Mapping Urban Green Spaces

CS 701 Final Project, Middlebury College, Fall 2021

Sam Marshall
Middlebury College
gsmarshall@middlebury.edu

Kaylen Rivers
Middlebury College
krivers@middlebury.edu

ABSTRACT

The scarcity of forested land in urban areas reduces quality of life, poses a health threat for millions of people, and worsens existing inequalities in housing, life expectancy, and mental health. Accurate maps of inequities in public access to green space can help activists and government officials address these issues. Tree Equity Score already provides a free, easy to understand web map of green space access for the entire United States. But while their overall workflow is public, they do not publish their analysis code or the tree cover data that underpins their analysis. Their analysis method differs significantly from those found in academic literature, and they provide little discussion of the theory behind it. We aim to replicate the Tree Equity Score analysis for New England using public satellite imagery and host the results on a public website. Our goals were to (1) create an easily accessible and fully open-source version of the analysis, (2) contribute to an assessment of whether or not the Tree Equity Score methodology is accurate, and (3) provide an open-source template that will make it easier for others to improve on this analysis or implement it for other regions. Although we did not meet our goal of publishing a website to present our findings, our results show broad agreement between the scores we calculated and those published on the original Tree Equity Score website.

1 INTRODUCTION

Urban green space refers to places like parks, playing fields, and gardens in urban areas. Green spaces are crucial for improving quality of life in urban areas as well as managing the environmental impacts of paved surfaces, such as excess water runoff and temperature buildup (known as the urban heat island effect). [1]. Green spaces benefit those who live near them by reducing stress levels, decreasing suicide rates, reducing rates of heart disease, improving air pollution, and increasing biodiversity [2]. Urban green spaces are also key in providing recreation areas for children, whose mental and physical health are greatly improved by access to nature.

While cities in general do not have sufficient green space, poor and nonwhite urban areas often face a disproportionate lack of green space [3]. The astonishing lack of green spaces in impoverished areas is often called the green-gray effect, where satellite imagery shows a distinct difference between the grayer, poorer neighborhood and the greener, wealthier neighborhood. Efforts to address this disparity are underway in many cities, and these projects require accurate maps of inequities in public access to green space.

In this paper we posit a formal problem statement identifying how we tackled the problem of sparse urban green spaces. We will provide a background of related works that influenced our satellite image processing and data analysis, describe the methods that we settled on for analysis, and present the results and our finished product.

2 PROBLEM STATEMENT

Tree Equity Score is a project created by American Forests to map the equity of green space distribution in US cities [4]. They have designed a simple, easily understandable metric to quantify the equity of green space distribution and have performed this analysis for all US cities. Their method quantifies green space equity on a scale from 0 to 100, with higher values indicating a more equitable distribution of green space, and is intended to help neighborhood advocates, government officials, and urban planners understand and improve green space access.

The tree equity score metric (TES) is a combination of tree canopy coverage with socioeconomic, demographic, health, and climate data. While Tree Equity Score does a good job of describing their methods and publishing their results as well as the demographic data they use, the canopy cover data that underpins their analysis is private and they do not provide code or a complete analysis workflow. Our goal is to improve understanding of the Tree Equity Score metric by creating an open-source replication of their analysis for urban areas in New England. This involves reimplementing the Tree Equity Score analysis based on their published methods, comparing our findings to the original findings, and making our methods and the results of this comparison publicly accessible.

Our project involves three main subproblems: 1) measuring the prevalence of urban green space, 2) identifying the equitable amount of green space based on local populations, and 3) displaying our data in a publicly accessible format.

We identify urban green space using publicly available satellite imagery, and use the workflow described on the Tree Equity Score website to quantify green space equity.

To display our data, we create a web-page with an interactive map displaying the TES of all urban areas in New England, as well as a graph showing the demographics of the local population. The interactive map shows scores in chunks of census block groups, the smallest region we could analyze with accurate data from the census. Each area is then colored according to its greenness, where a whiter color is associated with a lower score and a greener color with a higher score. The user can pan and zoom on the interface in order to see an location in New England included in our analysis. The radar chart displays the percentages of local nonwhite, senior, youth, unemployed, and impoverished populations, and will accurately display a radar chart for the census block group that is clicked from the map. The TES is also displayed as an integer.

In order to identify areas in need of green space development, we use satellite and census data to provide a TES, highlighting any disproportionately affected populations while displaying what urban areas are in need of the most development.

3 RELATED WORK

There is no single method for mapping access to green space that is ubiquitous in academic literature. Geographic studies related to the topic often have a specific focus within this broader problem, and we summarize a few of these methods based on a cursory literature review. Since we are replicating Tree Equity Score's methods rather than producing our own, we aim to provide context for how their methods are conceptually similar to and different from existing ways of solving this problem.

