COMPARING URBAN VEGETATION COVER WITH SUMMER LAND SURFACE TEMPERATURE IN THE SALT LAKE VALLEY

by

Joshua Deshawn Reynolds

A thesis submitted to the faculty of
The University of Utah
in partial fulfillment of the requirements for the degree of

Master of Science

Department of Geography

The University of Utah

May 2017

The University of Utah Graduate School

STATEMENT OF THESIS APPROVAL

The thesis of **Joshua Deshawn Reynolds**

has been approved by the following supervisory committee members:

Philip Dennison	, Chair	**01/05/2017** Date Approved
Simon Brewer	, Member	**01/05/2017** Date Approved
Ran Wei	, Member	**01/05/2017** Date Approved

and by **Andrea Brunelle** , Chair/Dean of

the Department/College/School of **Geography**

and by David B. Kieda, Dean of The Graduate School.

ABSTRACT

The climate of urban areas is influenced by the composition and configuration of different land cover types. Urban forests increase human comfort in urban areas by cooling the environment through evapotranspiration and shade. A tradeoff of urban forests in semiarid and arid climates is that they require large quantities of irrigation water to maintain. This study aimed to quantify the relationship between urban vegetation and land surface temperature (LST). Datasets derived from high-resolution lidar and National Agriculture Imagery Program (NAIP) orthoimagery were used in a random forest algorithm to classify urban vegetation, human-made and natural surfaces at a 1-meter scale, in the Salt Lake Valley of Utah. The resulting classification accuracy was 94%. LST was retrieved from an Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) scene captured on a hot, summer day. Percentages of each land cover class were calculated per ASTER pixel. These composition variables were compared to LST using Pearson's correlation analysis and were also used to create a multiple linear regression model. Percent deciduous tree cover was the variable most strongly correlated with LST, with a correlation coefficient of -0.55. Irrigated low-stature vegetation was also negatively correlated with LST (-0.33). Residuals from the multiple linear regression model varied over space, and additional dates of ASTER imagery are needed to determine whether these second-order spatial patterns are persistent.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

I was fortunate to be able to work with Dr. Philip Dennison after switching thesis projects midway through my master's studies. He provided just the right amount of guidance and direction during times of uncertainty. Also, he was always available if I had a question or needed advice. I would like to thank Austin Coates, Brent Lloyd, Sean Reid, and Melanie Cooke for devising the idea for this project and giving me a platform from which to launch my study. Special thanks to Austin for preprocessing the datasets.

Additionally, I would like to thank my undergraduate advisor and mentor, Dr. Steven Prager. He continues to provide great, objective advice and served as a fantastic role model during my academic career. It is safe to say that I would have never considered attending graduate school had I not met him.

I sincerely appreciate Dr. Phoebe McNeally and the DIGIT Lab for providing an environment for me to grow as a GIS professional. The experience I gained at the DIGIT Lab will prove invaluable in the field of GIS for years to come. In addition, I am eternally grateful for the additional semester of funding granted to me. Without this funding, it would have been tremendously difficult to graduate in a timely manner.

Finally, I am blessed to have a supportive family in my Mother, Aunt, Grandmother, and Grandfather. Despite living hundreds of miles away, they are present whenever I need verbal, moral, or financial support. Most importantly, they continue to encourage me to surpass my own expectations of myself.

INTRODUCTION

As of 2013, a large percentage (52%) of the world's population now resides in urban areas (Population Reference Bureau, 2013). This percentage continues to increase as populations grow and people migrate from rural areas to seek employment. Human comfort in urban areas is often linked with ambient temperatures. As urban areas expand, existing vegetation is often replaced with human-made surfaces. This has important implications for air and surface temperatures within an urban environment.

One consequence of urbanization is the urban heat island (UHI) effect. When natural vegetation is replaced with human-made surfaces, the area's ability to moderate temperatures worsens. Human-made surfaces like asphalt are less reflective than natural surfaces, leading to increased absorption of solar radiation. Additionally, decreasing the amount of natural vegetation in an area leads to a decline in total evapotranspiration (Akbari, 2009). As a result, air and surface temperatures tend to be several degrees higher in urban areas than their surrounding areas (Huang et al., 2011). Studies related to the UHI are critical for a few reasons. First, UHI increases demand for air conditioning during warmer months. This results in increased energy use. Second, UHI negatively impacts human health by increasing ground level ozone (Akbari et al., 1996). Third, UHI subjects a city's inhabitants to a health hazard known as heat stress. Evidence from previous research suggests that more hospitalizations and emergency calls occur during heat waves (Kinney et al., 2001).

Cities can mitigate the effects of UHI by maintaining an urban forest (Dimoudi & Nikolopoulou, 2003). Urban forests are known to have an "oasis" effect on a city's climate during the warmer months (Dwyer et al., 1992; Nowak & Dwyer, 2007). Large amounts of solar radiation are absorbed by trees and green vegetation and used for evapotranspiration (Akbari, 2009). Evapotranspiration lowers nearby air temperatures by absorbing energy through the latent heat of vaporization (Grimmond & Oke, 1991). Tall vegetation like trees also cool the environment by providing shade, thus reducing the solar heating of nearby surfaces. When combined, these effects have the ability to reduce air temperatures by as much as 5 degrees Celsius (Akbari, 2009). Urban forests also benefit the environment by reducing energy use and decreasing air pollution (Akbari, 2009; Dwyer et al., 1992). One tradeoff of maintaining an urban forest is that it requires a large quantity of water for irrigation purposes. This is an issue in semiarid climates where water resources are limited.

The goal of this research was to quantify the relationship between urban vegetation and land surface temperature (LST) in the Salt Lake Valley. This will be accomplished in two parts. Part one will involve classifying the valley into several land cover classes. Because of the size of the study area and characteristics of the input data, a random forest classifier will be used. Random forest processes high-dimensional datasets quickly and resists overfitting to the training data (Breiman, 2001; Rodriguez-Galiano et al., 2012). Therefore, the first research question answered by this thesis was how effective is random forest for classifying an urban area using orthoimagery and lidar data? This question was addressed by analyzing measures of accuracy related to using random forest such as variable importance, out-of-bag error, and a confusion matrix. Part

two involved quantifying the relationship between established urban vegetation classes and LST. Hence, the second research question was how is the cover of urban vegetation types in the Salt Lake Valley related to land surface temperature? This question was addressed by comparing land cover classes to LST, using Pearson's correlation analysis and multiple linear regression.

BACKGROUND

Urban Vegetation and LST

Measuring the impact of urban vegetation on daytime LST using remotely sensed data has been the subject of numerous studies. The normalized difference vegetation index (NDVI) has been frequently used to quantify urban greenness (Buyantuyev & Wu, 2010; Weng et al., 2003; Yuan & Bauer, 2007). Yuan and Bauer (2007) used NDVI and percent impervious surfaces from a linear spectral mixing model to analyze LST variations. They concluded that the relationship between NDVI and LST was nonlinear and strongly affected by season. They found that percent impervious surface had a very strong linear relationship with LST. Weng et al. (2004) noted in their study that the relationship between NDVI and LST was very weak. They found NDVI to be a decent predictor, but alluded that it does not necessarily measure the amount of vegetation in an area. Finally, Yuan and Bauer (2007) concluded that NDVI is not a suitable predictor of LST on its own, but is useful for supplementing analyses with other variables.

Other studies assessed additional methods of quantifying urban vegetation and have tested their ability to estimate LST. Weng et al. (2004) compared NDVI to vegetation fraction from a spectral mixture model in terms of each metric's ability to indicate LST. Unmixed vegetation fraction showed a slightly stronger negative correlation with LST than NDVI at all spatial resolutions examined (30 meters to 960 meters). Zhou et al. (2011) compared the use of land cover composition metrics and land

cover configuration metrics to describe LST. Composition refers to the variety and relative abundance of land cover features in an area (Gustafson, 1998; Turner, 2005).

Alternatively, configuration refers to the spatial arrangement or distribution of land cover features. Coarse vegetation (included trees and shrubs) and fine vegetation (included herbs and grasses) were among urban vegetation classes considered. Composition variables such as percent coarse vegetation, building, and paved surfaces were strongly correlated to LST. Inversely, percent fine vegetation, percent water, and percent soil were weakly correlated with LST. Areas composed of smaller percentages of coarse vegetation and fine vegetation had stronger positive relationships with LST. Incidentally, these areas also had larger percentages of building and paved surfaces. Areas composed of larger percentages of coarse vegetation and fine vegetation had stronger negative relationships with LST (Zhou et al. 2004).

In regards to configuration of urban vegetation, the significance and effect of each metric varied by land cover class (Zhou et al. 2004). The authors found that decreases in LST generally corresponded to increases in edge density, shape complexity, and variability in coarse vegetation. This suggested that an evenly distributed pattern of coarse vegetation decreases LST more effectively than a clustered pattern. Overall, after testing five linear models, the authors concluded that composition was found to be a more important indicator of LST than configuration (Zhou et al. 2011). Connors et al. (2013) also evaluated the ability of several configuration metrics in predicting LST. Their study took place in Phoenix, Arizona, which has an arid climate. Three microscale environments were examined: mesic residential (above average irrigation), xeric residential (little to no irrigation), and industrial. They found that the configuration

variables of edge complexity and patchiness affected LST in the industrial area. Results indicated that less complex shapes of grass and increases in patch density of impervious cover help cool the environment more efficiently. The big takeaway from their research was that landscape composition and configuration influence LST; however, these relationships are not consistent across every land use type. Also they found that the impacts of the configuration variables are codependent and none of them stood out from each other (Connors et al. 2013).

Du et al. (2016) reopened the question of whether composition or configuration was more important in predicting LST. The authors used multilevel models to examine the problem in a nontypical manner. The premise of these models is that explanatory variables have multiple levels with lower level variables being nested into higher ones. The authors compared the ability to predict LST of an ordinary least squares regression model (a single-level model) versus that of three different multilevel models. The authors rationalized that land cover configuration provides context for composition, therefore it was used as a level two variable with composition as a level one variable. In their analysis, the multilevel model outperformed the ordinary least squares regression model by having more accurate results, smaller residuals, and less spatial autocorrelation (Du et al. 2016). Also they found that configuration metrics had stronger coefficients with LST than composition metrics, which directly contrasts with the findings of Zhou et al. (2011). Their reasons were that the multilevel approach accounted for autocorrelation, multiple levels in the analysis were used, and that their study took place in a more modern city with evenly distributed buildings separated by vegetation patches. They concluded that it may be more efficient for a city to optimize configuration of human-made and vegetated

surfaces in order to reduce LST.

Previous studies also indicated that scale seems to be an important aspect of UHI. Weng et al. (2004) performed a fractal analysis on their vegetation fraction and LST data, and found that image texture increased with pixel aggregation from scales 30 meters to 120 meters, then decreased afterwards. The largest correlations were observed at a spatial resolution of 120 meters, which happens to be the operational scale of most temperature variations. Zhou et al. (2011) also noted that their study was performed on individual patches of their study area (average area of the patches was 76,000 square meters) and that greater correlations likely existed at coarser scales (e.g., the entire study area). These findings suggest that the areal unit of measure has a direct correspondence with the radiative, thermal, and moisture properties of the surface that determine LST. According to Gluch et al. (2006), performing fine-scale studies on LST response to land cover is equally important as coarse-scale studies because radiative transfer can vary significantly over space due to the diversity of urban land cover types and their respective physical properties.

Urban Land Cover Classification Using Random Forest

In order to accurately compare urban land cover to LST, it is necessary to obtain a high-resolution land cover dataset by using a classifier of some kind. Due to its ability to use a large variety of predictor variables and minimal need for fine-tuning, random forest has proven to be a valuable classifier in land cover studies (Gislason et al., 2006; Pal, 2005). In one such study, Gislason et al. (2006) used remotely sensed data and geographic data in a random forest to classify a mountainous area in Colorado. The

authors included Landsat Multispectral Scanner four-band imagery, elevation, slope, and aspect for use as predictors in the analysis. In their study, they compared the result of the random forest classification to several other classifiers, including classification and regression trees method, and bagging and boosting-based classifiers. Random forest proved more accurate than the basic CART implementation by 4.5%. In addition, it had comparable accuracies to the bagging and boosting classifiers, with much greater speeds. The authors noted that the advantage of using random forest is that it requires less human guidance than other methods. It also does not need to be tweaked, although one can add or remove input variables to adjust accuracy. The authors concluded that random forest was a very desirable method for classification using data from multiples sources.

