Performing a Supervised Land Use Land Cover Change

Analysis to Quantify Urban Expansion in the

Katy Prairie in Harris County, Texas

by

John Moore, B.S.F.

A Directed Research Submitted to Texas
State University in partial fulfillment
of the requirements for the degree of
Master of Applied Geography
GIS
August 2022

Committee Members:

Dr. Samantha Krause

Dr. Edwin Chow

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, John Moore, authorize duplication of this work, in whole or in part, for educational or scholarly purposes only.

ACKNOWLEDGEMENTS

I would like to thank my professors Dr. Krause and Dr. Chow for helping me and always being there for me when I have professional or personal questions and moments of doubt regarding the validity of my research or self. Thank you for giving me the opportunity to impart some wisdom and knowledge of our field onto me.

I also would like to express my gratitude to my friends and family who encouraged me along the way, believing in my decision.

TABLE OF CONTENTS

1. Problem Identification

Urbanization is the process of changing the landscape to facilitate a metropolitan region and sustain larger populations, though urban sprawl requires changing the land cover from herbaceous or forested land to developed land with impervious cover over the soil (Minnig 2017). Impervious cover is defined as anything that inhibits water infiltration and percolation. Examples include but are not limited to concrete or asphalt (Guo 2017). Land use and land cover (LULC) dynamic shifts are occurring across the globe, creating negative environmental impacts such as increased flooding, habitat degradation, and declining groundwater levels (Doyle 2021, Nath 2021, Zope 2017). All of these changes described above affect developing regions such as China, India, and Africa as functions of Gross Domestic Product (GDP) and urban population growth, while regions like North America urban land expansion is a driving factor (Seto 2011). In Texas, urban land expansion is focused on cities like Houston, the Dallas/Fort Worth Metroplex, and the Interstate 35 (I-35) Corridor encompassing San Antonio and Austin. To better understand urban land expansion and all the negative environmental impacts associated with this process, remote sensing scientists have developed several key methods to quantify the dynamic over time. LULC change analysis is one of these robust and well documented method, in which remotely sensed raster data is used to quantify the land cover change. This enables researchers to visualize the changing landscape and the negative ecological impacts that arise from increased impervious cover. This type of analysis is important for understanding ecosystem changes, increased flood hazards, and overall human health and vulnerability at a variety of scales which is evident from the work of Zope and coauthors in Mumbai, India and Wang in Hong Kong, China. Both studies LULC change analysis to identify how the regions have changed over time and what impacts that has had on the ecosystem, water cycle, and stormwater infrastructure (Zope et al. 2017, Wang et al. 2021).

Texas is home to the third largest county in the United States, Harris County. This county serves as an excellent example of urban expansion, as it has averaged 57,813 new residents annually from 1980 to 2021 (World Population Review, March 2022). This population growth has increased urban development within it as well surrounding counties as residents move outward into adjacent counties such as Fort Bend and Waller. Such adjacent counties have experienced average population increases of 17,758 and 949 people per year, across the same time. The assumption is that the LULC changes that sustain this growing population are having negative ecological and hydrologic effects in the Katie Prairie, which has portions of the undeveloped regions being used as detention ponds for the Houston Area (Garner 2020).

1.1 Research Questions

The questions driving this directed research are:

R1. What is the LULC change between 1989 and 2021 in the region?

R2. What are the changes in LULC specifically in the Kate Prairie from 1989 to 2021 that stem from the increased urbanization driven by population?

R3. If there is a change in impervious cover to the study area, how badly will the built environment be impacted during extreme weather and flood events? What types of flood models are appropriate for future analysis of this region?

To answer these questions, I hypothesize that that over the course of 32 years, impervious cover and the built environment has significantly increased into the Katy prairie and changed the landscape from a predominantly herbaceous landcover to that of urban and suburban land uses, thus changing the ecology and hydrology of the region. This work is important to many stakeholders and residents within the Houston-Galveston Area Council region because this project can utilize the methodology to recreate and compare the results of neighboring areas to determine the change. This project also highlights the expansion of an urban environment which brings other hydrologic problems to the area when discussing impervious cover in flood prone areas like Harris County. The Katy prairie (Figure 3) has historically been an area that allows for heavy infiltration due to the soil and vegetation type of the coastal prairie and replacing that landcover with concrete or other impervious cover disrupts that process and shifts the job of water displacement and retention onto municipalities and stormwater infrastructure (Garner 2020). Organizations such as Flood Control or Drainage Districts, Emergency Management departments, and local governments will benefit from having this knowledge and understanding the importance of flood mitigation, retention and detention ponds, and stormwater infrastructure to mitigate the increase of impervious cover caused by an increase in the region's population. This information will be placed on an interactive web map where users will be able to swipe over the satellite imagery, the classified images, and the hydrologic models to see how the community has shifted over the last 32 years. This website can also use National Land Cover Database (NLCD) data for a more in-depth land cover classification and in a 5-year time step to garner a more cohesive image in how the landscape has changed compared to the individual work done in this project of two classifications at the beginning and end of the time period in question.

2. Literature Review

2.1 Land Use Land Cover

LULC analysis of raster data is an important tool for quantifying how humans are manipulating the landscape around them, the term can be broken into two phrases that describes the ecosystem of the land, "land cover", and how the land is being utilized in the human sphere, which is called "land use" (J.R. Jensen 2015). Land use examples can be cotton fields, urban areas, or pine forests for timber production, whereas land cover may define these same areas in more vague terms describing herbaceous vegetation, forested land, and impervious cover (Das 2020). This categorization of land cover is necessary in determining the impacts that are associated with the increasing impervious cover from development and negative ecological impacts.

The process of documenting and quantifying LULC change can be laid out in a multistep approach. The first step is to determine the study area, timeframe, and goal of the research. This step allows researchers to determine the correct raster data to gather and know the limitations of their data. The second step is to collect the data for the LULC analysis from the appropriate sources depending on step one. This data can be sourced from numerous different earth observation satellites that are maintained by multiple different countries, including but not limited the United States of America, Japan, China, and Russia (Navin 2020). For this study, we will be using data gathered from Landsat 5 and Landsat 8, which can be downloaded by using the United States Geological Survey

(USGS) Earth Explorer web tool. The third step in the LULC analysis process is to preprocess the images. This involves georectifying the images, applying convolution filters, or smoothing filters to fix imaging errors, correct for cloud coverage in the image, or even create a similar pixel size between different images (Navin 2020). The importance of fixing these errors comes down to creating a reliable and consistent data set that can create repeatable studies in other regions of the world and reduce noise in the image. Next, the processed images need to be classified using LULC classification methods, this can be done using an already existing classification schema or by creating one specific to the region. This classification schema is then validated using a reference image and the appropriate classification logic is performed (Jensen 2015). Finally, after the appropriate LULC classification logic has been performed, the researcher validates and either accepts the results for interpretation, redoes the classification using another method to achieve the desired results, or rejects the initial hypothesis (Jensen 2015). Research done by Xiaoyong Li in 2020 highlights the ability to accurately perform LULC analysis over urban areas to map urbanization over time and then perform additional analysis to infer the environmental impacts. Concluding that with an increase in rainfall intensity the efficacy of current urban surface runoff regulation systems in place, like stormwater systems or urban green spaces, are decreased (Xiaoyong 2020). This is due to an overall increase in population and urban surfaces in Hong Kong with a decrease in urban green spaces, increasing the runoff. Noting that with new construction in the built-environment methods to challenge these impacts are made and interspersed green zones are recommended to be established to facilitate more recharge and regulation zones (Xiaoyong 2020). Similar studies have been performed by Doyle et. al. (2021) in Belize in an agricultural setting using Google Earth Engine, and by Wang et. al (2021) for the greater Hong Kong region in China, again for urban areas. These studies show the different GIS and remote sensing tools that are available to complete similar styles of research. Using a binary change detection method over a specific AOI small enough that does not require mosaicking, there can be quantifiable change detected in pixel value that can be used to document the LULC change occurring due to urbanization surrounding Houston Texas (Jensen 2015).

2.2 Hydrology

Hydrology is the study of water and its behavior in and around the Earth. Hydrological processes are a key factor for humanity because although water is a necessity to sustain life, it also creates incidents and disasters that cause flooding or drought conditions (Kim 2012). Understanding the unique hydrology of a region and the relationship water has with a changing environment is a quintessential part of this study and others like it which look to assess how land use changes and urbanization shift the historic water cycles.

