5-2017

Remotely sensed heat : variation and change in surface urban heat islands in a temperate eco-region of the United States.

Jeremy Sandifer
University of Louisville

REMOTELY SENSED HEAT: VARIATION AND CHANGE IN SURFACE URBAN HEAT ISLANDS IN A TEMPERATE ECO-REGION OF THE UNITED STATES

By

Jeremy Sandifer
B.S., University of Louisville, 2013

A Master's Thesis
Submitted to the Faculty of the
College of Arts and Sciences of the University of Louisville
In Partial Fulfillment of the Requirements
for the Degree of

Master of Science in Applied Geography

Department of Geography and Geosciences
University of Louisville
Louisville, KY

May 2017

REMOTELY SENSED HEAT: VARIATION AND CHANGE IN SURFACE URBAN HEAT

ISLANDS IN A TEMPERATE ECO-REGION OF THE UNITED STATESBY

By
Jeremy Sandifer
B.S., University of Louisville, 2013

A Thesis Approved on

April 21, 2017

by the following Thesis Committee:

Thesis Director

Dr. Andrea Gaughan

Co-Director or Second Thesis Committee Member

Dr. Keith Mountain

Third Thesis Committee Member

Dr. Margaret Carreiro

ACKNOWLEDGEMENTS

I would like to acknowledge all of those that have contributed in a substantial way to my development as an applied geographer and aspiring spatial scientist. For introducing me to the primary tools and methods involved in geographic analysis, I would like to thank Bob Forbes, DJ Biddle, Demetrio Zourarakis, and Mike Callahan. For providing in-depth guidance in utilizing the various applied geographic analyses, I want to thank Dr. Charlie Zhang, Dr. Wei Song, Dr. David Howarth, Dr. Andrew Day, Dr. Carol Hanchette, Dr. Margath Walker, and Dr. Jafar Hadizadeh. . Big thanks for the students and Faculty involved with the Friday brownbag meetings where great ideas were born and doused on a regular basis. Additional thanks for Ms. Sharon O'Bryan for many years of great customer service, the Department of Geography and Geosciences for the research assistantship opportunities, and the UofL Graduate Student Council for conference and travel expenses

Special thanks to my thesis committee including Dr. Keith Mountain, Dr. Margaret Carreiro, and Dr. Andrea Gaughan, in particular. Many thanks to Dr. Forrest Stevens for all of the technical assistance throughout the entire thesis project.

Finally, thanks to my family for all of the encouragement and support throughout, especially my wife Amy.

ABSTRACT

REMOTELY SENSED HEAT: VARIATION AND CHANGE IN SURFACE URBAN HEAT ISLANDS IN A TEMPERATE ECO-REGION OF THE UNITED STATES

Jeremy Sandifer

April 21st, 2017

Urban heat island (UHI) is a term used to describe increased surface and atmospheric temperatures in an urban core relative to surrounding non-urbanized areas. To examine the variability introduced into derived estimates of the surface UHI, this study constructs and compares multiple remotely sensed indicators of the surface UHI for major metropolitan cities of a temperate eco-region of the United States. The Moderate Resolution Imaging Spectroradiometer (MODIS) 8-day, 500-meter product (MOD11A2) is the source data used to calculate six different RS-derived UHI indicators for the year 2002 to 2012. The different SUHI indicators are evaluated using the Spearmans Rho rank-order correlation statistic to assess agreeability for 2012 and consistency over time 2002 to 2012. Inconsistencies exist in monthly rankings between indicators, and the degree to which the indicators detect change over time. Results suggest that land cover based indicators are highly correlated compared to urban heat island driven indicators in terms of magnitude and change over time.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. GENERAL INTRODUCTION

Introduction to the Urban Heat Island

Since the Industrial Revolution, humanity has acquired the technology and energy resources to sustain high levels of economic activity, resulting in large-scale changes to the surface of the Earth, with perhaps the most obvious being our local weather (Oke 1997). One of the most prevalent impacts from the changes to Earth's surface is the subsequent increase in impervious surfaces coincident with the increased number people living in urban areas. Less than 1 billion people lived in urban areas in 1950 and that increased to 3.9 billion by 2014, corresponding to 54% of the global population residing in all urban areas with some of the highest growth rates observed in developing regions of the African, Asian, and South American continents (U.N. 2015). While the meaning of the term *urban* will vary according to where in the world one is looking, in the U.S., systemically designated urbanized areas are defined by the United States Census Bureau as any spatial clustering of populations greater than 50,000 people and the adjacent supporting infrastructure and commercial developments to be included. These urban developments have generally replaced predominantly vegetation-covered environments with a range of impervious building materials, such as concrete and asphalt that are effective sinks (absorbers) of daytime solar energy and efficient radiators of that energy as well, generally resulting in an overall increase in temperatures. As a result, the term urban heat island (UHI) was coined to describe the diurnal-to-seasonally variable effect of air and surface temperature difference between the urban core and its surrounding areas (Oke 1982; Hung et al. 2006). It is generally agreed upon that the most important underlying driver of the UHI effect is the amount of impervious surface relative to the surrounding areas (Yuan and Bauer 2007; Stone, Hess, and Frumkin 2010), and that increasing vegetation densities effectively dampen the magnitude of the effect (Shi, Tao, and Liu 2014; Zhang, Wu, and Chen 2010).

Voogt and Oke (2003) distinguish two classes of UHI phenomena, atmospheric UHI- a warming of the air of both the urban canopy layer and boundary layer circulation, and surface UHI- representing only the emitted thermal energy or 'skin' temperature of the urban surface. Active investigations of the urban heat island effect begin to appear starting in the 19th century (e.g. Howard 1833; Gordon 1921) and continue consistently on a global basis and using a variety of spatial scales and approaches. Local ground based observations (measurements taken from less than 2 meters from the surface) are generally used to observe local atmospheric temperatures and regional stations are commonly combined to generate more widespread empirical studies using this kind of field data (Kopec 1970; Westendorf, Leuhart, and Howarth 1989). Recently, the magnitude of UHI effects in major U.S. metropolitan areas with urban populations greater than one (1) million were calculated using meteorological data drawn from the Global Historical Climatological Network (GHCN) and found to be increasing over time, particularly in the Southeast and Mid-west regions (Stone 2007). Louisville, Kentucky, in particular, has been ranked among the largest UHI signals in the U.S. (Kenward et al. 2014) as well as the highest decadal rate of increase in UHI magnitude from 1961 to 2010 (Stone 2012).

The UHI effect for larger metropolitan areas across the United States is increasingly viewed as a health and environmental problem with the highest UHI signals measured for cities such as Phoenix, AZ, Indianapolis, IN, and Atlanta, GA (Stone 2012; Kenward et al. 2014). Indeed, previous research has shown that the greatest potential temperature difference between urban areas and the surrounding landscape exist within the temperate forested eco-regions, precisely, where the majority of urbanization took place in the U.S. (Imhoff et al. 2010; Stone 2012). The growth in urbanization displaced some of the richest agriculturally productive land areas and had a surprisingly disproportionate impact on the environment (Shi, Tao, and Liu 2013). For perspective, consider that the conversion of only 3% of the U.S. natural land areas to urbanized areas essentially off set the coincident gains in the Net Primary Productivity (NPP) associated with the conversion of another 29% of the natural land areas to intensive crop agriculture during that same period (Imhoff et al. 2010). Seemingly, small changes to the environment, in this case a small change in proportional land cover, can potentially have outsize impacts, especially in the temperate mixed forest ecoregions. As cities continue to expand (spatially) in this region of the U.S., it is important to observe and understand how localized changes compare to other places and to understand how the changes are directly affecting the local environmental conditions. Negative impacts associated with the urban heat island effect are expected to increase along with the rise in temperatures due to climate change (Altman 2012) including the loss of native biodiversity (Alberti 2005; Ernstson et al. 2010), reduced air quality (Stone, 2008) and other impacts to human health (Tan et al. 2010). The specific focus here on the temperate eco-region ensures the differences detected are attributable to the unique physical characteristics of the individual study cities as opposed to other underlying, large-scale influences such as predominate weather patterns or extreme elevation, for example, that would otherwise complicate comparisons between locations.

Modeling the Urban Heat Island Effect

Conceptually, a model of the UHI as a physical phenomenon is an appropriate simplified representation of the real-world inter-relationships that exist between the constituent parts of the environment that collectively drive the distribution of surface temperatures in an urban area (Mirzaei 2015). There are various model types and each type necessarily designed to highlight some particular attribute of the UHI at the expense of others, all in an effort to minimize the number of complexities associated with representing the "real-world" ground conditions. The finest spatial scale or "micro-scale" includes the use of localized urban geometries to describe the urban heat dynamics that result from, among other things, restricted horizontal air advection and increased convective forcing from adjacent surfaces and how they impact levels of human comfort and building performance (Mirzaei 2015). The information derived from the use of these models is used to design and measure the efficiency of various strategies for mitigating the undesirable impacts of the UHI effect, green roofs or reflective concrete, for example (Stone 2012). While very high in spatial detail and information, the relationships quantified using micro-scale models are typically very limited in the spatial extent at which they are to remain accurate and useful. The computational effort can be quite large too and require specialized equipment.

Meso-scale models, on the other hand, generally apply to larger geographic extents and represent more broadly the generalized spatial relationships that exist between ambient environmental conditions (total portion of forestland cover or elevation, for instance) and the distribution of urban heat. The relationships described using meso-scale models broadly describe common behaviors of the UHI process across large areas and can be summarized using statistical functions and applied directly to estimate the distribution of urban heat where data may not be available, both spatially and temporally. Many contemporary examples employ a range of meso-scale models in conjunction with other spatial data to describe changes in surface temperatures based on changes in, for example, land-cover (Yuan and Bauer 2007; Zhang, Wu, and Chen, 2010), land-use (Middel et al. 2012), and levels of urbanization (Zhou, Huang, and Cadenasso 2011; Clinton and Gong 2013; Zheng, Myint, and Fan 2014). This utility comes at a price, however, in the form of a decreased ability to detect unique or localized influences specific to relatively few locations.

Most recently, efforts have involved the use of a combination of models that encompass co-variant information at many different spatial and temporal scales. The widespread adoption of geographic information systems (GIS) paralleled with improvements in computational efficiency have allowed for easy access to enormous amounts of data that can be accurately placed within the appropriate geographic context and overlaid with various other data types of different spatial extents and resolutions to be summarized using any of a number of statistical functions. In theory, generated functions can "couple" or fill the void between the different spatial scales of the layered data, and so offer potential for modeling information not readily apparent using one of the datasets alone (Mirzaei 2015). The models are important and especially useful because the parameters of the generated models can then be manipulated to project changes in urban heat distribution when given a quantity of change in the modeled co-variant data. Other contemporary meso-scale studies have projected changes in urban heat distribution by incorporating measures of spatial-temporal variation of socio-economic attributes of the population (e.g. Buyantuyev and Wu 2010) and effects of proposed land-use policy of human mortality (e.g. Stone et al. 2014). This diverse range of studies and models has provided general confirmation that the UHI is a consistent modern environmental process across a wide range of geographic settings.

Remote Sensing-Derived Estimates of SUHI Effect

While air temperature comparisons made using meteorological station data is the most direct measure of the UHI (Voogt and Oke 2003), it is often the case that permanent spatially dispersed meteorological stations lack sufficient spatial resolution for regional comparative studies. Studies utilizing station-based data typically use a single representative meteorological station within the city center and compare readings from that single location to one or more proximate rural stations to assess the observed differences in air temperature, the urban heat island. While valuable in terms of tracking long-term trends at regional and continental scales (i.e. Hanson et al. 2001), it lacks a spatially explicit component and can only be expected to represent the land areas immediately upwind of the station (Voogt and Oke 2003). Extrapolating point-based measurements to larger areas necessarily, but arbitrarily, simplifies heterogeneity in surface conditions and leaves in question the actual relative differences between and within cities in terms of the UHI effect, especially for geographically complex locations. Remote sensing of the UHI, in contrast, can potentially yield more spatially explicit detail regarding the distribution of signal intensity quantified using a variety of surface UHI indicators (Schwarz, Lautenbach, and Seppelt. 2011).

Remote sensing based studies of the surface urban heat island effect have provided a basis of evaluation in terms of understanding the dynamics of these complex environmental phenomena. For example, Jin, Dickinson, and Zhang (2005) utilized the global coverage of the Moderate Resolution Imaging Spectro-radiometer (MODIS) to characterize the SUHI for most continental land areas between approximately 30 and 60-degree N latitude at 5km spatial pixel resolution. Urban and rural areas delineated using the 250m MODIS land cover product confirm the constant presence of the SUHI and support generally the known effects of the urban heat island, that larger cities tend to have higher temperatures, as much as 1 degree Celsius per 100 km^2 of developed land area, for example. More significantly, that increased land area also has disproportionately higher near surface atmospheric temperatures during heat waves; values amplified by the concentration of developed land cover of as much as 56% (~ 0.5º C) or nearly double the additional contribution of 29% (~ 0.25º C) increase due to anthropogenic heat releases (Chen, Wang, and Zhu 2014). Furthermore, the casual relationships were significantly stronger in the temperate mixed forested ecological settings compared to coastal or grassland settings, for example. Interesting ecological anomalies include northern latitude urban areas that are cooler than the surrounding rural lands during the summer daytime (Ontario, Canada), and the little or reversed UHIs in arid environments where the relatively moist vegetated center cities surrounded by generally high albedo (Tempe, AZ), but shrubby and transpiration-limited desert surroundings (Imhoff et al. 2010).

While each of these SUHI modeling approaches have something to contribute to our overall understanding of the generation of the SUHI phenomena, they can lack consistency in their respective estimates when used for large geographic extents (Schwarz et al. 2012) and across different varying ecological contexts (Imhoff et. 2010). These inconsistencies are explained in large part by project oversights in explicitly defining how the representative areas (urban vs non-urban) are delineated, accounting for urban area spatial extents, and differences in the particular time of year analyzed (Stewart 2011). Among the most important (and commonly overlooked) considerations of the remote sensing approach for deriving estimates of SUHI intensity involves the delineation of representative urban and non-urban land areas from which surface temperatures are aggregated and summarized.

Schwarz, Lautenbach, and Seppelt (2011) distinguished two primary classes of surface UHI indicators, 1) UHI-driven as relative difference in mean LSTs of the 'urban core' versus the surrounding 'rural' areas and 2) land-cover driven measures that quantify differences in mean LST of representative land covers. These conceptual models differ primarily in how the non-urban or 'rural' is defined and the types of additional input data needed to calculate the UHI measure. The land-cover based model (i.e. Tomlinson et al. 2010) use *a priori* definitions of land cover to differentiate the urbanized land areas (built-up, impervious cover) from the non-urban areas (natural or agricultural land) where the difference in mean LST of each representative class is used to calculate the UHI magnitude. The UHI-driven model (i.e. Zhang and Wang 2008) involves creating a single representative measure of the 'urban core' LST and then subtracting the surrounding 'rural' or non-urban land area mean LST to calculate UHI magnitude. The

UHI-driven measurements most closely resemble those derived using data from permanent meteorological stations and help to explain the spatially-explicit nature of increased urban temperatures and account for the full spatial extent of the affected areas. Given the inconsistencies, noted above, between the various SUHI indicators implemented across studies, a major contribution of this current study is to systemically compare estimates of SUHI intensity obtained using a set of common SUHI indicators across multiple locations to quantify the degree of variability in results.

This study uses multiple approaches to constructing the surface UHI indicators (Table 1) for each of the cities while systematically controlling for urban spatial extent and the real fractional representation of land cover types. Controlling for these variables makes it easier to assess the variation that may be attributable to the selection of a specific SUHI indicator. I compare six (6) different SUHI indicators found in the literature (Table 1) that rely on the MODIS sensor for regional analysis. This particular group of indicators are previously utilized in similarly situated regional comparative studies investigating the variation introduced by use of different SUHI metrics for continental Europe (Schwarz, Lautenbach, and Seppelt 2011) and so gives good context for assessing the results for this current study

Indicator	Unit	Description	Approach	References
Difference urban-rural (DUR)	C	Difference in mean LST of urban boundary and buffered distance (35km)	UHI-driven	Imhoff et al. 2010; Zhange et al. 2010
Hot-island area (HIA)	%	Area within urban boundary with LST greater than the mean plus one standard deviation	UHI-driven	Zhang and Wang 2008
Magnitude (MAG)	C	Difference between max and mean LST within urban boundary	UHI-driven	Rajasekar and Weng 2009
Difference urban-agriculture (DUA)	C	Difference in mean LST between urban land and cropland	Landcover-driven	Jin et al. 2005
Difference urban-other (DUO)	C	Difference in mean LST between urban land and all others	Landcover-driven	Tomlinson et al. 2010; Zhou et al. 2010
Micro-island area (MIC)	%	Percentage of area with urban boundary with LST higher than the max LST of forested land cover	Landcover-driven	Aniello et al. 2005

Table 1. Surface Urban Heat Island Indicators with associated references listed.

The importance of the various characteristics of each study location varies according to the indicator selected and all indicators may not always be appropriate for all selected locations, even within the same ecological context. To avoid the kind of bias propagated by the use of a single indicator, combinations of multiple indicators better accommodate site-specific spatial heterogeneity in surface urban heat island conditions across observations (Schwarz, Lautenbach, and Seppelt 2011). Land cover driven comparative indicators compare representative urban versus a surrounding buffer: *difference urban-rural* (Tomlinson et al. 2010), or the difference between local representative land cover types: *difference urban – agricultural* (Jin, Dickinson, and Zhang 2005) and *difference urban-other* (Zhou et al. 2010) and so highlight the broadest range of conditions for each area. Urban heat island driven indicators efficiently summarize the distribution of the values and indicate single highest values such as *magnitude* (Rajasekar and Weng 2009) or the amount of land area impacted by increased heat such as, *hot island area* (Zhang and Wang 2008), and the *micro island* (Aniello et al. 2005).

Research Question

The overarching research question addresses the spatiotemporal distribution of surface urban heat island effects for large metropolitan areas within a temperate eco-region of the conterminous United States. As detailed above, the temperate mixed forested eco-region is especially sensitive to changes in the composition of the landscape, even when the overall change in spatial extent seems relatively minor. This research is assessing the useful of the various indicators for systemically comparing multiple locations in order to determine relative urban heat island intensity and to determine the degree to which the indicators consistently measure any change over time. This overarching research question further divided into three parts for clarity:

1. What is the spatial variation in SUHI values based on individual indicators for each metropolitan area aggregated by month for the year 2012?
2. When ranked according to SUHI value, do these rankings remain relatively consistent throughout the year and across each of the SUHI indicators?
3. Do the derived surface UHI indicators suggest a pattern of overall increases in monthly SUHI intensity from 2002 to 2012 for Louisville, KY and other locations within the same temperate region?

Recent studies suggest that cities such as Louisville, KY, Atlanta, GA, and Philadelphia, PA are undergoing significant increases in urban heat compared to other U.S. metropolitan areas (Stone 2012; Kentward et al. 2014) and I will assess whether this increase is consistently detected regardless of the type of indicator used to measure that change. In a similarly organized integrated study, surface and atmospheric temperatures were mostly consistent (~ 0.1ºC difference) during the day and only modestly divergent (within 2ºC) at night (Schwarz et al. 2012). Furthermore, when the locations (corresponding pixel and associated station location) are ranked and evaluated using Pearson coefficients, the relationship is significant, though only for the immediate sampling area. The larger an area used to aggregate the LST sample caused the correlation of the ranks (LSTs vs air) to decrease proportional to size of the sampling area. By evaluating the various indicators, I will assess whether or not consistently in measurements exists across each of the indicators as would be suggested by the documented relationship between LST and air temperatures.

Objectives and Hypotheses

In order to address these research questions, this study focuses on two primary objectives. The study objectives are to 1) determine the distribution of SUHI values for the study location by month and over time and 2) examine the degree to which the various indicators produce consistent results, both, on an annual basis and over time. The following hypotheses are addressed: 1) that the monthly value rankings for metropolitan areas are consistent across each SUHI indicator for the year 2012. In other words, does each indicator produce similar rankings for each time step? 2) That the SUHI value rankings for each metropolitan area are consistent across each month for the year 2012 regardless of SUHI indicator used for analysis. In other words, for a given indicator, do you get the same order each month of the year in 2012? 3) That the monthly SUHI indicators are consistent over time for all study locations for 2002 to 2012. Do we see any significant change during this time period for any of the locations for any of the indicators?

Rationale for Study

The urban heat island is a direct result of anthropogenic changes to the physical composition of the landscape over time and, as such, serves as a vivid example of man-made climate change (Stone 2012). The global human population is continuing to urbanize rapidly (U.N. 2015) and continued transformation of land surface from vegetative to impervious materials will contribute to increases in urban heat island intensity. Given this reality, the significance of the study involves the practical matter of creating a record of observation (baseline) of the UHI and the recent rate of growth in intensity as it relates to using the various remotely sensed metrics for the assessment of mitigation strategies. Identification of past and on-going patterns of urban heat development can assist urban managers and decision makers in coping with the uncertainty associated with planning for the impacts of future developments. In addition, through the comparison of the different remote-sensing metrics, this study will potentially yield important insights regarding the selection of appropriate measures of the surface UHI for further study within this temperate eco-region to determine the precise mix of casual factors that lead to the development of extreme urban heat. Lastly, this study will help us understand what exactly each UHI indicator is able to tell us about the size, intensity, and, more importantly, the impact (inferred from spatial-temporal distribution) of the UHI effect across this biologically important region.

CHAPTER 2. METHODS

Study Area

The UHI process of urban heating is driven primarily by the conversion of naturally vegetated landscapes to impervious surfaces and subsequent reductions in the latent heat flux (i.e. reduction in evapotransipirative cooling potential) compared to the surrounding (and still vegetated) hinterlands (Oke 1982). The magnitude of the urban heating is therefore not only a function of the amount of land cover change (i.e. size of urban area), but of the pre-existing ecological context (eco-regions) within which this change occurs. Therefore, to minimize the influence from differing climatic and vegetation regimes, the study region (Figure 1) is constrained to the land area contained within the temperature broadleaf and mixed forest eco-region of the U.S. as defined by Olson et al. (2001).

Within this temperate eco-region, the metropolitan areas selected for analysis are non-coastal urban areas with populations greater than 750,000 persons, for a total of 26. As the UHI effect is primarily a manifestation of urban development associated with increases in human population (Oke 1982), it is appropriate to use estimates of population density for delimiting urban settlements. United States Census Bureau designations of urbanized areas (UAs) include only those contiguous census tracts with densely settled populations of 50,000 people or more (Ratcliffe et al. 2016) including areas containing adjacent supporting infrastructure such as roads and shopping centers. Land areas falling within the delineated boundaries serve as the urban core for each of our study cities, while the land area falling beyond these boundaries serve as the non-urban or 'rural' areas.

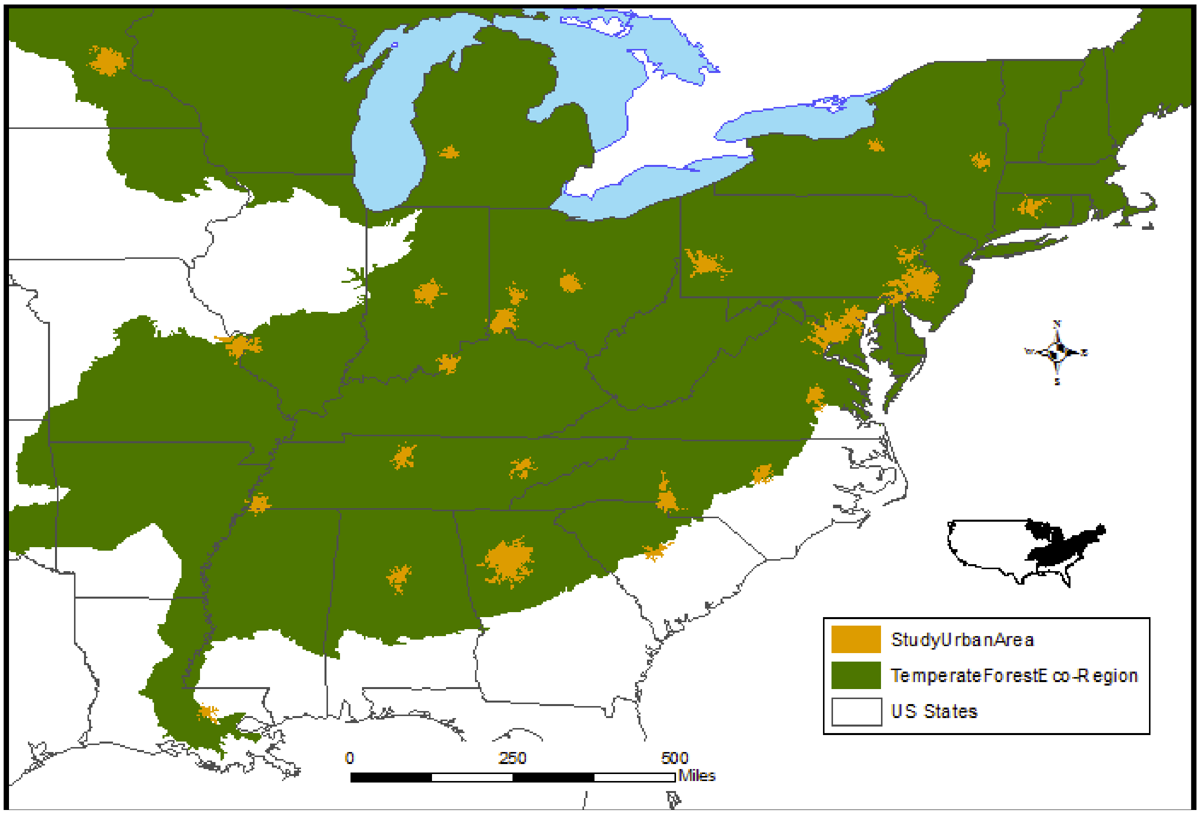


Figure 1. Study region is constrained to U.S. Temperate Broadleaf and Mixed Forest Eco-region (Olson et al. 2001) and urban areas with populations of at least 750,000.

Land Surface Temperature (LST) Data

This study uses the Moderate Resolution Imaging Spectroradiometer (MODIS) MOD11A2 8-day mean composite data for day-time land surface temperatures (LSTs) resampled to a spatial resolution of 500 meters using the cubic convolution resampling method before being zonally summarized. Many remote sensing-based studies have proven the suitability of the MODIS LST data products for use in urban heat analysis across a wide range of geographic settings and focused at regional spatial scales (i.e. Hung et al. 2006; Tomlinson et al. 2010; Imhoff et al. 2010; Schwarz, Lautenbach, and Seppelt. 2011). The larger extent of the MODIS imagery allows for the simultaneous measurements of much more land area at once and provides coincident measurements of each study location, potentially reducing the uncertainty associated with the observations.

The data was captured by NASA's Terra satellite over the study region at approximately 10:30-11 EST (15:30-16 UTC) and 2:30-3:30 EST (19:30-20 UTC). MODIS data level-1 pre-processing includes a generalized split window algorithm (Wan and Dozier 1996) that corrects for absorption and emission effects of the atmosphere. The MODIS composite data has proven consistent with *in situ* LST measurements with root mean squared differences of less than 0.5 K (Wan 2008). It has shown, however, that temporally aggregating MODIS data can inflate the difference in LSTs between urban and rural areas with the largest increases observed for summer months (Hu and Brunsell 2013).

Data procurement procedure was adapted from pre-existing work using the R programming language, exploiting the ftp download capabilities associated with the MODIS archive and evaluating the quality of each image before generating date specific image mosaics (Stevens 2015). Calculations on each 8-day MOD11A2 image layer implemented using the ArcPy package in Python resulted in 8-day SUHIx database tables. Data summaries compiled and hypothesis evaluations performed using a range of statistical packages and adapted code available for the R programming language. Complete references and adapted code attached in the appendix for reference.

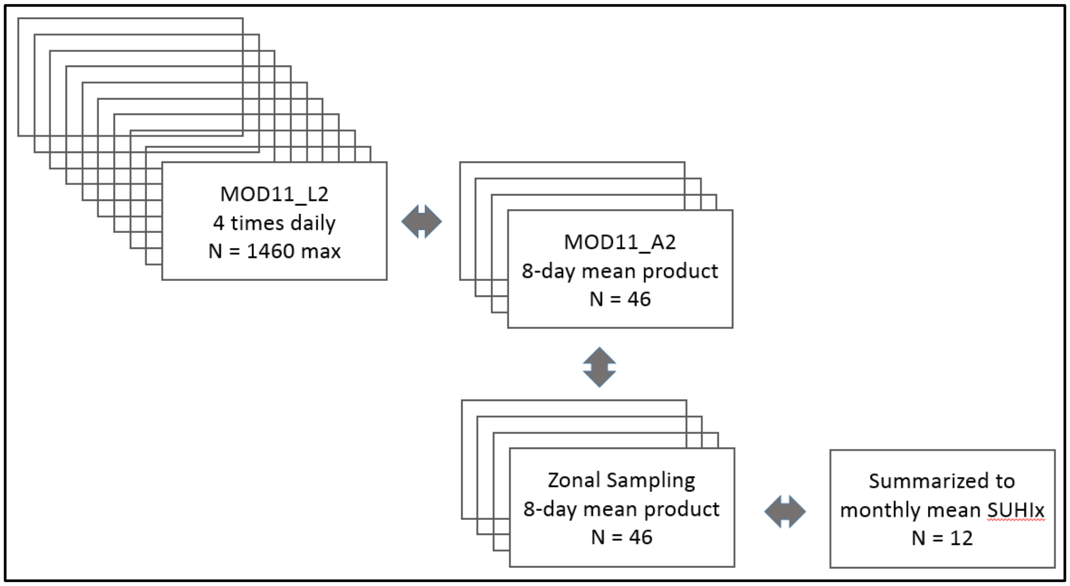


Figure 2. Diagram illustrates (approx.) the processing of the MODIS LST images to create the mean SUHI values used in the analysis for 2012 calendar year.

Land Cover Classification Data

Many of the SUHI indicators utilized in this study are derived from comparisons between zonal aggregations of LST pixels that fall within given land cover types. To represent the land cover characteristics of the study region, this study utilizes the 500m MODIS MCD12Q1 Collection 5 global land cover product derived from MODIS satellite multi-spectral imagery, ground-based 'truth-ing' data, and a range of supporting ancillary information. The land cover estimates are derived from ensemble decision-tree classification algorithms that are fed additional information describing the a-priori probability of land cover classes based on the prior years' data. Overall classifications accuracies are reported above 75% (Friedl et al. 2010) and generally within 85% agreement with the commonly utilized National Land Cover Dataset (NLCD) (Imhoff et al. 2010).