Light detection and ranging (lidar) data have been used in random forest analyses in a few studies for the purpose of improving urban area classification. Chehata et al. (2009) used random forest to select predictors from lidar-derived datasets. Twenty-one predictors (height-based, echo-based, eigenvalue-based, local plane-based, and full-waveform-based) derived from lidar data were reduced to six predictors using an iterative backward feature elimination technique. This is performed by repeatedly removing predictor variables with the least significance from the variable importance analysis and rerunning the classification until the lowest out-of-bag error is achieved. They concluded that each individual land cover class has its own combination of predictors with differing amounts of importance.

Guan et al. (2012) completed a similar study in which they evaluated the performance of random forest for selecting predictors for an urban area classification using lidar data and orthoimagery. The authors began with 48 total predictors in their

random forest analysis and narrowed the set to 15 by also using the iterative backwards feature elimination technique. The most important predictors were the normalized digital surface model, eigenvalue anisotropy, and intensity grey-level co-occurrence matrix measures. The authors concluded that random forest is a useful tool for selecting orthoimagery and lidar-based predictors for urban classification. They also concluded that using a large number of predictor variables in the random forest model does not necessarily lead to increased accuracy.

Finally, Guo et al. (2011) discussed in greater detail the ability of random forest to classify an urban environment using both orthoimagery and lidar data. This study along with others concluded that height difference derived from lidar was the most important predictor for all classes because it allows off-ground objects to be distinguished from on-ground objects (Chehata et al., 2009; Guan et al., 2012; Guo et al. 2011). The next most important predictors were the red band, blue band, amplitude of returns, and cross section of returns. Random forest classified impervious surfaces (i.e., "human-made ground and building classes) with minimal error. However, the algorithm had difficulties with distinguishing natural ground and human-made ground. It was purported by the authors that a lack of training data was to blame. However, an addition of a near-infrared band might have aided the analysis. After several runs using different predictor variables, it was determined that neither orthoimagery nor lidar are sufficient enough on their own to classify an urban landscape. Both are needed for an accurate urban land cover classification.

This research presents a chance to test random forest's ability to classify an urban landscape by incorporating orthoimagery and lidar data. Landscape classification in

urban heat island studies is often performed using object-based classification (Connors et al., 2013; Du et al., 2016; Zhou et al., 2011). However, Weng et al. (2004) and Fu and Weng (2016) used decision tree classifiers like random forest in their analysis and obtained accurate results. This research also presents an opportunity to learn more about the relationship between urban land cover and LST in a semiarid environment. Gluch et al. (2006) previously performed an urban heat island study in the Salt Lake area by obtaining descriptive statistics about LST for each land cover class. This research seeks to expand upon that knowledge by implementing random forest and by using composition variables to model urban vegetation cover's relationship with LST.

METHODS

Study Area

This study took place in the Salt Lake Valley in Utah. The valley is a dry lakebed left by Lake Bonneville and its natural surfaces are composed of salt desert, wetlands, foothills, and agricultural lands (Gluch et al., 2006). Native vegetation is sparse and resistant to drought because of the valley's low annual precipitation (30-38 cm). The valley is about 1,300 square kilometers in size and the average elevation is about 1,319 meters. The valley is surrounded by the Great Salt Lake to the north, the Oquirrh Mountains to the west, and the Wasatch Mountains to the east. This area encompasses Salt Lake City and nearby cities such as Draper, Murray, West Jordan, and West Valley City. Interstate 15 runs north-south through the location and serves as a primary route of transportation. As of July 2015, the valley has an estimated total population of about 1.1 million and is experiencing rapid urban and population expansion (Lowry Jr., 2009; U.S. Census Bureau, 2015). Extensive irrigation has made it possible for the valley to have a distinct and well-maintained urban forest that would otherwise not exist in a semiarid climate. This urban forest results in cooler daytime summer temperatures, at the cost of limited water resources for irrigation. The study boundary for this analysis is a result of the intersecting extents of the datasets described in the *Data* section and is approximately 840 square kilometers in size. A map depicting the study boundary and its location in a larger geographic context is shown in Figure 1.

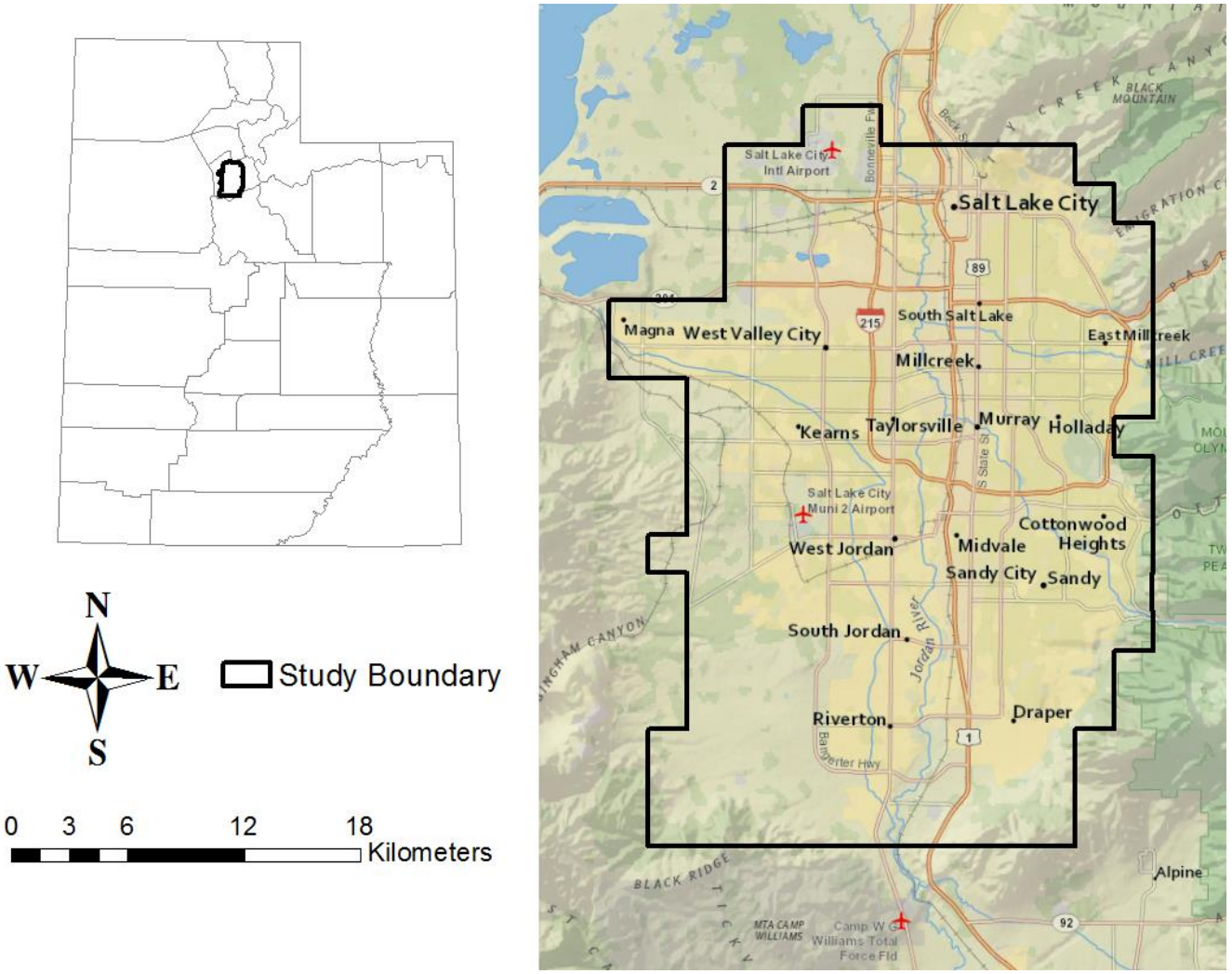

Figure. 1. Map of Salt Lake Valley. Base map courtesy of ESRI.

Data

In order to classify the study area, datasets that described and captured the differences between individual land cover classes were obtained. The predictor variables used in this study originated from high-resolution orthoimagery and lidar height data. Additional variables obtained include vegetation indices created from the orthoimagery and neighborhood rasters created from both the orthoimagery and lidar data. Each predictor variable is described in detail below.

NAIP Orthoimagery

The National Agricultural Imagery Program (NAIP) captured high-resolution aerial photography for the entire state of Utah during the summer of 2014. The orthorectified imagery has bands spanning four wavelengths (blue, green, red, and near-infrared) and was captured at a spatial resolution of 1 meter. NAIP image bands represent contrast adjusted relative brightness values. This dataset was acquired from Utah's Automated Geographic Reference Center (AGRC), the state's resource for GIS datasets (Utah AGRC, 2015). Each of the four bands was used as a predictor variable.

Vegetation indices were also created using the bands of the NAIP imagery. NDVI is one of the most widely used vegetation indices. The index is based on the premise that healthy vegetation canopies strongly absorb light in the red portion of the spectrum and strongly reflect light in the near-infrared portion of the spectrum (Rouse Jr et al., 1974; Weier & Herring, 2011). The formula for calculating NDVI is expressed as:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

High values (0.6 to 0.9) indicate dense vegetation in the pixel, moderate values (0.2 to

0.5) indicate sparse or senescing vegetation, and low values (0.1 to -1.0) indicate soil, rock, snow, or human-made surfaces (Weier & Herring, 2011). The green-red vegetation index (GRVI) was selected to compliment NDVI (Motohka et al., 2010). This particular index is useful for distinguishing green vegetation, soils, water, and snow. The formula for calculating GRVI is expressed as:

$$GRVI = \frac{GREEN - RED}{GREEN + RED}$$

The major difference that GRVI has from NDVI is it utilizes the green band rather than the near infrared band. This allows it to recognize leaves in different stages of phenology, in contrast to measuring greenness (Motohka et al., 2010).

Lidar

Lidar is a valuable active remote sensing technique for obtaining georeferenced information about the shape and surface (including heights) of objects on the Earth at high spatial resolution (Schmid et al., 2008). This technique uses laser light pulses collected by plane to create a densely packed point cloud. The location of each point in the cloud is obtained from onboard GPS (Gatziolis & Andersen, 2008). The heights of the points (also called returns) are determined by measuring the time delay between the emission of the pulse and return of the signal. In order to generate surface models, the point cloud is filtered by return number (e.g., first return, second return, bare earth) using outside software. Two advantages of lidar data are its high spatial resolution and ability to penetrate tree canopies.

Lidar data used in this research were collected during fall 2013 and spring 2014 by Quantum Spatial through partnerships developed by the Utah Geological Survey and

the AGRC (Utah AGRC, 2015). In this particular dataset, the point cloud was preprocessed and downloaded as first return and bare earth raster surfaces. In addition, a last return surface model was processed from the point cloud using the OpenTopography web application (OpenTopography, 2014). The initial spatial resolution of the data was 0.5 meters, but the data were resampled to 1 meter using spatial averaging to match the spatial resolution of the NAIP dataset.

