Urban Canopy Equity: Assessing Socioeconomic Disparities in Tree Cover and Quantifying Urban Forest Changes in Baltimore

Maya Clark

B.S. Geography and Environmental Planning, Towson University

CLA Summer Research Grant 2025

Faculty Mentor: Dr. Carter Wang

September 30, 2025

Abstract

This study analyzes the long-term trends of tree canopy health and distribution in Baltimore city while also exploring if sociodemographic factors have any influence on these trends. The project quantifies spatiotemporal disparities by analyzing percent tree canopy cover (PTC) and vegetation health (Normalized Difference Vegetation Index or NDVI) from 2011 to 2021 using Landsat imagery, U.S. Census data, and the Tree Canopy Cover (TCC) dataset from the U.S. Forest Service. Spatial statistical methods, including the Mann-Kendall test, Global Moran's *I*, and Getis-Ord *Gi** analysis were used to identify temporal trends and spatial patterns, while stepwise regression determined the relationship between PTC and sociodemographic factors. Findings revealed a 17% decline in Baltimore's PTC over the 11-year period, with statistically significant canopy loss concentrated in the northwestern quadrant of the city. The regression analysis revealed a statistically significant negative relationship between PTC and both population density and the percentage of vacant housing units. The disparity was particularly pronounced in more affluent neighborhoods, where higher levels of tree canopy were observed, indicating that socioeconomic status may be a key factor in the inequitable distribution of green spaces. These results provide insights to inform targeted reforestation strategies from local stakeholders in efforts to ensure the longevity of Baltimore's urban tree canopy for all residents.

1. Introduction

As urbanization continues to rise globally, with more than half the world's population now living in cities, the need for urban green spaces is becoming more critical (Kabisch, 2011). Urban trees offer numerous ecological and socio-economic benefits, including cooling, cost savings, and improved public health outcomes, and should be accessible to all urban residents regardless of their sociodemographic status (Ulmer et al., 2016; Yin et al., 2024). However, Baltimore's history is marked by persistent inequalities, rooted in factors like income and race, which continue to influence the city's landscape (Brown et al., 2019). A significant manifestation of this is the disparity in tree canopy cover.

In 2009, a comprehensive study conducted by Grove and O'Neil-Dunne (2009) analyzed Baltimore's tree canopy and modeled its realistic potential based on the city's existing land cover. The study's findings, using 2007 data, reported that Baltimore's tree canopy stood at 24.7%, while a significant portion of the city was comprised of hard surfaces (43%) and grassland (19%) (Grove & O'Neil-Dunne, 2009). Based on these results, the city of Baltimore set an ambitious goal to increase its tree canopy to 40% by 2037. This study is designed to assess if Baltimore is on track to achieve this goal and will do so by building upon the previous methods developed by Grove and O'Neil-Dunne (2009) to provide a more robust and current assessment of the city's progress.

Research has consistently shown that marginalized communities, often the direct result of historical redlining practices, experience lower percentages of tree canopy and are frequently excluded from urban reforestation initiatives (Locke et al., 2021). Due to Baltimore's history of inequity and redlining, an additional objective was introduced: to assess whether disparities in tree canopy distribution exist across the city based on sociodemographic factors such as race, ethnicity, income, and housing. By examining the relationships between sociodemographic characteristics and tree cover changes, the project aims to reveal environmental justice disparities that may exist and provide insight into where equitable reforestation efforts could take place within the city.

This project's findings will empower Baltimore's local communities, government agencies, and non-profit organizations to strategically focus tree canopy restoration efforts in areas experiencing or projected to experience significant canopy loss. This targeted approach will not only promote long-term ecological sustainability but also ensure a more equitable distribution of urban forest benefits, contributing to the overall health and equity of Baltimore's green spaces.

2. Study Area

Baltimore City is a major urban center located in the state of Maryland, covering an area of 80.95 square miles. Its position as a key port along the Chesapeake Bay and proximity to the Mid-Atlantic coast gives it a temperate four-season climate, with average daily temperatures ranging from 37°F in the winter to 77°F in the summer. As of 2021, the conclusion of the study period, Baltimore's population was 565,239, reflecting a 3.5% decline since 2010. The racial demographics are predominantly Black (60%) and White (27.4%), with smaller but growing populations of Hispanic/Latino (7.9%), Asian (2.5%), and other races (2.2%). While the city's median household income is $59,623, significant economic disparities persist. A staggering 20.1% of Baltimore's population lived at or below the poverty line, a figure substantially higher than the national average of 11.6%. These disparities are particularly pronounced along racial lines, with data from the 2021 American Community Survey indicating that the poverty rate for Black residents is 22.4%, compared to 14.2% for White residents. Historically, Baltimore's neighborhoods were significantly shaped by redlining, a discriminatory practice from the mid-20th century where federal agencies rated neighborhoods based on racial composition. This led to a systemic lack of investment in communities inhabited by Black residents, which is still visible today particularly in West and East Baltimore neighborhoods. These historically redlined areas often correspond to neighborhoods with less green space and lower tree canopy cover (Jung et al., 2024), directly contributing to the environmental disparities that exist in the city today.

3. Data and Methods

3.1 Data

3.1.1 PTC

Tree canopy cover change was analyzed using the annual Percent Tree Canopy Cover (TCC) dataset, produced by the U.S. Forest Service Geospatial Technology and Application Center (USDA Forest Service, 2023 and Housman *et al*. 2023). This dataset, downloaded from the Multi-Resolution Land Characteristics Consortium website (mrlc.gov), quantifies percent tree canopy density across the contiguous United States from 2011 to 2021 at a 30-meter spatial resolution. It is derived from the National Land Cover Database's Land Use/Land Cover change classifications and includes a layer specifying the standard error for each pixel.

3.1.2 Landsat Imagery

Tree health was assessed using NDVI (Normalized Difference Vegetation Index) derived from Landsat imagery, a dataset from the longest-running Earth observation program that has monitored our planet's environment and resources since 1972. For this study Landsat 5 and Landsat 8 data were utilized from the years 2011 to 2021 to capture tree health during peak growth periods, June to September respectively. The USGS Analysis Ready Data (ARD) product was downloaded from USGS EarthExplorer (earthexplorer.usgs.gov), which provides standardized, pre-processed imagery that is geometrically and radiometrically corrected (Earth Resources Observation and Science Center, 2021). This allowed for analysis of tree health over time with a standard baseline and metric for change.

3.1.3 U.S. Census Data

To analyze the spatial distribution of sociodemographic characteristics in relation to the conclusion of the study period in 2021, American Community Survey 5-year estimates from for 2021 were utilized and produced by the U.S. Census Bureau. Because this study is concerned with equitable reforestation efforts in the future, only the most recent sociodemographic data was necessary for analysis. Additionally, four data profiles were examined for all tracts within Baltimore city. Variables were chosen that describe the characteristics of both Baltimore's population and built environment and were analyzed for their relationship with PTC and NDVI. These variables have been chosen because in the literature they have demonstrated strong correlations with urban tree canopy over other sociodemographic variables (Schwarz et al., 2015). See Table 1 for a complete list of selected variables.

