Pre-Publication Manuscript

Landscape and Urban Planning 125 (2014) 76-88

SPATIAL STATISTICAL ANALYSIS AND SIMULATION OF THE URBAN HEAT ISLAND IN HIGH-DENSITY CENTRAL CITIES

B. Chun[1], J.-M. Guldmann[2*]

1. Center for Geographic Information Systems, Georgia Institute of Technology
Atlanta, Georgia 30332
2. Department of City and Regional Planning, The Ohio State University,
Columbus, Ohio 43210

Abstract

The Urban Heat Island (UHI) is a mounting problem in built-up areas, leading to increased temperatures, air pollution, and energy consumption. This paper explores the urban determinants of the UHI, using two-dimensional (2-D) and three-dimensional (3-D) urban information as input to spatial statistical models. The research involves: (a) estimating land surface temperatures, using satellite images, (b) developing a 3-D city model with LiDAR data, (c) generating urban parameters with 2-D/3-D geospatial information, and (d) conducting spatial regression analyses. The data are captured over three grids of square cells−480m, 240m, and 120m− and characterize the center of Columbus, Ohio. The results show that solar radiations, open spaces, vegetation, building roof-top areas, and water strongly impact surface temperatures, and that spatial regressions are necessary to capture neighboring effects. The best regression is obtained with the General Spatial Model (GSM), which is then used to simulate the temperature effects of different greening scenarios (green roofs, greened parking and vacant lots, vegetation densification) in the center of Columbus. The results demonstrate the potential of such models to mitigate the UHI through design and land-use policies.

Acknowledgments – The comments and suggestions of the Associate Editor and three referees on two earlier drafts are very much appreciated.

[*] Author for correspondence: Professor Jean-Michel Guldmann, Department of City and Regional Planning, The Ohio State University, 276 West Woodruff Ave., Columbus, Ohio 43210
Email: guldmann.1@osu.edu; Tel: 614-292-2257; Fax: 614-292-7106

1. INTRODUCTION

Projections by the United Nations suggest that 60% of the world's population will reside in urban regions by 2030. The resulting expansion of impervious surfaces is likely to intensify the urban heat island (UHI), whereby temperatures in urban core areas are higher than in surrounding suburban and rural areas. The UHI induces heat stress, tropospheric ozone formation, and resulting health problems. Higher temperatures lead to increased electricity demand for air conditioning, which, in turn, raises power plant pollution and greenhouse gas emissions. In addition, the UHI may increase water temperatures, resulting in water ecosystems impairment. It is clear that mitigating UHI impacts requires comprehensive planning strategies accounting for the effects of urban morphology, infrastructure, and greening, on the UHI. However, a lack of understanding of these effects has been a primary obstacle to implementing such strategies.

The objectives of this paper are to: (1) specify and estimate novel statistical regression models that account for both spatial neighborhood effects and the simultaneous effects of several urban characteristics on land surface temperatures derived from satellite imagery; (2) assess which grid scale used to capture the data yields the statistical model with the best explanatory and predictive power; and (3) use the best regression model to numerically simulate UHI-mitigating strategies in an urban design and planning context. To model the statistical relationship, both two-dimensional (2-D) and three-dimensional (3-D) information is used to represent the complex geometric structure of urban centers, with an application to the urban core of Columbus, Ohio. Earlier UHI research has primarily used 2-D information, such as land uses delineated with satellite imagery and building ground floor boundaries produced by Geographic Information Systems (GIS). In the case of homogeneous land uses, this data may be sufficient to predict surface temperatures (Carlson and Arthur, 2000). However, 3-D information is necessary to analyze more complex sites, including dense building clusters (Unger et al., 2004; Unger, 2006). LiDAR data are used to generate 3-D urban geometry characteristics. A hierarchy of grids, with cell sizes of 480m, 240m, and 120m, is used to integrate all the data. A spatial regression model capturing neighborhood effects and the relationship between surface temperatures and the geometric characteristics and other factors of the built environment is formulated and estimated for these different grids. The best model is then used to simulate

different greening scenarios for the center of Columbus, illustrating its potential for mitigating the UHI.

The remainder of the paper is organized as follows. Section 2 consists in a review of the relevant literature. Section 3 describes the study area, data sources, and data processing. Regression models are formulated and estimated in Section 4. The simulations of greening scenarios are presented in Section 5. Section 6 discusses remaining issues and areas for further research. Section 7 concludes the paper.

2. LITERATURE REVIEW

There has been an explosive growth in the research literature on the UHI in recent years. Using the search engine Google Scholar and typing in "urban heat island" returns 25,300 distinct journal articles, book chapters, books, working papers, and professional reports, as of October 9, 2013, with 86.5% of these works produced since 2001. Basic reviews of research on the energetic basis of the UHI and on satellite-derived thermal remote sensing of urban areas can be found in Oke (1982), Roth et al. (1989), Voogt and Oke (2003), and Arnfield (2003). There are two streams of modeling research attempting to explain the observed UHI patterns: numerical and statistical. Numerical models involve basic physics equations related to the conservation of mass and energy and the ideal gas law. See Lynn et al. (2009) for an example of such a model used to simulate UHI mitigation strategies. In order to inform the statistical methodology presented here, the focus of this review is on statistical analyses of the UHI.

Previous research shows that vegetated areas have lower surface temperatures than impervious ones (Owen et al., 1998). Oke (1988) reports a 2°C air temperature difference. Li et al. (2010) show that locations further away from a highway are associated with lower surface temperatures. Ca et al. (1998) report that the surface temperatures of a grass field in a park are 19°C lower than those of impervious surfaces. Landsberg and Maisel (1972) were among the first to show that impervious materials lead to a 1~2°C difference between rural and urban areas.

Yuan and Bauer (2007) find a very strong linear relationship (R^2>0.95) between mean surface temperature and the percentage of impervious areas. Xiao et al. (2007) also find a high correlation (>0.8) between impervious surfaces and surface temperatures. Trees on impervious surfaces help decrease temperatures by increasing the leaf-air vapor pressure gradient. The

transpiration model by Kjelgren and Montague (1998) shows that evaporation from leaves helps decrease urban temperatures.

Remotely-sensed imagery has been increasingly used to investigate the effects of 2-D surface characteristics on urban temperatures. Chen et al. (2006), Katpatal et al. (2008), and Amiri et al. (2009) generate maps displaying land use/land cover (LU/LC) patterns based on images captured by Landsat Thematic Mapper (TM or ETM+), and find that urban expansion reduces the amount of biomass helping to control surface temperatures. Jenerette et al. (2007) show that surface temperature differences within a city are correlated with the amount of vegetation. Using temporal ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) satellite images and LU/LC classification, Kato and Yamaguchi (2007) report that heat storage capability is larger in highly urbanized areas than in sprawling residential areas.

Urban geometry relating to the height and spacing of buildings is another factor for UHI formation. This height-spacing relationship is also called the urban canyon effect, affecting air circulation, wind flow, and thermal energy absorption (Oke, 1981; Bärring et al., 1985; Oke, 1988; Eliasson, 1996; Bottyán and Unger, 2003). Dense built-up areas cannot easily release heat energy into the atmosphere, due to the lack of open space resulting from building obstructions. Tall buildings impede wind flows and their cooling effects, and isolate hot air within the canyon. Moreover, when solar energy reaches building surfaces, it is absorbed into the walls. The temperature of the air surrounding these walls is then increased. The Sky View Factor (SVF) has been widely used to measure the visible sky (Unger et al., 2004; Grimmond, 2007; Gál et al., 2009; Unger, 2009). Grimmond (2007) indicates that reducing SVF increases absorption of solar radiation onto surfaces, decreases terrestrial radiation loss, decreases total turbulent heat transport, and reduces wind speeds, which all directly lead to increased surface temperatures. Unger et al. (2004) found that the difference in mean temperature between the maximum SVF (=1) and a minimum SVF (=0.66) is around 4.4°C. Gál et al. (2009) show that there is a strong linear and negative relationship between annual mean temperatures and SVF. The size and shape of building structures are prominent features explaining surface temperatures. Giridharan et al. (2004) show that areas of dense and tall buildings have 1~1.5°C higher temperatures (measured at 1.5m to 2m above the ground) than other areas.

The shortcomings in previous UHI research are as follows. First, it mostly relies on 2-D information, such as building footprints, land uses, and census information, while not accounting

for 3-D physical factors. Little research on the UHI has taken place in 3-D space because of data and other technical problems. Second, earlier research has used a small set of variables. For example, Carlson and Arthur (2000) use a vegetation index to explain the UHI. Unger *et al*. (2004) explain UHI intensity with the SVF variable, at the exclusion of other variables. Third, earlier research involves simple Ordinary Least Squares (OLS) regression models, excluding spatial effects from adjacent areas (Jauregui, 1990; Ca et al., 1998; Shashua-Bar and Hoffman, 2000). Spatial phenomena often involve spatial dependence, with values observed at one place being related to values at adjacent places. In order to avoid the estimation bias inherent in OLS models and to consider spatial and neighboring effects, it is necessary to specify and estimate spatial regression models.

