A SPATIAL ASSESSMENT OF GREEN INFRASTRUCTURE AND ITS POTENTIAL EFFECTIVENESS ON STREAMFLOW

A Dissertation

by

HYE KYUNG LEE

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

DOCTOR OF PHILOSOPHY

Chair of Committee,	Walter Gillis Peacock
Co-Chair of Committee,	Samuel D. Brody
Committee Members,	Ming-Han Li
	Wesley E. Highfield
	Francisco Olivera
Head of Department,	Forster Ndubisi

December 2015

Major Subject: Urban and Regional Science

ABSTRACT

As rapid urban development with its concomitant conversion of open space and added imperviousness increases the potential for adverse hazard impacts, the possible benefits of green infrastructure as a tool for hazard mitigation has become an emerging topic in landscape architecture and urban planning. However, research on this topic has been limited by the lack of effective tools for the identification and measurement of the specific dimensions of green infrastructure and the balance between green infrastructure and urban development particularly within highly developed urban environments. Consequently, there has been little empirical research conducted on the potential benefits of green infrastructure for reducing streamflow, an indicator of runoff and potential flooding mitigation.

This study seeks to further research green infrastructure as a potential tool for hazard mitigation by examining its consequences for streamflow over a 2-year period in 2004 and 2010 for two key urban areas subject to flooding in Texas (the Austin and Houston Metropolitan areas) by 1) utilizing high resolution (1-meter) imagery to develop fine resolution assessments of the amount, form, type and placement of green infrastructure in dense urban environments and then 2) utilizing these measures in panel models to assess the effectiveness of green infrastructure for reducing runoff as assessed by using streamflow gauge data predicting annual peak flow and mean flow.

More specifically, this study first identified an approach to employ the National Agriculture Imagery Program (NAIP) and Normalized Difference Vegetation Index

(NDVI) to identify and develop high-resolution measures of green infrastructure particularly germane for assessments within dense urban environments. Second, the statistical models developed utilizing these new measure of the extent and spatial patterns of green infrastructure suggested that green infrastructure indeed has consequences for streamflow reduction, particularly with respect to annual peak flow, in urban watersheds. Moreover, the analyses explained that green infrastructure in the Austin metropolitan area appears to be more effective on peak annual flow when compared to the Houston metropolitan area, suggesting that green infrastructure has elevated consequences in areas with greater topographical diversity. These variations perhaps imply that depending on different geographical characteristics, diverse guidelines for green infrastructure implementation should be applied.

The effectiveness of green infrastructure in critical places will help make a guideline for the balanced urban development with implementation of green infrastructure. This dissertation shows the utility of the new data for developing high-resolution measures of green infrastructure as a different approach compared to the conventional approaches. The consequences of green infrastructure for streamflow and potential flooding were clearly suggested. Also, this study begins to provide data that may well be used to establish guidelines for green infrastructure and effective runoff mitigation, and provides support for utilizing these data to guide research into green infrastructure spatial characteristics and hazard mitigation. Overall, the outcomes of this study will be helpful for the strategic planning and implementation of green infrastructure with streamflow issues, thus building community resilience.

DEDICATION

To my son, Minjoon

To my husband, Hwanyong,

and

To my family

ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Peacock, for his constant guidance and support, especially across the Pacific Ocean. I would also like to thank my co-chair, Dr. Brody, for his constant advice and encouragement. You both have been great chairs and mentors. I would also like to thank Dr. Highfield who accepted my request to be my committee member right before my proposal defense and answered my questions. Thanks also go to Dr. Li for advising me to think about a way to apply this study to design. I would also like to thank Dr. Olivera, who offered his hydrologic expertise for this study.

I would like to thank my family for their endless encouragement, support, and sacrifice throughout the course of this study. Especially, many thanks go to Joohae for taking good care of Minjoon.

Finally, I would like to thank my husband, Hwanyong for supporting me during our long academic years in Texas. Without his constant encouragement and motivation, I could not complete this dissertation.

NOMENCLATURE

CAPCOG	Capital Area Council of Government
C-CAP	Coastal Change Analysis Program
CRS	Community Rating System
DEM	Digital Elevation Model
FEMA	Federal Emergency Management Agency
GI	Green Infrastructure
GYRATE_AM	Radius of Gyration
HGAC	Houston-Galveston Area Council
LULC	Land Use Land Cover
NAIP	National Agriculture Imagery Program
NASA	National Aeronautics and Space Administration
NDVI	Normalized Difference Vegetation Index
NFIP	National Flood Insurance Program
NHD	National Hydrography Dataset
NIR	Near-infrared
NLCD	National Land Cover Database
NP	Number of Patches
PD	Patch Density
PERMH	Maximum Permeability
PERML	Minimum Permeability

STATSGO	State Soil Geographic Database
TNRIS	Texas Natural Resources Information System
USDA	U.S. Department of Agriculture
US EPA	U.S. Environmental Protection Agency
USGS	U.S. Geological Survey
VIF	Variance Inflation Factor

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 Background

Rapid urban development with its concomitant conversion of open space and added imperviousness increases the potential for stormwater runoff and flooding, resulting in increases in potential for adverse hazard impacts such as economic damage (Brody, Gunn, Peacock, & Highfield, 2011). Given these circumstances, the benefit of green infrastructure has become an emerging topic of study, and its potential use as a hazard mitigation tool is receiving more attention in the urban planning discipline. However, there are few studies that have directly assessed specific dimensions of green infrastructure and the balance between green infrastructure and urban development for enhancing community resilience. Moreover, little empirical research has been conducted on the potential benefits of green infrastructure to reduce streamflow.

1.2 Research Statement

Since green infrastructure is being touted as a potential hazard mitigation tool, it is important to study the relationship between green infrastructure and streamflow. Green infrastructure has been studied primarily at the community scale, utilizing very coarse measurements. At the site scale, green infrastructure has been applied with very site-specific details and characteristics. What has been missing is a more broadly based assessment employing refined measures that will allow for the examination of the particular forms and place-specific integration of green infrastructure within a community (Young, 2011). For example, previous research has considered pervious surfaces, such as different types of wetlands and undeveloped land use/land cover, to examine the impact on flood losses (Brody, Blessing, Sebastian, & Bedient, 2014; Brody, Peacock, & Gunn, 2012) at the county level. Additionally, the adverse hazard impacts from rapid urban development have been studied, specifically the effects of urbanization on runoff and impacts of development patterns on flooding (Brody, Kim, & Gunn, 2013; Olivera & DeFee, 2007). However, the consequences of specific green infrastructure patterns and forms within urban areas for streamflow reduction have not been studied.

This study seeks to further the research on green infrastructure as a potential tool for hazard mitigation by examining its consequences for streamflow over a 2-year period in 2004 and 2010 for two key urban areas subject to flooding in Texas, the Austin and Houston Metropolitan areas by 1) utilizing high resolution (1-meter) imagery to develop fine resolution assessments of the amount, form, type and placement of green infrastructure in dense urban environments and then 2) utilizing these measures in panel models to assess the effectiveness of green infrastructure for reducing runoff as assessed using streamflow gauge data predicting annual peak flow and mean flow. More specifically, in order to understand its potential effectiveness for reducing streamflow in urban areas, a spatial assessment of green infrastructure at the community/watershed scale was developed. This study seeks to measure various dimensions, such as the continuity and connectivity of green infrastructure, in order to assess the degree to which these spatial patterns are related to streamflow measurements.

This dissertation is expected to show the utility of the new high-resolution data for developing high-resolution measures of green infrastructure. The results of this study are also expected to provide a better understanding of the effect of green infrastructure integration on the reduction of urban streamflow discharge and potentially flooding. This dissertation provides support for utilizing these data to guide research in green infrastructure spatial characteristics and hazard mitigation. Furthermore, I hope that the findings will provide additional decision support tools and options for promoting streamflow reduction and flood mitigation for urban planners, policy makers, and community residents as they evaluate existing green infrastructure in communities and make decisions regarding the implementation of green infrastructure to reduce streamflow and mitigate flood damages for increased community resilience.

1.3 Research Objectives and Question

The overall purpose of this study is to improve understanding of the spatial assessment of green infrastructure in dense urban environments and its potential effects on streamflow. The research question for this study is: w*hat is the relationship between the spatial patterns of green infrastructure and streamflow measurement at watershed scale?* The specific objectives of this study are:

- To identify the distributed amount of the green infrastructure that can potentially moderate the adverse impacts of streamflow discharge by utilizing high resolution (1-meter) imagery
- To develop measures of the form, structure, and strategic placement of green infrastructure at the community/watershed scale.
- To assess the consequences and effectiveness of these dimensions of green infrastructure for reducing streamflow.

1.4 Research Significance

This research is significant and can be justified for two reasons. As population and urban development increase rapidly in Texas, conversion of open space and added imperviousness are closely related to adverse hazard impacts, such as stormwater runoff and flooding, resulting in increases in potential economic damage.

Second, there is a lack of empirical research assessing specific dimensions of green infrastructure and the balance between green infrastructure and urban development. Also, little research has been conducted on the potential benefits of green infrastructure to reduce streamflow in urban areas.

1.5 Dissertation Structure

This dissertation consists of seven Sections. Section 1 explains this study's background, research statements, research objectives and question, and the significance of this research. Section 2 reviews previous research about green infrastructure that is related to the research objectives and question, and exposes a research gap based on the literature review.

Since this study utilizes a new method of green infrastructure measurement involving a 1-meter high resolution imagery, Section 3 is devoted to explaining how this new measurement is applied. This section also addresses the effectiveness of utilizing this new measurement for green infrastructure research by comparing its results to other datasets. Section 4 addresses the research methods of this study. In this section, a conceptual model for this study containing dependent, independent, and control variables is presented to further understand the effectiveness of green infrastructure on streamflow. Section 4 also describes the study area, sample selection, concept measurements, and data analysis.

Section 5 presents the results of this study explaining the overall effect of green infrastructure on annual peak flow and mean annual flow. Section 6 discusses the results of this study and provides policy implications based on these results. Section 7 summarizes the key findings of this dissertation and explains limitations and future research.

2. LITERATURE REVIEW

2.1 Green Infrastructure

Green infrastructure has its origin in planning and conservation theory and is concerned with the linkage between parks and other green spaces and their potential benefits for people. It is also rooted in conservation efforts that deal with habitat fragmentation and link natural areas with benefit biodiversity (Benedict & McMahon, 2002). For example, at the end of the 19th century, Frederick Law Olmsted and Ebenezer Howard presented the continuity of green belts and parkway as an important concept in the history of green structures in urban areas (Madureira, Andresen, & Monteiro, 2011). In 1921, Benton Mackaye proposed the development of the Appalachian Trail, a 2,100-mile-long recreational linkage running along the Appalachian Mountains from Maine to Georgia, to serve as a buffer against development from Eastern cities (Minteer, 2006). He advocated the need for recreation with the use of green space corridors that follow natural landforms (Benedict & McMahon, 2002).

The term “green infrastructure” as well as the implementation of its concepts and values, has been applied frequently since 1999 (Benedict & McMahon, 2002). At that time, the President’s Council on Sustainable Development identified green infrastructure as one of five strategies that facilitate more efficient and sustainable land use and development patterns. It was also identified as a key approach to protecting ecosystems (President's Council on Sustainable Development, 1999). In 1999, a working group of the Conservation Fund and the U.S. Department of Agriculture (USDA) Forest Service developed a program to make green infrastructure an integral part of local, regional, and state plans and policies. This group defined green infrastructure as an interconnected network of green space that maintains natural ecological processes, sustains air and water resources, and contributes to the health and quality of life in communities (Benedict & McMahon, 2002). In addition to this commonly accepted definition, green infrastructure has been diversely defined depending on the respective professional discipline or the scale at which green infrastructure is implemented (Allen, 2012). Considering rapid population growth and expanding residential areas, green infrastructure is distinguished from conventional open space planning because green infrastructure addresses issues related to land development, growth management, and built infrastructure planning (Allen, 2012).

Green infrastructure has also become an emerging topic as a hazard mitigation tool. Its use is receiving more attention in the urban planning discipline due to positive outcomes that facilitate more effective reduction of flood impact in rapidly expanding residential areas. Brody et al.(2011) found that increasing percentages of sprawling, low-intensity development exacerbate flood losses due to conversion of open space. Rapidly expanding residential development results in conversion of open space and added imperviousness, thus increasing adverse hazard impacts, such as stormwater runoff and flooding; in response to these developing issues, the benefits of green infrastructure are beginning to be studied (Brody et al., 2011). To reduce stormwater runoff and flood damage from rapid development, strategic management and placement of green infrastructure should be studied. It is important for rapidly expanding residential areas to investigate how green infrastructure at the community level might facilitate more effective reduction of streamflow and flood impacts.

Before looking at previous studies at community scale, it is necessary to define different scales of green infrastructure implementation. The United States Environmental Protection Agency (US EPA) acknowledges green infrastructure at three different scales - regional or watershed scale, community and neighborhood scale, and site scale. According to these respective categories, at regional or watershed scale, green infrastructure is defined as the interconnected network of preserved or restored natural lands and waters with ecological functions (Benedict & McMahon, 2002). At the community and neighborhood scale, green infrastructure includes planning strategies to implement compact, mixed-use development, and urban forestry in order to reduce imperviousness. At site scale, green infrastructure involves the mimicry of natural hydrology systems to manage stormwater runoff using, for example, low impact development strategies. My goal in this study is to link watershed and community level scales to examine the potential effects of green infrastructure on streamflow measurement.

2.2 Green Infrastructure and Structural/Nonstructural Mitigation

Green infrastructure is sometimes thought of merely as a catchy buzzword. Because the adjective “green” can give a false sense of sustainability or resiliency, the term “green infrastructure” has often been used without concrete definitions or referents.

In order to implement green infrastructure effectively, it is important to be clear on its meaning; this facilitates the opportunity to consider how green infrastructure is distinguished from and yet fits with other mitigation approaches, both structural and nonstructural approaches. It is also necessary to improve general understanding on what makes green infrastructure important for hazard and resiliency research. Because streamflow is closely related to flooding, green infrastructure and structural/nonstructural mitigation is focused on flooding mitigation.

2.2.1 Structural/Nonstructural Mitigation

To better understand how green infrastructure fits with more traditional notions of structural and nonstructural mitigation, I begin by examining what is specifically meant by "structural and nonstructural" mitigation approaches for flooding mitigation. Structural mitigation requires the use of engineered safety features for creating protection from disaster impacts (Lindell, Perry, Prater, & Nicholson, 2006). At the community level, structural approaches include levees, floodwalls, fills, dikes, dams, detention basins, reservoirs, and straightening or widening waterway channels. For example, Brody et al. (2008) considers the performance of the number of dams in each county in order to show that structural mitigation significantly reduces flood damage. Also, the role of retention and detention ponds as a form of structural mitigation is strongly recommended based on its positive effect in reducing runoff (Brody, Highfield, & Kang, 2011). However, regardless of the significant progress in reducing flooding losses through structural mitigation, several limitations have also been realized. Burby

(1998a) explains that structural protection provides only partial protection up to the limits of its design, resulting in potentially significantly higher damages when these design limits are exceeded. For example, massive flood damages behind levees in the upper Mississippi and Missouri River basins in 1993 were due to the failure of levees, structures which did not provide complete protection (Burby, 1998a). Such excessive losses are partially due to structural mitigation techniques providing a false sense of safety and security, thus encouraging new development in hazard prone areas. Additionally, structural mitigation can create adverse impacts to environment, has tremendous costs, and requires long time frames for construction and maintenance. (Brody, Bernhardt, Zahran, & Kang, 2009).

To overcome the limitations of structural approaches, nonstructural mitigation has been suggested as a potential solution. The main goal of nonstructural mitigation is to direct development away from hazard prone areas. Nonstructural approaches include land use planning tools, education and training, environmentally sensitive area protection, forecasting, and other emergency and recovery policies for flood mitigation. For example, the acquisition of undeveloped floodplains and their subsequent dedication to open space is a prime illustration of nonstructural mitigation (Lindell et al., 2006). In the 1960s, Gilbert White (Platt, 1998) suggested that enforcing requirements for land use controls and building standards to reduce future losses from flooding would make the National Flood Insurance Program (NFIP), the most widely implemented nonstructural flood mitigation, feasible. Since then, many other nonstructural flood mitigation approaches have been based on this federal program. However, the NFIP is not without its critics who claim that affordable flood insurance can act to encourage more development in flood prone areas.

2.2.2 Green Infrastructure and Nonstructural Mitigation

Looking at regional scale, there are similarities between green infrastructure and nonstructural mitigation approaches. Green infrastructure implementation at regional scale includes land use planning for major parks, strategic open space, development ordinances, and incentive programs. Specifically, strategic land acquisition and conservation easements, open space preservation policies in comprehensive plans, and incentives and regulations for protection of floodplains, wetlands, and other natural resources are important components of green infrastructure; they help keep development out of floodplains and they preserve wetland areas and the ecosystem services they provide to help reduce flooding (Allen, 2012). Here, then the “nonstructural” policy of preserving green infrastructure helps preserve the “structural” integrity of wetlands and allows them to play their important ecological functions, providing critical ecosystem services with their mitigation benefits. When directly compared with these two traditional mitigation approaches (structural and nonstructural), many aspects of green infrastructure overlap with nonstructural mitigation. For example, nonstructural mitigation and green infrastructure both emphasize the strategic redirection of intensive development away from hazard prone areas (Burby, 1998b; Burby & Dalton, 1994) and stress the importance of ecological functions to mitigate hazards. Therefore, when conditions are appropriate, nonstructural mitigation with green infrastructure should be emphasized for effective streamflow reduction and flood mitigation.

2.2.3 Green Infrastructure and Structural Mitigation

On the other hand, there are also similarities between green infrastructure and structural mitigation approaches when looking at green infrastructure implementation at site scale. At site scale, green infrastructure is defined as an implementation of features that simulate natural hydrologic systems to manage stormwater runoff. This site scale green infrastructure also includes low impact development such as rain gardens, porous pavement, and green roofs (United States Environmental Protection Agency, 2010). Indeed, there are cases in which wetlands are reengineered or recreated to help create or restore the ecosystem services that help reduce flooding. These low impact development strategies require some engineered features in order to reduce runoff, thus implementing structural mitigation. Also, looking at the regional scale, once nonstructural mitigation is applied to an area, some types of engineered features for creating green infrastructure (i.e. construction of a new park) are often necessary. This also means that, just as with structural mitigation, there are some high upfront costs for green infrastructure implementation such as costs for land acquisition (Beatley, 2009). Compared to structural mitigation involving the construction of engineered protection from flooding, green infrastructure promotes ecological functions to absorb, store, and release runoff.

2.2.4 Attributes of Green Infrastructure for Hazard Mitigation

One of the factors that makes green infrastructure important for hazard mitigation is that it emphasizes the use of ecological functions to mitigate flooding. In regards to uncertain disturbances such as hurricanes and flooding, natural ecosystems and green infrastructure are suggested as some of the most effective defenses against natural disaster (Beatley, 2009). Examples of such long-term hazard defenses include coastal marshes and wetlands that absorb floodwaters, dune and beach systems that act as natural seawalls, and trees and tree canopies that protect properties from wind. Also, green infrastructure can be implemented at different scales including regional, community, and site scale, providing more opportunities for application. Green infrastructure as a larger pattern of integrated networks of wetlands, forests, and green spaces implemented at various scales provides extensive ecological services to mitigate adverse impacts from hurricanes and coastal storms (Beatley, 2009). Community resilience to disasters can be improved by action at a number of design scales (specifically at site level, city level, and regional level), including utilizing land use planning to minimize development in high-hazard areas, as well as by the preservation of a green infrastructure (Beatley, 2011).

2.3 Major Techniques and Practices of Green Infrastructure

An avoidance strategy is one of the major techniques and practices of green infrastructure for flood mitigation. Brody and Highfield (2013) show that when controlling for environmental, socioeconomic, and policy-related variables, open space protection significantly decreases flood damage in floodplain areas. They statistically studied the relationship between the performance of open space dedicated for flood mitigation under FEMA’s Community Rating System (CRS) program and the reduced insured flood damages for 450 local communities over an eleven-year period from 1999-2009. Open space preservation in the 100-year floodplain is one of the 18 creditable mitigation activities of the CRS. In this way, local jurisdictions are encouraged to prepare for flood management because high CRS class ratings can provide residents with NFIP premium discounts of up to 45 percent. Their findings suggest that using protected, open space as a land use tool for flood mitigation is a promising policy at the community level. Recommendations from their study include the fee simple purchase of entire parcels, conservation easements, and overlay zones, as well as the implementation of incentives such as transfer of development rights, density bonuses, and special taxing districts. However, this study only considers the total points of open space preservation content that is creditable under the CRS program. A more detailed analysis of the characteristics of the open space, such as its position in the floodplain, size of the designation, integration with wetland protection, and surrounding land use, will enhance understanding as to how green infrastructure implementation of open space protection may contribute to flood-resilient communities (Brody & Highfield, 2013).

Pervious surfaces, such as setbacks from or buffers around riparian areas, provide holding capacity for streamflow. These linear, protected areas can be implemented as the horizontal equivalent of freeboard to direct development away from floodplain areas and reduce damages to people and structures (Brody & Highfield, 2013). Preserving naturally occurring wetlands is another major technique of green infrastructure for flood mitigation. Previous research about the impact of wetland alteration on flood damage reveals the importance of green infrastructure. Godschalk et al. (1999) show that only 13 million acres of wetlands, which is 3% of the upper Mississippi watershed, would have been needed to prevent the catastrophic flood of 1993. More recent studies explain that alteration or elimination of naturally occurring wetlands due to rapid urbanization increases run off volume, peak discharges, and associated flood magnitudes (Brody et al., 2008). Their results suggest that wetland alterations, located mostly within the 100-year floodplain, can have significant effects on the amount of property damage resulting from a flood. This study emphasizes the importance of ecological functions to absorb, store, and release of wetlands. Conducting a cross- sectional time series regression analysis for the relationship between peak annual streamflow and wetland alteration emphasizes the importance of wetlands to decrease peak annual flow (Highfield, 2012).

Spatial patterns of green infrastructure such as connectivity and continuity are also important techniques for flood mitigation. Previous research shows that connectivity and compactness of the urban development is related to runoff conveyance (Olivera & DeFee, 2007). Compared to urban development patterns, consideration of spatial patterns of green infrastructure at community level is important to reduce streamflow. Analysis using landscape metrics, such as number of patches and correlation length for spatial patterns of green infrastructure, will help increase understanding of how green infrastructure relates to streamflow. The proximity of green infrastructure to waterways can be another indicator for the analysis. Shandas and Alberti (2009) explored vegetation volume and distribution as a way to mitigate the impact of urban development on stream systems. They examined the role of watershed vegetation patterns to explain variations in aquatic conditions in the Puget Sound lowland in Washington State. This study considered 100-meter buffers adjacent to the stream channel for riparian zones, using the 'buffering' distance recommended by the Washington State Department of Ecology. Findings suggest a strong relationship between the amount of riparian and watershed vegetation and instream biological conditions. Results of this study also recommended further research on the role of vegetation fragmentation, the combination of the amount of riparian vegetation, and the contiguity of upland vegetation. They suggested that the manipulation of vegetation by connecting fragmented patches of forest may help to regulate runoff frequency, volumes, and peak flow rates across the whole watershed (Shandas and Alberti, 2009). In Norway, Syversen (2005) examined several design criteria that influence the effect of surface runoff buffer zone, including buffer zone width, amount of surface water runoff into the buffer zone, seasonal variation, and vegetation type. One of the results was a significantly higher runoff removal efficiency from a 10 meter wide buffer zone, as compared to a 5 meter wide buffer zone. The results show no significant differences between forest buffer zones and grass buffer zones regarding their retention efficiency for nitrogen and phosphorus.

2.4 Normalized Difference Vegetation Index (NDVI) to Capture Green Infrastructure

In order to study the relationship between green infrastructure and streamflow, it is important to appropriately measure green infrastructure. Previous empirical research measures green infrastructure using different types of datasets, such as the National Land Cover database or the Coastal Change Analysis Program (C-CAP) Land Use Land Cover (LULC) Data (Brody et al., 2014). However, using 30 meter LULC data is not precise enough to capture detailed green infrastructure, especially in urban areas. These datasets sometimes overlook green infrastructure that is not captured due to the coarse resolution, and they do not consider green areas besides those in green-related categories. For example, green areas in high-intensity developed LULC type are not considered for the analysis of the relationship between LULC type and streamflow measurement. As different land uses have different degrees of development, variation in green infrastructure, such as canopy cover, exists across land use types (Hill, Dorfman, & Kramer, 2010). Zoning ordinances for regulating land use include tree canopy cover in some municipalities. For instance, Chesapeake, Virginia required through its zoning ordinance the maintenance of 10% canopy cover on parcels in non-residential zones, 15% canopy cover in multi-family residential zones, and 20% canopy cover in single-family residential zones (Hartel, 2003). In order to measure these variations in green infrastructure across different land uses, an empirical study with more detailed green infrastructure measurement should be followed. Further explanation to capture distributed green infrastructure in detail will be addressed in the next Section.

2.5 Landscape Metrics to Measure Green Infrastructure Spatial Patterns

Landscape metrics quantify specific spatial characteristics of individual patches and the spatial relationship among multiple patches (Gustafson, 1998). Landscape metrics have emerged as an important method of quantifying landscape patterns in order to gain a better understanding of the relationships and changes in landscapes through time and space (Park, Hepcan, Hepcan, & Cook, 2014). Two fundamental aspects of landscape structure are composition and configuration (Leitão et al., 2006). Composition refers to the quantification of the variety and abundance of patch types within the landscape; it explains the number, type, and extent of landscape elements without explicit consideration to their spatial distribution. Configuration refers to the spatial character, arrangement, position, or orientation of landscape elements (Leitão et al., 2006). For example, landscape composition measures the number of patch types (i.e., patch richness), the proportional abundance of each patch type (i.e., class area proportion), and the overall diversity of patch types (e.g., Shannon's and Simpson's diversity indices). On the other hand, landscape configuration measures patch shape and compactness, as well as the distance between patches of the same class (i.e., nearest neighbor distance) (Leitão et al., 2006).

Several researchers have analyzed green infrastructure patterns by using different sets of landscape metrics. Fernandes et al. (2011) used landscape metrics to measure the spatial configuration, isolation, inter-connectivity, and distribution of patches of three riparian classes (trees, shrubs, and herbaceous), assessing riparian vegetation structure and the influence of land use (Fernandes, Aguiar, & Ferreira, 2011). Shandas and Alberti (2009) quantified the total amount and fragmentation of vegetation in the riparian zones and watersheds by using landscape metrics. “Percent land” as the composition metric is used to explain the amount of vegetation relative to other land cover categories, whereas, “Aggregation index (AI)” is used to describe the fragmentation of vegetation (Shandas & Alberti, 2009). Hepcan (2013) analyzed the patterns and connectivity of urban green spaces in Turkey, measuring proportion of landscape, number of patches, mean patch size, and connectance index in order to quantify the patterns and the connectivity of green infrastructure. However, because landscape metrics has the major constraint of multicollinearity problems among different types of metrics, the resulting coefficient sign is sometimes counter-intuitive to the hypothesized relationship (Jones et al., 2001). Careful statistical model construct and interpretation are required to analyze the relationship between landscape metrics and streamflow.

Most landscape metrics for previous research have focused on the description and quantification of spatial patterns. There is a lack of empirical research on how these landscape patterns affect ecological processes (Giulio, Holderegger, &Tobias, 2009). As NDVI facilitates the detailed identification of distributed green infrastructure in urban areas and the differentiation of various types of green infrastructure, spatial patterns of these different types of green infrastructure will test their effects on streamflow based on landscape metrics. Effects of green infrastructure in different spatial locations can also be analyzed using landscape metrics, and this study will empirically analyze the relationship between green infrastructure patterns and their effect on streamflow.

2.6 Summary

Green infrastructure has been defined diversely depending on professional disciplines or the scale at which green infrastructure is implemented. Considering rapid population growth and the expansion of residential areas, green infrastructure should be distinguished from conventional open space planning. There are few studies that have directly assessed specific dimensions of green infrastructure, specifically the balance between green infrastructure and urban development for enhancing community resilience. Moreover, little empirical research has been conducted on the relationship between streamflow and specific green infrastructure patterns and forms within urban areas.

3. MEASURING GREEN INFRASTRUCTURE USING HIGH RESOLUTION NATIONAL AGRICULTURE IMAGERY PROGRAM (NAIP) DATA

For this study, I utilized here-to-for unutilized data to assess green infrastructure at an exceptionally high resolution in urban areas. As explained in Section 2, conventional approaches employed the National Land Cover Database (NLCD), Coastal Change Analysis Program (C-CAP) Land Cover data, which are gathered at a 30 by 30 meter resolution. However, this study utilized satellite imagery data produced by the National Agriculture Imagery Program (NAIP) to compute a Normalized Difference Vegetation Index (NDVI) providing 1 by 1 meter resolution data. Considering the importance of the new measurement, this section is devoted to discuss new green infrastructure measurement using high resolution imagery and compare this new measurement to the conventional approach. As key variables in this study were based on utilizing green infrastructure measurement, it is better to explain further about this new measurement, which provides highly detailed information of green infrastructure, especially for urban environments.

3.1 National Agriculture Imagery Program (NAIP)

According to the United States Department of Agriculture (USDA), the purpose of the National Agriculture Imagery Program (NAIP) is to acquire aerial imagery during the peak growing season at a one-meter resolution in the continental United States. Since

2002, the year NAIP pilot projects started, and 2003, the year NAIP officially began to acquire imagery, NAIP has continued to grow. The default spectral resolution is natural color – red, green and blue. However, since 2007, some states have been provided with four bands of data containing red, green, blue, and near infrared bands. Beginning in 2004, the state of Texas has acquired 1-meter resolution color infrared NAIP. The second phase in Texas for acquiring 1-meter resolution color infrared imagery was 2010.