The hydrologic cycle is defined as the “circulation of water through the atmosphere, land, lakes and oceans” implying that any transformative action involving urbanization or land use changes affect the water cycle and the quality of that cycled water (Guo 2017). Multiple studies have been conducted comparing how land use land cover change or urbanization has affected the study area’s groundwater and flood potential. There are two distinct types of studies being performed surrounding this topic: discrete and continuous. In the discrete studies, researchers look to solve issues surrounding flood events and posit effects of increased urbanization without hydrologic or stormwater planning (Zope 2017, Anni 2020, Ewane 2020). The other type of study, continuous, research how the land use change or status quo affects the hydrological cycle daily. These types of studies research current urban structures in place and their efficacy, groundwater recharge for a region, or current mitigation zones (Minnig 2017, Xiaoyong 2020, Garner 2020, Alexandres 2014). The methods for conducting this research are also varied and dependent on the overarching goals of the researchers. Methods like a surface water analysis (SWAT), quantitative precipitation estimation (QPE), or water budget calculation are useful when studying datasets over a period to project certain steps of the hydrologic cycle over various temporal scales (Eshtawi et al. 2016, Minnig 2017, Woodson et al. 2019). The other type of hydrologic analysis is discrete events where the current region is tested with extreme events. This is normally performed using modeling systems and inputting flood risk data that is region specific to create models (Zope 2017, Natarajan 2019). These types of studies still require external datasets of climate and weather commonly using duration-based step data modeled off of frequency values collected from national or local rain gauges found within the AOI, but utilize the modeling systems to simulate flood or extreme events. This information can be obtained from sites like the National Oceanic and Atmospheric Administration's (NOAA) Hydrometeorological Design Studies Center (HDSC) Precipitation Frequency Data Server with durations ranging from 5 minutes to 60 days and average recurrence intervals ranging from 1/1 year to 1/1000 years. Precipitation amounts on the table for Houston range from 0.502 inches at a 5 minute duration and 1-year recurrence to 45.1 inches at a 60 day duration and 1/1000-year recurrence (NOAA Hydrometeorological Design Studies Center July 2022).

A commonality between both methods is visualizing the data using a hydrologic modeling software. There are a few different types of software on the market that are in current use today, the United State Army Corps of Engineers has created multiple platforms that are utilized for different types of hydrologic modeling. The Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS) and River Analysis System (HEC-RAS) are two popular applications that are used for modeling hydrologic

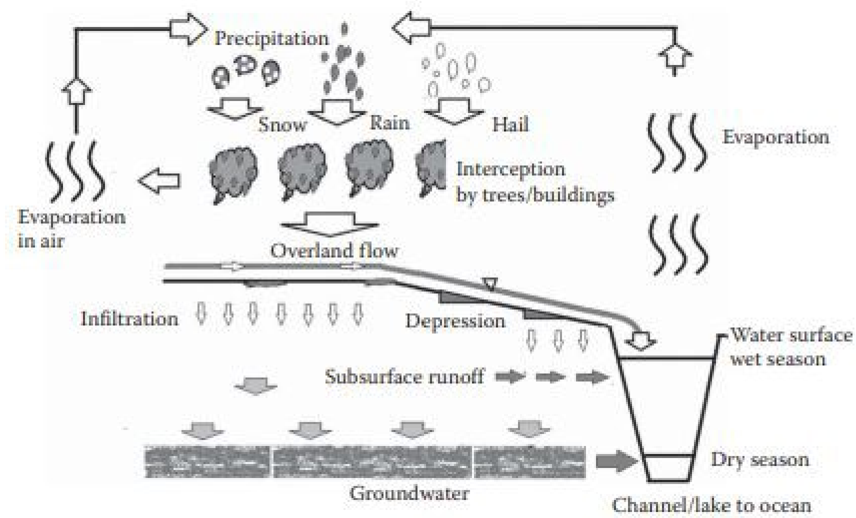

Figure 1. Hydrologic Cycle (Guo 2017).

systems because of their robust nature and stability in producing accurate results (Garner 2020, Natarajan 2019). HEC-HMS is an often-used program for this type of modeling because it offers a combination of common GIS functionalities with hydrology data to create unique flood models (Kabeja et al 2020).

3. Site and Situation

3.1 Economic and Population Growth

The greater Houston area is experiencing a large influx of new residents, with counties on average increasing in population by 222% from 1980 to 2021 projections (World Population Review May 2021). The population growth Houston is experiencing is well-documented and has been occurring since the 1950s with the population doubling from

1950 to 1960, going from 807,000 to 1.4 million residents (Stevens 1981). An expansive port hub, booming oil and gas economy, and one of the best healthcare centers in the nation have brought millions of people to the region looking for work and a better life over the past 50 years (Greater Houston Partnership April 2022, Forbes June 2022). This population increase requires expansion into regions that were previously agricultural and turning them into urban regions with impervious cover that inhibits the infiltration step in the hydrologic cycle, shown in Figure 1 and Figure 2, by increasing runoff and diverting water out of the region using storm drains and bayous (Nath 2021). This increase in built up area negatively impacts the environment by increasing the flood potential during storm events as shown by the work from Zope et al. (2017) where urbanization in Mumbai increased the flood peak discharge and decreased flood lag times and disrupts the hydrologic cycle in a region that historically is an important recharge zone (Zope 2017, Garner 2020). This disruption may have unforeseen consequences that cannot be accurately assessed until after the damage is done. For this reason, it is important to quantify and model the LULC and hydrology to create a baseline for future research to theorize impacts during flood.

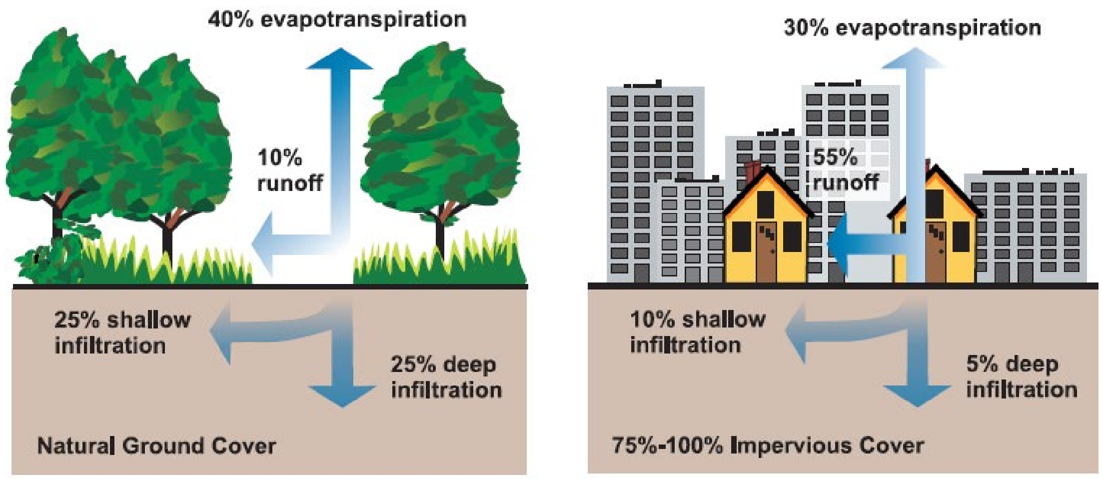

Figure 2. Infiltration compared between a forested or natural ground cover compared to majority impervious cover (ConstructionJunkie June 2022).

3.2 General Physical Description

Described to be a humid subtropical climate, Houston metro area averages an annual 52 inches of rainfall and an average temperature of 70°F, with temperatures exceeding 100°F in the summer and monthly highs being recording in the low to mid 90s as seen in Table 1 (Weather.gov March 2022). Harris County lies within the Gulf Coast Prairie and Marsh ecoregion of Texas, a flat low-lying region with elevation around 150 feet (Texas Parks and Wildlife Department July 2022). The region is inundated with streams and rivers all snaking into the gulf from deeper into the region with marshes, estuaries, and bays converging on the coast to act as a barrier between the ocean and land (Texas Parks and Wildlife Department July 2022). The native vegetation for the region is tallgrass prairies and live oak woodlands which act as habitat and spawning zones for migratory bird, fish, and shrimp (Texas Parks and Wildlife Department July 2022). The water table for the region ranges from 10-30' below land surface depending on the land surface altitude of the area in question with higher altitudes having a deeper water table (Noble et al. 1996). Harris County Flood Control District manages the drainage networks within the county and maintains a network of over 800 miles to mitigate flood and high-water events (Harris County Flood Control District July 2022). Common hazards for the region are hurricanes coming in from the Gulf of Mexico, related wind events such as tornados during these storms, and most importantly Flooding (Harris County Flood Control District July 2022). A recent example is in 2017, in which Hurricane Harvey brought devastating flooding to the entire region surrounding Houston and Harris County, dropping a record 60 inches of rain onto the region, causing an estimated $125 billion in damages, and being the direct cause of 68 deaths (Blake 2018). And future development will only bring more negative hydrological impacts if there is no change to the status quo. Additionally with the onset of climate change and pumping groundwater for use, the oceans along the Gulf of Mexico are rising, Galveston Beach specifically has risen 18 inches since 1950, and is now rising at a rate of 1 inch per year (Climate.gov and sealevelrise.org July 2022).