Table 1 – Variables selected from the American Community Survey 2021 5-Year Estimates for regression analysis.

Variable Name
Percent 25+ with a bachelor's degree
Median family income (dollars)
Percent employed
Percent vacant housing units
Percent built 2010 or later
Median age (years)
Percent White
Percent Hispanic or Latino
Population Density (Persons per square mile)

3.2 Methods

3.2.1 PTC Trends

The dataset used had already undergone significant pre-processing from the U.S. Forest Service's Geospatial Technology and Application Center, which allowed for immediate analysis. To determine the spatiotemporal trends in PTC, I employed zonal statistics within ArcGIS Pro (Version 3.3.0; ESRI, Redlands, CA) to calculate the mean PTC value for each census tract within Baltimore for each year in the dataset using the TCC dataset. A single mean annual value for every year from 2011 to 2021 was exported to an Excel spreadsheet for further statistical analysis in Minitab (See section 3.2.3).

3.2.2 NDVI Trends

To determine trends in tree health, the Normalized Difference Vegetation Index (NDVI) was calculated for all available Landsat images in ArcGIS Pro. NDVI is an index quantifying vegetation health and density by analyzing red and near-infrared light reflectance from surface features originally developed in 1974 (Rouse et al., 1974). NDVI is formulated as:

$$\frac{\rho NIR - \rho Red}{\rho NIR + \rho Red}$$

where ρ_{NIR} represents surface reflectance of the near-infrared light and ρ_{Red} represents surface reflectance of the red light. NDVI value ranges from -1 and 1, with a value of 0.6 and above representing healthy vegetation. Historically, it's been crucial for measuring tree health and vegetation biomass, serving as a key tool in many urban tree canopy health assessments (Xiao *et al.*, 2005).

The initial data processing involved using the quality band included in the NDVI data bundle to remove cloud-contaminated pixels with the Raster Calculator tool in ArcGIS Pro. After removing cloud-contaminated pixels, the NDVI image was masked with the PTC layer to extract tree pixels. The resulting image contains tree pixels and the corresponding NDVI values only. Zonal statistics were then employed using the Baltimore City census tracts shapefile as the zones to summarize mean NDVI values for each census tract. This process was carried out for multiple images annually, though varying dependent upon data availability and factors such as excess cloud coverage and Landsat's temporal resolution of 16 days, which provides around 2 images per month. Annual data was aggregated into a single mean value, and mean values from 2011 to 2021 were exported into an Excel spreadsheet for further statistical analysis (See section 3.2.3).

3.2.3 Mann Kendall Test

The Mann-Kendall statistical test is a non-parametric method used to detect monotonic trends in time series data. This test is particularly well-suited for diverse environmental datasets, including those with outliers, as it does not assume a specific data distribution. The Mann-Kendall test was favored over linear regression because it is robust against outliers, and it avoids the assumption that the change over time must follow a straight line This technique involves assessing the presence and direction of a trend by comparing the relative magnitudes of data points. It calculates the Sen's slope, where a positive value indicates an increasing trend and a negative value suggests a decreasing trend. The statistical significance of these trends is determined by the *p*-value. This test was carried out on both PTC and NDVI and the results of this analysis produced a Sen's slope value for every census tract in Baltimore City demonstrating the trend of increase or decrease. Additionally, absolute and relative changes for both PTC and NDVI were calculated using the mean from 2011 and 2021.

3.2.4 Global Moran's I

To reveal any potential patterns in trend values of PTC and NDVI, Global Moran's *I* (GMI) spatial autocorrelation was employed in ArcGIS Pro. It is an inferential statistic used to measure spatial autocorrelation, indicating the degree to which a feature's value is related to its neighbors (Moran, P. A. P., 1950). The index ranges from -1.0 to +1.0, where positive values signify clustering of similar values (positive spatial autocorrelation), negative values indicate dispersal of dissimilar values (negative spatial autocorrelation), and values near zero suggest a random spatial pattern. Accompanying the index are a *z*-score and *p*-value, which collectively assess the statistical significance of the observed pattern, allowing researchers to determine if the spatial arrangement is a result of true clustering/dispersion rather than random chance.

3.2.5 Getis-Ord Gi Hot Spot Analysis*

To identify hot spots of spatially significant trend values, the Getis-Ord Gi* Hot Spot Analysis was employed. This analysis is a local spatial statistic that identifies statistically significant clusters of high values ("hot spots") and low values ("cold spots") (Getis, A., & Ord, J. K., 1992). Unlike GMI, which is a global statistic and provides an overview of spatial patterns, Getis-Ord pinpoints specific areas by calculating a z-score and p-value for each feature locally, where intense clustering of similar values (high or low) occurs more than random chance would predict. This statistic was chosen to determine areas where reforestation should take place as the "cold spots" revealed in this analysis are most vulnerable to PTC loss at the highest rate.

3.2.6 Analysis of Demographic and Socioeconomic Factors

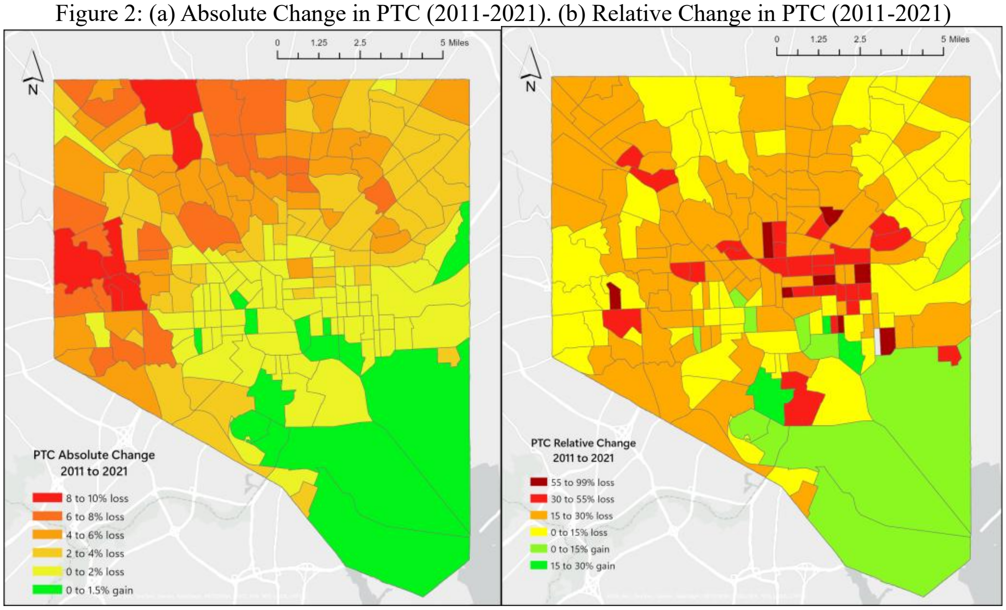
A stepwise linear regression analysis was employed to analyze the relationship between the chosen census predictor variables based on two tree canopy metrics which served as the dependent variables: the PTC 5-Year Mean (2017-2021) and the NDVI 5-Year Mean (2017-2021). The stepwise procedure was used to select the optimal set of predictors by iteratively adding or removing census variables to determine their individual and combined significant contribution to the explained variance of tree canopy metrics.

