A Spatial Statistical Framework for Evaluating Landscape Pattern and Its Impacts on the Urban Thermal Environment

by

Chao Fan

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved June 2016 by the
Graduate Supervisory Committee:

Soe W. Myint, Chair
Wenwen Li
Sergio J. Rey

ARIZONA STATE UNIVERSITY

August 2016

ABSTRACT

Urban growth, from regional sprawl to global urbanization, is the most rapid, drastic, and irreversible form of human modification to the natural environment. Extensive land cover modifications during urban growth have altered the local energy balance, causing the city warmer than its surrounding rural environment, a phenomenon known as an urban heat island (UHI). How are the seasonal and diurnal surface temperatures related to the land surface characteristics, and what land cover types and/or patterns are desirable for ameliorating climate in a fast growing desert city? This dissertation scrutinizes these questions and seeks to address them using a combination of satellite remote sensing, geographical information science, and spatial statistical modeling techniques.

This dissertation includes two main parts. The first part proposes to employ the continuous, pixel-based landscape gradient models in comparison to the discrete, patch-based mosaic models and evaluates model efficiency in two empirical contexts: urban landscape pattern mapping and land cover dynamics monitoring. The second part formalizes a novel statistical model called spatially filtered ridge regression (SFRR) that ensures accurate and stable statistical estimation despite the existence of multicollinearity and the inherent spatial effect.

Results highlight the strong potential of local indicators of spatial dependence in landscape pattern mapping across various geographical scales. This is based on evidence from a sequence of exploratory comparative analyses and a time series study of land cover dynamics over Phoenix, AZ. The newly proposed SFRR method is capable of

producing reliable estimates when analyzing statistical relationships involving geographic data and highly correlated predictor variables. An empirical application of the SFRR over Phoenix suggests that urban cooling can be achieved not only by altering the land cover abundance, but also by optimizing the spatial arrangements of urban land cover features. Considering the limited water supply, rapid urban expansion, and the continuously warming climate, judicious design and planning of urban land cover features is of increasing importance for conserving resources and enhancing quality of life.

Dedicated to my parents – Fan Ying and Sun Ye

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my advisor and committee chair, Professor Soe Myint. I have been amazingly fortunate to have an advisor who gave me the freedom to explore ideas on my own, and at the same time advice to recover when my steps faltered. Over the past six years, his consistent support have helped me overcome many difficulties in my academic development. I am truly grateful for his mentorship.

Second, I would like to thank my committee members, Professors Wenwen Li and Sergio Rey, for their invaluable advice and guidance throughout the process of planning, performing, and drafting my research for this dissertation. Their time, instructions, and suggestions are greatly appreciated.

During my PhD study, I received countless support, assistance, and encouragements from many friends and colleagues at Arizona State University. I was lucky to make many great friends, including Baojuan Zheng, Yiying Xiong, Yujia Zhang, and Xiaoxiao Li. I also would like to mention current and former members of the Remote Sensing and Geoinformatics Lab – Shai Kaplan, Chris Galletti, John Connors, Karina Benessaiah, and Jordan Smith for providing a positive working environment. Finally, I give sincere appreciation to Professors Anthony Brazel and Elizabeth Wentz, who provided suggestions and insight numerous times throughout my graduate studies.

Funding for this dissertation came from the National Aeronautics and Space Administration (NASA) under grant number NNX12AM88G. I also received financial support from the Anthony Brazel Research Award, Completion Fellowship from the

Graduate College at Arizona State University, and the Central Arizona-Phoenix Long Term Ecological Research (NSF grant number DEB-1026865).

Last but not least, I want to give my deepest gratitude to some of the most important people in my life. To my parents, Fan Ying and Sun Ye, for always encouraging and believing in me despite being 6,138 miles away. To my auntie and uncle, Rona Fan and Lin Liu, for providing immense love and support while I pursue my PhD degree in a foreign country. Finally, to my husband, Fengyu Wang – you have been a constant source of love, care, patience, and inspiration that carried me through some of the most difficult periods of my graduate studies. We just got married in May and I look forward to our journey of life together.

TABLE OF CONTENTS

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER 1

INTRODUCTION

1.1 Background

Urbanization is a complicated natural and social process involving increased proportion of population, spatially extended urban land cover, and dramatically growing economic activities (Ma et al., 2012). Occupying only 3% of the Earth's surface, urban areas are holding 54% of the current world population, which is projected to reach 66% by 2050 (UN, 2014). Rapid urban sprawl has caused fundamental changes to the Earth's socio-ecological systems well beyond the city boundaries (Wu et al., 2011). With roughly 3% of the Earth's landmass, urban areas are responsible for 60% of total residential water consumption, more than 78% of carbon release, and about 76% of wood mainly used for industrial purposes (Brown, 2001; Grimm et al., 2008). Some worldwide environmental issues that are triggered or aggravated by urban sprawl include loss of biodiversity, climate change, and degraded water and air quality (Seto and Fragkias, 2005). Land modification in the urbanization process to sustain increasing urban population has also resulted in other types of change to the Earth's environment (Grimm et al., 2008).

Among the many negative environmental impacts of urbanization is the raised temperatures in the densely built urban core with respect to its rural environment—a phenomenon commonly referred to as the urban heat island (UHI) effect. Due to the loss of vegetation and soil in the urban areas, much of the solar energy reaching the surface is absorbed and retained in the urban materials instead of being evaporated, leading to elevated surface temperatures. Urban geometry and heat emitted from anthropogenic activities contribute additional warmth to the urban environment.

The UHI has various impacts on the urban ecosystems, energy use, and residents' quality of life. Intensified UHI can lead to increased demands for water and energy consumption (Guhathakurta and Gober, 2007), elevated concentration of air pollutants (Lai and Cheng, 2009; Sarrat et al., 2006), and degraded water quality (Arnold Jr and Gibbons, 1996). More threateningly, the heat released from the urban infrastructure at night can increase the duration and magnitude of heat waves, causing a heightened risk of mortality and other heat-related health problems (Clarke, 1972; Whitman et al., 1997). Since its first discovery in 1818, the UHI has been a primary focus in climatology and urban ecology (Arnfield, 2003; Howard, 1833). There is an ever-increasing number of studies focusing on various aspects of the UHI, including its formation, development, as well as mitigation strategies to counter it (Jauregui, 1997; Oke, 1982; Oke, 2002; Onishi et al., 2010; Rizwan et al., 2008).

Among the many factors that affect the form and magnitude of the UHI, human-induced land cover change is an important and decisive one. Land cover modifications during the urbanization fundamentally alter the radiation and energy balance of the city and its surrounding area, ultimately leading to a heat island. Of particular relevance for the surface UHI is the composition of a landscape, i.e., the abundance and variety of land cover types. A vast amount of research has focused on the UHI variation in response to changes in the land composition, especially vegetation and built-up areas (Essa et al., 2013; Kaufmann et al., 2003; Sandholt et al., 2002; Weng et al., 2004). For example, Middel et al. (2012) examine the urban surface energy balance using the Local-Scale Urban Meteorological Parameterization Scheme (LUMPS) and found that urban vegetation can lower surface temperatures through reducing storage heat flux density.

Another study corroborates the positive impacts of green vegetation on heat island mitigation using a multiple linear regression analysis (Yuan and Bauer, 2007).

Expansion of urban green space comes with a cost. Water is an inevitable and critical limiting factor that impedes wide practice of this strategy. The cost-benefit trade-offs of green space expansion are important considerations for urban designers and land managers. A recent study suggests that while urban vegetation produces a cooling effect at night, it results in a ~20% increase in the outdoor water use (Gober et al., 2012). Another study looks into the trade-offs between nighttime temperatures and irrigation and determined that there exists a threshold of the water-to-temperature ratio beyond which increasing the frequency and intensity of irrigation has limited impact on the nighttime temperatures (Gober et al., 2009).

If increasing the area of vegetation is not possible for every city, can changing the spatial configuration of vegetation patches help cool down the city? In their study, Yokohari et al. (1997) show that the shape, size, and segmentation level of paddy fields can affect both the surface and air temperatures based on a case study at the urban fringe of Tokyo, Japan. More recently, a proliferation of research has considered the spatial configuration of urban land cover features (e.g., vegetation, impervious surfaces, soil) in relation to urban warming, and suggested that compact, large, and contiguous green spaces are desirable for lowering the surface temperatures (Cao et al., 2010; Zhang et al., 2009).

Of interest in this research is the associations between a city's thermal environment (surface temperature in particular) and its various land cover patterns (land composition and spatial configuration) at varying geographical scales. Fundamental to

this issue is the development of analytical tools for effective and accurate quantification of landscape pattern. Most if not all of the existing studies resort to the patch mosaic paradigm, which assumes the landscape as a mosaic of discrete patches separated by edges. Simple to understand and easy to implement, the patch mosaic model gains wide recognition and adoption among researchers from a variety of disciplines, notably after the release of the free software package FRAGSTATS (McGarigal and Marks, 1995). In spite of the conceptual simplicity and software accessibility, the patch mosaic model is now facing serious challenges in the landscape ecology community, primarily with regard to the representation of the inherently continuous landscape with discrete patch-based indices (McGarigal et al., 2005; Pearson, 2002; Southworth et al., 2004).

In an effort to understand the land cover impacts on the UHI, selection of statistical tools is of critical importance. The OLS is arguably the most widely employed statistical model and has been utilized to explore the UHI-land cover relations for decades. Two major issues arise, however, when conducting the OLS without considering its basic assumptions. First, the observations of interest (e.g., land surface temperature) are almost never independent and the associations among observations depend to a great extent on their geographical locations. Because of the spatial dependence, the OLS estimates are no longer consistent and/or unbiased. Second, the predictor variables in a multiple linear regression model are usually highly correlated (aka multicollinearity). Although moderate correlation does not cause much concern for model fitting and prediction, contribution of some predictor to the response variable is inevitably affected by the inclusion (and omission) of a correlated predictor, ultimately leading to inaccurate parameter estimates.

1.2 Research Objectives

Given the challenges to model the associations between UHI and land cover patterns, this dissertation has four objectives. The first objective is to develop tools that enable continuous representation of landscape structure and compare the utility of this method with the discrete patch-based paradigm. The second objective is to evaluate the usefulness of the continuous indices in monitoring long-term land cover dynamics in a rapidly expanding city. Coupling continuous models with thermal satellite imagery, the third objective is to examine the seasonal and diurnal surface temperatures in relation to the abundance and spatial arrangement of green vegetation in a desert city. The variability of this association at different geographical scales will also be explored. Last but not least, the fourth objective is to identify the impacts of spatial dependence and multicollinearity on model estimation and prediction and to address them from a spatial econometric perspective.

1.3 Outline of Dissertation

This dissertation is composed of six chapters: an introductory chapter, four chapters each serves as a first-authored research article, and a concluding chapter.

Chapter 2 compares the usefulness of continuous models with the traditional patch mosaic model in characterizing landscape structure over central Phoenix. Two local indicators of spatial autocorrelation—Getis-Ord *G* and local Moran's *I*—are employed as measures of land composition and spatial configuration of land cover features, respectively. Descriptive statistics are computed for both the discrete and continuous models for each land use type in Phoenix. This chapter was published in *Landscape and Urban Planning* in February 2014.

Chapter 3 integrates image time series, continuous spatial indices, and non-parametric regression into a spatiotemporal study of land cover dynamics over the Phoenix metropolitan area from 1991 to 2010. Time series maps of Getis-Ord *G* and local Moran's *I* are created for two important land cover types in the region: vegetation and built-up areas. The modified Mann-Kendall test is performed to detect and describe the monotonic trends in the quantity and spatial arrangement of these two land cover types. This chapter was submitted to *International Journal of Applied Earth Observation and Geoinformation* in April, 2016.

Using the method developed in Chapter 2, Chapter 4 examines the thermal pattern over central Phoenix in relation to the spatial arrangement of green vegetation in this region. Satellite images are utilized in conjunction with local spatial autocorrelation indices to quantify the spatial configuration of vegetation cover and evaluate its variable influences on seasonal and diurnal surface temperatures. This chapter was published in *Progress in Physical Geography* in March 2015.

Chapter 5 develops a methodological framework for estimating the land cover impacts on the summer daytime surface temperatures over central Phoenix. Spatial autoregressive models and ridge regression are integrated to address the spatial dependence and multicollinearity in the regression model. A combination of Monte Carlo simulation and an empirical study over Phoenix are carried out and the performance of the proposed model is compared to OLS, spatial regression models, and ridge regression model. This chapter was accepted to *Transactions in GIS* in June 2016.

Chapter 6 summarizes the main achievements of this dissertation and suggests future research directions.

CHAPTER 2

A COMPARISON OF SPATIAL AUTOCORRELATION INDICES AND LANDSCAPE METRICS IN URBAN LANDSCAPE MAPPING

2.1 Introduction

One of the critical topics that landscape ecology addresses is the interaction between landscape patterns and the underlying ecological processes (Turner et al., 2003; Turner, 1990; Uuemaa et al., 2009; Wu, 2008; Wu & Hobbs, 2002). Quantitative characterization of spatial pattern is a crucial step in determining this interaction (Buyantuyev & Wu, 2010; Luck & Wu, 2002). Remote sensing coupled with landscape pattern analysis provides a framework within which landscape structure can be quantified and studied. However, limitations of traditional land cover classifications are universal and present a challenge for ecologists and urban planners in accurately and realistically representing landscape patterns (Baatz & Schape, 1999; Campbell, 2007; Congalton, 1991; Foody, 1996; Ju et al., 2003; Turner et al., 2001).

While most current spatial pattern analysis programs adopt a patch mosaic paradigm where discrete patches are generated through a classification process, concerns have been growing regarding the limitations of discrete representation. One such limitation is that errors associated with land cover classification are pervasive and may result in less accurate and sometimes even misleading scientific results (Shao & Wu, 2008). Factors that have the potential to introduce errors into the analysis include the levels of detail of remote sensing data (e.g., spatial resolution, spectral resolution, radiometric resolution, etc.), rectification and registration procedures, locational errors, temporal variations, classification methods, and sampling schemes as well as lack of local

area knowledge and expertise of the analyst. Furthermore, the uncertainty associated with classification can compromise the reliability of landscape metrics derived from the thematic maps. Classification accuracy, therefore, is a crucial part in determining the overall quality of the final results. Recent studies found that the accuracy of landscape metrics is strongly linked to the classification accuracy of the associated maps (Li & Wu, 2004; Shao et al., 2001). However, it is well acknowledged that it is difficult to achieve reasonably high classification accuracy in urban landscapes that are dominated by complex boundaries and substantial spectral details. This makes it both time consuming and expertise demanding to obtain true landscape heterogeneity that can then be linked to the underlying ecological processes. It also poses a significant challenge for landscape ecologists in assessing the urban ecological consequences of urban sprawl.

Another limitation of traditional classification paradigm is that landscapes are assumed to be a collection of homogeneous units separated by clearly identified boundaries (Pearson, 2002). Such representation is more likely to be useful for landscapes where features can be distinctly identified, such as large-scale agriculture, than those where boundaries of the patches are difficult to determine, such as urban landscapes (McIntyre & Barrett, 1992). The process of assigning spectrally similar pixels to homogeneous geometric units also eliminates much of the within-class variability such that fine-scale features cannot be properly characterized (McGarigal & Cushman, 2005).

In addition, the all-as-one-class strategy allows only the binary evaluation of land cover conversions, thereby failing to capture partial changes or subtle land cover alterations that are equally important (Foody & Boyd, 1999; Lambin, 1997; Nepstad et al., 1999). This is of major concern for quantitative research on urban landscape structure

dynamics, which provides an important basis for the study of urban ecosystems. The discrete representation of continuous variables is also a major flaw for environmental studies wherein degradations or rehabilitations are of particular interest (Foody, 2001).

Finally, most pixel-by-pixel classifiers focus primarily on the spectral dimension of remotely sensed data. However, the spatial context of the data also includes useful information that has not been well exploited (Cihlar, 2000; De Jong & Burrough, 1995; Read & Lam, 2002). The spatial arrangement provides interpretations of features with greater detail and complexity which may be treated as an additional information source (Herold et al, 2003; Wulder & Boots, 1998). It is recommended that employing measures containing spatial and textural information via the analysis of brightness values be applied as a methodology of effectively characterizing spatial patterns (Musick & Grover, 1991). Efforts have been made to test the usefulness of spatial context based measures in characterizing landscape patterns. Pearson (2002) examined the utility of a local measure of spatial autocorrelation (Geary's *C*) in modeling structure in northern Australia savanna landscapes where features cannot be clearly identified and thus entity-based modeling is inappropriate. LeDrew et al. (2004) demonstrated the usefulness of another spatial autocorrelation indicator (Getis statistic) in detecting land cover modifications over time. Southworth et al. (2004) examined the potential of a continuous analysis using the local indicator of spatial association (LISA) in comparison to a discrete standard analysis on forest modeling and fragmentation. It was suggested that while continuous data products may encounter limitations regarding development and interpretation, they provide guidelines for future work on fragmentation analyses. McGarigal and Cushman (2005) discussed the limitations associated with the widely adopted patch mosaic model, and

introduced the landscape gradient model which is based on continuous heterogeneity rather than discrete representation, with efforts made in describing the utility of surface metrics on landscape ecological applications (McGarigal et al, 2009).

Awareness has been growing among landscape ecologists that the gradient or continuous model can provide accurate representation of landscape heterogeneity. Nevertheless, the application of such models remains very rare and there has been no direct comparison of the utility of continuous models and discrete patch mosaic paradigm. It is acknowledged that spatial autocorrelation indices can be used to measure the spatial dependency of features in space that implies landscape heterogeneity. We believe that the applications of such indices can advance our understandings on urban landscape structure and the associated ecological processes from a gradient perspective. The purpose of this study was to examine the utility of local spatial autocorrelation in characterizing landscape fragmentation in comparison to the analysis of landscape metrics. One of the major contributions of this study to the landscape ecology community is that instead of generating detailed land cover classes from fine-resolution commercial satellite data that are expensive, time consuming, labor intensive and expertise demanding, we employ medium-resolution satellite data (e.g., Landsat TM) that are available worldwide at no cost to characterize landscape fragmentation over large urban area effectively and efficiently. We compared FRAGSTATS metrics computed from a high-resolution (QuickBird) land cover classification map to the use of continuous indices derived from medium-resolution satellite data (Landsat TM), which are readily attainable. Two local spatial autocorrelation measures, Getis statistic (Getis & Ord, 1992) and local Moran's *I* (Anselin, 1995) were computed and their relationships with widely

adopted landscape metrics were evaluated statistically. The research questions we addressed are: 1) Can the spatial arrangement of urban land cover features be addressed within a continuous framework? 2) How does the use of continuous indices on medium-resolution data compare to class- and landscape-level metrics derived from high-resolution imagery in characterizing urban landscape fragmentation? 3) Do spatial autocorrelation indices provide accurate and reliable information in terms of urban landscape fragmentation for major land use types? 4) What are the strengths and challenges of utilizing spatial autocorrelation indices in landscape ecological applications?

2.2 Methods

2.2.1 Urban Landscape and Fragmentation

Urban landscapes are a mixture of natural and anthropogenic components. In the last few decades, urban sprawl has occurred at an unprecedented rate in U.S. cities and across the globe. A major concern of this rapid urbanization is its profound impacts on urban ecosystems and biodiversity due to fragmentation. This has prompted a growing interest among researchers to investigate the structure of urban landscapes and the ecological consequences of extensive urban sprawl.

The patch mosaic model gained recognition among landscape ecologists for quantifying landscape structure owing to its effectiveness evidenced in previous work (Buyantuyev & Wu, 2010; Griffith et al., 2000; Liu & Weng, 2008; Luck & Wu, 2002; Southworth et al., 2002). However, the presence of substantial spectral confusion and complex shapes of urban features has raised concerns regarding the utility of discrete

metrics in characterizing urban landscape patterns. The complexity residing in urban landscapes reduces the accuracy of land cover classification, which is inherently subjective, and causes problems in landscape pattern analysis that uses the discrete abstraction as inputs. The limitations of patch mosaic paradigm on urban landscape fragmentation led us to seek an alternative perspective with which the continuous characteristics of landscape heterogeneity can be effectively addressed.

2.2.2 Study Area

Phoenix is taken as a test site representative of most urbanizing regions of the U.S. Southwest. It is located in the northern Sonoran desert and is characterized by hot summers and low rainfall (Figure 2.1). The average temperature in the summertime is over 38 °C and the annual precipitation is about 195 mm. There is a great deal of diversity in land use and land cover (LULC) classes, including residential, commercial, cultivated grass, agriculture, desert, vegetation, unmanaged soil and water. Since the beginning of the last century, the population has exponentially increased resulting in rapid urban expansion that has encroached on desert and agricultural lands (Jenerette & Wu, 2001; Knowles-Yanez et al., 1999; Redman & Kinzig, 2008). Characterizing landscape patterns in the Phoenix urban area serves as an important first step toward understanding the ecological and biological consequences of the rapidly growing city.

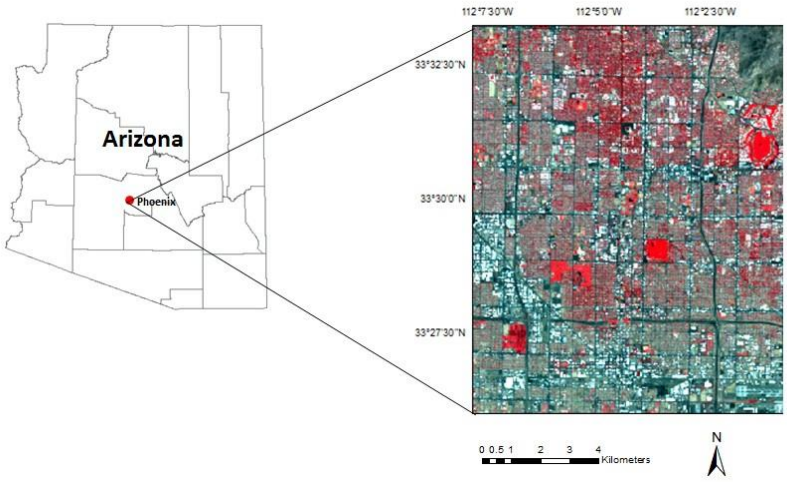

Figure 2.1 Study Area Located in Phoenix, Arizona.

2.2.3 Image Processing

The data used in this study consist of two satellite images—a QuickBird high-resolution multispectral image and a Landsat Thematic Mapper (TM) medium-resolution image. The QuickBird image was acquired on May 24, 2007 over the Phoenix urban area under cloud-free conditions and classified into six classes: buildings, trees/shrubs, grass, unmanaged soil, other impervious surfaces (e.g., roads, parking lots, airports) and water. The Landsat TM image acquired on May 17, 2007 was used to examine the continuous urban landscape heterogeneity. The two images were obtained close in time such that the likelihood of significant land cover changes is substantially reduced. Both images were orthorectified and co-registered using nearest neighbor resampling scheme with a root mean square error of less than 0.5.

The concept of spatially distinct regions with similar characteristics is well suited to landscape pattern analysis. Therefore, we used an object-oriented approach that aggregates pixels into discrete image objects where spectral variability is reduced (Herold

et al., 2003; Johnsson, 1994; Meddens et al., 2008). The creation of image objects takes advantage of both spectral and spatial information (Blaschke & Strobl, 2001). Different sets of parameters (shape, compactness, smoothness) and feature spaces (bands, indices, level of scales) were tested and hierarchical classification rules (nearest neighbor, membership functions, expert system) were implemented. After identifying land cover classes separately at different scales, we combined them using a GIS overlay function (Myint et al., 2011).

In addition to the object-oriented analysis, the most widely used supervised decision rule—maximum likelihood classifier—was employed to thoroughly evaluate the effectiveness of the object-based approach. Stratified random sampling with a minimum of 50 points per class was conducted to assess the accuracy of the per-pixel and object-oriented classifications respectively. For precision comparison purposes, the same sample points were applied to the output classification map generated by the two classification techniques. The object-based approach produced a substantially higher overall accuracy (90.40%) than the maximum likelihood classifier, which achieved an accuracy of 67.70 % (Myint et al., 2011). We have provided the error matrix, producer's accuracies, user's accuracies, overall accuracy, and kappa coefficient of the urban land cover map generated by the object-oriented classifier (Table 2.1).

Table 2.1 Classification Accuracy of the Urban Land Cover Map Generated by Object-oriented Approach

	Reference								
Classified	Buildings	Grass	Trees/ Shrubs	Other Impervious	Unmanaged Soil	Water	Total	Producer's Accuracy	User's Accuracy
Buildings	73	1	1	3	2	0	80	83.91%	91.25%
Grass	6	68	8	2	2	0	86	95.77%	79.07%
Trees/Shrubs	0	2	56	8	0	0	66	86.15%	84.85%
Other Impervious	1	0	0	87	0	0	88	83.65%	98.86%
Unmanaged Soil	6	0	0	3	70	1	80	94.59%	87.50%
Water	1	0	0	1	0	98	100	98.99%	98.00%
Total	87	71	65	104	74	99	500		

Overall Accuracy = 90.40%
Overall Kappa Statistics = 0.89

2.2.4 Spatial Pattern Analysis

The spatial pattern analysis was conducted based on the detailed land cover map. We used two approaches: 1) moving window analysis on class- and landscape-level metrics based on high-resolution imagery; and 2) continuous fragmentation analysis based on local spatial autocorrelation indices derived from medium-resolution Landsat TM data, using vegetation, built-up and soil index as inputs.

2.2.4.1 Moving Window Analysis Using Landscape Metrics

Landscape metrics have been extensively used in a variety of fields (Herzog et al., 2001; Ji et al., 2006; Peng et al., 2010; Seto & Fragkias, 2005; Southworth et al., 2002). Despite the abundant information landscape metrics have provided, the proliferation of metrics presents confusion to scientists and practitioners as to which metrics are appropriate in effectively characterizing relevant landscape components (Cushman et al., 2008; Griffith et al., 2000; Haines-Young & Chopping, 1996). For this study, we have

selected a suite of metrics available in FRAGSTATS software package (McGarigal & Marks, 1995) for both class- and landscape-level studies (Table 2.2). FRAGSTATS includes the composition metrics Percentage of Landscape (PLAND) and Largest Patch Index (LPI) that measure features using the composition of the landscape or abundance of certain classes within the landscape; and the configuration metrics Patch Density (PD), Aggregation Index (AI), Contagion (CONTAG) and Landscape Shape Index (LSI) that measure complexity, arrangement and proximity.

The two major land cover classes that we examined were vegetation and built-up area classes. The vegetation class was generated by combining the original land cover classes of trees/shrubs and grass. The built-up class was defined by merging buildings and other impervious surfaces as identified in the original classification map. Moving window analysis was employed for calculating metrics where each pixel was assigned an attribute value based on the neighboring pixels within a predefined spatial extent. The window size of 330 m was empirically selected following existing spatial pattern studies in Phoenix (Buyantuyev et al., 2010). A similar window size was also suggested by Myint et al. (2010) for the study of vegetation cover in effectively lowering surface temperatures over the Phoenix metropolitan area. The output is a raster grid with the value of each pixel representing the resultant metric value for that location.

Table 2.2 Descriptions of Landscape Metrics Used in the Study

Landscape metric	Description
Percentage of landscape (PLAND)	The percentage the landscape comprised of the corresponding patch type (Class level)
Largest patch index (LPI)	The percentage of the landscape comprised by the largest patch of the corresponding patch type (Class level)
Patch density (PD)	The number of patches per 100 hectares (Class level)
Aggregation index (AI)	Degree of aggregation of cells of the focal class (Class level)
Contagion (CONTAG)	Measures the extent to which patch types are aggregated or clumped (Landscape level)
Landscape shape index (LSI)	A perimeter-to-area ratio that measures the overall geometric complexity of the landscape (Landscape level)

2.2.4.2 Fragmentation Analysis Using Local Spatial Autocorrelation Indices

Given the limitations of patch mosaic paradigm, we evaluated the performance of spatial autocorrelation indices under a continuous framework. Three images of normalized indices—normalized difference vegetation index (NDVI); principle component analysis band 1 (PCA 1) and near infrared band (NIR) based built-up index (PNBI); and normalized difference soil index (NDSI) were created upon which spatial autocorrelation indices were applied. All three indices were employed in the landscape-level analysis whereas only the NDVI and PNBI were utilized in the class-level analysis representing the abundance of vegetation and built-up area respectively. The formulas for

calculating NDVI, PNBI and NDSI are

$$NDVI = \frac{Band\ 4 - Band\ 3}{Band\ 4 + Band\ 3} \quad (2.1)$$

$$PNBI = \frac{PCA\ 1 - Band\ 4}{PCA1 + Band\ 4} \quad (2.2)$$

$$NDSI = \frac{Band\ 5 - Band4}{Band\ 5 + Band\ 4} \quad (2.3)$$

The NDVI is effective at identifying green vegetation biomass due to a strong absorption in the red region (Band 3) and a strong reflection at near-infrared band (Band 4) (Jensen, 1996; Tucker, 1979). The NDSI proposed by Kearney et al. (2002) for identifying soil end members in the spectral mixture model was employed here to capture soil reflectance. In order to delineate urban built-up areas accurately, a new index combining PCA 1 and NIR and hereafter referred to as PNBI was developed and tested. Since PCA 1 generally represents man-made features and NIR captures active vegetation, the PNBI can be expected to effectively depict urban areas by enhancing built-up areas while suppressing the spectral confusion stemming from vegetation cover. The performance of the PNBI was compared with the normalized difference built-up index (NDBI) developed by Zha et al. (2003), originally developed for built-up area extraction from medium-resolution satellite imagery. For high-resolution urban mapping, PNBI outperformed NDBI because it provided clearly identified boundaries with less spectral mixing.

Two local measures of spatial dependency, the Getis statistic and the local Moran's *I* were applied on the normalized index images (NDVI, PNBI, NDSI) as alternative indicators of landscape pattern. In contrast with global measures of spatial autocorrelation that outline the degree of spatial association in a single value, local indicators assess the extent to which observations of similar and dissimilar values are

clustered for each location (Anselin, 1995). This fits well into our landscape pattern analysis because the degree of fragmentation can be measured within a defined spatial extent.

