# Mapping Past Land Use Changes and Modeling Future Urbanization Patterns in a Mid-Sized American City

A Thesis Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Duncan McCurry

May 2023

© by Duncan McCurry, 2023
All Rights Reserved.

# Acknowledgments

I want to thank my mentor and major advisor, Dr. Jie Zhuang. He has helped me every step of the way and given me all of the support I need to complete this project and to gain important academic and personal skills along the way. I would also like to thank the National Science Foundation (NSF) for funding this research through the UTK Food Energy Water for Sustainable Urban Systems (FEWSUS) program. Additionally, I appreciate all of the work that the FEWSUS program has done to allow me to expand my network and gain valuable experiences throughout my master's program. Next I want to thank my committee members: Dr. Mark Radosevich and Dr. David Hughes for providing help and advice along the way. Finally, I want to thank my parents, my brother, my girlfriend, and my dog for giving me all of the support and care that I could ever ask for.

# Abstract

On a global scale, humans have been moving from rural to urban areas at an increasingly fast rate in a process called urbanization. It is crucial to understand and predict the impacts of urbanization because city expansion encroaches on farmland and woodland, reducing regional ecosystem services. Therefore, urbanization must be well planned in a forecasting manner. The objective of this study is to map the land use change in Knox County, Tennessee, from 2001-2019 to identify the urbanization trends and understand the dynamic geospatial relationships between the urban and rural areas. The geospatial mapping was done using USGS land use data by classifying land into urban, agricultural, non-agricultural, and waterbody areas. Then, geographic information system (GIS) software was used to analyze the change of each land area in relation to infrastructure change. The results show a decrease in agricultural and non-agricultural land area and an increase in urban use. In addition, normalized difference vegetation index (NDVI) and elevation data, which were collected from the USGS Landsat database, were used to analyze the changes in vegetation, impervious surface area, and surface reflectance during the urban growth. The results were used to create a suitability index model from an environmental view in order to predict the areas in Knox County that are suitable for urban development. Finally, a machine learning model was implemented to compare areas identified as suitable with areas of predicted urbanization. These two models can enable urban planners to develop policies that geospatially optimize urban infrastructure while minimizing environmental footprints.

# Table of Contents


**1 Literature Review** **1**

1.1 Introduction . . . 1

1.1.1 Economic Development . . . 2

1.1.2 Demographic Changes . . . 3

1.1.3 Social Transformations . . . 4

1.1.4 Urban Space Modifications . . . 4

1.2 Land Use and Land Cover Changes . . . 5

1.2.1 Land Cover Classifications . . . 6

1.2.2 Past Land Use Change Research . . . 7

1.3 Environmental Impacts of Urbanization . . . 9

1.3.1 Urban Heat Island Effect . . . 9

1.3.2 Impervious Surface Area Buildup and Flood Risk . . . 11

1.3.3 Vegetation Loss . . . 12

1.4 Modeling Urbanization Patterns . . . 13

1.4.1 Identifying Driving Forces of Urbanization . . . 13

1.4.2 Applications of Predictive Modeling . . . 14

1.5 Purpose of Research . . . 15

1.5.1 Research Hypotheses . . . 16

**2 Mapping Land Use Changes** **17**

2.1 Introduction . . . 17

2.2 Overview of Study Area . . . 17

2.3 Materials and Methods . . . 20

2.4 Results . . . 21
2.5 Discussion . . . 42
2.5.1 Land Use Changes . . . 42
2.5.2 Drivers for Urbanization . . . 42
2.6 Conclusion . . . 43

**3 Environmental Footprint of Urbanization 44**
3.1 Introduction . . . 44
3.2 Materials and Methods . . . 45
3.2.1 Surface Temperatures . . . 45
3.2.2 Impervious Surfaces . . . 46
3.2.3 Vegetation Losses . . . 46
3.3 Results and Discussion . . . 49
3.3.1 Urban Heat Island Effect . . . 49
3.3.2 Impervious Surface Build-Up . . . 56
3.3.3 Vegetation Losses . . . 56
3.4 Conclusions . . . 68

**4 Modeling to Identify Suitability and Predict Development 69**
4.1 Introduction . . . 69
4.2 Materials and Methods . . . 70
4.2.1 Suitability Model . . . 70
4.2.2 Machine Learning Model . . . 74
4.2.3 Restricting Suitability . . . 76
4.3 Results . . . 76
4.3.1 Suitability . . . 76
4.3.2 Machine Learning Model . . . 82
4.4 Discussion and Conclusion . . . 91

**5 Conclusion 92**
5.1 Importance of Research Findings . . . 92

5.2 Future Research Directions . . . . . 92
Bibliography . . . . . 93

**Vita** **97**

# List of Tables


2.1 Area of Land Cover Classifications in Knox County (square miles) . . . . . 28
2.2 Land area for each classification of urban land intensity in pre-existing (up to 2001) urban areas . . . . . 34
2.3 Distance between new urban land and geospatial factors . . . . . 39
2.4 Area of new urban land developed on optimal, satisfactory, marginal, and unsatisfactory slope terrain. . . . . 41

3.1 Surface Temperatures (Fahrenheit) urban vs. non-urban land use types . . . 50
3.2 Surface Temperatures (Fahrenheit) in Various Areas in Knox County, May 2001 50
3.3 Surface Temperatures (Fahrenheit) in Various Areas in Knox County, June 2019 51
3.4 Impervious Surface Area . . . . . 63
3.5 Average NDVI change . . . . . 67

4.1 Standard of Evaluation for buffer rings. . . . . 72
4.2 Change in developed status from 2001-2019 . . . . . 83
4.3 Training and Test Data Cell Count . . . . . 86
4.4 Prediction Matrix Results . . . . . 87
4.5 Accuracy of First Prediction . . . . . 88
4.6 Prediction Matrix Results after Downsampling . . . . . 89
4.7 Accuracy of Training Data after Downsampling . . . . . 90

# List of Figures


2.1 Knox County, Tennessee . . . . . 19

2.2 Land Usage in 2001 in Knox County, Tennessee. Urban land is red, non-agricultural is green, cultivated and agricultural is yellow, water is blue. . . . 24

2.3 Land Usage in 2011 in Knox County, Tennessee. Urban land is red, non-agricultural is green, cultivated and agricultural is yellow, water is blue. . . . 25

2.4 Land Usage in 2019 in Knox County, Tennessee. Urban land is red, non-agricultural is green, cultivated and agricultural is yellow, water is blue. . . . 26

2.5 Rapid urbanization in West Knoxville near Fort Loudon Lake. . . . . 27

2.6 Change in total area for general land classifications from 2001-2019. . . . . 29

2.7 Urban area expansion from 2001-2019 . . . . . 30

2.8 Urban land density in 2001. Darker red indicates higher intensity. . . . . 31

2.9 Urban land density in 2011 with 2001 extent. Darker red indicates higher intensity. . . . . 32

2.10 Urban land density in 2019 with 2001 extent. Darker red indicates higher intensity. . . . . 33

2.11 This graph shows the change in urban land percentage from 2001-2019 for each census tract (labeled by its ID). The ten highest and ten lowest-urbanizing tracts will be used to compare and contrast throughout this thesis. . . . . 35

2.12 Ten slowest-urbanizing census tracts by change in percentage of urban area from 2001-2019. . . . . 36

2.13 Ten fastest-urbanizing census tracts by change in percentage of urban area from 2001-2019. . . . . 37

2.14 Median household income for the five fastest-urbanizing tracts (green) vs. the five slowest-urbanizing tracts (red). . . . . 38
2.15 Digital Elevation Model and slope map for Knox County, TN . . . . . 40

3.1 Flow chart describing surface temperature analysis methodology. . . . . 47
3.2 Normalized Difference Vegetation Index (NDVI) formula. . . . . 48
3.3 Temperature map for developed land use types May 30, 2001. . . . . 52
3.4 Temperature map for undeveloped land use types May 30, 2001. . . . . 53
3.5 Temperature map for developed land use types June 21, 2019. . . . . 54
3.6 Temperature map for undeveloped land use types June 21, 2019. . . . . 55
3.7 Impervious Surface Area in 2001. Darker color represents higher percentage of impervious surfaces represented by the pixel. . . . . 57
3.8 Impervious Surface Area in 2011. Darker color represents higher percentage of impervious surfaces represented by the pixel. . . . . 58
3.9 Impervious Surface Area in 2019. Darker color represents higher percentage of impervious surfaces represented by the pixel. . . . . 59
3.10 Impervious Surface Area for the fastest-urbanizing census tracts in 2001. Darker color represents higher percentage of impervious surfaces represented by the pixel. . . . . 60
3.11 Impervious Surface Area for the fastest-urbanizing census tracts in 2011. Darker color represents higher percentage of impervious surfaces represented by the pixel. . . . . 61
3.12 Impervious Surface Area for the fastest-urbanizing census tracts in 2019. Darker color represents higher percentage of impervious surfaces represented by the pixel. . . . . 62
3.13 Impervious surface area expressed as a percentage of the region's total area. 64
3.14 Change in NDVI value from 2001-2019. . . . . 65
3.15 Change in NDVI value from 2001-2019 for the fastest-growing tracts in the county. . . . . 66

4.1 General suitability formula for determining which areas in the county are most suitable for urban development. . . . . 71
4.2 Example map showing the multiple buffer rings for each point. Every white symbol shown represents one school, and the colored rings surrounding the point represent the different "grades" associated with proximity to school. . . 73
4.3 Standard Suitability Map for Knox County. Grey area represents existing urban land. . . . . 78
4.4 Suitability Map for West Knox County. Grey area represents existing urban land. . . . . 79
4.5 Overall Suitability without urban extent overlay. . . . . 80
4.6 Suitability Map for Knox County, restricted by high NDVI values. . . . . 81
4.7 Ranking of variable importance according to the ML model. . . . . 84
4.8 ML model prediction for development in Census Tract 57.09. . . . . 85

# Chapter 1

# Literature Review

## 1.1 Introduction

The world has been rapidly urbanizing in the 20th and 21st centuries. According to the Population Division of the United Nations (UN), in 1930, around 33 percent of the global population lived in cities. By 2018, over 55 percent of people. inhabited cities worldwide. Humanity has passed the 50 percent threshold, indicating that urban areas are the main habitat for the human species. The UN projects that by 2050, this trend of rural to urban movement will continue with over two-thirds of the global population living in cities. Specifically, North America has the highest urbanized population of the world at 82 percent (UN.org). Many countries, like the United States (US), have mostly completed the beginning processes of urbanization. However, other countries are just now starting to experience rapid urbanization. Urbanization is critical to study and understand for this reason– in a way, urbanization research can provide opportunities for developing countries to learn from the unintended mistakes or consequences that urbanized areas have faced in the past. Before discussing the contents of this thesis project, it is necessary to first have a baseline understanding of urbanization as simultaenous processes and transformations and why this type of research can help to achieve greater synergy between the built environment and the natural environment.

At the most basic level, urbanization is the transition of a region from predominantly agricultural to non-agricultural and is often accompanied by an increase in population [13]. It

is common to see the agricultural sector of a rapidly-urbanizing area dwindle as the population flows away from the rural area. This can virtually eliminate the traditional agriculture industry and merge any remaining agricultural activities with the modern economy [18]. Because of this, urbanization is often accompanied by the merging of various industries into a larger connected network. Urbanization is an assortment of changes that happen within a city, and understanding each change is critical to understanding the dynamic nature of urbanization, which is the first step towards lessening the environmental footprint associated with city growth. These processes take place simultaneously and at varying rates depending on the characteristics and attributes of the urban area itself. Dr. Chaolin Gu, in "Urbanization: Processes and Driving Forces", compartmentalized the urbanization process into four main processes: economic development, demographic changes, social transformations, and urban space modifications [12].

### 1.1.1 Economic Development

Urban areas supply numerous economic and commercial benefits [20]. Urbanization contributes to economic growth through the accumulation of physical capital, knowledge capital, and human capital [17]. This economic growth then can further increase urbanization rates by encouraging population agglomeration in a mutually beneficial interaction [17]. Historically, economic factors drove people to cities, as distance from local markets affected income and transportation costs. This distance can be thought of as a cost-factor that can be quantified and calculated [11] as a way to model population increase in cities.

Generally speaking, the most urbanized countries tend to have higher per capita gross domestic product (GDP) [12], and globally there is a positive correlation between urbanization rate and GDP [13]. On average, per capita GDP of the most urbanized countries is over 40,000 USD, while the least-urbanized countries have per capita GDPs of less than 1000 USD [12]. These least-urbanizing countries are located in Sub-Saharan Africa and Southeast Asia, and typically havd the proportion of urban inhabitants below 25 percent. The most-urbanizing countries are located in North America, Western Europe, and Asia. This stark difference in relative wealth is indicative of the intimate relationship between urbanization and economic development. Gu iterates that "between 2006 and 2012, the 750 largest cities in the world

created 87.7 million new sector jobs, amounting to 58 percent of all new jobs in 129 countries" [12]. These jobs have replaced the traditional urban jobs of manufacturing and assembly line work, which are relegated to developing countries. Now urban areas host high-skilled work like law practices, advertising, insurance, finance, research and development, and consulting jobs [12].

### 1.1.2 Demographic Changes

Urbanization as a demographic change involves an increase in urban population numbers and an increase in population density [12]. The rate of movement into cities can outpace economic development, leading to urban poverty and the formation of urban slums [22]. Olalekan, in an urban slum study on Nigerian cities, notes that this rapid urbanization creates a more vulnerable population- increased unemployment, more exposure to environmental risks, disease outbreaks, poor sanitation, social exclusion, elevated crime rates, pollution, etc. [22]. Typically, these issues are most prevalent in cities in largely un-urbanized countries where rapid urbanization processes are in their early phases.

While some African countries are projected to experience rapid-urbanization (over 50 percent change) in the coming years, others are experiencing a gradual decline as the region enters the late phases of growth. An interesting demographic trend happens when regions become predominantly urbanized. France and Germany are examples of countries that are on the opposite end of the spectrum from Sub-Saharan Africa in terms of urbanization. Urbanization rates in these countries have slowed down considerably and have even seen negative urban growth, classifying the regions as super aging urban societies [12]. Aging urban countries include the US, China, Japan, South Korea, Singapore, France, and Germany to name a few. In these areas, total fertility rates among the urban population have dropped significantly and the proportion of the population older than 65 years has increased [12]. Simplified, total fertility rate is the average number of children a woman is expected to birth in her lifetime (World Health Organization). One of the most drastic examples of this trend is in China, where total fertility rate fell from 6.0 to 1.60 between the years of 1949 and 2016, and only increased to 1.62 after the implementation of the two-child policy in 2016. The

sharp decrease in this case is in part due to governmental policies like the One Child Policy, but this trend is evident in many aging urban societies.

### 1.1.3 Social Transformations

One trademark social issue related to urbanization is the gentrification process. Gentrification is "the conversion of socially marginal and working-class areas of the central city to middle-class residential use" [27]. Dr. Sharon Zukin argues that this trend has been prevalent in American cities since the 1960s. When looking at years of gentrification research, it appears that gentrification can be thought of as a "geographic reshuffling" where high-skilled workers relocate closer to economic centers at the expense of displacing members of the low income working class [27]. This has negative consequences including residential segregation and social exclusion as more racially and ethnically-diverse populations are driven out of these areas [12]. However, gentrification also leads to a restoration of historical areas [27] and a revitalization of urban centers [12]. Because gentrification is largely driven by economic development and corporate investment, the process of urbanization can affect the social and demographic make-up of the cityscape.

### 1.1.4 Urban Space Modifications

The economic, demographic, and social trends associated with urbanization are interconnected with the physical and geospatial patterns of urban land use. An interesting example of this is the trend for European cities to be socially and demographically divided into the east (poor) and west (wealthy) side of the city. This division is so widespread in Europe because of prevailing south-westerly winds, which blows coal smoke from the south-west areas of cities to the north-east [12]. The presence of coal smoke has human health consequences, so as a result wealthy high-skilled laborers tended to move into the west side of cities and displace low-income workers to the east. This is an example of the well-documented phenomenon known as environmental injustice, or environmental racism in some cases. While this relates to social transformations, it also shows that urbanizing cities experience a reshuffling of the population for geographic reasons.

In recent years, urban areas in some countries have seen the rise of edge cities, or industrial parks on the periphery of the city. The Chinese government is one of the largest contributors to this trend, allocating millions of acres to these types of geographical economy-boosting sectors [26]. In China, there are 1568 of these parks as of 2017. They occupy less than 0.1 percent of China's total land area but contribute to around 10 percent of China's total GDP [26]. Research done by Zheng et al. indicates that these benefits are maximized because of the high amount of available space surrounding some major Chinese cities [26]. However, in areas with more pre-existing urban structures this is less viable due to the costs associated with converting the existing urbanities into industrial parks.

The rise of industrial parks is one facet of a larger overall trend: the formation of mega-regions. A mega-region is "an extended network of metropolitan centers and their surrounding areas, crossing county and state lines, linked by integrated labor markets, land use systems and transportation and communication infrastructure" [23]. In the US, this region can include up to 50 cities which are interconnected in part through interstates and highways [12]. There are 40 such regions globally, with 12 of them located in the US [19]. Not only do these areas contribute to the increased area of economic influence that a city has, but also the overall geographic space and land use area that the city encroaches on. Cities are no longer bound by the limits around the urban centers, instead they can grow to encompass smaller peripheral towns that depend on them economically.

## 1.2 Land Use and Land Cover Changes

As mentioned in the urban space modification section (1.1.4), urbanization is not only a demographic and socioeconomic shift within a city but also a geospatial change in land use and land cover. Land cover refers to "the attributes of a part of the Earth's land surface and immediate subsurface, including biota, soil, topography, surface and groundwater, and human structures" [15]. Land use, on the other hand, is a term that classifies how the land is used by humans. An important facet of land use research is the idea of "intensification". Land use intensification is the process of making land area more productive and can be quantified by either the amount of production added or the amount of resources allotted to a given area

of land [15]. In part because of the phenomenon of intensification, land classifications are dynamic and can be changed by human activity.

There are many different organizations that have varying definitions for land cover classifications, but the research presented in this thesis utilizes only the National Land Cover Database classifications.

### 1.2.1 Land Cover Classifications

It is necessary to define each type of land use classification used by the USGS, as these classifications are used in the data analyzed for this thesis project. A full list of land use classifications and their definitions can be found on the Multi-Resolution Land Classifications Consortium website. Classifications relevant to this research are listed and defined below:

- Developed or Urban - Areas of intensive land use where much of the land is covered by structures.
  - Developed, open space - Areas with a mixture of some constructed materials, but mostly vegetation in the form of lawn grasses. Impervious surfaces account for less than 20 percent of total cover.
  - Developed, low intensity - Areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 20 percent to 49 percent percent of total cover.
  - Developed, medium intensity - Areas with a mixture of constructed materials and vegetation. Impervious surfaces account for 50 percent to 79 percent of the total cover.
  - Developed, high intensity - Highly developed areas where people reside or work in high numbers. Impervious surfaces account for 80 percent to 100 percent of the total cover.
- Barren land - areas of bedrock, desert pavement, scarps, talus, slides, volcanic material, glacial debris, sand dunes, strip mines, gravel pits and other accumulations of earthen material. Generally, vegetation accounts for less than 15 percent of total cover.

- Deciduous Forest - areas dominated by trees generally greater than 5 meters tall, and greater than 20 percent of total vegetation cover. More than 75 percent of the tree species shed foliage simultaneously in response to seasonal change.

- Grassland/Herbaceous - areas dominated by gramanoid or herbaceous vegetation, generally greater than 80 percent of total vegetation. These areas are not subject to intensive management such as tilling, but can be utilized for grazing.

- Pasture/Hay - areas of grasses, legumes, or grass-legume mixtures planted for livestock grazing or the production of seed or hay crops, typically on a perennial cycle. Pasture/hay vegetation accounts for greater than 20 percent of total vegetation.

- Cultivated Crops - areas used for the production of annual crops, such as corn, soybeans, vegetables, tobacco, and cotton, and also perennial woody crops such as orchards and vineyards. Crop vegetation accounts for greater than 20 percent of total vegetation. This class also includes all land being actively tilled.

### 1.2.2 Past Land Use Change Research

Past research has identified land cover and land use change mapping as a crucial tool to better understand the dynamic interactions between humans and the environment, while providing opportunities to inform monitoring, policy, and reporting [10]. Regarding policy, accurate land use maps allow for policymakers to manage natural resources and anthropogenic activities like deforestation through geospatial imaging and analysis. Recent advances in remote sensing technology have also increased accuracy of this data by providing real-time data showing land cover changes [6]. Land-use and land-cover changes are a focus of research because they affect many environmental and landscape attributes including the quality of water, land and air resources, ecosystem processes and function, and the climate system itself through greenhouse gas fluxes and surface albedo effects [15].

One recent study on urban expansion in the Abuja city-region in Nigeria used GIS and land use change analysis to investigate the city-region's recent urbanization. Results of this analysis showed that urban and barren land expansion was extensive, while land

classified as forest or areas of high vegetation were dramatically reduced [8]. The study identified a discrepancy between the region's urban city plan and observed land cover changes, indicating that urbanization was occurring un-checked and contrary to the plan set forth by the development committee [8]. This unregulated development caused degradation of ecologically important forest land as it affected areas not identified as sites of potential development by the planning committee. Unfortunately, this is likely not an isolated incident. Many developing countries have large informal settlements and slum communities, which are responsible for a majority of urban land expansion in Sub-Saharan Africa [14]. This case study highlights the ability of land cover analysis to accurately visualize urban growth and identify instances where actual urban expansion does not follow plans set by the development committees. Additionally, it showcases land cover analysis as a tool for policymakers to better address the expansion of urban peripheral settlements.

Another case study that exemplifies the great potential for land use research is a study of the mega-region of Manila, Philippines [21]. The area is projected to experience a rapid population increase in the coming years, so the researchers analyzed past urban land use patterns to project how the population increase will change the future footprint of the city. They found that the elevated resource demand of the growing city population caused urban and agricultural area to encroach on natural ecosystem land cover. Transportation infrastructure developments driven by this increasing demand caused ecological disruptions like habitat fragmentation. Some of this expansion affected the protected natural land in the region as well. The overall trend of land use transitions was forest land to vegetative land, vegetative land to agricultural land, and then agricultural land to urban land. These results indicate that land use change is not a single, concrete event but rather a gradual process that typically results in ecologically-important areas transitioning to urban uses.

This past research shows that land cover studies, remote sensing technology, and geographic information systems (GIS) can increase cities' understanding of the broad range of impacts that their expansion causes. These methods allow for the quantification of land cover changes as well as a historical catalog of past urbanization patterns.

## 1.3 Environmental Impacts of Urbanization

Despite the aforementioned social and economic benefits of urbanization, regions of urban growth are often associated with dire environmental consequences. This is because urban areas share an intimate link with the surrounding environment: through ecosystem services, rural agricultural production, urban resource use and emissions, and even affecting the local climate. Among the most significant interactions are the urban heat island effect, reduction of carbon storage potential, and increased risk of urban flooding. These will be the main focuses of this thesis in regards to quantifying the environmental footprint of the study area's urbanization. However, it is important to note that there are many more direct and indirect consequences of urbanization that are not accounted for in this work.

### 1.3.1 Urban Heat Island Effect

One of the most critical externalities of urbanization is the urban heat island effect. According to the United States Environmental Protection Agency (EPA), the urban heat island effect occurs when, "cities replace natural land cover with dense concentrations of pavement, buildings, and other surfaces that absorb and retain heat" (EPA.gov). The EPA notes that the urban heat island effect (UHI) leads to higher energy costs, intensified air pollution levels, and increased heat-related mortality and illnesses. It is paramount when studying urbanization to determine how city growth affects surface temperatures.

The UHI has been extensively studied in China, specifically in rapidly-urbanizing regions. One such study was conducted in the Zhejiang Province in Eastern China. The researchers were mainly interested in the negative environmental externalities, like extreme heat events and air pollution, both of which are associated with the region's urban growth. By using climate data, NDVI measurements, and GIS, the study determined that the UHI caused a significant increase in both the average number of hot days (days where the temperature was in the 95th percentile of temperature) and extreme heat intensity in urban areas when compared to rural areas [25]. This trend is even more evident when looking at the highest-urbanized areas, which experience the most extreme heat events [25]. The study was also one of the first to examine urban night-time temperatures with regards to urbanization at such a

scale in China. Night-time temperatures were found to remain unusually high after extreme day-time temperatures. The implications of these findings indicate that rapid urbanization can exacerbate heat stress for urban populations by extending the duration of the extreme heat event throughout the night. These findings support the idea that the UHI needs to be studied extensively in a variety of urbanization scenarios as its effects will only worsen as climate change brings about more intense weather events.

The severity of the urban heat island effect is largely predicated on the shape, density, and land use types of the city. Debbage and Shepherd explore geospatial patterns and city morphology that could cause the UHI to worsen in a 2015 article. The authors begin by highlighting the two main modes of urban growth: sprawl and densification [3]. Urban sprawl refers to the expansion of city limits into previously un-urbanized or peri-urban areas. Densification, on the other hand, relates to a city more efficiently using the existing urban land to accommodate an increase in population, commercial, or industrial usages [3]. Generally speaking, proponents of sustainable urban planning tend to support densification as it reduces land use conversions and ecological destruction.

However, the authors indicate that in regards to the UHI, both sprawl and densification have been shown to exacerbate urban surface temperatures. The study uses climate data and spatial metrics to classify urban areas from the top 50 most-populated US states and then analyzes urban spatial patterns and surface temperatures to determine which characteristics actually increase UHI. The results show that urban contiguity, or uninterrupted urban area, is the greatest indicator for the UHI [3]. This means that both sprawl and densification can intensify the UHI and the often-advocated densification strategy does not necessarily mitigate urban temperatures as densification typically causes an increase in urban contiguity. The discussion of these results also yields interesting revelations relevant to this thesis. The idea that urban contiguity is the driving factor for the UHI indicates that any other type of land usage found in cities would likely decrease contiguity and therefore UHI [3]. For example, Central Park in New York City decreases the urban heat island effect in the area because the natural cover does not absorb and retain as much heat as building materials. The evapotranspiration via plant vegetation also reduces air temperatures. While this park

reduces the UHI, it only slightly reduces contiguity as it is only a relatively small area in an otherwise massive city.

Therefore, the results of this study could be applied to this area by creating many smaller parks spread throughout the city. While each park in this case does not cool the air as much as Central Park, they work in conjunction to reduce urban contiguity and may prove to be a successful UHI mitigation strategy for urban planners.

### 1.3.2 Impervious Surface Area Buildup and Flood Risk

Not only does urbanization affect surface temperatures, it also has a large effect on the hydrology of the area. When urban land use increases in area, the area covered by impervious surface material (pavement and building materials) increases as well. This diminishes the ability for soil in the area to infiltrate rainfall during and after a storm event because the soil is covered by anthropogenic materials, which in turn causes more stormwater runoff and urban flooding [9].

One important aspect of urban flooding is the present and impending effects of global climate change. According to the 2018 IPCC Report, there has been a global trend that the proportion of total rainfall from heavy rain events has increased over the last 50 years and this is likely to continue to worsen.

A 2015 study investigated the effects of increased impervious surface area (ISA) on flood risk at Longhua Bay, China. The area has seen rapid urbanization with more than 63 percent of natural land converted to urban [5]. The authors of the study found that this conversion increased the flood volume by 162 percent and peak discharge from precipitation events by 140 percent, which consequentially intensified the urban flood risk. The mechanism for this risk increase is driven by land use change and the relative capacity of natural and agricultural land to intercept, store, and transport rainfall when compared to that of impervious surfaces. Interestingly, the risk is also controlled by the location of the ISA. A risk index was created to compare upstream and downstream sub-basins, or small watersheds. The index value of 0.43 percent for downstream sub basins was much smaller than the value of 5.91 percent for upstream basins, indicating that ISA buildup for upstream regions led to a magnified flood risk. Lastly, the authors also mentioned that green roofs and permeable pavement can work

to mitigate flood risk by incorporating characteristics of natural land cover and intercept or slow runoff of precipitation [5].