The land cover imagery downloaded for all available dates from 2002 to 2012, so land cover-specific measurements were constrained to those years only. The land cover classification product algorithms are based on an evolving ensemble of models, ancillary data and iterative training samples and have been shown to be 75% accurate in forest dominated areas, class-specific accuracies are varied (Friedl et al. 2010). Land cover classification labels are probabilistic and mostly appropriate for coarse representation of spectrally distinct classes, so small changes in area or in density may not be properly represented within. The global vegetation classification scheme (IGBP) layer extracted and reclassified from the original 17 classes into the four (4) composite classes and utilized to calculate the indicators in Table 2. Land cover proportions were quantified for each of the included study cities and incorporated below as attributes (or characteristics) of the observations. The land covers quantified include agriculture, forest, urban, and an 'other' category containing all other land covers, with water excluded.

Land cover proportions for each location are summarized for each annual time step by simply counting the number of pixels classified as each land cover type that falls within each urban boundary. The difference in land cover totals between the year 2012 and 2002 is used to describe any land change trajectories taking shape in the study region.

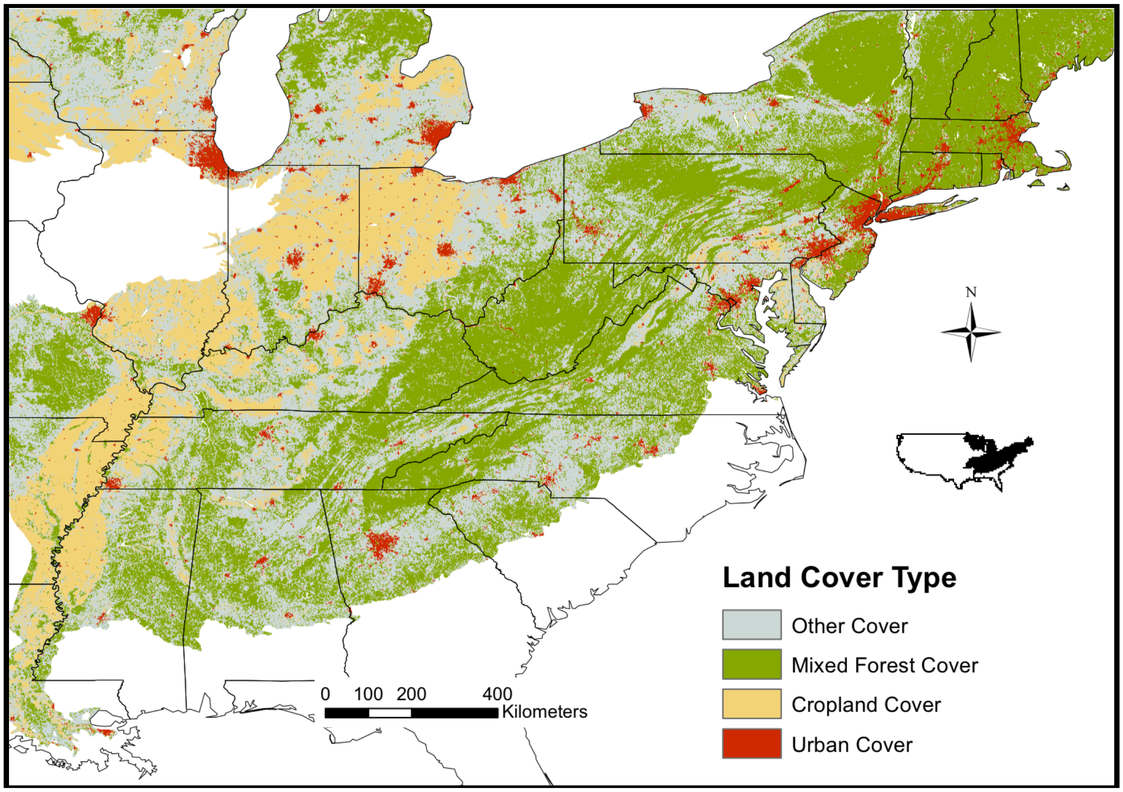


Figure 3. MODIS-derived land cover distribution across the land area within the temperate mixed forested eco-region at 250-meter spatial resolution (2012 is shown for reference).

Calculation of the Surface Urban Heat Island Effect 2012

For this study, six remote-sensing (6) indices were constructed from each MODIS LST image and then aggregated to monthly mean surface urban heat island (SUHI) effect for each of the twenty-six (26) largest major metropolitan areas within the temperate mixed forest eco-region for the year 2012. For each study location (N=26) six (6) different SUHI indicators are calculated; three based on the urban heat island-driven approach and three based on the land cover driven approach for each time step, and then summarized by month. Specific discussion of these indicators found in above sections highlight significant differences between them while original references provided in Table 1 provide additional source overview.

The land cover driven SUHI indicators (Table 1) differ primarily in how they represent the differences in land surface temperatures between representative ‘urban’ spaces versus ‘natural’ or ‘rural’ environments and each provides a different perspective (i.e. mean of area vs max value observed) Each of the land cover based calculations conveniently expressed as:

Mean LST of urban area – Mean LST of rural area = SUHIx (Eq. 1) or

Mean LST of urban cover – Mean LST of non-urban cover = SUHIx (Eq. 2),

where, the difference (SUHIx) in land surface temperatures (LST) between the urban and rural (non-urban) environments is aggregated by month for each of the indicators (SUHIx) for time period. The UHI driven indicators are statistical evaluations of the all observations that highlight highest individual occurrences (MAG) or measure the spatial extent (HIA, MIC) of increased LSTs and percentages (%) of total land area that meet some summary threshold expressed simply:

(CountP = = SC) / TotalP * 100 = SUHIx (Eq. 3),

where, all of the land area pixels (CountP) that meet a certain numeric threshold (==SC) are measured as a fraction of the total number of land area pixels (TotalP) that make up each study location. The fractional value is multiplied by 100 to convert to new value of percentage units (%). Calculations carried out for calendar year 2012 to evaluate hypotheses 1 and 2 to coincide with the latest available MODIS global land cover dataset, while the full length of the dataset (2002 to 2012) was considered for hypothesis 3. Complete code attached in the appendix for reference.

Calculation of Change over time 2002 to 2012

Much of the interest in this project involves the temporal patterns of SUHI intensity for each location in terms of change over time. In testing for significant change over time, each of the SUHI indicators that use only boundary designations (HIA, MAG, and DUR) are composed of LST values from the period during January 2002 to December 2012. The land cover driven indicators relied on MODIS land cover estimates that were not available for the full temporal extent of the MODIS LST data, so these indicators (DUA, DUO, and MIC) consider only January 2002 to December 2012. A simple linear model (SUHI ~ time) applied using R:

$$Y_m = C + T\beta_m + \varepsilon \qquad \text{(Eq. 4)}$$

where T is the length of time of each location's ($_m$) LST record (Y) and β equals the increase in Y per time step in T. C is the y-intercept constant and the error (ε) is assumed to be zero for the purposes here. The generated Beta coefficients (β) describe the approximate increase in SUHI over time with significance evaluated according to calculated p-values ($p < 0.05$). This simple analysis applied to all of the locations separately for each month and each indicator.

CHAPTER 3. RESULTS

Descriptive Overview of Surface Temperatures 2012

Here a summary is presented to outline the relevant measurements used to derive the estimates of the SUHI effect so that the reader gets a sense of the numeric distribution of values for this study region. Figure 3 below provides a good example of the way LSTs vary according to the distribution of the land cover. In the bottom left of the image is an automotive factory surrounded by impervious surfaces and, consequently, has higher temperatures relative to the still developing residential areas in the upper right of the image.

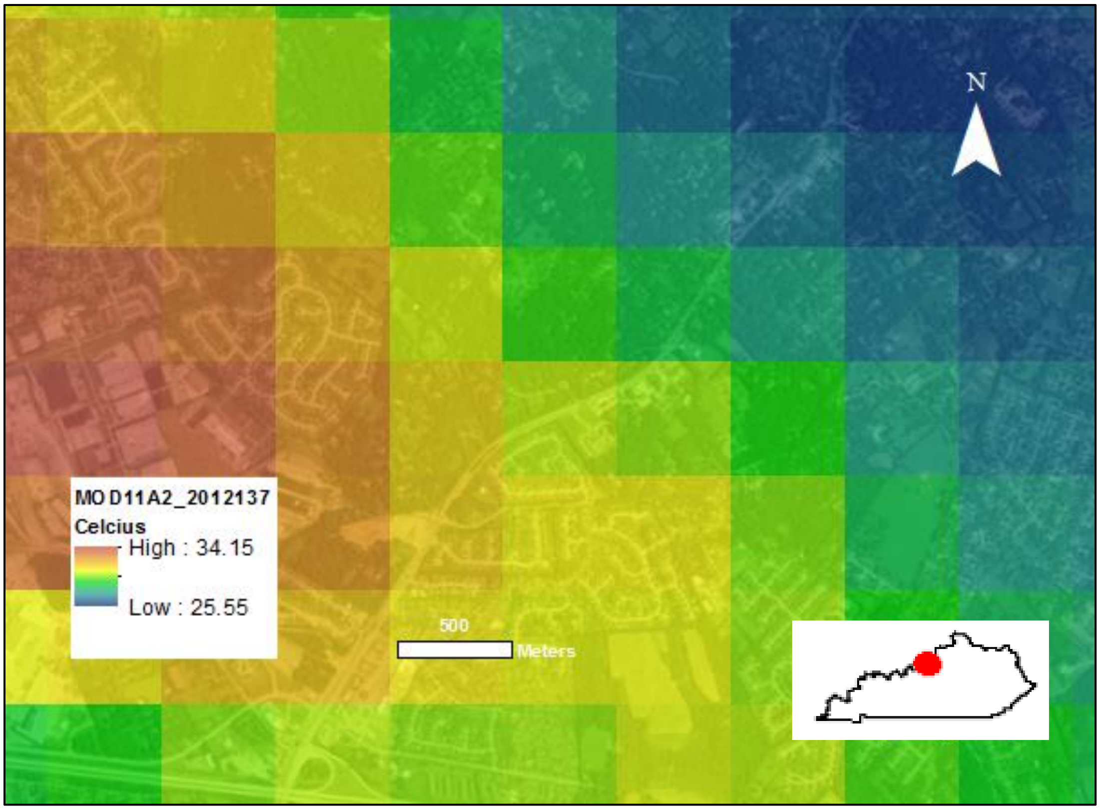


Figure 4. MODIS land surface temperature data (500m) with underlying RGB natural color image illustrates how temperatures graduate between land cover types. Image of northeast Louisville, KY in April 2012.

Figure 4 highlights the seasonal nature of mean LSTs for both the inner urban boundary as well as the mean LSTs of an associated 20km buffer representing the surrounding “rural” comparison. The relative difference in magnitude between the mean urban and rural temperatures is at a maximum during the summer months and converges significantly during the winter months. Annually, the urban and rural area mean LSTs for all of the study locations range from 5 - 35ºC. Figure 5 highlights the similar seasonal trends for other statistical summaries of LSTs within the urban boundaries of the study locations, including the minimum, maximum, and calculated range. The seasonal nature of the summary values is evident as it coincides with the peak observations in the summer months and are significantly cooler during the winter months. Rural and urban means essentially follow the same pattern as the summary statistics sharing coincident maximum and minimum mean temperatures trends.

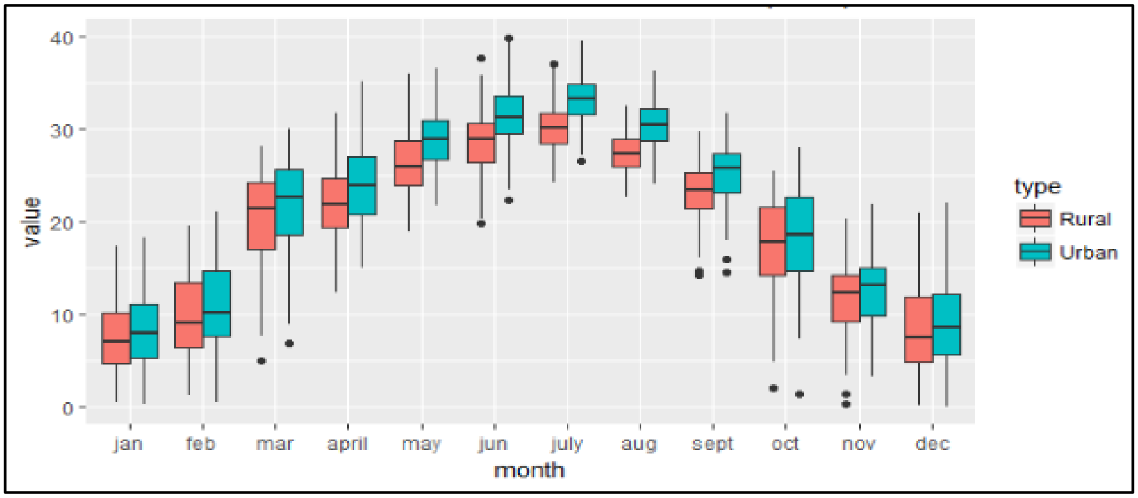


Figure 5. Box plot distributions of the LSTs within the urban boundary areas versus the surrounding rural area summarized by each month of 2012.

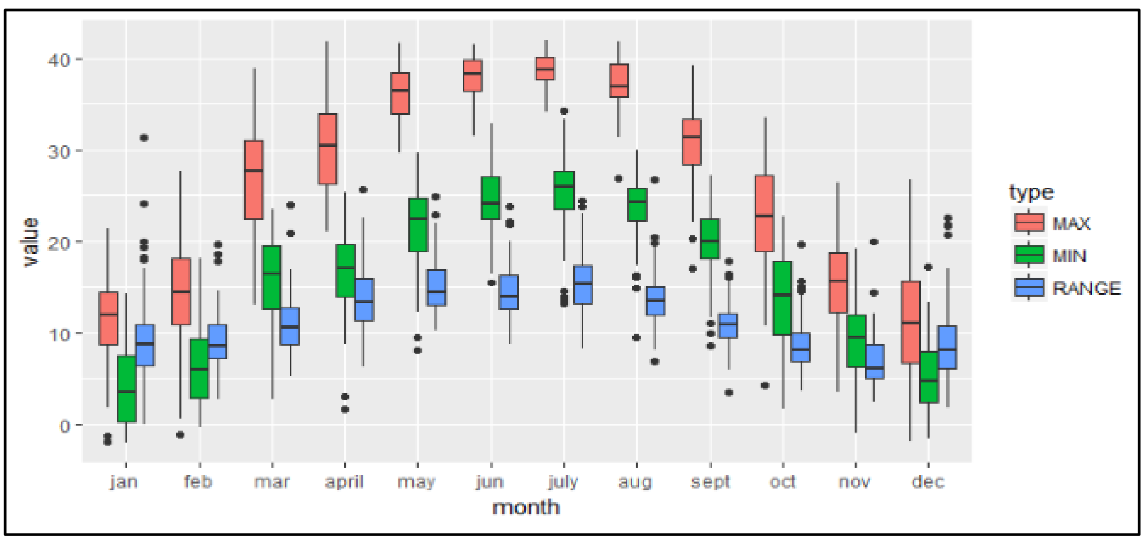


Figure 6. Descriptive summaries for within urban area boundaries by month for 2012. The maximum, minimum, and range shown as MAX, MIN, RANGE, respectively.

Land cover portions of land cover type for each urban area are quantified and summarized in Figure 6. Proportions of each land area vary considerably across each location, with a few areas having minimal coverage for some land types, minimal agricultural land in Hartford, CT or the sparsity of forested land cover in Indianapolis, IN for example. The land cover product does not make a distinction between different types of built-up areas, classified pixels tend to be permanent and are classified as urban when the amount of all impervious built areas (parking lots, building roofs, etc.) contained within the pixel boundary reaches a certain size (>50% of the overlying pixel). This type of land cover summary allows us to put into context the observed SUHI values for each location, for example, areas containing more classified urban pixels would presumably have highest SUHI values compared to an area dominated by forest. The land cover classes can be a bit limiting in terms of detail as a wide range of potential urban area types with, very different physical characteristics, are included in the same class, heavy industrial locations and more pedestrian locations like city-parks, for example. It is the case too that the newest urban developments do not show up in the data in terms of the generation of new urban classified pixels.

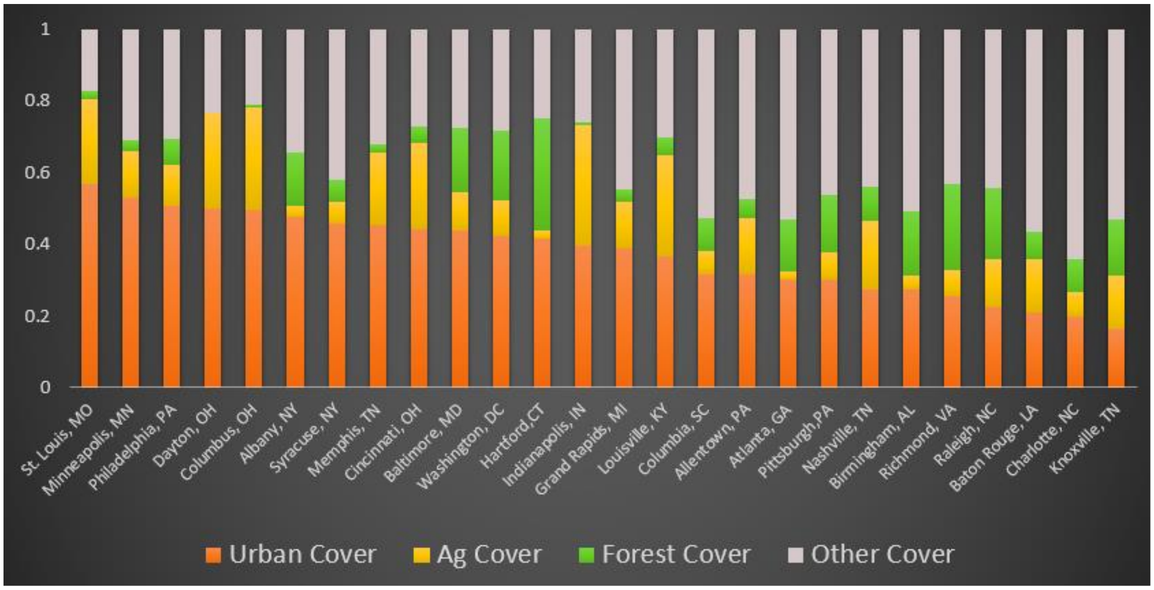


Figure 7. Land cover distribution (2012) calculated from the MODIS land cover data (250m) for each of the study areas. Original IGBP data are reclassified from the original 17 classes into 4 land cover classes needed as input for SUHI indicators.

Surface temperatures are aggregated spatially according to each land cover type and the difference taken between the mean values from each class (urban, ag, etc.) to obtain the relative differences between particular land cover types (*difference urban-agricultural* or *difference urban-other*, for example) as discussed in the methods above. Land surface temperature values, shown in Figure 7, show a similar seasonal pattern to summary statistics above in terms of the timing of peak values, seasonal trends in magnitude of the values, and the decreased intra-class variance during the summer months. Timing is especially important in this region because of the wide range in latitude and longitude included in this study region, particularly as it relates to the timing of the greening up of vegetation, as southern cities like Baton Rouge, LA are likely to begin the process earlier than Minneapolis, MN. This differential timing presumably is accounting for the higher variance in SUHI values leading up to the summer where the variance is considerably less. Surprisingly, classified urban land cover is generally cooler than agriculture lands for each time step, though this difference gets smaller as the year progresses to June and July. Agricultural land covers are the warmest throughout the year followed closely by classified urban, the other covers, while forested areas are the coolest as expected.

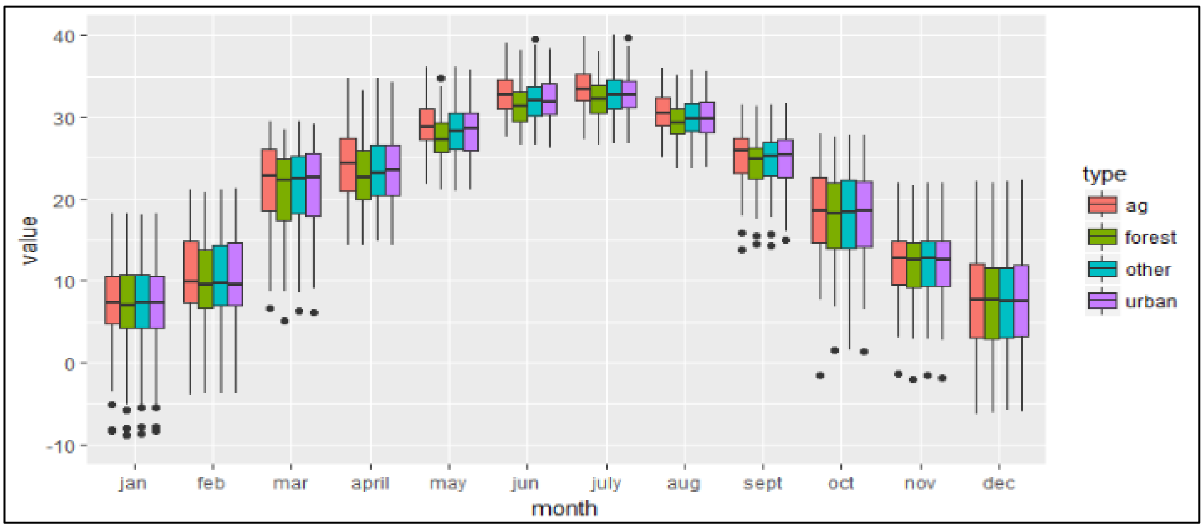


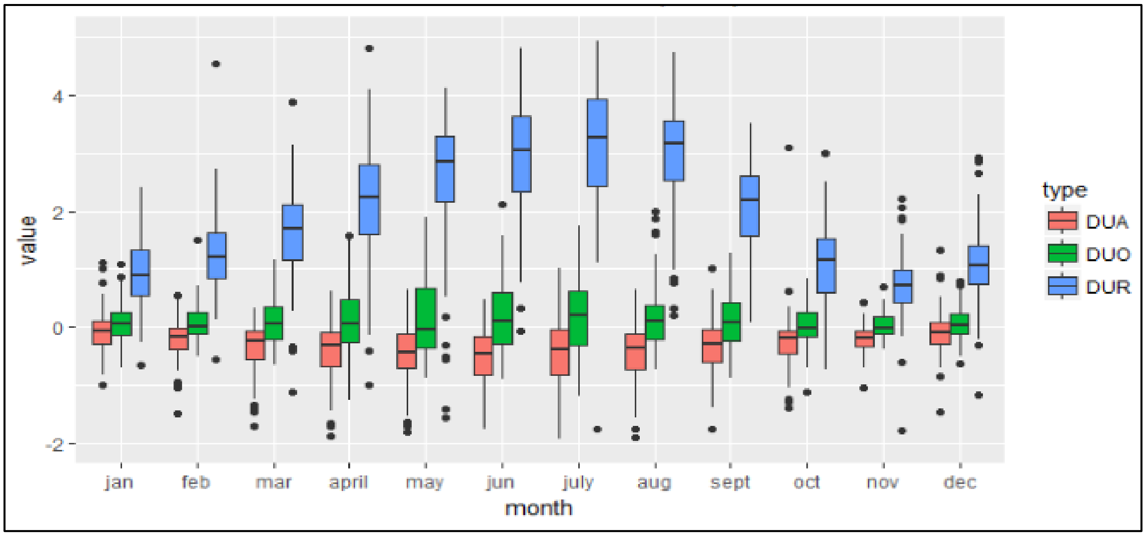
Figure 8. Box plot distributions of classified land cover-specific LSTs values for all locations and summarized by each month of 2012.

Descriptive Overview of Surface Urban Heat Island 2012

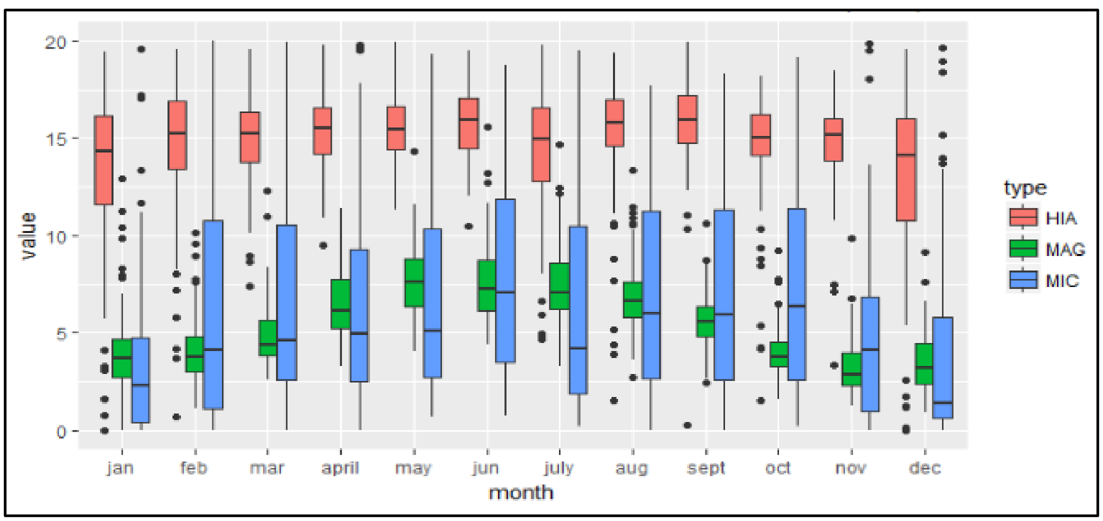
Calculation of the surface urban heat island were summarized into monthly mean values for the year 2012 for each of the indicators described in Table 1. The ***land cover driven*** indicators are highlighted together in Figure 8. This approach looks at the numeric distribution of LSTs for each representative land cover within each study location, per Equation 1 above. Each of the indicators follow the same general seasonal behavior as exhibited by the various land cover samples with smaller ranges in the winter months and the greatest variance in the summer times. The DUA indicator too highlights the higher temperatures observed year round for agricultural land, indicated by the continuous negative values, particularly during the summer months.

The ***urban heat island- driven indicators*** were summarized into monthly mean values for the year 2012 for each of the SUHIx indicators. Only the land areas within the urban boundaries are considered when aggregating surface temperature values for 2012, shown in Figure 9. In contrast to the land cover based comparisons treated above that indicate a highly seasonal pattern of variation, two of the three here do not reflect that pattern, instead, the HIA indicator is remarkably consistent throughout while the more-land-cover-sensitive MIC indicator follows multi phase pattern of high and low variance, presumably, coinciding with leaf on/off conditions. The timing of the phase shifts generally coincide with the start of the growing season and at the end of the growing season when most areas are cleared. The HIA is relatively consistent over the course of the year and does not share the same seasonal patterns seen with the previous indicators.

Overall, the DUR indicator is higher than the other indicators every month of the year and peaks in July, as do the urban and rural land area LSTs. Each of the remaining indicators peak earlier in the year in May, long before the peaks in LSTs for the respective land cover types. This behavior is shared among all types of generated values (land cover LSTs and calculated SUHIs) and each has the minimum variance in the peak annual warming (June-July) and substantially greater variance during spring and fall months, again coinciding with the leaf on/off transition periods.



<u>Figure 9</u>. Box plot distributions of calculated SUHI indicators *difference urban-ag, difference urban-other, and difference urban-rural*, shown in legend as DUA, DUO, and DUR, respectively and summarized by month.



<u>Figure 10.</u> Box plot distributions of calculated SUHI indicators *hot-island area, magnitude, and micro-island area*, shown in legend as HIA, MAG, and MIC, respectively and summarized by month.

Ranking Locations According to SUHIx for 2012

The numeric data represented in the graphs above are converted to rankings in order to compare the cities according to how large the respective SUHI is for each area. Rankings for each of the land cover driven indicators are shown in Tables 2-5 and illustrate very well the seasonal nature of the rankings. For example, if you look at Baton Rouge, LA the position of the DUR ranking shifts vry abruptly from near the bottom (19^{th}) in August to the top (1^{st}) for the next 3 months. Albany, NY ranks at the top (DUR) and near the bottom for each of the other indicators during the first several months of the year (February – June), highlighting again the differences in how each SUHI describes different aspects of each location. The behavior witnessed for other study locations like Birmingham, AL, Louisville, KY, and Nashville, TN, show abrupt changes in rank position connected to the seasonal nature of agricultural activities and the leaf on/off patterns of deciduous-dominated forest. The changes in SUHI rankings from month to month are likely influenced by site-specific localized and seasonal atmospheric conditions (weather) and associated environmental conditions (ecological context) and fall within the same context as the differential timing of the leaf on-off transitions.

	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	
Albany, NY	8	1	4	2	3	2	2	6	17	4	4	16	**Difference Urban - Rural (DUR)**
Allentown, PA	10	15	6	6	6	5	5	9	3	7	19	2	
Atlanta, GA	22	17	11	11	13	14	20	24	11	13	14	24	
Baltimore, MD	17	8	24	22	18	17	13	21	20	10	6	14	
Baton Rouge, LA	9	2	2	1	1	11	24	19	1	1	2	6	
Birmingham, AL	11	11	7	14	10	12	21	22	10	5	3	18	
Charlotte, NC	20	25	18	17	16	20	26	25	21	15	15	19	
Cincinnati, OH	26	20	23	16	17	19	9	5	14	24	20	20	
Columbia, SC	4	10	8	7	11	13	19	13	13	8	10	15	
Columbus, OH	19	9	12	20	23	21	1	1	6	20	9	5	
Dayton, OH	1	21	25	25	25	26	3	2	25	21	25	1	
Grand Rapids, MI	18	26	22	10	20	22	17	16	22	26	23	26	
Hartford, CT	5	6	5	12	8	7	8	15	2	9	17	9	
Indianapolis, IN	13	14	26	26	26	25	4	3	26	22	24	13	
Knoxville, TN	6	7	13	5	2	1	18	7	5	3	8	10	
Louisville, KY	15	4	10	4	5	3	11	14	4	6	12	7	
Memphis, TN	7	5	3	15	15	4	23	8	15	14	7	4	
Minneapolis, MN	21	24	19	24	24	24	7	12	24	25	26	3	
Nashville, TN	12	12	15	13	7	15	22	23	8	2	16	17	
Philadelphia, PA	23	19	14	21	19	16	10	20	18	17	13	12	
Pittsburgh, PA	25	13	16	18	14	10	14	17	16	16	21	21	
Raleigh, NC	24	22	21	23	22	23	25	26	23	23	18	8	
Richmond, VA	14	16	9	9	9	8	12	11	9	12	5	25	
St. Louis, MO	2	18	17	8	21	18	15	10	19	18	22	22	
Syracuse, NY	3	3	1	3	4	6	6	4	7	19	1	11	
Washington, DC	16	23	20	19	12	9	16	18	12	11	11	23	
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	
Albany, NY	14	26	26	26	26	26	26	26	24	26	26	26	**Difference Urban - Ag (DUA)**
Allentown, PA	25	18	23	23	21	22	21	21	22	22	9	25	
Atlanta, GA	9	15	17	14	14	16	7	5	12	17	18	8	
Baltimore, MD	12	24	24	24	25	24	23	23	23	25	24	22	
Baton Rouge, LA	17	2	2	1	3	2	9	16	2	2	1	5	
Birmingham, AL	3	1	3	9	4	3	4	3	5	3	2	3	
Charlotte, NC	6	19	14	18	10	5	3	6	10	15	15	9	
Cincinnati, OH	18	13	18	16	15	12	17	19	18	10	19	15	
Columbia, SC	21	5	10	5	7	8	12	15	3	12	8	10	
Columbus, OH	11	9	5	3	9	7	2	1	7	9	12	1	
Dayton, OH	*NA*	8	7	4	5	6	10	9	11	8	4	12	
Grand Rapids, MI	13	11	9	10	12	19	18	17	13	6	16	13	
Hartford, CT	22	25	25	25	24	25	25	25	26	24	22	21	
Indianapolis, IN	7	16	1	8	13	14	8	10	8	5	6	2	
Knoxville, TN	8	10	13	13	8	17	15	8	9	18	13	17	
Louisville, KY	4	6	4	7	6	10	14	7	4	4	5	4	
Memphis, TN	15	14	19	12	11	13	5	13	16	16	7	18	
Minneapolis, MN	19	21	21	20	20	18	19	18	21	20	23	7	
Nashville, TN	1	4	6	2	1	1	11	4	1	7	3	6	
Philadelphia, PA	24	23	20	22	22	23	24	22	25	23	25	23	
Pittsburgh, PA	10	20	12	17	16	15	16	14	15	19	20	19	
Raleigh, NC	5	3	8	6	2	4	1	2	6	13	10	14	
Richmond, VA	16	12	11	11	17	11	13	12	14	14	11	11	
St. Louis, MO	23	7	16	15	18	9	6	11	17	11	14	16	
Syracuse, NY	2	17	22	21	23	20	22	24	19	1	21	24	
Washington, DC	20	22	15	19	19	21	20	20	20	21	17	20	
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	

Table 2. Surface urban heats island (SUHI) rankings for all study locations for each month of the year. Noted here are the difference urban-rural (DUR) area (top) and difference urban-ag (DUA) land cover (bottom) indicators.