4. Results

4.1 PTC Change

Baltimore City experienced a notable decline in its overall PTC between 2011 and 2021, decreasing from 16.4% to 13.6% (Figure 1, appendix). This represents an absolute loss of 2.8% (Figure 2a) and a relative loss of 17% of its original tree canopy (Figure 2b). Most census tracts exhibited significant tree canopy decline as indicated by a negative Sen's slope value, with only a few showing modest increases (Figure 3). The northwestern areas of the city experienced the most substantial decreases in their original tree canopy, despite having some of the highest initial PTC percentages. In contrast, three census tracts in the south-central and western regions, along with some scattered internal patches, demonstrated more stable or slightly increasing tree canopy cover (Figure 3). It is particularly noteworthy that the census tracts exhibiting the highest PTC gains, though still modest, are predominantly found in largely industrial zones, such as Curtis Bay to the west and the Canton Industrial Area to the east (Figure 2a).

Figure 2: (a) Absolute Change in PTC (2011-2021). (b) Relative Change in PTC (2011-2021)



(a) (b)

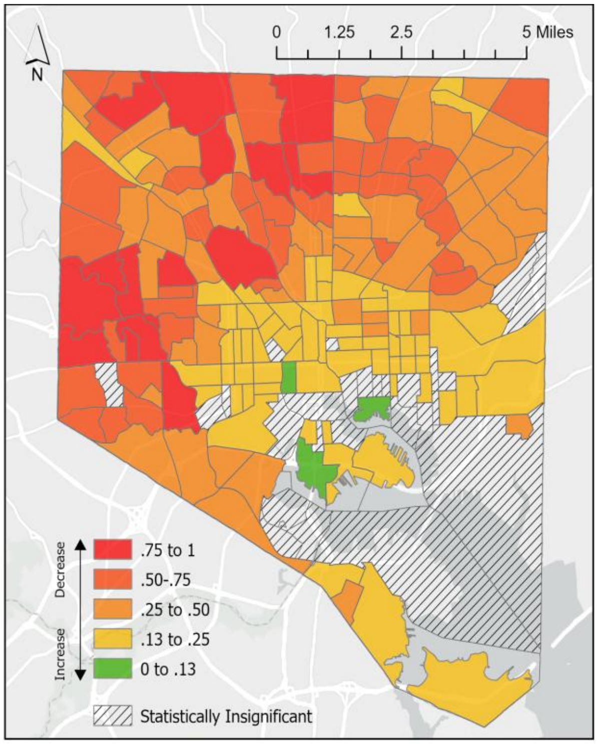


Figure 3: PTC Trend Analysis, Sens's Slope (2011-2021).

A GMI analysis of these Sen's slope values across all census tracts revealed a statistically significant positive spatial autocorrelation, with a Moran's *I* value of 0.5457 (Table 3, appendix). This indicates that the observed pattern of tree canopy change trends is more clustered than would be expected under a random distribution. The highly significant *z*-score of 22.0820 (p = 0.000) confirms that there is less than a 1% probability that this clustered pattern occurred by random chance. This result suggests that census tracts with similar trends in PTC change are spatially clustered together.

The Getis-Ord Gi* Hot Spot analysis of PTC change revealed statistically significant spatial clustering of both canopy loss and gain, indicating that these changes are not randomly distributed. Two prominent cold spots, representing areas of a significant PTC decrease with 99% confidence, were identified in the north and western regions of the city in several residential neighborhoods including Gwynn's Falls/Leakin Park, Mount Washington, and Cylburn arboretum (Figure 4). Conversely, two distinct hot spots, showing PTC increases with 90% confidence, were identified in the southern part of the city, adjacent to the Baltimore Harbor in the Cherry Hill Residential Area (Figure 4). However, it is important to note that these areas do not meet the conventionally accepted threshold for statistical significance of 95% and are therefore not accepted as true hot spots of PTC gain. The predominance of "Not Significant" areas across the remaining urban landscape suggests a more random pattern of canopy change outside of these identified clusters.

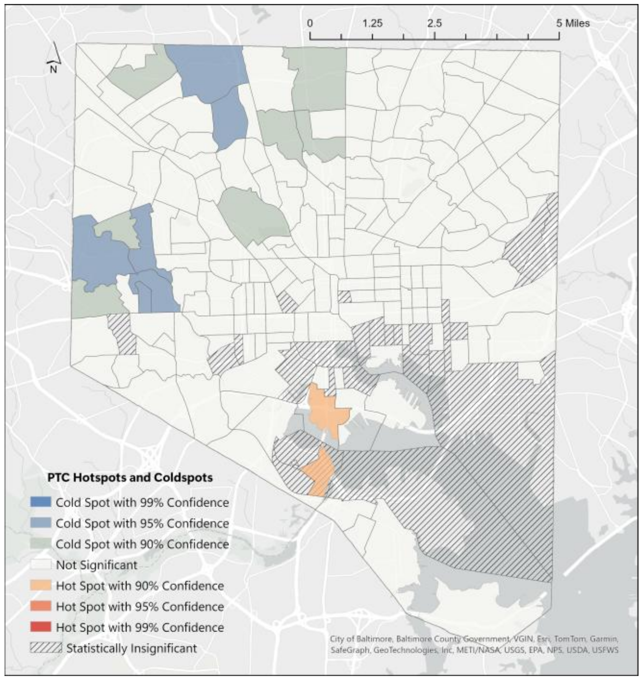


Figure 4: Hotspot Analysis of PTC Sen's Slope (2011-2021).

4.2 NDVI Change

The overall mean NDVI in Baltimore City exhibited a slight decrease from 0.165 in 2011 to 0.160 in 2021 (Figure 9, appendix), with modest increases over the majority of census tracts and NDVI decline concentrated in the northwest quadrant of the city (Figure 5). It is important to note that NDVI values in this range (0.160-0.165) are considered very low. Typically, NDVI values below zero indicate impervious surfaces, while values of 0.60 or higher are characteristic of healthy, lush vegetation. This low NDVI range suggests that the existing tree canopy in Baltimore City is generally sparse and may lack the robust environmental benefits associated with denser vegetation. The Mann-Kendall test indicated a statistically significant, modest increase in NDVI values across the municipality, with the exception of a significant downward trend observed specifically within the north-west quadrant (Figure 6).

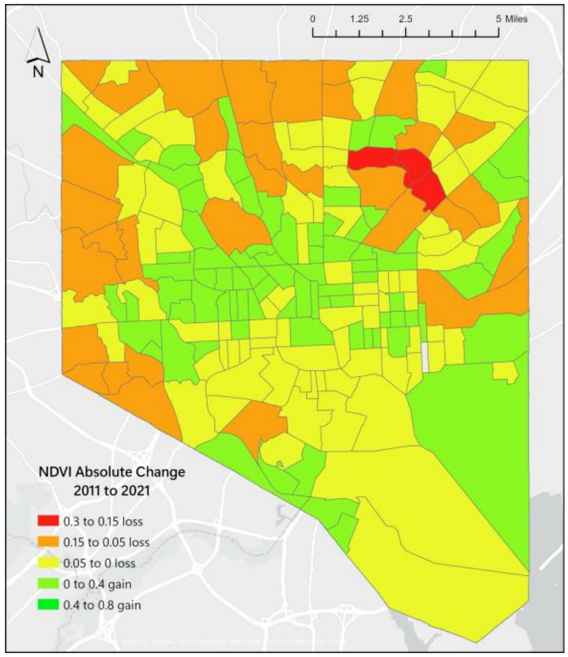


Figure 5: NDVI Absolute Change (2011-2021).