The original lidar datasets were not used as predictor variables because they were biased by the elevation gradient present within the valley. However, three normalized derivatives using those datasets were created for use in the model. First, a First Return-Bare Earth Difference model (FBD), which highlights tall features and remove background topography, was created by subtracting the bare earth surface from the first return surface. This surface shows the difference between the highest point of an object and the ground beneath it. Taller objects have higher values and shorter objects have lower values. Second, a Last Return-Bare Earth Difference model (LBD), which highlights buildings, was created by subtracting the bare earth surface from the last return surface. Because lidar is unable to penetrate rooftops, last return values for buildings are situated higher than other objects. The difference values between the last return and bare earth surface for buildings are much larger than for other objects. Finally, a First Return-Last Return Difference model (FLD) that distinguished tree types was created by subtracting the last return surface from the first return surface. Lidar is much less likely to penetrate through coniferous trees to the ground, than deciduous trees (especially during the fall and early spring). Therefore, the difference values between the first return and last return were greater for deciduous trees than coniferous trees.

Neighborhood Metrics

Neighborhood range and standard deviation rasters for 3 meter by 3 meter windows were generated from each of the individual bands, vegetation metrices, and lidar derivatives. Focal range was selected to capture cells that had high amounts of variation in 3x3 windows. In this surface, the value of each middle cell in a 3x3 window is calculated using the range of the surrounding cells (highest value minus the lowest value). For example, a window that contained both a tall object and a ground-level object would have a higher range value in the FBD raster. A window with only tall objects or only ground-level objects would have lower range value. Focal standard deviation was also used to capture variation in a 3x3 neighborhood. In this surface, the value of each middle cell in a 3x3 window is calculated to the standard deviation of the surrounding cells. This dataset helped locate values that differed significantly from the mean, thereby identifying outliers. Neighborhood metrics for 5 meter by 5 meter windows were tested but did not improve classification accuracy as much as the 3x3 metrics, therefore they were not used.

ASTER Land Surface Temperature

A raster surface depicting LST was obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER). ASTER is an instrument within the Terra platform that measures thermal infrared radiance at a spatial resolution of 90 meters. The processed image is show in Figure 2. LST is described as the radiative skin temperature of the ground. LST is not the same air temperature. This product, also known as (AST_08), was calculated using the five thermal infrared bands between the 8 and 12

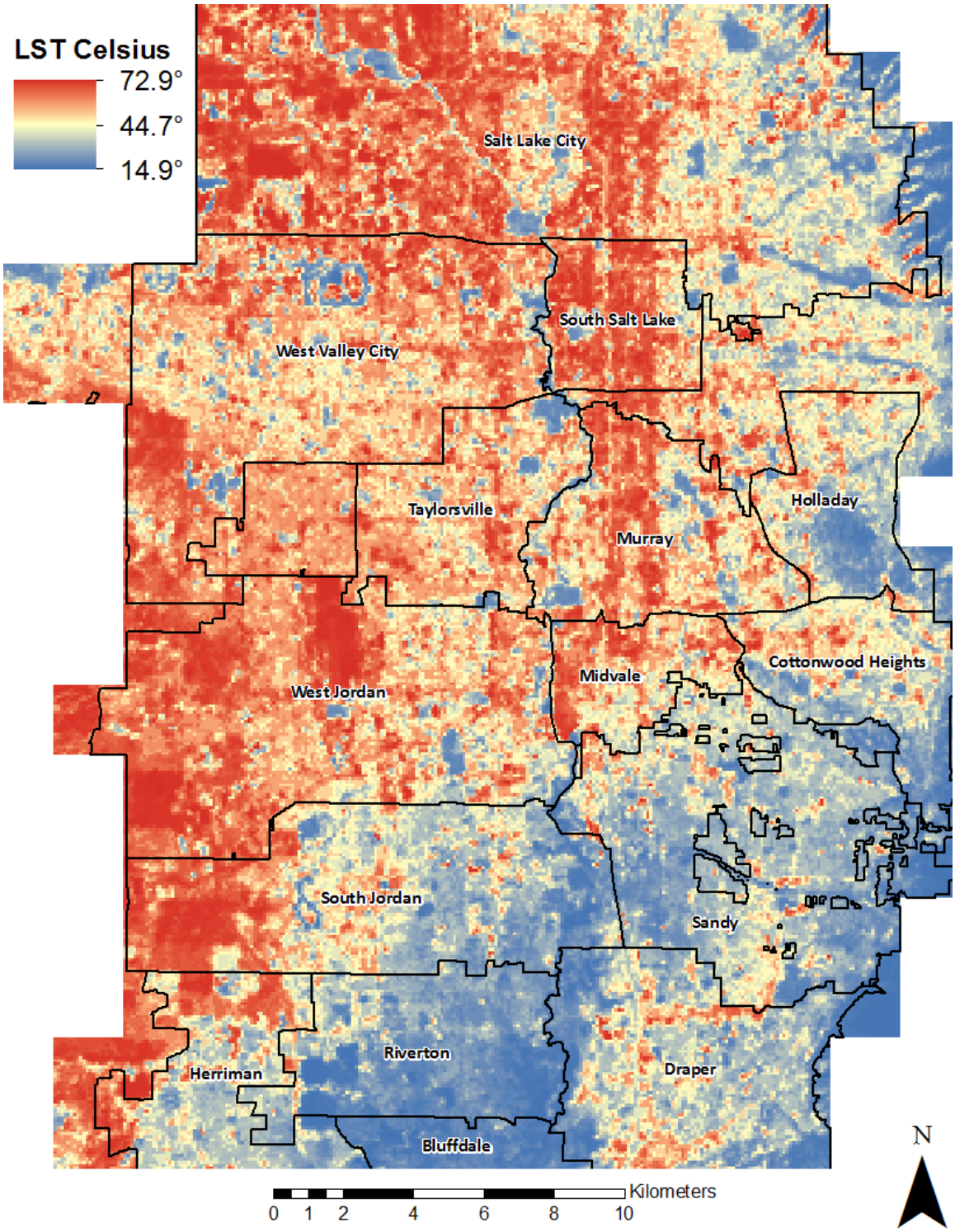

Figure. 2. ASTER Land Surface Temperature

wavelength (NASA LP DAAC, 2001). LST was measured in degrees Kelvin and is accurate within 1.5 degrees (Gillespie, 1998). The data were later converted to degrees Celsius. The ASTER image was captured on June 23rd, 2015. On this particular day, the mean air temperature was 25.5 degrees Celsius (about three degrees higher than average). Records indicated that it had not rained within the last 7 days, thus reducing the effect of evaporation on wet soil. The product clearly shows that the central business district of Salt Lake City, I-15 corridor, and barren areas are warmer than other areas on the map. The suburban areas south of downtown were slightly cooler. Areas close to the Jordan River and in the canyons of the Wasatch Mountains are much cooler. The advantage of using ASTER over similar products is that LST is retrieved from multiple thermal bands allowing for the separation of LST from emissivity (Gillespie, 1998).

Training and Testing Data

Training data for the classification were identified and collected with the assistance of both the NAIP and lidar datasets using both the ArcGIS and ENVI software. Python scripting was used to process the resulting training datasets. Seven land cover classes were selected for the study: Coniferous, Deciduous, Irrigated Low-Stature Vegetation (ILV), Non-Irrigated Low-Stature Vegetation/Soil (NLV/SOIL), Water, Impervious, and Building. They are described in further detail in Table 1.

Ten randomly selected tiles, one to two from each arbitrarily delineated (e.g. commercial, industrial, old residential, and new residential) land use type in the Salt Lake Valley, were selected from the study area for training data collection. This stratification was performed to capture any microscale variation in LST relationships. From each tile, 10

Table 1. Descriptions of each land cover class of the study

Class:	Description:
Coniferous	Includes canopies of trees with dense, evergreen foliage (typically height >1 meter).
Deciduous	Includes canopies of trees that drop their leaves in the winter (typically height > 1 meter).
Irrigated Low-Stature Vegetation (ILV)	Includes any irrigated form of low-stature vegetation such as lawns, parks, golf courses, and green agricultural crops (typically height < 1 meter).
Non-Irrigated Low-Stature Vegetation/Soil (NLV/SOIL)	Includes nonirrigated forms of low stature vegetation such as dry grasses and also includes bare soil. NLV and soil were combined into one class because it was difficult to distinguish between the two using the data available.
Impervious	Includes sidewalks, roads, and freeways
Building	Includes the rooftops of various structures including homes, buildings, and storage units (e.g., water and gas)
Water	Water: Includes pools, ponds, streams, rivers, and lakes.

instances of each land cover type were digitized in ArcGIS. Coniferous and deciduous trees were pinpointed by loading the lidar derivatives as multiband images in ENVI. Conifer trees are usually more difficult to penetrate with lidar, and thus last returns tend to be at a higher height. As a result, coniferous canopies tended to have lower values in the FLD layer and higher FBD values compared to other nontree surfaces. Last returns in deciduous trees were usually at a lower height due to greater canopy penetration. Thus, deciduous canopies tended to have higher values on the LBD and higher FBD values compared to other nontree surfaces. Features that were both in shadow and not in shadow were collected in order to make the model more robust.

In order to prepare the training polygons for use in the random forest, the NAIP imagery, lidar data, and additional indices for the particular site were stacked into a single composite image. Python scripting was used to organize the polygon data and extract the band values to each pixel within them. These data were imported into RStudio as a data frame for use in the random forest model. Using the same data from the training dataset to test the classification often leads to overestimates of classification accuracy. To account for this, an independent testing dataset was collected using 10 new tiles (Congalton, 1991). This testing dataset was used to generate a more unbiased accuracy estimate in the RStudio.

Classification Using Random Forest

Random forest is an ensemble learning technique in which a large number of decision trees are generated for regression or classification purposes. When using random forest for land cover classification, the algorithm uses a random subset (70%) of the

training data to create each tree in the model and uses the rest (30%) to test the model (Pal, 2005). Random selections of each predictor variable are used to split the nodes of each of the trees. Finally, the trees classify each pixel in the study area by voting for the most popular land cover class (Breiman, 2001).

There are several advantages to using random forest over other statistical classification algorithms. Random forest is a nonparametric statistical technique, therefore the data in the model do not have to be assumed normal. Also, the algorithm can handle a large number of predictor variables. Finally, the classifier is highly resistant to overfitting due to each tree in the often very large forest being randomly generated (Gislason et al., 2006; Pal, 2005). Random forest was an excellent selection as a classifier for this study because there were many candidate predictor variables being considered. The variables used in the study are outlined in Table 2. Iterative backwards feature elimination technique was performed to achieve the optimal selection of predictor variables. Training and testing of the random forest model took place in RStudio.

Out-of-bag (OOB) error was used as an initial test of the random forest model's accuracy. OOB error is a generalization of error of the whole model obtained by averaging error rates of out-of-bag samples permuted through the trees. The closer the value is to zero, the more accurate the model. Typically, OOB error is stated to be an unbiased estimate of true prediction error though some studies have demonstrated that this is not always true (Mitchell, 2011).

Variable importance is also used to assess a random forest model and is almost always reported when using this model. This statistic makes the model less of a black box by providing details about the contributions of each predictor variable to the model. It is

Table 2. Predictor Variables

Predictor:	Source:
Blue Band	NAIP
Green Band	NAIP
Red Band	NAIP
NIR Band	NAIP
NDVI	NAIP
GRVI	NAIP
FBD	Lidar
LBD	Lidar
FLD	Lidar
Blue (Focal Std)	NAIP
Green (Focal Std)	NAIP
Red (Focal Std)	NAIP
NIR (Focal Std)	NAIP
NDVI (Focal Std)	NAIP
FBD (Focal Std)	Lidar
LBD (Focal Std)	Lidar
FLD (Focal Std)	Lidar
Blue (Focal Range)	NAIP
Green (Focal Range)	NAIP
Red (Focal Range)	NAIP
NIR (Focal Range)	NAIP
NDVI (Focal Range)	NAIP
FBD (Focal Range)	Lidar
LBD (Focal Range)	Lidar
FLD (Focal Range)	Lidar

determined by fitting the random forest to the data, recording the OOB error for each data point, then averaging out that error for the entire forest. The importance of each variable is measured by permuting the values of that variable among the training data, then calculating the OOB error a second time. Finally, the importance score is calculated by averaging the difference in OOB error before and after the permutation over all of the trees and normalized by the standard deviation of the differences. Variables with higher scores are considered more important than variables with lower scores. This can be performed for the entire model or for individual land cover classes. This statistic helped decide which predictors to include in the final analysis, and revealed how different variables contribute to the discrimination of classes.