3. STUDY AREA, DATA SOURCES, AND DATA PROCESSING

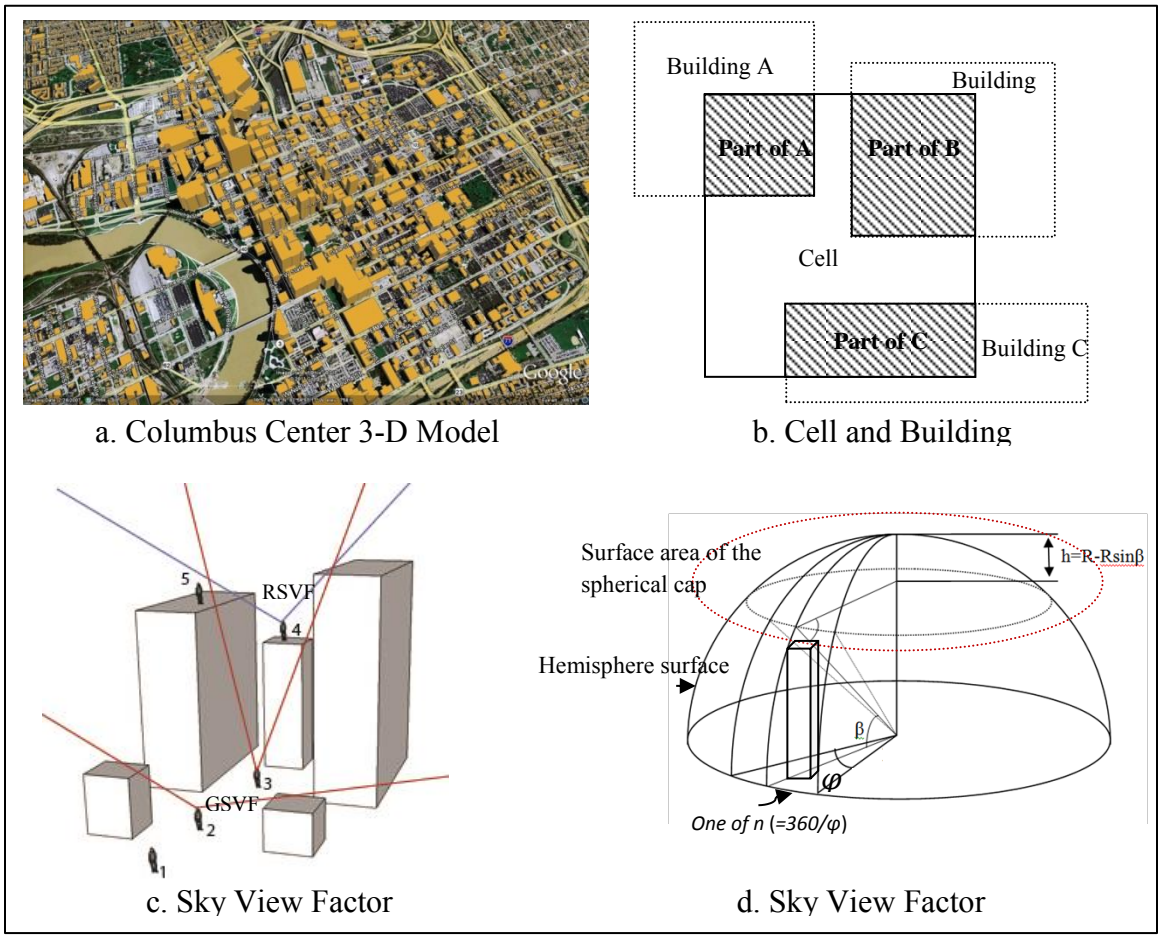
This research focuses on a densely-built section of the City of Columbus, Ohio, an area of 46.5 km^2 including the Central Business District (CBD). The Scioto and Olentangy rivers flow from north to south, and merge into one channel near the CBD. There are parks and outdoor recreation sites in the southern part of the study area, which also includes residential areas. A continuous population growth since 1950 has led to much building construction, resulting in increased surface temperatures by at least 5˚C since 1986. The proposed UHI models are based on the following data: (1) building features (building ground floor area, height); (2) factors that are directly derived from the 3-D geometry of the city (sky view and solar radiations); (3) 2-D land-use information (NDVI and water); and (4) surface temperatures. This information is extracted from geospatial data sources, such as LiDAR, Landsat TM, and existing GIS datasets. These various datasets are then integrated into grids of various scales. The base year of the analysis is 2007.

3.1. Buildings

In order to represent the topology and geometry of built-up environments, it is first necessary to develop a 3-D city model. Buildings are delineated as box-shaped, which is an acceptable approximation of real building shapes in the case of Columbus CBD (Figure 1.a). Three geospatial data sources are used: (1) LiDAR data to estimate each building height, (2) GIS building footprints to describe their ground-level boundaries, and (3) a Digital Terrain Model (DTM) to estimate terrain elevation from sea level. The LiDAR data have been gathered over the

period 2006-2008. The 3-D city model is developed in two steps: filtering LiDAR data and data integration.

First, filtering LiDAR data by using its properties (e.g., pulse, intensity, and classification) removes unnecessary data and reduces the size of the LiDAR dataset. A LiDAR pulse helps detect single and multiple surface layers. Although a tree canopy has multiple layers of leaves and branches, most building roofs are composed of a single layer. From the intensity of a LiDAR pulse, it is possible to distinguish rigid materials from other types of surfaces.



a. Columbus Center 3-D Model

b. Cell and Building

c. Sky View Factor

d. Sky View Factor

Figure 1: Building and 3-D Factors

The process of filtering the intensity of LiDAR data requires numerous trials because intensity values are relative. Pulses with intensity greater than 50 are assumed to detect hard surfaces. This filtering process helps classify and extract non-ground data, such as building roofs and the tops of trees. Although the GIS footprints of the buildings could be used alone to cookie-cut the LiDAR data to generate 3-D models, filtering these data first removes outlier observations and helps reduce GIS computational time.

Second, data integration involves the merging of LiDAR data and the GIS building footprints (from appraisal database) that define building boundaries. The spatial joining capability of ArcGIS is used to perform this procedure. The highest point among the LiDAR points within each building footprint is used to measure its height. The actual building height is then estimated by excluding the influence of topography, using the DTM.

The building ground floor area, BGFA, is the sum of the ground floor areas of all buildings in a given unit area (cell), and is equal to the total building rooftop areas because buildings are all box-shaped, with:

$$\mathrm{BGFA} = \sum_{i=1}^{n} a_i \qquad (1)$$

where a_i is building i ground floor area in a given cell, and n the number of building footprints. Figure 1.b illustrates this computation, and Figure 2.c displays the BGFA values over the study area.

3.2. Sky View Factor and Solar Radiations

Sky View Factor

The Sky View Factor (SVF) is computed as the ratio of visible sky within a given reference circle, and is used to measure the extent of 3-D open space. See Figures 1.c and 1.d. Three types of SVF measures are computed - Ground SVF (GSVF), Roof SVF (RSVF), and Total SVF (TSVF). GSVF and RSVF are illustrated in Figure 1c: GSVF pertains to points 1, 2, and 3, and RSVF to points 4 and 5. GSVF measures SVF at ground-level, and represents the pedestrian viewpoint. RSVF measures SVF on the roof of buildings. Both GSVF and RSVF vary from one point to another due to the variety of building locations, shapes, and sizes. ArcView SVF extension is used to calculate each SVF value. Technical details and trigonometric formulas are available in Li et al. (2004) and Gal and Unger (2008). Comparisons of this method with alternative approaches are also discussed in Gal et. al (2009) and Hammerle

et. al (2011). The total number of SVF measurement points over the study area is 9,152, including 7,682 points at ground level and 1,470 at building-roof level. As there are 2,288 cells of 120m size, there is an average of 4 measurement points for each such cell. Let Ng and Nr be the numbers of ground-level and roof-top SVF measurement points within a cell. TSVF for this cell is a summary index of SVF and is calculated as follows:

$$TSVF = \sum_{i=1 \to Ng} GSVF_i + \sum_{i=1 \to Nr} RSVF_i \quad (2)$$

where $RSVF_i$ or $GSVFi$ is the SVF value at point i. The patterns of TSVF computed for different grids are illustrated on Figure 3.b.

Solar Radiations

A solar radiation map is generated by the area solar radiation tool in ArcGIS Spatial Analyst extension, requiring the rasterization of the 3D city model and the identification of the latitude of the site (39.962). The tool accounts for the shading of buildings. Solar radiation calculations are performed in four steps: (1) generating a viewshed at each observation point; (2) applying a sunmap, tracking the position of the sun to estimate direct radiation; (3) applying a skymap for indirect radiation; and (4) overlaying the viewshed with the sunmap and the skymap (ESRI, 2009). The time is set for all hours on July 6, 2007. Diffusivity and transmissivity are set according to clear-sky conditions (the default value), corresponding to the weather conditions on July 6, 2007. The observation interval for solar radiation (*SR*) calculations is 30ft (9m), and the total number of solar radiation values is 394,774. Total solar radiation (*TSR*) and average solar radiation (*ASR*) are computed as follows:

$$TSR = \sum_{q=1}^{P} SR_q \quad (3)$$

$$ASR = \frac{TSR}{P} \times A \quad (4)$$

where SR_q is the solar radiation at observation point q and P is the number of observation points in a cell. A is the area of the cell. The unit is watt hours per square meter (WH/m^2). The values of ASR over the study area are displayed on Figure 2.d.

3.3. Land-Use Information

Normalized Difference Vegetation Index (NDVI)

The NDVI has been extensively used to measure vegetation cover. As discussed in Section 2, the NDVI has been used in many UHI statistical studies, because of the cooling effect of vegetation related to latent heat vaporization, and the convective cooling of tree canopies that are well ventilated and readily lose heat by wind. The NDVI is computed as a ratio involving different bands reflecting the percentage of vegetative ground cover, in particular the visible wavelength band (Red: 0.6μm~0.7μm) and the Near Infra-Red band (NIR: 0.7μm~1.1μm) of Landsat TM, because vegetation has a high spectral reflectivity of solar radiation on the NIR band. NDVI ranges from -1 to 1. Values between 0.5 and 1 commonly indicate dense vegetation cover (Carlson and Ripley, 1997). Higher NDVI values indicate healthy vegetation coverage, whereas lower NDVI values point to water and impervious materials. The equation defining NDVI is:

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{Red}}{\mathrm{NIR} + \mathrm{Red}} \tag{5}$$

The basic NDVI values ($\in [-1, +1]$) are proxies for the amounts of vegetation in the 30m cells. The NDVI values for larger cells are obtained by summing up the NDVI values of the 30m cells included in the larger cell. Therefore, the NDVI values used in the three grids are denoted TNDVI (Total NDVI). Figure 2.b illustrates the variations of NDVI for 30m cells over the study area. Red cells stand for vegetation-covered areas, while blue cells represent non-vegetated areas. The map points to the lack of vegetation in the CBD, but also to many vegetated areas around it, corresponding to residential areas and parks.