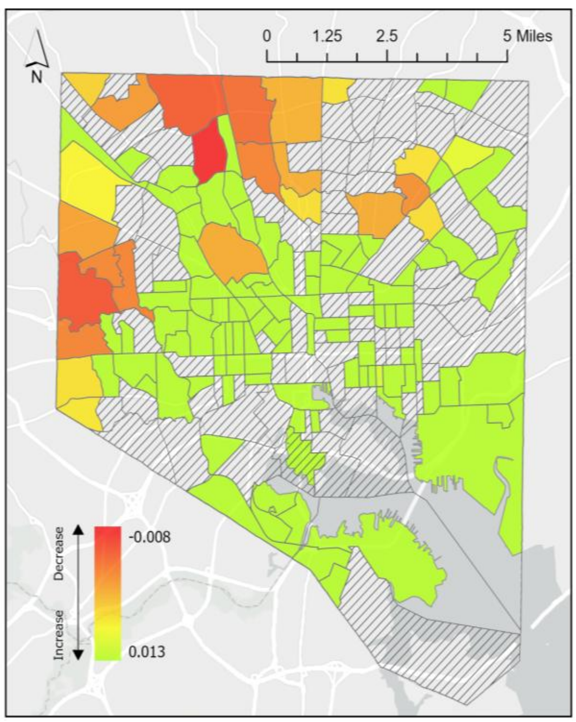


Figure 6: NDVI Mann-Kendall Trend Analysis, Sen’s Slope (2011-2021).

Similar to the patterns observed in PTC changes, a GMI analysis of the Sen's slope for NDVI within census tracts revealed a statistically significant positive spatial autocorrelation. With Moran's *I* index of 0.4336, the observed trends in NDVI change are more spatially clustered than would be expected under a random distribution (Table 4, appendix). The z-score of 8.0576 and a p-value of 0.00 indicate a less than 1% probability that this clustered pattern is a result of random chance. This strongly suggests that census tracts experiencing similar changes in NDVI are concentrated together geographically.

Furthermore, a similar spatial pattern to PTC was observed in the Getis-Ord analysis for NDVI, which identified several census tracts with statistically significant NDVI loss (at 95% confidence or higher) concentrated in residential neighborhoods in the northwestern quadrant of Baltimore. The highest NDVI losses were found within Cylburn Arboretum (Figure 7). In contrast, two isolated tracts showing NDVI gain were identified in Baltimore's southernmost region, within the Cherry Hill residential area, which encompasses several primary schools and Middle Branch Park along the harbor (Figure 7).

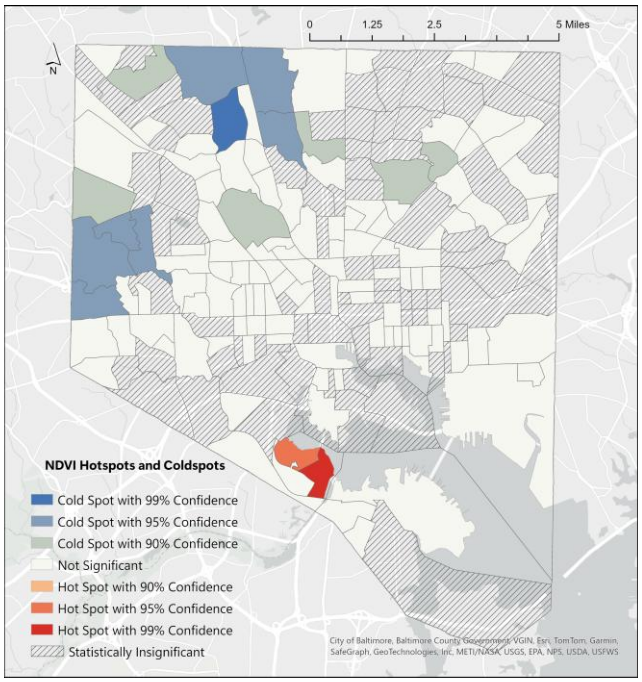


Figure 7: NDVI Mann-Kendall Trend Analysis, Sen's Slope (2011-2021).

4.3 Relationships between Census Variables, PTC, and NDVI

4.3.1 Pearson Correlation

The Pearson correlation analysis (Table 2) revealed that population density and percent vacant housing units are the strongest correlates for both PTC and NDVI, each demonstrating a significant negative association (p<0.05). A statistically significant, though weaker, negative correlation was also observed between PTC and percent housing units built in 2010 or later. In contrast, a weak positive relationship was found between PTC and median family income, suggesting a modest increase in tree canopy with rising income. The remaining variables, including demographics such as race, education, and employment, did not show a statistically significant linear relationship with either PTC or NDVI.

Table 2 – Variables selected from the American Community Survey 2021 5-Year Estimates for regression analysis.

Variable Name	**Correlation coefficient between variable and PTC-5 Year Mean (p-value)**	**Correlation coefficient between variable and NDVI 5-Year Mean (p-value)**
Percent 25+ with a bachelor's degree	-0.038 (p=0.593)	-0.034 (p=0.640)
Median family income (dollars)	0.175 (p=0.017)	0.118 (p=0.108)
Percent employed	-0.049 (p=0.499)	-0.048 (p=0.503)
Percent vacant housing units	-0.367 (p=0.00)	-0.327 (p=0.00)
Percent built 2010 or later	-0.195 (p=0.006)	-0.133 (p=0.064)
Median age (years)	0.148 (p=0.039)	0.098 (p=0.170)
Percent White	-0.061 (p=0.397)	-0.036 (p=0.616)
Percent Hispanic or Latino	-0.073 (p=0.308)	-0.035 (p=0.628)
Population Density (Persons per square mile)	-0.522 (p=0.00)	-0.584 (p=0.00)

4.3.2 Regression Results for PTC

The final regression model for 5-year mean PTC is formulated as:

PTC 5YR Mean *35.22 + 0.0001176 Median family income (dollars) - 0.0771 Percent Employed*
- 0.3610 Percent Vacant housing units - 0.409 % Built 2010 or later
- 0.000960 Pop Density (sq mi 2020) - 0.062 Median age (years)
- 0.1345 Percent White - 0.0295 Percent Hispanic or Latino

