2020

Testing Park Cool Island GIS Analysis Methods For Use In Semi-Urban Conservation Planning

David Heacock

TESTING PARK COOL ISLAND GIS ANALYSIS METHODS

FOR USE IN SEMI-URBAN CONSERVATION PLANNING

by

David Heacock

A thesis submitted to the Faculty of Salem State University in partial fulfillment of the
requirements for the degree of Master of Science in Geo-Information Science

Summer 2020

Signature and Approval Form

The following thesis has been read and approved by:

Marcos Luna   August 13, 2020

Professor Marcos Luna, Ph.D., Department of Geography, Thesis Director   Date

Stephen Young   8-25-2020

Professor Stephen Young, Ph.D., Department of Geography, Second Reader   Date

Abby Hardy-Moss   August 24, 2020

Abby Hardy-Moss, M.A., Department of Geography, Third Reader   Date

Student Name: David Heacock

Student ID#:   S0333199

Table of Contents

List of Figures

List of Tables

List of Acronyms

ARD – Landsat Analysis Ready Data

BT – brightness temperature

CS – cooling strength

DN – digital number

GIS – geographic information systems

GST – green surface temperature

LSI – landscape shape index

LST – land surface temperature

NCI – normalized compactness index

NDVI – normalized difference vegetation index

PCI – park cool island

PVSI – park vegetation and shape index

TOA – top of atmosphere

UHI – urban heat island

Abstract

The human and environmental impacts of urban heat islands (UHI) have become an increasingly relevant issue to city planners. This topic has spurred research into the relationships between land cover, ambient temperature, and the role of greenspace in emitting cooler air to its surrounding area, now known as the 'park cool island' effect (PCI). While ample research has been given to this phenomenon in dense urban areas, much less has been dedicated to semi-urban communities who may wish to inform their development practices as they expand their footprints. This research used satellite-derived Landsat Level-2 Provisional Surface Temperature data, MassGIS 2016 Land Cover / Land Use data, and MassGIS Standardized Assessors' Parcels data to analyze parcels in Essex County, Massachusetts, for PCI intensity and the influence of land cover and parcel characteristics on PCI. LST data from July, 2016, were used to evaluate the mean temperature difference between parcels and their surrounding area to derive PCI. Replicating methods demonstrated by Cao et al. [Landscape and Urban Planning, 96(4):224-231 (2010)], linear regression analyses were undertaken to determine the relationships between PCI, parcel land cover and geometry. The 500 meter buffer distance used by Cao et al. to calculate PCI was also analyzed. Twenty iterations of the linear regression model were run based on a changing buffer value to calculate PCI. Two sensitivity analyses were performed based on these model iterations: 1) change in model performance, as expressed by its R^2 value, across the range of PCI buffer distances and 2) the change in the coefficient strengths of the independent variables across the range of PCI buffer distances. The linear regression model underperformed as compared to Cao et al.'s study, however, it affirmed the 500 meter buffer distance as a parameter for calculating PCI, with that model iteration returning the highest R^2 value (0.587). Buffer distances greater than 500 meters performed relatively well, however, smaller buffer

values were associated with weak model performance. Among land cover coefficients, there were scale-sensitivities observed, with some variables changing in strength and polarity across the model iterations. It was determined that PCI could effectively evaluate cooling intensity in the study area, however, using it as a dependent variable within a linear regression model had only moderate performance. This was due to heterogeneity among the makeup of land cover within parcel buffer areas which inhibited the regression model's ability to build consistent relationships between land cover and PCI.

Introduction

The impacts of climate change are becoming more and more apparent in urban communities. An excess amount of impervious surface creates an environment where much of the sun's energy is retained, leading to higher ambient temperatures. This temperature increase can be anywhere from 2-5°C in comparison to rural surroundings, and is now commonly referred to as the urban heat island (UHI) effect (Taha 1997).

Heat islands have a human impact, as they increase the prevalence of heat-related illness, and can have environmental impacts if divergent microclimates develop (Shishegar 2014). In addition, higher temperatures accelerate smog formation, and can increase peak electricity demand by 5-10% as residents need to further offset raised temperatures with air conditioning (Mihalakakou et al. 2004; Akbari et al., 1992). In the long term, these reactions perpetuate the atmospheric changes that are at the source of climate change. While the human and environmental costs of elevated urban temperatures are evident, it can be a challenge to develop strategies aimed at mitigating the heat island effect.

Vegetation plays an important role in mitigating urban heat. Planting street trees can be a helpful way of preventing sunlight from reaching impervious surfaces in the first place by creating shade over buildings and streets. In addition, evaporation of moisture from soil and leaves cools ambient temperatures, a process known as evapotranspiration (EPA 2019). However, the success rate of saplings growing into maturity is often low because of the challenges of an urban environment (Roman and Scatena 2011). When they do reach maturity, their canopy will still be below tall buildings which limits their effectiveness in many urban areas. Despite these challenges, planting initiatives have been shown to decrease city temperatures significantly and are a cost-effective approach (The Nature Conservancy 2016).

Green and cool roofs are often able to provide the coverage that street trees cannot. Green roofs deliver similar shading and evaporative benefits to street trees, but are more easily planted on the roofs and walls of existing structures. In places where green roofs are not a viable option, cool roofs, which are made using synthetic reflective materials, can be installed to provide similar results. Use of these roofing alternatives have shown to be worthwhile investments for both urban planners and building managers (Sharma et al. 2016).

Considering the cooling effect that vegetation provides, it becomes apparent that urban parks can offer a place of refuge from high temperatures. Depending on the land cover and environmental conditions, urban parks have been measured to be up to 7°C cooler than their urban surroundings, earning the term "park cool islands" (PCI) (Jauregui, 1990). Additionally, if the vegetation is of high enough density, this cooling effect extends beyond the boundaries of the park and lowers the ambient air temperature of the surrounding neighborhood. While the magnitude is wind-dependent, there are observed cases of parks providing a cooling effect up to 840 meters away (Doick, Peace and Hutchings 2014). This ability brings added value to urban forest areas, as their removal would further elevate surrounding ambient temperatures. This has paved the way for many research projects, as scientists seek to understand the phenomenon so that these cooling effects can be fully utilized.

A number of studies have analyzed the cooling effects of urban parks around the world using similar methods to measure the influence of land cover on temperature. Far fewer studies, however, have applied such methods to semi-urban or suburban study areas that may not immediately face urban heat island challenges, but could if there is a future increase in impervious surface. Understanding heat islands in this conservationist context would help to make proactive rather than reactive decisions in the face of climate change.

This analysis applies methods used to analyze PCI intensity in dense, urban settings to a study area composed of semi-rural communities to answer the questions: (1) Do analysis methods commonly used to study PCI within urban areas produce similar outcomes in semi-urban areas? (2) How sensitive are PCI results to the buffer parameter around parks?

Literature Review

The Impacts of Urban Heat Islands:

The role that urban heat islands play in influencing local air temperature has been observed for decades (Akbari and Kolokotsa 2016). This has largely been attributed to dense amounts of impervious surface which are used in construction of urban infrastructure, and a lack of vegetation to offset this change in land cover (Taha 1997). Impervious surface refers to materials, such as concrete, brick and asphalt, which resist the absorption of water and create excess runoff (Leopold 1968). Besides their hydrological impacts, these materials also share the characteristic of absorbing a large proportion of the sunlight they receive (Morabito, Crisci, and Georgiadis 2017), which in turn significantly increases the ambient air temperature in the area (Arnfield 2003; Ahmed 2018).

As a result, cities have been observed to be a much as 12°C (22°F) hotter than their surrounding rural area due to their large volume of impervious surface (EPA 2019), which has been shown to have negative impacts on human health (Heaviside, Macintyre, and Vardoulakis 2017; Smargiassi et al. 2009) and species diversity (Čeplová, Kalusová, and Lososová 2017; McGlynn et al. 2019), heightens the intensity of air pollution (Sarrat et al. 2005), and increases peak electricity demand (Santamouris and Kolokotsa 2015).

The health risks associated with urban heat islands have been well documented across a number of different geographic contexts (Medina-Ramón et al. 2006; Smargiassi et al 2009; Heaviside, Macintyre, and Vardoulakis 2017), and have become a global concern in light of the threats of climate change (Gosling, McGregor, and Páldy 2007). The presence of urban heat islands increases the severity of naturally occurring heat waves, bringing with them greater occurrence of illness such as heat stroke, and an overall increase in mortality (Semenza et al. 1996). It has

been observed that city residents in areas of low vegetation have up to a 6% higher risk of morbidity or mortality as compared to residents in cooler areas (Schinasia et al. 2018). An increasing frequency of heat waves due to climate change is projected to have a significant impact on heat-related mortality over the next century (McGeehin and Mirabelli 2001), with the northeast United States facing up to a three-fold increase in mortality as compared to current levels (Petkova et al. 2013).

The negative health impacts of urban heat islands have been shown to disproportionately impact vulnerable and marginalized populations (Reid et al. 2009; Sampson et al. 2013). Elderly and chronically ill populations face a higher risk of heat-related mortality (Gronlund et al. 2014), in some cases up to 12% greater risk than the general population (Zanobetti et al. 2013). Residents living in neighborhoods with greater racial diversity, less education, and extreme poverty have been shown to have significantly higher heat-related mortality (Voelkel et al. 2018). These neighborhoods tend to have lower park accessibility and overall less vegetation coverage to aid in cooling the impervious surface (Rigolon, Browning, and Jennings 2018), and in general, have less resources available to develop strategies around urban heat island mitigation (Zukin et al. 2009). Because of these aspects, the topics of urban heat island formation and mitigation have emerged as points of focus in the discussion of environmental justice (Chakraborty 2017).

Observing Land Surface Temperatures:

UHI effects can be measured in a variety of ways. Traditionally, they have been observed in the field using thermometers at fixed stations, or mounted in mobile vehicles that capture temperature readings as they traverse a study area (Bowler et al. 2010; Song and Li, 2010). While it is preferable to capture the air temperature directly, this approach can be labor intensive and is unlikely to provide comparative measurements across many locations at a single time. Many heat island researchers have instead moved towards the use of land surface temperature (LST) data collected from the thermal-infrared band of satellite sensors because of the low cost and temporal advantages (Zhou et al. 2018).

While thermal-infrared satellite data can be an effective stand-in for ambient air temperature, it requires a conversion process from the digital number (DN) value captured by the sensor before being used in analysis (Pelta, Chudnovsky, and Schwarts 2016). The data also require correction according to atmospheric conditions, especially when performing interpretation between captures by different sensors (Chander, Markham, and Barsi, 2007). The following formulas are those laid out by NASA's Landsat program for use in converting thermal data for use in LST analysis (U.S. Geological Survey 2014).

Conversion to Top of Atmosphere (TOA) Radiance:

$$L_\lambda = M_L Q_{cal} + A_L \quad (1)$$

where L_λ is TOA spectral radiance, M_L is the band-specific multiplicative rescaling factor, A_L is the band-specific additive rescaling factor, and Q_{cal} is the quantized and calibrated DN pixel value. The brightness temperature value in Celsius can then be derived using the inverse Planck's function.

Conversion to At-Satellite Brightness Temperature (Celsius):

$$T = \frac{K_2}{\ln\left(\frac{K_1}{L_\lambda} + 1\right)} - 272.3 \quad (2)$$

where T is temperature in Celsius, K_1 is calibration constant 1, K_2 is calibration constant 2, ln is the natural logarithm, and L_λ is the TOA spectral radiance. The resulting data from these algorithms can then be used to represent the land surface temperature in imagery and spatial analysis.

There is some concern about whether using satellite land surface temperature data as a stand-in for air temperature is sufficiently accurate. A study by Lin et al. in 2014 aimed to assess the suitability of land surface temperature data for use in analyzing urban heat islands. In the study, Lin collected ambient air temperature at 24 on-site observatories and compared these samples with satellite data at those same locations to test their interchangeability (Lin et al. 2014). A statistical analysis revealed they were related to each other with a correlation coefficient of 0.81. This helped to reinforce the efficacy of using LST data as a proxy for air temperature to observe urban heat islands.

Some of the unexplained variation in Lin's study may have to do with the nature of the way LST data are collected. Each pixel value represents an average of thermal-infrared reflectance within a pixel's area of observation, while also being influenced by the thermal reflectance of surrounding pixels (Li 2017). Despite these errors in thermal-infrared LST data, they still provide sufficient accuracy for use in remote sensing applications, so long as they are not pushed beyond their spatial resolution limits.

Some analytical contexts, such as within public health, may require a greater degree of accuracy than can be derived from thermal infrared satellite imagery. For instance, there may be particular air temperature thresholds that are statistically significant when trying to prevent heat stroke illness (White-Newsome et al. 2013). In such an application, the 0.81 correlation coefficient cited by Lin et al. (2014) may not be accurate enough. These levels of accuracy or precision are not requirements of analyzing urban heat island effects at a parcel scale, however, which has allowed remotely sensed data to see widespread use in geospatial research (Zhang et al. 2017; Li et al. 2017; Lipping 2003). While site specific observations must be carefully quality-checked, land surface temperature can provide accuracy across geographic regions which makes it suitable for studying relative land cover relationships.