3.2 Normalized Difference Vegetation Index (NDVI)

NDVI is an index to generate an image of vegetation distribution. This index is derived from the red and near-infrared reflectance ratio of the difference between near-infrared and red reflectance to the sum of near-infrared (NIR) and red reflectance. It is calculated using the formula:

$$NDVI = (NIR - Red) / (NIR + Red).$$

The resulting NDVI values range from -1 to 1. This calculation is based on the fact that chlorophyll absorbs red, whereas the mesophyll leaf structure scatters near-infrared (Pettorelli et al., 2005). Positive NDVI values indicate green or vegetated surfaces, whereas negative NDVI values represent non-vegetated surfaces such as clouds, water, and snow. According to the National Aeronautics and Space Administration (NASA), very low values of NDVI (0.1 and below) correspond to barren areas of rock, sand, or snow. Moderate values represent shrub and grasslands (0.2 to

0.3), while high values indicate temperate and tropical rainforests (0.6 to 0.8). Using NDVI, Holben (1986) quantified several levels of green vegetation – dense green-leaf, medium green-leaf, and light green-leaf vegetation – and separated them from other components such as clouds, bare solid, rocks, and surface water. This separation is done as a function of the NDVI by stratification of cover classes.

Depending on NDVI values, it is also possible to differentiate land cover types and ecosystem functional types such as forest, trees, or shrubs (Pettorelli et al., 2005). This ability to differentiate among these different functional types is important when studying the relationship between green infrastructure and streamflow measurement because different vegetation conditions indicate different infiltration capacity, which closely relates to the amounts of water captured at streamflow gage stations. Kays (1980) tested infiltration rate for various land types in Sudbury Watershed, Charlotte, NC. This study revealed that the medium aged pine-mixed hardwood forest had a mean final constant infiltration rate of 31.56 cm/hr, whereas when the forest understory and loaf litter was removed, the resultant residential lawns had a mean infiltration rate of 11.20 cm/hr (Kays, 1980). In addition to these two land types, Table 1 displays infiltration rates of other land types. Another study shows that average non-compacted infiltration rates range from 37.7 to 63.4 cm/hr for natural forest and from 63.7 to 65.2 cm/hr for planted forest, as well as 22.5 cm/hr for pasture sites (Gregory, Dukes, Jones, & Miller, 2006). The importance of these studies is that they make clear that our ability to identify different land cover types provide information on different infiltration rates, and subsequently different abilities to absorb rainfall, reducing runoff.

Land Type	Mean Final Constant Infiltration Rate (cm/hr)
Medium aged pine-mixed hardwood forest with leaf litter	31.56
Slightly disturbed soils with lawns and large trees preserved	11.20
Slightly disturbed soils, previously cultivated filed, lawns and few young trees	4.78
Slightly disturbed soils, previously cultivated field with plow pan, lawns and few trees	0.70
Highly disturbed fill soils, lawns and few young trees	1.25
Highly disturbed cut soils, lawns and few young trees	0.67
Highly disturbed cut and compacted soils, sparse grass, no trees	0.45

Table 1. Infiltration Rates by Land Type (Kays, 1980)

To analyze the effectiveness of green infrastructure on streamflow reduction, it is important to appropriately measure green infrastructure. Conventional approaches utilized the National Land Cover Database (NLCD), Coastal Change Analysis Program (C-CAP) Land Cover data, which are captured at a 30 by 30 meter resolution. However, using 30-meter LULC data is not precise enough to capture detailed green infrastructure, especially in densely developed urban areas. These datasets sometimes ignore green infrastructure that is overlooked due to the coarse resolution, and they do not consider green areas besides those in green-related categories. For example, green areas in high-intensity developed LULC type are not considered for the analysis of the relationship between LULC type and streamflow. As different land uses have different degrees of development, variation in green infrastructure, such as canopy cover, exists across land use types (Hill et al., 2010). In order to measure these variations in green infrastructure across different land uses, an empirical study based on higher resolution data what will provide more detailed green infrastructure measurement should be undertaken.

3.3 Comparison of NDVI Green Infrastructure New Measurement to Other Datasets

Compared to previous research using LULC datasets, the Normalized Difference Vegetation Index (NDVI) provides an ability to capture green infrastructure in greater detail. The National Agriculture Imagery Program (NAIP), provided by the United States Department of Agriculture (USDA) offers 1-meter high-resolution imagery for the continental United States, thus capturing significantly more detailed green infrastructure. Table 2 compares different types of dataset – NAIP, Land Use, NLCD Land Cover, C-CAP Land Cover, and NDVI – for a part of Harris County, Texas. It shows that NDVI using NAIP, 1-meter resolution imagery, displays more detailed green infrastructure information compared to other datasets.

Types	Images	Legends	Resolution
NAIP			1-meter
Land Use			Vector Data

Table 2. Comparison NDVI to Other Datasets

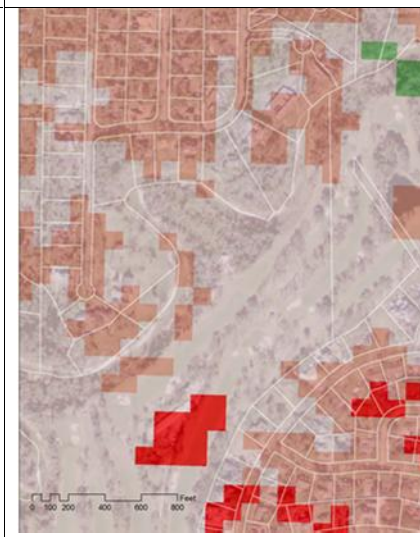
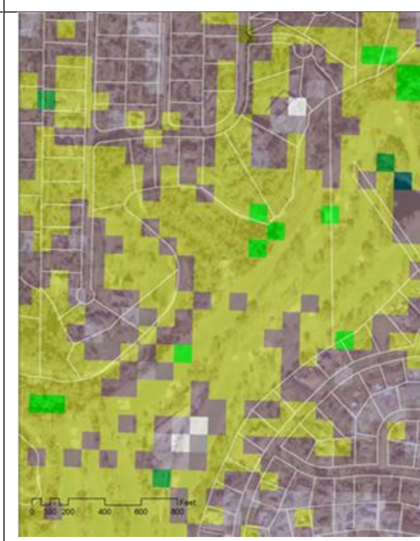
Types	Images	Legends	Resolution
NLCD Land Cover		Deciduous Forest Developed, Low Intensity Developed, Medium Intensity Developed, Open Space	30-meter
CCAP Land Cover		Deciduous Forest Developed, High Intensity Developed, Low Intensity Developed, Medium Intensity Developed, Open Space Mixed Forest Palustrine Forested Wetland Scrub/Shrub	30-meter

Table 2. Continued

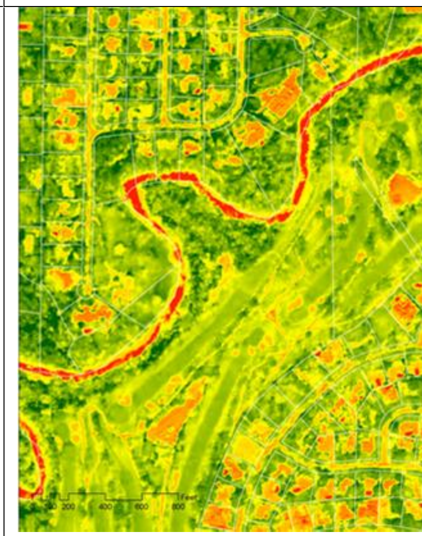
Types	Images	Legends	Resolution
NDVI			1-meter

Table 2. Continued

As can easily be seen by visually examining the above images, which are of the same location within Harris County, Texas, in comparison to the NDVI data image provided in the lower panel of the table, those of the CCAP and NLCD, not to mention the simple land use data, provides much higher resolution image. The NDVI data has an inherent ability to provide ability to differentiate not only different variations in the vegetation associated with different land-uses, but also variations in vegetation in smaller and refined areas of cover interspersed through-out this highly developed around this river/stream. This also enhanced my ability to assess patterns and connectivity among land uses.

To compare the effectiveness of using high-resolution aerial imagery for green infrastructure measurement, descriptive statistics were analyzed. The NLCD legend includes 20 classifications to measure green infrastructure as defined for this study; these classifications were sorted based on legend description. Table 3 shows sorted NLCD classifications for green infrastructure and non-green infrastructure. Among those 20 classifications, low, medium, and high intensity developed classes, open water, and barren land were all excluded for green infrastructure measurement. Focusing on classifications of relevance for green infrastructure classifications they included developed open space, forest, shrub land, herbaceous, planted/cultivated, and wetlands. To calculate the percentage of relevant green infrastructure classifications, the total number of selected green infrastructure cells was divided by the total number of cells.

	NLCD Classification
Green Infrastructure Classification	Open space developed, Forest, Shrub land, Herbaceous, Planted/Cultivated, and Wetlands
Non Green Infrastructure Classification	Open water, Low, medium, and high intensity developed, Barren Land (Rock/Sand/Clay)

Table 3. Green Infrastructure and Non Green Infrastructure Classification

To measure green infrastructure from 1-meter high resolution NAIP imagery, NAIP was downloaded from the Texas Natural Resources Information System (TNRIS) for county level. As NDVI values range from -1 to 1, green infrastructure was reclassified as “1” for the raster cells with a NDVI value greater than 0.1. After this reclassification process, to calculate green infrastructure percentage, the total number of cell reclassified as “1” was divided by the total number of cells. Figure 1 shows an image of green infrastructure based on NDVI in 2004 and NLCD classification in 2006 for Harris County, TX. The NLCD data are only available for 2001, 2006, and 2010, so 2006 data was employed, allowing for the most direct comparison. A visual comparison suggests that the NDVI data, because of the higher resolution, is able to capture and identify more types of “green” land uses in Harris County. Indeed, the total green infrastructure percentages for Harris County based on 1-meter high resolution was found to be 61.5 percent of the area, compared to the 51.5% based on the NLCD. Hence the finer resolution data and our new measurement of green infrastructure is a different approach at capturing detailed green infrastructure information, especially for urban environments such as Harris County.

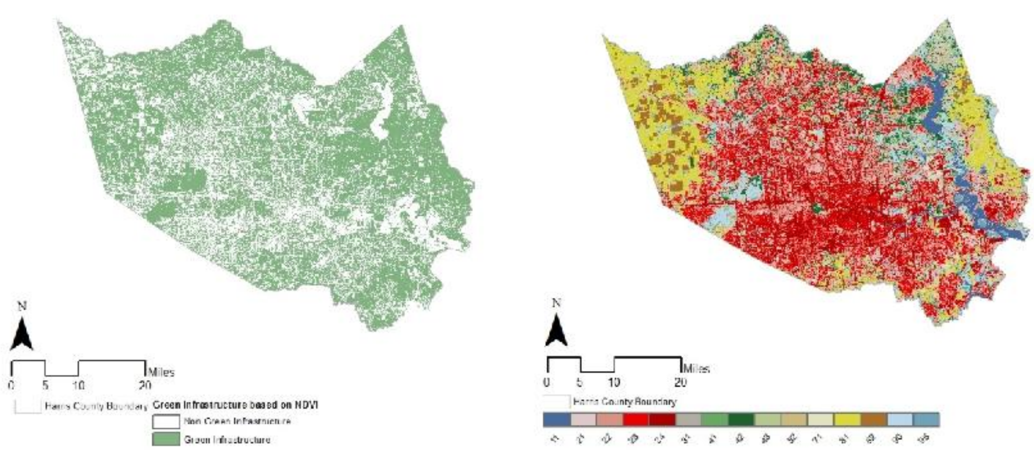
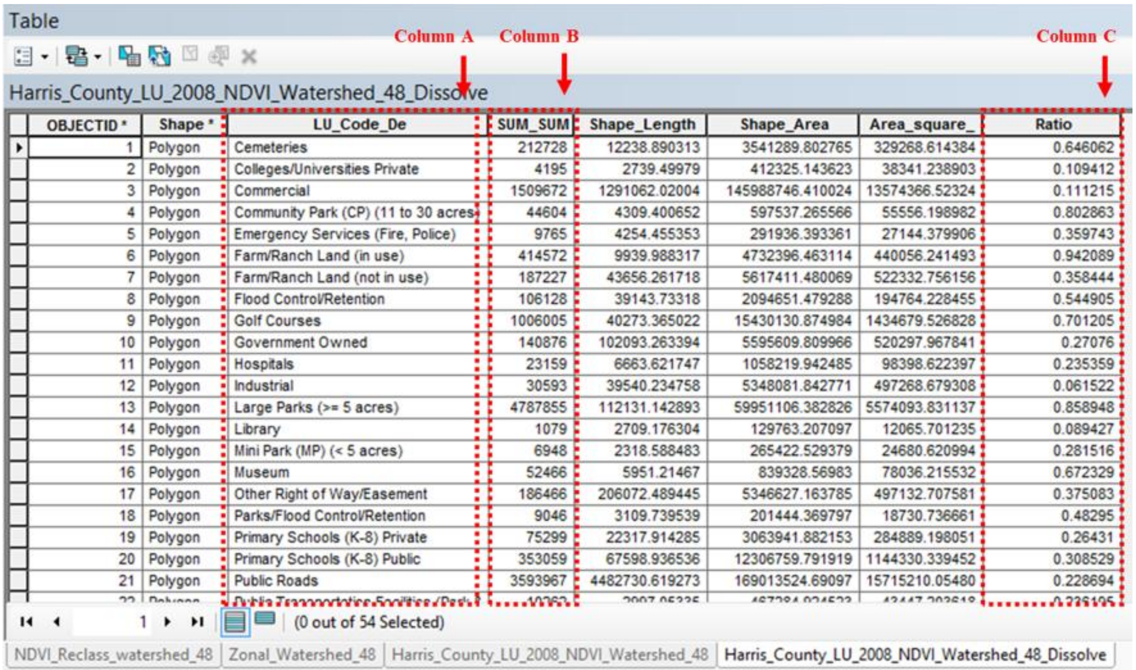


Figure 1. Normalized Difference Vegetation Index (NDVI) and National Land Cover Data (NLCD)

3.4 Green Infrastructure New Measurement Potential Implementation

In addition to providing higher resolution information regarding green infrastructure, these new data and measurements provide opportunities for refining out understanding of green infrastructure within existing land use classifications as well as the ability to capture more refined spatial patterns of green infrastructure. For example, this new measurement allows for the calculation of total areas of green infrastructure for each individual parcel in an urban area (if land use information is available for each parcel). Since NDVI is calculated based on 1-meter resolution imagery, one pixel equals 1 square meter. The total number of pixels with the value "1" represents the total area of green infrastructure in square meters for each parcel. Combining the results of NDVI in parcel layers with associated land use information, it is possible to compare the total ratio of green infrastructure by land use types. For example, the total ratio of green infrastructure for residential land use can be compared to the corresponding ratio for other land use types. Figure 2 is to further clarify how to measure total ratio of green infrastructure relative to other land use types. It shows land use types (Column A), total areas in square meters of green infrastructure across land use types (Column B), and total ratio of green infrastructure by each land use in a sample (Column C).

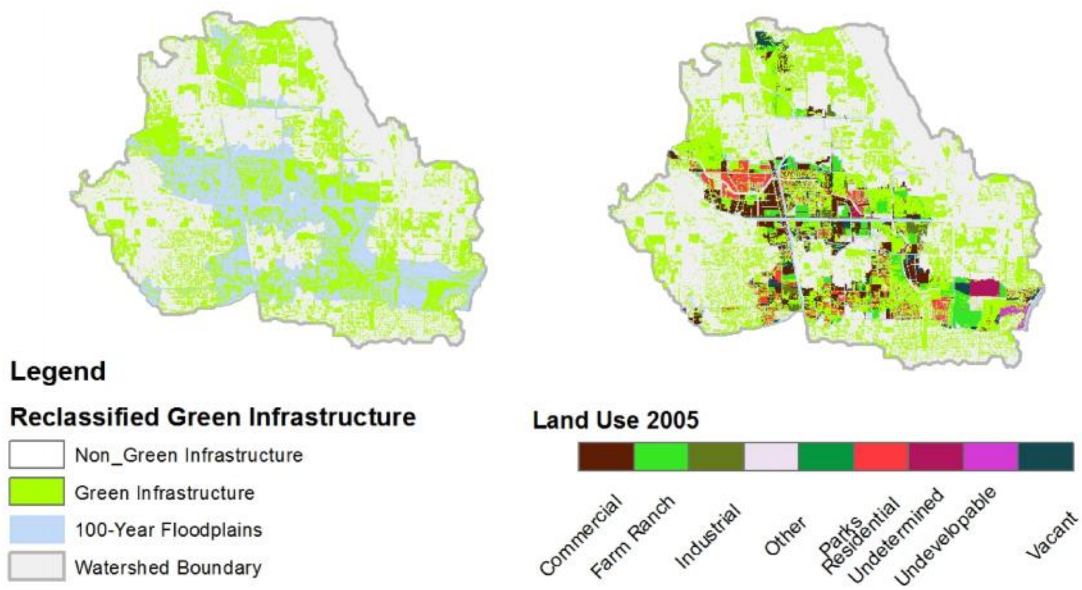


OBJECTID *	Shape *	LU_Code_De	SUM_SUM	Shape_Length	Shape_Area	Area_square_	Ratio
1	Polygon	Cemeteries	212728	12238.890313	3541289.802765	329268.614384	0.646062
2	Polygon	Colleges/Universities Private	4195	2739.49979	412325.143623	38341.238903	0.109412
3	Polygon	Commercial	1509672	1291062.02004	145988746.410024	13574366.52324	0.111215
4	Polygon	Community Park (CP) (11 to 30 acres	44604	4309.400652	597537.265566	55556.198982	0.802863
5	Polygon	Emergency Services (Fire, Police)	9765	4254.455353	291936.393361	27144.379906	0.359743
6	Polygon	Farm/Ranch Land (in use)	414572	9939.988317	4732396.463114	440056.241493	0.942089
7	Polygon	Farm/Ranch Land (not in use)	187227	43656.261718	5617411.480069	522332.756156	0.358444
8	Polygon	Flood Control/Retention	106128	39143.73318	2094651.479288	194764.228455	0.544905
9	Polygon	Golf Courses	1006005	40273.365022	15430130.874984	1434679.526828	0.701205
10	Polygon	Government Owned	140876	102093.263394	5595609.809966	520297.967841	0.27076
11	Polygon	Hospitals	23159	6663.621747	1058219.942485	98398.622397	0.235359
12	Polygon	Industrial	30593	39540.234758	5348081.842771	497268.679308	0.061522
13	Polygon	Large Parks (>= 5 acres)	4787855	112131.142893	59951106.382826	5574093.831137	0.858948
14	Polygon	Library	1079	2709.176304	129763.207097	12065.701235	0.089427
15	Polygon	Mini Park (MP) (< 5 acres)	6948	2318.588483	265422.529379	24680.620994	0.281516
16	Polygon	Museum	52466	5951.21467	839328.56983	78036.215532	0.672329
17	Polygon	Other Right of Way/Easement	186466	206072.489445	5346627.163785	497132.707581	0.375083
18	Polygon	Parks/Flood Control/Retention	9046	3109.739539	201444.369797	18730.736661	0.48295
19	Polygon	Primary Schools (K-8) Private	75299	22317.914285	3063941.882153	284889.198051	0.26431
20	Polygon	Primary Schools (K-8) Public	353059	67598.936536	12306759.791919	1144330.339452	0.308529
21	Polygon	Public Roads	3593967	4482730.619273	169013524.69097	15715210.05480	0.228694

Figure 2. Total Ratio of Green Infrastructure across Land Use Types

Compared to previous green-related land use information that only included such areas as parks and open spaces, this new measurement suggests a different approach of understanding and potential interpretation of how green infrastructure might be considered within land use categories not previously analyzed. This identified green infrastructure with its land use information can be used to examine the relationship between the amount of green infrastructure across different land use types and environmental hazards in urban areas.

Also, green infrastructure reclassifications can be converted land use data by combining reclassified green infrastructure information with land use data. This process provides detailed green infrastructure spatial patterns across different land use types. Figure 3 shows the potential results, indicating the difference between the reclassification (non-green infrastructure and green infrastructure) and the combined process. These detailed green infrastructure classifications sorted by land use types can be used to measure landscape metrics and, test the relationship between the amount of green infrastructure in various land use types and urban natural hazards.



Reclass_tif1__NDVI_NAIP_101

OBJECTID *	Value	Count
1	0	56697149
2	1	25165637

combine_luin_vat_raster_LU_08076000_IN

raster_LU_080	combine_l	combine_luin.vat:COU
Vacant	0	1409628
Farm Ranch	0	2559929
Undetermined	0	716480
Other	0	2226568
Industrial	0	777099
Commercial	0	3366896
Residential	0	2361387
Undevelopable	0	202344
Parks	0	42
Vacant	1	1483236
Farm Ranch	1	2953158
Undetermined	1	193800
Other	1	1405039
Industrial	1	185282
Commercial	1	996949
Residential	1	1773681
Undevelopable	1	268142
Parks	1	49

Figure 3. Reclassified Green Infrastructure with Land Use Data

3.5 Summary

Conventional approaches employed the National Land Cover Database (NLCD), Coastal Change Analysis Program (C-CAP) Land Cover data, which are gathered at a 30 by 30 meter resolution. Compared to these datasets, measuring green infrastructure using 1-meter high resolution NAIP imagery to compute a Normalized Difference Vegetation Index (NDVI) provides a different approach to capture green infrastructure, especially in urban areas. Utilizing this high resolution NAIP imagery is helpful in analyzing the effectiveness of green infrastructure on streamflow and potentially flood mitigation in urban areas. This new measurement for green infrastructure will provide support to guide research in to green infrastructure spatial characteristics and hazard mitigation.

4. RESEARCH METHODS

4.1 Conceptual Model

Figure 4 presents an overview of the conceptual model shaping the analysis employed in this research that represents the dependent, independent and control variables for analyzing the potential effectiveness of various dimension of green infrastructure on streamflow – annual peak flow and mean annual flow. Green infrastructure spatial pattern factors that were employed include assessments of 1) the overall green infrastructure in a watershed, 2) various measures of green infrastructure location, and 3) a number of measures of the spatial patterns of green infrastructure. In order to obtain sound estimates for the consequences of these dimensions of green infrastructure a host of natural environmental factors and built environmental factor were included in the statistical models developed as controls. The following discusses the measurement of each of the measures included within this conceptual model. However, before addressing these specifics, it is important to understand the process undertaken to capture and measure the key dependent variable for this analysis, streamflow, the selection of stream gauges capturing flow, and hence, the establishment of the basic units of analysis – watersheds -- and ultimately the sample selection of watersheds utilized in this research.

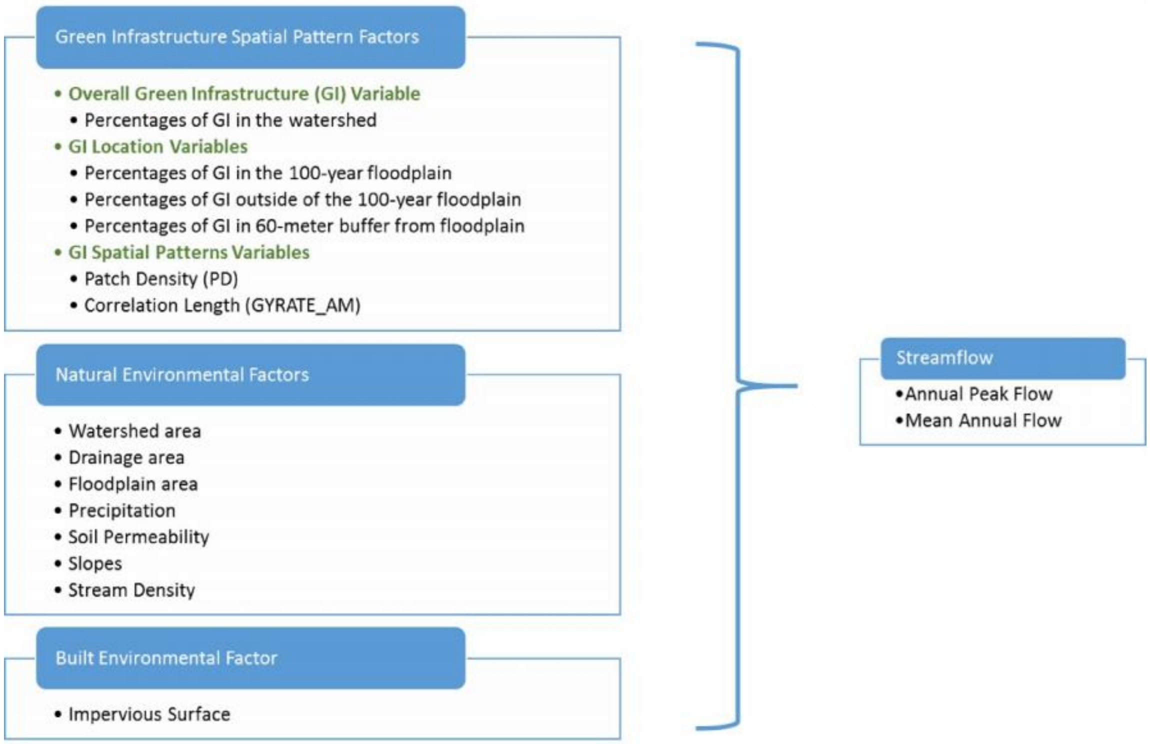


Figure 4. Conceptual Model

4.2 The Essential Steps in Sample Selection and Measurement

The key first step for this research was the selection of stream gage stations that provided the streamflow data. Once these stations had been identified, the watershed delineation associated with each station could be undertaken to define basic units of analysis. After delineating the watersheds associated with each station, NDVI was measured based on the NAIP imagery and key independent variables were constructed by utilizing the NDVI. NDVI reclassification allowed for the measurement of distributed green infrastructure. With these green infrastructure variables, FRAGSTATS was used to analyze spatial patterns of green infrastructure for each delineated watershed. Before running FRAGSTATS for the spatial patterns, resampling process from 1-meter resolution to 2-meter resolution was applied. Statistical analysis between these green infrastructure variables and streamflow was employed to assess the consequences of green infrastructure on streamflow. Figure 5 shows these steps. Finally, additional analysis and data collection of each watershed were undertaken to measure key control variables that must be included in the analysis in order to obtain valid and reliable estimates of the key independent variables.

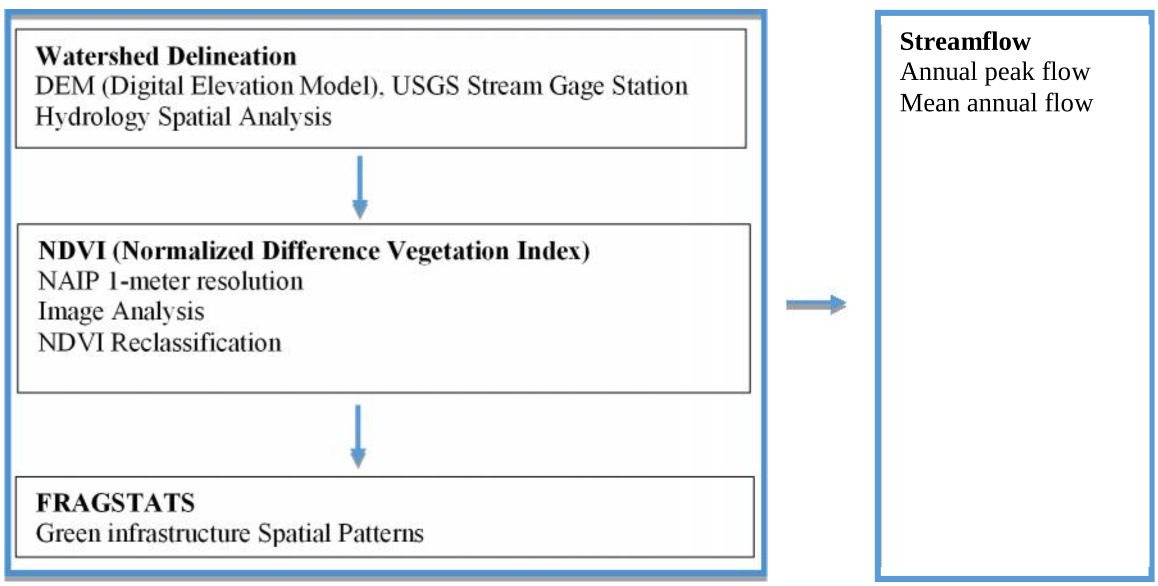


Figure 5. Research Method Flow

4.3 Sample Selection

As discussed earlier, many researchers have studied the consequences of land use and land use change at the community or county scale, by using aggregated data based on jurisdictions. For this study, however, rather than using the jurisdictional level, community or county level data, intra-jurisdictional data at the watershed level was employed. Specifically, watersheds were delineated based on locations of stream gage stations as a unit of analysis because ecological boundaries were more helpful in analyzing the effectiveness of green infrastructure for streamflow reduction.

For this study, the metropolitan areas of Houston and Austin in Texas were considered as the study area. Although these metropolitan areas are located in geographically different environments, these two areas both have stormwater runoff and flooding issues due to urban development and imperviousness and they are both located in the Gulf region of Texas. The Houston metropolitan area is an ideal area to analyze the relationship between green infrastructure and streamflow, which is one of the indicators for flooding. The number of fatalities from floods in Harris County, Texas (the county in which Houston is located) between 1960 and 2008 is the largest among all coastal counties in Texas (Brody et al., 2011). From 1996 to 2001, this county was one of the top 10 jurisdictions in the nation in approving land conversion for development (Brody et al., 2013). With the aim of strategically implementing green infrastructure dealing with open space conversion and rapid urban development, Harris County is an ideal urban area to study. The terrain characteristic of Harris County is very flat (Bedient, Holder, & Vieux, 2002), and it’s extremely flat terrain and intense rainfall patterns are closely linked to stormwater and flooding.