Water

Water bodies are considered separately from the NDVI. Although negative NDVI values may identify water, they may also represent impervious surfaces. The water polygons are identified by GIS, and represent the Olentangy and Scioto rivers. Water covers 3.61% of the study site. While this area is small, its impact on temperatures may be significant.

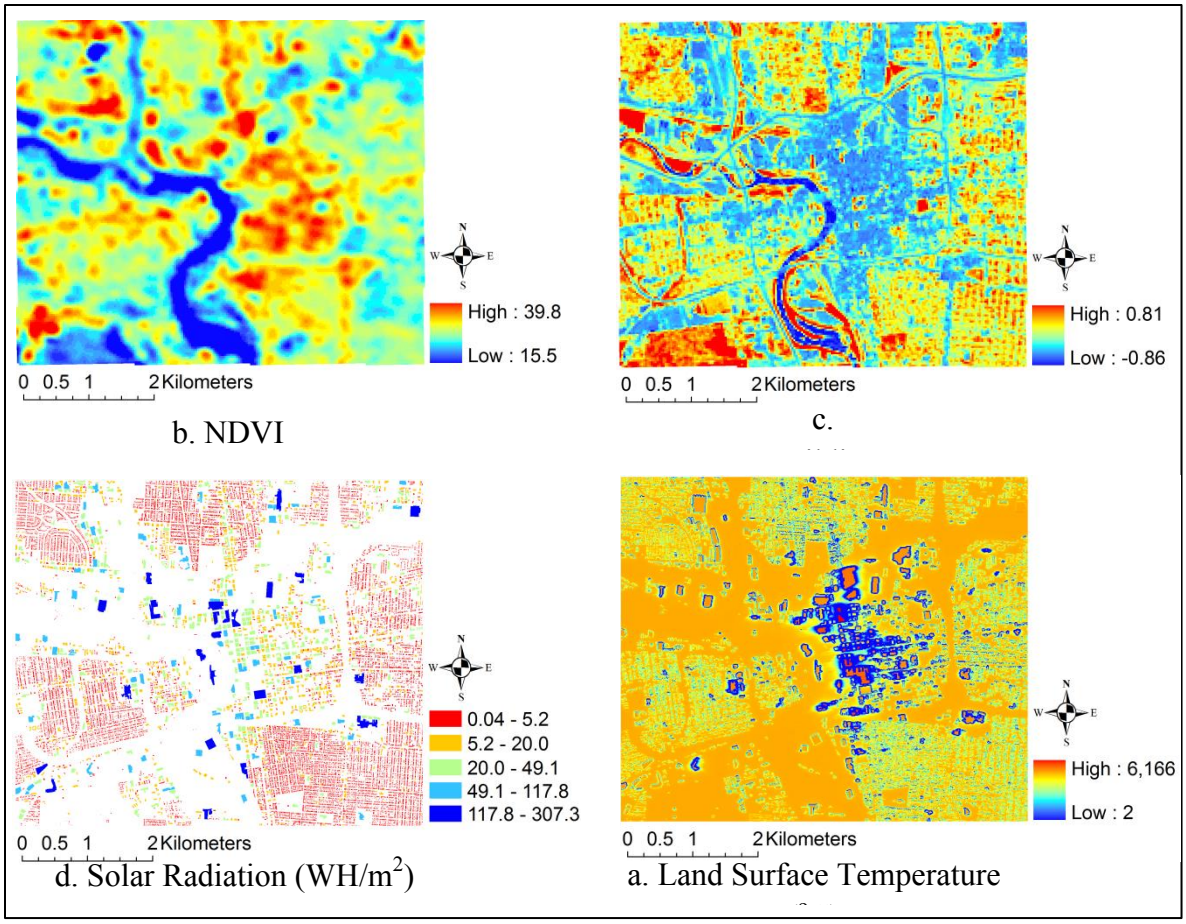


Figure 2: Land Surface Temperature, NDVI, Building Footprint, and Solar Radiation

3.4. Land Surface Temperature (LST)

LST represents the average surface temperature in a given unit area. This information is derived from the thermal band #6 captured by Landsat TM (July 6, 2007), with a 30m resolution. Although the original spatial resolution of the thermal (infrared) band is 120m, Ohioview has distributed 30m thermal images to match the resolution of the other bands by re-sizing the images (www.ohioview.org). These 30m LST values are averaged out over larger cells to create larger grids (e.g., a 120m×120m cell involves averaging out over sixteen 30m×30m cells).

The thermal band digital number (DN) must be numerically converted to a radiometric scale. The numerical conversion method proposed by the U.S. Geological Survey (USGS) has been widely used to create temperature maps, and has been updated by the Earth Resource Observation System Data Center (EROS). Equations (7) and (8) are used to perform this conversion (Chander and Markham, 2003):

$$T = K_2 / ln\left(\frac{K_1}{L_\lambda} + 1\right), \quad (6)$$

$$L_\lambda = G_{\text{rescale}} \times DN + B_{\text{rescale}}, \quad (7)$$

where T=surface temperature, L_λ= spectral radiance at the sensor's aperture, K_2= 1260.56°K, K_1= 607.76 W/(m$^2 \cdot sr \cdot \mu$m), sr=steradian (SI unit of solid angle), G_{rescale} = 0.055158 W/(m$^2 \cdot sr \cdot \mu$m)/DN, and B_{rescale}= 1.2378 W/(m$^2 \cdot sr \cdot \mu$m).

Equations (6)-(7) estimate surface temperatures under the assumption of uniform emissivity. To correct for the effects of variations in land cover emissivity, the emissivity-corrected surface temperature T_ε is computed as follows:

$$T_\varepsilon = T / \left\{1 + \left[\frac{\lambda \cdot T}{\rho} \cdot \ln(\varepsilon)\right]\right\} \quad (8)$$

$$\varepsilon = 0.004 P_V + 0.986, \quad (9)$$

$$P_V = \left(\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}\right)^2, \quad (10)$$

where λ is the wavelength of emitted radiance (11.5μm), ρ=1.438×100^{-2}m K, ε the emissivity, and P_V the vegetation proportion. For more details, see Carlson and Ripley (1997), Stathopoulou and Cartalis (2009), and Kim (2013). Figure 2.a represents the pattern of surface temperatures computed over the study area. LST ranges between 16 ^{o}C (water in the two rivers and very high-density vegetation clusters) and 40 ^{o}C (tall buildings and impervious areas).

3.5. Data Integration and Hierarchical Grid Structure

The various data sources have different spatial resolutions. For example, satellite images captured by Landsat TM have a 30m spatial resolution. Vector images of building footprints (LiDAR) have a higher resolution (less than 3m). Data overlay on uniform grids makes it possible to combine these varying input datasets, and to create a unique database, which then provides the input to statistical model estimation (Figure 3.a). The need for increasingly fine geographical scales emerges in the literature on the impacts of urban characteristics on the UHI. While some of the earlier research has led to meaningful statistical relationships (e.g., NDVI and temperature), the large spatial units generally used cannot properly reflect detailed urban

characteristics and spatial relationships. To examine such relationships, a small grid is necessary to properly reflect variability over space. In this research, three different grid cell sizes -480m, 240m, and 120m- are considered, as illustrated in the case of TSVF on Figure 3.b.