Park Cooling Effect Research:

The 'park cooling effect' has been studied by many researchers over the years to better understand the relationships between park characteristics and their cooling intensity. These studies have utilized a variety of different data to analyze parks, ranging from bicycle-mounted air temperature sensors, to weather stations, airplane mounted infrared sensors, and satellite imagery. Regardless of the data used, their observations have helped to frame park cooling effect discussions in regards to land cover, climate, moisture, and temporal changes. Researchers continue to ask: what influence does greenspace have on cooling urban environments, and what factors are most critical to this phenomenon?

Jauregui's 1990 study exemplified the park cool island effect at Chapultepec Park in Mexico City (Jauregui 1990). The author recorded air temperature, humidity, pollution, and rainfall data at weather stations both inside and outside the park, and analyzed their hourly and seasonal changes. By collecting these data, the study sought to observe the park cooling effect, but also how much of the effect could be attributed to vegetation and what influence the external environment may be having.

 The author found that minimum park temperatures tended to be 3 – 4°C cooler than the surrounding area, and its cooling influence extended approximately the distance of the park's width. This relationship between park width and cooling extent has been frequently cited by other researchers and served as a default value in their analyses (Spronken-Smith and Oke 1998; Hien and Yu 2004; Cao et al. 2010).

It was also found that the vegetated park had more rapid heating and cooling rates each day as compared to the neighboring impervious surface. While the built-up urban areas were slower to

warm up in the morning sun, they were also much slower to come down from peak temperatures in the afternoon. Jauregui found that peak temperatures were actually slightly higher in the vegetated park, but once that peak was passed, ended up cooling the adjacent area as temperatures dropped. The author also noted the influence of tall tree cover on reducing low-level wind speeds and creating turbulence. This in turn created an environment conducive to small-scale convective precipitation, which the study found to be more common within the park area.

Spronken-Smith and Oke's 1998 analysis of two parks located in Sacramento and Vancouver, BC, placed increased attention on various land cover types, and the influence that each had on cooling (Spronken-Smith and Oke 1998). Land cover was classified into the following categories: grass, grass with tree border, savannah, golf course, garden, multiuse (developed open space), and forest.

The study performed traverses on automobiles and bicycles, collecting continuous air temperature data beginning from the center of a park to its perimeter, adjacent urban neighborhoods, and outer rural area. A helicopter equipped with an infrared scanner was used to capture land surface temperature data across these same areas, at a spatial resolution of 3 meters. The authors defined park cooling intensity as the maximum urban temperature minus the minimum park temperature captured during the first half of a traverse, represented as the formula:

$$PCI = T_u - T_p \quad (1)$$

where T_p is the minimum air temperature collected on a traverse inside of a park, and T_u is the maximum air temperature collected as that traverse passed through the neighboring urban area.

The urban heat island effect was then calculated using the second half of a traverse, comparing urban to rural temperatures:

$$UHI = T_u - T_r \quad (2)$$

where T_r is the minimum air temperature collected at the end of a traverse in the undeveloped, rural area, and T_u is the maximum urban temperature. These values were generated for each traverse, so that they could be used to compare the land cover attributes of each park.

Similar to Jauregui's study, the authors observed cooling extents of parks approximate to their widths and that the effect varied by the time of the day. In addition, they were able to demonstrate the cooling impacts of the different land cover types. It was found that tree canopies provide maximum cooling effect in the afternoon (4°C), gardens and savannahs at dusk (3.2°C), and open grass at sunrise (2.9°C).

The authors put forward the theory that the moisture content of a land cover type has a significant impact on the timing of its cooling effect. Whereas irrigated greenspace has a relatively stronger cooling effect in the afternoon, dry grass parks cool down much more quickly after sunset due to their low moisture content. They point out that the influence of moisture is critical to understanding land cover, but also that the park cooling effect will function differently based on the macroscale climate of the study region.

As continued studies replicated these and satellite LST data became more available (allowing for easier comparison of many parks at a single capture time), researchers focused less on observation studies of singular parks, and more on comparisons among large numbers of parks within a geographic setting. Cao's paper "Quantifying the cool island intensity of urban parks

using ASTER and IKONOS data" compared the characteristics of 92 parks in Nagoya, Japan and their influence on park cooling intensity (Cao et al. 2010).

Cao et al. (2010) used Spronken-Smith and Oke's (1998) definition of park cooling intensity as the temperature difference between the inside and outside of a park, but used a 500 meter buffer zone (with water and other parks removed), and substituted air temperature with land surface temperature. The authors do not provide particular reasoning for this 500 meter parameter, but they express that their intention is to include thermal information from surrounding roads, residential and business buildings, and parking spaces. This would indicate that the aim of this method is to analyze the area beyond the cooling extent of a park, so that the thermal levels from impervious surface areas can be fully contrasted with the vegetated land covers of parks. One could presume that if this buffer zone were much narrower, such as at a distance of 100 meters or less, there could be a high degree of cooling influence from a park's vegetation on the thermal readings.

The study also incorporated a 'landscape shape index' that quantified the compactness of a park using its perimeter length and area. The formula used was:

$$\text{LSI} = \left(\frac{P_t}{2\sqrt{\pi \times A}} \right)$$

where P_t is the length of the perimeter of a parcel, and *A* is the area of a parcel. The index expresses the relationship between the size of a polygon and the length of its border, such that identically shaped polygons receive the same value regardless of their area. This enables Cao's analysis to use park size and shape as distinct variables within their statistical model, allowing their influence on park cooling to be assessed separately.

Areal coverage of various land covers were calculated using satellite imagery. The authors used a 4-meter spatial resolution IKONOS image with four bands in the visible and near-infrared bands (for spectral analysis of different vegetation types) to manually classify the land covers. The resulting land cover classes were: trees, grass, shrubs, soil, water, low albedo surfaces, high albedo surfaces, buildings, and shadow cover. Low albedo surfaces refer to those that absorb more electromagnetic radiation than they reflect, such as concrete and other impervious surfaces. These land cover values were expressed as percentages of coverage within a park. This is necessary for the park size variable to be included in the model.

With these data as independent variables, the authors performed a linear regression analysis with park cooling intensity as the dependent variable. It was found that cooling intensity is greater in larger parks, and only those larger than 2 ha (4.942 acres) consistently provided cooling benefits. The coefficients of determination (R^2) for the regression model were 0.888, 0.888, and 0.798 for spring, summer, and fall, with the most important variables including area of trees (2.315), shrubs (1.165), grass (-1.972), and shape index (-1.406). The authors felt strongly enough about these few variables, that they created a new "park vegetation and shape index" (PVSI) defined as:

$$\text{PVSI} = \log_{10}\left(\frac{A_{tree} + A_{shrub}}{LSI}\right)$$

where A_{tree} and A_{shrub} are the areas trees and shrubs cover within a park, and LSI is the landscape shape index of a park. Using leave-one-out cross validation they found this index predicted park cooling effect with a root mean square errors of 0.98, 0.94, and 0.71 for spring, summer, and fall, and concluded that these few variables were predictive of park cooling intensity in their study region.

Performing analysis at this scale, and with such large a number of parks, produced results that the authors felt would help urban planners to better understand what conditions are necessary for park cool islands to form. This includes what the temporal advantages and disadvantages are of different land covers for locations sharing the similar climates to the study area (i.e., Nagoya, Japan).

Others have used remote sensing techniques to analyze the configuration of green space within a city and its influence on city temperatures. Zhang's 2017 study explored this dynamic in Beijing, China, using satellite imagery from Chinese satellite Gaofen-2. The study looked at the influence of green patch configuration on their cooling effect, by investigating relationships between green surface temperature (GST), normalized difference vegetation index (NDVI), and the normalized compactness index (NCI). The authors then used linear regression with GST as the dependent variable, revealing a negative relationship to NCI with an R^2 value of 0.3156, which would indicate that if green patches in the city were more compact their temperatures would be cooler. They also found a negative relationship to NDVI with an R^2 value of 0.4328, indicating that higher NDVI values (as a result of more vegetation) would also reduce temperatures.

At the same time, the authors found that the cooling strength (CS) and NCI were negatively correlated with an R^2 value of 0.4689 and coefficient of -2.4661, which would indicate that when green patches are more compact, their cooling extent decreases. The authors also expressed that the compactness relationship to vegetation cooling effect changed depending on the area of the green patches, noting that patches ranging in size from 2 to 16 hectares (4.94 to 39.53 acres) were more sensitive to compactness than those that fell outside of this range. This study reveals a complicated relationship between vegetation and impervious surface within an urban

environment, but is important to consider when trying to leverage park cooling effect research

towards practical use in urban or conservation planning.

Methodology

As mentioned previously, the study of heat islands has predominantly been focused in locations with dense, urban development. These kinds of analyses have been helpful to reinforce the importance of greenspace on urban environments, measure the thermal influence of impervious surface, and inform development practices. While these aspects are especially critical to large urban centers, they are also of concern to smaller cities that are expanding their footprint, with opportunities to either preserve or develop existing greenspace. Current and predicted threats from climate change have become more of a concern to communities, and government grants have incentivized regional planners to take more proactive measures. This makes the management of heat island effects more relevant than ever to developing cities.

For this study, I looked to replicate the methods commonly used in current urban heat island analysis literature, and apply them to Essex County, Massachusetts. This approach helps to establish baseline relationships between regional land cover types and land surface temperature, and identify parcels which are particularly critical to cooling urban heat island areas. In addition, the buffer distance parameter surrounding study parcels was tested for a wide range of values to observe model sensitivity. The buffer distance for the park cooling effect evaluation of a parcel is often 500 meters, however, there is not always a robust explanation for this distance besides precedent in other literature. Seeing as the study area of this research differs from those in most other heat island literature, this parameter seemed important to explore in order to understand what role it plays within the model.

Study Area:

Essex County is the third most populous county in Massachusetts with a population of 789,034, and overall area of 1,339.58 km^2 (Census.gov 2019). It ranks fourth for population density in Massachusetts with 618.5 people per km^2. The county includes the cities of Salem, Lynn, Haverhill, Lawrence, Gloucester, and Newburyport. Weather records indicate an annual average of 1219 mm of rain and 1244 mm of snow per year (Northeast Regional Climate Center 2020). The area experiences a range of seasonal temperatures, with an average daily high of 27º C (80.6° F) in the summer and average highs of 2º C (35.6° F) in the winter.

The commonwealth of Massachusetts as a whole has seen a consistent increase in mean annual temperature, and is projected to experience a significant increase in days over 38º C (100° F) on the daytime heat index (Peery 2019), which takes into account humidity. According to Mass Audubon, the state experienced just 6 days a year with highs over 32.2° C (90° F) in the year 1960, but in 2020 that value has increased to 10 days a year (Mass Audubon 2020). They predict by 2050 this value will nearly double to 19 days a year on average. Such an increase in extreme heat days may bring higher prevalence of heat-related public health issues, such as heat stroke.

Essex County has a broad mix of land use, including dense residential and commercial, to agriculture and rural uses. According to the 2016 Land Cover dataset developed by MassGIS (MassGIS 2020) and the National Oceanic and Atmospheric Administration's (NOAA) Office of Coastal Management (OCM), the county is roughly 55.5% forested land cover (deciduous, evergreen, palustrine forested wetland, and estuarine forested wetland), 13.55% impervious surface, 9.87% developed open space (i.e. lawns, golf courses, soccer fields), with the remaining 21% a mix of palustrine and estuarine wetlands, cultivated, grasslands, and water.

Table 1: Essex County Land Cover

Land Cover	km^2	Percent
Deciduous Forest	491.82	36.71%
Impervious	181.58	13.55%
Developed Open Space	132.28	9.87%
Palustrine Forested Wetland	126.23	9.42%
Evergreen Forest	125.67	9.38%
Estuarine Emergent Wetland	69.52	5.19%
Water	50.24	3.75%
Palustrine Emergent Wetland	35.08	2.62%
Pasture/Hay	29.39	2.19%
Grassland	22.26	1.66%
Palustrine Scrub/Shrub Wetland	17.85	1.33%
Scrub/Shrub	15.69	1.17%
Cultivated	14.46	1.08%
Unconsolidated Shore	13.09	0.98%
Bare Land	11.02	0.82%
Palustrine Aquatic Bed	3.30	0.25%
Estuarine Scrub/Shrub Wetland	0.05	0.00%
Estuarine Forested Wetland	0.04	0.00%
Total	1,339.58	100.00%

Data:

The Landsat satellite system's multispectral scanners provide a large volume of data as they capture each point on earth roughly every 16 days. A wide range of light wavelengths are captured, from the visible spectrum (0.43 - 0.53 µm) up through the thermal infrared bands (10.6 – 12.51 µm) (NASA 2020). The spatial resolution varies depending on the wavelength of the band, from 15 meter resolution in the panchromatic band to 30 meters for bands within the visible spectrum, and 100 meters for thermal infrared. In addition to making these data freely and publicly available, the Landsat program has begun pre-processing data products for distribution

for use by researchers, referred to as the U.S. Landsat Analysis Ready Data (ARD) product bundle.

