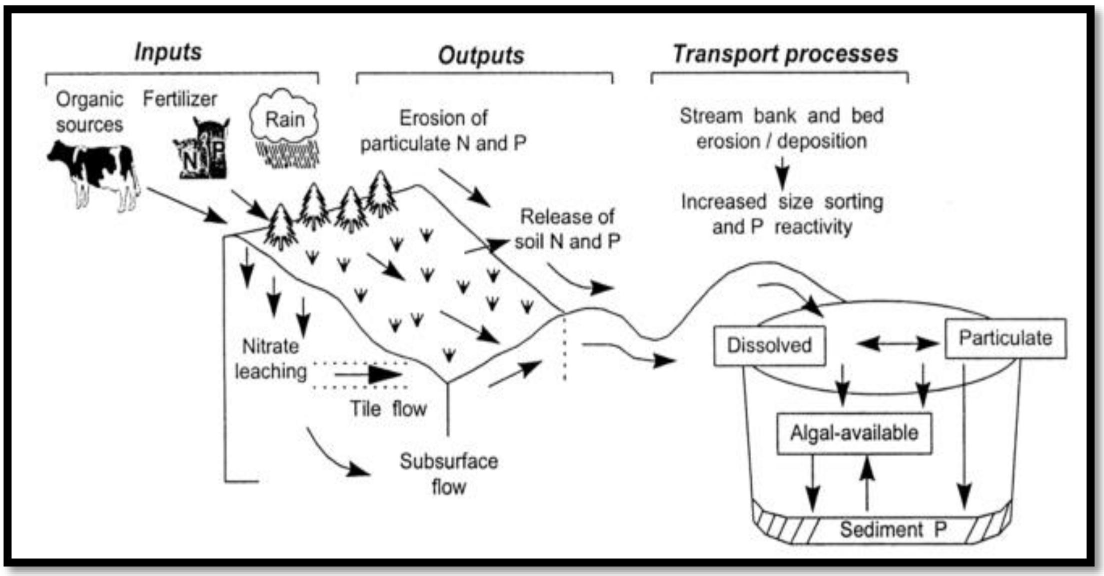


Figure 19: Process of nutrient transport from land. This figure is taken from Carpenter et al. (1998).

As the fertilizers and organic sources of pollution run off the land, they enter into nearby water sources. Some nutrients like nitrate are leached into the soil, while extra nitrogen and phosphorus enter the water through erosion. These nutrients, once in the water, become food for algae. As more nutrients enter the water, algal populations in the same water source rapidly increase. After they die, these algae are decomposed by bacteria which use oxygen during the decomposition process, thus leaving the water oxygen deficient.

Another aspect of using fertilizers on lawns has to do with the value homeowners place on having a green lawn. A recent study in Georgia showed that the majority of homeowners sampled considered it “somewhat important” to “very important” to have their lawn look as good as their neighbors and to have a green yard throughout the entire year (Varlamoff et al., 2001). To maintain a green yard throughout the year requires use of fertilizers and water, which could be a potential source of the phosphorus levels that enter the streams and rivers from urban lawns. Additionally, education geared towards homeowners about how much fertilizer their lawns actually require is sparse. Lack of awareness leads to homeowners over fertilizing their lawn and creating even higher phosphorus levels in their surrounding rivers and streams (Varlamoff et al., 2001). This is a fundamental problem, because in order to decrease the amount of phosphorus runoff coming from lawns there needs to be a decrease in the amount of phosphorus-based fertilizers used in lawns.

7. POLICY IMPLICATIONS

There are multiple policy recommendations that could be derived from this study for the Galveston Bay Estuary region. The two main categories are:

1. Create a more aggregated landscape of low intensity development patches
2. Decrease the amount of phosphorus based fertilizers applied to urban lawns.

Each of the two main categories will be addressed in the following paragraphs. In addition, the specific planning tools and policy implications that could be utilized to assist in reaching the desired outcome will be discussed. Desired locations where these planning tools and policies could be implemented will be discussed as well.

7.1. Create a more aggregated landscape of low intensity development patches

The following are low intensity development class metrics ideal for minimizing total phosphorus levels in streams:

1. Lower number of patches per area within watershed (lower patch density)
2. High percent of like adjacencies (increase the aggregation of patches)
3. High contiguity index
4. Larger average patch area

There are many ways that these different class metrics could be manipulated to reduce phosphorous levels in rivers and streams. One of the main drivers of the relationship between phosphorus and low intensity development class metrics (to generalize) is the level of fragmentation of the landscape. For instance, having a landscape that has all of the patches aggregated and very connected (high percent of like adjacencies, high average contiguity index) will help to minimize the phosphorus levels in the stream. This means that instead of having many different low intensity development subdivisions, being able to aggregate these subdivisions into a single continuous unit of housing will help to minimize the amount of phosphorus that enters nearby streams.

There are multiple planning tools that could assist in creating this less-fragmented low intensity development landscape. Four of these are development clustering, taxation, creating and maintaining urban growth boundaries, and utilizing transfer of development rights.

7.1.1. Development clustering

As mentioned previously, having more clustered low intensity development patches helps reduce the total phosphorus levels in the streams. For this reason, clustering of development will leave large plots of land open and untouched while concentrating development in specific parts of the watershed. Having development clustered will allow development to occur in larger patches but only on one portion of the site. This way the patch area of development patches is larger with high contiguity and high patch density. The other portions of the site where there is no development will be less-altered from future development and therefore the negative impacts of low intensity development that are seen on the total phosphorus levels (which are an indicator of water quality) will be minimized.

One off-shoot of clustering is planned unit developments. Planned unit developments simply means that each new development must have a specific plan that has to be approved by the administration (Beatley, 2012) If the administration takes into account the knowledge that sprawling intensity development is a problem to the phosphorus levels they will apply this knowledge when approving planned unit developments. Developments in close proximity to each other to reduce fragmentation and increase average patch size of development are preferable.

Clustering of development is fairly easy to implement because the transactions take place all on one parcel. Development clustering has been used effectively in other portions of the county and would most likely be transferable to the Galveston Bay Estuary. While in other states there has been different motivations behind clustering of development, the result has been the same and can be beneficial for minimizing total phosphorus levels as well. New York is one great case study (Brabec, 2001):

Southampton, New York has a long history of agricultural production. When the railroad came into the town there became an increase in tourism and therefore an increase in development for wealthy residents to travel in and out of the town via train. A comprehensive plan was written in 1970 to only develop specific places that would implement cluster development. This development clustering would leave 80% of the farmland untouched and the development would be clustered *"along a greenbelt park which would preserve the watershed's ponds, streams, and areas of high water table"*

(Brabec, 2001). When the subdivisions for clustering where approved, the amount of farmland protected jumped from 39.3% to 59.9% (Brabec, 2001).

7.1.2. Taxation incentives

For land that has not been developed, taxation can be used to incentivize developers to not develop and instead develop around the already developed plots of land (Beatley et al., 1994). Taxation has been used as a planning policy for other issues such as hazard mitigation – such as tax incentives for minimizing certain land uses in areas that are very hazard prone. This same theory can be used to aggregate the development and leave the rest of the undeveloped watershed unaltered by development. If the cost to develop on undeveloped and open land is more expensive then around lands that already have low intensity development, then developers will be more likely to develop around the existing development. One way to use tax incentives to encourage development in specific locations is by giving tax breaks to people who develop in certain areas (Beatley, 2012).

7.1.3. Urban growth boundaries

Another option is to have urban growth boundaries. If urban growth boundaries were imposed in the Galveston Bay Estuary then there would be a concrete way to make sure low intensity development did not sprawl past these boundaries and were therefore kept in a contained area. Urban growth boundaries have been successful in states such as

Oregon where the whole state requires that cities must have urban growth boundaries (Beatley et al., 1994).

A good example of an urban growth boundary is in the 4 counties that make up the Vancouver, Washington and Portland, Oregon cities in the pacific northwest of the United States. The main purpose of the urban growth boundaries in these cities are to contain urban sprawl and preserve forests and farms (Chang et al., 2014; Kline et al., 2014). Due to a bill passed in Oregon in 1973, by 1981 the land use plan in Portland, OR required that new development be contained within the urban growth boundary of Portland (Kline et al., 2014). The urban growth boundary in Vancouver, WA came years later (1991) but was created to have the same effect on the city as the urban growth boundary that was already established in Oregon (Kline et al., 2014). Kline et al. (2014) shows that growth outside of the Vancouver urban growth boundary is greater than outside of the Portland urban growth boundary because of the time at which each urban growth boundary was implemented. In fact, between 1974 and 2005 in the three Oregon counties that make up parts of the urban growth boundary and are studied in Kline et al. (2014) there was an 18.2% increase in urban development inside the urban growth boundary and a 0% increase outside of the urban growth boundary. However, when referring to the low density development there was a 4.8% increase within the urban growth boundary and a 4.2% increase outside of the urban growth boundary. The low density development is defined differently in Kline et al. (2014) but also refers to a more rural landscape. Another benefit of the Portland urban growth boundary was defined in a way that no headwaters are included within the urban growth boundary (Chang et al., 2014). Overall, the successfulness of the urban growth boundary in the Portland/Vancouver region being able to contain development within the urban growth boundary is relatively successful this type of planning tool could potentially be beneficial in the Galveston Bay Estuary.

Urban Growth boundaries are effective in limiting the sprawl of urban development (Perlman and Miller, 2005). With the increased development and population growth that has already been seen in the Houston/Galveston region, having an urban growth boundary will encourage higher and denser developments to occur. Because the low intensity development is what is driving the increased phosphorus levels in the streams, having higher and more concentrated development (all within the urban growth boundary) this would be an effective way to limit the amount of low intensity development in the region. These urban growth boundaries can be monitored over time and assessed periodically to see if they need to be expanded to accommodate the growing population (Perlman and Miller, 2005).

One example of where urban growth boundaries would be beneficial is in a watershed is located in the northeast portion of the watershed (watershed number 18697) (figure 20).

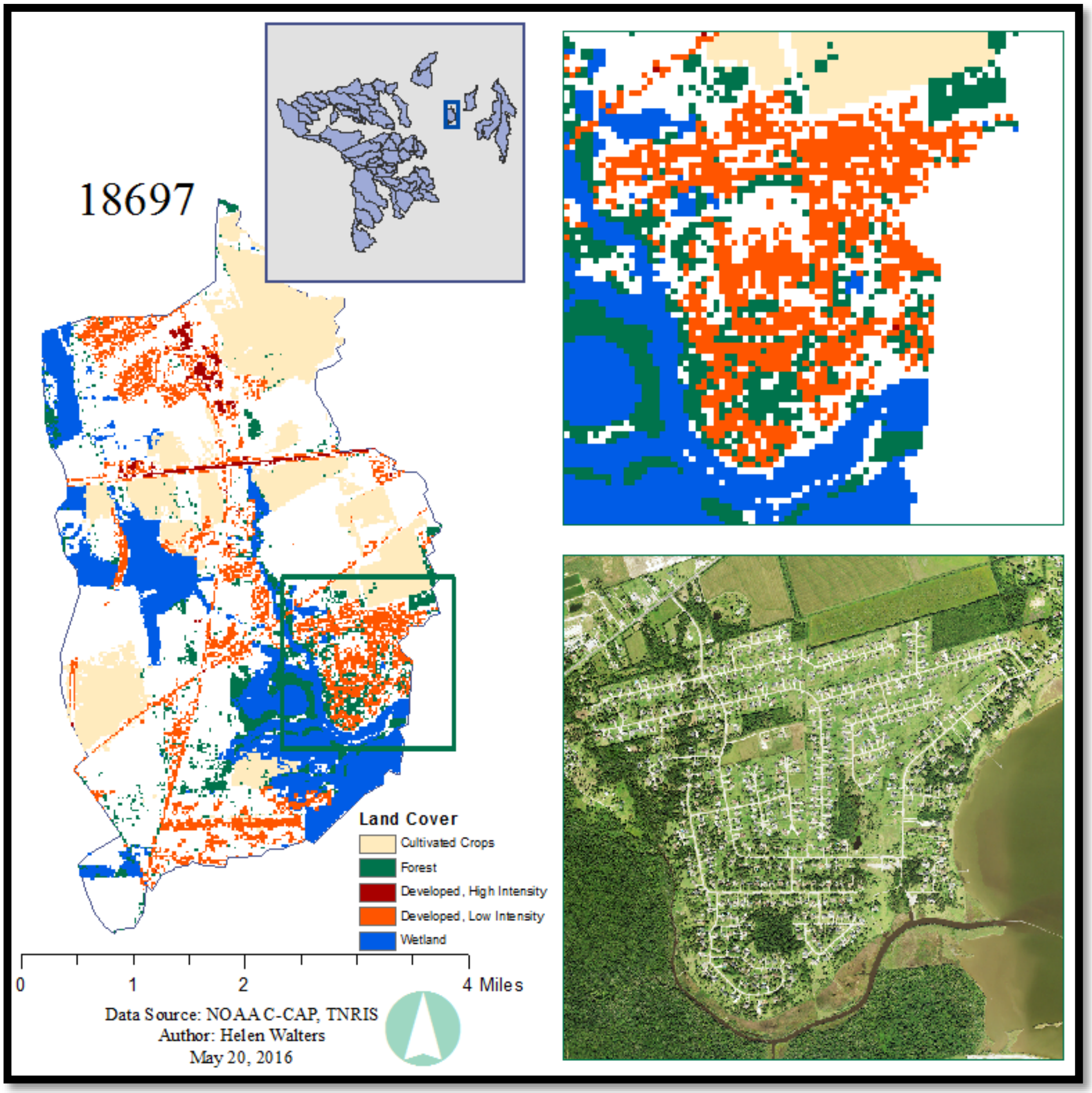


Figure 20: Watershed number 18697 located in the northeast portion of study area. The subset map in grey shows where in the study area this watershed is located, the watershed shows the land cover and the zoomed in area shows land cover in the top subset and the aerial imagery in the lower subset.

There are a couple different models that are looked at for this watershed: low intensity patch size and low intensity development patch density. The formula for the low intensity development average patch size patch area model is shown in equation 1 in Appendix G. Table 8 (also included in Appendix G) shows how, when all other variables are held constant, the total phosphorus levels change when low intensity development patch area is increased in the watershed. Because these are based on rates of change, the change rates for this watershed will stay constant for all of the other watersheds. When the average patch size is increased by 0.25 hectares, which is 2500 square meters (and therefore almost 3 pixels) the phosphorus will decrease by 0.239 mg/l.

The difference between watersheds is seen in the percentage of change. In a watershed where the total phosphorus levels are already higher, the percentage change will be much lower. For instance in the watershed shown in figure 20 (number 18697) with an increase in 0.25 hectares, for the average patch size, there will be a decrease by 12% in the phosphorus levels within the watershed (on averaged, based on this study and holding all other variables constant.

The watershed in figure 20 (#18697) is a good example of where urban growth boundaries could be used appropriately. In the case of this watershed, the zoomed in area is the majority of low intensity development. Around this watershed there is a lot of wetland area with some forested areas mixed in. In the case of this watershed it would be beneficial to be able to set the urban growth boundary in a manner that excludes the wetlands so that the wetlands and forests remain undeveloped. If the wetlands and forests are left as they are (by using the urban growth boundary) then the low intensity development will become more dense and concentrated but more of the natural ecosystem will not be destroyed.