3.1 Mapping Green Space

Many studies that map access to urban green space identify green space using existing government data [5–7]. These studies often use property ownership records or coarse land use maps such as the National Land Cover Database to identify green space, and often focus on the value of large public green spaces such as parks for recreation [6, 7]. This approach is reliable and provides a specific understanding of the value of green space, but does not capture small patches of green space such as planted road medians and ignores the aggregate value of these small patches for things like heat regulation.

Other studies use satellite imagery to classify urban land cover [8–10]. These studies vary in scope and analysis scale and use a variety of machine learning and image processing techniques, but largely focus on simply mapping urban tree cover rather than mapping access to green space. These approaches are able to capture finer details of urban green space distribution, but make no distinctions between public and private green spaces or large patches and small patches. This enables such approaches to quantify the general value of green space but not the specific value that each type of green space provides.

Some studies aim to combine the above approaches, using statistical models to make distinctions between types of green space based on a combination of remote sensing data and other data sources such as Open Street Map[11].

3.2 Quantifying Access to Green Space

Quantifying access to green space requires tangibly defining concepts like 'access' and 'equity'. Many studies quantify access to green space by measuring the distance residents must travel to get to green space or the population within a certain distance of green space [5]. Other studies add more nuance to this approach by using some version of a gravity model of spatial interaction, which is a common way to model the potential interaction between a target service provider (i.e. a hospital providing health care or a park providing opportunity for recreation) and the users of that service in different areas [6, 7]. This approach is able to capture differences in how people interact with various kinds of green space, but requires detailed information defining each target feature and is not well suited to capturing the value of small patches of green space.

3.3 Tree Equity Score Model

Tree Equity Score uses remotely sensed tree cover data which maps tree cover accurately at a fine scale but makes no distinctions between types of green space [12]. In contrast to the methods outlined above, their analysis measures the distribution of green space across a city and conceptualizes equity by asserting that areas with a higher potential for socioeconomic vulnerability have a greater need for green space. This approach incorporates and combines the many different services provided by urban green space, but does not specifically measure any one of them.

4 METHODS

All of our code, data, and results can be viewed and downloaded from the project Github page.

4.1 Platform

Google Earth Engine (GEE) is a cloud based geospatial analysis platform that is available for use in independent web apps as a JavaScript API. We implement the Tree Equity Score workflow with GEE and connect this back-end analysis with a front-end interface built with NodeJS. We use TidyCensus 0.11.4 in R to query data from the U.S. Census API and QGIS 3.16 to compile and visualize our results.

4.2 Website

In order to display our findings, we created a web-page. The page features an interactive map with the TES of every census block group of urban areas in New England, a radar chart displaying the demographics of the selected block group, and the TES itself. For our interactive map we employed the Google Maps API in conjunction with Google Earth Engine, which allowed us to display our census block groups atop the base layer of the map. To create our webpage we worked in HTML, relying on JavaScript for analysis and interaction with our APIs.

4.3 Original Workflow

We begin by describing the data and methods used in the original Tree Equity Score analysis as they appear on their website. Tree Equity Score quantifies the equity of green space distribution using a multi-criteria analysis that incorporates tree canopy cover, demographic and socioeconomic data from the U.S. Census, health survey data from the CDC, and temperature data from the USGS. They perform their analysis at the census block group level, which is the smallest geographic unit for which 5-year survey data is available, and use census Incorporated Places to define a region of analysis for each city [13]. Tree Equity Score states on their website that they update scores yearly and use the most recent data available to calculate scores, though they do not in all cases specify when they update scores or exactly which data versions they use [14].

4.3.1 Data Sources. The original analysis uses tree cover data donated by EarthDefine, a private geospatial data and analysis company. These data represent tree canopy extent for the entire contiguous United States at 60cm resolution, with a reported 97.3% accuracy rate in urban areas and a one year revisit rate [15].

Demographic and socioeconomic data is derived from 5-year American Community Survey (ACS) estimates provided by the U.S. Census. The ACS consists of a survey completed by about two million households each year, for which data is aggregated in 1-, 3-, and 5-year intervals [16, 17]. Tree Equity Score uses 5-year estimates from 2014-2018, and although they do not specify which variables they use, this information is relatively easy to infer from their descriptions of the indicators and the table identifiers [12]. It is unclear how they define urban areas based on census Incorporated Places, and this ambiguity was a significant source of error in our replication.

Health data is derived from CDC Places census tract data, which uses the Behavioral Risk Factor Surveillance System (BRFSS) survey as well as census data to provide information about a variety of chronic health indicators [18]. These data are available on a yearly basis and are constructed from the most recent data available, which may include BRFSS and census data from different years, and are only available at or above the census tract level [18]. Tree Equity Score does not specify which year of Places data they used. However, as of the time we performed our analysis 2020 was the most recent year available, and these data appear to match those used in the original analysis [19].