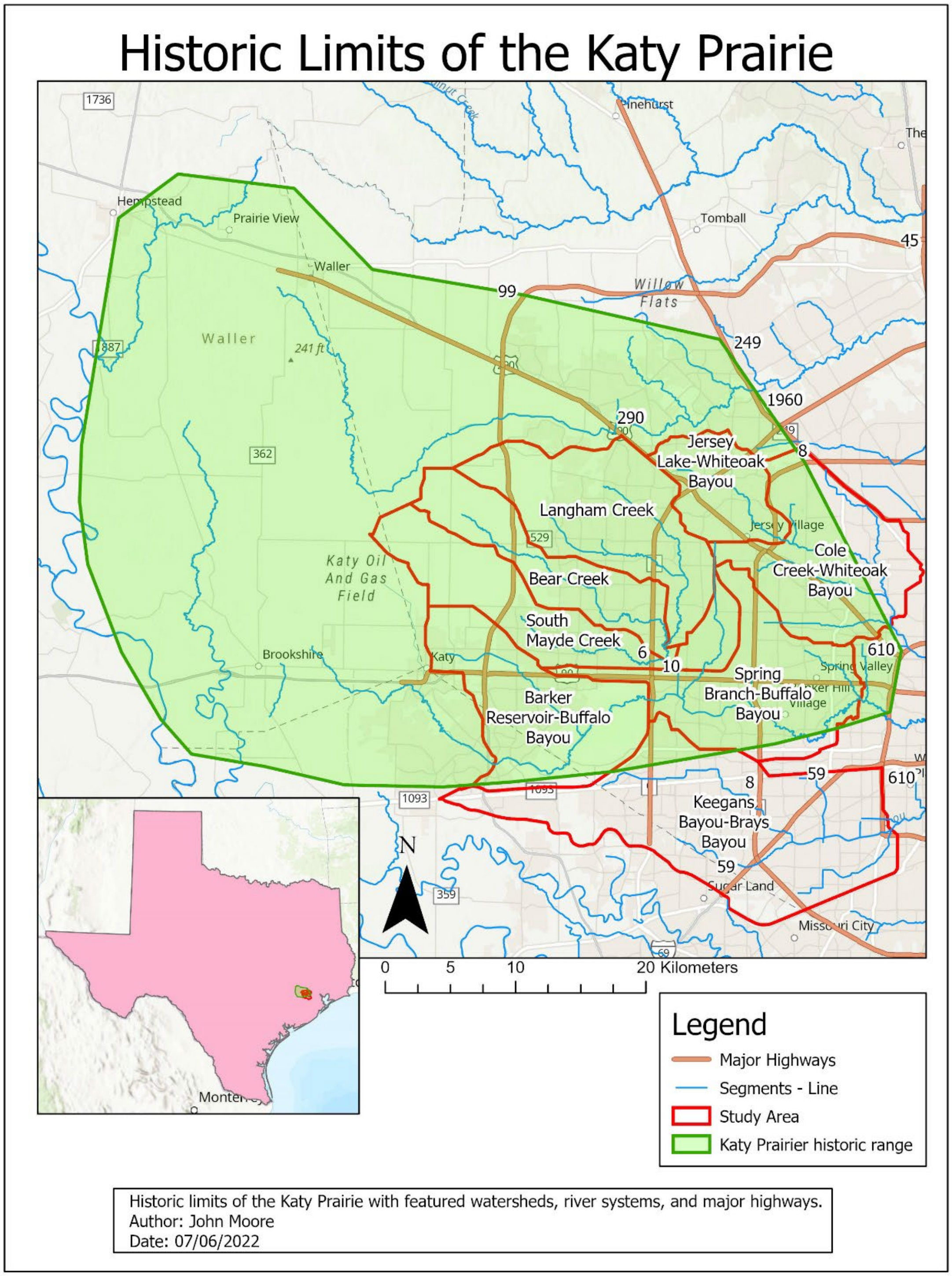

Figure 3. Historical Range of the Katy Prairie.

3.3 Specific Study Area Description

Coastal Prairies are a subset of the larger prairie system that extends North from Mexico all the way North into Canada and through the United State's Midwest. These coastal prairies historically followed the gulf coast from Mexico through Texas and into Louisiana (Garner 2020). Differing from other prairie systems in more northern states, coastal prairies are considered a subtropical climate receiving just under 52 inches of rainfall and averaging 70 °F annually from 1991 to 2020 (Weather.gov July 2022). Coastal prairies such as the Katy Prairie are often referred to as a prairie marsh system, due to the slow permeable clay, little to no elevation changes, and the ability to retain water after storm events (Texas Parks and Wildlife, Walker and Miears 1957). These systems contain a diverse micro-climate due to the soft rolling mounds and sloughs that can be found in the region (Hale et al. 2014). The retention capacity of these coastal prairies are also defined by the herbaceous vegetation found on site. The root structure of the species found in the region have deep fibrous root systems that increase infiltration through soil pores created by the structure (Dunne et al. 1991).

The location for this research is a roughly 1614 sq. km. area that spans across three counties, Fort Bend, Harris, and Waller (Figure 4). The east boundary begins at the beltway 8 in Houston, traveling west following Interstate Highway 10 with the west boundary stopping in Brookshire, Texas in in the middle of Waller County. This area was chosen due to the increased urban expansion that is directly visible over the past 40 years, with perceivably half of the area changing from agricultural land to urban area or subdivisions. The AOI completely covers or partially intersects cities such as Brookshire, Cinco Ranch, Fulshear, Katy, Jersey Village, and Mission Bend. These are all cities that have experienced or are currently experiencing large expansions in the last 40 years due to the sprawling urbanization stemming from development and increased migration to Houston and its surrounding suburbs. The Katy Prairie is specifically of interest for its ability as a retention and detention zone for stormwater and flood mitigation (Garner 2020). The Coastal Prairie Conservancy, formerly the Katie Prairie Conservancy, advocates for the reclamation, conservation, and preservation of the remaining acreage of the Katie Prairie, as well as having long term goals of protecting upwards of 50,000 acres (Coastal Prairie Conservancy June 2022). They maintain that the Katie Prairie and other coastal prairies like it provide unique ecological and monetary benefits to the surrounding communities that the urban environment does not provide. These benefits come in the form of carbon sequestration, increasing species diversity in both flora and fauna, and increasing the detention and retention of water during and after rain events. The degradation and fragmentation of this ecoregion specifically would increase storm water and flooding to all the land cover replacing the prairie. This has negative impacts for both the ecology and humans residing in this area due to habitat destruction, increased flooding, urban heat island effects, and economic damage suffered by the residents replacing the Katy prairie.

Table 1. Monthly 30 year normals from 1991 to 2020 measuring rain total, mean temperature, average high, and average low (Weather.gov, July 2022).

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Rain Totals (in)	3.76	2.97	3.47	3.95	5.01	6.00	3.77	4.84	4.71	5.46	3.87	4.03
Mean Temp (°F	53.80	57.70	63.80	70.00	77.40	83.00	85.10	85.20	80.50	71.80	62.00	55.40
Avg High (°F)	63.80	67.80	74.00	80.10	86.90	92.30	94.50	94.90	90.40	82.80	72.60	65.30
Avg Low (°F)	43.70	47.60	53.60	59.80	67.80	73.70	75.70	75.40	70.60	60.90	51.50	45.60

4. Methods

4.1 Data

The data used for this project are Landsat 5 and 8 Raster images sourced from The USGS Earth Explorer online tool, watershed data sourced from the USDA, and DEM Data sourced from Texas Natural Resource Information System (TNRIS). The Landsat Imagery was acquired for the same calendar month if possible and with as minimal cloud cover (<10%) to reduce interference with the results of the LULCC analysis. The Raster images are clipped to the bounding box of the watersheds that are the focus of the study so there is no excess data.