Table 3 – PTC Regression Analysis Results

Number of Observations (N)	**Standard Error**	**R^2**	**Adj. R^2**	**F-statistic**	**p-Value**
185	9.25534	51.85%	49.66%	23.69	0.000
Variable	**Coefficient**	**SE Coefficient**	**T-Statistic**	**P-Value**	**VIF**
Constant	35.22	7.09	4.96	0.000	
Median family income	0.000118	0.000027	4.39	0.000	3.03
Percent Employed	-0.0771	0.0773	-1.00	0.320	1.95
Percent Vacant housing units	-0.3610	0.0635	-5.68	0.000	1.38
Built 2010 or later	-0.409	0.115	-3.55	0.000	1.14
Population Density	-0.000960	0.000112	-8.56	0.000	1.18
Median age	-0.062	0.117	-0.53	0.598	1.33
Percent White	-0.1345	0.0422	-3.18	0.002	3.41
Percent Hispanic or Latino	-0.0295	0.0981	-0.30	0.764	1.32

The final regression model for PTC 5YR Mean was found to be statistically significant, as indicated by an F-statistic of 23.69 and a corresponding p-value of 0.000. This indicates that the independent variables collectively have a statistically significant effect on the dependent variable, PTC 5-Year Mean. The model explains with the adjusted R-squared value of 49.66% that nearly half of the variability in PTC 5YR Mean is explained by the census variables included in this analysis. The regression equation is based on an observation of 185 samples (census tracts) and the variation inflation factor (VIF) of all variables is less than 5, which indicates a low risk for multicollinearity.

Based on this analysis, median family income showed a positive relationship with PTC, indicating that more affluent communities experience higher percentages of tree canopy. All other significant variables—percent vacant housing units, housing built in 2010 or later, population density, and percent white—demonstrated a negative relationship with PTC. The strongest association was observed with the percentage of housing units built in 2010 or later. However, interpretation should be made with caution, as the lack of standardization across all variables may bias the perceived strength of the relationships. The remaining variables—percent employed, median age, and percent Hispanic/Latino—did not have a statistically significant relationship with PTC (p>0.05), suggesting they are not strong predictors of PTC in this model.

4.3.3 Regression Results for NDVI

The final regression model for 5-year mean NDVI is formulated as:

NDVI 5YR Mean	=	*0.3720 + 0.000000566 Median family income (dollars)* *- 0.000556 Percent Employed - 0.002256 Percent Vacant housing units* *- 0.002410 % Built 2010 or later - 0.000008892 Pop Density (sq mi 2020)* *- 0.001290 Median age (years) - 0.000467 Percent White* *- 0.000454 Percent Hispanic or Latino*

Table 3 – NDVI Regression Analysis Results

Number of observations (N)	Standard Error	R^2	Adj. R^2	F-Statistic	p-Value
185	0.0748371	48.30%	45.95%	18.40	0.000

Variable	Coefficient	SE Coefficient	T-Statistic	P-Value	VIF
Constant	0.3720	0.0574	6.49	0.000	
Median family income	0.000001	0.000000	2.61	0.010	3.03
Percent Employed	-0.000556	0.000625	-0.89	0.375	1.95
Percent Vacant housing units	-0.002256	0.000514	-4.39	0.000	1.38
Built 2010 or later	-0.002410	0.000931	-2.59	0.010	1.14
Population Density	-0.000009	0.000001	-9.80	0.000	1.18
Median age	-0.001290	0.000942	-1.37	0.173	1.33
Percent White	-0.000467	0.000342	-1.37	0.173	3.41
Percent Hispanic or Latino	-0.000454	0.000793	-0.57	0.568	1.32

The regression model for NDVI 5YR Mean was found to be statistically significant, as indicated by an F-statistic of 18.40 and a corresponding p-value of 0.000. This result confirms that the independent variables collectively have a statistically significant effect on the dependent variable, NDVI 5-Year Mean. The model explains with the adjusted *R*-squared value of 45.95% that nearly half of the variability in PTC 5YR Mean is explained by the census variables included in this analysis. The regression equation is based on an observation of 185 samples (census tracts) and the VIF of all variables is less than 5, which indicates a low risk for multicollinearity.

Based on the regression analysis, median family income demonstrated a positive association with NDVI, indicating that more affluent neighborhoods experience healthier tree canopy. Conversely, an inverse relationship with NDVI was observed for several variables, specifically percent vacant housing units, the percent housing constructed in 2010 or later, and population density. This finding indicates that lower tree health is statistically associated with areas characterized by recent development, higher vacancy rates, and increased population density. The remaining variables—percent employed, median age, percent White, and percent Hispanic/Latino—did not exhibit a statistically significant relationship with NDVI and were therefore not considered to have a predictive effect in the model.

5. Discussion

5.1 Implications for Baltimore's Tree Canopy

This analysis reveals a consistent, city-wide decline in Baltimore's tree canopy, with the most significant rates of loss concentrated in the northwestern quadrant (Figure 3). The most pronounced "cold spots" of PTC decline were observed in historically established neighborhoods in the northwest quadrant of the city (Figure 8), which had been positively affected by Baltimore City Forester Charles A. Young Jr.'s comprehensive urban tree census and directed tree planting initiative, which occurred from 1956 to 1959 (Merse et al., 2009). The age of these neighborhoods suggests that their mature tree populations may be reaching the end of their natural lifespans, a factor that could contribute to the observed decline in both PTC and NDVI. In addition to tree aging, other potential drivers of PTC loss include ongoing urban development and the effects of climate change. Continued development along the city's periphery can lead to the removal of existing tree cover to make way for new construction. Furthermore, rising summertime temperatures, exacerbated by both climate change and the urban heat island effect, are known to stress urban vegetation. Research has shown that such thermal stress can weaken trees through excess moisture loss and pest damage, a factor that could contribute to Baltimore's tree canopy decline (Cregg & Dix, 2001)

The regression analysis indicates that both PTC and NDVI are influenced by key demographic and built environment variables, with statistically significant models that explain roughly half of the observed variance. The strongest and most consistent predictors of reduced green space were directly tied to the built environment of the census tracts. Specifically, population density emerged as the most significant negative predictor for NDVI, which aligns with findings in urban ecology that denser, more urbanized areas containing a higher proportion of impervious surfaces demonstrate lower environmental health indicators (Yan et al., 2019). Similarly, the negative relationship with the percent of housing units built in 2010 or later was the strongest predictor for PTC. This result suggests that recent development practices, which often involve extensive land clearing, may not adequately incorporate or protect mature tree canopies, leading to a net loss of vegetation in newly developed areas. Conversely, median family income showed a weak but statistically significant positive relationship with both PTC and NDVI. This finding supports the concept that wealthier communities often have the resources to invest in and maintain urban green infrastructure, such as parks and street trees, and have more influence in the government regulatory bodies that enable the creation of more urban green spaces. (Richards et al., 2017). The negative association with percent vacant housing units was also a statistically significant finding, which is a nuanced result suggesting that the presence of unmanaged, vacant properties could indicate communities that lack the proper resources to create and maintain purposeful green spaces. The remaining variables—percent employed, median age, percent White, and percent Hispanic/Latino—did not have a statistically significant relationship with either PTC or NDVI in this model. While a key objective of this study was to assess if marginalized communities in Baltimore experience lower percentages of tree canopy, the regression analysis did not identify a statistically significant relationship between the racial demographic variables and the percent tree canopy or NDVI. Though the analysis did reveal that median family income and percent vacant housing units emerged as statistically significant predictors of tree canopy success, it is important to note that this regression analysis can only identify relationships between variables and is not suitable for determining causation. Therefore, while the results suggest a link between these socioeconomic factors and green space, they do not prove that one variable directly causes a change in the other.