One of the other models in this study looks at patch density of low intensity development patches. The equation for low intensity development patch density is shown in Equation 2 (Appendix G). When all variables are held constant and the low intensity development patch density is decreased by 1 unit the phosphorus will decrease by 0.066 mg/l (which in the case of this watershed comes out to about a 4% change). In the case of this watershed as well, having an urban growth boundary will the patch density of the low intensity development patches. Because there will be no way to develop the low intensity development outside of the urban growth boundary, the low intensity development patches will merge which will decrease the number of patches within as shown in this study, having decreased patch area is beneficial to the phosphorus levels in the water.

In this watershed if the percent of like adjacencies increases (from its current value of 56%) then the phosphorus levels will decrease. Increasing the percent of like adjacencies by 1% in this specific watershed will decrease the phosphorus values by 2%. This increase in percent of like adjacencies can help in reducing the fragmentation of the low intensity development within the watershed.

7.1.4. Transfer of development rights

Another option to encourage low intensity development in certain areas and not others in order to maximize the amount of aggregation of low intensity development landscape patches is transfer of development rights. Transfer of development rights essentially has one plot of land that receives the development rights (the area that is going to be developed on) and one plot that gets rid of development rights (the sending area). The plot that is receiving development rights is the plot that gets developed while the sending area has no development because there are no development rights there. The land that has had development rights transferred cannot be developed on (Perlman and Miller, 2005). As Perlman and Miller (2005) states, the way to make transfer of development rights work is to incentivize the landowners in the area sending the development rights. This means that the most valuable land in the area that is sending development rights will have the most number of development credits (Perlman and Miller, 2005).

One watershed where transfer of development rights could be utilized is a watershed in the north-central part of the study (#15864). As shown in figure 21, this watershed contains a lot of development and only a few patches of the natural ecosystem left. It would be beneficial to have the wetland areas transfer their development rights to all around the existing development, which would leave the majority of the areas that are currently not developed as is. As shown in the cutout above, there is a lot of low intensity development intermixed with forest land. One way to increase the average patch size and therefore decrease fragmentation would be to aggregate all of the development into this region instead would be to develop all around the existing development and stop development in the rest of the watershed.

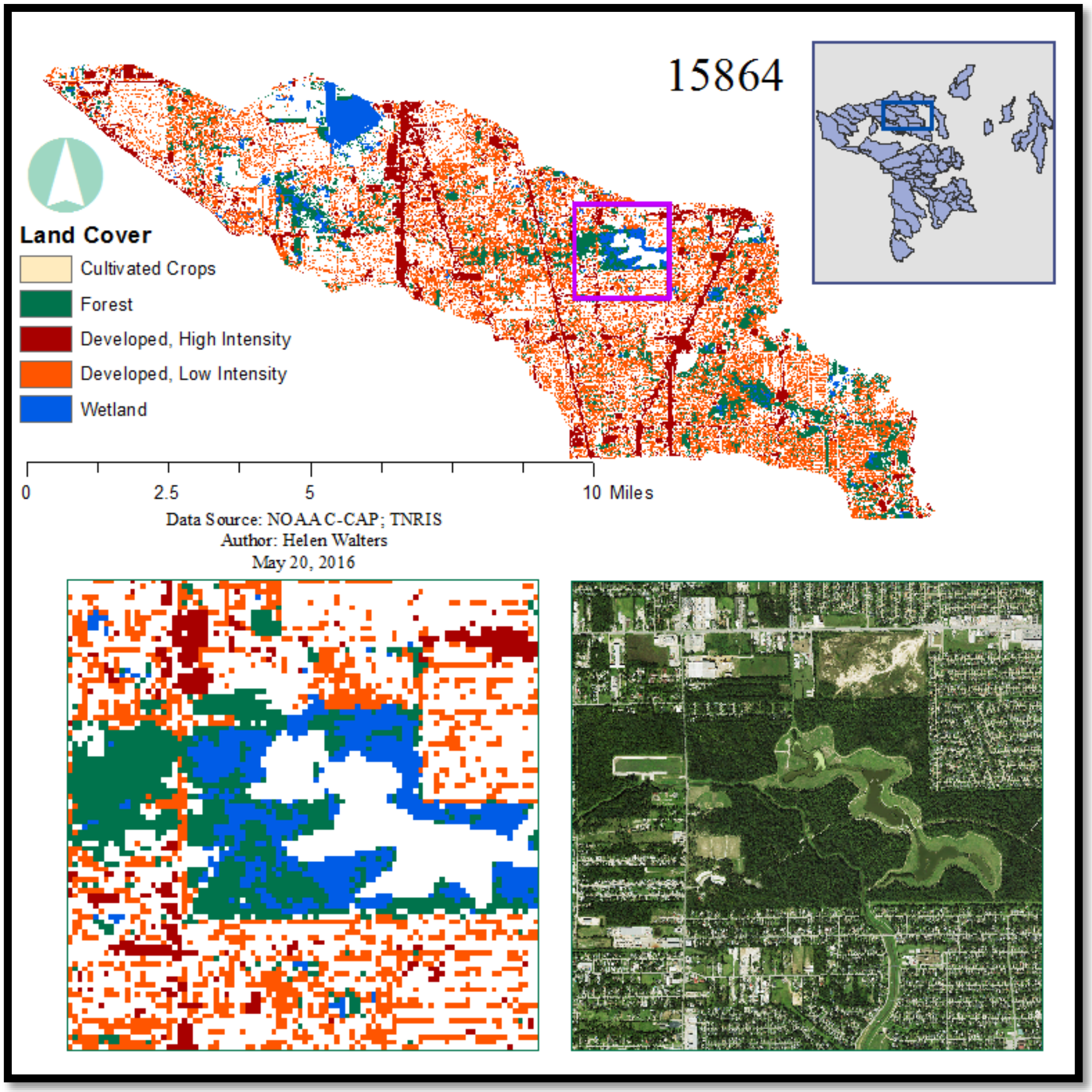


Figure 21: Watershed number 15864 in the north central part of the study. Watershed number 15864 in the north central part of the study. The subset in grey shows where the watershed is located within the study area. The watershed is shown with land cover types, the left subset shows the land cover of the specific area to implement the transfer of development rights, and the right subset is the aerial imagery in the location where the transfer of development rights will be implemented.

In the case of this watershed (15864) when the equation in Appendix G is used to calculate the phosphorus level in relation to average patch size it is shown that for a 0.25

hectares increase in average low intensity development patch area there is a decrease by 0.239 mg/l of phosphorus.

In terms of patch density, the model shows that increasing the average patch density increases by 1 patch per km (or 1 patch per 100 hectares) will increase the phosphorus levels by 0.0656. Table 9 (Appendix G) uses the patch density equation to calculate the effect of patch density on the phosphorus levels. In terms of percent change for this watershed, when the patch density is decreased by 1 patch per watershed (on average, holding all other terms in the model constant) the phosphorus is decreased by 19%. When the percent of like adjacencies in this model is increase by 1% (the normal is 52.5%) the decrease of phosphorus is 6% (Calculations and equations provided in Appendix G and H).

7.2. Decrease the amount of phosphorus based fertilizers applied to urban lawns

The second part of policy recommendations has to do with decreasing the amount of phosphorus-based fertilizers that enter the streams of the GBE. There are multiple avenues that could be explored in order to accomplish this objective. Simply decreasing the amount of phosphorus that is in phosphorus-based fertilizers available at stores would decrease the amount of phosphorus entering the streams. There could also be regulations that are placed on the amount of fertilizer that can be purchased and applied. This would ensure that there are less phosphorus-based fertilizers being applied to the lawns in the watershed.

7.2.1. Education and outreach

One key aspect of getting home owners to reduce the amount of fertilizers that are put on their lawns is education about the negative impacts of phosphorus on water quality. The public has to be aware in order to be able to act and change the way that they treat their lawns. As previously mentioned, many homeowners feel that it is important to have their lawns be green compared to their neighbors (Varlamoff et al., 2001). Holding workshops where the affects of lawn fertilizers on the surrounding ecosystems are discussed would be a good way help the public understand the negative repercussions that come out of having a green lawn. As Salm (2000) states, "awareness plays a major role in public support and in the general success of conservation."

Another type of education is in the form of posting signs at the stores that sell phosphorus-based fertilizers. In some states it is required that the phosphorus-based fertilizers be displayed in their own display and have educational signs that discuss the potential negative impacts of use of this fertilizer (Vt. Stat. Ann. tit. 10, § 1266b (2015)). The Maine Legislature also has a law that prohibits the sale of phosphorus-based fertilizers unless there is a sign states that it is not fit for use on "nonagricultural lawns or turf" because of the adverse effects it can have on water quality (Mn. Stat. Ann. tit. 38, § 419 (2007)).

7.2.2. Implement laws that limit or prohibit phosphorus based fertilizers

Limiting the amount of phosphorus based fertilizer that can be applied to lawns is a good way to reduce the runoff from low intensity development as well. In Vermont (Vt. Stat. Ann. tit. 10, § 1266b (2015)) no one is allowed to apply phosophorus-based fertilizers unless *"phosphorus fertilizer necessary for application to turf that is deficient in phosphorus as shown by a soil test performed no more than 18 months before the application of the fertilizer; or phosphorus fertilizer that is labeled as starter fertilizer and that is intended for application to turf when a property owner or an agent of a property owner is first establishing grass in turf via seed or sod procedures and the application of starter fertilizer is limited to the first growing season".* This is one way to make sure that the use of phosphorus fertilizers are minimized.

The Illinois General Assembly has a law that limits all phosphorus-based fertilizer application unless the soil is deemed deficient in phosphorus from a soil test (415 Ill. Comp. Stat. § 65/1-8). This means that in lawns where there is no phosphorus deficiency and the homeowner simply wants to see greener grass, no application allowed.

As of 2012 (Miller, 2012) there are 12 states in the United States that have some bans or restrictions on the use of phosphorus-based fertilizers. It is apparent that these states have evidence to support the bans and the negative impacts that there can be on water quality. Implementing some of these bans and being able to reduce the amount of phosphorus-based fertilizers used in Texas and specifically on low intensity development areas in the Galveston Bay Estuary would be beneficial to the surrounding streams and rivers.

Lehman et al. (2011) conducted a study that (over a three year time period) had an ordinance to reduce the amount of fertilizer that was put on lawns. After this three year period of reducing fertilizers, water samples were analyzed and it was shown that phosphorus levels dropped 11-35% (Lehman et al., 2011). As a result of this, the study area (Ann Arbor, Michigan) no longer sell phosphorus fertilizers in stores. It would be beneficial to test out a study similar to Lehman et al. (2011) in the Galveston Bay Estuary because the results of this study show that a decrease in phosphorus based fertilizer might help decrease the phosphorus levels in the streams.

There is a need for decreased use of phosphorus-based fertilizer in the Galveston Bay Estuary. A study done by Gronberg and Spahr (2012) for the United States Geologic Survey calculated the levels of farm and non-farm phosphorus-based fertilizer use for the nation from 1987-2006. While this time frame is before this study takes place, the results from Gronberg and Spahr (2012) show an increase in the use of non-farm phosphorus fertilizer within the counties that encompass the Galveston Bay Estuary (figure 22). In fact, every county that is within the Galveston Bay Estuary has an increase in the phosphorus-based fertilizers that are applied to non-farm areas. Harris county has the highest on average phosphorus levels during 1987 and 2006 (as shown by the map in figure 22). The calculations and data analysis processes that went into this data is from the USGS and can be found in more detail in (Gronberg and Spahr, 2012). In addition, it was noted in the report that *"states with a higher nonfarm-to-total* *fertilizer ratios for nitrogen and phosphorus tended to have higher urban land-use percentages" (Gronberg and Spahr, 2012).* This reinforces the relationship between phosphorus fertilizer and urban development within the Galveston Bay Estuary.

The potential policy implication of implementing lawns to reduce or remove phosphorus based fertilizer does not discuss nitrogen fertilizers or the affects that nitrogen can have on water quality. This is an important consideration for future research as nitrogen is an important nutrient to understand in relation to nutrient loading and eutrophication. This relationship is something that should be studied in future research due to its importance. The following section discusses validity threats as well as research limitations followed by future research and conclusions.

Non-farm phosphorus-based fertilizer in the counties of the Galveston Bay Estuary 1987-2006

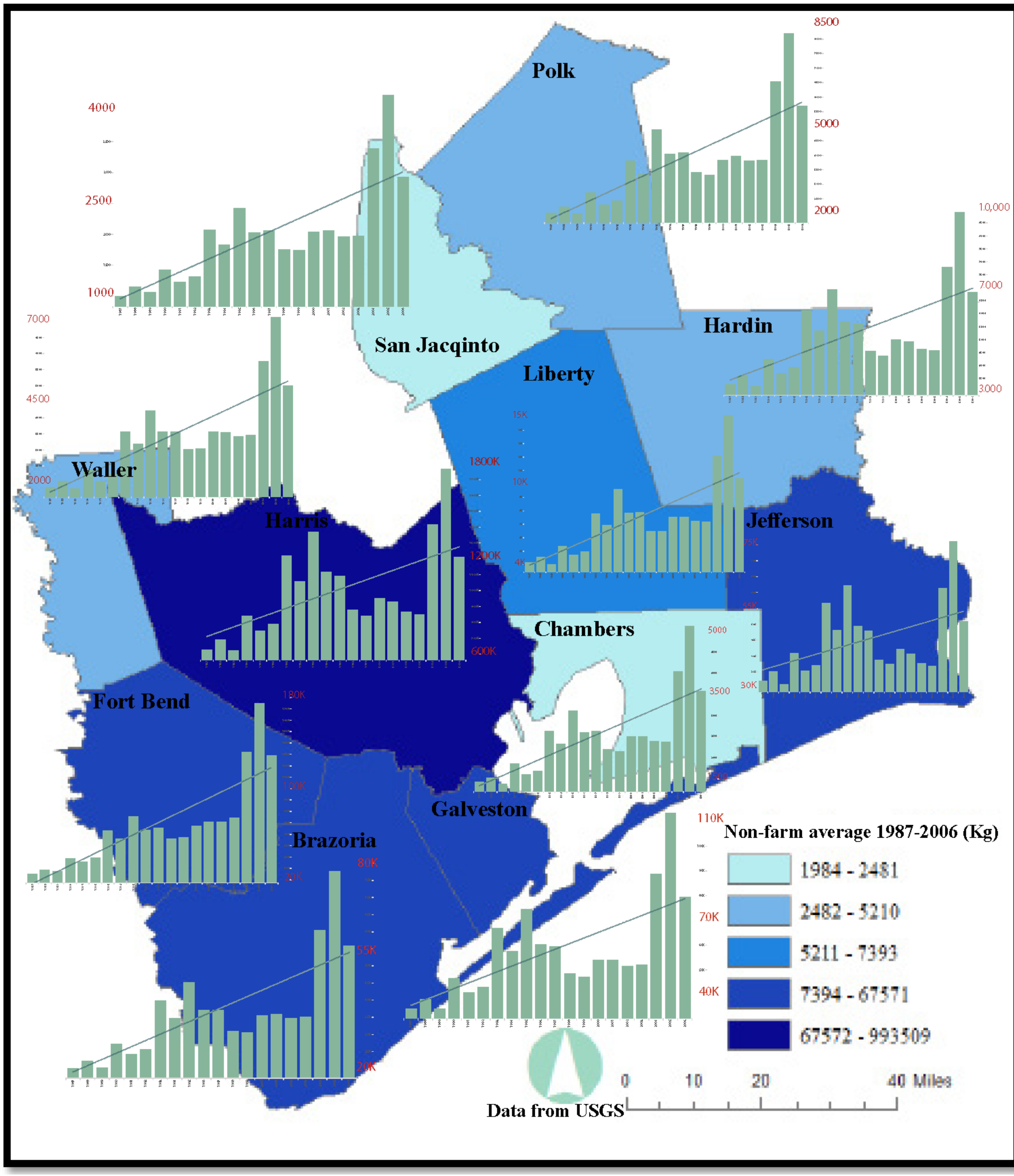


Figure 22: Phosphorus-based fertilizer across the eleven counties of the Galveston Bay Estuary from 1987-2006. The data was obtained from the United States Geologic Survey and based on input from the American Plant Flood Control Officials fertilizer sales data, Census of Agriculture fertilizer expenditures, and the United States Census Bureau. The histograms show the trend of non-farm phosphorus based fertilizer use across the specified counties during 1987-2006. The map and legend show the average phosphorus-based fertilizer use (measured in kilograms). A rigorous data analysis and processing process was conducted by the USGS to obtain the results reported and can be further assessed in Gronberg and Spahr (2012).