Confusion matrices were also used to assess model accuracy. They are the most common way to represent classification accuracy of remotely sensed data. They are defined as a square array of numbers set out in rows and columns that express the number of sample units assigned to a particular category relative to the actual category as verified on the ground (Congalton, 1991). Overall accuracy is obtained by dividing the total correct (diagonals) by the total number of pixels in the matrix. Producer's accuracy represents how well reference pixels of the ground cover type are classified. It is obtained by dividing the number of correctly classified pixels of each class by its respective column total. User's accuracy represents the probability that a pixel classified into a given category actually represents that category on the ground. It is obtained by dividing the number of correctly classified pixels of each class by its respective row total (Congalton, 1991). When classifying with a random forest, a confusion matrix is automatically generated using the data from the training sample. To avoid potential bias,

a confusion matrix was created using the independent testing dataset. This table was used to identify the least accurately classified land cover classes. Additional training samples were collected for the least accurate classes to improve model training.

Statistical Analyses

Vegetation composition variables obtained from the random forest classification result were compared to LST. Because the resolution of the ASTER LST data (90 meter) was larger than that of the land cover (1 meter), the land cover data were overlaid with the LST data. Percentages of each land cover type were calculated within each ASTER pixel, resulting in the following urban vegetation variables: percent coniferous, deciduous, ILV, NLV/SOIL, all trees, and all urban vegetation. In addition, mean tree height for each ASTER pixel was calculated by isolating pixels classified as deciduous and coniferous using the LBD surface obtained from the lidar and compared to LST. Human-made composition variables were also compared to LST. These included percent impervious, building, and all human-made surfaces.

Most common statistical analyses assume that data collected for the study are independent from one another. When working with spatial data, it is important to check for spatial autocorrelation. This common issue occurs when the outcome of two observations is related because of their distance. Because landscape metrics are typically autocorrelated, variograms were generated for each of the composition variables (Du et al., 2016). An example variogram created from the ASTER LST image is shown in Figure 3.

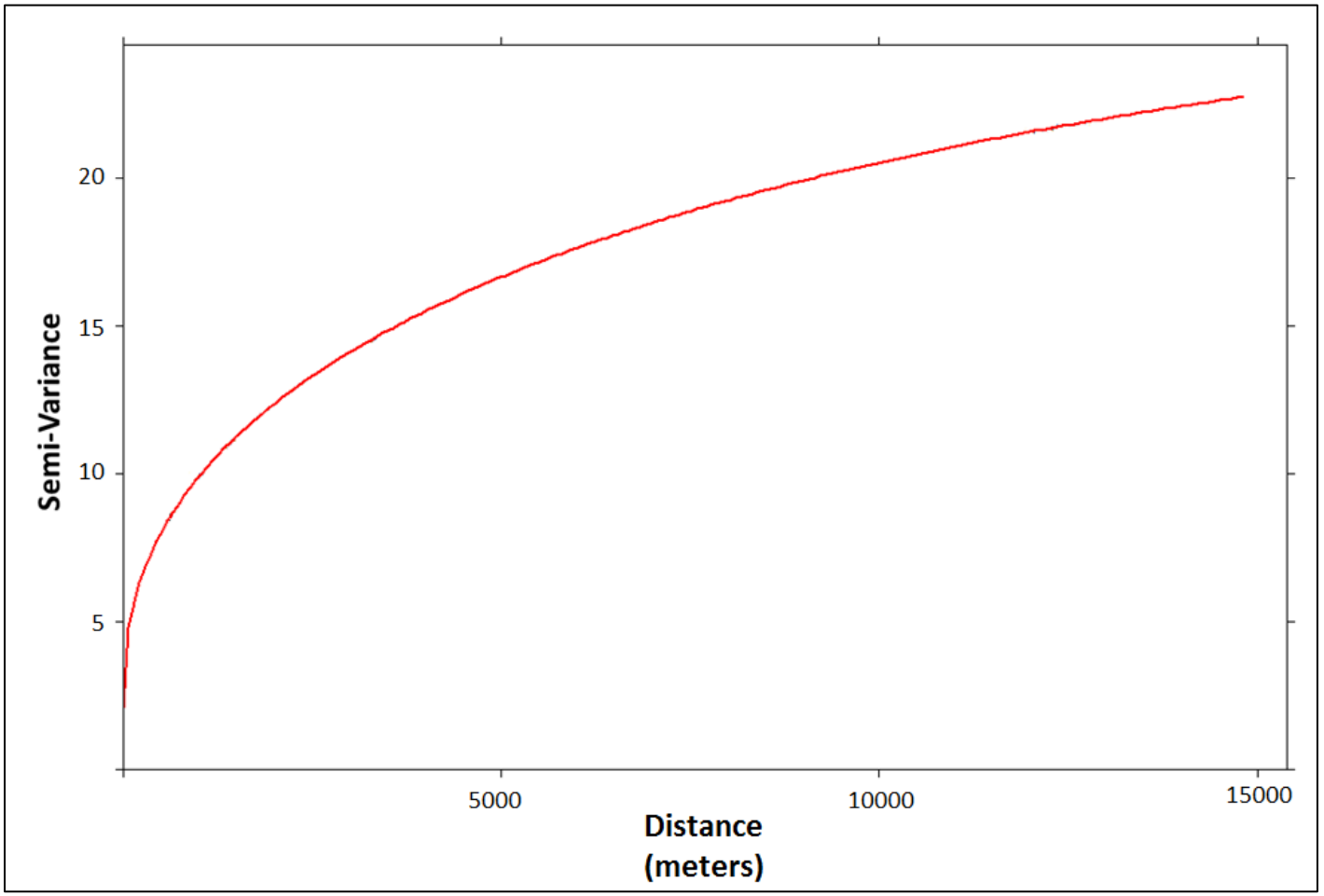

Figure 3. Semivariogram created using the ASTER LST

Variograms show variance in the data as it relates to distance between all pairs of samples in the population. The part of the variogram where the line levels out is known as the range; values in this portion are not spatially autocorrelated. The sill is the value on the y-axis where the range begins. Using the variogram of the ASTER LST image (the dataset used to aggregated the composition variables), a distance to stratify sample points was selected. Because the data were very autocorrelated, the range value was unsuitable by which to stratify sample observations. Using a distance of 15,000 meters would not have yielded many observations in each sample. Therefore, in order to partially address the issue of spatial autocorrelation, a distance of 2,500 meters (the distance where variance begins to slightly flatten) was used to stratify the observations. As a result, 1,000 random samples (each containing about 10% of the total observations) of the final dataset containing observations stratified by 2,500 meters were created. This was done by using the ArcGIS Create Random Points tool recursively in a Python script. An example of this sampling scheme is shown in Figure 4. New variograms were generated for a few of these samples to assess whether the stratification decreased the amount of autocorrelation. The amount of spatial autocorrelation in the new samples decreased to some extent. The statistical analyses that will be described in the coming sections were performed on each of the 1,000 samples. The final result of each analysis represents the median of those 1,000 runs.

Pearson's correlation analysis was performed in order to analyze urban vegetation's composition relationship with LST in the Salt Lake Valley. Correlation analysis is a technique for quantifying the strength of the relationship between two continuous variables. First, scatterplots of each variable were produced to examine if

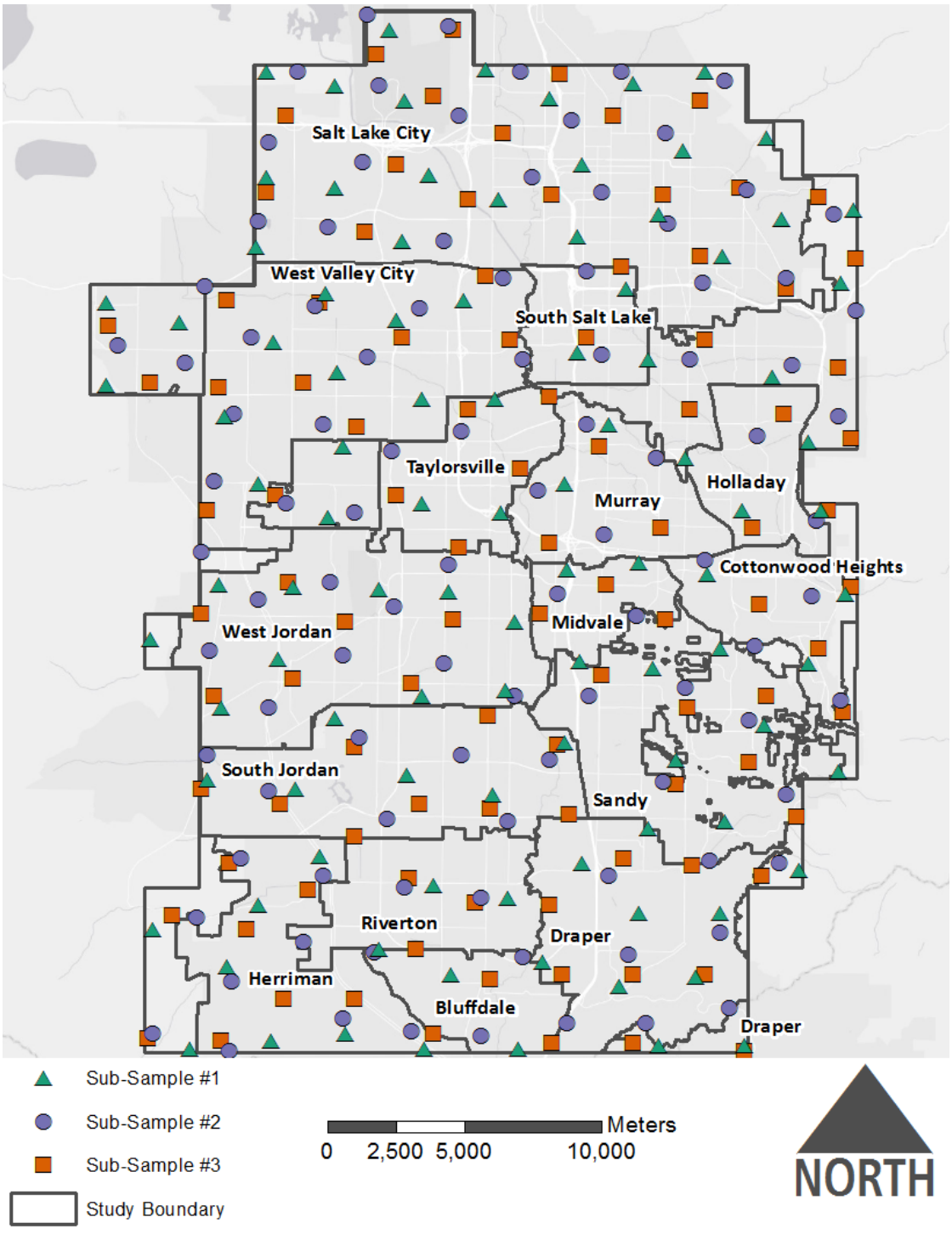

Figure 4. Map depicting the stratified sampling scheme. There are 1,000 subsamples (three shown). Each symbol is from a different random subsample. Each observation within its subsample is at least 2,500 meters apart from all other observations. Each symbol represents the center of one ASTER pixel. All observations are from the same master dataset.

there is a linear relationship with LST. Next, correlation coefficient values or *R* values were produced for variables that exhibited a linear relationship with LST. *R* values range between 1 and -1. If the value is positive, then both variables increase together. If the value is negative, one variable increases as the other variable decreases. Values closer to 1 or -1 indicate a strong relationship and values closer to zero indicate a weak or nonexistent relationship. The significance of each *R* value was tested through the calculation of *p* values. A value above 0.05 indicates the results are not significant, values between 0.01 and 0.05 are significant, and values less than 0.01 are highly significant. Plotting, correlation testing, and significance testing were performed in RStudio.