	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	Difference Urban - Other (DUO)
Albany, NY	9	22	21	22	21	23	22	20	17	13	18	21	
Allentown, PA	22	9	20	18	16	17	17	12	18	21	12	16	
Atlanta, GA	17	16	15	10	11	11	8	10	10	10	19	13	
Baltimore, MD	16	24	22	24	26	24	24	24	23	26	23	23	
Baton Rouge, LA	20	6	9	7	12	6	9	13	9	8	8	9	
Birmingham, AL	3	2	3	6	6	4	3	3	3	4	1	2	
Charlotte, NC	4	4	2	3	2	2	1	2	5	3	4	1	
Cincinnati, OH	15	18	26	23	25	20	20	23	24	18	22	19	
Columbia, SC	14	5	5	2	4	7	10	8	6	7	3	11	
Columbus, OH	7	12	12	12	17	16	5	9	16	14	15	7	
Dayton, OH	*NA*	11	10	9	9	10	15	15	13	16	10	15	
Grand Rapids, MI	11	10	13	14	10	9	11	7	11	17	9	17	
Hartford, CT	23	26	24	26	24	26	26	26	25	24	24	22	
Indianapolis, IN	8	20	7	15	18	18	19	17	14	11	13	14	
Knoxville, TN	2	1	1	1	1	1	2	1	1	1	2	3	
Louisville, KY	5	8	8	11	8	15	13	16	8	9	11	6	
Memphis, TN	18	21	23	20	20	22	14	19	22	20	17	24	
Minneapolis, MN	19	15	16	13	15	19	23	22	19	19	25	26	
Nashville, TN	1	3	4	4	3	3	6	4	2	2	6	4	
Philadelphia, PA	25	25	25	25	23	25	25	25	26	25	26	25	
Pittsburgh, PA	12	19	14	16	14	12	12	14	15	15	16	10	
Raleigh, NC	10	7	6	5	5	5	4	5	4	5	5	5	
Richmond, VA	13	13	11	8	7	8	7	6	7	6	7	8	
St. Louis, MO	24	14	19	21	19	14	16	18	20	12	21	18	
Syracuse, NY	6	17	17	17	13	13	18	11	12	22	14	12	
Washington, DC	21	23	18	19	22	21	21	21	21	23	20	20	

	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	Micro-Island (MIC)
Albany, NY	19	15	21	22	15	15	14	18	23	16	21	15	
Allentown, PA	14	13	20	19	9	13	12	20	9	13	16	9	
Atlanta, GA	22	25	13	20	17	22	21	22	19	17	17	20	
Baltimore, MD	17	9	14	12	14	21	19	14	14	11	15	13	
Baton Rouge, LA	13	12	16	16	21	16	13	19	20	14	23	17	
Birmingham, AL	5	8	4	6	4	7	11	11	4	2	4	3	
Charlotte, NC	21	22	19	18	19	23	20	17	15	22	9	26	
Cincinnati, OH	10	4	11	13	16	8	6	6	12	15	7	14	
Columbia, SC	18	17	18	15	22	18	16	15	18	21	20	18	
Columbus, OH	6	6	7	3	6	9	8	7	6	4	10	6	
Dayton, OH	1	1	1	1	1	1	1	1	1	1	1	2	
Grand Rapids, MI	7	7	9	7	3	4	4	4	7	6	5	8	
Hartford, CT	24	21	23	23	25	24	22	25	26	26	25	19	
Indianapolis, IN	2	3	3	2	2	2	2	2	2	3	3	5	
Knoxville, TN	20	20	22	21	20	19	15	16	16	19	13	11	
Louisville, KY	8	5	5	8	7	3	5	3	5	7	6	16	
Memphis, TN	3	11	6	4	5	5	10	5	3	5	2	1	
Minneapolis, MN	16	19	17	11	10	12	9	10	10	20	18	21	
Nashville, TN	11	10	8	9	8	11	17	12	8	9	8	7	
Philadelphia, PA	25	26	26	24	26	25	26	24	24	23	24	24	
Pittsburgh, PA	23	23	25	25	23	26	25	26	25	24	22	23	
Raleigh, NC	9	14	10	14	13	14	18	13	11	10	12	10	
Richmond, VA	15	18	15	17	18	17	24	21	17	18	19	22	
St. Louis, MO	4	2	2	5	12	6	3	9	13	12	11	4	
Syracuse, NY	12	16	12	10	11	10	7	8	21	8	14	12	
Washington, DC	26	24	24	26	24	20	23	23	22	25	26	25	
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	

Table 3. Surface urban heats island (SUHI) rankings for all study locations for each month of the year. Noted here are the difference urban-other land cover (top) and micro-island area (bottom) indicators.

Hot Island Area (HIA)

	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec
Albany, NY	19	10	5	6	10	7	4	5	26	18	13	11
Allentown, PA	4	4	8	15	3	4	2	2	4	21	16	19
Atlanta, GA	15	20	16	4	11	17	18	25	19	14	20	20
Baltimore, MD	9	7	25	21	6	14	12	17	16	23	15	13
Baton Rouge, LA	3	5	6	1	1	8	25	26	5	3	2	12
Birmingham, AL	5	19	15	24	18	21	21	24	20	13	7	3
Charlotte, NC	16	24	23	16	26	16	19	23	8	8	18	4
Cincinnati, OH	25	3	3	9	17	9	13	6	14	11	6	1
Columbia, SC	2	12	26	26	22	26	23	14	2	10	21	5
Columbus, OH	18	21	12	23	20	23	11	8	22	16	12	25
Dayton, OH	21	23	20	17	19	12	14	11	12	2	19	17
Grand Rapids, MI	22	25	18	7	8	22	6	7	11	20	24	24
Hartford, CT	8	11	19	5	7	3	3	4	7	19	17	2
Indianapolis, IN	24	15	21	13	25	24	10	15	9	4	8	18
Knoxville, TN	17	9	14	11	23	5	22	12	24	17	3	9
Louisville, KY	13	1	1	3	2	2	15	3	3	5	22	8
Memphis, TN	12	17	13	8	5	10	24	19	25	1	14	14
Minneapolis, MN	23	26	24	25	13	13	7	9	21	7	26	23
Nashville, TN	1	22	10	12	14	18	26	13	13	25	9	10
Philadelphia, PA	6	16	11	14	21	20	16	21	15	26	11	6
Pittsburgh, PA	26	2	22	18	15	19	5	10	17	9	23	22
Raleigh, NC	11	18	7	22	16	6	17	22	18	12	5	15
Richmond, VA	10	13	9	10	12	11	8	18	6	6	4	7
St. Louis, MO	14	8	4	20	24	25	20	16	23	22	25	21
Syracuse, NY	20	6	2	2	4	1	1	1	1	24	1	26
Washington, DC	7	14	17	19	9	15	9	20	10	15	10	16

Magnitude (MAG)

	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec
Albany, NY	3	24	19	19	17	17	10	17	26	18	11	9
Allentown, PA	13	22	26	20	22	12	11	8	10	19	18	18
Atlanta, GA	10	1	8	3	7	7	12	9	3	2	3	2
Baltimore, MD	8	18	18	5	6	3	2	4	5	6	20	16
Baton Rouge, LA	22	4	11	22	24	10	26	15	16	5	8	13
Birmingham, AL	15	9	5	4	5	4	14	18	1	3	17	17
Charlotte, NC	17	3	13	8	11	16	15	16	8	14	12	11
Cincinnati, OH	4	16	6	16	12	14	16	21	17	13	15	24
Columbia, SC	11	10	21	10	16	21	24	11	7	23	7	19
Columbus, OH	16	23	20	21	26	26	19	25	25	25	19	20
Dayton, OH	24	25	24	25	21	24	22	26	23	26	16	26
Grand Rapids, MI	23	26	23	26	25	25	23	24	24	16	22	25
Hartford, CT	12	13	17	17	20	6	5	3	14	8	14	23
Indianapolis, IN	2	21	9	23	23	22	13	13	12	24	26	8
Knoxville, TN	19	19	16	13	4	20	25	22	9	17	25	21
Louisville, KY	18	8	15	9	8	11	9	5	18	21	2	14
Memphis, TN	20	11	4	12	19	18	7	6	6	1	4	15
Minneapolis, MN	9	14	1	24	18	19	17	23	22	11	5	1
Nashville, TN	21	20	7	14	9	5	18	2	15	10	13	22
Philadelphia, PA	5	7	3	1	3	1	1	1	2	7	6	6
Pittsburgh, PA	1	5	12	7	10	2	6	14	4	15	10	3
Raleigh, NC	25	12	25	6	13	23	21	19	21	20	21	10
Richmond, VA	26	17	22	15	1	13	20	7	20	12	24	12
St. Louis, MO	7	2	2	2	14	9	3	12	11	4	1	5
Syracuse, NY	6	6	10	18	15	15	4	10	19	22	9	4
Washington, DC	14	15	14	11	2	8	8	20	13	9	23	7
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec

Table 4. Surface urban heats island (SUHI) rankings for all study locations for each month of the year. Noted here the hot-island area (top) and magnitude (bottom) indicators.

Do the indicators give us the same monthly rankings 2012 (H 1)?

The next item involves the degree to which the rankings derived from each indicator agree at each time step. Here, The Spearman's Rho matrix is used to compare the correlation in the monthly rankings for each of the indicators. Results are summarized in Tables 5- 7.

Overall, the indicators do not consistently provide the same rankings, though July is a noteworthy exception. In July of 2012, the rankings generated from all of the indicators are correlated with one another except the MIC, likely because this is the warmest part of the year and the difference in LST is maximized between land covers and so too the SUHIs derived from them. This lends support to most urban heat island studies choosing to focus on this part of the year and this finding suggest that the chance of varying results due to choice of indicator is minimized for this period.

Regarding matches among particular indicators, the DUR and HIA are correlated during the months of February – June, August, October, and November. This pairing certainly makes sense given that high HIA values indicate a great portion of the local land area is relatively warmer and so likely to be still warmer than the rural areas. Each of these indicators describe a slightly different aspect of the SUHI, on the difference between the urban and rural (DUR) and the total land area that is impacted by LST greater than the local mean (HIA) and both of them generally reinforce each other in terms of the physical processes involved in the UHI. This reinforces the idea that a larger percentage of urban land as a total portion of land area drives a larger difference in the overall mean temperatures in that urban area, relative to the surroundings. Essentially, the large warm core of LST increases the mean of the urban area, leading to a larger difference between those local urban and rural areas. The HIA indicator is also often correlated (January, October, and November) with the MAG variable, which makes sense too considering that higher large spatial clusters of warmer temperatures (HIA) and likely to have at least a few extreme values, which are particularly emphasized by the MAG indicator, especially during the cooler months of the year.

Another pair frequently correlated with one another are the DUA and the MIC indicators, significant in February to April and September to November. The relationship here though is less clear, though, it seems that the seasonal changes in vegetation, including forest leaf-on/off conditions and the agricultural cycle, are likely the factors here. As noted above, locations with heavy portions of agricultural lands, such as Birmingham, AL and Indianapolis, IN, lack corresponding areas of forest. Other indicators seem to have often matches, such as the correlation between the DUA and DUO indicators, or the high correlations between the MIC and MAG indicators; however, they are not always consistent throughout the year compared to the instances just noted.

	DUA	DUO	DUR	HIA	MAG	MIC	value	month
DUA	1	0.83	-0.07	-0.23	-0.13	0.35	Correlation	january
DUO	0.83	1	0.04	-0.17	-0.12	0.23		
DUR	-0.07	0.04	1	0.02	-0.33	0.33		
HIA	-0.23	-0.17	0.02	1	-0.46	-0.19		
MAG	-0.13	-0.12	-0.33	-0.46	1	-0.18		
MIC	0.35	0.23	0.33	-0.19	-0.18	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.7451	0.2737	0.5381	0.0863	p-value	
DUO	0		0.8409	0.4123	0.5603	0.2737		
DUR	0.7451	0.8409		0.9089	0.1005	0.0976		
HIA	0.2737	0.4123	0.9089		0.0193	0.3444		
MAG	0.5381	0.5603	0.1005	0.0193		0.3859		
MIC	0.0863	0.2737	0.0976	0.3444	0.3859			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.78	0.05	-0.1	0.09	0.5	Correlation	february
DUO	0.78	1	-0.06	-0.19	0.02	0.2		
DUR	0.05	-0.06	1	0.54	0.06	0.08		
HIA	-0.1	-0.19	0.54	1	0.22	0.04		
MAG	0.09	0.02	0.06	0.22	1	-0.39		
MIC	0.5	0.2	0.08	0.04	-0.39	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.8228	0.643	0.6719	0.0096	p-value	
DUO	0		0.7843	0.355	0.9195	0.3256		
DUR	0.8228	0.7843		0.0049	0.7716	0.6866		
HIA	0.643	0.355	0.0049		0.2775	0.8306		
MAG	0.6719	0.9195	0.7716	0.2775		0.0506		
MIC	0.0096	0.3256	0.6866	0.8306	0.0506			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.76	-0.2	0	-0.08	0.52	Correlation	march
DUO	0.76	1	-0.08	-0.18	-0.2	0.25		
DUR	-0.2	-0.08	1	0.39	0.03	-0.22		
HIA	0	-0.18	0.39	1	0.08	0.21		
MAG	-0.08	-0.2	0.03	0.08	1	0.11		
MIC	0.52	0.25	-0.22	0.21	0.11	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.3307	0.9934	0.6964	0.0066	p-value	
DUO	0		0.6866	0.3766	0.3392	0.2221		
DUR	0.3307	0.6866		0.0467	0.867	0.276		
HIA	0.9934	0.3766	0.0467		0.7063	0.3074		
MAG	0.6964	0.3392	0.867	0.7063		0.6075		
MIC	0.0066	0.2221	0.276	0.3074	0.6075			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.69	-0.03	-0.08	-0.16	0.55	Correlation	april
DUO	0.69	1	0.12	-0.18	-0.02	0.15		
DUR	-0.03	0.12	1	0.58	-0.04	-0.22		
HIA	-0.08	-0.18	0.58	1	-0.23	-0.14		
MAG	-0.16	-0.02	-0.04	-0.23	1	-0.3		
MIC	0.55	0.15	-0.22	-0.14	-0.3	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.8775	0.6915	0.4241	0.0034	p-value	
DUO	0		0.5593	0.3803	0.9247	0.4622		
DUR	0.8775	0.5593		0.0021	0.8618	0.2775		
HIA	0.6915	0.3803	0.0021		0.2495	0.483		
MAG	0.4241	0.9247	0.8618	0.2495		0.1367		
MIC	0.0034	0.4622	0.2775	0.483	0.1367			
	DUA	DUO	DUR	HIA	MAG	MIC		

Table 5. Spearman's Rho rank correlation values and associated p-values for January thru April for the difference urban-agriculture, difference urban-other, difference urban-rural, hot-island, magnitude, and micro-island indicators.

	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.72	-0.02	-0.24	-0.08	0.32	Correlation	may
DUO	0.72	1	0.22	-0.25	0.15	0.14		
DUR	-0.02	0.22	1	0.5	0.24	-0.35		
HIA	-0.24	-0.25	0.5	1	-0.13	0.03		
MAG	-0.08	0.15	0.24	-0.13	1	-0.44		
MIC	0.32	0.14	-0.35	0.03	-0.44	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.9195	0.2453	0.7088	0.1059	p-value	
DUO	0		0.276	0.2118	0.4767	0.483		
DUR	0.9195	0.276		0.0097	0.2356	0.0756		
HIA	0.2453	0.2118	0.0097		0.5281	0.8827		
MAG	0.7088	0.4767	0.2356	0.5281		0.0233		
MIC	0.1059	0.483	0.0756	0.8827	0.0233			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.72	-0.34	-0.26	-0.19	0.34	Correlation	june
DUO	0.72	1	-0.1	-0.13	-0.21	0.06		
DUR	-0.34	-0.1	1	0.55	0.33	-0.26		
HIA	-0.26	-0.13	0.55	1	0	-0.05		
MAG	-0.19	-0.21	0.33	0	1	-0.51		
MIC	0.34	0.06	-0.26	-0.05	-0.51	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.0918	0.192	0.3603	0.0864	p-value	
DUO	0		0.6382	0.5392	0.2946	0.7614		
DUR	0.0918	0.6382		0.0038	0.0998	0.1944		
HIA	0.192	0.5392	0.0038		0.996	0.8074		
MAG	0.3603	0.2946	0.0998	0.996		0.0075		
MIC	0.0864	0.7614	0.1944	0.8074	0.0075			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.75	-0.49	-0.63	-0.41	0.23	Correlation	july
DUO	0.75	1	-0.59	-0.54	-0.62	-0.04		
DUR	-0.49	-0.59	1	0.71	0.25	0.34		
HIA	-0.63	-0.54	0.71	1	0.34	0		
MAG	-0.41	-0.62	0.25	0.34	1	-0.2		
MIC	0.23	-0.04	0.34	0	-0.2	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.0116	0.0006	0.0383	0.2685	p-value	
DUO	0		0.0015	0.0047	0.0007	0.8618		
DUR	0.0116	0.0015		0	0.2261	0.0904		
HIA	0.0006	0.0047	0		0.0932	0.9881		
MAG	0.0383	0.0007	0.2261	0.0932		0.3173		
MIC	0.2685	0.8618	0.0904	0.9881	0.3173			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.68	-0.2	-0.4	-0.23	0.31	Correlation	august
DUO	0.68	1	-0.22	-0.23	-0.24	0.06		
DUR	-0.2	-0.22	1	0.61	-0.31	0.48		
HIA	-0.4	-0.23	0.61	1	-0.06	0.26		
MAG	-0.23	-0.24	-0.31	-0.06	1	-0.34		
MIC	0.31	0.06	0.48	0.26	-0.34	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0.0001	0.3307	0.0431	0.2495	0.1249	p-value	
DUO	0.0001		0.2867	0.264	0.2411	0.7538		
DUR	0.3307	0.2867		0.0009	0.1285	0.0141		
HIA	0.0431	0.264	0.0009		0.7894	0.1981		
MAG	0.2495	0.2411	0.1285	0.7894		0.0918		
MIC	0.1249	0.7538	0.0141	0.1981	0.0918			
	DUA	DUO	DUR	HIA	MAG	MIC		

Table 6. Spearman's Rho rank correlation values and associated p-values for May thru August for the difference urban-agriculture, difference urban-other, difference urban-rural, hot-island, magnitude, and micro-island indicators.

	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.81	0.14	0.15	-0.03	0.44	Correlation	september
DUO	0.81	1	0.16	0.15	-0.04	0.22		
DUR	0.14	0.16	1	0.31	0.07	-0.27		
HIA	0.15	0.15	0.31	1	-0.07	-0.12		
MAG	-0.03	-0.04	0.07	-0.07	1	-0.19		
MIC	0.44	0.22	-0.27	-0.12	-0.19	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.4914	0.4746	0.8696	0.0238	p-value	
DUO	0		0.448	0.4622	0.8332	0.2867		
DUR	0.4914	0.448		0.1258	0.7437	0.1873		
HIA	0.4746	0.4622	0.1258		0.7437	0.5458		
MAG	0.8696	0.8332	0.7437	0.7437		0.3462		
MIC	0.0238	0.2867	0.1873	0.5458	0.3462			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.42	-0.12	0.28	-0.25	0.65	Correlation	october
DUO	0.42	1	0.33	0.29	-0.12	0.14		
DUR	-0.12	0.33	1	-0.11	0.26	-0.18		
HIA	0.28	0.29	-0.11	1	-0.13	0.15		
MAG	-0.25	-0.12	0.26	-0.13	1	-0.25		
MIC	0.65	0.14	-0.18	0.15	-0.25	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0.0311	0.5593	0.1736	0.2156	0.0003	p-value	
DUO	0.0311		0.0998	0.1534	0.5481	0.5042		
DUR	0.5593	0.0998		0.5866	0.2017	0.3803		
HIA	0.1736	0.1534	0.5866		0.5303	0.452		
MAG	0.2156	0.5481	0.2017	0.5303		0.2143		
MIC	0.0003	0.5042	0.3803	0.452	0.2143			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.71	0.05	0.2	-0.1	0.52	Correlation	november
DUO	0.71	1	0.23	0.3	-0.32	0.36		
DUR	0.05	0.23	1	0.57	0.02	-0.24		
HIA	0.2	0.3	0.57	1	-0.46	-0.03		
MAG	-0.1	-0.32	0.02	-0.46	1	-0.12		
MIC	0.52	0.36	-0.24	-0.03	-0.12	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0	0.8048	0.3156	0.6169	0.0066	p-value	
DUO	0		0.2611	0.1386	0.1082	0.0738		
DUR	0.8048	0.2611		0.0024	0.9326	0.2397		
HIA	0.3156	0.1386	0.0024		0.0185	0.8827		
MAG	0.6169	0.1082	0.9326	0.0185		0.5638		
MIC	0.0066	0.0738	0.2397	0.8827	0.5638			
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA	1	0.56	0.05	0.29	-0.09	0.1	Correlation	december
DUO	0.56	1	0.01	0.55	-0.3	0		
DUR	0.05	0.01	1	-0.05	-0.06	0.12		
HIA	0.29	0.55	-0.05	1	-0.14	-0.38		
MAG	-0.09	-0.3	-0.06	-0.14	1	-0.33		
MIC	0.1	0	0.12	-0.38	-0.33	1		
	DUA	DUO	DUR	HIA	MAG	MIC		
DUA		0.0033	0.815	0.1667	0.6741	0.6317	p-value	
DUO	0.0033		0.968	0.0036	0.1348	0.996		
DUR	0.815	0.968		0.7982	0.7841	0.5803		
HIA	0.1667	0.0036	0.7982		0.4978	0.0561		
MAG	0.6741	0.1348	0.7841	0.4978		0.0998		
MIC	0.6317	0.996	0.5803	0.0561	0.0998			
	DUA	DUO	DUR	HIA	MAG	MIC		

Table 7. Spearman's Rho rank correlation values and associated p-values for Sept thru December for the difference urban-agriculture, difference urban-other, difference urban-rural, hot-island, magnitude, and micro-island indicators.

Do the rankings stay the same during 2012 (H 2)?

Here I answer the second hypothesis that deals with whether or not the indicators will generate the same order of rankings each month, given a particular indicator. As before, I make use of the Spearman's Rho ranking order correlation statistic for comparing the ranking of each indicator for correlation at each monthly time step. Results are summarized in Tables 8-10.

The land cover based indicators (DUA, DUO, MIC) that exploit the differences between land cover types show almost perfect stability throughout the year with significant rankings for each time step, as evidenced by the p-values. This suggest that the difference between each of the land cover types is fixed for a given location in a categorical sense and has little to no variation during the year. This further suggests that the relative difference in temperatures for a given set of land cover types is fairly consistent across cities, and that perhaps these LST are a function of the physical attributes of the location (i.e. relative portion of built land area).

The urban heat island based indicators (DUR, HIA, MAG) that compare only the distribution of values within the urban boundary is less consistent during the year. The DUR indicator, for instance, is less stable compared to the land cover indicators whereas the mid-year months do not correlate to the periods before and after. The summer months of June and July are not correlated with the other months, so that apparently, the rankings in the spring are no indication of the rankings to be observed in the summer. The HIA indicator is also is less stable during the first months of the year, but does show some consistency in the summer months when potential inter-class land cover differences are maximized. It certainly makes sense that the locations with the overall warmest surface features are likely to remain as such during the warmest months of the year. The HIA indicator follows the same general pattern as the MAG and does not correlate consistently with the other months.

	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan	1	0.49	0.44	0.46	0.36	0.36	0.12	0.42	0.26	0.39	0.26	0.28	DUR	Pearson Rho
feb	0.49	1	0.71	0.56	0.65	0.68	0.13	0.26	0.6	0.63	0.68	0.32		
mar	0.44	0.71	1	0.73	0.75	0.76	-0.06	0.06	0.72	0.58	0.68	0.21		
april	0.46	0.56	0.73	1	0.84	0.72	-0.12	0.08	0.67	0.6	0.51	-0.12		
may	0.36	0.65	0.75	0.84	1	0.87	-0.18	-0.09	0.79	0.83	0.65	-0.03		
jun	0.36	0.68	0.76	0.72	0.87	1	-0.04	0.07	0.7	0.71	0.62	0.02		
july	0.12	0.13	-0.06	-0.12	-0.18	-0.04	1	0.74	-0.04	-0.27	-0.19	0.3		
aug	0.42	0.26	0.06	0.08	-0.09	0.07	0.74	1	0.03	-0.25	-0.07	0.35		
sept	0.26	0.6	0.72	0.67	0.79	0.7	-0.04	0.03	1	0.67	0.54	0.14		
oct	0.39	0.63	0.58	0.6	0.83	0.71	-0.27	-0.25	0.67	1	0.6	0.04		
nov	0.26	0.68	0.68	0.51	0.65	0.62	-0.19	-0.07	0.54	0.6	1	-0.02		
dec	0.28	0.32	0.21	-0.12	-0.03	0.02	0.3	0.35	0.14	0.04	-0.02	1		
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan		0.0116	0.026	0.017	0.0691	0.0714	0.5436	0.0305	0.1932	0.0471	0.206	0.1617		P- value
feb	0.012		0	0.0029	0.0003	0.0002	0.5347	0.2042	0.0011	0.0005	1E-04	0.1098		
mar	0.026	0		0	0	0	0.7716	0.7538	0	0.0017	2E-04	0.309		
april	0.017	0.0029	0		0	0	0.557	0.6915	0.0002	0.0013	0.008	0.5729		
may	0.069	0.0003	0	0		0	0.3896	0.655	0	0	3E-04	0.8984		
jun	0.071	0.0002	0	0	0		0.8332	0.7162	0	0	8E-04	0.9089		
july	0.544	0.5347	0.7716	0.557	0.3896	0.8332		0	0.8618	0.177	0.344	0.133		
aug	0.031	0.2042	0.7538	0.6915	0.655	0.7162	0		0.8958	0.2208	0.749	0.0824		
sept	0.193	0.0011	0	0.0002	0	0	0.8618	0.8958		0.0002	0.004	0.4957		
oct	0.047	0.0005	0.0017	0.0013	0	0	0.177	0.2208	0.0002		0.001	0.854		
nov	0.206	0.0001	0.0002	0.0079	0.0003	0.0008	0.3444	0.7488	0.0043	0.0011		0.93		
dec	0.162	0.1098	0.309	0.5729	0.8984	0.9089	0.133	0.0824	0.4957	0.854	0.93			
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan	1	0.36	0.46	0.39	0.5	0.41	0.41	0.57	0.56	0.51	0.31	0.41	DUA	Pearson Rho
feb	0.36	1	0.78	0.88	0.88	0.85	0.69	0.7	0.86	0.76	0.82	0.64		
mar	0.46	0.78	1	0.9	0.84	0.78	0.67	0.73	0.89	0.75	0.79	0.81		
april	0.39	0.88	0.9	1	0.91	0.84	0.72	0.74	0.92	0.73	0.83	0.78		
may	0.5	0.88	0.84	0.91	1	0.88	0.75	0.81	0.93	0.65	0.83	0.71		
jun	0.41	0.85	0.78	0.84	0.88	1	0.84	0.79	0.86	0.68	0.78	0.74		
july	0.41	0.69	0.67	0.72	0.75	0.84	1	0.89	0.73	0.51	0.66	0.7		
aug	0.57	0.7	0.73	0.74	0.81	0.79	0.89	1	0.79	0.47	0.63	0.75		
sept	0.56	0.86	0.89	0.92	0.93	0.86	0.73	0.79	1	0.72	0.81	0.79		
oct	0.51	0.76	0.75	0.73	0.65	0.68	0.51	0.47	0.72	1	0.66	0.64		
nov	0.31	0.82	0.79	0.83	0.83	0.78	0.66	0.63	0.81	0.66	1	0.63		
dec	0.41	0.64	0.81	0.78	0.71	0.74	0.7	0.75	0.79	0.64	0.63	1		
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan		0.0785	0.02	0.0519	0.0111	0.0405	0.0426	0.0031	0.0033	0.0093	0.137	0.0397		P- value
feb	0.079		0	0	0	0	0	0	0	0	0	0.0005		
mar	0.02	0		0	0	0	0.0002	0	0	0	0	0		
april	0.052	0	0		0	0	0	0	0	0	0	0		
may	0.011	0	0	0		0	0	0	0	0.0004	0	0		
jun	0.041	0	0	0	0		0	0	0	0.0001	0	0		
july	0.043	0	0.0002	0	0	0		0	0	0.0073	2E-04	0		
aug	0.003	0	0	0	0	0	0		0	0.0146	5E-04	0		
sept	0.003	0	0	0	0	0	0	0		0	0	0		
oct	0.009	0	0	0	0.0004	0.0001	0.0073	0.0146	0		2E-04	0.0005		
nov	0.137	0	0	0	0	0	0.0002	0.0005	0	0.0002		0.0006		
dec	0.04	0.0005	0	0	0	0	0	0	0	0.0005	6E-04			
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		

Table 8. Spearman's Rho rank correlation values, associated p-values (shaded) for each month of 2012 for the difference urban-rural (DUR) and the difference urban-agriculture (DUA) indicators. Significant change denoted ($p < 0.05$).