In addition to the Houston metropolitan area, the Austin metropolitan area was included in the study area. The Austin metropolitan area is located approximately 150 miles northwest of the Gulf of Mexico in south-central Texas (Veenhuis & Gannett, 1986). Travis County, Texas (the county in which Austin is located) is among the top 10 percent of flood prone communities as reported by the Federal Emergency Management Agency (FEMA)[1]. Travis County was among the top 20 Texas counties with the highest flood causalities between 1997-2001 (Zahran, Brody, Peacock, Vedlitz, & Grover, 2008). The rainfall and terrain characteristics of Travis County make the region vulnerable to stormwater and flooding. As it is located near the Gulf Coast of Texas, heavy precipitation from tropical storms can be produced. Additionally, steep topography with rapid urban development patterns is linked to flash flooding (Looper & Vieux, 2012). These two metropolitan areas have different regional characteristics, and these two regions have a wide variation of environmental conditions such as slope and soil type. These variations help analyze the effectiveness of green infrastructure on streamflow for those rapidly urbanizing areas in Texas.

The U.S. Geology Survey (USGS) provides different scales of hydrologic units. Both the Austin and Houston metropolitan areas are part of the Texas-Gulf region. However, instead of using these predefined hydrologic units, I delineated new watershed boundaries for this study based on available gage stations in the Austin and Houston metropolitan areas. To analyze the relationship between streamflow data and green infrastructure, it was more reasonable to delineate watersheds pertinent to each gage station, because stream gage stations can be used as outlet points for watershed delineation.

---

[1] https://www.traviscountytx.gov/tnr/swmp/floodplain-maps

To select stream gage stations, I first considered all stream gage stations in the Austin and Houston metropolitan areas including 23 counties. Figure 6 shows all of these counties overlaid with stream gage station information. Second, I considered the stream gage stations that were not located at the outlet of a dam or reservoir. Finally, I considered stream gage stations providing at least one of two datasets: daily discharge, or peak streamflow for the study period for 2004 and 2010. The red points in Figure 6 refer to the available stream gage stations based on the selection parameters for the study period for the water year of 2004 and 2010.

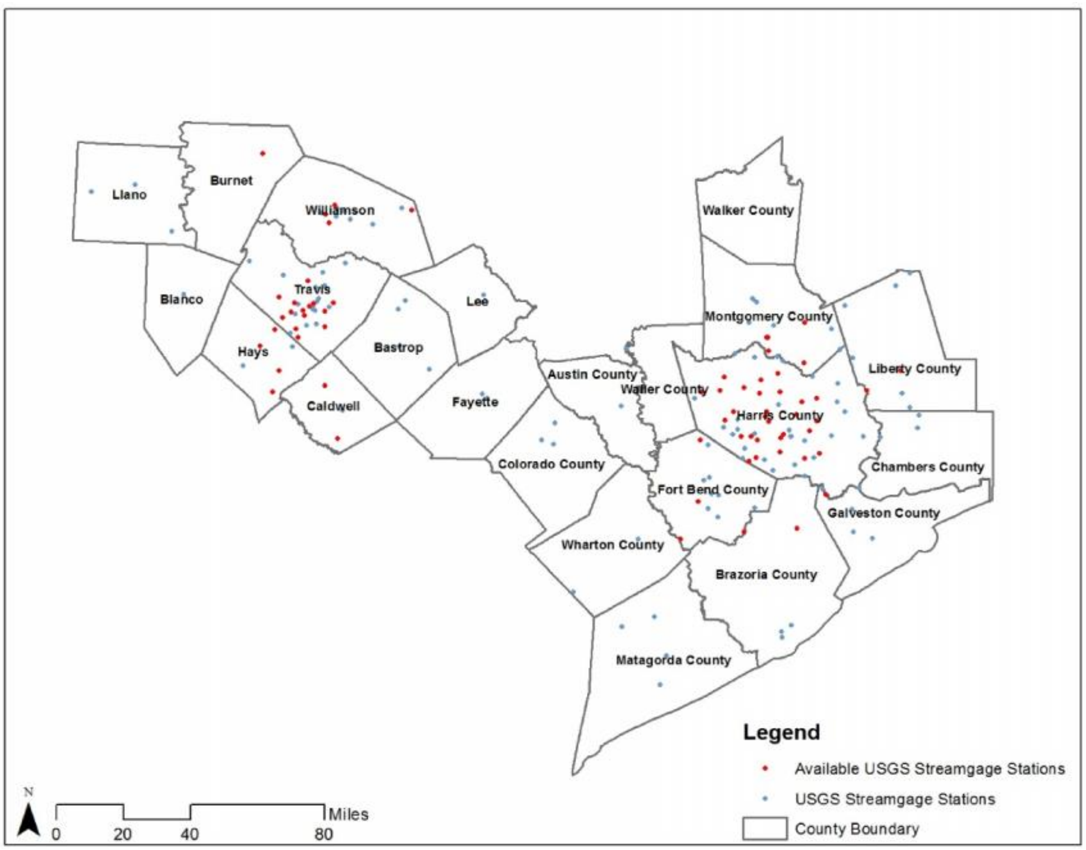


Figure 6. Available USGS Stream Gage Stations

4.3.1 Stream Gage Station Selection

Since the unit of analysis of this study was based on streamflow measurements gathered at selected stream gage stations, the selection of these station was an important step for this study. First, I narrowed my selection to stream gage stations in 13 counties from the Houston-Galveston Area Council (HGAC) and 10 counties from the Capital Area Council of Government (CAPCOG). These counties are located in different geographic settings and are, as seen here, governed by two different councils. This was advantageous to this study since, in order to analyze the effect of green infrastructure on streamflow measurement, it was better to have variances. The second criteria for selection was that the gage station should not be located at the outlet of a dam or reservoir. Third, daily stream gage measurement data must be available for more than 90 percent of the total number of days in 2004 and 2010. The above selection parameters resulted in a set of 66 stream gage stations available for the analysis of peak streamflow and 54 stream gage stations for daily discharge streamflow study. Since this study conducted data analysis for two-time period panel data analysis, the total number of data points collected were 132 for peak flow analysis, and 108 for daily discharge analysis.

4.3.2 Watershed Delineation

To delineate watersheds based on the selected stream gage stations, hydrologic analysis by the ArcGIS Spatial tool was used with the Digital Elevation Model (DEM) downloaded from the National Hydrography Dataset (NHD) Plus. Using the NHD Plus dataset, I checked each stream gage station to see if it was located on the NHD flow line. Additional snapping processing was conducted to manually relocate stream gage stations to intersect the NHD flow line. Once each watershed was delineated, watersheds with enough streamflow measurement data for the study period (2004 and 2010) were selected. Using the NHD Plus dataset based on the available stream gage stations, 66 watersheds across the two metropolitan areas were delineated. Ideally, it was best to have continuous watersheds as units of analysis. However, due to lack of stream gage data for the study area in 2004 and 2010, it was not available to delineate watersheds continuously. Figure 7 shows delineated watersheds based on their stream gage selection parameters.

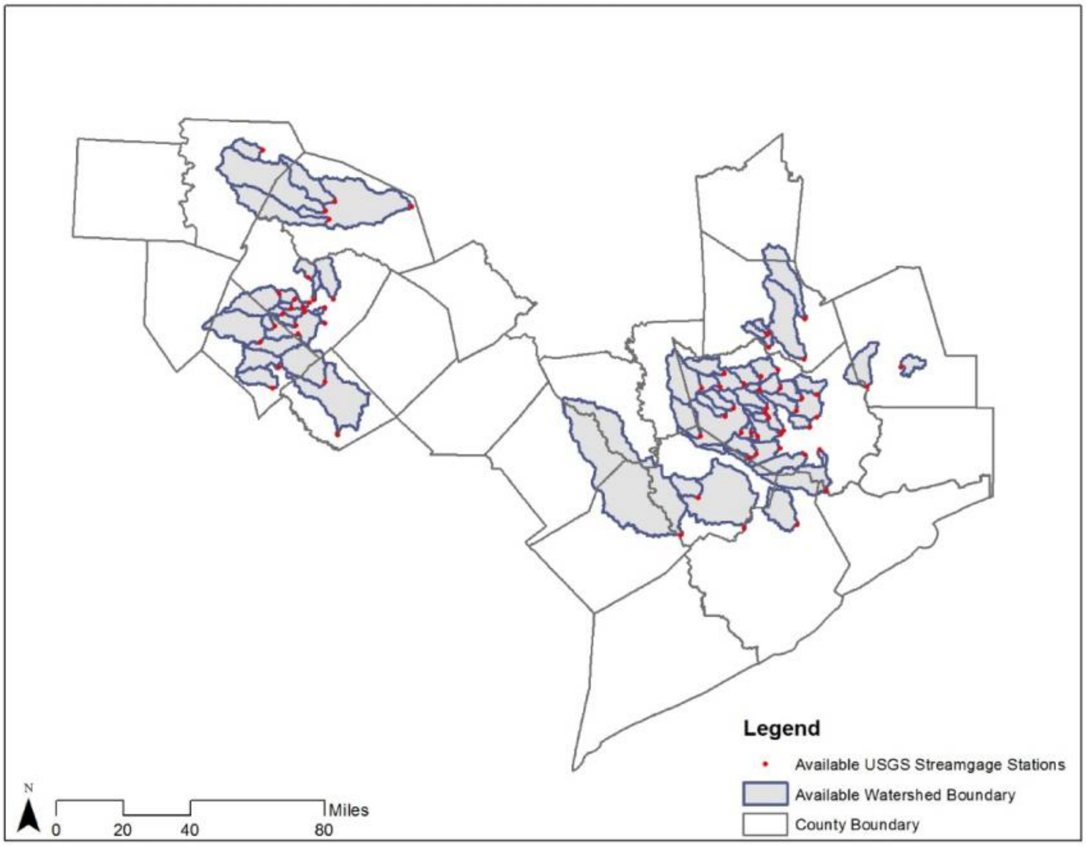


Figure 7. Watershed Delineation

4.4 Concept Measurement

4.4.1 Dependent Variables

The dependent variables in the conceptual model were streamflow discharge – annual peak flow and mean annual flow. This study sought to further the research on green infrastructure and its potential effect on reducing streamflow in urban environments over a 2-year period in 2004 and 2010. Therefore, streamflow peak and mean discharge data were collected for 2004 and 2010. The water year is different from the calendar year, since October 1st usually records the lowest annual levels for hydrological systems. The U.S. Geological Survey (USGS) defines a water year as the 12-month period extending from October 1st of any given year through September 30th of the following year, thus the dependent variables were collected for the water years of 2004 and 2010. Annual peak flow was the maximum discharge recorded each water year at each individual gage station, and mean annual flow was the average of all discharge recorded over the water year at each individual gage station.

For the purpose of this study, streamflow discharge data for the 2-year period in 2004 and 2010 were selected. I originally hoped to consider more recent years' data. However, 2004 was the first year that NAIP imagery was available for the study area in the proper format for NDVI measurement (1-meter high resolution and color infrared), and 2010 was the most recent year with imagery available in the same resolution and format. Due to the difficulty acquiring proper NAIP imagery gathered in more recent years, only the time period between 2004 and 2010 was considered.

The NAIP is typically scheduled to be acquired during the peak growing seasons in the continental U.S. However, due to unusual weather patterns, storms, cloud cover, fires (smoke), and other factors, the imagery is not always acquired at the same peak growing season from year to year. Not surprisingly, NDVI values, which show vegetation distribution can differ depending on when the NAIP imagery was acquired. For example, in the same year, NDVI values collected in summer are usually higher than corresponding values in winter. For Harris County, 2004 NAIP imagery was acquired on four different dates between August 13th and December 10th, whereas 2010 NAIP imagery was acquired on May 3rd (See Appendix C). Although the NAIP imagery acquisition dates for the study area in 2004 and 2010 were less than optimal, since their acquisition dates varied, they were likely to vary much less than in areas with more clearly defined seasons and they still provided unique data allowing for and assessment of green infrastructure within and between these periods.

Annual peak flow and mean annual flow variables were developed by collecting daily stream discharge data for each gage station. Annual peak flow was the maximum recorded discharge each water year, and it represented the high flow event for each year at each individual gage station. Mean annual flow was the average of all discharge measurement at each gage station recorded over the water year. The 12-month period water year defined by USGS were used for this study. The two dependent variables was calculated from the daily streamflow data recorded at each stream gage station for the two water years studied. Therefore, to ensure workable results, the data at each station should be individually checked to ensure there was no missing data. If more than 10 percentage of the total days in each water year were discovered to be missing, that stream gage station was not included. To measure mean annual flow, daily mean discharge streamflow for each gage station was gathered. The annual mean discharge was calculated by summing all recorded streamflow discharge and dividing it by the total number of recorded days. These dependent variables should be normally distributed for data analysis. For this reason, log transformation was conducted for the dependent variables. Figure 8 shows log transformed dependent variables to ensure normal distribution.

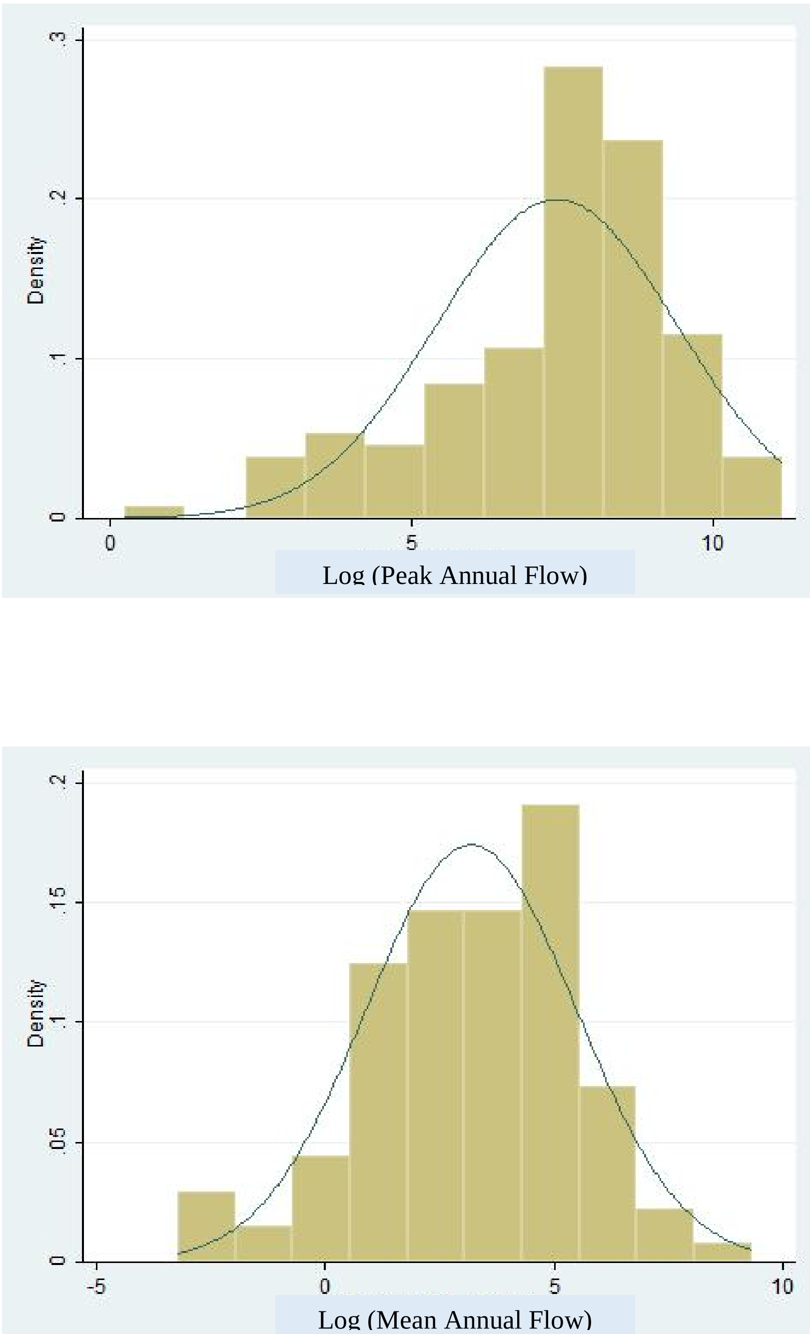


Figure 8. Log Transformed Dependent Variables

4.4.2 Independent Variables

As discussed in the research statement and literature review, several studies have analyzed the effects of green infrastructure. However, the effects of the spatial location and patterns of green infrastructure within urban areas have not been studied, particularly with a 1-meter high resolution imagery. The following discusses various measures of the spatial location and patterns of green infrastructure that were employed in this dissertation.

4.4.2.1 Green Infrastructure Spatial Patterns

Overall Green Infrastructure

First, overall green infrastructure was defined as the percentages of green infrastructure captured from NAIP imagery at watershed scale. Previous research suggests that naturally occurring land cover related to grass and forestland is effective for capturing and slowing surface runoff (Brody et al., 2014). Highly detailed measurements of green infrastructure, particularly within complex urban environments, were also expected to have a negative effect on annual peak flow and mean annual flow.

Hypothesis 1. The higher the percentages of green infrastructure in a watershed, the lower the annual peak flow and mean annual flow.

To measure the overall green infrastructure in the delineated watersheds, the reclassification of NDVI values was conducted to assign values of either “0” or “1”;

where “1” represented green infrastructure, and “0” signified anything else that was not green infrastructure. According to the National Aeronautics and Space Administration (NASA), very low values of NDVI (0.1 and below) correspond to barren areas of rock, sand, or snow. Moderate values represent shrub and grasslands (0.2 to 0.3), while high values indicate temperate and tropical rainforests (0.6 to 0.8). Based on these thresholds, each cell with an NDVI value less than 0.1 was reclassified as “0”, whereas cells with an NDVI value greater than 0.1 were reclassified as “1” (See Table 4).

NDVI values were calculated by using image analysis in ArcGIS. The NDVI was analyzed based on 1-meter high resolution NAIP imagery. First, NAIP imagery for the study area for both 2004 and 2010 was clipped for each delineated watershed. Second, the NDVI was calculated by selecting NDVI tab on the image analysis window in ArcGIS. The formula for the calculation is (NIR –RED) / (NIR + RED). Since the 2004 NAIP imagery was supplied in 1-meter color infrared format, for correct NDVI calculation, the red band should be assigned as “band 2,” and the infrared band should be assigned as “band 1”. Next, raster images with NDVI value were exported. Finally, reclassification for green infrastructure classification was conducted by using the reclassify tool in ArcGIS. NDVI reclassification allowed for the measurement of the total amount of green infrastructure.

	Green Infrastructure	
NDVI value	-1 – 0.1	0.1 – 1
Reclassify	0	1
Classification	Non Green Infrastructure	Green Infrastructure

Table 4. Reclassification for Categorizing Green Infrastructure

Green Infrastructure Location Measures

It may not simply be the amount of green infrastructure within a watershed that is key, but also the location. For example, previous studies suggest that green infrastructure has the ecological capability to absorb, store, and slowly release water, thus decreasing runoff and streamflow (Brody et al., 2008). Green infrastructure implementation such as setbacks from or buffers around waterways in the 100-year floodplains reduces adverse impacts from flooding. These linear protected areas can be implemented as the horizontal equivalent of freeboard, a margin of safety added to the base flood elevation to direct development away from floodplain areas and reduce damages to people and structures (Brody & Highfield, 2013). Other research also explains the effectiveness of wetlands and stream vegetated buffers in decreasing the rate of water flow (Castelle, Johnson, & Conolly, 1994). More specifically, some research has considered the width of a buffer to analyze the effectiveness of stream buffers. Shandas and Alberti (2009) considered a 100-meter buffer adjacent to the stream channel for riparian zones, which is the ‘buffering’ distance recommended by the Washington State Department of Ecology.

The Bayou Greenways program in Texas has suggested a minimum width of 60-meters (200-ft) in the development of environmental corridors purposed for flood mitigation and water quality improvements throughout the region. Since the Bayou Greenways program is located in part of this study area, the same 60-meter (200-feet) buffer zones from floodplains were considered for this study; the percentages of green infrastructure in these defined buffer zones were calculated. For this measurement, only a 60-meter buffer area around the 100-year floodplain was considered.

In light of these findings, a number of locational dimensions of green infrastructure were considered. Specifically three measures capturing the location of green infrastructure relative to the 100-year floodplain boundary were considered. These measures were: the percentage of green infrastructure in the 100-year floodplains in each watershed, the percentage of green infrastructure outside the 100-year floodplains, and the percentage of green infrastructure in the 60-meter buffer around the floodplain. My hypotheses relative to these measures are:

Hypothesis 2. The higher the percentages of green infrastructure in the 100-year floodplain in a watershed, the lower the mean annual flow and annual peak flow.

Hypothesis 3. The higher the percentages of green infrastructure outside of the 100-year floodplain in a watershed, the lower mean annual flow and annual peak flow.

Hypothesis 4. The higher the percentages of green infrastructure within the 60-meter buffer zones around the floodplain in a watershed, the lower the mean annual flow and annual peak flow.

4.4.2.2 Green Infrastructure Spatial Pattern Measures

Landscape metrics measure spatial patterns of the different landscape types such as landscape composition and configuration for patch, class, and landscape level (Leitão et al., 2006). They also measure the number, size, distribution, connectivity, and configuration (Brody et al., 2013; Olivera & Defee, 2007). Relatively homogeneous area that differs from its surroundings is defined as a patch. Patch-level landscape metrics measure characteristics of individual patches. A class is composed of a set of patches of the same type. Class-level metrics measures the configuration of a particular patch type, such as total extent, average patch size and degree of aggregation. A landscape refers a set of all patches within the study area. Therefore landscape-level metrics measure the overall composition and configuration of the patch mosaic, regardless of class value (Leitão et al., 2006). Tischendorf (2001) explains that class level metrics have stronger statistical relationships with response variables than landscape level metrics. For this reason, class level of patch density and correlation length were included as independent variables. As there are several different sets of landscape metrics identified in the literature, it was important to select appropriate landscape metrics to measure the spatial patterns of green infrastructure.

In order to measure spatial distribution of green infrastructure and how the composition changes, patch density (PD) was used as one of green infrastructure's spatial pattern variables. As land use patterns have changed as a result of rapid urban development, distribution of green infrastructure across land use types has also been altered. In order to measure spatial distribution of green infrastructure and how the composition has changed, it was logical to consider landscape metrics focusing on diversity – number of patches (NP) and patch density (PD). However, direct comparison of the number of patches (NP) for varying watershed areas has its problems, which does not consider how different each watershed area is. As Patch density (PD) normalizes the number of patches considering watershed size, Patch density (PD) metric was included as one of spatial pattern measures. (Leitão et al., 2006). To further help analyze landscape configuration, correlation length (GYRATE_AM), a variable that explains patch extensiveness, was included for this study (Leitão et al., 2006). Including these landscape metrics may help capture how particular forms of spatial patterns of green infrastructure can have consequences on reducing streamflow. Independent variables of landscape metrics were measured by utilizing the FRAGSTATS program, with the reclassified green infrastructure data gathered using the NDVI. For this process, resampling cell resolution from 1-meter to 2-meter was conducted before running the FRAGSTATS.

Patch Density (PD)

Patch density measures spatial patterns of green infrastructure, because it can explain how green infrastructure is distributed on a per unit area basis. In detail, Patch density (PD) is achieved by dividing Patch Number (PN) by total landscape area. Patch Density is calculated as follows:

$$Patch\ Density = \frac{PN}{A} \times (10{,}000\ m^2/ha.) \times 100$$

A = total area of watershed in m^2

PN = Patch Number

Because the total area of each watershed for this study varied, Patch Density was considered as a more appropriate measure of landscape composition than other landscape metrics since it facilitated comparisons among landscapes of varying size. A larger number of patches within each watershed explains a denser pattern of green infrastructure (Brody, Kim & Gunn, 2012). Dense patches of green infrastructure relate to a less impervious surface that could reduce streamflow. In light of these observations my hypothesis is:

Hypothesis 5. The higher the Patch density, the lower the mean annual flow and annual peak flow.

GYRATE (Correlation Length)

Radius of Gyration (GYRATE_AM) measures landscape configuration providing a good indicator of patch extensiveness by calculating the mean distance between each cell in a patch and the patch's centroid (Leitão et al., 2006). For example, with all else being constant, a larger patch means a greater radius of gyration. Also, more elongated shape of patches shows a greater radius of gyration. GYRATE measures the landscape connectivity, and the measured radius can be interpreted as the average distance that water can move within a single patch (Leitão et al., 2006). Additionally, GYRATE_AM, known as correlation length, measures the average distance within a patch from a random starting point and moving in a random direction (Leitão et al., 2006). Therefore, a large value of GYRATE_AM means more connectivity and hence the ability to influence run-off. Consequently my hypothesis with respect to this measure is:

Hypothesis 6. The higher the radius of gyration, the lower the mean annual flow and annual peak flow.

4.4.3 Control Variables

To analyze the relationship between green infrastructure and streamflow measurements, several control variables that are related to the effect on streamflow were considered. These variables were considered based on previous research showing significant relationship between each variable and streamflow measurements.

Furthermore, these variables provided the opportunity to statistically control for factors that were related to streamflow, hence allowing the analysis to obtain better estimates of the effects of green infrastructure, spatial patterns and location on the two dependent variables. The general sets of control measure that were included area related to 1) natural environmental variables (watershed area, drainage area, floodplain area, stream density, slopes, precipitation, and soils) and 2) impervious surface measures. The following sections discuss each of these sets of control measures.

4.4.3.1 Natural Environmental Variables

Watershed Area

According to Brody et al.(2011), watershed area is a significant factor affecting discharge. A larger watershed area means more discharge and streamflow. Watershed area was calculated for each delineated watershed in ArcGIS by using calculate geometry.

Drainage Area

When I compared each watershed area to the contributing drainage area provided by USGS, nested watersheds showed different values. Nested watersheds for the study area were inspected by overlaying stream gage stations and NHD flow lines in ArcGIS. For nested watersheds, there are two ways to control the effects of drainage area. One is to add upstream watershed area to downstream area. Another one is to include contributing drainage area as a separate control variable. For this study, contributing drainage area was included as one of the control variables in addition to watershed area.

Floodplain Area

The 100-year floodplain refers to land with a 1 percent chance of inundation each year due to flooding. Floodplain area was calculated as a percentage of the delineated watershed in the 100-year floodplain in ArcGIS based on FEMA floodplains Q3 dataset.

Precipitation

Precipitation is considered to be the most important factor contributing to local flooding. Brody et al.(2012) explain that more rainfall means more chances for overflow of streams and rivers due to excessive runoff. The mean annual precipitation for each watershed was calculated based on PRISM dataset. A monthly grid dataset was downloaded for water year 2004 and water year 2010. This twelve-month grid dataset for each water year was summed by using image analysis in ArcGIS. Zonal statistics as table was used to calculate mean precipitation, and this value was divided by 12 for annual mean precipitation.

Slope

The slope of a watershed has a significant effect on the temporal concentration and the amount of water storage; the steeper the watershed slope, the more rainfall concentration and the higher the stream peaks (Brody et al., 2011). Average percent slope in each sample was calculated by using the slope spatial analysis tool in ArcGIS with 10m Digital Elevation Models (DEMs). Mean slopes for watersheds were calculated by using the spatial analyst tool in ArcGIS based on Digital Elevation Model (DEM). Zonal statistics as table tool created a table showing mean slope as percent for delineated watershed boundary.

Stream Density

Stream density was calculated for each delineated watershed in ArcGIS by using calculate geometry and the field calculator. First, the total length of the stream in each watershed was calculated, and then this value was divided by total watershed area.

Soil Permeability

Soil permeability is significantly related to infiltration, runoff, and a corresponding streamflow. Average soil permeability, based on inches per hour of water infiltration, was calculated for each watershed by using the State Soil Geographic Database (STATSGO) (Brody et al., 2012). The STATSGO provides two values of the permeability rate. One is the maximum permeability (PERMH), and the other is the minimum permeability (PERML). The mean permeability was measured by averaging the sum of the maximum and minimum permeability. Since there were different soil compositions for each watershed, area-weighted average soil permeability was calculated for each watershed in ArcGIS.

4.4.3.2 Built Environmental Variables

Impervious Surfaces

Impervious surfaces are significantly related to surface runoff and peak discharges (Paul and Meyer, 2001; Brezonik and Stadelman, 2002). Percent impervious surface was measured using the National Land Cover Database for 2006 and 2011. This dataset consists of 30-meter resolution coverage, and it provides 20 classes of land cover. To calculate percent impervious surface for this study, three land cover classes – low, medium, and high intensity developed land cover – were aggregated within each watershed.

	Variable	Type	Measurement	Hypothesized Effects	Source
Streamflow	Annual Peak Flow	Dependent Variable	Maximum annual flow at each stream gage station		USGS National Water Information System
	Mean Annual Flow	Dependent Variable	Average discharge of the water year at each stream gage station		USGS National Water Information System
Overall Green Infrastructure	% of GI in the watershed	Independent Variable	Area of GI/Area of watershed	-	USDA NAIP Imagery
Green Infrastructure Location Measures	% of GI in the 100-year floodplain	Independent Variable	Area of GI outside of the 100-year floodplain/Area of outside of the 100-year floodplain	-	USDA NAIP Imagery
	% of GI outside of the 100-year floodplain	Independent Variable	Area of GI outside of the 100-year floodplain/Area of outside of the 100-year floodplain	-	USDA NAIP Imagery
	% of GI in the 60-meter buffer around floodplain	Independent Variable	Area of GI in the 60-meter buffer around floodplain/Area of 60-meter buffer around floodplain	-	USDA NAIP Imagery
Green Infrastructure Spatial Pattern Measures	Patch Density (PD)	Independent Variable	Number of patches per 100 hectare	-	USDA NAIP Imagery
	GYRATE_AM (Correlation Length)	Independent Variable	Mean distance between each cell in a patch and the patch's centroid cells	-	USDA NAIP Imagery
Natural Environmental Factors	Watershed Area	Control Variable	Area of watershed	+	
	Drainage Area	Control Variable	Area of contributing drainage area	+	
	Floodplain Area	Control Variable	Percentage of area within the FEMA-defined 100-year floodplain	+	HGAC (Harris Galveston Area Council)
	Stream Density	Control Variable	Total length of stream/Area of watershed	+	National Hydrology Dataset
	Slopes	Control Variable	Average percent slope	+	Digital Elevation Models (DEMs)
	Precipitation	Control Variable	Annual average rainfall	+	PRISM dataset
	Soils	Control Variable	Average soil permeability	-	STATSGO (State Soil Geographic Database)
Built Environmental Factor	Impervious Surface	Control Variable	Percent of Impervious Surface	+	National Land Cover Dataset

Table 5. Conceptual Measurement and Hypothesized Effects

4.5 Data Analysis Plan

The goal of this study is again to undertake a much more refined assessment of green infrastructure of two key indicators of the potential flood related consequences of run-off, stream peak flow and mean discharge, in a set 66 of urban watersheds for two points in time, 2004 and 2010. This analysis was made possible by generating extremely high-resolution assessments of green infrastructure employing the NAIP data and NDVI as discussed above. This dissertation's analysis represents a novel opportunity to employ highly refined data to explore the consequences of detailed assessment of various dimensions of green infrastructure, but there are also limitations. In the following I will briefly outline the data analysis approaches that made the best usage of these data to explore the consequences of green infrastructure, pushing the analyses as far as is possible, but also making clear the limitations. I first will address the general statistical approach to be undertaken to assess the consequences of dimensions of green infrastructure for the same group of watershed in different years and then outline the general analysis strategy to assess the consequences of green infrastructure and other variables of interest.