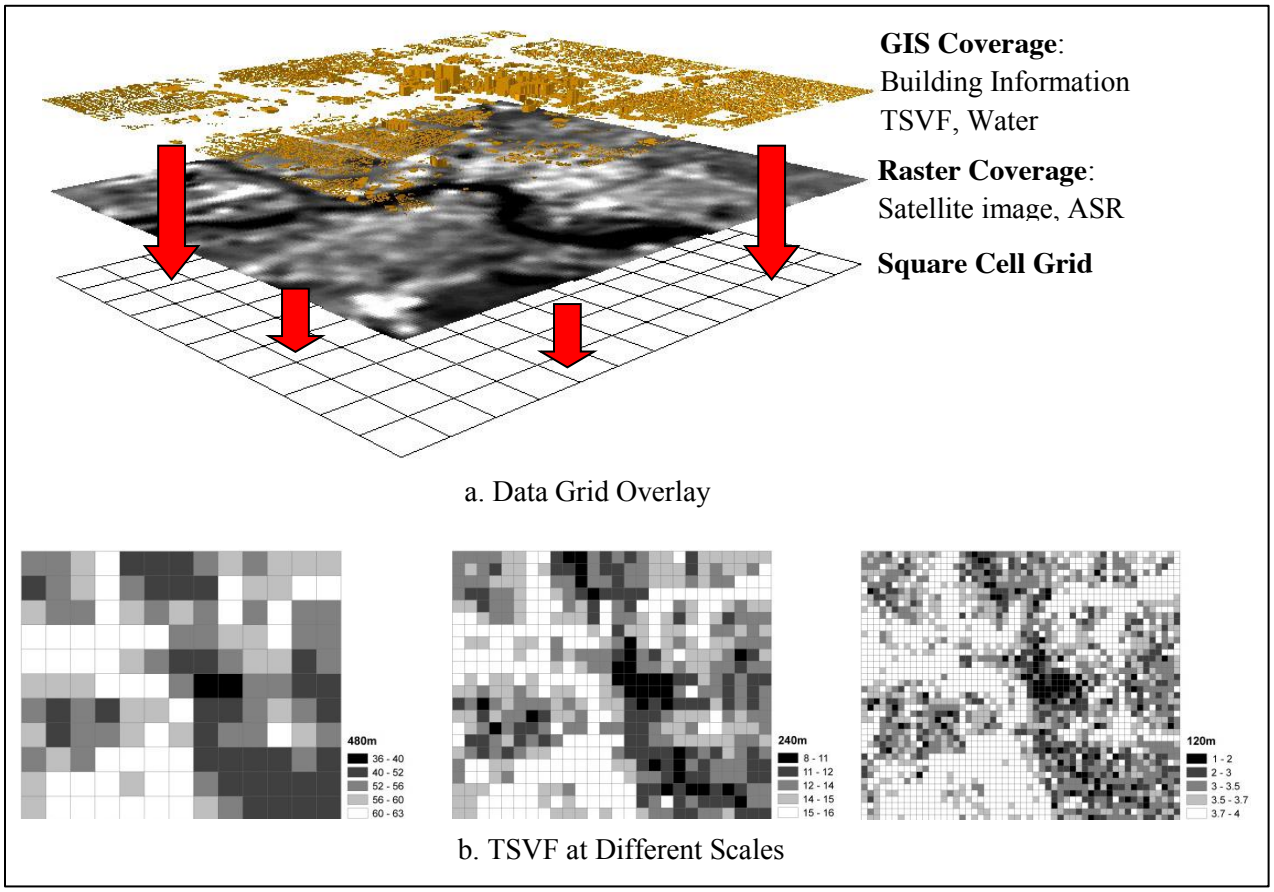


Figure 3: Data Grid Overlay and Hierarchical Scales

Descriptive statistics for all the variables and the three grids are presented in Table 1. When shifting from one grid to the next larger grid (e.g., from 120m to 240 m), the values of the area variables, BGFA and WATER, are summed up over all the smaller cells included in a given cell of the larger grid. Therefore, their mean values are necessarily in the ratio 4:1 for a given grid and the next smaller one. The same is true for the variables TNDVI and TSVF because of the additive way in which they are computed. In contrast, the variables LST and ASR, which are averaged out over any grid, have necessarily the same mean values for the three grids, but their ranges increase when shifting from larger to smaller cells.

Table 1: Descriptive Statistics for All Variables and Grids

Grid	Variable	Mean	Minimum	Maximum	Coefficient of Variation	Unit
480m	LST	30.89	25.20	34.49	0.06	˚C
	BGFA	36,489.71	699.56	104,086.24	0.06	m^2
	ASR	5,426	3,407	5,749	0.06	Wh/m^2
	TSVF	56.5	36.8	63.8	0.09	N/A
	TNDVI	49.5	-5.4	138.4	0.51	N/A
	WATER	7,463.74	0.00	101,182.74	0.26	m^2
240m	LST	30.89	22.90	37.71	0.07	˚C
	BGFA	9,122.43	0.00	44,780.19	0.06	m^2
	ASR	5,425	2,954	5,752	0.07	Wh/m^2
	TSVF	14.1	8.5	16.0	0.10	N/A
	TNDVI	12.4	-8.0	42.1	0.62	N/A
	WATER	1,865.96	0.00	55,354.05	0.35	m^2
120m	LST	30.89	21.22	39.17	0.08	˚C
	BGFA	2,280.58	0.00	13,339.39	0.08	m^2
	ASR	5,425	2,221	5,753	0.08	Wh/m^2
	TSVF	3.5	1.3	4.0	0.13	N/A
	TNDVI	3.1	-3.3	11.7	0.76	N/A
	WATER	466.47	0.00	14,399.97	0.42	m^2

*LST=averaged land surface temperatures; BGFA=building ground footprint area; ASR=Averaged solar radiation; TSVF=total sky view factor; TNDVI=total normalized difference vegetation index; WATER=water area.

4. STATISTICAL MODELING OF THE UHI

4.1. Model Estimation

Regression analysis is used to explore the relationship between LST and the urban characteristics discussed in Section 3 over each grid structure, and to determine the estimation method that best represents the UHI, including the Ordinary Least Squares (OLS) Model, the Spatial Lag (Autoregressive) Model (SAR), and the General Spatial Model (GSM), a combination of the SAR and SEM (Spatial Error Model) models. The temperature observed at a given location (cell) is likely to be spatially correlated with the temperatures observed at adjacent or close-by locations, due to atmospheric flows and exchanges. These spatial dependencies can be accounted for by estimating SAR and GSM models (LeSage and Pace, 2009).

The first step is to estimate an OLS model, and to use the Moran I test to assess whether spatial autocorrelation (SA) exists in the residuals (error terms). The logarithm of the dependent variable, ln(LST), is used, because it reduces the heteroscedastic effect of wide-ranging dependent variables. The selected model specification is:

$$\ln(LST) = a_0 + a_1 BGFA + a_2 ASR + a_3 TSVF + a_4 TNDVI + a_5 WATER + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I) \tag{11}$$

All the coefficients are significant at the 99% level and have signs consistent with the results of earlier research. Summary results for the three grids are presented in Table 2: R^2=0.835 for the 480m grid, 0.779 for the 240m grid, and 0.647 for the 120m grid. The decline in the explanatory power of the OLS model is due to (1) the wider range of variations in the dependent variable LST over smaller cells, (2) the larger sample size for smaller cells grids (N=2,288 for 120m, 572 for 240m, and 143 for 480m), and (3) the closer proximity of smaller cells, which are likely to be more sensitive to the influence of neighboring cells (the spatial autocorrelation effect) than larger cells. The OLS model cannot capture this effect, hence the lower explanatory power. This is indeed confirmed by the Moran I values (Table 2), which are statistically significant, implying that the OLS error terms are spatially autocorrelated. In order to reduce this SA, spatial lag models (SAR) are next estimated to capture the effects of LST in neighboring grid cells, with:

$$\ln(LST) = a_0 + \rho W \ln(LST) + a_1 BGFA + a_2 ASR + a_3 TSVF + a_4 TNDVI + a_5 WATER + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I) \tag{12}$$

Table 2 presents summary results for the SAR model. The first-order contiguity (queen) matrix *W* has the highest log-likelihood value in all three grids, and is retained for further analyses. See Appendix B for details about the construction of this matrix. The coefficients of the independent variables have the same signs as in the OLS models and are all significant at the 1% level. Positive spatial lag coefficient (ρ) estimates imply that higher temperatures in adjacent cells increase the central cell's temperature. It is notable that ρ increases from 0.21 to 0.78 when shifting from the 480m grid to the 120m grid, implying that the smaller the distance between adjacent cells, the stronger their mutual influences. This result is reasonable. However, the Lagrange Multiplier (LM) test of SA yields statistically significant LM values, implying that there is still SA among the residuals.

Table 2. OLS and SAR Model Summary

	Grid Size	480m (N=143)	240m (N=572)	120m (N=2,288)
OLS	R^2	.835	.779	.647
	Moran-I	0.240***	0.275***	0.438***
SAR	ρ	0.21***	0.40***	0.78***
	Log-likelihood	389.41	1,392.62	5,498.32
	LM-test	42.19***	491.57***	27,899.33***

N=sample size; ρ=spatial lag term; LM=Lagrange Multiplier test
***: statistically significant at the 1 % level

The next and final step is then to estimate a general spatial model (GSM) with the first contiguity matrix:

$$\ln(LST) = a_0 + \rho W \ln(LST) + a_1 BGFA + a_2 ASR + a_3 TSVF + a_4 TNDVI + a_5 WATER + u, \quad (13)$$
$$u = \lambda \mathrm{W}u + \varepsilon, \varepsilon \sim \mathrm{N}(0, \sigma^2 \mathrm{I})$$

The estimation results for the 480m, 240m, and 120m grids are presented in Table 3. The coefficients ρ and λ are statistically significant at the 1% level, and positive. As the grid becomes smaller, both ρ and λ become larger. The fits of the GSM models are better than those of the SAR models, as indicated by higher log-likelihood and R^2 values. It is also notable that the three R^2 values are similar, with a slight edge for the 120m grid (0.88). This suggests that smaller grid cells provide the basis for better UHI models. The GSM coefficient estimates are similar to those in the SAR models. The spatial lag coefficients ρ, while declining from the larger to the smaller grids, are smaller than those in the SAR models. The neighboring LST values have therefore less effects in the GSM model, as compared to the SAR model. This can be explained by the GSM spatial autocorrelation effects captured, in part, by the unobserved variables included in the error term (coefficient λ). The effects of the independent variables are further explored in the following section.

Table 3: GSM Model Results- First Contiguity Matrix

Variable	GRID					
	480m		240m		120m	
	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
BGFA	8.26E-08***	4.91	3.18E-07***	27.76	4.22E-07***	21.65
ASR	6.22E-05***	5.21	4.50E-05***	12.25	1.47E-05***	19.46
TSVF	-0.0035***	-4.02	-0.0073***	-4.98	-0.0082***	-4.86
TNDVI	-0.0011***	-10.30	-0.0036***	-17.78	-0.0078***	-23.80
WATER	-1.42E-07***	-13.58	-4.79E-07***	-49.17	-1.03E-06***	-37.34
Constant	3.034***	150.29	2.361***	20.46	1.1166***	273.65
ρ	0.085***		0.28***		0.67***	
λ	0.482***		0.420***		0.490***	
Log-likelihood	396.54		1410.64		5,578.48	
R^2	.868		.842		.878	

***P<0.01

4.2. Impact Analysis

Assessing the impacts of the independent variables in a spatial regression model is more complex than in the case of an OLS model, where the partial derivative of the dependent variable with regard to an independent variable can be used directly. Because of the spatial autocorrelation of the dependent variable, it is necessary to consider the effects of the independent variables in the neighboring cells. As proposed by Lesage and Pace (2009), the direct, indirect, and total impacts of any independent variable are computed using the matrix $(I-\rho W)^{-1}$. The direct impact is related to the impact of variables within the cell itself. The indirect impact represents the effects of the variables in neighboring cells, as defined by the contiguity matrix W. Finally, the total impact is the sum of the direct and indirect impacts. These impacts vary from cell to cell. The formulas are adjusted to compute elasticities instead of derivatives.