The Getis statistic measures local concentration by evaluating the sum of attribute values within a radius of distance from the original point as a proportion of the sum of attribute values of all the points within the entire region under investigation (Getis & Ord, 1992). It is defined as

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j}{\sum_j x_j} \tag{2.4}$$

where x_j denotes the variate value of a pixel at location j. $w_{ij}(d)$ is a binary spatial weights matrix where ones (1) are assigned to pixels within distance of *d* around the center pixel and zeros (0) to others. The spatial weights matrix is row-standardized for computation purposes. The standardized Getis statistic is further developed for remote sensing applications to assess the significance of the Getis statistic under the null hypothesis of no spatial dependency assuming that the Getis statistic is normally distributed (Wulder & Boots, 1998). The standardized version of the Getis statistic is

$$G_i^*(d) = \frac{\sum_j w_{ij}(d)x_j - W_i^*\bar{x}}{s[W_i^*(n - W_i^*)/(n-1)]^{1/2}} \tag{2.5}$$

where $W_i^* = \sum_j w_{ij}(d)$. In Eq. (2.5), $\bar{x}$ and *s* are mean and standard deviation, respectively, of all pixels in the entire image. A high positive value of the standardized Getis statistic indicates a spatial clustering of values that deviate substantially from the mean in the positive direction, and a low negative value implies a spatial clustering of values far from the mean along the negative direction. To be consistent with the moving window analysis conducted above, the distance *d* was determined to be 150 m, which led

to a window size of 330 m with the radius being 150 m (150×2+30 = 330).

The local Moran's *I* differs from the Getis statistic in that the covariances rather than the sums are computed. The local Moran's *I* is defined as

$$I_i(d) = \frac{n(x_i - \bar{x})}{\sum_i (x_i - \bar{x})^2} \sum_j w_{ij}(d)\left(x_j - \bar{x}\right) \tag{2.6}$$

where x_i denotes the variate value at location *i* and $\bar{x}$ represents the average value of all the pixels in the study region. The same spatial weight matrix was created and used here. An area is interpreted as clustered or homogeneous (high-value clustering or low-value clustering) when the local Moran's *I* is significantly higher than the mean. Dispersed or heterogeneous areas are indicated by a significantly low value of local Moran's *I*. A specialized routine was developed to compute the Getis statistic and the local Moran's *I* in the MATLAB software package and the output raster grids were obtained with each pixel representing the calculated value of the statistic. Our hypothesis is that the Getis statistic and local Moran's *I* can be expected to glean useful information on landscape pattern and can be used as important indicators of urban landscape fragmentation.

2.2.5 Statistical and Comparative Analyses

A sample of 1000 randomly generated points was utilized to evaluate the relationship between landscape metrics and spatial autocorrelation indices. Four vegetation and built-up area class metrics were compared with Getis of NDVI and Getis of PNBI respectively and statistical models were developed to estimate landscape-level metrics using continuous indices generated from Landsat TM imagery.

In addition, the performance of the discrete landscape metrics and continuous

spatial indices was evaluated based on their capability of quantifying spatial patterns for different LULC types. The land uses were extracted from the pre-existing land use map of the Central Arizona-Phoenix Long-Term Ecological Research (CAP LTER) region completed by Stefanov et al (2001). Major land uses that we are particularly interested in are: mesic (moderate water use) residential, xeric (low water use) residential, commercial, cultivated grass, vegetation, desert and water. Local Moran's *I* of NDVI was utilized as a measure of spatial heterogeneity at the landscape level. The map of local Moran's *I* was divided into five zones using the Jenks natural breaks method that seeks to minimize within-class variance while maximizing the inter-class variance to achieve the optimal arrangement of pixels (Jenks, 1977). Table 2.3 lists the lower and upper bounds of each zone along with the average values. For each zone, the percentage of each LULC class was calculated and compared with landscape-level metrics.

Table 2.3 Statistics for Five Zones of Local Moran's *I* of NDVI

	Minimum	Maximum	Mean
Zone 1	-3.09	0.47	-1.31
Zone 2	0.47	1.98	1.23
Zone 3	1.98	5.55	3.77
Zone 4	5.55	13.21	9.38
Zone 5	13.21	24.44	18.83

2.3 Results

2.3.1 Landscape Pattern Analysis at the Class Level

A comparison was made between the spatial pattern of the Phoenix urban area measured by the vegetation and built-up class metrics with that measured by the local

spatial autocorrelation indices. Four class-level metrics (PLAND, LPI, PD and AI) were selected (Figure 2.2). PLAND and LPI revealed a similar pattern of vegetation patches where high values resided mainly in mesic residential areas and cultivated grass. Both LPI and PD suggested that xeric residential areas have the most fragmented landscape with moderate vegetation cover. The patterns indicated by AI resembled those suggested by PLAND that vegetation was aggregated predominately in cultivated grass, mesic residential and xeric residential areas. The Getis of NDVI (Figure 2.4a) exhibited a pattern that was very similar to FRAGSTATS metrics for vegetation with high density of vegetation found in mesic residential areas and cultivated grass. Low values of Getis of NDVI (cold spots) were mostly concentrated in commercial regions where vegetation was very sparse. Close inspection of the results implied a high correlation between FRAGSTATS metrics and spatial autocorrelation indices at the class level. This was further confirmed by the comparison of the FRAGSTATS metrics for the built-up class to Getis of PNBI as shown in Figures 2.3 and 2.4b.

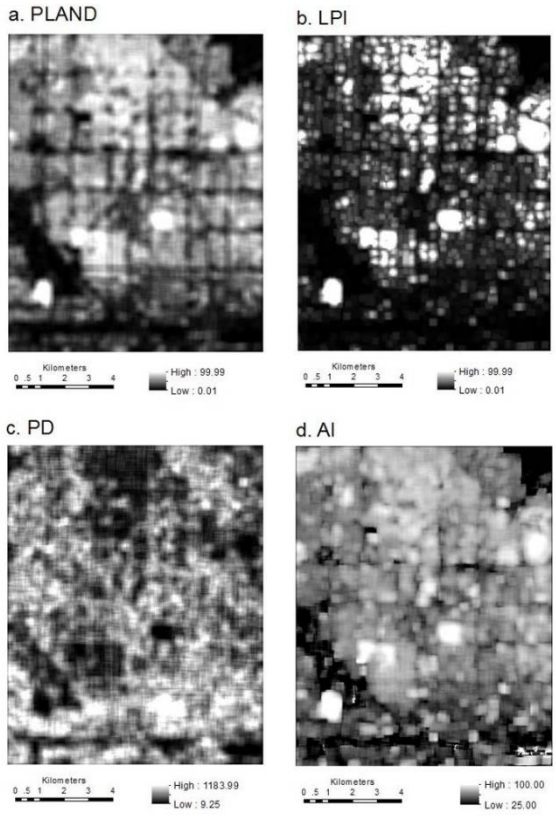

Figure 2.2 Output Raster Grids of Moving Window Analysis for Vegetation Class Metrics Derived from QuickBird Imagery for Phoenix Urban Area on May 24, 2007: a. PLAND; b. LPI; c. PD; d. AI.

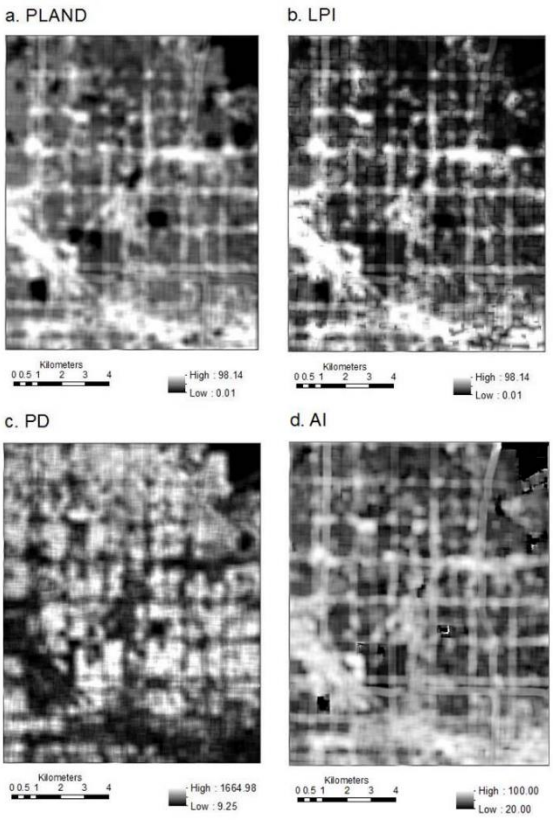

Figure 2.3 Output Raster Grids of Moving Window Analysis for Built-Up Class Metrics Derived from QuickBird Imagery for Phoenix Urban Area on May 24, 2007: a. PLAND; b. LPI; c. PD; d. AI.

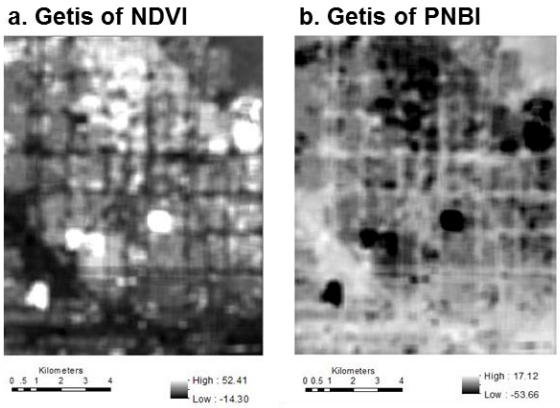

Figure 2.4 Getis-transformed Maps from Landsat TM Imagery for Phoenix Urban Area: a. Getis of NDVI; b. Getis of PNBI.

The relationship between class-level metrics (vegetation and built-up) and local spatial autocorrelation indices are shown in scatterplots (Figure 2.5). The PLAND and LPI were well related with the Getis statistic as evidenced by the fairly high coefficients of determination for both classes. Hump-shaped patterns were found in the relationship between the PD with respect to the Getis statistic for both classes. The PD measures the number of patches per 100 hectares. It was demonstrated as one of the best indicators of landscape fragmentation because it implied how a particular class is fragmented within the landscape. Figure 5c demonstrated a typical situation where increasing vegetation abundance led to a greater degree of fragmentation, but it dropped as vegetation gradually dominated the landscape. It is noteworthy that both low and high values of the Getis indicated spatially clustered patterns (either low-value clustering or high-value clustering), thus a homogeneous landscape; whereas Getis values around zero signified a maximum degree of landscape fragmentation and could be treated as the critical point for evaluating landscape heterogeneity. The AI for vegetation showed a moderate non-linear

relationship with Getis of NDVI and a saturation point was reached when Getis of NDVI was around 20. The trend of AI, however, was different for the built-up class, where the built-up patches continued to aggregate without reaching a saturation point.

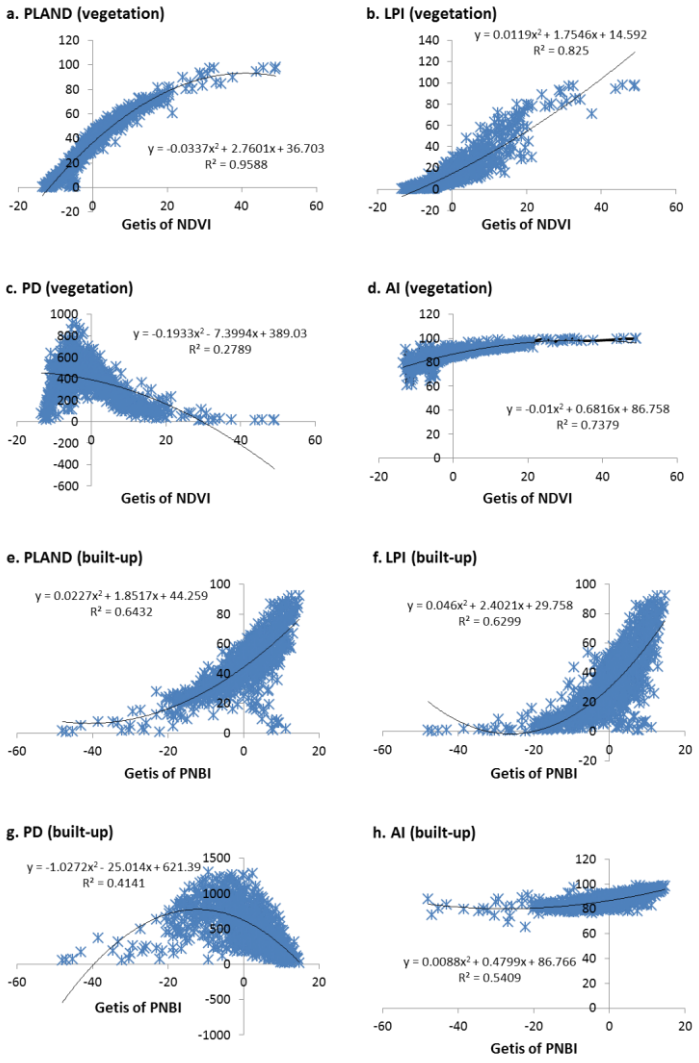

Figure 2.5 Scatterplots of Class Metrics vs. Getis statistics: a-d. Vegetation Class Metrics vs. Getis of NDVI; e-h. Built-Up Class Metrics vs. Getis of PNBI.

2.3.2 Landscape Pattern Analysis by Land Use Type

Local autocorrelation indices, such as local Moran's *I*, can glean extra information

on landscape fragmentation beyond the traditional clustering indices (Southworth et al., 2004). We hereby examined the utility of local Moran's *I* in comparison to landscape metrics in characterizing landscape fragmentation for different LULC types. We divided the local Moran's *I* of NDVI into five zones using the natural breaks classification method (Table 2.3) and for each zone, the percentage of every LULC type was presented in a pie chart (Figure 2.6). For comparison purposes, descriptive statistics for the two landscape metrics – CONTAG and LSI were generalized for every LULC type (Table 2.4). Since there was only one sample point, the water class was excluded. In Figure 2.6, xeric residential was the most fragmented landscape throughout the six land use types as it took up the highest proportion in Zone 1, a zone representing the lowest level of the local Moran's *I*. Zone 2 typified a random spatial pattern with the local Moran's *I* around zero and was primarily dominated by commercial areas. As the local Moran's *I* increased, clustered land use types such as mesic residential and vegetation gradually took up larger portions of the pie chart with mesic residential taking 55.26% of Zone 3 and vegetation taking 42.86% of Zone 4. Grass was indicated as the most homogeneous landscape with the highest proportion found in Zone 5, a zone signifying the highest level of the local Moran's *I*. A high degree of fragmentation was found in xeric residential areas as suggested by low mean CONTAG and high mean LSI. Additional information that was gleaned by the landscape metrics not implied by the local Moran's *I* was that desert showed higher degree of fragmentation than grass. This is in part due to the fact that the presence of trees and bushes that grow on the bare soil can be clearly captured by the land cover classification of fine-resolution imagery.

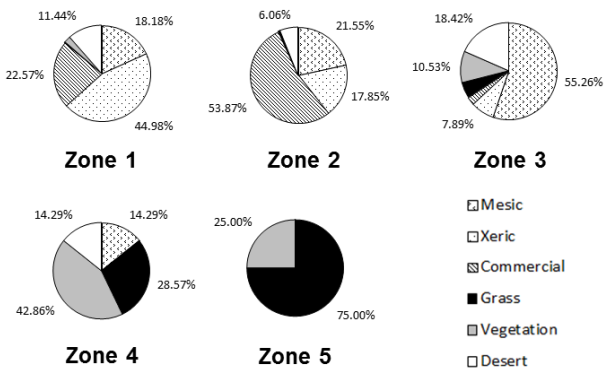

Figure 2.6 Pie Charts Showing the Percentages of LULC Types in the Five Zones of Local Moran's *I* of NDVI.

Table 2.4 Descriptive Statistics of Landscape Metrics by Land Use Type

Land use type	Mean		Standard deviation		Coefficient of variation	
	CONTAG	LSI	CONTAG	LSI	CONTAG	LSI
Mesic residential	47.23	10.76	9.23	2.08	0.20	0.19
Xeric residential	39.72	12.14	5.44	1.70	0.14	0.14
Commercial	42.37	10.40	6.83	1.97	0.16	0.19
Grass	69.33	5.27	16.84	2.72	0.24	0.52
Vegetation	55.09	8.07	16.89	3.30	0.31	0.41
Desert	52.40	9.20	17.69	3.64	0.34	0.40

2.4 Discussion

2.4.1 Comparing Continuous Framework with Discrete Representation

Discrete landscape models have found important merit in representing a wide range of landscapes (Griffith et al., 2000; Liu & Weng, 2008; Southworth et al., 2002). However, the applicability and reliability of such models are undermined by the limitations associated

with the subjective categorization of continuous landscape variables into discrete classes. Geostatistical techniques coupled with remote sensing were first applied to landscape ecological applications. This is conceptually in concert with the landscape gradient model (McGarigal & Cushman, 2005) which allows a wide variety of texture-based image processing techniques to be incorporated into landscape ecology studies. Local spatial autocorrelation indices were evaluated in comparison with discrete landscape metrics in quantifying the landscape patterns of Phoenix and surprisingly great analogy was found in the two sets of indices.

The behavior of the Getis statistic is conceptually in line with that of composition metrics which measure characteristics related to the abundance of particular land cover classes within the landscape. The conceptual consistency was confirmed by our experimental results that the Getis statistic showed strong relationship with composition metrics (PLAND and LPI) for both vegetation and built-up classes. On the other hand, the spatial pattern described by the local Moran's *I* showed remarkable similarity to that measured by configuration metrics for major LULC types.

The complexity residing in classification has long been a hot topic and there are no efficient and accurate classification systems that are readily available and applicable to most, if not all types of landscapes. One of the main contributions of continuous indices to the landscape ecology community is that no classification is required to understand landscape patterns and the underlying ecological processes, as continuous indices use raw data. The utilization of continuous raw data also enables us to eliminate arbitrariness in the selection of thematic resolution as well as a substantial amount of uncertainty associated with discrete representation, especially for urban landscapes.

2.4.2 Scaling Effects

It is widely known that landscape pattern is highly scale-dependent (O'Neill et al., 1996; Turner et al.,1989; Wu et al., 2002), and that the patterns observed under different scales tend to vary. Following previous studies (Lam & Quatrochi, 1992; Wu, 2004), we refer scale to "grain" (the spatial resolution or pixel size of an image) and "extent" (the spatial extent under which patterns are observed). In this experiment, indices derived from images of different grain sizes were compared under the same spatial extent (window size). This is in line with our goal of using readily available coarse-resolution data to achieve high quality results comparable to those obtained by fine-resolution commercial imagery. There were concerns, however, associated with this experimental design as spatial patterns are subject to scale change. We addressed these concerns by 1) demonstrating and comparing the landscape patterns characterized by the Getis statistic derived from images of different grain sizes (QuickBird and Landsat TM), and 2) understanding the impacts of scale on the correlations between two composition metrics (PLAND and LPI) and Getis of NDVI in response to changing grain size and spatial extent.

We calculated Getis of NDVI within a 330 m window size using QuickBird and Landsat TM, respectively and examined the impact of different grain sizes on the resulting maps. Despite the remarkable inconsistency in the magnitude of the attribute values, the two maps showed substantial analogy with hardly any discrepancy between them. Statistically, the correlation between the two maps exceeded 0.99 with a sample of 1000 randomly selected points. This not only justified our initial assumption that the use of identical window size should be able to smooth out some of the differences in the level

of details in the two images, which makes them comparable, but also ignited our interest in understanding how varying scales (grain size and spatial extent) might affect the statistical results we had obtained so far.

Our preceding experiment suggested that changing grain size may not have significant effects on Getis of NDVI, and thus the correlation between the Getis statistic and landscape metrics. We expanded our experiment by evaluating the correlation between the two sets of indices—Getis of NDVI versus vegetation class metrics for changing grain size (from 30 m to 300 m); this time keeping the grain size consistent for both indices (Figure 2.7). The experiment was conducted for six ascending window sizes, from 840 m to ~6000 m.

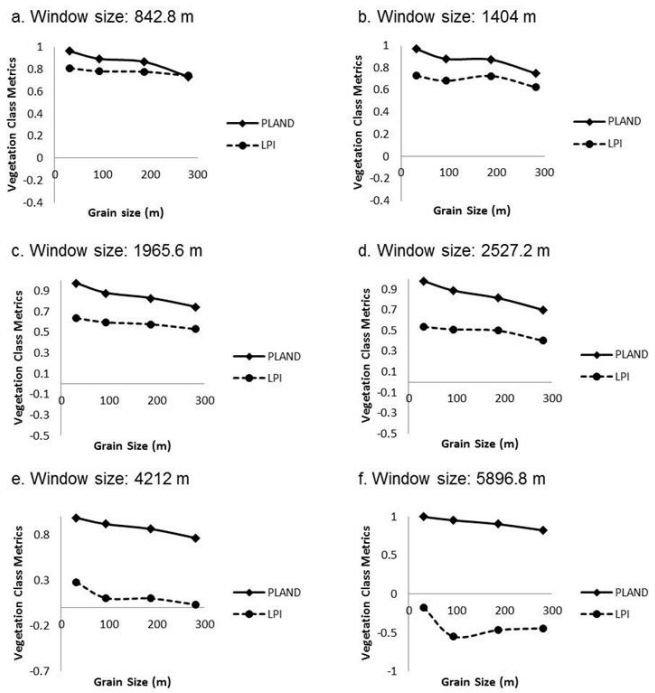

Figure 2.7 Scalograms Showing the Effects of Changing Grain Size on the Correlation between Getis of NDVI and Vegetation Class Metrics. Six Window Sizes were Investigated: a. 842.8 m; b. 1404 m; c. 1965.6 m; d. 2527.2 m; e. 4212 m; f. 5896.8 m.

It was noticed that both PLAND and LPI decreased mildly in their correlation with Getis of NDVI as grain size increased, regardless of the variation in window size. The correlation across the six window sizes stayed consistent for PLAND whereas it dropped dramatically for LPI as the window size increased, with the correlation falling below zero when the window size reached 6000 m. In general, the correlation between Getis of NDVI and vegetation class metrics was not very sensitive to changing grain size. That is, controlling for the window size, analogous spatial patterns can be identified by the two paradigms and it holds for varying levels of grain size.

While changing grain size did not affect the relationship significantly, the

correlation demonstrated different patterns with respect to changing spatial extent (Figure 2.8). We tested the correlation between the same variables as we did for changing grain size, with the window size varying from 200 m to 6000 m. Results showed that while a slowly mounting pattern was suggested for PLAND, the curve for LPI presented a dramatic decline in its correlation with Getis of NDVI, the value of which dropped to negative as the window size became sufficiently large. The decline may be in part due to little variation in the values of LPI, as a result of the small map size caused by the large window sizes. It seems that Getis of NDVI showed more robustness in representing the composition of a landscape than measuring the abundance of dominant patches, as was evidenced by the consistently high correlation with PLAND regardless of scales. More experiments are needed to address the effects of scale on the correlation between the Getis statistic and landscape metrics for other classes, as well as the responses of the relationship between the local Moran's *I* and configuration metrics to varying scales.

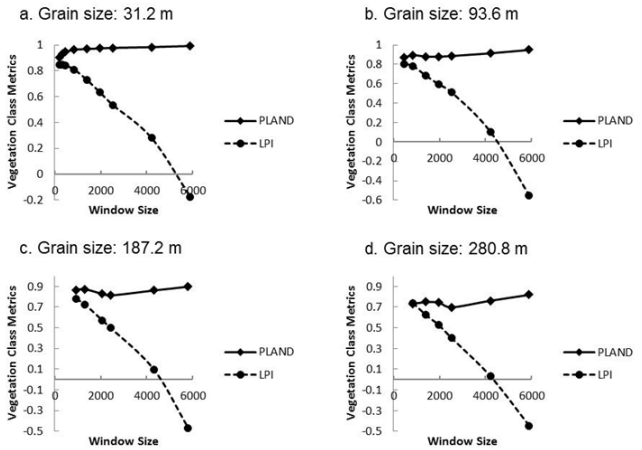

Figure 2.8 Scalograms Showing the Effects of Changing Spatial Extent on the Correlation between Getis of NDVI and Vegetation Class Metrics. Four Grain Sizes were Investigated: a. 31.2 m; b. 93.6m; c. 187.2 m; d. 280.8 m.

2.5 Conclusion

Statistical analyses suggest that there is a clear relationship between the Getis statistic and class metrics for vegetation and built-up classes. PLAND and LPI are well related with the Getis while PD presents a hump-shaped curve with respect to the Getis. We found that the Getis value around zero is a critical point for identifying landscape fragmentation. The multiple regression models developed in this paper can be used effectively to describe spatial pattern without creating detailed LULC maps or classifying high-resolution imagery. As predictors, Getis of NDVI and Getis of PNBI are significant in predicting important landscape metrics. Furthermore, the local Moran's *I* shows great analogy to the FRAGSTATS metrics in evaluating landscape fragmentation for different LULC types with xeric residential areas consistently demonstrating the highest degree of fragmentation across the region, followed by commercial and mesic residential areas. In terms of the scaling effects, it is found that the Getis statistic shows more robustness in measuring landscape composition than the abundance of dominant patches within the landscape identified using the patch mosaic paradigm. While PLAND demonstrates consistently high correlation with the Getis statistic regardless of the scales, LPI is more sensitive to changing extent than changing grain size.

The continuous indices presented in this paper show great potential in effectively characterizing the spatial patterns of any land cover type or a combination of different land cover types. Future research should consider the use of local spatial autocorrelation indices for a diversity of landscapes including forest, agricultural, coastal zones, deserts and grasslands as the use of continuous indices better addresses the dynamics associated

with gradual land transformations such as deforestation and land degradation. The methodological framework outlined here can be extended if additional continuous indices are developed to enhance the ability to accurately represent specific landscape characteristics. It remains for landscape ecologists to demonstrate how the proposed indices are useful in linking landscape pattern with the underlying processes, which is a central question of landscape ecology that needs to be clearly addressed by all ecological applications.

CHAPTER 3

TIME SERIES ANALYSIS OF URBAN DYNAMICS USING SEQUENTIAL LANDSAT IMAGERY AND SPATIAL STATISTICAL MODELING

3.1 Introduction

Human influences on the natural world have been a central topic in the literature of recent decades (An et al., 2005; Stem, 1993; Turner et al., 2003). Urban growth, from regional sprawl to global urbanization, is the most rapid, drastic, and irreversible form of human modification to the Earth's surface. Urban areas are holding 54% of the world's population, and this number is projected to reach 66% by 2050 (United Nations, 2014). Extensive replacement of the natural surface with impervious, manmade materials has profoundly affected ecological functioning and services well beyond city boundaries (Wu et al., 2011). While urban areas occupy only 3% of the Earth's landmass, they are responsible for 60% of total residential water consumption, more than 78% of carbon release, and about 76% of wood mainly used for industrial purposes (Brown, 2001; Grimm et al., 2008). Land modification in the urbanization process to sustain increasing urban population has also resulted in other types of change to the Earth's environment (Grimm et al., 2008).

Accurate and updated knowledge of the spatiotemporal pattern of urbanization is important for understanding urban growth and its various environmental consequences in a rapidly changing world (Ma et al., 2012; Zhang and Seto, 2011). Recent advancement of remote sensing technologies has stimulated a proliferation of research devoted to improve our understanding of the human impacts on the natural environment (Kaufmann et al., 2007; Martin, 2008; Voogt and Oke, 2003). Satellite imagery, collected from

space-borne sensors, provides repeatable and consistent representation of the Earth's surface at various temporal, spatial, and spectral scales (Begue et al., 2011). Long-term remote sensing records also facilitate examination and quantification of changes to landscape patterns through repetitive collection of data at multiple spatial scales and time intervals (Gillanders et al., 2008; Waylen et al., 2014). Key to the land change science is the ability of remote sensing to generate useful environmental parameters (e.g., vegetation and built-up indices) that reflect biophysical and socioeconomic properties of the earth's surface. Such indices are useful for identifying and extracting land cover features and thus are increasingly employed to characterize spatial and temporal patterns of landscape change.

Current research on land change science has relied heavily on thematic land use land cover (LULC) maps. Most common is the coupling of classified satellite imagery at multiple time points with a suite of spatial metrics characterizing various properties relevant to the land composition and spatial configuration of landscape at the patch (land cover/use object) level (Herold et al., 2003). The effectiveness of patch-based landscape model is evidenced by a large body of research designed to examine both cross-sectional and spatiotemporal patterns of various types of landscape not restricted to urban (Deng et al., 2009; Luck and Wu, 2002; Read and Lam, 2002; Seto and Fragkias, 2005). Some well-established tools (e.g. FRAGSTATS) and analytical methodologies (e.g. regression, ANOVA) facilitates the experimental design, analysis, and interpretation of this simple organizational framework, making it one of the most adopted tools for landscape pattern mapping over the last few decades.

Continuous landscape models emerge in response to the incapability of the patch mosaic model to detect within-class variations and the substantial subjectivity and arbitrariness introduced from the land cover classification analyses (McGarigal and Cushman, 2005). Additionally, continuous measures containing spatial information via analysis of brightness values provide unique sources of information relevant to the spatial arrangement and heterogeneity of landscape not available in the spectral domain (Musick and Grover, 1991). Existing research has compared the usefulness of various continuous measures with the patch mosaic model and shows great potential of continuous models in forest modeling (Southworth et al., 2004), savanna mapping (Pearson, 2002), and urban fragmentation analysis (Myint et al., 2015). Major advantages of continuous analysis can be briefly summarized as: (1) a closer association with the original satellite data as no land cover categorization is required; (2) greater flexibility to incorporate changes at the pixel level, which allows detection of subtle and gradual land cover modifications; (3) possibility to assess global and local spatial associations via evaluating spatial concentration and arrangement using continuous spatial indicators.

Long-term satellite imagery offers sufficient temporal sampling frequency and duration that permit detection of changes with the presence of substantial variation (De Beurs and Henebry, 2005). Frequently used methods to detect trends include regression trend analysis (Kogan and Zhu, 2001; Ma et al., 2012), harmonic analysis (Immerzeel et al., 2005; Jakubauskas et al., 2001), PCA (Hayes and Sader, 2001; Lasaponara, 2006; Shabanov et al., 2002), and ARIMA models (Jiang et al., 2010; Linthicum et al., 1999). Some of these high-powered statistical methods require considerable data transformation and thus are more suitable for detecting changes in a relatively short time series.

Computational demands also increase dramatically for increased spatial extent under study.