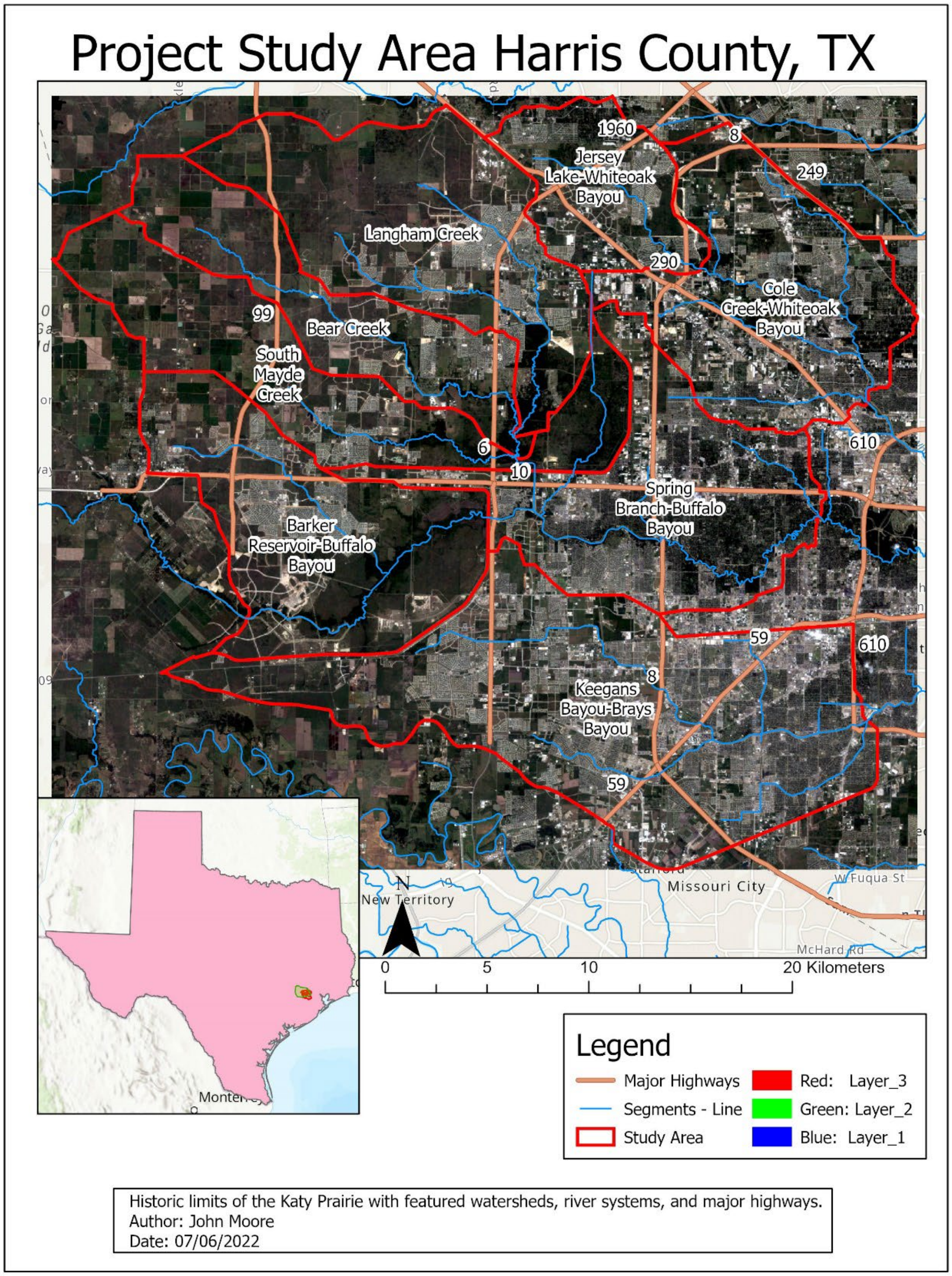

Figure 4. Map of the study area to be classified with major highways and waterways denoted.

4.2 Analysis

The two raster images that are being used for the LULCC Analysis were first imported into ERDAS Imagine 2.8 and compiled there using the layer stack tool. The stacked images were then opened into ArcGIS Pro to be clipped to the bounding box of the eight watersheds that are the study area. After being clipped the raster images from Landsat 8 and Landsat 5 were processed in ERDAS Imagine 2018 (16.5) using the Normalized Difference Vegetation Index (NDVI) tool to visualize the vegetative production present in the images. The NDVI Function compares the near-infrared band and red band to produce an image that compares vegetation and nonvegetative objects (example: water or roads) ranging from a pixel value of 1 to -1.

$$NDVI = (NIR - Red) \div (NIR + Red)$$

By producing images that highlight productive vegetation this creates a visualization of the area of interest for the user that increases differentiation of vegetated regions versus built-up regions which is useful during the training process. After this process is completed for both images training data was created using the signature editor tool. The tool is used by creating polygon vectors of similar pixels contained on the raster. For the purposes of this project, five classes were delineated: urban, suburban, forest, herbaceous vegetation, and soil. These classes were defined based on prior knowledge by the researcher of the area as well as pixel values from the satellite imagery and NDVI Images. This was completed in an ad-hoc manner with the Landsat imagery being the source material for differentiating between classes such as urban or suburban and herbaceous vegetation or forest. For the Urban class a cluster of commercially zoned infrastructure was used as the basis for the model and a series of neighborhoods were used for the suburban class dataset. Forested land within the Barker reservoir was used as the basis for the forest class and various agricultural fields were used when training the herbaceous vegetation class. Water and clouds were initially included in the classes for the images, but both classes were being misrepresented with other classes and diluting the efficacy of the project so they were removed. The training data was selected using groups of similar pixels that researchers determined were quality representations of the classes presented above. After the training data was completed, a supervised classification was performed. This tool utilizes the training data and expands the classes to the rest of the raster image to produce a new image coloring the pixels with the determined class for later evaluation. For this project, the classification was performed with decision rules as the non-parametric rule being parallelpiped and the parametric rule following the maximum likelihood. The parallelpiped ruling allows for pixels not explicitly defined within the training data to be considered for that classification with the maximum likelihood ruling and assumes that a particular pixel belongs to a particular class. By using these two methods within the supervised classification, the tool creates a robust classification schema that encompasses the raster images.

Once the supervised classification is performed on both images an accuracy assessment is performed. This assessment is necessary to the model because it compares how the researchers perceive the points compared to how the classification tool perceives the points and determines if the model fits the raster image. The number of points used in the accuracy assessment to determine a significant sample size is based on the multinomial distribution equation. This equation is used when there are more than two groups being observed (Jenson 2015).

$$N = \frac{B\pi_i(1-\pi_i)}{b_i^2}$$

Where Π_i is the proportion of a population that has the proportion closest to 50% in the *i*th class out of *k* classes, b_i is the desired precision and B_i is the upper (α/k) x 100^{th} percentile of the chi square (x^2) distribution with 1 degree of freedom(df) (Jenson 2015). *B* is determined from the x^2 table with 1 df and 1- α/k with (Jenson 2015). With five classes and a desired confidence interval of 80% the formula to determine *B* is As $X^2_{(1,0.96)} = 4.21788$:

$$1 - \frac{\alpha}{k} = 1 - \frac{0.20}{5} = 0.96$$

$$N = \frac{4.21788(.30)(1-0.30)}{0.05^2} = 354.312$$

This result means to be 80% confident in the sample size of the accuracy assessment at least 354 points are required. To make the points even this number was rounded up to 360 points, or 72 points per class. The accuracy assessment was conducted using a stratified random sampling with a minimum number of points collected in each class being 50 to ensure that each class was represented as a safety measure (Jenson 2015). The accuracy assessment was performed for both images with the researchers matching the designated pixel with the class they deemed appropriate. Once the assessments were completed an accuracy report was generated that contained tables of the error matrix, accuracy totals, and kappa statistics. The purpose of these tables and data outputs from the accuracy assessment compare the efficacy of the computer-generated classification and the overall quality of the training data. If the kappa coefficient and accuracy totals are too low, the process must be redone with new training data to have classes with differentiation that is cohesive across the raster image.

Once the accuracy assessments were completed and the kappa coefficient and overall classification accuracy were at acceptable levels, an image difference change detection function was used to determine what pixels changed in value between the two images with a 10% change in rgb spectral reflectance up or down. Additionally, a union matrix tool was utilized between both images to determine what class changed from the 1989 image to the 2021 image. This was performed on the two classified images to highlight differences between the two and visualize what land cover types have changed to if change occurred.

5. Results

5.1 Classification Results

With the supervised classification performed over both images, the class sizes and percentages of each raster image can be calculated based on pixel size and total area of the raster images. For the 1989 Landsat 5 raster image are classified as: Bare Soil 31.11 km^2 (2%), Forest 247.66 km^2 (15%), Herbaceous vegetation 735 km^2 (46%), Suburban 300.68 km^2 (19%), and Urban 299.52 km^2(19%) totaling 1614.28 km^2. For this image, the dominant land cover classification was the Herbaceous vegetation, encompassing both agricultural land, municipal right of ways, golf courses, and other non-wooded vegetated areas. Ties for the second largest land cover type are the Urban and Suburban cover types each making up 19% of the image. The 2021 Landsat 8 raster image are classified as: Bare Soil 62.05 km^2(4%), Forest 207.19 km^2(13%), Herbaceous vegetation 149.47 km^2 (9%), Suburban 784.53 km^2(49%), and Urban 405.23 km^2(25%) totaling 1608.47 km^2.

Table 1. Supervised classification results for the 1989 Landsat 5 and 2021 Landsat 8 imagery sorted by class order.