The original analysis also uses surface temperature data from Landsat 8 to quantify urban heat island effects. Landsat 8 thermal bands measure surface temperature at 30m spatial resolution with a revisit rate of 16 days, but it is unclear how exactly this data is used. The data dictionary for the Tree Equity Score output data defines a temperature variable that is "the average temperature of the block group on a hot summer's day" but gives no other details as to how this is defined [20]. It would be reasonable to assume that this value is some average over each block group that is then normalized by citywide extremes, but it is unclear.

4.3.2 Gap Score. The equity score is composed of a canopy cover Gap score and a demographic Priority score. For a given block group, the Gap score is defined as the difference between the existing percent canopy cover and a canopy cover goal that is dependent on the biome in which the block group is located, normalized over the urban area that contains the block group.

The canopy cover goal is a general target for percent canopy cover based on the biome (such as forest, desert, or grassland) in which each city is located. This goal is then multiplied by an adjustment for population density; areas with lower population density receive a higher goal, while areas with higher population density receive a lower goal (see the indicator construction guide on the project Github page for more details[1]). Tree Equity Score does not specify how they determine city biomes, and the actual goal values are omitted from the output data in any areas where tree cover data was provide by EarthDefine. The original methods list a general goal of 40% forest cover, and this value is consistent with the goal values found in TES output data for Rhode Island where canopy cover is based on public data.

The Gap score is calculated by comparing the existing canopy cover with the canopy cover goal and normalizing this difference by the maximum difference value for the whole city:

$$GAP = (GOAL - EC)/(GOAL - EC)_{max} \tag{1}$$

Where $GOAL$ is the density-adjusted canopy cover goal for a particular block group, EC is existing percent canopy cover for that block group, and $(GOAL - EC)_{max}$ is the maximum difference value for the incorporated place that the block group lies within. The result is a score between zero and one, where zero indicates adequate tree cover and one indicates a need for greater tree cover.

4.3.3 Priority Score. The Priority score is a measure of how vulnerable the population of a black group is to social, economic, and climate inequities, and is defined as the average of six normalized indicators: unemployment rate, income relative to the poverty line, age dependency ratio, nonwhite population, health status, and urban heat island severity. Unemployment, income, dependency ratio, and nonwhite population are defined as percentages of the overall population. Tree Equity Score does not specify exactly which variables they use or how they calculate these percentages from the raw population data, but their descriptions of these indicators give enough information to reconstruct them, and we were able to independently match the values found in their output data. See the data dictionary on the project Github page for more detail on variable definitions[2].

The health status indicator incorporates the prevalence of poor physical health, poor mental health, asthma, and coronary heart disease as defined by the CDC and represented by the percent of the population that experiences each condition [**?**]. Tree Equity Score does not specify how they construct a single health index from those measures, but we were able to independently match the values found in their output data by normalizing each of them and averaging the results. See the indicator construction guide on the project Github page for more details[3].

Tree Equity Score does not specify how they calculate urban heat island severity.

Each of the above indicators is normalized according to Equation 2 and combined according to Equation 3:

$$N_i = (x_i - x_{i_m in})/(x_{i_m ax} - x_{i_m in}) \tag{2}$$

$$PRIORITY = \sum_{i=1}^{n} N_i/n \tag{3}$$

Where for each indicator N_i, x_i is the value of that indicator for a single block group, x_{imin} is the minimum value citywide for that indicator, and x_{imax} is the maximum value citywide for that indicator [13]. These normalized values are then averaged together to create a single Priority score between zero and one, where zero indicates the lowest potential for inequity and one indicates the highest potential for inequity in a particular city.

4.3.4 Final Tree Equity Score. The final Tree Equity Score is then calculated as the product of the Gap and Priority scores:

$$TES = 100(1 - GAP * PRIORITY) \tag{4}$$

[1]Located in data/metadata/indicator_construction.txt

[2]Located in data/metadata/replication_data_dictionary.txt

[3]See note 1 on page 3

Where the final TES value lies between zero and 100, with zero indicating the greatest need for improved tree cover in a particular city and 100 indicating the least need in that city.

4.4 Our Adaptation

*maybe switch around the ordering of these topics, maybe split into subsections if needed (but probably not)

We hew as close as possible to the original methods for our analysis, using the same data and the same units of analysis wherever possible. We deviate from the original analysis primarily by using our own land cover classification, omitting temperature data from the priority score, and performing our analysis on Census Designated Places rather than Incorporated Places.