The Mann-Kendall test (Kendall, 1955; Mann, 1945) offers a simple and effective tool for describing monotonic trends in a long-term time series without recourse to computationally demanding techniques. Originally derived from a rank-based correlation test, the Mann-Kendall test has seen extensive applications in hydrological and climatology studies (Déry et al., 2009; Dunn, 1996; Gough et al., 2004; Han et al., 2010; Mohsin and Gough, 2010). Typical in a time series is the serial dependence among successive observations. As Cox and Stuart (1955) pointed out, positive serial dependence can lead to a higher chance of significant answer, even when the trend is absent. The modified Mann-Kendall test (Hamed and Rao, 1998) was proposed to reduce the effect of serial correlation on the significance of the trend test through modifying the variance of the Mann-Kendall statistic. After correcting for the serial autocorrelation, the modified Mann-Kendall test is robust to outliers, non-normality, serial dependence, and missing values, making it suitable for detecting trends in an image time series.

This paper presents our efforts to integrate image time series, continuous spatial indices, and non-parametric regression into a spatiotemporal study of urban dynamics. The utility of this method is tested on a 20-year image time series over Phoenix. Application of this method permits spatially explicit, pixel-based assessment of landscape pattern changes across the urban landscape. Additionally, findings from this study improve our understanding of the human impacts on the natural environment, which is of paramount importance for sustainable urban development in the context of rapid global change.

3.2 Materials and Methods

3.2.1 Study Area

This study focuses on the Phoenix metropolitan area, located in the northern part of the Sonoran desert in Arizona (Figure 3.1). Situated at the confluence of Salt and Gila Rivers, this area is characterized by a subtropical desert climate with hot and dry summers and warm winters. The average annual temperature is ~22 °Celsius (C), with the average summer and winter temperature 40 °C and 3 °C, respectively. This region receives an average of 32 days of precipitation each year, and the average annual rainfall is ~203 millimeters (mm) (ADWR, 2013). Our study area covers the Central Arizona-Phoenix Long-Term Ecological Research (CAP LTER) site, which is the subject of a multitude of urban ecology research focusing on the various relationships between urbanization and the socio-ecological systems using interdisciplinary approaches (Grimm and Redman, 2004). Since last century, Phoenix has witnessed a tremendous land transformation primarily from agriculture and natural lands to urban areas (Knowles-Yánez et al., 1999; Luck and Wu, 2002). The entire Phoenix metropolitan area consists of 23 cities and is home to more than four million residents. It is also among the fastest growing metropolitan areas in the nation with an estimated population growth rate of 45.3% from 1990 to 2000 (USCB, 2001).

Figure 3.1 Map of the Phoenix Metropolitan Area.

3.2.2 Data Acquisition and Processing

Our data set is a yearly Landsat TM image time series spanning from 1991 to 2010. The acquisition date of the imagery ranges from June to July and images of the best quality were selected. We used images acquired at approximately the same time of the year in the summertime to minimize the spectral variation brought by phenological effects and to make sure that vegetation is active and detectable. We identified clouded areas and shadows in the images using the Fmask tool developed in Zhu and Woodcock (2012) and replaced the values in these areas with a constant value similar to those of neighboring pixels. As areas with clouds and shadows are located in the mountainous regions well beyond the urban extent, we believe a constant is sufficient for this particular study without resorting to more complicated algorithms.

Vegetation and impervious surfaces are important land cover types in an urban environment. Hence, the spatial and temporal patterns of vegetation and built-up areas are of particular relevance for analyzing urban dynamics, especially in desert regions. We

employ the Normalized Difference Vegetation Index (NDVI), which is a common indicator of green biomass and has been extensive used in phenological studies (Lee et al., 2002; Reed et al., 1994; Studer et al., 2007), biomass production (Paruelo et al., 1997; Tieszen et al., 1997; Todd et al., 1998), crop mapping (Lunetta et al., 2010; Thenkabail et al., 2005), and land cover identification (DeFries and Townshend, 1994; Hansen et al., 2000). For anthropogenic areas, we utilize a recently developed built-up index based on a combination of the first band of the principal component analysis (PCA1) and the near infrared band (B4), hereafter referred to as PNBI (Fan and Myint, 2014). The PNBI takes the same format as the NDVI, with B4 and B3 in the NDVI replaced with PCA1 and B4, respectively. A comparison of the PNBI with the widely used Normalized Difference Built-up Index (NDBI) shows that PNBI is more sensitive to the spectral distinction between manmade features and open soil, making it particularly suitable for distinguishing built-up areas from natural desert lands and xeric residential areas in an arid region.

3.2.3 Landscape Pattern Analysis

Spatial structures of landscapes are closely associated with the land composition and spatial configuration of landscape elements. Land composition refers to the abundance and variety of land cover objects within a landscape, while configuration focuses on the spatial character, position, arrangement, and orientation of land cover patches (McGarigal, 1995). Unlike most of existing studies adopting the patch mosaic models, our study benefits from a continuous framework which offers a more detailed and accurate characterization of landscape heterogeneity. The utility of continuous representation of landscape structure was evidenced in a number of studies (Pearson,

2002; Qi and Wu, 1996; Southworth et al., 2004). We herein adopts the spatial autocorrelation approach detailed in (Fan and Myint, 2014) as an effective means to measure the concentration and spatial arrangement of land cover types in a rapidly growing desert region.

3.2.3.1 Land Composition

We use the Getis-Ord *G* as our measure of land composition. The Getis-Ord *G* measures the degree of local concentration of the attribute of interest. It is computed as the sum of attribute values within a predefined radius distance from the original point as a proportion of the sum of values for all the observations within the entire study region (Getis and Ord, 1992). The radius distance is set to 150 m which is the minimum radius encompassing all the major land cover types in the urban area. The same window size was employed in Buyantuyev et al. (2010) in a landscape pattern analysis over Phoenix. A high value of the *G* statistic indicates a spatial clustering of high values and vice versa. The significance of the *G* depends on the distribution of the statistic. Under the assumption of normal distribution, a hot spot (or a cold spot) is detected with strongly high positive (low negative) Z scores of the *G* statistic. While the significance test is illuminating for making statements about local pockets, it increases the difficulty in the interpretation of the final results. To maintain the temporal and spatial continuity of the data set, we use the *G* statistic solely as a spatial indicator of land composition without regard to its statistical significance. Respectively, we compute the *G* statistic for the NDVI and PNBI maps derived in Section 3.2.2 for a continuous assessment of the local concentration of vegetation and built-up features.

3.2.3.2 Spatial Configuration

Spatial configuration includes a variety of properties concerning the spatial character, orientation, proximity, and interspersion of land cover features. In this study, we focus on one of these characters—spatial arrangement. We use another spatial autocorrelation index—the local Moran's *I* as our measure of spatial arrangement (Anselin, 1995). Unlike the *G* statistic, the local Moran's *I* assesses the correlation between the focal observation and its proximate observations. A high positive local Moran's *I* indicates a pattern of spatial clustering (regardless of its property) whereas a low negative local Moran's *I* suggests a dispersed spatial arrangement. The local Moran's *I*'s capability to identify clustered and dispersed spatial patterns makes it a useful indicator for landscape structure mapping (Fan et al., 2015; Qi and Wu, 1996; Zheng et al., 2014). We have three candidate maps for which the local Moran's *I* can be calculated: Tasseled Cap brightness, PNBI, and NDVI. We decided to use the NDVI maps over the other two mainly for two reasons. First, the Tasseled Cap brightness values for vegetation and some of the impervious surfaces are very close to the mean of the entire image. As the local Moran's *I* is incapable of distinguishing spatial clusters of medium values from random pattern, it is very likely that the local Moran's *I* is insignificant for vegetation and some of the built-up areas. Second, due to the high albedo of desert lands, unmanaged soil and other undisturbed lands usually receive higher scores of PNBI than most of the impervious surfaces. Therefore, the local Moran's *I* for PNBI will likely pick up the expansive open soil in the mountainous areas rather than the manmade features in the urban area. NDVI is more appropriate in this case as it clearly shows the spatial pattern of the landscape by means of highlighting vegetated surfaces in the region. As with the *G*

statistic, we consider the local Moran's *I* as an indicator of spatial arrangement of urban features with no regard for its statistical significance.

3.2.4 Mann-Kendall Test

The original Mann-Kendall (MK) test is a non-parametric statistical test frequently used for analysis of trend in hydrologic and climatologic data. It statistically assesses whether a monotonic increasing or decreasing trend of a variable is present over time. Unlike parametric regression analysis, the MK test does not require the trend to be linear, nor does it require normality in the observed data. Mathematically, the MK test statistic *S* is calculated by subtracting the number of positive differences in the data from the number of negative differences, or more formally

$$S = \sum_{k-1}^{n-1} \sum_{j-k+1}^{n} sgn(x_j - x_k) \tag{3.1}$$

where x_j and x_k are observations obtained at times 1, 2, …, n, respectively. *S* is positive if data collected later in time tend to be greater than those obtained earlier, and vice versa. Under the assumption of independent observations, the MK statistic can be approximated by a normal distribution when the number of observations exceeds 10, with mean and variance given by

$$E(S) = 0 \tag{3.2}$$

$$var(S) = n(n-1)(2n+5)/18 \tag{3.3}$$

In a time series, the likelihood of violating the assumption of independent observations is extremely high because data in a time series tend to be autocorrelated, i.e., the value on the observation obtained at the current time point is conditional on that obtained at a previous time point. The serial correlation can increase the chance of

identifying significant trends, even if the trends are absent. Hamed and Rao (1998) proposed a modified MK test in which the variance is modified to account for the serial correlation. The modified MK test computes the autocorrelation function between the ranks of the observed data after removing a non-parametric trend estimate from the data. Autocorrelations significant at a certain significance level (e.g., 5%) are then used in the evaluation of the modified MK statistic. The adjusted variance of the MK statistic is

$$var_m(S) = var(S) \cdot \frac{n}{n_S^*} \tag{3.4}$$

where

$$\frac{n}{n_S^*} = 1 + \frac{2}{n(n-1)(n-2)} \sum_{i=1}^{n-1} (n-i)(n-i-1)(n-i-2)\rho_S(i) \tag{3.5}$$

Here n is the total number of observations in the time series. n_S^* is the effective sample size after adjusting for the serial correlation. $\rho_S(i)$ is the autocorrelation function between the ranks of the data for lag i. Only autocorrelations that are significant are included in the calculation. The significance of the trend is determined by comparing the standardized test statistic with the critical value of the standard normal distribution at a predefined significance level.

3.2.5 Land Cover Dynamics at the Metropolitan and Municipality Levels

We analyzed the spatiotemporal dynamics of vegetation and built-up areas in reference to the total urban land in 2009, i.e., the spatial extent of our analysis is restricted to the urban area in 2009. The urban extent map was created from a land cover classification map in 2009 from the CAP LTER website (caplter.asu.edu). A maximum likelihood classifier was employed with a total of six classes created: agricultural soil,

residential/disturbed, asphalt, cultivated vegetation, desert/undisturbed, and water. A total of 300 sample points was used for the accuracy assessment and the overall classification accuracy achieved 89.7%. The final map of urban extent was created by combining residential/disturbed, asphalt, and cultivated vegetation into a single urban class. We performed the modified MK test for the abundance and spatial arrangement of vegetation and built-up areas in the region and summarized the total number of pixels (total area) and percent of lands showing (significantly) increasing and decreasing patterns in the value of the two spatial indicators.

In addition to the analysis at the metropolitan level, zonal statistics were derived for major municipalities in the Phoenix metropolitan area. We used the city boundaries in 2010 as our reference city boundary data, because the spatial extents they encompass are the largest over the entire study period. The city boundary data are available through the TIGER/Line product by the US Census Bureau. Similar to the analysis at the metropolitan level, we calculated for each city the total area and percent of lands with (significantly) increasing and decreasing trends in the land composition of vegetation and built-up areas. The MK test based on the local Moran's *I* of NDVI also reveals the monotonic trend in the development of spatial heterogeneity for each city over the last twenty years.

3.3 Results and Discussion

3.3.1 Spatiotemporal Dynamics of Vegetation

Table 3.1 shows the statistics for vegetation dynamics derived from the modified MK test on the *G* of NDVI. Within the study area, the total area of land with increases in

vegetation abundance is 1120.72 km^2, of which 469.84 km^2 shows significant vegetation increases. The number of pixels showing an increasing and significant increasing vegetation cover accounts for 38.98% and 16.34% of the urban land in 2009. Comparatively, more areas report declining trends of vegetation cover with 1719.1 km^2 (59.79%) and 879.16 km^2 (30.58%) showing a declining and significant declining pattern of vegetation abundance, which is about 1.5 times the area with vegetation increases.

Table 3.1 Area and Percent Area Changes in the *G* of NDVI from 1991 to 2010

	Original	Significant*
Area increase (km^2)	1120.72	469.84
Percent area increase (%)	38.98	16.34
Area decline (km^2)	1719.1	879.16
Percent area decline (%)	59.79	30.58

* Significant at the 0.05 level.

Landscape change maps provide a visual tool for discerning the location of significant changes across the landscape. Figure 3.2 shows the spatial distribution of statistically significant vegetation changes in the Phoenix metropolitan area. We found the majority of vegetation increase occurring in the northeast part of the region where new residential areas emerge as a result of land conversion from natural desert lands. The second largest portion of vegetation increases is located within some of the urbanized areas where more and more vegetation is brought to the residential areas. On the other hand, most vegetation declines are seen in the highly developed regions, notably in the central and southeast part of the area. Overall, there are more areas with vegetation declines than vegetation increases in this region.

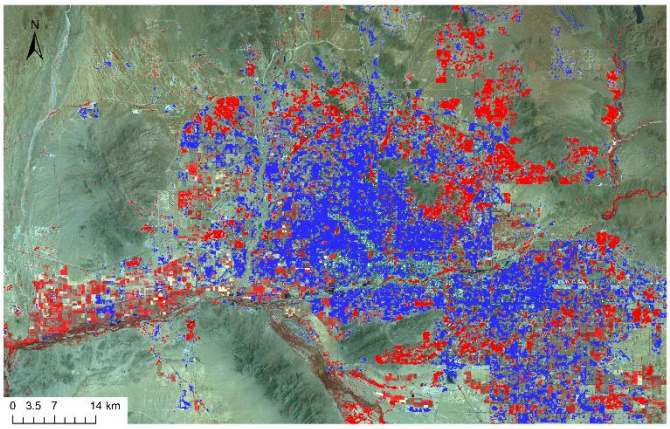

Figure 3.2 Spatiotemporal Changes in the *G* of NDVI over the Phoenix Metropolitan Area (Red: Significant Increase; Blue: Significant Decrease).

The relationship between urbanization and vegetation dynamics is a complex, variable, and nonlinear one. The probability and degree of vegetation changes is not only location-specific, but also vary at different stages of urbanization. While declines of vegetation cover is likely during urban expansion, it is not necessarily true for every city (Zhou et al., 2014). Liu et al. (2015) investigates the degree of vegetation degradation with respect to urbanized areas across 50 cities in the world and shows that while vegetation tends to decline in general, some large cities can experience vegetation restoration as opposed to degradation at some point of their urban development. In the desert environment, our observed vegetation dynamics manifest a distinct pattern which differs from most cities studied so far. Instead of losing vegetation to impervious and other manmade features, urbanization in desert cities accompanies vegetation plantation and irrigation, which presumably results in an overall increase in the abundance of

vegetation. The substantial vegetation declines, however, require further evaluation with regard to the specific land conversion processes behind these changes.

3.3.2 Spatiotemporal Dynamics of Built-up Areas

The total area showing an increasing trend in the *G* of PNBI is 1449.58 km^2, and 637.29 km^2 of them are significant (Table 3.2). The number of pixels with (significant) increases accounts for 50.41% (22.16%) of the total number of pixels that are identified as urban in 2009. We observed a comparable area of land with decreases in the *G* of PNBI (1387.33 km^2), and the area showing significant declines (24.98%) is only slightly greater than that with significant increases (22.16%).

Table 3.2 Area and Percent Area Changes in the *G* of PNBI from 1991 to 2010

	Original	Significant*
Area increase (km^2)	1449.58	637.29
Percent area increase (%)	50.41	22.16
Area decline (km^2)	1387.33	718.26
Percent area decline (%)	48.25	24.98

* Significant at the 0.05 level.

An important issue that complicates the analysis of built-up dynamics is that open soil tends to have higher PNBI than most built-up areas due to their higher albedo scores. Therefore, a decline in the value of PNBI can indeed be a useful indication of land conversion from open soil to urban land use, regardless of land use type. Figure 3.3 shows the spatial distribution of significant changes in the *G* of PNBI. We observed a pattern nearly inverse to that of vegetation changes where increases in the built-up index are seen in the urbanized areas in the central and southeast part and declines on the

outskirts of the region. Increases in PNBI are evident in places undergoing changes from non-commercial areas (e.g., cropland, residential) to commercial land use, which usually accompany a big vegetation loss. The areas with a decreasing trend in the built-up index show great correspondence to those with vegetation increases, collectively delineating the areas experiencing land development from natural desert lands.

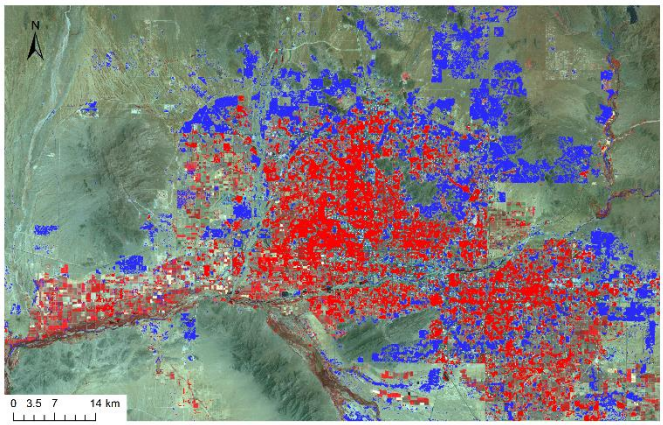

Figure 3.3 Spatiotemporal Changes in the *G* of PNBI over the Phoenix Metropolitan Area (Red: Significant Increase; Blue: Significant Decrease).

3.3.3 Changes in the spatial pattern of the landscape

The spatial arrangement of the urban landscape was assessed by using the MK test on the local Moran's *I* time series. Based on Table 3.3, the total area of land with increases in the local Moran's *I* is 1200.4 km^2, and 458.52 km^2 is statistically significant. A greater number of pixels, on the other hand, reports (significant) declines in the local Moran's *I*, suggesting that a relatively large portion of the landscape has become spatially more dispersed as the cities evolve.

Table 3.3 Area and Percent Area Changes in the Local Moran's *I* from 1991 to 2010

	Original	Significant*
Area increase (km^2)	1200.4	458.52
Percent area increase (%)	41.75	15.95
Area decline (km^2)	1634.91	703.54
Percent area decline (%)	56.86	24.47

* Significant at the 0.05 level.

Figure 3.4 shows the spatial distribution of significant changes in the local Moran's *I*. The trend of decreasing local Moran's *I* dominates the entire landscape, notably in the central and southeast part of the region which is occupied with large residential areas. Declines in the local Moran's *I* indicate elevated level of spatial complexity, primarily due to the extensive land transitions from agriculture and desert lands to spatially more complex urban land uses. Areas with increased local Moran's *I* are located in the northeast part of the region where desert lands were converted into cultivated vegetation, commercial, and mesic residential areas.

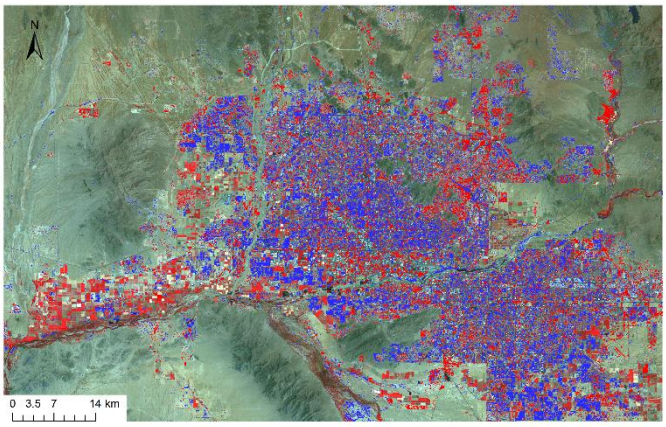

Figure 3.4 Spatiotemporal Changes in the Local Moran's *I* over the Phoenix Metropolitan Area (Red: Significant Increase; Blue: Significant Decrease).

Land fragmentation is a phenomenon of land form change that accompanies most forms of urbanization. Through interspersing natural landscapes with developed land features, land fragmentation significantly alters the urban ecosystems, leading to a number of detrimental consequences including loss of biodiversity, decreased productivity of agricultural and forest management, and elevated costs for provision of public services (York et al., 2011). Quantifying landscape structure, especially land fragmentation, is not only important for mapping the trajectories of urban growth and change, but also essential for understanding the profound impacts of land conversions on the bio-socio-ecological systems. As a local indicator of spatial dependence and association, the local Moran's *I* provides an effective means for characterizing the spatial complexity of urban landscapes from a spatial statistical perspective (Myint et al., 2015; Zheng et al., 2014). The use of satellite images without creating LULC maps makes the spatial indicators particularly appealing in longitudinal analysis of landscape pattern changes.

3.3.4 Landscape Pattern Changes at the City Level

City-level analysis of land cover dynamics provides specific information about each city's individual growing pattern. Figure 3.5 shows the area and percent of urban lands with significant vegetation increases and declines for major municipalities in the region. Note first that for almost every city, the areas with significant vegetation declines are always greater than those with significant vegetation increases, indicating an overall tendency of vegetation loss (Figure 3.5a). Of all the cities, Phoenix and Tolleson have the largest and smallest area of vegetation increases and the counterpart cities for vegetation

declines are Phoenix and Guadalupe, respectively. A different story is told when considering the area changes with respect to city size as there is a great variation in the city size in this metropolitan area, varying from 2 km^2 (Guadalupe) to 1342 km^2 (Phoenix). When percent of urban land is considered, Paradise Valley and Fountain Hills stand out as cities with the largest and second largest area of vegetation increases (Figure 3.5b). As a majority of their current urban areas is developed from natural desert rather than croplands, these two cities have seen significant growths in vegetation, a great amount of which is brought by local residents. Tolleson features a unique form of urban expansion in which 56.92% of its current urban lands reports vegetation declines and only 0.01% shows vegetation increases, a 56.91% difference.

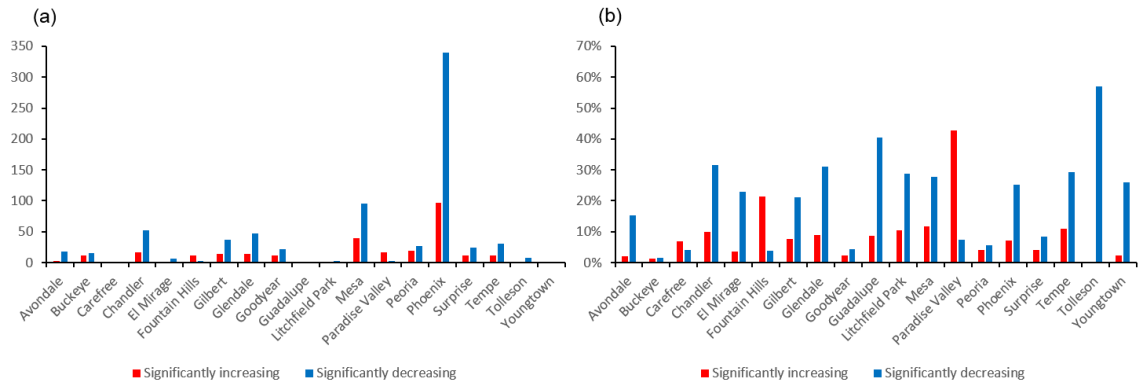

Figure 3.5 Vegetation Dynamics for Major Municipalities over the Phoenix Metropolitan Area: a. Area Changes; b. Percent Area Changes.

When it comes to built-up areas, Phoenix continues to lead the game with the largest area of PNBI increases (276.54 km^2), followed by Mesa (69.66 km^2) and Glendale (42.81 km^2) (Figure 3.6a). Unlike vegetation, the area showing significant declines in the built-up abundance is not consistently greater than that with significant increases. The

city with most area of PNBI declines is Phoenix, suggesting extensive land conversions from open soil to urban land use in this city. Similar patterns are observed for Paradise Valley, Fountain Hills, and Carefree, three cities showing the highest percentages of PNBI decreases (Figure 3.6b). The column chart highlights Tolleson as the city with the most PNBI increases (51%) and smallest PNBI decreases (0%), indicating that nearly all the current urban areas in Tolleson are developed from active or fallow croplands rather than natural desert.

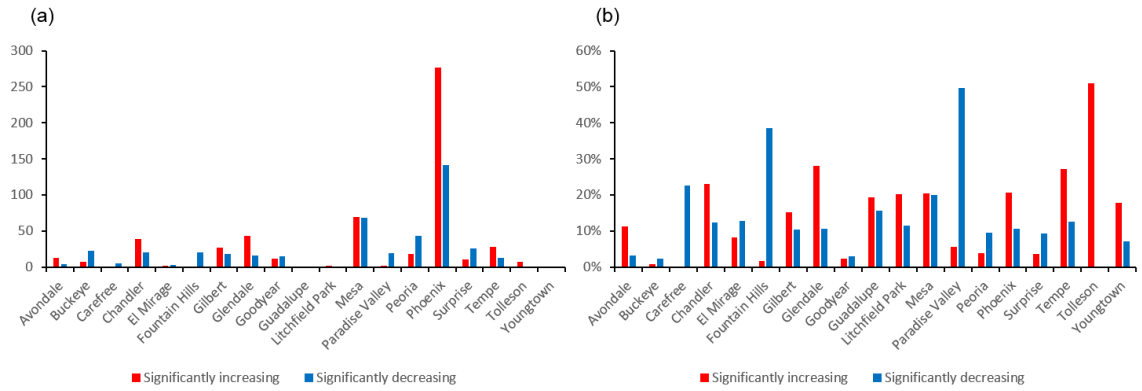

Figure 3.6 Built-up Area Dynamics for Major Municipalities over the Phoenix Metropolitan Area: a. Area Changes; b. Percent Area Changes.

Besides changes in the composition of vegetation and built-up areas, there is a significant alteration in the spatial structure of some cities in this region, exemplified by Phoenix, Chandler, and Mesa to name a few (Figure 3.7). Changes in the landscape structure are closely associated with land transition processes. Particularly relevant transition types include land conversions from open soil/croplands to residential/commercial, from residential to commercial, and from water intensive residential (mesic) to water efficient residential (xeric) or vice versa.

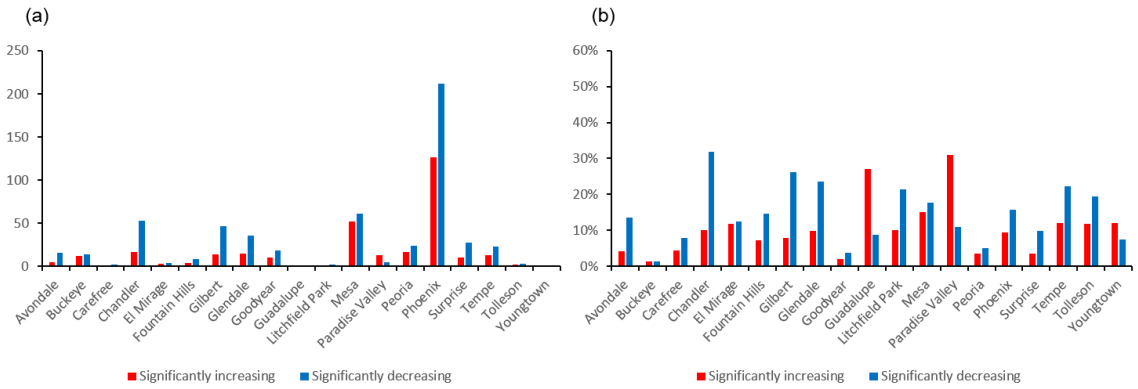

Figure 3.7 Changes in the Local Moran's *I* for Major Municipalities over the Phoenix Metropolitan Area: a. Area Changes; b. Percent Area Changes.

3.3.5 Linking Land Cover Changes to Land Use Conversions

We link the various types of land cover change identified using our approach with the specific types of land use conversions to gain a better understanding of the processes that underlie the changes in vegetation and built-up areas. We selected 100 sample points for each category of land cover change (e.g., vegetation increases) and identified the land conversion process occurred at each point from 1991 to 2010. We only show results for vegetation and decide to drop those for built-up areas due to the substantial overlapping patterns reflected in the dynamics of these two land cover types. Tables 3.4 and 3.5 show the land use conversion matrix for vegetation increases and declines from 1991 to 2010. In the areas showing vegetation increases, 57% of the sample points indicates land conversions from other land use types to residential areas and natural desert lands are responsible for 28% of these conversions. We also observed vegetation increases in the residential areas over the last 20 years, which likely occurs at the early phase of urban expansion when trees and shrubs were planted as part of the development plans.

Table 3.4 Land Use Conversion Matrix for Areas with Vegetation Increases

1991	2010					Total
	Active cropland	Residential	Commercial	Cultivated vegetation	Riparian	
Active cropland		1		1		2
Fallow cropland	5	5		4		14
Open soil	1	28	7	11	3	50
Residential		22				22
Commercial		1	2			3
Riparian					9	9
Total	6	57	9	16	12	100

In the areas reporting vegetation declines, 11 and 8 out of 100 sites have changed from agriculture to residential and commercial areas, respectively (Table 3.5). Similar findings are reported in Keys et al. (2007) in a study of urban dynamics in the city of Phoenix from 1970 to 2000 where they found that more than half of the croplands in 1970 had become urban land uses, primarily residential in 2000. Changes from croplands to urban land uses also corroborate the findings in Fan et al. (2014) which identifies a substantial loss of agriculture to urban areas in the time period from 1995 to 2010.

Table 3.5 Land Use Conversion Matrix for Areas with Vegetation Declines

1991	2010			Total
	Fallow cropland	Residential	Commercial	
Active cropland	3	11	8	22
Fallow cropland		1	4	5
Open soil		4	13	17
Residential		46	9	55
Cultivated vegetation			1	1
Total	3	62	35	100

We found a general decline of vegetation in the residential areas, evidenced by 46 of 100 sample sites (Table 3.5). The remarkable vegetation declines—as pointed out in a previous section—can be an outcome of extensive xeriscaping practices in this region. Residents are encouraged to replace their "green" yards with less water-demanding and more drought-resistant landscaping. In an effort to conserve water, some cities (e.g., Tempe, Mesa, and Chandler) have initiated "Grass-to-Xeriscape" programs that offer rebates to their residents for converting their lawns into low-water-use landscaping. While xeriscaping seems to reduce the need for water use from irrigation, it can render the living environment much hotter, which in turn elevates the demands for water and energy (Ruddell and Dixon, 2014).