Lastly, it is important to consider the differences in methodology between this study and the original study by Grove and O'Neil-Dunne (2009), which utilized land cover classifications derived from high-resolution aerial photography to analyze the city's tree canopy. There may arise discrepancies in results simply due to a difference in methods. This previous study's major conclusion was that residential properties "*control the largest percentage of the city's tree canopy*." Therefore, public outreach on the importance of planting and maintaining trees on private properties should be a central component of future reforestation efforts.

5.2 Recommendations

These recommendations outline a targeted reforestation strategy, focusing on "cold spots"—areas experiencing the most significant decline in tree canopy. While these efforts are crucial, it is important to recognize that tree canopy loss is a widespread issue throughout Baltimore. Therefore, a successful long-term strategy requires planting efforts across the entire city, not just in these priority areas. Achieving the city's long-standing goal of 40% tree canopy (TreeBaltimore, n.d.) requires a multi-faceted approach involving collaboration between the city's Forestry Division and local non-profit organizations, such as Blue Water Baltimore, Baltimore Tree Trust, and other independent groups invested in maintaining the city's tree canopy. Community involvement in the city's re-greening efforts holds great potential for success, as it is linked to providing psychological benefits and fostering a sense of community for participating neighborhood residents (Konu et al., 2024).

To combat continued degradation, Baltimore's reforestation strategy needs to be carefully tailored to the specific conditions of different neighborhoods. For example, while Northwest Baltimore has experienced significant tree canopy loss, these neighborhoods still possess a relatively high amount of existing canopy compared to more central regions. The primary goal in these areas should therefore be strategic planting and maintenance to mitigate further loss and bolster the health of the already substantial tree canopy.

The neighborhoods identified as statistically significant spots of PTC loss include (Figure 8):

- **To the east:** Dickeyville, Wakefield, Franklintown, Windsor Hills, Gwynn's Falls/Leakin Park, Edmonson Village, and Edgewood.
- **To the north:** Mount Washington and Coldspring.

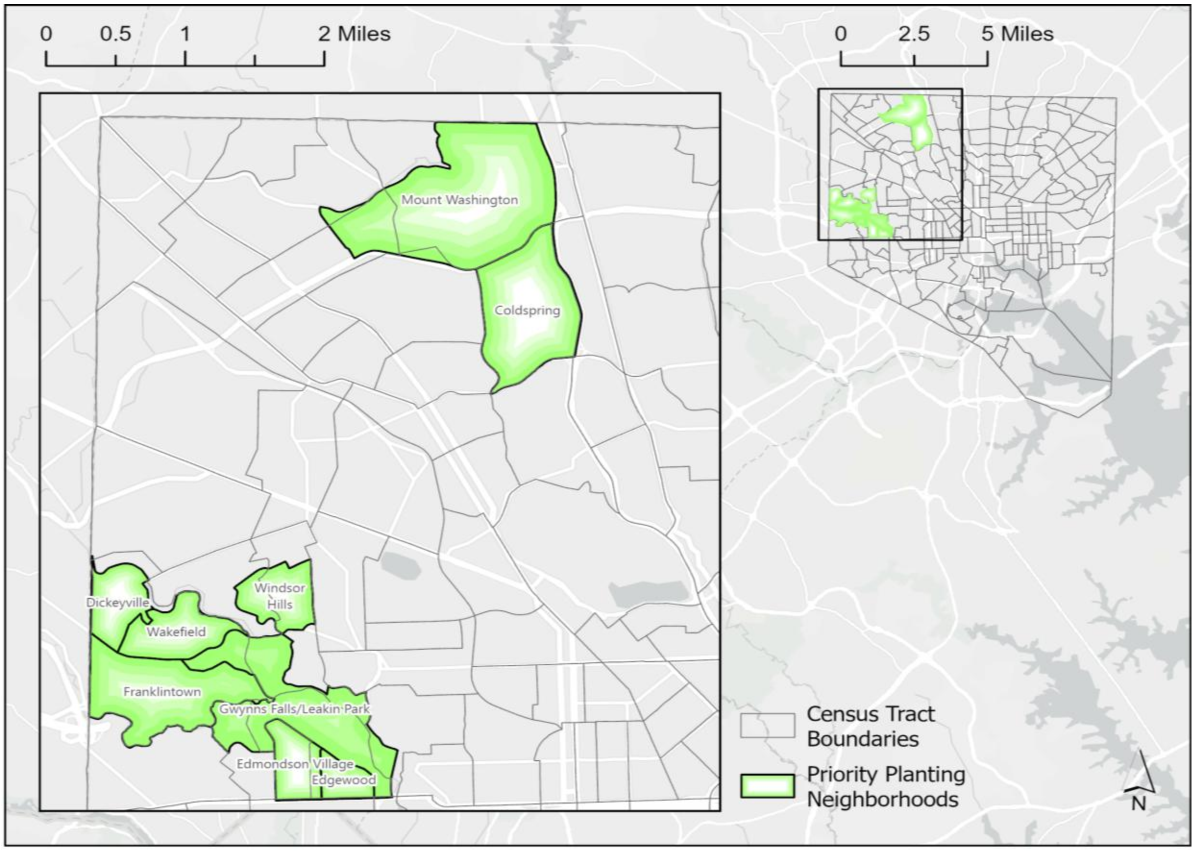


Figure 8: Recommended Neighborhoods for Priority Planting.

In contrast, central Baltimore faces a more critical challenge. Here, the data reveals a rapid loss of the very limited tree canopy that remains in many areas closer to the city center. For these areas, the term "reforestation" may not be entirely applicable, as there is often insufficient pervious cover to support new tree growth. A viable strategy for these neighborhoods would involve addressing the underlying land cover issues first, perhaps through urban planning initiatives that create more green spaces and permeable surfaces before planting efforts can begin.

Additionally, the City of Baltimore should implement a robust annual monitoring program for tree canopy density and health in order to protect the longevity and continuity of tree canopy for Baltimore residents. This study shows that areas with a lower existing PTC are more susceptible to tree loss. Therefore, a strategic approach is needed to prioritize not only planting new trees but also maintaining the existing canopy, especially in these vulnerable low-PTC areas. By protecting its current canopy, the city can create a more stable foundation that will likely facilitate successful reforestation efforts and better support its overall green infrastructure goals.

5.3 Limitations

This study's findings are subject to several limitations primarily related to the spatial resolution and classification accuracy of the datasets used. A key limitation of the USDA's TCC dataset lies in its temporal and spatial resolution, which presents a challenge for future replicability of this study. The dataset, currently offering annual data only from 2011 to 2021 at a 30-meter resolution, may not accurately reflect current conditions or capture the precise tree cover in highly developed areas where trees are sparser. Furthermore, this 30-meter resolution poses a challenge in accurately capturing fine-scale urban tree features, such as individual street trees or small clusters within residential yards. In highly heterogeneous areas, including urban and fragmented suburban landscapes, this resolution can result in mixed pixels that do not precisely represent the actual tree canopy fraction.