Table 4 presents the mean values of these impacts. Indirect impacts are always smaller than direct impacts for the 480m and 240m grids. However, the difference between these impacts in the 120m grid is much smaller, supporting the hypothesis that neighborhood impacts increase with smaller cells.

Table 4: Mean Values for the Direct, Indirect, and Total Impacts

Grid	480m			240m			120m		
Var.	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total
BGFA	0.034	0.029	0.063	0.032	0.022	0.054	0.011	0.01	0.021
ASR	0.353	0.299	0.652	0.252	0.169	0.421	0.083	0.073	0.156
TSVF	-0.207	-0.175	-0.312	-0.106	-0.072	-0.177	-0.03	-0.027	-0.057
TNDVI	-0.057	-0.048	-0.105	-0.046	0.031	-0.077	-0.025	-0.022	-0.048
WATER	-0.012	-0.01	-0.022	-0.01	-0.007	-0.017	-0.005	-0.005	-0.01

First, the building ground footprint area (BGFA) has positive direct and indirect impacts in all cells, indicating that increasing building roof-top areas always lead to increased LST. Impervious materials on the roofs of buildings absorb thermal energy from incoming solar radiations without any buffer. On average, a 1% increase in BGFA leads to a temperature increase of 0.06% for the 480m grid, 0.05% for the 240m grid, and 0.02% for the 120m grid. On average, the direct impact represents about 75% of the total impact in the 480m and 240m grids. However, in the 120m grid, direct impacts are strictly positive in about 40% of the cells, because zero impacts characterize cells where there are no buildings.

Second, solar radiations (ASR) have uniformly positive direct and indirect impacts, enhancing the UHI. While incoming shortwave solar radiations vary with solar altitude and atmosphere conditions, they lead to increased LST during daytime. The ASR impact is larger than the impacts of all the other variables: 0.65%, 0.42%, and 0.16% for the 480m, 240m, and 120m grids, respectively.

Third, TSVF, representing urban canyon effects, always exerts negative direct and indirect impacts on LST. The negative effect of TSVF is related to the fact that a more visible open sky enhances air circulation in densely-built environments. Urban canyons surface temperatures are higher than in open areas, pointing to heat trapping as an overriding effect. The results suggest that the SVF on both the ground and building roofs plays a role in mitigating the UHI. The impact of TSVF is second only to ASR: a 1% increase in TSVF leads to LST decreases of 0.31%, 0.18%, and 0.06% for the 480m, 240m, and 120m grids, respectively.

Fourth, an increase in TNDVI, representing vegetation and biomass, leads to decreased LST. The average impact of TNDVI decreases, in absolute terms, from 0.10% to 0.08% to 0.05% when shifting from the 480m grid to the 240m grid to the 120m grid.

Finally, WATER impacts are always negative or nil. The zero impacts are simply due to the fact that the rivers flowing through the study area cover a limited number of cells. The presence of a water body helps in mitigating high daytime surface temperatures in Summer.

5. SIMULATION OF GREENING SCENARIOS

What are the impacts of alternative greening actions on surface temperatures? To answer this question, the 120m GSM model presented in Table 3 is used in a simulation mode, computing the temperature of each cell under a specific greening scenario. Before describing the proposed scenarios, the basic computational steps of the simulation are first presented. Assuming that the error terms u and ε are equal to 0, Eq. (13) can be rewritten in matrix form as:

$$\ln(LST) = [I-\rho W]^{-1} a_0 + [I-\rho W]^{-1} a_1 BGFA + [I-\rho W]^{-1} a_2 ASR + [I-\rho W]^{-1} a_3 TSVF + [I-\rho W]^{-1} a_4 TNDVI + [I-\rho W]^{-1} a_5 WATER \quad (14)$$

W is a matrix NxN, with N=2,288. So, of course, is [I-ρW] and its inverse, which must be computed first. The term a_0 is a column vector of dimension N, where all the elements are equal to the intercept; the terms a_1BGFA, a_2ASR, a_3TSVF, a_4TNDVI, and a_5WATER are also all column vectors of dimension N. Hence, each product with the matrix $[I-\rho W]^{-1}$ is a column vector of dimension N, and the sum on the right side of Eq. (14) is a vector of dimension N, equal to the vector ln(LST). Applying the exponential transformation to both sides of Eq, (14) leads to the generalized relationship:

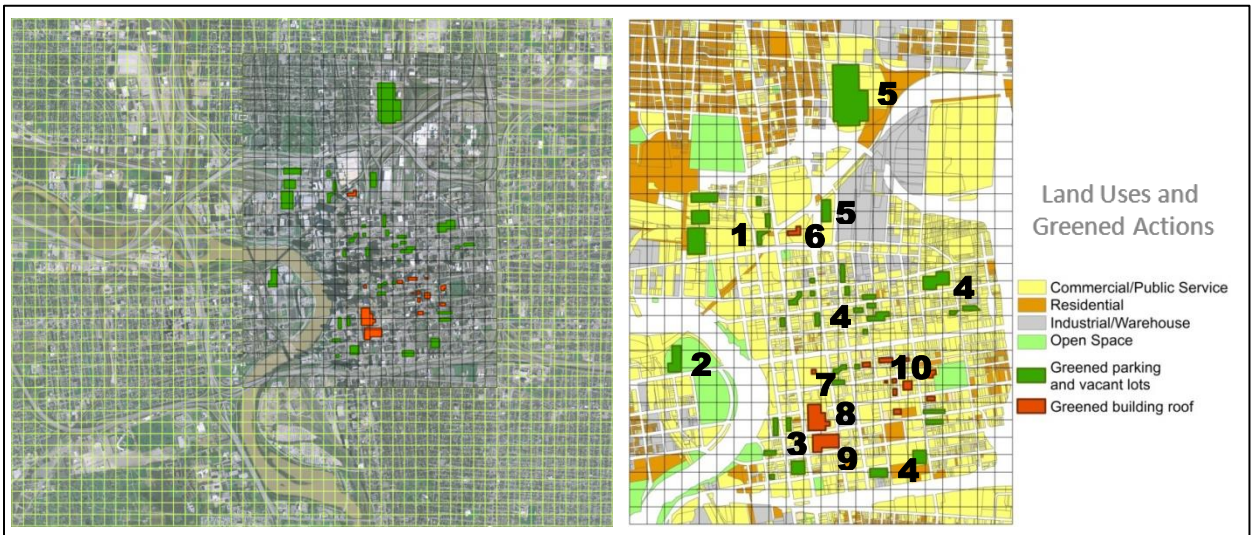
$$LST = F(BGFA, ASR, TSVF, TNDVI, WATER) \quad (15)$$

Let the subscript "o" represent before-greening actions, and the subscript "a" after-greening actions. The change in temperature ΔLST is then:

$$\Delta LST = F(BGFA_a, ASR_a, TSVF_a, TNDVI_a, WATER_a) - F(BGFA_o, ASR_o, TSVF_o, TND\ VI_o, WATER_o) \quad (16)$$

Three scenarios are considered. Scenarios 1 and 2 involve greening roofs, parking spaces, and vacant lots in a sub-area of the study area, as illustrated in Figure 4. Figure 4.a

presents the general location of the sub-area, while Figure 4.b provides the detailed locations of the greening actions. The roofs were selected based on a detailed analysis of their sizes. Note that there were no green roofs at the time of this analysis (2007). However, a few green roofs have been actually constructed since 2007 in the Columbus metropolitan area, including the Lazarus building (former department store), the Franklin County Courthouse, and the Columbus Commons underground parking garage, all within the study area. These changes have not been incorporated in the simulation, which reflects the situation in 2007. The total area of greened roofs is 82,030 m^2, and the total area of greened parking and vacant lots is 457,970 m^2.



a. General Location of Greening Actions b. Detailed Locations of Greening Actions

1. Parking spaces near Nationwide Arena; 2. Parking space near COSI; 3. Parking spaces near S. Front St.; 4. Small parking space in the downtown; 5. Vacant lot; 6. Greater Columbus Convention Center roof; 7. City Hall roof; 8. City Center roof; 9. Greyhound building roof; 10. Small building roofs