Increases in the local Moran's *I* mainly occur as a result of the transitions from other land use types to residential and commercial areas (Table 3.6). Vegetation increases in residential, for example, tend to elevate NDVI in the focal pixel and its neighboring pixels, causing higher covariance and thus the local Moran's *I*. This likely explains the increasing local Moran's *I* in the residential areas. The land conversion from residential to commercial is another major cause for elevated local Moran's *I*. While this process is not as common as reciprocal switches between residential, it results in a substantial reduction in the spatial complexity across the region.

Table 3.6 Land Use Conversion Matrix for Areas with Increasing Local Moran's *I*

1991	2010						Total
	Active cropland	Open soil	Residential	Commercial	Cultivated vegetation	Riparian	
Active cropland				1	2		3
Fallow cropland	2		4		2		8
Open soil	1	4	14	16	5		40
Residential			31	16			47
Riparian						2	2
Total	3	4	49	33	9	2	100

The process of urbanization accompanies significant increases in the structural complexity of the landscape (Table 3.7). Our analysis has identified reciprocal changes between residential areas as the most important cause for raised spatial heterogeneity, followed by conversions from active croplands, open soil, and fallow croplands to residential. The practice of xeriscaping, for example, increases the spatial complexity by turning vegetation dominant landscapes into ones with open soil interspersed with vegetation and other land cover features. Urban development from open soil and croplands is another important process that complicates the spatial structure of the urban landscape. Our findings are in concert with Wu et al. (2011) that landscape in the Phoenix metropolitan area is becoming increasingly more fragmented in structure, more complex in shape, and more diverse in land use types. Buyantuyev et al. (2010) reaches the similar conclusion by showing a 15% decrease in the value of contagion, a landscape ecological indicator measuring the degree of patch clumpiness.

Table 3.7 Land Use Conversion Matrix for Areas with Declining Local Moran's *I*

1991	2010					Total
	Fallow cropland	Residential	Commercial	Cultivated vegetation	Riparian	
Active cropland	2	23	3			28
Fallow cropland		9				9
Open soil		19				19
Residential		34				34
Commercial			4			4
Cultivated vegetation		1		3		4
Riparian					2	2
Total	2	86	7	3	2	100

3.3.6 Long-term Landscape Mapping from a Spatial Statistical Perspective

This study presents a quantitative framework for analyzing long-term urbanization-induced land cover changes in a desert metropolitan region. Benefiting from a combination of environmental variables, local spatial indicators, and non-parametric regression, this study systematically quantifies the spatiotemporal patterns of vegetation and built-up features from a continuous landscape modeling perspective. What differs this study from existing LULC change research is the use of successive satellite images spanning 20 years as opposed to selecting several scenes and detecting changes occurred between each pair of them. While the plethora of data enables land change analyses at a much finer temporal resolution, it poses unprecedentedly critical challenges for computational capabilities, storage capacity, and processing efficiency (Wu et al., 2014). Long-term landscape change analysis requires sophisticated algorithms to perform land cover classification for every satellite scene in the study period (Li et al., 2015). The process of generating long-term LULC maps is not only a time-consuming and labor

intensive project, but also introduces errors and uncertainty into the system, which ultimately affect the accuracy of subsequent change analysis (Fan and Myint, 2014). Landscape change assessment through continuous environmental variables and spatial indicators provides a brand new perspective for urban landscape mapping. It embraces instead of avoids the fact that urban landscape is an inherently heterogeneous entity and bases the analysis on the original satellite signatures without any data transformation and/or simplification (e.g., LULC classification). In addition to detecting changes in the land composition, our approach manages to identify alterations in the spatial configuration of the landscape using local indicators of spatial association. The modified MK test used in this study permits evaluation of monotonic (not necessarily linear) trends in the abundance and spatial arrangement of LULC features at the per-pixel level. As a non-parametric test, the modified MK test is particularly useful for detecting changes when (1) linear trends are not substantiated, (2) several observations are missing, and (3) power transformations are either not desirable or fail to produce normality.

3.4 Conclusion

The Phoenix metropolitan area has undergone dramatic landscape changes since the 1970s. The rapid urban sprawl has profoundly altered the local and regional socio-ecological systems, leading to a number of negative environmental impacts such as loss of productive farmland, fragmented wildlife habitats, elevated city temperature, and reduced biodiversity (Seto et al., 2011). Our study presents a systematic analytical framework for evaluating the spatiotemporal patterns of vegetation and manmade features in this desert region. It is also one of the early attempts that integrates satellite

remote sensing, spatial autocorrelation indices, and statistical modeling into the analysis of urban dynamics.

Our results show that over the last two decades, the total abundance of vegetation has increased in the northeast part of the region, where new residences were converted and built from natural desert lands. The central and southeast part, on the other hand, have seen extensive vegetation declines largely driven by the practice of xeriscaping to lower water use for irrigation. The spatial distribution of the built-up index is inverse to that of vegetation dynamics where increases are most evident in the central and southeast and declines at the outskirts. Due to the high albedo of desert sands, reductions in the built-up index is not an indication of loss of built-up areas but rather land conversions from desert and other open lands to urban land uses. We found that land conversions from agriculture and desert lands are important factors responsible for elevated spatial heterogeneity. Another process that complicates the landscape is xeriscaping where a mixture of land use types emerge in replace of the originally vegetation dominant landscape. Our analysis at the city level highlights Phoenix as the city experiencing the most dramatic urbanization, which is evidenced by significant land conversions from both desert and croplands into a diversity of residential, commercial, and cultivated vegetation. Our analysis also highlights cities with distinct urbanization patterns, with Paradise Valley and Fountain Hills exemplifying pure development from desert lands whereas nearly 100% of Tolleson's current urban areas have been converted from active or fallow agriculture and nothing else.

Several issues need to be scrutinized in pursuit of the next stage of this study. First, while the built-up index (PNBI) used in this study is capable of distinguishing built-

up areas from bare soil, it fails to locate built-up features in the high positions of its numerical range. This complicates the interpretations of PNBI-based built-up dynamics as declines in PNBI do not explicitly suggest loss of built-up areas. Improvements on the index is certainly needed while an active search and evaluation of established built-up indices is also helpful. Such indices may include the index-based built-up index (IBI) (Xu, 2008), the enhanced built-up and bareness index (EBBI) (As-Syakur et al., 2012), and the biophysical composition index (BCI) (Deng and Wu, 2012). Second, the results of our analysis should be interpreted with regard to the specific spatial scale without overgeneralization. The well-known Modifiable Areal Unit Problem (MAUP) has been extensively discussed in the literature (Fotheringham, 1989; Openshaw and Openshaw, 1984; Openshaw and Taylor, 1979) and it is important to evaluate the results of any statistical analysis in the context of specific grain size and areal unit (Jelinski and Wu, 1996; Turner et al., 1989). The scale and zoning effect on the results of our analysis is another direction for future research.

CHAPTER 4

QUANTIFYING SPATIAL ARRANGEMENTS OF URBAN VEGETATION AND ITS IMPACTS ON SEASONAL SURFACE TEMPERATURES

4.1 Introduction

Urbanization replaces natural surface materials with manmade materials. Replacing these natural landscape components significantly alters the radiative, thermal, moisture, and aerodynamic properties of the environment. As a result, urban areas are often much warmer than their surrounding rural areas – a phenomenon known as the urban heat island (UHI) effect (Lo and Quattrochi, 2003; Voogt and Oke, 2003). The UHI effect has significant implications for ecosystems, energy demands, and human well-being. For example, excessive heat may increase water and air-conditioning demands, thereby threatening the sustainability of water and energy provision (Gober et al., 2011; Yuan and Bauer, 2007). Increased UHI intensity can raise the concentration of urban pollutants, such as ground-level ozone, which detrimentally affects human health (EPA, 2012; NJDEP, 2006). Due to the greater long-wave radiant heat load and reduced wind speed, nighttime UHI is likely to increase the duration and magnitude of heat waves, thereby elevating the risk of heat-related mortality (Clarke, 1972). Low-income urban residents are generally more vulnerable to these extreme heat impacts due to their lack of social resources to tackle the heat stress (Connors et al., 2013; Harlan et al., 2006).

UHI effects are commonly identified by measuring air temperature at the urban canopy layer (UCL) and the urban boundary layer (UBL) (Oke 1976, Voogt and Oke, 2003). The advancement of thermal remote sensing technologies has made it possible to measure thermal surface patterns, commonly referred to as surface UHI (SUHI) (Roth et

al., 1989). SUHI differs from atmospheric UHI in terms of underlying causes, measurement techniques, and mitigation strategies. For instance, air temperature measurements are taken *in situ* using fixed thermometers or automobile traverses with instruments (e.g., thermometer, barometer, and hygrometer) mounted on a vehicle. In contrast, SUHI is measured by the upwelling thermal radiance received by a sensor mounted on a platform, typically a satellite or an aircraft. As SUHI is measured indirectly, it is strongly influenced by the conditions of the intervening atmosphere.

The intensity and spatial pattern of the SUHI also displays a much stronger dependence on land surface characteristics, such as land use and land cover (LULC) type (Arnfield, 2003; Buyantuyev and Wu, 2010). For instance, vegetative cover, such as trees, shrubs, and grass, cools the surrounding environment by altering the surface energy balance through evapotranspiration (Gallo et al., 1993; Owen et al., 1993; Taha, 1997). In addition, shading provided by trees and shrubs reduces surface energy absorption, which contributes significantly to the cooling effect (Akbari, 2002; Chow et al., 2012). Vegetation benefits human health by altering wind speeds and dispersing air pollutants (Kinney, 2008; Nowak et al., 2006). Urban planners have long recognised these positive effects and have widely adopted the use of green areas in urban environments. In fact, this strategy is understood to be one of the most effective strategies for UHI mitigation (EPA, 2012).

Numerous studies have considered the characteristics of urban vegetation in relation to its positive role in mitigating urban warming (Kaufmann et al., 2003; Sandholt et al., 2002; Van de Griend and Owe, 1993), including a large body of work on UHI mitigation through urban forestry in Phoenix (Buyantuyev and Wu, 2009; Myint et al.,

2010; Jenerette et al., 2007). One study examined the urban surface energy balance in response to land cover variations using a Local-Scale Urban Meteorological Parameterization Scheme (LUMPS) (Middel et al., 2011). The investigation determined that urban vegetation can potentially reduce storage heat flux density and therefore plays an important role in UHI mitigation. A more recent study used high-resolution, remotely sensed data to investigate thermal pattern variations with respect to detailed urban land cover characteristics in Phoenix (Myint et al., 2013). The study concluded that an inverse relationship, which was most pronounced during summer daytime, existed between vegetation fraction and surface temperature. Chow et al. (2012) provide a comprehensive review of the UHI research in Phoenix.

While the effectiveness of urban vegetation in ameliorating the UHI effect is well-understood, the cost-benefit trade-offs of increasing urban vegetation are also important considerations for land managers. Among these trade-offs is the potential of increased water demands, which is a major challenge for expanding green areas in an urban environment, especially for arid and semi-arid cities with limited water supplies. Recent attempts to address the trade-off between urban cooling and water usage include a study by Gober et al. (2012), which demonstrated that increasing urban vegetation enhances urban nighttime cooling but also elevates outdoor water use by approximately 20%. Another study delved into the interrelationships among irrigated landscapes, nighttime temperatures, and outdoor water use and found that a threshold of the water-to-temperature ratio can be determined beyond which increased water consumption had limited effects on UHI mitigation (Gober et al., 2010).

Given these challenges, optimising the spatial pattern of urban vegetation in desert cities becomes particularly important for ameliorating UHI while conserving resources. While the inverse relationship between vegetation fraction and land surface temperature (LST) is well-documented (Chen et al., 2006; Li et al., 2011; Weng et al., 2004; Yuan and Bauer, 2007), the potential impacts of the spatial arrangement of vegetation on the urban thermal environment at various geographic scales is less thoroughly characterised. We hypothesise that with a fixed amount of vegetative cover, the spatial pattern (e.g., clustered, dispersed) of green vegetation has impacts on LST and these impacts are scale-dependent.

Recent studies demonstrate that the fragmentation of urban greenspace does indeed influence air and surface temperatures (Connors et al., 2013; Li et al., 2012; Yokohari et al., 1997; Zhang et al., 2009). Furthermore, it was found that a combination of factors including shape, size, and segmentation level of vegetation patches influences the LST with lower LST usually associated with large, contiguous, and compact vegetation patches (Cao et al., 2010; Li et al., 2012; Zhang et al., 2009). Most if not all of these studies adopted a patch mosaic model, which represents a landscape as a number of homogeneous patches generated via a classification system (Forman, 1995; McGarigal and Cushman, 2005).

The patch mosaic model is effective as it provides a simplified framework for describing a heterogeneous landscape. However, this model often fails to represent landscape heterogeneity accurately, as it suppresses within-class variation, thereby causing a substantial loss of information and potential error propagation (McGarigal and Cushman, 2005). For instance, the model is limited in its ability to identify subtle land

cover modifications, such as deforestation and land degradation (Foody and Boyd, 1999; Lambin, 1997). A comprehensive discussion on the limitations of discrete landscape representation is available in Fan and Myint (2014).

Another issue with this simplified approach pertains to the utilisation of established tools (e.g., FRAGSTATS) under the patch mosaic paradigm. One of the most commonly used measures of landscape heterogeneity is patch density (PD), a built-in metric in the FRAGSTATS software package representing the total number of patches in a unit area. While PD reflects the fragmentation of a landscape, which is an important indicator in a number of ecological applications (Griffith et al., 2000; Liu and Weng, 2008; Luck and Wu, 2002; Xie et al., 2006), it fails to capture the actual spatial arrangement (e.g., clustered or dispersed) of land cover patches. The term "spatial pattern" is not confined to the segmentation level but has profound implications for how things are arranged in space.

There is a growing awareness of potential issues associated with the discrete modelling approach, and many researchers have started to explore the possibility of characterising continuous landscapes within a gradient framework. One study utilised a measure of spatial autocorrelation—Geary's C—in modelling the landscape structure of savanna landscapes in Australia, where objects are difficult to identify using traditional discrete models (Pearson, 2002). McGarigal et al. (2009) introduced a suite of continuous metrics that describe various aspects of a landscape, such as surface diversity, surface roughness, and texture. More recently, Fan and Myint (2014) examined the utility of geostatistical techniques for measuring urban landscape fragmentation in comparison to the use of discrete models and found that continuous models had the potential to

characterise the spatial patterns of land cover features more effectively with medium-resolution satellite data.

In this paper, we employ a spatial autocorrelation approach to represent the spatial heterogeneity of urban features accurately. Spatial autocorrelation measures spatial dependency within a gradient framework and has been demonstrated to be effective in characterising the spatial pattern of major urban features with remotely sensed data (Fan and Myint, 2014). Using this methodology, we aim to (1) determine the spatial arrangement of green vegetation cover using continuous spatial autocorrelation indices coupled with high-resolution satellite data; (2) examine the role of urban vegetation, particularly in terms of its spatial arrangement, on seasonal and daytime/nighttime LST, adjusting for the effects of vegetation abundance; and (3) investigate the sensitivity of vegetation-LST relationships to varying geographic scales.

4.2 Methods

4.2.1 Study Area

Our study is confined to the central part of Phoenix, Arizona, covering an area of ~178 km^2. The city is located in the northern Sonoran desert and is characterized by hot and dry summers and warm winters. The average maximum temperature for this region exceeds 38°C during summertime with an average annual precipitation of approximately 195 mm. The Phoenix urban area consists of a variety of LULC classes, including commercial, industrial, residential, cultivated grassland, unmanaged soil, desert, and water. This region was historically dominated by agricultural lands and desert, most of which have been converted to residential and commercial lands since the early 20th

century as the population increased rapidly (Jenerette and Wu, 2001; Knowles-Yánez et al., 1999).

The intense land cover modification has affected the energy balance and ecological functioning of the entire region, extending well beyond the Phoenix metropolitan area (Grimm et al., 2008). Given the limited space in the city, a better understanding of vegetation-LST relationships, particularly the relationship between the spatial arrangement of urban vegetation cover and LST, can provide significant insights for sustainable urban management and the mitigation of UHI.

4.2.2 Image Processing

We used a cloud-free QuickBird image acquired on 24 May 2007 to derive the vegetation land cover map. The satellite image has a spatial resolution of 2.4 m with four spectral channels: blue, green, red, and near-infrared. Three principal component analysis (PCA) bands were generated as auxiliary data (Stevens, 2009). Instead of using pixel-based classification, we employed an object-oriented approach that groups pixels into discrete image objects by selected properties, primarily spectral similarity (Herold et al., 2003; Johnsson, 1994; Meddens et al., 2008). This object-oriented classification was conducted using the eCognition software package (Definiens 2009).

Grass, trees, and non-vegetation classes were identified using a nearest neighbour classifier with ten training samples selected for each class. The nearest neighbour classifier allows users to add and/or modify training samples after each round of classification, until a satisfactory classification map is obtained (Ivits and Koch, 2002). Based on the user's expert knowledge coupled with a series of trial-and-error tests, five bands in the feature space were determined as inputs to the classifier due to their

sensitivity to green biomass. These bands include the visible green and red bands, the near-infrared band, the ratio PCA band 1, and the normalised difference vegetation index (NDVI) band (Myint et al., 2011).

The ratio PCA band 1 is the ratio of PCA band 1 to the summation of all seven bands in the feature space. The NDVI band was calculated as the ratio of the difference between the near-infrared and red bands to the summation of the same two bands. The producer's accuracy and user's accuracy were 80% and 81.67% for trees, and 93.33% and 92.67% for grass (Myint et al., 2011). Two binary maps, for grass and trees, respectively, were produced by assigning a value of one to vegetation pixels and a value of zero to non-vegetation pixels.

4.2.3 Land Surface Temperature

LST data were derived from Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) imagery. Four images acquired on 06 July 2005, 22 August 2005, 27 February 2007, and 05 March 5 2007 were used to examine Phoenix's seasonal and daytime/nighttime LST characteristics (Figure 4.1).

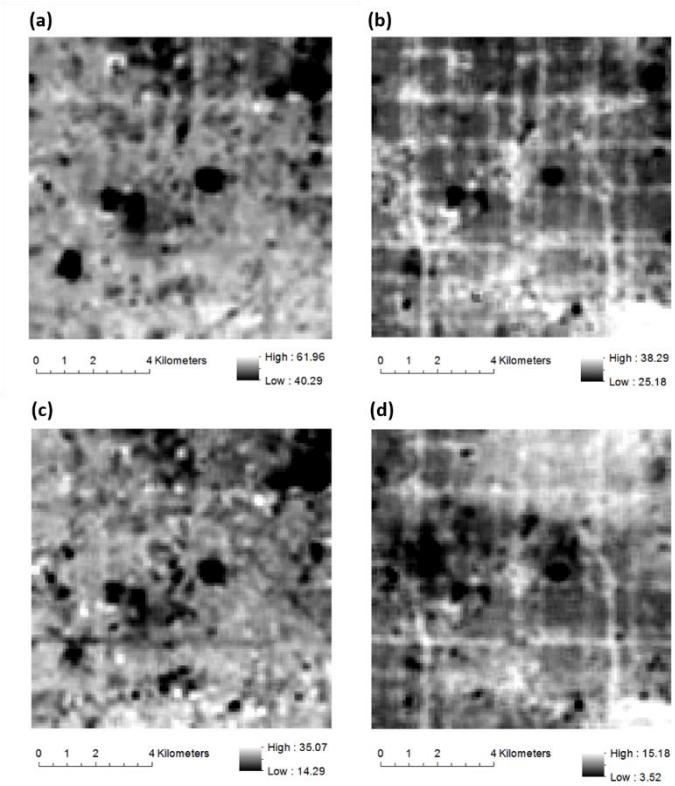

Figure 4.1 LST Maps Derived from ASTER Imagery: a. 06 July 2005 (Daytime); b. 22 August 2005 (Nighttime); c. 27 February 2007 (Daytime); d. 05 March 2007 (Nighttime).

A temperature-emissivity separation (TES) algorithm with an absolute accuracy of 1 – 4 K and a relative accuracy of 0.3 K (JPL, 2001) was employed to derive land surface kinetic temperatures from the five thermal infrared channels of the ASTER imagery. Table 4.1 shows the descriptive statistics of the LST maps. ASTER LST data were spatially resampled to different window sizes.

Table 4.1 Descriptive Statistics of the ASTER LST Data (℃).

Acquisition date	Acquisition time	Mean	SD	Maximum	Minimum
06 July 2005	11:20 am (MST)	55.81	2.87	61.96	40.29
22 August 2005	10:34 pm (MST)	31.56	1.75	38.29	25.18
27 Februrary 2007	11:15 am (MST)	29.71	1.48	35.07	14.29
05 March 2007	10:35 pm (MST)	10.44	1.65	15.18	3.52

4.2.4 Fraction and Spatial Pattern of Vegetation

Vegetation fraction, defined as the ratio of the number of pixels classified as vegetation to the total number of pixels within a predefined window size, was used to quantify the abundance of green vegetation. Efforts have been devoted to explore the most effective spatial extent of the vegetation-temperature relationship using a variety of techniques including remote sensing (Myint et al., 2010) and climate modeling (Kormann and Meixner et al., 2001). Collectively, a window size of ~200 m was determined to be optimal for maximum cooling effects. We followed previous empirical studies and used 204 m (2.4 m × 85 = 204 m) as the spatial extent in calculating the vegetation fraction, as well as the spatial autocorrelation index, which characterises the spatial pattern or arrangement of vegetation.

We used the local Moran's *I* (Anselin, 1995) as a continuous index to measure the spatial heterogeneity of vegetation. As one of the local indicators of spatial association (LISA), the local Moran's *I* is effective at characterising the spatial arrangement of particular land cover types at a local scale (Fan and Myint, 2014). It enables assessment of the degree to which similar and dissimilar observation values cluster around locations of interest.

To represent the spatial patterns of a particular land cover type and minimise the influence of other types of cover, a vegetation binary map was created to act as a filter on the local Moran's *I* map. Specifically, the local Moran's *I* of vegetation (local Moran's I^v) was calculated as

$$I_{ij}^{v}(d) = \frac{\sum_{k=1}^{d}\sum_{l=1}^{d} I_{[(i-1)d+k][(j-1)d+l]} V_{[(i-1)d+k][(j-1)d+l]}}{n_{ij}} \tag{4.1}$$

where i and j represent the spatial dimensions of the output map, and k and l represent the spatial dimensions of the input maps—*I* and *V*. The input consists of two maps with the same dimension. *I* is the output map of the local Moran's *I*, and *V* is a vegetation binary map with ones denoting vegetation pixels and zeroes denoting the others. The product of *I* and *V* ensures that only the spatial pattern of vegetation is considered. n_{ij} is the total number of vegetation pixels within a window size of *d*. The output from Eq. (4.1) is a degraded map of I^v using a local average filter with a window size of *d*. The value of each pixel represents the spatial arrangement of vegetation within an area of 2.4 *d* by 2.4 *d*. In general, consistent high/low values of the local Moran's I^v indicate clustered/dispersed areas, respectively.

Using a set of hypothetical scenarios for different spatial patterns with fixed quantities of "1"s, we tested the utility of the local Moran's I^v (Figure 4.2). It is important to note that given the same fraction (51%), the local Moran's I^v varies with different patterns of "1"s. Specifically, while clustered and scattered patterns receive positive and negative scores, respectively, the value of the local Moran's I^v approximates zero when no systematic pattern is observed.

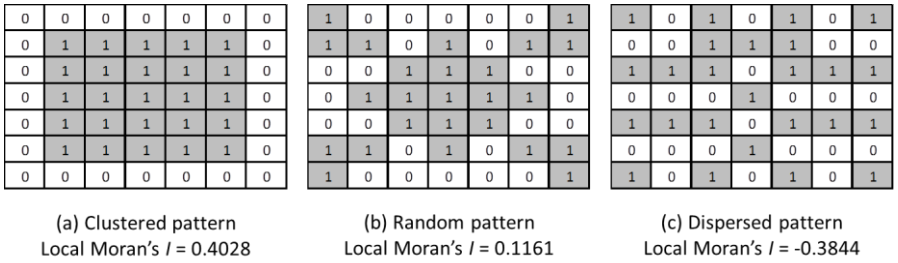

Figure 4.2 Hypothetical Scenarios Demonstrating Various Spatial Patterns of "1"s with Fixed Percentage: 25/49 = 51%.

4.2.5 Statistical Analyses

We investigated the bivariate relationships between LST and the spatial arrangement of vegetation indicated by the local Moran's I^v using ordinary least square (OLS) regression. Regression models were fit for the four scenes (summer daytime, summer nighttime, winter daytime, and winter nighttime) to determine the influence of the spatial pattern of urban vegetation on seasonal and daytime/nighttime LST.

Noting that higher amounts of vegetation cover present a more clustered pattern than lower amounts of vegetation cover, we investigated the effects of spatial arrangement by controlling vegetation fraction to minimise the effects of varying vegetation abundance on LST. We divided the entire dataset (2,704 pixels) into ten categories, using increments of 10% vegetation fraction. Due to the limited number of samples with denser vegetation, we combined classes with > 60% fraction for grass as one single category (>30% fraction for trees). We then developed a regression model for each category, which resulted in seven models for grass and four for trees.

To clearly elucidate the relative importance of the spatial arrangement of greenspace with respect to LST in comparison to composition, multiple linear regression

models were developed with the predictor variables: local Moran's I^v and vegetation fraction. Given the inherent association between composition and spatial configuration, multicollinearity between the two predictors could be an issue in building the regression model. Highly correlated predictors can compromise the estimation accuracy of the OLS model (Kutner et al., 2005). In addition, independent interpretation of the regression coefficients is no longer appropriate due to the confounded impacts on the coefficients. To remedy this effect, we employed a ridge regression model as an alternative to the OLS model.

Ridge regression is designed to improve estimation accuracy by virtue of including a penalty of constant (Hoerl and Kennard, 1970). The standardized ridge estimator can be calculated from

$$\mathbf{b}_R = (\mathbf{r}_{XX} + c\mathbf{I})^{-1}\mathbf{r}_{YX} \tag{4.2}$$

where $\mathbf{r}_{XX}$ and $\mathbf{r}_{YX}$ are correlation matrices composed of simple correlations among predictor variables, and between the response variable and each predictor variable respectively. c is a bias constant reflecting the bias in the regression estimators and can be determined using the ridge trace – a plot of the ridge standardised coefficients against various values of c. The optimal bias constant is the smallest value of c when the coefficients cease to decrease or increase and start to become stable. When the standardised coefficients were obtained, they were transformed back to the original variables for interpretation.

To examine how vegetation-LST relationships vary at different geographical scales, we employed and tested an array of window sizes from 104 m to 504 m in increments of 100 m. The factors under observation include the coefficients of

determination (R^2) of bivariate regression models, the R^2 of ridge regression models, and the coefficients of the local Moran's I^v in the ridge regression models. Scalograms were created to show the trends of the factors with respect to varying scales for both grass and trees, and for all of the four scenes. This investigation provided important indications of the optimum window size at which maximum cooling effect is achieved.

4.3 Results and Discussion

4.3.1 Bivariate Relationships between Spatial Pattern of Vegetation and LST

Figure 4.3 shows the bivariate relationships between the local Moran's I^v and LST. The local Moran's I^v showed consistently negative relationships with respect to seasonal LST for all four scenes regardless of vegetation type. The spatial pattern of grass shows more variation in daytime LST than nighttime LST, and has significantly stronger correlation with summer LST than winter LST (Figures 4.3a and b). Similar trends held for trees (Figures 4.3c and d), where summer and daytime LSTs showed a much stronger association with the spatial arrangement of trees compared to winter and nighttime LSTs.

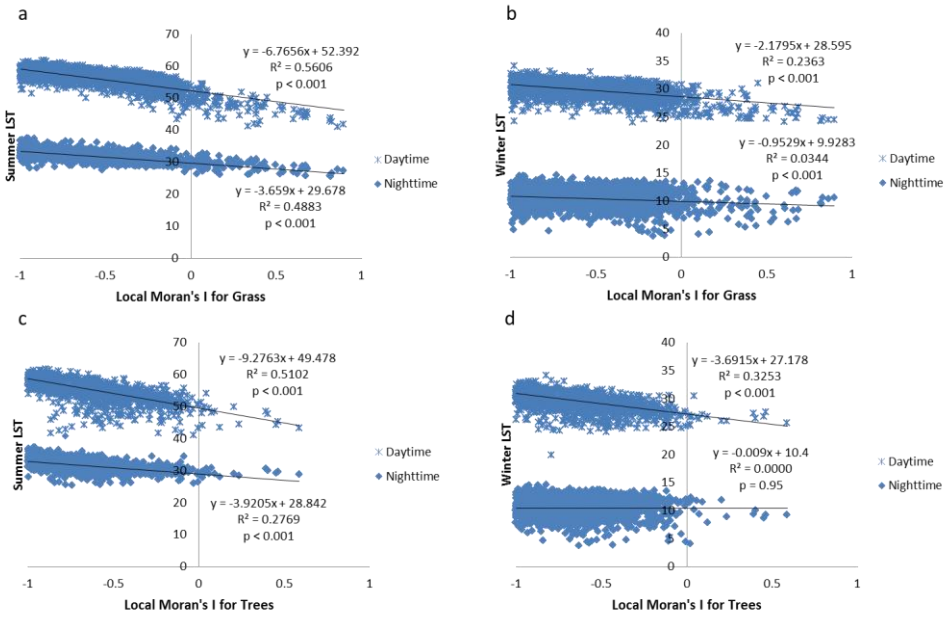

Figure 4.3 Bivariate Relationships between the Local Moran's I^v and Seasonal LST: a. Local Moran's I^v of Grass vs. Summer LST; b. Local Moran's I^v of Grass vs. Winter LST; c. Local Moran's I^v of Trees vs. Summer LST; d. Local Moran's I^v of Trees vs. Winter LST (Relationships for Daytime and Nighttime are Presented Separately on Each Diagram).