In the same vein as the urban heat island effect, the debate of sprawl vs. densification has implications for flood risk. This is explored in a 2011 study in Britain. The primary purpose of the study was to understand how Britain's projected population increase and changes in urbanization will affect flood risk and carbon storage [7]. The results from this study indicate that densification of the urbanities will result in 1.7 million people living in areas that will have a 10 percent elevated peak flows [7]. This is caused by both a reduction of natural land cover and a population increase in the same region. Conversely, the sprawl scenario resulted in no significant change in flood risk. These findings are relevant to mitigation strategies and urban planning, especially because densification tends to be recommended to combat a host of environmental problems caused by urbanization.

While a review of past research can imbue a pessimistic outlook on city growth and flood risk, it is important to highlight countries that have employed strategies to combat this seemingly ubiquitous trend. The most notable example of this is China's Sponge City construction program. The program, which started in 2015, aims to turn cities in China into regions that can absorb, hold, reuse, and purify water as needed [16]. This is done by utilizing green infrastructure and low-impact development, while also managing surrounding land use types and working to reduce the urban heat island effect [16]. A study on China's exposure and vulnerability to urban flooding found that China's efforts to manipulate the urban space (like the Sponge City program) have actually led to a significant decrease in fatalities and socioeconomic vulnerability [4]. However, the authors note that this decrease is not the end of work on this front; as climate change brings about more intense and unpredictable precipitation events the need for flood mitigation strategies will continue in China and other quickly-urbanizing countries.

### 1.3.3 Vegetation Loss

Many studies have linked increasing urbanization and urban sprawl area to a decrease in carbon sequestration potential and an increase in carbon emissions. This is due to a multitude of factors: deforestation reducing the amount of carbon-storing vegetation, soil

degradation caused by land use intensification, and the tendency for urban peoples to consume a disproportionately high amount of non-renewable fossil fuels. Yan et al. (2018) attempted to use fine-resolution (30 meters) land imaging to calculate the effects of urbanization on net primary productivity (NPP) in the Pearl River Delta, China from 2001-2009 [24]. The authors use fine resolution imagery and compare the results to 500 meter imagery to understand the extent that large-resolution pixels can underestimate NPP reductions. The 30-m resolution NPP calculation indicates that around 350 Gg C was lost while the 500-m resolution only accounted for around 150 Gg C lost due to urban sprawl. The vegetation loss found with the 30-m resolution is equivalent to a reduction in absorption of 570.5 Gg CO2 [24]. This concerning reduction in carbon sequestration potential means that urbanization is actively intensifying anthropogenic greenhouse gas emissions and reducing the natural environment's ability to buffer this intensification. While this study was conducted in China, there is no reason that urban sprawl is not impeding carbon sequestration globally.

## 1.4 Modeling Urbanization Patterns

While it is useful to map past and current land use changes, research has also been done to extrapolate this data to predict future urbanization trends. Predictive models can be useful because they can evaluate the effects of different urbanization scenarios and help planners identify impending land use complications before they happen.

### 1.4.1 Identifying Driving Forces of Urbanization

In order to build a model for urban growth, it is necessary to first identify which characteristics of the city and the surrounding area influence the urbanization process. These characteristics are used as the inputs for the model, as they drive the geospatial expansion and distribution of urban land.

Researchers attempted to create a simulation of metropolitan growth using the city of Lahore, Pakistan as an example [1]. The inputs were identified from a literature review and then selected based on a quantitative equation to measure the association between the driving factor and the urbanization rate. The inputs for this simulation were categorized

into biophysical, infrastructure, and socioeconomic. Biophysical factors include terrain characteristics like slope, elevation, and distance to streams or canals. Infrastructure includes distance to roads, schools, existing urban land, hospitals, and percentage of households with accessible drinking water. Socioeconomic includes employment rate, population growth, and literacy rate.

Another example of identifying urbanization drivers for predictive modeling is a 2007 Master's Thesis that analyzes the same study area as this thesis project. Each factor's influence on urbanization rates was quantified using statistical analysis and ranked in order of its relative impact. The ranking of the most impactful variables is as follows (most impactful to least): Proximity to existing urban land, proximity to roads, proximity to urban node, proximity to existing commercial area, existing land cover type, and school rankings. Other factors, such as terrain slope, proximity to schools, and home values, were noted to impact urbanization but to a much less significant extent than the six ranked variables. Once again, these variables are not solely geospatial parameters or solely demographic characteristics, but instead are a conglomeration of all aspects of the urban area.

### 1.4.2 Applications of Predictive Modeling

Perhaps the most relevant example of how to apply predictive modeling techniques is a masters thesis examining land use changes in Knox county, TN [2]. The work served as the foundation for this thesis project, with the results presented in this thesis acting as a follow-up study to the previous thesis.

The masters thesis used 1992 and 2001 NLCD data to project and modeled how urbanization will affect the future footprint of Knoxville. The projections utilized both rule-based modeling and statistical models to extrapolate land use data and population data for future scenarios [2]. His results showed extensive urban growth in western Knox county and a slowing of growth in the central sector of the county due to the fact that it is largely urbanized already. Future research directions suggested mapping the destruction of natural land and the build up of impermeable surfaces in the county.

## 1.5 Purpose of Research

While there have been numerous studies addressing rapid urbanization, particularly in regions where growth is in the early phases, there is a research gap regarding smaller, moderately urbanizing cities that are in the later stages of development. This knowledge gap is crucial to understand, as the world urbanizes and more societies surpass the rapid urbanization stage and move into a slower form of urbanization like Knoxville. Assessing Knoxville's urbanization over the last two decades involves understanding changes in the city's urban footprint, changes in it's interactions with the environment, and then using this information to model the city's future growth. The questions to be addressed to fill this gap are as follows:

1. How has the proportion of urban, agricultural, and natural land use types changed during Knoxville's recent urbanization?

2. How does infrastructure and topography affect the geospatial pattern of developed land?

3. What environmental consequences are associated with urbanization in Knoxville?

4. How can knowledge of past urbanization and current environmental consequences inform urban and city planners to reduce Knoxville's footprint in the future?

The overall technical objectives of this project, each of which corresponds to a section of the thesis, are to:

1. Quantify Knoxville's recent urbanization through land use changes.

2. Determine the environmental consequences of urbanization.

3. Use land use information, geospatial factors, and environmental inputs to create a model predicting where urbanization will happen and scenarios for urbanization that reduce environmental harm.

### 1.5.1 Research Hypotheses

Based on the synthesis of past research, it is expected that: the proportion of urban land in the study area will increase, while agricultural and natural land area will decrease as a consequence.

Additionally, these land use changes will cause a reduction in NPP in the fastest-urbanizing areas, as measured through the study as a decrease in average NDVI values over the date range of the study. The urbanization also is expected to increase surface temperatures in Knox county over the same time period due to an increase in the area of heat-retaining surfaces. Overall, the urban growth is expected to worsen the county's environmental footprint and the vulnerability of the city to extreme weather events like heat waves and intense precipitation.

1. The proportion of urban land in the study area will increase, while agricultural and natural land area will decrease as a consequence. This urban land build-up, if it follows trends from the literature, will typically form in close proximity to major interstates, waterways, schools, hospitals, and the local topography.

2. These land use changes will cause a reduction in NPP in the fastest-urbanizing areas, as measured through the study as a decrease in average NDVI values over the date range of the study. The urbanization also is expected to increase surface temperatures in Knox county over the same time period due to an increase in the area of heat-retaining surfaces.

3. The areas with the highest probability of development are those areas that are near infrastructure, existing urban land, favorable terrain, and in tracts with higher housing prices.

# Chapter 2

# Mapping Land Use Changes

## 2.1 Introduction

This section is primarily concerned with analyzing changes in Knox county's land use and urbanization. This section will address the following research questions:

1. How has the proportion of urban, agricultural, and natural land use types changed during Knoxville's recent urbanization (2001-2019)?
2. How do infrastructure and topography affect the geospatial pattern of developed land?
3. As Knoxville urbanizes, is the city increasing urban sprawl, increasing in density, or both?

A review of literature has led to the hypotheses that urban land has likely expanded and encroached on natural land. In terms of geospatial factors driving urbanization, it is predicted that proximity to roadways and terrain slope dictate the spatial pattern of urban land.

## 2.2 Overview of Study Area

This study is focused on Knox County, Tennessee. A map of the state of Tennessee with Knox county can be found in Figure 2.1.

Knoxville, TN is the third largest city in Tennessee, preceded by Nashville and Memphis. The city has a population of 190,740 and Knox county as a whole has a population of 478,971 as of the 2020 census. Knoxville is located on the Tennessee River, just west of the Great Smoky Mountains and the Great Smoky Mountains National Park. In fact, Knoxville's geographical positioning and railroad development led to its growth as a manufacturing and distribution city as it was close to Nashville as well as the Appalachian Region.

Since 1940, Knox county has experienced strong, steady growth in population at a rate of around 2 percent per year (knoxplanning.org). Interestingly, the share of this growth from 2000-2010 was distributed so that only 10 percent of new residents were moving to the city of Knoxville while the rest moved to other areas in the county. However, from 2010-2017, this balance changed as 29 percent of new residents moved to the city of Knoxville and 71 percent moved to other areas in the county. This trend shows that more people are moving into the Knox county area now than they were in the past.

Projections indicate that the county population could increase to around 650,000 residents by 2075 and the county plus the surrounding region could experience population growth to 1.3 million (knoxplanning.org). These projections demonstrate a need for Knoxville as a city to understand the dynamics of this population growth, not only in terms of economic and social changes, but the interactions of the city and the surrounding areas.

In order to proactively plan for the projected urbanization in the city, it is paramount to first understand the trends and patterns of the city's past urbanization. One way to analyze urbanization patterns is through land use changes. Every land area can be classified as having one land use type, like "cultivated cropland" or "deciduous forest". These classifications change as land is transformed from one land use type into another. This thesis project first maps land use changes in the county over the years of 2001-2019. By analyzing the trends in land usage over the past 20 years in Knox county, it can be possible to make predictions on how these changes will continue in the future.

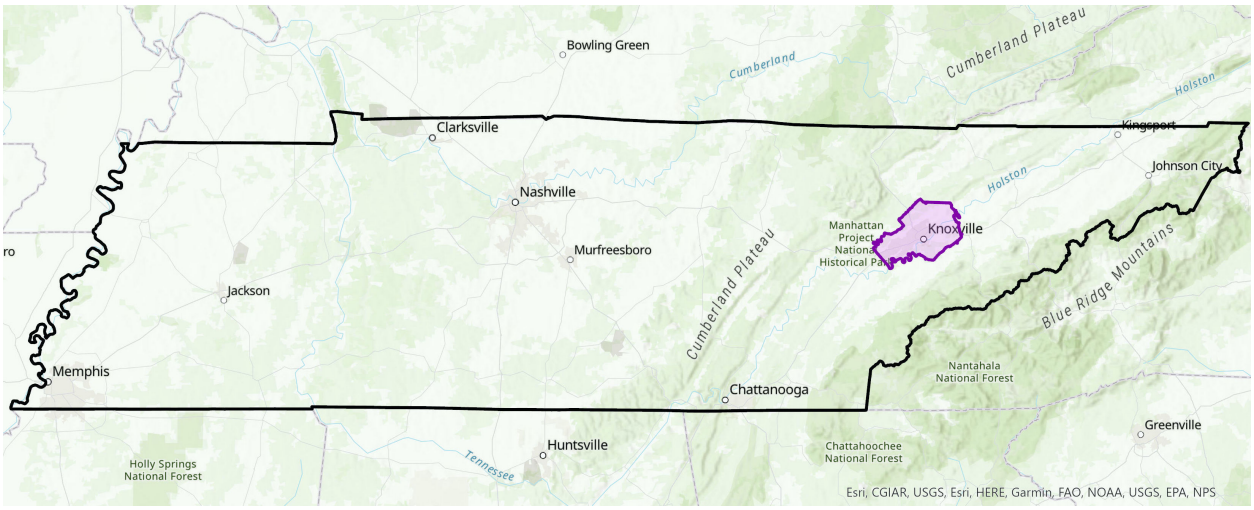


**Figure 2.1:** Knox County, Tennessee

## 2.3 Materials and Methods

In order to map land use changes in Knox county, land use data was sourced from the National Land Cover Database (NLCD), provided by United States Geological Survey (USGS). NLCD land cover data depicts high-resolution information, with one pixel equating to a 30 meter square area. Land cover data was used for the years 2001, 2011, and 2019. Land classifications were reclassified to simplify the process in some instances. Then, land use maps for the county were created for 2001, 2011, and 2019. These maps were analyzed using ArcGIS Pro to count the number of pixels for each land use type and then multiplying the count by 900 square meters then converting to square miles to calculate the total area.

Considering the impact of slope and terrain on urban development, a slope map was created. The digital elevation model (DEM) and slope map can be seen in Figure 2.15. Classifications for slope suitability were referenced from the US Department of Housing and Urban Development.

Because this study is primarily concerned with urban development, a map was also created to showcase the growth in urban land over the study time period. By using ArcGIS Pro's Extract by Mask tool, only areas that were classified as were included in the map. The base 2001 urban land use map was created first, then the map of urban growth from 2001-2011 was created by extracting the 2011 urban area from the 2001 urban area and only including the newly-urban areas. This process was done one more time to include the areas of urban growth from 2011 to 2019. After the maps were created, the "newly urban areas", or those areas that became urban after 2001 were clipped into a separate map. This new map layer was combined with the slope map in order to determine the proportion of new urban development that occurred on each slope classification.

Urban density data was used from the NLCD database as well to analyze the changes in built up land intensity. Each dataset (2001, 2011, and 2019) was clipped to the 2001 urban extent using the extract by mask geoprocessing tool in ArcGIS Pro. The reasoning for clipping each dataset to the 2001 extent is to compare the change in intensity over a uniform baseline area, which in this case is the pre-existing 2001 urban land.

Additionally, census tract boundaries were superimposed on the land use maps as a way to analyze smaller regions within the county. Census tracts were chosen because they can be used with census data. This is important because the maps not only show land use changes, but they can contextualize these changes through population and socioeconomic data. Census data was sourced from census.gov in order to obtain median household income data by census tract. Percent changes in urban land area from 2001-2019 were calculated for each census tract and the tracts were ranked by percent change in order to identify areas of both high and low urban growth. Throughout this study, the areas of high and low growth are compared and contrasted as small case studies within the study area to better understand which parameters cause urbanization rates to differ. These calculations were largely done in Microsoft Excel and then the Select by Attribute geoprocessing tool in ArcGIS Pro was used to create maps of each type of tract (fast or slow urbanization).

In order to assess which generally-accepted drivers for urbanization have the most influence in this area, the map of newly-urbanized areas was used. Each newly-urbanized area was converted from raster to polygon in ArcGIS Pro. A map of Tennessee major roads was taken from ESRI. A water body layer was created by converting the "Open water" land classification areas into polygons. The same method was used to convert the 2001 urban extent into polygons. After that was completed, the distance from each newly-urban area to the closest road, water body, and area of existing urban land was calculated using the ArcGIS Pro geoprocessing tool called "Near". Running statistics on these results allowed for the calculation of the mean nearest distance, minimum, maximum, and standard deviation for each geospatial parameter.

## 2.4 Results

The results of the analysis of the changes in land cover from 2001 to 2019 show that urban area has expanded at the expense of other land types. The changes in land use area proportions is shown in Table 2.1 and Figure 2.6. When examining the entire county, the land cover types that are most predominant are developed, deciduous forest, and hay/pasture. In terms of changes, developed land increased in area by 21.65 square miles over the study time period.

Additionally, barren land, shrub/scrub, herbaceous, and cultivated crops grew slightly in area at changes of 0.29, 1.61, 2.43, and 0.05 square miles respectively. Conversely, many land cover types experienced a reduction in area. Most notably, deciduous forests, evergreen forests, and hay/pasture shrunk by 8.88, 2.28, and 13.41 square miles respectively. The land use maps were created to identify areas of rapid growth in urban area. These maps can be seen in Figure 2.2, Figure 2.3, Figure 2.4, and Figure 2.5. A legend for these maps is located in the appendix of this thesis. The area that has seen the most urbanization during the study period appears to be western Knox County based on these maps.

Another application of the land cover dataset is analyzing the intensity of urban or developed land. There are four main classifications for urban land: "developed, open space", "developed, low intensity", "developed, medium intensity", and "developed, high intensity". The urban intensity maps were created for each study year and can be seen in Figure 2.8, Figure 2.9, and Figure 2.10. A breakdown of how these intensities have changed over time can be seen in Table 2.2. The table shows that a majority of the urban land in Knox County is classified as developed, open space or developed, low intensity. However, these are trending down in favor of medium and high intensity development which is indicated by the increasing area values for these two classifications over time.

Using the census tract boundaries, the urban land build up was calculated and sorted by highest to lowest change in urban land percentage from 2001-2019, as shown in Figure 2.11. The top ten fastest-urbanizing tracts are shown in Figure 2.13, while the slowest-urbanizing tracts are shown in Figure 2.12. Throughout this thesis, the top ten will be referred to as "Urbanizing tracts" while the bottom ten will be referred to as "Non-urbanizing tracts". While the term "non-urbanizing" is not exactly true, these tracts have seen negligible urban growth. It is also important to note that some of the tracts with no urban growth are already completely urban. These fully-urban tracts were excluded from the selection of non-urbanizing tracts because they have already completed the urbanization process.

Considering western Knox County is an area of rapid urbanization, it is unsurprising that a majority of the fastest-urbanizing tracts are located in this area. The slowest-urbanizing tracts seem to be on the periphery of the county. These tracts are large in area and have very little pre-existing urban infrastructure.

By linking census data with the land use maps, a graph was created comparing the fastest-urbanizing tracts with the slowest in terms of their median household income. This graph can be seen in Figure 2.14. The average median household income for the fast-urbanizing tracts was $103,815.20 USD in 2019 and just $47,390.60 USD for the slowest-urbanizing tracts.

The map layer for newly-urbanized areas in Knox county was also clipped to the slope map to understand the impact that terrain slope has on development. The analysis of land area for newly-urbanized areas on varying slopes shows that 9.29 square miles of urban land was built on "optimal" slopes (0-6%), 8.91 square miles on "satisfactory" slopes (6-12%), 3.93 square miles on "marginal" slopes (12-18%), and 2.24 square miles on "unsatisfactory" slopes (greater than 18%).

The results for assessing the relative proximity of each geospatial parameter (roadways, water bodies, and existing urban land) can be seen in Table 2.3. The mean distance from each area of recent urbanization to the various infrastructures shows that existing urban land was on average just under 14 meters away, roads were around 925 meters away, and water bodies were around 1150 meters away. Each factor was bordering the new urban land in at least on instance, as indicated by the minimum distances. Road ways had the largest maximum distance and standard deviation, which shows that there were areas of urbanization that did not occur near major roads. This is an interesting result, because proximity to major roadways is a driving factor for urbanization. Perhaps the data layer used in this project did not contain all of the major roads, only the most important highways and interstates. A map layer that also has smaller roads could help to increase the accuracy of this data, because it would show that the areas urbanizing seemingly far away from major roadways are actually still connected to the main city through smaller roads. Furthermore, the primary reason that urban proximity has such a small mean distance is because the urban extent is so large and widespread within the city. This means that every single point in Knox county is close to an urban area in some capacity. Contrast this with roadways and water, both of which are plentiful in the county but neither one covers as much land area as urban development.

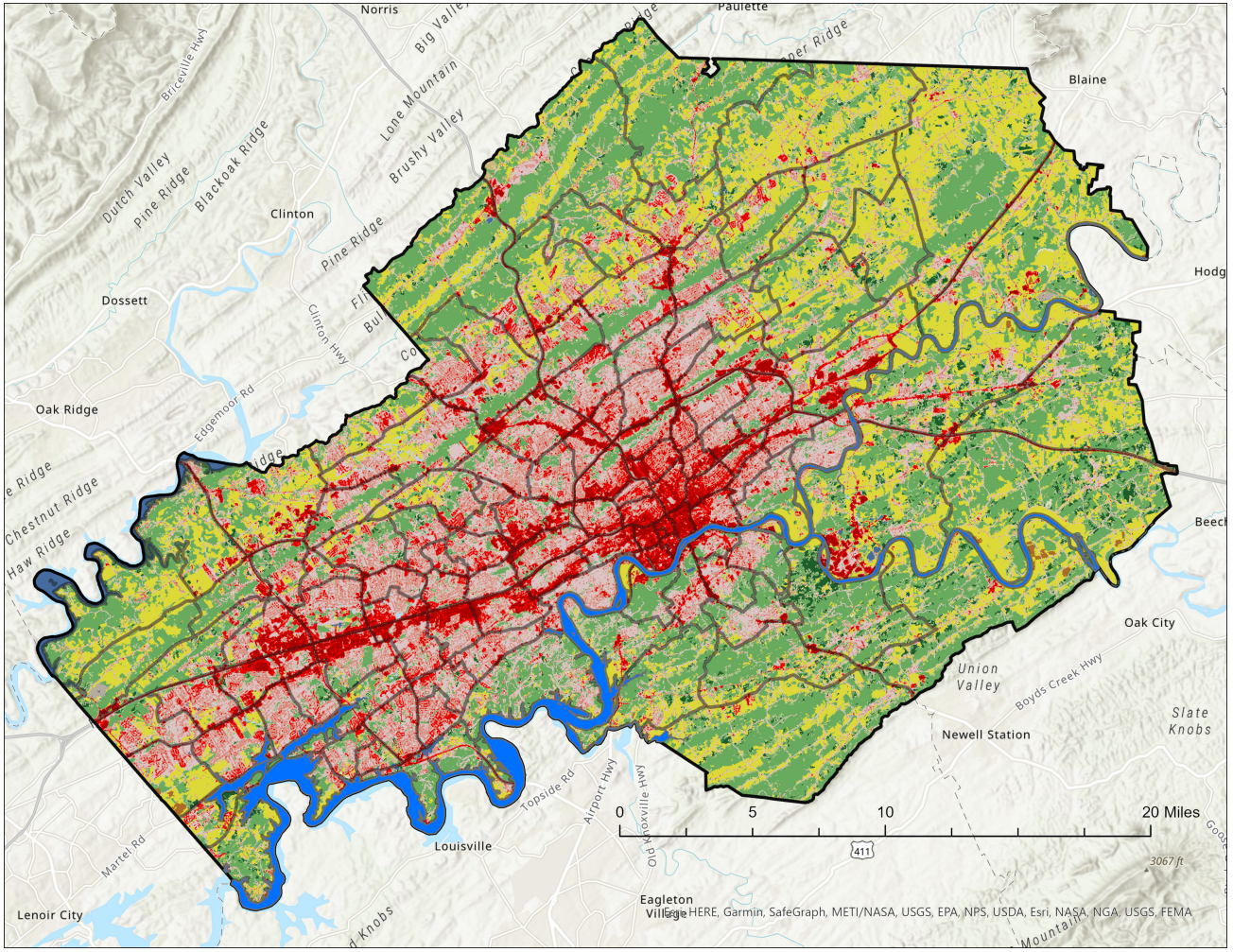


**Figure 2.2:** Land Usage in 2001 in Knox County, Tennessee. Urban land is red, non-agricultural is green, cultivated and agricultural is yellow, water is blue.

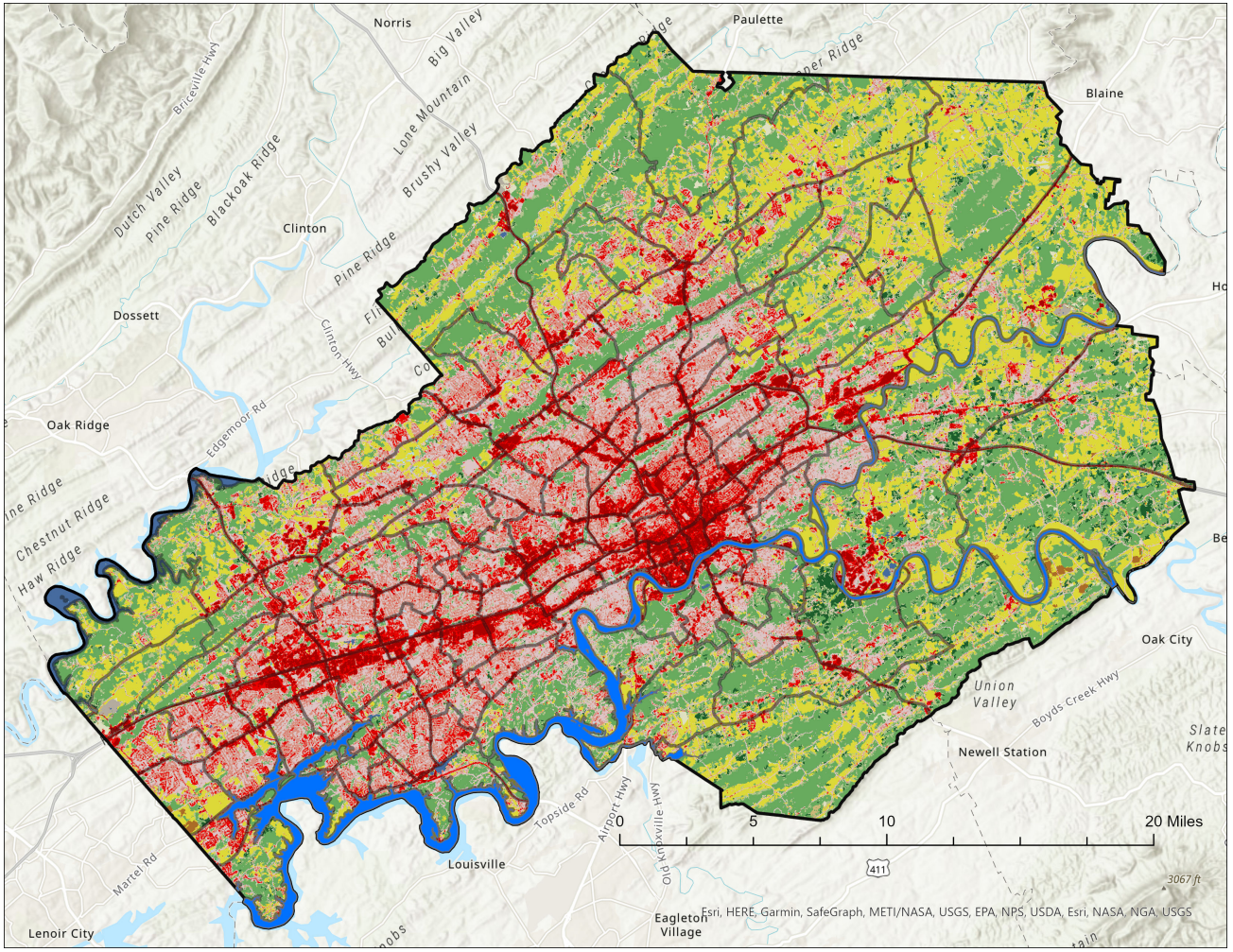


**Figure 2.3:** Land Usage in 2011 in Knox County, Tennessee. Urban land is red, non-agricultural is green, cultivated and agricultural is yellow, water is blue.