4.4.1 Land Cover Classification. We identify tree cover by using a pixel-based random forest classifier with aerial imagery from the National Agriculture Imagery Program (NAIP) accessed through Google Earth Engine (GEE). We first classify general land cover into six categories: trees, grass and non-tree vegetation, buildings, roads and other paved surfaces, dirt, and water. We then reclassify these categories into a boolean classification of tree canopy cover and all other land cover. This improves the accuracy of our final classification by creating finer divisions between the spectral features that define each class.

The input imagery has one-meter spatial resolution with four spectral bands (visible red, green, blue, and near infrared), and for most of our study region of New England this imagery is available for 2012, 2014, 2016, and 2018 [21]. For each image tile from any year that intersects a given city study region, we calculate the Normalized Difference Vegetation Index (NDVI):

$$NDVI = (NIR - RED)/(NIR + RED) \tag{5}$$

Where NIR is the reflectance value for the near-infrared band and RED is the reflectance value for the red band. We then reduce this multi-year image stack into a single image by calculating the mean, maximum, 20th percentile, and standard deviation of the value of each band. The final input to our classifier is a composite image with 20 bands which represents the aggregation of all available imagery for a particular location. While using multiple years of imagery may not accurately represent areas where tree cover changes significantly, this method ensures spatial coverage for all city regions and gives the classifier more information to work with in areas where tree cover is relatively stable.

Training data for the classifier consists of 40 hand-collected training points per class, yielding 240 total points. This collection is then randomly split into training and validation sets, with 70% (160 points) used for training and 30% (80 points) used for accuracy assessment. Training data were collected for an initial test region of Manchester, NH.

4.4.2 Data Acquisition. We use the same demographic data used in the original analysis. We compile this data by using the R package TidyCensus to query the U.S. Census API for all variables necessary to calculate the priority indicators, as well as the block group geometries and Census Designated Place (CDP) geometries for our region of interest. We then attach each block group to its corresponding CDP by spatially joining the two layers together, yielding a single vector layer where each entry is represents a block group and contains its spatial geometry, census population estimates for each variable, and a unique identifier for the CDP it lies within. Block groups that lie outside of a CDP are dropped from the joined layer, and this layer is then uploaded to Google Earth Engine.

We compiled health status data by downloading 2020 CDC Places Census Tract data from the CDC website, which covers the entire United States, and then using QGIS to select only data for New England states.

We omit the heat island data from our analysis because of time constraints and difficulty understanding how it was incorporated into the original analysis.

4.4.3 Unplanned Deviations to Data Inputs. Due to difficulty scaling our analysis to all of New England and difficulty understanding the exact definition of urban areas that Tree Equity Score uses, we calculate scores based on Census Designated Places rather than Incorporated Places as in the original analysis.

4.4.4 Implementation in GEE. Google Earth Engine (GEE) encodes vector data layers as GeoJSON feature collections, and the input to our analysis consists of one feature collection for the census block group indicators and a separate feature collection for the census tract health status data. Preprocessing begins by joining the health data to the census data by matching census tracts to census block groups according to their census GEOID, which are unique identifiers given to each census geometry. These identifiers are nested according to the census geometry hierarchy, meaning that the GEOID for a block group contains the GEOID for the census tract (as well as county and state) it lies within. This geometry hierarchy also means that health data for each census tract is joined to multiple block groups, but there is no block group that might receive data from multiple census tracts. At this stage the four health indicators we are interested in (physical health, mental health, asthma, and coronary heart disease) are selected and the remaining variables dropped.

We then split the resulting feature collection into a list of feature collections containing one collection for each city (Algorithm 1). We do this by creating a list of unique CDP identifiers from the feature collection, then iterating through the whole collection and checking the value of each block group's CDP GEOID against each value in this list.

Algorithm 1 Splitting Input Feature Collection

Given: $combined\_indicators$, CDP_names
$places\_list$ = empty list
for all $name \in CDP\_names$ **do**
  $name\_fc$ = empty feature collection
  for all $feature \in combined\_indicators$ **do**
    if $feature.cdp\_name == name$ **then**
      Add $feature$ to $name\_fc$
    end if
  end for
  Add $name\_fc$ to $places\_list$
end for
Output: $places\_list$

This is among the most computationally intensive components of our analysis, involving a large amount of data transfer and having a potential runtime of $O(M * N)$ where M is the number of unique CDPs and N is the total number of block groups [1]

At this stage we have our data organized into a list where each entry is a single feature collection representing a city unit, and we perform the remainder of the analysis on a per-city basis by mapping our analysis function over this list. Algorithm 2 represents the analysis workflow for a single feature collection; we perform our analysis for the entire study region of New England by mapping this algorithm over the list of city feature collections produced by Algorithm 1.