Class	**1989 Area (km^2)**	**1989 Percent**	**2021 Area (km^2)**	**2021 Percent**
Urban	299.52	19%	405.23	25%
Suburban	300.68	19%	784.53	49%
Forest	247.66	15%	207.19	13%
Herbaceous Vegetation	735.31	46%	149.47	9%
Bare Soil	31.11	2%	62.05	4%
Total	1614.28		1608.47	

Unlike the Landsat 5 imagery, the Suburban land cover class represents nearly half of the image as seen in table 1. Urban land cover represents one-quarter of the image, and bare soil, forested land, and herbaceous vegetation represents just over one-quarter of the image combined. Due to pathing differences between the two images, there are a total of 1605.31 km^2 of overlap between the two images according to the error matrix.

5.2 Accuracy Assessment

This step in the process of performing a LULC change analysis is imperative to determine the fitment of the produced models and how well the training data models the rest of the raster images. Both images utilized the same methodology of 360 points with a stratified random sample selected for this step. Each class has a producer and user accuracy, determining the sensitivity of the class (Producer) and reliability to determine what class the location belongs to (User). The overall accuracy is determined by dividing the number of locations deemed correct by the total number of locations created on the raster image.

$$\frac{Number\ of\ points\ Correct}{Total\ Number\ of\ points} = \frac{258}{360} = 71.67\%$$

Beginning with the assessment of the Landsat 5 image, the overall classification accuracy is 71.67%. The user accuracy ranges from forest being the lowest at 55.22% to herbaceous vegetation being the highest at 78%. Additionally, the overall kappa statistic for this accuracy assessment is 0.5619, which is considered moderate in terms of agreement with the model with individual kappa coefficients ranging from forest with 0.5294 at the lowest to urban with 0.6033 at the highest. Comparatively, the Landsat 8 image achieved an overall classification accuracy of 66.67% with 240 locations correctly with user accuracy ranging from soil being the lowest at 54% to herbaceous vegetation at the highest accuracy with 72.41%. The overall kappa statistics for the Landsat 8 image are 0.5721, which equates to a moderate agreement with the classification model, and individual coefficients ranging from soil with a coefficient of 0.5012, to herbaceous vegetation with a coefficient of 0.6587. Table 6 indicates the classification change from 1989 to 2021, with the rows indicating the 1989 classes, and columns indicating the 2021 classes.

Table 2. Accuracy totals for the 1989 imagery post-supervised classification accuracy assessment.

Class name	Reference Totals	Classified Totals	Number Totals	Producers Accuracy	Users Accuracy
Urban	62	67	45	72.58%	67.16%
Suburban	74	67	42	56.76%	62.69%
Forest	54	55	33	61.11%	60.00%
Herbaceous Vegetation	160	164	125	78.13%	76.22%
Bare Soil	10	7	4	40.00%	57.14%
Totals	360	360	249		
		Overall Classification Accuracy =			69.17%

Table 5. Kappa (K^) statistics for both imagery organized by class order.

Class Name	1989 Kappa	2021 Kappa
Urban	0.6033	0.5246
Suburban	0.5303	0.5667
Forest	0.5294	0.6222
Herbaceous Vegetation	0.572	0.6587
Bare Soil	0.5592	0.5012
Overall Kappa Statistics	0.5619	0.5721

Table 3. Post-Supervised Classification error matrix for the 1989 image.

Classified Data	Urban	Suburban	Forest	Herbaceous Vegetation	Bare Soil
Urban	45	14	0	4	4
Suburban	10	42	4	11	0
Forest	0	2	33	20	0
Herbaceous Vegetation	5	16	16	125	2
Bare Soil	2	0	1	0	4
Column Total	62	74	54	160	10

Table 4. Post-Supervised Classification error matrix for the 2021 imagery.

Classified Data	Urban	Suburban	Forest	Herbaceous Vegetation	Bare Soil
Urban	50	21	2	5	1
Suburban	27	78	5	1	0
Forest	1	10	43	8	0
Herbaceous Vegetation	0	1	15	42	0
Bare Soil	4	3	3	13	27
Column Total	82	113	68	69	28

Table 6. Matrix union data from 1989 to 2021 describing the number of square kilometers that transferred from one land cover type to another.

	Urban (km^2)	Suburban (km^2)	Forest (km^2)	Herbaceous Vegetation (km^2)	Bare Soil (km^2)	Total (km^2)
Urban (km^2)	141.79	143.58	5.67	6.74	1.58	
Suburban (km^2)	57.23	230.15	6.79	5.06	1.28	
Forest (km^2)	34.26	84.00	101.48	17.59	9.06	
Herbaceous Vegetation (km^2)	163.67	313.73	90.56	113.82	46.46	
Bare Soil (km^2)	8.06	12.45	1.62	5.04	3.67	
Sum (km^2)	405.00	783.91	206.12	148.24	62.04	1605.31

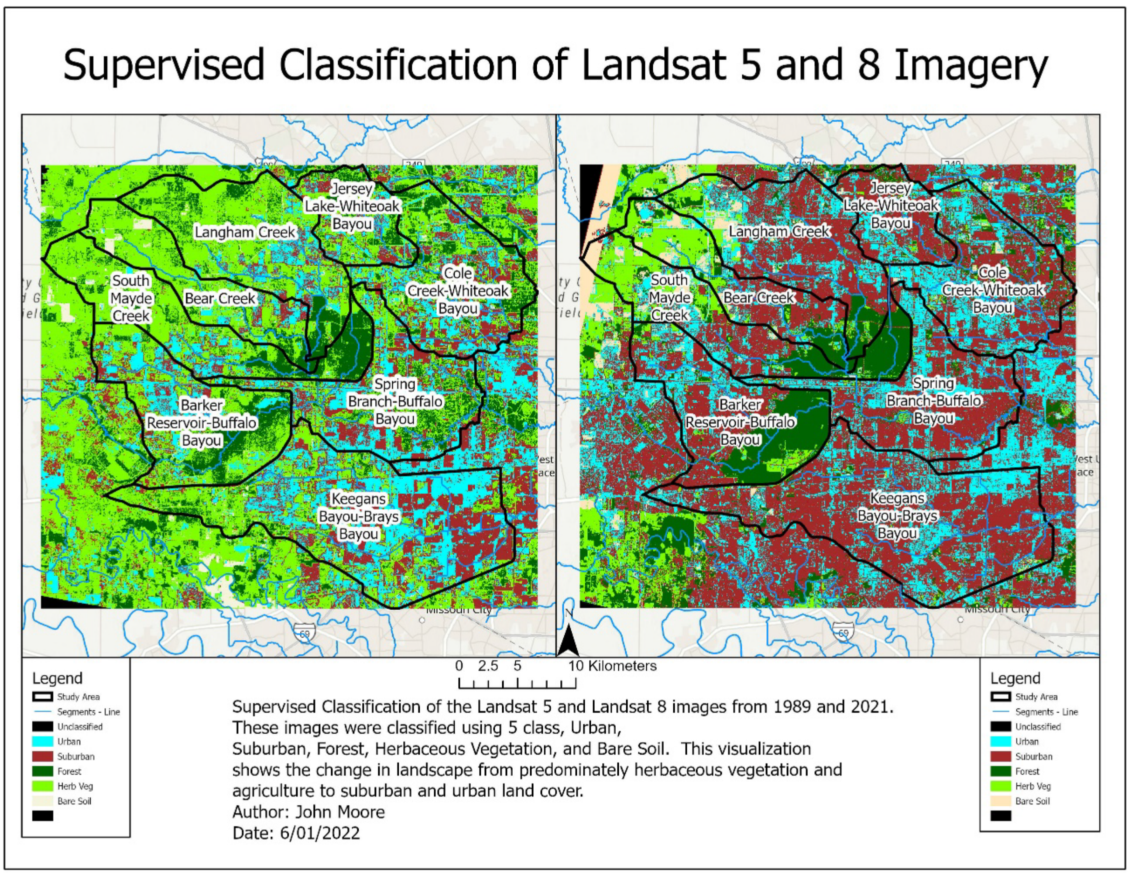

Figure 5. Supervised Classification of the Landsat Imagery from 1989 and 2021.

6. Discussion

6.1 Major Findings

From the LULC change analysis (Figure 5), 1605.31km^2 were comparable due to pathing differences with the satellites and a total of 1023.51 km^2 changed in cover type from 1989 to 2021. The major land cover shift comes from overall herbaceous vegetation decreasing by 580km^2 with 65.5% of that shifting to urban or suburban land uses. This change can be recognized in the satellite imagery and classified images in the western side of the study area where Houston's suburbs have expanded into former agricultural land. Suburbs like Sugar Land have grown in population dramatically from 1989 to 2021 over quadrupling in size, with other master planned communities experiencing similar pressures of urban growth which can be seen in Figure 7 (World Population Review, April 2021).