Among these include the Level-2 Provisional Surface Temperature product, which delivers land surface temperature in Kelvin at 30 meter resolution. Normally one would have to convert the digital number (DN) value of a Landsat image to spectral radiance (L), then to temperature (in Kelvin), and finally to Celsius. The Provisional Surface Temperate product preprocesses the first two steps, with only the temperature conversion needing to be performed.

A Landsat image captured on July 13th, 2016 at 10:26 a.m., was selected for use in this study. Landsat's satellite system utilizes a sun-synchronous, near-polar orbit, meaning every point on earth is captured at roughly this same time locally. While it may be preferable to capture land surface temperature data for this study at peak air temperature time in the afternoon, the thermal patterns should remain consistent among land cover types in the morning, and thus their relationships should remain stable as well. It should be noted, however, that one may observe different land cover relationships to land surface temperature at night, where forests retain heat longer than grassy fields due to differences in moisture content (Jauregui 1990). This particular image was selected for its timing in the summer months featuring full vegetation cover, it's relatively warm peak temperature of 31° C (88° F), and minimal cloud cover. The thermal infrared image from the ARD product series was selected, and converted to Celsius.

Detailed land use data were available from MassGIS in their 2016 Land Cover / Land Use product. This was a cooperative effort between MassGIS and the National Oceanic and Atmospheric Administration's (NOAA) Office of Coastal Management (OCM). The product provides high resolution land cover data for the entire state, following the standards set by the National Land Cover Database. Land cover types were derived using a mix of supervised

classification of a 1-meter spatial resolution raster image, MassDOT Roads spatial data, manual editing using 2013 orthoimagery, and feedback from local experts. Through a quality control process of validating 446 on-the-ground points, OCM staff determined the data accuracy to be 94.2% (Office for Coastal Management 2020).

The land use data documented 19 different land cover types, some distinct, and some representing a mix of multiple types. These are: Impervious, Developed Open Space, Cultivated Land, Pasture/Hay, Grassland, Deciduous Forest, Evergreen Forest, Scrub/Shrub, Palustrine Forested Wetland, Palustrine Scrub/Shrub Wetland, Palustrine Emergent Wetland, Estuarine Forested Wetland, Estuarine Scrub/Shrub Wetland, Estuarine Emergent Wetland, Unconsolidated Shore, Bare Land, Open Water, Palustrine Aquatic Bed, and Estuarine Aquatic Bed.

Parcel boundary data were made available from MassGIS' Level 3 standardized assessors' parcel program. This new standard was put into practice in October 2013, and has allowed for more accurate parcel boundaries and attributes, as well as simplified merging of data between towns. Statewide 'LOC_ID's are generated for each parcel using their centroid coordinates, which provides unique ID's. This helps to avoid querying parcels using their 'MAP_PAR_ID' or 'PROP_ID' which are derived from map, block, and lot codes and can be redundant if referenced across multiple towns. Data from the thirty four cities and towns of Essex County were downloaded from MassGIS, and merged into one layer. The land use attribute table was also joined to the parcels using the land use code for reference. These parcels were then pared down to a smaller number for inclusion into the analysis, creating a selection of 'parks' similar to those studied by Cao (Cao et.al 2010).

While this project aimed to replicate Cao's methods as closely as possible, implementing these

methods in a semi-urban study area required some adaptation. Cao's study did not require a park

selection process since the study area of Nagoya, Japan, is entirely developed, and the parks of

interest to be studied are self-evident through their land cover and public designation. Most green

areas that exist are likely due to intentional planning decisions and ongoing management by city

administration. While similar parks exist within Essex County, there are many more privately

owned green parcels adjacent to urban heat island areas that are at risk of being developed into

impervious surface. Seeing as the context of this study is inherently conservation-minded, this

approach was used to guide the parcel selection process.

In addition to the conservation backdrop behind this study, a second purpose behind this

selection process was to limit the scope of the analysis only to those parcels which were adjacent

to heat island areas. This was necessary in order to appropriately utilize the PCI formula within

this study area. The dependent variable of Cao's model (Cao et al. 2010), PCI, is calculated by

the difference between the internal and surrounding land surface temperatures of a parcel. In

dense, urban areas where this variable has been developed, the land cover is generally

homogenous, being composed of buildings and other impervious surface. Urban parks and

greenspace exist as the exceptions. In such a context the PCI value is calculated under a uniform

scenario – the study park includes a mix of vegetation and other land covers, and the area around

it is almost exclusively impervious surface.

This uniformity means that vegetation within a park will generally be shown as having some

cooling effect, and at worst will result in a value near zero. The independent variables are tied to

park characteristics that are able to establish relationships to a PCI value that is derived from a

predictable ratio. The challenge of using these methods in a study area where impervious surface

isn't dominant is that PCI doesn't just become a function of the composition of land cover within a parcel, but of the land cover surrounding it as well. For example, if a parcel in a dense, urban area was comprised entirely of greenspace, it would likely receive a positive PCI value. If this same parcel was instead situated in a rural area abutting other greenspace, it may receive a PCI value near zero, as there would be minimal temperature difference between it and its adjacent parcels. This aspect is shown in Figure 1, where two largely forested parcels receive considerably different PCI scores based on the different levels of impervious surface in their buffer area. The parcel selection process was created with interest in utilizing the PCI variable within its proper context.

Figure 1: Rural forest parcel and urban forest parcel

First, parcels were limited to towns in the county that have the densest impervious surface in the county. These towns included Andover, Haverhill, Lawrence, Lynn, Methuen, North Andover, Peabody, Salem, and Saugus. Several of these towns have expressed heat related threats due to climate change in their Municipal Vulnerability Preparedness reports. Second, a field was created to calculate the acreage of each parcel. As referenced in Cao's study, the minimum urban

park size to be observed to experience a cooling effect was 2 ha (4.942 acres). This value was used as a minimum threshold.

Next, a 500-meter buffer was generated around each parcel and the mean land surface temperature was calculated within that buffer. Parcels were selected if their buffers were one standard deviation or warmer than the mean. This standard deviation approach to define heat island areas has been used in several studies (Effat, Taha, and Mansour 2014; Y. Ma, Y. Kuang, and N. Huang. 2010). This resulted in a final sample set of 393 parcels that are located in communities facing heat island threats, are of large enough size that they have the potential to provide a cooling effect, and are adjacent to heat island areas. These selection methods help to appropriately utilize the PCI formula in this study area, and highlight parcels that could provide conservation value in relation to urban cooling.

PCI Analysis:

In order to analyze the park cool intensity of parcels the following methods were performed: (1) Evaluate study parcels for PCI intensity. (2) Calculate the sum of different land cover types, as well as the LSI and acreage for each parcel. (3) Perform multivariate regression using land cover, LSI, and parcel size variables to observe what influence land cover within a parcel has on cooling its surrounding area. (4) Re-run the analysis using a number of different buffer distances to observe how the model responds to this changing parameter.

These methods were scripted using the Arcpy package within Python 2.7 in order to ensure that they were replicable, as well as to increase the ease of performing the buffer distance sensitivity analysis. For this study, buffer distances from 50 to 1000 meters were analyzed in increments of 50 meters, for a total of 20 iterations.

Parcels were first buffered according to the specified distance, outside of polygon only. This created 'donut-hole' polygons surrounding each study parcel that enabled the calculation of PCI intensity. Parcel 'LOC_ID's provided the unique ID key for joining data back to the original parcel dataset.

PCI intensity was calculated using the formula utilized in Cao's study. That is:

$$PCI = T_U - T_P$$

where T_P is the mean land surface temperature within a parcel, and T_U is the mean land surface temperature of the buffer area surrounding a parcel. Land surface temperature data in the buffer area that falls within water land cover types, or that falls in another study parcel is excluded from the analysis, and does not contribute to the mean land surface temperature reading of T_U. This was performed with interest in adhering to Cao's methods as closely as possible.

The land surface temperature data were first converted from Kelvin to Celsius using the formula:

$$T_{(°C)} = T_{(K)} - 273.15$$

where $T_{(°C)}$ is the temperature in Celsius, and $T_{(K)}$ is the temperature in Kelvin. Once temperature values were processed, the data were converted from a raster to vector format using the ArcGIS Raster to Point (Conversion) tool.

With each point representing an individual pixel from the LST image, and the value of each representing a pixel's temperature, a spatial join was performed to select all points falling within a parcel. The values of each point were averaged in order to generate the mean LST in Celsius for each parcel. This same process was replicated for the parcel buffers. After removing points

that fell within water land cover areas or neighboring parcels, the mean LST values of each buffer polygon were joined back to their respective parcels, and PCI was calculated for each.

This process could have instead been performed using a zonal statistics tool to derive mean LST directly from the raster data, however it was found that in cases of overlapping polygons, the tool failed to return accurate results for all features. Spatially joining points was found to be consistently accurate, though LST pixels were only assessed if their centroid fell within a parcel, as that's where the points are generated. One could instead convert the raster to a polygon grid in order evaluate subsets of individual pixels, however, calculating the LST mean across a parcel did not require this level of precision. A vector-based approach also allowed for easy querying of feature attributes, though it was relatively more computationally heavy.
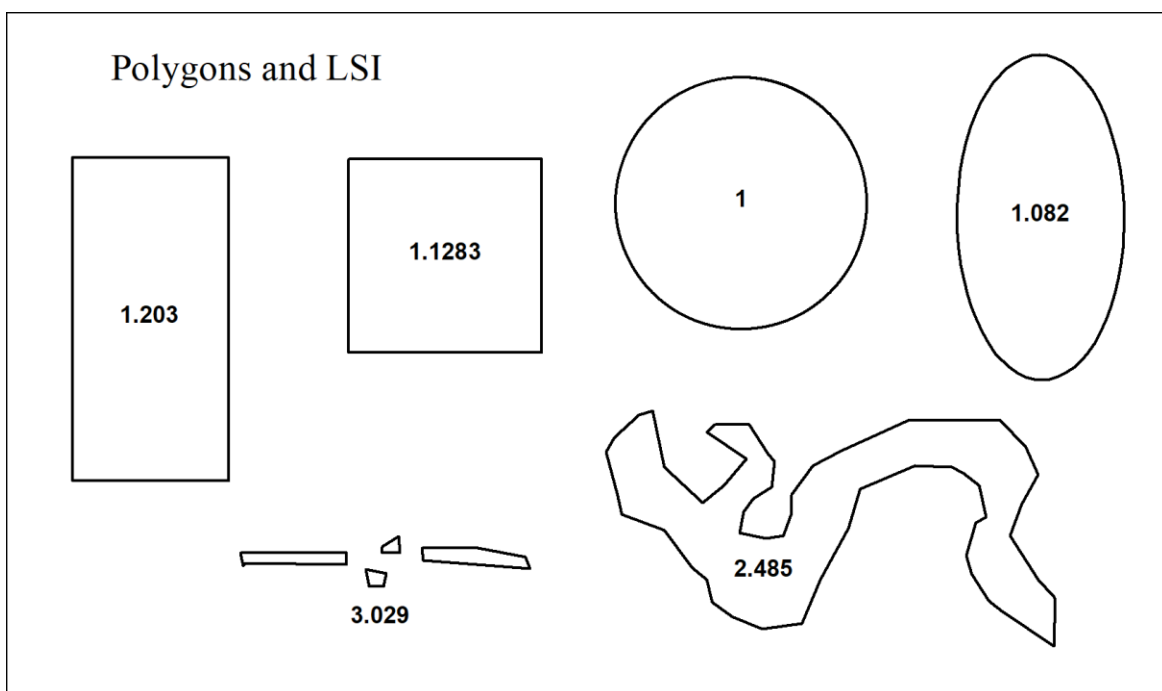
Once PCI was calculated for each parcel, the coverage of various land cover types was calculated and summed for each. This was performed by splitting the 2016 Land Cover dataset by cover type, intersecting each with the parcel layer, calculating the sum in acres, and joining the value back to the parcel layer. These values were then used to generate the percentage of coverage of each land cover type within a parcel.

After performing exploratory regression models using different collections of land cover types contained in the 2016 Land Cover dataset as variables, it was found that there was significant multicollinearity among them. This was, at times, due to too large a number of land cover variables being included in the model. It was also due to many of the land cover types sharing similarities, or being expressions of a number of other variables.

Palustrine and estuarine wetland types had these challenges, which could not be resolved by combining them, or selectively excluding them. Seeing as these land cover types were also not

the most relevant to this study of urban greenspace, they were excluded from the model. Other

land cover types, such as cultivated land and pasture/hay occurred at such limited frequency that

they were also removed. In the end, the model was limited to the seven land cover types of:

deciduous forest, developed open space, evergreen forest, grassland, impervious, scrub/shrub,

and water.