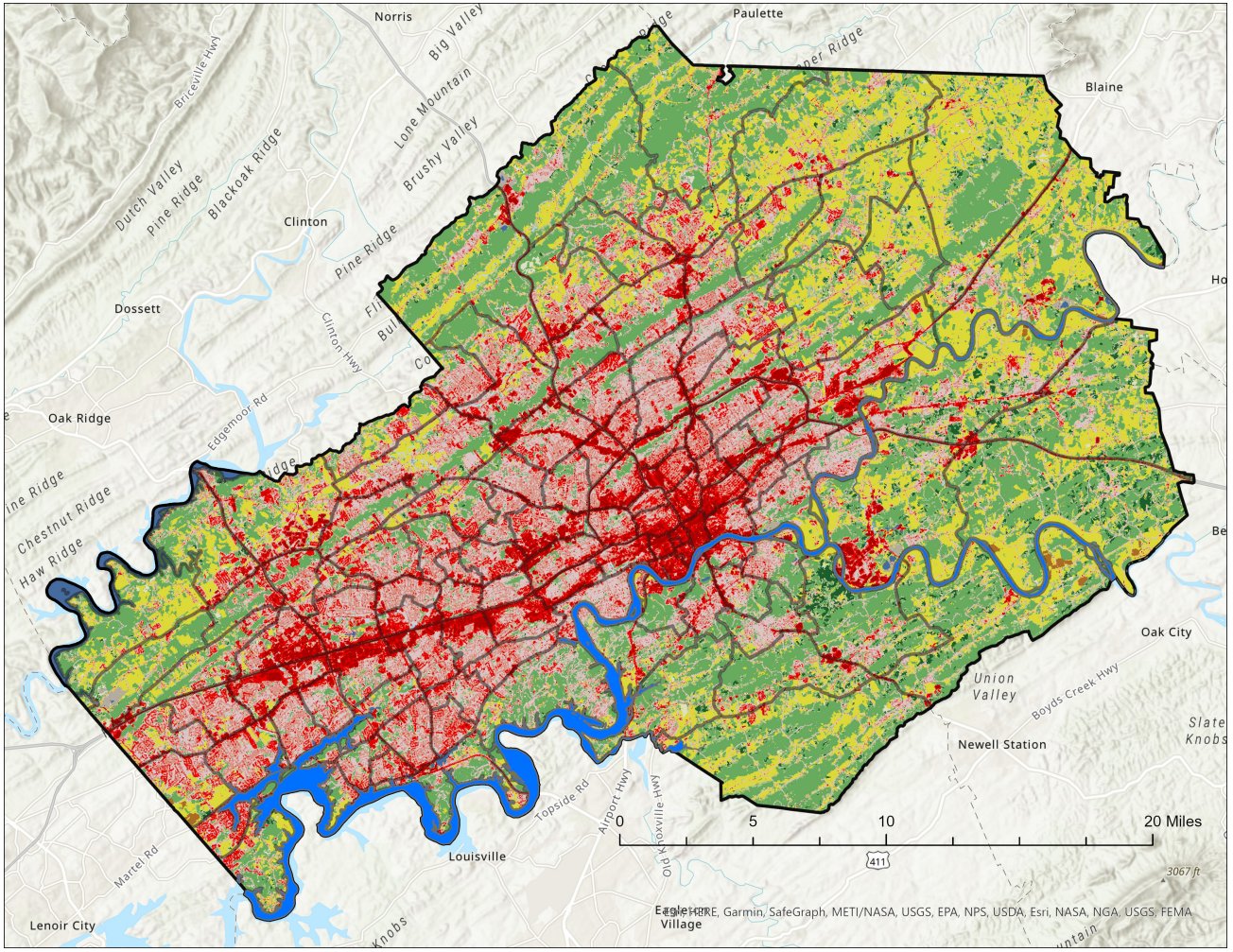


**Figure 2.4:** Land Usage in 2019 in Knox County, Tennessee. Urban land is red, non-agricultural is green, cultivated and agricultural is yellow, water is blue.

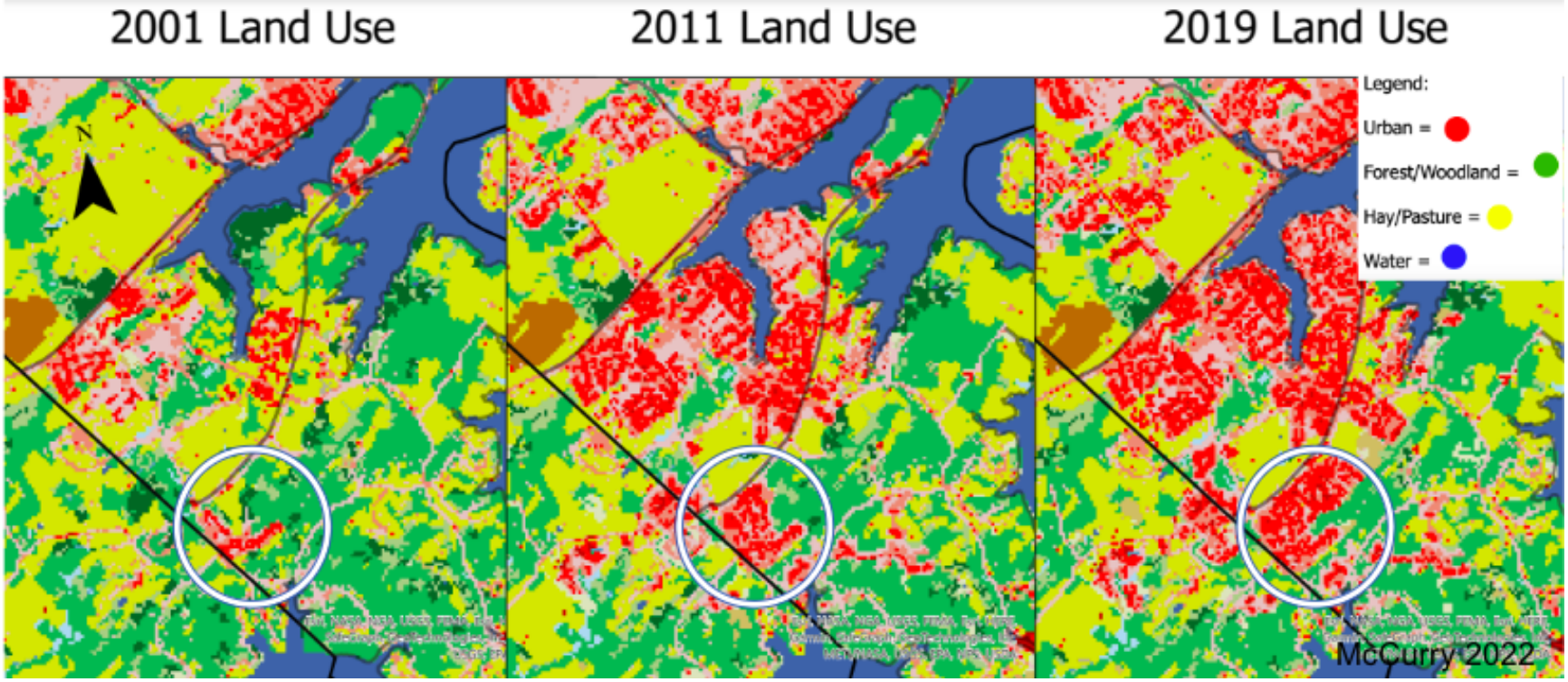


**Figure 2.5:** Rapid urbanization in West Knoxville near Fort Loudon Lake.

**Table 2.1:** Area of Land Cover Classifications in Knox County (square miles)

| Class | 2001 | 2011 | 2019 | Net Change |
|---|---|---|---|---|
| Open Water | 17.75 | 17.62 | 17.65 | -0.09 |
| Developed | 206.31 | 223.96 | 227.97 | +21.65 |
| Barren Land | 0.98 | 0.86 | 1.28 | +0.29 |
| Deciduous Forest | 143.35 | 135.51 | 134.48 | -8.88 |
| Evergreen Forest | 11.53 | 9.52 | 9.25 | -2.28 |
| Mixed Forest | 28.75 | 27.40 | 27.42 | -1.33 |
| Shrub/Scrub | 0.93 | 1.99 | 2.54 | +1.61 |
| Herbaceous | 1.63 | 5.02 | 4.06 | +2.43 |
| Hay/Pasture | 112.39 | 101.77 | 98.98 | -13.41 |
| Cultivated Crops | 0.68 | 0.70 | 0.73 | +.05 |
| Woody Wetlands | 1.41 | 1.39 | 1.39 | -.02 |
| Emergent Herbaceous | 0.13 | 0.12 | 0.12 | -.01 |

*Changes in land area of each NLCD land cover type is shown above for the years of 2001, 2011, and 2019 in Knox County. The total change in area represents the difference in land area from 2001 to 2019.*

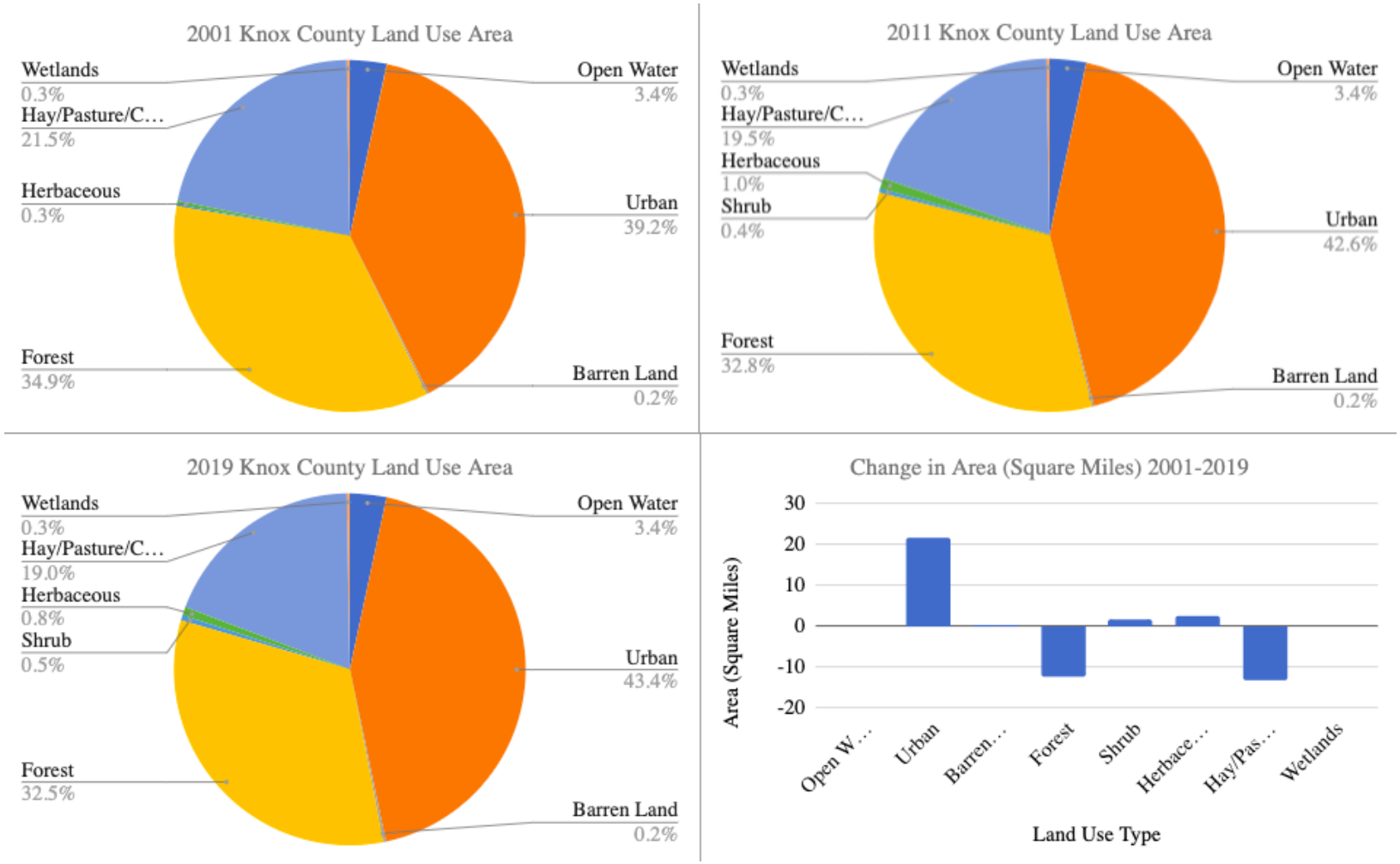


**Figure 2.6:** Change in total area for general land classifications from 2001-2019.

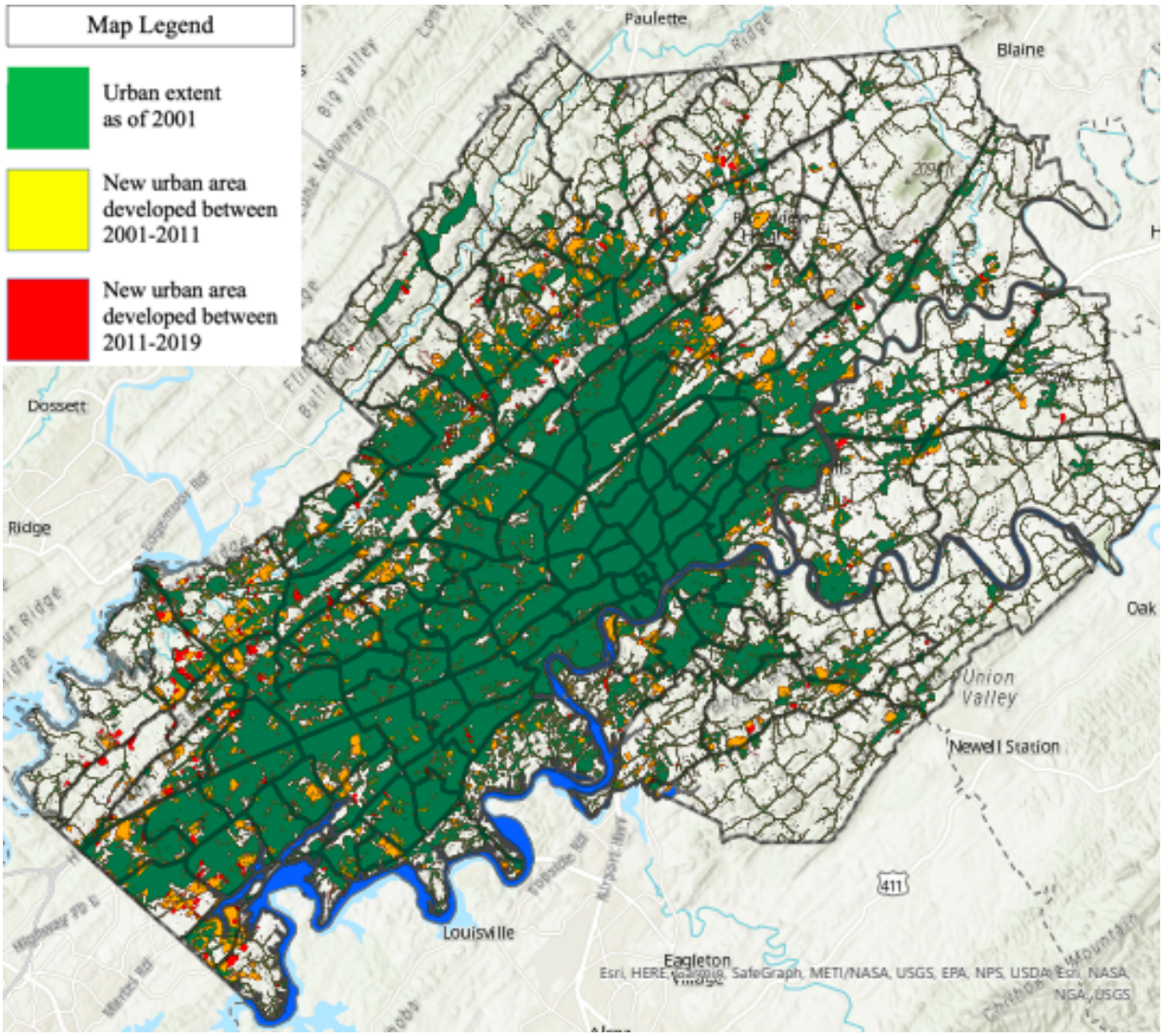


**Figure 2.7:** Urban area expansion from 2001-2019

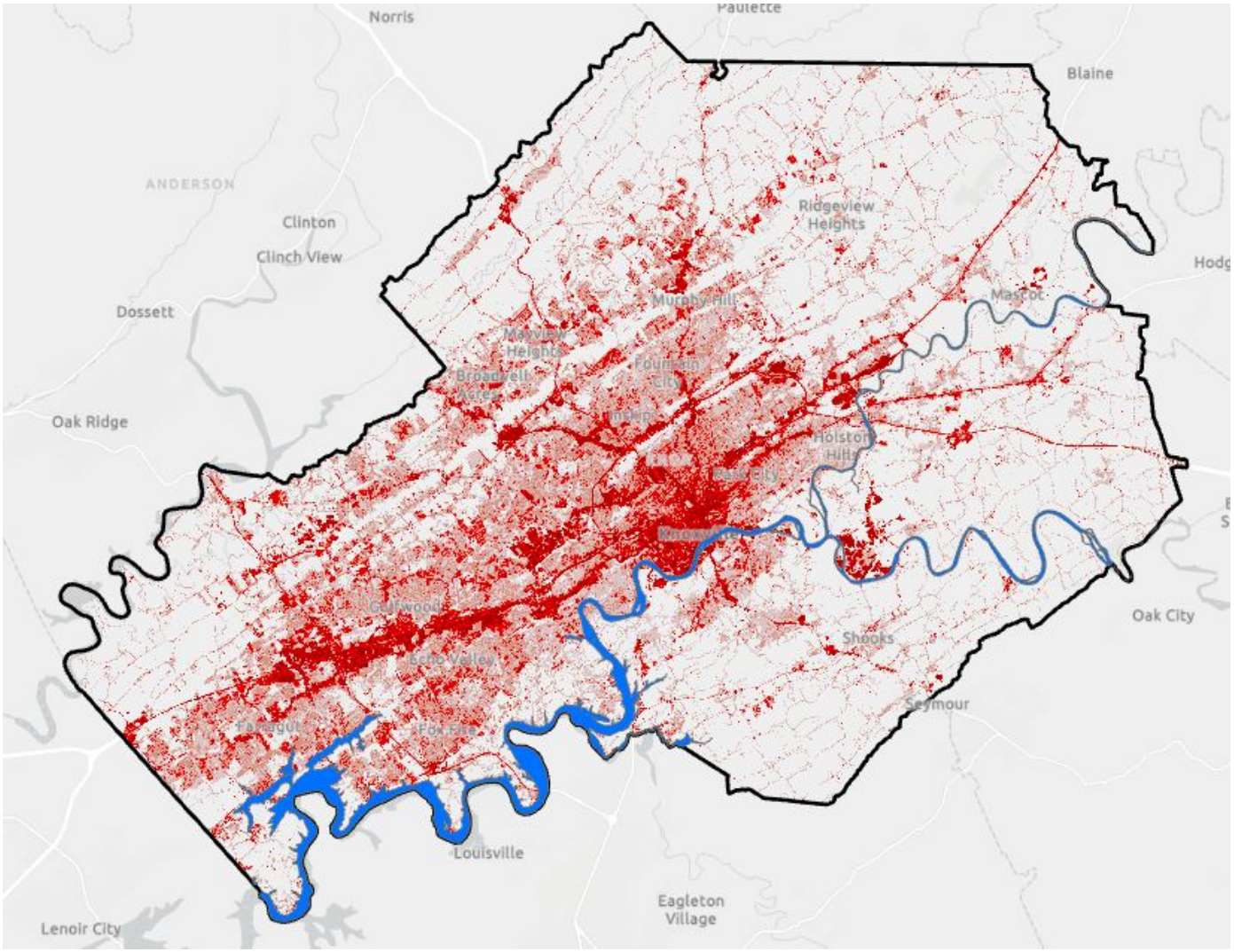


**Figure 2.8:** Urban land density in 2001. Darker red indicates higher intensity.

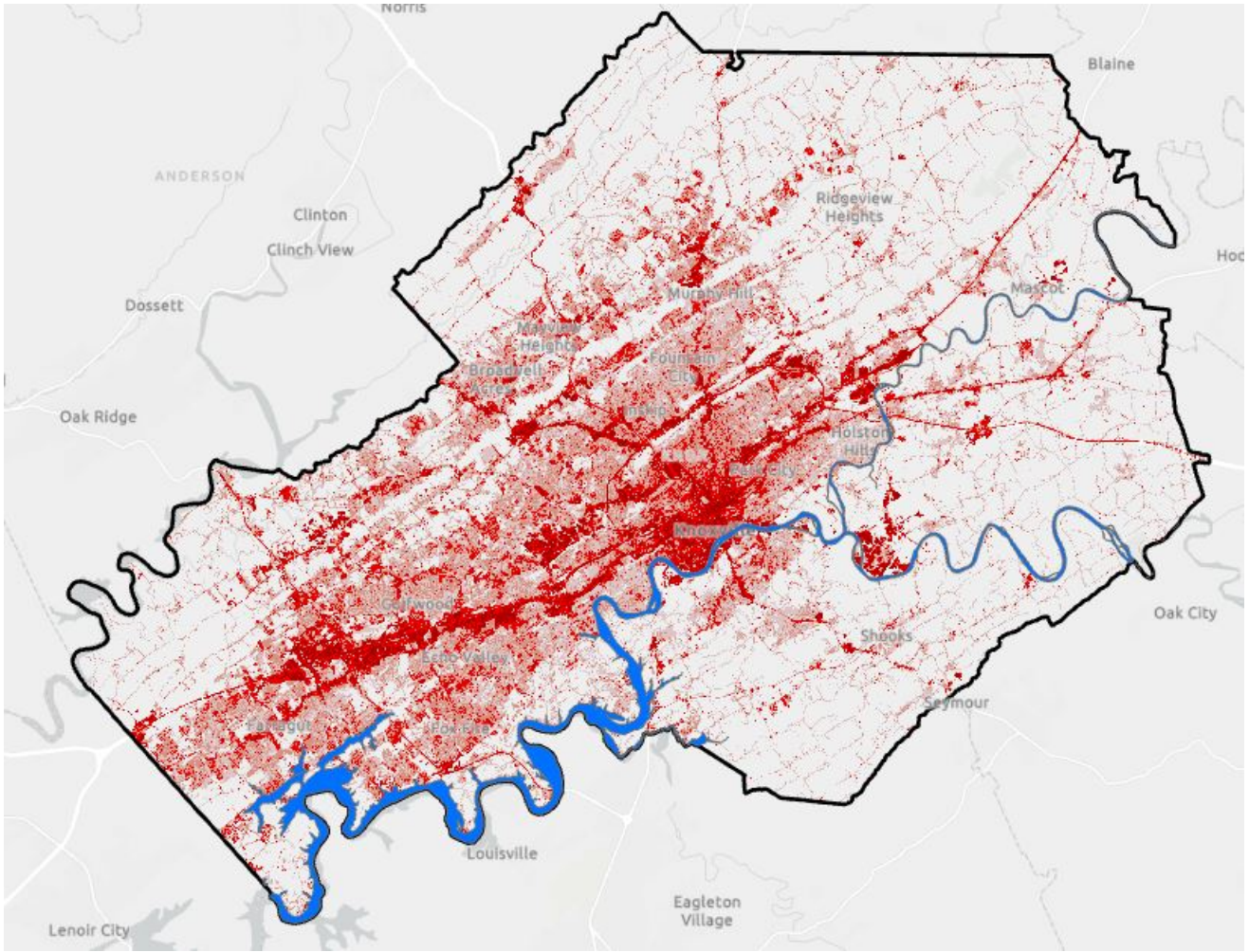


**Figure 2.9:** Urban land density in 2011 with 2001 extent. Darker red indicates higher intensity.

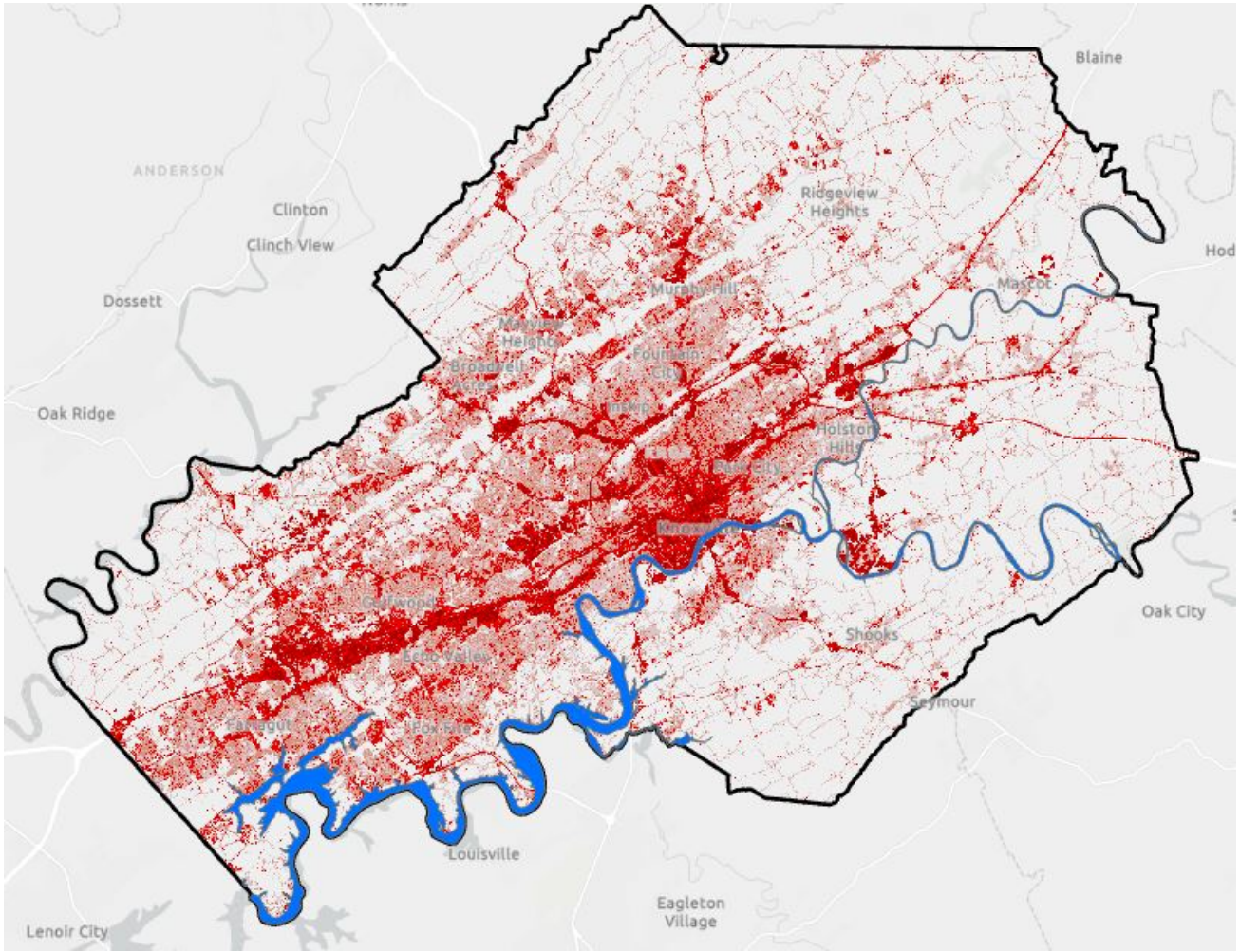


**Figure 2.10:** Urban land density in 2019 with 2001 extent. Darker red indicates higher intensity.

**Table 2.2:** Land area for each classification of urban land intensity in pre-existing (up to 2001) urban areas

| | **2001** | **2011** | **2019** |
|---|---|---|---|
| Developed, open space | 84.53 | 81.02 | 78.14 |
| Developed, low intensity | 70.79 | 70.11 | 69.55 |
| Developed, medium intensity | 36.28 | 39.47 | 42.27 |
| Developed, high intensity | 14.58 | 15.57 | 16.22 |

*This table shows the land area (square miles) of each type of developed land classification.*

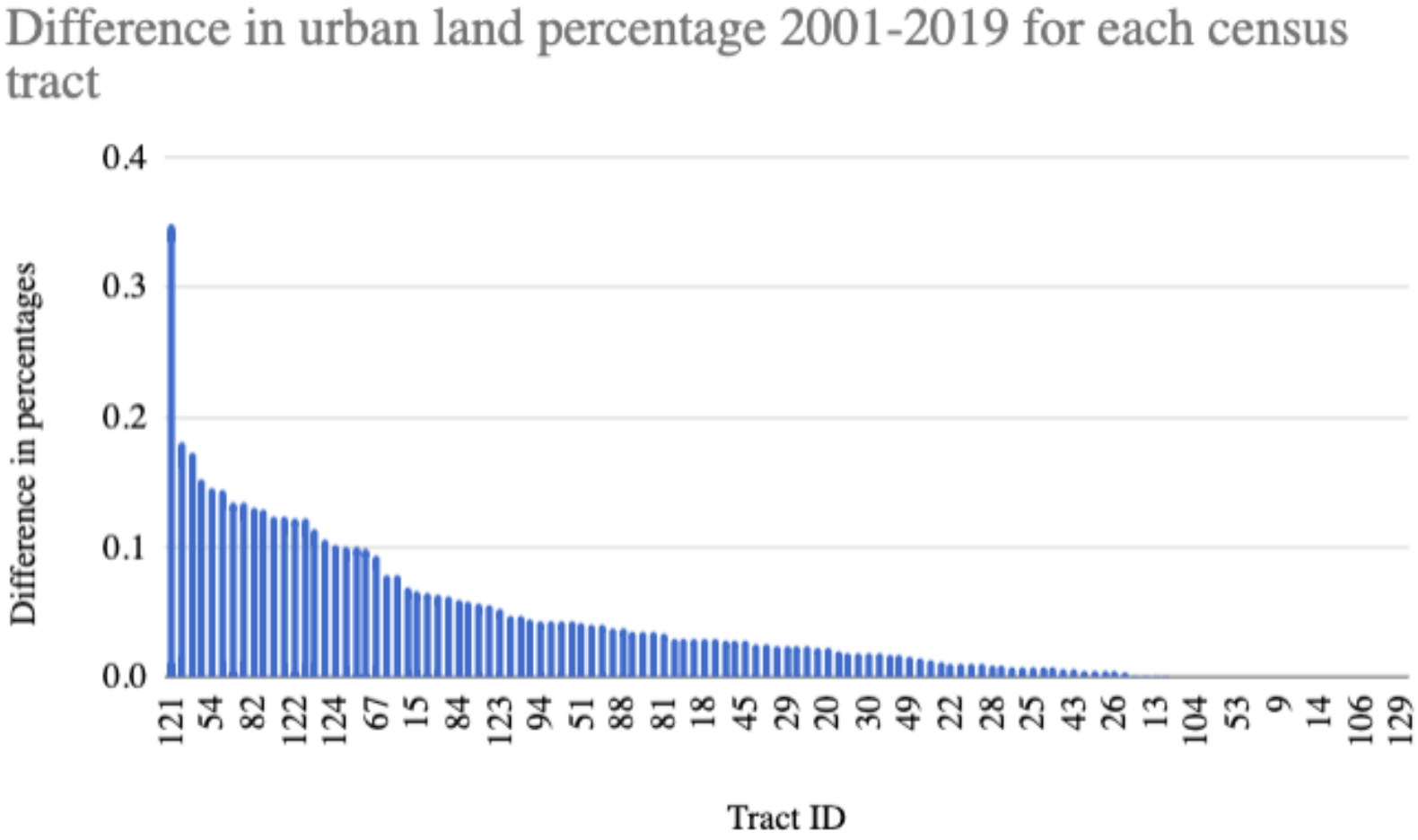


**Figure 2.11:** This graph shows the change in urban land percentage from 2001-2019 for each census tract (labeled by its ID). The ten highest and ten lowest-urbanizing tracts will be used to compare and contrast throughout this thesis.