There are a variety of statistical approaches that might be utilized to statically model streamflow characteristics for the 66 watersheds with measures on key independent and control variables for two points in time. One simple approach might be to treat these at two cross-sectional datasets and analyze them for the two points in time, and then examine for differences in the two models. Yet another would be to treat them as two independent samples and simply pool the data, and undertake a single analysis of streamflow with and without a temporal dummy variable assessing for the consequences of the key measures, and perhaps assessing for the appropriateness of pooling the data. However each of these approaches would suffer for a number of problems, not the least of which is that they would not be taking full advantage of having repeated measures on the same observations at two points in time, allowing for a more causal assessment of the relationships. Given the nature of these data, to make maximum use of their structure allowing for an more refined assessment of the key independent variables, while at the same time appropriately taking into account correlated error due to repeated measures, a series of panel models were analyzed (Wooldridge, 2009). The general structure these panel models is as follows:

$$y_{it} = \beta_1 x_{it} + \beta_2 x_{it} + \cdots\cdots\cdot + \alpha + u_{it} + \varepsilon_{it},$$

Where *i* denotes the watershed, and *t* = 1 or 2, where 1 corresponds to 2004, and 2 corresponds to 2010. The independent variables included measure of green infrastructure developed as discussed above, along with an appropriate set of controls. These models allowed me to appropriately model the error, addressing the consequences of having repeated measures for each of the watersheds at two points in time, while making full use of the data to obtain good estimates of the effects of green infrastructure and other key independent and control measures. While panel models are the most appropriate techniques to employ given the nature of my data, one can find major divergence and disagreement in the literature over what types of panel models are appropriate given different data structures and theoretical issues. This disagreement revolves around the use of random versus fixed effect models. In many respects these disagreements are often between many in the econometrics and business literature that hold that fixed effect model yield statistically superior estimates versus many in the broader social science and even environmentally oriented literature that suggests that there are both sound empirical and strong theoretical rationales for employing random effect models. In the case of the former, the argument suggests that it is only significant to address within observation variation, hence individual observational factors should be treated as fixed, while in the latter, the argument suggests that variation within and between observations are significant, hence focusing on only within variation is far too limiting and does not properly specify the models. The issue basically boils down to whether or not time invariant measures should be addressed in the models. In the social science literature these are often measures such as gender, race/ethnicity, and other theoretically significant measures. In the present case, the issue is whether or not time invariant measures associated with watersheds, such as slope, drainage area, or watershed area itself should be considered in the models. Clearly, from an engineering and environmental perspective, ignoring such measures would be impossible, and yet from a statistical approach where the focus is on within watershed variation is critical, these factors could be treated as "fixed." Since there is no clear consensus in the broader literature as to which technique is preferable, both fix and random effect panel models were utilized and presented in this dissertation. Hence, my discussions often focus on obtaining robust assessment of the potential consequences of green infrastructure on streamflow, implying consistency across the two forms of models.

4.5.1 Analysis Strategy

While the goal of this study is to undertake an assessment of the effects of green infrastructure on of two key indicators of the potential flood related to the consequences of run-off, stream peak flow and mean discharge, utilizing significantly improved assessments of that infrastructure made possible using the NAIP data and NDVI based measurement in set 66 of urban watersheds for two points in time, 2004 and 2010, I must now address one of the most important limitations of this dissertation -- sample size – and its consequences for my analysis strategy. For the proposes of this dissertation I have undertaken the significant activities of learning to use the NAIP data and NDVI measures, proposed and developed various measures of green infrastructure using these data, identified urban stream gauges with useful data for two points in time in two urban areas subject to flooding and identified each gauge's associated watersheds and drainage areas, generating a sample of annual peak flow with 66 observations and a sample of mean annual flow with 54 observations at two points in time, yielding 132 and 108 observations in the panel model. As I worked with the a here-to-for never employed high resolution data I was able to develop measures a host of green infrastructure related to size, location, and spatial patterns, that do display high correlations resulting in major issues of multicollinearity, given the small number of observations in my sample.[2] Specifically, as will be seen below, there were high levels of multicollinearity among the multiple measures of green infrastructure I would like to include in each of the models predicting streamflow. Technically speaking, this is not a problem of estimation,

[2] Correlation matrix of dependent, independent and control variables is available in Appendix B.

assuming other areas of model specification assumptions have been met, the estimated coefficients are statistically sound; however, statistical testing of these coefficients is compromised by inflated standard errors (Wooldridge, 2009). The only actual solution to this problem is to increase the sample size, which was not practicable within the context of this dissertation research, but something I do intend on addressing in the future. Therefore, my solution was to run a series of iteratively developed models, including sets of green infrastructure measures that did not display multicollinearity issues. While this was a less than satisfactory solution, it allowed me to 1) assess the utility of these new data to assess the consequences of green infrastructure in urban environments and 2) to explore the potential consequences of various dimensions of green infrastructure on streamflow. It is my hope that this analysis will show, albeit tentative, the utility of these data and point to potential consequences of not only green infrastructure in general for streamflow, but also highlight dimensions of green infrastructure for future research.

In light of the above, my analysis strategy was as follows. First, my analysis was broken into two parts, the first of which addressed *peak streamflow*, while the second addressed *mean flow levels*. Within each of these primary analyses two phases of analyses preformed. Phase 1 focused on specific models that tested the consequences of 1) overall green infrastructure described as the percentage of green infrastructure in watershed, 2) locational dimensions of green infrastructure (i.e. green infrastructure in the100-year floodplains, outside of the 100-year floodplains, or in 60-meter buffer zones from floodplains), and 3) landscape metrics for green infrastructure spatial pattern – patch density and correlation length. In general, these individual models focused on only one dimension (general, spatial, and landscape metrics), but where possible combinations of these measures were included simultaneously. Throughout both fixed and random effects models were employed, allowing for some or all (time variant and invariant) control measures to be included in the models.

In Phase 2, the models for both phases were reanalyzed testing for statistically significant difference across the two urban areas included in this analysis: The Austin and the Houston metropolitan areas. This phase enabled me to assess whether or not there might be important regional variations in the consequences of green infrastructure across the two urban areas. These two areas vary considerably with the terrain characteristics such as slope and soil types which may have consequences for the effectiveness of GI when addressing annual peak flow. The procedure for undertaking this testing was straightforward. First a regional or metropolitan dummy variable were included in each of the models previously estimated in Phase 1, along with a set of interactions between the dummy variable and each of the key GI variables included in the model. Specifically, a regional dummy variable labeled Houston, which equals one for Houston region and zero for Austin region, was included along with interactions between this variable and each of the key GI measures included in a specific model. The dummy tested for "level" differences between the two areas, while the interactions tested for incremental variations in the effects for each GI measure. Statistical testing was performed to test for improvements in the models (implying metropolitan variations), and for incremental and net effects for each GI measure. It should be noted that this analysis was only performed using the random effect models, because the “dummy” variable is time invariant.

It should be noted that for all of the above analysis diagnostics for cross-sectional dependence, heteroskedasticity, and spatial autocorrelation test were performed. Friedman’s test for cross sectional dependence showed that there was no cross-sectional dependence. The result of heteroskedasticity test recommended the use of robust for robust standard errors, hence robust standard errors were utilized and were presented in all tables. Global Moran’s I test for regression residual showed that there was no spatial autocorrelation (See Appendix A). The regression residual of peak annual flow had a Moran’s *I* value of 0.194 (p=0.194). As the data for this study was for only two years with gaps, the serial correlation test was not necessary. The multicollinearity issue was tested by running variance inflation factor (VIF) after conducting the least squares dummy variable model with the same variables. As indicated above there were significant issues found, hence the models presented represent those in which inclusion of specific subsets of independent variables did not present multicollinearity issues. Again, this was less than a satisfactory solutions, compared to increasing the sample size, but it was a workable one given constraints.

Finally it should also be noted that Hausman Test assessing whether or not there were significant variations in the fixed and random effect estimates for time variant variables were conducted throughout the analysis. Such test were always significant with $Prob(\chi^2) < 0.05$. For those predisposed to only focusing on within observation differences these results are interpreted to mean that fixed effect model should be preferred. In this context time invariant control variables such as watershed area, drainage area, floodplain percentages, soil permeability, slope percentages, and stream density which are important factors to for predicting streamflow should be excluded from the models. However, as noted above, since random effect panel models allow for the inclusion of time-invariant variables as explanatory variable and allow for a more theoretically appropriate model specification when considering both within and between observation variations, random effect models were presented as well. When considering the overall results, I paid particular attention to assessments of green-infrastructure that were significant across both types of models, since they were much more likely to be robust.

5. ASSESSING THE EFFECT OF GREEN INFRASTRUCTURE ON PEAK ANNUAL FLOW AND MEAN ANNUAL FLOW

This section presents the various panel models addressing the hypotheses discussed in the previous section. I begin with a brief discussion of the descriptive statistics for the variables in the analysis with a specific focus on the green infrastructure measures.

5.1 Descriptive Statistics

Table 6 presents the descriptive statistics for all of the dependent variables (both logged and level variables) and for independent and control variables. I have displayed the descriptive statistics for the sample of 2004, 2010, and the combined sample of 2004 and 2010.

Over two year period in 2004 and 2010, it should perhaps not be surprising that in these two rapidly growing areas I have seen an overall decline in the absolute, as well as percentage loss in green infrastructure. Figure 9 displays data on the percentage point change in green infrastructure for each of the 66 watersheds between 2004 and 2010. Of the 66 watersheds, 51 experienced loss in green infrastructure, with many experiencing losses in excess of 30 percentage points. On the other hand, a much smaller number, 15, did experience gains, with 5 having gains of greater than 10 percentage points. However, the overall the average decreased amount of green infrastructure was 5.9 percentage points.

Table 7 provides data on the percentage point change in green infrastructure between 2004 and 2010 classified in terms of its location relative to the 100-year floodplain. Interestingly, while both urban areas in this study have seen a loss in green infrastructure overall in their watersheds, there are considerable differences in these losses, depending upon location. The smallest loss, on average, occurred within the 100-year floodplain, where the average was 2.7 percentage point loss in green infrastructure over the 6 year period. In the 60-meter buffer around the floodplain, the losses averaged 4 percentage points. The highest average percentage point losses occurred in watershed areas well outside the 100-year floodplain. In these areas the average percentage point loss was 6.5. On the whole then, these findings suggest that greater losses occurred well outside the 100-year floodplain in watersheds located in the Austin and Houston metropolitan areas.

Variable	2004					2010					Combined				
	N	Mean	Standard Deviation	Min	Max	N	Mean	Standard Deviation	Min	Max	N	Mean	Standard Deviation	Min	Max
Peak Annual Flow (ft3 / s) (log-transformed)	66	9300.03 (8.64)	11796.2 (0.96)	672 (6.51)	71100 (11.17)	66	1859.34 (6.19)	2823.01 (2.01)	1.3 (0.26)	14700 (9.60)	132	5566.71 (7.41)	9327.57 (2.00)	1.3 (0.26)	71100 (11.17)
Mean Annual Flow (ft3 / s) (log-transformed)	54	400.77 (4.49)	1543.23 (1.47)	3.95 (1.60)	11115.27 (9.32)	54	59.51 (2.38)	179.42 (1.72)	0.04 (0.04)	958.68 (6.87)	108	230.14 (3.43)	1106.79 (1.91)	0.04 (0.04)	11115.27 (9.32)
GI% in Watershed	66	59.91	8.96	37.67	81.29	66	54.00	12.58	26.52	83.54	132	56.96	11.27	26.52	83.54
GI% in Floodplain	66	67.36	10.77	40.35	91.03	66	64.70	15.67	31.82	89.16	132	66.03	13.46	31.82	91.03
GI% outside of Floodplain	66	58.50	8.91	34.63	76.52	66	52.01	12.53	23.25	82.66	132	55.26	11.31	23.25	82.66
GI% in 60m Buffer around Floodplain	66	61.91	10.54	26.74	83.78	66	57.90	15.11	14.34	86.58	132	59.91	13.13	14.34	86.58
Patch Density (Number of Patches/ 100 ha.)															
Watershed	61	701.86	261.75	177.49	359.32	61	145.88	150.33	770.21	1293.72	122	530.59	272.21	150.32	1293.72
In Floodplain	65	514.55	287.87	133.30	332.04	65	161.65	79.21	813.19	1259.28	130	423.29	249.94	79.21	1259.28
Outside of Floodplain	63	740.64	269.41	194.25	378.38	63	150.53	181.35	821.87	1354.31	126	559.51	283.39	181.35	1354.31
In Floodplain Buffer	65	1967.48	592.84	918.03	1469.66	65	472.88	740.16	3827.75	4020.14	130	1718.57	589.69	740.16	4020.14
GYRATE_AM (Mean distance in meter between each cell in a patch and the patch's centroid cells)															
Watershed	61	947.52	647.71	153.17	882.32	61	827.45	111.02	4290.07	2667.86	122	914.92	740.68	111.02	4290.07
In Floodplain	65	564.59	393.81	121.61	631.36	65	459.94	68.75	2105.41	1776.17	130	597.97	427.81	68.75	2105.41
Outside of Floodplain	63	557.80	301.05	124.68	478.07	63	356.86	83.14	1948.81	1287.65	126	517.93	331.24	83.14	1948.81
In Floodplain Buffer	65	181.96	101.83	13.88	171.60	65	108.04	37.72	528.37	775.48	130	176.78	104.7	13.88	775.48
Watershed Area (sq.mi)	66	66.81	102.82	2.66	724.05	66	66.81	102.82	2.66	724.05	132	66.81	102.82	2.66	724.05
Drainage Area (sq.mi)	66	1673.02	7516.89	6.3	45339	66	1673.02	7516.89	6.3	45339	132	1673.02	7516.89	6.3	45339
Floodplain %	66	16.44	9.78	3.03	46.73	66	16.44	9.78	3.03	46.73	132	16.44	9.78	3.03	46.73
Precipitation (mm)	66	931.21	120.98	751.78	1167.09	66	316.63	96.15	177.68	600.82	132	623.92	327.10	177.68	1167.09
Soil Permeability (mm/hour)	66	29.02	20.1	9.91	123.88	66	29.02	20.1	9.91	123.88	132	29.02	20.1	9.91	123.88
Impervious %	66	34.07	29.79	0.007	86.82	66	36.60	30.68	0.007	87.80	132	35.33	30.15	0.0073	87.8
Slope %	66	2.57	2.74	0.25	11.06	66	2.57	2.74	0.25	11.06	132	2.57	2.74	0.25	11.06
Stream Density (mi/sq.mi)	66	2.30	0.66	0.99	3.96	66	2.30	0.66	0.99	3.96	132	2.30	0.66	0.99	3.96

Table 6. Descriptive Statistics for Variables in 2004, 2010, and Combined

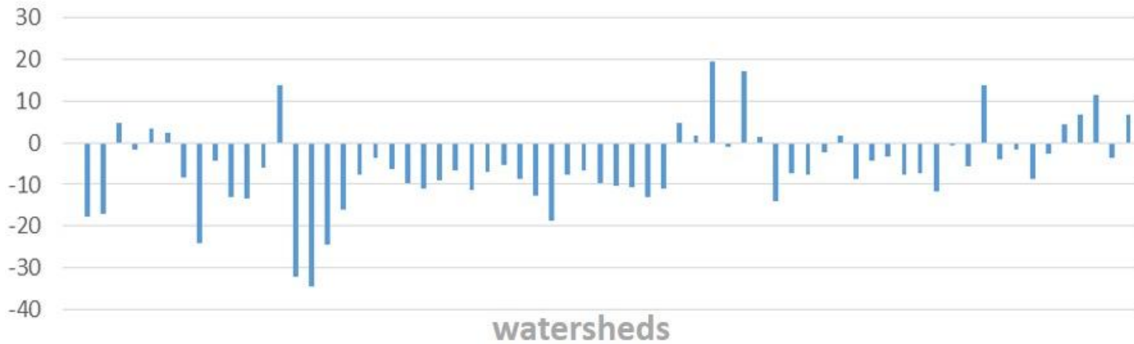


Figure 9. Percentage Point Changes in Green Infrastructure between 2004 and 2010 (%2010 - %2004)

Green Infrastructure Location	Percentage Point Change in Green Infrastructure
Watershed	-5.9 %
In the 100-year floodplain	-2.7 %
Outside the 100-year floodplain	-6.5 %
60-m buffer around the 100-year floodplain	-4.0 %

Table 7. Percentage Changes of Green Infrastructure between 2004 and 2010

5.2 The Effect of Green Infrastructure on Peak Annual Flow

Table 8 presents a series of fixed and random effect panel models assessing the consequences of green infrastructure on peak annual flow, focusing on the overall percentage of green infrastructure within a watershed and with respect to the location of green infrastructure relative to the floodplain. Unfortunately due to multicollinearity problems between total green infrastructure and its locational measures, separate sets of models were analyzed. Specifically, three sets of models are presented in Table 8. The 1A Models were a baseline set of fixed and random effect models predicting peak annual flow employing only the basic control variables capturing critical factors generally thought to influence peak annual flows related to a watershed's precipitation and other salient characteristics. The 2A Models included a set of fix and random effect models predicting peak flow with the baseline controls and a measure of the overall percentage of green infrastructure within the watershed. The 3A Models included three sets of random and fixed effect models with baseline controls and each of the green infrastructure locational measures: the percentage of green infrastructure in the 100-year floodplain, in the 60-meter buffer around the floodplain, and outside the floodplain. Finally, the 4A Models pushed the analysis a bit, by presenting a set of models, including pairs of the locational measures that was not completely free of multicollinearity issues, but perhaps were borderline.

Focusing first on the baseline set of models (1A Models) I can see that, in general, the controls worked as expected. In the fixed effect model the two time variant indicators of precipitation and imperviousness had significant effects on peak annual flow. Specifically, precipitation had a significant positive effect on peak annual flow, suggesting that every millimeter of precipitation increased peak annual flow by .54 percent. Also, every percentage point of impervious surface in a watershed also increased peak annual flow by 36%. The random effects model suggests that watershed and drainage area also had significant positive effects while soil permeability significantly reduced peak annual flow across watersheds through time. Every square mile of watershed area increased peak flow by 0.27%, while every square mile of drainage area increased peak flow by 0.006%. On the whole, these patterns of findings with respect to the baseline control measures hold across all other model sets, with the exception of soil permeability and slope. Soil permeability tended to become insignificant once green infrastructure measures were introduced into the model while the watershed's slope tended to become positively albeit only marginally (.1) significant after green infrastructure measures were introduced into the model.

Returning to the baseline set of models one can see that the R^2 for the fixed model is .6057, suggesting nearly 61% of the within observation variation was on average accounted for by the baseline model. The random effect's model's overall R^2 suggest that just over 56% of the total variation in peak annual flow was accounted for by the basic model. Table 8 also presents the findings for the Hausman Test associated with each set of panel models estimated. This test assesses significant differences between the coefficients associated with the time-variant variables estimated by the fixed and random effect models. Perhaps not surprisingly, all of these tests were statistically significant, suggesting that there were indeed significant differences between these sets of estimates across all model sets. For some, these results suggest that the "fixed" effect results were preferred. However, since the random effects models allow for theoretically significant and substantive time-invariant variables (watershed area, drainage area, floodplain percentages, soil permeability, slope percentages, and stream density) to be considered as explanatory/control variables both sets of models were presented.

5.2.1 Overall Green Infrastructure and Annual Peak Flow

The 2A Models added to the baseline models the percentage of watershed area associated with green infrastructure and in both cases adding overall green infrastructure significantly increased the respective R^2's of the models, in comparison to the appropriate base model. Whether considering the fixed or random effects model, in each case the consequences of green infrastructure was, statistically significant and negative, having controlled for the baseline factors. These findings are consistent with the first hypothesis that overall green infrastructure should significantly reduce peak annual flow. However, as anticipated by the Hausman test results discussed above, the magnitudes of the effects of green infrastructure differed considerably between the two models. In the fixed effect model, the results suggest that the peak annual flow decreased by 7.2%[3] for every percent increase in overall green infrastructure, while in the random effects model there was a 2.9% decrease for every percent increase in overall green infrastructure.

[3] $100(e^{-.0744748} - 1) = -7.2\%$

Both cases represent rather significant reductions in peak flow for the period under consideration.

5.2.2 Green Infrastructure Location and Annual Peak Flow

The 3A Models assessed the consequences of the specific locations of green infrastructure within the watershed location variables on annual peak flow. In light of multicollinearity issues I first presented three sets with only one of the locational variables in the model, then a set with two of the locational measures.[4] The first model set in 3A included only the percent of green infrastructure in the 100-year floodplain. This measure was negative and statistically significant, in both the fixed and random effects model ($P < .01$ in the fixed and $P < 0.05$ in the random). The effect in the fixed effects model suggest that with every percent increase in GI within the 100-year floodplain, peak annual flow decreased by 7.7%. Controlling for time invariant measures reduced the effect to 2.1%. As can be seen in the next set of models, the effects of GI outside the floodplain appeared to be quite comparable, in that both measures were significant, have negative effects, and their magnitudes in the fixed (-7.1%) and the random (-2.5%) were similar to those for GI in the floodplain. The only differences between these two sets, were with respect to the R^2 values where they were slightly higher in the models including the percent GI within the floodplain (fixed = .6985; random = .5787), as opposed to the percent GI outside the floodplain (fixed = .6447, random = .5745).

[4] Models with all three measures displayed features suggesting major issues with multicollinearity.

The final set of models in this series included measures of green infrastructure in the 60-meter buffer around the floodplain. In this set, the coefficient for green infrastructure was only significant in the fixed effects model. In the fixed effects model the percent of GI in the 60-meter buffer around the floodplain had a significant negative effect, as anticipated, and the coefficient's magnitude suggests that with every percentage point increase in GI in this buffer, peak annual flows decreased by 8.9%. However, when other time invariant measures are controlled for, this effect became insignificant.

The 4A models pushed the analysis a bit further by including sets of the GI locational measures within the same model. The 4A models included both green infrastructure in the 100-year floodplain and outside of the 100-year floodplain variables. Fixed effects model suggests that green infrastructure in the 100-year floodplain had a negative and statistically significant effect, with a 10.9% reduction in annual peak flow for every percentage point increase in GI within the floodplain. But, green infrastructure outside of the 100-year floodplain was not statistically significant. Furthermore, both measures in the random effects model were not statistically significant.

5.2.3 Green Infrastructure Spatial Patterns and Annual Peak Flow

Table 9 presents a series of fixed and random effect panel models assessing the consequences of green infrastructure on peak annual flow, focusing on the overall percentage of green infrastructure within a watershed and with respect to the location of green infrastructure relative to the floodplain, and also considering one of the spatial patterns of green infrastructure, patch density. Each set of models included pairs of the locational measures and patch density assessments of green infrastructure. Due to multicollinearity issues already explained for Table 8, separate sets of models were analyzed. Specifically, three sets of models are presented in Table 9 in that each set of models with the measure of %GI overall or with respect to location includes the associated the patch density measures – hence each model assessed not only the percentage of GI overall or with respect to particular locations within the watershed, but also the patch density of that percentage of GI. In general, the larger the patch density measure, the more contiguously clustered the %GI. The 2B Models included a set of fixed and random effect models predicting peak flow with the baseline controls and a pair of a measure of the overall percentage of green infrastructure within the watershed and a measure of patch density, within the watershed. The 3B models included three sets of fixed and random effect models with baseline controls and pairs of the green infrastructure locational measures and the spatial pattern measure. The locational measures included the percentage of green infrastructure in the 100-year floodplain, in the 60-meter buffer around the floodplain, and outside the floodplain with each model including appropriate spatial pattern (patch density) measures associated with each GI location measure. Finally, the 4B models pushed the analysis a bit by presenting pairs of the locational measures and the spatial pattern measures.

Focusing first on the 2B Models, I can see that the controls again worked as expected. In the fixed effect model the two time variant indicators of precipitation and imperviousness indeed had significant effects on peak annual flow. Precipitation had a significant positive effect on peak annual flow, suggesting that every millimeter of precipitation increased peak annual flow by .40 percent. Also, every percentage point of impervious surface in a watershed also increased peak annual flow by 29%. The random effects model suggests that watershed and drainage area also had significant positive effects while soil permeability significantly reduced peak annual flow across watersheds through time. Every square mile of watershed area increased peak flow by 1.04%, while every square mile of drainage area increased peak flow by 0.008%. On the whole, these patterns of findings of control measures tended to hold across all other model sets, with the exception soil permeability and slope. Soil permeability tended to become insignificant once green infrastructure measures were introduced into the model.

Hausman Test associated with each set of panel models estimated were also included in Table 9. Again I find that all of these tests were statistically significant, suggesting that there were indeed significant differences between these sets of estimates across all model sets. For some, these results suggest that the “fixed” effect results were to be preferred. However, since random effect panel data analysis allows for theoretically significant and substantive time-invariant variables (watershed area, drainage area, floodplain percentages, soil permeability, slope percentages, and stream density) to be considered as explanatory/control variables both sets of models are presented.

The 2B models added to the baseline models the percentage of watershed area associated with green infrastructure and patch density as a spatial pattern of green infrastructure in the watershed. Adding overall green infrastructure and patch density in the watershed significantly increased the respective R^2's of the models, in comparison to the appropriate base model. Whether considering the fixed or random effects model, in each case the consequences of green infrastructure was, statistically significant and negative, having controlled for the baseline factors. These findings are consistent with the first hypothesis that overall green infrastructure should significantly reduce peak annual flow, and the fifth hypothesis that higher the patch density should significantly reduce peak annual flow. However, the magnitudes of the effects of green infrastructure differed considerably between two models. In the fixed effect model, the results suggest that the peak annual flow decreased by 12.2% for every percent increase in overall green infrastructure, while in the random effect model there was a 5.6% decrease for every percent increase in overall green infrastructure. The results of spatial pattern, patch density shows the magnitudes of the effects of green infrastructure also differed between the fixed and random effect models. In the fixed effect model, the results suggest that the peak annual flow decreased by 0.42% for every number of GI patch increase per 100 hectare, while in the random effect model there was a 0.19% decrease for every number of GI patch increase per 100 hectare.

The 3B and 4B models assessed the consequences of the specific locations of green infrastructure within the watershed location variables and patch density as one of the spatial pattern variables of green infrastructure on annual peak flow. As noted above given, multicollinearity issues I have first presented three sets with only one of the locational variables in the model and the spatial pattern variables, then a set (4B models) with two of the locational measures and the spatial pattern measures. The first model set in 3B included only the percent of green infrastructure in the 100-year floodplain and patch density of green infrastructure in the 100-year floodplain. Only the percentage of green infrastructure in the 100-year floodplain was negative and statistically significant in the fixed effect model ($P < .05$). The effect in the fixed effects model suggests that with every percent increase in green infrastructure within the 100-year floodplain, peak annual flow decreased by 8.0%. As can be seen in the next set of models, the effect of green infrastructure outside the floodplain appeared to be statistically significant and negative in both the fixed and random effects models, although in the random effects model the effects were only marginally significant at the .1 level of significance. The effect in the fixed effects model suggests that with every percent increase in green infrastructure outside the 100-year floodplain decreased 12.6%, and every number of GI patch increase per 100 hectare decreased 0.42%. Controlling for time invariant measures reduced the effect to 5.0% and 0.16%, respectively.

The final set in the 3B series included measures of green infrastructure within a 60-meter buffer around the floodplain. In this set, the coefficient for green infrastructure was only significant in the fixed effects model. In the fixed effects model the percent of GI in the 60-meter buffer around the floodplain had a significant negative effect, as anticipated, and the coefficient's magnitude suggests that with every percentage point increase in GI in this buffer, peak annual flows decreased by 9.8%. Patch density of GI in the 60-meter buffer around the floodplain also had a significant negative effect and the coefficient's magnitude suggests that with every number of GI patch increase per

100 hectare in this buffer, peak annual flows decreased by 0.08%. However, when other time invariant measures were controlled for as they were in the random effects model, both the %GI in the buffer and its density became insignificant.

The 4B models again pushed the analysis including a set of GI locational measures and spatial pattern measures within the same model. In this case the 4B set included both green infrastructure in the 100-year and outside of the 100-year floodplain as well as their associated patch density measures. Fixed effects model suggests that only patch density of green infrastructure outside the 100-year floodplain had a negative and statistically significant effect, with a 0.28% reduction in peak flow for every number of GI increase per 100 hectare outside the 100-year floodplain. On the other hand, random effects model suggests that green infrastructure outside the 100-year floodplain and patch density of green infrastructure outside the 100-year floodplain had a negative and statistically significant effect, although the former was only marginally significant. The effect in the random effects model suggests that with every percent increase in green infrastructure outside the 100-year floodplain decreased 4.0%, and every number of GI patch increase per 100 hectare decreased 0.22%.