Figure 2: Examples of LSI

The total acreage and landscape shape index value (LSI) was then calculated for each parcel. The

LSI formula used was:

$$LSI = \left(\frac{P_t}{2\sqrt{\pi \times A}} \right)$$

where P_t is the length of the perimeter of a parcel, and *A* is the area of a parcel. The resulting

outcome of this formula is that a perfect circle returns a value of 1.00, a square returns

approximately 1.13, with shapes of increasing complexity returning larger values. A long,

snakelike polygon, for instance, might return a value of 5.00. This formula helps to generalize

shape complexity into a useable value within a regression analysis, where higher values correlate to increased complexity. Within the context of parcels and urban forest, this variable helps to capture the influence of vegetation 'compactness' on cooling the surrounding area.

With these attributes calculated per parcel, they were then entered in as variables into a multivariate regression model. PCI was set as the dependent variable, with the seven land cover percentage values, parcel acreage, and LSI as independent variables. Twenty different regression models were analyzed to reflect the changing calculation of PCI according to the different parcel buffer distances, with the independent variables remaining constant across all model iterations.

Figure 3: Histogram of parcel acreage
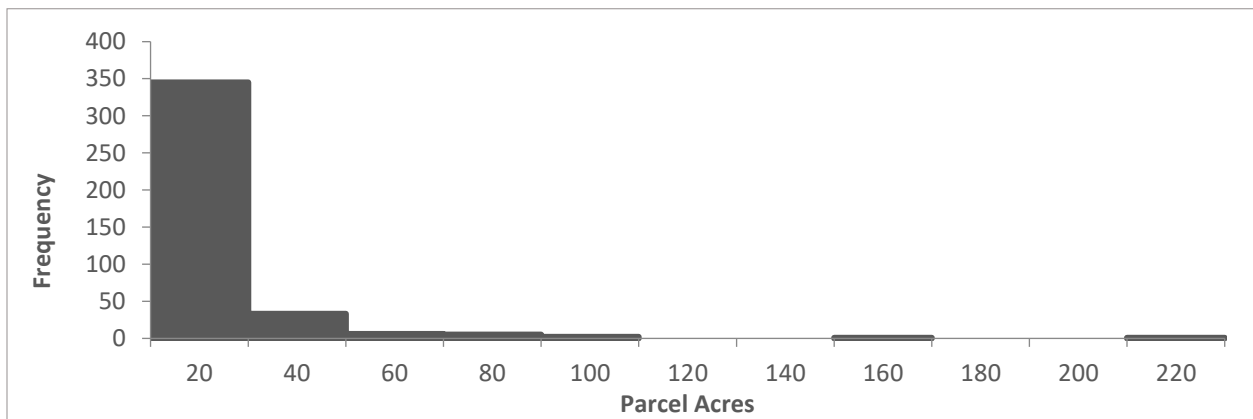
Figure 4: Histogram of parcel LSI values
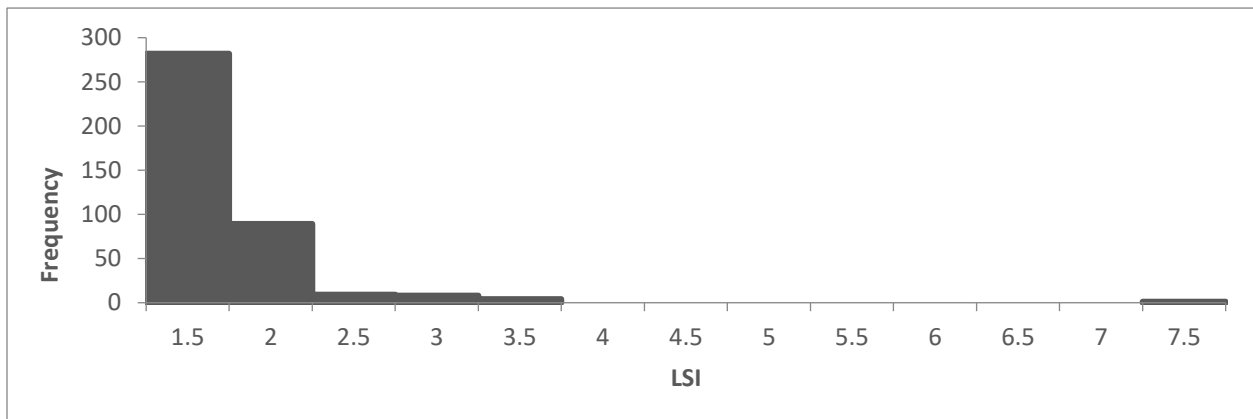
Results

After running regression models of the study parcels for all buffer distances, the relationships between land cover, LSI, park size, and park cool intensity can be observed, as well as the influence of the buffer distance parameter. First, the 500 meter buffer results will be reviewed as a default value in comparison to results from Cao's study (Cao et al. 2010):

500 Meter Buffer:

Table 2: Model Results

Independent Variables	Std. coefficient	Sig.
Deciduous Forest	2.593194	**0.000192***
Developed Open Space	-1.788128	**0.008900***
Evergreen Forest	-1.301609	0.525553
Grassland	-0.049073	0.978853
Impervious Surface	-7.211488	**0.000000***
Scrub/Shrub	0.526811	0.787307
Water	3.882023	0.050770
Parcel Acres	-0.010017	0.124640
LSI	-0.228931	0.285846
Model summary		
Multiple R^2 / Adjusted R^2	0. 5876 / 0.5780	
Std. error of estimation (C)	2.482684	

The regression results in Table 2 show that three of the nine independent variables were determined to be statistically significant in the model. Deciduous forest, developed open space, and impervious surface resulted in p-values smaller than 0.05, and among the strongest coefficient values of the model at 2.593, -1.788, and -7.211 respectively.

Those variables considered statistically insignificant included evergreen forest, grassland, scrub/shrub, water, parcel acres, and LSI. Of these independent variables, evergreen forest,

grassland, parcel acres, and LSI showed negative relationships with PCI (-1.301, -0.049, -0.010, -0.229) and scrub/shrub and water showed positive relationships (0.526, 3.882). The model as a whole returned an R^2 value of 0.587 with a standard error of estimation of 2.482.

This would seem to show that generally, natural land covers are associated with positive cooling effects, and impervious surface within a parcel is predictive of a negative cooling effect. Interestingly, evergreen forest results show a negative relationship with cooling as compared to deciduous forest. Likewise, developed open space shows a significant negative relationship to PCI despite being composed of vegetation. These aspects may be worth exploring to determine whether they have to do with different biological processes or correlate to different landscaping practices.

In this model, water is shown as having a cooling effect, which has not always been the case in previous studies. Parcel acreage doesn't produce a strong predictive relationship within this model, and could be due to not setting a minimum vegetation threshold for parcels entered into the model. The LSI variable does seem to show that increased shape complexity has some negative effect on cooling, though this relationship is generally weak. As shown in Figure 3 and Figure 4, the distribution of acres and LSI among parcels are both strongly right-skewed, and the range of values is narrow for these variables, so they are likely to have smaller coefficients.

Figure 5 shows the mapped residuals of each of the study parcels using a Jenk's (natural breaks) classification. Parcels symbolized in blue represent those that had a stronger cooling effect (higher PCI) than predicted by the model, and parcels in red representing those that had a weaker cooling effect (lower PCI) than predicted. A local Moran's I analysis was performed with the residuals as the input to test for spatial autocorrelation, which was found to be significant with a p-value of 0 and a z-score of 16.909. This indicates that results of the regression are clustered,
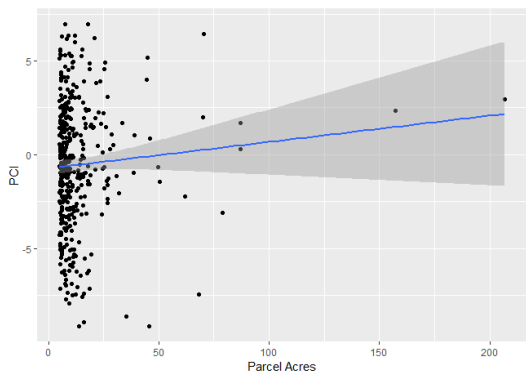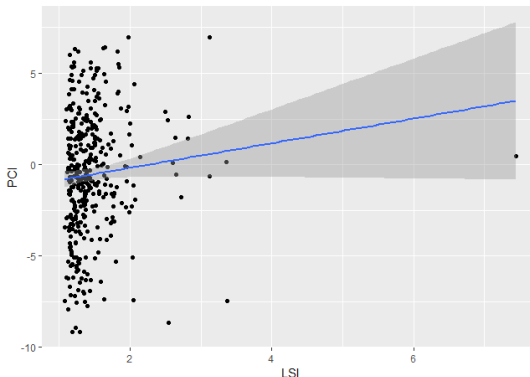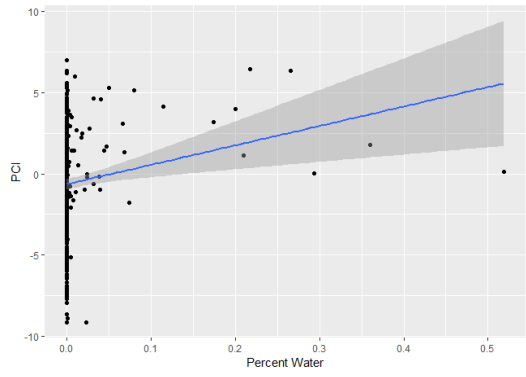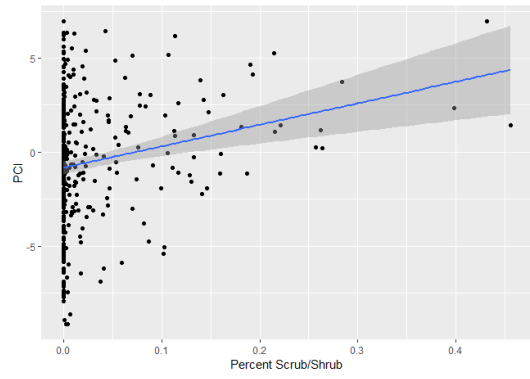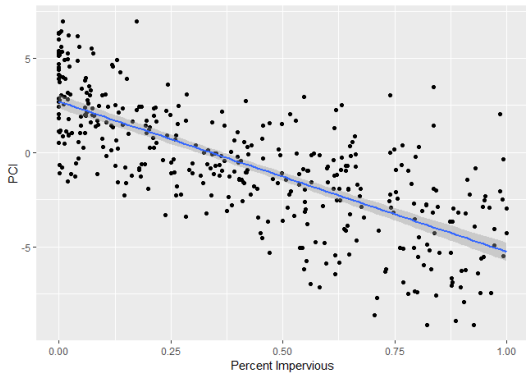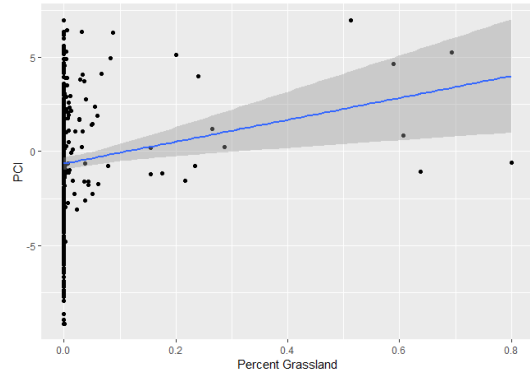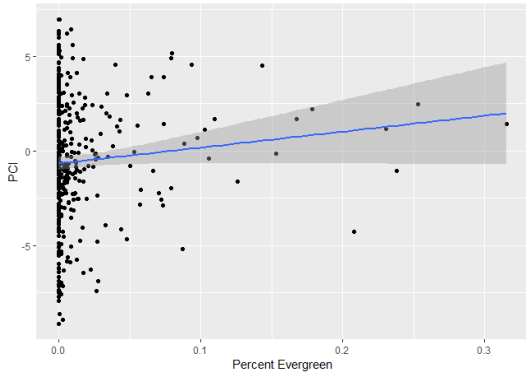
and that the chance that the data are randomly distributed is less than 1%. This may be due to an absent geographic variable.