Algorithm 2 City Level Analysis Workflow

```
Given: place_collection ▷ Collection representing a single city
Classify composite image covering the city
Remap the classification to trees vs other
for all feature ∈ place_collection do
    Calculate priority indicators
end for
Calculate estimated canopy cover, canopy cover goal, and GAP
  score (Eq. 1)
Calculate citywide max and min for each indicator
for all feature ∈ place_collection do
    Normalize each indicator (Eq. 2)
    Calculate tree equity score (Eq. 4)
end for
Output: places_collection with an additional column for each
  indicator as well as the gap, priority, and tree equity scores.
```

The bulk of the analysis is made up of simple mathematical operations performed for all elements of a collection[2]. In GEE we write each of these calculations as a function that operates on a single feature and map these functions over the whole feature collection.

For each collection, we create a boolean tree cover map that covers the extent of the city (see section 4.4.1 for details). We then construct the input priority indicators from the raw census variable estimates: for example, we calculate unemployment rate by dividing the unemployed population by the total population. See the indicator construction guide on the project Github page for details on how each indicator is calculated[4].

Next, we calculate the percent tree cover of each block group. This is the most computationally intensive part of the analysis, and in order to prevent it from failing because of server-side time or memory limits we aggregate the image to 4m pixels before performing this step.

Finally, we calculate the citywide maximum and minimum values for each indicator, normalize the indicators, and use the resulting values to calculate the final tree equity score (see sections 4.3.2 through 4.3.4 for details).

[1]In practice, GEE performs all server-side computation in parallel, so this only a loose assessment of runtime.
[2]These are known as 'field calculations' in many desktop GIS programs
[4]See note 1 on page 3

	Actual Class							
	Land Cover Category	Tree	Grass	Building	Road	Dirt	Water	Total
Assigned Class	Tree	12						12
	Grass		13					13
	Building			13				13
	Road				10			10
	Dirt					17		17
	Water						15	15
	Total	12	13	13	10	17	15	80

Table 1: Confusion matrix for land cover classification, where entries along the diagonal indicate correctly classified points and entries off of the diagonal would indicate misclassified points. Overall accuracy is the count of all correctly classified points divided by the total number of points, or in this case (12+13+13+10+17+15)/80 = 1

5 RESULTS

This section should include the following:

- Details and explanations of results obtained (which itself could have sub-sections). This is where you should provide tables, graphs, and/or figures that illustrate your results. If you have created a new application, include screenshots of it in action. You should also provide a link to your application if it is web-based.

You should entitle these sections and sub-sections with names that describe the key points (for example, instead of "methods we use", the heading "Statistical-Based Learning" would be more informative). The Methods and Results sections should together be approximately 3–4 pages in length.

5.1 Classification Accuracy

We attempted to assess the accuracy of our classifier for our test region of Manchester, NH by building a confusion matrix for both our land cover classification and our boolean tree cover map. However, the limited number of validation points and the way the points were collected produced a listed accuracy of 100%, which is undoubtedly incorrect (Table 1).

Our classifier frequently confuses buildings with roads, light colored buildings with dirt, and building shadows with water (Figure 1). However, it generally does a good job at its intended purpose of distinguishing trees from all other categories. This is likely because buildings, roads, dirt, and water can all be visually and spectrally similar in the right circumstances, but all are quite visually and spectrally different from vegetation. In addition, including NDVI in the classifier input but not including other band ratios (such as the bare soil index) biases the classifier towards identifying vegetation.

5.2 Analysis Results

In order to assess the accuracy of our results, we visualize the difference between the original Tree Equity Scores and our replication. The original Tree Equity Score data can be downloaded on a state-by-state basis from their website, and these data include the final score, the priority score, and the demographic data that goes into the priority score, but do not include the estimated canopy cover. After downloading the Tree Equity Score results for each state in New England, we used QGIS 3.16 to merge these layers together