Table 10 presents a series of fixed and random effect panel models assessing the consequences of green infrastructure on peak annual flow, focusing on the overall percentage of green infrastructure within a watershed and with respect to the location of green infrastructure relative to the floodplain, and the one of spatial patterns of green infrastructure, GYRATE_AM. The pattern of equations in this table was similar to those in Table 9 in that each set of models included pairs of the locational measures and the spatial pattern of green infrastructure. The 2C Models included a set of fixed and random effect models predicting peak annual flow with the baseline controls and a pair of a measure of the overall percentage of green infrastructure within the watershed and a measure of spatial pattern (GYRATE_AM), within the watershed. The 3C models included three sets of fixed and random effect models with baseline controls and pairs of the green infrastructure locational measures again with their appropriate GYRATE_AM measure. Finally, the 4C models pushed the analysis utilizing pairs of the locational measures and the spatial pattern measures.

The 2C models added to the baseline models the percentage of watershed area associated with green infrastructure and GYRATE_AM as the spatial pattern of green infrastructure in the watershed. Whether considering the fixed or random effects model, in each case the consequences of green infrastructure in the watershed was, statistically significant and negative, having controlled for the baseline factors. These findings are consistent with the first hypothesis that overall green infrastructure should significantly reduce peak annual flow. As I have seen in previous analyses, the magnitudes of the effects of green infrastructure differed. In the fixed effect model, the results suggest that the peak annual flow decreased by 9.7% for every percent increase in overall green infrastructure, while in the random effect model there was a 3.2% decrease for every percent increase in overall green infrastructure. Unlike the case with the patch density measures GYRATE_AM was statistically insignificant in both models.

The 3C models assessed the consequences of the specific locations of green infrastructure within the watershed location variables and GYRATE_AM as one of the spatial pattern variables of green infrastructure on annual peak flow. Again, because of multicollinearity issues have followed the same pattern as before, this model estimated three separate models with sets of locational measures with their appropriate spatial pattern measure. The first model set in 3C included only the percent of green infrastructure in the 100-year floodplain and GYRATE_AM in the 100-year floodplain. Only the percentage of green infrastructure in the 100-year floodplain was negative and statistically significant in the fixed effect model ($P < .05$), suggesting that every percent increase in green infrastructure within the 100-year floodplain, peak annual flow decreased by 5.9%. However, it was not significant in the random effects model. Furthermore and most important for these models, the spatial pattern variables (GYRATE_AM) was not statistically significant in either model. As can be seen in the next set of models, the effect of green infrastructure outside the floodplain again was statistically significant and negative in both the fixed and random effects models. The effect in the fixed effects model suggest that with every percent increase in green infrastructure outside the 100-year floodplain decreased 10.1%. The peak annual flow decreased by 0.27% for every one meter distance increase in GI patch between each cell in a patch and the patch's centroid. Controlling for time invariant measures reduced the effect to 4.3% and 0.13%, respectively. Most interestingly, and perhaps disconcertingly, the spatial effects of GYRATE_AM were statistically significant in these models, albeit only marginally significant in the fixed effect model, but their effects were positive. The final set in this series included measures of green infrastructure within a 60-meter buffer around the floodplain. In this set, the coefficient for green infrastructure was only significant in the fixed effects model. In that model the percent of GI in the 60-meter buffer around the floodplain had a significant negative effect, as anticipated, and the coefficient's magnitude suggest that with every percentage point increase in GI in this buffer, peak annual flows decreased by 10.2%. However, when other time invariant measures were controlled for in the random effects model, the GI within the 60-meter buffer became insignificant.

The 4C models included a set of GI locational measures and spatial pattern measures within the same model. This included both green infrastructure in the 100-year floodplain and outside of the 100-year floodplain variables. Fixed effects model suggests that green infrastructure in the 100-year floodplain had a negative and statistically significant effect, with an 11.1% reduction in peak flow for every percentage point in GI increase in the 100-year floodplain. On the other hand, random effects model suggests that green infrastructure outside the 100-year floodplain and GYRATE_AM of green infrastructure in and outside the 100-year floodplain had a negative and statistically significant effect. The effect in the random effects model suggest that with every percent increase in green infrastructure outside the 100-year floodplain decreased 3.9%, and every one meter distance increase in GI patch between each cell in a patch and the patch's centroid increased 0.18%.

Fixed and Random Panel Model Predicting Peak Annual Flow

	1A Models		2A Models		3A Models						4A Models	
Baseline control variables	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random
Watershed Area		.0026***		.0023***		.0027***		.0022**		.0027***		.0025**
		(.0009)		(.0009)		(.0009)		(.0009)		(.0008)		(.0010)
Drainage Area		.00005***		.00006***		.00006***		.00006***		.00006***		.00006***
		(.00001)		(.00001)		(.00001)		(.00001)		(.00001)		(.00001)
Floodplain %		.0146		.0183		.0117		.0161		.0107		.0129
		(.0116)		(.0119)		(.0124)		(.0115)		(.0121)		(.0125)
Precipitation	.0054***	.0040***	.0055***	.0043***	.0052***	.0041***	.0056***	.0043***	.0053***	.0041***	.0049***	0.0042***
	(.0007)	(.0004)	(.0007)	(.0005)	(.0006)	(.0004)	(.0007)	(.0005)	(.0006)	(.0005)	(.0008)	(.0004)
Soil permeability		-.0110**		-.0056		-.0077		-.0063		-.0083*		-.0066
		(.0045)		(.0051)		(.0047)		(.0052)		(.0049)		(.0053)
Impervious %	.3576***	.2072***	.2202*	.0137**	.2179**	.0163***	.2314**	.0147***	.1928*	.0174***	.2605**	.0150***
	(.1127)	(.0046)	(.1105)	(.0056)	(.1048)	(.0051)	(.1105)	(.0055)	(.1051)	(.0055)	(.1039)	(.0056)
Slope %		.0567		.1216*		.1058*		.1113*		.0872		.1161*
		(.0578)		(.0654)		(.0614)		(.0640)		(.0670)		(.0643)
Stream Density		.2773		.0940		.1042		.1347		.2214		.0867
		(.2403)		(.2452)		(.2562)		(.2356)		(.2473)		(.2495)
Overall GI variable												
Percentages of GI												
in the watershed			-.0745***	-.0292**								
			(.0220)	(.0122)								
GI Location variables												
Percentages of GI												
in the 100-year floodplain					-.0799***	-.0214**					-.1148***	-.0169
					(.0188)	(.0092)					(.0377)	(.0127)
outside the 100-year floodplain							-.0737**	-.0253**			.0604	-.0096
							(.0239)	(.0123)			(.0501)	(.0171)
60 meter buffer around floodplain									-.0931***	-.0155		
									(.0211)	(.0145)		
Constant	-8.5670*	3.2013***	.4280	4.9967***	1.7720	4.9270***	-.1997	4.6975***	2.8627	4.2186***	-.5778	5.1300***
	(4.3842)	(.7053)	(4.7297)	(.9407)	(4.5300)	(.9493)	(4.7144)	(.9106)	(4.4093)	(1.1320)	(4.51093)	(.9492)
N	132	132	132	132	132	132	132	132	132	132	132	132
R-squared												
within	0.6057	0.5633	0.6519	0.6035	0.6985	0.6148	0.6447	0.5943	0.6865	0.5940	0.7071	0.6151
between		0.5665		0.5351		0.5151		0.5384		0.5259		0.5165
overall		0.5644		0.5785		0.5787		0.5745		0.5697		0.5795
Hausman Test: Prob (χ2)	0.0308		0.0069		0.0000		0.0085		0.0000		0.0000	

Notes: Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 8. Fixed and Random Effect Panel Model Predicting Peak Annual Flow 1

Fixed and Random Panel Model Predicting Peak Annual Flow

	2B Models		3B Models						4B Models	
Baseline control variables	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random
Watershed Area		.0104***		.0027		.0062**		.0023		.0072**
		(.0032)		(.0024)		(.0029)		(.0023)		(.0028)
Drainage Area		.00008***		.00006***		.00005***		.00006***		.00005***
		(.00002)		(.00001)		(.00001)		(.00001)		(.00001)
Floodplain %		.0129		.0114		.0074		.0107		.0029
		(.0126)		(.0126)		(.0127)		(.0122)		(.0138)
Precipitation	.0040***	.0034***	.0052***	.0043***	.0039***	.0036***	.0048***	.0041***	.0039***	.0036***
	(.0008)	(.0005)	(.0008)	(.0005)	(.0008)	(.0005)	(.0007)	(.0005)	(.0009)	(.0005)
Soil permeability		-.0085**		-.0066		-.0073		-.0085*		-.0060
		(.0037)		(.0047)		(.0048)		(.0050)		(.0050)
Impervious %	.2546**	.0147**	.2289**	.0172***	.2644**	.0149**	.2094*	.0170***	.2846**	.0175***
	(.1221)	(.0061)	(.1106)	(.0048)	(.1185)	(.0062)	(.1057)	(.0055)	(.1157)	(.0059)
Slope %		.1494**		.0984		.1244*		.0853		.1206*
		(.0746)		(.0607)		(.0708)		(.0724)		(.0689)
Stream Density		.0796		.1564		.2061		.2283		.2747
		(.2870)		(.2523)		(.2664)		(.2499)		(.2692)
Overall GI variable										
Percentages of GI										
in the watershed	-.1302***	-.0574**								
	(.0262)	(.0226)								
GI Location variables										
Percentages of GI										
in the 100-year floodplain			-.0832**	-.0067					-.0993	.0080
			(.0416)	(.0135)					(.0606)	(.0209)
outside the 100-year floodplain					-.1348***	-.0510**			.0208	-.04040*
					(.0251)	(.0206)			(.0568)	(.0217)
60 meter buffer around floodplain							-.1033***	-.0152		
							(.0213)	(.0147)		
GI Spatial Patterns variables										
PD (Patch Density)										
in the watershed	-.0042***	-.0019*								
	(.0014)	(.0010)								
in the 100-year floodplain			-.0002	.0010					.0007	.0018
			(.0020)	(.0007)					(.0021)	(.0011)
outside the 100-year floodplain					-.0043***	-.0016*			-.0028**	-.0023**
					(.0013)	(.0008)			(.0014)	(.0009)
60 meter buffer around floodplain							-.0008*	-.00001		
							(.0005)	(.00035)		
Constant	4.8992	7.8787***	1.5469	3.2166*	5.0704	7.2884***	4.4306	4.2567***	1.1351	5.4752***
	(5.3019)	(1.9558)	(6.2753)	(1.3227)	(4.9833)	(1.7456)	(4.6506)	(1.4729)	(6.2023)	(2.0175)
N	122	122	130	130	126	126	130	130	126	126
R-squared										
within	0.6637	0.6058	0.6977	0.6160	0.6794	0.6110	0.7052	0.5914	0.7292	0.6465
between		0.5529		0.5130		0.5446		0.5197		0.5276
overall		0.5870		0.5788		0.5870		0.5659		0.6029
Hausman Test: Prob ($\chi 2$)	0.0101		0.0001		0.0021		0.0000		0.0002	

Notes: Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 9. Fixed and Random Effect Panel Model Predicting Peak Annual Flow 2

Fixed and Random Panel Model Predicting Peak Annual Flow

	2C Models		3C Models						4C Models	
Baseline control variables	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random
Watershed Area		.0077**		.0032		.0038		.0023		.0071**
		(.0032)		(.0028)		(.0025)		(.0025)		(.0032)
Drainage Area		.00009***		.00007***		.00006***		.00006***		.00007***
		(.00002)		(.00001)		(.00001)		(.00001)		(.00001)
Floodplain %		.0153		.0205		.0118		.0106		.0290**
		(.0120)		(.0146)		(.0117)		(.0115)		(.0140)
Precipitation	.0058***	.0042***	.0049***	.0040***	.0056***	.0042***	.0054***	.0041***	.0046***	.0040***
	(.0008)	(.0005)	(.0008)	(.0004)	(.0008)	(.0005)	(.0007)	(.0005)	(.0009)	(.0005)
Soil permeability		-.0078**		-.0062		-.0063		-.0084*		-.0020
		(.0038)		(.0050)		(.0046)		(.0050)		(.0051)
Impervious %	.2351**	.0181***	.1985	.0137**	.2112*	.0217***	.2013*	.0171***	.2396*	.0182***
	(.1167)	(.0062)	(.1235)	(.0059)	(.1217)	(.0062)	(.1077)	(.0056)	(.1233)	(.0062)
Slope %		.1233*		.0953		.0969		.0838		.0906
		(.0682)		(.0608)		(.0729)		(.0664)		(.0719)
Stream Density		.0814		.1452		.2313		.2289		.2605
		(.2544)		(.2517)		(.2578)		(.2488)		(.2626)
Overall GI variable										
Percentages of GI										
in the watershed	-.1022***	-.0326**								
	(.0327)	(.0134)								
GI Location variables										
Percentages of GI										
in the 100-year floodplain			-.0604**	-.0131					-.1179**	.0041
			(.0241)	(.0115)					(.0458)	(.0148)
outside the 100-year floodplain					-.1065***	-.0437***			.0550	-.0394**
					(.0358)	(.0145)			(.0600)	(.0193)
60 meter buffer around floodplain							-.1070***	-.0155		
							(.0248)	(.0170)		
GI Spatial Patterns variables										
GYRATE_AM (Correlation Length)										
in the watershed	.0012	.0002								
	(.0008)	(.0002)								
in the 100-year floodplain			-.0013	-.0006					-.0005	-.0013**
			(.0012)	(.0005)					(.0013)	(.0006)
outside the 100-year floodplain					.0027*	.0013**			.0015	.0018***
					(.0021)	(.0006)			(.0020)	(.0007)
60 meter buffer around floodplain							.0027	.00006		
							(.0028)	(.00187)		
Constant	-.4107	4.8126***	1.9915	4.5608***	.6648	4.6803***	2.7198	4.2251***	-.0532	4.3418***
	(5.3174)	(.9393)	(5.2306)	(.9481)	(5.4379)	(.9405)	(4.6475)	(1.1260)	(5.5106)	(1.0663)
N	122	122	130	130	126	126	130	130	126	126
R-squared										
within	0.6309	0.5771	0.7067	0.6206	0.6454	0.5963	0.7168	0.5912	0.6913	0.6383
between		0.5598		0.5050		0.5612		0.5201		0.5782
overall		0.5709		0.5791		0.5855		0.5659		0.6145
Hausman Test: Prob (χ2)	0.0348		0.0000		0.0383		0.0000		0.0009	

Notes: Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 10. Fixed and Random Effect Panel Model Predicting Peak Annual Flow 3

5.2.4 Regional Variations in Green Infrastructure Effects on Annual Peak Flow

Before proceeding to my examination for the next dependent variable, mean annual flows, it may be important to consider if the effects found for green infrastructure are equivalent across the two principal sample areas. Specifically, as noted above, I have drawn my sample of watersheds from two major metropolitan areas in Texas that are subject to flooding, the Austin and Houston metropolitan areas. As noted above, these two areas vary considerably with the terrain characteristics such as slope and soil types which may have consequences for the effectiveness of GI when addressing annual peak flow.

Table 11 presents the five random effect panel models assessing the consequences of the percentage of green infrastructure in the watershed and for specific locations on peak annual flow that were originally presented in Table 8. Each model now included the Houston dummy variable and the associated Houston GI interaction variables. The last row on this table (in blue) presents the statistical test for the joint effects of adding the dummy and associated interaction term(s) to the model, again implying significant variations between the two metropolitan areas with respect to how GI is working in their watersheds.[5]

The first model (2A) in Table 11 assessed variations in the consequences for the overall percent of green infrastructure within the watershed between the Houston and Austin. Not surprisingly, since the Houston dummy and interaction coefficients were not

[5] It is safer to perform the overall F-test or a test for the significance of the combined joint effects of including the dummy and associated interaction(s), rather than depending on individual t-test for the dummy and interaction terms because one or more multiple t-test may be significant due to random error and multicollinearity issues may obscure the individual t-test as well.

significant, the test for the joint effects was not significant. This finding indicates that the significant negative effect of %GI in the watershed (see Table 8) held in both Houston and Austin. The 3A models tested for variations in locational aspects of GI between the two regions. The tests for whether or not the dummy and interaction term were jointly significant in both the first and third models in the 3A set were significant (Prob(χ^2) < .05) suggesting that there were differences in the way the %GI performs in the floodplain and within the 60-meter buffer. The first model in this series suggests that the effect of %GI in the floodplain for the Austin was significant and negative; with a 8.6% decrease in annual peak flow for every percent increase in GI within the floodplain.[6] However, the interaction coefficient for differential effect of %GI in the floodplain for the Houston area was positive and significant, resulting in the net effect being .000475, which was not significant.[7] In the third model the effect of %GI in the 60-meter buffer for the Austin area was again significant and negative suggesting a 7.8% reduction in annual peak flow for every percent increase in GI within the 60-meter buffer around the floodplain. However, yet again the interaction coefficient for the differential effect in Houston was positive and significant (.0859), yielding a net effect (.0051), which was not significant.[8] The results for the final model (4A) simply substantiated the findings of the first model in 3A, only now controlling for the %GI outside the floodplain, in that there was a significant reduction in annual peak flow of

[6] $100(e^{-.09008712} - 1) = -8.6\%$

[7] The net coefficient is -0.0900872 + 0.09056208 = .00047496. Its standard error is .0075974, which is not significant.

[8] The net coefficient is -.08084283 + .08589418 = .00505135 with a standard error of .0127549, which is not significant.

9.8% for every % increase in GI within the floodplain in Austin, but the net effect in the Houston area was not significant.

Table 12 presents random effect panel models assessing the consequences of green infrastructure on peak annual flow with a regional dummy variable (Houston), focusing on the percent green infrastructure within a watershed, with respect to the location of green infrastructure relative to the floodplain and the spatial pattern related the patch density of green infrastructure. Each set of models included the standard set of controls plus the percent of GI within the watershed or for specific locations within the watershed, associated patch density measure of each location, the regional dummy (Houston) and a set of interaction terms between the regional dummy and each GI locational and patch density measure. Unlike the results for the percent GI models presented in Table 11, the test for joint effects were consistent for each of these models, implying that the consequences of GI, when both location and patch density are included in models, were significantly different between the two regions. Indeed, also unlike Table 11, in these models the dummy variable’s coefficients were all significant and negative indicating that there were substantial differences in annual peak flow between these two regions after controlling for the other factors in these models.

The overall findings for the analyses presented in Table 12, suggest when patch density was controlled for, the substantial and significant negative consequences for green infrastructure persisted in watersheds located in the Austin region, however the effects essentially disappeared for watersheds in the Houston region. More specifically, whether considering the %GI in the watershed or in various locations relative to the floodplain (in, out, or in a 60-meter buffer) the level coefficients associated with these measures were always statistically significant and negative, indicating there was always a negative impact on annual peak flow. Furthermore, for watersheds in the Austin region, there were additional benefits in reducing peak annual flow by increasing the patch density for GI in the entire watershed (see model 2B) and outside the watershed (see second model in the 3B series). These findings were also supported by the results in Model 4B.

On the other hand, the interaction terms associated with the %GI and patch density measures, which assessed for differential effects of each of these measures for Houston's watersheds were often statistically significant and positive. The resulting net effects for GI were consequently zeroed out or left with a slight positive effect. For example, in model 2B the level effect for %GI in the watershed was -.14014048, while the interaction term's coefficient was .16492852, yielding a net effect of 0.02478804. This was marginally significant at the .1 level, suggesting an increase of 2.5% per percent of GI in Houston. Furthermore, the consequences of the patch density for the GI within the watershed yielded similar findings for model 2B. The statistically significant level coefficient was -.0048, indicating a negative effect for patch density in Austin's watersheds, but the statistically significant interaction coefficient was .0073, yielding a statistically significant net coefficient of .0026, suggesting that annual peak flow increased by .26% for every percent increase in patch density. Similar patterns held for the 3B models and were generally supported by the findings associated with model 4B.

Table 13 presents random effect panel models assessing the consequences of green infrastructure on annual peak flow with a regional dummy variable (Houston), focusing on the percent green infrastructure within a watershed, with respect to the location of green infrastructure relative to the floodplain and the one of spatial patterns of green infrastructure, GYRATE_AM. The pattern of equations in this table was similar to those in Table 12 in that each set of models included the standard set of controls plus the percent of GI within the watershed or for specific locations within the watershed, associated GYRATE_AM measure of each location, the regional dummy (Houston) and a set of interaction terms between the regional dummy and each GI locational and GYRATE_AM measure. The statistical test for the joint effects were consistent for each of these models except 2A Model. It implies that significant variations between the two metropolitan areas with respect to how GI was working, when both location and GYRATE_AM were included in models. The dummy variable's coefficients in these models except 2A Model were also all significant and negative indicating that substantial differences in annual peak flow between these two regions after controlling for the other factors in these models are presented.

The overall findings for the analyses presented in Table 13, suggest when GYRATE_AM was controlled for, the significant negative consequences for green infrastructure persisted in watersheds located in the Austin region. However, the effects essentially disappeared for in the 100-year floodplain and the 60-meter buffer around the floodplain in the Houston region. More specifically, whether considering the %GI in the watershed or in various locations relative to the floodplain (in, out, or in a 60-meter buffer) the level coefficients associated with these measures area always were statistically significant and negative, indicating there was always a negative impact on annual peak flow. However, for watersheds in the Austin region, there were not any additional benefits in reducing annual peak flow by increasing GYRATE_AM for GI.

The interaction terms associated with the %GI and GYRATE_AM measures assessed for differential effects of each of these measures for Houston's watersheds. The results of the interaction terms were statistically significant and positive in the 100-year floodplain and 60-meter buffer around the floodplain (see first and third models in the 3B series). The resulting net effects for GI were consequently zeroed out, however were not significant. For example, the first model in the 3B series the level effect for %GI in the 100-year floodplain was -.0730884, while the interaction term's coefficient was .07225284, resulting in the net effect being -.0008356, which was not significant.[9] In the third model the effect of %GI in the 60-meter buffer for the Austin area was again significant and negative suggesting a 9.5% reduction in annual peak flow for every percent increase in GI within the 60-meter boundary. However, yet again the interaction coefficient for the differential effect in Houston was positive and significant (.0887), yielding a net effect (-.0108), which was not significant.[10] For watersheds in the Houston region, there were not any significant additional benefits in reducing annual peak flow by increasing GYRATE_AM for GI.

[9] The net coefficient is -.0730884+ .07225284= -.0008356. Its standard error is .0104309, which is not significant.
[10] The net coefficient is -.09945667 + .0886632 = -.0107935 with a standard error of .0202912, which is not significant.

Random Panel Data Analysis with a Dummy Variable Interaction Predicting Peak Annual Flow

	2A Model	3A Models			4A Model
Baseline control variables					
Watershed Area	.0023**	.0023**	.0022**	.0030***	.0021*
	(0.0009)	(0.0011)	(0.0010)	(0.0008)	(0.0012)
Drainage Area	.00006***	.00006***	.00006***	.00006***	.00006***
	(0.00001)	(0.00001)	(0.00001)	(0.00001)	(0.00001)
Floodplain %	.0249*	.0248**	.0214*	.0256**	.0234*
	(0.0127)	(0.0121)	(0.0124)	(0.0108)	(0.0124)
Precipitation	.0042***	.0036***	.0043***	.0038***	.0036***
	(0.0005)	(0.0004)	(0.0005)	(0.0005)	(0.0005)
Soil permeability	-.0068	-.0104*	-.0062	-.0110*	-.0078
	(0.0059)	(0.0054)	(0.0061)	(0.0057)	(0.0067)
Impervious %	.0151***	.0212***	.0156***	.0221***	.02180***
	(0.0058)	(0.0055)	(0.0058)	(0.0055)	(0.0061)
Slope %	.1384	.1248*	.1118	.1380*	.0754
	(0.0892)	(0.0699)	(0.0904)	(0.0813)	(0.0861)
Stream Density	.0424	.1916	.0961	.2133	.3116
	(0.2317)	(0.2166)	(0.2300)	(0.2109)	(0.2353)
Regional dummy variable					
Houston					
	-2.2339	-6.3230***	-1.1326	-5.5780***	-5.2707***
	(1.4629)	(1.8484)	(1.5462)	(1.9897)	(2.0368)
Overall GI variable					
GI% in Watershed	-.0554**				
	(0.0233)				
GI% in Watershed*Houston	.0361				
	(0.0278)				
GI Location variables					
GI% in Floodplain		-.0901***			-.1033***
		(0.0242)			(0.0350)
GI% in Floodplain*Houston		.0906***			.1141***
		(0.0262)			(0.0356)
GI% out of Floodplain			-.0378		.0317
			(0.0248)		(0.0352)
GI% out of Floodplain*Houston			.0167		-.0493
			(0.02880)		(0.0394)
GI% in Floodplain Buffer				-.0808***	
				(0.0281)	
GI% in Floodplain Buffer*Houston				.0859***	
				(0.0319)	
Constant	6.6494***	9.6250***	5.5217***	8.3490***	8.5800***
	(1.6064)	(1.9279)	(1.6258)	(2.1334)	(1.9524)
N	132	132	132	132	132
R-squared					
within	0.5991	0.6750	0.5904	0.6225	0.6953
between	0.5569	0.5218	0.5503	0.5561	0.5041
overall	0.5842	0.6185	0.5763	0.5987	0.6245
Test for Joint Effect: Prob (χ2)	0.2752	0.0023	0.7032	0.0195	0.0101

Note: This model includes a dummy variable(Houston), regional dummy variable. It equlas one for Houston, and zero for Austin.
Test for joint effect is for testing joint effect of a regional dummy variable and a dummy interaction.
Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 11. Random Effect Panel Model with a Dummy Variable Predicting Peak Annual Flow Phase 1

Random Panel Data Analysis with a Dummy Variable Interaction Predicting Peak Annual Flow

	2B Model		3B Models		4B Model
Baseline control variables					
Watershed Area	.0088**	-.0010	.0059*	.0021	.0051*
	(0.0038)	(0.0024)	(0.0027)	(0.0023)	(0.0029)
Drainage Area	.00007***	.00007***	.00005***	.00005***	.00006***
	(0.00002)	(0.000008)	(0.00001)	(0.000007)	(0.00002)
Floodplain %	.0109	.0235*	.0100	.0289**	.0143
	(0.0128)	(0.0135)	(0.014)	(0.0126)	(0.0131)
Precipitation	.0039***	.0039***	.0040***	.0041***	.0034***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Soil permeability	-.0109**	-.0076	-.0098**	-.0130***	-.0111**
	(0.0047)	(0.0065)	(0.0048)	(0.0050)	(0.0055)
Impervious %	.0155**	.0207***	.0150**	.0232***	.0196***
	(0.0069)	(0.0056)	(0.0067)	(0.0055)	(0.0062)
Slope %	.1871**	.0907	.1657*	.1537**	.1168
	(0.0931)	(0.0680)	(0.0885)	(0.0752)	(0.0781)
Stream Density	.1366	.3497	.1837	.2711	.2767
	(0.2507)	(0.2482)	(0.2319)	(0.2279)	(0.2503)
Regional dummy variable					
Houston					
	-13.2989***	-13.7455***	-13.1407***	-9.0939***	-19.3203***
	(1.6017)	(4.2832)	(1.6420)	(2.0210)	(3.9244)
Overall GI variable					
GI% in Watershed	-.1401***				
	(0.0247)				
GI% in Watershed*Houston	.1649***				
	(0.0235)				
GI Location variables					
GI% in Floodplain		-.1263**			-.1484***
		(0.0519)			(0.0556)
GI% in Floodplain*Houston		.1655***			.1571***
		(0.0497)			(0.0572)
GI% out of Floodplain			-.1349***		-.0484
			(0.0244)		(0.0422)
GI% out of Floodplain*Houston			.1616***		.0705
			(0.0242)		(0.0444)
GI% in Floodplain Buffer				-.0757***	
				(0.0286)	
GI% in Floodplain Buffer*Houston				.0899***	
				(0.0304)	
GI Spatial Patterns variables					
PD (Patch Density)					
PD in Watershed	-.0047***				
	(0.0009)				
PD in Watershed*Houston	.0073***				
	(0.0010)				
PD in Floodplain		-.0029			-.0038
		(0.0025)			(0.0025)
PD in Floodplain*Houston		.0054**			.0038
		(0.0025)			(0.0028)
PD out of Floodplain			-.0044***		-.0028***
			(0.0008)		(0.0009)
PD out of Floodplain*Houston			.0071***		.0049***
			(0.0009)		(0.0014)
PD in Floodplain Buffer				-.0008	
				(0.0005)	
PD in Floodplain Buffer*Houston				.0019***	
				(0.0005)	
Constant	14.1482***	13.0892***	13.7661***	9.2951***	20.0120***
	(2.0408)	(4.8342	(1.976)	(2.1111)	(4.2493)
N	122	130	126	130	126
R-squared					
within	0.7736	0.7431	0.7788	0.7002	0.8190
between	0.5967	0.4823	0.6061	0.5590	0.6099
overall	0.7127	0.6442	0.7181	0.6485	0.7463
Test for Joint Effect: Prob (χ2)	0.0000	0.0058	0.0000	0.0000	0.0000

Note: This model includes a dummy variable(Houston), regional dummy variable. It equlas one for Houston, and zero for Austin.
Test for joint effect is for testing joint effect of a regional dummy variable and a dummy interaction. Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 12. Random Effect Panel Model with a Dummy Variable Predicting Peak Annual Flow Phase 2

Random Panel Data Analysis with a Dummy Variable Interaction Predicting Peak Annual Flow

	2C Model	3C Models			4C Model
Baseline control variables					
Watershed Area	.0064**	.0023	.0030	.0018	.0056
	(0.0032)	(0.0028)	(0.0023)	(0.0025)	(0.0034)
Drainage Area	.00008***	.00006***	.00005***	.00006***	.00006***
	(0.00002)	(0.00001)	(0.00001)	(0.00001)	(0.00002)
Floodplain %	.0227*	.0267*	.0286**	.02382**	.0380***
	(0.0124)	(0.0137)	(0.0126)	(0.0104)	(0.0133)
Precipitation	.0042***	.0036***	.0042***	.0038***	.0034***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Soil permeability	-.0092**	-.0094*	-.0054	-.0102*	-.0050
	(0.0046)	(0.0057)	(0.0055)	(0.0058)	(0.0059)
Impervious %	.0202***	.0200***	.0256***	.0235***	.02884***
	(0.0070)	(0.0063)	(0.0065)	(0.0056)	(0.0073)
Slope %	.1331	.1169	.0737	.1096	.0446
	(0.0915)	(0.0771)	(0.0890)	(0.0783)	(0.0843)
Stream Density	.0490	.2040	.1529	.2121	.3636
	(0.2483)	(0.2262)	(0.2363)	(0.2174)	(0.2590
Regional dummy variable					
Houston	-2.5088	-5.6693***	-3.3195**	-6.3287***	-6.9082***
	(1.6707)	(1.7006)	(1.5702)	(2.3447)	(1.6429)
Overall GI variable					
GI% in Watershed	-.0647**				
	(0.0269)				
GI% in Watershed*Houston	.04040				
	(0.0321)				
GI Location variables					
GI% in Floodplain		-.0731***			-.0947**
		(0.0257)			(0.0371)
GI% in Floodplain*Houston		.0723***			.1086***
		(0.0276)			(0.0389)
GI% out of Floodplain			-.08415***		-.0043
			(0.0260)		(0.0420)
GI% out of Floodplain*Houston			.0472		-.0284
			(0.0322)		(0.0460)
GI% in Floodplain Buffer				-.0995***	
				(0.0361)	
GI% in Floodplain Buffer*Houston				.0887**	
				(0.0441)	
GI Spatial Patterns variables					
GYRATE_AM (Correlation Length)					
GYRATE_AM in Watershed	.0003				
	(0.0003)				
GYRATE_AM in Watershed*Houston	-.00002				
	(0.0003)				
GYRATE_AM in Floodplain		-.0010			-.0013
		(0.0012)			(0.0011)
GYRATE_AM in Floodplain*Houston		.0010			.0007
		(0.0012)			(0.0010)
GYRATE_AM out of Floodplain			.0017**		.0017
			(0.0008)		(0.0011)
GYRATE_AM out of Floodplain*Houston			-.0001		.00003
			(0.0008)		(0.0012)
GYRATE_AM in Floodplain Buffer				.0020	
				(0.0023)	
GYRATE_AM in Floodplain Buffer*Houston				.0021	
				(0.0032)	
Constant	6.7254***	9.0222***	7.0917***	9.3630***	9.5642***
	(1.6999)	(1.6588)	(1.6725)	(2.4208)	(1.8425)
N	122	130	126	130	126
R-squared					
within	0.5701	0.6944	0.5900	0.6297	0.7321
between	0.5926	0.5031	0.6362	0.5524	0.5781
overall	0.5775	0.6241	0.6056	0.6020	0.6738
Test for Joint Effect: Prob (χ2)	0.4538	0.0003	0.0626	0.0067	0.0000

Note: This model includes a dummy variable(Houston), regional dummy variable. It equlas one for Houston, and zero for Austin.
Test for joint effect is for testing joint effect of a regional dummy variable and a dummy interaction. Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 13. Random Effect Panel Model with a Dummy Variable Predicting Peak Annual Flow Phase 3

5.2.5 Summary of Effectiveness of Green Infrastructure on Peak Annual Flow

The series of fixed and random effect panel models assessing the consequences of green infrastructure on annual peak flow suggests that the percent of green infrastructure in the watershed was a significant negative determinant of annual peak flow. Furthermore, there was evidence that its locational features with respect to the floodplain may well have consequences, and spatial attribute related to patch density were significant as well. However, the GRYATE measure did not seem to have much consequence in the majority of the models predicting annual peak flow.