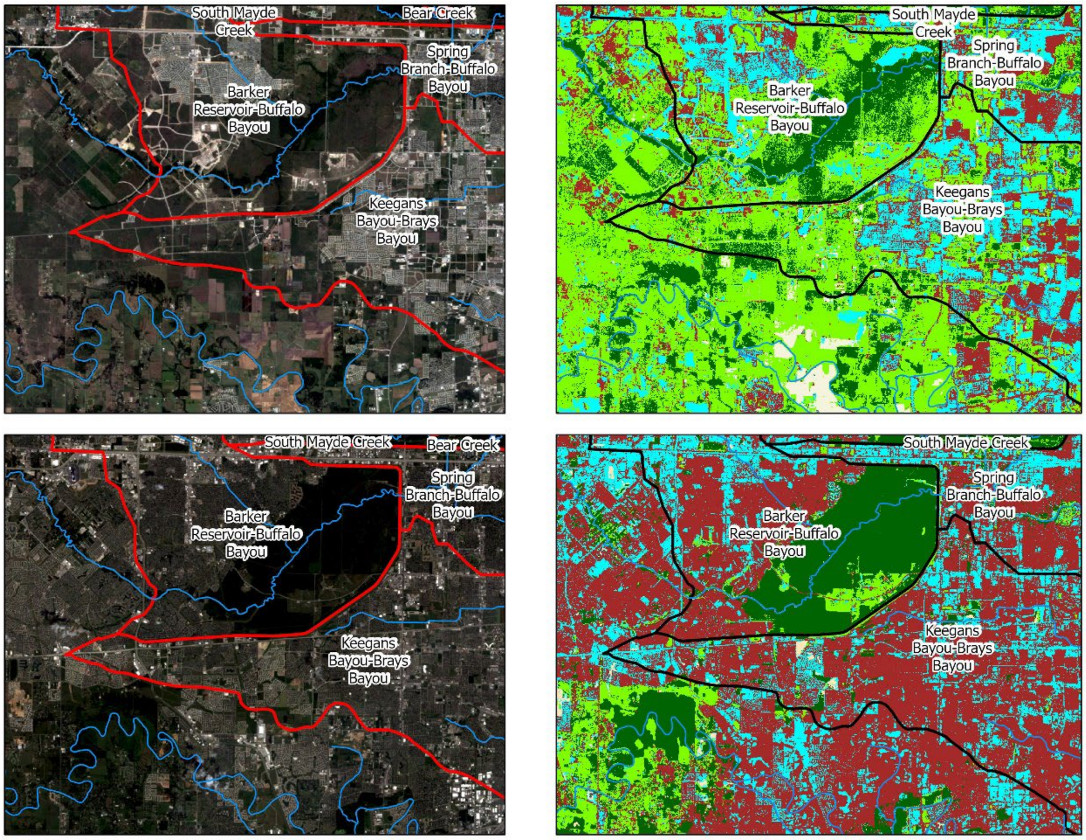

Figure 6. Landsat 5 (top) and Landsat 8 (bottom) Imagery highlighting urban expansion around the Barker Reservoir.

west of the Addicks-Barker Reservoir systems. These communities that have changed the landscape are primarily single-family homes for Houston commuters as seen in Figure 6. This growth is evident by the expansion starting along the major highways of Highway 290, Interstate 10, and Highway 69 which are the three highways visible heading northwest, west, and southwest, respectively. These routes allow for daily commuting to be possible and facilitated additional land use change over the last 40 years. This growth and expansion outwards from Houston into agricultural land and converting the land causes changes to the region's hydrology by decreasing infiltration and increasing runoff potential. The amount of infiltration is already reduced from gulf coast prairie levels though when the land was converted to crop land because of the change in long term flora and traditional agricultural practices removing the native, deep rooted, floral and disturbing the soil, changing the porosity and damaging the clay structure (Garner 2020). This shift from the native landscape to one centered around agriculture began to increase the runoff and decrease the soil infiltration capacity reducing the overall efficacy of the prairier ecosystem. This change to the region initially onset by agricultural practices is exacerbated by paving over these agricultural fields and shifting the land use to a residential land use with suburban, single-family homes and road networks, increasing impervious cover creates an increased flood hazard potential across the region. This increase comes from all the excess water that is not permeated through the soil, becoming overland flow more than what the region is capable of naturally mitigating. The current response to this has been to create and mitigate the storm water run off by the installation of storm drainage networks, detention and retention ponds, and utilizing the bayou network found in Harris County (Garner 2020, Xiaoyong 2020). These changes to the landscape inevitably increase peak discharge amounts and frequency during and after flood events as water that historically once infiltrated the soil is displaced into running off the surface into the bayou network system (Konrad nodate). This increased discharge has implications not only for low recurrence events such as 100-year or 500-year events, but even moreso with higher recurrence events such as 2-year events. According to Konrad from the USGS, peak discharge for a 2-year event is more likely to increase by 100 to 600% after urban development has occurred, which means there is a 50% chance of a region experiencing that peak discharge every year (Konrad nodate).

Another shift in the ecology is the loss of forested land to urban and suburban land uses. Forested land accounted for 34.26 km^2 or 8% of urban land shifts and 84 km^2 or 11% of suburban land to 2021. Some of this can be accounted for due to similarities in forested land and the tree creating a canopy in some older neighborhoods of Houston. This overstory creates a visual effect that is like an overstory of a forest canopy which causes some errors to the model mistaking what is there with visible pixel values. This can be corrected via utilizing ground truthing to verify or rectify what is on the ground compared to what is viewed by the computer model.

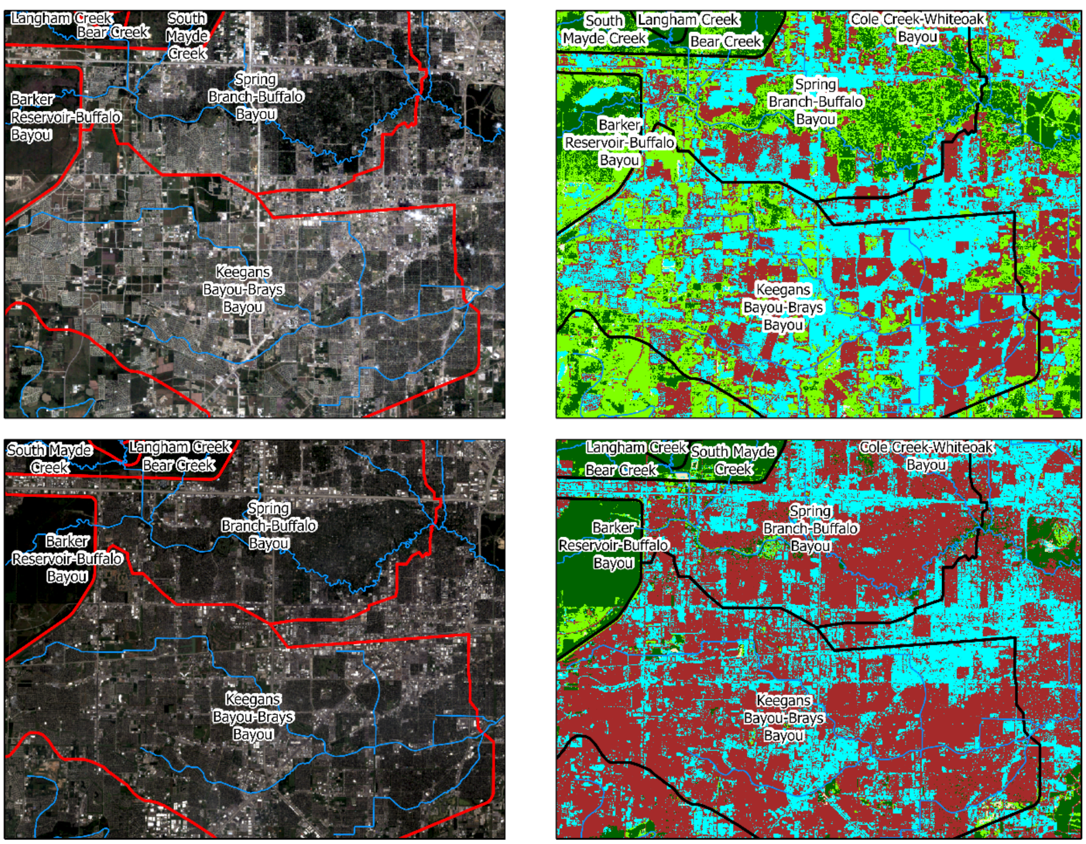

Figure 7. Landsat 5 (top) and Landsat 8 (bottom) Imagery showing neighborhoods in West Harris County with densely vegetated overstory.