Before creating a linear model with multiple predictor variables, it is important to ensure the variables are not multicollinear. Multicollinearity is when independent variables are correlated with other independent variables in the model. Calculating variance inflation factors (VIF) allows one to learn how collinear their variables are (O'Brien, 2007). Once calculated, interpretation is simple: the higher the value is, the more collinear the variable is. A stepwise VIF test was used to find any multicollinear variables and remove them if necessary. This version of VIF is more robust because it recursively calculates VIF values for each variable, removes collinear variables if necessary, then recalculates VIF values until the optimal set of variables is achieved. This test was performed in *R*.

In order to further quantify the relationship between the land cover composition variables and LST, a multiple linear regression analysis was performed. This analysis revealed how much variation in the dependent variable could be explained by the independent variables. It also revealed which independent variables (should the others

remain constant) have the largest impact on the dependent variable. One such model was constructed in RStudio using percentages of each land cover class as the independent variables and LST as the dependent variable. The model was analyzed by interpreting its coefficients.

RESULTS

Random Forest Classification

A random forest model was created in RStudio using Breiman and Cutler's *randomForest* package and was used to classify the study area. A total of 172,960 sample points were collected using the sampling technique described in the training data section and were used to train the model. Each point had a digital number at the specific location for each predictor variable for use as the predictor and land cover classification for use as the response. A total of 501 trees were generated when creating the model. This amount was selected because a large number of predictor variables (27) was included in the model. After charting the amount of OOB error for each run, it was clear that error had stabilized. The model took about 10.5 minutes to generate in RStudio.

Approximately 840 square kilometers of the Salt Lake Valley were classified in the analysis. An example of the classified study area is shown in Figure 5. Approximately 29.4% of the study area was classified as NLV/SOIL. Twenty-seven percent was classified as impervious, 20.1% was classified as deciduous, 13.3% was classified as buildings, and 8% was classified as ILV. Only 1.3% was classified as coniferous trees, 0.4% was classified as water, and 0.2% remained unclassified.

The importance of each variable was assessed using mean decrease in model accuracy metric, which is generated by the random forest algorithm. By far the two most important predictor variables were GRVI standard deviation (78.55) and NDVI (77.1).

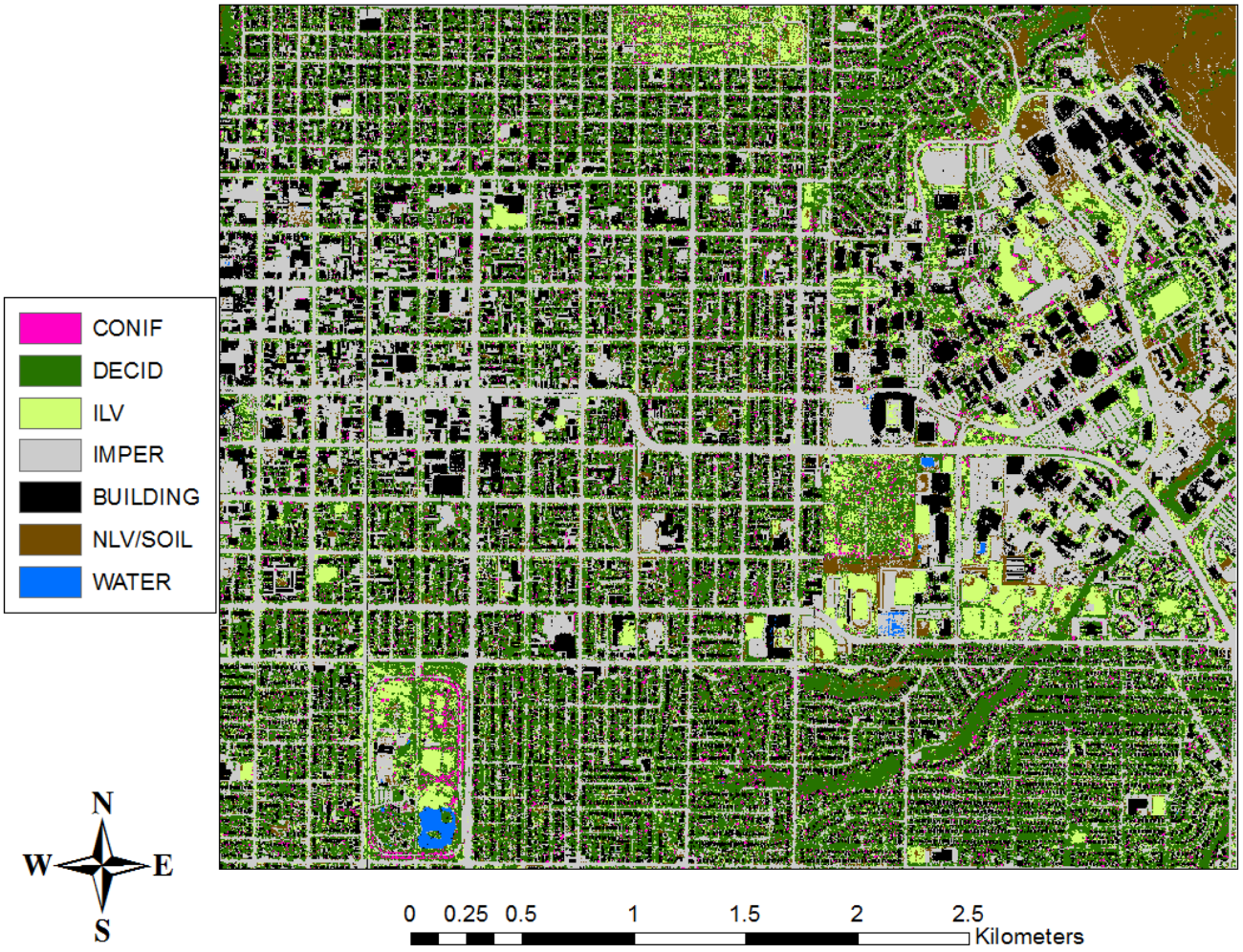

Figure 5. Example of classification result. Downtown SLC is to the west. The University of Utah is to the east. Liberty Park is to the southwest.

GRVI standard deviation was instrumental because it helped distinguish trees (coniferous and deciduous) from all other land cover types. Trees typically had higher values in this dataset than all other land cover types due to variability in index values within tree canopies. NDVI was also a very important predictor variable because it helped distinguish densely vegetated, sparsely vegetated surfaces, and nonvegetated surfaces from each other. The most important lidar variable (FLD) was only the sixth most important variable over all. FLD was important for distinguishing coniferous and deciduous trees, because coniferous trees tended to have higher values in this dataset due to lower penetration of the canopy. The two least important variables were the red standard (19.84) deviation and red range (17.42). Variable importance for each predictor is shown in Figure 6. Multiple tests without the least important variables were performed. However, model accuracy did not see significant improvements. Thus, no variables were removed from the model using iterative backwards feature elimination.

The OOB Error of the model was 1.88. This meant the model had an overall classification accuracy of about 98.2% from tests using the out-of-bag samples of the training data. The resulting kappa value was also 0.98. The internally generated confusion matrix is shown in the Appendix. Because internal accuracy tests are known to have some bias, a more representative accuracy test was performed using the independent testing dataset described earlier. The confusion matrix from the independent test is shown in Table 3. In this test, overall accuracy of the model was 94% with a kappa value of 0.92. The building class had the highest producer's accuracy (98.2%) of all the classes. This was due to buildings having distinct spectral and height characteristics. Rarely, buildings were misclassified as impervious surfaces. However, it was usually easy to

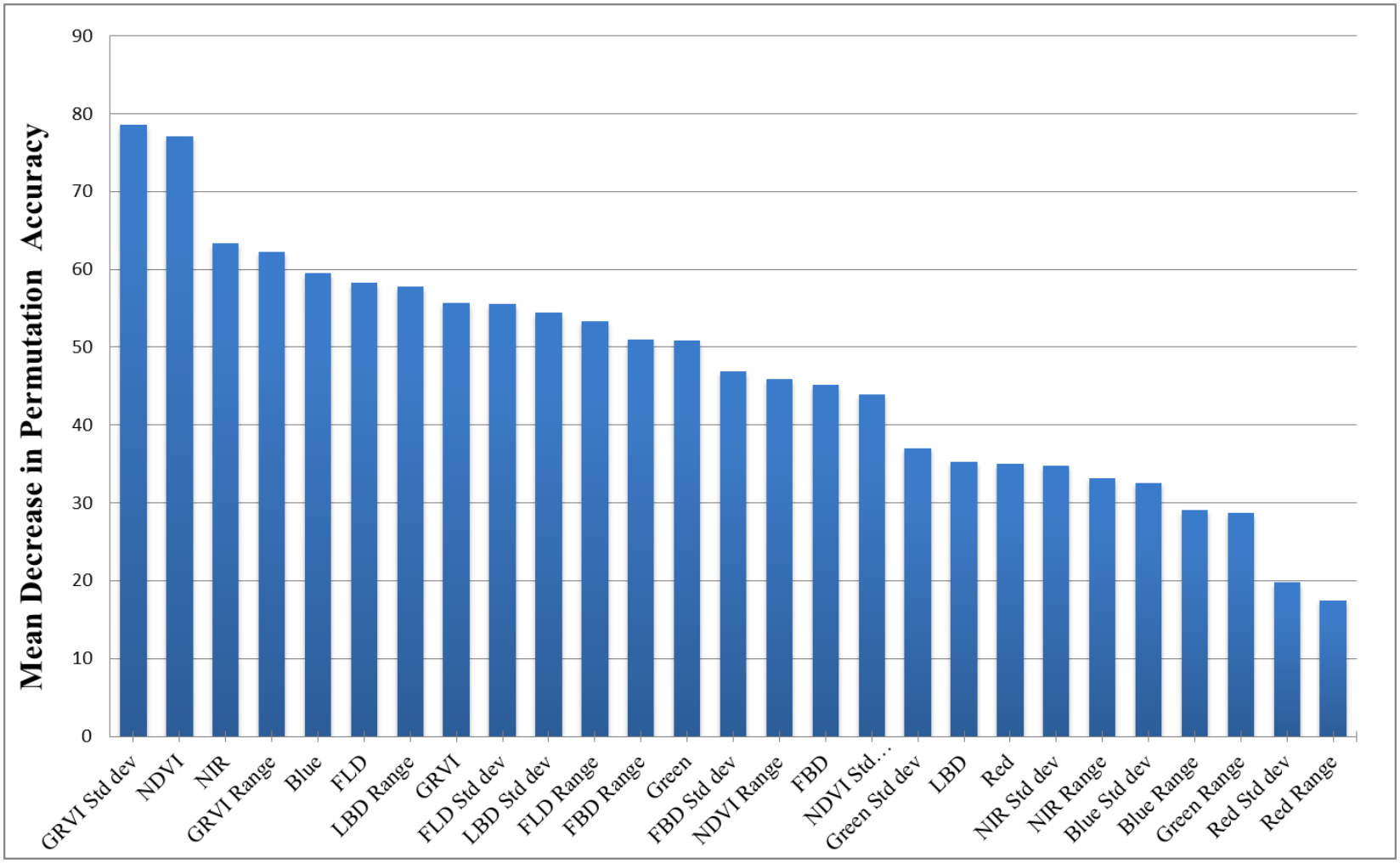

Figure 6. Variable importance of land cover predictors by mean decrease in permutation accuracy.

Table 3: Confusion matrix generated using independent testing dataset.