8. VALIDITY THREAT/RESEARCH LIMITATIONS

There are multiple aspects of this study that should be noted for potential validity threats. First, the land cover data from NOAA-CCAP is stated as 2010, but based on how the land cover data is created, there are multiple years that are combined to make complete picture. This is because the data source that the land cover is obtained from creates land cover for the entire United States coasts, which is not possible to create in one year. This means that there might be up to three years of data in the land cover imagery that is for 2010. When the satellite imagery is compiled into one land cover raster, multiple satellite images are used to compute these values which can potentially cause some validity threat issues.

Another validity threat to this study is the TCEQ water quality data. The years that were utilized in this study: 2010-2013 aren't always measured at the same time and number of times per year. There are years where certain gauges are not measured and therefore not represented in the data record. This is why this study utilizes an average of four years of phosphorus data. This four-year average will remove some of the variability and validity issues that arise from this data. However, not all of the error can be phased out in this averaging process. The stations used for the study have slightly differing levels of sample collection during the four years. For certain observations there are only 80 observations, but for others there are over 100.

A third validity threat to this study is statistical bias. As previously mentioned, there is some slightly higher than acceptable variance inflation (VIF) in the low intensity development patch density model. Because the variable of interest is patch density, there is no way to remove the variable from the model and therefore it has to be noted that there is this slight violation. In addition, in the high intensity development models, average patch size removes one of the control variables because of high collinearity, which might be problematic because that variable is not accounted for. There may also be lingering heteroscedasticity in the models. Spatial lag models tend to remove heteroscedasticity, but in some cases all of it cannot be removed.

Another aspect of potential validity threat comes from the septic systems data that is utilized from the Houston-Galveston Area Council (H-GAC). The data is developed by H-GAC by geo-referencing permits with septic systems and then matching the addresses to the centroid of each parcel. When there was a latitude and longitude coordinate these were left as a location for each septic tank. The database contains permitted OSSF permits and any septic systems before 1989 are grandfathered in and do not need a permit (however, there are some that are listed in the database, which is why the database dates back to 1984). Due to the way this data was collected and analyzed, there could be some issue with the validity and accuracy of the data.

Another aspect of validity threats within this study is the inclusion of all control variables. Excluding controls that are essential to the ecosystem will cause internal validity within the quasi-experiment. There are a few variables that could be included in the future that might help explain the relationship between phosphorus levels and development patterns within the Galveston Bay Estuary. These variables include golf courses (nitrogen and phosphorus based fertilizers are used frequently on golf courses (Baris et al., 2010)) and toxic release sites (there may be sites that release nutrients such as phosphorus and ammonia). There may be other control variables that could be used in the future if the data was available.

9. CONCLUSIONS AND FUTURE RESEARCH

This study is unique in that it uses ecological landscape metrics to examine low intensity and high intensity development patterns. The development is split up into these two categories based on percent of impervious surfaces (Brody et al., 2013; Brody et al., 2011). There are only a few other studies that take the traditional ecologically based landscape metrics and apply them to anthropogenic development patterns (Alberti et al., 2007; Carle et al., 2005; Sun et al., 2013). This study shows that there is a relationship between the patterns of low intensity development and total phosphorus levels within the Galveston Bay Estuary. This relationship is driven by four low intensity development metrics: patch density ($P<0.01$), percent of like adjacencies ($P<0.05$), average patch size ($P<0.05$), and contiguity ($P=0.06$). The only significant high intensity development metric is average patch area ($P<0.05$), which could be due to the dominance of low intensity development within our study area.

The implications of these results can be showcased through multiple planning tools and policy implications including development clustering, tax incentives, urban growth boundaries, transfer of development rights, education and outreach, and laws to limit or ban phosphorus-based fertilizers. Each of the planning tools and polices can encourage low intensity development patterns that maximize connectedness and minimize fragmentation. In addition, reducing phosphorus-based fertilizers used in the Galveston Bay Estuary will reduce the amount of phosphorus in the rivers of the GBE.

This study provides a good starting point for understanding the complex relationship between phosphorus levels and low intensity development metrics in the Galveston Bay Estuary. However, additional research still needs to be done to fully understand this relationship. There are some aspects of this study that drive future research. First, it would be beneficial to have a larger sample size that incorporates more of the high intensity development within the region including areas such as downtown Houston. This might provide a better picture of the relationship between phosphorus and low intensity and high intensity development metrics observed in this study.

In addition to examining a larger sample size and area, examining other water quality variables would be beneficial to understanding the effect of development metrics on water quality. While total phosphorus is a good indicator of water quality, there are many other indicators that could help piece together a better understanding of the water quality in the area. It would be helpful to include other nutrients such as nitrogen, as well as total suspended solids and biological indicators that are representative of stream health. Another avenue of future research that could be explored is using a water quality model such as the Agricultural Nonpoint Source Pollution Model (AGNPS), which analyzes the relationships between land cover and the nutrient component of runoff (Corbett et al., 1997; He et al., 2000). This water quality model includes control variables like precipitation and elevation in addition to land cover and water quality variables.

The part of this study that involves the effects of phosphorus-based fertilizers (from low intensity development) on the phosphorus levels in the rivers is something that could be investigated further. There is a lot of previous research that shows the problems behind using phosphorus-based fertilizers, although none to our knowledge that has been conducted in the Galveston Bay Estuary. Therefore, this piece of work reinforces the notion that the percentage of low intensity development in a watershed increases the phosphorus levels in the streams and these region-specific results could be used to influence policy. Studying further these effects and other cases where studies that have been conducted result in changing the local or state-wide ordinances could help make these results applicable to the region and influence change.

An alteration to the study that would be beneficial to future research in this field is selecting a sample that is not spatially nested. The nesting issue with this dataset means that, while a spatial lag model removes a lot of the spatial autocorrelation and multicollinearity within the model, there is some spatial autocorrelation and multicollinearity that might not be removed. High multicollinearity can increase the standard errors that result from the models and may cause a lack of significance (Chang et al., 2014). If the samples are spatially independent, this would make the process of removing spatial autocorrelation much simpler and create results completely devoid of this type of correlation. One way to have a sample that was not nested would be to use the NHD Hydrologic Unit Code 12 (HUC12) watersheds and then take a weighted average of the water quality monitoring gauge points within the pre-designated watersheds based on the flow direction and flow accumulation for each monitoring point within the HUC12s. This would be a way to have a non-nested sample but also take into account the water quality that is upstream of the water quality monitoring point.

Another variable that could be potentially interesting to include in this analysis as a control is the number and/or location of golf courses in the study area. Golf courses are a very large consumer of phosphorus-based fertilizers and it has been shown to be a water quality problem (Baris et al., 2010). According to one study that looked at golf courses across the United States, there was a high level of phosphorus exceedence (based on the EPA recommended phosphorus levels by Ecoregion) (Baris et al., 2010).

Another important aspect of future research is to look at other water quality variables. Many other water quality variables have been examined in the literature previously and studying some of these other variables in this study would be important to understand the entire health of the ecosystem. The water quality part of the literature review stated other water quality variables that have been analyzed and these are all ones for future research: total suspended solids (Ahearn et al., 2005; Coulter et al., 2004), fecal coliform (Nagy et al., 2012; Tong and Chen, 2002), total nitrogen and total phosphorus (Chang, 2008; Halstead et al., 2014; Tong and Chen, 2002), pH (Chang, 2008; Coulter et al., 2004; Tong and Chen, 2002; Zampella et al., 2007), taxa richness (Booth et al., 2004), temperature (Chang, 2008; Coulter et al., 2004), and sodium, cadmium, lead, and conductivity (Tong and Chen, 2002). Nitrogen (including ammonia, nitrate, and nitrite) are all of particular interest because of the importance in nutrient loading and eutrophication.

One expansion area to this study is relating the results of this study based on the effect that low intensity and high intensity development metrics have on phosphorous levels. Water pollution is a chronic hazard that can have long term effects (of which phosphorus loading is one portion). This study could be expanded to look at multiple hazards and what low intensity and high intensity development patterns minimize these hazards. For instance, flooding is dependent on landscape metrics (Brody et al., 2013; Brody et al., 2011). Brody et al. (2011) looks at both low intensity development and high intensity development using the same land cover data as this study (NOAA C-CAP). Having more clustered development results in lower amounts of flood damage and low intensity development (because it is by nature more spread out) is more flood prone (Brody et al., 2011). Another study (Brody et al., 2013) shows that the more aggregated development is, the more it reduces flood losses. Some of the development metrics used in this study include patch number, total area of development, and patch density. All of these metrics could be used across both flooding damages and water quality variables to see if there are development patterns that benefit both. It would be interesting to take this knowledge about flooding into account and because these study show that a more aggregated landscape is beneficial to phosphorus and flood damage to combine both and see what development landscape maximizes both would be interesting. This may mean that by minimizing total phosphorus levels in the watershed (based on development metrics) would also minimize flooding in the area. Coupling different topics like this could really help create a landscape that is maximized to benefit multiple aspects of the system.

REFERENCES

415 Ill. Comp. Stat., USA, Illinois.

2007. Maine Statutes: Cleaning Agents and lawn turf fertilizer containing phosphate banned. http://www.mainelegislature.org/legis/statutes/38/title38sec419.html.

2015. Vermont Statutes Annotated. http://legislature.vermont.gov/statutes/section/10/047/01266b.

Aguilera, F., Valenzuela, L.M., Botequilha-Leitão, A., 2011. Landscape metrics in the analysis of urban land use patterns: A case study in a Spanish metropolitan area. Landscape and Urban Planning 99, 226-238.

Ahearn, D.S., Sheibley, R.W., Dahlgren, R.A., Anderson, M., Johnson, J., Tate, K.W., 2005. Land use and land cover influence on water quality in the last free-flowing river draining the western Sierra Nevada, California. Journal of Hydrology 313, 234-247.

Alberti, M., 2005. The effects of urban patterns on ecosystem function. International regional science review 28, 168-192.

Alberti, M., Booth, D., Hill, K., Coburn, B., Avolio, C., Coe, S., Spirandelli, D., 2007. The impact of urban patterns on aquatic ecosystems: An empirical analysis in Puget lowland sub-basins. Landscape and Urban Planning 80, 345-361.

Alberti, M., Marzluff, J.M., Shulenberger, E., Bradley, G., Ryan, C., Zumbrunnen, C., 2003. Integrating humans into ecology: opportunities and challenges for studying urban ecosystems. BioScience 53, 1169-1179.

Allan, J.D., 2004. Landscapes and riverscapes: The influence of land use on stream ecosystems. Annual Review of Ecology Evolution and Systematics 35, 257-284.

Baris, R.D., Cohen, S.Z., Barnes, N.L., Lam, J., Ma, Q., 2010. Quantitative analysis of over 20 years of golf course monitoring studies. Environmental toxicology and chemistry 29, 1224-1236.

Beatley, T., 2012. Planning for coastal resilience: best practices for calamitous times. Island Press.

Beatley, T., Brower, D., Schwab, A., 1994. An Introduction to Coastal Zone Management. Island Press, Washington, D.C.

Bedan, E.S., Clausen, J.C., 2009. Stormwater Runoff Quality and Quantity From Traditional and Low Impact Development Watersheds(1). Journal of the American Water Resources Association 45, 998-1008.

Bernhardt, E.S., Band, L.E., Walsh, C.J., Berke, P.E., 2008. Understanding, managing, and minimizing urban impacts on surface water nitrogen loading. Year in Ecology and Conservation Biology 2008 1134, 61-96.

Booth, D.B., Karr, J.R., Schauman, S., Konrad, C.P., Morley, S.A., Larson, M.G., Burges, S.J., 2004. Reviving urban streams: Land use, hydrology, biology, and human behavior. Journal of the American Water Resources Association 40, 1351-1364.

Brabec, E., 2001. An evaluation of the effectiveness of cluster development in the Town of Southampton, New York. Urban Ecosystems 5, 27-47.

Brody, S., Kim, H., Gunn, J., 2013. Examining the Impacts of Development Patterns on Flooding on the Gulf of Mexico Coast. Urban Studies 50, 789-806.

Brody, S.D., Gunn, J., Peacock, W., Highfield, W.E., 2011. Examining the Influence of Development Patterns on Flood Damages along the Gulf of Mexico. Journal of Planning Education and Research 31, 438-448.

Bruno, D., Belmar, O., Sanchez-Fernandez, D., Guareschi, S., Millan, A., Velasco, J., 2014. Responses of Mediterranean aquatic and riparian communities to human pressures at different spatial scales. Ecological Indicators 45, 456-464.

Carle, M.V., Halpin, P.N., Stow, C.A., 2005. Patterns of watershed urbanization and impacts on water quality. Journal of the American Water Resources Association 41, 693-708.

Carpenter, S.R., Caraco, N.F., Correll, D.L., Howarth, R.W., Sharpley, A.N., Smith, V.H., 1998. Nonpoint pollution of surface waters with phosphorus and nitrogen. Ecological Applications 8, 559-568.

Chan, F., Barth, J., Lubchenco, J., Kirincich, A., Weeks, H., Peterson, W.T., Menge, B., 2008. Emergence of anoxia in the California Current large marine ecosystem. Science 319, 920-920.

Chang, H., 2008. Spatial analysis of water quality trends in the Han River basin, South Korea. Water Research 42, 3285-3304.

Chang, H., Thiers, P., Netusil, N.R., Yeakley, J.A., Rollwagen-Bollens, G., Bollens, S.M., Singh, S., 2014. Relationships between environmental governance and water quality in a growing metropolitan area of the Pacific Northwest, USA. Hydrology and Earth System Sciences 18, 1383-1395.

Clement, C., Bricker, S., Pirhalla, D., 2001. Eutrophic Conditions in Estuarine Waters. Silver Spring, MD: National Oceanic and Atmospheric Administration.

Corbett, C.W., Wahl, M., Porter, D.E., Edwards, D., Moise, C., 1997. Nonpoint source runoff modeling A comparison of a forested watershed and an urban watershed on the South Carolina coast. Journal of Experimental Marine Biology and Ecology 213, 133-149.

Correll, D.L., 1998. The role of phosphorus in the eutrophication of receiving waters: A review. Journal of Environmental Quality 27, 261-266.

Coulter, C.B., Kolka, R.K., Thompson, J.A., 2004. Water quality in agricultural, urban, and mixed land use watersheds. Journal of the American Water Resources Association 40, 1593-1601.

Council, H.-G.A., 2014. 2014 Basin Highlights Report.

Dauer, D.M., Ranasinghe, J.A., Weisberg, S.B., 2000. Relationships between benthic community condition, water quality, sediment quality, nutrient loads, and land use patterns in Chesapeake Bay. Estuaries 23, 80-96.