Within the forest classification 49% did not change because it resides within the reservoirs which act as detention ponds to hold water during storm events and to manage the bayou system levels so that there is a decreased change of flooding in downtown Houston. Since Hurricane Harvey, the Harris County Flood Control District has undertaken multiple projects to expand channels, construct new detention ponds, conducting buy-out programs for homes in flood prone neighborhoods, and most recently discussing a $30 Billion Tunnel system that is proposed to go under Harris County and detain excess runoff stormwater to mitigate the flood potential during high rain events (Harris County Flood Control District, 2022). These actions work to mitigate the issue of increased urbanization within the county and to mitigate economic damage incurred during high intensity rain events that are common to the region. By performing these projects, the flood control district is mitigating an issue caused by the outward expansion and population growth experienced by the region. This can be seen as a byproduct from other studies such as where detention basins and improvements to current stormwater systems aid in the transporting or storing water to prevent excessive flooding. Urbanization also comes with some concern regarding the groundwater levels and particulate matter inside such as hazardous chemicals from vehicles, minerals from concrete that leach, and larger debris that is moved into the bayou system (Nath et. al 2021, Zope et al. 2017, Xiaoyong 2020).

6.2 Limitations and sources of error:

Some limitations of this study are the potential size of the study area and satellite imagery data used in the study. With the study area being so large, the pixel size of the 1989 data is 30m x 30m. This causes the creation of mixed pixels, which is a pixel that has a digital number representing the average energy reflected from multiple surfaces within that area represented by the single pixel (Jensen 2015). This causes some errors in the classifications that were described earlier, such as confusing more established neighborhoods with forested land. This confusion can be rectified by having ground truth data or knowledge of the site, but if that was not the case then misconceptions about the overall area would arise. Other methods to remedy this issue would be to utilize a fuzzy classifier which would use an algorithm like a maximum likelihood method to determine what the best fit for the mixed pixel would be or object based image analysis which looks at clusters of pixels to determine similarity between other clusters for classification.

6.3 Past Research

Previously, this study area was a topic of research for the author when creating python tools to automate the Normalized Difference Vegetation Index (NDVI) function, to automate an unsupervised classification of multiple raster images, and to perform a change detection analysis over the classified images in an iterative manner. In this study, the same data was used as a means of comparing the efficacy of the python scripts with a user completing the project by hand and testing different methods for a more hands-on approach. The results of that study indicated change occurred, but due to analysis and processing changes it is not quantitatively comparable to this study.

7. Conclusions and Future Directions

During the last 32 years the eight watersheds studied have shifted from a predominately herbaceous vegetation and forest land cover with an agricultural land use to a predominately suburban land use with a mix of urban and forested land with a small percentage being covered by herbaceous vegetation or bare soil. This shift is ultimately driven by the increasing population of new residents coming to the county to work within the Houston area metropolitan area in industries such as oil and gas, engineering, and medicine. The switch from an herbaceous centric aoi with 49% of the land cover in 1989 to suburban centric covering 46% of the aoi in 2021 is evidence that the boom in population is sprawling out into the available countryside and away from Houston's downtown.

These changes over the 32 year period have increased the impervious surfaces to the region by decreasing the percentages of herbaceous and forested land cover replacing them with suburban and urban land cover, thus decreasing the overall soil infiltration and shifting a percentage of that water towards runoff. This increase in runoff was initially seen in the region when the land use shifted from historical coastal prairie to agricultural, but with this further shift into a built environment the hydrologic cycle is impacted even more. This is evident from the increase in flooding the region experiences during high rain events such as hurricane Harvey and tropical storm Beta. Events like these have caused governmental agencies like HCFCD to increase their mitigation efforts utilizing methods like increasing the bayou network, dredging and increasing current bayou depths and widths, performing home buy-back programs in flood plains, and even proposing a new tunnel system to detain water during intense flood events to prevent overland flooding.

The last research question posited at the outset of this project with a sense of vagueness, it depends. There is a definite change to the land use and land cover of the region, and from the FEMA map in figure 8 there is a definite impact and flood hazard for the study area, but this was not quantified in a specific dollar or by amount of people impacted within this study. This study is a great bounding point to answer this research question though, utilizing modeling applications such as HEC-HMS, Google Earth Engine, or even an ESRI extension of ArcHydro to model and properly quantify the persons affected by an extreme weather event.

The first location that would be a good start to test the hydrologic modeling for these watersheds would be to examine the Keegans Bayou-Brays Bayou watershed and specifically the region around Bellaire, Texas. According to the FEMA Flood Insurance Risk Map (FIRM) data, Bellaire, and the surrounding incorporated Houston, Texas is within the 1% annual flood chance. This area experiences frequent roadway flooding during storm events and during extreme storm events like Hurricane Harvey experienced significant flooding. This would make the region a good candidate for beginning the hydrologic modeling at a smaller scale with the ability to upscale to other watersheds once the initial analysis is complete.

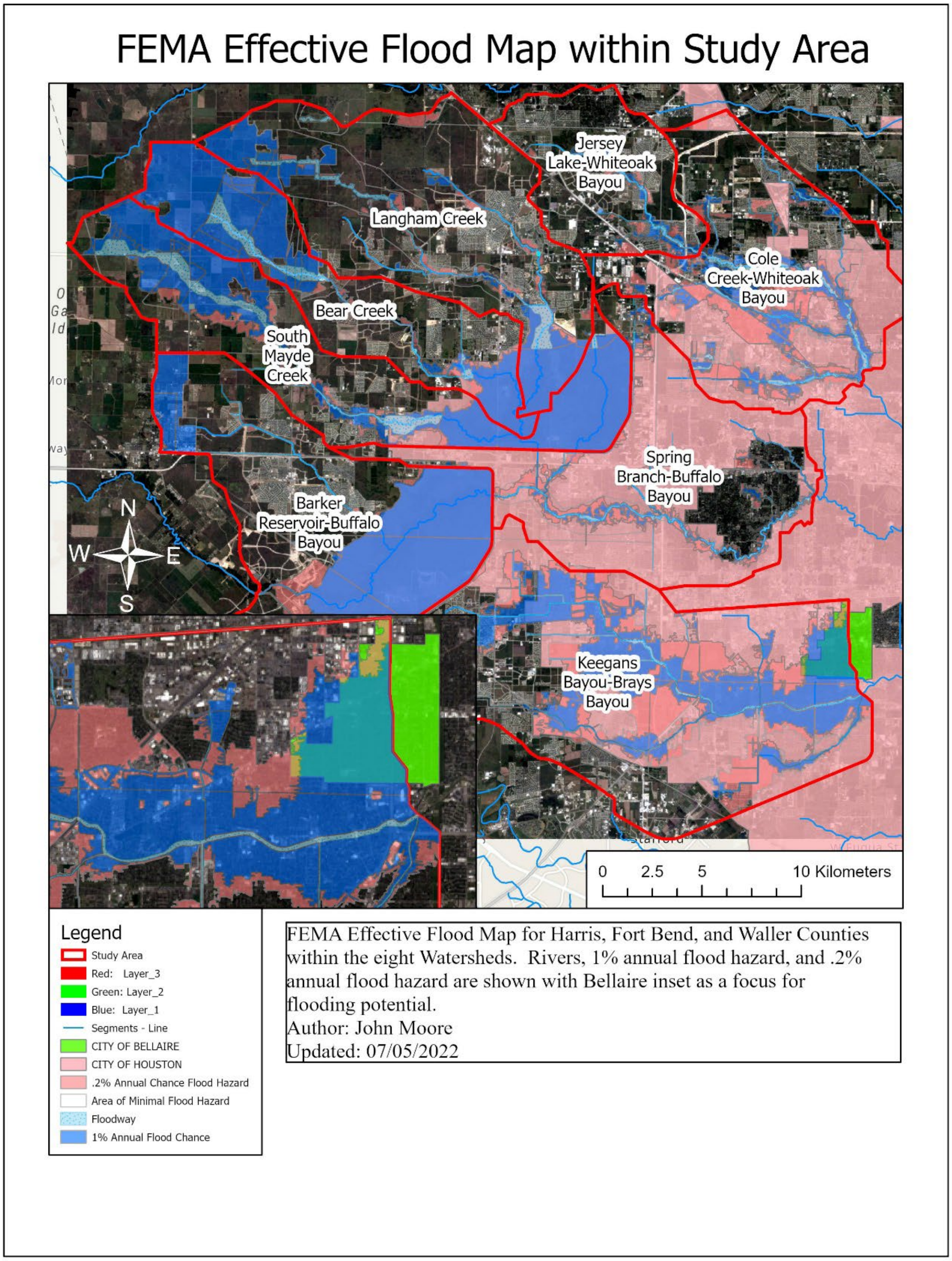

Figure 8. FEMA FIRM Mapping to assess need for Hydrologic modeling within the Study Area.