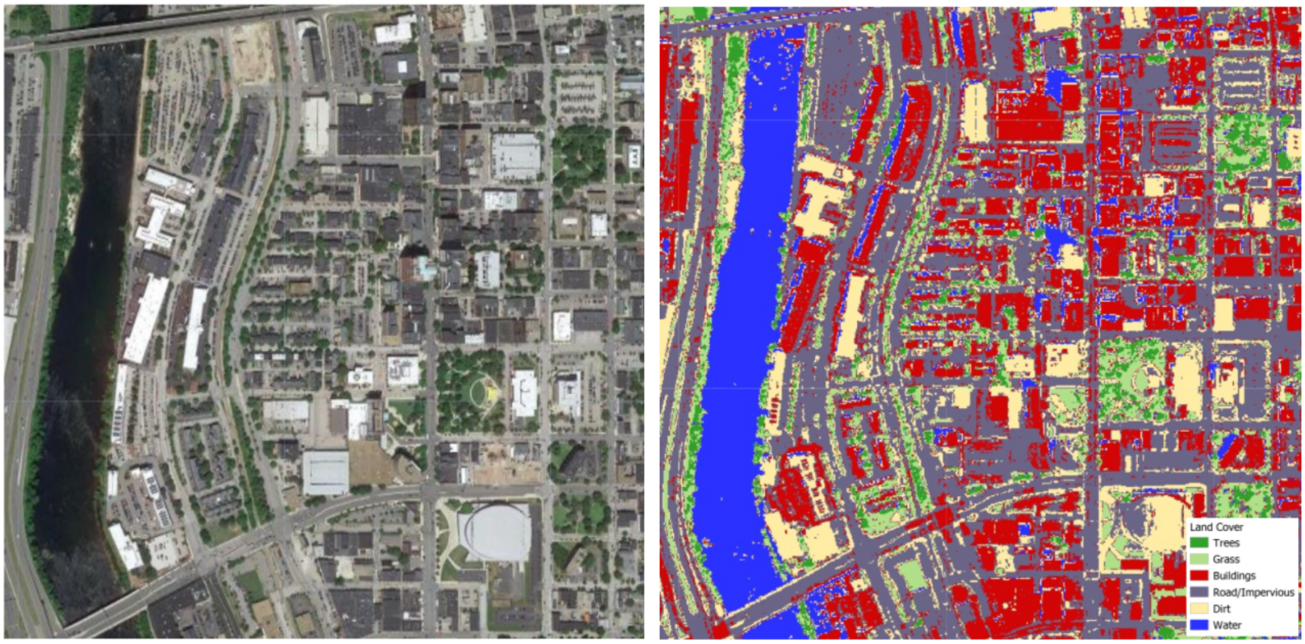


Figure 1: Left: True color satellite image of downtown Manchester, NH. Right: Sample classified image for the same region. Note the errors in distinguishing buildings from roads and areas that are misclassified as water.

and join the original data to our replication data by block group GEOID. We then subtracted the original scores from our replicated scores.

The results of our analysis show broad agreement between the original Tree Equity Score data and our replication (Figure 2). In a sample region of Boston, Massachusetts, our scores were within 10 points of the original scores for 87% of block groups (Figure 2c)>. For the entire study region, this number was 75% (Figure 2d).

Comparing the priority scores in the same way yields similar results: 71% of our priority scores were within a 10% error margin of the original scores. Given that we used the same demographic, socioeconomic, and health data, this suggests that the urban heat island data in combination with the different units of analysis has a modest influence on priority scores in New England.

Although the accuracy of our classifier is unknown, it appears to be serviceable but limited. Given this limited accuracy and our deviations from the original workflow, our results suggest that the Tree Equity Score metric is robust to changes in input data.

6 DISCUSSION

Our results show general agreement between the original Tree Equity Score analysis and our reproduction, though differences in the input data and units of analysis complicate this comparison. Further work is needed in order to investigate exactly where our analysis differs from the original analysis. It may be possible to reconstruct the original canopy cover data from the provided gap scores, which would shed more light on how much of the difference in results is due to differences between canopy cover estimates versus differences between priority scores. In addition, re-running our analysis with the correct Incorporated Place units would make it easier to directly compare our results with the original results.

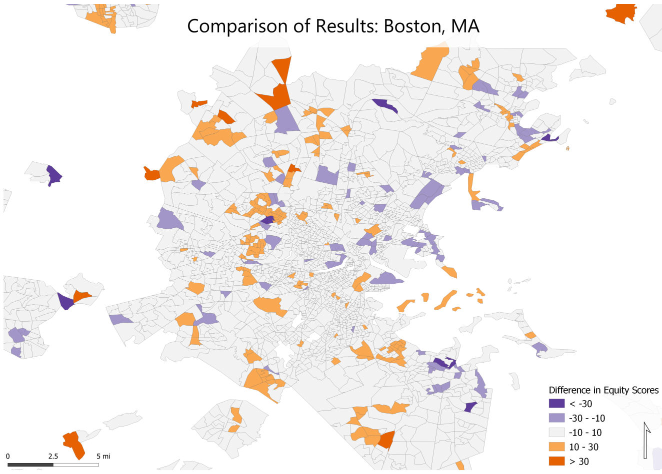
There were several directions we would have taken our project if our scope had not been limited by the semester. In accordance with our goal of displaying as much of our data as possible, we wished to add togglable layer to the map to display our satellite data to convey how our image processing works. Furthermore, we would like to feature all our raw data on the website, available for download., including the results of our classification and the accuracy of our analysis.