	Reference Data								
Class	CONIF	DECID	ILV	IMPER	BUILDING	NLV/SOIL	WATER	TOTAL	**UA(%)**
CONIF	5113	43	0	0	22	0	1	5179	98.7
DECID	559	5193	147	10	8	3	0	5920	87.7
ILV	46	211	4972	5	0	10	3	5247	94.8
IMPER	21	32	63	12357	242	285	1116	14116	87.5
BUILDING	102	8	0	308	17955	69	27	18469	97.2
NLV/SOIL	0	24	166	405	44	10260	11	10910	94
WATER	0	0	25	5	7	0	3661	3698	99
TOTAL	5841	5511	5373	13090	18278	10627	4819	63539	-
PA(%)	87.5	94.2	92.5	94.4	98.2	96.5	76	-	-

Overall Accuracy = 94%

Kappa Coefficient = 0.92

discriminate these classes because they had vastly differing heights. NLV/SOIL had the next highest producer's accuracy (96.5%). This class also had distinct spectral and height characteristics. NLV/SOIL typically had low NDVI and GRVI values because pixels of this class did not contain dense, green vegetation. Also, pixels of this class were almost always ground-level. Sometimes NLV/SOIL was mistaken for impervious surfaces, likely because some soils were greyish in color and appeared to resemble concrete.

The two classes with the lowest producer's accuracies were coniferous trees and water. The model had some trouble distinguishing coniferous trees from deciduous trees, resulting in a producer's accuracy for coniferous trees of 87.5%. Pixels where there were dense populations of deciduous trees overlapping each other were sometimes classified as coniferous trees. In addition, the model had trouble classifying the less dense outer edges of the canopy of coniferous trees. An interesting thing to note is that there were far fewer coniferous trees in the study area than most other classes. Because of this, the margin for error in classification was much smaller. Water had by far the lowest producer's accuracy at 76%. This was due to a large amount of water pixels being mistaken for dark, impervious surfaces because they have similar spectral and height characteristics. Water also happened to be the least present class in the reference data.

Water also had the highest user's accuracy (99%) of all the classes despite having the lowest producer's accuracy. This is likely due to it being the least common land cover class that was predicted by the classification. It is also possible that the testing samples for the water class mostly consisted of easy-to-classify examples of water and hence, the model had no difficulties. Coniferous trees had the second highest user's accuracy (98.7%), despite having the second lowest producer's accuracy. Like water, the testing

samples selected for coniferous trees likely consisted of clear easy-to-classify (e.g., coniferous trees isolated in open areas) examples of coniferous trees.

The two classes with the smallest user's accuracy were deciduous trees and impervious surfaces. Both of these classes were similar because the model tended to overclassify them. For example, deciduous trees were occasionally predicted where coniferous trees, the class with the most similar spectral and height characteristics, were supposed to be. Impervious surfaces were occasionally predicted where NLV/SOIL and buildings were supposed to be. Sometimes low-lying buildings were mistaken for impervious surfaces. Additionally, grey soils were also sometimes mistaken for concrete.

Pearson's Correlation Analysis

The percentages of each land cover type within each ASTER pixel were obtained and recorded in a table along with the pixel's LST value. Maps depicting percentage of urban vegetation and human-made land cover are shown in Figures 7 and 8. In both maps, it is apparent that there are a large percentage of human-made surfaces near downtown SLC and western SLC and the I-15 corridor. Most of the residential areas are composed of a mix of both urban vegetation and human-made surfaces. The areas on the western edge of the valley are mostly composed of NLV/SOIL. The areas on the eastern edge as you approach the Wasatch Mountains become more and more vegetated. Any ASTER pixel that contained more than 2.5% "no data" values was excluded from the analysis. In total, only 0.4% of the data was removed. Using the percentages of each land cover type as the independent variable and LST as the dependent variable,

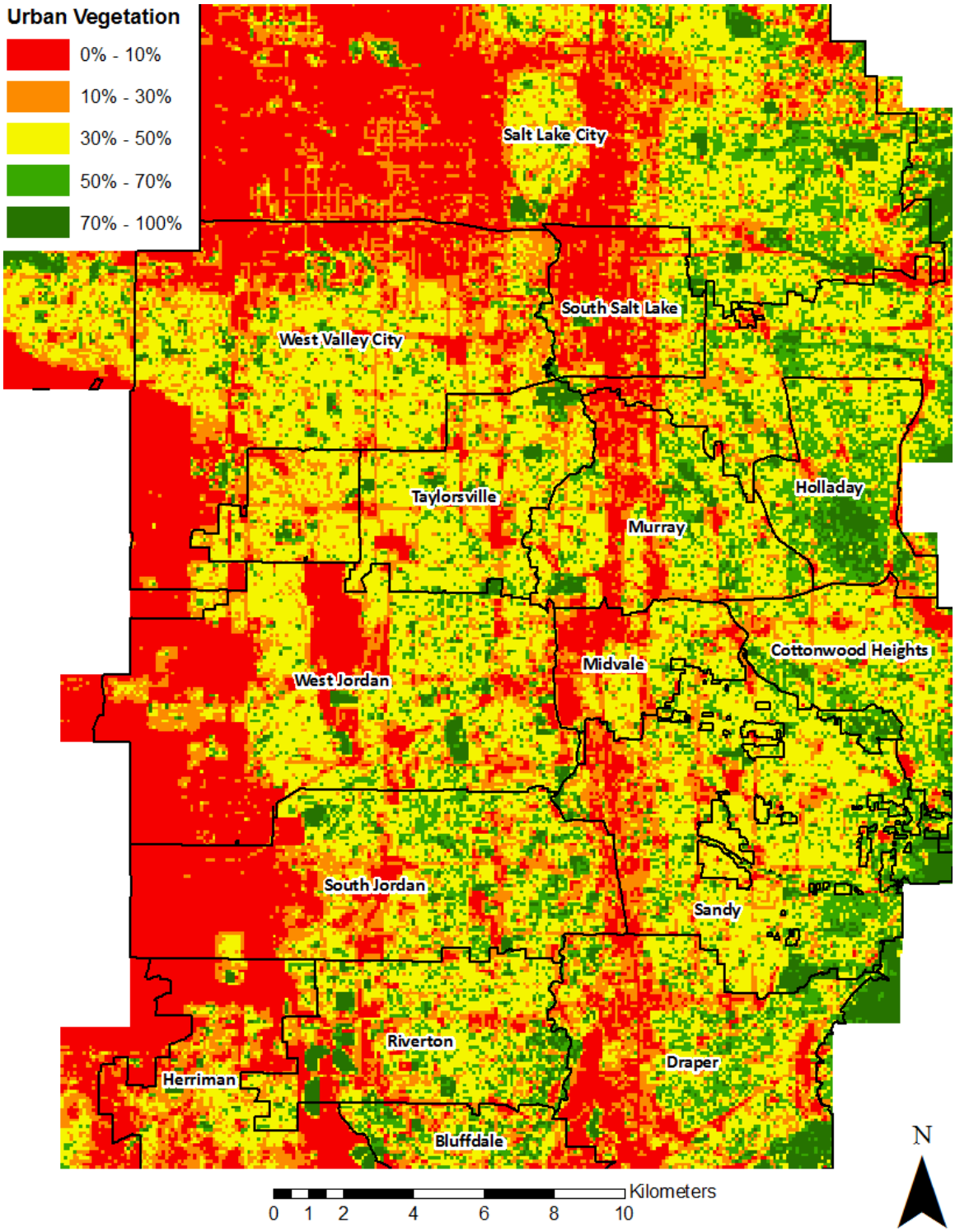

Figure 7. Percentage urban vegetation (includes Deciduous, ILV, and Coniferous) per ASTER pixel in the study area.

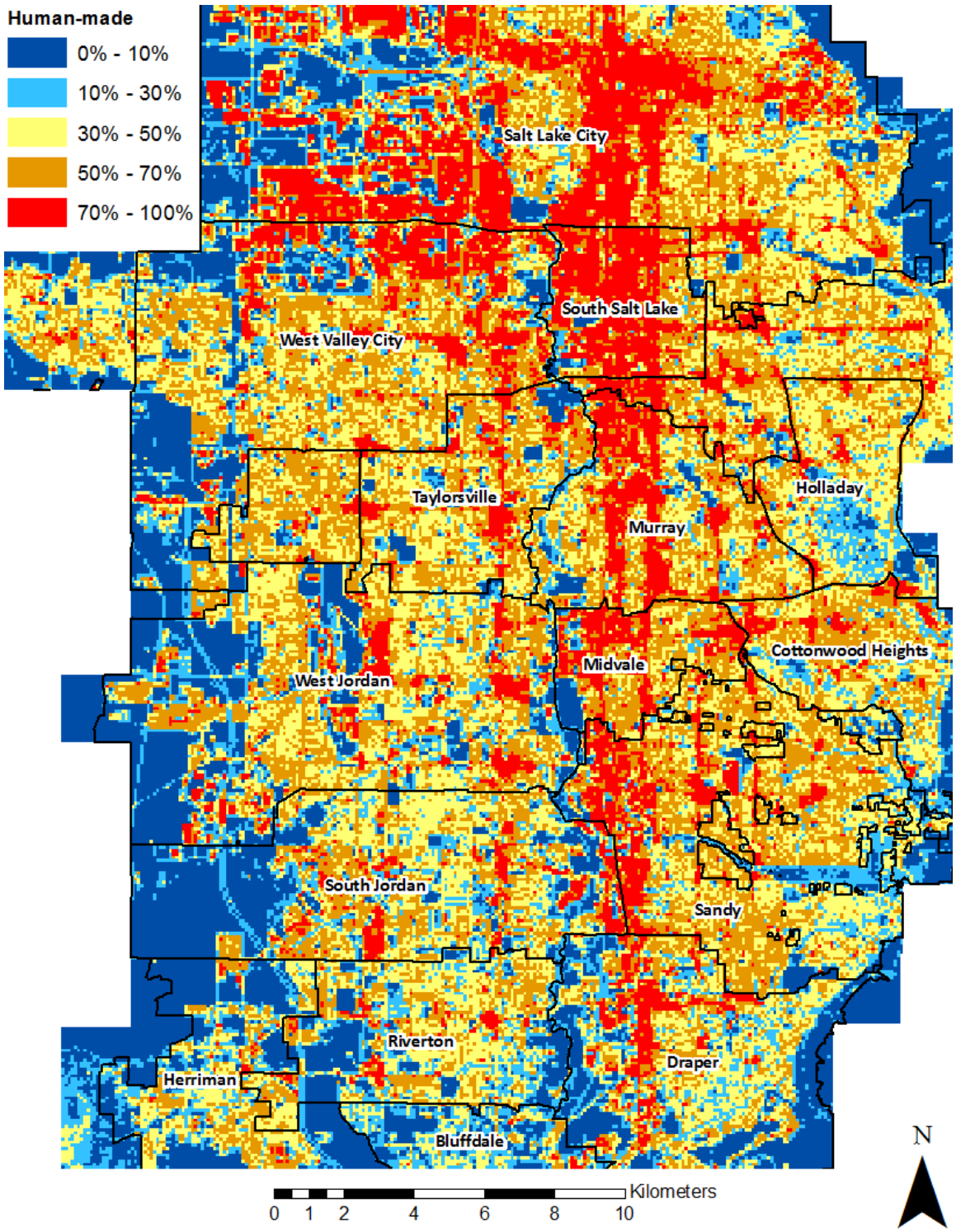

Figure 8. Percentage human-made (includes Building and Impervious) per ASTER pixel in the study area.

scatterplots were generated and Pearson's correlation analysis was performed. The scatterplots are shown in Figure 9. In the plots for percent deciduous, ILV, tree, and urban vegetation, there are clear negative trends in LST. There are clear, slight positive trends in LST for the following variables: percent NLV/SOIL, impervious, building, and human-made. The *p* values for all correlation coefficients were all less than 0.01, indicating that the results were highly significant.

Percent deciduous had the strongest negative correlation with LST of the individual percentage variables, with a coefficient of -0.55. Percent ILV had the next strongest negative correlation with LST with a coefficient of -0.33. Percent coniferous had a weak correlation with LST with a coefficient of -0.16. Mean tree height did not have a correlation with LST. It had a coefficient of -0.02. The variable that had the strongest positive correlation with LST was percent impervious with a coefficient of 0.26. Percent NLV/SOIL had the next strongest positive correlation with a coefficient of 0.24. Finally, percentage building had a weak positive correlation with LST with a coefficient of 0.16. Table 4 shows the correlation coefficient of each land cover class when compared with LST.