Diaz, R.J., Rosenberg, R., 2008. Spreading dead zones and consequences for marine ecosystems. science 321, 926-929.

Dietz, M.E., Clausen, J.C., 2008. Stormwater runoff and export changes with development in a traditional and low impact subdivision. Journal of Environmental Management 87, 560-566.

Dodds, W.K., 2006. Nutrients and the "dead zone": the link between nutrient ratios and dissolved oxygen in the northern Gulf of Mexico. Frontiers in Ecology and the Environment 4, 211-217.

Dudley, B., May, L., 2007. Estimating the phosphorus load to waterbodies from septic tanks.

EPA, 2013. BASINS (Better Assessment Science Integrating point and Non-point Sources).

EPA, 2015. Commercial Fertilizer Purchased, p. Nutrient Policy Data.

Gburek, W.J., Folmar, G.J., 1999. Flow and chemical contributions to streamflow in an upland watershed: a baseflow survey. Journal of Hydrology 217, 1-18.

Gove, N.E., Edwards, R.T., Conquest, L.L., 2001. Effects of scale on land use and water quality relationships: A longitudinal basin-wide perspective. Journal of the American Water Resources Association 37, 1721-1734.

Gritzner, J., 2006. Identifying Wetland Depressions in Bare-Ground LIDAR for Hydrologic Modeling, 2006 ESRI International User Conference.

Gronberg, J.M., Spahr, N.E., 2012. County-lvel estimates of nitrogen and phosphorus from commerical fertilizer for the Conterminous United States, 1987-2006, in: U.S. Geological Survey Scientific Investigations Report 2012-5207, p. (Ed.).

Gustafson, E.J., 1998. Quantifying landscape spatial pattern: What is the state of the art? Ecosystems 1, 143-156.

Halstead, J.A., Kliman, S., Berheide, C.W., Chaucer, A., Cock-Esteb, A., 2014. Urban stream syndrome in a small, lightly developed watershed: a statistical analysis of water chemistry parameters, land use patterns, and natural sources. Environmental Monitoring and Assessment 186, 3391-3414.

Hart, M.R., Quin, B.F., Nguyen, M.L., 2004. Phosphorus runoff from agricultural land and direct fertilizer effects: A review. Journal of Environmental Quality 33, 1954-1972.

He, C.S., Malcolm, S.B., Dahlberg, K.A., Fu, B.J., 2000. A conceptual framework for integrating hydrological and biological indicators into watershed management. Landscape and Urban Planning 49, 25-34.

Helsel, D.R., Hirsch, R.M., 1992. Statistical Methods in Water Resources. Elsevier.

Hogan, D.M., Jarnagin, S.T., Loperfido, J.V., Van Ness, K., 2014. Mitigating the effects of landscape development on strems in urbanizing watersheds Journal of the American Water Resources Association 50, 163-178.

Huang, J.L., Klemas, V., 2012. Using Remote Sensing of Land Cover Change in Coastal Watersheds to Predict Downstream Water Quality. Journal of Coastal Research 28, 930-944.

Irwin, E.G., Bockstael, N.E., 2007. The evolution of urban sprawl: Evidence of spatial heterogeneity and increasing land fragmentation. Proceedings of the National Academy of Sciences 104, 20672-20677.

Johnston, C.A., 1991. Sediment and nutrient retention by fresh-water wetlands-effects on surface-water quality Critical Reviews in Environmental Control 21, 491-565.

Kemp, W.M., Boynton, W.R., Adolf, J.E., Boesch, D.F., Boicourt, W.C., Brush, G., Cornwell, J.C., Fisher, T.R., Glibert, P.M., Hagy, J.D., 2005. Eutrophication of Chesapeake Bay: historical trends and ecological interactions. Marine Ecology Progress Series 303, 1-29.

King, K.W., Balogh, J.C., Agrawal, S.G., Tritabaugh, C.J., Ryan, J.A., 2012. Phosphorus concentration and loading reductions following changes in fertilizer application and formulation on managed turf. Journal of Environmental Monitoring 14, 2929-2938.

Kline, J.D., Thiers, P., Ozawa, C.P., Yeakley, J.A., Gordon, S.N., 2014. How well has land-use planning worked under different governance regimes? A case study in the Portland, OR-Vancouver, WA metropolitan area, USA. Landscape and Urban Planning 131, 51-63.

Kroll, C., Luz, J., Allen, B., Vogel, R.M., 2004. Developing a watershed characteristics database to improve low streamflow prediction. Journal of Hydrologic Engineering 9, 116-125.

Lee, S.W., Hwang, S.J., Lee, S.B., Hwang, H.S., Sung, H.C., 2009. Landscape ecological approach to the relationships of land use patterns in watersheds to water quality characteristics. Landscape and Urban Planning 92, 80-89.

Lehman, J.T., Bell, D.W., Doubek, J.P., McDonald, K.E., 2011. Reduced additions to river phosphorus for three years following implementation of a lawn fertilizer ordinance. Lake and Reservoir Management 27, 390-397.

Leitao, A.B., Miller, J., Ahern, J., McGarigal, K., 2006. Measuring Landscapes: A planner's handbook. Island Press, Washington D.C.

Lenat, D.R., Crawford, J.K., 1994. Effects of land-use on water-quality and aquatic biota of 3 North Carolina piedmont streams Hydrobiologia 294, 185-199.

Liu, A.J., Cameron, G.N., 2001. Analysis of landscape patterns in coastal wetlands of Galveston Bay, Texas (USA). Landscape Ecology 16, 581-595.

Liu, Z., Li, Y., Li, Z., 2009. Surface water quality and land use in Wisconsin, USA - a GIS approach. Journal of Integrative Environmental Sciences 6, 69-89.

Maranon, E., Cermeno, P., Huete-Ortega, M., Lopez-Sandoval, D.C., Mourino-Carballido, B., Rodriguez-Ramos, T., 2014. Resource Supply Overrides Temperature as a Controlling Factor of Marine Phytoplankton Growth. Plos One 9, 8.

Matte, A.L.L., Muller, S.C., Becker, F.G., 2015. Forest expansion or fragmentation? Discriminating forest fragments from natural forest patches through patch structure and spatial context metrics. Austral Ecology 40, 21-31.

May, C.W., Horner, R.R., Karr, J.R., Mar, B.W., Welch, E.B., 1999. Effects of urbanization on small streams in the Puget Sound ecoregion. Watershed Protection Techniques 2, 79.

McGarigal, K., 2015. Fragstats- Spatial Pattern Analysis Program for Quantifying Landscape Structure Version 4.0.

McKay, L., Bondelid, T., Dewald, T., Johnson, J., Moore, R., Rea, A., 2012. NHDPlus Version 2 User Guide.

Miller, K., 2012. State Laws Banning Phosphorus Fertilizer Use. OLR Research Report.

Moore, J.W., Schindler, D.E., Scheuerell, M.D., Smith, D., Frodge, J., 2003. Lake eutrophication at the urban fringe, Seattle region, USA. AMBIO: A Journal of the Human Environment 32, 13-18.

MPCA, 1999. Effects of Septic Systems on Ground Water Quality - Baxter Minnesota. Minnesota Pollution Control Agency - Ground Water and Toxics Monitoring Unit, St. Paul, MN, p. 47.

Muscutt, A.D., Harris, G.L., Bailey, S.W., Davies, D.B., 1993. Buffer zones to improve water-quality - a review of their potential use in UK agriculture. Agriculture Ecosystems and Environment 45, 59-77.

Nagy, R.C., Lockaby, B.G., Kalin, L., Anderson, C., 2012. Effects of urbanization on stream hydrology and water quality: the Florida Gulf Coast. Hydrological Processes 26, 2019-2030.

Nelson, E.J., Booth, D.B., 2002. Sediment sources in an urbanizing, mixed land-use watershed. Journal of Hydrology 264, 51-68.

NHDPlusV2, 2015. NHDPlus Version2 User's Manual, in: USEPA (Ed.).

NOAA-CCAP, 2010. Regional Land Cover Classification Scheme. National Oceanic and Atmospheric Administration (NOAA) Office for Coastal Management, Charleston, SC, pp. Data collected 1995-present.

Olmanson, L.G., Brezonik, P.L., Bauer, M.E., 2013. Airborne hyperspectral remote sensing to assess spatial distribution of water quality characteristics in large rivers: The Mississippi River and its tributaries in Minnesota. Remote Sensing of Environment 130, 254-265.

Örnólfsdóttir, E.B., Lumsden, S.E., Pinckney, J.L., 2004. Nutrient pulsing as a regulator of phytoplankton abundance and community composition in Galveston Bay, Texas. Journal of Experimental Marine Biology and Ecology 303, 197-220.

Ou, Y., Wang, X.Y., 2011. GIS and ordination techniques for studying influence of watershed characteristics on river water quality. Water Science and Technology 64, 861-870.

Paerl, H.W., Hall, N.S., Peierls, B.L., Rossignol, K.L., 2014. Evolving paradigms and challenges in estuarine and coastal eutrophication dynamics in a culturally and climatically stressed world. Estuaries and coasts 37, 243-258.

Paul, M.J., Meyer, J.L., 2001. Streams in the urban landscape. Annual Review of Ecology and Systematics 32, 333-365.

Perlman, D., Miller, J., 2005. When Humans and Nature Collide, Practical Ecology for planners, developers, and citizens. Island Press, Washington, D.C., pp. 36-49.

Peterjohn, W.T., Correll, D.L., 1984. Nutrient dynamics in an agricultural watershed-observations on the role of a riparian forest Ecology 65, 1466-1475.

Qiu, Z.Y., Prato, T., Wang, H.M., 2014. Assessing long-term water quality impacts of reducing phosphorus fertilizer in a US suburban watershed. Water Policy 16, 917-929.

Randolph, J., 2003. Ecosystem and Watershed Management, Environmental Land Use Planning and Management. Island Press, Washington D.C., pp. 253-260.

Rapport, D.J., Costanza, R., McMichael, A.J., 1998. Assessing ecosystem health. Trends in Ecology and Evolution 13, 397-402.

Rebich, R.A., Houston, N.A., Mize, S.V., Pearson, D.K., Ging, P.B., Evan Hornig, C., 2011. Sources and Delivery of Nutrients to the Northwestern Gulf of Mexico from Streams in the South-Central United States1. JAWRA Journal of the American Water Resources Association 47, 1061-1086.

Reddy, K.R., Kadlec, R.H., Flaig, E., Gale, P.M., 1999. Phosphorus retention in streams and wetlands: A review. Critical Reviews in Environmental Science and Technology 29, 83-146.

Salm, Clark, 2000. Protected Areas for Beaches, Marine and Coastal Protected Areas: A Guide for Planners and Managers, pp. 231-238.

Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsa, D., Anthemidis, A., Sofoniou, M., Kouimtzis, T., 2003. Assessment of the surface water quality in Northern Greece. Water Research 37, 4119-4124.

Sliva, L., Williams, D.D., 2001. Buffer zone versus whole catchment approaches to studying land use impact on river water quality. Water Research 35, 3462-3472.

Smith, S.V., Swaney, D.P., Talaue-Mcmanus, L., Bartley, J.D., Sandhei, P.T., McLaughlin, C.J., Dupra, V.C., Crossland, C.J., Buddemeier, R.W., Maxwell, B.A., 2003. Humans, hydrology, and the distribution of inorganic nutrient loading to the ocean. BioScience 53, 235-245.

Soldat, D.J., Petrovic, A.M., 2008. The Fate and Transport of Phosphorus in Turfgrass Ecosystems. Crop Science 48, 2051-2065.

STATA, 2013. Stata User Guide Release 13, 13 ed. Stata Press.

Su, S.L., Hu, Y.N., Luo, F.H., Mai, G.C., Wang, Y.P., 2014. Farmland fragmentation due to anthropogenic activity in rapidly developing region. Agricultural Systems 131, 87-93.

Sun, R.H., Chen, L.D., Chen, W.L., Ji, Y.H., 2013. Effect of Land-Use Patterns on Total Nitrogen Concentration in the Upstream Regions of the Haihe River Basin, China. Environmental Management 51, 45-58.

Sun, Y.W., Guo, Q.H., Liu, J., Wang, R., 2014. Scale Effects on Spatially Varying Relationships Between Urban Landscape Patterns and Water Quality. Environmental Management 54, 272-287.

SWQM, 2012. 2012 Guidance for Assessing and Reporting Surface Water Quality in Texas. Surface Water Quality Monitoring Program; Monitoring and Assessment Section; Water Quality Planning Division, p. 139.

Tong, S.T.Y., Chen, W.L., 2002. Modeling the relationship between land use and surface water quality. Journal of Environmental Management 66, 377-393.

Tran, C.P., Bode, R.W., Smith, A.J., Kleppel, G.S., 2010. Land-use proximity as a basis for assessing stream water quality in New York State (USA). Ecological Indicators 10, 727-733.

Trends, T.L., 2014. Texas Land Facts, in: Resources, t.A.M.I.o.R.N. (Ed.).

Tu, J., 2011. Spatial and temporal relationships between water quality and land use in northern Georgia, USA. Journal of Integrative Environmental Sciences 8, 151-170.

Turner, R.E., Rabalais, N.N., 1991. CHANGES IN MISSISSIPPI RIVER WATER-QUALITY THIS CENTURY. Bioscience 41, 140-147.

U.S. EPA, O.o.w., 2010. Texas Water Quality Report. U.S. Environmental Protection Agency, Washington, DC.

USDA-NRCS, U.a.E., The Watershed Boundary Dataset (WBD) was created from a variety of sources from each state and aggregated into a standard national layer for use in strategic planning and accountability, in: Coordinated effort between the United States Department of Agriculture-Natural Resources Conservation Service (USDA-NRCS), t.U.S.G.S.U., and the Environmental Protection Agency (EPA) (Ed.).

USEPA, 2005. Stormwater Phase II Final Rule: Small Construction Program Overview, in: water, U.U.S.E.P.A.O.o. (Ed.), p. 5 pp.

Varlamoff, S., Florkowski, W.J., Jordan, J.L., Latimer, J., Braman, K., 2001. Georgia homeowner survey of landscape management practices. Horttechnology 11, 326-331.

Walsh, C.J., Roy, A.H., Feminella, J.W., Cottingham, P.D., Groffman, P.M., Morgan, R.P., 2005. The urban stream syndrome: current knowledge and the search for a cure. Journal of the North American Benthological Society 24, 706-723.

Williams, M., Hopkinson, C., Rastetter, E., Vallino, J., Claessens, L., 2005. Relationships of land use and stream solute concentrations in the Ipswich River basin, northeastern Massachusetts. Water Air and Soil Pollution 161, 55-74.