	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan	1	0.58	0.72	0.62	0.64	0.6	0.63	0.7	0.75	0.65	0.69	0.75	DUO	Pearson Rho
feb	0.58	1	0.84	0.89	0.89	0.9	0.83	0.85	0.87	0.81	0.87	0.81		
mar	0.72	0.84	1	0.94	0.9	0.88	0.81	0.83	0.93	0.86	0.88	0.87		
april	0.62	0.89	0.94	1	0.95	0.92	0.86	0.87	0.94	0.86	0.87	0.83		
may	0.64	0.89	0.9	0.95	1	0.93	0.84	0.9	0.95	0.84	0.89	0.83		
jun	0.6	0.9	0.88	0.92	0.93	1	0.9	0.92	0.93	0.85	0.88	0.86		
july	0.63	0.83	0.81	0.86	0.84	0.9	1	0.93	0.86	0.85	0.84	0.89		
aug	0.7	0.85	0.83	0.87	0.9	0.92	0.93	1	0.91	0.8	0.9	0.87		
sept	0.75	0.87	0.93	0.94	0.95	0.93	0.86	0.91	1	0.89	0.92	0.89		
oct	0.65	0.81	0.86	0.86	0.84	0.85	0.85	0.8	0.89	1	0.81	0.85		
nov	0.69	0.87	0.88	0.87	0.89	0.88	0.84	0.9	0.92	0.81	1	0.85		
dec	0.75	0.81	0.87	0.83	0.83	0.86	0.89	0.87	0.89	0.85	0.85	1		
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan		0.003	0	0.0009	5E-04	0.0014	0.0008	0.0001	0	0.0004	1E-04	0		P- value
feb	0.003		0	0	0	0	0	0	0	0	0	0		
mar	0	0		0	0	0	0	0	0	0	0	0		
april	9E-04	0	0		0	0	0	0	0	0	0	0		
may	5E-04	0	0	0		0	0	0	0	0	0	0		
jun	0.001	0	0	0	0		0	0	0	0	0	0		
july	8E-04	0	0	0	0	0		0	0	0	0	0		
aug	1E-04	0	0	0	0	0	0		0	0	0	0		
sept	0	0	0	0	0	0	0	0		0	0	0		
oct	4E-04	0	0	0	0	0	0	0	0		0	0		
nov	1E-04	0	0	0	0	0	0	0	0	0		0		
dec	0	0	0	0	0	0	0	0	0	0	0			
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan	1	0.9	0.93	0.94	0.88	0.93	0.84	0.88	0.86	0.92	0.84	0.86	MIC	Pearson Rho
feb	0.9	1	0.85	0.85	0.77	0.87	0.84	0.83	0.76	0.83	0.75	0.79		
mar	0.93	0.85	1	0.94	0.84	0.87	0.76	0.85	0.82	0.89	0.85	0.78		
april	0.94	0.85	0.94	1	0.88	0.88	0.82	0.92	0.86	0.89	0.84	0.8		
may	0.88	0.77	0.84	0.88	1	0.89	0.79	0.85	0.91	0.92	0.85	0.81		
jun	0.93	0.87	0.87	0.88	0.89	1	0.92	0.92	0.84	0.86	0.81	0.78		
july	0.84	0.84	0.76	0.82	0.79	0.92	1	0.9	0.71	0.75	0.72	0.71		
aug	0.88	0.83	0.85	0.92	0.85	0.92	0.9	1	0.83	0.82	0.87	0.71		
sept	0.86	0.76	0.82	0.86	0.91	0.84	0.71	0.83	1	0.83	0.89	0.75		
oct	0.92	0.83	0.89	0.89	0.92	0.86	0.75	0.82	0.83	1	0.83	0.87		
nov	0.84	0.75	0.85	0.84	0.85	0.81	0.72	0.87	0.89	0.83	1	0.75		
dec	0.86	0.79	0.78	0.8	0.81	0.78	0.71	0.71	0.75	0.87	0.75	1		
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan		0	0	0	0	0	0	0	0	0	0	0		P- value
feb	0		0	0	0	0	0	0	0	0	0	0		
mar	0	0		0	0	0	0	0	0	0	0	0		
april	0	0	0		0	0	0	0	0	0	0	0		
may	0	0	0	0		0	0	0	0	0	0	0		
jun	0	0	0	0	0		0	0	0	0	0	0		
july	0	0	0	0	0	0		0	0	0	0	0		
aug	0	0	0	0	0	0	0		0	0	0	0		
sept	0	0	0	0	0	0	0	0		0	0	0		
oct	0	0	0	0	0	0	0	0	0		0	0		
nov	0	0	0	0	0	0	0	0	0	0		0		
dec	0	0	0	0	0	0	0	0	0	0	0			
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		

Table 9. Spearman's Rho rank correlation values, associated p-values (shaded) for each month of 2012 for the difference urban-other (DUO) and the micro-island area (MIC) indicators. Significant change denoted ($p < 0.05$).

	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan	1	0.05	0.08	-0.09	0.23	0.01	-0.43	-0.38	0.28	-0.24	0.21	0.43	HIA	Pearson Rho
feb	0.05	1	0.44	0.32	0.36	0.43	0.19	0.33	0.27	-0.08	0.19	0.15		
mar	0.08	0.44	1	0.5	0.31	0.47	-0.01	0.22	0.03	-0.18	0.42	0.03		
april	-0.09	0.32	0.5	1	0.54	0.55	0.12	0.22	0.27	0.02	0.25	0.08		
may	0.23	0.36	0.31	0.54	1	0.61	0.32	0.27	0.3	-0.03	0.05	-0.16		
jun	0.01	0.43	0.47	0.55	0.61	1	0.28	0.38	0.21	0.05	0.34	0.15		
july	-0.43	0.19	-0.01	0.12	0.32	0.28	1	0.65	0.21	-0.22	-0.13	-0.38		
aug	-0.38	0.33	0.22	0.22	0.27	0.38	0.65	1	0.2	-0.24	-0.2	-0.22		
sept	0.28	0.27	0.03	0.27	0.3	0.21	0.21	0.2	1	0.08	0.1	0.16		
oct	-0.24	-0.08	-0.18	0.02	-0.03	0.05	-0.22	-0.24	0.08	1	-0.05	0.1		
nov	0.21	0.19	0.42	0.25	0.05	0.34	-0.13	-0.2	0.1	-0.05	1	0.26		
dec	0.43	0.15	0.03	0.08	-0.16	0.15	-0.38	-0.22	0.16	0.1	0.26	1		
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan		0.7945	0.694	0.6744	0.264	0.956	0.0271	0.0537	0.1714	0.2342	0.2978	0.0282		P- value
feb	0.7945		0.023	0.1114	0.073	0.029	0.3532	0.1005	0.177	0.7137	0.3621	0.4705		
mar	0.694	0.023		0.0095	0.118	0.015	0.9643	0.276	0.8932	0.3711	0.0305	0.9011		
april	0.6744	0.1114	0.0095		0.004	0.004	0.5661	0.2899	0.1792	0.9405	0.2156	0.7038		
may	0.264	0.0732	0.118	0.004		0.001	0.1163	0.1861	0.1377	0.8696	0.8254	0.4359		
jun	0.9564	0.0293	0.0154	0.004	0.001		0.1617	0.0556	0.3026	0.8048	0.0884	0.4643		
july	0.0271	0.3532	0.9643	0.5661	0.116	0.162		0.0003	0.3026	0.2852	0.5172	0.059		
aug	0.0537	0.1005	0.276	0.2899	0.186	0.056	0.0003		0.3341	0.2397	0.3173	0.2821		
sept	0.1714	0.177	0.8932	0.1792	0.138	0.303	0.3026	0.3341		0.6891	0.624	0.4439		
oct	0.2342	0.7137	0.3711	0.9405	0.87	0.805	0.2852	0.2397	0.6891		0.7945	0.6193		
nov	0.2978	0.3621	0.0305	0.2156	0.825	0.088	0.5172	0.3173	0.624	0.7945		0.1981		
dec	0.0282	0.4705	0.9011	0.7038	0.436	0.464	0.059	0.2821	0.4439	0.6193	0.1981			
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan	1	0.2	0.43	0.19	0.04	0.33	0.62	0.16	0.3	0.03	0.26	0.47	MAG	Pearson Rho
feb	0.2	1	0.53	0.65	0.33	0.49	0.33	0.34	0.54	0.48	0.66	0.58		
mar	0.43	0.53	1	0.31	0.23	0.42	0.39	0.2	0.4	0.59	0.53	0.45		
april	0.19	0.65	0.31	1	0.72	0.61	0.46	0.44	0.71	0.49	0.32	0.36		
may	0.04	0.33	0.23	0.72	1	0.58	0.26	0.31	0.47	0.37	-0.03	0.27		
jun	0.33	0.49	0.42	0.61	0.58	1	0.61	0.63	0.62	0.66	0.29	0.28		
july	0.62	0.33	0.39	0.46	0.26	0.61	1	0.56	0.4	0.35	0.35	0.45		
aug	0.16	0.34	0.2	0.44	0.31	0.63	0.56	1	0.47	0.38	0.35	0.16		
sept	0.3	0.54	0.4	0.71	0.47	0.62	0.4	0.47	1	0.53	0.21	0.21		
oct	0.03	0.48	0.59	0.49	0.37	0.66	0.35	0.38	0.53	1	0.33	0.23		
nov	0.26	0.66	0.53	0.32	-0.03	0.29	0.35	0.35	0.21	0.33	1	0.39		
dec	0.47	0.58	0.45	0.36	0.27	0.28	0.45	0.16	0.21	0.23	0.39	1		
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		
jan		0.3173	0.0296	0.3603	0.857	0.099	0.0008	0.4419	0.1321	0.8879	0.1932	0.0152		P- value
feb	0.3173		0.0054	0.0003	0.099	0.011	0.0968	0.0891	0.0043	0.0139	0.0002	0.0018		
mar	0.0296	0.0054		0.1172	0.269	0.031	0.0519	0.3273	0.0404	0.0014	0.0052	0.0212		
april	0.3603	0.0003	0.1172		0	0.001	0.0172	0.025	0	0.0117	0.1106	0.0708		
may	0.8566	0.099	0.2685	0		0.002	0.1981	0.1249	0.0157	0.06	0.8775	0.1804		
jun	0.099	0.0113	0.0311	0.001	0.002		0.0009	0.0005	0.0007	0.0002	0.1454	0.1585		
july	0.0008	0.0968	0.0519	0.0172	0.198	9E-04		0.0027	0.0451	0.0792	0.0831	0.0201		
aug	0.4419	0.0891	0.3273	0.025	0.125	5E-04	0.0027		0.0156	0.059	0.0818	0.4419		
sept	0.1321	0.0043	0.0404	0	0.016	7E-04	0.0451	0.0156		0.0049	0.3074	0.2978		
oct	0.8879	0.0139	0.0014	0.0117	0.06	2E-04	0.0792	0.059	0.0049		0.0968	0.2596		
nov	0.1932	0.0002	0.0052	0.1106	0.878	0.145	0.0831	0.0818	0.3074	0.0968		0.0471		
dec	0.0152	0.0018	0.0212	0.0708	0.18	0.159	0.0201	0.4419	0.2978	0.2596	0.0471			
	jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec		

Table 10. Spearman's Rho rank correlation values, associated p-values (shaded) for each month of 2012 for the hot-island area (HIA) and the magnitude (MAG) indicators. Significant change denoted ($p < 0.05$).

Do the SUHI indicators show change 2002 to 2012 (H 3)?

The land cover driven indicators relied on MODIS land cover data that were not available for the full temporal extent of the MODIS LST data, so the analysis here is constrained to the 2002 to 2012 timeframe. Table 11 highlights the major trends in land cover change for each of the study locations. The most significant trend observed involves the almost universal loss of agricultural land compared to 2002 levels, such as the steep declines in Louisville, Syracuse, NY, and Washington, D.C. of 40%, 63%, and 36%, respectively. The major exception in terms of agriculture land loss is Atlanta’s gain of 84% and other comparatively modest increases in Memphis, TN and Baton Rouge, LA, at 6% and 1.5%, respectively. The amount of urban land cover is nearly all consistent (0%) with small increase observed in Pittsburg, PA of 0.2% and the only registered negative growth in urban land in Minneapolis, MN at -0.1% compared to 2002 levels. The other significant trend in the land cover is the large increase in forested land cover, including a tremendous increase of 180% in Syracuse, NY and other substantial increases in Columbus, OH, Indianapolis, IN, Philadelphia, PA, at 122%, 56%, and 46%, respectively. This increase in classified forested land cover is not as widespread as the loss of cropland, however, and some places like Louisville, KY, Atlanta, GA, and Raleigh, NC have significant decreases in overall forested land cover of 10%, 48%, and 21%, respectively.

While this study is not directly concerned with the specific mix of casual factors driving the variation in surface urban heat island values, observing the trajectory of land cover change lends additional context within which I can interpret the findings. Based on the observed trends above, I have some reasonable expectations as to how the SUHIx values have likely changed during the time 2002 to 2012.

	type	2002	2012	change%		type	2002	2012	change%
Albany, NY	agriculture	92	29	-68.5	Indianapolis, IN	agriculture	2504	1915	-23.5
	forest	467	1086	132.5		forest	68	106	55.9
	other	1077	520	-51.7		other	1927	2480	28.7
	urban	1499	1500	0.1		urban	2939	2940	0.0
Allentown, PA	agriculture	578	407	-29.6	Knoxville, TN	agriculture	683	424	-37.9
	forest	186	281	51.1		forest	723	497	-31.3
	other	1729	1805	4.4		other	2426	2911	20.0
	urban	1145	1145	0.0		urban	752	752	0.0
Atlanta, GA	agriculture	634	1168	84.2	Louisville, KY	agriculture	1419	862	-39.3
	forest	4042	2097	-48.1		forest	250	224	-10.4
	other	14746	16155	9.6		other	1527	2114	38.4
	urban	8377	8381	0.0		urban	1856	1856	0.0
Baltimore, MD	agriculture	835	523	-37.4	Memphis, TN	agriculture	1058	1121	6.0
	forest	1346	1505	11.8		forest	107	113	5.6
	other	2106	2251	6.9		other	1674	1605	-4.1
	urban	3335	3339	0.1		urban	2353	2353	0.0
Baton Rouge, LA	agriculture	581	589	1.4	Minneapolis, MN	agriculture	1437	658	-54.2
	forest	292	102	-65.1		forest	361	532	47.4
	other	2181	2362	8.3		other	3477	4177	20.1
	urban	802	803	0.1		urban	5975	5972	-0.1
Birmingham, AL	agriculture	215	215	0.0	Nashville, TN	agriculture	1147	819	-28.6
	forest	988	549	-44.4		forest	573	470	-18.0
	other	2796	3235	15.7		other	2624	3056	16.5
	urban	1506	1506	0.0		urban	1638	1639	0.1
Charlotte, NC	agriculture	550	523	-4.9	Philadelphia, PA	agriculture	2427	1311	-46.0
	forest	726	678	-6.6		forest	1455	2117	45.5
	other	5088	5150	1.2		other	6430	6867	6.8
	urban	1565	1567	0.1		urban	10648	10654	0.1
Cincinnati, OH	agriculture	1987	1109	-44.2	Pittsburgh, PA	agriculture	737	466	-36.8
	forest	369	357	-3.3		forest	1546	1940	25.5
	other	2238	3124	39.6		other	4398	4268	-3.0
	urban	3652	3656	0.1		urban	2858	2865	0.2
Columbia, SC	agriculture	271	259	-4.4	Raleigh, NC	agriculture	728	632	-13.2
	forest	353	301	-14.7		forest	1076	846	-21.4
	other	2121	2182	2.9		other	2401	2726	13.5
	urban	1267	1267	0.0		urban	1208	1209	0.1
Columbus, OH	agriculture	1542	989	-35.9	Richmond, VA	agriculture	371	333	-10.2
	forest	31	69	122.6		forest	1253	1312	4.7
	other	1125	1640	45.8		other	2226	2204	-1.0
	urban	2658	2658	0.0		urban	1331	1332	0.1
Dayton, OH	agriculture	968	651	-32.7	St. Louis, MO	agriculture	2291	1875	-18.2
	forest	7	8	14.3		forest	212	267	25.9
	other	846	1161	37.2		other	1673	2035	21.6
	urban	1822	1823	0.1		urban	5511	5511	0.0
Grand Rapids, MI	agriculture	383	230	-39.9	Syracuse, NY	agriculture	120	45	-62.5
	forest	100	169	69.0		forest	130	365	180.8
	other	1329	1413	6.3		other	860	706	-17.9
	urban	1159	1159	0.0		urban	943	945	0.2
Hartford, CT	agriculture	133	83	-37.6	Washington, DC	agriculture	1357	875	-35.5
	forest	1699	2484	46.2		forest	2750	2667	-3.0
	other	1355	619	-54.3		other	3910	4475	14.5
	urban	2253	2254	0.0		urban	5920	5925	0.1
	type	2002	2012	change%		type	2002	2012	change%

Table 11. Change in land cover portions (2002 to 2012) for each of the land cover classes used to derive the SUHI indicators are tabulated and the difference describes the trajectory of change.

Tables 12-14 below highlight the significant changes detected using the simple linear model (SUHI ~ time). The beta coefficients represent the modeled change in SUHI per time step (2002-2012?)for a given indicator for a given location.

Very evident among the land urban heat island-based indicators (DUR, HIA, MAG) are small, but consistent increases over time for most of the locations, particularly during the spring months of March, April May, coinciding with the early part of the temperate growing season. Many of the significant changes for these indicators are in geographically dispersed locations and are statistically significant during this time including Albany, NY, Atlanta, GA, and Charlotte, NC, each with significant increases in the spring months. It is certainly noteworthy that none of the locations register a decrease in the DUR indicator over time, no matter the situation with land cover trajectories as discussed in the previous section, perhaps suggesting a the link between vegetation increase and SUHI is less important at this particular time. Conversely, the same locations that have experienced large increases in forest cover have experienced a commensurate decrease in the HIA indicator, which is more a measure of the area impacted by high temperatures, as opposed to a measure of the potential difference between urban and non-urban areas. This is important to consider further, especially in terms of measuring the efficiency of common mitigation efforts undertaken by public agencies, particularly tree planting efforts. The MAG indicator is less telling and shows no discernible trends in terms of timing or direction of change. Only sporadic significant change is detected such as small increase in Louisville, KY in February and Albany, NY in June, and most locations indicate none at all.

Highly prominent among the land cover-based indicator (DUA, DUO, MIC) results is a consistent negative growth in the DUA indicator, which means the potential LST difference between the land cover classes in increasing, agricultural lands are apparently warming up compared to the urban areas, as evidence by the consistent significance. These increases, many of which are statistically significant, are particularly found during the growing season, both in early spring (March) and in the peak of summer growing season (July). This behavior is observed not only in places where agricultural land is decreasing such as Charlotte, NC and Louisville, KY, but also in the few locations that have increased their agricultural land holdings, such as Atlanta, GA and Baton Rouge, LA. The other land cover based indicators are more consistent and stable over time while the DUO indicator is less conclusive in terms of any popular trending or change over time. In most cases, the DUO indicator shows no change over time, is not consistent in terms of direction of any change, and is not statistically significant in most cases. The "other" land cover category is a kind of catch for land cover pixel falling outside of the major categories, so lack of systemic behavior is not unexpected. The MIC indicator likewise is highly variable in terms of strength of measured change, the direction of change, timing of change, and the statistical significance of that change. Many site specific exceptions are found, however, with particularly strong examples of increases over time for this indicator found at Birmingham, AL, Cincinnati, OH, and Indianapolis, IN that all show significant increases during several months of the year. Conversely, Grand Rapids MI, and Albany, NY, Columbia, SC each register significant decreases for multiple months of the year, particularly during the summer.

		jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	
Albany, NY	DUR	0	0.02	0.04 *	0	0.01	0.02 *	0.01 *	0.01	0	0	0	0	
Albany, NY	DUO	0	0	0	0	0	0	0	0	0	0	0	0	
Albany, NY	DUA	0	0	-0.0 *	0	0	0	-4.6	0	0	0	0	0	
Albany, NY	MIC	0.02	0.04	0	-0.5 *	-0.5 *	-0.5 *	-0.4 *	-0.2	-0.3 *	-0.3	0	0	
Albany, NY	MAG	0	0	-0.0 *	0	0.01	0.04 *	0.03	-0.0 *	0	0	0.01	0.04	
Albany, NY	HIA	1.01	-5.3	0.15 *	0	0.01	0	0	0.07	-0.6	0.01	0.05	-2.1	
Allentown, PA	DUR	0	0.01	0.03 *	0.01	0.02 *	0	0	0.01	0	0	0	0.01	
Allentown, PA	DUO	0	0	0	0	0	0.01 *	0.00 *	0.00 *	0.00 *	0	0	0	
Allentown, PA	DUA	0	0	-0.0 *	-0.0 *	0	0	0	0	0	0	0	0	
Allentown, PA	MIC	0.06	0.05	0.08	-0.6 *	-2.3	0.03	-0.4 *	-0.3	-0.6 *	-0.2	-0.2	-0.1	
Allentown, PA	MAG	0	0	-0.0 *	0.03	0.03	0.01	0.01	0	0	0	0	0	
Allentown, PA	HIA	2.12	-0.6	0.04	0	0.06	0.24	0.02	0.08 *	0	0	0.04	0	
Atlanta, GA	DUR	0	0	0.01 *	0.01 *	0	0.01 *	0	0	0.01 *	0	0	0	
Atlanta, GA	DUO	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	
Atlanta, GA	DUA	0	0	0	0	-0.0 *	-0.0 *	0	0	0	0	0	4.72	
Atlanta, GA	MIC	0.07	0.02	0.19 *	0.06	0	0.02	0	0	0.12	0.24 *	0.09	0	
Atlanta, GA	MAG	0	0.03	0.01	0	0.01	0	0	0	0	0	0.04	0.03	
Atlanta, GA	HIA	0	0	-0.3	0.03	0.01	0.03	0.26	0	-0.5	0.06	0.01	0	
Baltimore, MD	DUR	0	0	0	0	0.01 *	0	0.01	0	0	0.01	0.01	0	
Baltimore, MD	DUO	0	0	0	0	0	0	0	0	0	-0.0 *	0	0	
Baltimore, MD	DUA	0	0	0	0	0	0	0	0	0	0	0	0	
Baltimore, MD	MIC	0.14	0.09	0	-0.1	0.14	0.18	0	0	0.04	0.15	0.25 *	-0.1	
Baltimore, MD	MAG	0.01	0	0	0.01	0.02	0	0	0.01	-0.0 *	0	0	0	
Baltimore, MD	HIA	0	0	0.89	0.12	-2.2	0.01	0.05	0	0.01	0.01	0.06	0	
Baton Rouge, LA	DUR	0	0.01	0.02 *	0.01 *	0.01	0.02 *	0.01	0	0.01	0	0	0.01	
Baton Rouge, LA	DUO	-0.0 *	0	0	0	-0.0 *	0	0	-0.0 *	0	-0.0 *	0	0	
Baton Rouge, LA	DUA	0	0	0	0	0	0	-0.0 *	-0.0 *	0	0	0	0	Beta Coefficient
Baton Rouge, LA	MIC	0.19	-0.7	0.12	0.18	0.01	0.29 *	-1.9	0.31 *	0.04	0.22	0.03	0.03	
Baton Rouge, LA	MAG	0	0	0	-0.0 *	0	0.03	-0.0 *	0.01	0.01	0.04 *	0	0.01	
Baton Rouge, LA	HIA	0.09	0.09	0.07	0.04	0.02	0.06	-0.7	-4.6	0	-0.0 *	0.02	0	
Birmingham, AL	DUR	0	0.01 *	0.01	0	0	0.02 *	0	0	0	0.01	5.75	0	
Birmingham, AL	DUO	0	0	0	0	0	0	0	-0.0 *	0	0	0	0	
Birmingham, AL	DUA	0	0	0	0	0	-0.0 *	0	0	0	0	0	0	
Birmingham, AL	MIC	0.46 *	0.43 *	0.51 *	0.46	0.60 *	0.57 *	0.02	0.31	0.48 *	0.96 *	0.33	0.45	
Birmingham, AL	MAG	0	0	8.3	0	0.01	0.03	0	0	0.01	0	0	0	
Birmingham, AL	HIA	0	0.06	0.03	0.31	0	0.01	-0.1	0	-0.3	0.09	0	0	
Charlotte, NC	DUR	0	0	0	0.03 *	0.01 *	0.02 *	0.01 *	0.01 *	0	0	0	0	
Charlotte, NC	DUO	-0.0 *	0	-0.0 *	0	-0.0 *	-0.0 *	0	0	-0.0 *	0	0	0	
Charlotte, NC	DUA	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	
Charlotte, NC	MIC	0	-0.1	-0.9	-0.2	-0.1	-0.1	0	-0.3 *	-0.1	-0.2	-0.1	0	
Charlotte, NC	MAG	0.01	0.01	0.02	0	0.02	0	0	0.01	0	-0.0 *	0	0.01	
Charlotte, NC	HIA	0	-0.1 *	0.01	-0.2	-0.1	0	2.78	0	0.02	0.02	0	0	
Cincinnati, OH	DUR	0	0	0	0.01	0	0.01	0.01 *	0.02 *	0	0	0	0.01	
Cincinnati, OH	DUO	0.01	0	-0.0 *	0	-0.0 *	0	-0.0 *	-0.0 *	0	0	0	0	
Cincinnati, OH	DUA	0.01	0	-0.0 *	0	0	0	-0.0 *	0	0	0	0	0	
Cincinnati, OH	MIC	0.11	0.14	0.22	0.17	0.27	0.37	0.43	0.57 *	0.46	0.06	0.19	0	
Cincinnati, OH	MAG	0	0.03	0.01	0	0	0	0	0	0	0	0	0.02	
Cincinnati, OH	HIA	-0.1	0.02	0.24 *	0	0.66	0.02	0	0.06	0	0	0.10 *	0.69	
Columbia, SC	DUR	0	0	0	0.02 *	0.03 *	0.03 *	0.01	0.02 *	0	0	0	0	
Columbia, SC	DUO	0	0	0	0	0	0	0	0	0	0	0	0	
Columbia, SC	DUA	0	0	0	0	0	-0.0 *	-0.0 *	-0.0 *	0	0	0	0	
Columbia, SC	MIC	0.03	0.08	0.16	-0.1	-0.3 *	0.03	0.1	0.05	0	-0.1	0.06	-0.1	
Columbia, SC	MAG	0	0	0	0	0.01	0	0	0	0	0	0.01 *	0.02 *	
Columbia, SC	HIA	0.09 *	0	0	0	-2.6	0	0	0	0.06 *	0.03	0	0	
		jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	

Table 12. Regression generated beta coefficients (β) estimate change over each time step of the record for each indicator. Significant change denoted (*). N ≈ 48 for each iteration.

		jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	
Columbus, OH	DUR	0.01	0	0.01	0	0	0	0.01	0.01	0	0	0	0	Beta Coefficient
Columbus, OH	DUO	0	0	0	0	-0.0 *	0	0	0	-0.0 *	0	0	0	
Columbus, OH	DUA	0	0	0	0	-0.0 *	0	0	0	0	0	0	0	
Columbus, OH	MIC	-0.5	-32	0.45	0.34	0.79	-8.6	0.63	0.98	0.22	1.01	-0.4	-1.4	
Columbus, OH	MAG	0	0	0	0.01	0	0.01	0	0	0	0	0	0.03	
Columbus, OH	HIA	0.01	0.08	0.02	0	0.01	0	0.12	0.03	0	0.03	-0.1	0.07	
Dayton, OH	DUR	0.02 *	0.03 *	0	0	0	0.01	0.02 *	0.02 *	0	0.03 *	0	0.01	
Dayton, OH	DUO	0	0	0	-0.0 *	0	0	-0.0 *	-0.0 *	-0.0 *	0	0	0	
Dayton, OH	DUA	0	0	0	0	0	0	0	0	0	0	0	0	
Dayton, OH	MIC	584	413	535	601	508	531	541	541	527	511	587	414	
Dayton, OH	MAG	0	0.03	0	0.02	0	0	0	0	0	0	0	0	
Dayton, OH	HIA	0	-2.4	0.07	0.01	0.02	0.06 *	0	0.05	0	0.09	0	4.14	
Grand Rapids, MI	DUR	0.04	0	0	0.02	0	0.01	0	0	0	0	0	0	
Grand Rapids, MI	DUO	0	-0.0 *	-0.0 *	0	0	0	0	0	0	0	-0.0 *	0	
Grand Rapids, MI	DUA	0	0	0	-9.7	0	0	0	0	0	0.00 *	0	-0.0 *	
Grand Rapids, MI	MIC	-7.3	1.53	0.56	-0.5	-0.3	-1.6 *	-1.2 *	-1.2 *	-1.5 *	-4.1	0.12	-24	
Grand Rapids, MI	MAG	0	0	0	0	0	0	-6.4	0.01	0	0.02	0	0	
Grand Rapids, MI	HIA	0.17 *	0.83	0.57	0	0	0.01	0	0	0	0.01	0.05	1.44	
Hartford, CT	DUR	0	0	0.03 *	0.01	0	0	0	0	0	0	0	0	
Hartford, CT	DUO	0	0	0	-0.0 *	0	-0.0 *	-0.0 *	-0.0 *	-0.0 *	0	0	0	
Hartford, CT	DUA	0	0	0	0	0.01	0	0.01 *	0	0	0	0	0	
Hartford, CT	MIC	0	0	0.01	0	-0.1 *	-0.1 *	-0.0 *	-0.0 *	-0.1 *	0	-0.0 *	0	
Hartford, CT	MAG	0	0.01	0	0.02 *	0	0.02	0.01	0.01	0	0.01	0.01	0	
Hartford, CT	HIA	0	-0.1	0.07	0	0	0	0.02	0.04	0.01	0	0	0.08	
Indianapolis, IN	DUR	0	0.02	0	0	0	0	0.01	0	0	0	0	0	
Indianapolis, IN	DUO	0	0	0	0	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	0	0	-6.6	
Indianapolis, IN	DUA	0	0	0	0	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	0	0	0	
Indianapolis, IN	MIC	8.89	-21	0.6	1.69 *	1.76 *	2.64 *	1.69 *	1.21 *	0.65	1.57 *	0.75	6.4	
Indianapolis, IN	MAG	0.03	0.01	0	0.03 *	0.01	0	0.01	0	0.01	0.01	0	0.03	
Indianapolis, IN	HIA	0.87	-2.8	0.07	0	0	0.33	-0.0 *	0	0	0	0.02	0	
Knoxville, TN	DUR	0	0.01 *	0	0	0.02 *	0.01	0	0.02 *	0	0	0	0	
Knoxville, TN	DUO	0	0.01 *	0.01 *	0.01 *	0.01 *	0.01 *	0	0.01 *	0.00 *	0	0	0	
Knoxville, TN	DUA	0	0	0	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	0	-0.0 *	-0.0 *	
Knoxville, TN	MIC	0	0.07	0.04	0.05	0.18 *	0.09	0.06	0.01	0.11	0.14	0.07	0	
Knoxville, TN	MAG	0	0	0.01	0	0.03	0	0	0	0.02	0	0	0.03	
Knoxville, TN	HIA	0	0	0.02	0.01	0	0.03	0	0.03	0	0.09 *	0.02	-0.1	
Louisville, KY	DUR	0	0	0	0.01	0	0.01 *	0.01	0	0	0	0	0.01	
Louisville, KY	DUO	0	0	0	0	0	-0.0 *	0	0	0	0	0	0	
Louisville, KY	DUA	0	0	0	0	0	-0.0 *	-0.0 *	-0.0 *	0	-0.0 *	0	0	
Louisville, KY	MIC	0.08	-0.1	0.19	0.33	0.03	-0.1	0.17	0.41	-0.1	0	-1	-0.2	
Louisville, KY	MAG	0	0.06 *	0	0.01	0	0.01	0	0.01	0	0	0.02	0.04	
Louisville, KY	HIA	0.04	-0.3	0.12 *	0	0.01	0.03	0.49	0.02	0.01	0	0.02	0.04	
Memphis, TN	DUR	0	0	0.01	0	0	0.02 *	0	0	0	0	0	0	
Memphis, TN	DUO	0	0	0	0	0	-0.0 *	0	-0.0 *	-0.0 *	0	-0.0 *	-0.0 *	
Memphis, TN	DUA	0	0	0	-0.0 *	-0.0 *	-0.0 *	0	-0.0 *	-0.0 *	-0.0 *	-0.0 *	-0.0 *	
Memphis, TN	MIC	-1.9	-0.4	0	-0.4	-0.7	0.77	-0.5	0.04	-0.7	-1	0.55	-0.4	
Memphis, TN	MAG	0	0	0	0	0	0.04	0.01	0	0	0.04 *	0	0.01	
Memphis, TN	HIA	0.02	0.01	0	0	0.01	0.03	0	0	0	0.07	0	0	
Minneapolis, MN	DUR	0.02	0	0	0.01	0	0	0	0	0	0	0	0.01	
Minneapolis, MN	DUO	0	0	0	0	0	0	0	0	0	0	0	0	
Minneapolis, MN	DUA	0	0	0	0	-6	0	0	0.00 *	-0.0 *	0	0	0.01	
Minneapolis, MN	MIC	0.01	0	-0.2	-0.1	0	0.26 *	0	0.18	0.08	0	0.03	0	
Minneapolis, MN	MAG	0	-0.0 *	0.01	0	0	0	0	0	0	0	0.01	0	
Minneapolis, MN	HIA	-9.2	1.46	-0.4	-0.5	-0.4	0.01	0	0.02 *	0	0.18 *	20.1	0	
		jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	