Figure 5: Mapped residuals of regression model (500 meter buffer iteration)
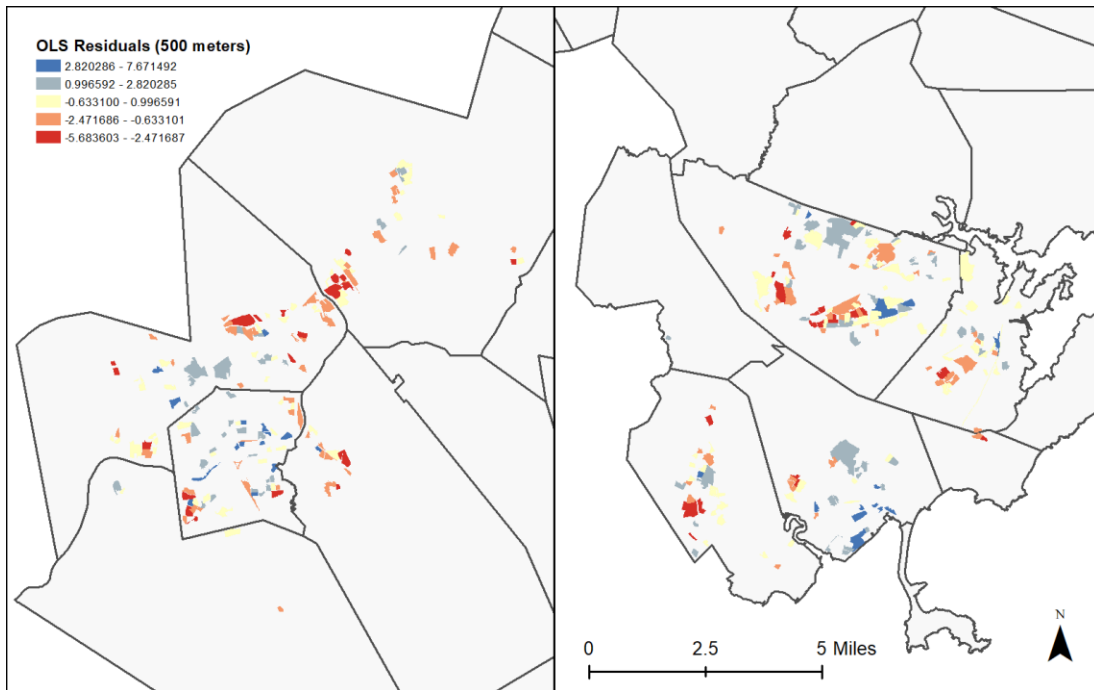
Figure 6: Plots of model independent variables
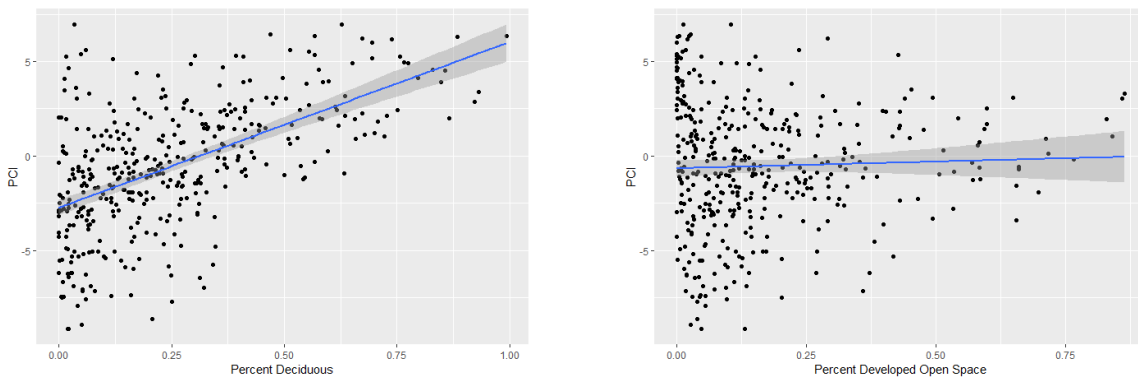
Comparing Parcel Buffer Values:

After performing regression analyses for each of the parcel buffer distances, and while leaving the independent variables unchanged, the results can be presented side-by-side in order to observe the sensitivity of the model to this parameter. To be specific, this parameter is the PCI calculation of each parcel at a given buffer distance. A total of twenty buffer distance values were assessed, from 50 meters to 1000 meters in increments of 50 meters.

Looking first at the model performance as viewed through its multiple R^2 value, it can be seen that model reaches its peak at 500 meters (0.5876), leading next to a slight descent towards the 1000 meter mark (seen in Figure 7). While model performance seems to greatly increase from the 50 meter iteration to the 400 meter model, it does not decrease so rapidly once moving past the 500 meter peak.

Figure 7: R^2 value across each model
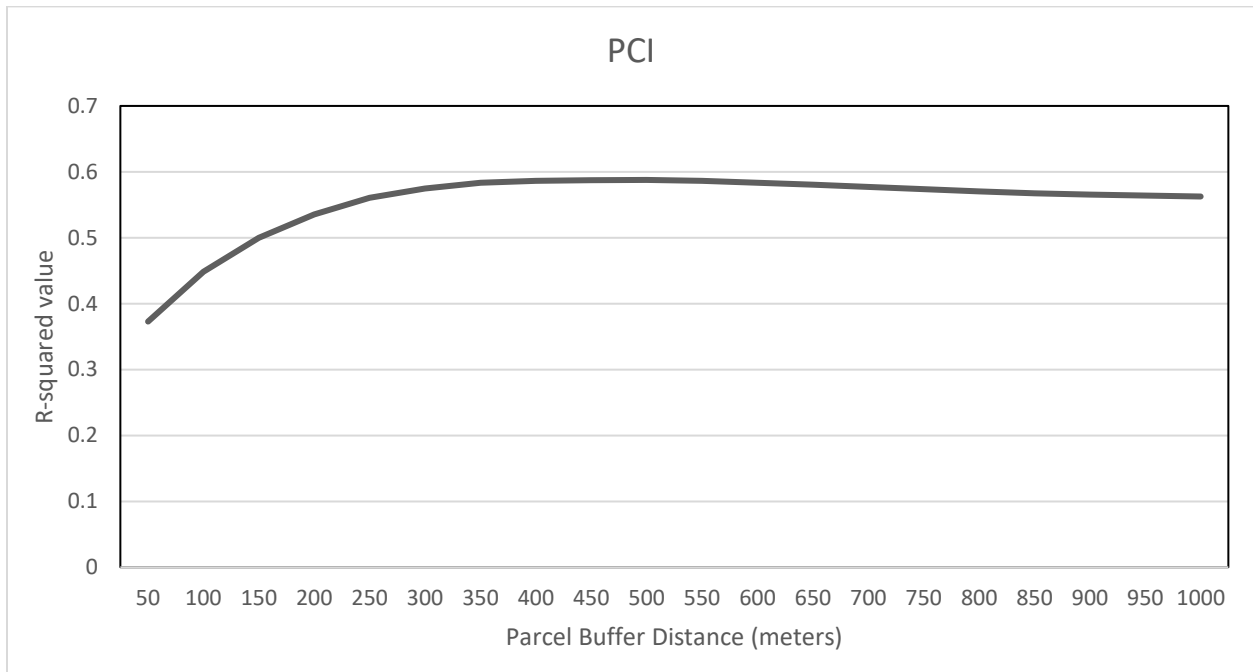
Figure 8: Variable coefficients across each model
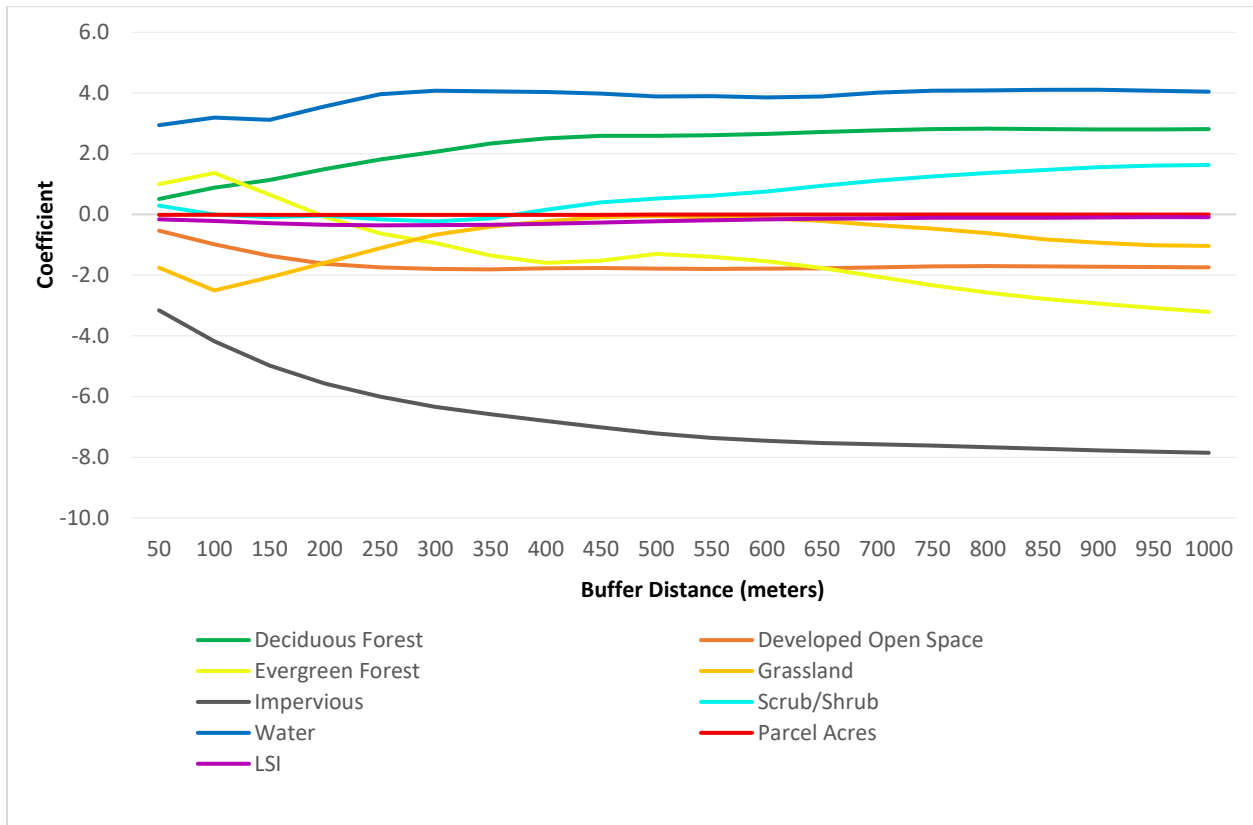
Throughout each iteration of the model most variables remained relatively stable, however a few stand out as being highly influenced by buffer distance. Impervious surface shows a moderately negative relationship with PCI at 50 meters (-3.160), but quickly decreases towards being the strongest negative variable at around 550 meters (-7.364). This may suggest that the influence of percent imperviousness in predicting PCI increases with buffer distance up to the 550 meter mark where it plateaus towards 1000 meters, and small buffer distances may underestimate the predictive influence of impervious surface within the PCI formula.

The grassland coefficient seems to experience the most fluctuation, beginning at -1.754 it then dips below -2, ascends towards a plateau near 0 at a buffer distance of 450 meters, and descends again to -1 around 1000 meters. Likewise, the evergreen coefficient begins near 1 at a buffer distance of 50 meters, and after a slight increase in value at 100 meters descends to -3.210 at a

buffer distance of 1000 meters. Similar to impervious surface, this would suggest that a model utilizing a small buffer distance may not be capturing accurate predictive relationships of some of these land cover types. Impervious surface was the only variable to return statistically significant p-values across the entire range of buffer distances. Developed open space and evergreen forest returned significant p-values from 150 meters to 1000 meters.

Discussion

This study aimed to apply methods previously used to analyze the cooling properties of greenspace in urban areas to parcels within the study area of Essex County. While the threat of urban heat islands is not currently as severe as in other geographies, it is a growing concern to planning organizations in the region. Implementation of existing scientific methods to study the urban heat island effect and the impacts of vegetative cooling could help to inform city planning, and equip conservation agents to take preventative actions towards the heat island effect. Alongside this implementation of methods, this study sought to test the robustness of Cao's model (Cao et al. 2010), and perform a sensitivity analysis on the buffer distance used in the calculation of PCI to see how it might respond differently in this geographic context.

This process involved the selection of parcels experiencing heat island threats, producing buffers around them from a range of 50 to 1000 meters (at increments of 50 meters), and calculating PCI values for them using land surface temperature data. High resolution land cover data were intersected with the parcels to produce percentage values of coverage for each land cover type, and these were entered as independent variables into a linear regression model. Additionally, the acreage and landscape shape index (LSI) of each parcel were included in the model to determine what impacts size and compactness may have on cooling. These independent variables remained constant across each of the twenty different models that were run. The dependent variable, PCI, changed according to the buffer distance used in its calculation.

It was found that Cao's use of a 500 meter buffer distance in the PCI formula resulted in the strongest performance across all iterations of the model with a multiple R^2 of 0.587 (Cao et al. 2010). Impervious surface, deciduous forest, and developed open space emerged as statistically

significant through most iterations, with deciduous forest having a positive relationship with PCI, and impervious surface and developed open space having a negative relationship with PCI.

These results indicate that PCI is able to provide some prediction of the cooling effect of parcels. They seem to be in line with previous studies contrasting impervious surface and forested land cover, but also highlight differences in vegetation as developed open space (includes lawns, soccer fields, and golf courses) was shown to be correlated with heating rather than cooling. Grass and open field land covers seem to return different cooling effects (both positive and negative) based on the environmental factors in previous studies, so it is interesting to observe these differences in this model. This could be due to the selection of study parcels included, or the timing of the thermal imagery capture in the morning rather than peak afternoon temperatures.