Second, the methodology of the PTC estimates may systematically undercount sparse or small trees, particularly in dense urban environments dominated by impervious surfaces. Conversely, the classification process risks potential error by misclassifying dense shrublands or tall non-tree vegetation as tree canopy. It is also important to note that the classification accuracy of PTC estimates is known to vary regionally based on the quality of training data and source imagery.

Finally, the NDVI dataset used in this analysis contains substantial gaps due to frequent cloud cover over the study area of Baltimore. Future work should consider integrating alternative data sources, such as Sentinel-2 or ASTER, to obtain more consistent and complete measurements of vegetation health and density.

To address the spatial limitations of the TCC/PTC dataset, there is high potential for improved accuracy using object-based classification methods that utilize spatial relationships and often incorporate Light Detection and Ranging (LiDAR) technology (King et al., 2013). LiDAR offers a significant advantage over traditional monitoring by providing highly detailed 3D measurements of a tree canopy's height, density, and volume, which would generate a more comprehensive and accurate dataset for future studies.

6. Conclusion

This study aims to analyze the spatial and temporal patterns of tree canopy in Baltimore City and determine their correlation with sociodemographic variables. This analysis of Baltimore's urban tree canopy from 2011 to 2021 revealed a significant decline, with a 17% loss in PTC that was particularly concentrated in the northwestern quadrant of the city. While overall tree health, as measured by NDVI, showed only a minor average decrease from 0.165 to 0.160, this decline was also most pronounced in the upper northwest, with other areas generally maintaining or improving in health.

The study also explored the relationship between tree canopy/NDVI and demographic and socioeconomic factors, finding an inverse correlation between PTC and both population density and vacant housing, while a positive relationship was observed with median income. Despite this inconclusiveness on the socioeconomic aims, the definitive finding of an alarming rate of tree canopy decline highlights the urgent need for strategic intervention. Therefore, this study recommends targeted planting efforts in the residential neighborhoods in the northwestern quadrant, while also advocating for a more holistic city-wide approach that includes community involvement and outreach to private landowners, given that the majority of the city's tree canopy is on residential land.

Building on these findings, future research should focus on refining tree canopy estimation through the incorporation of LiDAR data to capture detailed 3D structure and improve accuracy, particularly for fine-scale urban features. To address the inconsistent temporal coverage of the NDVI dataset, it is recommended to utilize alternative high-resolution satellite imagery (e.g., ASTER) to achieve more consistent vegetation monitoring at a finer resolution. Furthermore, subsequent studies should consider exploring the direct impact of localized community re-greening initiatives on tree canopy success and social metrics, employing a mixed-methods approach that integrates quantitative spatial analysis with qualitative data on community engagement and psychological wellbeing. This will allow for the development of more meaningful and impactful urban forestry initiatives.

Acknowledgements

I would like to thank the College of Liberal Arts at Towson University for supporting this work through funding from the 2025 CLA Summer Research Grant. I am also deeply grateful to Dr. Carter Wang for his invaluable guidance throughout this project.

References

Baltimore City Department of Parks and Recreation. (2007). Tree Baltimore. https://www.treebaltimore.org/

Brown, L., Buccino, D. L., Casiano, M., Collins, S., Darrow, S., Durington, M., ... & Stocks, S. (2019). *Baltimore revisited: Stories of inequality and resistance in a US city*. Rutgers University Press.

Cregg, B. M., & Dix, M. E. (2001). Tree moisure stress and insect damage in urban areas in relation to heat island effects. *Arboriculture & Urban Forestry (AUF)*, *27*(1), 8-17.

Earth Resources Observation and Science (EROS) Center. (2021). Landsat 4-9 U.S. Analysis Ready Data, Collection 2 [dataset]. U.S. Geological Survey. https://doi.org/10.5066/P960F8OC.

Getis, A., & Ord, J. K. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, *24*(3), 189–206. https://doi.org/10.1111/j.1538-4632.1992.tb00261.x

Grove, J. M., & O'Neil-Dunne, J. P. M. (2009). *A Report on the City of Baltimore's Existing and Possible Urban Tree Canopy*. USDA Forest Service, Northeastern Research Station & University of Vermont, Spatial Analysis Laboratory.

Housman, I. W., Schleeweis, K., Heyer, J. P., Ruefenacht, B., Bender, S., Megown, K., Goetz, W., & Bogle, S. (2023). *National Land Cover Database Tree Canopy Cover Methods v2021.4* (GTAC-10268-RPT1). U.S. Department of Agriculture, Forest Service, Geospatial Technology and Applications Center. https://www.mrlc.gov/sites/default/files/docs/TCC_v2021-4_Methods.pdf

Jung, M. C., Yost, M. G., Dannenberg, A. L., Dyson, K., & Alberti, M. (2024). Legacies of redlining lead to unequal cooling effects of urban tree canopy. *Landscape and Urban Planning*, *246*, 105028.

Kabisch, N., & Haase, D. (2011). Diversifying European agglomerations: Evidence of urban population trends for the 21st century. *Population, space and place*, *17*(3), 236-253.

Kendall, M. G. (1975). *Rank Correlation Methods* (4th ed.). Charles Griffin.

King, K. L., & Locke, D. H. (2013). A comparison of three methods for measuring local urban tree canopy cover. Arboriculture & Urban Forestry. 39 (2): 62-67., 39(2), 62-67

Konu, B., White, P., & Pescott, O. L. (2024). Does community involvement in tree planting improve success? And is community tree planting only proven to improve psychological wellbeing? A systematic review. *EcoHealth*. Advance online publication. https://doi.org/10.1089/eco.2024.0038

Locke, D. H., Hall, B., Grove, J. M., Pickett, S. T., Ogden, L. A., Aoki, C., ... & O'Neil-Dunne, J. P. (2021). Residential housing segregation and urban tree canopy in 37 US Cities. *NPJ Urban Sustainability*, *1*(1), 15.

Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, 245-259.

Merse, C. L., Buckley, G. L., & Boone, C. G. (2009). Street Trees And UrbanRenewal: A BaltimoreCase Study. *The Geographical Bulletin*, *50*(2), 1.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika*, *37*(1/2), 17–23. https://doi.org/10.2307/2332142

Multi-Resolution Land Characteristics (MRLC) Consortium. (2025). *NLCD Tree Canopy Cover*. https://www.mrlc.gov/data/type/nlcd-tree-canopy-cover

Richards, D. R., Passy, P., & Oh, R. R. (2017). Impacts of population density and wealth on the quantity and structure of urban green space in tropical Southeast Asia. *Landscape and Urban Planning*, *157*, 553-560.

Schwarz, K., Fragkias, M., Boone, C. G., Zhou, W., McHale, M., Grove, J. M., ... & Cadenasso, M. L. (2015). Trees grow on money: urban tree canopy cover and environmental justice. *PloS one*, *10*(4), e0122051.