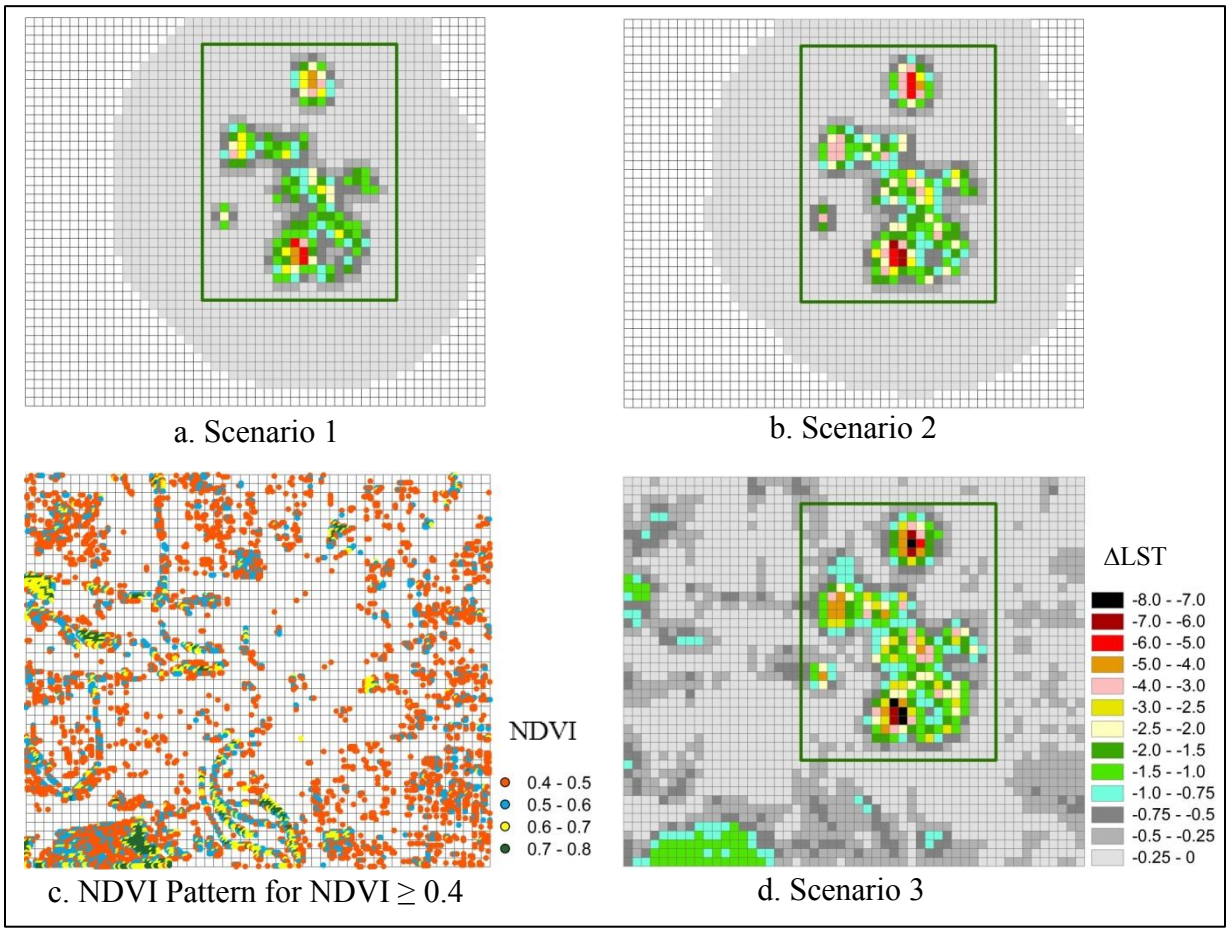
Figure 4: Simulated Greening Actions in the Center of Columbus

Scenarios 1 and 2 are characterized by NDVI values of 0.6 and 0.8 for the selected greened roofs and parking/vacant lots, respectively. The grid-based data are modified as follows to reflect the greening actions. The ground areas of greened parking/vacant lots have their NDVI changed to 0.6 or 0.8, and these areas are then apportioned over the relevant 120m cells. In the case of greened roofs, their prior-to-greening BGFA values are deleted from the database, and their new NDVI values are apportioned over the relevant cells. Scenario 3 combines Scenario 2

with an increase by 25% of the NDVI of all the areas (except those included in Scenario 2) with NDVI ≥ 0.4. NDVI statistics for all cells in the study area and for those with NDVI ≥ 0.4 are presented in Table 5. Cells with NDVI ≥ 0.4 have an average NDVI of 0.51, and a maximum NDVI of 0.79. When all cells are considered, the NDVI varies between -0.86 and 0.59, with a mean of 0.19. The NDVI pattern for the selected cells is also presented in Figure 5.c.

Table 5: NDVI Descriptive Statistics

Coverage	Average	Standard Deviation	Minimum	Maximum	25%	Median	75%
All cells	0.19	0.19	-0.86	0.79	0.03	0.18	0.33
Cells with NDVI ≥ 0.4	0.51	0.09	0.40	0.79	0.44	0.48	0.56



Scenario 1: NDVI=0.6 for greened roofs and parking/vacant lots
Scenario 2: NDVI=0.8 for greened roofs and parking/vacant lots
Scenario 3: Scenario 2 + increasing by 25% all cells with NDVI ≥ 0.4

Figure 5: UHI Mitigation Scenarios and Temperature Changes ΔLST (^{o}C)

The resulting temperature change ΔLST for Scenarios 1, 2, and 3, are illustrated in Figures 5.a, 5.b, and 5.d, respectively. While the largest temperature changes are centered on the greened roofs and parking/vacant lots, it is clear that these effects spread out beyond these roofs and lots, due to the spatial autocorrelation effects. This spread expands from Scenario 1 to Scenario 3, as expected. In the cases of Scenarios 1 and 2, ΔLST spreads beyond the sub-area with green borders on Figures 5.a and 5.b, but at a low level (by a quarter of a Celsius degree at most). However, in the case of Scenario 3, the increased NDVI for all cells on Figure 5.c leads to more widespread ΔLST. Within the green-boundary sub-area, some cells are characterized by ΔLST in the range of 7-8 ^{o}C, and clusters with ΔLST in the range 0.75-1.5 ^{o}C appear in the southern and western parts of the study area. Table 6 provides further evidence of these patterns of temperature changes. Temperature decreases greater than or equal to 3^{o}C characterize 11 cells under Scenario 1, 27 cells under Scenario 2, and 39 cells under Scenario 3. Temperature decreases in the range 1-3 ^{o}C characterize 93 cells under Scenario 1, 105 cells under scenario 2, and 170 cells under Scenario 3. These temperature changes can become inputs to further impact analyses, as discussed in Section 6.

Table 6: Number of Cells by Temperature Change Interval

Temperature Change Intervals	Scenario 1	Scenario 2	Scenario 3
[-8 - -7]	0	0	4
[-7 - -6]	0	3	4
[-6 - -5]	3	5	1
[-5 - -4]	4	1	15
[-4 - -3]	4	18	15
[-3 - -2.5]	11	8	21
[-2.5 - -2]	11	25	23
[-2 - -1.5]	29	31	31
[-1.5 - -1]	42	41	95
[-1 - -0.75]	27	34	91
[-0.75 - -0.5]	56	53	170
[-0.5 - -0.25]	71	73	581
[-0.25 – 0]	1241	1238	1237
[0]	789	758	0
Total	2288	2288	2288

6. DISCUSSION

Several issues related to the proposed methodology and its simulation implementation are first discussed, including the relationship between surface and air temperatures, the role of water, the measurement of vegetation, and the assessment of UHI-mitigating strategies. Extensions of the research approach are next outlined.

What is the relationship between satellite-derived surface temperature and air temperature, which most directly affect people? Based on their review of the available literature, Voogt and Oke (2003) conclude that the relationship depends upon the horizontal transport of heat by wind (advection), which depends upon the morphology of the urban surface. In urban canyons, insolation shading may reduce the total energy input, but long-wave energy exchanges between buildings, even when in shade, may maintain high surface and air temperatures, depending upon the sky view. Such effects should be captured, at least partially, by the GSM because it incorporates the TSVF variable. At the micro-scale, the correlation between the two temperatures can be explained by atmospheric mixing, mean wind velocity, and the thermal properties of surface materials. At larger scales typical of satellite sensors, the correlation is poorer, with substantial noise. However, at night the correlation is much improved, particularly over vegetated surfaces with reduced micro-scale advection. More recent research suggests a stronger correlation between the two temperatures. For instance, Nichol et al. (2006), using 545 temperature data pairs in Hong-Kong over a variety of surface materials, at the surface and at 1m above it, obtain a correlation coefficient of 0.97 for hourly averages. Kloog et al. (2012), using 2003 data for Massachusetts, obtain R^2=0.85 between the two temperatures. They further improve the fit to R^2=0.95 by regressing the air temperature over the surface temperature and other variables, such as wind speed, elevation, NDVI, and impervious surfaces. Thus, provided that the necessary air temperature data are made available, such a relationship could be estimated in the case of Columbus center, and could be coupled with the surface temperature GSM model to compute the air temperature changes resulting from a UHI-mitigating action.

The water of the two rivers flowing through Columbus downtown has a significant temperature reduction effect. However, it is not clear that the same effect would be obtained with non-flowing water basins that could be purposefully constructed to reduce the UHI. Further, this effect is most likely to be reversed in Winter, when the water may have a warming effect.

While the NDVI is the most common measure of vegetation in UHI studies, and has turned out to be a significant variable in the GSM model, alternative measures could be considered. For instance, Nichol et al. (2006) have proposed vegetation density (VD) to measure vegetation amount in all vertical layers above ground. VD is considered to better represent the total biomass. Vegetation height and density are likely to better account for shading effects. Another measure is the vegetation cover (VC), representing the percentage of ground covered by vegetation. Carlson and Ripley (1997) propose the Leaf Area Index (LAI), while Jenerette et al. (2007) use a soil-adjusted vegetation index (SAVI). Yuan and Bauer (2007) show that using the percentage of impervious surface area (ISA) yields better results than the NDVI. Other indices based on different satellite bands have been proposed by Chen et al. (2006), including the Normalized Difference Water Index (NDWI), Built-up Index (NDBI), and Bareness Index (NDBaI). All these variables could be tested, and might improve model estimation. Finally, surface albedo, which represents reflectance of short-wave radiation, is known to affect the UHI, and should be incorporated into statistical models when such data are available.