Our results suggest that clustered patterns of grass or trees lower surface temperatures more effectively than dispersed patterns. In other words, less fragmented patterns of urban vegetation reduce seasonal LST more effectively, especially during summer daytime. Our findings are consistent with previous studies showing that higher LST is associated with higher PD, which corresponds to a more fragmented pattern (Li et al., 2012; Zhang et al., 2009; Zhou et al., 2011). The superiority of the local Moran's I^v resides in its ability to provide continuous representation of the true heterogeneity of the

landscape that could not be conveyed by the discrete landscape metrics such as PD, which only considers the total number of patches within a unit area.

The results also show that clustered patterns of trees lower surface temperatures more effectively than aggregated grass patches in summer and winter daytime. This finding may be observed because clustered tree canopies shade the surface by absorbing incipient solar radiation, which substantially lowers the surface temperature. In addition, clustered trees may increase latent heat fluxes through evaporation, thereby reducing the sensible heat emitted from the surface. Spatial patterns show stronger impacts on daytime LST than nighttime LST. The greatest rates of urban surface cooling generally occur right after sunset and do not last long (Oke, 1982), which likely explains the weak relationships for nighttime LST, as images taken after 10:30 pm would hardly detect the large intra-urban temperature magnitudes.

4.3.2 Relationships between Spatial Pattern and LST by Vegetation Abundance

The spatial arrangement of vegetation is highly dependent upon the level of vegetation abundance. It is possible that the effects of spatial patterns on LST may be confounded by vegetation fraction had we not controlled for it. Figure 4.4 shows the results of regression analyses between the local Moran's I^v of grass and summer LST for every 10% increment of grass fraction. In general, the variance of LST explained by the local Moran's I^v increased as the level of grass abundance incremented during daytime. The spatial pattern of grass was strongly correlated with summer LST when grass fraction was greater than 50%. Note that the slope was steeper for grass fractions of 40%-50% and 50%-60% during daytime. This suggests that manipulating the spatial configuration of grass at the abundance level of ~50% is most effective in regulating the local climate.

Nighttime LST was weakly correlated with the local Moran's I^v for all fraction groups, except the one with more than 60% fraction. Overall, there was no strong relationship between the local Moran's I^v and summer LST when grass fraction was less than 40%. For landscapes with grass fraction greater than 40%, daytime LST was more sensitive to the variation in the spatial arrangement of grass than nighttime LST. Winter LST was weakly related to the local Moran's I^v for all fraction groups (results not shown here), suggesting that the spatial pattern of grass does not have significant effects on winter LST.

Figure 4.5 shows the relationship between the local Moran's I^v of trees and summer LST for every 10% increase in tree fraction. In contrast to grass, the variance in daytime LST explained by the local Moran's I^v did not increase as tree fraction increased. The predictive power of the models was comparatively weak as evidenced by the magnitude of the R^2. It is noteworthy that unlike grass, the regression coefficients for trees during daytime showed remarkable stability (between -9 to -11) for all fraction groups. This suggests that clustered pattern of trees lowers daytime LST effectively regardless of the amount of tree cover. Similar to grass, nighttime LST does not have a significant relationship with the spatial pattern of trees for all fraction levels.

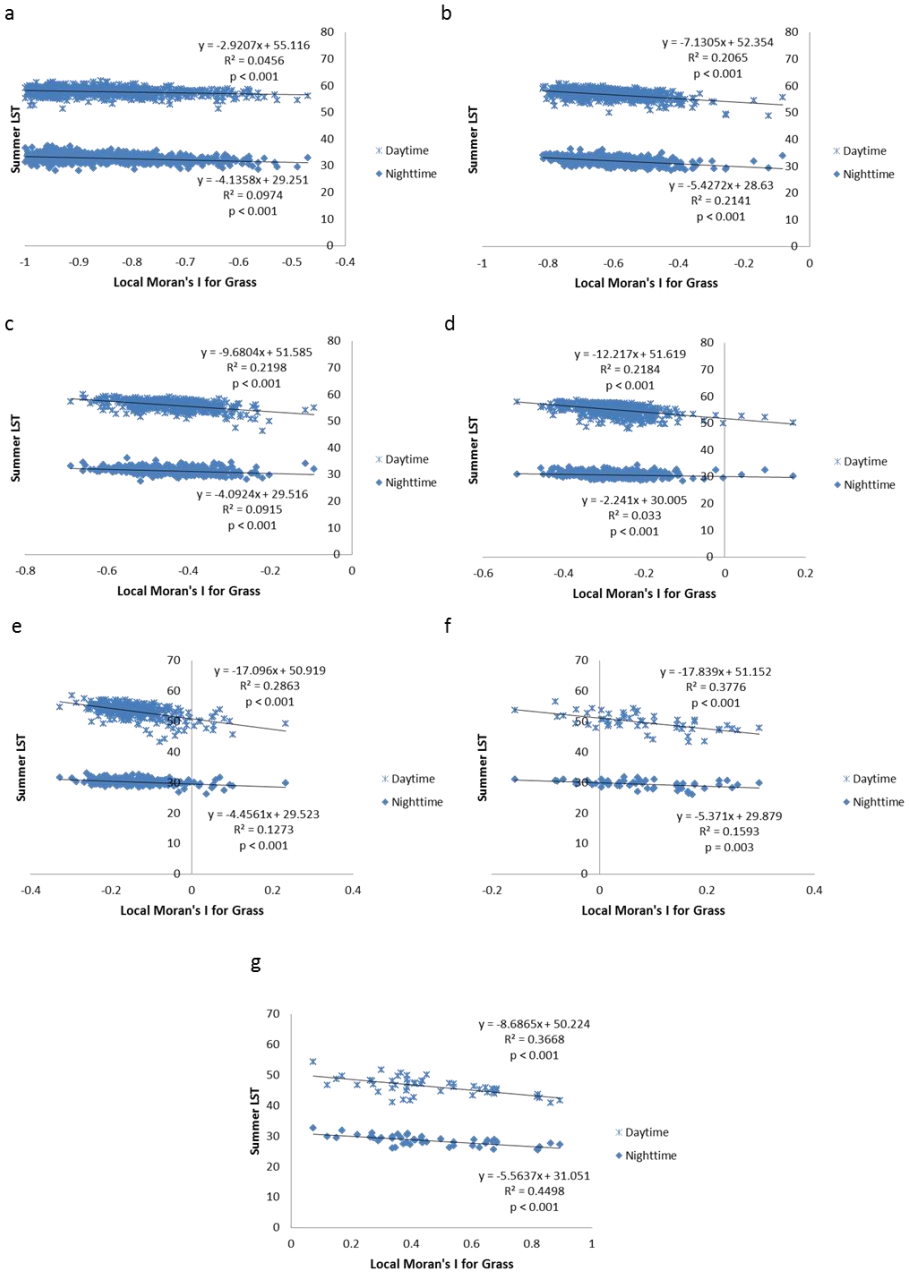

Figure 4.4 Bivariate Relationships between the Local Moran's I^v of Grass and Summer LST by Grass Fraction: a. Grass Fraction of 0-10%; b. Grass Fraction of 10%-20%; c. Grass Fraction of 20%-30%; d. Grass Fraction of 30%-40%; e. Grass Fraction of 40%-50%; f. Grass Fraction of 50%-60%; g. Grass Fraction of 60%-100% (Relationships for Daytime and Nighttime are Presented Separately on Each Diagram).

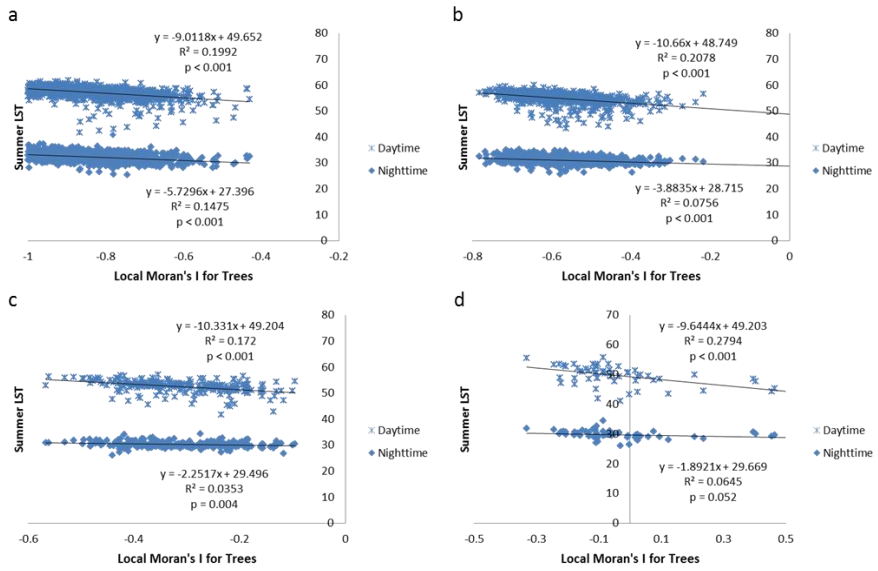

Figure 4.5 Bivariate Relationships between the Local Moran's I^v of Trees and Summer LST by Tree Fraction: a. Tree Fraction of 0-10%; b. Tree Fraction of 10%-20%; c. Tree Fraction of 20%-30%; d. Tree Fraction of 30%-100% (Relationships for Daytime and Nighttime are Presented Separately on Each Diagram).

These results highlight that aggregated patterns of urban vegetation lower local LST most effectively during summer daytime. Significant relationships were found between the spatial patterns of vegetation and LST, controlling for the percentage of vegetation cover. In general, our findings are in agreement with previous studies showing that less disturbed vegetation surfaces, rather than fragmented patterns, can effectively reduce the magnitude of LST (Li et al., 2012; Yokohari et al., 1997; Zhang et al., 2009). However, Zhou et al. (2011) quantified the spatial pattern of woody vegetation using mean nearest neighbour distance (MNND), and reported that a clustered pattern of woody vegetation tends to elevate LST. The contradictory results may be attributed to the confounding effects caused by the high correlation between MNND and the percentage of

vegetation (Zhou et al., 2011) and/or ignorance of the effects of vegetation abundance on the relationships between spatial pattern and LST. Thus, we suggest considering both factors to conduct a thorough evaluation of the pattern-LST relationship.

4.3.3 Results of Ridge Regression

To determine the combined and individual effects of spatial composition and configuration on seasonal LST, we developed multiple regression models. The results of the ridge regression for grass suggested that the local Moran's I^v and grass fraction jointly explained as high as 52% of the variance in summer LST during daytime, followed by 45.2%, 21.3% and 3.2% for summer nighttime, winter daytime, and winter nighttime, respectively (Table 4.2). These findings suggest that the combined effects of spatial pattern and composition account for more variance in LST than any predictor alone. The regression models also indicate that in spite of seasonal variation, LST had significant negative relationships with the local Moran's I^v and grass fraction. A key message conveyed here is that the absolute amount of grass plays a more important role in lowering seasonal LST than the spatial arrangement. This is indicated by the higher values of regression coefficients associated with grass fraction. Given a fixed amount of grass, higher values of the local Moran's I^v lowered the LST more effectively, highlighting the positive effects of aggregated patterns of grass in mitigating the UHI effect.

Table 4.2 Ridge Regression Results for Grass (n = 2,670).

	Coefficients	Std Error	*t* -statistic	*p* -value
Summer daytime (R^2 = 0.52)				
Intercept	55.344	0.045	1229.688	< 0.001
Local Moran's I^v	-2.891	0.055	-52.839	< 0.001
Grass fraction	-4.526	0.105	-43.196	< 0.001
Summer nighttime (R^2 = 0.452)				
Intercept	31.264	0.028	1121.433	< 0.001
Local Moran's I^v	-1.572	0.034	-46.374	< 0.001
Grass fraction	-2.423	0.065	-37.324	< 0.001
Winter daytime (R^2 = 0.213)				
Intercept	29.535	0.027	1078.049	< 0.001
Local Moran's I^v	-0.906	0.033	-27.806	< 0.001
Grass fraction	-1.356	0.062	-21.718	< 0.001
Winter nighttime (R^2 = 0.032)				
Intercept	10.397	0.033	316.732	< 0.001
Local Moran's I^v	-0.321	0.036	-8.787	< 0.001
Grass fraction	-0.675	0.07	-9.653	< 0.001

We also found negative correlations for both spatial pattern and composition with respect to LST for trees (Table 4.3). However, the total variance in LST explained by the local Moran's I^v and vegetation fraction together was less than that for grass. The highest R^2 was found in summer daytime, followed by winter daytime, summer nighttime, and winter nighttime (Table 4.3). The R^2 for winter nighttime was almost zero because nighttime LST during winter is consistently low regardless of the variation in either vegetation fraction or spatial patterns. The regression coefficients for fractional abundance are generally two times larger than those for the local Moran's I^v in all four

models. Clustered patterns of trees led to a decrease in LST when holding tree fraction constant.

Table 4.3 Ridge Regression Results for Trees (n = 2,648).

	Coefficients	Std Error	*t* -statistic	*p* -value
Summer daytime (R^2 = 0.476)				
Intercept	53.892	0.066	811.524	< 0.001
Local Moran's I^v	-4.112	0.087	-47.124	< 0.001
Tree fraction	-8.877	0.24	-37.039	< 0.001
Summer nighttime (R^2 = 0.256)				
Intercept	30.677	0.045	675.953	< 0.001
Local Moran's I^v	-1.765	0.06	-29.601	< 0.001
Tree fraction	-3.635	0.164	-22.19	< 0.001
Winter daytime (R^2 = 0.294)				
Intercept	28.983	0.034	843.943	< 0.001
Local Moran's I^v	-1.533	0.045	-33.784	< 0.001
Tree fraction	-3.32	0.125	-26.642	< 0.001
Winter nighttime (R^2 = 0.000)				
Intercept	10.44	0.042	246.58	< 0.001
Local Moran's I^v	-0.026	0.055	-0.474	0.636
Tree fraction	-0.164	0.151	-1.082	0.279

Our ridge regression models suggest significant negative relationships between the spatial arrangement of vegetation and LST. These results generally agree with previous findings using different approaches with a diverse selection of study areas (Li et al., 2012; Yokohari et al., 1997; Zhang et al., 2009). In addition, we demonstrate that the impacts of the spatial pattern of vegetation peak during summer daytime, whereas the effects are negligible during winter nighttime regardless of vegetation type. Our findings

have important implications for sustainable urban planning, especially for desert cities with limited water supplies. While it is difficult to increase the total area of greenspace in desert cities, optimising spatial patterns of vegetation will act as an alternative strategy to moderate local climate. This is especially applicable and valuable during the summer when excessive heat threatens cities and stresses water and energy services.

4.3.4 Scaling Effects

Figure 4.6 shows the statistical results for grass at varying geographical scales. We observed similar trends of R^2 for bivariate models and the multiple regression models as the window size changes. For example, the R^2 increases accordingly for summer daytime, summer nighttime, and winter daytime as the spatial extent increases. R^2 stayed fairly weak during winter nighttime across all the window sizes. The pattern of coefficients associated with the local Moran's I^v resembled that of R^2, showing an increment in their absolute magnitude (a decrease in value) for increasing scale. It is noticeable that for all three scalograms, the slope for summer daytime was steep between the window sizes of 100 m to 200 m and flattened afterward, suggesting that a window size of ~200 m is preferred in examining vegetation-LST relationships.

Because R^2 increases as the window size increases, we handled window sizing with caution as an overly large window could potentially lead to a substantial loss of information. This is because degrading both the local Moran's I^v and LST maps will likely result in biased inferences and inaccurate interpretation (Kutner et al., 2005).

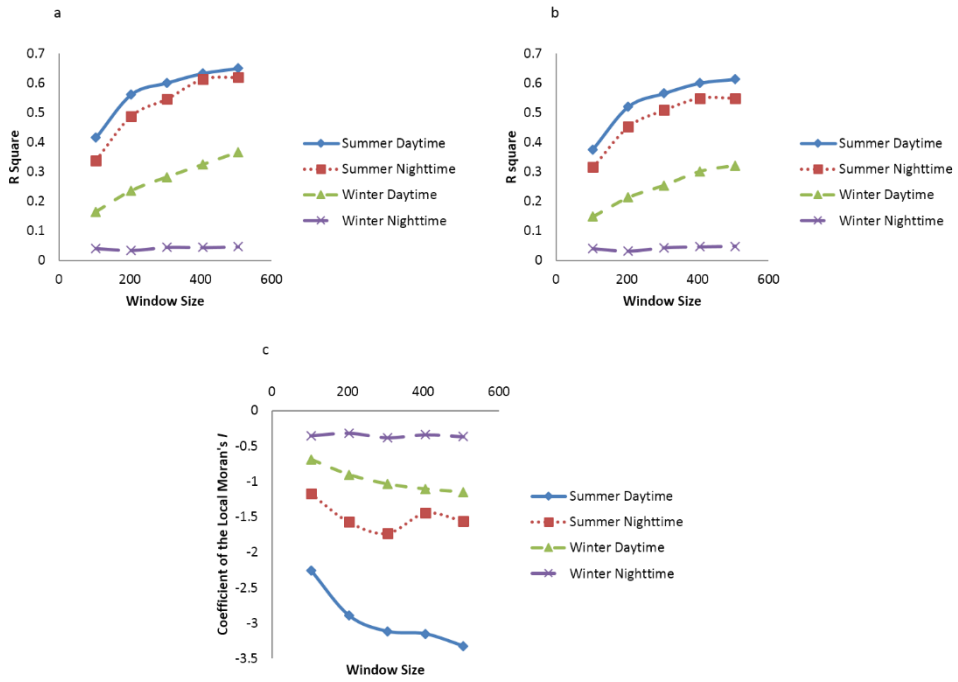

Figure 4.6 Scalograms Showing the Effects of Changing Window Size on the Relationship between the Local Moran's I^v of Grass and Seasonal LST: a. R^2 of the Bivariate Model vs. Window Size; b. R^2 of the Ridge Regression Model vs. Window Size; c. Coefficient of the Local Moran's I^v of Grass in the Ridge Regression Model vs. Window Size.

While the magnitudes of the R^2 for the model between the local Moran's I^v of grass and LST are much higher for summer nighttime than for winter daytime, they are similar for trees (Figure 4.7). Interestingly, the relationship recorded for winter daytime stayed consistently stronger than summer nighttime across all window sizes for trees (Figures 4.7a and b). This implies that the spatial pattern of trees plays a more important role in shaping the local climate during winter daytime than summer nighttime. Similarly to grass, the relationships for winter nighttime were very weak with low R^2 (~0) for both

the bivariate and ridge regression models. The coefficients of the local Moran's *I* signified a very interesting pattern with a trough at the window size of 200 m. This again confirms that a reasonable window size for studying vegetation-LST relationships is ~200 m. The scaling effects examined here are consistent with previous findings that an area of ~200 m is optimal when examining the cooling effects of urban vegetation on micro-scale climate.

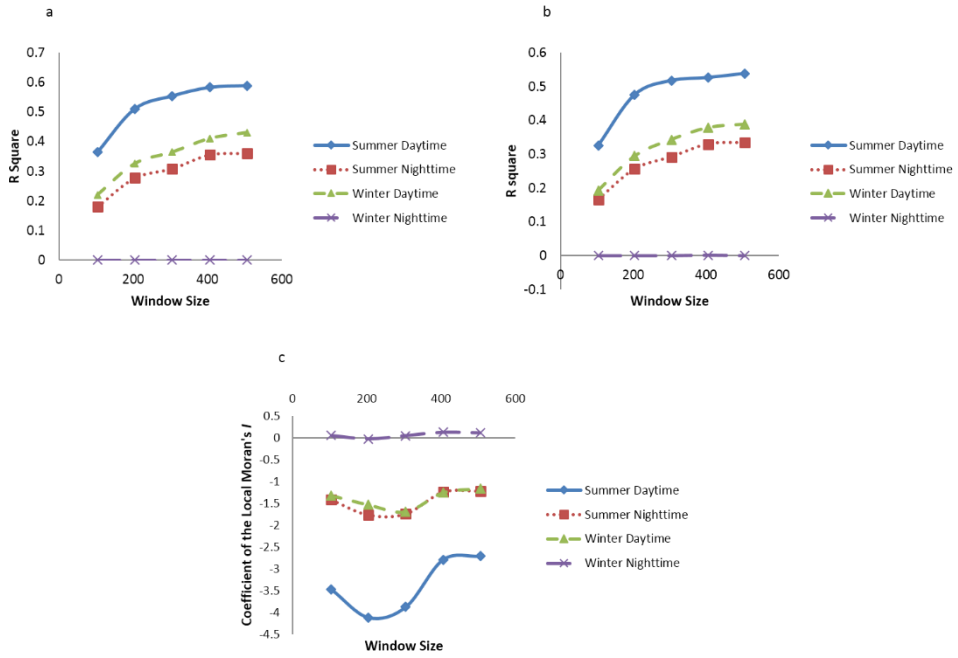

Figure 4.7 Scalograms Showing the Effects of Changing Window Size on the Relationship between the Local Moran's I^v of Trees and Seasonal LST: a. R^2 of the Bivariate Model vs. Window Size; b. R^2 of the Ridge Regression Model vs. Window Size; c. Coefficient of the Local Moran's I^v in the Ridge Regression Model vs. Window Size.

4.3.5 Implications for Management and Urban Planning

Vegetation plays a critical role in moderating urban climate. Greenspace is more crucial for desert regions with high temperatures and sparse rainfall. While increasing vegetation abundance can be an effective strategy for UHI mitigation, it is problematic in arid regions as increased vegetation cover also causes increased municipal water consumption, which contradicts long-term sustainability goals. To conserve water resources, some desert cities such as Las Vegas have recently initiated "cash-for-grass" programs, rewarding homeowners for replacing their "green" yards with less water-demanding landscapes (Sovocool et al., 2001). However, this landscape conversion strategy might result in elevated air temperatures that are likely to cause increased demands on water and energy use (Ruddell and Dixon, 2013). Given that expanding green areas or xeriscaping alone are proving to be insufficient UHI mitigation strategies, optimising the spatial configuration of urban vegetation is an alternative for sustainable urban design.

Our results show negative relationships between local Moran's I^v and seasonal LST for both grass and trees, implying that less fragmented or spatially compact vegetation mosaics are effective for reducing urban temperatures. Furthermore, the impact of spatial vegetation patterns on LST are most pronounced during summer daytime, suggesting that the judicious design of vegetation patterns can produce significant summer cooling effects. This is particularly desirable in desert cities as summer cooling greatly reduces the huge demands on water and energy use. This research shows that planning large and compact patches of vegetated surfaces in urban systems can be one of the most effective and easy-to-implement strategies of UHI

mitigation in comparison to other options. For example, it is impractical and cost prohibitive to install cool roofs or convert existing roofs to high-albedo surfaces across an entire urban area. In addition, the effectiveness of high-albedo roofs may be reduced over time given that dust storms frequently occur in desert regions (Chow et al., 2012).

UHI solutions vary due to numerous factors largely determined by the geographical location of the city. For example, cities in temperate or tropical regions can increase urban greenery without stressing their water supply, as this is not the most critical issue in these regions. Some considerations shape the final decision more than others. Therefore, trade-offs between temperature, water, and energy use are always worth investigating. It is highly recommended that cost-benefit analyses be conducted before implementing any mitigation programs. A few other considerations such as maintenance, pest control, and safety are also warranted.

4.3.6 Quantifying Spatial Pattern from a Continuous Perspective

Spatial pattern analysis is an important and useful technique for identifying the arrangement of geographical features in space. Using patch mosaic models, recent efforts have been devoted to the assessment of landscape patterns in relation to local climate (Liu and Weng, 2008; Zhou et al., 2011; Li et al., 2012), spread of disease (Liu and Weng, 2009), water quality (Griffith et al., 2002; Lee et al., 2009), and residential value estimation (Geophegan et al., 1997). Although discrete models appear to be effective in characterising different landscape patterns, the limitations associated with the classification of continuous environmental variables into discrete categories greatly undermine the reliability and applicability of such models (Fan and Myint, 2014; McGarigal and Cushman, 2005).

As continuous models more accurately represent landscape heterogeneity, they are more suitable for quantifying landscape structure with the potential of addressing pattern-process interactions. Recent studies have demonstrated the effectiveness of continuous models in a number of landscape ecological applications. It is beneficial to apply these models to a broader range of fields with the aim of addressing the linkage between spatial patterns of land cover features and the underlying processes that are potentially influenced by the variation in landscape pattern. Future studies should expand on our study by utilising continuous models for understanding the potential impacts of landscape pattern on a wider range of biophysical and socio-economic factors. For example, a similar approach could be used to investigate the relationships between the spatial arrangement of vegetation and head height temperature or humidity, which could elucidate UHI mitigation and sustainable urban design.

4.4 Conclusion

Previous research has shown that urbanisation is one of the main drivers of climate change (Quattrochi and Ridd, 1994; Zhou et al., 2011) and that green vegetation plays a key role in moderating urban environments and mitigating the UHI effect. This research expands on previous studies investigating the cooling effect of vegetation abundance on LST by considering the impacts of the spatial pattern of green vegetation, especially grass and trees, on urban surface temperatures at varying geographical scales in Phoenix, Arizona. Negative relationships between the local Moran's I^{v} and LST suggest that clustered or less fragmented patterns of vegetation lower seasonal LST more effectively than dispersed patterns. Throughout the four scenes, LST shows more sensitivity to spatial patterns of green vegetation during summer daytime whereas winter

nighttime LST is weakly correlated to the local Moran's I^v. Our results indicate that the spatial arrangement of vegetation significantly affects the area's seasonal LST, regardless of vegetation type. The results from multiple regression analysis also show that both vegetation abundance and spatial arrangement play important roles in lowering seasonal LST. Our experiments on the scaling effect suggest that a reasonable window size for examining the vegetation-LST relationship is ~200 m.

From these results, we conclude that optimising the spatial arrangement of green vegetation improves the urban environment and effectively mitigates UHI and that despite seasonal variation, aggregated, rather than dispersed, patterns of grass and trees are preferred for cooling the environment. Furthermore, planning clustered patches of vegetation conforms to city sustainability efforts as it ameliorates temperatures without requiring substantial consumption of urban resources. It is important to remember, however, that vegetation-LST relationships are likely to vary from location to location due to differing climate conditions, land cover characteristics, landforms, and topography. Therefore, it would be interesting to conduct similar studies in multiple cities with diverse physical conditions. The spatial autocorrelation technique proposed is an early attempt to understand the pattern-process relationship using continuous models. It would be beneficial to expand the applications of this technique by linking the spatial patterns of other land cover types (e.g., buildings, pavements, and soil) to numerous biophysical parameters (e.g., air temperature, air quality, water use, evapotranspiration, and humidity) and socio-economic factors (e.g., human comfort, income, housing value, and disease).

CHAPTER 5

SPATIALLY FILTERED RIDGE REGRESSION (SFRR): A REGRESSION FRAMEWORK TO UNDERSTANDING IMPACTS OF LAND COVER PATTERN ON URBAN MICROCLIMATE

5.1 Introduction

Urbanization is one of the most drastic forms of human modifications to the environment. The replacement of natural landscapes with impervious man-made infrastructure has dramatically altered the surface energy balance of urban environment, causing elevated temperatures in the urban area relative to its rural surroundings; this is referred to as an urban heat island (UHI) (Lo and Quattrochi, 2003; Oke, 1987). The UHI has important and extensive implications for urban ecosystems, energy demands, and residents' quality of life. Increased temperature in urban areas can give rise to increased demands for water and energy use (Arnfield, 2003; Guhathakurta and Gober, 2007;), elevated concentration of air pollutants (Lai and Cheng, 2009; Sarrat et al., 2006), and a heightened risk of heat-related mortality (Clarke, 1972; Conti et al., 2005; Ellis et al., 1975; Tan et al., 2010; Whitman et al., 1997). Since its first discovery in 1818, the UHI has been a primary focus in climatology and urban ecology (Arnfield, 2003; Howard, 1833). There is an ever-increasing number of studies focusing on various aspects of the UHI, including its formation, development, as well as mitigation strategies to counter it (Oke, 1982; Onishi et al., 2010; Rizwan et al., 2008).

The UHI can be broadly categorized into two types: atmospheric UHI and surface UHI. The two types of UHI differ in the way they form and develop, and different sets of techniques are required to measure them. Atmospheric UHI is measured at local scales by virtue of *in situ* instruments, while surface UHI is usually identified at regional and global scales using indirect measurement such as remote sensing. The advancement of remote sensing techniques has provided an opportunity for measuring surface UHI at an unprecedentedly large spatial scale. Meanwhile, it allows research into the magnitude and pattern of surface UHI as ecological responses to a variety of biophysical and socioeconomic indicators at a per-pixel basis.

UHI, particularly surface UHI, is closely related to the land surface characteristics, or more specifically, the land composition and spatial configuration of land use/land cover (LULC) features (Arnfield, 2003; Buyantuyev and Wu, 2010; Rotem-Mindali et al., 2015; Voogt and Oke, 2003). Green vegetation and impervious surfaces are important land cover types that affect the local climate. Green vegetation ameliorates surface temperatures through evapotranspiration and shading (Gallo et al., 1993; Taha, 1997; Tayyebi and Jenerette, 2016; Weng et al., 2004). Impervious surfaces, such as pavements covered by asphalt, brick, and concrete, absorb short-wave radiation during the day and slowly release the heat at night, making it slow to cool down the city (Mallick et al., 2013; Xiao et al., 2007; Yuan and Bauer, 2007). Not only land cover abundance affects the surface temperatures, the spatial arrangements of land cover features matter as well. Several studies have looked into the associations between the spatial patterns of different LULC types and surface temperatures (Li et al., 2012; Maimaitiyiming et al., 2014) with suggestion that urban temperature varies as a function

of the shape, size, and spatial configuration of land cover patches, with stronger relationships found for vegetation and anthropogenic features (Fan et al., 2015; Li et al., 2012; Yokohari et al., 1997; Zheng et al., 2014; Zhou et al., 2011).

While a vast amount of research has considered urban LULC patterns as important contributors to the UHI, findings from these studies do not always converge. In fact, a comprehensive review of the UHI literature reveals considerable inconsistency in the strength and direction of the relationship between LULC patterns and land surface temperatures (LST) (Zhou et al., 2011; Zhang et al., 2009; Li et al., 2012). We attribute the discrepancy in part to the idiosyncratic characteristics (e.g., climate condition, economic status, policy) of each city or study region. More importantly, a large portion of the inconsistency is due to the varying methods employed in different case studies. The most commonly used method in UHI studies is the ordinary least square (OLS) regression. While the simplicity of the OLS makes it a tempting method, its basic assumptions are often not taken into consideration, leading to unreliable parameter estimates and poor predictions.