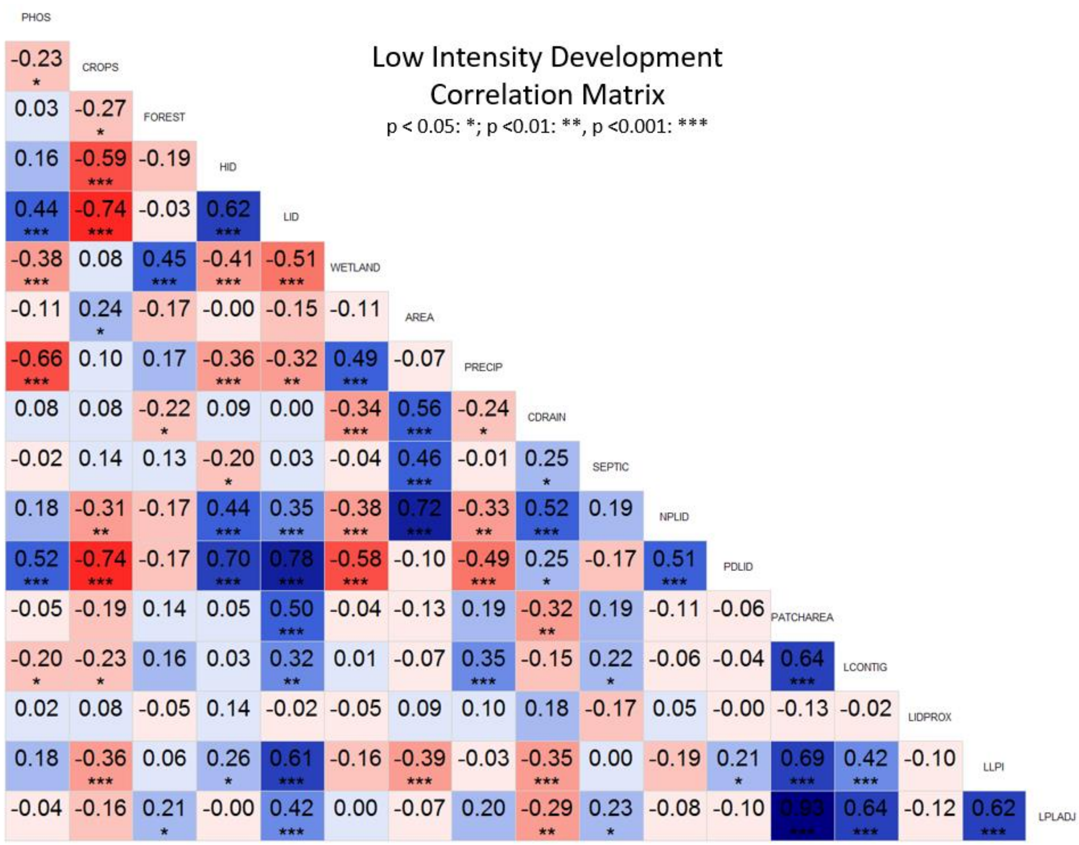
Williams, R.L., 2000. A note on robust variance estimation for cluster-correlated data. Biometrics 56, 645-646.

Withers, P.J.A., Jarvie, H.P., Hodgkinson, R.A., Palmer-Felgate, E.J., Bates, A., Neal, M., Howells, R., Withers, C.M., Wickham, H.D., 2009. Characterization of Phosphorus Sources in Rural Watersheds. Journal of Environmental Quality 38, 1998-2011.

Zampella, R.A., Procopio, N.A., Lathrop, R.G., Dow, C.L., 2007. Relationship of land-use/land-cover patterns and surface-water quality in the Mullica River basin. Journal of the American Water Resources Association 43, 594-604.

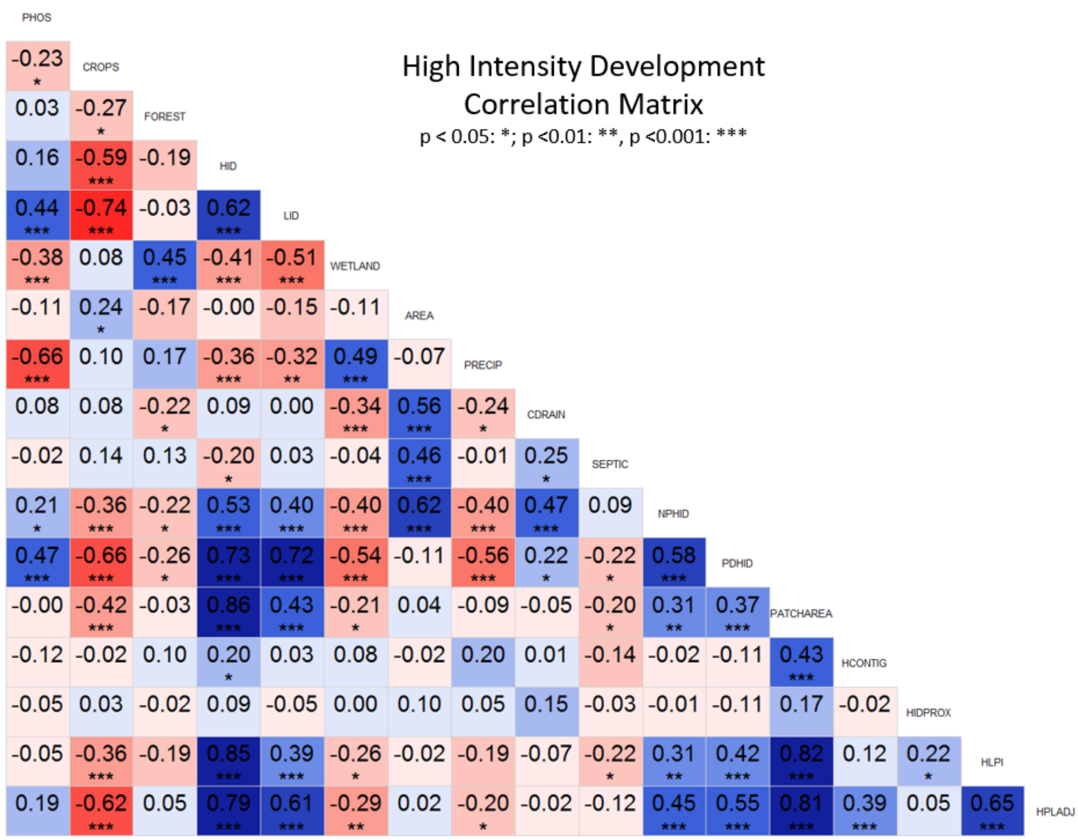
APPENDIX A

CORRELATION MATRIX



	PHOS	CROPS	FOREST	HID	LID	WETLAND	AREA	PRECIP	CDRAIN	SEPTIC	NPLID	PDLID	PATCHAREA	LCONTIG	LIDPROX	LLPI
CROPS	-0.23 *															
FOREST	0.03	-0.27 *														
HID	0.16	-0.59 ***	-0.19													
LID	0.44 ***	-0.74 ***	-0.03	0.62 ***												
WETLAND	-0.38 ***	0.08	0.45 ***	-0.41 ***	-0.51 ***											
AREA	-0.11	0.24 *	-0.17	-0.00	-0.15	-0.11										
PRECIP	-0.66 ***	0.10	0.17	-0.36 ***	-0.32 **	0.49 ***	-0.07									
CDRAIN	0.08	0.08	-0.22 *	0.09	0.00	-0.34 ***	0.56 ***	-0.24 *								
SEPTIC	-0.02	0.14	0.13	-0.20 *	0.03	-0.04	0.46 ***	-0.01	0.25 *							
NPLID	0.18	-0.31 **	-0.17	0.44 ***	0.35 ***	-0.38 ***	0.72 ***	-0.33 **	0.52 ***	0.19						
PDLID	0.52 ***	-0.74 ***	-0.17	0.70 ***	0.78 ***	-0.58 ***	-0.10	-0.49 ***	0.25 *	-0.17	0.51 ***					
PATCHAREA	-0.05	-0.19	0.14	0.05	0.50 ***	-0.04	-0.13	0.19	-0.32 **	0.19	-0.11	-0.06				
LCONTIG	-0.20 *	-0.23 *	0.16	0.03	0.32 **	0.01	-0.07	0.35 ***	-0.15	0.22 *	-0.06	-0.04	0.64 ***			
LIDPROX	0.02	0.08	-0.05	0.14	-0.02	-0.05	0.09	0.10	0.18	-0.17	0.05	-0.00	-0.13	-0.02		
LLPI	0.18	-0.36 ***	0.06	0.26 *	0.61 ***	-0.16	-0.39 ***	-0.03	-0.35 ***	0.00	-0.19	0.21 *	0.69 ***	0.42 ***	-0.10	
LPLADJ	-0.04	-0.16	0.21 *	-0.00	0.42 ***	0.00	-0.07	0.20	-0.29 **	0.23 *	-0.08	-0.10	0.93 ***	0.64 ***	-0.12	0.62 ***

A- 1: Low intensity development correlation matrix



A- 2: High intensity development correlation matrix

APPENDIX B

STATISTICAL MODELS

The VIFs are based on a simple OLS regression to determine if there is multicollinearity. Then, if there is, the control variable that has the high multicollinearity is removed and the regression is re-run. Then, the spatial regression is run to get the coefficients.

Low Intensity Development Models

Proximity to stream	
	VIF
PLANDLID	4.85
PLANDCC	4.82
PLANDWetla	2.8
PLANDHID	2.58
Area_KM	2
Contrib_dr	1.92
PLANDFores	1.75
OSSF_count	1.72
Precip_in	1.51
PX_LID_ALL	1.27

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set              : updateddata2
Spatial Weight        : forGEODA2.gwt
Dependent Variable    :  logphos101  Number of Observations:   99
Mean dependent var    :   -0.414403  Number of Variables   :   12
S.D. dependent var    :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)    :    0.596849

R-squared             :    0.776040  Log likelihood        :    -77.5977
Sq. Correlation       : -            Akaike info criterion :     179.195
Sigma-square          :    0.266952  Schwarz criterion     :     210.337
S.E of regression     :    0.516674

-----------------------------------------------------------------------
    Variable    Coefficient      Std.Error       z-value    Probability
-----------------------------------------------------------------------
W_logphos101     0.5968495       0.081239       7.346834     0.00000
    CONSTANT      9.535249       1.358823       7.017287     0.00000
   precip_in    -0.1967445     0.02512273      -7.831336     0.00000
  contrib_dr  0.0005654664   0.0008748909      0.6463279     0.51807
     count_1 -9.790453e-005  0.0001549223     -0.6319589     0.52741
     area_km  0.0007132766    0.001077379       0.662048     0.50794
     plandcc  -0.006755696     0.01191034     -0.5672128     0.57057
  plandfores    0.02898637     0.01484218       1.952973     0.05082
    plandlid    0.02646883     0.01583821       1.671201     0.09468
  plandwetla   -0.01933878    0.009661821      -2.001566     0.04533
    plandhid   -0.04490164     0.01099338      -4.084426     0.00004
  px_lid_all     0.2668581      0.1942178       1.374014     0.16944
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                        DF       VALUE        PROB
Breusch-Pagan test                         10         34.8613     0.00013

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                        DF       VALUE        PROB
Likelihood Ratio Test                       1         33.2649     0.00000
```

B- 1: Low intensity development proximity to stream regression VIF and spatial lag model.

LPI LID	
Variable	**VIF**
PLANDLID	7.33
PLANDCC	5.01
PLANDWetla	2.9
LPILID	**2.48**
PLANDHID	2.43
Area_KM	2.17
Contrib_dr	1.87
PLANDFores	1.69
OSSF_count	1.63
Precip_in	1.45

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)  :    0.566033

R-squared           :    0.774596  Log likelihood        :    -77.6017
Sq. Correlation     : -            Akaike info criterion :     179.203
Sigma-square        :    0.268672  Schwarz criterion     :     210.345
S.E of regression   :    0.518336

-----------------------------------------------------------------------
       Variable     Coefficient      Std.Error        z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.5660331      0.08656615       6.538735      0.00000
    CONSTANT       9.019893        1.360128       6.631652      0.00000
   precip_in     -0.1878135         0.02466      -7.616119      0.00000
  contrib_dr   0.0006112731    0.0008690023      0.7034195      0.48179
     count_1  -0.0001402901    0.0001516418     -0.9251409      0.35489
     area_km   0.0003545097     0.001121957      0.3159746      0.75202
     plandcc    0.003505052      0.01217644      0.2878551      0.77346
  plandfores     0.03342176      0.01466426       2.279131      0.02266
    plandlid     0.04774238      0.02004728       2.381489      0.01724
  plandwetla    -0.01517538     0.009967497      -1.522486      0.12789
    plandhid    -0.03995046      0.01071905      -3.727053      0.00019
      lpilid    -0.08134963      0.05831634      -1.394971      0.16302
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                      DF       VALUE        PROB
Breusch-Pagan test                        10        33.8928     0.00019

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                      DF       VALUE        PROB
Likelihood Ratio Test                     1         27.6420     0.00000
```

B- 2: Low intensity development largest patch index regression VIF and spatial lag model.

Percent of Like Adjacencies	
Variable	**VIF**
PLANDLID	7.92
PLANDCC	5.45
PLANDWetla	2.9
PLANDHID	2.42
PLADJLID	**2.33**
Area_KM	2.05
Contrib_dr	1.93
PLANDFores	1.82
OSSF_count	1.63
Precip_in	1.61

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)  :    0.563443

R-squared           :    0.780670  Log likelihood        :    -76.2246
Sq. Correlation     : -            Akaike info criterion :     176.449
Sigma-square        :    0.261432  Schwarz criterion     :     207.591
S.E of regression   :    0.511304

-----------------------------------------------------------------------
       Variable     Coefficient      Std.Error        z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.5634432      0.08413907       6.696571      0.00000
    CONSTANT       9.551916        1.324078       7.214012      0.00000
   precip_in     -0.1742668      0.02485657      -7.010894      0.00000
  contrib_dr   0.0003581713    0.0008699778      0.4117016      0.68056
     count_1  -0.0001345807    0.0001492703     -0.9015904      0.36727
     area_km     0.00113727     0.001079645       1.053374      0.29217
     plandcc    0.009563797      0.01260281       0.758862      0.44793
  plandfores     0.04140838      0.01509961       2.742347      0.00610
    plandhid    -0.04257313      0.01055168      -4.034727      0.00005
    plandlid     0.05924875      0.02082798       2.844671      0.00445
  plandwetla    -0.01373509     0.009822143      -1.398381      0.16200
    pladjlid    -0.03599565      0.01669712        -2.1558      0.03110
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                      DF       VALUE        PROB
Breusch-Pagan test                        10        34.0347     0.00018

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                      DF       VALUE        PROB
Likelihood Ratio Test                     1         29.1229     0.00000
```

B- 3: Low intensity development percent of like adjacencies regression VIF and spatial lag model.

Contiguity	
Variable	**VIF**
PLANDLID	5.25
PLANDCC	4.46
PLANDWetla	2.79
PLANDHID	2.42
Area_KM	2
Precip_in	1.82
Contrib_dr	1.8
PLANDFores	1.71
OSSF_count	1.68
CONTIG_MNL	**1.65**

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)  :    0.565953

R-squared           :    0.778281  Log likelihood        :    -76.7849
Sq. Correlation     : -            Akaike info criterion :      177.57
Sigma-square        :    0.264279  Schwarz criterion     :     208.711
S.E of regression   :    0.514081

-----------------------------------------------------------------------
    Variable     Coefficient      Std.Error        z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.5659529     0.08462869       6.687483     0.00000
    CONSTANT       9.365093       1.335351       7.013205     0.00000
   precip_in     -0.1691829     0.02648925       -6.38685     0.00000
  contrib_dr   0.0007451914   0.0008435895      0.8833578     0.37704
     count_1  -0.0001039065   0.0001526501     -0.6806839     0.49607
     area_km   0.0007568146    0.001072398      0.7057218     0.48036
     plandcc  -0.0004320295     0.01139074    -0.03792814     0.96974
  plandfores     0.03540113     0.01470723       2.407057     0.01608
    plandlid     0.04140915     0.01684277       2.458571     0.01395
  plandwetla    -0.01819058    0.009643477      -1.886309     0.05925
    plandhid    -0.04109071     0.01060452      -3.874829     0.00011
  contig_mnl      -8.382077       4.455249      -1.881394     0.05992
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                         DF        VALUE        PROB
Breusch-Pagan test                           10        36.2932      0.00007

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                         DF        VALUE        PROB
Likelihood Ratio Test                         1        28.9883      0.00000
```

B- 4: Low intensity development contiguity regression VIF and spatial lag model.