Table 13. Regression generated beta coefficients (β) estimate change over each time step of the record for each indicator. Significant change denoted (*). N ≈ 48 for each

		jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	Beta Coefficient
Nashville, TN	DUR	0	0	0	0	0.01	0.01	0	0	0	0	0	0	
Nashville, TN	DUO	0	0	0	0	0	0	0	0	0	-0.0 *	0	0	
Nashville, TN	DUA	0	0	0	0	0	0	0	0	0	-0.0 *	0	0	
Nashville, TN	MIC	0	0	0.05	0.05	0.03	0.2	0.03	0.03	-0.2	0	0.07	0.13	
Nashville, TN	MAG	0	0.02	0	0.01	0.03 *	0.01	0	0.01	0	0	0	0.02	
Nashville, TN	HIA	0	0	0.02	0.32	1.45	0.01	-0.1	0.04	0	0.02	0.03	0	
Philadelphia, PA	DUR	0	0	0.01 *	0	0	0	0	0	0	0	0	0	
Philadelphia, PA	DUO	0	0	0	0	0	0	0	0	0	0	0	0	
Philadelphia, PA	DUA	0	0	0	0	0	0	0	0	0	0	-3.9	0	
Philadelphia, PA	MIC	0	0	0	0	0	0	0	0	0	0.03	0	0	
Philadelphia, PA	MAG	0.02	0	0	0	0	0	0	0	-0.0 *	0	0	0	
Philadelphia, PA	HIA	0.06	-9.4	0.06	0.02	-0.1	0.01	0.01	0.16	-0.1	0	0.02	0.26	
Pittsburgh, PA	DUR	0	0.01	0	0	0.01	0.01	0.01 *	0.01 *	0	0	0	0.02	
Pittsburgh, PA	DUO	0	0	0	0	0	0	0.00 *	0	0	0.00 *	0	0	
Pittsburgh, PA	DUA	0	0	0	0	-0.0 *	0	0	0	0	0	0	0	
Pittsburgh, PA	MIC	0	0.01	0	0	0.05	0.01	0.03	0.03	0	0.08	0.07	0	
Pittsburgh, PA	MAG	0.02	0	0	0.02	0	0.03	0	0	0	-0.0 *	0	0	
Pittsburgh, PA	HIA	-2.7	-1	0.99	0	0	0	0.01	0.14	0	0.05	0.73	-2.8	
Raleigh, NC	DUR	0	0	0.01 *	0	0.01 *	0	0.02 *	0.01	0	0.01	0	0	
Raleigh, NC	DUO	0	0	0.00 *	0.00 *	0	0	7.14	0	0	0	0	0	
Raleigh, NC	DUA	0	0	0	0	0	0	-0.0 *	0	0	-0.0 *	-0.0 *	0	
Raleigh, NC	MIC	0.22	0.25	0.23	0.02	0	0.08	0	0.03	0	0.07	0.37	0	
Raleigh, NC	MAG	0	0	0	0.03	0.03	0	0	0.01	0	0	0	0.01	
Raleigh, NC	HIA	0.03	0	0.03	0	0.02	0.01	0	0.21	0	0.07	0	0	
Richmond, VA	DUR	0	0	0	0.01	0.02 *	0.02 *	0.01 *	0.01	0	0	0	0	
Richmond, VA	DUO	0	0	0	0	0	0	0	0	-8.9	0	0	0	
Richmond, VA	DUA	0	0	0	0	0	0	0	0	0	0	0	0	
Richmond, VA	MIC	0	0	0	0	0.06	0.04	-0.2	0.06	0.1	0.19	0	-0.1	
Richmond, VA	MAG	0	0.01	0	0.03	0.03	0	0	0.03	0	0	0	0	
Richmond, VA	HIA	-4.5	0.01	0	0.04	-2.6	0.02	0.04	0	0	0	0.01	0	
St. Louis, MO	DUR	0.02	0	0	0.01	0	0	0.01	0	0	0	0	0.01	
St. Louis, MO	DUO	0	0	0	0	0	-0.0 *	0	0	-0.0 *	0	0	0	
St. Louis, MO	DUA	0	0	0	0	0	0	0.00 *	0	0	0	0	0	
St. Louis, MO	MIC	0.15	0.14	-0.7	-0.6	-1.1 *	-1.7 *	-0.5	-0.7 *	-0.4	-0.8	0.39	-1.7	
St. Louis, MO	MAG	0	0.05	0.03	0	0	-0.0 *	0.02	0	0	0.03	0.02	0.03	
St. Louis, MO	HIA	4.02	0.24	0.07	0	-1.8	0	0.01	0.08	0.01	0	7.56	0	
Syracuse, NY	DUR	0	0.01	0	0	0	0	0	0	0	0	0	0	
Syracuse, NY	DUO	0.02 *	0	0	0	0	0	0	0	0	0	0	0	
Syracuse, NY	DUA	0.02 *	0	0.01	0	0	0	0	0	0	0.01	0	0	
Syracuse, NY	MIC	-0.1	0.2	-0.4	-0.9	-0.6	-0.9 *	-1.1 *	-0.9 *	-0.5	0.42	-0.8 *	-2.2	
Syracuse, NY	MAG	0	0.01	0	0	0.01	0.01	0.01	0	0	-0.0 *	0	0.01	
Syracuse, NY	HIA	1.18	2.08	4.16	0.04	0.03	0	0	0.02	0	0	-1.8	-4.1	
Washington, DC	DUR	0	0	0	0	0.02 *	0.01	0	0	0	0	0	0	
Washington, DC	DUO	0	0	0	0	-0.0 *	0	0	0	0	-0.0 *	0	0	
Washington, DC	DUA	0	0	0	0	0	0	0	0	0	0	0	0	
Washington, DC	MIC	0	0	0.02	0	0.02	0.09	0.04	0.04	0.01	0	0	0	
Washington, DC	MAG	0	0	0	0.03	0.02	0.01	0.01	0	0	0	0	0	
Washington, DC	HIA	0.10 *	0	0.01	0	-0.3	0	0	0	0	0.02	0.02	0.01	
		jan	feb	mar	april	may	jun	july	aug	sept	oct	nov	dec	

Table 14. Regression generated beta coefficients (β) estimate change over each time step of the record for each indicator. Significant change denoted (*). N ≈ 48 for each iteration.

In order to see the specific areas of change, I simply apply the linear model to each of the pixels in Louisville, for example, against time (same analysis perform for H3) to see how each individual has changed from 2002 to 2012. Figure 13 indicates the distribution of calculated p-values for the month of March (2002 to 2012), where values less than 0.05 indicate significant change. It is evident that the areas most impacted by the conversion of agriculture land to some other class is where the majority of the significant pixels lie. Figure 10 shows the distribution of the corresponding beta values and indicate the magnitude of change experienced during 2002 to 2012, in some places as much as 0.2 per time step. While all pixels in the Louisville area are calculated to have some positive change in LSTs, the areas with the largest changes in land cover are the most likely to be of any significance, particularly the areas in the northeast section of the Louisville urban boundary where high population growth has driven that change in land cover, presumably.

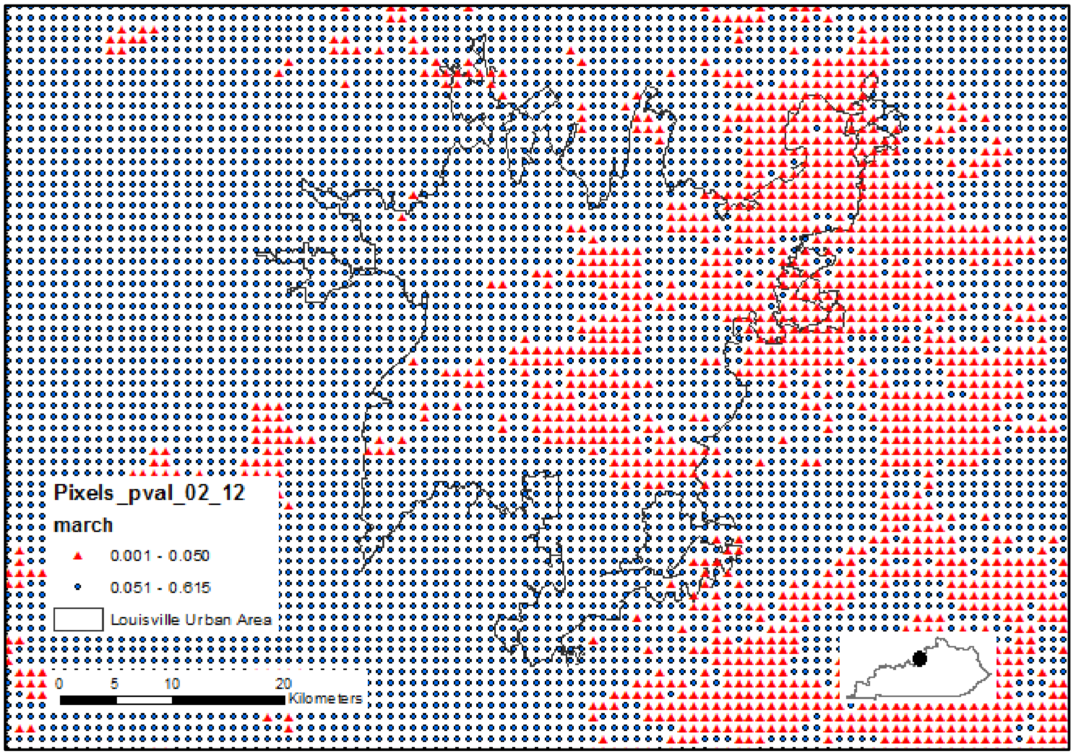


Figure 11. Pixel based analysis of LST changes for the Louisville, Kentucky area. P-values of 0.05 or lower indicate significant change over time (March 2002 to 2012).

CHAPTER 4. DISCUSSION

I have examined the seasonal distribution of LSTs for *urban* and *rural* areas and for classified land cover types *urban, agriculture, forest, and other* to derive estimates of the SUHI *DUA, DUO, DUR, MAG, MIC, and HIA* for 26 U.S. cities with populations greater than 750k. Land cover portions were also summarized for each location on an annualized basis from 2002 to 2012. This study has answered all of the stated hypothesis to the degree that I can generalize the results and understand what they mean in terms of better understanding *spatial* and *temporal* variation in the distribution of LSTs and SUHIs in this region. All locations were ranked according to the observed SUHI on a monthly basis for the year 2012, those rankings were subjected to a Spearman's Rho Ranking Correlation analysis to quantify the degree to which the indicators agreed (H1) and the whether or not those rankings remained consistent throughout the year (H2). SUHI values were calculated for each location (2002-2012) and those temporal records were subjected to simple linear regression against time to check for significant change over time (H3). Here I discuss how the various results further an understanding of the overarching research question involving the spatial and temporal distribution of the SUHI.

Comparison of the Different SUHI Approaches 2012

All classes of measurements show the same general annual variation for 2012, though the timing of the maximum values varies according to the land covers involved in the calculation. Maximum for most LST and SUHI values were observed in July and the minimum occurred in January coinciding with the seasonal nature of Earth-Sun geometry confirming at least that these indicators are responding to broad changes in the physical environment in a consistent manner. Overall, the Spearman's Rho test suggest that not all of the SUHI indicators are generating the same rankings each time, in line with previous studies comparing the similarities between various indicators (Schwarz, Lautenbach, and Seppelt. 2011; Steutker 2002), however, each indicator is genuinely describing some aspect of the SUHI that will affected by various underlying factors. Comparisons between locations using different time steps would not be appropriate, especially considering the examples of cities holding positions at both the top and bottom positions simultaneously. This study area is vast and so has an extensive north-south component and so too a significant timing differential between when the seasonal vegetation will begin and end. based on this difference in latitude. Previous studies have noted changes in the variance of LSTs and the direct link to latitude dependent seasonality (Imhoff et al. 2011; Jin, Dickinson, and Zhang 2005).

Interestingly, all indicators share a pattern where the classified agricultural land cover is consistently warmer than the other land cover types, inconsistent with expectations based on previous results (i.e. Jin, Dickinson, and Zhang 2005). One possible reason for this could be the timing of the LSTs measurements of the day-time MOD11A2 product, which occur before the peak warming of the diurnal cycle and thereby are cooler than the maximum potential at the time of the measurement. On a global basis, agricultural land cover has a lower maximum temperature than urban area Jin, Dickinson, and Zhang 2005), but an induced warming effect of 0.5ºC was detected when considering the impact of increased use of irrigation on daily minimum temperatures in eastern China (Shi, Tao, and Liu 2013). The irrigated farmland has a higher minimum surface temperature (i.e. nighttime temperature) due to its specific capacity to absorb energy during the day and therefore responds faster to increases in daily insolation, relative to the classified urban land areas. This study was unable to distinguish between the various land use activities taking place within the classified agricultural areas, such as irrigated versus rain fed farming or the small urban developments that may have occurred. We know too, from the correlation ranking results, that the rankings derived from land cover based indicators DUA and DUO are consistent during the year and that in Louisville, KY, for example, but the difference in urban and agriculture cover is getting incrementally smaller over time. This would further suggest that we are missing activities that are taking place within the classified land cover. Here, the small increases in residential developments (urbanization) that are in fact lowering the capacity of the land cover to absorb and retain heat, in line with previous indications of the potential for urbanization in former farmland to reduce maximum temperatures (Shi, Tao, and Liu 2013).

Recall that almost no urban areas experienced an increase in urban areas, yet each cities would have accommodated some increase in population at the expense of some preexisting land cover, presumably agricultural land. The mixing of the agricultural land cover with the new developments should serve to lower temperatures and is supported by the decrease in the difference between classified urban and agricultural land cover. At the same time, the area based DUR indicator is mostly stable over time for most locations, suggesting the processes are at work in both the urban and surrounding areas, and therefore, the net change in area-based measurements is minimal. Essentially, the difference in surface temperatures between land cover types (DUA, DUO) is the same in all places (~ 1ºC), but the aggregate SUHI indicators (DUR) for each location is more a factor of the amount of each of the land cover present in each location and so has a larger variance than would be suggested otherwise.

The HIA and MIC indicators have very different distributions compared to the other indicators and so offer perhaps the most informative perspective of the SUHI that indicates how much of the total land areas of each location is actually impacted by extreme LSTs. The fact that these two indicators lack a seasonal component similar would seem to indicate that the areas impacted by extreme heat is fixed for a given area, and this is partially supported by the almost perfect monthly correlations of the MIC rankings for 2012, though the HIA monthly ranking are not as stable. These results are consistent with previous findings derived from continental Europe that compared these same indicators for the months of June and January (Schwarz, Lautenbach, and Seppelt. 2011).

Spatial Distribution of Surface Urban Heat Island 2012

The rankings resulted in a mixed bag, at times producing simultaneous high and low rankings for a single locations including Albany, NY in the summer months (as the most extreme example) and other cities like Syracuse, NY, Baton Rouge, LA, and Louisville, KY periodically holding positions at both ends of the spectrum. The correlation data generated for H1 highlights the moderate degree of agreeability among the *land cover-*

derived indicators, though H2 shows these indicators to be very consistently ranked during the year, except for the winter months (where potential differences are at a minimum). The *urban heat island driven indicators,* shown by H1 to do not agree among each other, yet are highly correlated with other individual indicators and complement our understanding of the SUHI, including HIA to DUR (relationship between urban cover percent and increased LSTs) and the MIC to MAG (*inverse* relationship of forested cover to increased LSTs). The urban heat island driven indicators are shown by H2 to be considerably less consistent throughout the year, though the summer months are relatively consistent and presumably follow the trajectory of the growing season.

As an example of the growing season based timing differential, the cities of Baton Rouge, LA is among the highest ranked location for the DUR indicator and the city of Minneapolis, MN is ranked near last, until the positions flip later in the growing season, presumably once the vegetation in MN has had a chance to grow.

Surface Urban Heat Island Change over Time 2002 to 2015

Evaluation of H3 also gives mixed results where many locations have SUHI values that are actually decreasing over time; some locations even have simultaneous positive and negative growths depending on the indicator considered. Many locations show significant increases over time, especially in the spring and summer, include Philadelphia, PA, Columbus, SC, and St. Louis, MO of up to 0.1, 0.2, and 0.2ºC/year, respectively. Perhaps more importantly, many other locations such as Baltimore, MD, and Cincinnati, OH are indicating significant summer increases in the portion of total land area impacted by high LSTs of anywhere from 0.02 to 0.08 % per year, while other SUHIs for those locations may indicate little or even negative growth during that same time. Interestingly, Louisville, KY has shown a significant increase in the overall LST difference between the urban areas relative to the surrounding 20km buffer summer months of up to 0.01ºC/year (June) while the LST difference between urban and agricultural land has been decreasing by 0.01ºC/year during a co-incident period. Furthermore, this location is actually experiencing a decrease for total land area impacted by extreme heat (greater than 1 standard deviation above the local mean) in March, the start of the growing season.

While this current study did not attempt to address the individual factors contributing to changes in LSTs, a simple examination of the landscape change trajectories in the study locations can clarify why we may be witnessing sometimes-contradictory trends across the various indicators. Recall from Figure 7 that agricultural lands demarcated by the MODIS land cover product have the highest mean temperatures for most of the year. Compared to 2012, nearly all locations had considerably more agricultural land areas in 2002 within the urban boundary areas. A significant decrease in the amount of agricultural land, as per the mean land cover temperatures, would necessarily result in some decrease in the aggregate LST, therefore, a drop in the agriculture land within the urban boundaries lowers LSTs overall. However, a larger coincident decrease in agriculture land for the corresponding rural area would then necessarily experience an even larger decrease in LST and result in a comparatively smaller relative difference between the urban and rural land areas.

Limitations in Current Study

While the remote sensing approach, as detailed above, has proven especially useful for quantifying and integrating spatially explicit distributions of environmental parameters (e.g. LST and land cover type), there are other important limitations to consider for this study. First, this study relies on the use of SUHI indicators that emphasize different features and may not accurately portray all of the contributing elements important to the formation of the SUHI (Stewart 2011). This over simplification of SUHI formation is necessary due to the complexity and effort involved in assembling a comprehensive dataset that mimics the spatial scale of the individual processes contributing to SUHI formation (Mirzaei and Haghighat 2010). Simplification creates additional uncertainty because influences at differing scales (i.e. regional atmospheric circulation versus urban geometry) are not included, so any emerging results offer an incomplete picture at best. This study emphasizes a range of indicators based on urban area boundaries and classified land cover because land cover, both proxy measures of the anthropogenic influences on the physical environment, to development more comprehensive understanding of the SUHI at a regional spatial scale (Stone 2012). The use of multiple types of SUHI indicators compensates for the specific focus of the individual indicators (Schwarz, Lautenbach, and Seppelt 2011) and yields a more diverse perspective than could be obtained otherwise.

The primary limitation of the MODIS land cover data in this study is the large surface area aggregated for each pixel (500 meters) that obscures the small-scale features within this complex area limiting the degree of detail represented in the dataset. Whether or not a small (but significant) patch of forested land cover is represented in the land cover data depends on the exact alignment of the product pixel boundaries and the underlying land cover orientation. The individual patches are sometimes bisected and therefore less likely to meet the surface area threshold (majority) required to label that pixel as forest.

The significant impact or this current study is that many features and land cover patches (such as newly planted tree clusters) will be underrepresented in the final product (Friedl et al. 2010) and possible not detected at all in the coarse land cover product. Land cover products with better spatial resolution exist, such the National Land Cover Dataset (30 meters), but the temporal resolution is every 5 years and so fails then to capture the year to year variation in land cover distribution. This study compromises high spatial resolution of land cover types for temporal consistency due to the short interval of time we have to work with (2002 to 2012).

Another issue limiting the potential usefulness of the research is that the imagery collected was specific to the daytime overpass. Many previous studies focus on the daytime due to the greatest potential for differences between representative urban and non-urban sites (Oke and Voogt 2013; Schwarz, Lautenbach, and Seppelt. 2011), yet the greatest potential risk to human health due to local elevated temperatures is at night, when elevated temperatures prevent needed cooling (Altman 2012). The overall magnitude of the SUHI is generally smaller at night, especially for temperate mid latitude locations (Imhoff et al. 2010), but the impact to public health of seemingly small increase is potentially large for urban residents (Stone 2012). Future efforts for this project will focus on integrating nighttime measurements for a more complete understanding of the dynamics of the SUHI.

This current study has illustrated several approaches to quantifying the status and trajectory of the SUHI for many of the largest U.S. cities and reported seemingly contradictory outcomes, highlighting the importance of properly constructing the research objectives to align with the information provided by the selected indicator. For example, recall that Albany, NY experienced an increase in forested land cover of 132% from 2002 and 2012, yet observed a co-incident increase in temperature difference between the urban and surrounding rural areas. The impact of the canopy increases on the change in temperature difference between urban and rural is unclear, but trending up while the micro-island indicator clearly shows that the total area impacted by high temperatures decreased significantly during this time. The perspective regarding how well the increased forest area has helped to mitigate increased temperatures could be interpreted in multiple ways depending upon the perspective of the question, the difference between magnitudes of SUHI as oppose to extent of, for example. This finding reinforces earlier findings (Schwarz et al 2011) and further highlights the importance of proper selection of one or more indicators for SUHI comparisons in alignment with research objectives.

CHAPTER 5. CONCLUSION

Within the context of the research question posed and the results generated from the evaluation of each hypothesis, three main conclusions are drawn from this study: 1) that choice of indicator and 2) timing of observation affects the outcomes (sometimes), and 3) the amount of change detected for any place is dependent on the choice of indicator.

The choice of indicator can substantially influence the SUHI effect observed for a given location. When ranked according to value magnitude, some locations occupied conflicting positions within those rankings, holding both high and low positions simultaneously. This behavior is witnessed across multiple locations and has been reported previously in similar regional based studies (Schwarz et al. 2012: Schwarz, Lautenbach, and Seppelt. 2011). The occurrence of that behavior here too suggest that the rankings based on any one kind of indicator (i.e. Stone 2012; Kenward et al. 2014) should be considered as only a partial explanation the overall dynamics of the urban heat island effect.

Just as important, in terms of consistent measuring of the SUHI, the timing of the measurement matters more for some indicators than it does for others. In the case of the *land cover driven SUHIs*, the relative LST difference between particular land cover types remains consistent throughout the year for a given location and therefore timing is less important for capturing differences using the SUHI rankings. The *urban heat island* *driven indicators*, on the other hand, show considerably less correlation over time and tend to produce varying rankings for each new time step, and so are likely to give different results each observation. The temperature mixed forested eco-region is spatially and temporally complex, mostly due to the seasonal nature of the natural landscape (leaf on/off) and the differences in timing of these cycles across the multiple study locations and so should be treated using a range of approaches. The first two parts of the research question posed here involve the usability of the indicators for fair and consistent measuring LST temperature differences within a given ecoregion and over time. The results generated here inform the use of these indicators by showing which indicators are correlated, and so provide the same information, and those indicators that are not correlated, and so provide additional, or even complementary, information about each study location.

Several locations from the subset of US cities in the study have some form of statistically significant change in SUHI over time, though the direction of change was not always consistent and could be misleading. For example, urban Louisville, KY seemingly is experiencing a significant increase in temperatures over time (2000 to 2015) relative to the surrounding areas (DUR), however, another indicator (DUA) instead indicates a relative "cooling" of the surrounding lands due to a decrease in agricultural land, which we know from above to be consistently the warmest land cover type. This example further highlights the necessity of providing for a range of indications when quantifying and comparing values across locations that are essentially unique individuals.

Finally, this project has been successful in terms of generating results that help us understand the kind of information we can derive from the use of surface urban heat island indicators for monitoring of change over time. The primary contribution involves findings that highlight how the indicators can behave in less than intuitive ways and produce misleading outcomes if extreme care is not taken in the selection proper indicators. In terms of future assessments of the effectiveness of ongoing mitigation efforts at combating the impacts of the SUHI. Multiple indicators provide validation and a better overall understanding of the forces at work in the urban heat island. The apparent conflict in findings clearly highlighted for Louisville, KY to illustrate one of the biggest challenges to understanding regional change over time, particularly the misinterpretation of results. Louisville is seemingly experiencing an increase in the overall difference between urban and non-urban areas, but the overall land area impacted by warmer temperatures has actually been decreasing over time. It is clear that relying on a single approach to conceptualizing the SUHI is problematic and will not provide sufficient understanding of the real world conditions. Decision makers and community stakeholders could potentially benefit from the findings generated here by better understanding the important of considering multiple persoective when assessing issues of the public benefit.

REFERENCES

Alberti, M. 2005. The effects of urban patterns on ecosystem function. *International Regional Science Review,* 28**,** 168-192.

Altman, P. 2012. Killer Summer Heat: Projected Death Toll from Rising Temperatures in America Due to Climate Change. Natural Resources Defense Council Issue Brief. May 2012 IB:12-05-C. https://www.nrdc.org/sites/default/files/killer-summer-heat-report.pdf

Aniello, C., K. Morgan, A. Busbey and L. Newland 1995. Mapping micro-urban heat islands using Landsat TM and GIS. – Computers and Geosciences 21 (8): 965-969.

Buyantuyev, A., and J. Wu. 2010. Urban heat islands and landscape heterogeneity: linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. *Landscape Ecology,* 25**,** 17-33.

Carreiro, M. M., and C. Tripler. 2005. Forest Remnants Along Urban-Rural Gradients: Examining Their Potential for Global Change Research. *Ecosystems,* 8**,** 568-582.

Chen, X., H. Zhao, P. Li, and Z. Yin. 2006. Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote Sensing of Environment* 104.

Chen, D., Wang, X., and Zhu, W. (2014). WRF simulations of the urban heat island under hot-weather synoptic conditions: the case of Hangzhou City, China. *Atmospheric Research*, 138,364-377.

Clinton, N., and P. Gong. 2013. MODIS detected surface urban heat islands and sinks: Global locations and controls. *Remote Sensing of Environment* 134**,** 294-304.

Ernstson, H., S.E. van der Leeuw, C.L. Redman, D. Meffert, G. Davis, C. Alfsen, and T. Elmqvist. 2010. Urban Transitions: On Urban Resilience and Human-Dominated Ecosystems. In *AMBIO*. Royal Swedish Academy of Science.

Friedl, M. A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., andHuang, X. 2010. MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment, 114, 168–182

Gordon, J. H. 1921. Temperature survey of the Salt River Valley, Arizona. *Monthly Weather Review,* 49**,** 271-274.

Hansen, J., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl . 2001. A closer look at United States and global surface temperature change, Journal of. Geophysical. Review., 106(20): 23947–23963, doi:10.1029/2001JD000354.

Howard L. 1833. The climate of London: deduced from meteorological observations made in the metropolis and at various places around it. Harvey and Darton: London.

Hu L., N.A. Brunsell. 2013 The impact of temporal aggregation of land surface temperature data for surface urban heat island (SUHI) monitoring, Remote Sensing of Environment, Vol.134:162-174. doi.org/10.1016/j.rse.2013.02.022.

Hung, T., D. Uchihama, S. Ochi, Y. Yasuoka. 2006. Assessment with satellite data of the urban heat island effects in Asian mega cities. *International Journal of Applied Earth Observation and Geoinformation,* 8**,** 34-48.

Imhoff, M., P. Zhang, R. Wolfe, and L. Bounoua. 2010. Remote sensing of the urban heat island effect across biomes in the continental USA. *Remorte Sensing of Environment,* 114**,** 504-513.

Jin, M., R.E. Dickinson, and D. Zhang. 2005. The Footprint of Urban Areas on Global Climate as Characterized by MODIS. *Journal of Climate,* 18**,** 1551-1565.

Kenward, A. D. Yawitz, T. Sanford, and R. Wang. 2014. Summer in the City: Hot and Getting Hotter. Princeton, NJ 08542: Climate Central.

Kopec, R. 1970. Further observations of the urban heat island in a small city. *Bulletin American Meteorological Society,* 51**,** 602-606.

Li, Y., H. Zhang, and W. Kainz. 2012. Monitoring patterns of urban heat islands of the fast-growing Shanghai metropolis, China: Using time-series of Landsat TM/ETM+ data. *International Journal of Applied Earth Observation and Geoinformation,* 19**,** 127-138.