Otherwise, we were always hoping to cover the entire continental United States in our data and interactive map. Due to time limitations, we were only able to expand to include New England from our original sample of Manchester, NH. One factor preventing scalability is the need for more and more diverse training data so that our image classifier a) becomes more accurate and b) can process satellite data from outside of New England, as all our current training data resides in New Hampshire now.

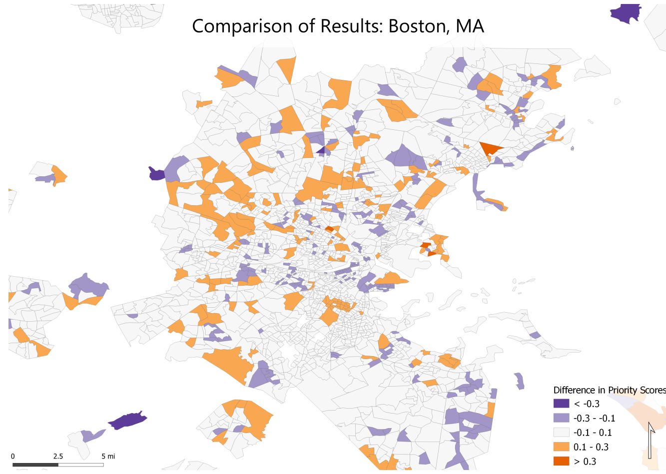
Finally, in order to perfectly replicate Tree Equity Score's analysis, we need to include temperature data. This data emphasizes the Priority Score in areas with a large urban heat island effect, raising the temperatures and worsening living conditions. This analysis was outside of our scope, and we were able to accurately reproduce their findings without it.

REFERENCES

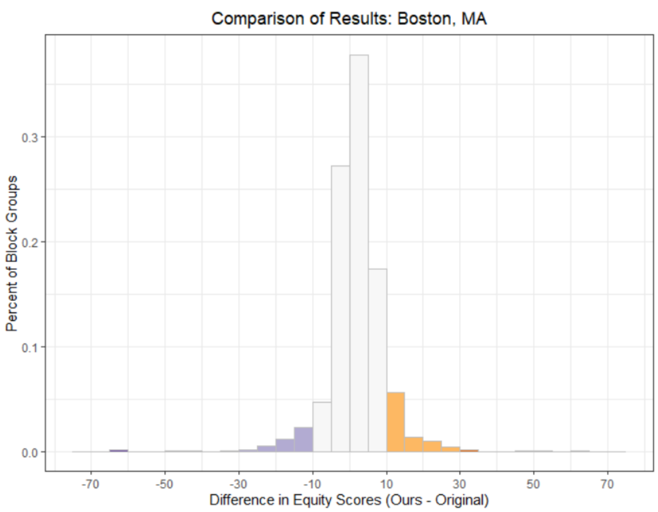
[1] C. Yu and W. N. Hien, "Thermal benefits of city parks," *Energy and Buildings*, vol. 38, no. 2, pp. 105–120, 2006. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778805000794



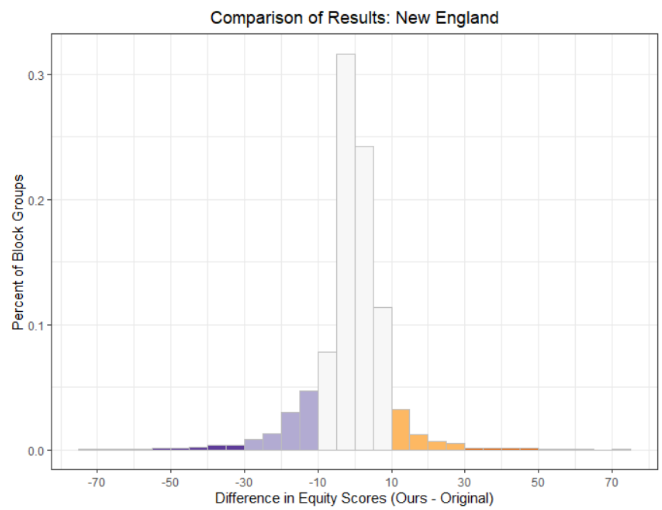
(a) ***Difference between original and replicated tree equity scores.***



(b) ***Difference between original and replicated priority scores.***



(c) ***Distribution of tree equity score difference in Boston (n = 1821).***



(d) ***Distribution of tree equity score difference for New England (n = 6475).***

Figure 2: Comparing our replication results to the original Tree Equity Score data. 2a, 2b, and 2c compare results for a sample region of Boston, MA, while 2d compares results for the entire study region of New England. All differences are calculated as Difference = Our Score - Orginal Score