Average Patch Size	
Variable	**VIF**
PLANDLID	9.95
PLANDCC	6.01
PLANDWetla	3.02
AREA_MNLID	**2.98**
PLANDHID	2.43
Area_KM	2.02
Contrib_dr	1.96
PLANDFores	1.75
Precip_in	1.7
OSSF_count	1.63

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)  :    0.557419

R-squared           :    0.781912  Log likelihood        :    -75.8864
Sq. Correlation     : -            Akaike info criterion :     175.773
Sigma-square        :    0.259952  Schwarz criterion     :     206.914
S.E of regression   :    0.509855

-----------------------------------------------------------------------
    Variable     Coefficient      Std.Error        z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.5574187     0.08427078       6.614615     0.00000
    CONSTANT       8.186455       1.387826        5.89876     0.00000
   precip_in     -0.1695903     0.02524739      -6.717141     0.00000
  contrib_dr   0.0002570295   0.0008768785      0.2931187     0.76943
     count_1  -0.0001351555   0.0001486894     -0.9089789     0.36336
     area_km    0.001029495    0.001068869      0.9631627     0.33547
     plandcc     0.01348471     0.01321632       1.020307     0.30758
  plandfores     0.03915899     0.01474205       2.656279     0.00790
    plandhid    -0.04277051      0.0105203      -4.065523     0.00005
    plandlid     0.06975987     0.02332737       2.990473     0.00279
  plandwetla    -0.01128662     0.01001122      -1.127397     0.25957
  area_mnlid     -0.9564817      0.4146952      -2.306469     0.02108
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                         DF        VALUE        PROB
Breusch-Pagan test                           10        32.2421      0.00036

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                         DF        VALUE        PROB
Likelihood Ratio Test                         1        28.4606      0.00000
```

B- 5: Low intensity development average patch size regression VIF and spatial lag model.

In the patch density model, the variable that has a higher than accepted VIF (greater than 10) is the variable of interest: low intensity development patch density, this variable cannot be removed from the model. Instead, it should be noted that the variance inflation is slightly higher than commonly accepted for this variable.

Patch Density	
Variable	**VIF**
PDLID	**11.11**
PLANDCC	9.53
PLANDLID	4.75
PLANDWetla	3.63
PLANDHID	2.42
Contrib_dr	2.18
Area_KM	2.05
PLANDFores	1.95
Precip_in	1.86
OSSF_count	1.69

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set              : updateddata2
Spatial Weight        : forGEODA2.gwt
Dependent Variable    :   logphos101  Number of Observations:   99
Mean dependent var    :    -0.414403  Number of Variables   :   12
S.D. dependent var    :      1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)    :     0.478399

R-squared             :     0.794087  Log likelihood        :    -72.3943
Sq. Correlation       : -             Akaike info criterion :     168.789
Sigma-square          :      0.24544  Schwarz criterion     :      199.93
S.E of regression     :     0.495419

-----------------------------------------------------------------------
    Variable     Coefficient      Std.Error        z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.4783989     0.08959285         5.3397      0.00000
    CONSTANT       5.632466       1.636535       3.441703      0.00058
   precip_in     -0.1558397     0.02528767      -6.162676      0.00000
  contrib_dr  -0.0006247775   0.0009133515     -0.6840494      0.49394
     count_1  -8.003529e-005  0.0001468598     -0.5449775      0.58577
     area_km    0.001531313    0.001052651       1.454721      0.14575
     plandcc     0.04200597     0.01646336       2.551483      0.01073
  plandfores     0.05377803     0.01524041       3.528648      0.00042
    plandlid     0.03140959     0.01495981       2.099598      0.03576
  plandwetla     0.00329369     0.01089058      0.3024347      0.76232
    plandhid    -0.04315562     0.01021197      -4.225985      0.00002
       pdlid     0.06558664     0.01817219       3.609176      0.00031
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                       DF        VALUE          PROB
Breusch-Pagan test                        10         28.6678        0.00141

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                       DF        VALUE          PROB
Likelihood Ratio Test                      1         20.0688        0.00001
```

B- 6: Low intensity development average patch density regression VIF and spatial lag model.

Patch Number	
Variable	**VIF**
Area_KM	8.14
NPLID	**8.03**
PLANDCC	6.25
PLANDLID	4.72
PLANDWetla	3.03
PLANDHID	2.42
OSSF_count	1.81
Contrib_dr	1.8
PLANDFores	1.7
Precip_in	1.55

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)  :    0.590101

R-squared           :    0.773701  Log likelihood        :    -78.0403
Sq. Correlation     : -            Akaike info criterion :     180.081
Sigma-square        :     0.26974  Schwarz criterion     :     211.222
S.E of regression   :    0.519365

-----------------------------------------------------------------------
    Variable    Coefficient     Std.Error       z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.590101     0.08288883      7.119186      0.00000
    CONSTANT      8.851434       1.401716      6.314712      0.00000
   precip_in    -0.1826547      0.0253605     -7.202331      0.00000
  contrib_dr  0.0008451649   0.0008502169      0.994058      0.32019
     count_1  -0.000101903   0.0001592553    -0.6398718      0.52226
     area_km  -0.001138004    0.002179699    -0.5220922      0.60161
     plandcc   0.005276962     0.01362741     0.3872316      0.69858
  plandfores     0.0335334      0.0147091      2.279772      0.02262
    plandhid   -0.04117609     0.01070054     -3.848037      0.00012
    plandlid    0.03024313     0.01570762      1.925379      0.05418
  plandwetla   -0.01550415     0.01015547     -1.526679      0.12684
       nplid  7.591946e-005  7.60059e-005     0.9988627      0.31786
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                        DF      VALUE        PROB
Breusch-Pagan test                         10       32.7002      0.00031

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                        DF      VALUE        PROB
Likelihood Ratio Test                       1       31.8526      0.00000
========================= END OF REPORT ==============================
```

B- 7: Low intensity development patch number regression VIF and spatial lag model.

High Intensity Development

Proximity to Stream	
Variable	**VIF**
PLANDLID	4.72
PLANDCC	4.45
PLANDWetla	2.79
PLANDHID	2.49
Area_KM	2
Contrib_dr	1.84
PLANDFores	1.7
OSSF_count	1.62
Precip_in	1.44
PX_HID_ALL	**1.07**

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.    (Rho) :    0.601432

R-squared           :    0.772006  Log likelihood        :    -78.5309
Sq. Correlation     : -            Akaike info criterion :     181.062
Sigma-square        :    0.271759  Schwarz criterion     :     212.203
S.E of regression   :    0.521305

-----------------------------------------------------------------------
    Variable     Coefficient      Std.Error        z-value     Probability
-----------------------------------------------------------------------
W_logphos101      0.6014325      0.08202788        7.33205      0.00000
    CONSTANT       9.206784        1.354813        6.79561      0.00000
   precip_in     -0.1883986      0.02482494      -7.589088      0.00000
  contrib_dr   0.0008628738    0.0008624948       1.000439      0.31710
     count_1  -0.0001492869     0.000151461     -0.9856457       0.32431
     area_km   0.0007397421     0.001086813      0.6806524      0.49609
     plandcc  -0.002021007      0.01155313     -0.1749315      0.86113
  plandfores     0.03260158      0.01478301       2.205341      0.02743
    plandlid     0.02977681      0.01576912       1.888298      0.05899
  plandwetla    -0.01859305     0.009762634      -1.904512      0.05684
    plandhid    -0.04121203      0.01090021      -3.780848      0.00016
  px_hid_all    0.009043653       0.1593587     0.05675029      0.95474
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                      DF       VALUE        PROB
Breusch-Pagan test                        10       33.4701      0.00023

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                      DF       VALUE        PROB
Likelihood Ratio Test                     1        33.3514      0.00000
```

B- 8: High intensity development proximity to stream regression VIF and spatial lag model.

Variable	VIF
PLANDHID	8.21
LPIHID	**4.86**
PLANDCC	4.75
PLANDLID	4.73
PLANDWetla	2.79
Area_KM	2.01
Contrib_dr	1.91
PLANDFores	1.7
OSSF_count	1.61
Precip_in	1.48

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.    (Rho) :    0.543573

R-squared           :    0.776669  Log likelihood        :    -76.9357
Sq. Correlation     : -            Akaike info criterion :     177.871
Sigma-square        :    0.266202  Schwarz criterion     :     209.013
S.E of regression   :    0.515948

-----------------------------------------------------------------------
    Variable     Coefficient      Std.Error        z-value     Probability
-----------------------------------------------------------------------
W_logphos101      0.5435734      0.08836701       6.151316      0.00000
    CONSTANT       8.930782        1.360942       6.562206      0.00000
   precip_in     -0.1861519      0.02477222      -7.514539      0.00000
  contrib_dr   0.0004518769    0.0008761637      0.5157448      0.60603
     count_1  -0.0001791985    0.0001506404      -1.189577       0.23421
     area_km   0.0008840934     0.001077953      0.8201593      0.41213
     plandcc    0.003164988      0.01180283       0.268155      0.78858
  plandfores     0.03412396      0.01462682       2.332972      0.01965
    plandlid     0.03152499      0.01569888       2.008104      0.04463
  plandwetla    -0.01693918      0.00969753      -1.746751      0.08068
    plandhid  -0.009759716      0.02027998     -0.4812489      0.63034
      lpihid    -0.06856714      0.03731362       -1.83759      0.06612
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                      DF       VALUE        PROB
Breusch-Pagan test                        10       28.7509      0.00137

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                      DF       VALUE        PROB
Likelihood Ratio Test                     1        24.6672      0.00000
```

B- 9: High intensity development largest patch index regression VIF and spatial lag model.

Percent of Like Adjacencies	
Variable	**VIF**
PLANDLID	4.77
PLANDCC	4.46
PLANDHID	4.07
PLADJHID	**3.57**
PLANDWetla	2.83
Area_KM	2.15
Contrib_dr	1.87
PLANDFores	1.83
OSSF_count	1.62
Precip_in	1.44

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)  :    0.587056

R-squared           :    0.776097  Log likelihood        :    -77.4814
Sq. Correlation     : -            Akaike info criterion :     178.963
Sigma-square        :    0.266883  Schwarz criterion     :     210.104
S.E of regression   :    0.516607

-----------------------------------------------------------------------
    Variable    Coefficient      Std.Error       z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.587056     0.08287271       7.083829     0.00000
    CONSTANT      8.855091       1.364777       6.488306     0.00000
   precip_in    -0.1944668     0.02484001      -7.828774     0.00000
  contrib_dr   0.001120848    0.000862124         1.3001     0.19357
     count_1 -0.0001384109   0.0001501524     -0.9218032     0.35663
     area_km  0.0003349686    0.001114225      0.3006292     0.76370
     plandcc -0.0005110207     0.01145635    -0.04460589     0.96442
  plandfores    0.02652915     0.01515792       1.750184     0.08009
    plandlid    0.02817245     0.01568854       1.795734     0.07254
  plandwetla   -0.01666473    0.009760408      -1.707381     0.08775
    plandhid    -0.0540665     0.01381343      -3.914052     0.00009
    pladjhid    0.01389977    0.009536208       1.457578     0.14496
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                       DF      VALUE        PROB
Breusch-Pagan test                         10       35.4590     0.00010

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                       DF      VALUE        PROB
Likelihood Ratio Test                       1       31.8852     0.00000
```

B- 10: High intensity development percent of like adjacencies regression VIF and spatial lag model.

Contiguity	
Variable	**VIF**
PLANDCC	4.93
PLANDLID	4.89
PLANDWetla	2.84
PLANDHID	2.83
Area_KM	2.01
Contrib_dr	1.86
PLANDFores	1.81
OSSF_count	1.66
Precip_in	1.58
CONTIG_MNH	**1.34**

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set            : updateddata2
Spatial Weight      : forGEODA2.gwt
Dependent Variable  :  logphos101  Number of Observations:   99
Mean dependent var  :   -0.414403  Number of Variables   :   12
S.D. dependent var  :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)  :     0.59587

R-squared           :    0.772306  Log likelihood        :    -78.4055
Sq. Correlation     : -            Akaike info criterion :     180.811
Sigma-square        :    0.271401  Schwarz criterion     :     211.952
S.E of regression   :    0.520962

-----------------------------------------------------------------------
    Variable    Coefficient      Std.Error       z-value    Probability
-----------------------------------------------------------------------
W_logphos101     0.5958699     0.08249821       7.222822     0.00000
    CONSTANT      9.374466       1.386101       6.763191     0.00000
   precip_in    -0.1927804     0.02604599      -7.401541     0.00000
  contrib_dr 0.0007896924   0.0008672915      0.9105271     0.36254
     count_1 -0.0001390613    0.000153314     -0.9070358     0.36439
     area_km 0.0007717909    0.001087317      0.7098122     0.47782
     plandcc  -0.004041901      0.0121553     -0.3325216     0.73950
  plandfores    0.03067815     0.01523537       2.013614     0.04405
    plandlid    0.02858621     0.01603145       1.783133     0.07456
  plandwetla   -0.01915365    0.009841519      -1.946209     0.05163
    plandhid   -0.04338173     0.01161123      -3.736188     0.00019
  contig_mnh     0.7114457       1.402501      0.5072693     0.61197
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                       DF      VALUE        PROB
Breusch-Pagan test                         10       33.7164     0.00021

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                       DF      VALUE        PROB
Likelihood Ratio Test                       1       32.3723     0.00000
```

B- 11: High intensity development contiguity regression VIF and spatial lag model.

Average Patch Size	
Variable	**VIF**
PLANDHID	12.46
AREA_MNHID	**7.73**
PLANDCC	5.64
PLANDLID	4.9
PLANDWetla	2.8
PLANDFores	2.21
Area_KM	2.17
Precip_in	1.87
Contrib_dr	1.87
OSSF_count	1.7

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set             : updateddata2
Spatial Weight       : forGEODA2.gwt
Dependent Variable   :  logphos101  Number of Observations:   99
Mean dependent var   :   -0.414403  Number of Variables   :   11
S.D. dependent var   :     1.09177  Degrees of Freedom    :   88
Lag coeff.   (Rho)   :    0.601801

R-squared            :    0.758714  Log likelihood        :   -81.3397
Sq. Correlation      : -            Akaike info criterion :    184.679
Sigma-square         :    0.287603  Schwarz criterion     :    213.226
S.E of regression    :    0.536286

-----------------------------------------------------------------------
    Variable    Coefficient      Std.Error       z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.6018008     0.08287466      7.261578      0.00000
    CONSTANT       7.779427       1.343536      5.790266      0.00000
   precip_in     -0.1656022     0.02521154     -6.568509      0.00000
  contrib_dr   0.0007312674   0.0008872889     0.8241593      0.40985
     count_1  -0.0001550754    0.000158991    -0.9753717      0.32938
     area_km   0.0008424821     0.00114982     0.7327077      0.46374
     plandcc     0.01145925     0.01089817      1.051485      0.29304
  plandfores     0.04739477       0.014749      3.213423      0.00131
    plandlid     0.03239387     0.01631401      1.985648      0.04707
  plandwetla    -0.01594248    0.009982851     -1.596987      0.11027
  area_mnhid     -0.3458972      0.1198194     -2.886822      0.00389
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                       DF       VALUE        PROB
Breusch-Pagan test                         9         32.1530     0.00019

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                       DF       VALUE        PROB
Likelihood Ratio Test                      1         32.1765     0.00000
```

Remove PLANDHID because the VIF is too high.