Middel, A., A. Brazel, P. Gober, S.W. Myint, H. Chang, and J. Duh. 2012. Land cover, climate, and the summer surface energy balance in Phoenix, AZ, and Portland, OR. *International Journal of Climatology,* 32**,** 2020-2032.

Mirzaei, P., and F. Haghighat. 2010. Approaches to study Urban Heat Island e Abilities and limitations. *Building and Environment,* 45**,** 2192-2201.

Mirzaei, P. 2015. Recent challenges in modeling of urban heat island, Sustainable Cities and Society, Vol 19: 200-206.

Nichol, J., and M. Wong. 2009. High Resolution Remote Sensing of Densely Urbanised Regions: a Case Study of Hong Kong. *Sensors,* 9**,** 4695-4708.

Oke, T. R. 1982. The Energetic Basis of the Urban Heat Island. *Quaterly Journal of the Royal Meteorological Society,* 108**,** 1-24.

Oke, T. R.. 1997. Urban climates and global environmental change. 273-287. In R.D. Thompson and A. Perry (ed.) Applied climatology: Principals and practice. Routledge, London, U.K.

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., Kassem, K. R. 2001. Terrestrial ecoregions of the world: a new map of life on Earth. Bioscience 51(11):933-938.

Rajasekar U, and Q. Weng. 2008. Urban heat island monitoring and analysis using a non-parametric model: A case study of Indianapolis. ISPRS Journal of Photogrammetry and Remote Sensing 64: 86–96, DOI:10.1016/j.isprsjprs.2008.05.002

Ratcliffe, M., C. Burd, K. Holder, and A. Fields, "Defining Rural at the U.S. Census Bureau," ACSGEO-1, U.S. Census Bureau, Washington, DC, 2016.

Schwarz, L., S., Lautenbach, and S. Seppelt. 2011. Exploring indicators for quantifying surface urban heat islands of European cities with MODIS land surface temperatures. *Remore Sensing of Environment,* 115**,** 3175-3186.

Schwarz, N., U. Schlink, U. Franck, and K. Großmann. 2012. Relationship of land surface and air temperatures and its implications for quantifying urban heat island indicators—An application for the city of Leipzig (Germany). *Ecological Indicators,* 18**,** 693-704.

Shi, W., Tao, F. and Liu, J. (2014), Regional temperature change over the Huang-Huai-Hai Plain of China: the roles of irrigation versus urbanization. Int. J. Climatol., 34: 1181–1195. doi:10.1002/joc.3755

Stewart, I, 2011. A systematic review and scientific critique of methodology in modern urban heat island literature. International Journal of Climatology., 31, 200–217.

Stone, B. 2007. Short Communication: Urban and rural temperature trends in proximity to large US cities: 1951-2000. *International Journal of Climatology,* 27**,** 1801-1807.

Stone, B. 2008. Urban Heat and Air Pollution: An Emerging Role for Planners in the Climate Change Debate. J*ournal of the American Planning Association,* 71**,** 13-25.

Stone, B. 2012. *City and the Coming Climate Change*. Cambridge University Press.

Stone, B., J. Hess, and H. Frumkin. 2010. Urban form and extreme heat events: Are sprawling cities more vulnerable to climate change than compact cities? *Environment Health Prospective,* 118**,** 1425-1428.

Stone B, Vargo J, Liu P, Habeeb D, DeLucia A. 2014. Avoided Heat-Related Mortality through Climate Adaptation Strategies in Three US Cities. *PLOS ONE* 9(6): e100852. https://doi.org/10.1371/journal.pone.0100852

Tan, J., Y. Zheng, X. Tang, C. Guo, L. Li, G. Song, X. Zhen, D. Yuan, A.J. Kalkstein, F. Li, and H. Chen. 2010. The urban heat island and its impact on heat waves and human health in Shanghai. *International Journal of Biometeorology,* 54**,** 75-84.

Tomlinson, C.J., L. Chapman, J.E. Thornes, C.J. Baker. 2012. Derivation of Birmingham's summer surface urban heat island from MODIS satellite images. *International of Climatology,* 32**,** 214-224.

United Nations, Department of Public Information, The 2015 Revision of World Population Prospects. SC/11018, 29 July 2015, https://esa.un.org/unpd/wpp/

Voogt, J.A. and T.R. Oke. 2003. Thermal remote sensing of urban climates. *Remote Sensing of Environment,* 86**,** 370-384.

Wan, Z.: New refinements and validation of the MODIS land surface temperature/emissivity products, Remote Sensing of the Environment, 112, 59–74, doi:10.1016/j.rse.2006.06.026, 2008

Wan, Z. and Dozier, J.: A generalized split-window algorithm for retrieving land-surface temperature from space, IEEE Geosciences and Remote Sensing, 34, 892–905, doi:10.1109/36.508406, 1996

Westendorf, K., C. Leuhart, and D. Howarth. 1989. A preliminary assessment of the Louisville urban heat island. *Transactions of the Kentucky Academy of Science,* 50**,** 86-93.

Yuan, F., and M. Bauer. 2007. Comparison of impervious surface area and normalized difference vegetation index as indicators of surface urban heat island effects in Landsat imagery. *Remote Sensing of Environment,* 106**,** 375-386.

Zhang, X., P. Wu, and B. Chen. 2010. Relationship between vegetation greenness and urban heat island effect in Beijing City of China. *Procedia Environmental Sciences,* 2**,** 1438-1450.

Zhang, J., and Y. Wang. 2008. Study of the relationships between the spatial extent of surface urban heat islands and urban characteristic factors based on Landsat ETM plus data. *Sensors,* 8**,** 7453-7468.

Zheng, B., S. Myint, and C. Fan. 2014. Spatial configuration of anthropogenic land cover impacts on urban warming. *Landscape and Urban Planning,* 130**,** 104-111.

Zhou, W., G. Huang, M. Cadenasso. 2011. Does spatial configuration matter? Understanding the effects of land cover pattern on land surface temperature in urban landscapes. *Landscape and Urban Planning,* 102**,** 54-63.

Zhou, J., J. Li, and J. Yue. 2010. Analysis of urban heat island (UHI) in the Bejing metropolitan area by time-series MODIS data. *IEEE International Geoscience and Remote Sensing Symposium* 3327-3330.

APPENDIX ITEMS

A1. Code implemented in Python and ArcPy to sample MODIS data

'20/2017 www.planetb.ca/projects/syntaxHighlighter/popup.php

```python
# Jeremy Sandifer (University of Louisville)
# Spring 2017 UHI Analysis
# ---------------------------------------------------------------------------
# Spatial.py
# Created on: 2016-12- 29 20:36:15

# Description: The steps needed to generate each UHI indicator have
# scripted out to ensure reproducibility.  The user has to supply
# three (3) pcs of data:  Study cities, landcover data, and LST data.
# ---------------------------------------------------------------------------

# Import arcpy module
#import os, sys
import arcpy


# Check out any necessary licenses
arcpy.CheckOutExtension("spatial")
arcpy.CheckOutExtension("GeoStats")

# Set Geoprocessing environments
arcpy.env.workspace = "C:\\GISdata\\NewUHI\\Data\\MOD_11a2"
arcpy.env.overwriteOutput = "TRUE"

import time

timer = time.time()
timer1 = time.localtime(timer)
timer2= time.asctime(timer1)
print(timer2)

# Global variables:
UrbanArea = "C:\\GISdata\\NewUHI\\data\\Boundaries\\UrbanProject.shp"
RuralArea = "C:\\GISdata\\NewUHI\\data\\Boundaries\\Rural20km.shp"
outpath = "C:\\GISdata\\NewUHI\\output\\tables\\"
outfolder = "C:\\GISdata\\NewUHI\\output\\"
#Lou ="C:\\GISdata\\NewUHI\\data\\Lou_sample.shp"
Lou ="C:\\GISdata\\NewUHI\\output\\SampleParcels.shp"
LOUpts = "C:\\GISdata\\NewUHI\\Lou_pts.shp"

datasetList = arcpy.ListRasters("*", "TIF") ## all rasters 2000 to 2016
##[94:722]
################################################## Start Urban
##############################################################
for dataset in datasetList:
    newfieldname =  "urban_" + dataset[8:16] + ".dbf"
    print(newfieldname)
    ### resample LST data to 500m
    lst_tmp_tif = arcpy.Resample_management(dataset, "re_tmp", "500 500", "CUBIC")
    #### Get the mean Urban LST for each UA boundary
    urb_tab = arcpy.gp.ZonalStatisticsAsTable_sa(UrbanArea, "NAME10", lst_tmp_tif, newfieldname, "DA
    ####  Need to Add new field to adjust the precision
    arcpy.AddField_management(urb_tab, "URBAN", "Float", 9, 2, "", "", "", "")
    #### Now Populate the field with the extracted data
    arcpy.CalculateField_management(urb_tab, "URBAN", "((!MEAN! *0.02)-273.15)", "Python_9.3","")
    #### Output to .dbf for later processing
    arcpy.TableToTable_conversion(urb_tab, outpath, newfieldname)
    print("Got it!!")
##############################################################
##################################################  END Urban


################################################## Start Rural
##############################################################
for dataset in datasetList:
    newfieldname = "rural" + dataset[8:16] + ".dbf"
    print(newfieldname)
    ### resample LST data to 500m
```

tp://www.planetb.ca/projects/syntaxHighlighter/popup.php 1/

```
    lst_tmp_tif = arcpy.Resample_management(dataset, "re_tmp", "500 500", "CUBIC")
    #### Get the mean Urban LST for each UA boundary
    rur_tab = arcpy.gp.ZonalStatisticsAsTable_sa(RuralArea, "NAME10", lst_tmp_tif, newfieldname, "DA
    ####  Need to Add new field to adjust the precision
    arcpy.AddField_management(rur_tab, "Rural", "Float", 9, 2, "", "", "", "")
    #### Now Populate the field with the extracted data
    arcpy.CalculateField_management(rur_tab, "Rural", "((!MEAN! *0.02)-273.15)", "Python_9.3","")
    #### Output to .dbf for later processing
    arcpy.TableToTable_conversion(rur_tab, outpath, newfieldname)
    print("Moving on to the next !!")

###########################################################
###############################################  END Rural


############################################### Start Hot Island Area
###########################################################

## Indicator 2 [UHI2]: Hot Island Area (HIA) is the % of total Area within urban boundary with LST g
## than or equal to the mean plus one standard deviation.  Note that the "Area" attribute of the STU
## shapefile uses the square kilometer, while the "Shape_Area" is square meter.

#Local variables:
HIA = outfolder + "HIAtmp.tif"
UA_HIA = outfolder + "UA_HIA"
Study2 = outfolder + "UrbanProject2.shp"
HotIslandname = outfolder + "HotIsland.tif"



for dataset in datasetList:
    newfieldname =  "hia_" + dataset[8:16] + ".dbf"
    #RasterLST = dataset

    print(dataset)
    ## Resample LST data to Cubic Convolution 250 meter
    #lst_tmp_tif = arcpy.Resample_management(dataset, "re_tmp", "500 500", "CUBIC")
    print("resampled")

    ## Calculate zone Mean and STD, send to raster
    ZonalMean = arcpy.sa.ZonalStatistics(UrbanArea,"NAME10",dataset, "MEAN", "DATA")
    a=ZonalMean.save(outfolder + "ZonalMean.tif") ## must be saved to disk
    print("got mean saved")

    ZonalSTD = arcpy.sa.ZonalStatistics(UrbanArea,"NAME10",dataset, "STD", "DATA")
    b=ZonalSTD.save(outfolder + "ZonalSTD.tif") ## must be saved to disk
    print("got STD saved")

    ## Raster calculations (its ugly, but its what I was able to make work!)
    f1 = arcpy.sa.Raster(dataset)
    f2 = arcpy.sa.Raster(outfolder + "ZonalMean.tif")
    f3 = arcpy.sa.Raster(outfolder + "ZonalSTD.tif")
    print(" saved rasters reloaded as variables")

    outraster = (f1 - f2 - f3)
    c = outraster.save(outfolder + "HIA_tmp.tif") ## must be saved to disk
    f4 = arcpy.sa.Raster(outfolder + "HIA_tmp.tif")
    print("executed raster calculation")

    ## clean up a few things that were causing errors..
    del(f1)
    del(f2)
    del(f3)

    ##   HotIsland Mark is the Annual Mean LST - The Mean - The STD.  This raster is reclassed to sh
    ##   only positive values as 1 and all others as NODATA.  The SUM of these is
    ##   now the amount of area within th urban areas with this condition.
    arcpy.gp.Reclassify_sa(f4, "Value", "-1000 0 NODATA;0 1000 1", HotIslandname, "DATA")  ## may ne
    del(f4)
```

```
    print ("reclassified the LST")

    arcpy.gp.ZonalStatisticsAsTable_sa(UrbanArea, "NAME10",HotIslandname, UA_HIA, "DATA", "SUM")
    print ("Zonal stats on the LST reclass")

    ## Create a copy of the Study Area to join and allow for calculationinvolving the original AREA
    arcpy.CopyFeatures_management(UrbanArea, Study2 ,"", "", "", "")
    arcpy.JoinField_management(Study2, "NAME10", UA_HIA, "NAME10","")
    arcpy.AddField_management(Study2, "HotArea", "Float", 9, 2, "", "", "", "")
    arcpy.CalculateField_management(Study2, "HotArea", "(!SUM!/!ARKM!)*100 ", "Python_9.3","")
    print("got joins and field calculations completed")

    ## Output to .dbf for later processing
    arcpy.TableToTable_conversion(Study2, outpath, newfieldname)
    print("Got it!!  Starting to clean up old rasters...")

    ## clean house to avoid problems
    arcpy.Delete_management(outfolder + "ZonalMean.tif")
    #os.remove(outfolder + "ZonalMean.tif")
    arcpy.Delete_management(outfolder + "ZonalSTD.tif")
    arcpy.Delete_management(outfolder + "HIA_tmp.tif.tif")
    arcpy.Delete_management(outfolder + "HotIsland.tif")
    print("All clear..  Moving on to the next LST image...")

####################################################################
##############################################  END Hot Island Area


############################################## Start LandCovers
################################################################
  This section has to be run for each year that LandCover is available.
  The land cover data has been preprocessed in ArcMap into
  shapefiles that make it easy to do zonal stats.

 grab the LC  shapefiles (pre-existing for 2002 to 2012)
LC = "C:\\GISdata\\NewUHI\\data\\MCD\\lc5_2012s_Intersect.shp"  ## Change here!
## set new workspace and list rasters
arcpy.env.workspace = "C:\\GISdata\\NewUHI\\Data\\MOD\\2012"  ## and change here! Dont forget the na
LCdata = arcpy.ListRasters("*", "TIF")

for dataset in LCdata:
    RasterLST = dataset
    print(dataset)
    ## Resample LST data to Cubic Convolution 250 meter
    lst_tmp_tif = arcpy.Resample_management(RasterLST, "re_tmp", "500 500", "CUBIC")
    print("resampled")
    #### Get the mean Urban LST for each UA boundary  #####
    Zonal_dbf = arcpy.gp.ZonalStatisticsAsTable_sa(LC, "FID_lc5_20", lst_tmp_tif, "zo_tmp.dbf", "DAT
    print("zonal stats")
    ####  Need to Add new field to adjust the precision
    arcpy.JoinField_management(Zonal_dbf, "FID_lc5_20", LC, "FID_lc5_20", "")
    print("joined")
    ### Complete statistics for the summary
    Name_Stats_dbf = outpath + dataset[8:16] + "LC_2012" + ".dbf" ## be sure to change this name eac
    arcpy.Statistics_analysis(Zonal_dbf, Name_Stats_dbf, "MEAN MEAN", "NAME10;GRIDCODE")
    print("stats, and finished.  Moving to next LST image...")



################################################################
################################################  END LandCovers

############################################## Start MAX and Range
################################################################

##  Magnitude and Range- the difference between MAX
##  and MEAN LST and the difference btw MAX and MIN, respectively
##  within Urban Area boundaries of the locations..
print("Starting MAX and Range ...")
```

```
##
###Local Variables
ZoneStats = outfolder + "ZoneAll"

## run loop like always...
for dataset in datasetList:
    newfieldname = "max_range_" + dataset[8:16] + ".dbf"
    RasterLST = dataset
    print(dataset)
    ## Resample LST data to Cubic Convolution 250 meter
    #lst_tmp_tif = arcpy.Resample_management(RasterLST, "re_tmp", "500 500", "CUBIC")
    ##  Run All statististic to get MAX and Mean, and calculate the difference btw them
    arcpy.gp.ZonalStatisticsAsTable_sa(UrbanArea, "NAME10", dataset, ZoneStats, "DATA", "ALL")
    print("finished the stat calculations.  Adding new field...")
    arcpy.AddField_management(ZoneStats, "MaxUHI", "Float", 9, 2, "", "", "", "")
    ## the below step first de-scales MODIS value so that the result is degrees K..
    arcpy.CalculateField_management(ZoneStats, "MaxUHI", "([MAX]*0.02) - ([MEAN]*0.02)", "VB","")
    print("New field should contain the new UHI value.  Exporting record now...")
    ## Output to .dbf for later processing
    arcpy.TableToTable_conversion(ZoneStats, outpath, newfieldname)
    print("Got it!!  Moving to the next LST image..")

####################################################################
####################################################  END MaxRange

################################################# Start Micro Island
###################################################################

###### Percentage of area within urban boundary with LST
###### higher than the MAX of Forested Landcover
print("Starting Micro Island...")

## set new workspace and list rasters
arcpy.env.workspace = "C:\\GISdata\\NewUHI\\Data\\MOD\\2012"  ##  Change here!
LCdata = arcpy.ListRasters("MOD*", "TIF")
## grab the LC  annual image
MCD = "C:\\GISdata\\NewUHI\\data\\MCD\\LC2012\\MCD12Q1.A2012001.Land_Cover_Type_1.tif"  ## Change he

# Local Variables
MCD_forest = outfolder + "MCD_forest"
F_LST = outfolder + "F_lst.tif"
MCD_F = outfolder + "MCD_F_lst"
Tmp = outfolder + "TmpMOD.tif"
MOD_for = outfolder + "MOD_forest"
Study6 = outfolder + "UrbanProject6.shp"

## and the loop of course, standard fare by now..
for dataset in LCdata:
    newfieldname = "micro_" + dataset[8:16] + ".dbf"  ## dont include the outpath here!  The export
    RasterLST = dataset
    print(dataset)
    ## Resample LST data to Cubic Convolution 250 meter
    lst_tmp_tif = arcpy.Resample_management(RasterLST, "re_tmp", "500 500", "CUBIC")
    print("resampled")

    ## Creat copy of study area to permit joins and such..
    arcpy.CopyFeatures_management(UrbanArea, Study6,"","","","")
    #Reclassify Forest land cover =1, all others = NODATA,
    #Sample to get the MAX Forest LST.
    arcpy.gp.Reclassify_sa(MCD, 'VALUE', '0 NODATA;1 1;2 1;3 1;4 1;5 1;6 NODATA;7 NODATA;8 NODATA;9
    print("copied study area and reclassified the land cover the data")
    arcpy.gp.ExtractByMask_sa( lst_tmp_tif, MCD_forest, F_LST)
    ForestMAX = arcpy.sa.ZonalStatistics(UrbanArea,"NAME10",F_LST, "MAXIMUM", "DATA")
    ForestMAX.save(outfolder + "MCD_F_MAX")
    print("Extract LST pixels and summarized data")
    ## The raster calculations are touchy.  It would seem that I could do better
    ## than the next 6 lines, but this is what I could get to work..
    ## Raster calculation is  (MOD 11 - MaxForest LST = Difference btw the 2)
    d1= outfolder + "MCD_F_MAX"
```

```
    d2=arcpy.sa.Raster(lst_tmp_tif)
    outraster1 = ((d2 - d1)*0.02)
    outraster1.save(outfolder + "MOD_fmax.tif")
    d3= outfolder + "MOD_fmax.tif"
    print("value obtained")
    ## reclassify to show resulting LST pixels that met condition ( >= MAX Forest LST)
    arcpy.gp.Reclassify_sa(d3, "Value", "-50 0 NODATA;0 50 1", Tmp, "DATA")
    print("reclassified LST data")
    del (d1)
    del (d2)
    del (d3)
    ### The study area file changes because of duplicated fields in the previous zonal steps
    arcpy.gp.ZonalStatisticsAsTable_sa(Study6, "NAME10",Tmp, MOD_for, "DATA", "SUM")
    arcpy.JoinField_management(Study6, "NAME10", MOD_for, "NAME10","")
    arcpy.AddField_management(Study6, "MicArea", "Float", 9, 2, "", "", "", "")
    arcpy.CalculateField_management(Study6, "MicArea", "(!SUM!/!ARKM!)*100 ", "Python_9.3","")
    ##arcpy.gp.ZonalStatisticsAsTable_sa(StudyArea, "NAME10",Tmp, MOD_for, "DATA", "SUM")
    ##arcpy.AddField_management(MOD_for, "UHI6", "Float", 9, 2, "", "", "", "")
    ##arcpy.CalculateField_management(MOD_for, "UHI6", "(!SUM!/!COUNT!)* 100", "Python_9.3","")
    print("New field should contain the new UHI value.  Exporting record now...")
    ## Output to .dbf for later processing
    arcpy.TableToTable_conversion(Study6, outpath, newfieldname)
    print("Got it!!  Moving to the next LST image..")

print("finished")
timer = time.time()
timer1 = time.localtime(timer)
timer2= time.asctime(timer1)
print(timer2)


##################################################################
##################################################  END Micro Island


############################################### Start Louisville Samples
###########################################################

####  Magnitude and Range- the difference between MAX
####  and MEAN LST and the difference btw MAX and MIN, respectively
####  within Urban Area boundaries of the locations..
print("Louisville Sampling...")

#Local Variables
ZoneStats = outfolder + "ZoneAll_"

## run loop like always...
for dataset in datasetList:
    newfieldname = dataset[8:16] + "Lou" + ".dbf"
    RasterLST = dataset
    print(dataset)
    ##    ## Resample LST data to Cubic Convolution 250 meter
    lst_tmp_tif = arcpy.Resample_management(RasterLST, "re_tmp", "250 250", "CUBIC")
    print("resampled")
    ##  Run All statistics (or just what you need)
    arcpy.gp.ZonalStatisticsAsTable_sa(Lou, "name", lst_tmp_tif, ZoneStats, "DATA", "ALL")
    print("finished the stat calculations")
    ## Output to .dbf for later processing
    arcpy.TableToTable_conversion(ZoneStats, outpath, newfieldname)
    print("Got it!!  Moving to the next LST image..")

###########################################################
################################################  END Louisville Samples

### UHI Indicator 7: The closest approximation of the methods
### used by Stone and Kenward et al..  Difference in mean LST
### Urban and mean LST for (245-55km) buffered Rural.
 Local Variables
Stone= "C:\\UHI\\Data\\Boundaries\\Stone_rural.shp"
```

```
Urban7 = "C:\\UHI\\Output\\urban7"
Rural7 = "C:\\UHI\\Output\\rural7"

## Get the mean Urban LST for each UA boundary and put that value in field called URBAN7
arcpy.gp.ZonalStatisticsAsTable_sa(StudyArea, "NAME10", MOD11, Urban7, "DATA", "MEAN")
arcpy.AddField_management(Urban7, "Urban7", "Float", 9, 2, "", "", "", "")
arcpy.CalculateField_management(Urban7, "Urban7", "(!MEAN!*0.02)-273.15", "Python_9.3","")
## Get the mean LST for each buffered rural zone, place in field called RURAL7
### Note that Dayton, OH is not included in the output!! :(
arcpy.gp.ZonalStatisticsAsTable_sa(Stone, "NAME10", MOD11, Rural7, "DATA", "MEAN")
arcpy.AddField_management(Rural7, "Rural7", "Float", 9, 2, "", "", "", "")
arcpy.CalculateField_management(Rural7, "Rural7", "(!MEAN!*0.02)-273.15", "Python_9.3","")
## Join the URBAN7 and RURAL7 Tables and do calculation
arcpy.JoinField_management(Urban7, "NAME10", Rural7, "NAME10","")
arcpy.AddField_management(Urban7, "UHI7", "Float", 9, 2, "", "", "", "")
arcpy.CalculateField_management(Urban7, "UHI7", "(!URBAN7! - !RURAL7!)", "Python_9.3","")
## Create Excel File from INFO table derived previous step, name of table is appended to workspace p
arcpy.TableToExcel_conversion(Urban7, "UHI7.xls")

print("You should have 7 new excel files...")
print("Remind Amy about Cassie's stuff")




######## Point based UHI points
################################################## Start Lou Pts of LST
#######################################################
for dataset in datasetList:
    newfieldname = outpath + dataset[8:16] + "_lst_lou" + ".dbf"
    print(newfieldname)
    ## Resample ndvi data to Nearest Neighbor 250 meter
    #lst_tmp_tif = arcpy.Resample_management(dataset, "re_tmp", "500 500", "NEAREST")
    print("didn't resample LST")
    # Execute ExtractValuesToTable
    arcpy.ExtractValuesToTable_ga(LOUpts, dataset, newfieldname, "", "")
    print("Pixel values extracted!!")
    print("Got it and moving on to the next!!")
################################################## End Lou Pts of LST
#######################################################
```

A2. Code implemented in R for month to month correlations (H1)

Import all packages

```
##############################################################
##### Install needed libraries and packages  ################

#install.packages("spacetime")
require("spacetime")
#install.packages("xts")
require(xts)
#install.packages("tidyr")
require("tidyr")
library("dplyr")
#install.packages("reshape2")
require("reshape2")
require(foreign)
require(ggplot2)
#install.packages("season")
require(scales)
require(Hmisc)
library(stats)
```

Set a few global variables. This section will be expanded later to reduce the number of overall code blocks. Need a little help in making my blocks fit within functions that I can call over and over.

```
### Set pathway to be appended to each output.
path <- "C:/GISdata/NewUHI/output/"
## Create a master list that will hold all of the data
AllData <- data.frame()
## subset to 2012 only
time1<-as.Date("2012-01-01")
time2<-as.Date("2012-12-26")
```

We will go through the steps necesary to import the data generated from the .py sampling doc. Here we import all of the .dbfs and start to summarize and visualize the data in support of our research objectives.

1. Import Urban LSTs for each location

```
##########################################################################################
##############
##########################################
####  Start Urban Metric  ################

### set directory by appending above pathway. Then create
### vector contining the names of each of the visible files in the working directory.
setwd(paste(path, "/UrbanTables/", sep=""))
files <- list.files(pattern = "\\.dbf$")

## Create a new vector to retain combined urban data
urban_data <- data.frame()

##### set the looping structure to cycle through each file returned from dir()
for (file in files) {
  ##  extract day of year (DOY) from file name
  DOY <- substr(file, 8, 14)
```

```
  ## read in the data
  table_data <- read.dbf(file, as.is = TRUE)
  ## grab the variables we need
  new_data <- table_data[,c(1,6)]
  ## create new column called 'year' to store the label (factor?) we created above
  new_data[,"DOY"] <- DOY
  ## create new column for ZONE type
  new_data[,"ZONE"] <- "Urban"
  ## append this newly created data frame to the master we created earlier
  urban_data <- rbind(urban_data, new_data)
}
#head(urban_data)
# clean house (just in case)
rm(DOY)
rm(table_data)
rm(new_data)
rm(file)
rm(files)

# rename variables to your liking
names(urban_data) <- c("name", "value", "date", "type")
## recognize the date
urban_data[,"date"] <- as.Date(urban_data$date, format="%Y%j")

## merge with the master sheet
AllData <- rbind(AllData, urban_data)
rm(urban_data) ## clean up to conserve memory


##########################################################################
######################################## End Urban Metric
```

2. Import Rural Buffer LSTs for each location

```
#############################################
####  Start Rural Metric ################

### set new directory and list files
setwd(paste(path, "/RuralTables/", sep=""))
files <- list.files(pattern = "\\.dbf$")

## Create a new vector to retain combined rural data
rural_data <- data.frame()

##### set the looping structure to cycle through each file returned from dir()
for (file in files) {
  ##  extract day of year (DOY) from file name
  DOY <- substr(file, 7, 13)
  ## read in the data
  table_data <- read.dbf(file, as.is = TRUE)
  ## grab the variables we need
  new_data <- table_data[,c(1,6)]
  ## create new column called 'year' to store the label (factor?) we created above
  new_data[,"DOY"] <- DOY
  ## create new column for ZONE type
  new_data[,"ZONE"] <- "Rural"
  ## append this newly created data frame to the master we created earlier
  rural_data <- rbind(rural_data, new_data)
}
```

```
#head(rural_data)

# clean house (just in case)
rm(DOY)
rm(table_data)
rm(new_data)
rm(file)
rm(files)

# rename variables to your liking
names(rural_data) <- c("name", "value", "date", "type")

## recognize the date
rural_data[,"date"] <- as.Date(rural_data$date, format="%Y%j")

## merge with the master sheet
AllData <- rbind(AllData, rural_data)
rm(rural_data) ## clean up to conserve memory

##########################################################################
####################################### End Rural Metric
```

3. Calculate the difference between the Urban and Rural zones for each location.

```
#####################################################################################
####  Start Calculation of UHI "Urban-Rural Difference" Metric  ################

## http://seananderson.ca/2013/10/19/reshape.html

## convert "long" data to "wide" using reshape2 package
AllData_wide <- dcast(AllData, date + name ~ type, value.var = "value")
## do the calculations
AllData_wide$DiffUHI <- AllData_wide$Urban - AllData_wide$Rural
## reform the data as "long" format
AllData_melted <- melt(AllData_wide, id.vars = c("name", "date"))
## Re-select the variables in the order we want
AllData <- AllData_melted[,c(1,4,2,3)]
## rename the columns accordingly
names(AllData) <- c("name", "value", "date", "type")

## clean up for memory's sake..
rm(AllData_melted)
rm(AllData_wide)
```

5. Summarize by month and visualize that distribution

```
####  Summarize the data according to class label
#subsetData <- filter(AllData, date > time1 & date < time2 & name == "Louisville/Jeffer
son County, KY--IN" & type =="DiffUHI")
subsetData <- filter(AllData, date > time1 & date < time2 & type=="Urban"|type=="Rural"
)

## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                       levels = c("01", "02", "03", "04", "05", "06",
                                  "07", "08", "09", "10", "11", "12"),
                       labels = c("jan", "feb", "mar", "april", "may",
                                  "jun", "july", "aug", "sept",
                                  "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
```

```
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))
subsetData$Temp<- subsetData$value

## vectorize as df for plotting
df <- subsetData

#########################################
#####  Simple Boxplot of all three variables used for the SUHI calculations
ggplot(df, aes(x=Month, y=value, fill=type)) + geom_boxplot() + ylim(0,40) +
    ggtitle("All Locations SUHI: Urban, Rural LST Means (2012)")

## Warning: Removed 1530 rows containing non-finite values (stat_boxplot).
```