We spend the next section discussing the two fundamental issues arising from but not constrained to the studies of the UHI and provide clear reasons why the most popular OLS is flawed when it is used for spatial data coupled with highly correlated independent variables. In section three, we introduce an integrated methodological framework whose validity and effectiveness are tested using a series of simulation experiments. Section four reports the results of the simulation. For two forms of spatial dependence, our method is compared to three candidate models based on three common criteria for estimators. Section five describes a case study where the new method is applied to an empirical data

set as a simple demonstration of our method to address the LULC-LST relationship in a desert city. Section six summarizes the work and discusses the limitations and future directions.

5.2 Background

The OLS is arguably one of the most widely employed statistical tools by researchers from a broad range of disciplines. It is a standard method for estimating parameters in a linear model. The OLS estimator has desirable properties (e.g., unbiased, efficient) when all the standard regularity conditions are met. When one or more conditions are violated, some of the properties do not hold, making OLS a less preferred and sometimes an inferior estimator.

Spatial dependence is a unique characteristic of geographic data. It can be considered as the functional relationship between the attributes of a phenomenon that occurs at one location and those that occur elsewhere (Anselin, 1988). For instance, remote sensing satellites record natural radiation emitted or reflected from continuous surfaces, which is later sampled into regularly spaced and equally sized lattices. The values on these lattices depend to a great extent on their geographical locations. As stated in the first law of geography, "everything is related to everything else, but near things are more related than distant things." (Tobler, 1979).

Spatial effects arise as a result of spatial data manipulation (e.g., spatial aggregation, arbitrary delineation of unit boundaries) frequently performed in applied work. Because of the spill-over effect, errors in one spatial unit are related to those in adjacent units. This makes the assumption of no error correlation invalid. A second cause of spatial dependence pertains to the fundamental properties inherent in a variety of

spatial interaction processes. As a result, what happens at one location is determined by what happens at other locations in the system (Anselin, 1988). This form of spatial dependence is often addressed by including a function of the dependent variable observed at other locations as an additional regressor. Because the autoregression term is determined by disturbances at other locations, it cannot be considered as fixed in repeated samples. The assumption of fixed independent variables in repeated samples is therefore violated.

Spatial dependence in geographic data poses both an opportunity and a challenge. Spatial autocorrelation in the data constitutes an important source of auxiliary information that has been evidenced to be useful in image classification (Wulder and Boots, 1998), noise deduction (Switzer and Ingebritsen, 1986), feature selection (Warner and Shank, 1997), and sensor calibration (Bannari et al., 2005). Conversely, the presence of spatial effects causes the standard conditions of OLS invalid, making the OLS estimator a less efficient estimator (Jokar Arsanjani et al., 2013).

Several solutions have been suggested by ecologists to address the spatial effects in the data. Legendre (1993) proposed two methodological frameworks—the "raw data approach" and the "matrix approach"—to incorporate spatial structure into ecological modeling. The "raw data approach" utilizes partial regression analysis to model the species-environment relationship for individual species and constrained ordination analyses for multivariate cases such as community analysis. The "matrix approach", on the other hand, models the spatial structure by a distance matrix (e.g., a connection matrix), and calculates the correlation between environment and species using a partial Mantel test (Manly, 1986). The two approaches are effective for broad-scale

autocorrelation (i.e., spatial trend), but have limited ability to account for fine-scale autocorrelation, which is more often encountered in ecological studies (Lichstein et al., 2002).

The class of spatial autoregressive models can be utilized to analyze the relationship when fine-scale autocorrelation is present (Cressie, 1993; Haining, 1990; Anselin, 1988). It incorporates spatial dependence via a spatial weights matrix that specifies the strength of spatial interaction between each spatial unit and its neighbors (Anselin, 1988; Anselin and Rey, 1991; Cressie, 1993). Spatial autoregressive models can take two general forms depending on different processes that cause spatial dependence. From a substantive perspective, if the process under consideration includes direct interactions between observational units in space, the relevant specification is the so-called spatial lag model. This affords the modeling of these spillovers between observations at neighboring locations through the inclusion of a function of the dependent variable observed at other locations as an additional explanatory variable. A second form of spatial dependence specification is the spatial error model, which is appropriate when forms of measurement errors or omitted/unobservable variables create spatial autocorrelation between proximal observations. Here the spatial dependence is handled through the error term of the model rather than as an additional regressor. Despite their applications to statistics, regional science, and other social sciences, spatial autoregressive models have yet to be embraced by the environmental science community, with a few exceptions (Ji and Peters, 2004; Li et al., 2012; Lichstein et al., 2002; Miller et al., 2007; Overmars et al., 2003).

A second problem of using OLS in UHI studies relates to the issue of multicollinearity, which refers to a phenomenon where independent variables in a regression model are highly correlated. When multicollinearity is present, both the OLS estimator and R^2 remain unbiased. However, the variances of OLS estimators associated with collinear variables are inflated. Because of the inflated variances, significant variables may appear insignificant (Ohlemüller et al., 2008). Also due to the high correlations among predictors, independent interpretation of coefficients is no longer valid. Very few studies have made attempts to address multicollinearity in their analyses despite the relevance of the issue (exceptions include Chestnut et al., 1998, Hart and Sailor, 2009, and Myint et al., 2010), in part due to (1) unawareness of how multicollinearity affects parameter estimation, (2) uncertainty about how to remediate the problem, and (3) inaccessible software tools (Dormann et al., 2013).

The problem of multicollinearity is intractable due to the inability of statistical means to discriminate highly correlated predictor variables. Several approaches have been developed to achieve robust parameter estimation and accurate model prediction. Dormann et al. (2013) reviewed a variety of approaches to multicollinearity problems, ranging from predictor clustering (e.g., principle component analysis (PCA), Hoeffding/Ward-clustering), through latent variable methods (e.g., principle component regression (PCR), partial least squares (PLS)) to shrinkage techniques (e.g., ridge regression, LASSO), and compared their performances using a series of simulations. Among all the methods, shrinkage regression was found to perform particularly well for models with different complexity levels of functional relationship and different degrees of multicollinearity. Lazaridis et al. (2010) compared standard regression methods to

shrinkage regression techniques in predicting tree mortality using Moderate-Resolution Imaging Spectroradiometer (MODIS) imagery. The study showed strong potential of shrinkage regression methods in dealing with multicollinearity in remote sensing data sets, such as collinearity among hyperspectral satellite bands and high correlations among biophysical variables.

Our article presents a novel methodological framework called spatially filtered ridge regression (SFRR) that investigates and addresses the preceding problems. Through systematic integration of spatial autoregressive models and ridge regression, SFRR is capable of producing more reliable parameter estimates than the commonly adopted statistical models. We test the utility of SFRR using a series of simulation experiments, followed by an empirical study over central Phoenix. By comparing and contrasting the performances of four candidate models, three important questions are addressed: (1) what are the consequences of ignoring spatial dependence and/or multicollinearity when modeling the UHI-land cover relationship? (2) How does SFRR overcome these problems and how does it perform compared to other candidate models in terms of parameter estimation and model adequacy? (3) Using the proposed method, what are the respective impacts of land composition and spatial configuration of LULC features on the surface temperatures in central Phoenix?

5.3 Methodology

5.3.1 Spatial Autoregressive Models

As mentioned previously, the spatial lag model extends the classic regression model to include an additional variable that expresses the spillover between neighboring locations. More formally:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \boldsymbol{\mu} \tag{5.1}$$

where $\mathbf{y}$ is an $n \times 1$ vector of observations on the dependent variable, ρ is the spatial lag coefficient, and $\mathbf{W}\mathbf{y}$ is the spatial lag with $\mathbf{W}$ an $n \times n$ spatial weights matrix expressing the neighbor relations between each pair of observations. $\mathbf{X}$ is an $n \times k$ design matrix with observations on the traditional explanatory variables with associated parameter vector β. $\boldsymbol{\mu}$ is a well-behaved error term: $\boldsymbol{\mu} \sim N(0, \sigma^2\mathbf{I})$ with $\mathbf{I}$ being the identity matrix. The inclusion of the spatial lag on the right hand side of this specification introduces a form of endogeneity—correlation between predictor variables and the error term—that needs to be taken into account by an appropriate estimator, such as maximum likelihood or generalized method of moments.

The spatial error model structures the spatial autocorrelation as part of the error term:

$$\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\mu} \tag{5.2}$$

where now

$$\boldsymbol{\mu} = \lambda\mathbf{W}\boldsymbol{\mu} + \epsilon \tag{5.3}$$

or

$$\boldsymbol{\mu} = (\mathbf{I} - \lambda\mathbf{W})^{-1}\epsilon \tag{5.4}$$

In contrast to the well-behaved error term in the lag model above, the error term here has a variance-covariance matrix:

$$E[\boldsymbol{\mu}\boldsymbol{\mu}'] = \sigma^2(\mathbf{I} - \lambda\mathbf{W})^{-1}(\mathbf{I} - \lambda\mathbf{W}')^{-1} \tag{5.5}$$

Due to the non-spherical nature of this covariance matrix, OLS coefficient standard error estimates will be biased; although, the coefficient estimates themselves remain unbiased.

5.3.2 Ridge Regression

Ridge regression shrinks the regression coefficients by applying a constraint to the OLS estimates (Hoerl and Kennard, 1970). It involves the addition of a penalty of constant to the size of coefficient estimate while minimizing the sum of squared residuals. Ridge regression increases precision of parameter estimation, but because of the constant, ridge estimator is no longer unbiased.

Our point of departure is the correlation transformation, which is a simple standardization procedure that facilitates comparison of regression coefficients by expressing them in the same units. Mathematically, the correlation transformation is given by

$$Y^* = \frac{1}{\sqrt{n-1}}\left(\frac{Y - \bar{Y}}{s_Y}\right) \tag{5.6}$$

$$X_k^* = \frac{1}{\sqrt{n-1}}\left(\frac{X_k - \overline{X_k}}{s_k}\right) \tag{5.7}$$

where s_Y and s_k are the respective standard deviations for Y and each of the predictor variable. k ranges from 1 to $p-1$ where p is the total number of parameters. It can be shown that the $\mathbf{X}'\mathbf{X}$ matrix for the transformed variables is identical to the correlation matrix of pairwise correlations between each predictor variable, or $\mathbf{X}'\mathbf{X} = \mathbf{r}_{XX}$,

and the $\mathbf{X}'\mathbf{Y}$ matrix for the transformed variables is simply a vector of correlation between the dependent variable and each predictor variable, or $\mathbf{X}'\mathbf{Y} = \mathbf{r}_{YX}$. Therefore, the least squares normal equations are now given by

$$\mathbf{r}_{XX}\mathbf{b} = \mathbf{r}_{YX} \tag{5.8}$$

where **b** is the vector of the point estimates of the standardized regression coefficients.

The ridge standardized estimators are obtained by adding a bias constant *c* to the normal equations:

$$(\mathbf{r}_{XX} + c\mathbf{I})\mathbf{b}_{\boldsymbol{R}} = \mathbf{r}_{YX} \tag{5.9}$$

Solution of Eq. (5.9) yield the point estimates of the standardized ridge regression coefficients:

$$\mathbf{b}_R = (\mathbf{r}_{XX} + c\mathbf{I})^{-1}\mathbf{r}_{YX} \tag{5.10}$$

The value on the bias constant *c* needs to be sought before parameter estimation. Several strategies have been suggested for determining the optimal value of the bias constant, such as minimizing the mean squared error of estimation (Kasarda and Shih, 1977) and minimizing the prediction error on a new data set via bootstrapping or cross-validation (Golub et al., 1979). Another commonly adopted approach makes use of a ridge trace (Hoerl and Kennard, 1970), which is a plot of the ridge coefficients versus the bias constant. Generally, the size of regression coefficients tends to approach zero as the bias constant increases. The optimal bias is the smallest value of bias when ridge coefficients start to stabilize. This is the approach we used in the simulations and the empirical study. Once the bias constant is determined, the ridge standardized coefficients

can be obtained using Eq. (5.10). The standardized coefficients need to be transformed back to the original variables for interpretation purposes.

5.3.3 SFRR

The SFRR is a simple combination of two techniques: spatial regression and ridge regression. It effectively addresses spatial dependence and multicollinearity within a single framework. We developed this method for both spatial lag and spatial error dependence, the use of which is determined by the specification search outlined in Florax et al. (2003). Simply put, the specification search used in this study makes use of a set of decision rules in search of the alternative model. As straightforward as they are, the decision rules have been shown to be effective based on considerable evidence from many simulation experiments (Anselin and Rey, 1991; Florax et al., 2003).

5.3.3.1 SFRR for Spatial Lag Dependence

Recall Eq. (5.1) for the general form of a spatial lag model. Alternatively, it can be considered as removing spatial autocorrelation effect via a spatial filter ($\boldsymbol{I} - \rho\mathbf{W}$):

$$(\mathbf{I} - \rho\mathbf{W})\mathbf{y} = \mathbf{X}\beta + \boldsymbol{\mu} \tag{5.11}$$

The endogeneity in the model specification requires the adoption of other estimation methods, such as maximum likelihood and spatial two stage least squares. Once an appropriate estimate of ρ is obtained, it is substituted into $(\mathbf{I} - \rho\mathbf{W})\mathbf{y}$, which is now a vector of observations on the spatially filtered dependent variable. The transformed variable is treated as the new dependent variable in the ridge regression that follows. The standardized SFRR estimator for spatial lag dependence can be expressed as

$$\mathbf{b}_{SFRRL} = (\mathbf{r}_{XX} + c\mathbf{I})^{-1}\mathbf{r}_{Y'X} \tag{5.12}$$

where $\mathbf{r}_{YX}$ is a vector of Pearson's correlation between the spatially filtered dependent variable and each predictor variable. The bias constant is determined following the description in Section 5.3.2.

5.3.3.2 SFRR for Spatial Error Dependence

After substituting the reduced form (Eq. (5.4)) into the general form (Eq. (5.2)), the spatial error specification with a spatial autoregressive process (SAR) process can be expressed as

$$\mathbf{y} = \mathbf{X}\beta + (\mathbf{I} - \lambda\mathbf{W})^{-1}\epsilon \tag{5.13}$$

or further

$$(\mathbf{I} - \lambda\mathbf{W})\mathbf{y} = (\mathbf{I} - \lambda\mathbf{W})\mathbf{X}\beta + \epsilon \tag{5.14}$$

As with the spatial lag model, $(\mathbf{I} - \lambda\mathbf{W})$ serves as a spatial filter that removes the structure (spatial autocorrelation) from the error term. This functions the same way as a Cochrane-Orcutt estimation to remove serial correlation in a time series model (Anselin and Rey, 2014). Commonly used methods to estimate coefficients for a spatial error specification include maximum likelihood and generalized method of moments. The estimated λ is substituted back into $(\mathbf{I} - \lambda\mathbf{W})\mathbf{y}$ and $(\mathbf{I} - \lambda\mathbf{W})\mathbf{X}$, which are treated as the dependent and independent variables in the subsequent ridge regression model. The standardized SFRR estimator for spatial error dependence is

$$\mathbf{b}_{SFRRE} = (\mathbf{r}_{X'X'} + c\mathbf{I})^{-1}\mathbf{r}_{Y'X'} \tag{5.15}$$

where X' and Y' represent the spatially filtered independent and dependent variables, respectively. The workflow of the SFRR framework for spatial lag and spatial error dependence is depicted in Figure 5.1.

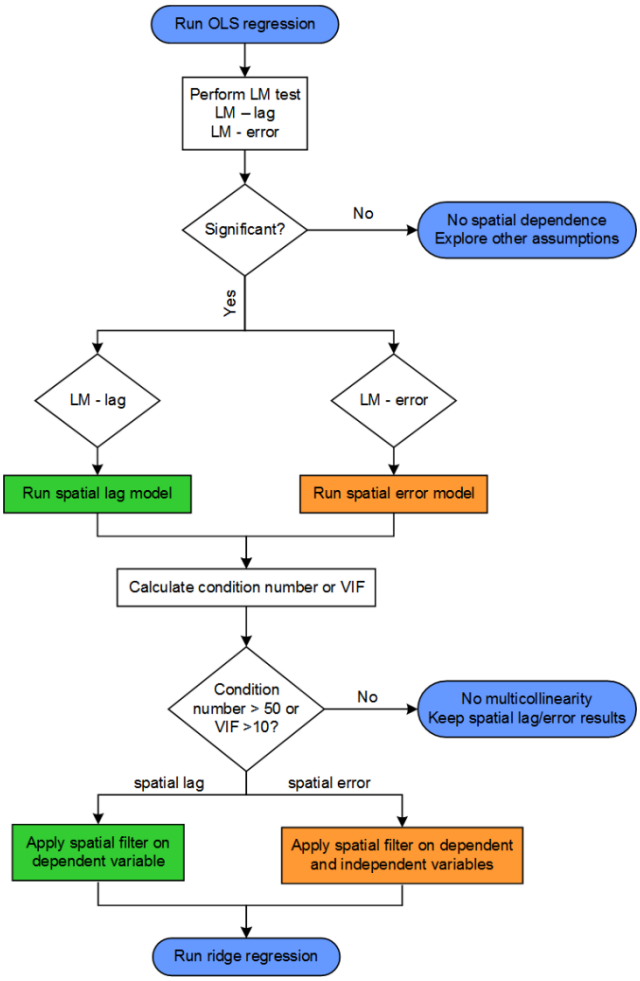

Figure 5.1 Overview of the SFRR Framework.

5.3.4 Simulation Experiments

5.3.4.1 Independent Variables

We used ten LULC variables as our independent variables. These variables were derived from a cloud-free QuickBird image acquired on May 24, 2007 over central Phoenix. The spatial resolution of the image is 2.4 meters (m). We employed an object based image analysis, which yielded six LULC categories: trees/shrubs, grass, buildings, other impervious surfaces, unmanaged soil, and water (Figure 5.2). We decided to

exclude water from our analysis because it is weakly related to the surface temperatures within the study area (Myint et al., 2013). The error matrix for the object-based classifier was provided in Table 5.1.

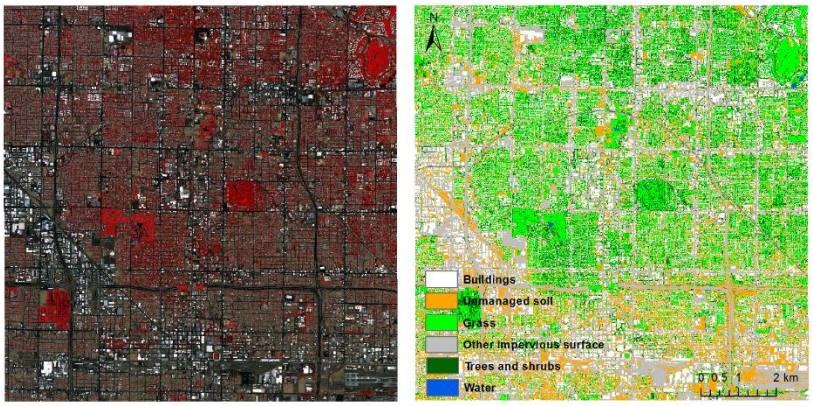

Figure 5.2 a. QuickBird Imagery over Central Phoenix, AZ; b. Classified Output of the QuickBird Imagery.

Table 5.1 Error Matrix of the Urban Land Cover Classification Produced by the Object-based Classifier

Classified	Reference						Producer's	User's
	Buildings	Grass	Trees/shrubs	Other impervious	Unmanaged soil	Total	accuracy (%)	accuracy (%)
Buildings	73	1	1	3	2	80	84.88	91.25
Grass	6	68	8	2	2	86	94.44	79.07
Trees/shrubs	0	3	56	8	0	67	86.15	83.58
Other impervious	1	0	0	87	0	88	84.47	98.86
Unmanaged soil	6	0	0	3	70	79	94.59	88.61
Total	86	72	65	103	74	400		

Overall accuracy = 88.5%

For each LULC category, we created two landscape variables—land cover fraction and local Moran's *I*—that measure the composition and configuration of each class within the landscape, respectively. Land cover fraction was used as an indicator of land composition. By definition, land cover fraction measures the proportion of a LULC

type in a predefined spatial extent. Existing literature suggests that 200 m was an optimal extent for maximized effects of LULC features on the urban climate (Fan et al., 2015; Kormann and Meixner, 2001; Myint et al., 2010). We followed this criterion and computed land cover fractions within a window size of 204 m (2.4 m × 85 = 204 m).

The local indicator of spatial association (LISA) evaluates the existence of spatial clusters and degree of clustering for a given variable at a local scale (Anselin, 1995). A number of studies have demonstrated the potential of LISA to measure spatial arrangements of land cover features within desert landscapes (Fan and Myint, 2014; Fan et al., 2015; Myint et al., 2015; Zheng et al., 2014). In this study, we utilized a refined version of LISA as our spatial configuration variable.

We calculated the local Moran's *I* based on the binary map created for each LULC category. The same set of binary maps also functioned as a local average filter and was applied to the map of the local Moran's *I* to remove the impacts from other types of cover. The final output was a raster grid with each pixel representing the spatial arrangement of a particular LULC type within an area of 204 m by 204 m (See Fan et al., 2015 for discussions, formulas, and detailed computation procedures). In general, high and low values of local Moran's *I* indicate clustered and dispersed patterns of land cover, respectively, with values near zero indicating random patterns (note: clusters of medium values do not happen in the binary case).

The ten LULC variables—five composition variables and five configuration variables—constitute our predictor set in the simulation. As area proportions of different land cover types sum up to a fixed quantity, we would expect certain degree of multicollinearity among the ten predictors. Furthermore, the inherent association between

land composition and configuration aggravates the multicollinearity because the higher the abundance of a LULC type, the more likely that this type of land cover tends to cluster.

The degree of multicollinearity was evaluated using condition number, which measures the degree to which predictor variables have a linear relationship. The condition number is defined by

$$k(\mathbf{X}'\mathbf{X}) = \sqrt{\frac{\lambda_{max}(\mathbf{X}'\mathbf{X})}{\lambda_{min}(\mathbf{X}'\mathbf{X})}} \tag{5.16}$$

where $\lambda_{max}(\mathbf{X}'\mathbf{X})$ and $\lambda_{min}(\mathbf{X}'\mathbf{X})$ are the largest and smallest eigenvalues of $\mathbf{X}'\mathbf{X}$, respectively ($\mathbf{X}$ is the design matrix). Note that the condition number was calculated after standardizing the design matrix such that the norm of each column is 1 (Anselin and Rey, 2014). Generally, a condition number of higher than 30 or 50 indicates the presence of mulcollinearity. Our predictor variables produced a condition number of 328.05, which is well-beyond the threshold above, indicating severe multicollinearity.

5.3.4.2 Dependent Variable

We considered two forms of spatial dependence (spatial lag and spatial error effect) as we created the dependent variable. In the spatial lag specification, the dependent variable $\mathbf{y}$ can be generated from the reduced form:

$$\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\beta + (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{u} \tag{5.17}$$

where β was set to be a vector of ones (i.e., equal impact from each predictor variable), and $\mathbf{u}$ is random disturbance with Gaussian distribution (StDev = 1.0).

For spatial error dependence, we limited our method to the case where the disturbance follows a SAR error process. The response variable **y** with spatial error dependence can be generated as:

$$\mathbf{y} = \mathbf{X}\beta + (\mathbf{I} - \lambda\mathbf{W})^{-1}\epsilon \tag{5.18}$$

where β and ϵ were created in the same manner as the β and **u** in the spatial lag case.

5.3.4.3 Simulation Parameters

We tested varying degrees of spatial dependence by using an array of values of ρ (in the spatial lag case) and λ (in the spatial error case) as we created the synthetic data. The value of the spatial autoregressive coefficient ranged from 0.1 to 0.9 with an increment of 0.2. For each level of spatial dependence, we evaluated the performances of four estimation methods: OLS, spatial regression with corresponding model specification, ridge regression, and SFRR. Each fitting method was replicated 1000 times to provide a total of 48 sampling distributions (2 forms of spatial dependence × 6 autoregressive coefficients × 4 estimation methods = 48) for every regression coefficient in the model.

Bias and variance are common criteria for choosing an estimator. Among the most desirable estimators is the unbiased estimator with the minimum variance, aka the best unbiased estimator. Despite its popularity with many researchers, this criterion tends to omit estimators with slight amount of biases and extremely low variances. Mean square error (MSE) is another criterion that accounts for the trade-off between bias and efficiency. As a special case of the weighted square error criterion, MSE is equivalent to minimizing the expectation of the squared error loss (Kennedy, 1998). Statistically, MSE is the sum of the variance of the estimator and the squared bias, which reduces to the

variance for unbiased estimators. Although the MSE criterion is less adopted than the best unbiased criterion, it shows great potential when the best unbiased criterion fails to produce stable estimates. In this regard, the multicollinearity among predictors makes MSE a more reliable criterion than the simple measures of bias and variance.

The simulation experiments and the case study were implemented using Python 2.7.3. We used the spreg module in Python Spatial Analysis Library (PySAL) (Rey and Anselin, 2010) for the OLS and spatial models and NumPy for the ridge regression and SFRR.

5.4 Simulation Results

We summarized the statistical properties of parameter estimation for varying degrees of spatial lag/error dependence. Tables 5.2-5.4 show the averaged bias, variance, and MSE calculated across all the regression coefficients estimated from the OLS, spatial regression model, ridge regression, and SFRR. We excluded the intercept in calculating the averages because properties of the intercept are not the major focus of this study. Moreover, because the sampling distributions of the intercept often deviate dramatically from those of the slopes, inclusion of the intercept can be distracting and impedes us from discovering the true patterns of the coefficient distributions.

The average bias for parameters estimated from the OLS was quite small for models with low spatial dependence (ρ from 0 – 0.3) (Table 5.2). As the spatial effect increased, the size of bias (regardless of the direction) increased quickly and reached 0.554 when ρ is 0.9. This supports the statement in Anselin (1988) that ignoring the spatial lag dependence can cause biased estimation, which grows as the spatial effect develops. Comparatively, the estimation bias from the spatial lag model was fairly stable

and was not affected by the increasing autoregressive coefficient. The mean bias was phenomenal for the ridge regression and SFRR. This was expected because we intentionally introduced the bias to improve the precision of parameter estimates. As with the spatial lag model, the size of the bias was stable throughout all levels of ρ for both the ridge regression and SFRR.

Table 5.2 Mean Bias of the Ten Regression Coefficients Estimated Using OLS, Spatial Lag/Error Model, Ridge Regression, and SFRR

ρ / λ	Spatial lag dependence				Spatial error dependence			
	OLS	SLM[a]	RR[b]	SFRR	OLS	SEM[c]	RR	SFRR
0	-0.098	-0.132	-0.955	-0.962	0.099	0.046	-0.954	-0.958
0.1	0.029	0.155	-0.955	-0.956	-0.082	-0.098	-0.957	-0.958
0.3	0.098	0.169	-0.959	-0.952	0.056	-0.122	-0.96	-0.955
0.5	-0.191	0.132	-0.959	-0.954	0.044	-0.125	-0.959	-0.954
0.7	-0.426	0.193	-0.952	-0.962	0.031	-0.024	-0.955	-0.953
0.9	-0.554	0.115	-0.935	-0.96	-0.104	-0.064	-0.956	-0.951

[a] SLM: spatial lag model.
[b] RR: ridge regression.
[c] SEM: spatial error model.

When the errors follow a SAR process, the assumption of uncorrelated disturbance is no longer valid. Both the OLS and spatial error model have a small bias even in the presence of strong spatial error dependence. In other words, neither structured disturbances nor the collinearity among the predictors affect the unbiasedness of the OLS estimator. Again, the ridge regression and SFRR estimators were biased, and the size of the bias was not sensitive to changing spatial effects.

Table 5.3 shows the averaged variance of the slopes for the four candidate models. When spatial lag dependence was present, the mean variance for the OLS grew quickly as ρ increased, ranging from 50.13 when there was no spatial effect to 228.49

when ρ reached 0.9. Clearly, failing to consider spatial effect not only affects the biasedness, but also influences the estimation efficiency. The spatial lag model prevented the variance from increasing, yet the size of the variance remained fairly large (~50). By contrast, a remarkably lower estimation variance was observed for the ridge regression and SFRR. The mean variance remained consistently low—about 11% of the true slope value—for the SFRR whereas it grew slowly for the ridge regression as the spatial effect developed.

Table 5.3 Mean Variance of the Ten Regression Coefficients Estimated Using OLS, Spatial Lag/Error Model, Ridge Regression, and SFRR

ρ / λ	Spatial lag dependence				Spatial error dependence			
	OLS	SLM[a]	RR[b]	SFRR	OLS	SEM[c]	RR	SFRR
0	50.13	49.332	0.115	0.113	50.392	48.058	0.112	0.116
0.1	54.766	52.211	0.114	0.116	50.339	48.132	0.115	0.115
0.3	56.473	50.665	0.134	0.113	53.605	52.697	0.124	0.108
0.5	67.119	50.392	0.151	0.107	65.587	52.01	0.152	0.102
0.7	89.849	48.497	0.214	0.108	86.45	50.44	0.224	0.102
0.9	228.49	47.138	0.548	0.107	243.592	47.561	0.565	0.099

[a] SLM: spatial lag model.
[b] RR: ridge regression.
[c] SEM: spatial error model.

While spatial error dependence does not raise serious concern for biasness, it becomes more of an issue when it comes to the estimation efficiency. We observed a growing pattern in the mean variance for the OLS as the spatial error effect intensified, especially when λ was greater than 0.5. The variance for the spatial error model fluctuated between 47 and 53 and was not affected by the changing spatial effect. Similar to spatial lag dependence, the ridge regression model and SFRR achieved a much lower

estimation variance than the other two models, with the SFRR being the only estimator with consistently low variance regardless of strength of the spatial error effect.

MSE is a criterion that accounts for the trade-off between low bias and high efficiency. When bias and variance are considered together, the SFRR stood out with substantially lower MSE than the OLS and spatial lag/error model (Table 5.4). Furthermore, the size of MSE for the SFRR was not affected by the degree of spatial dependence as the ridge regression was and this property held for both forms of spatial dependence. Because the sampling distributions of the coefficients are closely concentrated around the expected values, the chance of obtaining estimates close to the true value is much higher for the SFRR than for the OLS and spatial regression models. The MSE criterion highlights SFRR as the most accurate and reliable estimator among the four estimators. Therefore, the use of SFRR should be widely encouraged in place of the conventional techniques when both spatial dependence and multicollinearity are present in the model.