Random effect panel models assessing regional variations in green infrastructure effects on annual peak flow suggest that while there did not appear to be differences in the consequences of %GI in the overall watersheds between Austin and Houston areas, in that green infrastructure significantly reduced peak annual flows in both areas, there were some variations with respect to GI's locational features. In particular, the %GI in the floodplain and in the buffer around the floodplain appeared to be more effective in watersheds located in the Austin region in comparison to the Houston region. Specifically, these findings suggest that the consequences of green infrastructure whether considering the percent in the entire watershed or with respect to specific locations, remained significant and negative – reducing annual peak flow for watersheds within the Austin area. Furthermore, increasing the patch density of GI, at least outside the floodplain could have added benefits. With respect to Houston watershed however, the result, after controlling for patch density, brought into question the benefits of GI both within the whole watershed, as well as for specific locations, for reducing peak annual flow level. Furthermore patch density did not help reducing peak annual flow. Also, spatial pattern related to GYRATE_AM did not show statistically significant results on reducing annual peak flow for both the Austin and Houston areas.

5.3 The Effect of Green Infrastructure on Mean Annual Flow

I now turn my attention to the analysis of mean annual flow. The analysis proceeded as was undertaken when predicting peak annual flow. Table 14 presents a series of fixed and random effect panel models assessing the consequences of green infrastructure on mean annual flow, focusing on the overall percentage of green infrastructure within a watershed and with respect to the location of green infrastructure relative to the floodplain. The same three sets of models as Table 8 are presented. The 1A Models was a baseline set of fixed and random effect models predicting mean annual flow employing only the basic control variables. The 2A Models included a set of fix and random effect models predicting mean flow with the baseline controls and a measure of the overall percentage of green infrastructure within the watershed. The 3A Models included three sets of fixed and random effect models with baseline controls and each of the green infrastructure locational measures: the percentage of green infrastructure in the 100-year floodplain, in the 60-meter buffer around the floodplain, and outside the floodplain. Finally, the 4A Models, again pushed the analysis a bit, by presenting a set of models, including pairs of the locational measures that were not completely free of multicollinearity issues.

Focusing first on the baseline set of models (1A Models), I can see that, in general, the controls worked as expected. In the fixed effect model the two time variant indicators of precipitation and imperviousness indeed had significant effects on mean annual flow. Specifically, precipitation had a significant positive effect on annual mean flow, suggesting that every millimeter of precipitation increased annual mean flow by .41 percent. Also, every percentage point of impervious surface in a watershed also increased annual mean flow by 20.3%. The random effects model suggests that watershed and drainage area also had significant positive effects. Every square mile of watershed area increased mean flow by 0.51%, while every square mile of drainage area increased mean flow by 0.009%. Again as I saw with the models predicting peak annual flow, these patterns of findings with respect to the baseline control measures held across all other model sets, however in these models there was only one exception rather than two, and in this case it was the variable slope. The watershed's slope tended to become positively albeit mostly only marginally (.1) significant after green infrastructure measures were introduced into the model. One can see that the R^2 for the fixed model was .8278, suggesting nearly 83% of the within observation variation was on average accounted for by the baseline model. The random effect's model's overall R^2 suggest that just over 69% of the total variation in annual mean flow was accounted for by the basic model (as well as 81% of the within observation variation).

Table 14 also presents the findings for the Hausman Test associated with each set of panel models estimated. Again, this test assesses for significant differences between the coefficients associate with the time-variant variables estimated by the fixed and random effect models. Unlike in the case of annual peak flow, where all of these test were significant across all models, in this case only two models presented in Table 14 had significant Hausman tests. Indeed, if you examined the within variance R^2's they were often quite similar in magnitude for most models. Yet again, since random effect panel data analysis allowed for theoretically significant and substantive time-invariant variables (watershed area, drainage area, floodplain percentages, soil permeability, slope percentages, and stream density) to be considered as explanatory/control variables both sets of models are presented.

On the whole there was remarkably little support found for my general hypotheses that green infrastructure, at least in terms of the percent of green infrastructure in the watershed or in various locations with respect to the floodplain had any consequence for mean annual flows. The results for the 2A models, whether considering the fixed or random effects model, showed that the percent of green infrastructure in the watershed was insignificant, having controlled for the baseline factors. The results for the 3A models which assessed the consequences of the % of green infrastructure for specific locations within the watershed on annual mean flow were mixed at best.[11] The first model set in 3A included the percent of green infrastructure in the 100-year floodplain. In this set, the coefficient for green infrastructure was only significant and negative in the fixed effects model where as anticipated, and the coefficient's magnitude suggest that with every percentage point

[11] As was the case with the annual peak flow models, the mean peak flow Model showed major issues with multicollinearity, hence the 3A models include only one each of these location measures although in the 4A models I include at least two of the variables.

increase in GI in the 100-year floodplain, annual mean flows decreased by 2.2%. However, when other time invariant measures were controlled for, this effect became insignificant. As can be seen in the next set of models, the effects of GI outside the floodplain were insignificant in both models. The final set of models in this series included measures of green infrastructure in 60-meter buffer around the floodplain. Whether considering the fixed or random effects model, in each case the consequences of green infrastructure was, statistically significant and negative, having controlled for the baseline factors. These findings are consistent with the fourth hypothesis that green infrastructure in the 60-meter buffer around the floodplain should significantly reduce annual mean flow. The magnitudes of the effects of green infrastructure were consistent between the two models. The results suggest that the annual mean flow decreased by 2.5% (fixed) or 2.3% (random) for every percent increase in 60-meter buffer green infrastructure around the floodplain.

The results with the 4A models which included GI locational measures in the 100-year floodplain and outside of the 100-year floodplain, were consistent with the findings for the first and second 3A models. The fixed effects model suggests that green infrastructure in the 100-year floodplain had a negative and statistically significant effect, with a 4.4% reduction in mean flow for every percentage point increase in GI within the floodplain. But, green infrastructure outside of the 100-year floodplain was not statistically significant. However, both measures in the random effects model were not statistically significant.

5.3.1 Green Infrastructure Spatial Patterns and Mean Annual Flow

Table 15 presents a series of fixed and random effect panel models assessing the consequences of the overall percentage of green infrastructure within a watershed and with respect to the location of green infrastructure relative to the floodplain, and one of the spatial patterns of green infrastructure, patch density. The pattern presented here is similar to that seen in Table 9 due to the same multicollinearity issues requiring three sets of models.[12] The 2B Models included a set of fixed and random effect models predicting mean flow with the baseline controls and a pair of a measure of the overall percentage of green infrastructure within the watershed and a measure of spatial pattern, specifically patch density, within the watershed. The 3B models included three sets of fixed and random effect models with baseline controls and pairs of the green infrastructure locational measures and the appropriate spatial pattern measure. The locational measures included the percentage of green infrastructure in the 100-year floodplain, outside the floodplain, and in the 60-meter buffer around the floodplain. The spatial pattern measures included patch density of green infrastructure in the 100-year floodplain, outside the floodplain, and in the 60-meter buffer around the floodplain. Finally, the 4B models included pairs of the locational and the spatial pattern measures.

Focusing first on the control variables in the 2B Models, we can see that, in general, the controls tended to work as expected. In the fixed effect model the only precipitation, time variant indicator, indeed had significant effects on annual mean flow.

[12] As discussed above, running separate models does not solve the problem of multicollinearity; indeed, it introduces issues of model specification. The only real solution is to increase sample size, allowing estimating standard errors with more limited variance inflation issues. However, by running separate models I can seek to understand the potential differences in the effects of GI.

Specifically, precipitation had a significant positive effect on annual mean flow, suggesting that every millimeter of precipitation increased annual mean flow by .36 percent. The random effects model suggests that watershed, drainage area, floodplain, precipitation, and slope also had significant positive effects. Every square mile of watershed area increased mean flow by 1.7%, while every square mile of drainage area increased mean flow by 0.012%. Every percentage point increase in slope increased mean flow by 16.5%, while every one percentage point increase in floodplain increased mean flow by 2.5%. Every millimeter of precipitation increased peak annual flow by .34 percent. On the whole, these patterns of findings of control measures held across all other model sets, with the exception floodplain. Floodplain was only significant, albeit marginally significant, in the 2B Models. It should also be noted that yet again the Hausman test for the 2B models was not significant, indicating similar effects for the time invariant measures in both models. Now turning my attention to the consequences of GI, the 2B models added to the baseline models the percent of the watershed in green infrastructure and the patch density of GI in the watershed as the spatial pattern measure. Whether considering the fixed or random effects model, in each case the consequences of percent green infrastructure as well as its patch density was not statistically significant.

The 3B models assessed the consequences of the specific locations of green infrastructure within the watershed location variables and patch density within the same location on annual mean flow. The first two model set of models focused on percent and patch density of GI in and outside the 100-year floodplain. The results suggest that GI locational and patch density measures did not have the expected negative significant effects whether considering the fixed or random effect models. Patch density in the 100-year floodplain was only marginally significant, but positive in the random effects model. The final set in this series included measures of green infrastructure within a 60-meter buffer around the floodplain and associated patch density measures. In this set, the coefficient for green infrastructure location measure was significant in the fixed and random effects models. In the fixed effects model the percent of GI in the 60-meter buffer around the floodplain had a significant negative effect, as anticipated, and the coefficient's magnitude suggest that with every percentage point increase in GI in this buffer, annual mean flows decreased by 2.8%. The random effect model suggests that the percent of GI in the 60-meter buffer around the floodplain also had a significant negative effect and the coefficient's magnitude suggests that with every percentage point increase in GI in this buffer, annual mean flows decreased by 2.4%. Patch density of GI in the 60-meter buffer around the floodplain was not statistically significant in either the fixed or random effects models.

Finally, the 4B models including a set of GI locational measures and spatial pattern measures within the same model. In particular, this model included both green infrastructure in and outside the 100-year floodplain. Whether considering the fixed or random effects model, in each case the consequences of green infrastructure were statistically insignificant, having controlled for the baseline factors. None of the GI measures, whether locational or patch density, were statistically significant and negative in these two models. Indeed, only the patch density for green infrastructure in the 100-

year floodplain was marginally statistically significant (.1 level) in the random effect model. However, the effect was positive suggesting that as patch increases per 100 hectare in the 100-year floodplain, annual mean flows increased by 0.12%.

On the whole, the findings with respect to green infrastructure, when considering the percent of the watershed or various locations relative to the floodplain and patch density, showed relatively little consequence with respect to mean annual flow. The only consistent finding with respect to both the fixed and random effects models was that the percent of a 60-meter buffer around the floodplain significantly reduced annual mean flow.

Table 16 presents a series of fixed and random effect panel models assessing the consequences of green infrastructure on annual mean flow, focusing on the overall percentage of green infrastructure within a watershed, the percent of particular locations relative to the floodplain, and the GYRATE_AM spatial pattern measures. The pattern followed in this table is the same as the last. The 2C Models included a set of fixed and random effect models predicting mean flow with the baseline controls, a measure of the overall percentage of green infrastructure within the watershed and the spatial pattern GYRATE_AM measure. The 3C models included three sets of fixed and random effect models with baseline controls, the green infrastructure locational measures and the spatial pattern measure. The locational measures included the percentage of green infrastructure in the 100-year floodplain, outside the floodplain, and in the 60-meter buffer around the floodplain. The spatial pattern measures included in each model were the GYRATE_AM associated with location. Finally, the 4C models pushed the analysis a bit by presenting pairs of the locational measures and the spatial pattern measures.

The 2C models added to the baseline models the percentage of watershed area associated with green infrastructure and GYRATE_AM as the spatial pattern of green infrastructure in the watershed. Whether considering the fixed or random effects model, in each case the consequences of green infrastructure in the watershed was, statistically significant and negative, having controlled for the baseline factors and spatial pattern related to GYRATE_AM. These findings are consistent with the first hypothesis that overall green infrastructure should significantly reduce annual mean flow and represents the first time that we have seen significant negative effects for overall GI. In the fixed effect model, the results suggest that the annual mean flow decreased by 2.7% for every percent increase in overall green infrastructure, while in the random effect model there was a comparable 3.2% decrease for every percent increase in overall green infrastructure. Interestingly, the result of spatial pattern, GYRATE_AM was only significant in the random effect model, but the effect was positive. The effect suggests that the annual mean flow increased by 0.08% for every one meter distance increase in GI patch between each cell in a patch and the patch's centroid.

The 3C models assessed the consequences of the specific locations of green infrastructure within the watershed location variables and associated spatial attributes of the GI within these locations assessed by GYRATE_AM on annual mean flow. The first model set in 3C included only the percent of green infrastructure in the 100-year floodplain and GYRATE_AM in the 100-year floodplain. Whether considering the fixed or random effects model, in each case the consequences of green infrastructure in the watershed was, statistically insignificant. However, the GYRATE_AM in the 100-year floodplain infrastructure was negative and significant only in the fixed effect model. The coefficient suggests that annual mean flow decreased by 0.1% for every one meter distance increase in GI patch between each cell in a patch and the patch's centroid. As can be seen in the next set of models, the effect of green infrastructure outside the floodplain and its spatial attributes (GYRATE_AM) were statistically significant in the random effect model; however the effects were in opposite directions. The effects indicate that with every percent increase in green infrastructure outside the 100-year floodplain there was a 3.3% reduction in mean annual flow, but with every one meter distance increase in GI patch between each patch cell and the patch's centroid, mean annual flow increased by 0.15%. The final set in this series included measures of green infrastructure within a 60-meter buffer around the floodplain. In this set, the coefficients for green infrastructure were significant in the fixed and random effects models. In the fixed effects model the percent of GI in the 60-meter buffer around the floodplain had a significant negative effect, as anticipated, and the coefficient's magnitude suggests that with every percentage point increase in GI in this buffer, annual mean flows decreased by 2.6%. The spatial attributes of GI in this buffer did not have a statistically significant effect.

Finally, the 4C models again pushed the analysis by including set of GI locational and spatial pattern measures for areas within and outside the floodplain within the same model. On the whole the findings suggest that the percent of GI in and outside the floodplain had no effects, whether examining the fixed or random effects models. Furthermore, the GYRATE_AM measures were neither significant in the fixed effect models. The GYRATE_AM measures were only significant in the random effect models; however, the signs and magnitudes differed considerably. Random effect model results suggest that GYRATE_AM in the 100-year floodplain had a negative and statistically significant effect, with a 0.08% reduction in mean flow for every one meter distance increase in GI patch between each cell in a patch and the patch's centroid in the 100-year floodplain. On the other hand, random effects model suggest that GYRATE_AM outside the 100-year floodplain had a positive and statistically significant effect. The effect in the random effects model suggest that with for every one meter distance increase in GI patch between each cell in a patch and the patch's centroid in the 100-year floodplain increased 0.18%.

Fixed and Random Panel Model Predicting Mean Annual Flow

	1A Models		2A Models		3A Models						4A Models	
Baseline control variables	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random
Watershed Area		.0051***		.0049***		.0051***		.0048***		.0051***		.0051***
		(0.0015)		(0.0015)		(0.0015)		(0.0015)		(0.0016)		(0.0015)
Drainage Area		.00009***		.00010***		.00010***		.00009***		.00010***		.00010***
		(0.00002)		(0.00002)		(0.00002)		(0.00002)		(0.00002)		(0.00002)
Floodplain %		.0228		.0258		.0212		.0244		.0199		.0218
		(0.0148)		(0.0159)		(0.0160)		(0.0154)		(0.0150)		(0.0168)
Precipitation	.0041***	.0035***	.0041***	.0036***	.0040***	.0035***	.0041***	.0036***	.0040***	.0036***	.0038***	.0035***
	(0.0003)	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0002)	(0.0004)	(0.0003)	(0.0003)	(0.0002)	(0.0004	(0.0003)
Soil permeability		-.0010		.0026		.0009		.0022		.0031		.0014
		(0.0063)		(0.0072)		(0.0066)		(0.0073)		(0.0069)		(0.0076)
Impervious %	.1846***	.0166*	.1433**	.0121*	.1454**	.01432*	.1567**	.0127*	.1302*	.0120*	.1924***	.0137*
	(0.0638)	(0.0068)	(0.0689)	(0.0072)	(0.0702)	(0.0074)	(0.0678)	(0.0070)	(0.0708)	(0.0073)	(0.0595)	(0.0070)
Slope %		.0853		.1299*		.1158*		.1226*		.1372**		.1208*
		(0.0608)		(0.0709)		(0.0682)		(0.0706)		(0.0680)		(0.0715)
Stream Density		-.0908		-.2295		-.2009		-.2028		-.2119		-.2121
		(0.2999)		(0.3123)		(0.3186)		(0.3034)		(0.3105)		(0.3091)
Overall GI variable												
Percentages of GI												
in the watershed			-.0181	-.0183								
			(0.0128)	(0.0119)								
GI Location variables												
Percentages of GI												
in the 100-year floodplain					-.0222**	-.0133					-.0452**	-.0116
					(0.0105)	(0.0090)					(0.0209)	(0.0164)
outside the 100-year floodplain							-.0128	-.0156			.0405	-.0038
							(0.0136)	(0.0125)			(0.0274)	(0.0228)
60 meter buffer around floodplain									-.0248*	-.0228**		
									(0.01250)	(0.0103)		
Constant	-4.4752**	-0.0613	-2.2059	1.1267	-1.7665	1.0343	-2.9378	0.9206	-1.2811	1.497	-3.7867*	1.1259
	(2.0532)	(1.0528)	(2.5204)	(1.0941)	(2.6340)	(1.2319)	(2.4622)	(1.0456)	(2.6522)	(1.1272)	(2.2605)	(1.1247)
N	108	108	108	108	108	108	108	108	108	108	108	108
R-squared												
within	0.8278	0.8110	0.8333	0.8232	0.8425	0.8274	0.8302	0.8187	0.8390	0.8301	0.8503	0.8269
between		0.6285		0.6198		0.6070		0.6234		0.6164		0.6087
overall		0.6966		0.6959		0.6894		0.6964		0.6963		0.6903
Hausman Test: Prob (χ2)	0.0949		0.2855		0.0181		0.3214		0.2585		0.0084	

Notes: Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 14. Fixed and Random Effect Panel Model Predicting Mean Annual Flow 1

Fixed and Random Panel Model Predicting Mean Annual Flow

	2B Models		3B Models						4B Models	
Baseline control variables	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random
Watershed Area		.0167***		.0103***		.0114***		.0099***		.0121***
		(0.0022)		(0.0020)		(0.0026)		(0.0019)		(0.0026)
Drainage Area		.0001***		.00009***		.00008***		.00009***		.00008***
		(0.000008)		(0.00002)		(0.00003)		(0.00002)		(0.00002)
Floodplain %		.0250*		.0179		.0175		.0176		.0136
		(0.0145)		(0.0152)		(0.0138)		(0.0143)		(0.0158)
Precipitation	.0036***	.0034***	.0043***	.0038***	.0036***	.0035***	.0038***	.0036***	.0038***	.0036***
	(.0006)	(0.0004)	(.0005)	(0.0003)	(.0005)	(0.0004)	(.0004)	(0.0003)	(.0006)	(0.0004)
Soil permeability		.0006		.0022		.0015		.0036		.0016
		(0.0057)		(0.0058)		(0.0066)		(0.0062)		(0.0069)
Impervious %	.1190	.0195***	.1490**	.0204***	.1308*	.0189***	.1291*	.0161**	.1851***	.0220***
	(.0762)	(0.0070)	(.0727)	(0.0071)	(.0731)	(0.0067)	(.0722)	(0.0069)	(.0636)	(0.0069)
Slope %		.1528**		.1199*		.1339*		.1630**		.1153
		(0.0705)		(0.0665)		(0.0724)		(0.0695)		(0.0748)
Stream Density		-.2668		-.1410		-.1721		-.2526		-.0730
		(0.2886)		(0.30328138)		(0.2944)		(0.2994)		(0.3075)
Overall GI variable										
in the watershed	-.0221	-.0155								
	(.0177)	(0.0184)								
GI Location variables										
Percentages of GI										
in the 100-year floodplain			-.00004	.0064					-.0417	.0072
			(.01821)	(0.0118)					(.0281)	(0.0171)
outside the 100-year floodplain					-.0229	-.0162			.0598	-.0047
					(.0174)	(0.017)			(.0370)	(0.0282)
60 meter buffer around floodplain							-.0279**	-.0240**		
							(.0127)	(0.0106)		
GI Spatial Patterns variables										
PD (Patch Density)										
in the watershed	-.0006	-.0003								
	(.0008)	(0.0007)								
in the 100-year floodplain			.0012	.0011*					.0009	.0012*
			(.0009)	(0.0006)					(.0008)	(0.0007)
outside the 100-year floodplain					-.0007	-.0002			.00010	-.00042
					(.0007)	(0.0006)			(.00080)	(0.0007)
60 meter buffer around floodplain							-.0002	-.00002		
							(.0003)	(0.00026)		
Constant	-1.1130	0.5291	-4.1447	-1.5427	-1.1889	0.6793	-0.7027	1.2868	-5.6898*	-1.1581
	(3.2514)	(1.5536)	(3.0911)	(1.2959)	(3.0148)	(1.4396)	(2.7794)	(.9762)	(3.0087)	(1.7009)
N	98	98	106	106	102	102	106	106	102	102
R-squared										
within	0.8132	0.8057	0.8422	0.8284	0.8223	0.8132	0.8377	0.8276	0.8519	0.8289
between		0.6381		0.6361		0.6543		0.6438		0.6406
overall		0.7036		0.7081		0.7110		0.7126		0.7078
Hausman Test: Prob (χ2)	0.8398		0.0769		0.6233		0.2473		0.0076	

Notes: Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 15. Fixed and Random Effect Panel Model Predicting Mean Annual Flow 2

Fixed and Random Panel Model Predicting Mean Annual Flow

	2C Models		3C Models						4C Models	
Baseline control variables	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random	Fixed	Random
Watershed Area		.0113***		.0106***		.0081***		.0097***		.0098***
		(0.0026)		(0.0019)		(0.0027)		(0.0023)		(0.0030)
Drainage Area		.000127***		.00009***		.00009***		.00009***		.00009***
		(0.000007)		(0.00002)		(0.00002)		(0.00002)		(0.00002)
Floodplain %		.0255**		.0282		.0220*		.0169		.0369**
		(0.0125)		(0.0182)		(0.0123)		(0.0142)		(0.0164)
Precipitation	.0040***	.0037***	.0037***	.0034***	.0039***	.0036***	.0040***	.0036***	.0035***	.0035***
	(.0004)	(0.0003)	(.0004)	(0.0003)	(.0004)	(0.0003)	(.0004)	(0.0002)	(.0005)	(0.0003)
Soil permeability		.0006		.0027		.0029		.0038		.0062
		(0.0052)		(0.0062)		(0.0061)		(0.0063)		(0.0069)
Impervious %	.1283*	.0280***	.1156	.0163**	.1336*	.0243***	.1290*	.0161**	.1730**	.0213***
	(.0717)	(0.0071)	(.0784)	(0.0083)	(.0723)	(0.0061)	(.0732)	(0.0074)	(.0699)	(0.0067)
Slope %		.1204**		.1301*		.1178		.1582**		.1133
		(0.0599)		(0.0676)		(0.0725)		(0.0703)		(0.0718)
Stream Density		-.1113		-.2035		-.1243		-.2553		-.0799
		(0.2670)		(0.3161)		(0.2868)		(0.3096)		(0.2956)
Overall GI variable										
Percentages of GI										
in the watershed	-.0273*	-.0329***								
	(.0143)	(0.0106)								
GI Location variables										
Percentages of GI										
in the 100-year floodplain			-.0083	-.0054					-.0533	.0147
			(.0120)	(0.0103)					(.0323)	(0.0190)
outside the 100-year floodplain					-.0147	-.0330***			.0580	-.0415
					(.0155)	(0.0128)			(.0383)	(0.0255)
60 meter buffer around floodplain							-.0258*	-.0262**		
							(.0144)	(0.0120)		
GI Spatial Patterns variables										
GYRATE_AM (Correlation Length)										
in the watershed	.0006	.0008***								
	(.0004)	(0.0003)								
in the 100-year floodplain			-.0010**	-.0005					-.0002	-.0008*
			(.0005)	(0.0004)					(.0006)	(0.0004)
outside the 100-year floodplain					.0003	.0015***			-.0003	.0018***
					(.0006)	(0.0005)			(.0006)	(0.0006)
60 meter buffer around floodplain							.0002	.0004		
							(.0020)	(0.0015)		
Constant	-2.2345	0.1169	-1.1267	0.3148	-2.4840	0.4033	-1.3137	1.3229	-3.5784	-0.1088
	(2.7770)	(1.0267)	(2.9193)	(1.2181)	(2.7418)	(0.9988)	(2.7501)	(1.1635)	(2.7054)	(1.0829)
N	98	98	106	106	102	102	106	106	102	102
R-squared										
within	0.8185	0.8100	0.8500	0.8347	0.8198	0.8056	0.8350	0.8273	0.8505	0.8161
between		0.7200		0.6177		0.7267		0.6451		0.7293
overall		0.7552		0.6989		0.7546		0.7133		0.7600
Hausman Test: Prob ($\chi2$)	0.4485		0.0079		0.3064		0.7847		0.0086	

Notes: Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 16. Fixed and Random Effect Panel Model Predicting Mean Annual Flow 3

5.3.2 Regional Variations in Green Infrastructure Effects on Mean Annual Flow

Table 17 presents the five random effect panel models assessing the consequences of the percentage of green infrastructure in the watershed and for specific locations on mean annual flow that were originally presented in Table 14. Each model now included the Houston dummy variable and the associated Houston GI interaction variables. The last row on this table (in blue) presents the statistical test for the joint effects of adding the dummy and associated interaction term(s) to the model, again implying significant variations between the two metropolitan areas with respect to how GI was working in their watersheds.[13]

The first model (2A) in Table 17 assessed variations in the consequences for the overall percent of green infrastructure within the watershed between the Austin and Houston. This model explains that there were not statistically significant effects of overall percent of green infrastructure on reducing mean annual flow. The 3A models tested for variations in locational aspects of GI between the two regions. The joint effect test in the first model in the 3A set was significant (Prob(χ^2) < .01) suggesting that there were differences in the way the %GI performed in the 100-year floodplain. The first model in this series suggests that the effect of %GI in the floodplain for the Austin was significant and negative; with a 5.8% decrease in mean annual flow for every percent increase in GI within the floodplain.[14] However, the interaction coefficient for

[13] It is safer to perform the overall F-test or a test for the significance of the combined joint effects of including the dummy and associated interaction(s), rather than depending on individual t-test for the dummy and interaction terms because one or more multiple t-test may be significant due to random error and multicollinearity issues may obscure the individual t-test as well.