The next steps of this study will be to perform a hydrologic model over the eight watersheds in the study area to limit test the current bayou system in place and determine impacts of severe flood events. The current plan is to work on this within the HEC-HMS ecosystem, but Google Earth Engine is also being looked at as a potential application to use, because of the interconnection with online cloud datasets and customization capabilities. Another option is to use higher spatial resolution data to differentiate between more acute elevation changes within the AOI. This will help accuracy in watershed delineation and flow modeling.

Finally, once the hydrologic analysis is complete the project will be published online through ESRI's ArcGIS Online platform for users to view the changes to the region through the classified images and how it compares to current hydrologic datasets and flood models. This hub will provide the general public, local governments, and other interested geographers a unique opportunity to delve into the data and see some of the side effects of Houston's growth has had on the region.

Overall, the landscape has dramatically shifted over the past 32 years in Harris County, expanding the built environment outwards and encroaching on farmland to sustain the growing population of Houston and its neighboring suburbs. There is evidence to support the claim that an increase in the built environment has expanded into historically herbaceous vegetated land cover and shifted the land use to an Urban and Suburban classification. This work has answered both research questions, seeking out to quantify the LULC within the AOI and also to determine how this growth has affected the Katy Prairie and its ability to perform as a coastal prairie. These habitats provide a unique ecological role within Texas through flood mitigation and unique fauna habitat that is dangerously close to being lost to human intervention.

Works Cited

Abera, W., M. Assen, P. Satyal. 2018. Spatio-temporal land use/cover dynamics and its implication for sustainable land use in Wanka watershed, northwestern highlands of Ethiopia. Modeling Earth Systems and Environment. 5: 571-581. doi: 10.1007/s40808-018-0547-5.

Alexandres, C. G. 2014. Modeling Stormwater Sewer Systems using High Resolution Data. University of Texas at Austin.

Alturk, B., F. Konukcu. 2019. Modeling land use/land cover change and mapping morphological fragmentation of agricultural lands in Thrace Region/Turkey. Environment, Development and Sustainability (2020) 22:6379–6404. doi:10.1007/s10668-019-00485-3.

Anni, A. H., S. Cohen, S. Praskievicz. 2020. Sensitivity of Urban Flood Simulations to stormwater. Journal of Hydrology 588 (2020) 125028. Doi:10.1016/j.jhydrol.2020.125028

Blake, E. S., Zelinsky, D. A. 2017. National Hurrican Center Tropical Cycle Report: Hurricane Harvey. National Hurricane Center.

Climate.gov. “Climate Change: Global Sea Level”. (July. 05. 2022).

Construction Junkie. Watch this Permeable Concrete “Drink” 1000 Gallons of Water in a Minute. Date Accessed: June 2022).

Das, S., D. P. Angadi. 2021. Land use land cover change detection and monitoring of urban growth using remote sensing and GIS techniques: a micro-level study. GeoJournal. doi:10.1007/s10708-020-10359-1

Doyle, C. T, Beach. S, Luzzadder-Beach. 2021. Tropical Forest and Wetland Losses and the Role of Protected Areas in Northwestern Belize, Revealed from Landsat and Machine Learning. MDPI. Remote Sensing (2021). 13. 379. doi: 10.3390/rs13030379.

Dunne, T., Zhang, W., and Aubry, B. F. (1991). "Effects of Rainfall, Vegetation, and Microtopography on Infiltration and Runoff." Water Resources Research, 27(9), 2271–2285

Elaji, A., W. Ji. 2020. Urban Runoff Simulation: How Do Land Use/Cover Change Patterning and Geospatial Data Quality Impact Model Outcome?. MDPI. Water 2020, 12, 2715; doi:10.3390/w12102715.

Garner, M. 2020. Understanding the Impact of Changes to Coastal Prairie Landscapes on Watershed Response and Urban Flood Mitigation: A Case Study of the Katy Prairie in Houston, Texas. Rice University.

Gilmer, Bill. 2018. Proximity Counts: How Houston Dominates the Oil Industry. Forbes. Date Accessed: June 2022.

Guo, J. C.Y., 2017. Urban Flood Mitigation and Stormwater Management. Taylor & Francis Group. ISBN: 9781138198142

Hale, C., Gori, A., and Blackburn, J. (2014). Ecosystem Services of the Mid-Texas Coast. 57.

Han, D. M. J. Currell, G. Cao, B. Hall. 2017. Alterations to groundwater recharge due to anthropogenic landscape change. Journal of Hydrology 554 (2017) 545-557. Doi: 10.1016/j.jhydrol.2017.09.018.

Harris County Flood Control District. “Drainage Networks”. (July. 05. 2022).

Jensen, J. R. 2015. Introductory Digital Image Processing A Remote Sensing Perspective. Pearson. ISBN: 0-1340-5816-X

Konrad, C. P.. 2013. Effects of Urban Development on Floods. U.S. Geological Survey:Fact Sheet 076-03. (July. 05. 2022).

Minnig, M., Moeck, C., Radny, D., Schirmer, M., Impact of urbanization on groundwater recharge rates in Dübendorf, Switzerland, Journal of Hydrology (2017), doi: doi.org/10.1016/j.jhydrol. 2017.09.05.

Nath B., W. Ni-Meister., R. Choudhury. 2021. Impact of urbanization on land use and land cover change in Guwahati city, India and its implication on declining groundwater level. Groundwater for Sustainable Development 12 (2021) 100500. doi: https://doi.org/10.1016/j.gsd.2020.100500.

Nath B., Z. Wang., Y. Ge, K. Islam, R. P. Singh, Z. Niu. 2020. Land Use and Land Cover Change Modeling and Future Potential Landscape Risk Assessment Using Markov-CA Model and Analytical Hierarchy Process. ISPRS. International Journal of Geo-Information. 2020, 9, 134; doi:10.3390/ijgi9020134.

National Oceanic and Atmospheric Administration. Hydrometeorological Design Studies Center – Precipitation Frequency Data Server. (July.12.2022).

Navin, M. S., L. Agilandeeswari. 2020. Multispectral and hyperspectral images based land use / land cover change prediction analysis: an extensive review. Multimedia Tools and Applications (2020) 79:29751–29774. doi:10.1007/s11042-020-09531-z.

Noble, J.E., P.W. Bush, M. C. Kasmarek, and D.L. Barbie. 1996. Estimated depth to the water table and estimated rate of recharge in outcrops of the Chicot and Evangeline aquifers near Houston, Texas. Water Resource Investigation Report 96-4018. DOI: 10.3133/wri964018. (July. 05. 2022).

Paul S., K. G. Saxena, H Nagendra, N. Lele. 2021. Tracing land use and land cover change in peri-urban Delhi, India, over 1973–2017 period. Environ Monit Assess (2021) 193: 52. Doi: 10.1007/s10661-020-08841-x.

Sea Level Rise. (.n.d.). "Texas' Sea Level is Rising" (July. 12. 2022).

Seto KC, Fragkias M, Güneralp B, Reilly MK (2011) A Meta-Analysis of Global Urban Land Expansion. PLOS ONE 6(8): e23777. https://doi.org/10.1371/journal.pone.0023777

Stevens W.K., (1981) "Houston , Fastest Growing Big City, Showing Signs of Having Hit Prime". The New York Times. (July. 12. 2022).

Texas Parks and Wildlife. (n.d.). "Texas-Louisiana Coastal Prairie — Texas Parks & Wildlife Department." (June. 06. 2022).

Walker, R. K., and Miears, R. J. (1957). "The Coastal Prairie region in southwestern Louisiana and in southeastern Texas includes about 8 million acres. It almost parallels the coast of the Gulf of Mexico. A long, narrow band of coastal marshland lies be- tween the prairie and the gulf." 4.

Wanga, S.S. and Ndambuki, J.M. (2017) Accuracy Assessment of Land Use/Land Cover Classification Using Remote Sensing and GIS. *International Journal of Geosciences*, 8, 611-622.https://doi.org/10.4236/ijg.2017.84033

Wang, X., P. Cong, Y. Jin, X. Jia, J. Wang, Y. Han. 2021. Assessing the Effects of Land Cover Land Use Change on Precipitation Dynamics in Guangdong–Hong Kong–Macao Greater Bay Area from 2001 to 2019. Remote Sens. 2021, 13, 1135. https:// doi.org/10.3390/rs13061135.

Yirsaw, E., W. Wu, X. Shi, H. Temesgen, B. Bekele. 2017. Land Use/Land Cover Change Modeling and the Prediction of Subsequent Changes in Ecosystem Service Values in a Coastal Area of China, the Su-Xi-Chang Region. MDPI. Sustainability 2017, 9, 1204; doi:10.3390/su9071204.

Zope, P. E., T. I. Eldho, V. Jothiprakash. 2017. Nat Hazards (2017) 87:1267–1283 DOI: 10.1007/s11069-017-2816-4