Table 5.4 Mean MSE of the Ten Regression Coefficients Estimated Using OLS, Spatial Lag/Error Model, Ridge Regression, and SFRR

ρ / λ	Spatial lag dependence				Spatial error dependence			
	OLS	SLM[a]	RR[b]	SFRR	OLS	SEM[c]	RR	SFRR
0	50.175	49.384	1.049	1.062	50.433	48.063	1.047	1.059
0.1	54.779	52.31	1.05	1.053	50.366	48.162	1.055	1.06
0.3	56.5	50.74	1.078	1.039	53.622	52.727	1.071	1.046
0.5	67.223	50.455	1.1	1.037	65.612	52.051	1.091	1.033
0.7	90.441	48.635	1.154	1.053	86.454	50.45	1.162	1.033
0.9	230.412	47.228	1.486	1.047	243.639	47.57	1.503	1.03

[a] SLM: spatial lag model.
[b] RR: ridge regression.
[c] SEM: spatial error model.

5.5 Case Study

The simulation experiments show the SFRR's capability to produce accurate parameter estimates in the presence of spatial effects and multicollinearity. In this section, we apply this method to an empirical data set in an effort to address the fundamental question raised at the beginning of the paper: what are the respective impacts of composition and spatial configuration of various land cover features on surface temperature and UHI mitigation?

The case study is confined to central Phoenix, which is situated at the northern part of the Sonoran desert. As a typical desert city in the southwest United States, Phoenix has faced a series of heat-related issues that are inescapably associated with the phenomenal heat island effect in this area (Golden et al., 2008; Kalkstein and Sheridan, 2007; Guhathakurta and Gober, 2007). The city has excessive sunshine all year around with the average maximum temperature of 30.4°Celsius (C) (US Climate Data, 2016). The average annual precipitation is ~203 millimeters (mm), making irrigation a major source for maintaining greenspace in the city (ADWR, 2013).

In this section, we report our effort of an empirical study as a demonstration of the SFRR in solving a real-world problem. Understanding how different composition and spatial arrangements of land cover features affect surface temperatures would be beneficial for the mitigation of the heat island effect and its associated heat-related problems without depleting the water and energy consumption.

Following the procedure in Section 5.3.4, one composition and one configuration variable were created for each of the five LULC categories: trees/shrubs, grass, buildings, other impervious surfaces, and unmanaged soil. Instead of creating the synthetic data, we

used the surface temperature data over central Phoenix as observations on the dependent variable. The LST product was derived from an Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) image acquired on July 6, 2005 (Figure 5.3). Note that this is a summer daytime image taken at 11:20 am local time. We then fit this data set using the SFRR following the procedure detailed in Section 5.3.3 and obtained the parameter estimates. We also fit the OLS to the same data set for comparison purposes.

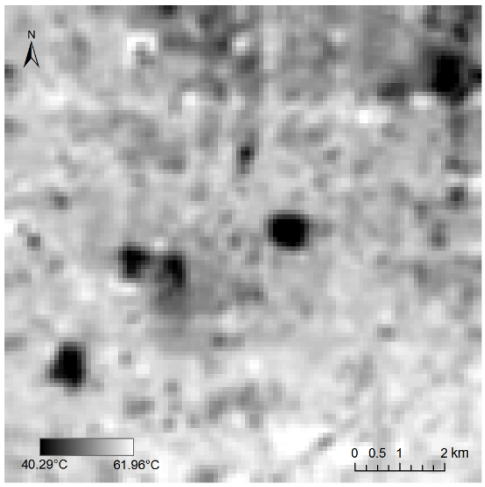

Figure 5.3 Summer Daytime LST (°C) Map of Central Phoenix.

Table 5.5 shows the diagnostics for spatial dependence after the OLS regression. The specification search identified spatial lag as the alternative model and thus a spatial lag model was fit to the data. We then substituted the ρ (0.135) into$(\mathbf{I} - \rho\mathbf{W})y$, which was treated as the dependent variable in the ridge regression.

Table 5.5 Diagnostics for Spatial Dependence

Test	Value	*p*-value
Moran's I (error)	1.486	0.1373
Lagrange Multiplier (lag)	10.299	0.0013
Robust LM (lag)	9.66	0.0019
Lagrange Multiplier (error)	2.049	0.1523
Robust LM (error)	1.41	0.235

Table 5.6 shows the coefficients of the LULC variables estimated from the OLS and SFRR. For the OLS regression, all five composition variables were positively associated with the surface temperature. While we would normally expect positive (warming) effects from buildings, parking lots, and sometimes soil on the surface temperatures, the warming effects from trees, shrubs, and grass in a late summer morning is highly suspicious and is most likely due to estimation error. Because of the enormous multicollinearity (condition number: 110.47), the OLS estimates were quite unstable, and therefore the likelihood of obtaining estimates close to the true value is substantially reduced. In fact, the coefficient estimate of any predictor variable depends to a great extent on whether or not other correlated predictor variables are included in the model (Kutner et al., 2005). Thus, the coefficient of any LULC variable does not reflect its individual association with the surface temperature, but rather a marginal effect given other variables in the model. The OLS estimates are unreliable also because the spatial lag effect was ignored in the data fitting process. This causes the OLS estimates to be biased and inconsistent.

Table 5.6 Coefficient Estimates for the Ten Land Composition and Configuration Variables Obtained from the OLS and SFRR

Predictor variables	OLS	SFRR
Constant	41.448	47.013
Composition variables		
Trees/shrubs	7.892	-3.063
Grass	1.612	-1.588
Buildings	9.812	3.831
Other impervious surfaces	2.884	0.379
Unmanaged soil	5.125	3.036
Configuration variables		
Trees/shrubs	-7.124	-2.084
Grass	-4.142	-1.293
Buildings	-3.014	1.246
Other impervious surfaces	-2.059	-0.012
Unmanaged soil	-0.187	1.285
Pseudo R^2	0.633	0.609

On the contrary, the SFRR gives more reliable estimates by accounting for both the spatial dependence and multicollinearity. From the simulation experiments in Section 5.4, SFRR allows us to produce substantially more precise estimates at the cost of small amount of biases (also minimal MSE). In light of the SFRR estimates, both composition and configuration of land cover features affect the UHI intensity. Specifically, there was a negative association between vegetation fraction and LST, with trees/shrubs having a larger impact on the LST than that of grass. Conversely, we observed a positive relationship between LST and composition of buildings and open soil. Other impervious surfaces, such as parking lots, roads, and rooftops, have smaller impacts on the LST. These results are in line with existing literature concerning the cooling and warming

effects from vegetated surfaces and anthropogenic features, respectively (Mallick et al., 2013; Taha, 1997; Weng et al., 2004; Yuan and Bauer, 2007).

Spatial configuration of LULC classes also affects LST. There was a negative relationship between the local Moran's *I* of vegetation and LST, indicating that in a desert urban environment, clustered pattern of vegetation are more desirable for ameliorating temperatures than fragmented and dispersed patterns. A consistent statement was suggested in Li et al. (2012) using a correlation analysis, Zhang et al. (2009) a simple regression analysis, Maimaitiyiming et al. (2014) a normalized mutual information measure, and Fan et al. (2015) a modified local spatial statistic. Zhou et al. (2011), however, found a positive association between spatial pattern of woody vegetation and surface temperatures, where a simple patch based index was used to represent the configuration of vegetation patches. The discrepancy in the results could be due to the wide variance of the sampling distribution of the parameters, given that no effort was made to address the spatial effect and the multicollinearity. The spatial configuration of buildings and open soil were positively related to LST. Clustered patterns of buildings and soil can produce aggregate warming effects that further elevate LST. Other impervious surfaces, on the other hand, were weakly related to the LST. Because paved surfaces absorb short-wave radiation in the daytime and release long-wave radiation at night, their impacts on the surface temperature are expected to be significant at nighttime.

We reported the pseudo R^2 for both methods as an approximate measure of fit (Table 5.6). It is defined as the squared zero-order correlation between the observed and fitted values. We observed a moderately high pseudo R^2 for both the OLS and SFRR, indicating that neither spatial effect nor multicollinearity affects the goodness of fit. Note

that different from the R^2, the pseudo R^2 cannot be interpreted as the proportion of variance explained by the predictor variables (Anselin and Rey, 2014).

5.6 Discussion and Conclusion

The dominance of the OLS model in the environmental science community is facing serious challenges. In spite of the conceptual simplicity and accessibility via most statistical software packages, improper use of the OLS without evaluating its regularity conditions can lead to erroneous inclusion of predictor variables, imprecise parameter estimates, and poor predictions. Spatial dependence is a fundamental facet of geographic phenomena that has yet received full awareness from many geographers. Because the existence of either form of spatial dependence (substantive spatial process or a nuisance) violates the basic assumptions of a classical linear regression model, OLS estimates are no longer consistent and/or unbiased.

Multicollinearity poses another issue. Most geophysical and ecological studies involve estimation of multiple parameters in a single regression model. When predictors are highly correlated, the parameter estimates tend to vary widely from sample to sample. The large sample variability can lead to inflated estimation variance, causing significant predictors insignificant. Moreover, the contribution of each predictor to the response is affected by the inclusion (non-inclusion) of the other predictors in the model.

SFRR was proposed in this paper as an alternative to cope with the two issues by virtue of systematic integration of spatial filtering and ridge regression. While spatial filter removes the spatial autocorrelation effect, ridge regression takes care of the large estimation variance. Simulation experiments highlight the SFRR estimator as an accurate and reliable estimator in comparison with the OLS, spatial lag/error, and ridge regression

estimators. With the minimal MSE, the likelihood for a SFRR estimator to be near the true value is substantially greater than the other three models, and this property holds regardless of form and strength of spatial dependence.

We learned from the case study that multicollinearity does not hinder our ability to achieve a good fit, nor does it affect statistical inferences on the mean response and predictions if they are made within the geographical or temporal range of the sampled observations (Dormann et al., 2013). To demonstrate this, we computed the averaged Akaike's Information Criterion (AIC) for the four models (using simulated data) as a measure of fit (Figure 5.4). The AIC is a log-likelihood measure of model adequacy that allows comparison between spatial models and standard regression models (Anselin and Rey, 2014). The smallest AIC indicates the best model.

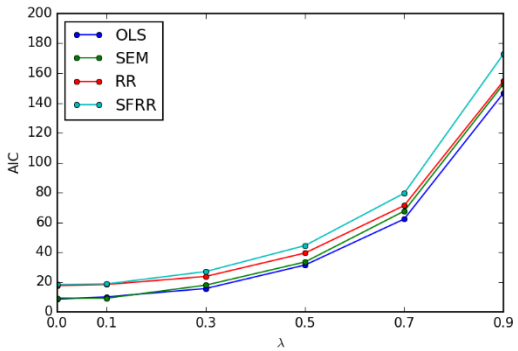

Figure 5.4 Averaged AIC for OLS, Spatial Error Model, Ridge Regression, and SFRR (SEM: Spatial Error Model; RR: Ridge Regression)

As λ grows, the AIC increased for all four models. When λ is fixed, the AIC was smallest for the OLS, followed by the spatial error model, ridge regression, and SFRR. Consistent with the case study above, the multicollinearity among predictors is not so much of a concern when the purpose is to obtain good fits or make predictions.

We show from the simulation and the case study that the SFRR estimator is superior to the OLS, spatial regression, and ridge regression estimators in the presence of spatial dependence and multicollinearity. A major drawback of the SFRR, however, is that statistical inference on the SFRR estimators is difficult to undertake because distributions of ridge estimators (and thus SFRR estimators) are usually unknown (Kennedy, 1998). Previous statistical literature suggest that ridge regression provides the same *t* and central *F* distributions as the least square regression (Coutsourides and Troskie, 1979; Obenchain, 1977). Therefore, the same *t* and *F* tests should be utilized to assess the significance of ridge estimators. Another study evaluates two non-exact *t* tests based on two ridge estimators and shows that both tests perform better than the least square *t* test for models with both large and small standard errors (Halawa and El Bassiouni, 2000). The non-exact tests were later applied to linear ridge regression and extended to logistic ridge regression in Cule et al. (2011). Future research should investigate these significant tests in the context of SFRR in terms of validity, robustness, and computational cost.

Another limitation is that the SFRR alone provides no explicit guidance as to whether a predictor should be included or dropped from the model. This can be done with the aid of certain variable selection technique once valid significance test is developed. A popular approach called stability selection serves this purpose via applying a dimension reduction technique to subsamples of the data and identifying predictors that are consistently significant (Meinshausen and Bühlmann, 2010). The integration of the SFRR and variable selection is another direction for future research.

CHAPTER 6

CONCLUSIONS

6.1 Summary of Dissertation

Accurate and effective quantification of landscape structure serves as a first and fundamental step toward addressing the linkages between spatial pattern and the underlying ecological processes (Luck and Wu, 2002). Despite the dominance of the patch mosaic model in the community of landscape ecology, a growing amount of work has been dedicated to evaluating the potential of continuous models (aka landscape gradient models) in landscape pattern mapping (McGarigal and Cushman, 2005; Southworth et al., 2004; Zheng et al., 2014).

Chapter 2 examined the utility of a family of the landscape gradient models—spatial autocorrelation indices—in measuring landscape fragmentation in a desert city. An exploratory data analysis was performed to quantitatively evaluate the relations between the most widely used landscape metrics and the proposed continuous indices. Results suggest a certain amount of consistency between the two classes of models and highlight the Getis-Ord G and local Moran's I as useful alternatives to the classical patch mosaic model with the benefit that no land cover classification is required. The statistical relationships between the Getis-Ord G and land composition metrics are found to be more sensitive to changing areal extent than changing grain size.

The usefulness of continuous spatial indices is corroborated in Chapter 3 through a spatiotemporal study of land cover dynamics over the Phoenix metropolitan area. This research utilizes a combination of sequential satellite imagery, spatial autocorrelation indices, and non-parametric regression in an effort to evaluate changes in the quantity and

spatial distribution of vegetation and built-up areas in a 20-year time period. Results identify a significant loss of vegetation in the central and southeast part of the region whereas vegetation increases mainly occur in the northeast, featuring a pattern of urban encroachment into natural desert lands. Increases of built-up areas are phenomenal across the entire region as a result of land conversions from agriculture, desert, and riparian into urban land use types. There is an overall enhancement of spatial heterogeneity in the majority of this area. Besides urbanization itself, extensive practice of xeriscaping is another major cause of increased spatial complexity.

Another core objective of this dissertation is to develop spatial analytical tools that enable accurate estimates of land cover impacts on urban climate. Using the continuous spatial indices proposed in Chapter 2, Chapter 4 sets out by examining the seasonal and diurnal variability of surface UHI in relation to the quantity and spatial distribution of urban vegetation at various geographical scales. Results highlight the great potential of optimizing spatial configuration of grass and trees in ameliorating city temperatures with evidence supporting the positive impacts of planning spatially clustered vegetation patches rather than scattered and dispersed ones. This research also shows that the association between vegetation and surface temperature in an urban environment is strongest within a ~200 m × 200 m areal extent. A similar spatial extent has been suggested by Myint et al. (2010) using a correlation analysis and Kormann and Meixner (2001) using climate modeling.

Evaluation of the relative impacts from composition and configuration factors on urban climate requires inclusion of both types of variables in a regression model. Although unbiased, naive parameter estimates are subject to inflated estimation variance

due to the strong correlations between composition and configuration variables. Spatial dependence inherently associated with geographic data complicates the problem, leading to inefficient and/or biased estimates. Chapter 5 resolves around these two issues and develops a unified regression framework called spatially filtered ridge regression (SFRR) that integrates estimation procedures of spatial autoregressive models and ridge regression. Based on evidence from a sequence of Monte Carlo simulation and an empirical study over Phoenix, the SFRR estimator substantially reduces the estimation variance at the cost of a small bias, significantly improving the estimation accuracy in the presence of spatial effects and multicollinearity.

6.2 Future Work

6.2.1 Landscape Mapping

This research utilizes continuous spatial indices that characterize the landscape structure from a gradient perspective. While the indices themselves are illuminating, they can be further enhanced and extended through incorporating data from various resources. For instance, a joint use of spatial autocorrelation indices and Lidar data can provide additional information about the surface characteristics of a landscape, such as its height distribution and surface roughness. In addition, the ecological implications of landscape pattern quantified using gradient metrics still remain to be explored.

6.2.2 Spatiotemporal Landscape Modeling

Time is a critical component of landscape pattern analysis and there have been numerous efforts dedicated to monitoring long-term changes of landscape structure using various analytical techniques. Of particular concern is the central question of

spatiotemporal analysis - How to properly and effectively couple spatial and temporal domains and how to simultaneous address spatial dependence and serial correlation within a single analytical framework? The next phase of my research will examine this problem and explore possible solutions using knowledge from various disciplines such as spatial econometrics, statistics, ecology, and economics.

6.2.3 Land Cover Impacts on Other Factors

In addition to surface temperatures, landscape pattern is associated with a wide range of biophysical and socioeconomic factors. Potential questions of interest are: Does a clustered or dispersed pattern of vegetation use more water and/or air conditioning? What type of landscaping is most commonly found in a wealthy/poor neighborhood? Is a particular land cover type/pattern strongly linked to a particular type of disease? These and many other questions regarding land cover impacts on our socio-ecological system warrant future research.

REFERENCES

ADWR. (2013) Phoenix AMA Cultural Water Demand. Retrived from http://www.azwater.gov/AzDWR/StatewidePlanning/WaterAtlas/ActiveManagementAreas/Cultural/PhoenixAMA.htm (accessed Mar 17, 2016)

Akbari, H. (2002) Shade trees reduce building energy use and CO_2 emissions from power plants. *Environmental pollution* 116: S119-S126.

An, L., Linderman, M., Qi, J., Shortridge, A., and Liu, J. (2005) Exploring complexity in a human–environment system: an agent-based spatial model for multidisciplinary and multiscale integration, *Annals of the Association of American Geographers* 95(1): 54-79.

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*: Springer Science & Business Media.

Anselin, L. (1995). Local indicators of spatial association – LISA. *Geographical Analysis* 27(2), 93-115.

Anselin, L. (2004) Exploring spatial data with GeoDa™: a workbook. *Urbana* 51: 61801.

Anselin, L. and Rey, S. (1991) Properties of tests for spatial dependence in linear regression models. *Geographical analysis* 23: 112-131.

Anselin, L. and Rey, S. (2014) *Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL*: GeoDa Press LLC, Chicago, IL, USA.

Arnfield, AJ. (2003) Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. *International journal of climatology* 23: 1-26.

Arnold Jr, C.L. and Gibbons, C.J. (1996) Impervious surface coverage: the emergence of a key environmental indicator. *Journal of the American planning Association* 62: 243-258.

As-Syakur, A.R., Adnyana, I., Arthana, I.W., and Nuarsa, I.W. (2012) Enhanced Built-Up and Bareness Index (EBBI) for mapping built-up and bare land in an urban area, *Remote Sensing* 4(10): 2957-2970.

Baatz, M. and Schape, A. (1999) Object-oriented and multi-scale image analysis in semantic networks. *Proceedings of the 2nd international symposium on operationalization of remote sensing* 16–20 August 1999. Enschede: ITC.

Bannari, A., Omari K., Teillet P.M., and Fedosejevs, G. (2005) Potential of Getis statistics to characterize the radiometric uniformity and stability of test sites used for the calibration of Earth observation sensors. *Geoscience and Remote Sensing, IEEE Transactions on* 43: 2918-2926.

Begue, A., Vintrou, E., Ruelland, D., Claden, M., and Dessay, N. (2011) Can a 25-year trend in Soudano-Sahelian vegetation dynamics be interpreted in terms of land use change? A remote sensing approach, *Global Environmental Change* 21(2): 413-420.

Blaschke, T. and Strobl, J. (2001) What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. *GIS – Zeitschrift für Geoinformationssysteme* 6: 12-17.

Brown, L.R. (2001) Eco-economy, *Earth Policy Institute*.

Buyantuyev, A., Wu, J., and Gries, C. (2010) Multiscale analysis of the urbanization pattern of the Phoenix metropolitan landscape of USA: Time, space and thematic resolution. *Landscape and Urban Planning* 94(3-4): 206-217.

Campbell, J. (2007) *Introduction to Remote Sensing* (4th ed.). The Guilford Press, New York.

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

Chen, X.L., Zhao, H.M., Li, P.X., and Yin, Z.Y. (2006) Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. *Remote sensing of environment* 104: 133-146.

Chestnut, L.G., Breffle, W.S., Smith, J.B., and Kalkstein, L.S. (1998) Analysis of differences in hot-weather-related mortality across 44 US metropolitan areas. *Environmental Science & Policy* 1: 59-70.

Chow, W.T. and Brazel, A.J. (2012) Assessing xeriscaping as a sustainable heat island mitigation approach for a desert city. *Building and Environment* 47: 170-181.

Chow, W.T., Brennan, D., and Brazel, A.J. (2012) Urban Heat Island Research in Phoenix, Arizona: Theoretical Contributions and Policy Applications. *Bulletin of the American Meteorological Society* 93.

Cihlar, J. (2000) Land cover mapping of large areas from satellites: Status and research priorities. *International Journal of Remote Sensing* 21(6-7): 1093-1114.

Clarke, J.F. (1972) Some effects of the urban structure on heat mortality. *Environmental research* 5: 93-104.

Congalton, R.G. (1991) A review of assessing accuracy of classifications of remotely sensed data. *Remote Sensing of Environment* 37: 35-46.

Connors, J.P., Galletti, C.S., and Chow, W.T. (2013) Landscape configuration and urban heat island effects: assessing the relationship between landscape characteristics and land surface temperature in Phoenix, Arizona. *Landscape ecology* 28: 271-283.

Conti, S., Meli, P., Mineli G, Solimini, R., Toccaceli, V., Vichi, M., Beltrano, C., and Perini, L. (2005) Epidemiologic study of mortality during the Summer 2003 heat wave in Italy. *Environmental Research* 98: 390-399.

Coutsourides, D. and Troskie, C. (1979) *F* and *T* tests for a general-class of estimators. *South African Statistical Journal* 13: 113-119.

Cox, D.R., and Stuart, A. (1955) Some quick sign tests for trend in location and dispersion, *Biometrika* 42: 80-95.

Cressie, N.A. (1993) *Statistics for Spatial Data*. Revised edition: John Wiley & Sons, New York, NY, USA.

Cule, E., Vineis, P. and De Iorio, M. (2011) Significance testing in ridge regression for genetic data. *BMC Bioinformatics* 12: 372.

Cushman, S.A., McGarigal, K., and Neel, M.C. (2008) Parsimony in landscape metrics: Strength, universality and consistency. *Ecological Indicators* 8: 691-703.

De Beurs, K., Henebry, G. (2005) A statistical framework for the analysis of long image time series, *International Journal of Remote Sensing* 26(8): 1551-1573.

De Jong, S.M. and Burrough, P.A. (1995) A fractal approach to the classification of Mediterranean vegetation types in remotely sensed images. *Photogrammetric Engineering and Remote Sensing* 61: 1041-1053.

Definiens, A. (2009) Definiens eCognition Developer 8 User Guide. *Definens AG, Munchen, Germany*.

DeFries, R., Townshend, J. (1994) NDVI-derived land cover classifications at a global scale, *International Journal of Remote Sensing* 15(17): 3567-3586.

Deng, C. and Wu, C. (2012) BCI: A biophysical composition index for remote sensing of urban environments, *Remote Sensing of Environment* 127: 247-259.

Deng, J.S., Wang, K., Hong, Y., and Qi, J.G. (2009) Spatio-temporal dynamics and evolution of land use change and landscape pattern in response to rapid urbanization, *Landscape and Urban Planning* 92(3): 187-198.

Déry, S.J., Hernández-Henríquez, M.A., Burford, J.E., and Wood, E.F. (2009) Observational evidence of an intensifying hydrological cycle in northern Canada, *Geophysical Research Letters* 36(13).

Dormann, C.F., Elith, J., Bacher S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R., Gruber, B., Lafourcade, B., Leitão, P.J., and Münkemüller, T., McClean, C., Osborne, P.E., Reineking, B., Schröder, B., Skidmore, A.K., Zurell, D., and Lautenbach, S. (2013) Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36: 27-46.

Dunn, D.E. (1996) Trends in nutrient inflows to the Gulf of Mexico from streams draining the conterminous United States, 1972-93, US Geological Survey.

Ellis, F., Nelson, F. and Pincus, L. (1975) Mortality during heat waves in New York City July, 1972 and August and September, 1973. *Environmental Research* 10: 1-13.

EPA RUHI. (2012) Compendium of Strategies Urban Heat Island Basics, 2009. *Available on line at: http://www. epa. gov/hiri/resources/compendium. htm*.

Fan, C. and Myint, S. (2014) A comparison of spatial autocorrelation indices and landscape metrics in measuring urban landscape fragmentation, *Landscape and Urban Planning* 121: 117-128.

Fan, C., Myint, S.W., and Zheng, B. (2015) Measuring the spatial arrangement of urban vegetation and its impacts on seasonal surface temperatures, *Progress in Physical Geography* 39(2): 199-219.

Fan, C., Zheng, B., Myint, S.W., and Aggarwal, R. (2014) Characterizing changes in cropping patterns using sequential Landsat imagery: an adaptive threshold approach and application to Phoenix, Arizona, *International Journal of Remote Sensing* 35(20): 7263-7278.

Florax, R., Folmer, H. and Rey, S.J. (2003) Specification searches in spatial econometrics: the relevance of Hendry's methodology. *Regional Science and Urban Economics* 33: 557-579.

Foody, G.M. (1996) Approaches for the production and evaluation of fuzzy land cover classifications from remotely sensed data. *International Journal of Remote Sensing* 17: 1317-1340.

Foody, G.M. (2001) Monitoring the magnitude of land-cover change around the southern limits of the Sahara. *Photogrammetric Engineering and Remote Sensing* 67(7): 841-847.

Foody, G.M. and Boyd, D.S. (1999) Detection of partial land cover change associated with the migration of inter-class transitional zones. *International Journal of Remote Sensing* 20(14): 2723-2740.

Fotheringham, A.S. (1989) Scale-independent spatial analysis, *Accuracy of spatial databases*: 221-228.

Gallo, K., McNab, A., Karl T, Brown, J.F., Hood, J.J., and Tarpley, J.D. (1993) The use of a vegetation index for assessment of the urban heat island effect. *Remote Sensing* 14: 2223-2230.

Geoghegan, J., Wainger, L.A., and Bockstael, N.E. (1997) Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. *Ecological economics* 23: 251-264.

Getis, A. and Ord, J. (1992) The analysis of spatial association by distance statistics. *Geographical Analysis* 24(3): 189-206.

Gillanders, S.N., Coops, N.C., Wulder, M.A., Gergel, S.E., and Nelson, T. (2008) Multitemporal remote sensing of landscape dynamics and pattern change: describing natural and anthropogenic trends, *Progress in Physical Geography* 32(5): 503-528.

Gober, P., Brazel, A., Quay R, Myint, S., Grossman-Clarke, S., Miller, A., and Rossi, S. (2009) Using watered landscapes to manipulate urban heat island effects: how much water will it take to cool Phoenix? *Journal of the American Planning Association* 76: 109-121.

Gober, P., Middel, A., Brazel, A., Myint, S., Chang, H., Duh, J.D., and House-Peters, L. (2012) Tradeoffs between water conservation and temperature amelioration in Phoenix and Portland: implications for urban sustainability. *Urban Geography* 33: 1030-1054.

Golden, J.S., Hartz, D., Brazel A, Luber, G., and Phelan, P. (2008) A biometeorology study of climate and heat-related morbidity in Phoenix from 2001 to 2006. *International Journal of Biometeorology* 52: 471-480.

Golub, G.H., Heath, M., and Wahba, G. (1979) Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics* 21: 215-223.

Gough, W.A., Cornwell, A.R., and Tsuji, L.J. (2004) Trends in seasonal sea ice duration in southwestern Hudson Bay, *Arctic*: 299-305.

Griffith, J.A., Martinko, E.A., and Price, K.P. (2000) Landscape structure analysis of Kansas at three scales. *Landscape and Urban Planning* 52: 45–61.

Griffith, J.A., Martinko, E.A., Whistler JL, and Price, K.P. (2002) Interrelationships among landscapes, NDVI, and stream water quality in the US Central Plains. *Ecological Applications* 12: 1702-1718.

Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X., and Briggs, J.M. (2008) Global change and the ecology of cities, *Science* 319(5864): 756-760.

Grimm, N.B., Foster, D., Groffman P, Grove, J.M., Hopkinson, C.S., Nadelhoffer, K.J., Pataki, D.E., and Peters, D.P. (2008) The changing landscape: ecosystem responses to urbanization and pollution across climatic and societal gradients. *Frontiers in Ecology and the Environment* 6: 264-272.

Grimm, N.B., Redman, C.L. (2004) Approaches to the study of urban ecosystems: the case of Central Arizona—Phoenix, *Urban Ecosystems* **7**(3): 199-213.

Guhathakurta, S. and Gober, P. (2007) The impact of the Phoenix urban heat island on residential water use. *Journal of the American Planning Association* 73: 317-329.

Haines-Young, R. and Chopping, M. (1996) Quantifying landscape structure: A review of landscape indices and their application to forested landscapes. *Progress in Physical Geography* 20: 418-445.

Haining, R. (1990) *Spatial Data Analysis in the Social and Environmental Sciences*: Cambridge Univeristy Press, Cambridge, UK.

Halawa, A. and El Bassiouni, M. (2000) Tests of regression coefficients under ridge regression models. *Journal of Statistical Computation and Simulation* 65: 341-356.

Hamed, K.H. and Rao, A.R. (1998) A modified Mann-Kendall trend test for autocorrelated data, *Journal of Hydrology* 204(1): 182-196.

Han, W., Meehl, G.A., Rajagopalan, B., Fasullo, J.T., Hu, A., Lin, J., Large, W.G., Wang, J.-w., Quan, X.-W., and Trenary, L.L. (2010) Patterns of Indian Ocean sea-level change in a warming climate, *Nature Geoscience* 3(8): 546-550.

Hansen, M., DeFries, R., Townshend, J.R., and Sohlberg, R. (2000) Global land cover classification at 1 km spatial resolution using a classification tree approach, *International Journal of Remote Sensing* 21(6-7): 1331-1364.