[2] J. Maas, R. A. Verheij, P. P. Groenewegen, S. de Vries, and P. Spreeuwenberg, "Green space, urbanity, and health: how strong is the relation?" *Journal of Epidemiology & Community Health*, vol. 60, no. 7, pp. 587–592, 2006, publisher: BMJ Publishing Group Ltd _eprint: https://jech.bmj.com/content/60/7/587.full.pdf. [Online]. Available: https://jech.bmj.com/content/60/7/587

[3] J. R. Wolch, J. Byrne, and J. P. Newell, "Urban green space, public health, and environmental justice: The challenge of making cities 'just green enough'," *Landscape and Urban Planning*, vol. 125, pp. 234–244, May 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169204614000310

[4] American Forests, "Tree equity score: Ensuring tree cover in cities is equitably distributed," 2021, https://www.americanforests.org/tools-research-reports-and-guides/tree-equity-score/, Last Accessed on 2021-12-17.

[5] H. Wüstemann, D. Kalisch, and J. Kolbe, "Access to urban green space and environmental inequalities in Germany," *Landscape and Urban Planning*, vol. 164, pp. 124–131, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169204617300750

[6] Y. Shen, F. Sun, and Y. Che, "Public green spaces and human wellbeing: Mapping the spatial inequity and mismatching status of public green space in the Central City of Shanghai," *Urban Forestry & Urban Greening*, vol. 27, pp. 59–68, Oct. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1618866716303351

[7] J. Zhang, W. Yue, P. Fan, and J. Gao, "Measuring the accessibility of public green spaces in urban areas using web map services," *Applied Geography*, vol. 126, p. 102381, Jan. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0143622820314806

[8] C. Huang, J. Yang, N. Clinton, L. Yu, H. Huang, I. Dronova, and J. Jin, "Mapping the maximum extents of urban green spaces in 1039 cities using dense satellite images," *Environ. Res. Lett.*, vol. 16, no. 6, p. 064072, Jun. 2021, publisher: IOP Publishing. [Online]. Available: https://doi.org/10.1088/1748-9326/ac03dc

[9] A. Puissant, S. Rougier, and A. Stumpf, "Object-oriented mapping of urban trees using Random Forest classifiers," *International Journal of Applied Earth Observation and Geoinformation*, vol. 26, pp. 235–245, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0303243413000822

[10] T.-B. Ottosen, G. Petch, M. Hanson, and C. A. Skjøth, "Tree cover mapping based on Sentinel-2 images demonstrate high thematic accuracy in Europe," *International Journal of Applied Earth Observation and Geoinformation*, vol. 84, p. 101947, Feb. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0303243419306087

[11] C. Ludwig, R. Hecht, S. Lautenbach, M. Schorcht, and A. Zipf, "Mapping Public Urban Green Spaces Based on OpenStreetMap and Sentinel-2 Imagery Using Belief Functions," *ISPRS International Journal of Geo-Information*, vol. 10, no. 4, p. 251, Apr. 2021, number: 4 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2220-9964/10/4/251

[12] Tree Equity Score, "Data sources," 2021, https://treeequityscore.org/datasources, Last Accessed on 2021-12-15.

[13] ——, "Methodology," 2021, https://treeequityscore.org/methodology, Last Accessed on 2021-12-15.

[14] ——, "Faq," 2021, https://treeequityscore.org/faq/, Last Accessed on 2021-12-15.

[15] EarthDefine, "US Tree Map: Seamless high-resolution tree canopy data," 2021, https://www.earthdefine.com/treemap/, Accessed on 2021-12-16.

[16] US Census Bureau, "American Community Survey 5-Year Data (2009-2019)," 2021, https://www.census.gov/data/developers/data-sets/acs-5year.html, Accessed on 2021-12-15.

[17] US Census Bureau American Community Survey Office, "ACS Sample Size," 2021, https://www.census.gov/acs/www/methodology/sample-size-and-data-quality/sample-size/, Accessed on 2021-12-15.

[18] Centers for Disease Control and Prevention, "PLACES: Census Tract Data (GIS Friendly Format), 2020 release | Chronic Disease and Health Promotion Data & Indicators," eSRI Shapefile. [Online]. Available: https://chronicdata.cdc.gov/500-Cities-Places/PLACES-Census-Tract-Data-GIS-Friendly-Format-2020-/ib3w-k9rq

[19] Tree Equity Score, "Data: Raw data downloads of the geospatial data that powers tree equity score," 2021, https://treeequityscore.org/data, Last Accessed on 2021-12-08 as ESRI Shapefile.

[20] ——, "Data dictionary," 2021, https://treeequityscore.org/data/. Note: File data_dictionary.txt provided with New Hampshire data.

[21] US Department of Agriculture, "NAIP, NRI and Resource Information Maps," 2021, https://fsa.usda.gov/programs-and-services/aerial-photography/status-maps/index.