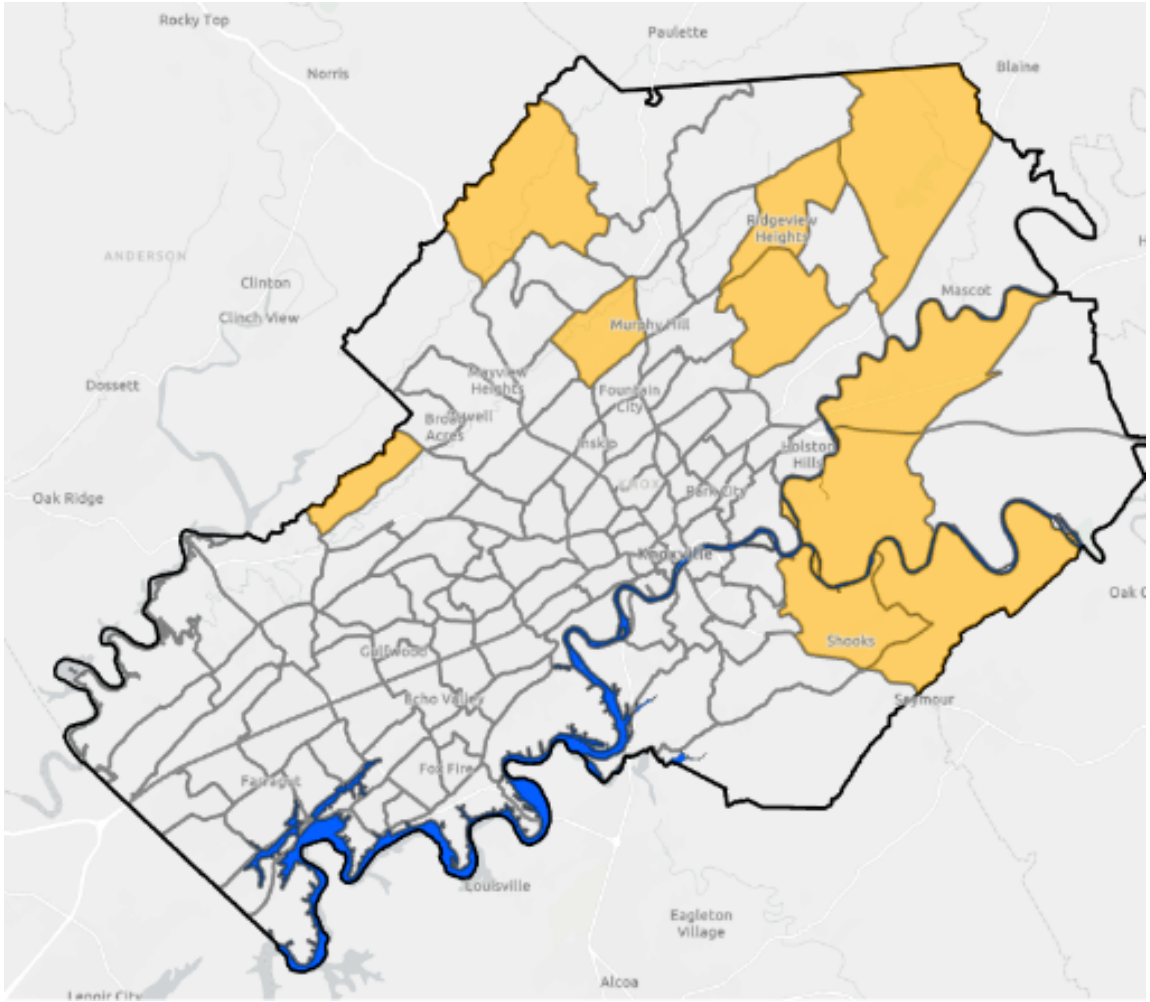


**Figure 2.12:** Ten slowest-urbanizing census tracts by change in percentage of urban area from 2001-2019.

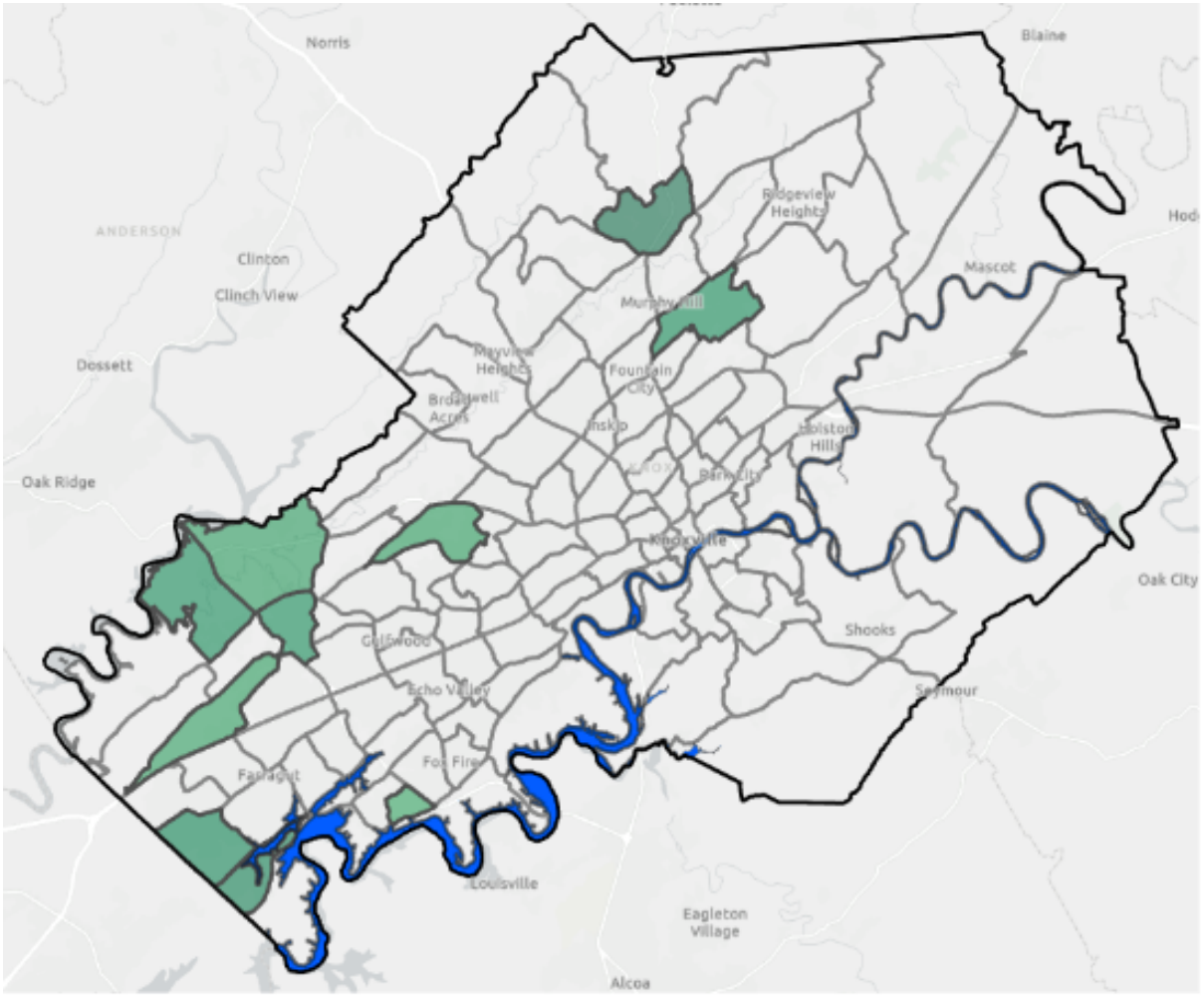


**Figure 2.13:** Ten fastest-urbanizing census tracts by change in percentage of urban area from 2001-2019.

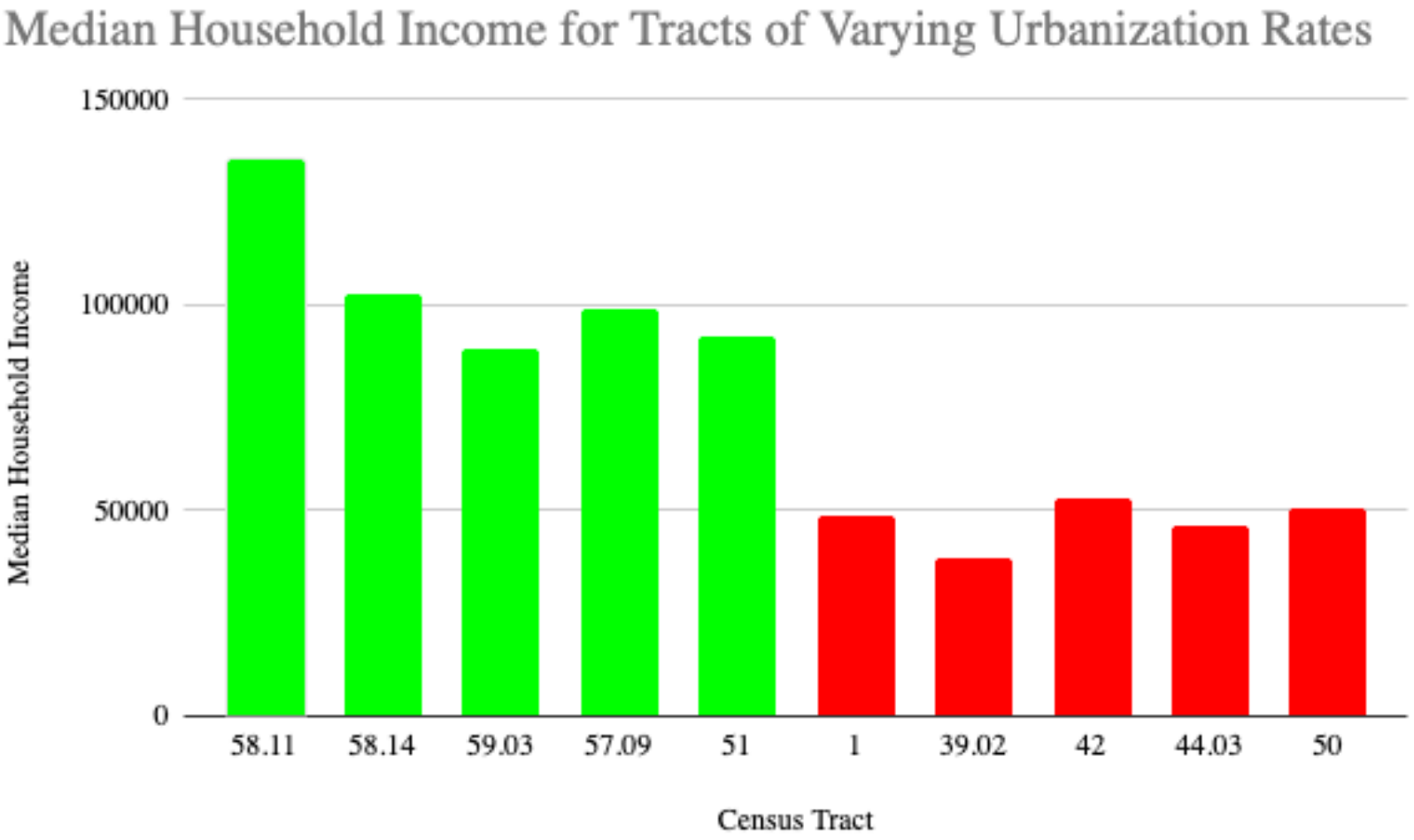


**Figure 2.14:** Median household income for the five fastest-urbanizing tracts (green) vs. the five slowest-urbanizing tracts (red).

**Table 2.3:** Distance between new urban land and geospatial factors

| | **Water polygons** | **Road lines** | **Urban polygon** |
|---|---|---|---|
| Mean distance | 1143.33 | 924.96 | 13.97 |
| Minimum distance | 0 | 0 | 0 |
| Maximum distance | 4739.61 | 7291.91 | 368.60 |
| St. deviation | 859.39 | 1138.87 | 33.76 |

*Mean, minimum, maximum, and standard deviations for each newly-urbanized area's distance (meters) from major geospatial drivers for urbanization.*

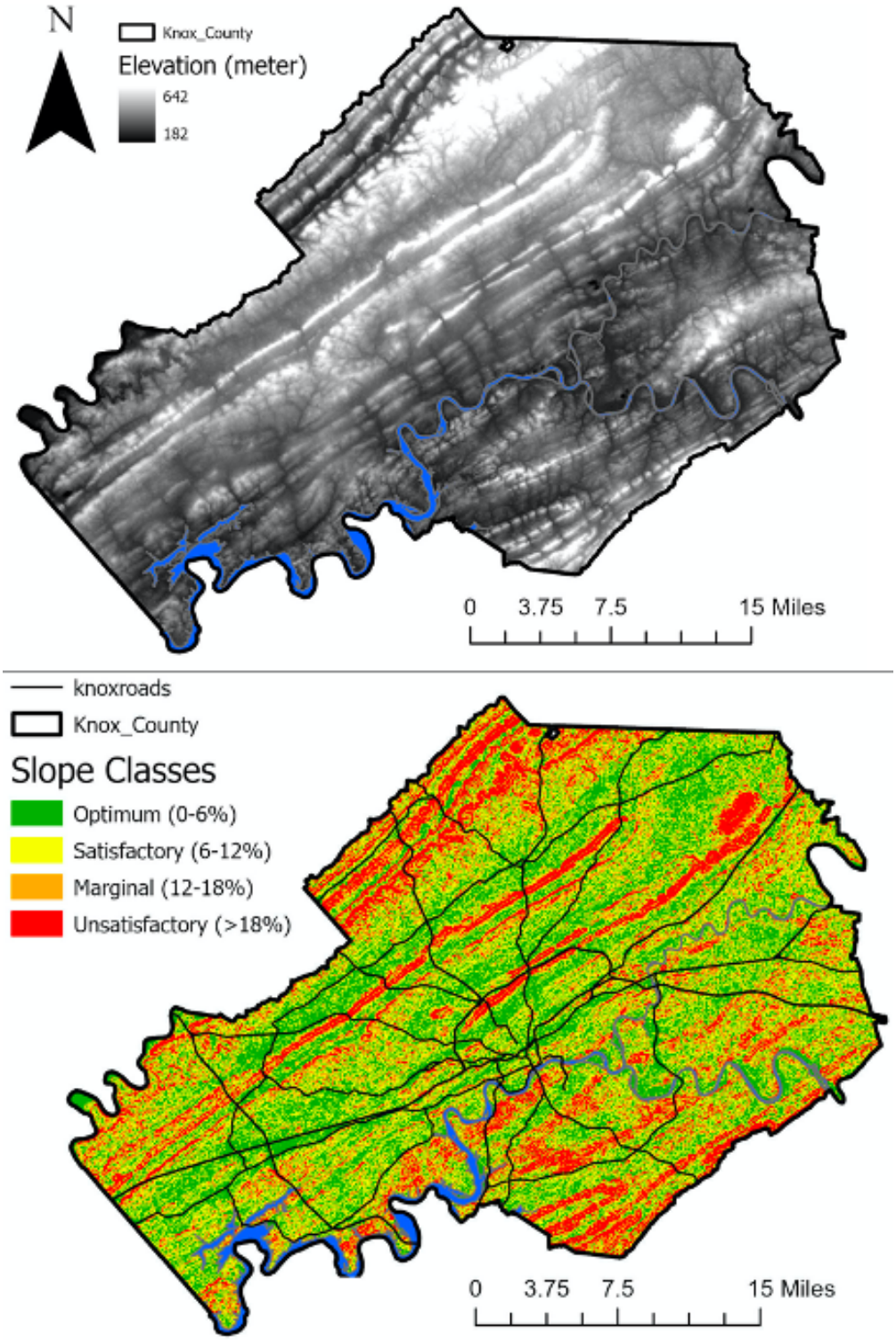


**Figure 2.15:** Digital Elevation Model and slope map for Knox County, TN

**Table 2.4:** Area of new urban land developed on optimal, satisfactory, marginal, and unsatisfactory slope terrain.

| **Percent Slope** | **Pixel Count** | **Area(square miles)** |
|---|---|---|
| Optimal (0-6%) | 26723 | 9.29 |
| Satisfactory (6-12%) | 25637 | 8.91 |
| Marginal (12-18%) | 11321 | 3.93 |
| Unsatisfactory (>18%) | 6455 | 2.24 |

*Breakdown of slope severity on which new urban land was developed. "New urban land" refers to land classified as developed in 2019 but was not classified as developed in 2001.*

## 2.5 Discussion

### 2.5.1 Land Use Changes

The results for the land use change analysis show that urban land has significantly increased in area since 2001. It had the most growth out of every land use classification. This expansion came at the expense of forest and agricultural (hay/pasture) land area. These trends follow similar patterns as studies mentioned in the literature review, in which urbanization causes forest land to transition into agricultural land, and agricultural land to transition into developed urban land [21].

Another key takeaway from this section is the build up and increased intensity of the urban land. As mentioned in Chapter 1, urbanizing cities typically increase the footprint of the city through urban sprawl and also become denser. Seeing as the urban land area has increased from 2001 to 2019, it can be concluded that urban sprawl is occurring in the county. Additionally, the urban intensity data here shows that Knoxville is becoming denser as it urbanizes as well. These morphological changes in the city structure can allow it to accommodate a growing population at the expense of agricultural and ecologically-important land.

### 2.5.2 Drivers for Urbanization

The main drivers for urbanization that were studied in this section were: terrain slope, proximity to existing urban land, proximity to water bodies, and proximity to major roads. The mean distance from each area of recent urbanization to the various infrastructures indicates that proximity to existing urban land is the dominant driving force associated with urbanization in this region. This result makes sense considering a large, central portion of the county is already urbanized. It shows that urban land tends to radiate outward from a central urban area and is generally unlikely to emerge far from any existing urban land. Surprisingly, proximity to roadways did not have as much as an impact as expected. One potential explanation for this could be that the NLCD classifies some highways and interstates as developed land, so the roadway "line" is surrounded by the developed land classification.

This would make proximity calculations show that new urban areas are closer to the urban classification than the road line, even though they are representing the same thing.

Slope also plays an important role in urban development. A majority of the newly-developed land took place on sites with optimal or satisfactory slopes. The classification with the harshest slope, "unsatisfactory", had the least amount of development. This trend is likely caused by the increased development costs associated with building on steep slopes.

## 2.6 Conclusion

To conclude, there have been many land use changes in Knox county in the 21st century, namely the extensive build up of urban land at the cost of a reduction of agricultural and forest land. As the population grows, Knoxville is expected to become denser and expand its city footprint. The new urban areas are spreading from existing urban land, but their geospatial distribution is also influenced by terrain slope, proximity to water bodies, and roadways. These are all important factors that should be considered as urbanization continues to occur in Knoxville.

# Chapter 3

# Environmental Footprint of Urbanization

## 3.1 Introduction

While it is important to catalog and quantify changes to the county's land use, it is arguably more important to understand the impacts and consequences associated with these changes. As stated in the literature review, urbanization can lead to many environmental issues such as the urban heat island effect and an increased flood risk. Cities and local planners must be cognizant of these issues both from the perspective of how past land use changes caused current environmental issues as well as how current and upcoming urban planning policies can cause even more environmental problems in the future.

The main technical objective of this section is to quantify some of the consequences of Knoxville's past land use changes in terms of their environmental impact. This section will address the following research questions:

1. How has Knoxville's urbanization from 2001-2019 affected the city's vulnerability to extreme heat events and urban flooding?

2. How have land use changes impacted the vegetation levels across the county?

## 3.2 Materials and Methods

### 3.2.1 Surface Temperatures

In order to analyze the urban heat island effect, it is necessary to compare surface temperatures from the urban areas to the non-urban areas in the county. Surface temperature data was sourced from Landsat 8 surface temperature datasets (ST_B10) from Earth Explorer and input into ArcGIS to create surface temperature maps for the entire county for summer days in 2001 and 2019. In order to obtain the temperature data in Fahrenheit, the remote sensing data had to be converted using the scale factor of (Value*0.00341802)+149.0 as indicated in the Landsat 8 product guide. This converted the data to Kelvin, which was then converted to Fahrenheit. Data conversions were done using the Raster Calculator geoprocessing tool in ArcGIS. This data is the remotely-sensed surface temperature as a pixel for each 30x30 meter area, much like the NLCD land use data used in chapter 2. Due to limitations on the amount of available without cloud cover, only one dataset, or day, was used for each year.

Once the county-wide surface temperature maps were created, the map layer for the 2019 urban extent was used to extract the surface temperatures for both 2001 and 2019. The reason that the 2001 surface temperatures were masked to the 2019 urban area polygon was to compare overall changes in temperature in a uniform area. This can help to show how surface temperatures not only increase in magnitude but also how elevated surface temperatures affect a larger area as the county becomes increasingly urban. This method created a layer for only urban surface temperatures. Then, another layer was created by excluding the area within the urban polygon and instead clipping the non-urban land with the surface temperature layer. This was done to understand the extent of which urbanization artificially changes surface temperatures in contrast with rural or non-urbanized land uses. Two other layers were used for masking the data: the top ten fastest-urbanizing tract layer and the bottom ten fastest-urbanizing tract layer. By extracting only surface temperatures from these areas, the changes in surface temperatures in urbanizing vs. non-urbanizing areas can be analyzed. These maps were analyzed using the statistical analysis geoprocessing tool to determine the minimum, maximum, mean, and standard deviation of the large dataset containing one temperature reading for each of the thousands of pixels at 30x30m resolution.

An overview of the methodology used for this section can be seen in Figure 3.1.

### 3.2.2 Impervious Surfaces

The next environmental issue to analyze is the build-up of impervious surfaces, which has been shown to increase urban flood risk [5], [7].

The first step for analyzing ISA increases is to obtain the NLCD impervious surface data layer for the US, which can be found on the ESRI living atlas. To use this data for 2001, 2011, and 2019, the ArcGIS Pro search catalog pane was used to find the dataset and add the data to the ArcGIS Pro project. Then, a definition query was used to select only for the relevant years. Next, this data was extracted by mask to the Knox county polygon for each year. From here, impervious surface area data layer was extracted by mask to the top ten fastest-urbanizing tract layer in order to narrow the scope to understand the extent of ISA build-up in west Knoxville, where urbanization is occurring most rapidly.

### 3.2.3 Vegetation Losses

The methodology for the NDVI analysis is similar to the surface temperature analysis. First, Landsat 4 ST_B03 and ST_B04 data was taken from Earth Explorer. These data layers represent the red and near infrared (NIR) bands. Once this data was imported, layers for band 3 and band 4 were converted from digitally-sensed values (DN) to band wavelength through the conversion formula (DN - 0.0000275) - 0.2, as explained in the Landsat 4 collection 2 level 2 product guide. Then, the raster calculator tool was used to find the NDVI values with the NDVI formula seen in Figure 3.2.

Once the data was converted, a difference map was created to show how the NDVI values for each 30x30m pixel changed from 2001-2019. This can be seen in Figure 3.14. Then, this overall map was extracted to just the highly-urbanizing tracts to understand the vegetation-urbanization dynamics for the fastest growing areas. This map can be seen in Figure 3.15.