The simulations presented in Section 5 provide a clear measure of the impacts of greening strategies. However, temperature decreases must be further analyzed in order to reach a comprehensive cost-benefit analysis of these strategies. In particular, the resulting decline in the need for air-conditioning must be assessed, together with the subsequent decreases in electricity consumption/production and in air pollution and greenhouse-gas emissions by power plants. Also, reduced temperatures help reduce the production of tropospheric ozone. These pollution reductions must then be translated into health benefits (reduced morbidity and mortality). Further, green roofs help reduce the heat influx into buildings, further decreasing cooling needs, and also retain rainwater, reducing peak surface runoff, nonpoint water pollution, and the need for stormwater facilities. These impacts are discussed in Tsang and Jim (2011) in the case of green roofs in Hong-Kong. More generally, ground-level greening would also provide these water-related benefits, but its cost must be assessed in terms of the benefits that other land uses would have provided to residents, businesses, and city government (foregone tax revenues). Greening parking lots, an unfortunate but ubiquitous occurrence in most American cities, might require building underground parking, at a high cost, and/or a shift to public transportation, which would enhance smart growth.

The spatial statistical model could also be used to assess other, non-greening, UHI mitigating strategies, such as installing water basins in central city areas, or expanding the beds of existing rivers. Interestingly, this is currently happening to the Olentangy River in Columbus, after removal of several dams. Other strategies could involve setting restriction on distances between buildings and on their heights, to increase the SVF and enhance air circulation. However, the resulting decrease in shading and increase in solar radiations over impervious surfaces might have opposite effects, which could all be assessed by the model.

There are several lines of research that could extend the methodology presented in this paper. First, the comparison of the three grids results suggests that the best model is obtained with the smaller 120m cells (Table 3). While all the data may be captured over smaller cells (e.g., 60m or 30m), the available software is not currently able to estimate spatial models over such grids, because of the necessary size of the W matrix. In the 120m case, there are 2,288 cells, and therefore the W matrix has dimensions 2,288x2,288 and 5.235 Million entries. In the 60m and 30m cases, W would have close to 21 and 84 Million entries, respectively. Inverting such matrices is not yet routinely possible, but might be so in the near future. Second, surface temperatures, solar radiation, and NDVI vary over the year. The UHI spatial statistical model could be estimated for different periods/seasons, accounting for such changes. A multi-period model could be used to design strategies that balance contradictory UHI effects, namely that temperatures should be reduced in the hot season to reduce cooling needs, but increased in the cold season to reduce heating needs. Nevertheless, one should bear in mind the limitation in using satellite images associated with their acquisition time. Since satellite orbital periods are constant, Landsat TM always passes over the State of Ohio with a 16-day repeat cycle at 10:00am. This makes it impossible to capture surface temperatures at other times, such as early morning, late afternoon, or nighttime. Finally, the methodology could be tested over different cities/locales with different topography, climate, morphology, and land-use structure, to assess the generality of the results obtained for the city of Columbus.

7. CONCLUSIONS

The objectives of this research were to specify and estimate improved statistical regression models of the UHI, determine the best scale of capture of the necessary data, and

illustrate the use of the model(s) for designing UHI-mitigating strategies. These objectives have been, by and large, achieved.

Satellite data have been used to estimate surface temperatures and vegetation coverage, GIS data to delineate building boundaries and land uses, and LiDAR data to closely approximating the real urban infrastructure and environment in 3-D space. The resulting 3-D city model has allowed for the computation of a comprehensive set of 3-D urban characteristics. The integration of these various datasets represents an advance over past research, which has generally focused on a smaller subset of urban characteristics.

A multi-step statistical analysis, involving regression models of increasing complexity, has been implemented, culminating in the selection of the General Spatial Model (GSM), which accounts for spatial autocorrelation in both the dependent variable (LST) and the error term. This result confirms that spatial interactions between cells within a specified neighborhood characterize the UHI. Biased results would be obtained by not accounting for these effects, and this problem is likely to characterize much of the earlier UHI statistical research. In addition, the GSM is able to account for the simultaneous effects of five urban characteristics captured over regular grids: building roof-top areas (BGFA), TNDVI, solar radiations (ASR), sky view factor (TSVF), and water. The impacts of these variables is consistent with past research: increasing BGFA and ASR lead to increased surface temperatures (LST), while increasing TNDVI, TSVF, and Water lead to decreasing LST. Finally, the best GSM results are obtained with the 120m grid, which suggests that using smaller cells better captures inter-cell effects. Unfortunately, software limitations do not currently allow for testing the GSM over smaller cells (e.g, 60m or 30m).

A regression model, specified within a generalized theoretical framework, lets the data speak, but cannot represent the detailed physical processes of the UHI. For instance, open paved areas subject to all-day insolation receive more heat input, as measured by the ASR variable, but this heat would be more likely to be lost through convection, as compared to an urban canyon. Such loss is captured by the TSVF variable. In contrast, while urban canyons may be subject to less solar heat input due to mutual shading effects, heat trapping may be an overriding effect, as measured by TSVF. Conceivably, the GSM could be used to analyze hypothetical urban layouts involving clearly delineated open paved areas and urban canyons, and assess which type of area contributes more to the UHI. However, there are counteracting factors that are not accounted for

by the model and may modify the above processes: for instance, urban canyons can accelerate air flows through the Venturi effect, depending on wind velocity. Another issue is the heating contribution of roof tops, as compared to ground-level paved areas, because a building several stories high could be better ventilated. However, this ventilation might depend upon the SVF of the roof top, which is a component of the TSVF index used in the GSM. To clarify this effect, further research might use the RSVF and GSVF factors separately. Finally, the model could be improved by accounting for the albedo of roof-top and ground-level areas, when such data become routinely available.

REFERENCES

Amiri, R., Weng, Q., Alimohammadi, A., and Alavipanah, S. (2009). Spatial-temporal dynamics of land surface temperature in relation to fractional vegetation cover and land use/cover in the Tabriz urban area, Iran. *Remote Sensing of Environment*, Vol. 113: 2606-2617.

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, Norwell, MA 0261, USA.

Arnfield, A.J., 2003, Two Decades of Urban Climate Research: A Review of Turbulence, Exchanges of Energy and Water, and the Urban Heat Island, *International Journal of Climatology*, 23(1), 1-26.

Bärring, L., Mattsson, J., Lindqvist, S. (1985). Canyon geometry, street temperatures and urban heat island in Malmö, Sweden. *Journal of Climatology*, Vol. 5 (4): 433-444.

Bottyán, Z. and Unger, J. (2003). A multiple linear statistical model for estimating the mean maximum urban heat island. *Theoretical and Applied Climatology*, Vol. 75 (3-4): 233-243.

Ca, V., Asaeda, T., and Abu, E. (1998). Reductions in air conditioning energy caused by a nearby park. *Energy and Buildings*, Vol. 29: 83-92.

Carlson, T., and Arthur, S. (2000). The impact of land use-land cover changes due to urbanization on surface microclimate and hydrology: a satellite perspective. *Global and Planetary Change*, Vol. 25: 49-65.

Carlson, T.N., and Ripley, D.A. (1997). On the relation between NDVI, fractional vegetation cover, and leaf area index. *Remote Sensing of Environment*, Vol. 62: 241-252.

Chander, G. and Markham, B. (2003). Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges. *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 41 (11): 2674-2677.

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

Eliasson, I. (1996). Urban nocturnal temperatures, street geometry and land use. *Atmospheric Environment*, Vol. 30 (3): 379-392.

ESRI (2009). *ArcGIS Desktop 9.3 Tutorials*. Environmental Systems Research Institute, Inc

Gál, T., and Unger, J. (2008). *An ArcView extension for SVF estimation based on vector format 3D urban surface database: User's manual*, Department of Climatology and Landscape Ecology, University of Szeged, Szeged, Hungary.

Gál, T., Lindberg, F., and Unger, J. (2009). Computing continuous sky view factors using 3D urban raster and vector databases: comparison and application to urban climate. *Theoretical and Applied Climatology*, Vol. 95: 111-123.

Giridharan, R., Ganesan, S., and Lau, S. (2004). Daytime urban heat island effect in high-rise and high-density residential developments in Hong Kong. *Energy and Buildings*, Vol.36: 525-534.

Grimmond, S. (2007). Urbanization and global environmental change: local effects of urban warming. *Geographical Journal*, Vol. 173: 75-79.

Hammerle, M., Gal, T., and Unger, J. (2011). Comparison of models calculating the sky view factor used for urban climate investigations. *Theoretical and Applied Climatology*, Vol. 105: 521-527.

Jauregui, E., 1990, Influence of a Large Urban Park on Temperature and Convective Precipitation in a Tropical City, *Energy and Buildings*, 15-16, 457-463.

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

Katpatal, Y., Kute, A., and Satapathy, D. (2008). Surface and air temperature studies in relation to land use/land cover of Nagpur urban area using Landsat 5 TM data. *Journal of Urban Planning and Development*, Vol. 134 (3): 110-118.

Kato, S. and Yamaguchi, Y. (2007). Estimation of storage heat flux in an urban area using ASTER data. *Remote Sensing of Environment*, Vol. 110: 1-17.

Kim, J.-P. (2013). Variation in the accuracy of thermal remote sensing. *International Journal of Remote Sensing*, Vol. 34 (2): 729-750.

Kjelgren, R., and Montague, T. (1998). Urban tree transpiration over turf and asphalt surfaces. *Atmospheric Environment*, Vol. 32 (1): 35-41.

Kloog, I., Chudnovsky, A., Koutrakis, P., and Schwartz, J. (2012). Temporal and spatial assessments of minimum air temperature using satellite surface temperature measurements in Massachusetts, USA. *Science of the Total Environment*, Vol. 432: 85-92.