It was interesting to observe the change of model performance across buffer values, and that model performance peaked at the 500 meter value. From this sensitivity analysis it seems that using this distance to evaluate PCI helps to maximize model performance in terms of R^2 values, and validates Cao's use of this distance as a parameter (Cao et al. 2010). Both the R^2 values and the independent variable coefficients seem to reach greater stability around the 500 meter distance which would further support Cao's use of this distance.

This outcome may help to eliminate uncertainty for other researchers regarding this parameter who are utilizing the PCI formula. At the same time, this study's 500 meter model underperformed in comparison to Cao's. Cao's model R^2 values ranged from 0.79 to 0.88 depending on the season in which the LST data were derived from, whereas the model peaked at 0.587 in this study.

The difference in outcomes between this study and Cao's could be the result of variance in climate, input data quality, land cover differences, or existing density of impervious surface. While each of these vary across studies, it is seems that differences in impervious surface density could be the largest contributing factor. While the input parcels of this study were limited to those of greatest heat island threat, the overall density of impervious surface in Essex County is much lower than that in other study areas. This aspect makes it more of a challenge for the PCI variable to build consistent relationships with other variables.

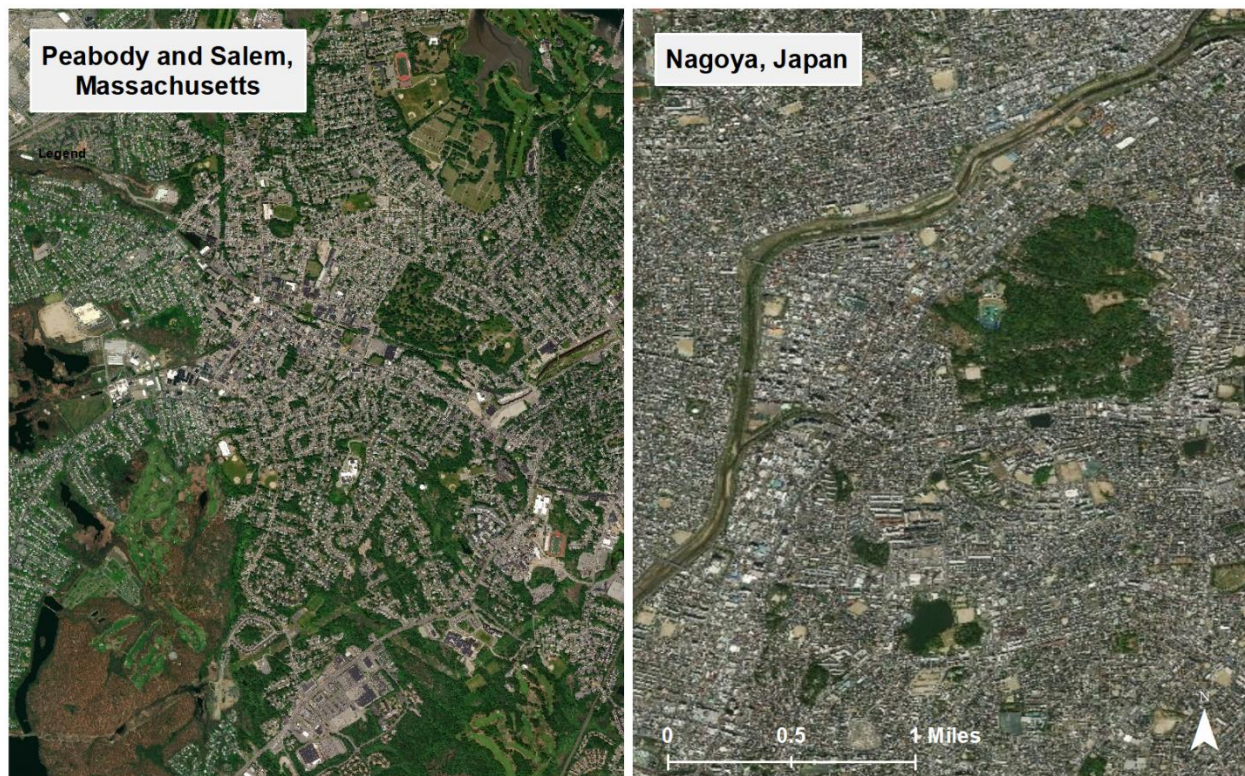**Figure 9: Essex County, MA and Nagoya, Japan at 1:40,000**

Figure 9 shows an orthographic image of Essex County next to Nagoya, Japan, which was studied by Cao (Cao et al. 2010). It is apparent that the green parks in Nagoya are distinct from their impervious surroundings, which means the mean LST values of their buffers are derived almost exclusively from impervious surface. Essex County cities have relatively more

vegetation, which is an asset from an urban cooling perspective, however it means that this land cover can contribute confusion or noise into a model that utilizes PCI as a dependent variable. This aspect could be worth further researching in order to determine if the PCI model is only suitable for dense cities, and to what extent the land cover of buffer areas must remain homogenously impervious surface. If use of PCI within a semi-urban setting is in fact a limitation, it may be worth investigation modifications to the PCI formula to be more applicable to developing regions, or including additional variables into the model in order to be more predictive. As was revealed in the Moran's I spatial autocorrelation test the results were significantly clustered, so exploring the use of additional geographic variables could be worthwhile to increase model performance in this region.

This study also highlighted the sensitivity of the land cover coefficients to buffer distance (Figure 8). While many remained fairly constant throughout the range of distances or stabilized upon reaching a plateau, evergreen forest had a positive coefficient value at 50 meters but became negative as the buffer distance increased. The grassland coefficient remained negative in all iterations of the model, however it displayed a wave-like fluctuation. This would indicate that some land cover coefficients experience a scale-sensitivity, and that their importance to the model changes according to the buffer distance used to calculate PCI. It's possible that this could be due to characteristics of this study area, or it could be that the importance of the variables change relative to each other depending on this distance. If these patterns are present in the results of other PCI analyses but are unobserved, it could be possible to make overly-simplistic conclusions about the land cover relationships. Testing the PCI model within other geographic contexts to observe changes in land cover coefficients across buffer distances could help to understand if the scale-sensitivity is unique to Essex County or is inherent to PCI.

Further steps could be taken to more fully explore the use of the PCI model within Essex County. This study only used one thermal infrared image as a source for LST. The land cover relationships revealed in the regression models then may be only relevant to the time of day and season of that image capture. While the urban heat island issue is only a threat to human health within the summer months in Essex County, analyzing cooling properties at different times of the day might produce different outcomes. Fully understanding these land cover relationships are especially necessary to effective urban planning surrounding the heat island issue.

While water land cover did not result as statistically significant in the model, it was very close to passing that threshold in terms of its p-value (0.05077). It also had a strong, positive coefficient throughout each of the model iterations. This would seem to indicate that it does play a role in cooling nearby landscapes in the study region, which has not always been the case in previous studies. This could be an area of further research to see whether this has to do with the type of water bodies present in Essex County, or perhaps the time of day or year that the thermal information was captured.

Beyond the PCI model, it could also be worth analyzing the compactness of greenspace as demonstrated by Zhang (Zhang 2017). Such an approach would not be as effective at highlighting individual parcels, but may provide a different look at the relationship between impervious surface and greenspace. Zhang's analysis may provide insight into the geographic distribution of these land covers, which would help to understand the impacts of development practices in regards to heat islands at a town scale, and avoid some of the challenges inherent to PCI. This approach could be of greater benefit to municipal planners who are interested in developing cities in a way to prevent the formation of urban heat islands, or are interested in cooling a city in a general sense with less concern about site-specific cooling benefits.

For use in Essex County, and similar communities looking to evaluate heat island threats and mitigation parcels, the PCI formula does seem to be an effective means of identifying parcels of high cooling intensity. It provides a straight-forward methodology to compare parcels within a study region without the need for first identifying heat island areas. The temperature difference between a parcel and its buffer area is an easily communicable value, and demonstrates the actual cooling intensity of a parcel. One must be aware of the spatial limitations of the land surface temperature data being used to calculate PCI, as small urban parks may not be captured in enough detail in images from Landsat or ASTER sensors. That being said, when used at the proper scale, PCI is an accurate tool for representing the park cool island value of a parcel.

The use of a regression model to predict PCI using land cover variables in Essex County is likely less effective than other approaches. As stated previously, the variance in the levels of impervious surface in a parcel buffer makes it a challenge for a model to establish relationships between PCI and land cover, ultimately decreasing model performance. While this variance in PCI score is a reflection of something true (a parcel must be situated near urban heat island areas in order to provide 'urban cooling'), it doesn't serve as a viable dependent variable in this region. Understanding the cooling benefits of different land cover types is still of importance for urban planning, and would be worthwhile to study using different methods. Such an example may be a grid-based analysis of land cover and LST, and observing these relationships in correlation matrices or a linear regression model.

Regardless of the next steps taken, these results demonstrate the impact that impervious surface has on air temperature and the environment. These observations can help to spur discussion at the local level in regards to development practices as well as conservation planning. Additionally, it points to deciduous forest as the most significant land cover type to preserve for

cooling benefits in this region, which is a natural resource that can be a challenge to restore (Drummond and Loveland 2010). It is also important to note that while dense urban development poses urban heat island threats, sprawling suburban development can bring other environmental and human hazards (Radeloff, Hammer, and Stewart 2005). Land cover studies, such as this one, can function as a tool for planners to make informed decisions surrounding the urban heat island issue at both regional and site-specific scales.

References

Ahmed, S. 2018. Assessment of urban heat islands and impact of climate change on
    socioeconomic over Suez Governorate using remote sensing and GIS techniques. The
    Egyptian Journal of Remote Sensing and Space Science. Volume 21, Issue 1. 15-25.

Akbari, H. 2005. Potentials of urban heat island mitigation. International Conference "Passive
    and Low Energy Cooling for the Built Environment", May 2005, Santorini, Greece.

Akbari, H., Davis, S., Dorsano, S., Huang, J., Winnett, S., 1992. Cooling Our Communities: A
    Guidebook on Tree Planting and Light-Colored Surfacing. Us Environmental Protection
    Agency. Office of Policy Analysis, Climate Change Division, 217pp.

Akbari, H., and Kolokotsa, D. 2016. Three decades of urban heat islands and mitigation
    technologies research. Energy and Buildings. Volume 133, Issue 1. 834-842.

Arnfield, A.J. 2003. Two decades of urban climate research: a review of turbulence, exchanges
    of energy and water, and the urban heat island. International Journal of Climatology.
    Volume 23, Issue 3.

Bowler, D. E., Buyung-Ali, L., Knight, T. M., and Pullin, A. S. 2010. Urban greening to cool
    towns and cities: A systematic review of the empirical evidence. Landscape and Urban
    Planning, 97(3), 147–155.

Cao, X., Onishi, A., Chen, J., Imura, H, 2010. Quantifying the cool island intensity of urban
    parks using ASTER and IKONOS data. Landscape and Urban Planning. 96, 224-231.

Čeplová, N., Kalusová, V., Lososová, Z. 2017. Effects of settlement size, urban heat island and
    habitat type on urban plant biodiversity. Landscape and Urban Planning, Amsterdam,
    Netherlands: Elsevier Science BV. Volume 159. 15-22

Chakraborty, J. 2017. Focus on environmental justice: new directions in international research.
    Environmental Research Letters. Volume 12. Number 3.

Chander, G., Markham, B., and Barsi, J. 2007. IEEE Geoscience and Remote Sensing Letters.
    Vol 4. No 3. 490-494.

Doick, K.J., A. Peace, and TR Hutchings. 2014. The role of one large greenspace in mitigating
    London's nocturnal urban heat island. Science of The Total Environment 493: 662-671.

Drummond, M., Loveland, T. 2010. Land-use pressure and a transition to forest-cover loss in the
    eastern United States. BioScience. Volume 60, Issue 4. Pages 286-298.

Effat, H., Taha, L. and Mansour, K. 2014. Change Detection of Land cover and Urban Heat
    Islands using Multi-Temporal Landsat Images, application in Tanta City, Egypt. Open
    Journal of Remote Sensing and Positioning. 1. 1-15. 10.15764/RSP.2014.02001.

Essex County Greenbelt Association. 2017. Essex County Greenbelt Strategic Plan June 2017.
    June. Accessed February 19, 2019.

Gosling, S.N., McGregor, G.R., Páldy, A. 2007. Climate change and heat-related mortality in six cities part 1: Model construction and validation. Int. J. Biometeorol. 51, 525–540.

Gronluand, C. Berrocal, V, White-Newsome, J., Conlon, K., O'Neil, M. 2015. Vulnerability to extreme heat by socio-demographic characteristics and area green space among the elderly in Michigan, 1990-2007. Environ Res.136: 449–461

Heaviside, C., Macintyre, H., and Vardoulakis, S. 2017. The Urban Heat Island: Implications for Health in a Changing Environment. Built Environment and Health. Curr Envir Health Rpt 4. 296-305.