[14] $100(e^{-.05978639} - 1) = -5.8\%$

differential effect of %GI in the floodplain for the Houston area was positive and significant, resulting in the net effect being .0097224, which was not significant.[15] In the third model the effect of %GI in the 60-meter buffer for the Austin area was again significant and negative suggesting a 4.4% reduction in mean annual flow for every percent increase in GI within the 60-meter boundary. However, yet again the interaction coefficient for the differential effect in Houston was positive and significant (.0355), yielding a net effect (-.0096), which was not significant.[16] The results for the final model (4A) simply substantiated the findings of the first model in 3A, only now controlling for the %GI outside the floodplain, in that there was a significant reduction in mean annual flow of 6.4% for every % increase in GI within the floodplain in Austin. Also, the interaction coefficient for differential effect of %GI in the floodplain for the Houston area was positive and significant, resulting in the net effect being .0375662, which was significant. It implies that controlling for the %GI outside the floodplain, there was a significant increase in mean annual flow of 3.8% for every % increase in GI within the floodplain in Houston.

Table 18 presents random effect panel models assessing the consequences of green infrastructure on mean annual flow with a regional dummy variable (Houston), focusing on the percent green infrastructure within a watershed, with respect to the location of green infrastructure relative to the floodplain and the spatial pattern related the patch density of green infrastructure. Each set of models included the standard set of

---

[15] The net coefficient is -0.05978639 + 0.06950877 = .0097224. Its standard error is .0086237, which is not significant.

[16] The net coefficient is -.04511327 + .03551387 = -.0095994 with a standard error of .011253, which is not significant.

controls plus the percent of GI within the watershed or for specific locational within the watershed, associated patch density measure of each location, the regional dummy (Houston) and a set of interaction terms between the regional dummy and each GI locational and patch density measure. Unlike the results for the percent GI models presented in Table 17, the test for joint effects were consistent for each of these models except the third model in 3B series, implying that the consequences of GI, when both location and patch density were included in models, were significantly different between the two regions.

Table 12 assessing the consequences of green infrastructure on annual peak flow suggests that when patch density was controlled for, the significant negative consequences for green infrastructure persisted in watersheds located in the Austin region, however the effects essentially disappeared for watersheds in the Houston region. However, Table 18 assessing the consequences of green infrastructure on mean annual flow explains that green infrastructure in the 100-year floodplain and 60-meter buffer around the floodplain only had the significant negative effects on reducing annual mean flow. Furthermore, for watersheds in the Austin region, there were not any additional benefits in reducing mean annual flow by increasing the patch density for GI across all the models.

On the other hand, the interaction terms associated with the %GI measure, assessing for differential effects of the measure for Houston's watersheds were often statistically significant and positive. The resulting net effects for GI were consequently zeroed out or left with a slight positive effect. For example, in the first model in 3B series, the level effect for %GI in the 100-year floodplain was -.04342452, while the interaction term's coefficient was .07917335, yielding a net effect of 0.0357488. This was significant at the .01 level, suggesting an increase of 3.6% in mean annual flow per percent of GI in Houston. However, there were not similar findings as the consequences of the patch density for the GI on reducing annual peak flow.

Table 19 presents random effect panel models assessing the consequences of green infrastructure on mean annual flow with a regional dummy variable (Houston), focusing on the percent green infrastructure within a watershed, with respect to the location of green infrastructure relative to the floodplain and the one of spatial patterns of green infrastructure, GYRATE_AM. The pattern of equations in this table was similar to those in Table 18 in that each set of models included the standard set of controls plus the percent of GI within the watershed or for specific locational within the watershed, associated GYRATE_AM measure each location, the regional dummy (Houston) and a set of interaction terms between the regional dummy and each GI locational and GYRATE_AM measure. The statistical test for the joint effects were consistent for each of these models except 2A Model and the second model in 3C series. It implies that there were significant variations between the two metropolitan areas with respect to how GI was working, when both location and GYRATE_AM were included in models.

The overall findings for the analyses presented in Table 19, suggest when GYRATE_AM was controlled for, the significant negative consequences for green infrastructure tended to persist in watersheds located in the Austin region. However, the effects essentially disappeared in the 100-year floodplain in the Houston region. More specifically, whether considering the %GI in the watershed or in various locations relative to the floodplain (in, or in a 60-meter buffer) except the outside the 100-year floodplain the level coefficients associated with these measures were always statistically significant and negative, indicating there was a negative impact on mean annual flow.

The interaction terms associated with the %GI and GYRATE_AM measures assessed for differential effects of each of these measures for Houston's watersheds. The results of the interaction terms were statistically significant and positive in the overall watershed and outside the floodplain (see 2C and third model in the 3C series). The first model in the 3B series the level effect for %GI in the 100-year floodplain was -.0432214, while the interaction term's coefficient was .0596651, resulting in the net effect being .0164436, which was not significant[17]. For watersheds in the Houston region, there were not any significant additional benefits in reducing mean annual flow by increasing GYRATE_AM for GI.

[17] The net coefficient is -0.0432214 + 0.0596651 = .0164436. Its standard error is .0100932, which is not significant.

Random Panel Data Analysis with a Dummy Variable Interaction Predicting Mean Annual Flow

	2A Model	3A Models			4A Model
Baseline control variables					
Watershed Area	.0049***	.0048***	.0048***	.0053***	.0042***
	(0.0016)	(0.0014)	(0.0016)	(0.0018)	(0.0014)
Drainage Area	.00009***	.00009***	.00010***	.00010***	.00010***
	(0.00002)	(0.00002)	(0.00002)	(0.00002)	(0.00002)
Floodplain %	.0221	.0245	.0189	.0200	.0262
	(0.0208)	(0.0188)	(0.0201)	(0.0197)	(0.0160)
Precipitation	.0036***	.0031***	.0036***	.0034***	.0031***
	(0.0003)	(0.0002)	(0.0003)	(0.0002)	(0.0003)
Soil permeability	.0001	-.0042	.0011	-.0007	.0010
	(0.0077)	(0.0076)	(0.0078)	(0.0079)	(0.0083)
Impervious %	.0117	.0172**	.0121*	.0137*	.0165**
	(0.0072)	(0.0076)	(0.0070)	(0.0074)	(0.0070)
Slope %	.1611*	.1546**	.1336	.1718**	.0897
	(0.0880)	(0.0742)	(0.0928)	(0.0753)	(0.0792)
Stream Density	-.2298	-.0693	-.1778	-.1659	.0953
	(0.3160)	(0.3582)	(0.2971)	(0.3213)	(0.3149)
Regional dummy variable					
Houston					
	-.4666	-4.3664***	.5915	-1.9614	-2.9552**
	(1.4840)	(1.4104)	(1.5494)	(1.3230)	(1.4649)
Overall GI variable					
GI% in Watershed	-.0274				
	(0.0247)				
GI% in Watershed*Houston	.0132				
	(0.0254)				
GI Location variables					
GI% in Floodplain		-.0598***			-.0659***
		(0.0160)			(0.0217)
GI% in Floodplain*Houston		.0695***			.1035***
		(0.0198)			(0.0261)
GI% out of Floodplain			-.0108		.0304
			(0.0281)		(0.0331)
GI% out of Floodplain*Houston			-.0062		-.0696**
			(0.0277)		(0.0337)
GI% in Floodplain Buffer				-.04511***	
				(0.0160)	
GI% in Floodplain Buffer*Houston				.0355*	
				(0.0201)	
Constant	1.5829	3.9918***	0.5529	2.7676**	2.4189*
	(1.4082)	(1.4840)	(1.4864)	(1.2924)	(1.3965)
N	108	108	108	108	108
R-squared					
within	0.8264	0.8604	0.8177	0.8396	0.8620
between	0.6157	0.6101	0.6281	0.6098	0.6505
overall	0.6945	0.7035	0.6989	0.6957	0.7295
Test for Joint Effect: Prob (χ2)	0.8008	0.0020	0.8725	0.2055	0.0009

Note: This model includes a dummy variable(Houston), regional dummy variable. It equlas one for Houston, and zero for Austin.
Test for joint effect is for testing joint effect of a regional dummy variable and a dummy interaction.
Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 17. Random Effect Panel Model with a Dummy Variable predicting Mean Annual Flow Phase 1

Random Panel Data Analysis with a Dummy Variable Interaction Predicting Mean Annual Flow

	2B Model	3B Models			4B Model
Baseline control variables					
Watershed Area	.0168***	.0097***	.0125***	.0111***	.0107***
	(0.0028)	(0.0022)	(0.0030)	(0.0023)	(0.0035)
Drainage Area	.00013***	.00009***	.00008***	.00008***	.00008***
	(0.000008)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Floodplain %	-.0060	.0132	-.0065	.0092	.0071
	(0.0210)	(0.0174)	(0.0212)	(0.0173)	(0.0169)
Precipitation	.0037***	.0034***	.0037***	.0035***	.0034***
	(0.0005)	(0.0003)	(0.0004)	(0.0003)	(0.0004)
Soil permeability	-.0108	-.0046	-.0086	-.0040	-.0059
	(0.0075)	(0.0073)	(0.0070)	(0.0070)	(0.0076)
Impervious %	.0200***	.0221***	.02103***	.0187**	.0245***
	(0.0075)	(0.0073)	(0.0072)	(0.0073)	(0.0064)
Slope %	.2254**	.1777**	.1778*	.2332***	.1116
	(0.1020)	(0.0786)	(0.1018)	(0.0803)	(0.0802)
Stream Density	-.1093	.0379	-.0157	-.1576	.1674
	(0.3071)	(0.3591)	(0.3056)	(0.3120)	(0.3003)
Regional dummy variable					
Houston					
	-3.8180	-5.0487**	-4.4902*	-2.2272	-4.2533
	(2.3471)	(2.2158)	(2.3023)	(1.7994)	(3.3753)
Overall GI variable					
GI% in Watershed	-.0191				
	(0.0352)				
GI% in Watershed*Houston	.0634*				
	(0.0372)				
GI Location variables					
GI% in Floodplain		-.0434*			-.0608
		(0.0239)			(0.0424)
GI% in Floodplain*Houston		.0792***			.0885*
		(0.0260)			(0.0465)
GI% out of Floodplain			-.0225		.0436
			(0.0343)		(0.0435)
GI% out of Floodplain*Houston			.0683*		-.0293
			(0.0360)		(0.0507)
GI% in Floodplain Buffer				-.0470***	
				(0.0162)	
GI% in Floodplain Buffer*Houston				.0431*	
				(0.0224)	
GI Spatial Patterns variables					
PD (Patch Density)					
PD in Watershed	-.0006				
	(0.0007)				
PD in Watershed*Houston	.0031***				
	(0.0008)				
PD in Floodplain		.0005			.0008
		(0.0012)			(0.0019)
PD in Floodplain*Houston		.0010			-.0011
		(0.0012)			(0.0022)
PD out of Floodplain			-.0006		.0002
			(0.0006)		(0.0007)
PD out of Floodplain*Houston			.0031***		.0019
			(0.0008)		(0.0014)
PD in Floodplain Buffer				-.00002	
				(0.0003)	
PD in Floodplain Buffer*Houston				.0002	
				(0.0004)	
Constant	.2670	1.7348	.7421	2.2667**	.2077
	(2.1207)	(1.8855)	(2.0571)	(1.1454)	(3.1777)
N	98	106	102	106	102
R-squared					
within	0.8525	0.8671	0.8568	0.8369	0.8862
between	0.6270	0.6300	0.6346	0.6544	0.6629
overall	0.7152	0.7182	0.7136	0.7227	0.7422
Test for Joint Effect: Prob ($\chi 2$)	0.0004	0.0014	0.0003	0.1686	0.0005

Note: This model includes a dummy variable(Houston), regional dummy variable. It equlas one for Houston, and zero for Austin.
Test for joint effect is for testing joint effect of a regional dummy variable and a dummy interaction. Standard errors are in parentheses.

***$P<0.01$ **$P<0.05$ *$P<0.1$

Table 18. Random Effect Panel Model with a Dummy Variable predicting Mean Annual Flow Phase 2

Random Panel Data Analysis with a Dummy Variable Interaction Predicting Mean Annual Flow

	2C Model	3C Models			4C Model
Baseline control variables					
Watershed Area	.0119***	.0111***	.0085***	.0108***	.0089**
	(0.0030)	(0.0024)	(0.0033)	(0.0027)	(0.0037)
Drainage Area	.0001***	.00009***	.00009***	.00008***	.00010***
	(0.000007)	(0.00003)	(0.00003)	(0.00003)	(0.00003)
Floodplain %	.0159	.0201	.0151	.0049	.0325**
	(0.0168)	(0.0170)	(0.0174)	(0.0170)	(0.0157)
Precipitation	.0036***	.0030***	.0035***	.0035***	.0029***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Soil permeability	-.0031	-.0053	.0003	-.0012	.0016
	(0.0059)	(0.0070)	(0.0072)	(0.0069)	(0.0081)
Impervious %	.0250***	.0173**	.0231***	.0204***	.0218***
	(0.0067)	(0.0084)	(0.0062)	(0.0078)	(0.0070)
Slope %	.1574*	.2138***	.1394	.2278***	.1021
	(0.0903)	(0.0817)	(0.0977)	(0.0869)	(0.0856)
Stream Density	-.0788	-.0517	-.0899	-.2213	.1790
	(0.2646)	(0.3462)	(0.2922)	(0.3297)	(0.2941)
Regional dummy variable					
Houston					
	.3116	-3.1598**	.1501	-1.2080	-3.2974**
	(1.2703)	(1.4963)	(1.5544)	(1.5571)	(1.6055)
Overall GI variable					
GI% in Watershed	-.0350*				
	(.0207)				
GI% in Watershed*Houston	.0136				
	(.0226)				
GI Location variables					
GI% in Floodplain		-.0432**			-.0659**
		(0.0202)			(0.0326)
GI% in Floodplain*Houston		.0597***			.1162***
		(0.0230)			(0.0351)
GI% out of Floodplain			-.0344		.0168
			(0.0276)		(0.0386)
GI% out of Floodplain*Houston			.0077		-.0684*
			(0.0290)		(0.0394)
GI% in Floodplain Buffer				-.0479**	
				(0.0220)	
GI% in Floodplain Buffer*Houston				.0113	
				(0.0319)	
GI Spatial Patterns variables					
GYRATE_AM (Correlation Length)					
GYRATE_AM in Watershed	.0009**				
	(.0004)				
GYRATE_AM in Watershed*Houston	-.0005				
	(.0004)				
GYRATE_AM in Floodplain		-.0007			-.0003
		(0.0007)			(0.0007)
GYRATE_AM in Floodplain*Houston		.0002			-.0006
		(0.0007)			(0.0008)
GYRATE_AM out of Floodplain			.0016***		.0012
			(0.0006)		(0.0009)
GYRATE_AM out of Floodplain*Houston			-.0004		-.000007
			(0.0009		(0.001125)
GYRATE_AM in Floodplain Buffer				.00008	
				(0.00184)	
GYRATE_AM in Floodplain Buffer*Houston				.0062	
				(0.0041)	
Constant	-.0652	2.5522*	.3424	2.4302	1.9977
	(1.4605)	(1.4866)	(1.5773)	(1.5635)	(1.6227)
N	98	106	102	106	102
R-squared					
within	0.8006	0.8687	0.8058	0.8486	0.8688
between	0.7553	0.6238	0.7289	0.6406	0.7257
overall	0.7730	0.7148	0.7562	0.7182	0.7767
Test for Joint Effect: Prob (χ2)	0.2632	0.0015	0.9332	0.0178	0.0005

Note: This model includes a dummy variable(Houston), regional dummy variable. It equlas one for Houston, and zero for Austin.
Test for joint effect is for testing joint effect of a regional dummy variable and a dummy interaction. Standard errors are in parentheses.

****P <0.01 **P <0.05 *P <0.1*

Table 19. Random Effect Panel Model with a Dummy Variable predicting Mean Annual Flow Phase 3

5.3.3 Summary of Effectiveness of Green Infrastructure on Mean Annual Flow

The series of models assessing the consequences of percent of green infrastructure within a watershed and for particular locations relative to the floodplain on mean annual flows suggests that I do not see the same negative significant consequences as were seen for peak annual flows. Indeed, the only consistently negative effect related to green infrastructure within the 60-meter barrier around the floodplain. There were also significant negative effects displayed for green infrastructure within the floodplain in the fixed effect models. Indeed, there was remarkably little support found for my general hypotheses that green infrastructure, at least in terms of the percent of green infrastructure in the watershed or in various locations with respect to the floodplain had any consequence for mean annual flows.

The findings with respect to green infrastructure, when considering the percent of the watershed or various locations relative to the floodplain and patch density, showed relatively little consequence with respect to mean annual flow. The only consistent finding with respect to both the fixed and random effects models was that the percent of the 60-meter floodplain buffer in GI significantly reduced annual mean flow. The results with respect to locational aspects of green infrastructure and spatial pattern measures on mean annual flow suggest that in general when only considering locational aspects, the percent of GI appeared to be weaker consequence for annual mean flow. There was some evidence of a negative significant effect for %GI in the floodplain, but only in the fixed effect model, and yet consistent evidence of a negative effect of %GI in the floodplain buffer in both fixed and random effect models. This consistent negative effect became somewhat stronger, when the patch density was controlled for (see Table 15) and when GYRATE_AM was controlled for as well (see Table 16). Interestingly, the %GI in the entire watershed also became significant and negative when GYRATE_AM was controlled for as well. On the whole, however, when compared to annual peak flow, the consequences of GI for annual mean flow appeared to be much weaker and certainly more inconsistent.

5.4 Overall Summary of Effectiveness of Green Infrastructure on Peak Annual Flow and Mean Annual Flow

Overall summary of effectiveness of green infrastructure on annual peak flow and mean annual flow suggests that green infrastructure had consequences for streamflow reduction, particularly with respect to annual peak flow, in urban watersheds. The series of fixed and random effect panel models assessing the consequences of green infrastructure on annual peak flow suggests that the percent of green infrastructure in the watershed was a significant negative determinant of annual peak flow. Furthermore, there was evidence that its locational features with respect to the floodplain may well have consequences and spatial attribute related to patch density were significant as well. However, the GRYATE_AM measure did not seem to have much consequence in the majority of the models predicting peak annual flow. In comparison to the significant effectiveness of green infrastructure for reducing annual peak flow, the consequences of GI for annual mean flow appeared to be much weaker and inconsistent. The series of models assessing the consequences of percent of green infrastructure within a watershed and for particular locations relative to the floodplain on mean annual flows suggests that there were not the same negative significant consequences as were seen for peak annual flows. Indeed, the only consistently negative effect related to green infrastructure appeared within the 60-meter barrier around the floodplain. There was remarkably little support found for my general hypotheses that green infrastructure, at least in terms of the percent of green infrastructure in the watershed or in various locations with respect to the floodplain had any consequence for mean annual flows.

Random effect panel models assessing regional variations in green infrastructure effects on annual peak flow suggest that while there did not appear to be different in the consequences of %GI in the overall watersheds between Austin and Houston areas, in that green infrastructure significantly reduced annual peak flows in both areas, there were some variations with respect to GI's locational features. In particular, the %GI in the floodplain and in the buffer around the floodplain appeared to be more effective in watersheds located in the Austin region in comparison to the Houston region. Furthermore, increasing the patch density of GI, at least outside the floodplain could have added benefits. With respect to Houston watershed however, the result, after controlling for patch density, brought into question the benefits of GI both within the whole watershed, as well as for specific locations, for reducing peak annual flow level. Spatial pattern related to GYRATE_AM did not show statistically significant results on reducing annual peak flow for both the Austin and Houston areas. In comparison to the assessment of regional variations in green infrastructure effects on annual peak flow, whether considering the percent in the entire watershed or with respect to specific locations, the consequences of green infrastructure remained mostly significant and negative – reducing mean annual flow for watersheds within the Austin area. However, spatial pattern measures –patch density and GYRATE_AM did not show statistically significant results on reducing mean annual flow for both the Austin and Houston areas.

6. DISCUSSION

6.1 Discussion of the New Approach to Measuring Green Infrastructure

This study employed here-to-for unutilized data to assess green infrastructure as a different approach in urban areas at an exceptionally high level of resolution. Conventional approaches utilized the National Land Cover Database (NLCD) and Coastal Change Analysis Program (C-CAP) Land Cover data, which are recorded at a 30 by 30 meter resolution. In contrast, this study used the 1-meter high resolution imagery data produced by the National Agricultural Imagery Program (NAIP), which provides a different approach to measure green infrastructure especially in urban areas. Because key variables in this study were based on utilizing a new measurement for green infrastructure, it was better to explain further the effectiveness of highly detailed information about green infrastructure on peak annual flow and mean annual flow; this is especially true with regards to green infrastructure in urban environments. Since its initial project in 2003, NAIP has acquired imagery during the growing seasons in the continental U.S.; using this high resolution imagery was helpful when analyzing green infrastructure on different temporal and spatial scales. Also, utilizing this high resolution NAIP imagery will help researchers analyze the effectiveness of green infrastructure on streamflow and potentially flooding mitigation in urban areas. However, to assess the usefulness of the new measurement of green infrastructure with a finer resolution data, a comparative analysis between the new measurement and the conventional approaches should be conducted.

6.2 Discussion of the Effectiveness of Green Infrastructure on Peak Annual Flow and Mean Annual Flow

When controlling for other control variables, the series of models assessing the consequences of green infrastructure indicated that green infrastructure had an important effect on streamflow. Based on the results of this study, I conclude that green infrastructure has consequences for streamflow reduction, particularly with respect to annual peak flow, in urban watersheds. The results of the various fixed and random effects models to test the consequences of the percent of green infrastructure for reducing annual peak flow showed that green infrastructure in the watershed, in the 100-year floodplain, and outside the 100-year floodplain were statistically significant for annual peak flow reduction. In the fixed effect model, the peak annual flow decreased by 7.2% for every percent increase in overall green infrastructure, while in the random effects model there was a 2.9% decrease. The effect in the fixed effects model suggest that with every percent increase in GI within the 100-year floodplain, peak annual flow decreased by 7.7%. Controlling for time invariant measures reduced the effect to 2.1%. The effects of GI outside the floodplain appeared to be quite comparable, in that both measures were significant, had negative effects, and their magnitudes in the fixed (-7.1%) and the random (-2.5%) were similar to those for GI in the floodplain. The only differences between these two sets, were with respect to the R^2 values where they were slightly higher in the models including the percent GI within the floodplain (fixed = .6985; random = .5787), as opposed to the percent GI outside the floodplain (fixed = .6447, random = .5745). While these differences were slight, particularly with respect to the random effects models, they perhaps suggest slightly greater consequences for preserving GI within floodplains when it comes to peak annual flows. In the fixed effects model the percent of GI in the 60-meter buffer around the floodplain had a significant negative effect, and the coefficient's magnitude suggests that with every percentage point increase in GI in this buffer, peak annual flows decreased by 8.9%. However, when other time invariant measures were controlled for, this effect became insignificant. The result of the 4A models including both green infrastructure in the 100-year floodplain and outside of the 100-year floodplain variables perhaps gave some weight to the relative importance for retaining or expanding GI within the floodplain, but again the findings were mixed. Based on these consequences of green infrastructure in the watershed as well as in different locations relative to the 100-year floodplain, it could be suggested that that preserving and implementing green infrastructure in the 100-year floodplains is most useful to reduce peak annual flow.

This study also tested a series of fixed and random effect panel models assessing the consequences of green infrastructure on streamflow, focusing on the overall percentage of green infrastructure within a watershed and with respect to the location of green infrastructure relative to the floodplain, and also considering the spatial patterns of green infrastructure, patch density and GYRATE_AM. The results also point to the effectiveness of green infrastructure on annual peak flow reduction. Specifically, the percent of a watersheds green infrastructure as well as the density of the patches of green infrastructure had consequences on peak annual flow. Increasing the percent of GI within the watershed reduced peak flow, and increasing the density also had mitigative consequences, reducing the peak flow as well. These findings suggest that while there were substantial and significant reductions in peak flow with increase of GI within a watershed, but these reductions could be enhanced by increasing their density.

When just examining the percent of GI in the watershed, it appeared that increasing GI both in and outside the floodplain had negative consequences (see Table 9). However, just based on the consistency of results between the fixed and random effect models, it appeared that both the percent of GI and its patch density of GI outside the floodplain consistently had negative and significant consequences on annual peak flow. This implies that percent of a green infrastructure as well as the density of the patches of green infrastructure had consequences on peak annual flow outside the floodplain. From a planning perspective, this means that not only the percent of green infrastructure, but also the spatial patterns of green infrastructure are important factors that should be considered for strategical green infrastructure implementation for outside of the 100-year floodplain.

In comparison to the consequence of patch density, GYRATE_AM as one of green infrastructure's spatial pattern measures, did not seem to have much consequence in the majority of the models predicting peak annual flow. Specifically they showed no effect when focusing on total GI as a percent of the entire watershed, within the floodplain, as well as within a 60-meter buffer. The spatial features showed some significance, but positive effects outside the floodplain, and simultaneously opposite effects when both spatial characteristics were considered for in and outside the floodplain. GYRATE_AM measure was only significant in the locational model examining GI outside the floodplain (displaying positive effects) and positive and negative effects, canceling each other out, when split between in and outside the floodplain. This is evidenced by the same pattern seen in the patch density model. GYRATE_AM measures how far across a landscape a patch extends, either the patch shape is elongated with a higher GYRATE value, or it is comprised of compact patch shapes of the same size. Therefore, GYRATE_AM provides insights into the average distance that a streamflow can move across a landscape. However, since this measure does not explain the direction of the patch shape, (for example, whether a patch is located perpendicular or parallel to a waterway), this spatial pattern may not have a strong statistical relationship with streamflow. Therefore, GYRATE_AM did not show significant effect on reducing peak annual flow, and this can be interpreted that GYRATE_AM is not a good measurement for predicting peak annual flow.

Based on the results of this study, green infrastructure is effective for streamflow reduction, and especially effective with regards to annual peak flow. In comparison to the results regarding annual peak flow, the series of fixed and random effects models assessing the consequences of percent of green infrastructure on mean annual flows did not present the same negative significant consequences as were seen for peak annual flows. The results with respect to locational aspects of green infrastructure on mean annual flow rates were quite variable and dependent on the set of other variables included in the analysis. In general when only considering locational aspects, the percent of GI in the entire watershed or outside the floodplain, there appeared to be no consequence for annual mean flow. There is some evidence of a negative significant effect for %GI in the floodplain, but only in the fixed effect model, and yet consistent evidence of a negative effect of %GI in the floodplain buffer in both fixed and random effect models. A one percent increase in green infrastructure in the 60-meter buffer around the floodplain translated to a 2.3 percentage reductions in annual mean flow. This consistent negative effect became somewhat stronger, when the patch density was controlled for (see Table 15) and when GYRATE_AM was controlled for as well (see Table 16). Interestingly, the %GI in the entire watershed also became significant and negative when GYRATE_AM was controlled for as well. On the whole, however, when compared to annual peak flow, the consequences of GI for annual mean flow average rates appeared to be much weaker and certainly more inconsistent. Since green infrastructure appeared to be more effective for reducing peak annual flow, strategic green infrastructure implementation should be applied to watersheds with high peak annual flow issues.

My sample of watersheds were drawn from two major metropolitan areas in Texas that are subject to flooding, the Austin and Houston metropolitan areas. Since these two areas vary considerably with the terrain characteristics, the consequences for the effectiveness of GI may vary across these two areas. The random effects models assessing these regional variations in the consequences for green infrastructure on annual peak flow suggest that while there appeared to be no differences in the consequences of

%GI in the overall watersheds between Austin and Houston areas, in that green infrastructure significantly reduced annual flows in both areas, there were some variations with respect to GI's locational features. In particular, the %GI in the floodplain and in the buffer around the floodplain appeared to be more effective in watersheds located in the Austin region in comparison to the Houston region. However these findings must be tempered by the results when controlling for spatial features as well. The regional variation analysis showed that the consequences of green infrastructure whether considering the percent in the entire watershed or with respect to specific locations, remained significant and negative – reducing peak annual flow for watersheds within the Austin area. Furthermore, increasing the patch density of GI, at least outside the floodplain can have added benefits. With respect to Houston watershed however, the result, after controlling for patch density, bring into question the benefits of GI both within the whole watershed, as well as for specific locations, for reducing peak annual flow level. Furthermore patch density did not help attenuating peak annual flow. These latter finding are obviously counter to the general expectations of this dissertation and are quite different from the previous analysis to this point. An obvious potential explanation is that I have pushed the analysis too far and issues of multicollinearity had been compounded leading to larger standard errors and less reliable estimates. This may well be the case, but the consequences for my expectations cannot be simply ignored nor dismissed.

Furthermore, GYRATE_AM did not show statistically significant results on reducing annual peak flow for both the Austin and Houston areas. In comparison to the assessment of regional variations in green infrastructure effects on annual peak flow, whether considering the percent in the entire watershed or with respect to specific locations, the consequences of green infrastructure remained mostly significant and negative – reducing mean annual flow for watersheds within the Austin area. Specifically, green infrastructure in the 100-year floodplain of Austin metropolitan area was consistently significant for mean annual flow reduction. However, spatial pattern measures –patch density and GYRATE_AM did not show statistically significant results on reducing mean annual flow for both the Austin and Houston areas. Especially, the consequences of spatial measure of GYRATE_AM did not indicate any significant effect on either peak annual flow or mean annual flow. Even the testing performed with a regional dummy did not find that GYRATE_AM was effective across the different models. This again emphasizes that this measurement is not significant enough to capture the relationship between the green infrastructure spatial pattern and streamflow.

In comparison to the assessment of regional variations in green infrastructure effects on annual peak flow, whether considering the percent in the entire watershed or with respect to specific locations, the consequences of green infrastructure remained mostly significant and negative – reducing mean annual flow for watersheds within the Austin area. The results of the random effects models assessing variations in the consequences for green infrastructure on streamflow implies that the geographical characteristics of the Austin metropolitan area (such as its steep slope, as compared to the flat slope of the Houston metropolitan area) is one of reasons green infrastructure works better to reduce both peak annual flow and mean annual flow.