Landsberg, H., and Maisel, T. (1972). Micrometeorological observations in an area of urban growth. *Boundary-Layer Meteorology*, Vol.2: 365-370.

LeSage, J. and Pace, K. (2009). *Introduction to spatial econometrics*. CRC Press/Taylor & Francis Group, Boca Raton, FL 33487, USA.

Li, S., Zhao, Z., Miaomiao, X., and Wang, Y. (2010). Investigating spatial non-stationary and scale-dependent relationships between urban surface temperature and environmental factors using geographically weighted regression. *Environmental Modeling & Software*, Vol. 25: 1789-1800.

Li, W., Putra, S.Y., and Yang, P.P. (2004). GIS Analysis for the Climatic Evaluation of 3D Urban Geometry – The Development of GIS Analysis Tools for Sky View Factor. In Proceedings of GIS in Developing Countries (GISDECO) 2004 Conference, Johor Bahru, Malaysia.

Lynn, B.H., Carlson, T.N., Rosenzweig, C., Goldberg, R., Druyan, L., Cox, J., Gaffin, S., Parshall, L., and Civerolo, K. (2009). A modification to the NOAH LSM to simulate heat mitigation strategies in the New York City metropolitan area. *Journal of Applied Meteorology and Climatology*, Vol.48: 199-216.

Nichol, J., Wong, M.S., Fung, C., and Leung, K.K.M. (2006). Assessment of urban environmental quality in a subtropical city using multispectral satellite images. *Environment and Planning B: Planning and Design*, Vol. 33: 39-58.

Oke, T. (1981). Canyon geometry and the nocturnal urban heat island: comparison of scale model and field observations. *Journal of Climatology*, Vol. 1 (3): 237-254.

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

Oke, T. (1988). Street design and urban canopy layer climate. *Energy and Building*, Vol. 11 (1-3): 103-113.

Owen, T, Carlson, T. and Gilles, R. (1998). An assessment of satellite remotely sensed land cover parameters in quantitatively describing the climatic effect of urbanization. *International Journal of Remote Sensing*, Vol. 19 (9): 1663-1681.

Roth M., Oke T.R. and Emery W.J., 1989, Satellite-Derived Urban Heat Islands from Three Costal Cities and the Utilization of such Data in Urban climatology, *International Journal of Remote Sensing*, 10, 1699–1720.

Shashua-Bar, L. and Hoffman, M.E., 2000, Vegetation as a Climatic Component in the Design of an Urban Street: An Empirical Model for Predicting the Cooling Effect of Urban Areas with Trees, *Energy and Buildings*, 31, 221-235.

Stathopoulou M. and Cartalis C., 2009, Downscaling AVHRR Land Surface Temperatures for Improved Surface Urban Heat Island Intensity Estimation, *Remote Sensing of Environment*, 113, 2592-2605.

Tsang, S.W., and Jim, C.Y. (2011). Game-theory approach for resident coalitions to allocate green-roof benefits. *Environment and Planning A*, Vol. 43(2): 363-377.

Unger, J., Bottyán, Z., Sümeghy, Z, and Gulyás, Á. (2004). Connections between urban heat island and surface parameters: measurements and modeling. *Quarterly Journal of the Hungarian Meteorological Service*, Vol. 108 (3): 173-194.

Unger, J. (2006). Modeling of the annual mean maximum urban heat island using 2D and 3D surface parameters. *Climate Research*, Vol.30: 215-226.

Unger, J. (2009). Connection between urban heat island and sky view factor approximated by a software tool on a 3D urban database. *Internal Journal of Environment and Pollution*, Vol. 36: 59-80.

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

Yuan, F., and Bauer, M. (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*, Vol. 106: 375-386.

Xiao, R., Ouyang, Z., Zheng, H., Li, W., Schienke, E., and Wang, X. (2007). Spatial pattern of impervious surfaces and their impacts on land surface temperature in Beijing, China. *Journal of Environmental Sciences*, Vol. 19 (2): 250-256.

Appendix A – Variables and Acronyms

ΔLST:	Land Surface Temperature Change
2-D:	Two-Dimensional
3-D:	Three-Dimensional
AGSVF:	Average Ground-Level Sky View Factor
ARSVF:	Average Roof-Level Sky View Factor
ASR:	Average Solar Radiation
BGFA:	Building Ground Floor Area
CBD:	Central Business District
DN:	Digital Number
DTM:	Digital Terrain Model
GIS:	Geographic Information System
GSM:	General Spatial Model
GSVF:	Ground Sky View Factor
Landsat TM:	Landsat Thematic Mapper
LiDAR:	Light Detection and Ranging
LM:	Lagrange Multiplier
LST:	Land Surface Temperature
NDVI:	Normalized Difference Vegetation Index
NIR:	Near Infra-Red
OLS:	Ordinary Least Squares
RSVF:	Roof Sky View Factor
SA:	Spatial Autocorrelation
SAR:	Spatial Autoregressive (Lag) Model
SEM:	Spatial Error Model
SR:	Solar Radiation

SVF: Sky View Factor
TNDVI: Total Normalized Difference Vegetation Index
TSR: Total Solar Radiation
TSVF: Total Sky View Factor
UHI: Urban Heat Island
W: Neighborhood Structure Matrix

Appendix B – Neighborhood Structure Matrix W

The matrix W reflects the concept of binary contiguity between spatial units. Consider the hypothetical grid made of 16 cells, as presented in Figure B.1.a. Assume that two cells are neighbors of each other if they have a common edge or vertex. Consider cell 1. Its neighbors are then cells 2, 5, and 6. These connections are reflected in the first line of the 16x16 matrix in Figure B.1.b., where a 1 appears in columns 2, 5, and 6, while all the other entries are set to 0. Likewise, the neighbors of cell 10 are cells 5, 6, 7, 9, 11, 13, 14, and 15, and the 1's in line 10 of the matrix reflect these relationships. Generally, the matrix in Figure B.1.b is converted to the standardized weight matrix in Figure B.1.c by dividing the elements in each line by their sum. In the case of line 1, the sum in Figure B.1.b is 3, hence the weights 0.333 in Figure B.1.c. Let X be the vector of a variable x taking values over the 16 cells. The product WX, a column vector of dimension16, then represents the average values of x over the neighbors of each of the 16 cells. The above contiguity is called queen contiguity. Rook contiguity characterizes connections through edges only, while bishop contiguity only involves vertex connections only. First-order contiguity involves only directly contiguous cells, while second-order contiguity involves, in addition, the cells contiguous to the first-order cells. Higher-order contiguity is defined similarly. There are, of course, other ways to define the matrix W. The maximum-distance approach involves all the cells whose centroids are within a given distance from the centroid of the cell considered. Also, the weights may be modified to reflect the declining effect of distance, as in the gravity model. For further insights into spatial econometrics methods, see, for instance, Anselin (1988).

1	5	9	13
2	6	10	14
3	7	11	15
4	8	12	16

a. Sample grid

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	1	0	0	1	1	0	0	0	0	0	0	0	0	0	0
2	1	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0
3	0	1	0	1	0	1	1	1	0	0	0	0	0	0	0	0
4	0	0	1	0	0	0	1	1	0	0	0	0	0	0	0	0
5	1	1	0	0	0	1	0	0	1	1	0	0	0	0	0	0
6	1	1	1	0	1	0	1	0	1	1	1	0	0	0	0	0
7	0	1	1	1	0	1	0	1	0	1	1	1	0	0	0	0
8	0	0	1	1	0	0	1	0	0	0	1	1	0	0	0	0
9	0	0	0	0	1	1	0	0	0	1	0	0	1	1	0	0
10	0	0	0	0	1	1	1	0	1	0	1	0	1	1	1	0
11	0	0	0	0	0	1	1	1	0	1	0	1	0	1	1	1
12	0	0	0	0	0	0	1	1	0	0	1	0	0	0	1	1
13	0	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0
14	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	1
15	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1
16	0	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0

b. First-order queen's contiguity relations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0	0.333	0	0	0.333	0.333	0	0	0	0	0	0	0	0	0	0
2	0.2	0	0.2	0	0.2	0.2	0.2	0	0	0	0	0	0	0	0	0
3	0	0.2	0	0.2	0	0.2	0.2	0.2	0	0	0	0	0	0	0	0
4	0	0	0.333	0	0	0	0.333	0.333	0	0	0	0	0	0	0	0
5	0.2	0.2	0	0	0	0.2	0	0	0.2	0.2	0	0	0	0	0	0
6	0.125	0.125	0.125	0	0.125	0	0.125	0	0.125	0.125	0.125	0	0	0	0	0
7	0	0.125	0.125	0.125	0	0.125	0	0.125	0	0.125	0.125	0.125	0	0	0	0
8	0	0	0.2	0.2	0	0	0.2	0	0	0	0.2	0.2	0	0	0	0
9	0	0	0	0	0.2	0.2	0	0	0	0.2	0	0	0.2	0.2	0	0
10	0	0	0	0	0.125	0.125	0.125	0	0.125	0	0.125	0	0.125	0.125	0.125	0
11	0	0	0	0	0	0.125	0.125	0.125	0	0.125	0	0.125	0	0.125	0.125	0.125
12	0	0	0	0	0	0	0.2	0.2	0	0	0.2	0	0	0	0.2	0.2
13	0	0	0	0	0	0	0	0	0.333	0.333	0	0	0	0.333	0	0
14	0	0	0	0	0	0	0	0	0.2	0.2	0.2	0	0.2	0	0	0.2
15	0	0	0	0	0	0	0	0	0	0.2	0.2	0.2	0	0.2	0	0.2
16	0	0	0	0	0	0	0	0	0	0	0.333	0.333	0	0	0.333	0

c. Standardized weight matrix

Figure B.1: Spatial Neighborhood Matrix