Variable	**VIF**
PLANDLID	4.8
PLANDCC	3.75
PLANDWetla	2.75
Area_KM	2.12
Contrib_dr	1.84
OSSF_count	1.68
PLANDFores	1.6
AREA_MNHID	**1.5**
Precip_in	1.36

B- 12 High intensity development average patch size regression VIF and spatial lag model. Because the percent of high intensity development in the watershed is to highly collinear with the average patch size, it is removed from the spatial regression.

Patch Density	
Variable	**VIF**
PDHID	**7**
PLANDCC	6.19
PLANDLID	4.79
PLANDWetla	2.83
PLANDHID	2.56
Precip_in	2.11
PLANDFores	2.1
Area_KM	2.07
Contrib_dr	2
OSSF_count	1.68

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set              : updateddata2
Spatial Weight        : forGEODA2.gwt
Dependent Variable    :  logphos101  Number of Observations:   99
Mean dependent var    :   -0.414403  Number of Variables   :   12
S.D. dependent var    :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)    :    0.573714

R-squared             :    0.773448  Log likelihood        :    -77.9282
Sq. Correlation       : -            Akaike info criterion :     179.856
Sigma-square          :     0.27004  Schwarz criterion     :     210.998
S.E of regression     :    0.519654

-----------------------------------------------------------------------
    Variable    Coefficient      Std.Error        z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.5737137      0.0856118        6.70134      0.00000
    CONSTANT       8.104495       1.694196       4.783681      0.00000
   precip_in     -0.1733249     0.02847095      -6.087782      0.00000
  contrib_dr   0.0005338615     0.00090336      0.5909731      0.55454
     count_1  -0.0001250865   0.0001538372     -0.8131093       0.41616
     area_km    0.001007936    0.001105028      0.9121362      0.36170
     plandcc    0.005962422     0.01360802      0.4381551      0.66127
  plandfores     0.04084868     0.01653402       2.470583      0.01349
    plandlid     0.02908388      0.0157645       1.844897      0.06505
  plandwetla    -0.01664897    0.009845197      -1.691075      0.09082
    plandhid      -0.044159     0.01100811      -4.011498      0.00006
       pdhid      0.0313843     0.02803386       1.119514      0.26292
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                        DF        VALUE         PROB
Breusch-Pagan test                         10        35.6534       0.00010

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                        DF        VALUE         PROB
Likelihood Ratio Test                       1         28.3964       0.00000
```

B- 13: High intensity development average patch density regression VIF and spatial lag model.

Patch Number	
Variable	**VIF**
NPHID	**5.89**
PLANDCC	5.69
Area_KM	5.64
PLANDLID	4.72
PLANDWetla	2.88
PLANDHID	2.45
OSSF_count	1.83
Contrib_dr	1.8
PLANDFores	1.73
Precip_in	1.66

```
SUMMARY OF OUTPUT: SPATIAL LAG MODEL - MAXIMUM LIKELIHOOD ESTIMATION
Data set              : updateddata2
Spatial Weight        : forGEODA2.gwt
Dependent Variable    :  logphos101  Number of Observations:   99
Mean dependent var    :   -0.414403  Number of Variables   :   12
S.D. dependent var    :     1.09177  Degrees of Freedom    :   87
Lag coeff.   (Rho)    :    0.595116

R-squared             :    0.772327  Log likelihood        :     -78.393
Sq. Correlation       : -            Akaike info criterion :     180.786
Sigma-square          :    0.271377  Schwarz criterion     :     211.928
S.E of regression     :    0.520939

-----------------------------------------------------------------------
    Variable    Coefficient      Std.Error        z-value    Probability
-----------------------------------------------------------------------
W_logphos101      0.5951164     0.08274726       7.191978      0.00000
    CONSTANT       8.955273       1.442849       6.206659      0.00000
   precip_in     -0.1839926     0.02616969      -7.030751      0.00000
  contrib_dr   0.0008712154   0.0008525502       1.021893      0.30683
     count_1  -0.0001229084   0.0001606399     -0.7651176       0.44420
     area_km  -2.813173e-005   0.001818748    -0.01546764       0.98766
     plandcc    0.001236102     0.01304097     0.09478602      0.92448
  plandfores     0.03386725     0.01492628       2.268968      0.02327
    plandlid     0.03002431     0.01576007       1.905087      0.05677
  plandwetla     -0.0175352    0.009921477      -1.767398      0.07716
    plandhid    -0.04178013     0.01080094      -3.868194      0.00011
       nphid  8.044446e-005   0.0001514641      0.5311123       0.59534
-----------------------------------------------------------------------

REGRESSION DIAGNOSTICS
DIAGNOSTICS FOR HETEROSKEDASTICITY
RANDOM COEFFICIENTS
TEST                                        DF        VALUE         PROB
Breusch-Pagan test                         10        33.5937       0.00022