6.3 Policy Implications

The results of this study indicate that green infrastructure implementation in urban areas have important influences on streamflow, especially annual peak flow. This conclusion underscores the importance of protecting and implementing green infrastructure in order to maintain existing ecosystem functions, as well as to attenuate streamflow in urban areas.

Even though green infrastructure has a significant effect on streamflow, it is both difficult and expensive to preserve the existing green infrastructure and implement green infrastructure in urban areas such as the Austin and Houston metropolitan areas. Therefore, several policy approaches should be followed to acquire green infrastructure in critical places vulnerable to runoff. These strategies include conservation easements, overlay zones, the transfer of development rights, and density bonuses (Brody & Highfield, 2013). For example, green infrastructure is one of the goals proposed in the "Imagine Austin comprehensive plan".[18] Austin, as a local jurisdiction, has proposed the use of green infrastructure to protect environmentally sensitive areas and integrate nature into the city. The City of Austin has been purchasing property to create the Water Quality Protection Lands, and they applied some of these strategies in their efforts to acquire new green infrastructure. In addition to this approach to implementing green infrastructure, several other policy implications will be helpful for preserving existing

[18] ftp://ftp.ci.austin.tx.us/npzd/Austingo/web_IACP_full_reduced.pdf

green infrastructure and to balancing urban development and green infrastructure implementation.

The results of this study also indicate that the spatial pattern of the green infrastructure significantly affects the amount of streamflow, even when controlling for multiple variables. The statistical models of the spatial patterns across different locations indicate that high patch density has an effect on streamflow. This study explains the varying levels of effectiveness of different spatial patterns of green infrastructure by analyzing the location of the green infrastructure. The results emphasize that not only floodplain management itself, but also management outside of the floodplain should be planned. For example, as a series of panel models of overall green infrastructure and different locational aspects of green infrastructure somehow showed that preserving green infrastructure in the 100-year floodplain would be most useful for reducing peak annual flow. To preserve existing green infrastructure, different types of policy options to encourage the development out of the 100-year floodplain can be suggested. These options are density bonuses, transfer of development rights, clustering and conservation easements (Brody et al.,2013). In addition to preserve existing green infrastructure, another policy option to implement green infrastructure can be suggested. To increase green infrastructure in the 100-year floodplains, zoning ordinance for regulating land use including tree canopy cover can be suggested.

7. CONCLUSION

7.1 Research Summary

First, this study explained how to employ 1-meter high resolution NAIP and NDVI to develop high-resolution measures of green infrastructure particularly germane for assessments within dense urban environments. Utilizing this new measurement allowed for a different approach of green infrastructure measurement, especially urban areas with rapid development and imperviousness. Based on the results of this study, green infrastructure did have consequences for streamflow reduction, particularly with respect to annual peak flow, in urban watersheds. Second, this study empirically evaluated the impacts of green infrastructure within a watershed and with respect to the location of green infrastructure relative to the floodplain, on peak annual flow and mean annual flow. The new measurement of green infrastructure as a different approach allowed me consider locational aspects with in the watershed. The results of this study will be used to establish guidelines for green infrastructure and effective runoff mitigation. The findings of this study will also help provide additional decision support tools for urban planners, policy makers, and community residents, as they evaluate the existing green infrastructure in communities and make decisions regarding the implementation of new green infrastructure to reduce streamflow and enhance community resilience. This study also explained the varying levels of effectiveness of different spatial patterns of green infrastructure by analyzing the green infrastructure variables related to location and spatial patterns. This research also provided guidelines regarding the appropriate amount and spatial patterns of green infrastructure to reduce streamflow in urban areas. Results from analysis with a regional dummy variable illustrated that green infrastructure in the Austin metropolitan area floodplain tended to be more effective for consistently reducing peak annual flow, as compared to the flat terrain of the Houston metropolitan area. Therefore, depending upon the area's geographical characteristics, diverse guidelines for green infrastructure implementation should be applied. The effectiveness of green infrastructure in critical places will help researchers create guidelines for balanced urban development incorporating the implementation of green infrastructure.

This study has several contributions to the research on green infrastructure and streamflow. First, this study allowed more locational aspects and spatial patterns of green infrastructure. In terms of locational aspects, it may not simply consider the amount of green infrastructure in the watershed, but also the locational measurement such as in the 100-year floodplain, outside the 100-year floodplain, and the 60-meter buffer around the floodplain. Also, since this study considered spatial patterns in these different locations.

In sum, this dissertation showed the utility of the new data available for developing high-resolution measurements of green infrastructure. The consequences of green infrastructure in affecting streamflow and potential flooding were clearly suggested. Moreover, this study begins to provide data that may well be used to establish guidelines for green infrastructure and effective runoff mitigation. Finally, this dissertation provided support for utilizing these data to guide research in green infrastructure's spatial characteristics and hazard mitigation. Overall, the outcomes of this study will be helpful in the strategic planning and implementation of green infrastructure with streamflow issues, thus building community resilience.

7.2 Limitations and Future Research

One limitation of this study is the potential for internal validity threats such as history threat and selection bias (Babbie, 2011). This study analyzed the relationship between green infrastructure and streamflow by using a longitudinal analysis to reduce the history threat. However, as this study considered only two time periods, water years 2004 and 2010, it did not include data generated between 2004 and 2010. Therefore, there is a potential internal history threat. Also, although there were strict requirements implemented for selecting stream gage stations to delineate new watershed boundaries, stream gage station selection could affect the internal validity of this study. A second internal limitation is that the NAIP acquisition dates were different in 2004 and 2010 (See Appendix C). However, a comparison of the total areas of green infrastructure in 2004 and 2010 showed that there was a significant decrease in green infrastructure in the study area; this was one of the assumptions of this study that urban developments have linked to decreased green infrastructure (See Appendix D).

Another limitation of this study is a potential external validity threat due to the fact that so far, the effectiveness of green infrastructure for streamflow reduction had only been observed in the Austin and Houston metropolitan urban areas. In order to generalize the effect of green infrastructure, further research in different urban settings should be pursued.

One of the major constraints of this dissertation was sample size due to the limited amount of available streamflow data and its consequences for my analysis strategy. Issues of multicollinearity also have been compounded leading to larger standard errors and less reliable estimates. In addition to the limited streamflow data and lack of stream gage stations (especially in urban area), watershed delineation was not possible in some parts of the study area. However, considering this area's urban sprawl patterns, the watershed delineations I was able to create did adequately cover this urban development and explained the effectiveness of green infrastructure on reducing streamflow, and especially annual peak flow.

Although this study provides important information about the relationship between green infrastructure and streamflow, future studies should endeavor to understand this relationship more fully. This study analyzed only two landscape metrics for green infrastructure spatial patters. Further research focused on testing additional green infrastructure spatial patterns will provide a better understanding of the effectiveness of green infrastructure on streamflow reduction in urban areas.

This study employed here-to-for unutilized data to assess green infrastructure in urban areas at an exceptionally high level of resolution in urban areas as a different approach to measure green infrastructure. To assess the effectiveness of this different measurement, future study is recommended to compare the consequences of the new measurement with a finer resolution to the conventional approaches.

Also, how local jurisdictions implement green infrastructure as a mean of effective non-structural mitigation for streamflow reduction is suggested for future research that will be analyzed by considering green infrastructure related Community Rating System (CRS) activities as a control variable. In this same way, CRS activities can be analyzed for their relationship to green infrastructure's effect implementation.

This research examined only a two-year time frame in 2004 and 2010. Future investigations should include a broader time frame by incorporating more recent year's green infrastructure and streamflow data. Also, new measurement of green infrastructure using 1-meter high resolution imagery should be analyzed with regards to different types of green infrastructure and land use information. Such future work will provide essential information about the usefulness of green infrastructure in urban areas, and how green infrastructure distribution across different land use types affects effectiveness of green infrastructure on reducing streamflow and flooding in urban areas.

REFERENCES

Allen, W. (2012). Environmental reviews and case studies: Advancing green infrastructure at all scales: From landscape to site. *Envionmental Practice, 14*(1), 17-25.

Babbie, E. (2011). *The basics of social research*. 5th edition. Belmont, CA: Wadsworth.

Beatley, T. (2009). *Planning for coastal resilience: Best practices for calamitous times.* Washington, DC: Island Press.

Beatley, T. (2011). Resiliency to disasters. In A.L. Dannen, H. Frumkin, & R. Jackson (Eds.), *Making Healthy Places: Designing and Building for Healthy, Well-being, and Sustainability* (pp. 244-258). Washington, DC: Island Press.

Bedient, P., Holder, A., & Vieux, B. (2002). A radar-based flood alert system (FAS) designed for Houston, Texas. Proceedings from *the Ninth International Conference on Urban Drainage.* Portland, OR.

Benedict, M., & McMahon, E. (2002). Green infrastructure: Smart conservation for the 21st century. *Renewable Resources Journal*, 20(3), 12-17.

Brezonik, P., & Stadelman, T. (2002). Analysis and predictive models of stormwater runoff volumes, loads, and pollutant concentrations from watersheds in the Twin Cities metropolitan area, Minnesota, USA, *Water Research*, 36(7), 1743-1757.

Brody, S., Bernhardt, S., Zahran, S., & Kang, J. (2009). Evaluating local flood mitigation strategies in Texas and Florida. *Built Environment*, 35(4), 492-515.

Brody, S., Blessing, R., Sebastian, A., & Bedient, P. (2014). Examining the impact of land use/land cover characteristics on flood losses. *Journal of Environmental Planning and Management*, 57(8), 1252-1265.

Brody, S., Gunn, J., Peacock, W., & Highfield, W. (2011). Examining the influence of development patterns on flood damages along the Gulf of Mexico. *Journal of Planning Education and Research*, 31(4), 438-448.

Brody, S., & Highfield., W. (2013). Open space protection and flood mitigation: A national study. *Land Use Policy*, 32, 89-95.

Brody, S., Highfield, W., & Kang. J. (2011). *Rising waters: The causes and consequences of flooding in the United States*. Cambridge, NY: Cambridge University Press.

Brody, S., Kim, H., & Gunn, J. (2013). Examing the impacts of development patterns on flooding on the Gulf of Mexico Coast. *Urban Studies*, 50(4), 789-806.

Brody, S., Peacock, W., & Gunn, J. (2012). Ecological indicators of flood risk along the Gulf of Mexico. *Ecological Indicators*, 18, 493-500.

Brody, S., Zahran, S., Highfield, W., Grover, H., & Vedlitz, A. (2008). Identifying the impact of the built environment on flood damage in Texas. *Disasters*, 32(1), 1-18.

Burby, R. (1998a). Natural hazards and land use: An introduction. In R. J. Burby(Ed.), *Cooperating with Nature: Confronting Natural Hazards with Land-Use Planning for Sustainable Communities* (pp. 1-26). Washington, DC: Joseph Henry Press.

Burby, R. (1998b). Policies for sustainable land use. In R. J. Burby(Ed.), *Cooperating with Nature: Confronting Natural Hazards with Land-Use Planning for Sustainable Communities* (pp. 263-292). Washington, DC: Joseph Henry Press.

Burby, R., & Dalton, L. (1994). Plans can matter! The role of land use plans and state planning mandates in limiting the development of hazardous areas. *Public Administration Review*, 54(3), 229-238.

Castelle, A., Johnson, A., & Conolly, C. (1993). Wetland and stream buffer size requirements – A review, *Journal of Environmental Quality*, 23(5), 878-882.

Fernandes, M., Aguiar, F., &Ferreira, M. (2011). Assessing riparian vegetation structure and the influence of land use using landscape metrics and geostatistical tools. *Landscape and Urban Planning*, 99(2), 166-177.

Giulio, M., Holderegger, R., & Tobias, S. (2009). Effects of habitat and landscape fragmentation on humans and biodiversity in densely populated landscapes. *Journal of Environmental Management*, 90, 2959-2968.

Godschalk, D., Beatley, T., Berke, P., Brower, D., & Kaiser, E. (1999). *Natural hazard mitigation: Recasting disaster policy and planning*. Washington, DC: Island Press.

Gregory, J., Dukes, M., Jones, P., & Miller, G. (2006). Effect of urban soil compaction on infiltration rate. *Journal of Soil and Water Conservation*, 61(3), 117-124.

Gustafson, E. (1998). Quantifying landscape spatial pattern: What is the state of the art? *Ecosystems*, 1(2), 143-156.

Hepcan, S. (2013). Analyzing the pattern and connectivity of urban green spaces: A case study of Izmir, Turkey. *Urban Ecosystems*, 16(2), 279-293.

Hartel, D. (2003). Tree canopy ordinances. Retrieved from http://www.urbanforestrysouth.org/resources/library/ttresources/tree-canopy-ordinances/fss_get/file

Highfield, W. (2012). Section 404 permitting in Coastal Texas: A longitudinal analysis of the relationship between peak streamflow and wetland alteration. *Environmental Management*, 49(4), 892-901.

Hill, E., Dorfman, J., & Gramer, E. (2010). Evaluating the impact of government land use policies on tree canopy coverage. *Land Use Policy*, 27(2), 407-414.

Holben, B. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, 7(11), 1417-1434.

Jones, K., Neale, A., Nah, M., Van Remortel, R., Wickham, J., Riitters, K., & O'Neill, R. (2001). Predicting nutrient and sediment loadings to streams from landscape metrics: A multiple watershed study from the United States Mid-Atlantic Region. *Landscape Ecology*, 16(4), 301-312.

Kays, B. (1980). Relationship of forest destruction and soil disturbance to increased flooding in the suburban North Carolina Piedmont. Proceedings from *Conference of the Metropolitan Tree Improvement Alliance*. Rutgers, NJ.

Leitão, A., Miller, J., Ahern, J., & McGarigal, K. (2006). *Measuring landscapes: A planner's handbook*. Washington DC: Island Press.

Lindell, M., Prater, C., Perry, R., & Nicholson, W. (2006). Hazard mitigation. In *Fundamentals of Emergency Management* (Chapter 7). Retrieved from http://training.fema.gov/hiedu/aemrc/booksdownload/fem/

Looper, J., & Vieux, B. (2012). An assessment of distributed flash flood forecasting accuracy using radar and rain gauge input for a physics-based distributed hydrologic model. *Journal of Hydrology*, 412, 114-132.

Madureira, H., Andresen, T., & Monteiro, A. (2011). Green infrastructure and planning evolution in Porto. *Urban Forestry & Urban Greening*, 10(2), 141-149.

Minteer, B. (2006). *The landscape reform: Civic pragmatism and environmental thought in America.* Cambridge, MA: The MIT Press.

Olivera, F., & DeFee, B. (2007). Urbanization and its effect on runoff in the Whiteoak Bayou Watershed, Texas. *Journal of American Water Resources Association*, 43(1), 170-182.

Park, S., Hepcan, C., Hepcan, S., & Cook, E. (2014). Influence of urban form on landscape pattern and connectivity in metropolitan regions: A comparative case study of Phoenix, AZ, USA, and Izmir, Turkey. *Environmental Monitoring and Assessment*, 186(10), 6301-6318.

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

Pettorelli, N., Vik, J., Mysterud, A., Gaillard, J., Tucker, C., & Stenseth, N. (2005). Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in ecology & evolution*, 20(9), 503-510.

Platt, H. (1998). Planing and Land Use Adjustments in Historical Perspective. In R. J. Burby(Ed.), *Cooperating with Nature: Confronting Natural Hazards with Land-Use Planning for Sustainable Communities* (pp. 29-56). Washington, DC: Joseph Henry Press.

President's Council on Sustainable Development. (1999). *Towards a sustainable America: Advancing prosperity, opportunity, and a healthy environment for the 21st century*. Retrieved from http://clinton2.nara.gov/PCSD/Publications/tsa.pdf

Shandas, V. & Alberti, M. (2009). Exploring the role of vegetation fragmentation on aquatic conditions: Linking upland with riparian areas in Puget Sound lowland streams. *Landscape and Urban Planning*, 90(1), 66-75.

Syversen, N. (2005). Cold-climate vegetative buffer zones as pesticide-filters for surface runoff. *Water Science and Technology*, 51, 63-71.

Tischendorf, L. (2001). Can landscape indices predict ecological processes consistently? *Landscape Ecology*, 16(3), 235-254.

United States Environmental Protection Agency Office of Wetlands, Oceans and Watersheds. (2010). *Green infrastructure case studies: Municipal policies for managing stormwater with green infrastructure* (EPA Publication No. EPA-841-F-10-004). Washington DC.

Veenhuis, J., & Gannett, D. (1986). *The effects of urbanization on floods in the Austin metropolitan area, Texas*: US Department of the Interior, Geological Survey.

Wooldridge, J. (2009). *Introductory econometrics : A modern approach*. Mason, Ohio: South Western, Cengage Learning.

Young, R. (2011). Planting the living city: Best practices in planning green infrastructure—Results from major US cities. *Journal of the American Planning Association*, 77(4), 368-381.

Zahran, S., Brody, S., Peacock, W., Vedlitz, A., & Grover, H. (2008). Social vulnerability and the natural and built environment: A model of flood casualties in Texas. *Disasters*, 32(4), 537-560.

APPENDIX A

CORRELATION MATRIX

Correlation Matrix

	Peak Annual Flow	Mean Annual Flow	Watershed Area	Drainage Area	Floodplain %	Precipitation	Soil Permeability	Impervious %	Slope %	Stream Density
Peak Annual Flow	1.000									
Mean Annual Flow	0.766	1.000								
Watershed Area	0.054	0.266	1.000							
Drainage Area	0.261	0.454	0.138	1.000						
Floodplain %	0.214	0.292	0.111	0.160	1.000					
Precipitation	0.663	0.633	-0.005	0.024	0.194	1.000				
Soil Permeability	-0.069	0.017	-0.045	-0.044	0.102	0.088	1.000			
Impervious %	0.221	0.187	-0.397	-0.048	0.120	0.104	0.094	1.000		
Slope %	-0.170	-0.230	-0.106	-0.039	-0.536	-0.248	-0.234	-0.350	1.000	
Stream Density	-0.172	-0.233	0.101	0.036	-0.227	-0.179	-0.198	-0.719	0.661	1.000
GI %										
in the watershed	-0.059	-0.005	0.053	0.062	-0.090	0.195	0.255	-0.453	0.328	0.303
in the floodplain	-0.139	-0.020	0.100	0.144	-0.244	0.011	0.149	-0.330	0.335	0.210
outside the floodplain	-0.052	-0.032	-0.007	0.013	-0.151	0.209	0.251	-0.446	0.363	0.324
in the 60m buffer	-0.132	-0.095	0.147	0.071	-0.353	0.018	0.151	-0.509	0.459	0.417
PD										
in the watershed	-0.388	-0.354	0.036	-0.033	-0.127	-0.662	-0.289	0.031	0.089	0.117
in the floodplain	-0.067	-0.199	-0.188	-0.117	0.107	-0.326	-0.237	0.203	-0.140	-0.100
outside the floodplain	-0.405	-0.359	-0.028	-0.061	-0.113	-0.667	-0.273	0.053	0.089	0.098
in the 60m beffer	-0.390	-0.402	-0.056	-0.033	-0.305	-0.555	-0.213	-0.295	0.485	0.386
GYRATE_AM										
in the watershed	-0.138	0.053	0.513	-0.069	-0.133	-0.029	0.067	-0.661	0.324	0.412
in the floodplain	-0.161	0.082	0.557	0.257	0.271	-0.056	0.184	-0.557	0.015	0.317
outside the floodplain	-0.023	0.062	0.396	-0.003	-0.148	0.055	0.061	-0.653	0.330	0.433
in the 60m buffer	-0.102	-0.029	0.375	0.064	-0.138	-0.039	0.023	-0.455	0.404	0.403

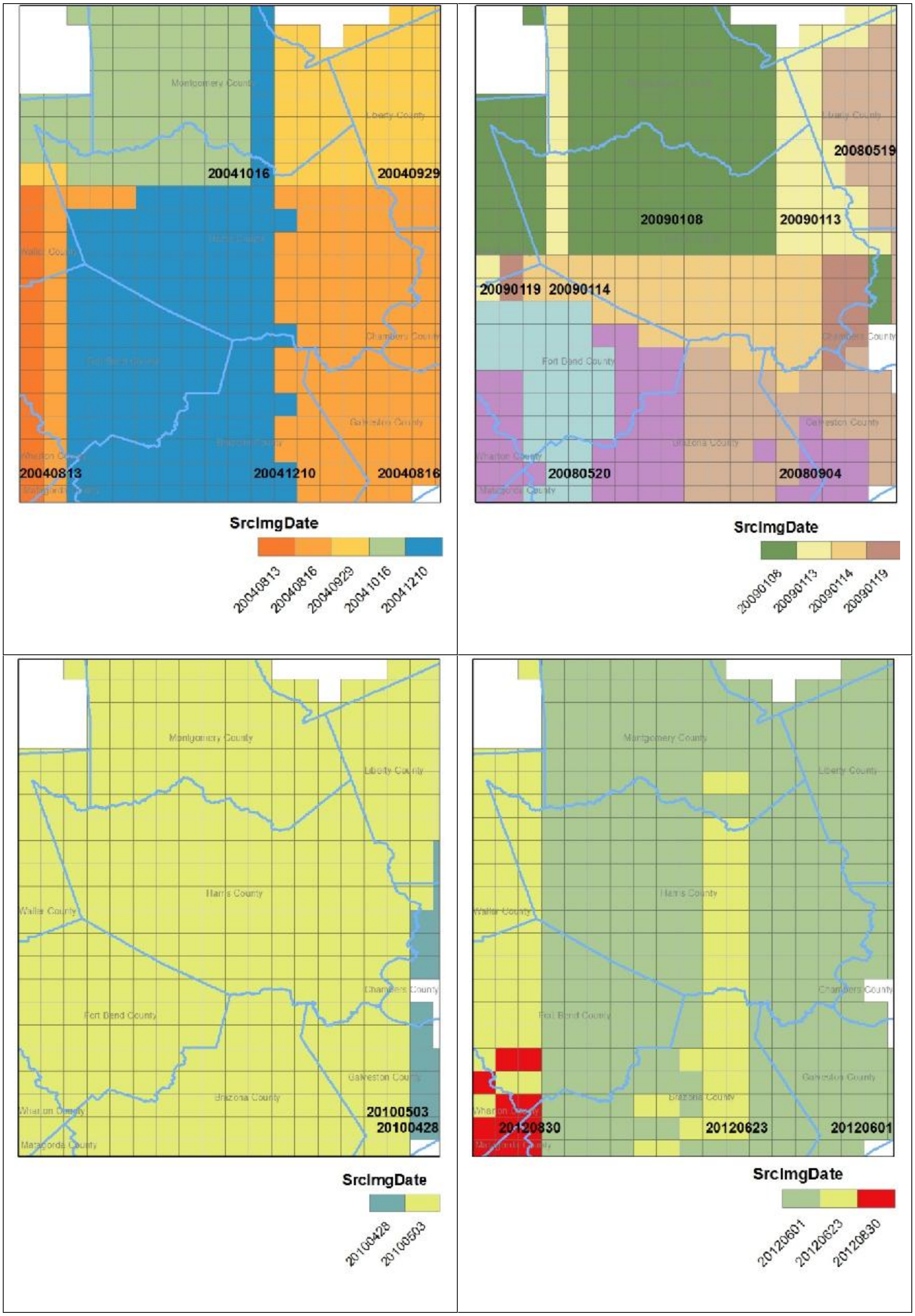
	GI% in the watershed	GI% in the floodplain	GI% outside the floodplain	GI% in the 60m buffer	PD in the watershed	PD in the floodplain	PD outside the floodplain	PD in the 60m buffer	GYRATE_AM in the watershed	GYRATE_AM in the floodplain	GYRATE_AM outside the floodplain	GYRATE_AM in the 60m buffer
GI %												
in the watershed	1.000											
in the floodplain	0.793	1.000										
outside the floodplain	0.981	0.730	1.000									
in the 60m buffer	0.821	0.825	0.809	1.000								
PD												
in the watershed	-0.631	-0.405	-0.617	-0.407	1.000							
in the floodplain	-0.710	-0.790	-0.664	-0.674	0.687	1.000						
outside the floodplain	-0.616	-0.371	-0.612	-0.399	0.996	0.660	1.000					
in the 60m buffer	0.003	0.146	0.005	0.139	0.508	0.244	0.520	1.000				
GYRATE_AM												
in the watershed	0.606	0.485	0.609	0.582	-0.213	-0.341	-0.211	0.140	1.000			
in the floodplain	0.522	0.517	0.428	0.487	-0.239	-0.441	-0.215	0.046	0.703	1.000		
outside the floodplain	0.651	0.483	0.665	0.609	-0.322	-0.389	-0.341	0.087	0.911	0.575	1.000	
in the 60m buffer	0.503	0.522	0.488	0.697	-0.325	-0.434	-0.321	0.053	0.617	0.477	0.658	1.000

APPENDIX B

NAIP ACQUISITION DATES

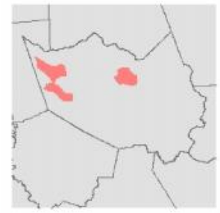
The NAIP acquisition is scheduled such that imagery is gathered during the peak growing seasons in the continental U.S. However, due to delays caused by unusual weather patterns, storms, cloud cover, fires (smoke), and other factors, imagery is not always acquired at the peak growing season. The table below shows each acquisition date for the study area for the years 2004, 2009, 2010, and 2012. For example, 2004 NAIP imagery was acquired on 4 different dates from August 13th through December 10th, whereas 2010 NAIP imagery was acquired on May 3rd. As this study considered a 2-year period in 2004 and 2010 and Normalized Difference Vegetation Index (NDVI) changes from 2004 to 2010, different acquisition dates can be one of limitations of this study.

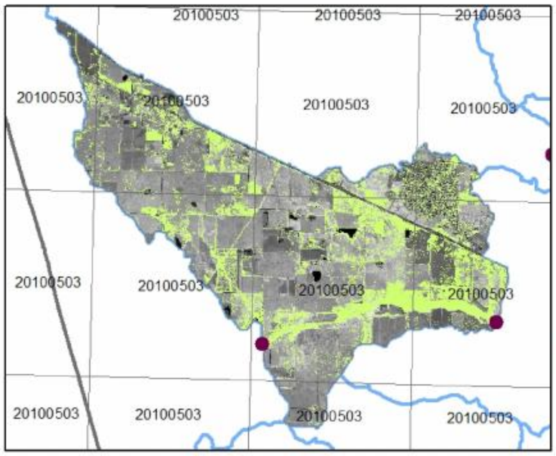
I calculated the NDVI for three watersheds for the years 2004 and 2010 to see if there were significant changes during the 6-year period, regardless of different acquisition dates. For this calculation, value 0 means non-green infrastructure, and 1 refers to green infrastructure. Count means total area in square meters for non-green infrastructure and for green infrastructure. Comparison of the total areas of green infrastructure between 2004 and 2010 for the three watersheds shows that there has been a significant decrease in green infrastructure.


20041016
20040929
20040813
20041210
20040816
SrcImgDate
20040813
20040816
20040929
20041016
20041210
20080519
20090108
20090113
20090119
20090114
20080520
20080904
SrcImgDate
20090108
20090113
20090114
20090119
Montgomery County
Liberty County
Harris County
Waller County
Chambers County
Fort Bend County
Galveston County
Brazoria County
Wharton County
Matagorda County
20100503
20100428
SrcImgDate
20100428
20100503
20120830
20120623
20120601
SrcImgDate
20120601
20120623
20120830

APPENDIX C

GREEN INFRASTRUCTURE COMPARISON BETWEEN 2004 AND 2010



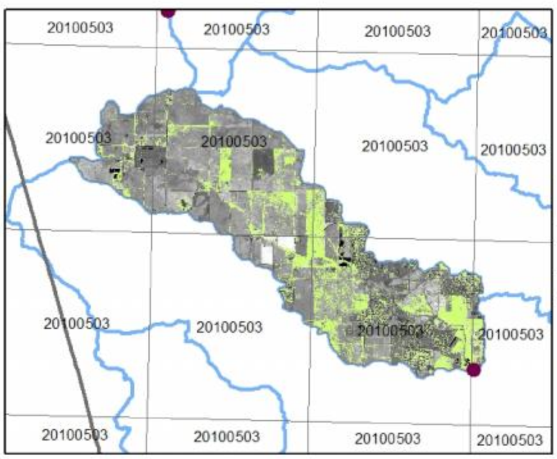


Reclass_NDVI_2004

OBJECTID *	Value	Count
1	0	6830654
2	1	3316027

Reclass_NDVI_2010

OBJECTID *	Value	Count
1	0	7867910
2	1	2278771

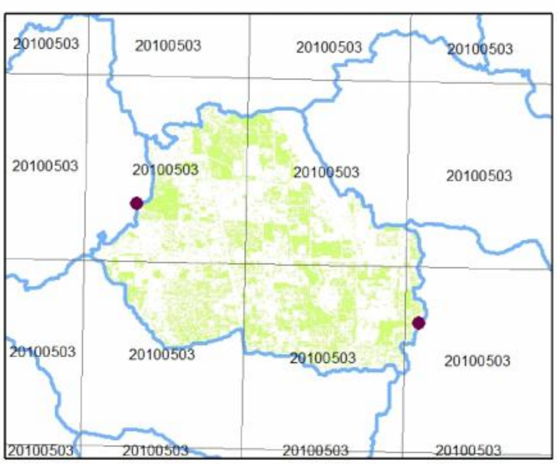


Reclass_NDVI_2004

OBJECTID *	Value	Count
1	0	4133817
2	1	2274609

Reclass_NDVI_2010

OBJECTID *	Value	Count
1	0	5229062
2	1	1179364



Reclass_tif1_NDVI_NAIP_04

OBJECTID *	Value	Count
1	0	4986799
2	1	3199479

Reclass_tif1__NDVI_NAIP_10

OBJECTID *	Value	Count
1	0	5869714
2	1	2516563