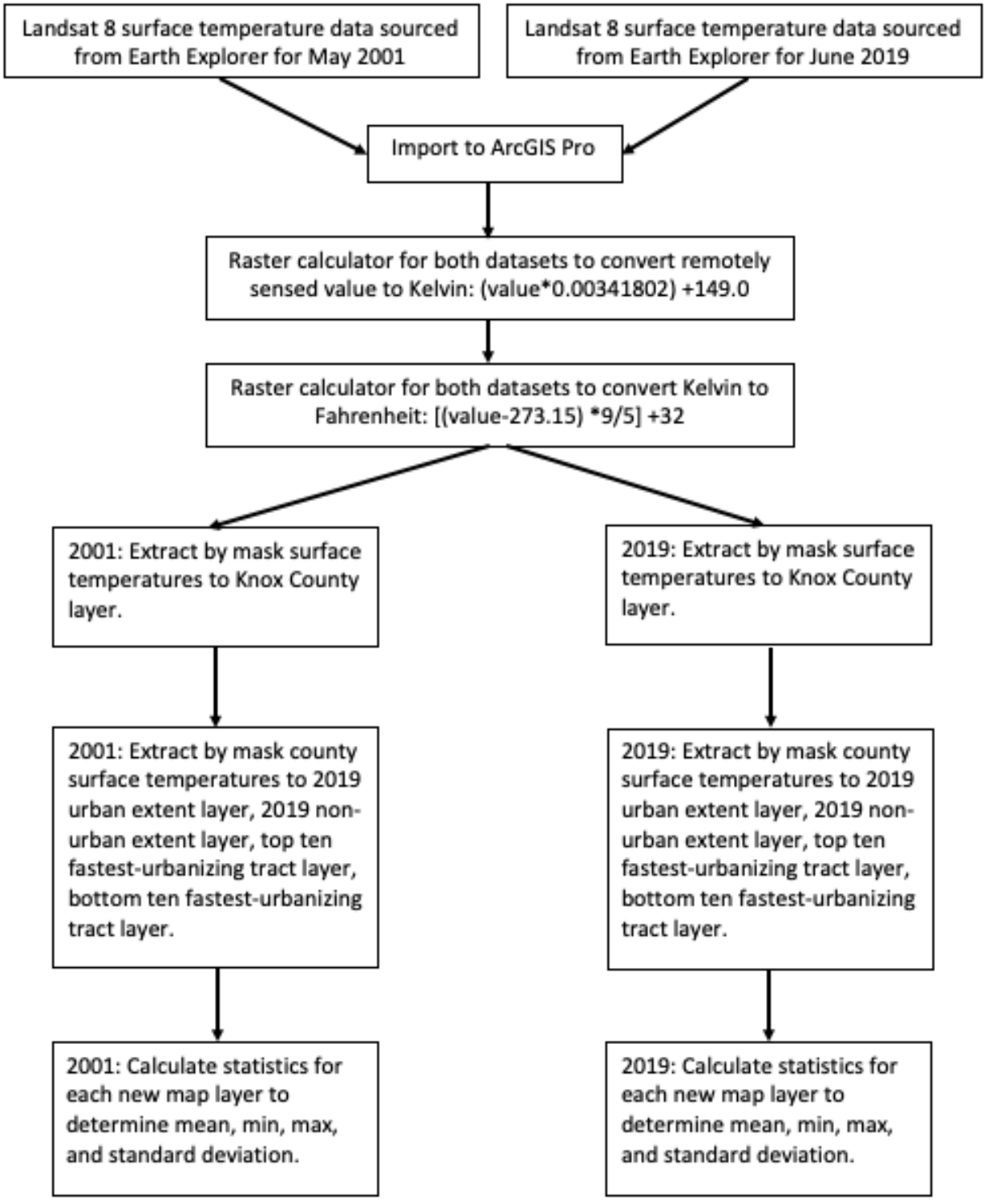


**Figure 3.1:** Flow chart describing surface temperature analysis methodology.

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$$

**Figure 3.2:** Normalized Difference Vegetation Index (NDVI) formula.

## 3.3 Results and Discussion

### 3.3.1 Urban Heat Island Effect

The analysis of surface temperatures for 2001 and 2019 show that there is a significantly higher average surface temperature in urban areas as opposed to areas with any other land use classification. Additionally, the fastest-urbanizing tracts have an average temperature of around 5.5 degrees F higher than the county average, and almost 10 degrees F higher than the slowest-urbanizing tracts for 2001. The urbanizing tracts have a slightly higher average temperature for 2019 when compared to the county average and the slowly-urbanizing tracts, as shown in Table 3.1 Table 3.2 Table 3.3. This data shows that the urban heat island effect is certainly affecting temperatures in Knox county. One limitation to this data is the presence of cloud cover. It is very difficult to find county-wide Landsat data that is completely free of clouds, and the clouds impact the temperature data. For the purposes of this research, many of the "minimum" temperature values seen are actually cloud temperatures and do not accurately reflect the temperature of the land below. Additionally, the extremely high temperatures seen in the "maximum" sections in the tables are likely large buildings that absorb and retain heat. This data is still relevant because the presence of urban land leads to more of these heat-retaining buildings and unlike the cloud cover, these do affect the temperature of the surrounding area. Due to limitations with the amount of data that could be obtained, it is difficult to say if the urban heat island effect is worsening. However, this data does show that urban areas in Knox county are associated with higher surface temperatures, so one could conclude that as the urban area expands, the urban heat island effect will increase its affected area as well.

**Table 3.1:** Surface Temperatures (Fahrenheit) urban vs. non-urban land use types

| | **Urban (2001)** | **Non-urban (2001)** | **Urban (2019)** | **Non-urban (2019)** |
|---|---|---|---|---|
| Mean | 87.99 | 77.84 | 93.83 | 84.27 |
| Minimum | 45.73 | 38.88 | 54.01 | 52.47 |
| Maximum | 131.51 | 122.55 | 141.14 | 128.01 |
| St. deviation | 8.12 | 7.34 | 9.07 | 6.42 |

*Mean, minimum, maximum, and standard deviations surface temperatures for areas classified as "urban" vs. all other land use types.*

**Table 3.2:** Surface Temperatures (Fahrenheit) in Various Areas in Knox County, May 2001

| | **County** | **Urbanizing** | **Non-urbanizing** |
|---|---|---|---|
| Mean | 81.82 | 87.23 | 77.68 |
| Minimum | 38.88 | 74.85 | 38.88 |
| Maximum | 131.51 | 128.81 | 114.50 |
| St. deviation | 9.12 | 5.36 | 8.17 |

*Surface temperature mean, minimum, maximum, and standard deviations for the county, top ten fastest-urbanizing tracts, bottom ten fastest-urbanizing tracts, urban land use, and non-urban land use areas, respectively.*

**Table 3.3:** Surface Temperatures (Fahrenheit) in Various Areas in Knox County, June 2019

| | **County** | **Urbanizing** | **Non-urbanizing** |
|---|---|---|---|
| Mean | 88.02 | 89.12 | 85.94 |
| Minimum | 52.47 | 56.53 | 54.73 |
| Maximum | 141.14 | 129.60 | 131.59 |
| St. deviation | 8.89 | 8.72 | 6.42 |

*Surface temperature mean, minimum, maximum, and standard deviations for the county, top ten fastest-urbanizing tracts, and bottom ten fastest-urbanizing tracts.*

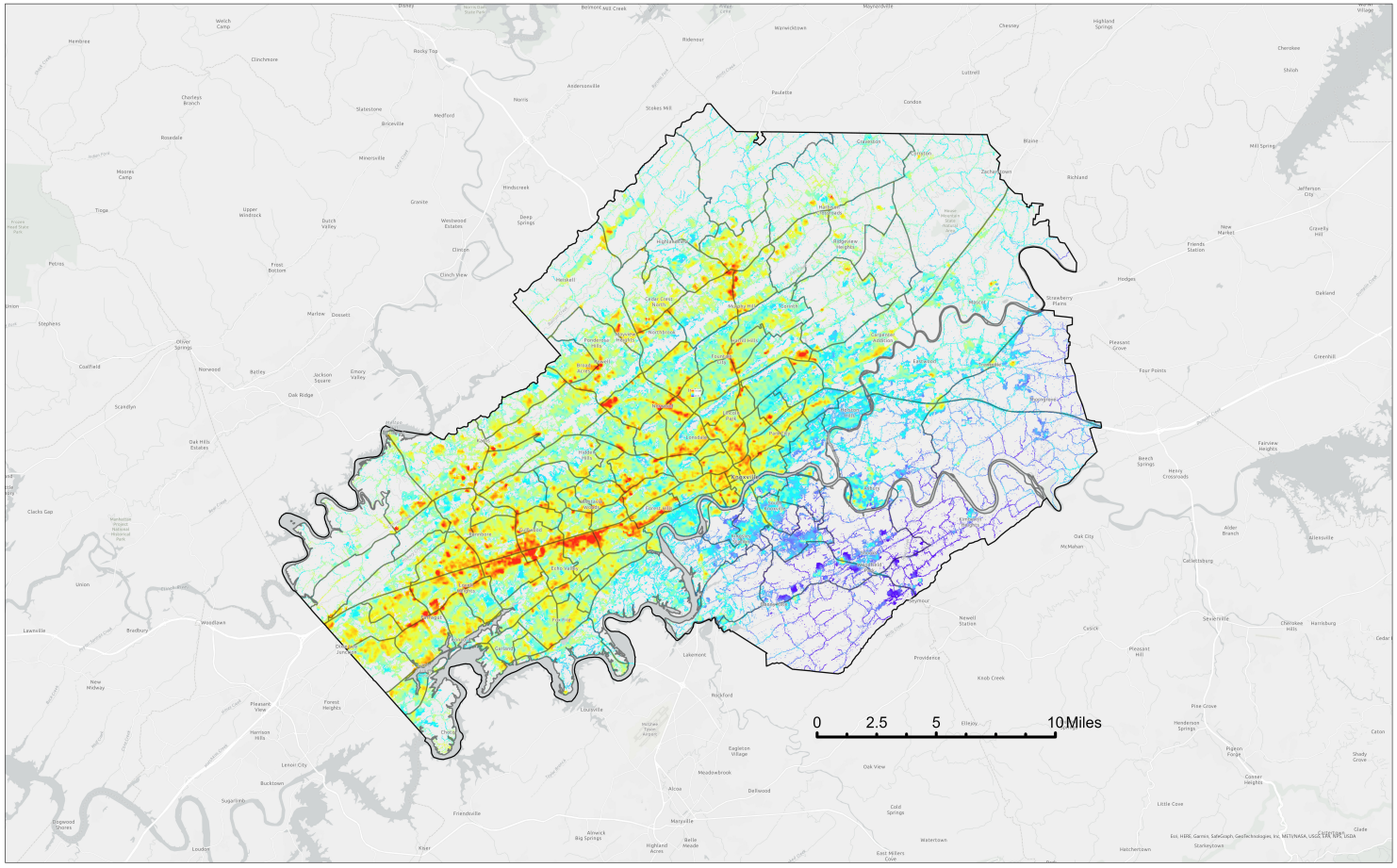


**Figure 3.3:** Temperature map for developed land use types May 30, 2001.

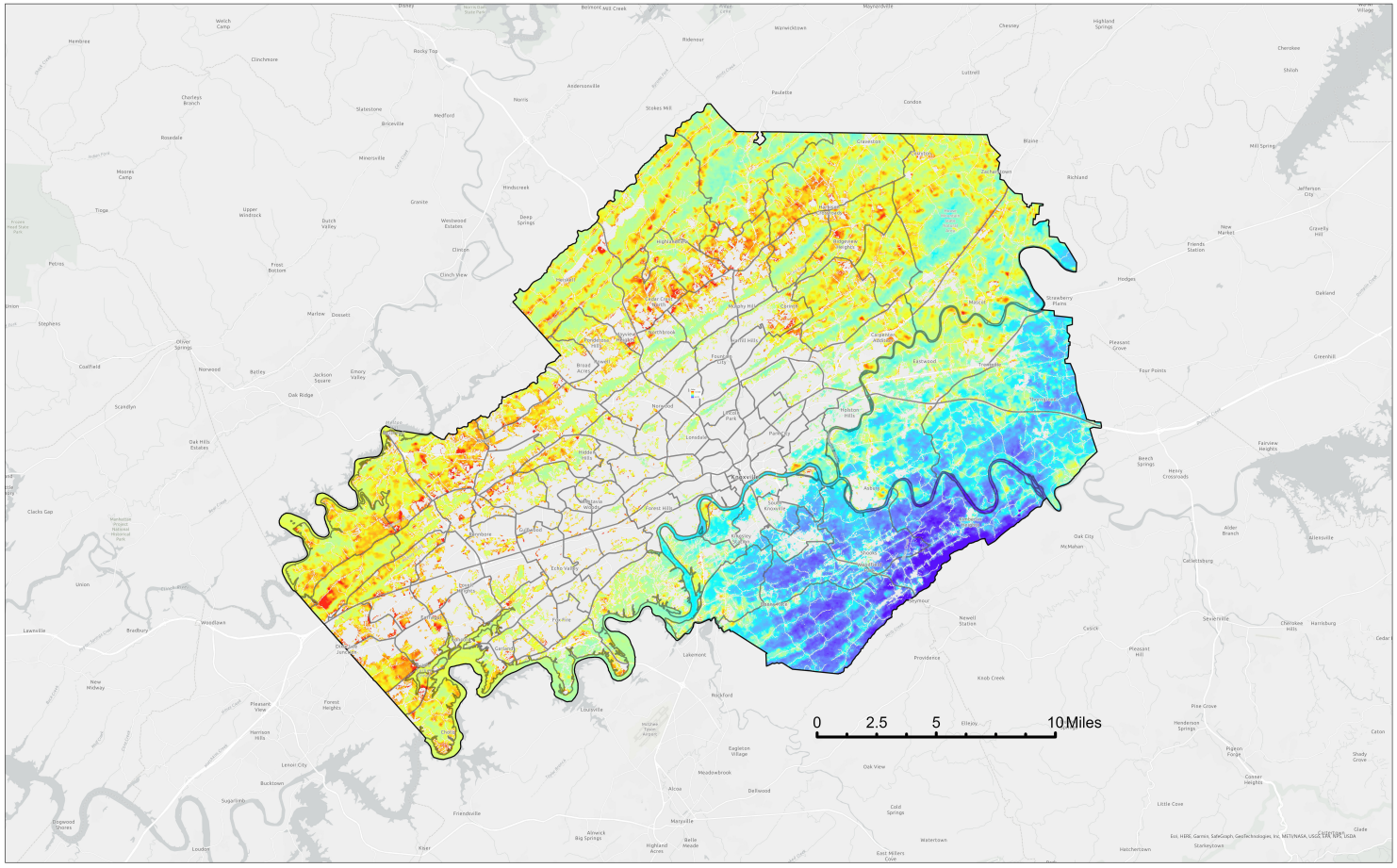


**Figure 3.4:** Temperature map for undeveloped land use types May 30, 2001.

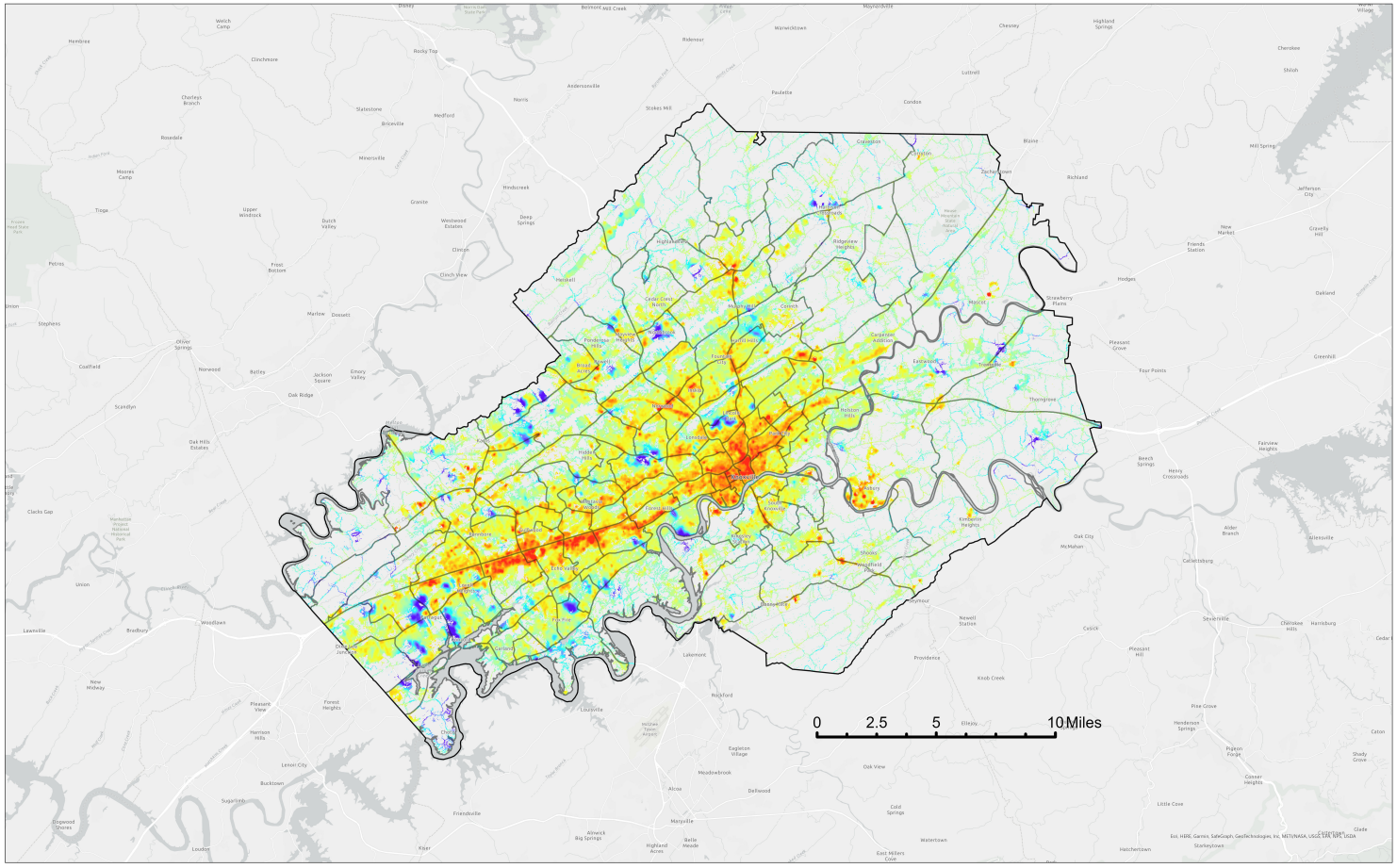


**Figure 3.5:** Temperature map for developed land use types June 21, 2019.

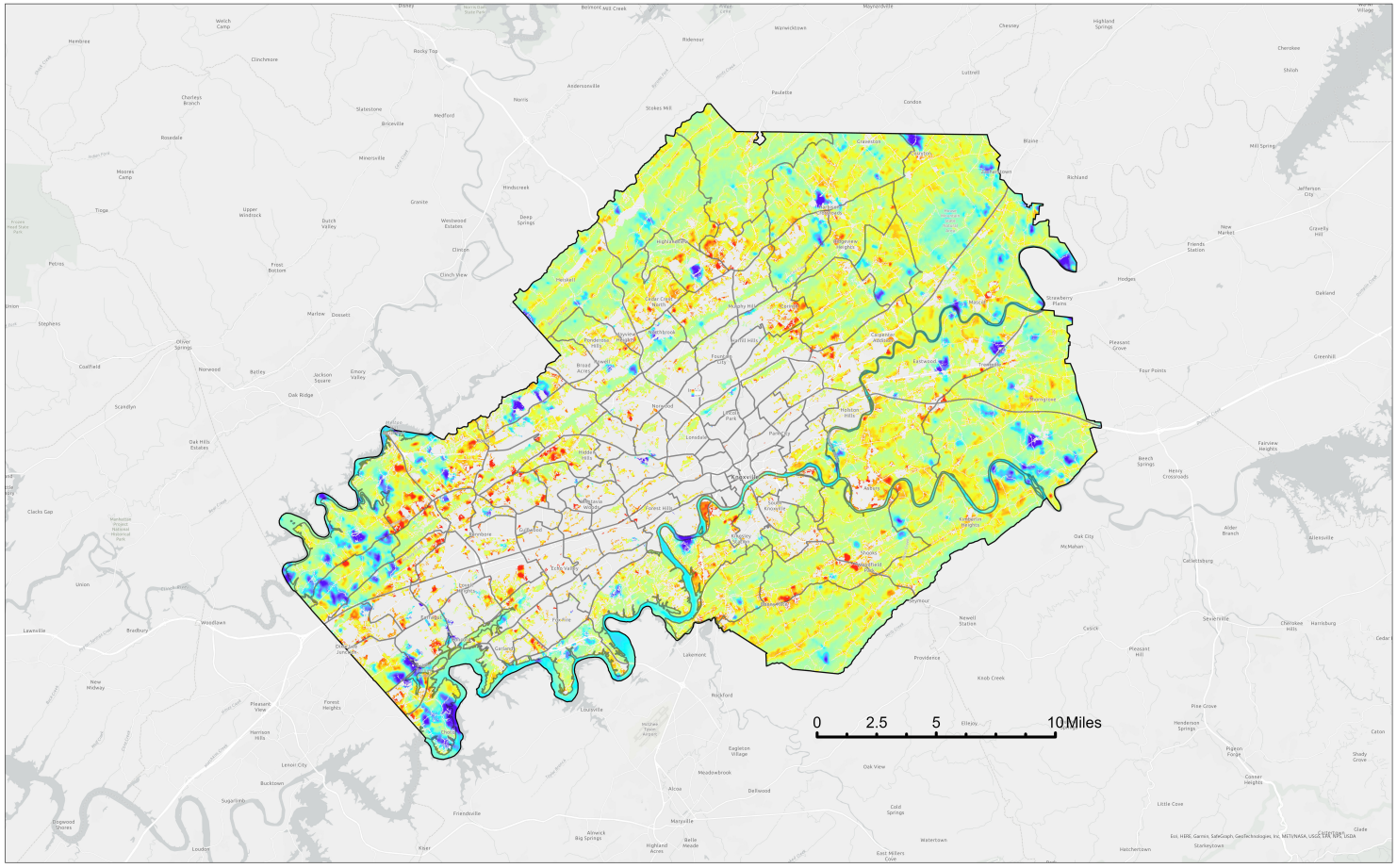


**Figure 3.6:** Temperature map for undeveloped land use types June 21, 2019.

### 3.3.2 Impervious Surface Build-Up

The maps for the impervious surface build-up analysis on a county level can be seen in Figure 3.7, Figure 3.8, and Figure 3.9. These maps show impervious surface build-up has been occurring most in Western Knox county and along the interstate connecting this region to the downtown center. Seeing as this build-up is widespread in western Knoxville, maps were created to highlight impervious surface changes in the rapidly-urbanizing census tracts in west Knoxville. These maps can be seen here: Figure 3.10, Figure 3.11, Figure 3.12.

Table 3.4 and Figure 3.13 show the build-up of impervious surface area over the study years. On the county level, ISA has increased from 65.85 square miles in 2001 to 78.43 square miles in 2019. This trend of ISA increase can also be seen in both the urbanizing and non-urbanizing tracts. When looking specifically at the ten tracts that are urbanizing the fastest, it is clear that the impervious surface area is positively correlated with urbanization rate. The percentage of ISA in these tracts increases from 12.19% in 2001 to 19.58% in 2019.

This has severe implications for Knoxville's vulnerability to urban flooding. As previously mentioned, urban flooding risk increases as ISA increases. This means that areas with rapid growth are also becoming more vulnerable to flooding.

### 3.3.3 Vegetation Losses

The NDVI maps show that while there has been a decrease in vegetation quantity in some areas, there have been unexpected gains in vegetation amounts as well. The NDVI map shows that the central part of the county south of the Tennessee River has experienced widespread vegetation losses. However, this trend does not seem to be consistent with urbanization patterns. In fact, Figure 3.12 shows that the fastest-developing census tracts have generally had an increase in vegetation.

Table 3.5 shows that the overall average changes in NDVI were very minimal. When examining the NDVI maps, this makes sense because there are areas with positive and negative change, so the overall change is minimal from year to year.

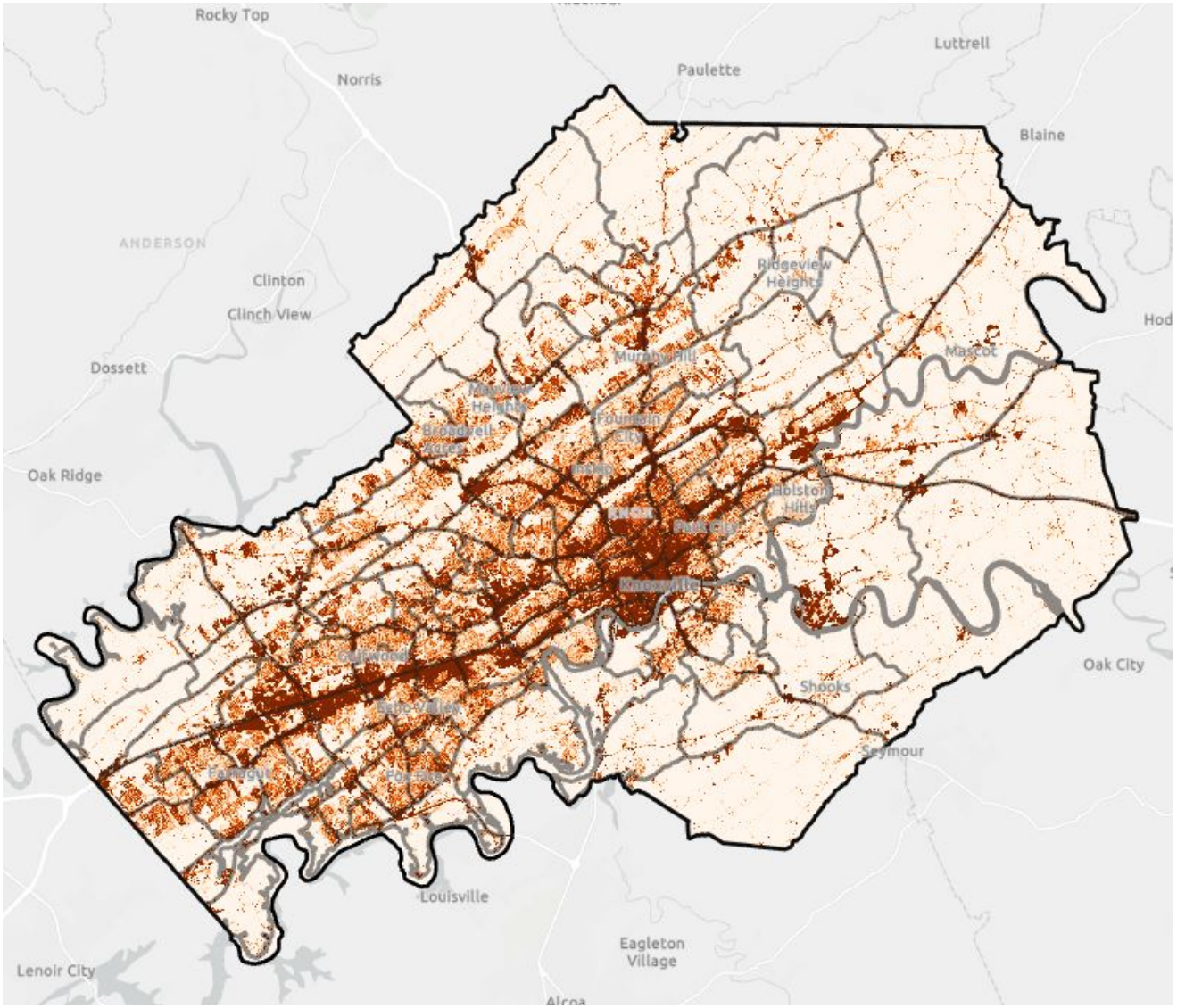


**Figure 3.7:** Impervious Surface Area in 2001. Darker color represents higher percentage of impervious surfaces represented by the pixel.