Several of the variables that logically fit together were combined into three new independent variables (percent urban vegetation, tree, and human-made). Percent urban vegetation (includes percent deciduous, coniferous, and ILV) had a strong negative correlation with LST with a coefficient of -0.63. Percent tree (includes percent deciduous and coniferous) had a moderate negative relationship with LST with a coefficient of -0.54. It is worth mentioning that deciduous trees comprise about 94% of tree cover in the study area.

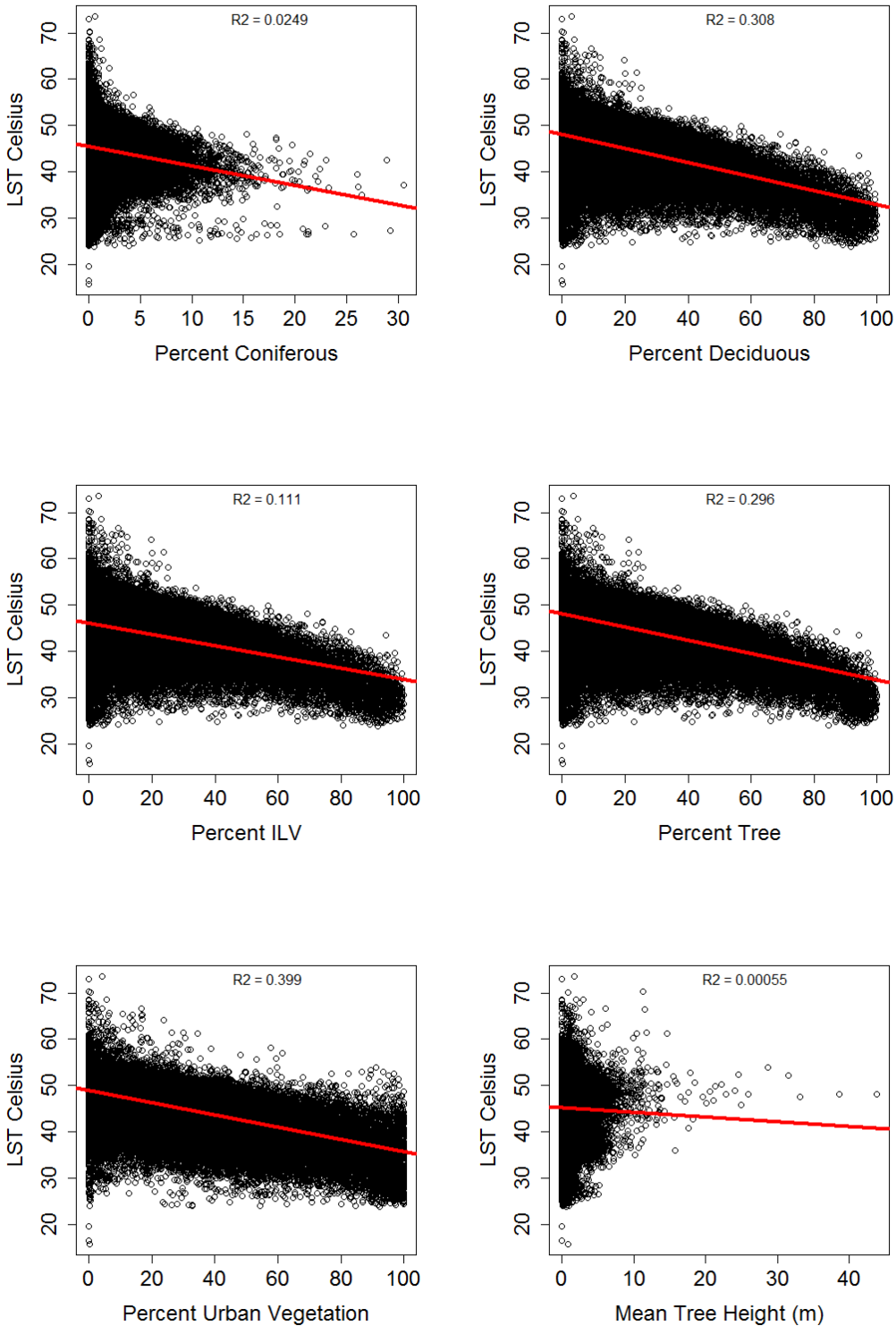

Figure 9. Scatterplots of each variable used to predict LST.

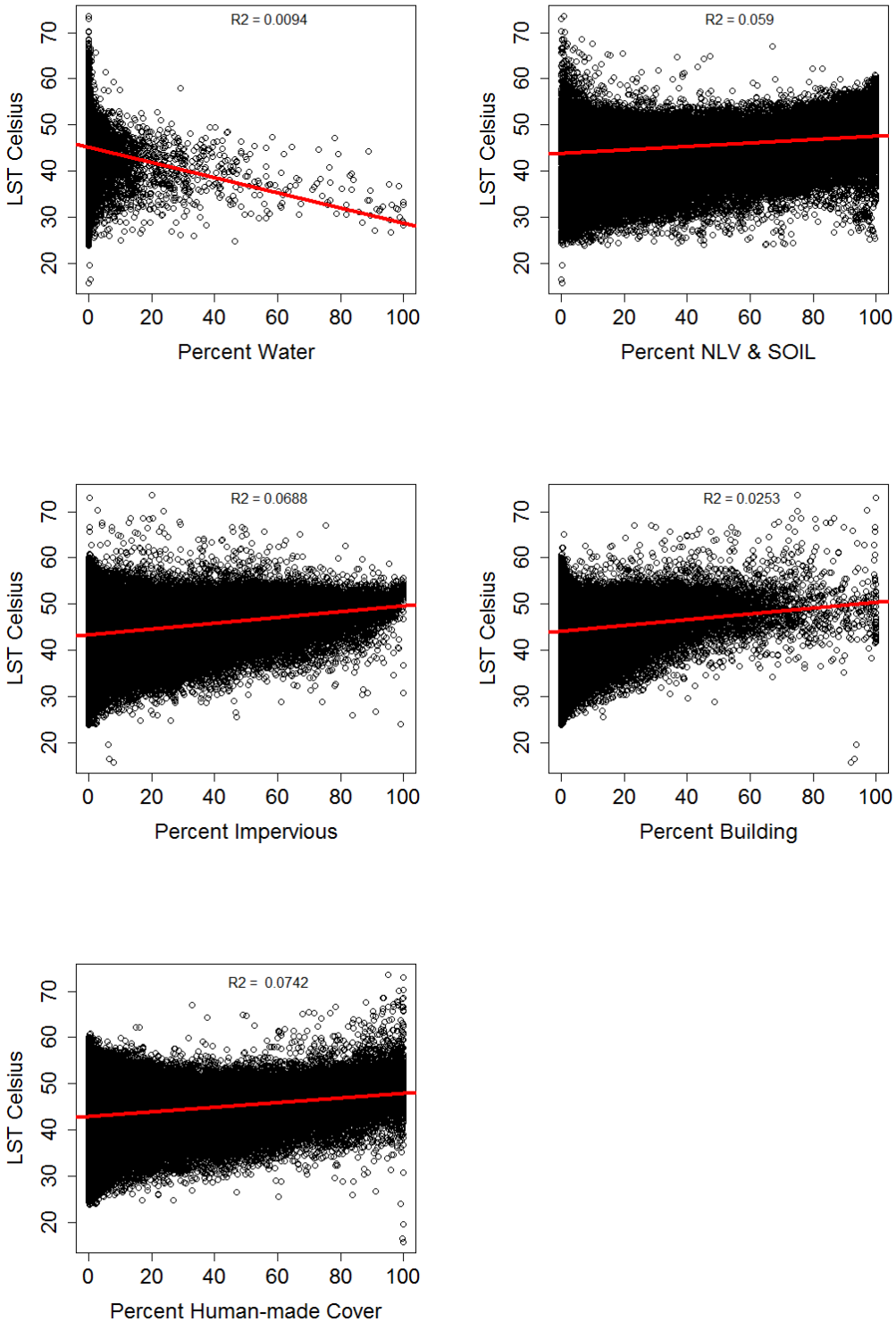

Figure 9. Continued.

Table 4. Pearson's correlation analysis results

Variable	Coefficient
% All Urban Vegetation	-0.63
% Deciduous	-0.55
% Tree	-0.54
% ILV	-0.33
% Coniferous	-0.16
% Water	-0.09
Mean Tree Height	-0.02
% Building	0.16
% NLV/Soil	0.24
% Impervious	0.26
% Human-Made	0.27

Percent human-made (includes percentage impervious and building) had a weak positive correlation with LST with a coefficient of 0.27.

Multiple Linear Regression

The stepwise VIF test revealed that percentage NLV/SOIL was a redundant or multicollinear variable, thus it was excluded from the model. The median residual value (0.46) was close to 0 and the first and third quartile were close in value (-1.9 and 2.4, respectively), indicating the data were normally distributed. The R^2 of the model was 0.56, indicating that about 56% of the variation in LST can be explained by the predictor variables listed above. The intercept of the model was 48.397. According to the model, if the percentage of all land cover classes was equal in the study area, the temperature would be 48.4 degrees Celsius. Only two of the coefficients for the predictor variables were significant. Percentage deciduous had the strongest negative effect with a coefficient of -0.180 followed by percentage ILV with a coefficient of -0.111. From the model, we can infer that; for every 1% increase in deciduous tree land cover, LST decreases by 0.18 degrees Celsius if the other variables remain constant. Also, for every 1% increase in ILV land cover, LST decreases by 0.11 degrees Celsius if the other variables remain constant. The *p* value for the *F* test, 1.77 E-15, indicates that the model containing the above predictor variables is more useful in predicting LST than a model that does not take them into account. Coefficients of the multiple linear regression model are shown in Table 5.

Table 5. Coefficients of the multiple linear regression

Intercept	% Coniferous	% Deciduous	% ILV	% Imper	% Building	% Water
48.397**	0.242	-0.180**	-0.111**	0.002	0.060	-0.200

$R^2 = 0.56$
Adj. $R^2 = 0.53$

* Coefficient is significant at the 0.05 level (two-tailed).
** Coefficient is significant at the 0.01 level (two-tailed).

DISCUSSION

Random Forest Classification Effectiveness

The research aimed to uncover the effectiveness of random forest for classifying an urban area using orthoimagery and lidar data. The results of this study agreed with other studies that random forest is an accurate classifier for urban environments (Chehata et al., 2009; Guan et al., 2012; Guo et al. 2011). Overall, random forest effectively and efficiently classified the Salt Lake Valley at a 1-meter spatial scale using the datasets available. The classification worked extremely well in the valley area. Individual features on the map were easily distinguishable. It was not as useful in the mountainous regions in the outskirts of the study area, because the dark canyons of the mountain were mistaken for water. This was acceptable because these areas were not the focus of the study. These regions were not common in the dataset so accuracy did not suffer much because of them.

In this study, 27 total predictor variables were used to train the random forest model. Typically when using random forest, one performs iterative backward feature selection to achieve the optimal set of predictors. A threshold at which variables should be removed from the model does not exist. In the literature, Guan et al. (2012) removed all variables with less than 20% mean decrease in accuracy. In this study, it was not necessary to remove any variables. The best method for determining which predictor variables to remove is to recreate the model minus those variables with low importance. If accuracy improves, keep them out of the final model build. Because subsequent runs of

the model did not yield substantially better results, none of the variables were removed.