Hien, W.N., Yu, C. 2004. The thermal effects of city greens on surroundings under the tropical climate. The 21st Conference on Passive and Low Energy Architecture. Eindhoven, The Netherlands. September 2004.

Jauregui, E, 1990. Influence of a large urban park on temperature and convective precipitation in a tropical city. Energy Build. 15-16, 457-463.

Leopold, L. B. 1968. Hydrology for urban land planning - A guidebook on the hydrologic effects of urban land use. U.S. Geological Survey. 554. 26.

Li, H., Zhou, Y., Li, X., Meng, L., Wang X., Sha, W., Soudi, S. 2017. A new method to quantify surface urban heat island intensity. Science of the Total Environment. 624, 262-272.

Lin, W., Yu, T., Chang, X., Wu, W., Zhang, Y. 2014. Calculating cooling extents of green parks using remote sensing: Method and test. Landscape and Urban Planning. 134, 66-75.

Lipping, Di. 2003. The development of remote-sensing related standards at FGDC, OGC, and ISO TC 211. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings. Vol. 1. 643-647.

Mass Audubon. 2020. Losing Ground: Nature's Value In A Changing Climate. https://www.massaudubon.org/our-conservation-work/advocacy/shaping-the-future-of-your-community/publications-community-resources/losing-ground/statistics/county/essex

MassGIS (Bureau of Geographic Information). 2020. Executive Office of Energy and Environmental Affairs (EOEEA). 2016 Land Cover/Land Use.

McGeehin M., Mirabelli, M. 2001. The potential impacts of climate variability and change on temperature-related morbidity and mortality in the United States. Environmental Health Perspectives. 109(Suppl 2): 185–189.

McGlynn, T., Meineke, E., Bahlai, C., Emily, E., Hartop, E., Adams, B., and Brown, B. 2019. Temperature accounts for the biodiversity of a hyperdiverse group of insects in urban Los Angeles. Proceedings of the Royal Society B: Biological Sciences. 286. 1912.

Medina-Ramón, M.. Zanobetti, A., Cavanagh, D. 2006. Extreme Temperatures and Mortality: Assessing Effect Modification by Personal Characteristics and Specific Cause of Death in a Multi-City Case-Only Analysis. Environmental Health Perspectives. 114(9): 1331–1336.

Mihalakakou, G., Santamouris, M., Papanikolaou, N., Cartalis, C., Tsangrassoulis, A., 2004. Simulation of the urban heat island phenomenon in Mediterranean climates. Pure and Applied Geophysics. 161, 529-451.

Morabito, M., Crisci, A. and Georgiadis, T. 2017. Urban Imperviousness Effects on Summer Surface Temperatures Nearby Residential Buildings in Different Urban Zones of Parma. Remote Sensing. 10. 26.

NASA. 2020. Landsat Science: Landsat 8. https://landsat.gsfc.nasa.gov/landsat-data-continuity-mission/

Nature Conservancy. 2016. Planting Healthy Air: A global analysis of the role of urban trees in addressing particulate matter pollution and extreme heat.

Northeast Regional Climate Center. 2020. Massachusetts Climate. nrcc.cornell.edu/regional/climatenorms/climatenorms.html

Office for Coastal Management. 2020. C-CAP Land Cover, Massachusetts, 2016, https://inport.nmfs.noaa.gov/inport/item/54917.

Peery, Lexi. 2019. Report: By 2100, Mass. Could Have 26 Days A Year That Feel Hotter Than 100 Degrees. https://www.wbur.org/earthwhile/2019/07/16/boston-heat-wave-climate-change.

Pelta, R., Chudnovsky, A., and Schwarts, J. 2016. The spatial and temporal behavior of brightness temperature in Tel-Aviv and its application to air temperature monitoring. Environmental Pollution. 208. 153-160.

Petkova, E., Horton, R., Bader, D., Kinney, P. 2013. Projected Heat-Related Mortality in the U.S. Urban Northeast. Int. J. Environ. Res. Public Health. 10. 6734–6747.

Radeloff, V.C., Hammer, R.B. and Stewart, S.I. 2005. Rural and Suburban Sprawl in the U.S. Midwest from 1940 to 2000 and Its Relation to Forest Fragmentation. Conservation Biology, 19: 793-805.

Reid, C., O'Neil, M., Gronlund, C., Brines, S., Brown, D., Diez-Roux, A. 2009. Mapping community determinants of heat vulnerability. 117(11): 1730–1736.

Rigolon, A., Browning, M., and Jennings, V. 2018. Inequities in quality urban park systems: An environmental justice investigation of cities in the United States. Landscape and Urban Planning.

Roman, Lara, and Scatena, Frederick. 2011. Street tree survival rates: Meta-analysis of previous studies and application to a field survey in Philadelphia, PA, USA. Urban Forestry and Urban Greening. Volume 10, Issue 4, 2011, Pages 269-274.

Sampson, N., Gronlund, C., Buxton, M., Catalano, L., White-Newsom, J., Conlon, K., O'Neil, M., McCormick, S., Parker, E. 2013. Staying cool in a changing climate: Reaching vulnerable populations during heat events. April ; 23(2): 475–484.

Santamouris, M., Kolokotsa, D. 2015. On the impact of urban overheating and extreme climatic conditions on housing, energy, comfort and environmental quality of vulnerable population in Europe. Energy and Buildings. 98. 125-133.

Sarrat, C., Lemonsu, A., Masson, V., Guedalia, D. 2005. Impact of urban heat island on regional atmospheric pollution. Atmospheric Environment 40. 1743-1758.

Schinasi, L., Benmarhnia, T., De Roos, A. 2018. Modification of the association between high ambient temperature and health by urban microclimate indicators: A systematic review and meta-analysis. Environ Research. 161:168-180.

Semenza, J., Rubin, C., Falter, K. Selanikio, J., Flanders, W., Howe, H., Wilhelm, J. Heat-related Deaths during the July 1995 Heat Wave in Chicago. 1996. New England Journal of Medicine. 335. 84–90

Sharma A., Conry P., Fernando HJS, Hamlet Alaln, Hellmann JJ, Chen F. 2016. Green and cool roofs to mitigate urban heat island effects in the Chicago metropolitan area: evaluation with a regional climate model. Environmental Research Letters. Lett 11 064004.

Shishegar, Nastaran. 2014. The Impacts of Green Areas on Mitigating Urban Heat Island. The International Journal of Environmental Sustainability Vol. 9 (Issue 1): 119-130.

Smargiassi, A., Goldberg, M., Plante, C., Fournier, M., Baudouin, Y., and Kosatsky, T. 2009. Variation of daily warm season mortality as a function of micro-urban heat islands. Journal of Epidemiology and Community of Health. 63. 8. 659-664.

Song, J. F., and Li, X. Y. 2010. A study on the cooling effects of greenery on the surrounding areas by computer simulation for green built environment. In K. Li, L. Jia, X. Sun, M. Fei, and G. W. Irwin (Eds.), Life system modeling and intelligent computing (pp. 653–661).

Spronken-Smith, R.A., Oke, T.R. 1998. The thermal regime of urban parks in two cities with different summer climates. International Journal of Remote Sensing, 19:11. 2085-2104.

Taha, H. 1997. Urban climates and heat islands: albedo, evapotranspiration, and anthropogenic heat. Energy Build. 25, 99-103

U.S. Census Bureau (2019). Population estimates, July 1, 2019, (V2019). Retrieved from https://www.census.gov/quickfacts/fact/table/essexcountymassachusetts.

U.S. Environmental Protection Agency. 2019. Using Trees and Vegetation to Reduce Heat Islands. epa.gov. July. https://www.epa.gov/heat-islands/using-trees-and-vegetation-reduce-heat-islands

U.S. Environmental Protection Agency. 2020. Heat Island Impacts. epa.gov. July.
  https://www.epa.gov/heatislands/heat-island-impacts.

U.S. Geological Survey. 2014. Using the USGS Landsat 8 Product.
  https://landsat.usgs.gov/using-usgs-landsat-8-product.

Voelkel J., Hellman D., Sakuma R., and Shandas V. 2018. Assessing Vulnerability to Urban
  Heat: A Study of Disproportionate Heat Exposure and Access to Refuge by Socio-
  Demographic Status in Portland, Oregon. International Journal of Environmental
  Research and Public Health. 15. 4. 640.

White-Newsome J., Brines S., Brown, D., Dvonch J., Gronlund C., Zhang K., Oswald E., O'Neill
  M. 2013. Validating Satellite-Derived Land Surface Temperature with in Situ
  Measurements: A Public Health Perspective. Environmental health perspectives. 121.
  10.1289/ehp.1206176.

Y. Ma, Y. Kuang, and N. Huang, Coupling urbanization analyses for studying urban thermal
  environment and its interplay with biophysical parameters based on TM/ETM+ imagery.
  International Journal of Applied Earth Observation and Geoinformation, vol. 12, no. 2,
  pp. 110–118, 2010.

Zanobetti, A., O'Neil, M., Gronlund, C. 2013. Susceptibility to Mortality in Weather Extremes:
  Effect Modification by Personal and Small Area Characteristics In a Multi-City Case-
  Only Analysis. Epidemiology. 24(6): 809–819.

Zhang, Y., Zhan, Y., Yu, T., Ren, X. 2017. Urban green effects on land surface temperature
  caused by surface characteristics: A case study of summer Beijing metropolitan region.
  Infrared Physics and Technology. 86, 35-43.

Zhou, D., Xiao, J., Bonafoni, S., Berger, C., Deilami, K., Zhou, Y., Frolking, S., Yao, R., Qiao,
  Z., Sobrino, J. 2019. Remote Sensing. 11. 1. 48.

Zukin, S., Trujillo, V., Frase, P., Jackson, D., Recuber, T., and Walker, A. 2009. New Retail
  Capital and Neighborhood Change: Boutiques and Gentrification in New York City. City
  and Community. 8. 1.

Appendix

Python Scripts:

The below scrips were written in Python 2.7 utilizing Esri's Arcpy package. There are two scrips used in this analysis. The first, 'LCLU_Process_Parcels.py' calculates the coverage of different land covers within a parcel dataset, and this function is called within 'RunHeatIslandAnalysis.py' which evaluates land surface temperature, LSI, PCI, and other variables according to user-submitted buffer distances, then runs linear regression models for each definition of PCI.

LCLU_Process_Parcels.py

```python
def LCLU_Process_Parcels(parcels, LCLU, homeEnv):
    # setup
    import arcpy, os
    from arcpy import env
    home_gdb = os.path.dirname(homeEnv)
    os.chdir(home_gdb)
    arcpy.env.workspace = home_gdb
    env.overwriteOutput = True

    # Create Geodatabase
    geodatabase = "LandCoverProcessingParcels.gdb"
    try:
        arcpy.CreateFileGDB_management(home_gdb, geodatabase)
        arcpy.AddMessage("Geodatabase Created")
```

```python
    except Exception:
        pass


    # set directory to newly create geodatabase
    home_gdb2 = home_gdb
    home_gdb = home_gdb + "\%s" % (geodatabase)
    os.chdir(home_gdb)
    arcpy.env.workspace = home_gdb


    tempparcels = "tempparcels"
    LCLU_fl = "LCLU_fl"


    arcpy.CopyFeatures_management(parcels, "LCLU_Parcels")
    arcpy.MakeFeatureLayer_management("LCLU_Parcels", tempparcel
s)
    arcpy.MakeFeatureLayer_management(LCLU, LCLU_fl)
    print "Feature Layers Created"



    # full list of land cover types, if needed:   landcover_list
 = ['Bare Land', 'Cultivated', 'Deciduous Forest', 'Developed Op
en Space', 'Estuarine Emergent Wetland', 'Estuarine Forested Wet
land', 'Estuarine Scrub/Shrub Wetland', 'Evergreen Forest', 'Gra
ssland', 'Impervious', 'Palustrine Aquatic Bed', 'Palustrine Eme
rgent Wetland', 'Palustrine Forested Wetland', 'Palustrine Scrub
/Shrub Wetland', 'Pasture/Hay', 'Scrub/Shrub', 'Unconsolidated S
hore', 'Water']
    landcover_list = ['Deciduous Forest', 'Developed Open Space'
, 'Evergreen Forest', 'Grassland', 'Impervious', 'Scrub/Shrub',
'Water']


    arcpy.AddField_management(tempparcels, "Parcel_Acres", "DOUB
LE")
    calcacres = "!shape.area@acres!"
```

```python
    arcpy.CalculateField_management(tempparcels, "Parcel_Acres",
 calcacres, "PYTHON_9.3")


    fieldlist = []

    idField = "LOC_ID"

    for x in landcover_list:

        query1 = "\"COVERNAME\" IN ('%s')" % (x)

        print query1

        arcpy.SelectLayerByAttribute_management(LCLU_fl, "NEW_SE
LECTION", query1)

        templayer1 = "templayer1"

        templayer2 = "templayer2"

        templayer3 = "templayer3"

        print x + " Selected"

        arcpy.MakeFeatureLayer_management(LCLU_fl, templayer1)

        if x == 'Water':

            arcpy.CopyFeatures_management(templayer1, "Water")

        arcpy.Intersect_analysis([templayer1,tempparcels], templ
ayer2)

        print x + " Intersected"

        arcpy.Dissolve_management(templayer2, templayer3, [idFie
ld])

        print x + " Dissolved"

        fieldname1 = str(x).replace(" ", "_").replace("/", "_")
+ "_ac"

        arcpy.AddField_management(templayer3, fieldname1, "DOUBL
E")

        calcacres = "!shape.area@acres!"

        arcpy.CalculateField_management(templayer3, fieldname1,
calcacres, "PYTHON_9.3")

        print x + " Acres Calculated"

        arcpy.JoinField_management(tempparcels, idField, templay
er3, idField, [fieldname1])
```

```
        print x + " Acres Field Joined"

        null_formula = "def updateValue(value):\\n  if value ==
None:\\n    return '0'\\n  else: return value"

        updatenull = "updateValue( !%s! )" % (fieldname1)

        arcpy.CalculateField_management(tempparcels, fieldname1,
 updatenull, "PYTHON_9.3", null_formula)

        fieldname2 = str(x).replace(" ", "_").replace("/", "_")
+ "_pc"

        exp1 = "(!%s! / !Parcel_Acres!)" % (fieldname1)

        arcpy.AddField_management(tempparcels, fieldname2, "DOUB
LE")

        arcpy.CalculateField_management(tempparcels, fieldname2,
 exp1, "PYTHON_9.3")

        fieldlist.append(fieldname2)

        print x + " Percentage Calculated"

    print fieldlist

    arcpy.JoinField_management(parcels, idField, tempparcels, id
Field, fieldlist)

    os.chdir(homeEnv)

    arcpy.env.workspace = homeEnv

    return home_gdb
```