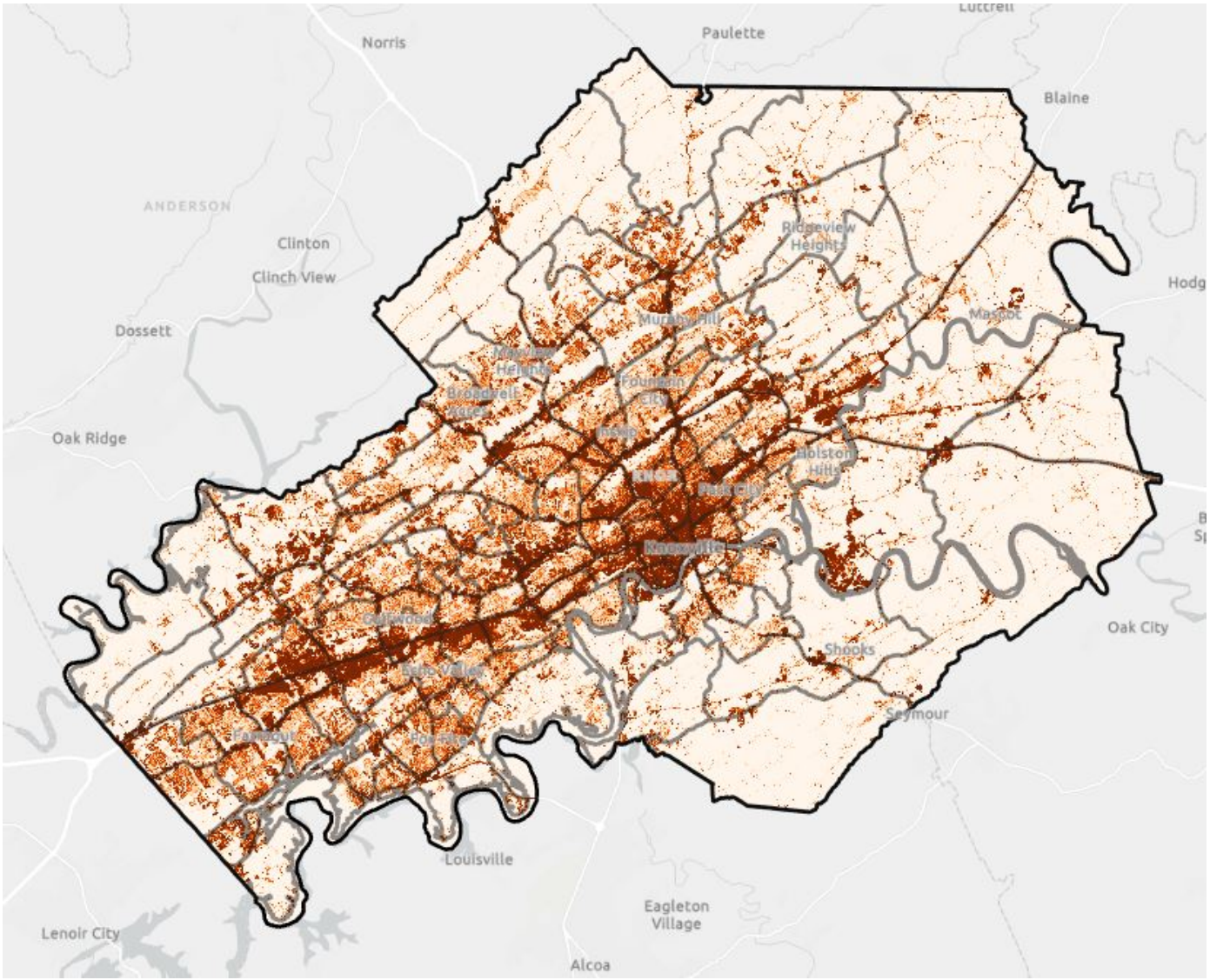


**Figure 3.8:** Impervious Surface Area in 2011. Darker color represents higher percentage of impervious surfaces represented by the pixel.

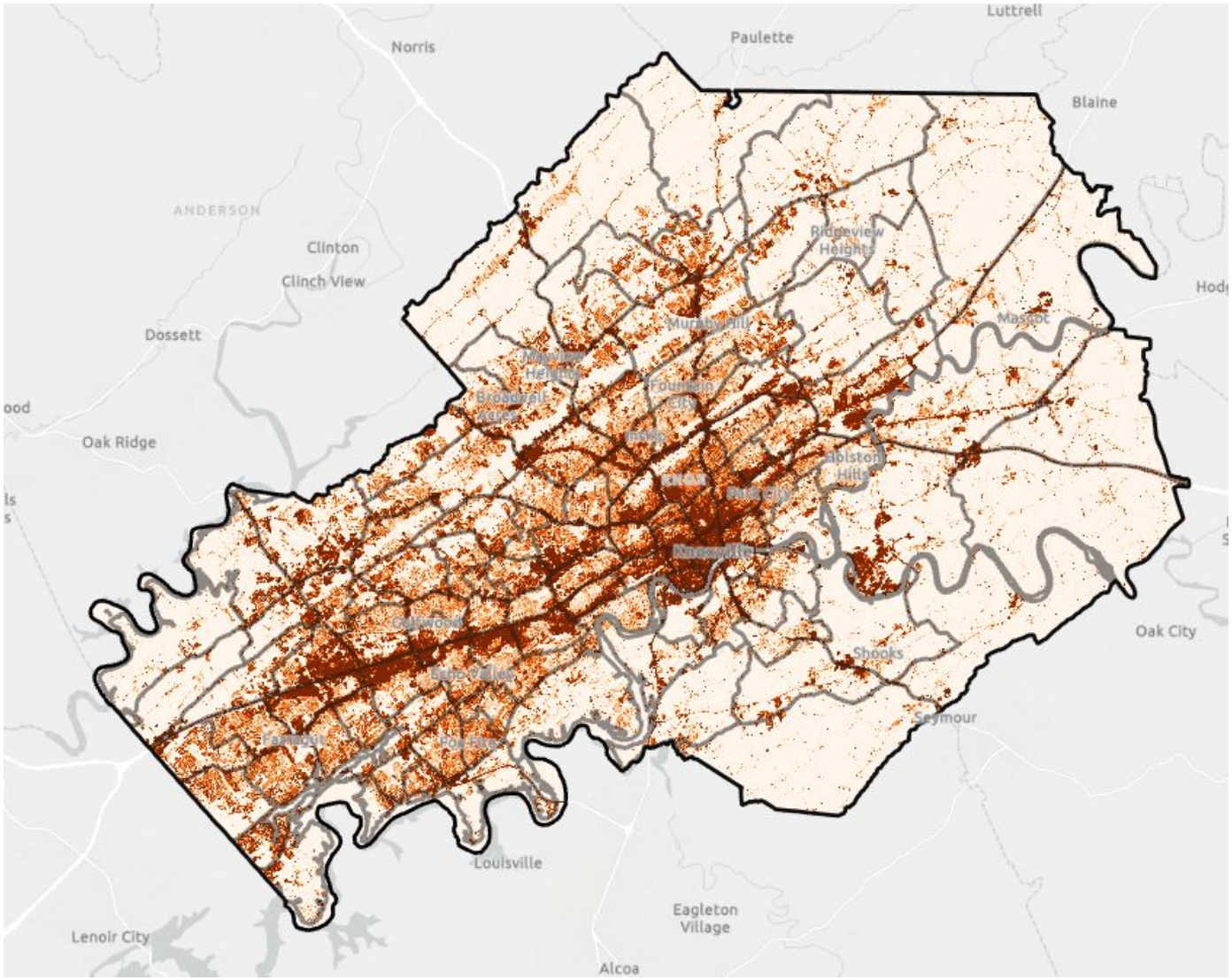


**Figure 3.9:** Impervious Surface Area in 2019. Darker color represents higher percentage of impervious surfaces represented by the pixel.

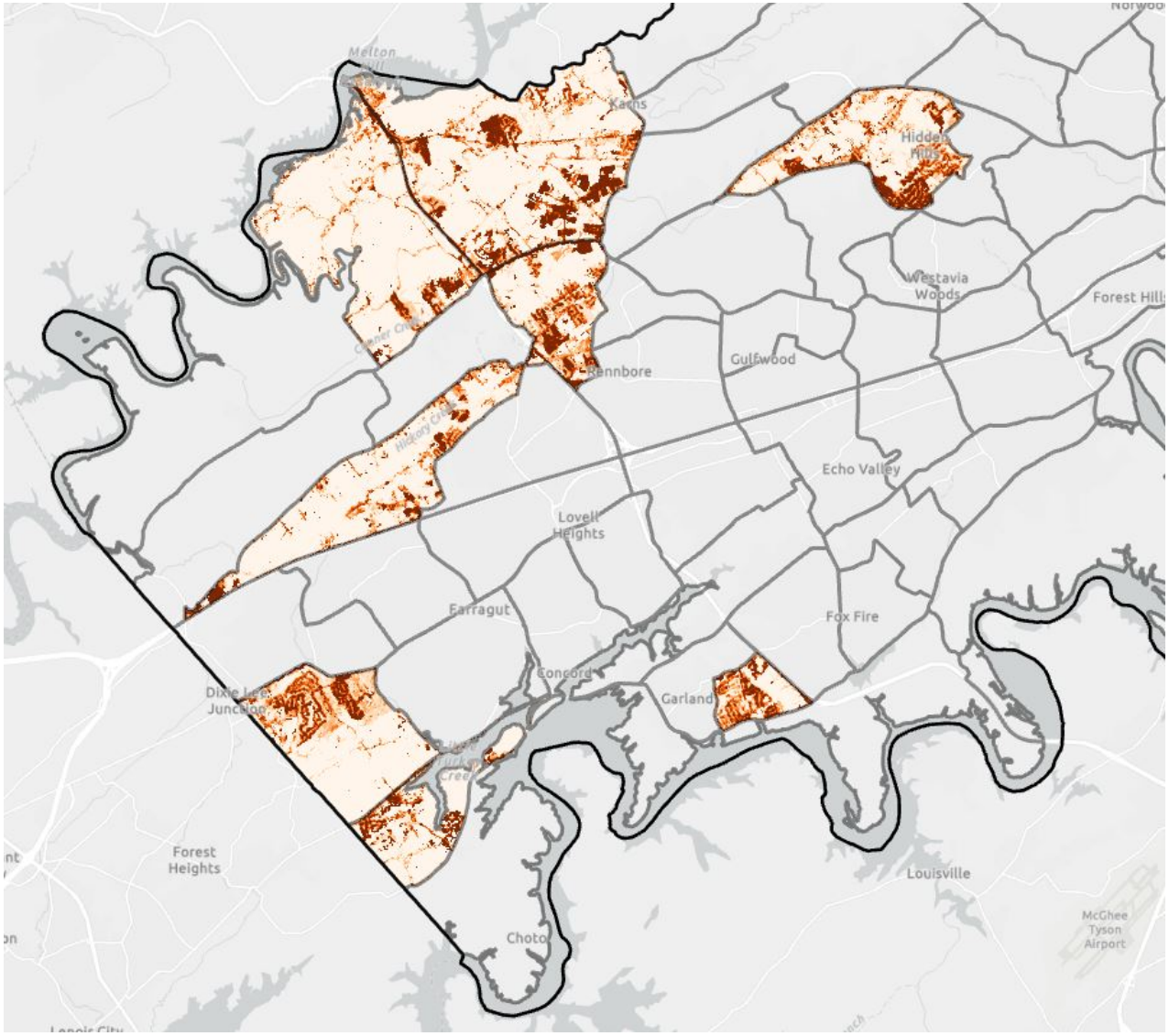


**Figure 3.10:** Impervious Surface Area for the fastest-urbanizing census tracts in 2001. Darker color represents higher percentage of impervious surfaces represented by the pixel.

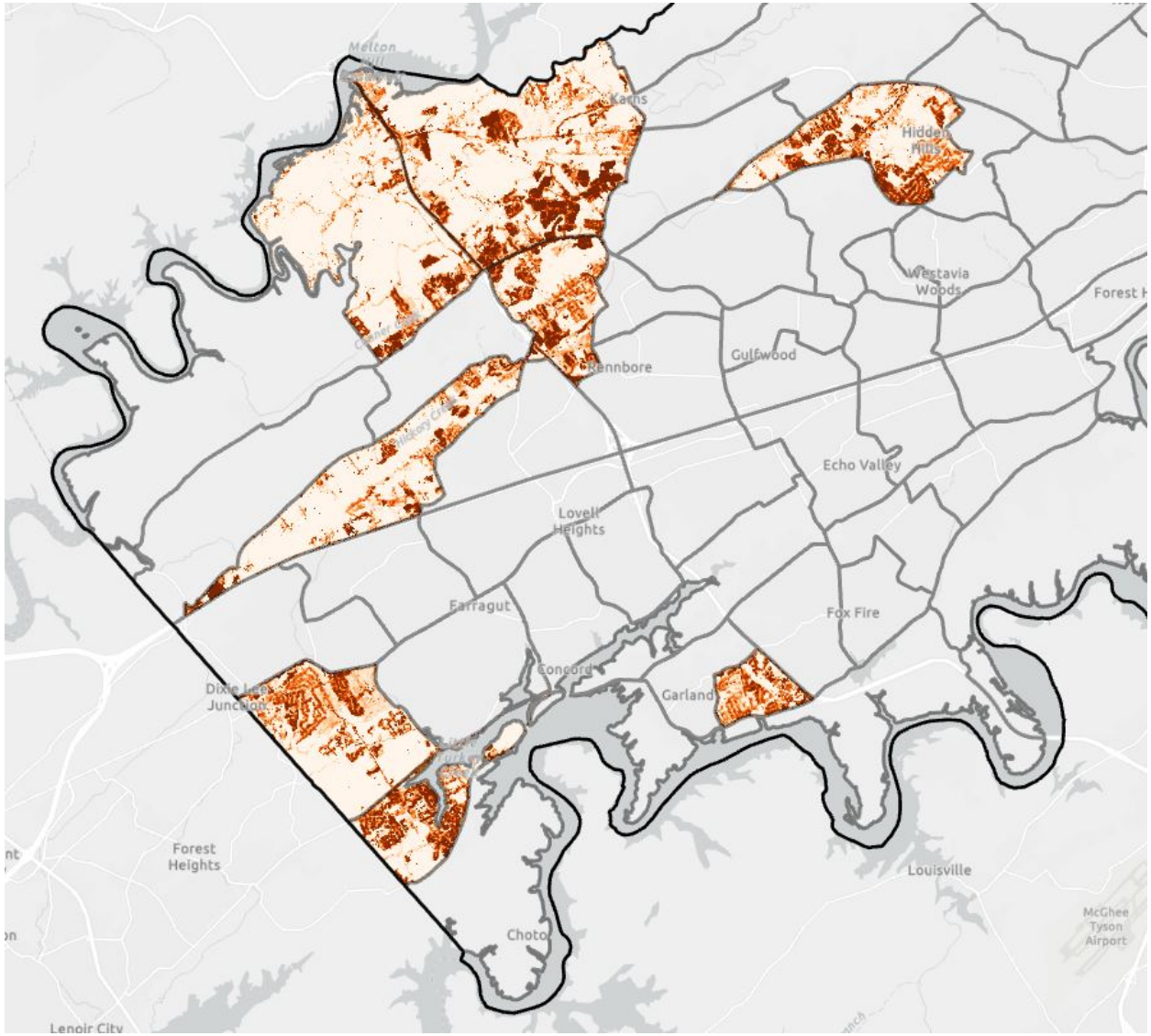


**Figure 3.11:** Impervious Surface Area for the fastest-urbanizing census tracts in 2011. Darker color represents higher percentage of impervious surfaces represented by the pixel.

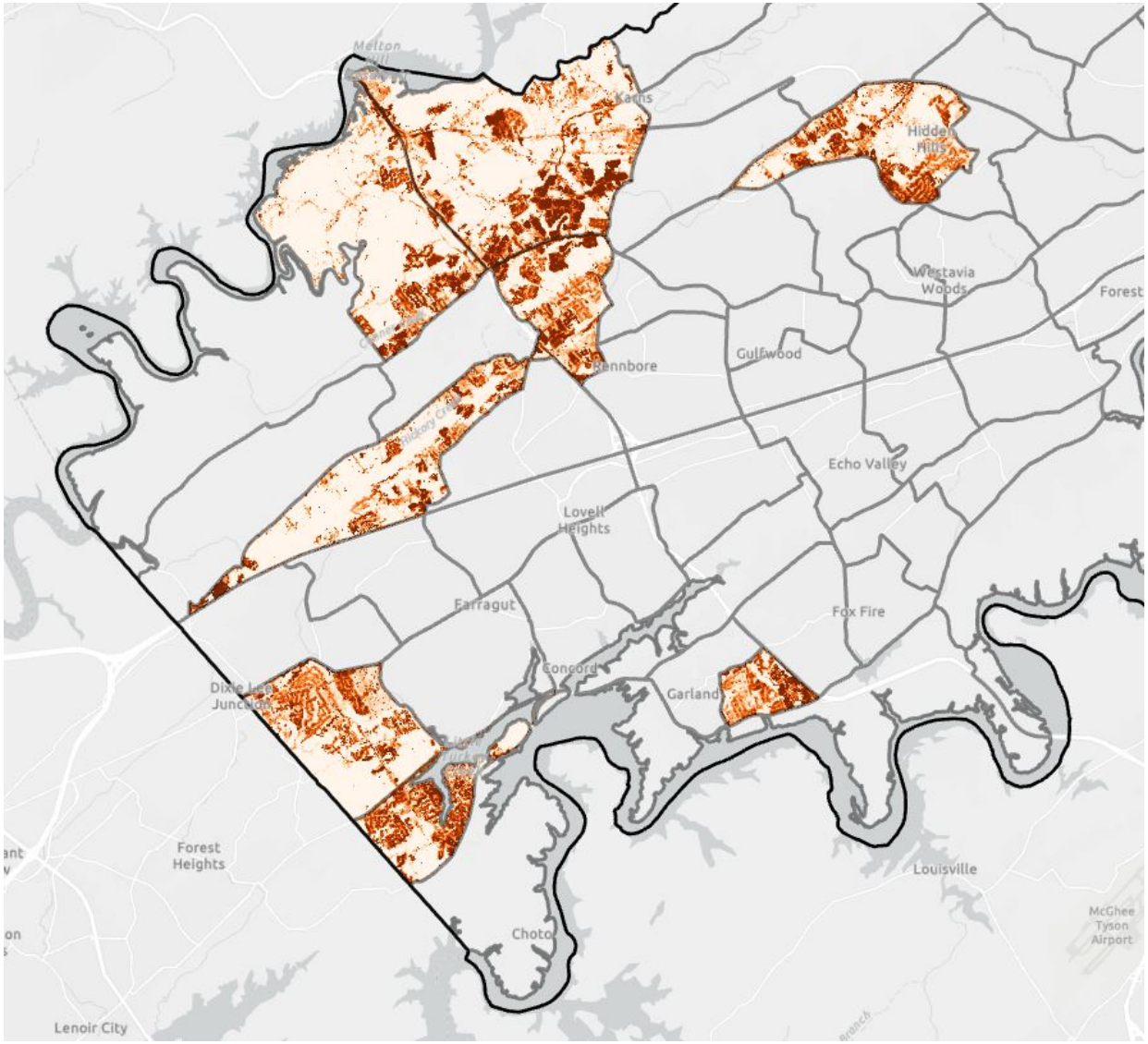


**Figure 3.12:** Impervious Surface Area for the fastest-urbanizing census tracts in 2019. Darker color represents higher percentage of impervious surfaces represented by the pixel.

**Table 3.4:** Impervious Surface Area

| | **2001** | **2011** | **2019** | **Total area of region** |
|---|---|---|---|---|
| County | 65.85 (12.52%) | 74.31 (14.13%) | 78.43 (14.91%) | 525.86 |
| Urbanizing | 4.79 (12.19%) | 6.75 (17.18%) | 7.69 (19.58%) | 39.28 |
| Non-urbanizing | 6.34 (5.52%) | 7.54 (6.57%) | 8.04 (7.00%) | 114.85 |

*Impervious Surface Area (square miles) for the county, top ten fastest-urbanizing tracts, and bottom ten fastest-urbanizing tracts. Percentages in parenthesis refer to the percent of the total region area that is impervious surface area.*

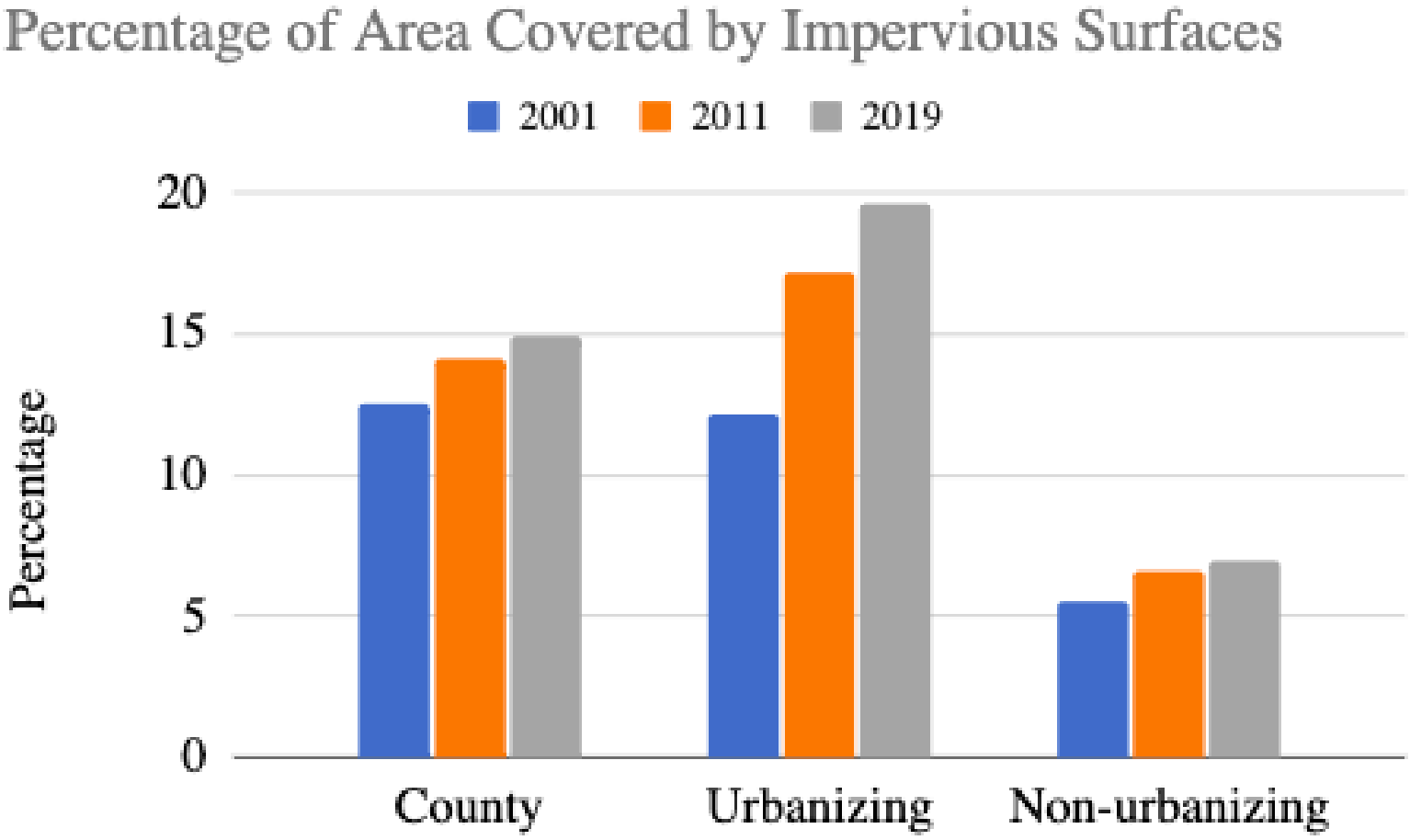


**Figure 3.13:** Impervious surface area expressed as a percentage of the region's total area.

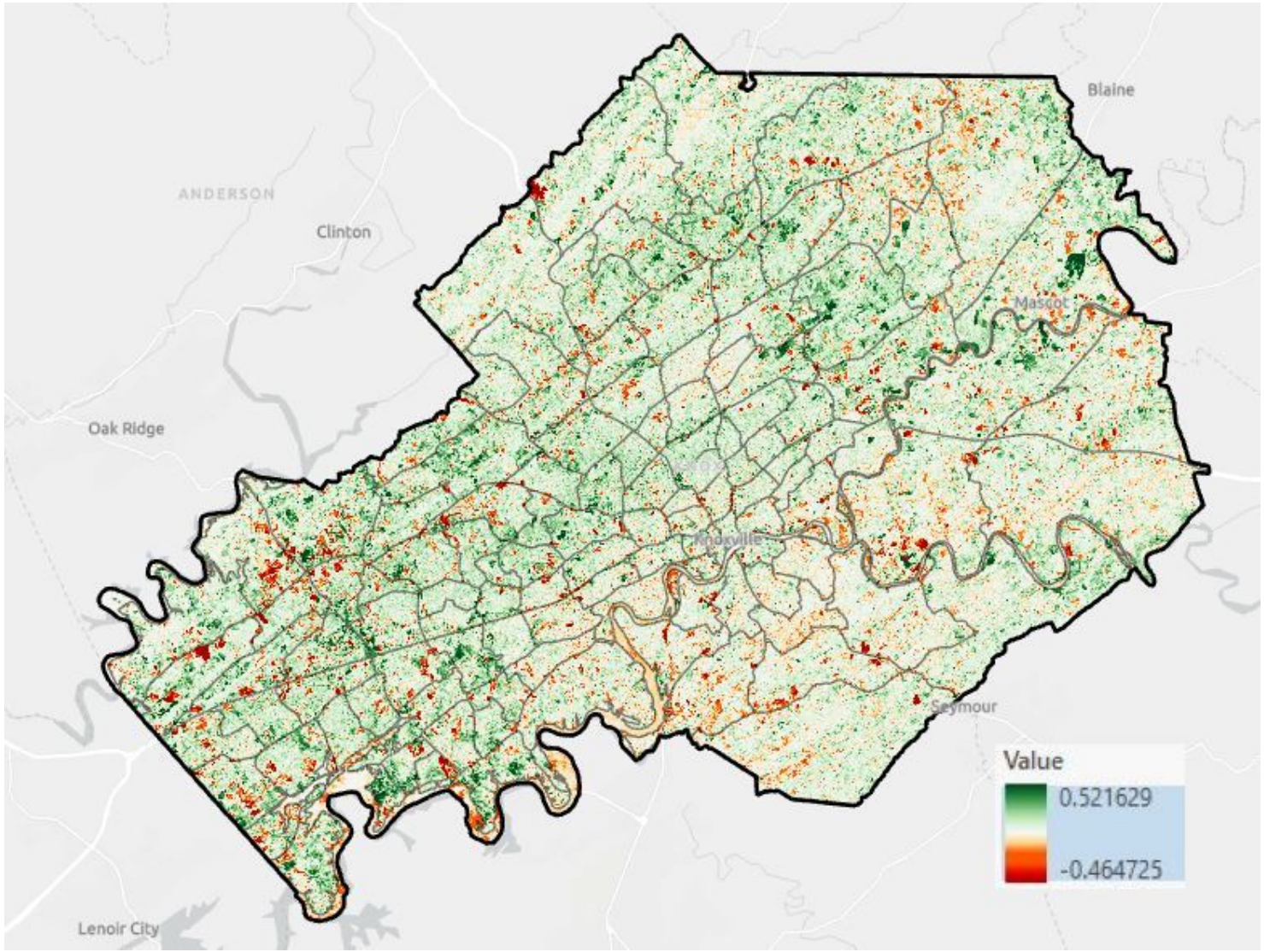


**Figure 3.14:** Change in NDVI value from 2001-2019.

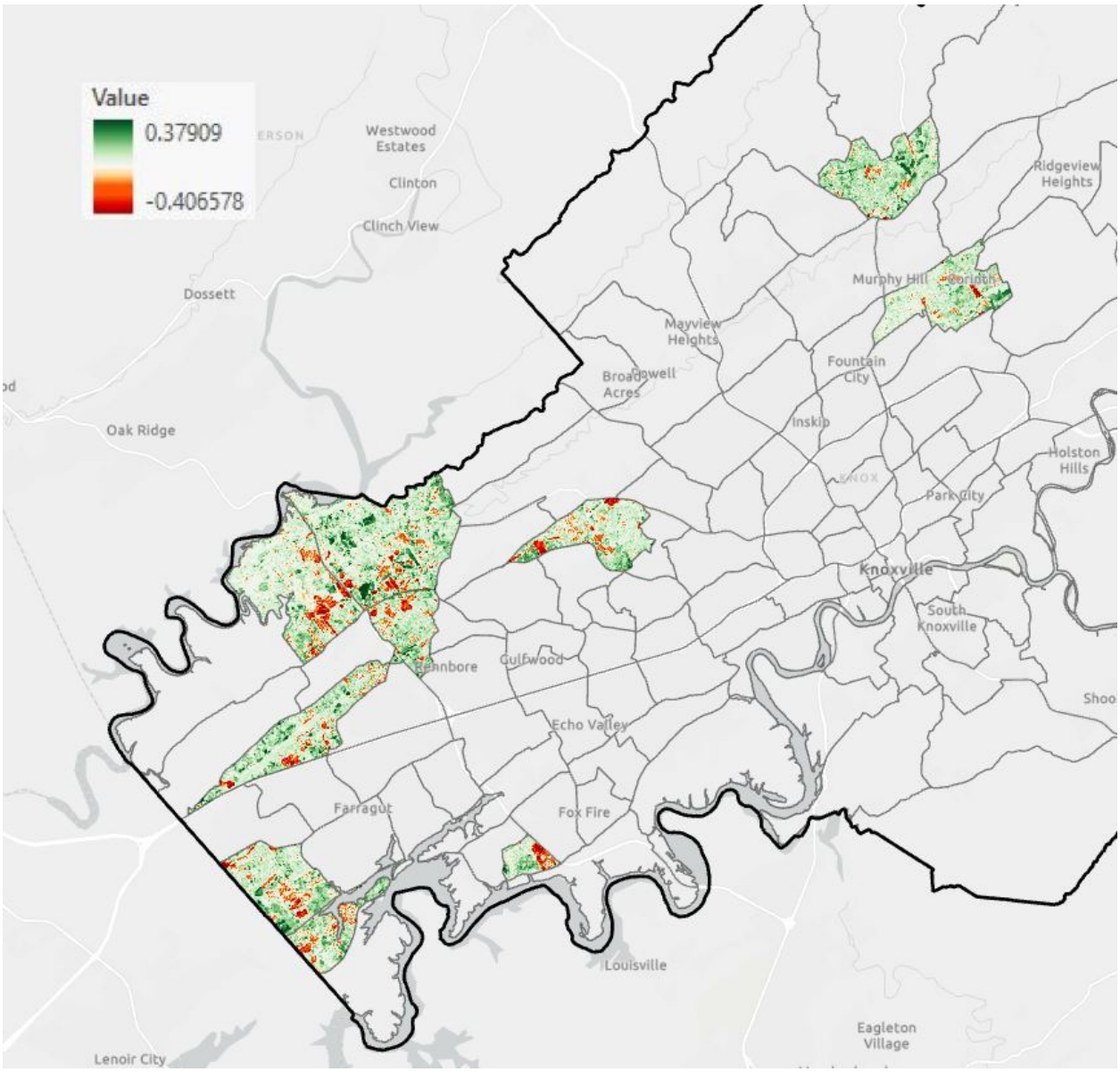


**Figure 3.15:** Change in NDVI value from 2001-2019 for the fastest-growing tracts in the county.

**Table 3.5:** Average NDVI change

| | **Urbanizing Tracts** | **Non-urbanizing Tracts** |
|---|---|---|
| Average NDVI Change | -0.0025 | -0.0062 |