Harlan, S.L., Brazel, A.J., Prashad L, Stefanov, W.L., and Larsen, L. (2006) Neighborhood microclimates and vulnerability to heat stress. *Social Science & Medicine* 63: 2847-2863.

Hart, M.A. and Sailor, D.J. (2009) Quantifying the influence of land-use and surface characteristics on spatial variability in the urban heat island. *Theoretical and Applied Climatology* 95: 397-406.

Hayes, D.J. and Sader, S.A. (2001) Comparison of change-detection techniques for monitoring tropical forest clearing and vegetation regrowth in a time series, *Photogrammetric Engineering and Remote Sensing* 67(9): 1067-1075.

Herold, M., Liu, X., and Clarke, K.C. (2003) Spatial metrics and image texture for mapping urban land use. *Photogrammetric Engineering and Remote Sensing* 69(9): 991-1001.

Herzog, F., Lausch, A., Muller, E., Thulke, H., Steinhardt, U., and Lehmann, S. (2001) Landscape metrics for assessment of landscape destruction and rehabilitation. *Environmental Management* 27(1): 91-107.

Hoerl, A.E. and Kennard, R.W. (1970) Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 12: 55-67.

Howard, L. (1833) *Climate of London Deduced from Meteorological Observations* (3rd edtion). London: Harver & Darton.

Immerzeel, W., Quiroz, R., and De Jong, S. (2005) Understanding precipitation patterns and land use interaction in Tibet using harmonic analysis of SPOT VGT-S10 NDVI time series, *International Journal of Remote Sensing* 26(11): 2281-2296.

Ivits, E. and Koch, B. (2002) Object-oriented remote sensing tools for biodiversity assessment: A European approach. *Proceedings of the 22nd EARSeL Symposium.* Mill press Science Publishers Rotterdam, Netherlands, 04-06.

Jakubauskas, M.E., Legates, D.R., and Kastens, J.H. (2001) Harmonic analysis of time-series AVHRR NDVI data, *Photogrammetric Engineering and Remote Sensing* 67(4): 461-470.

Jelinski, D.E. and Wu, J. (1996) The modifiable areal unit problem and implications for landscape ecology, *Landscape Ecology* 11(3): 129-140.

Jenerette, G.D., Harlan, S.L, Brazel, A., Jones, N., Larsen, L., and Stefanov, W.L. (2007) Regional relationships between surface temperature, vegetation, and human settlement in a rapidly urbanizing ecosystem. *Landscape ecology* 22: 353-365.

Jenerette, G.D. and Wu, J. (2001) Analysis and simulation of land use change in the

central Arizona – Phoenix region. *Landscape Ecology* 16: 611-626.

Jenks, G. (1977) *Optimal Data Classification for Choropleth Maps*. Occasional paper No. 2, department of geography, University of Kansas.

Jensen, J.R. (Ed.). (1996) *Introductory Digital Image Processing*. Prentice-Hall, Englewood Cliffs, NJ.

Ji, L. and Peters, A. (2004) A spatial regression procedure for evaluating the relationship between AVHRR-NDVI and climate in the northern Great Plains. *International Journal of Remote Sensing* 25: 297-311.

Ji, W., Ma, J., Twibell, R.W., and Underhill, K. (2006) Characterizing urban sprawl using multi-stage remote sensing images and landscape metrics. *Computers, Environment and Urban Systems* 30(6): 861-879.

Jiang, B., Liang, S., Wang, J., and Xiao, Z. (2010) Modeling MODIS LAI time series using three statistical methods, *Remote Sensing of Environment* 114(7): 1432-1444.

Johnsson, K. (1994) Segment-based land use classification from SPOT satellite data. *Photogrammetric Engineering and Remote Sensing* 60(1): 47-53.

Jokar, A.J., Helbich, M., Kainz, W., and Boloorani, A.D. (2013) Integration of logistic regression, Markov chain and cellular automata models to simulate urban expansion. *International Journal of Applied Earth Observation and Geoinformation* 21: 265-275.

JPL. (2001) ASTER Higher-Level Product User Guide, Advanced Spaceborne Thermal Emission and Reflection Radiometer. (accessed Oct 2013).

Ju, J., Kolaczyk, E.D., and Gopal, S. (2003) Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing. *Remote Sensing of Environment* 84(4): 550-560.

Kalkstein, A.J. and Sheridan, S.C. (2007) The social impacts of the heat-health watch/warning system in Phoenix, Arizona: assessing the perceived risk and response of the public. *International Journal of Biometeorology* 52: 43-55.

Kasarda, J.D. and Shih, W-FP. (1977) Optimal bias in ridge regression approaches to multicollinearity. *Sociological Methods & Research* 5: 461-470.

Kaufmann, R.K., Seto, K.C., Schneider, A., Liu, Z., Zhou, L., and Wang, W. (2007) Climate response to rapid urban growth: evidence of a human-induced precipitation deficit, *Journal of Climate* 20(10): 2299-2306.

Kaufmann, R., Zhou, L., Myneni, R., Tucker, C.J., Slayback, D., Shabanov, N.V., and Pinzon, J. (2003) The effect of vegetation on surface temperature: A statistical analysis of NDVI and climate data. *Geophysical Research Letters* 30: 2147.

Kearney, M.S., Rogers, A.S., Townshend, J.R. G., Stevenson, J.C., Rizzo, E., Stutzer, D., and Sundberg, K. (2002) Landsat imagery shows decline of coastal marshes in Chesapeake and Delaware bays. *Eos, Transactions, American Geophysical Union* 83(16): 173, 177-178.

Kendall, M.G. (1955) Rank correlation methods. Griffin, London.

Kennedy, P. (1998) *A Guide to Econometrics*: MIT Press, Cambridge, MA.

Keys, E., Wentz, E.A., and Redman, C.L. (2007) The spatial structure of land use from 1970–2000 in the Phoenix, Arizona, metropolitan area, *The Professional Geographer* 59(1): 131-147.

Kinney, P.L. (2008) Climate change, air quality, and human health. *American journal of preventive medicine* 35: 459-467.

Knowles-Yanez, K., Moritz, C., Fry, J., Redman, C.L., Bucchin, M., and McCartney, P.H. (1999) Historic Land Use: Phase 1 Report on Generalized Land Use. Central Arizona - Phoenix Long - Term Ecological Research (CAPLTER) Arizona State University, Tempe.

Kogan, F. and Zhu, X. (2001) Evolution of long-term errors in NDVI time series: 1985–1999, *Advances in Space Research* 28(1): 149-153.

Kormann, R. and Meixner, F.X. (2001) An analytical footprint model for non-neutral stratification. *Boundary-Layer Meteorology* 99: 207-224.

Kumar, A., Pandey, A.C., and Jeyaseelan, A.T. (2012) Built-up and vegetation extraction and density mapping using WorldView-II. *Geocarto International* 1-12.

Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W. (2005) *Applied linear statistical models*: McGraw-Hill New York.

Lai, L. and Cheng, W. (2009) Air quality influenced by urban heat island coupled with synoptic weather patterns. *Science of the Total Environment* 407: 2724-2733.

Lam N.S.-N. and Quattrochi, D.A. (1992) On the issues of scale, resolution, and fractal analysis in the mapping sciences. *Professional Geographer* 44: 88-98.

Lambin, E.F. (1997) Modelling and monitoring land cover change processes in tropical regions. *Progress in Physical Geography* 21(3): 375-393.

Lasaponara, R. (2006) On the use of principal component analysis (PCA) for evaluating interannual vegetation anomalies from SPOT/VEGETATION NDVI temporal series, *Ecological Modelling* 194(4): 429-434.

Lazaridis, D.C., Verbesselt, J., and Robinson, A.P. (2010) Penalized regression techniques for prediction: a case study for predicting tree mortality using remotely sensed vegetation indices This article is one of a selection of papers from Extending Forest Inventory and Monitoring over Space and Time. *Canadian Journal of Forest Research* 41: 24-34.

LeDrew, E.F., Holden, H., Wulder, M., Derksen, C., and Newman, C. (2004) A spatial statistical operator applied to multidate satellite imagery for identification of coral reef stress. *Remote Sensing of Environment* 91: 271-279.

Lee, R., Yu, F., Price, K., Ellis, J., and Shi, P. (2002) Evaluating vegetation phenological patterns in Inner Mongolia using NDVI time-series analysis, *International Journal of Remote Sensing* 23(12): 2505-2512.

Lee, S.W., Hwang, S.J., Lee, S.B., Hwang, H.S., and Sung, H.C. (2009) Landscape ecological approach to the relationships of land use patterns in watersheds to water quality characteristics. *Landscape and Urban Planning* 92: 80-89.

Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm? *Ecology* 74: 1659-1673.

Li, H. and Wu, J. (2004) Use and misuse of landscape indices. *Landscape Ecology* 19(4): 389-399.

Li, J., Song, C., Zhu, F., Meng, X., and Wu, J. (2011) Impacts of landscape structure on surface urban heat islands: A case study of Shanghai, China. *Remote Sensing of Environment* 115: 3249-3263.

Li, X., Gong, P., and Liang, L. (2015) A 30-year (1984–2013) record of annual urban dynamics of Beijing City derived from Landsat data, *Remote Sensing of Environment* 166: 78-90.

Li, X., Zhou, W., Ouyang, Z., Xu, W., and Zheng, H. (2012) Spatial pattern of greenspace affects land surface temperature: evidence from the heavily urbanized Beijing metropolitan area, China. *Landscape ecology* 27: 887-898.

Lichstein, J.W., Simons, T.R., Shriner, S.A., and Franzreb, K.E. (2002) Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs* 72: 445-463.

Linthicum, K.J., Anyamba, A., Tucker, C.J., Kelley, P.W., Myers, M.F., and Peters, C.J. (1999) Climate and satellite indicators to forecast Rift Valley fever epidemics in Kenya, *Science* 285(5426): 397-400.

Liu, H. and Weng, Q. (2008) Seasonal variations in the relationship between landscape pattern and land surface temperature in Indianapolis, USA. *Environmental Monitoring and Assessment* 144: 199-219.

Liu, H. and Weng, Q. (2009) An examination of the effect of landscape pattern, land surface temperature, and socioeconomic conditions on WNV dissemination in Chicago. *Environmental monitoring and assessment* 159: 143-161.

Liu, Y., Wang, Y., Peng, J., Du, Y., Liu, X., Li, S., and Zhang, D. (2015) Correlations between urbanization and vegetation degradation across the world's metropolises using DMSP/OLS nighttime light data, *Remote Sensing* 7(2): 2067-2088.

Lo, C. and Quattrochi, D.A. (2003) Land-use and land-cover change, urban heat island phenomenon, and health implications: a remote sensing approach. *Photogrammetric Engineering and Remote Sensing* 69: 1053.

Luck, M. and Wu, J. (2002) A gradient analysis of urban landscape pattern: A case study from the Phoenix metropolitan region, Arizona, USA. *Landscape Ecology* 17(4): 327-339.

Lunetta, R.S., Shao, Y., Ediriwickrema, J., and Lyon, J.G. (2010) Monitoring agricultural cropping patterns across the Laurentian Great Lakes Basin using MODIS-NDVI data, *International Journal of Applied Earth Observation and Geoinformation* 12(2): 81-88.

Ma, T., Zhou, C., Pei, T., Haynie, S., and Fan, J. (2012) Quantitative estimation of urbanization dynamics using time series of DMSP/OLS nighttime light data: A comparative case study from China's cities, *Remote Sensing of Environment* 124: 99-107.

Maimaitiyiming, M., Ghulam, A., Tiyip, T., Pla, F., Latorre-Carmona, P., Halik, Ü., Sawut, M., and Caetano, M. (2014) Effects of green space spatial pattern on land surface temperature: Implications for sustainable urban planning and climate change adaptation. *ISPRS Journal of Photogrammetry and Remote Sensing* 89: 59-66.

Mallick, J., Rahman, A., and Singh, C.K. (2013) Modeling urban heat islands in heterogeneous land surface and its correlation with impervious surface area by using night-time ASTER satellite data in highly urbanizing city, Delhi-India. *Advances in Space Research* 52: 639-655.

Manly, B.F. (1986) Randomization and regression methods for testing for associations with geographical, environmental and biological distances between populations. *Researches on Population Ecology* 28: 201-218.

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

Martin, R.V. (2008) Satellite remote sensing of surface air quality, *Atmospheric Environment* 42(34): 7823-7843.

McGarigal, K. and Cushman, S.A. (2005) The gradient concept of landscape structure. In Wiens, J. A. & Moss, M.R. (Eds.), *Issues and Perspectives in Landscape Ecology* (pp. 112-119). Cambridge University Press, Cambridge.

McGarigal, K. and Marks, B.J. (1995) FRAGSTATS: Spatial Pattern Analysis Program for Quantifying Landscape Structure. USDA Forest Service General Technical Report PNW-GTR-351, Pacific Northwest Research Station, Portland, OR, 122 pp.

McGarigal, K., Tagil, S. and Cushman, S.A. (2009) Surface metrics: An alternative to path metrics for the quantification of landscape structure. *Landscape Ecology* 24(3): 433-450.

McIntyre, S. and Barrett, F.W. (1992) Habitat variegation, an alternative to fragmentation. *Conservation Biology* 6: 146–147.

Meddens, A.J.H., Hudak, A.T., Evans, J.S., Gould, W.A., and Gonzalez, G. (2008) Characterizing forest fragments in boreal, temperate, and tropical ecosystems. *Ambio* 37: 569–576.

Meinshausen, N. and Bühlmann, P. (2010) Stability selection. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72: 417-473.

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

Miller, J., Franklin, J., and Aspinall, R. (2007) Incorporating spatial dependence in predictive vegetation models. *Ecological Modelling* 202: 225-242.

Mohsin, T., and Gough, W.A. (2010) Trend analysis of long-term temperature time series in the Greater Toronto Area (GTA), *Theoretical and Applied Climatology* 101(3-4): 311-327.

Musick, H.B. and Grover, H.D. (1991) Image textural measures as indices of landscape

pattern. In M.G. Turner, and R.H. Gardner (Eds.), *Quantitative Methods in Landscape Ecology* (pp. 77-103). SpringerVerlag, New York.

Myint, S.W., Brazel, A., Okin, G., and Buyantuyev, A. (2010) Combined effects of impervious surface and vegetation cover on air temperature variations in a rapidly expanding desert city. *GIScience & Remote Sensing* 47(3): 301-320.

Myint, S.W., Gober, P., Brazel, A., Grossman-Clarke, S., and Weng, Q. (2011) Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment* 115(5): 1145-1161.

Myint, S.W., Wentz, E.A., Brazel, A.J., and Quattrochi, D.A. (2013) The impact of distinct anthropogenic and vegetation features on urban warming. *Landscape ecology*: 1-20.

Myint, S.W., Zheng, B., Talen, E., Fan, C., Kaplan, S., Middel, A., Smith, M., Huang, H.-P., and Brazel, A. (2015) Does the spatial arrangement of urban landscape matter? Examples of urban warming and cooling in Phoenix and Las Vegas, *Ecosystem Health and Sustainability* 1(4): 1-15.

Nepstad, D.C., Veríssimo, A., Alencar, A., Nobre, C., Lima,E., Lefebvre,P., Schlesinger, P., Potter, C., Moutinho, P., Mendoza, E., Cochrane, M. and Brooks, V. (1999) Large-scale impoverishment of Amazonian forests by logging and fire. *Nature* 398: 505-508.

NJDEP. (2006) Weather and Air Quality. New Jersey Department of Environmental Protection.

Nowak, D.J., Crane, D.E., and Stevens, J.C. (2006) Air pollution removal by urban trees and shrubs in the United States. *Urban forestry & urban greening* 4: 115-123.

Obenchain, R. (1977) Classical F-tests and confidence regions for ridge regression. *Technometrics* 19: 429-439.

Ohlemüller, R., Anderson, B.J., Araújo, M.B., Butchart, S.H., Kudrna, O., Ridgely, R.S., and Thomas, C.D. (2008) The coincidence of climatic and species rarity: high risk to small-range species from climate change. *Biology Letters* 4: 568-572.

Oke, T.R. (1976) The distinction between canopy and boundary-layer urban heat islands. *Atmosphere* 14: 268-277.

Oke, T.R. (1982) The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society* 108: 1-24.

Oke, T.R. (1987) *Boundary Layer Climates*: Methuen, New York.

O'Neill, R.V., Hunsaker, C.T., Timmins, S.P., Timmins, B.L., Jackson, K.B. and Jones, K.B., Riitters, K.H., and Wickham, J.D. (1996) Scale problemsin reporting landscape pattern at the regional scale. *Landscape Ecology* 11: 169-180.

Onishi, A., Cao, X., Ito, T., Shi, F., and Imura, H. (2010) Evaluating the potential for urban heat-island mitigation by greening parking lots. *Urban Forestry & Urban Greening* 9: 323-332.

Openshaw, S. and Openshaw, S. (1984) The modifiable areal unit problem, Geo Abstracts University of East Anglia.

Openshaw, S. and Taylor, P.J. (1979) A million or so correlation coefficients: three experiments on the modifiable areal unit problem, *Statistical Applications in the Spatial Sciences* 21: 127-144.

Overmars, K., De Koning, G., and Veldkamp, A. (2003) Spatial autocorrelation in multi-scale land use models. *Ecological Modelling* 164: 257-270.

Owen, T., Carlson, T., and Gillies, R. (1998) An assessment of satellite remotely-sensed land cover parameters in quantitatively describing the climatic effect of urbanization. *International Journal of Remote Sensing* 19: 1663-1681.

Paruelo, J.M., Epstein, H.E., Lauenroth, W.K., and Burke, I.C. (1997) ANPP estimates from NDVI for the central grassland region of the United States, *Ecology* 78(3): 953-958.

Pearson, D.M. (2002) The application of local measures of spatial autocorrelation for describing pattern in north Australian landscapes. *Journal of Environmental Management* 64: 85-95.

Peng, J., Wang, Y., Zhang, Y., Wu, J., Li, W., and Li, Y. (2010) Evaluating the effectiveness of landscape metrics in quantifying spatial patterns. *Ecological Indicators* 10(2): 217-223.

Qi, Y. and Wu, J. (1996) Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices. *Landscape Ecology* 11: 39-50.

Quattrochi, D.A. and Ridd, M.K. (1994) Measurement and analysis of thermal energy responses from discrete urban surfaces using remote sensing data. *International Journal of Remote Sensing* 15: 1991-2022.

Read, J. M. and Lam, N.S.-N. (2002) Spatial methods for characterizing land cover and detecting land cover changes for the tropics. *International Journal of Remote Sensing* 23(12): 2457-2474.

Redman, C.L. and Kinzig, A.P. (2008) Water can flow uphill: A narrative of Central Arizona. In C.L. Redman, and D.R. Foster (Eds.), *Agricultural landscapes in transition: Comparisons of long-term ecological and cultural change* (pp. 238-271). New York, NY: Oxford University Press.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., and Ohlen, D.O. (1994) Measuring phenological variability from satellite imagery, *Journal of Vegetation Science* 5(5): 703-714.

Rey, S.J. and Anselin, L. (2010) PySAL: A Python library of spatial analytical methods. *Handbook of Applied Spatial Analysis.* Springer, 175-193.

Rizwan, A.M., Dennis, L.Y., and Chunho, L. (2008) A review on the generation, determination and mitigation of Urban Heat Island. *Journal of Environmental Sciences* 20: 120-128.

Rotem-Mindali, O., Michael, Y., Helman, D., and Lensky, I.M. (2015) The role of local land-use on the urban heat island effect of Tel Aviv as assesed from satellite remote sensing. *Applied Geography* 56: 145-153.

Roth, M., Oke, T., and Emery, W. (1989) Satellite-derived urban heat islands from three coastal cities and the utilization of such data in urban climatology. *International Journal of Remote Sensing* 10: 1699-1720.

Ruddell, D.M. and Dixon, P.G. (2014) The energy–water nexus: are there tradeoffs between residential energy and water consumption in arid cities?, *International Journal of Biometeorology* 58(7): 1421-1431.

Sandholt, I., Rasmussen, K., and Andersen, J. (2002) A simple interpretation of the surface temperature/vegetation index space for assessment of surface moisture status. *Remote Sensing of environment* 79: 213-224.

Sarrat, C., Lemonsu, A., Masson, V., and Guedalia, D. (2006) Impact of urban heat island on regional atmospheric pollution. *Atmospheric Environment* 40: 1743-1758.

Seto, K.C. and Fragkias, M. (2005) Quantifying spatiotemporal patterns of urban land-use change in four cities of China with time series landscape metrics. *Landscape Ecology* 20(7): 871-888.

Seto, K.C., Fragkias, M., Güneralp, B., and Reilly, M.K. (2011) A meta-analysis of global urban land expansion, *PloS One* 6(8): e23777.

Shabanov, N.V., Zhou, L., Knyazikhin, Y., Myneni, R.B., and Tucker, C.J. (2002) Analysis of interannual changes in northern vegetation activity observed in AVHRR data from 1981 to 1994, *Geoscience and Remote Sensing, IEEE Transactions on* 40(1): 115-130.

Shao, G., Liu, D., and Zhao, G. (2001) Relationships of image classification accuracy and variation of landscape statistics. *Canadian Journal of Remote Sensing* 27: 33-43.

Shao, G. and Wu, J. (2008) On the accuracy of landscape pattern analysis using remote sensing data. *Landscape Ecology* 23(5): 505-511.

Southworth, J., Munroe, D., and Nagendra, Harini. (2004) Land cover change and landscape fragmentation – Comparing the utility of continuous and discrete analyses for a western Honduras region. *Agriculture, Ecosystems and Environment* 101: 185-205.

Southworth, J., Nagendra, H., and Tucker C. (2002) Fragmentation of a landscape: Incorporating landscape metrics into satellite analyses of land-cover change. *Landscape Research* 27(3): 253-269.

Sovocool, K.A., Rosales, J.L., and Authority, S.N.W. (2001) A five-year investigation into the potential water and monetary savings of residential xeriscape in the Mojave Desert.

Stefanov, W.L, Ramsey, M.S., and Christensen, P.R. (2001) Monitoring urban land cover change: an expert system approach to land cover classification of semiarid to arid urban centers. *Remote Sensing of Environment* 77: 173-185.

Stem, P.C. (1993) A second environmental science: human-environment interactions, *Science* 260: 1897-1899.

Stevens, J. (2009) *Applied multivariate statistics for the social sciences*: Taylor & Francis US.

Studer, S., Stöckli, R., Appenzeller, C., and Vidale, P. L. (2007) A comparative study of satellite and ground-based phenology, *International Journal of Biometeorology* 51(5): 405-414.

Switzer, P. and Ingebritsen, S. (1986) Ordering of time-difference data from multispectral imagery. *Remote Sensing of Environment* 20: 85-94.

Taha, H. (1997) Urban climates and heat islands: albedo, evapotranspiration, and anthropogenic heat. *Energy and buildings* 25: 99-103.

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

Tayyebi, A. and Generette, G.D. (2016) Increases in the climate change adaption effectiveness and availabilty of vegetation across a coastal to desert climate gradient in metropolitan Los Angeles, CA, USA. *Science of the Total Environment* 548: 60-71.

Thenkabail, P.S., Schull, M., and Turral, H. (2005) Ganges and Indus river basin land use/land cover (LULC) and irrigated area mapping using continuous streams of MODIS data, *Remote Sensing of Environment* 95(3): 317-341.

Tieszen, L.L., Reed, B.C., Bliss, N.B., Wylie, B.K., and DeJong, D.D. (1997) NDVI, C3 and C4 production, and distributions in Great Plains grassland land cover classes, *Ecological applications* 7(1): 59-78.

Tobler, W. (1979) Cellular geography. *Philosophy in Geography.* Springer, 379-386.

Todd, S., Hoffer, R., and Milchunas, D. (1998) Biomass estimation on grazed and ungrazed rangelands using spectral indices, *International Journal of Remote Sensing* 19(3): 427-438.

Tucker, C.J. (1979) Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8: 127–150.

Turner, B.L., Matson, P.A., McCarthy, J.J., Corell, R.W., Christensen, L., Eckley, N., Hovelsrud-Broda, G.K., Kasperson, J.X., Kasperson, R.E., and Luers, A. (2003) Illustrating the coupled human–environment system for vulnerability analysis: three case studies, *Proceedings of the National Academy of Sciences* 100(14): 8080-8085.

Turner, M.G. (1990) Spatial and temporal analysis of landscape patterns. *Landscape Ecology* 4(1): 21–30.

Turner, M.G., Gardner, R.H., and O'Neill, R.V. (2001) *Landscape Ecology: in Theory and Practice: Pattern and Process*. Springer-Verlag, New York.

Turner, M.G., O'Neill, R.V., Gardner, R.H., and Milne, B.T. (1989) Effects of changing spatial scale on the analysis of landscape pattern. *Landscape Ecology* 3: 153-162.

Turner, R.E., Rabalais, N.N., Justic, D., and Dortch, Q. (2003) Global patterns of dissolved N, P and Si in large rivers. *Biogeochemistry* 64(3): 297–317.

United Nations (2014) World Urbanization Prospects 2014: Highlights, United Nations Publications.

United States Census Bureau (2001) Ranking tables for metropolitan areas: 1991 and 2000, Census 2000 PHC-T-3.

US Climate Data. (2016) Retrieved from http://www.usclimatedata.com/climate/phoenix/arizona/united-states/usaz0166 (accessed Feb 29, 2016).

Uuemaa, E., Antrop, M., Roosaare, J., Marja, R., and Mander, Ü. (2009) Landscape Metrics and Indices: An Overview of Their Use in Landscape Research. *Living Reviews in Landscape Research* 3(1), Retrieved June 19, 2012 from - http://www.livingreviews.org/lrlr-2009-1

Van de Griend, A. and Owe, M. (1993) On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces. *International Journal of Remote Sensing* 14: 1119-1131.

Voogt, J.A. and Oke, T.R. (2003) Thermal remote sensing of urban climates, *Remote Sensing of Environment* 86(3): 370-384.

Warner, T.A. and Shank, M.C. (1997) Spatial autocorrelation analysis of hyperspectral imagery for feature selection. *Remote Sensing of Environment* 60: 58-70.

Waylen, P., Southworth, J., Gibbes, C., and Tsai, H. (2014) Time series analysis of land cover change: Developing statistical tools to determine significance of land cover changes in persistence analyses, *Remote Sensing* 6(5): 4473-4497.

Weng, Q., Lu, D., and Schubring, J. (2004) Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote sensing of Environment* 89: 467-483.

Whitman, S., Good, G., Donoghue, E.R., Benbow, N., Show, W., and Mou, S. (1997) Mortality in Chicago attributed to the July 1995 heat wave. *American Journal of Public Health* 87: 1515-1518.

Wu, J. (2004) Effects of changing scale on landscape pattern analysis: Scaling relations. *Landscape Ecology* 19: 125-138.

Wu, J. (2008) Making the case for landscape ecology: An effective approach for urban sustainability. *Landscape Journal* 27(1): 41-50.

Wu, J. and Hobbs, R. (2002) Key issues and research priorities in landscape ecology: An idiosyncratic synthesis. *Landscape Ecology* 17(4): 355-365.

Wu, J., Shen, W., Sun, W., and Tueller, P.T. (2002) Empirical patterns of the effects of changing scale on landscape metrics. *Landscape Ecology* 17: 761-782.

Wu, J., Jenerette, G.D., Buyantuyev, A., and Redman, C.L. (2011) Quantifying spatiotemporal patterns of urbanization: The case of the two fastest growing metropolitan regions in the United States, *Ecological Complexity* 8(1): 1-8.

Wu, Q., Hu, D., Wang, R., Li, H., He, Y. Wang, M., and Wang, B. (2006) A GIS-based moving window analysis of landscape pattern in the Beijing metropolitan area, China. *International Journal of Sustainable Development and World Ecology* 13: 419-434.

Wu, X., Zhu, X., Wu, G.-Q., and Ding, W. (2014) Data mining with big data, *Knowledge and Data Engineering, IEEE Transactions on* 26(1): 97-107.

Wulder, M. and Boots, B. (1998) Local spatial autocorrelation characteristics of remotely sensed imagery assessed with the Getis statistic. *International Journal of Remote Sensing* 19: 2223–2231.

Xiao, R., Ouyang, Z., Zheng, H., Li, W.F., Schienke, E.W., and Wang, X.K. (2007) Spatial pattern of impervious surfaces and their impacts on land surface temperature in Beijing, China. *Journal of Environmental Sciences* 19: 250-256.

Xie, Y., Yu, M., Bai, Y., and Xing, X. (2006) Ecological analysis of an emerging urban landscape pattern—desakota: a case study in Suzhou, China. *Landscape ecology* 21: 1297-1309.

Xu, H. (2008) A new index for delineating built-up land features in satellite imagery, *International Journal of Remote Sensing* 29(14): 4269-4276.

Yokohari, M., Brown, R.D., Kato, Y., and Moriyama, H. (1997) Effects of paddy fields on summertime air and surface temperatures in urban fringe areas of Tokyo, Japan. *Landscape and Urban Planning* 38: 1-11.

York, A.M., Shrestha, M., Boone, C.G., Zhang, S., Harrington Jr, J.A., Prebyl, T.J., Swann, A., Agar, M., Antolin, M.F., and Nolen, B. (2011) Land fragmentation under rapid urbanization: A cross-site analysis of Southwestern cities, *Urban Ecosystems* 14(3): 429-455.

Yuan, F. and Bauer, M.E. (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(3): 375-386.

Zha, Y., Gao, J., and Ni, S. (2003) Use of normalized difference built-up index in automatically mapping urban areas from TM imagery. *International Journal of Remote Sensing* 24(3): 583-594.

Zhang, Q. and Seto, K.C. (2011) Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data, *Remote Sensing of Environment* 115(9): 2320-2329.

Zhang, X., Zhong, T., Feng, X., and Wang, K. (2009) Estimation of the relationship between vegetation patches and urban land surface temperature with remote sensing. *International Journal of Remote Sensing* 30: 2105-2118.

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

Zhou, D., Zhao, S., Liu, S., and Zhang, L. (2014) Spatiotemporal trends of terrestrial vegetation activity along the urban development intensity gradient in China's 32 major cities, *Science of the Total Environment* 488: 136-145.

Zhou, W., Huang, G., and Cadenasso, M.L. (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.

Zhu, Z. and Woodcock, C.E. (2012) Object-based cloud and cloud shadow detection in Landsat imagery, *Remote Sensing of Environment* 118: 83-94.