6. Convert each value a ranking and test for correlation for each month. The question being answered here involves whether or not the rankings are staying in roughly the same order each step. The code presented probably more complicated than it should be, but I didnt map this out ahead of time and so missed oppurtunities tom catch the individual variables being isolated here.

```
#######################################################################
######################  Start Ranking of DUR Values 2012   ########

## subset the above to include only the SUHI indicator
subsetData <- filter(AllData, date > time1 & date < time2 & type=="DiffUHI")
## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                      levels = c("01", "02", "03", "04", "05", "06",
                                 "07", "08", "09", "10", "11", "12"),
                      labels = c("jan", "feb", "mar", "april", "may",
                                 "jun", "july", "aug", "sept",
                                 "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))

## convert "long" data to "wide" using reshape2 package and
## calculate the mean in the process
subsetData_wide <- dcast(subsetData, name ~ Month, mean)
## check it out
head(subsetData_wide)

##                       name       jan      feb      mar    april      may
## 1 Albany--Schenectady, NY 1.2276340 2.240340 2.430183 3.287900 3.479167
## 2       Allentown, PA--NJ 1.1426017 1.137170 1.979775 2.669400 3.226800
## 3             Atlanta, GA 0.4722233 1.033645 1.671875 2.349850 2.891167
## 4           Baltimore, MD 0.8824700 1.554670 1.140550 1.454550 2.272100
## 5         Baton Rouge, LA 1.1924000 1.849325 2.463575 3.512275 3.847000
## 6          Birmingham, AL 1.0730633 1.259670 1.954600 2.283975 3.118400
##        jun     july      aug     sept      oct       nov       dec
## 1 3.885650 4.461900 3.710775 1.912425 1.618465 1.0904100 0.9443323
## 2 3.683500 3.983500 3.286225 2.616625 1.540538 0.4680367 1.6188533
## 3 3.166625 2.506100 2.086350 2.264825 1.179175 0.7013333 0.5352400
## 4 2.748200 3.234375 2.542400 1.797650 1.220695 1.0070233 0.9623733
## 5 3.274450 2.063700 2.696200 3.036775 2.194525 1.6617667 1.2911333
## 6 3.259100 2.500175 2.364850 2.267600 1.593850 1.2503333 0.8897333

## reform the data back to "long" format
subsetData_melted <- melt(subsetData_wide, id.vars = "name")
```

```
## Create a new attribute to rank the variables for each time step.
rankings <- subsetData_melted %>%
  group_by(variable) %>%
  mutate(yrrank = row_number(-value))

## convert "long" data to "wide" using reshape2 package
rankings_wide <- dcast(rankings, name ~ variable,value.var="yrrank")
## check the rankings

## remove the names
rankings_wide <- rankings_wide[,-1]
## generate covariate matrix using require "Hmisc"" package ((cormatrix$r) and (cormatr
ix$P) to access output)
cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')

# ### Generate Corrwlation visuale or Rhos or p-vals
# cordata = melt(cormatrix$P)
# ggplot(cordata, aes(x=Var1, y=Var2, fill=value)) +
#   geom_tile() + xlab("") + ylab("")

### Generate with all info included
## https://www.r-bloggers.com/spearman-correlation-heat-map-with-correlation-coefficien
ts-and-significance-levels-in-r/

abbreviateSTR <- function(value, prefix){  # format string more concisely
  lst = c()
  for (item in value) {
    if (is.nan(item) || is.na(item)) { # if item is NaN return empty string
      lst <- c(lst, '')
      next
    }
    item <- round(item, 2) # round to two digits
    if (item == 0) { # if rounding results in 0 clarify
      item = '<.01'
    }
    item <- as.character(item)
    item <- sub("(^[0])+", "", item)    # remove leading 0: 0.05 -> .05
    item <- sub("(^-[0])+", "-", item)  # remove leading -0: -0.05 -> -.05
    lst <- c(lst, paste(prefix, item, sep = ""))
  }
  return(lst)
}

cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')
cordata = melt(cormatrix$r)
cordata$labelr = abbreviateSTR(melt(cormatrix$r)$value, 'r')
cordata$labelP = abbreviateSTR(melt(cormatrix$P)$value, 'P')
cordata$label = paste(cordata$labelr, "n",
                      cordata$labelP, sep = "")
cordata$strike = ""
cordata$strike[cormatrix$P > 0.05] = "X"

txtsize <- par('din')[2] / 2.25  ## change demoninator for txt size
ggplot(cordata, aes(x=Var1, y=Var2, fill=value)) + geom_tile() +
  theme(axis.text.x = element_text(angle=90, hjust=TRUE)) +
  xlab("") + ylab("") +
  geom_text(label=cordata$label, size=txtsize) +
  geom_text(label=cordata$strike, size=txtsize * 4, color="red", alpha=0.4)
```

7. That brings us to the end of the DUR Indicator. Now, Import Land cover specific data..

```
########################################################################################
###
####################################################
####  Start Landcover Metrics ################

### set new directory and list files

setwd(paste(path, "/LandCoverTables/", sep=""))
files <- list.files(pattern = "\\.dbf$")


## Create a new vector to retain combined rural data
LC_data <- data.frame()
##### set the looping structure to cycle through each file returned from dir()
for (file in files) {
  ##  extract day of year (DOY) from file name
  DOY <- substr(file, 2, 8)
  ## read in the data
  table_data <- read.dbf(file, as.is = TRUE)
  ## create new column called 'year' to store the label (factor?) we created above
  table_data[,"DOY"] <- DOY
  ## grab the variables we need
  new_data <- table_data[,c(1,4,5,2)] ## remove this
  ## append this newly created data frame to the master we created earlier
  LC_data <- rbind(LC_data, new_data)
}

# clean house
rm(DOY)
rm(table_data)
rm(new_data)
rm(file)
rm(files)

## Make a few needed changes to the data
# rename variables to your liking
names(LC_data) <- c("name", "value", "date", "type")

# go ahead and unscale the value
LC_data$value <- LC_data$value * 0.02 - 273.15

# set type as factor
LC_data$type <- as.factor(LC_data$type)
LC_data$type <- factor(LC_data$type,
                    levels = c(1,2,3,4,5),
                    labels = c("forest", "other", "ag", "urban", "bare"))
#                    labels = c("forest", "other", "ag", "urban", "urban"))
## recognize the date
LC_data[,"date"] <- as.Date(LC_data$date, format="%Y%j")


head(LC_data)

##                        name      value       date   type
## 1 Albany--Schenectady, NY -3.2659852 2002-01-01 forest
## 2 Albany--Schenectady, NY -3.0555670 2002-01-01  other
## 3 Albany--Schenectady, NY -3.7444220 2002-01-01     ag
## 4 Albany--Schenectady, NY -3.2965133 2002-01-01  urban
```

```
## 5       Allentown, PA--NJ -0.2369192 2002-01-01 forest
## 6       Allentown, PA--NJ  0.3079781 2002-01-01  other
```

8. Visualize the distribution of LSTs by land cover type

```
#########################################################
####  Start Plot of Land Covers  ################
## subset data
subsetData <- filter(LC_data, date > time1 & date < time2)

## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                       levels = c("01", "02", "03", "04", "05", "06",
                                  "07", "08", "09", "10", "11", "12"),
                       labels = c("jan", "feb", "mar", "april", "may",
                                  "jun", "july", "aug", "sept",
                                  "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))
subsetData$Temp<- subsetData$value

## vectorize as df for plotting
df <- subsetData

#########################################
#####  Simple Boxplot of all three variables used for the SUHI calculations
ggplot(df, aes(x=Month, y=value, fill=type)) + geom_boxplot() +
    ggtitle("All Locations Land Cover LSTs: Forest, Other, Ag, Developed, Bare (2012)")
```

9. Calculate the relevant land cover driven SUHIs and then summarize by month

```
###############
## make the other necesary calculations using reshape2  ## http://seananderson.ca/2013/
10/19/reshape.html

## convert "long" data to "wide" using reshape2 package
LCdata_wide <- dcast(LC_data, date + name ~ type, value.var = "value")

## do the calculations for Difference Urban - Other UHI metric
LCdata_wide$Diff_UrbanOther <- LCdata_wide$urban - LCdata_wide$other
## do the calculations for Difference Urban - Ag UHI metric
LCdata_wide$Diff_UrbanAg <- LCdata_wide$urban - LCdata_wide$ag
## do the calculations for Difference Urban - Forest UHI metric
LCdata_wide$Diff_UrbanForest <- LCdata_wide$urban - LCdata_wide$forest

## reform the data as "long" format
LCdata_melted <- melt(LCdata_wide, id.vars = c("name", "date"))

## Re-select the variables in the order we want
LC_data_processed <- LCdata_melted[,c(1,4,2,3)]
## rename the columns accordingly
names(LC_data_processed) <- c("name", "value", "date", "type")

## merge with the master sheet
AllData <- rbind(AllData, LC_data_processed)

## resubset data to get all years
subsetData <- filter(LC_data_processed)
## Grab the variables we need for the summary
```

```
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                        levels = c("01", "02", "03", "04", "05", "06",
                                   "07", "08", "09", "10", "11", "12"),
                        labels = c("jan", "feb", "mar", "april", "may",
                                   "jun", "july", "aug", "sept",
                                   "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))
```

11. Generate DUA boxplots, rankings, and Rho

```
############################################
## Pull out only the DUA indicator and generate boxplot
df <- filter(subsetData, date > time1 & date < time2 & type=="Diff_Urba
nAg")

ggplot(df, aes(x=Month, y=value)) + geom_boxplot() +
    ggtitle("2012 All Locations SUHI: Difference Urban - Ag (DUA)")

## Warning: Removed 1 rows containing non-finite values (stat_boxplot).
```

```
###### Generate Rankings
## convert "long" data to "wide" using reshape2 package and
## calculate the mean in the process
df_wide <- dcast(df, name ~ Month, mean)

## reform the data back to "long" format
df_melted <- melt(df_wide, id.vars = "name")

## Create a new attribute to rank the variables for each time step.
rankings <- df_melted %>%
  group_by(variable) %>%
  mutate(yrrank = row_number(-value))

## convert "long" data to "wide" using reshape2 package
rankings_wide <- dcast(rankings, name ~ variable,value.var="yrrank")
## remove the names
rankings_wide <- rankings_wide[,-1]
## generate covariate matrix using require "Hmisc"" package ((cormatrix$r) and (cormatr
ix$P) to access output)
cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')
```

12. Generate DUO boxplots, rankings, and Rho

```
## Pull out only the DUA indicator and generate boxplot
df <- filter(subsetData, date > time1 & date < time2 & type=="Diff_UrbanOther")

ggplot(df, aes(x=Month, y=value)) + geom_boxplot() +
    ggtitle("2012 All Locations SUHI: Difference Urban - Other (DUO)")

## Warning: Removed 1 rows containing non-finite values (stat_boxplot).

###### Generate Rankings
## convert "long" data to "wide" using reshape2 package and
## calculate the mean in the process
df_wide <- dcast(df, name ~ Month, mean)
```

```
## reform the data back to "long" format
df_melted <- melt(df_wide, id.vars = "name")

## Create a new attribute to rank the variables for each time step.
rankings <- df_melted %>%
  group_by(variable) %>%
  mutate(yrrank = row_number(-value))

## convert "long" data to "wide" using reshape2 package
rankings_wide <- dcast(rankings, name ~ variable,value.var="yrrank")

## remove the names
rankings_wide <- rankings_wide[,-1]
## generate covariate matrix using require "Hmisc"" package ((cormatrix$r) and (cormatr
ix$P) to access output)
cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')



#########################################################################
####################################### End Land Covers
```

13. Visualize for change over time for the land cover SUHIs

```
## resubset data to get all years
subsetData <- filter(AllData, name == "Louisville/Jefferson County, KY--IN")
## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                       levels = c("01", "02", "03", "04", "05", "06",
                                  "07", "08", "09", "10", "11", "12"),
                       labels = c("jan", "feb", "mar", "april", "may",
                                  "jun", "july", "aug", "sept",
                                  "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))
subsetData$Temp<- subsetData$value

## vectorize as df for plotting
df <- subsetData


###################################################
## Complete summary of year and month measurements.  All Data repersented!
devData <- subset(df, type=="Urban", select=c(type, Year, Month, value))
meanDevData <- subset(df, type=="Rural", select=c(type, Year, Month, value))
medDevData <- subset(df, type=="DiffUHI", select=c(type, Year, Month, value))

ggplot(df,aes(Year,value,colour=type)) +
  #geom_point(data=devData,size=I(2),alpha=I(0.6)) +
  #geom_line(data=meanDevData,size=I(1.5),alpha=I(0.6)) +
  geom_line(data=medDevData,size=I(1.5),alpha=I(0.4)) +
  theme_grey(base_size=15) +
  theme(legend.title = element_blank(), legend.position=c(.5,.25), axis.title.y=element
_blank(),axis.text.x=element_blank()) +
  ggtitle("Louisville, Kentucky: DUR SUHI by Month (2000 to 2015)") + facet_grid(. ~ Mo
nth) +
  xlab("Years: 2000 to 2015")
```

14. Start Micro Island SUHI indicator..

```
####################################################################################
##############################################################
####  Start Micro Island Area % Metric ################

## This metric is an estimate of the portion of the total land
## area within the urban area is warming than the coolest forested
## LST for each time step.

### set new directory and list files

setwd(paste(path, "/MicroIslandAreaTables/", sep=""))
files <- list.files(pattern = "\\.dbf$")

## Create a new vector to retain combined rural data
MicroIsland_data <- data.frame()

##### set the looping structure to cycle through each file returned from dir()
for (file in files) {
  ##  extract day of year (DOY) from file name
  DOY <- substr(file, 8, 14)
  ## read in the data
  table_data <- read.dbf(file, as.is = TRUE)
  ## grab the variables we need
  new_data <- table_data[,c(2,14)]
  ## create new column called 'year' to store the label (factor?) we created above
  new_data[,"DOY"] <- DOY
  ## append this newly created data frame to the master we created earlier
  MicroIsland_data <- rbind(MicroIsland_data, new_data)
}

head(MicroIsland_data)

##                      NAME10   MicArea     DOY
## 1              Hartford, CT  0.220501 2002001
## 2            Pittsburgh, PA  0.963750 2002001
## 3             Baltimore, MD  4.212750 2002001
## 4 Washington, DC--VA--MD  4.408420 2002001
## 5            Birmingham, AL 35.939900 2002001
## 6 Cincinnati, OH--KY--IN  6.585990 2002001

# clean house
rm(DOY)
rm(table_data)
rm(new_data)
rm(file)
rm(files)

## create new column for  type
MicroIsland_data[,"type"] <- "MicroArea%"
# rename variables to your liking
names(MicroIsland_data) <- c("name", "value", "date", "type")
## recognize the date
MicroIsland_data[,"date"] <- as.Date(MicroIsland_data$date, format="%Y%j")

## merge with the master sheet
AllData <- rbind(AllData, MicroIsland_data)

## Subset only the variables we want to look at
```

```
subsetData <- filter(MicroIsland_data, date > time1 & date < time2 , type=="MicroArea%"
)
## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                       levels = c("01", "02", "03", "04", "05", "06",
                                  "07", "08", "09", "10", "11", "12"),
                       labels = c("jan", "feb", "mar", "april", "may",
                                  "jun", "july", "aug", "sept",
                                  "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))

## vectorize as df for plotting
df <- subsetData

#########################################
#####  Simple Boxplot of all three variables used for the SUHI calculations
ggplot(df, aes(x=Month, y=value, fill=type)) + geom_boxplot() +
    scale_y_continuous(limits=c(0,100), breaks=seq(0,100,10), expand = c(0, 0)) +
    ggtitle("All Locations: MicroIsland % (2012)")

##################################################### Generate Rankings
## convert "long" data to "wide" using reshape2 package and
## calculate the mean in the process
df_wide <- dcast(df, name ~ Month, mean)

## reform the data back to "long" format
df_melted <- melt(df_wide, id.vars = "name")

## Create a new attribute to rank the variables for each time step.
rankings <- df_melted %>%
  group_by(variable) %>%
  mutate(yrrank = row_number(-value))

## convert "long" data to "wide" using reshape2 package
rankings_wide <- dcast(rankings, name ~ variable,value.var="yrrank")

## remove the names
rankings_wide <- rankings_wide[,-1]
## generate covariate matrix using require "Hmisc"" package ((cormatrix$r) and (cormatr
ix$P) to access output)
cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')

#########################################################################
######################################## End Micro Island Area %
```

15. Start UHI-driven indicators..

```
#######################################################################################
##############
###########################################################
####  Start Max & Range Metric  ################

## The name of this metric is a bit misleading.  The Range is just that,
## the difference between the MIN and MAX.  The Max UHI, here, is defined
## as the difference between the mean of all LST and the absolute highest LST.

### Set pathway to be appended to each output.
path <- "C:/GISdata/NewUHI/output/"

### set new directory and list files
```

```
setwd(paste(path, "/MaxRangeTables/", sep=""))
files <- list.files(pattern = "\\.dbf$")

## Create a new vector to retain combined rural data
range_data <- data.frame()
##### set the looping structure to cycle through each file returned from dir()
for (file in files) {
  ##  extract day of year (DOY) from file name
  DOY <- substr(file, 12, 18)
  ## read in the data
  table_data <- read.dbf(file, as.is = TRUE)
  ## grab the variables we need
  new_data <- table_data[,c(1,15,5,6,7)]
  ## create new column called 'year' to store the label (factor?) we created above
  new_data[,"DOY"] <- DOY
  ## append this newly created data frame to the master we created earlier
  range_data <- rbind(range_data, new_data)
}

# clean house (just in case)
rm(DOY)
rm(table_data)
rm(new_data)
rm(file)
rm(files)

## Now make some conversions to make the data more readable
range_data$MIN <- range_data$MIN * 0.02 - 273.15
range_data$MAX <- range_data$MAX * 0.02 - 273.15
range_data$RANGE <- range_data$RANGE * 0.02

## rename variables to your liking
names(range_data) <- c("name", "MAG", "MIN", "MAX", "RANGE", "date") ## mag was called
maxuhi in python by accident!

## recognize the date
range_data[,"date"] <- as.Date(range_data$date, format="%Y%j")

## reform the data as "long" format using melt()
range_data_melted <- melt(range_data, id.vars = c("name", "date"))

## Re-select the variables in the order we want
range_data <- range_data_melted[,c(1,4,2,3)]

## rename the columns accordingly
names(range_data) <- c("name", "value", "date", "type")

## merge with the master sheet
AllData <- rbind(AllData, range_data)

## clean up for memory's sake..
rm(range_data_melted)


#########################################################################
####################################### End Max & Range Metric
```

16. Plot the Summary Stats of Min, Max, Range for All locations

```
#########################################################
####  Start Plot of Max Min Range Values   ################
```

```
## Subset only the variables we want to look at
subsetData <- filter(range_data, date > time1 & date < time2 , type=="MAX" | type=="MIN
" | type=="RANGE")

## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                       levels = c("01", "02", "03", "04", "05", "06",
                                  "07", "08", "09", "10", "11", "12"),
                       labels = c("jan", "feb", "mar", "april", "may",
                                  "jun", "july", "aug", "sept",
                                  "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))

## vectorize as df for plotting
df <- subsetData

#########################################
#####  Simple Boxplot of all three variables used for the SUHI calculations
ggplot(df, aes(x=Month, y=value, fill=type)) + geom_boxplot() +
    ggtitle("All Locations: Min, Max, Range Value Distribution (2012)")
```

17. Now Generate Range SUHI boxplots, rankings, and Rho

```
## Pull out only the DUA indicator and generate boxplot
df <- filter(subsetData, type=="RANGE")

ggplot(df, aes(x=Month, y=value)) + geom_boxplot() +
    ggtitle("2012 All Locations SUHI: Range")

###### Generate Rankings
## convert "long" data to "wide" using reshape2 package and
## calculate the mean in the process
df_wide <- dcast(df, name ~ Month, mean)

## reform the data back to "long" format
df_melted <- melt(df_wide, id.vars = "name")

## Create a new attribute to rank the variables for each time step.
rankings <- df_melted %>%
  group_by(variable) %>%
  mutate(yrrank = row_number(-value))

## convert "long" data to "wide" using reshape2 package
rankings_wide <- dcast(rankings, name ~ variable,value.var="yrrank")

## remove the names
rankings_wide <- rankings_wide[,-1]
## generate covariate matrix using require "Hmisc"" package ((cormatrix$r) and (cormatr
ix$P) to access output)
cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')

###############
```

18. Plot the Magnitude for All locations by month

```
#########################################################
####  Start Plot of Magnitude Values   ################

## Subset only the variables we want to look at
subsetData <- filter(range_data, date > time1 & date < time2 , type=="MAG")
```

```
## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                      levels = c("01", "02", "03", "04", "05", "06",
                                 "07", "08", "09", "10", "11", "12"),
                      labels = c("jan", "feb", "mar", "april", "may",
                                 "jun", "july", "aug", "sept",
                                 "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))

## vectorize as df for plotting
df <- subsetData

#########################################
#####  Simple Boxplot of all three variables used for the SUHI calculations
ggplot(df, aes(x=Month, y=value, fill=type)) + geom_boxplot() +
    ggtitle("All Locations SUHI: Magnitude  (2012)")
```

19. Now Generate Magnitude SUHI boxplots, rankings, and Rho

```
###### Generate Rankings
## convert "long" data to "wide" using reshape2 package and
## calculate the mean in the process
df_wide <- dcast(df, name ~ Month, mean)

## reform the data back to "long" format
df_melted <- melt(df_wide, id.vars = "name")

## Create a new attribute to rank the variables for each time step.
rankings <- df_melted %>%
  group_by(variable) %>%
  mutate(yrrank = row_number(-value))

## convert "long" data to "wide" using reshape2 package
rankings_wide <- dcast(rankings, name ~ variable,value.var="yrrank")


## remove the names
rankings_wide <- rankings_wide[,-1]
## generate covariate matrix using require "Hmisc"" package ((cormatrix$r) and (cormatr
ix$P) to access output)
cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')

###############
```

20. Bring in the Hot Island Area Data

```
#######################################################################################
##############
########################################################
####  Start Hot Island Area % Metric ################

### set new directory and list files

setwd(paste(path, "/HotIslandAreaTables/", sep=""))
files <- list.files(pattern = "\\.dbf$")

## Create a new vector to retain combined rural data
HotIsland_data <- data.frame()
```

```r
##### set the looping structure to cycle through each file returned from dir()
for (file in files) {
  ##  extract day of year (DOY) from file name
  DOY <- substr(file, 6, 12)
  ## read in the data
  table_data <- read.dbf(file, as.is = TRUE)
  ## grab the variables we need
  new_data <- table_data[,c(2,14)]
  ## create new column called 'year' to store the label (factor?) we created above
  new_data[,"DOY"] <- DOY
  ## append this newly created data frame to the master we created earlier
  HotIsland_data <- rbind(HotIsland_data, new_data)
}

# clean house
rm(DOY)
rm(table_data)
rm(new_data)
rm(file)
rm(files)

head(HotIsland_data)

##                      NAME10 HotArea     DOY
## 1            Hartford, CT 18.2281 2000065
## 2          Pittsburgh, PA 15.7133 2000065
## 3           Baltimore, MD 11.9101 2000065
## 4 Washington, DC--VA--MD 12.4524 2000065
## 5          Birmingham, AL 17.0660 2000065
## 6 Cincinnati, OH--KY--IN 16.2713 2000065

## create new column for ZONE type
HotIsland_data[,"type"] <- "HotArea%"
# rename variables to your liking
names(HotIsland_data) <- c("name", "value", "date", "type")
## recognize the date
HotIsland_data[,"date"] <- as.Date(HotIsland_data$date, format="%Y%j")

## merge with the master sheet
AllData <- rbind(AllData, HotIsland_data)

##############################################################################
####################################### End Hot Island Area % Metric
```

21. Plot the HIA data for 2012

```r
##########################################################
####  Start Plot of Hot Island Area Metric   ################

## subset
subsetData <- filter(HotIsland_data, date > time1 & date < time2, type=="HotArea%")

## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                           levels = c("01", "02", "03", "04", "05", "06",
                                      "07", "08", "09", "10", "11", "12"),
                           labels = c("jan", "feb", "mar", "april", "may",
                                      "jun", "july", "aug", "sept",
                                      "oct", "nov", "dec"))
```

```
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))

## vectorize as df for plotting
df <- subsetData

#########################################
#####  Simple Boxplot of all three variables used for the SUHI calculations
ggplot(df, aes(x=Month, y=value, fill=type)) + geom_boxplot() +
  scale_y_continuous(limits=c(0,25), breaks=seq(0,100,10), expand = c(0, 0)) +
    ggtitle("All Locations SUHI: Hot Island Area %  (2012)")

## clen up for memory's sake...
rm(HotIsland_data)
```

22. Now Generate Hot Island Area SUHI boxplots, rankings, and Rho

```
###### Generate Rankings
## convert "long" data to "wide" using reshape2 package and
## calculate the mean in the process
df_wide <- dcast(df, name ~ Month, mean)

## reform the data back to "long" format
df_melted <- melt(df_wide, id.vars = "name")

## Create a new attribute to rank the variables for each time step.
rankings <- df_melted %>%
  group_by(variable) %>%
  mutate(yrrank = row_number(-value))

## convert "long" data to "wide" using reshape2 package
rankings_wide <- dcast(rankings, name ~ variable,value.var="yrrank")

## remove the names
rankings_wide <- rankings_wide[,-1]
## generate covariate matrix using require "Hmisc"" package ((cormatrix$r) and (cormatr
ix$P) to access output)
cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')

###############
```

Now save off the data to save us up to this point!

```
## Save off a copy in case something happens...

## save the work history for this analysis
#save(AllData, file="C:/GISdata/NewUHI/data/UHI2017.R")

## save the dataframe off as a .csv or .dbf so that it can be shared wi
th non-R'ers
#write.dbf(as.data.frame(AllData), file="C:/GISdata/NewUHI/data/new_All
Data_saved.dbf")
```

A3. Code implemented in R for indicator to indicator correlation (H2)

1. Import all packages as always

```
##############################################################
##### Install needed libraries and packages  ################

#install.packages("spacetime")
require("spacetime")
#install.packages("xts")
require(xts)
## Install all of the needed packages here
# install.packages(c("tidyr", "dplyr", "ggplot2",
#                    "reshape2","foreign", "Hmisc"))

require("tidyr")
library("dplyr")
require("reshape2")
require(foreign)
require(ggplot2)
require(scales)
require(Hmisc)
library(stats)
```

2. Grab that data from the previous workflow to reveal each measurement type we have

```
## bring in the data saved from H1 and prep as before..
file="C:/GISdata/NewUHI/data/new_AllData_saved.dbf"
## read in the data
H1_data <- read.dbf(file, as.is = TRUE)
## make sure the variables are as you expect
# H1_data$type <- factor(H1_data$type,
#                        labels = c("ag", "bare", "built", "DUA", "DUF",
#                                   "DUO", "DUR", "forest", "HIA", "MAG",
#                                   "MAX", "MIC", "MIN", "other","RANGE",
#                                   "Rural","Urban"))
unique(H1_data$type)

##  [1] "Rural"            "Urban"            "DiffUHI"
##  [4] "forest"           "other"            "ag"
##  [7] "urban"            "bare"             "Diff_UrbanOther"
## [10] "Diff_UrbanAg"     "Diff_UrbanForest" "MicroArea%"
## [13] "MAG"              "MIN"              "MAX"
## [16] "RANGE"            "HotArea%"

## Make the variables we need for the summary
H1_data$month <- as.factor(substr(H1_data$date, 6, 7))
H1_data$month <- factor(H1_data$month,
                       levels = c("01", "02", "03", "04", "05", "06",
                                  "07", "08", "09", "10", "11", "12"),
                       labels = c("jan", "feb", "mar", "april", "may",
                                  "jun", "july", "aug", "sept",
                                  "oct", "nov", "dec"))
H1_data$day <- as.factor(substr(H1_data$date, 9, 10))
H1_data$year<- as.factor(substr(H1_data$date, 1, 4))
```

3. Remember, we are doing these test 1 month at a time, so subset 1 month and build the functions to work for that.

4. First thing bring over those functions we will be using. 1) Function to summarize, rank, and generate covariant matrix, and 2) Function to generate lots of cool additional info for the Spearman plots.

```
### Generate function for generating Spearman plots with all included info.
## https://www.r-bloggers.com/spearman-correlation-heat-map-with-correlation-coefficien
ts-and-significance-levels-in-r/

abbreviateSTR <- function(value, prefix){  # format string more concisely
  lst = c()
  for (item in value) {
    if (is.nan(item) || is.na(item)) { # if item is NaN return empty string
      lst <- c(lst, '')
      next
    }
    item <- round(item, 2) # round to two digits
    if (item == 0) { # if rounding results in 0 clarify
      item = '<.01'
    }
    item <- as.character(item)
    item <- sub("(^[0])+", "", item)    # remove leading 0: 0.05 -> .05
    item <- sub("(^-[0])+", "-", item)  # remove leading -0: -0.05 -> -.05
    lst <- c(lst, paste(prefix, item, sep = ""))
  }
  return(lst)
}


## Generate function that will Summarize, rank, and run the SPearmans test
RankTheData <- function(H4_data){  # format string more concisely
## start function
  ## convert "long" data to "wide" using reshape2 package and
  ## calculate the mean in the process
  subsetData_wide <- dcast(H4_data, name ~ newtype, mean) ## is either type or new type
  ## check it out
  #head(subsetData_wide)
  #write.csv(subsetData_wide, file="C:/GISdata/NewUHI/output/Value_data.csv")

  ## reform the data back to "long" format
  subsetData_melted <- melt(subsetData_wide, id.vars = "name")

  ## Create a new attribute to rank the variables for each time step.
  rankings <- subsetData_melted %>%
    group_by(variable) %>%
    mutate(yrrank = row_number(-value))

  ## convert "long" data to "wide" using reshape2 package
  rankings_wide <- dcast(rankings, name ~ variable,value.var="yrrank")
  ## check the rankings
  #write.csv(rankings_wide, file="C:/GISdata/NewUHI/output/r_data.csv")
  #write.csv(rankings_wide, file="C:/GISdata/NewUHI/output/p_data.csv")


  ## remove the names
  rankings_wide <- rankings_wide[,-1]
  ## generate covariate matrix using require "Hmisc"" package ((cormatrix$r) and (corma
trix$P) to access output)
```

```
  cormatrix = rcorr(as.matrix(rankings_wide), type='spearman')
  #print(cormatrix)
  ## save the dataframe off as a .csv or .dbf so that it can be shared with non-R'ers
  #write.csv(as.data.frame(cormatrix), file="C:/GISdata/NewUHI/ind2ind.csv")

  cordata = melt(cormatrix$r)
  cordata$labelr = abbreviateSTR(melt(cormatrix$r)$value, 'r')
  cordata$labelP = abbreviateSTR(melt(cormatrix$P)$value, 'P')
  cordata$label = paste(cordata$labelr, "n",
                        cordata$labelP, sep = "")
  cordata$strike = ""
  cordata$strike[cormatrix$P > 0.05] = "X"

  txtsize <- par('din')[2] / 1  ## change demoninator for txt size
  ggplot(cordata, aes(x=Var1, y=Var2, fill="white")) + geom_tile() +
  #ggplot(cordata, aes(x=Var1, y=Var2, fill=value)) + geom_tile() +
    theme(axis.text.x = element_text(angle=90, hjust=TRUE)) +
    xlab("") + ylab("") +
    geom_text(label=cordata$label, size=txtsize) +
    geom_text(label=cordata$strike, size=txtsize * 4, color="blue", alpha=0.4)
  #return(cormatrix)
## end function
}
```