## 3.4 Conclusions

To conclude, Knoxville's urbanization is likely increasing the extent of the urban heat island effect and the city's vulnerability to urban flooding. As mentioned in the literature review, one popular strategy to help mitigate the effect of the urban heat island phenomenon is to break up the urban continuity by creating parks, green roofs, and other types of green infrastructure.

As for the impervious surface build-up, this trend is occurring in almost every census tract. Seemingly, once impervious surfaces and urban land in general are built, they tend to stay for good. This poses problems for the city's resilience to flooding and can also exacerbate the urban heat island effect. The increase in ISA is most prominent in the rapidly-urbanizing census tracts, which makes sense as one of the ways urban/developed land is defined is the presence of ISA. However, most of this new developed land is replacing forest and pasture land. These two land uses that are becoming depleted were helping to absorb and infiltrate precipitation, and now they are instead creating additional water runoff. This could be of great concern because water runoff typically carries contaminants and pollutants, and most of the ISA build-up is happening near lakes and rivers. As urbanization continues, water contamination will likely increase as well.

The vegetation loss analysis showed that the overall change in vegetation density in the county is negligible. The analysis did not provide much useful information in part because of the nature of NDVI values. NDVI cannot easily distinguish between different types of vegetation, only density and abundance. There are also many factors that can affect NDVI readings, such as atmospheric disturbances like cloud cover and the time of year. Because of this, NDVI may not be a great indicator of areas that are ecologically important, at least at the resolution used in the project. Instead, land use type could be a more effective identifier.

All of the data collected in this section supports the idea that urban contiguity causes or exacerbates many problems in the urban area. By breaking up this continuity, the urban heat island effect would be lessened and water runoff could infiltrate inside of the city limits instead of causing additional runoff.

# Chapter 4

# Modeling to Identify Suitability and Predict Development

## 4.1 Introduction

This section will begin by explaining the goals, hypotheses, and objectives of the suitability model chapter. It is important to stress the need for a predictive urban model to include environmental factors such as the urban heat island effect, impervious surface build-up, and vegetation losses associated with urbanization.

The main technical objective of this section is to develop and compare two models for predicting urbanization patterns using geospatial parameters. Then, using the environmental factors from chapter 3, the models can be used to produce restricted urbanization patterns under scenarios of environmental protection policies.

The main research questions to answer in this chapter are:

1. Which non-urban areas are most likely or suitable to become urban in the near future?
2. How can the urbanization patterns be changed to protect important ecological areas?
3. How can urbanization patterns be changed to reduce the city's vulnerability to weather-related events?
4. What are the strengths and weaknesses of the two different models?

Creating a model to predict, or at the very least identify, areas of future urbanization can have a wide range of benefits to urban planners, policy-makers, and even the average person who is looking to buy real estate in the county. Then, the models can be adjusted to evaluate areas of high vegetation or high slope as a detractor from the overall suitability score. This would allow for planners to limit deforestation and ecosystem service loss as the city expands.

## 4.2 Materials and Methods

### 4.2.1 Suitability Model

In order to identify areas of potential urbanization, the first step is to create a suitability equation that includes relevant parameters. After those parameters are chosen, the average suitability is determined for each 30x30m pixel by taking into account its distance from each of the various factors.

Once the suitability map is created, this will serve as the predictive model to show where urbanization is likely to take place in the future. Then, the NDVI data layer can be used to restrict the future urbanization areas to only those which harm the natural ecosystems the least.

The suitability model is created by determining an average "suitability score" for each pixel (30x30m) in Knox county. This score incorporates a variety of factors that have been linked to urbanization rate as mentioned in the literature review section (Chapter 1).

Each pixel will be assigned a grade from 2-6 for each factor. Factors 1-7 are graded by the proximity of each pixel to the nearest location of the respective factor. Factors 8 and 9 are census-tract level data, so every pixel within a specific tract will have the same grade for these factors. Every proximity-driven factor uses the same standard of evaluation except for "Proximity to existing urban land", which assigns grades on 0.2 mile intervals (starting at 0.2) as opposed to 1 mile intervals. This was done because the existing urban land is widespread throughout the county and using the standard intervals would result in every cell receiving a grade of 6.

## Standard Urban Suitability Model

$$SI = F1L * F2M * F3N * F4O * F5P * F6Q * F7R * F8S * F9T$$

SI = suitability index

F1 = schools, L = grade for schools
F2 = hospitals, M = grade for hospitals
F3 = grocery stores, N = grade for grocery stores
F4 = parks, O = grade for parks
F5 = highways, P = grade for highways
F6 = waterways, Q = grade for waterways
F7 = pre-existing urban, R = grade for pre-existing urban
F8 = median household income, S = grade for median household income
F9 = median housing price, T = grade for median housing price

a = maximum grade (6)

$$F1L = (L-1)(a-1)$$
$$F2M = (M-1)(a-1)$$
$$F3N = (N-1)(a-1)$$
$$F4O = (O-1)(a-1)$$
$$F5P = (P-1)(a-1)$$
$$F6Q = (Q-1)(a-1)$$
$$F7R = (R-1)(a-1)$$
$$F8S = (S-1)(a-1)$$
$$F9T = (T-1)(a-1)$$

**Figure 4.1:** General suitability formula for determining which areas in the county are most suitable for urban development.

**Table 4.1:** Standard of Evaluation for buffer rings.

| **Standard of Evaluation (miles)** | **Suitability Grade** |
|---|---|
| $<1$ | 6 |
| $<2$ | 5 |
| $<3$ | 4 |
| $<4$ | 3 |
| $\geq 4$ | 2 |

*Assignment of suitability grades based on proximity to each potential explanatory factor.*

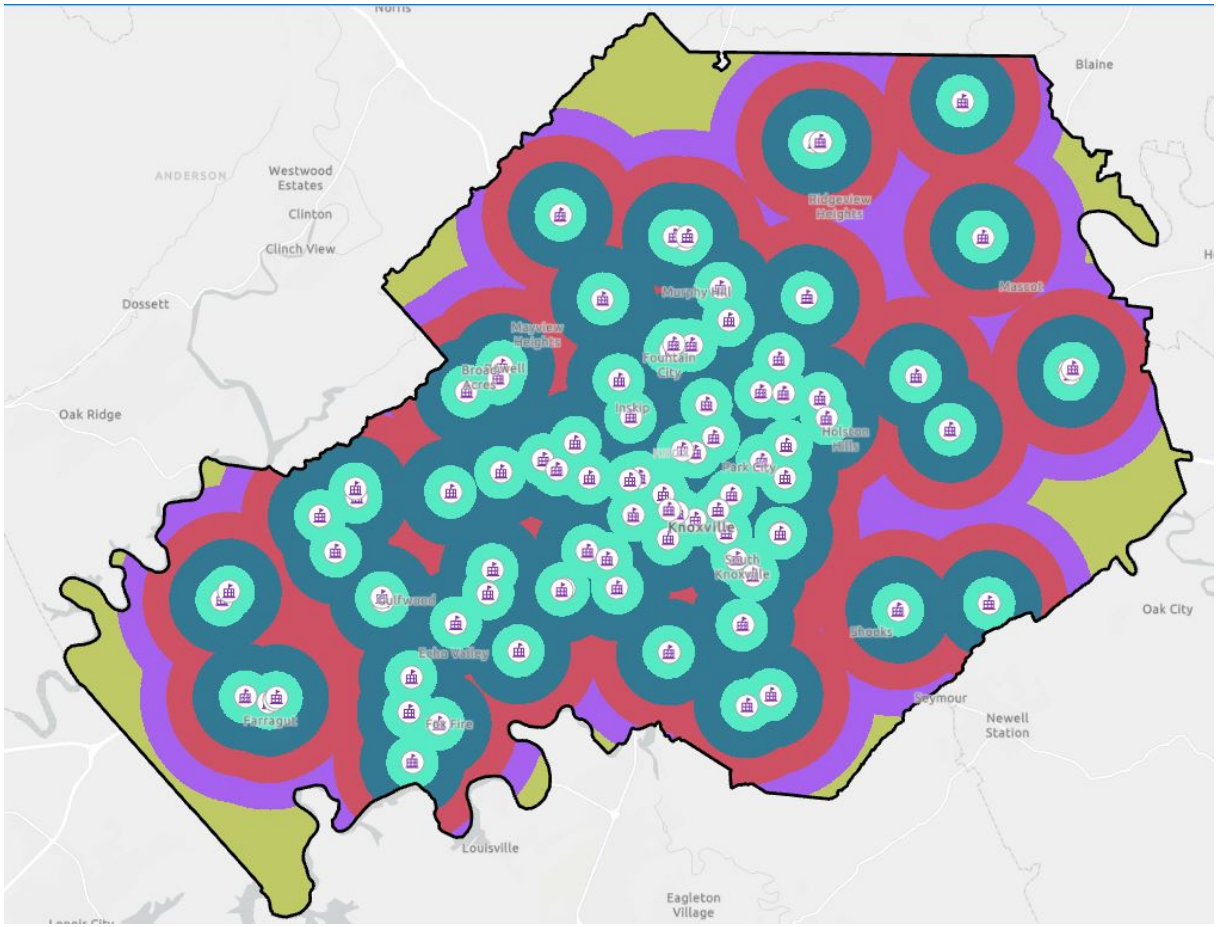


**Figure 4.2:** Example map showing the multiple buffer rings for each point. Every white symbol shown represents one school, and the colored rings surrounding the point represent the different "grades" associated with proximity to school.

It is important to note that this model assumes that each factor is equally responsible for increasing the area's suitability. While this may be true in some instances, it is not necessarily the case always. Therefore, this model was created in a way so that it can easily be changed to weight certain factors as more influential than others. Some example maps showing this flexibility can be seen in the results section.

### 4.2.2 Machine Learning Model

The second model used in this project is a machine learning model that can identify and rank the importance of certain parameters or inputs and predict the likelihood of urbanization in each part of the county. This model is based on the methodology employed by ESRI GIS analysts here: https://www.esri.com/arcgis-blog/products/arcgis-pro/analytics/a-deep-dive-into-predicting-urban-growth-using-arcgis-and-r/.

This model relies first on data from the 2001 and 2019 NLCD datasets, which were already utilized in Chapter 2. However, additional steps had to be taken to properly prepare the data for analysis. First, this particular model is primarily concerned with urban vs. non-urban classifications and identifying which cells have changed from 2001 to 2019. The NLCD datasets were reclassified so that all land classified as Developed (open space, low/medium/high intensity) was marked as "Urban", while all other land use types were considered Undeveloped. Then, the Detect Change geoprocessing tool was used to identify which areas changed in classification, either from undeveloped to developed or developed to undeveloped. It also noted which cells experienced no change. This data was subsequently converted to a point layer, which consisted of around 1.5 million points. Then, the buffer layers explained previously were used to append proximity and economic data to each cell using the Extract Multi Values tool. This created one comprehensive data layer that included a multitude of potential explanatory variables. The data appended includes distance from closest hospital, school, grocery store, roadway, waterway, existing urban land, and park. Similar to the suitability model, this data does not contain exact distances for each proximity category, but instead a grade (2-6) in which each grade is indicative that the cell falls within a certain range of distances. Additionally, median household income and slope percent were included as well.

Once this data layer was created, the .csv attribute table was exported from ArcGIS Pro and then imported into R studio to begin the statistical analysis. The initial part of the analysis is counting and inventorying the data from the attribute table. This data is classified by its grid code, which represents the change from 2001 to 2019. Each pixel was counted and grouped by its grid code. The next step was to omit areas that were developed in 2001 as the model is only concerned with changes from undeveloped to developed land. This included 593,726 cells that were urban for the start and end of the study year range. This also would omit any cells that were developed in 2001 and changed to undeveloped in 2019, but there were no cells that met this classification so this was unnecessary.

Then, this data was split into training and test datasets. The training data contained 85% of the entire dataset, while the testing data contained the other 15% to validate the training predictions and assess accuracy.

After counting the cells for both datasets, it became clear that the overall dataset was unbalanced as the amount of cells that were undeveloped in 2001 and still undeveloped in 2019 (no change in land use) were significantly higher than cells that changed to developed. As shown in Table 4.3, a significant amount of the data consists of cells that were undeveloped in 2001 and stayed undeveloped in 2019. This caused the model to predict that every cell would stay undeveloped in the future because so many cells stayed undeveloped in the training data. Additionally, this caused the model to report almost perfect accuracy despite never predicting an area would become urban.

One way to address this issue is by downsampling. Downsampling is the process of reducing one set of data to a size that is more comparable to the smaller dataset by randomly selecting cells to discard. Adhering to the methodology used in the ESRI project, the training data with a 0 grid code (undeveloped to undeveloped) was reduced to 3 times as much as the grid code 2 (undeveloped to developed) dataset.

Once the data was pre-processed, a random forest analysis was conducted through R Studio to create decision trees using the various inputs (eg. hospital distance, road distance, proximity to existing urban land). This analysis helped to value each factor by its importance in predicting development patterns.

Finally, the prediction results were imported back into ArcGIS Pro and appended to the point layer that was extracted at the beginning of the process. This was done using the Join Field tool, as the prediction results contained a common field with the original layer.

### 4.2.3 Restricting Suitability

After the model is created, the next step is to incorporate environmental sustainability into the future urbanization model. This is done not only to help conserve important environmental resources within the county, but also to reduce the city's vulnerability to flooding and heatwaves.

Therefore, areas with high NDVI values need to be protected, as they represent high vegetation and forested areas. Another way to incorporate this is through restricting suitability to certain land use types. For example, Chapter 2: Geospatial Analysis showed that agricultural and forested land were reduced the most as urban land expanded. This model can use the "forest" land use classification as well as high NDVI values as deterrents of suitability. This is done by assigning higher NDVI values to lower suitability grades, and assigning suitability grades to each land use type.

## 4.3 Results

### 4.3.1 Suitability

Figure 4.3 and Figure 4.4 show results of the standard suitability formula model. The gray area is the existing urban land, while the colored areas show the extent to which the model thinks the area is suitable for urban life. For clarification, any area on the map that has a non-gray color is a non-urban land use type. This model identifies southwest Knox county as most suitable for urban development, due to a combination of its proximity to various amenities and infrastructure. Unsurprisingly, the north and east sides of the county are still regarded as unsuitable. The only part of this region identified as "less suitable" are those pockets of land that are adjacent to existing urban land. Within West Knoxville specifically,

a few areas were identified as "most suitable". These areas are surrounded by existing urban land and area also very close to schools, major rivers, and lakes.

Figure 4.5 and Figure 4.6 both show the same model results with no urban overlay. These maps showcase the model's ability to correctly identify some areas as suitable, as already urban areas are marked as suitable. This can be seen by comparing Figure 4.3 with Figure 4.5, where the gray area represents urban in the first map and that existing urban area is mostly considered (visible in the second map). Figure 4.6 shows how the suitability identification can change, or become more restricted, as more data is added to the model. The NDVI values were added the same way as other variables, except high NDVI values lowered the overall suitability score. The suitable area in this map is much smaller than the standard map and more focused on a few main nodes: the downtown area and the rapidly-developing west knoxville area. In a way, this model is suggesting that Knoxville needs to focus on densification more than expansion, as much of the surrounding area became restricted when the NDVI values were added. One way that this model could be improved is by adjusting the grading for the NDVI values. Perhaps in the future this model would interpret low NDVI values as very favorable or suitable for urban development, so that the restricted suitability map could actually have certain areas that increase in suitability once the restrictions are factored in.

As mentioned in Chapter 3, NDVI values are not necessarily the most accurate way to assess a region's ecological importance. Another way to restrict the model would be to overlay land use data. If planners are concerned about maintaining an existing land use type, that type could be converted to a polygon and the whole shape could be considered "unsuitable" by the model. An ecological need for this would be if there is an endangered bird species that onyl inhabits one specific kind of forest. The model could assign very low suitability grades to areas that have this type of forest. This gives planners a tool to incorporate more data into the decision making process for future zoning and development plans.