Ulmer, J. M., Wolf, K. L., Backman, D. R., Tretheway, R. L., Blain, C. J., O'Neil-Dunne, J. P., & Frank, L. D. (2016). Multiple health benefits of urban tree canopy: The mounting evidence for a green prescription. *Health & place*, *42*, 54-62.

U.S. Census Bureau. (2021). Race, Bachelor's degree/ Graduate or professional degree, Median household income. [Data set]. Decennial Census 2020. Retrieved from https://data.census.gov/table/ACSST5Y2022.S0101?q=people, accessed March 31, 2025.

USDA Forest Service. (2023). USFS NLCD Percent Tree Canopy CONUS v2021-4. [data set] Sioux Falls, SD.

U.S. Geological Survey. (n.d.). *EarthExplorer*. Retrieved September 27, 2025, from https://earthexplorer.usgs.gov/

Rouse, J. W., Jr., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *Third Earth Resources Technology Satellite-1 Symposium*, *1*(A), 309-317.

TreeBaltimore. (n.d.). *TreeBaltimore | Baltimore's Tree Planting Partnership*. https://www.treebaltimore.org/

Xiao, Q., & McPherson, E. G. (2005). Tree health mapping with multispectral remote sensing data at UC Davis, California. *Urban Ecosystems*, *8*, 349-361.

Yan, Z., Teng, M., He, W., Liu, A., Li, Y., & Wang, P. (2019). Impervious surface area is a key predictor for urban plant diversity in a city undergone rapid urbanization. *Science of the Total Environment*, *650*, 335-342.

Yin, Y., Li, S., Xing, X., Zhou, X., Kang, Y., Hu, Q., & Li, Y. (2024). Cooling benefits of urban tree canopy: a systematic review. *Sustainability*, *16*(12), 4955.

Appendix

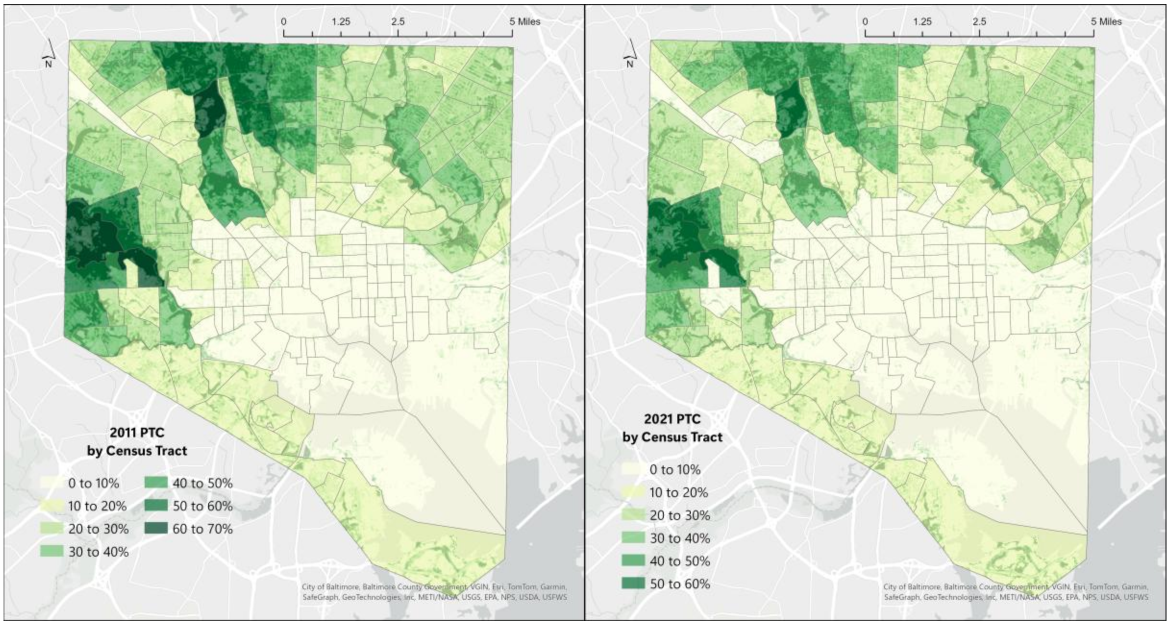


Figure 1: 2011 PTC (left) and 2021 PTC (right) at the Census Tract Level.

Table 3 – Global Moran's I Spatial Autocorrelation Results for PTC Sen's Slope

Moran's Index	0.545691
Expected Index	-0.005780
Variance	0.000624
z-score	22.081958
p-value	0.000000

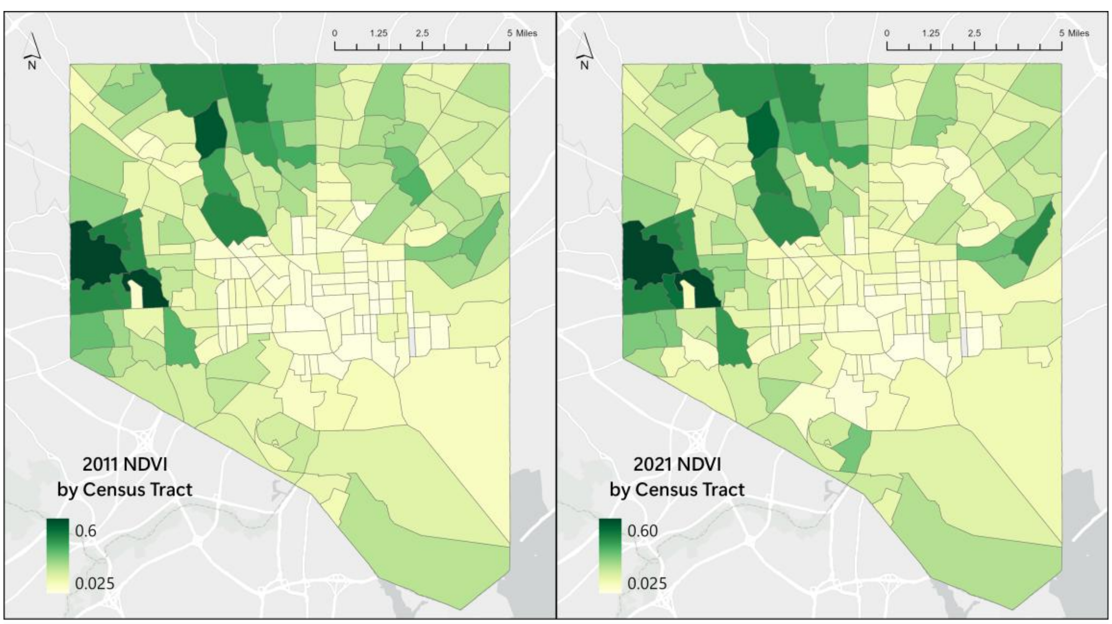


Figure 9: 2011 NDVI (left) and 2021 NDVI (right) at the Census Tract Level.

Table 4 – Global Moran's I Spatial Autocorrelation Results for NDVI Sen's Slope

Moran's Index	0.433597
Expected Index	-0.008197
Variance	0.003006
z-score	8.057646
p-value	0.000000