5. Use RankTheData() function to check for correlation between each set of rankings for each month of 2012. this will answer the First Hypothesis (Jan = Feb) Do the Monthly Rankings agree with each other each month when each indicator is used?
   Do you get the same rankings each month of the year when using each indicator?
   How else to say in easy to understand way?

```
## Only interested in the year 2012 for now.
H2_data <- filter(H1_data, year=="2012")
head(H2_data)
H2_data$newtype <- as.character(H2_data$type)


H3_data <- filter(H2_data, month=="jan")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)

H3_data <- filter(H2_data, month=="feb")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="mar")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="april")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
```

```
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="may")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="jun")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="july")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="aug")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="sept")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="oct")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
H3_data <- filter(H2_data, month=="nov")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
H3_data <- filter(H2_data, month=="dec")
H4_data <- filter(H3_data, newtype=="Diff_UrbanAg"|newtype=="Diff_UrbanOther"|newtype==
"DiffUHI"|
                    newtype=="MAG"|newtype=="HotArea%"|newtype=="MicroArea%")
H4_data$newtype <- as.factor(H4_data$newtype)
RankTheData(H4_data)
```

Make boxplots to highlight the land cover and area LSTs distributions for 2012.

```
## Only interested in the year 2012 for now.
H4_data <- filter(H1_data, year=="2012")

###########################
#### Boxplot of Urban Rural Only
H5_data <- filter(H4_data,type=="Urban"|type=="Rural")
ggplot(H5_data, aes(x=month, y=value, fill=type)) + geom_boxplot() + ylim(0,40) +
```

```r
    ggtitle("All Locations SUHI: Urban and Rural Land Areas (2012)")
###########################
#### Boxplot of Land Covers only
H6_data <- filter(H4_data,type=="urban"|type=="ag"|type=="forest"|type=="other")
ggplot(H6_data, aes(x=month, y=value, fill=type)) + geom_boxplot() + ylim(-10,40) +
    ggtitle("All Locations SUHI: Land Cover Areas (2012)")
###########################
#### Boxplot of SUHI land cover driven indicators of 2012
H7_data <- filter(H4_data,type=="DUA"|type=="DUO"|type=="DUR")
ggplot(H7_data, aes(x=month, y=value, fill=type)) + geom_boxplot() + ylim(-2,5) +
    ggtitle("All Locations: Land Cover Driven Indicators (2012)")
###########################
#### Boxplot of Numeric Stat Summaries (min, max, range, magnitude)
H8_data <- filter(H4_data,type=="MAX"|type=="MIN"|type=="RANGE")
ggplot(H8_data, aes(x=month, y=value, fill=type)) + geom_boxplot() + ylim(-2,42) +
    ggtitle("All Locations SUHI: Summaries for Urban Areas (2012)")
###########################
#### Boxplot of Distribution measures (HIA and MIC)
H9_data <- filter(H4_data,type=="HIA"|type=="MIC"|type=="MAG")
ggplot(H9_data, aes(x=month, y=value, fill=type)) + geom_boxplot() + ylim(0,20) +
    ggtitle("All Locations SUHI: Distribution Summaries for Urban Areas (2012)")
##########################
#### Same as above
ggplot(H9_data, aes(x=month, y=value, fill=type)) + geom_boxplot() + ylim(0,20) +
    ggtitle("All Locations SUHI: Distribution Summaries (2012)")
```

A4. Code implemented in R for Linear model (H3)

1. LM() for change over time at significant levels for all locations. The variables will seem to be Louisville specific because it originally written to accommodate only one location, and then later expanded to accommodate all of them. Will later make changes when defining functions. The difference simply replaces the individual pixels with individual cities. Simple.

```
## grab each pixel data and run 2 regression tests, 1) with all of
## the  months included
## subset to 2012 only
time1<-as.Date("2002-01-01")
time2<-as.Date("2012-12-26")
new_time <- filter(H1_data, date > time1 & date < time2)
lou_data <- filter(new_time, date > time1 & date < time2 &
                     type=="Diff_UrbanAg"|type=="Diff_UrbanOther"|
                     type=="DiffUHI"|type=="HotArea%"|type=="MicroArea%"|type=="MAG")
## tend to the data factors
lou_data$type <- as.factor(lou_data$type)
lou_data$name <- as.factor(lou_data$name)
lou_data$type <- factor(lou_data$type,
                       levels = c("Diff_UrbanAg", "Diff_UrbanOther", "DiffUHI",
                                  "HotArea%", "MicroArea%", "MAG"),
                       labels = c("DUA", "DUO", "DUR", "HIA", "MIC","MAG"))

#summary(lou_data)
#str(lou_data)
#head(lou_data)

## create a list of all pixels and the months to be analyzed
pixels <- unique(lou_data$name)
suhis <- unique(lou_data$type)
months <- unique(lou_data$month)
length(pixels)

## [1] 26

length(suhis)

## [1] 6

length(months)

## [1] 12

## create data.frame() to hold the change data as it comes in
lou_change <- data.frame(matrix(0, ncol = 15, nrow = 12))
names(lou_change) <- c("test_val", "jan", "feb", "mar", "april", "may",
                       "jun", "july", "aug", "sept",
                       "oct", "nov", "dec","name", "suhi")

n = 0
for (pixel in pixels){  ## cycle through each location
#  print(pixel)
  ## Grab only the data that for that pixel
  pixel_data <- filter(lou_data, name==pixel)

  for (suhi in suhis){
```

```
    suhi_data <- filter(pixel_data, type==suhi)
    n = n + 1

    for (j in months){  ## cycle through each month
      #j=1
      n2 = n * 2
      n1 = n2 - 1
      #    print(j) ## name the month
      month_data <- filter(suhi_data, month==j) ## pull out that SUHI's data
      tim <- c(1:length(month_data$value)) ############  THIS IS THE FEW CHANGES
      val <- month_data$value  ############  THIS IS THE FEW CHANGES
      #lou_change[n,j] <- lm(val~tim)$coefficients[2] ## beta-value
      lou_change[n1,j] <- summary(lm(val~tim))$coefficients[2] ## beta-value
      lou_change[n2,j] <- summary(lm(val~tim))$coefficients[8] ## p-value
      lou_change$name[n1] <- pixel
      lou_change$name[n2] <- pixel
      lou_change$test_val[n1] <- "beta"
      lou_change$test_val[n2] <- "pval"
      lou_change$suhi[n1] <- suhi
      lou_change$suhi[n2] <- suhi
    } ## close 3rd loop

  }
    #kept_data <- rbind(kept_data, ktm_change) ## grab the data
}

## check this out for reference of last test
summary(lm(val~tim))
```

Print the model for reference

```
##
## Call:
## lm(formula = val ~ tim)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7.1309 -0.8131  0.1417  1.8955  4.1662
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 14.52430    0.87894  16.525   <2e-16 ***
## tim          0.01234    0.03480   0.355    0.725
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.832 on 41 degrees of freedom
## Multiple R-squared:  0.003056,   Adjusted R-squared:  -0.02126
## F-statistic: 0.1257 on 1 and 41 DF,  p-value: 0.7247
```

```
head(lou_change)
```

```
##   test_val          jan          feb           mar        april
## 1     beta  0.000844642  0.023510544  0.0461446041 -0.003356396
## 2     pval  0.962007742  0.060405158  0.0074846060  0.832376953
## 3     beta -0.001161064 -0.003009231 -0.0041002354 -0.001181710
## 4     pval  0.808820812  0.310051621  0.0765268056  0.767252875
## 5     beta -0.013429536 -0.014362610 -0.0296394905 -0.008426230
## 6     pval  0.166040241  0.327535394  0.0005909365  0.391481291
##             may          jun         july          aug         sept
## 1  0.0137153659  0.026217851  1.127824e-02  0.015860437 -0.001548203
```

```
## 2  0.2060167357  0.011768201  2.828568e-02  0.061638356  0.859421067
## 3 -0.0004595315  0.001798006  3.243540e-03  0.001305083  0.004067680
## 4  0.8883082678  0.557304281  1.568293e-01  0.610767401  0.099353491
## 5 -0.0151146617 -0.011090038 -4.600096e-06 -0.002894554  0.008762724
## 6  0.0661550835  0.142259426  9.992784e-01  0.671750053  0.255485095
##            oct          nov           dec                  name suhi
## 1 0.0039162934 -0.018750187 -0.0197521332 Albany--Schenectady, NY  DUR
## 2 0.7570949297  0.056032951  0.1421881884 Albany--Schenectady, NY  DUR
## 3 0.0050173267  0.002129092 -0.0005774411 Albany--Schenectady, NY  DUO
## 4 0.2639546380  0.267734939  0.8670620103 Albany--Schenectady, NY  DUO
## 5 0.0003810412  0.003822383  0.0040171857 Albany--Schenectady, NY  DUA
## 6 0.9581102394  0.354787470  0.7154715851 Albany--Schenectady, NY  DUA
```

2. Now we simply select the ones that have p-values lower than 0.05 and mark them. That where the significant change lies..

```
#### save p-values off to the "all change data table"
pvals <- lou_change[lou_change$test_val== "pval",]
head(pvals)

##    test_val       jan        feb          mar        april          may
## 2      pval 0.9620077 0.06040516 0.0074846060 0.8323769534 0.206016736
## 4      pval 0.8088208 0.31005162 0.0765268056 0.7672528755 0.888308268
## 6      pval 0.1660402 0.32753539 0.0005909365 0.3914812909 0.066155084
## 8      pval 0.6040084 0.67480290 0.9797565451 0.0002879495 0.005216514
## 10     pval 0.9626739 0.18458747 0.0004887326 0.7491596096 0.324251600
## 12     pval 0.5599902 0.83887127 0.0010088720 0.5588767399 0.737957173
##            jun         july        aug       sept        oct        nov
## 2  0.0117682014 0.0282856770 0.06163836 0.85942107 0.75709493 0.05603295
## 4  0.5573042814 0.1568292804 0.61076740 0.09935349 0.26395464 0.26773494
## 6  0.1422594259 0.9992784274 0.67175005 0.25548509 0.95811024 0.35478747
## 8  0.0003380564 0.0005107186 0.14640751 0.01249958 0.07075597 0.45152875
## 10 0.0167375753 0.0785325241 0.03691197 0.30636813 0.73531817 0.27659342
## 12 0.9815336420 0.6257441843 0.06422276 0.41340524 0.81637595 0.21757023
##          dec                    name suhi
## 2  0.1421882 Albany--Schenectady, NY  DUR
## 4  0.8670620 Albany--Schenectady, NY  DUO
## 6  0.7154716 Albany--Schenectady, NY  DUA
## 8  0.5211228 Albany--Schenectady, NY  MIC
## 10 0.1638665 Albany--Schenectady, NY  MAG
## 12 0.6398247 Albany--Schenectady, NY  HIA

#all_change_data <- rbind(all_change_data, pvals)
#### save p-values off to the "all change data table"
betas <- lou_change[lou_change$test_val== "beta",]
head(pvals)

##    test_val       jan        feb          mar        april          may
## 2      pval 0.9620077 0.06040516 0.0074846060 0.8323769534 0.206016736
## 4      pval 0.8088208 0.31005162 0.0765268056 0.7672528755 0.888308268
## 6      pval 0.1660402 0.32753539 0.0005909365 0.3914812909 0.066155084
## 8      pval 0.6040084 0.67480290 0.9797565451 0.0002879495 0.005216514
## 10     pval 0.9626739 0.18458747 0.0004887326 0.7491596096 0.324251600
## 12     pval 0.5599902 0.83887127 0.0010088720 0.5588767399 0.737957173
##            jun         july        aug       sept        oct        nov
## 2  0.0117682014 0.0282856770 0.06163836 0.85942107 0.75709493 0.05603295
## 4  0.5573042814 0.1568292804 0.61076740 0.09935349 0.26395464 0.26773494
## 6  0.1422594259 0.9992784274 0.67175005 0.25548509 0.95811024 0.35478747
## 8  0.0003380564 0.0005107186 0.14640751 0.01249958 0.07075597 0.45152875
## 10 0.0167375753 0.0785325241 0.03691197 0.30636813 0.73531817 0.27659342
## 12 0.9815336420 0.6257441843 0.06422276 0.41340524 0.81637595 0.21757023
```

```
##           dec                      name suhi
## 2  0.1421882 Albany--Schenectady, NY  DUR
## 4  0.8670620 Albany--Schenectady, NY  DUO
## 6  0.7154716 Albany--Schenectady, NY  DUA
## 8  0.5211228 Albany--Schenectady, NY  MIC
## 10 0.1638665 Albany--Schenectady, NY  MAG
## 12 0.6398247 Albany--Schenectady, NY  HIA

#all_change_data <- rbind(all_change_data, betas)

## have a look at the data
#head(all_change_data) ## should see lst or ndvi as type
#tail(all_change_data) ## should see the other one as type

## save p-values off to dbf if needed
# write.dbf(pvals, file="C:/GISdata/NewUHI/output/pvals_change_00_12.dbf")
# write.csv(pvals, file="C:/GISdata/NewUHI/output/pvals_change_00_12.csv")
# write.dbf(betas, file="C:/GISdata/NewUHI/output/betas_change_00_12.dbf")
# write.csv(betas, file="C:/GISdata/NewUHI/output/betas_change_00_12.csv")

resulttable <- data.frame()
resulttable <- rbind(resulttable,pvals)
resulttable <- cbind(resulttable,betas)
#head(resulttable)

names(resulttable) <- c("test_val","jan","feb","mar","april","may","jun","july", ## 1-8
                        "aug","sept","oct","nov","dec","name", "suhi","test_val", ## 9-
16
                        "jan1", "feb1", "mar1","april1","may1","jun1" ,"july1" , ## 17
- 23
                        "aug1","sept1","oct1","nov1","dec1","name1","suhi")      ## 24
- 30


for (j in c(1:12)){
  nc  <- j + 1 ## jan pval at 2nd column
  nc1 <- nc + 15 ## jan beta at 17th colum
  nc2 <- nc1 + 14 ## jan sig-beta at 31st

  for (i in 1:length(resulttable$test_val)){
    val <- resulttable[i,nc] ## grab the pval
    bet <- resulttable[i,nc1] ## grab the beta
    #print(i)
    if (val <= 0.05) { ## evaluate the pval
      #print("yes")
      ## and record the beta with an (*) for ID purposes..
      resulttable[i,nc2] <- paste(substr(bet,1,4), "*", sep=" ") ## label as sig!
    } else {
      #print("no")
      resulttable[i,nc2] <- paste(substr(bet,1,4), "", sep="")} ## label with value
  }
}
##############  end the down direction
head(resulttable)

##    test_val       jan        feb          mar        april          may
## 2      pval 0.9620077 0.06040516 0.0074846060 0.8323769534 0.206016736
## 4      pval 0.8088208 0.31005162 0.0765268056 0.7672528755 0.888308268
## 6      pval 0.1660402 0.32753539 0.0005909365 0.3914812909 0.066155084
## 8      pval 0.6040084 0.67480290 0.9797565451 0.0002879495 0.005216514
## 10     pval 0.9626739 0.18458747 0.0004887326 0.7491596096 0.324251600
```

```
## 12     pval 0.5599902 0.83887127 0.0010088720 0.5588767399 0.737957173

## 12 -5.306835584  0.159796047 -0.021273458  0.0146329213  0.0006842445
##            july1         aug1        sept1          oct1         nov1
## 2   1.127824e-02  0.015860437 -0.001548203  0.0039162934 -0.018750187
## 4   3.243540e-03  0.001305083  0.004067680  0.0050173267  0.002129092
## 6  -4.600096e-06 -0.002894554  0.008762724  0.0003810412  0.003822383
## 8  -4.908760e-01 -0.243690297 -0.355080032 -0.3135967788 -0.086102843
## 10  3.086399e-02 -0.035737139 -0.012861875 -0.0057338936  0.018294425
## 12 -7.918499e-03  0.074839031 -0.603068191  0.0176384902  0.059109808
##             dec1                  name1 suhi.1  V31  V32    V33    V34
## 2  -0.0197521332 Albany--Schenectady, NY    DUR 0.00 0.02 0.04 *   -0.0
## 4  -0.0005774411 Albany--Schenectady, NY    DUO -0.0 -0.0   -0.0   -0.0
## 6   0.0040171857 Albany--Schenectady, NY    DUA -0.0 -0.0 -0.0 *   -0.0
## 8  -0.0346920436 Albany--Schenectady, NY    MIC 0.02 0.04   -0.0 -0.5 *
## 10  0.0408819088 Albany--Schenectady, NY    MAG -0.0 -0.0 -0.0 *   0.00
## 12 -2.1203214479 Albany--Schenectady, NY    HIA 1.01 -5.3 0.15 *   -0.0
##       V35    V36    V37    V38    V39  V40  V41  V42
## 2    0.01 0.02 * 0.01 *   0.01   -0.0 0.00 -0.0 -0.0
## 4    -0.0   0.00   0.00   0.00   0.00 0.00 0.00 -0.0
## 6    -0.0   -0.0   -4.6   -0.0   0.00 0.00 0.00 0.00
## 8  -0.5 * -0.5 * -0.4 *   -0.2 -0.3 * -0.3 -0.0 -0.0
## 10   0.01 0.04 *   0.03 -0.0 *   -0.0 -0.0 0.01 0.04
## 12   0.01   0.00   -0.0   0.07   -0.6 0.01 0.05 -2.1

#write.csv(resulttable, file="C:/GISdata/NewUHI/output/rework_pvals_change_00_12.csv")

rm(nc,nc1,nc2,val,bet,i,j)
```

A5. Code implemented for Louisville specific analysis

1. Pixel Based Measurements. Import Urban LSTs for Louisville pixels only

```
##########################################################################################
##############
##########################################
####  Start Louisville pixel data  ################

### set directory by appending above pathway. Then create
### vector contining the names of each of the visible files in the working directory.
setwd(paste(path, "/LouisvillePixelTables/", sep=""))
files <- list.files(pattern = "\\.dbf$")

## Create a new vector to retain combined urban data
pix_data <- data.frame()
##### set the looping structure to cycle through each file returned from dir()
for (file in files) {
  ##  extract day of year (DOY) from file name
  DOY <- substr(file, 2, 8)
  ## read in the data
  table_data <- read.dbf(file, as.is = TRUE)
  if (nrow(table_data) != 0){
    ## grab the variables we need
  new_data <- table_data[,c(1:3)]
  ## create new column called 'year' to store the label (factor?) we created above
  new_data[,"DOY"] <- DOY
  ## append this newly created data frame to the master we created earlier
  pix_data <- rbind(pix_data, new_data)
  } else {}
}

#head(urban_data)
# clean house (just in case)
rm(DOY)
rm(table_data)
rm(new_data)
rm(file)
rm(files)

# rename variables to your liking
names(pix_data) <- c("value", "ID", "raster", "doy")
## recognize the date
pix_data[,"date"] <- as.Date(pix_data$doy, format="%Y%j")

pix_data$month <- as.factor(substr(pix_data$date, 6, 7))
pix_data$month <- factor(pix_data$month,
                       levels = c("01", "02", "03", "04", "05", "06",
                                  "07", "08", "09", "10", "11", "12"),
                       labels = c("jan", "feb", "mar", "april", "may",
                                  "jun", "july", "aug", "sept",
                                  "oct", "nov", "dec"))
pix_data$day <- as.factor(substr(pix_data$date, 9, 10))
pix_data$year<- as.factor(substr(pix_data$date, 1, 4))
pix_data$lst <- pix_data$value * 0.02 - 273.15
head(pix_data)


## create a list of all pixels and the months to be analyzed
```

```
pixels <- unique(pix_data$ID)
months <- unique(pix_data$month)
length(pixels)
length(months)

## create data.frame() to hold the change data as it comes in
pixel_change <- data.frame(matrix(0, ncol = 14, nrow = 12))
names(pixel_change) <- c("test_val", "jan", "feb", "mar", "april", "may",
                         "jun", "july", "aug", "sept",
                         "oct", "nov", "dec","ID")
#kept_data <- data.frame()
n = 0
for (pixel in pixels){  ## cycle through each location
#  print(pixel)
  ## Grab only the data that for that pixel
  pixel_data <- filter(pix_data, ID==pixel)

  n = n + 1
  for (j in months){  ## cycle through each month
    #j=1
    n2 = n * 2
    n1 = n2 - 1
#     print(j) ## name the month
    month_data <- filter(pixel_data, month==j) ## pull out that SUHI's data
    tim <- c(1:length(month_data$lst))
    val <- month_data$lst
    #lou_change[n,j] <- lm(val~tim)$coefficients[2] ## beta-value
    pixel_change[n1,j] <- summary(lm(val~tim))$coefficients[2] ## beta-value
    pixel_change[n2,j] <- summary(lm(val~tim))$coefficients[8] ## p-value
    pixel_change$ID[n1] <- pixel
    pixel_change$ID[n2] <- pixel
    pixel_change$test_val[n1] <- "beta"
    pixel_change$test_val[n2] <- "pval"
  } ## close 3rd loop
    #kept_data <- rbind(kept_data, ktm_change) ## grab the data
}

rm(j,i,n,n1,n2,pixel,pixels, tim, val)

##########################################################################
################################################ End Louisville pixel data
```

3. Visualize for change over time (come back to this after addressing each of the indicators) for Louisville.

```
## resubset data to get all years
subsetData <- filter(H1_data, name == "Louisville/Jefferson County, KY--IN")
## Grab the variables we need for the summary
subsetData$Month <- as.factor(substr(subsetData$date, 6, 7))
subsetData$Month <- factor(subsetData$Month,
                         levels = c("01", "02", "03", "04", "05", "06",
                                    "07", "08", "09", "10", "11", "12"),
                         labels = c("jan", "feb", "mar", "april", "may",
                                    "jun", "july", "aug", "sept",
                                    "oct", "nov", "dec"))
subsetData$Day <- as.factor(substr(subsetData$date, 9, 10))
subsetData$Year<- as.factor(substr(subsetData$date, 1, 4))
subsetData$Temp<- subsetData$value

## vectorize as df for plotting
df <- filter(subsetData, type=="DUR"|type=="HIA")
```

```
##################################################
## By month of Measurment and Type across each year to highlight seasonality
cbPalette <- c( "#009E73", "#FF9900", "#0072B2", "#D55E00", "#CC79A7")
cbPalette <- c( "#009E73", "#FF9900")

ggplot(df,aes(Month,value,fill=type)) +
  geom_bar(stat = "identity", position=position_dodge()) +
  ggtitle("All Locations: DUR SUHI by Month by Type (2000 to 2015)") +
  theme_grey(base_size=15) + theme(legend.position=c(.5,.125), axis.title.y=element_bla
nk()) +
  scale_fill_manual(values=cbPalette) + ylim(0, 30)

###################################################
## Complete summary of year and month measurements.  All Data repersented!
DURData <- subset(df, type=="DUR", select=c(type, Year, Month, value))
HIAData <- subset(df, type=="HIA", select=c(type, Year, Month, value))
medDevData <- subset(df, type=="DiffUHI", select=c(type, Year, Month, value))

ggplot(df,aes(Year,value,colour=type)) + ylim(0, 7) +
  #geom_point(data=devData,size=I(2),alpha=I(0.6)) +
  #geom_line(data=meanDevData,size=I(1.5),alpha=I(0.6)) +
  geom_line(data=DURData,size=I(1.5),alpha=I(0.4)) +
  theme_grey(base_size=15) +
  theme(legend.title = element_blank(), legend.position=c(.5,.1), axis.title.y=element_
blank(),axis.text.x=element_blank()) +
  ggtitle("Louisville, Kentucky: DUR SUHI by Month (2000 to 2015)") + facet_grid(. ~ Mo
nth) +
  xlab("Years: 2000 to 2015")

ggplot(df,aes(Year,value,colour=type)) + ylim(0, 25) +
  #geom_point(data=devData,size=I(2),alpha=I(0.6)) +
  #geom_line(data=meanDevData,size=I(1.5),alpha=I(0.6)) +
  geom_line(data=DURData,size=I(1.5),alpha=I(0.4)) +
  theme_grey(base_size=15) +
  #theme(legend.title = element_blank(), legend.position=c(.5,.10), #axis.title.y=eleme
nt_blank(),axis.text.x=element_blank()) +
  ggtitle("Louisville, Kentucky: HIA SUHI by Month (2000 to 2015)") + facet_grid(. ~ Mo
nth) +
  xlab("Years: 2000 to 2015")
```

CURRICULUM VITA

NAME & RESEARCH INTERESTS

Jeremy Sandifer
Sandulan82@gmail.com

Integrated modeling of spatial processes; Human-environmental interactions; Geospatial technology (GIS, remote sensing, open source programming); Climate and environmental reconstructions; Land systems change dynamics; Drone tech; Geospatial education

EDUCATION

2013 – University of Louisville – *Louisville, Kentucky*

B.S. in Applied Geography *summa cum laude* , concentration in environmental analysis

Committee: Dr. Keith Mountain (advisor), Dr. Christopher Day

Thesis: *Characterization of the Urban Heat Island Intensity Within Jefferson County, Kentucky*

RESEARCH EXPERIENCE

2013-2015 – Graduate Research Assistantship, University of Louisville – *Louisville, Kentucky*. Dr. Keith Mountain and Dr.Andrea Gaughan

- ArcMap/ENVI/R/Python based spatio-temporal applied geographic research
- Developed computer based programs for statistical interrogation of spatial data
- Designed lab materials for instructing non-GIS users to implement derived program/scripts

2013-2014 – Field Research Assistant, Institutional and Faculty Developmental Grant, University of Louisville- *Louisville, Kentucky*. Dr. Christopher Day (PI)

- Assisted with dendrochronology-based analysis of historic environmental conditions in the Red River Gorge Geologic Area of Daniel Boone National State Park.
- Coordinated field work, specimen preparation and data processing with PI
- Co-authored publication in peer-review research journal on variance modeling of data

TEACHING EXPERIENCE

2015-2017 – Course Instruction, Kentucky State University – *Frankfort, Kentucky*

- *Introduction to Geographic Information Systems and Spatial Analysis (MES490)* focused on geospatial concepts and spatial data analysis with ESRI ArcMap and development of lab-based exercises for skill development

2013-2015 – Graduate Teaching, University of Louisville – *Louisville, Kentucky*

- Sole developer and lecturer of cyber-based course for, *Introduction to Geosciences and Earth Systems (GEOS200)*, focused on conceptualization of Earth processes and development of laboratory methodologies
- TA and lab instructor for *Remote Sensing of the Environment (GEOG590)*

2011-2013 – STEM Undergraduate Teaching, University of Louisville – *Louisville, Kentucky*

- UTA and lab instructor, *Introduction to Geosciences and Earth Systems (GEOS200)*, focus on student services and outcomes
- Tutor service provisions, Geography/Geosciences-based courses

2008-2010 – Training Coordinator, JPD Satellite Systems – *New Orleans, Louisiana*

- Technical proficiency testing and certification/compliance for field technicians
- Developed and implemented continuing technical education programs

REFEREED ARTICLES

Day, C.A. and **J. Sandifer.** 2015. An Annual Streamflow Reconstruction of the Red River, Kentucky Using a White Pine (*Pinus Strobus*) Chronology. Journal of Geography and Earth Sciences Vol 3(1): pp. 1-14 http://jgesnet.com/journals/jges/Vol_3_No_1_June_2015/1.pdf

Sandifer, J. 2013. Influence of Surface Land Cover on the Urban Heat Island Intensity within Metropolitan Jefferson County, Kentucky. Papers in Applied Geography Vol 36, pp. 323-331. http://applied.geog.kent.edu/AGCPapers/2013/P323-331/index.html

Sandifer, J. and Mountain, K. 2013. Remote Sensing and Identifying the Urban Heat Island from Space: An Overview. Sustain Magazine Vol 29. The Kentucky Institute for the Environment and Sustainable Development. University of Louisville, Louisville, KY. pp. 9-12. http://louisville.edu/kiesd/sustain-magazine/SUSTAIN-29rev.pdf

PROFESSIONAL SERVICE AND OTHER ACTIVITIES

- Association of American Geographers (AAG)
- Kentucky Academy of Sciences (KAS) and Kentucky Geographic Alliance (KGA)
- Kentucky Association of Mapping Professionals (KAMP):
    - 2012-2017 Education Committee; Demetrio Zourarakis (advisor)

CONTRIBUTED PRESENTATIONS / POSTERS AT CONFERENCES

- 2017 Eastern Kentucky Small farm Conference *Hindman, Kentucky* Integrated use of GIS in small farm setting
- 2017 Kentucky Fruit and Vegetable Growers Annual Meeting, *Lexington, Kentucky*, Jeremy Sandifer Drones, Satellite Data, and Geographic

Information Systems.

- 2016 Kentucky Academy of Sciences - *Louisville, Kentucky*, Assessment of Invasive Species in Forest Areas
- 2016 HBCU-UP/CREST National Science Foundation, Annual PI Conference, *Washington D.C.* Promoting STEM Education at Kentucky State University: Summer Apprenticeship Program
- 2016 KAMP GIS Conference – *Covington, Kentucky.* Spatial Distribution of Sustainable Agriculture Practices in Kentucky
- 2015 Association of American Geographers (AAG) Annual Conference – *Chicago, Illinois,* Spatio-temporal variation of the Surface Urban Heat Island 2000 to 2010.
- 2015 Kentucky Academy of Sciences - *Northern Heights, Kentucky*, Correlations between Land Cover and Socioeconomic Conditions using Satellite Imagery Data in Eastern Kentucky. Cynthia Rice, Andrew Gott, and Ken Bates
- 2015 Annual Conference- *Owensboro, Kentucky,* Parcel Data and its Value to Farmers. Cynthia Rice 2014 National Association of Regional Councils (NARC) Annual Conference- *Louisville, Kentucky*. Solving Urban Heat Island Challenges. Dr. Keith Mountain and Erin Thompson
- 2014 KAMP GIS Conference, Kentucky Association of Mapping Professionals (KAMP). Moderating: Hot Topics Panel: Kentucky K-12 GeoMentoring; State of the Practice
- 2013 KAMP GIS Conference – *Louisville, Kentucky.* Land Cover and Variation of Urban Heat Island Intensity
- 2013 Mid-Atlantic Association of American Geographers (MAAAG) Conference of Applied Geography- *Annapolis, Maryland*. The Impact of Land Cover on the Dist. of the UHI Effect