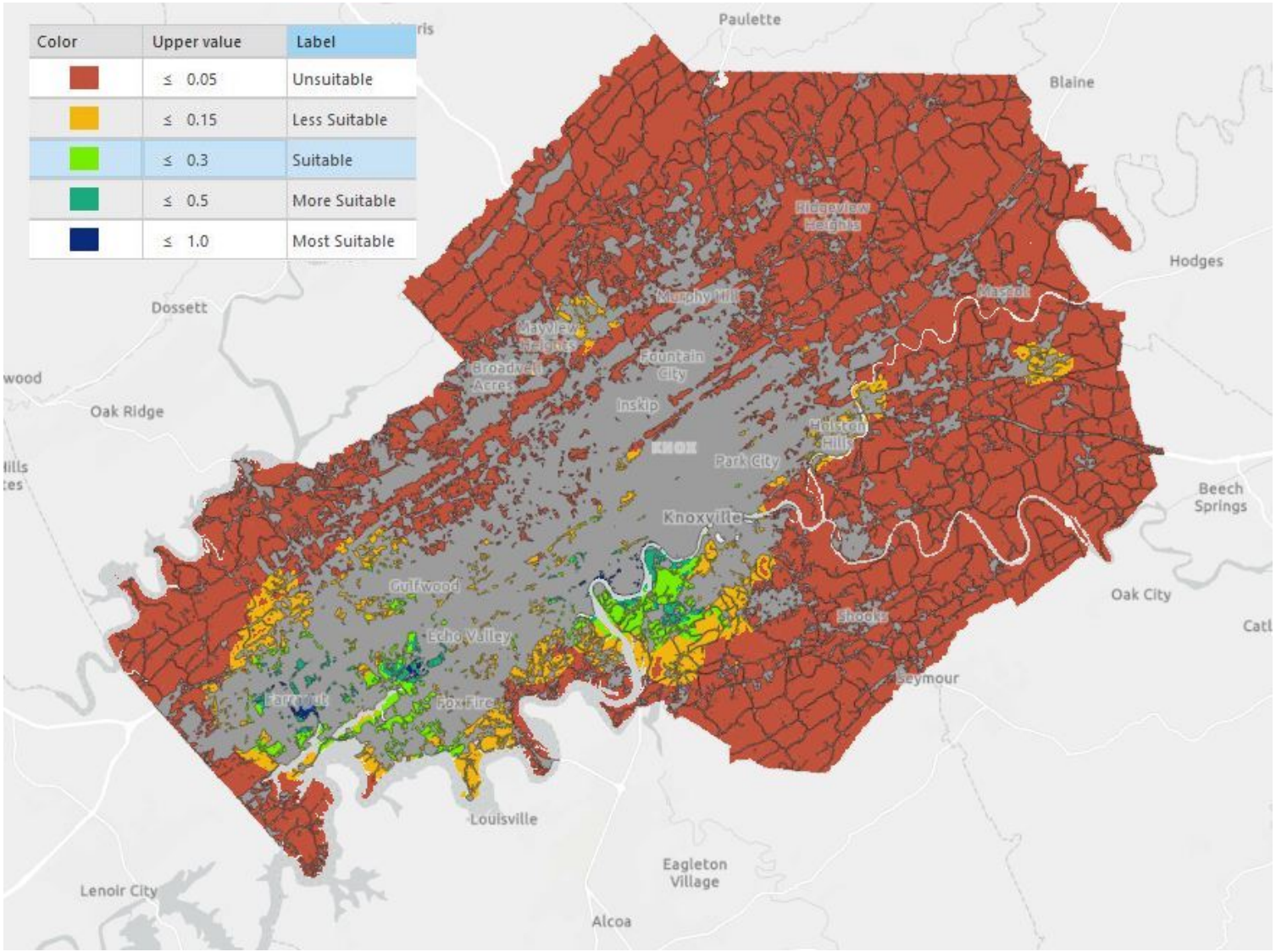


**Figure 4.3:** Standard Suitability Map for Knox County. Grey area represents existing urban land.

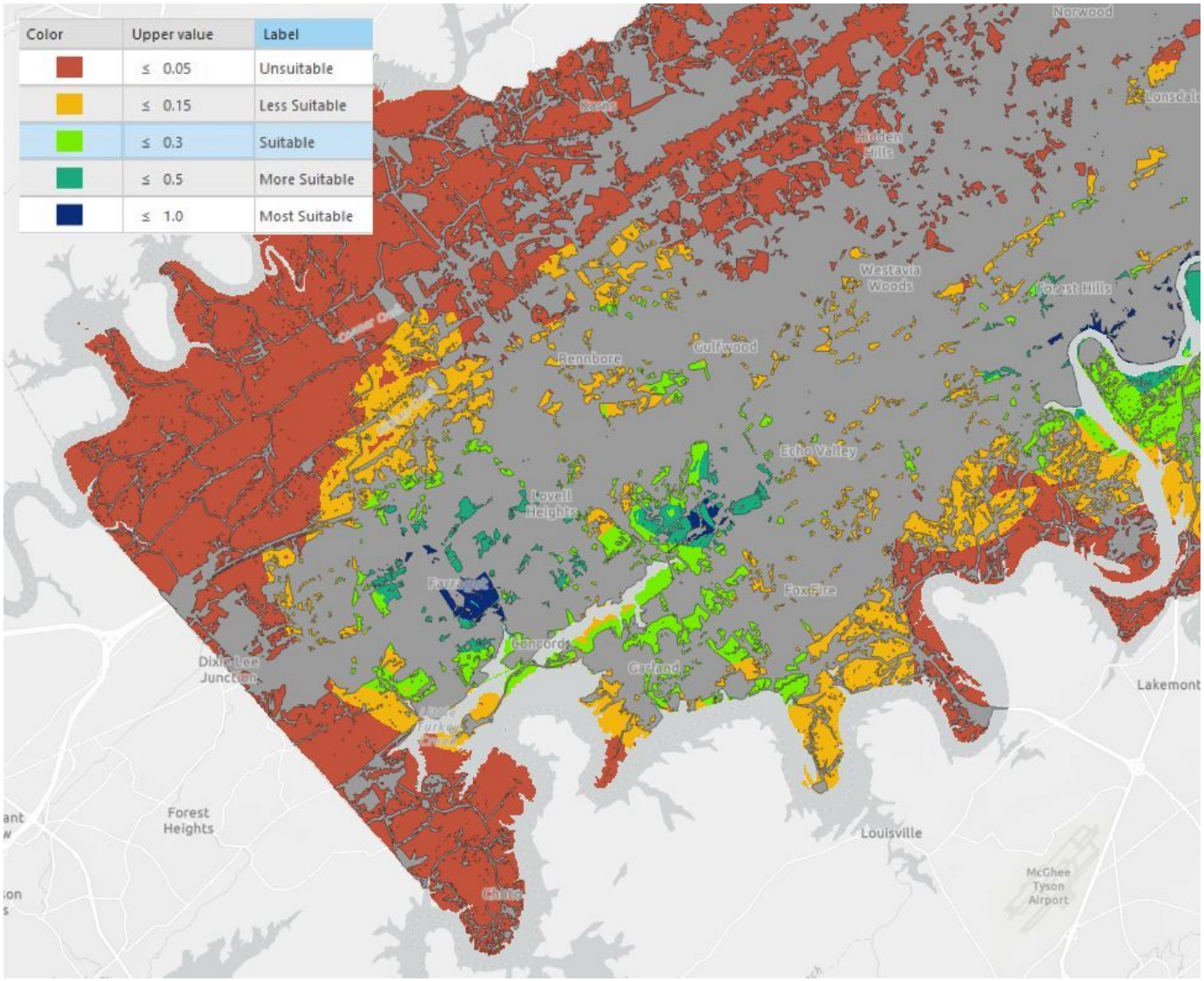


**Figure 4.4:** Suitability Map for West Knox County. Grey area represents existing urban land.

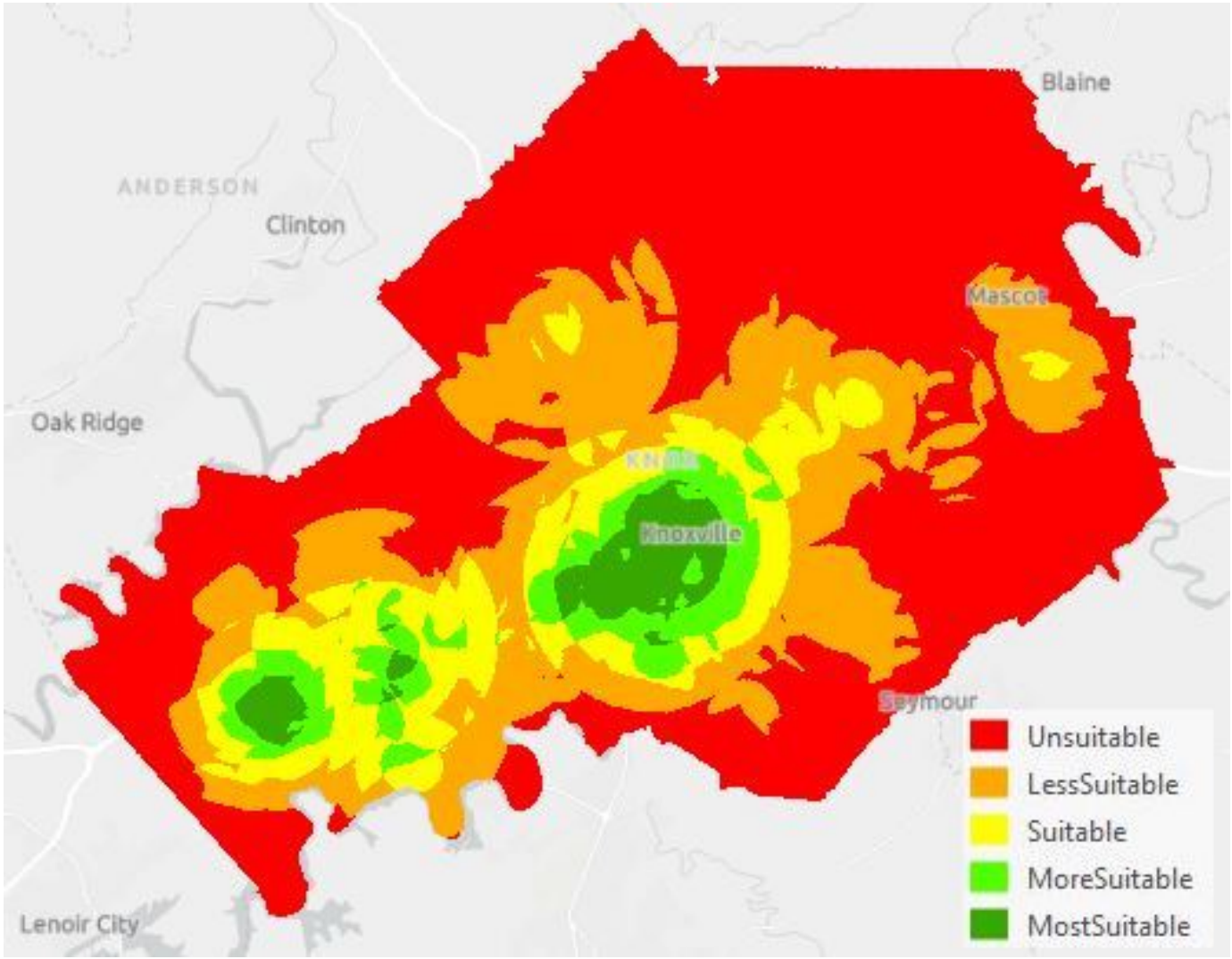


**Figure 4.5:** Overall Suitability without urban extent overlay.

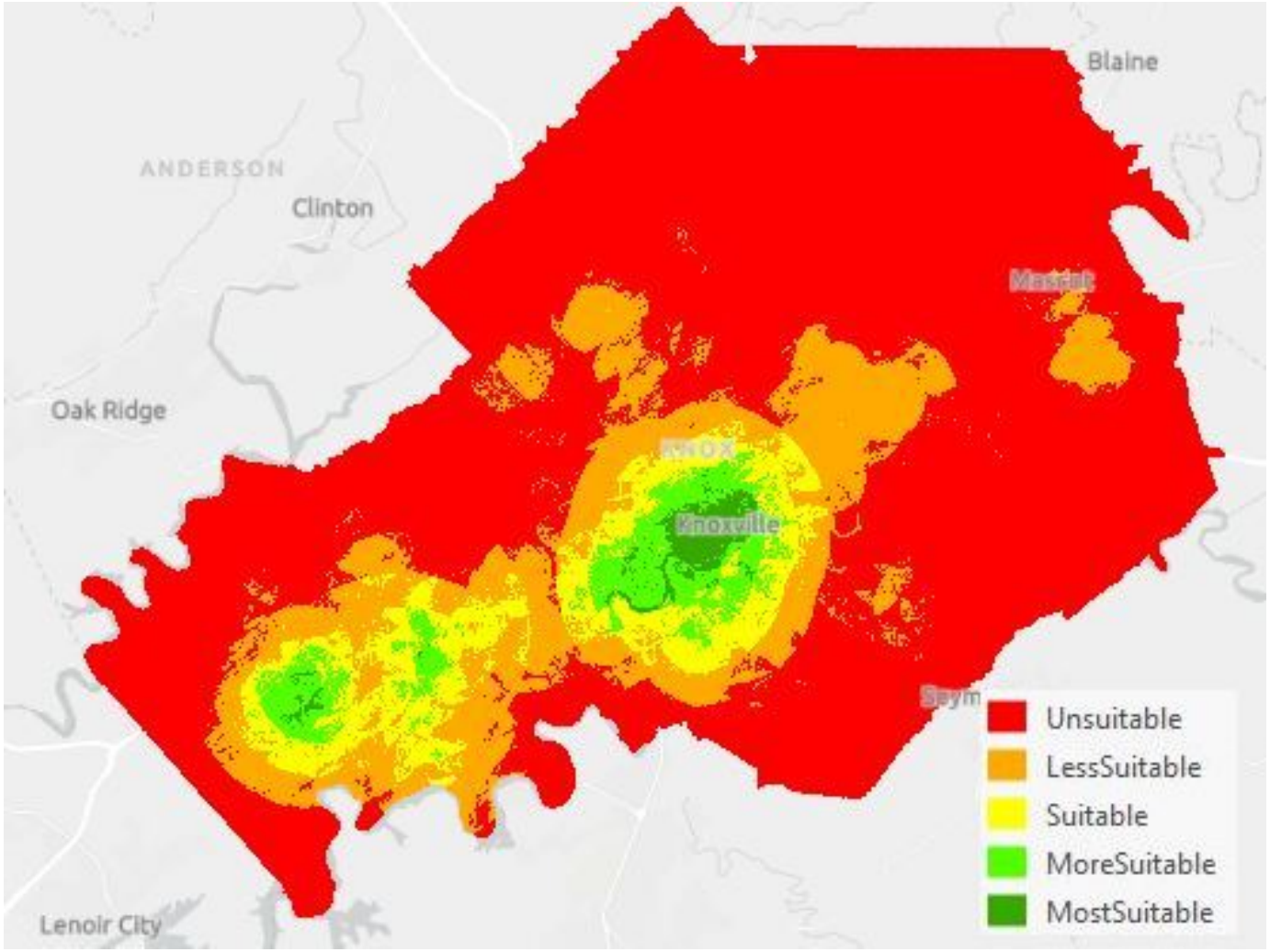


**Figure 4.6:** Suitability Map for Knox County, restricted by high NDVI values.

### 4.3.2 Machine Learning Model

The model's ranking of variable importance to predict development can be seen in Figure 4.7. Surprisingly, median household income was ranked as the most important variable. Median household income was a tract-level dataset and therefore did not have individual unique values for each cell within a tract. This shows how important economic factors are for driving urbanization. Proximity to grocery stores and percent slope were the next most important variables. Proximity to urban land was considered the least important, even though both models (suitability and ML) predict urbanization on the outskirts of the developed land. A reason for this is the urban extent was so large, most of the surrounding area was close enough to make the proximity in the closest range. Because all urban proximity values were very close together (generally a 4 or a 5 for the grade), it was not seen as an influential factor by the model. This could also be the case for the grocery store proximity being so influential; grocery stores are widespread and no area in the county is more than a mile or two from the closest one.

Figure 4.8 shows the predicted areas of urbanization (red) for Census Tract 57.09 (blue area), which is one of the fastest-urbanizing tracts. It is important to note that the training data used for this model was from 2019, while the basemap seen in the map is imagery from 2023. The discrepancy in these years allow for easy verification that the model is correctly predicting some areas for development. In Figure 4.8, some of the predicted areas of urbanization have already been developed since 2019. Another advantage of this model is the resolution at which it can predict development. Because each cell has data for every input variable, each cell receives its own prediction from the model. Therefore, this model can predict urbanization down to a small scale, which is crucial when planning development areas around protected land. Unfortunately, this mapping process could only be done on a small area like this census tract. The processing time is much too great to create this map for the whole county. This is mostly due to the large quantity of data points (1.5 million cells), which increase the time for the field join in ArcGIS Pro.

**Table 4.2:** Change in developed status from 2001-2019

| Grid Code | Change from 2001 to 2019 | Cell Count |
|---|---|---|
| 0 | Undeveloped to Undeveloped | 857257 |
| 1 | Developed to Developed | 593726 |
| 2 | Undeveloped to Developed | 62307 |
| 3 | Developed to Undeveloped | 0 |

*Grid codes for each type of change in land usage from 2001-2019. This also shows the count of 30x30m cells in the county that are associated with each change.*

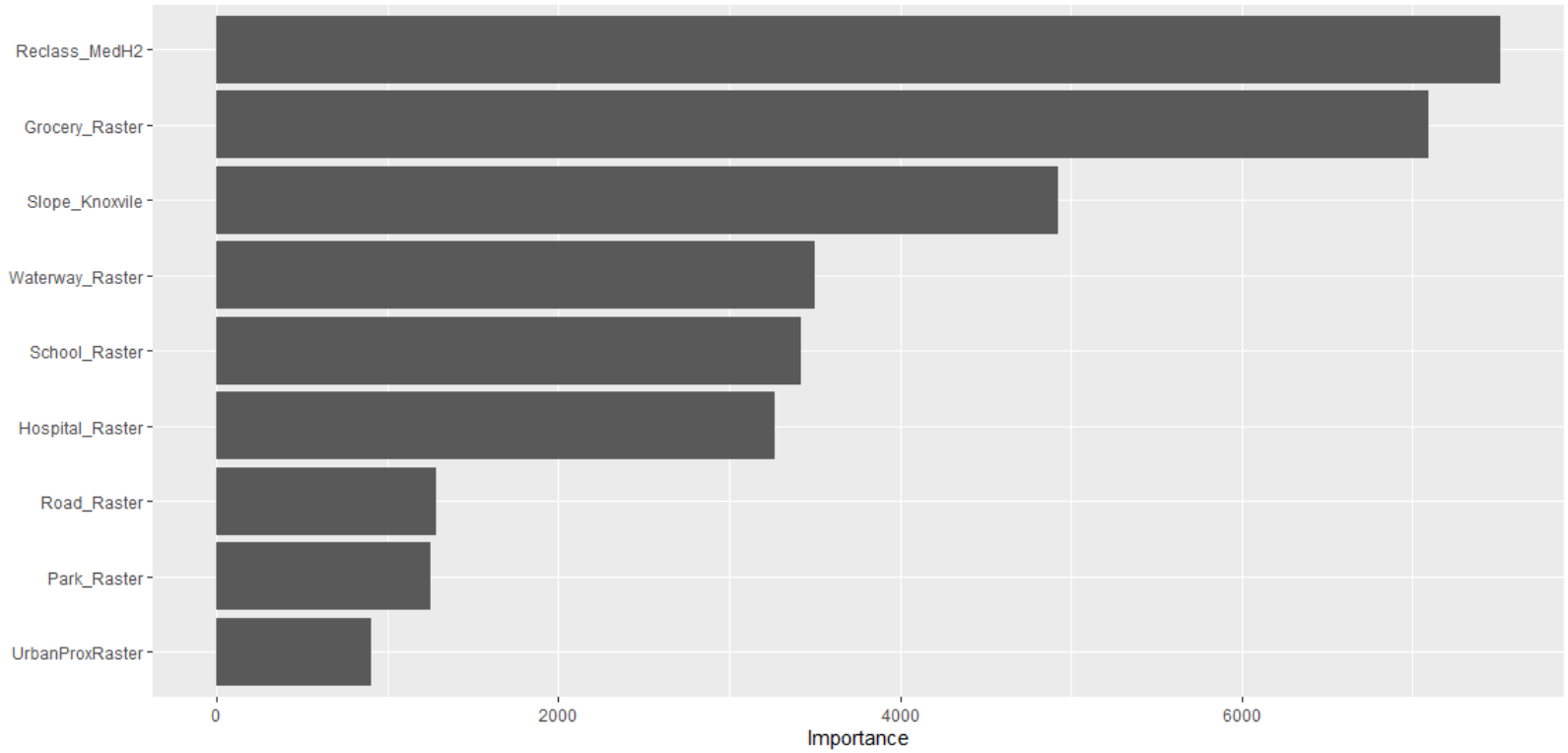


**Figure 4.7:** Ranking of variable importance according to the ML model.

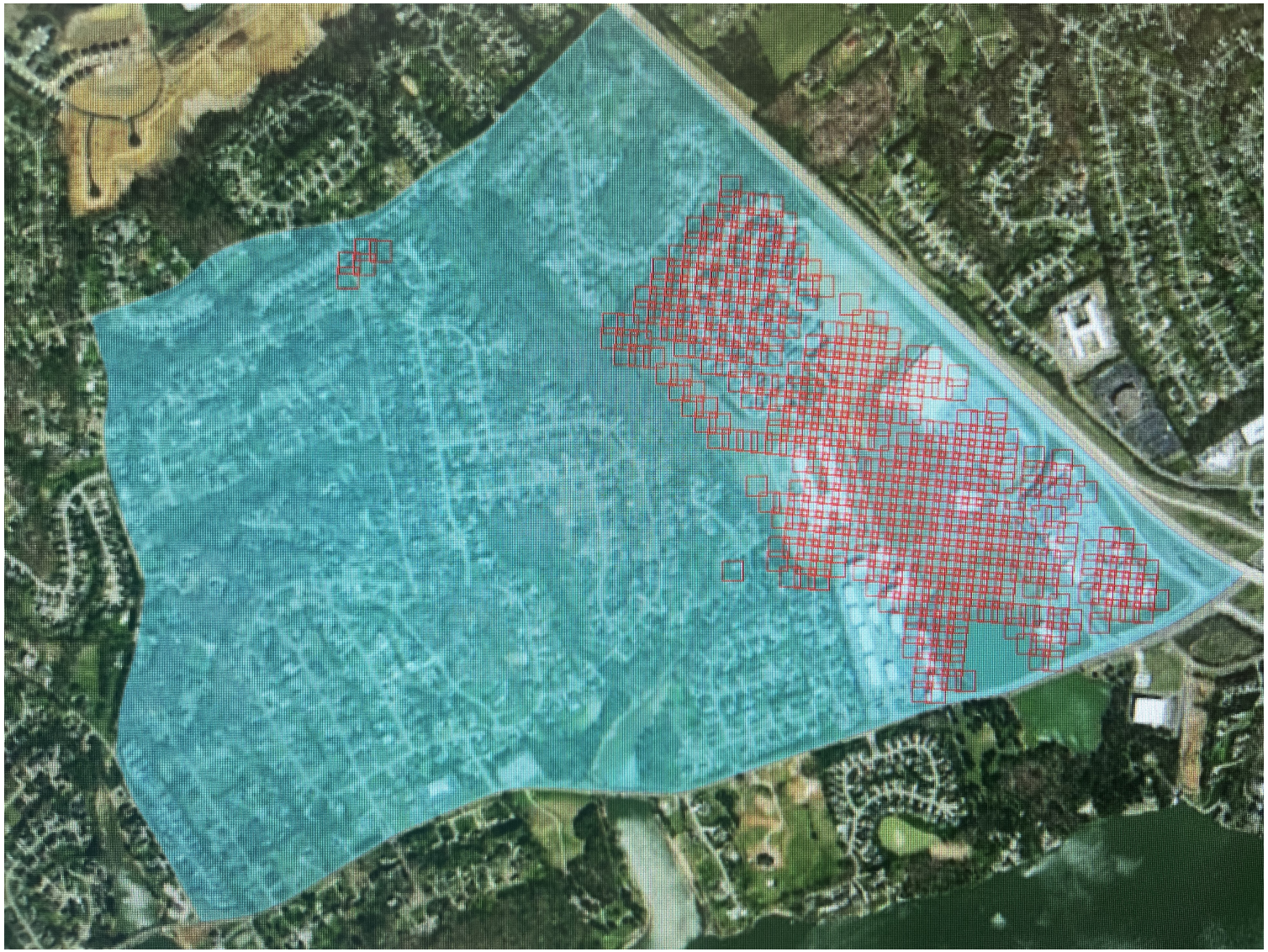

**Figure 4.8:** ML model prediction for development in Census Tract 57.09.

**Table 4.3:** Training and Test Data Cell Count

| Grid Code | Train Data | Test Data |
|---|---|---|
| **0** | 687766 | 121364 |
| **2** | 52599 | 9289 |

*This table shows the count of undeveloped and developed cells that the model used with 85 percent of the data used as training data.*

**Table 4.4:** Prediction Matrix Results

| | **True 0** | **True 2** |
|---|---|---|
| **Predicted 0** | 121364 | 9289 |
| **Predicted 2** | 0 | 0 |

*Results from first prediction. This table shows the count of undeveloped cells that the model predicted would stay undeveloped (grid code = 0) and the count that would change from undeveloped to developed (grid code = 2).*

**Table 4.5:** Accuracy of First Prediction

| | **Cell Count** | **Accuracy** |
|---|---|---|
| **Grid Code = 0** | 687766 | 1.00 |
| **Grid Code = 2** | 52599 | 0.00 |
| **Overall** | 740375 | 0.922 |

*Prediction Accuracy after first attempt.*

**Table 4.6:** Prediction Matrix Results after Downsampling

| | **True 0** | **True 2** |
|---|---|---|
| **Predicted 0** | 113883 | 6311 |
| **Predicted 2** | 7481 | 2978 |

*Results from prediction after downsampling. This table shows the count of undeveloped cells that the model predicted would stay undeveloped (grid code = 0) and the count that would change from undeveloped to developed (grid code = 2).*

**Table 4.7:** Accuracy of Training Data after Downsampling

| | **Cell Count** | **Accuracy** |
|---|---|---|
| **Grid Code = 0** | 157797 | 0.925 |
| **Grid Code = 2** | 52599 | 0.562 |
| **Overall** | 210396 | 0.899 |

*Amount of cells for each grid code and accuracy of prediction after downsampling.*

## 4.4 Discussion and Conclusion

When comparing the two models, it is clear that the suitability model can be used to identify areas suitable for development. However, there are some concerns with the validity of the data. Firstly, each input is weighted equally. It is rarely the case that each external factor driving urbanization has the same amount of influence as all of the other ones. This means that the model underestimated the influence of the most important factors and overestimated the importance of the less-relevant factors. Additionally, this model did not take into account all of the inputs that could be affecting urbanization, like zoning policy, quality of schools, drive time to downtown area, etc. The model is only as good as the data that is used, so incorporating every possible factor that could be driving urbanization would help to improve the model if its relative impact could be estimated.

On the other hand, the machine learning model used the data in a more rigorous and thorough fashion by omitting incomplete data and attempting to create a more balanced dataset through downsampling. The model has clear advantages over the suitability model: more detailed resolution, accuracy assessments, and variable importance ranking. The use of decision trees also allowed the model to understand which variables were actually best at predicting urbanization. This is a huge advantage over the suitability model, which either weighs each variable equally or can be adjusted if there is sufficient data to support weighing the variables differently. However, the sheer amount of data points needed to conduct analysis (1.5 million points for the county) makes this type of analysis much more difficult and slow to perform.

One potential solution that addresses issues with both models is to use the ML model to rank the importance of each input. Then, when the importance is quantified, coefficients can be added to the suitability formula to more accurately weight each factor. This way, one can eliminate the processing power limitations of the ML model and the unknowns of variable importance issue from the suitability model. These maps also do not need to include data points for land that was already urban in 2001, so these cells can be omitted to speed up the processing time.

# Chapter 5

# Conclusion

## 5.1 Importance of Research Findings

The findings from this project show that urbanization is a complex, multi-faceted process that is difficult to fully quantify. Knox county is experiencing a significant reduction in forest and agricultural land as a result of urban expansion. The environmental impact section reinforces the concept that city planners need to be aware of how urbanization processes affect the city's vulnerability to heat and flooding. The urbanization modeling shows that these envrironmental impacts are not going to lessen, as Knoxville will continue to grow in population and expand in area. The comparison of the urbanization models shows that models all have advantages and disadvantages and the model is only as good as the data.

## 5.2 Future Research Directions

There are many possible future directions to go with this research. Planners can use these models when creating scenario-based maps for future urbanization patterns. Knoxville, through the AdvanceKnox program, is creating these maps and polling stakeholders on their preferences. Each scenario is rated on its emphasis on education, recreating, natural land, and more. The suitability model can identify areas where focusing development would be the easiest, and then changing the factor weights for the suitability formula can pinpoint areas that meet the requirements for proximity needs.

Another potential follow-up to this project could be to optimize the suitability model for individual homeowner usage. For example, this model could be adapted into a web app. This app would assist families who want to move to Knoxville because they could identify and rank the aspects of home location that matter the most to them, like school or public park proximity. The suitability model could identify suitable areas that meet their requirements.

# Bibliography


[1] Saad Saleem Bhatti et al. “A multi-scale modeling approach for simulating urbanization in a metropolitan region”. In: *Habitat International* 50 (2015), pp. 354–365 (cit. on p. 13).

[2] Charles Steven Brown. “The Changing Landscape: The Urbanization of Knox County, Tennessee”. In: (2007) (cit. on p. 14).

[3] Neil Debbage and J Marshall Shepherd. “The urban heat island effect and city contiguity”. In: *Computers, Environment and Urban Systems* 54 (2015), pp. 181–194 (cit. on p. 10).

[4] Wei Ding et al. “A Review of Flood Risk in China during 1950–2019: Urbanization, Socioeconomic Impact Trends and Flood Risk Management”. In: *Water* 14.20 (2022), p. 3246 (cit. on p. 12).

[5] Shiqiang Du et al. “Quantifying the impact of impervious surface location on flood peak discharge in urban areas”. In: *Natural Hazards* 76 (2015), pp. 1457–1471 (cit. on pp. 11, 12, 46).

[6] Michel ED Chaves, Michelle CA Picoli, and Ieda D. Sanches. “Recent applications of Landsat 8/OLI and Sentinel-2/MSI for land use and land cover mapping: A systematic review”. In: *Remote Sensing* 12.18 (2020), p. 3062 (cit. on p. 7).

[7] Felix Eigenbrod et al. “The impact of projected increases in urbanization on ecosystem services”. In: *Proceedings of the Royal Society B: Biological Sciences* 278.1722 (2011), pp. 3201–3208 (cit. on pp. 12, 46).

[8] Evidence Chinedu Enoguanbhor et al. “Land cover change in the Abuja City-Region, Nigeria: Integrating GIS and remotely sensed data to support land use planning”. In: *Sustainability* 11.5 (2019), p. 1313 (cit. on p. 8).

[9] Yongqiang Fang et al. “Chinese built-up land in floodplains moving closer to freshwaters”. In: *International Journal of Disaster Risk Science* 12 (2021), pp. 355–366 (cit. on p. 11).

[10] Cristina Gómez, Joanne C White, and Michael A Wulder. “Optical remotely sensed time series data for land cover classification: A review”. In: *ISPRS Journal of photogrammetry and Remote Sensing* 116 (2016), pp. 55–72 (cit. on p. 7).

[11] Michael Gross. *The urbanisation of our species*. 2016 (cit. on p. 2).

[12] Chaolin Gu. “Urbanization: Processes and driving forces”. In: *Science China Earth Sciences* 62 (2019), pp. 1351–1360 (cit. on pp. 2–5).

[13] J Vernon Henderson et al. “Urbanization and economic development”. In: *Annals of economics and finance* 4 (2003), pp. 275–342 (cit. on pp. 1, 2).

[14] Somik Vinay Lall, J Vernon Henderson, and Anthony J Venables. *Africa’s cities: Opening doors to the world*. World Bank Publications, 2017 (cit. on p. 8).

[15] Eric FMDA Lambin, Mark DA Rounsevell, and Helmut J Geist. “Are agricultural land-use models able to predict changes in land-use intensity?” In: *Agriculture, Ecosystems & Environment* 82.1-3 (2000), pp. 321–331 (cit. on pp. 5–7).

[16] Hui Li et al. “Sponge city construction in China: A survey of the challenges and opportunities”. In: *Water* 9.9 (2017), p. 594 (cit. on p. 12).

[17] Wei Liang and Ming Yang. “Urbanization, economic growth and environmental pollution: Evidence from China”. In: *Sustainable Computing: Informatics and Systems* 21 (2019), pp. 1–9 (cit. on p. 2).

[18] Robert E Lucas Jr. “Life earnings and rural-urban migration”. In: *Journal of political economy* 112.S1 (2004), S29–S59 (cit. on p. 2).

[19] Joan Marull et al. “Emerging megaregions: A new spatial scale to explore urban sustainability”. In: *Land Use Policy* 34 (2013), pp. 353–366 (cit. on p. 5).

[20] Walaa Abou El-Haggag Mehanna and Wesam Abou El-Haggag Mehanna. "Urban renewal for traditional commercial streets at the historical centers of cities". In: *Alexandria Engineering Journal* 58.4 (2019), pp. 1127–1143 (cit. on p. 2).

[21] Binaya Kumar Mishra et al. "Implications of urban expansion on land use and land cover: Towards sustainable development of Mega Manila, Philippines". In: *GeoJournal* 86 (2021), pp. 927–942 (cit. on pp. 8, 42).

[22] Bakare Ganiyu Olalekan. "Urbanization, urban poverty, slum and sustainable urban development in Nigerian cities: Challenges and opportunities". In: *Developing Country Studies* 4.18 (2014), pp. 13–19 (cit. on p. 3).

[23] Henry L Vega and Leo Penne. "Governance and institutions of transportation investments in US mega-regions". In: *Transport* 23.3 (2008), pp. 279–286 (cit. on p. 5).

[24] Yuchao Yan et al. "Assessing the impacts of urban sprawl on net primary productivity using fusion of Landsat and MODIS data". In: *Science of the Total Environment* 613 (2018), pp. 1417–1429 (cit. on p. 13).

[25] Xuchao Yang et al. "Environmental consequences of rapid urbanization in Zhejiang province, East China". In: *International journal of environmental research and public health* 11.7 (2014), pp. 7045–7059 (cit. on p. 9).

[26] Siqi Zheng et al. "The birth of edge cities in China: Measuring the effects of industrial parks policy". In: *Journal of Urban Economics* 100 (2017), pp. 80–103 (cit. on p. 5).

[27] Sharon Zukin. "Gentrification: culture and capital in the urban core". In: *Annual review of sociology* 13.1 (1987), pp. 129–147 (cit. on p. 4).

# Vita

Duncan Paul McCurry was born in Knoxville, Tennessee, on June 18th, 1997. Duncan graduated from Ravenwood High School in Brentwood, Tennessee, in 2015, and began attending the University of Tennessee, Knoxville later that year. Duncan finished his undergraduate studies in 2019 with a degree in Environmental and Soil Science. In 2021, Duncan began his pursuit of a Masters Degree at the University of Tennessee, Knoxville. Duncan graduated with his Masters of Science in the spring of 2023.