RunHeatIslandAnalysis.py

```python
# setup

import arcpy, os

from arcpy import env

from LCLU_Process_Parcels import LCLU_Process_Parcels

home_gdb = r'C:\Users\David\Documents\SSU\MastersThesis\2_Thesis
Research\Data\Analysis\1.FinalAnalysis'

os.chdir(home_gdb)

arcpy.env.workspace = home_gdb

env.overwriteOutput = True



# INPUT DATA

InputParcels = r'C:\Users\David\Documents\SSU\MastersThesis\2_Th
esisResearch\Data\Parcels\AnalysisParcels.gdb\Parcels_Group10'
    # parcels

InputLSTRaster = r'C:\Users\David\Documents\SSU\MastersThesis\2_
ThesisResearch\Data\Landsat_Thermal\2016_7_13\LST_20160713.tif'
  # LST data

InputLCLU = r'C:\Users\David\Documents\SSU\MastersThesis\2_Thesi
sResearch\Data\lclu_gdb\MA_LCLU2016.gdb\LANDCOVER_LANDUSE_POLY'
  # LCLU data

study_area = r'C:\Users\David\Documents\SSU\MastersThesis\2_Thes
isResearch\Data\townssurvey_gdb\townssurvey.gdb\EssexCo'
  # study area boundary (Essex County)

buffer_value = [50, 100, 150, 200, 250, 300, 350, 400, 450, 500,
 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000]
  # buffer values for evaluating PCI
                    # buffer value

new_geodatabase = "UrbanCoolingAnalysis_v6_" + str(buffer_value[
-
1]).replace(" ", "") + "m"
 # new geodatabase name



print new_geodatabase
```

```python
# Create Geodatabase
geodatabase = new_geodatabase +".gdb"
try:
    arcpy.CreateFileGDB_management(home_gdb, geodatabase)
except Exception:
    pass
print "Geodatabase Created at: " + home_gdb + "\%s" % (geodataba
se)


# set directory to newly create geodatabase
home_gdb = home_gdb + "\%s" % (geodatabase)
os.chdir(home_gdb)
arcpy.env.workspace = home_gdb


# make copy of parcels
parcels_copy = "Analysis_Parcels"
arcpy.CopyFeatures_management(InputParcels, parcels_copy)
print "Parcels copied to Geodatabase"


# make feature layer of parcels
Parcels = "Parcels"
arcpy.MakeFeatureLayer_management(parcels_copy, Parcels)
print "Feature layer made of Parcels"


# pull largest buffer value and create string
buffer_value = sorted(buffer_value)
inputbuffervalue = str(buffer_value[-1]) + " meters"


# Buffer study_area by buffer_value
```

```python
study_area_buffered = "Study_Area_Buffered" + inputbuffervalue.r
eplace(" ", "")

arcpy.Buffer_analysis(study_area, study_area_buffered, inputbuff
ervalue)


# get bounding extent of study_area_buffered

def getFCextent(fc):

    desc = arcpy.Describe(fc)

    xmin = desc.extent.XMin

    xmax = desc.extent.XMax

    ymin = desc.extent.YMin

    ymax = desc.extent.YMax

    return "%s %s %s %s" % (xmin, ymin, xmax, ymax)

print "Calculated bounding area of study area buffered " + str(i
nputbuffervalue)


# clip LST raster to study_area_buffered

InputLSTRaster_clipped = "LST_Raster_Clipped"

arcpy.Clip_management(InputLSTRaster, getFCextent(study_area_buf
fered), InputLSTRaster_clipped, study_area_buffered, "", "Clippi
ngGeometry")

print "LST raster clipped to study_area_buffered"


# Calculate Coverage of Land Covers inside each Parcel using imp
orted function

print "Process Land Covers"

LCLU_GDB = LCLU_Process_Parcels(Parcels, InputLCLU, home_gdb)

# reset directory to home Geodatabase

os.chdir(home_gdb)

arcpy.env.workspace = home_gdb

print home_gdb

# assign water feater layer from LCLU layer
```

```python
LCLU_Water = os.path.join(LCLU_GDB, "Water")

# calculate parcel acreage

arcpy.AddField_management(Parcels, "Parcel_Acres", "DOUBLE")

calcacres = "!shape.area@acres!"

arcpy.CalculateField_management(Parcels, "Parcel_Acres", calcacr
es, "PYTHON_9.3")


# convert LST raster to points

print "Converting LST Raster to Points..."

LST_Raster_Points = "LST_Raster_Points"

arcpy.RasterToPoint_conversion(InputLSTRaster_clipped, LST_Raste
r_Points)

print "Convert LST Raster to Points"


# make feature layer of LST points

LST_Raster_Points_FL = "LST_Raster_Points_FL"

arcpy.MakeFeatureLayer_management(LST_Raster_Points,LST_Raster_P
oints_FL)

print "Feature Layer made of LST Points"


# calculate mean LST within parcels (using spatial join of point
s) and join to parcel feature layer

    # create function to take care of field mapping:

def CalculateLST(targetFeatures, joinFeatures, outfc, name_value
):

    # Create a new fieldmappings and add the two input feature c
lasses.

    fieldmappings = arcpy.FieldMappings()

    fieldmappings.addTable(targetFeatures)

    fieldmappings.addTable(joinFeatures)

    # get index of "grid_code" field
```

```
    grid_codeFieldIndex = fieldmappings.findFieldMapIndex("grid_
code")

    fieldmap = fieldmappings.getFieldMap(grid_codeFieldIndex)

    # Get the output field's properties as a field object

    field = fieldmap.outputField

    # Rename the field and pass the updated field object back in
to the field map

    field.name = name_value

    field.aliasName = name_value

    fieldmap.outputField = field

    # Set the merge rule to mean and then replace the old fieldm
ap in the mappings object

    # with the updated one

    fieldmap.mergeRule = "mean"

    fieldmappings.replaceFieldMap(grid_codeFieldIndex, fieldmap)

    #Run the Spatial Join tool, using the defaults for the join
operation and join type

    arcpy.SpatialJoin_analysis(targetFeatures, joinFeatures, out
fc, "#", "#", fieldmappings)


Parcels_LST_SJ = "Parcels_LST_SJ"

CalculateLST("Analysis_Parcels", LST_Raster_Points, Parcels_LST_
SJ, "LST_Mean")

print "Calculated Mean LST within each Parcel using spatial join
"


# Join LST value back to Parcels based on LOC_ID

arcpy.JoinField_management(Parcels, "LOC_ID", Parcels_LST_SJ, "L
OC_ID", ["LST_Mean"])

print "LST value joined back to Parcels"


    # calculate mean LST within parcel buffers (using spatial jo
in of points) and join to parcel feature layer
```

```
        # Buffer LST will be calculated only in non-
water areas, and those not inside other Parcels (per Cao analysi
s)

        # iterate through each parcel for individualized analysi
s, using dictionary to store values and write with an updateCurs
or

    # create feature layer of buffers

for buffer in buffer_value:


    # take input parcels and buffer by largest buffer_value (out
side of parcel only)

    inputbuffervalue = str(buffer) + " meters"

    Parcels_Buffered = "Parcels_Buffered" + (str(inputbuffervalu
e).replace(" ", ""))

    arcpy.Buffer_analysis(Parcels, Parcels_Buffered, inputbuffer
value, "OUTSIDE_ONLY")

    print "Parcels buffered for value: " + str(inputbuffervalue)


    Parcels_Buffered_fl = "Parcels_Buffered_fl"

    arcpy.MakeFeatureLayer_management(Parcels_Buffered, Parcels_
Buffered_fl)


    arcpy.CopyFeatures_management(Parcels_Buffered_fl, "ClipLST_
Temp1")                          # make copy of parcel buffer

    arcpy.Clip_analysis(LST_Raster_Points_FL, "ClipLST_Temp1", "
ClipLST_Temp2")              # clip LST points to parcel buf
fers

    arcpy.Erase_analysis("ClipLST_Temp2", Parcels, "ClipLST_Temp
3")                              # within buffers, erase LST poi
nts that fall within other parcels

    arcpy.Erase_analysis("ClipLST_Temp3", LCLU_Water, "ClipLST_T
emp4" )                          # within buffers, erase LST poi
nts that fall within water

    LST_Mean_Buffer_FieldName = "LST_Mean_Buffer_" + str(inputbu
ffervalue).replace(" ", "")      # buffer field name
```

```python
    CalculateLST("ClipLST_Temp1", "ClipLST_Temp4", "ClipLST_Temp
5", LST_Mean_Buffer_FieldName)   # use previously created functi
on to calculate Mean LST using a spatial join

    print "Calculated Mean LST within each Buffer using spatial
join"


    # Join LST value back to Parcels based on LOC_ID

    arcpy.JoinField_management(Parcels, "LOC_ID", "ClipLST_Temp5
", "LOC_ID", [LST_Mean_Buffer_FieldName])

    print "LST Buffer value joined back to Parcels"


    # calculate PCI

    PCI_FieldName = "PCI_" + (str(inputbuffervalue).replace(" ",
 ""))

    arcpy.AddField_management(Parcels, PCI_FieldName, "DOUBLE")

    PCI_Formula =  "!%s! - !LST_Mean!" % (LST_Mean_Buffer_FieldN
ame)

    arcpy.CalculateField_management(Parcels, PCI_FieldName, PCI_
Formula, "PYTHON_9.3")


    # delete temp layers:

    arcpy.Delete_management("ClipLST_Temp1")

    arcpy.Delete_management("ClipLST_Temp2")

    arcpy.Delete_management("ClipLST_Temp3")

    arcpy.Delete_management("ClipLST_Temp4")

    arcpy.Delete_management("ClipLST_Temp5")


    print "Calclulated LST for buffers: " + str(LST_Mean_Buffer_
FieldName)


# loop ended


# calculate LSI
```

```python
arcpy.AddField_management(Parcels, "LSI", "DOUBLE")

LSI_Formula =  "!Shape_Length! / (2 * math.sqrt(3.14159 * !Shape
_Area!) )"

arcpy.CalculateField_management(Parcels, "LSI", LSI_Formula, "PY
THON_9.3")


# with all data in place, now run regression analysis

# create integer field

arcpy.AddField_management(Parcels, "Int_Field", "SHORT")

arcpy.CalculateField_management(Parcels, "Int_Field", "!OBJECTID
!", "PYTHON_9.3")


Parcels = os.path.join(home_gdb, "Analysis_Parcels")

print Parcels

PDF_home = os.path.join(os.path.dirname(home_gdb), ("PDF_Outputs
_" + str(buffer_value[-1]).replace(" ", "") + "m"  )  )

os.mkdir(PDF_home)

print PDF_home


for value in buffer_value:

    outputfc = os.path.join(home_gdb, ("OLS_" + str(value) + "m"
))

    print outputfc

    dep_var = "PCI_" + str(value) + "meters"

    pdf_var = os.path.join(PDF_home , "OLS_" + str(value) + "m.p
df")

    print pdf_var

    ind_var = "Deciduous_Forest_pc;Developed_Open_Space_pc;Everg
reen_Forest_pc;Grassland_pc;Impervious_pc;Scrub_Shrub_pc;Water_p
c;Parcel_Acres;LSI"

    arcpy.OrdinaryLeastSquares_stats(Parcels, "Int_Field", outpu
tfc, dep_var, ind_var, "", "", pdf_var)
```