DIAGNOSTICS FOR SPATIAL DEPENDENCE
SPATIAL LAG DEPENDENCE FOR WEIGHT MATRIX : forGEODA2.gwt
TEST                                        DF        VALUE         PROB
Likelihood Ratio Test                       1         32.2123       0.00000
```

B- 14: High intensity development patch number regression VIF and spatial lag model.

APPENDIX C

WATER QUALITY ANALYSIS

Due to the limited accuracy of the TCEQ SWQMIS laboratory machines used in measuring these nutrients, there are predefined screening values put into place by the EPA and reported in the "2012 Guidance for Assessing and Reporting Surface Water Quality in Texas" (SWQM, 2012). This means that if there are values below the detection limit then the accuracy of the numeric value cannot be trusted. Because the laboratory machine cannot actually detect the value of the nutrient if the recorded value is below the detection limit, the current TCEQ protocol is to use half of the screening level value and use this as the official reported value. This processes is considered a simple substitution method and is used in many data analysis protocols (Helsel and Hirsch, 1992). The benefit of this methodology to obtain a more realistic value under the detection limit is that the maximum number of data points are retained and there is most likely neither an over or underestimate of collected water quality parameter sample value. The practice of halving the screening level is used frequently by TCEQ and accepted as an adequate practice. While it is acknowledged that there are more robust methods of analysis for this situation, this simple substitution method is the most efficient and is widely accepted.

Four years of water quality variables will be averaged in this analysis. The reason for this is to eliminate the seasonal and temporal anomalies that may occur. The four years being averaged are 2010-2013, which is the most recent complete dataset available. A portion of 2014 water quality data is available, but the entire year has some data gaps. The indicator utilized in this study is Total Phosphorus, wet method (TP) with a unique identification code of 00665 and a unit of measure: Mg/L as P.

The water quality parameters used in this study and obtained from SWQMIS also went a stringent cleaning process from Houston Advanced Research Center (HARC).

1. The initial dataset was pulled from the SWQM database onto an SQL Server.
2. Raw data from TCEQ, DSHS, and TPWD were filtered and added in by spatial identifiers.
3. The three datasets were merged into one, relevant fields were matched up, irrelevant fields were dropped, and a source of data field was added.
4. Duplicates were removed defined by the Station, Parameter Code, Value Date, Time, and Depth. All of the steps up to this point were conducted using SQL. All of the duplicates were whittled down to one value.
5. The Parameter Codes were queried to keep desired values and the dataset was imported into Access.
6. Then, the TCEQ stations were matched up and the stream types were classified into tidal stream, non-tidal stream, or estuary. The stations that fell on the fringe of being on the bay or in a watershed were manually updated to be put in the proper watershed. In locations where there were GBEP watersheds but the stations were designated as estuaries, (TCEQ stated it was an estuary and GBEP said it was a Tributary) the GBEP watersheds were removed and the TCEQ designations were the only ones used.

The shapefile that was used to join the water quality data to (by the unique identifier: Station_ID) is from the SWQMs database. The stations that are in this file are collected by TCEQ and partner agencies like the Clean River Program, USGS, TSSWCB, and TPWD. All of the unique stations contain latitude and longitudes that were then projected using the World Geodetic System 1984 datum (WGS84).

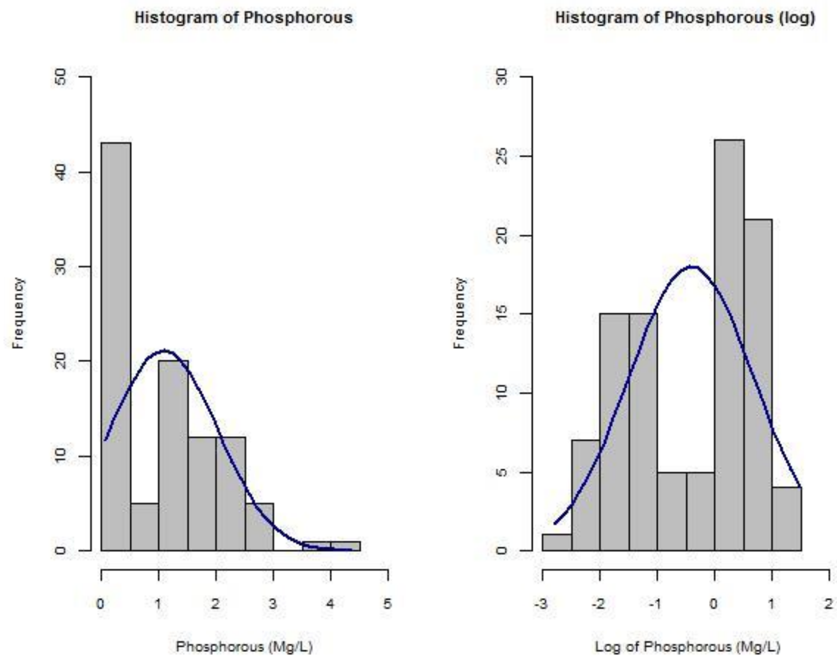
Once the cleaned data was obtained from HARC, some methods were undertaken to fit the data to the purpose of this study.

1. Firstly, all entries for years other then 2010-2013 were removed, leaving 20734 entries.
2. Then, the specific parameters that were necessary for this study were selected for (total phosphorus and ammonia). It should be noted that ammonia has not been included in the analysis of the study but was accounted for in the study.
3. All of the stations that had less than 20 samples for the 4 year cumulative period were removed so that the robustness of the nutrient averages over the 4 year period would be retained.
4. This file was then imported into ArcGIS where it was joined to the shapefile of stations projected in WGS84. All of the shapefiles were snapped to the flow accumulation raster that was obtained from NHDPlusV2 using a 90 meter (3 grid cells) snapping distance.
5. Watershed delineations were then run for the resulting 235 stations across the Galveston Bay Estuary Study Area.
6. Then, all watersheds above 110 square miles and below 10 square miles were removed to eliminate too much variation in watershed size and the resulting study sample was 99 watersheds.

The median of watersheds area within the study area is 34 square miles, which is equivalent to the average size of HUC12 watersheds (based on the National Hydrography Dataset Hydrologic Unit Codes). As stated above due to the need for spatial variation in watershed size the watersheds between 110 square miles and 10 square miles were selected for and the resulting dataset includes 99 water quality monitoring gauge points with adequate data and area.

Four years of water quality variables will be averaged in this analysis. The reason for this is to eliminate the seasonal and temporal anomalies that may occur. The four years being averaged are 2010-2013, which is the most recent complete dataset available. A portion of 2014 water quality data is available, but the entire year has some data gaps.

The total phosphorus levels are positively skewed which means that a log transformation was necessary to account for the skewness (figure C-1).



C- 1: Histogram of phosphorous and log of phosphorus (dependent variable)

APPENDIX D

LAND COVER RECLASSIFICATION SCHEME

The initial step in this process was to take NOAA-C-CAP land cover data for 2010 and reclassify it into the desired land covers in this study.

D- 1: Reclassification of NOAA-C-CAP land cover.

Reclassified Land Use Classification	***NOAA-CCAP Classifications***	***Value***
HID	High Intensity Development	3
MID	Medium Intensity Development	4
LID	Low Intensity Development	5
Forest	Deciduous, evergreen, mixed forest	10
Cultivated Crops	Cultivated crops	7
Pasture/Hay	Pasture and Hay	8
Grassland	Grassland/Herbaceous	9
Wetlands	Palustrine and estuarine forested wetlands, scrub/shrub wetlands, emergent wetlands	12
Other	Unclassified, developed open space, scrub/shrub, unconsolidated shore, bare land, open water, palustrine/estuarine aquatic bed	6, 11, 13, 14, 15

The patch metrics that are focused upon in this study are proximity of development patch to river, average area, contiguity, patch number, patch density, largest patch index, and percent of like adjacencies. All of these class metrics aside from proximity of development patch to river are calculated in FRAGSTATS. Proximity of development patch to river is calculated in ArcMap10.2 using the near table tool. The literature review show a visual display of each class metric.

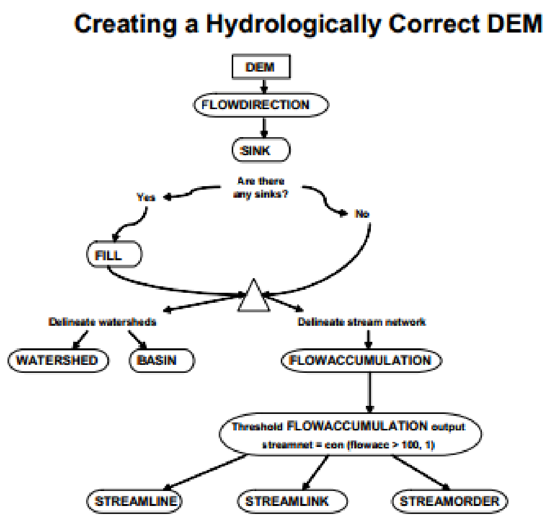
APPENDIX E

WATERSHED DELINEATIONS

Watershed delineations are used when point data has been collected and want to be used in a watershed analysis. The concept of watershed delineations is to create a boundary representing the area that contributes to the single point (based on the topography and hydrography of the area in question) (EPA, 2013). The watershed delineation approach that is used in this study is a Digital Elevation Model (DEM) based approach. The boundaries of the watershed are created automatically by the ArcMap10.2 software.

The DEM used for the watershed delineation is acquired from NHDPlusV2. The DEM used in this study is a combination of NHDPlusV2 DEMs. NHDPlusV2 is a dataset that incorporates data from both the NHD, National Elevation Dataset (NED), and the national Watershed Boundary Dataset (WBD) (McKay et al., 2012). The benefit of using the DEM from this source is that the data has been processed previously and does not contain any topological errors. There are many errors that can occur in a DEM; however this USGS and EPA DEM has not only filled the errors in not only the topology but uses burn components for hydro-enforcement. Using a DEM with the burn-in component can help solve some of the problems that come with inaccuracies of elevation data (EPA, 2013). The way that burn-in components solves these problems is by defining the location of the stream network by force.

There is a strict process that is followed to hydrologically correct DEMs. The reason that this process is so important is because these DEMs are used to generate flow direction and flow accumulation rasters. All of the processes to correct hydrologically correct DEMs are followed in this dataset (figure E-1). The DEM from NHDPlusV2 is on a 30 x 30 meter grid recorded in centimeters. From the NHDPlusV2 DEM the flow direction grid is created which then allows the flow accumulation and catchment grids are generated, both of which important components in delineating watersheds (NHDPlusV2, 2015).

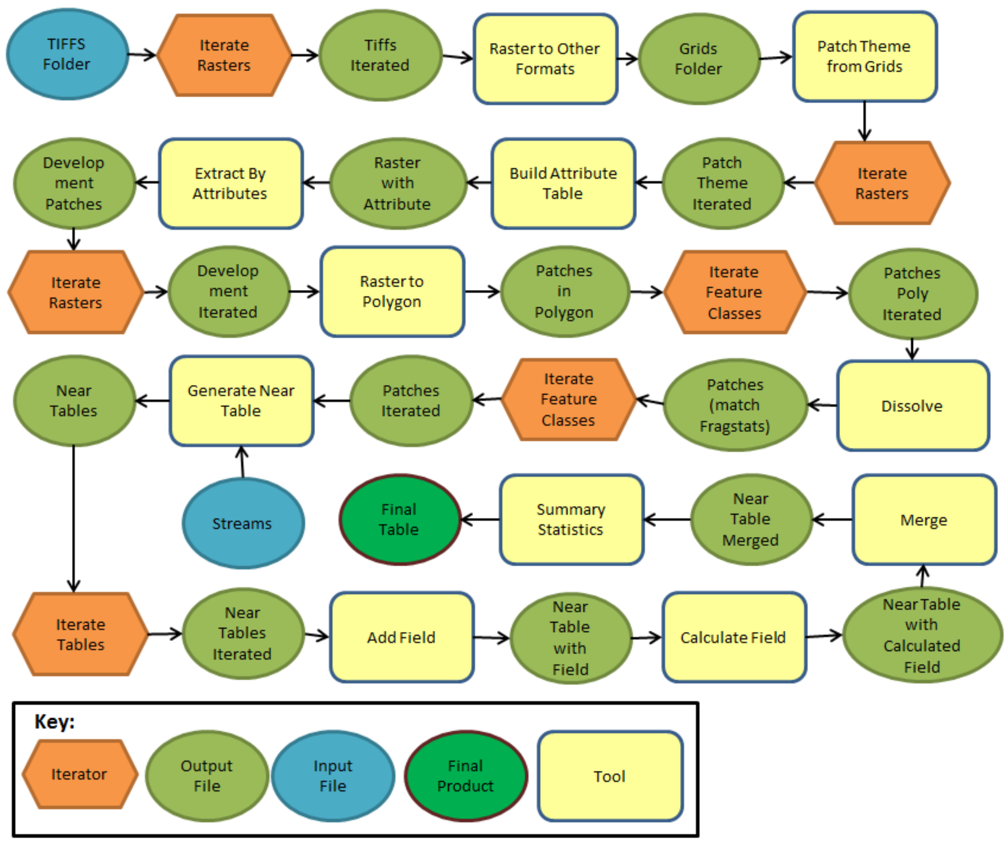


E- 1: Protocol to create hydrologically correct DEM (Gritzner, 2006).

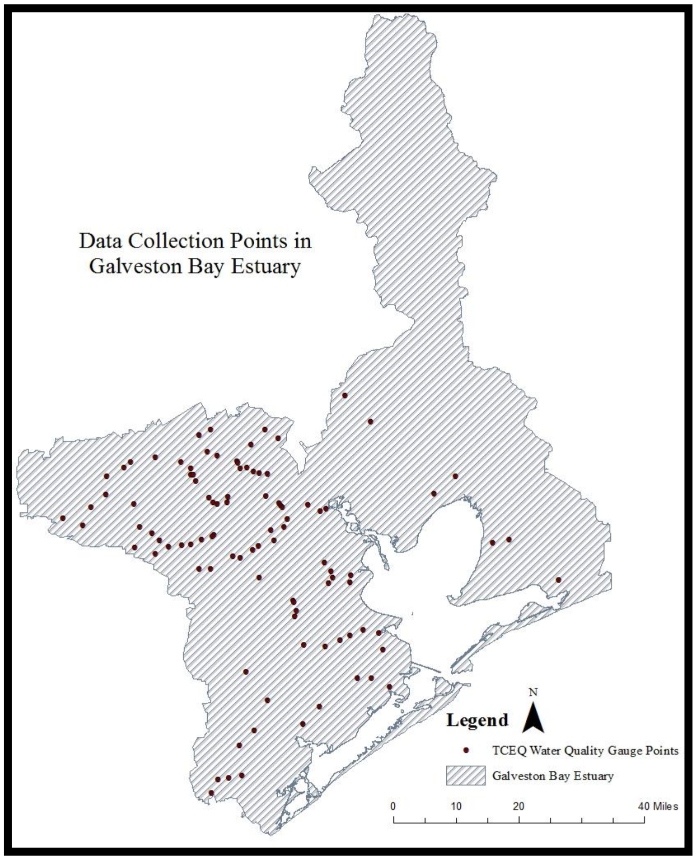
APPENDIX F

PROXIMITY CALCULATIONS

Proximity to nearest stream/river was calculated using ArcMAP10.2 figure F-1 shows the flow chart by which the proximity was calculated. Figure F-2 is a map showing where the locations are that are used for the watershed delineations.



F- 1: Average proximity to development calculation conducted in ArcMap10.2



F- 2: Water quality gauge points used for watershed delineations.

APPENDIX G

CALCULATION EQUATIONS

The formula for the low intensity development average patch size patch area model is shown in Equation 1. The calculations for how phosphorus will differ based on changes in average patch size areas in a specific watershed are based on this calculation.

Equation 1 (Patch Size):

$$\begin{aligned} y = 8.187 + {} & 0.5574(WeightLogPhos) + 0.0135(Crops) + 0.0392(Forest) \\ & - 0.0428(HID) - 0.01129(Wetland) + 0.0698(LID) \\ & + 0.0010(Area) - 0.1696(Precipitation) \\ & + 0.00026(Contributing\ Drainage) - 0.000135(OSSF\ Count) \\ & - 0.956(Average\ Patch\ Area\ LID) \end{aligned}$$

The formula for the low intensity development average patch density model is shown in Equation 2. The calculations for how phosphorus will differ based on changes in average patch density in a specific watershed are based on Equation 2.

Equation 2 (Patch Density):

$$\begin{aligned} y = 5.633 + {} & 0.4784(WeightLogPhos) + 0.0420(Crops) + 0.0538(Forest) \\ & - 0.0432(HID) + 0.0033(Wetland) + 0.0314(LID) \\ & + 0.00153(Area) - 0.1558(Precipitation) \\ & - 0.00063(Contributing\ Drainage) - 0.00008(OSSF) \\ & + 0.0656(Patch\ Density\ LID) \end{aligned}$$

The formula for the low intensity development average percent of like adjacencies model is shown in Equation 3. The calculation for how phosphorus levels differ based on the changes in the Equation 3 model are shown in Table G-1.

Equation 3 (Percent of Like Adjacencies):

$$
\begin{aligned}
y = 9.5519 + {} & 0.5634(WeightLogPhos) + 0.0096(Crops) + 0.0414(Forest) \\
& - 0.0426(HID) + 0.0137(Wetland) + 0.0592(LID) \\
& + 0.00114(Area) - 0.174(Precipitation) \\
& + 0.00036(Contributing\ Drainage) - 0.00013(OSSF) \\
& - 0.036(Patch\ Density\ LID)
\end{aligned}
$$

G-1 Changes of phosphorus levels with increasing low intensity development patch size, patch density, and average percent of like adjacencies by different amounts.

Watershed 18697	**Value of model**	**Change of Phosphorus relative to normal**	**Percent Change relative to normal**
Patch Area (1.46 normal)	-2.04		
Patch Area + .25 ha	-2.28	-0.24	12%
Patch Area + .5 ha	-2.52	-0.48	23%
Patch Area + .75 ha	-2.76	-0.72	35%
Patch density (7.3 normal)	-1.46		
Patch density + 1 patch/ha	-1.40	0.07	-4%
Patch density + 2 patch/ha	-1.33	0.13	-9%
Patch density + 3 patch/ha	-1.27	0.20	-13%
PLADJ (55.8 normal)	-1.75		
PLADJ+1%	-1.79	-0.04	2%
PLADJ+2%	-1.83	-0.07	4%
PLADJ+3%	-1.86	-0.11	6%
Watershed 15864	**Value of model**	**Change of Phosphorus relative to normal**	**Percent Change relative to normal**
Patch Area (1.214 currently)	0.41		
Patch Area + .25 ha	0.17	-0.24	-58%
Patch Area + .5 ha	-0.07	-0.48	-116%
Patch Area + .75 ha	-0.30	-0.72	-174%
Patch density (23.3 normal)	0.35		
Patch density + 1 patch/ha	0.41	0.07	19%
Patch density + 2 patch/ha	0.48	0.13	38%
Patch density + 3 patch/ha	0.55	0.20	56%
PLADJ (52.5 normal)	0.56		
PLADJ+1%	0.52	-0.04	-6%
PLADJ+2%	0.49	-0.07	-13%
PLADJ+3%	0.45	-0.11	-19%

APPENDIX H

CALCULATION FOR CHANGE IN APPENDIX G

The calculations were done by coefficient*value for the watershed (for every variable). The result was then the calculated phosphorus levels within the watershed. Then, the percent change was calculated and reported in the paper. An example of this table is shown in Appendix H-1 figure.

	Patch Density	Constant	WeightP	Crops	Forest	HID	Wetland	LID	Area	Precipitation	Contributing Drainag	OSSF Cou	Patch Density LID	Total	Change	% change
	coef	5.6325	0.4784	0.042006	0.05378	-0.04316	0.00329	0.03141	0.00153	-0.1558	-0.000625	-8.00E-05	0.06559			
	value		0.2441044	0.0073	9.7321	10.866	3.3782	28.2919	122.677	50.701399	122.677	1181	23.299101			
	coef*value	5.6325	0.1167795	0.0003066	0.5233923	-0.4689766	0.0111143	0.8886486	0.1876958	-7.899277964	-0.076673125	-0.0945272	1.528188035	**0.34917**		
	Patch Density	**Constant**	**WeightP**	**Crops**	**Forest**	**HID**	**Wetland**	**LID**	**Area**	**Precipitation**	**Contributing Drainag**	**OSSF Cou**	**Patch Density LID + 1**	**Total**	*Change*	
	coef	5.6325	0.4784	0.042006	0.05378	-0.04316	0.00329	0.03141	0.00153	-0.1558	-0.000625	-8.00E-05	0.06559			
	value		0.2441044	0.0073	9.7321	10.866	3.3782	28.2919	122.677	50.701399	122.677	1181	24.299101			
	coef*value	5.6325	0.1167795	0.0003066	0.5233923	-0.4689766	0.0111143	0.8886486	0.1876958	-7.899277964	-0.076673125	-0.0945272	1.593778035	**0.41476**	**0.06559**	**19%**
15864	**Patch Density**	**Constant**	**WeightP**	**Crops**	**Forest**	**HID**	**Wetland**	**LID**	**Area**	**Precipitation**	**Contributing Drainag**	**OSSF Cou**	**Patch Density LID + 2**	**Total**	*Change*	
	coef	5.6325	0.4784	0.042006	0.05378	-0.04316	0.00329	0.03141	0.00153	-0.1558	-0.000625	-8.00E-05	0.06559			
	value		0.2441044	0.0073	9.7321	10.866	3.3782	28.2919	122.677	50.701399	122.677	1181	25.299101			
	coef*value	5.6325	0.1167795	0.0003066	0.5233923	-0.4689766	0.0111143	0.8886486	0.1876958	-7.899277964	-0.076673125	-0.0945272	1.659368035	**0.48035**	**0.13118**	**38%**
	Patch Density	**Constant**	**WeightP**	**Crops**	**Forest**	**HID**	**Wetland**	**LID**	**Area**	**Precipitation**	**Contributing Drainag**	**OSSF Cou**	**Patch Density LID + 3**	**Total**	*Change*	
	coef	5.6325	0.4784	0.042006	0.05378	-0.04316	0.00329	0.03141	0.00153	-0.1558	-0.000625	-8.00E-05	0.06559			
	value		0.2441044	0.0073	9.7321	10.866	3.3782	28.2919	122.677	50.701399	122.677	1181	26.299101			
	coef*value	5.6325	0.1167795	0.0003066	0.5233923	-0.4689766	0.0111143	0.8886486	0.1876958	-7.899277964	-0.076673125	-0.0945272	1.724958035	**0.54594**	**0.19677**	**56%**

	Patch Density	Constant	WeightP	Crops	Forest	HID	Wetland	LID	Area	Precipitation	Contributing Drainag	OSSF Cou	Patch Density LID	Total	Change	% change
	coef	5.6325	0.4784	0.042006	0.05378	-0.04316	0.00329	0.03141	0.00153	-0.1558	-0.000625	-8.00E-05	0.06559			
	value		-0.6613451	18.8341	5.9476	0.9441	14.8903	10.6632	38.125198	56	38.125198	272	7.3165			
	coef*value	5.6325	-0.3163875	0.7911452	0.3198619	-0.0407474	0.0489891	0.3349311	0.0583316	-8.7248	-0.023828249	-0.0217709	0.479889235	**-1.46189**		
	Patch Density	**Constant**	**WeightP**	**Crops**	**Forest**	**HID**	**Wetland**	**LID**	**Area**	**Precipitation**	**Contributing Drainag**	**OSSF Cou**	**Patch Density LID + 1**	**Total**	*Change*	
	coef	5.6325	0.4784	0.042006	0.05378	-0.04316	0.00329	0.03141	0.00153	-0.1558	-0.000625	-8.00E-05	0.06559			
	value		-0.6613451	18.8341	5.9476	0.9441	14.8903	10.6632	38.125198	56	38.125198	272	8.3165			
	coef*value	5.6325	-0.3163875	0.7911452	0.3198619	-0.0407474	0.0489891	0.3349311	0.0583316	-8.7248	-0.023828249	-0.0217709	0.545479235	**-1.3963**	**0.06559**	**-4%**
18697	**Patch Density**	**Constant**	**WeightP**	**Crops**	**Forest**	**HID**	**Wetland**	**LID**	**Area**	**Precipitation**	**Contributing Drainag**	**OSSF Cou**	**Patch Density LID + 2**	**Total**	*Change*	
	coef	5.6325	0.4784	0.042006	0.05378	-0.04316	0.00329	0.03141	0.00153	-0.1558	-0.000625	-8.00E-05	0.06559			
	value		-0.6613451	18.8341	5.9476	0.9441	14.8903	10.6632	38.125198	56	38.125198	272	9.3165			
	coef*value	5.6325	-0.3163875	0.7911452	0.3198619	-0.0407474	0.0489891	0.3349311	0.0583316	-8.7248	-0.023828249	-0.0217709	0.611069235	**-1.33071**	**0.13118**	**-9%**
	Patch Density	**Constant**	**WeightP**	**Crops**	**Forest**	**HID**	**Wetland**	**LID**	**Area**	**Precipitation**	**Contributing Drainag**	**OSSF Cou**	**Patch Density LID + 3**	**Total**	*Change*	
	coef	5.6325	0.4784	0.042006	0.05378	-0.04316	0.00329	0.03141	0.00153	-0.1558	-0.000625	-8.00E-05	0.06559			
	value		-0.6613451	18.8341	5.9476	0.9441	14.8903	10.6632	38.125198	56	38.125198	272	10.3165			
	coef*value	5.6325	-0.3163875	0.7911452	0.3198619	-0.0407474	0.0489891	0.3349311	0.0583316	-8.7248	-0.023828249	-0.0217709	0.676659235	**-1.26512**	**0.19677**	**-13%**

H-1: Example of watershed specific calculation.