The model had an OOB error of 1.88 and an overall classification accuracy of 94 percent, which are impressive considering the scale and study area size. All of the land cover types of the study were classified well, some better than others. Random forest distinguished urban vegetation types very well. Errors of omission for both deciduous trees (0.058) and ILV (0.075) were very low. It was unexpected that GRVI standard deviation was the most important variable. The most likely reason for this was because this variable allowed the model to distinguish both tree classes (coniferous and deciduous) from all other land cover classes, most importantly ILV. Analysis of the GRVI standard deviation raster revealed that both tree types had high amounts of variability due to the effects of intercanopy shadowing. Other land cover features like ILV, impervious surfaces, and buildings typically had very low variability. Normally, the lidar rasters were relied on to make the distinction between ILV and the tree types. This layer made it possible to distinguish these land cover classes using a derivative of the orthoimagery. As expected, NDVI was a very important variable because it allowed for the model to distinguish between densely vegetated, sparsely vegetated, and nonvegetated surfaces. The lidar was then used to help discriminate between what was on and off the ground. The high importance of GRVI and NDVI suggests that these variables do complement one another and should be used in tandem as predictor variables.

Even with having lidar available, ILV was most often misclassified as deciduous trees. The reason for this was that the model seemed to classify pixels of ILV that were in shadows as deciduous trees. ILV in shadow appears similar spectrally to the darker canopy of deciduous trees. Also, deciduous trees were commonly misclassified as

coniferous trees. This is likely due to instances where tall deciduous tree canopies would overlap with shorter ones within a pixel. This resulted in last return values being higher, thus making the difference between first return and last return smaller. This is important because FLD values for coniferous trees were typically low. Coniferous trees were the least accurately predicted vegetation type (0.875) although error was still very low. The higher than average omission error is evident when the trees are viewed on the map. The model tended to locate and classify the inner canopy of coniferous trees very well but often classified the outer edges of their canopies as deciduous. This is likely due to there being fewer returns on the edge of canopies in the lidar data. NDVI, FBD range, and LBD were all instrumental in classifying this land cover class. NLV/SOIL was classified very accurately. NDVI and GRVI had very high importance ratings for this class and it allowed for this type to easily be distinguished from other ground-level types such as ILV and impervious surfaces.

Probably the greatest flaw of this classification was the fact that it likely overclassified deciduous trees in places where there are shadows. Initially, there were concerns on tuning the model to classify areas inside shadowy areas. Collecting additional training polygons of each land cover class in these areas improved the model's ability to classify those areas for certain land cover classes. After tuning the training data, impervious and NLV/SOIL pixels that were inside shadow were able to be more accurately classified. As mentioned earlier in the paragraph, ILV pixels that were in shadow were often misclassified as deciduous. Deciduous trees having the second lowest user's error supports that this class was overclassified.

Human-made land cover types were classified even more accurately. Buildings

were classified with extremely high producer's (0.982) and user's accuracy (0.972). The lidar surface models were instrumental in distinguishing this class with the impervious surface class. Impervious surfaces had a high producer's accuracy (0.944) but a slightly lower user's accuracy (0.875). Impervious surfaces were occasionally classified where water should have been. This was typical of shallow, dark bodies of water. Impervious surfaces were also misclassified as NLV/SOIL due to some soil types having a greyish coloring. They were also misclassified as buildings due to the model having issues with overpasses. Overpasses have similar spectral characteristics of impervious surfaces yet they are elevated from the ground, therefore the model classifies them as building. Attempts to train the model to recognize these overpasses did not yield the desired results.

Random forest was a highly suitable algorithm for classifying the Salt Lake Valley urban area. Overall accuracy was very high and all land cover types were classified very well. Incorporating a large amount of variables into the model was simple and the ability to interpret their importance was very beneficial to the analysis. Studies that did not have the near-infrared band available seemed to have trouble classifying certain features (Chehata et al., 2009; Guan et al., 2012; Guo et al. 2011). The addition of lidar and the near infrared band from the orthoimagery greatly facilitated the model's ability to have more levels of classification of this urban landscape. Lidar helped distinguish objects with differing vertical profiles and the near-infrared band helped distinguish vegetation from nonvegetation. Using the two together, it was possible to easily distinguish the vegetation classes and human-made classes. It was also possible to distinguish coniferous trees from deciduous trees. As evidenced by the fact that the model

was able to accurately predict to tiles not included in the training dataset, random forest did not overfit the data. Given the size and scale of the datasets involved, the model generated fairly quickly. It would be intriguing to compare the ability of random forest to classify this study area to that of other classification algorithms.

Urban Vegetation's Relationship With LST

This research also aimed to reveal relationships between urban vegetation types and LST in the Salt Lake Valley. Percentages of each land cover type were compared with LST of each ASTER pixel in the study area. The results of Pearson's Correlation Analysis heavily supported prior findings about urban vegetation's relationship with LST (Chen et al., 2006; Gluch et al., 2006; Weng et al., 2004;Yuan & Bauer, 2007; Zhou et al., 2011). As expected, percent deciduous had a moderate negative correlation (-0.55) with LST. Percent ILV also had a significant negative correlation (-0.33) with LST. This indicates that increasing the amount of deciduous trees and ILV like lawns, parks, and green agriculture in an area was associated with a cooling effect on the surrounding environment. Unexpectedly, percent coniferous had a weak negative correlation (-0.16) with LST. Studies of transpiration rates between the two tree types suggest that conifers and deciduous trees transpire at similar rates during the summer (Smith & Hinckley, 1995). The results of this study did not support that. Poor correlations between the abundance of the coniferous tree class and LST were likely due to there being only a small area of coniferous trees (1.3% of the study area) compared to deciduous trees, which were approximately one-fifth of the study area.

Percent tree (coniferous and deciduous) had a comparable correlation coefficient

with LST (-0.54) to other studies. For example, Zhou et al. (2011) reported a similar value for a correlation coefficient (-0.47) between percent woody vegetation (includes trees and shrubs) and LST. Weng et al. (2004) also reported strong negative correlations across multiple scales for their percent tree variable equivalent. Finally, percent urban vegetation had a strong negative correlation (-0.63) with LST. This directly supports that idea that transpiring vegetation such as trees and lawns helps cool the environment in the Salt Lake Valley (Quattrochi & Ridd, 1998). Mean tree height did not have a significant correlation (-0.02) with LST. The results of this study seem to indicate that an urban forest's ability to cool the environment hinges more on tree cover than the average height of the trees within it.

The results of the multiple linear regression further supported that the Salt Lake Valley's urban vegetation helps cool the environment during the summer. Although the only two coefficients that were significant were percent deciduous and percent ILV, evidence that they impacted LST was strong. The coefficients for the two variables were -0.18 and -0.11, respectively. Assuming that the composition of the other variables remains the same, LST would decrease by approximately 1 degree Celsius for every 5% increase of deciduous trees land cover type. LST would decrease about half of that amount for every 5% increase in ILV. Zhou et al. (2011) also created a multiple linear regression model using composition variables. The results of their model also suggested that percent woody vegetation and percent fine vegetation played a large role in cooling the environment.

Results of this study also seem to support that urban vegetation types have a greater cooling effect on the environment than nonirrigated vegetation. NLV/SOIL had a

weak positive correlation (0.24) with LST. Previous research in the study area has indicated that dry grasses and soils (the most common land cover type classified in the Salt Lake Valley) are associated with relatively warmer temperatures (Quattrochi & Ridd, 1998). Quattrochi and Ridd (1998) attribute this to underlying soils being so dry that there is little cooling due to evaporation or transpiration. In order to make any real conclusions on how percentage NLV affects temperature, NLV and soil would need to be classified separately to isolate each individual land cover class' relationship with LST.

It was unexpected that percent impervious surface and building had a weak positive correlation (0.24 and 0.16, respectively) with LST. This contrasts with the results of Zhou et al. (2011) who found that similar classes, percent pave and building, had correlation coefficients of 0.55 and 0.57, respectively. However, Zhou et al. (2011) studied LST in Baltimore, where urban forests are much closer to the native vegetation cover and impervious surface and buildings represent greatly reduced evapotranspiration. In Salt Lake City, impervious surface and buildings represent a smaller departure from the native vegetation, so percent impervious surface and building cover may be less strongly correlated with LST.

Use of Multiple Linear Regression Model

Assessment of the coefficients of the multiple linear regression model reveals that this model was an adequate method for predicting LST, with room for improvement. Figure 10 shows the spatial distribution of the residuals of the multiple linear regression when predicting LST for the entire study area. Although the spread of the residuals appears to be high in the legend, Figure 11 shows the actual distribution of the residuals

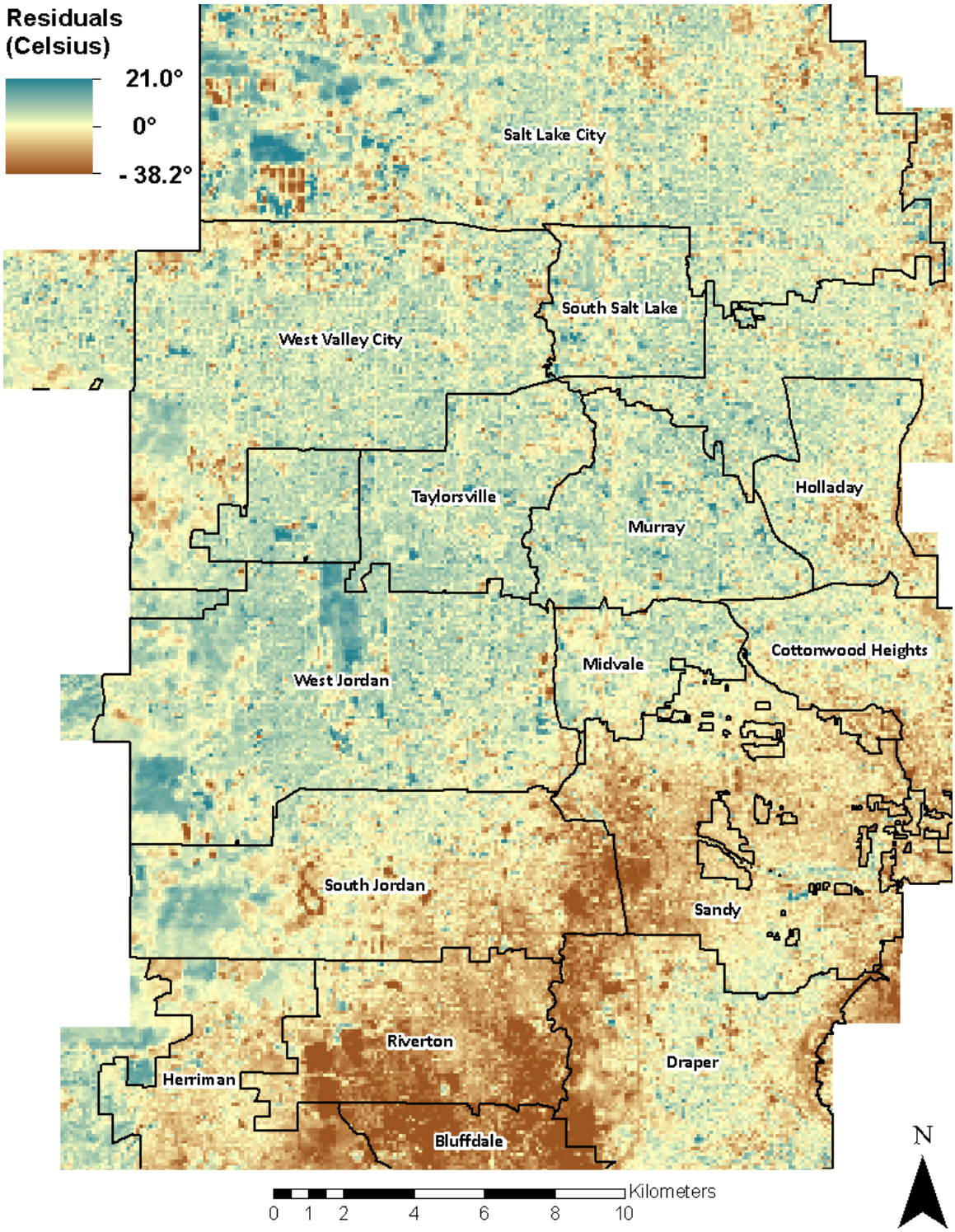

Figure 10. Mapped residuals of the multiple linear regression LST prediction.

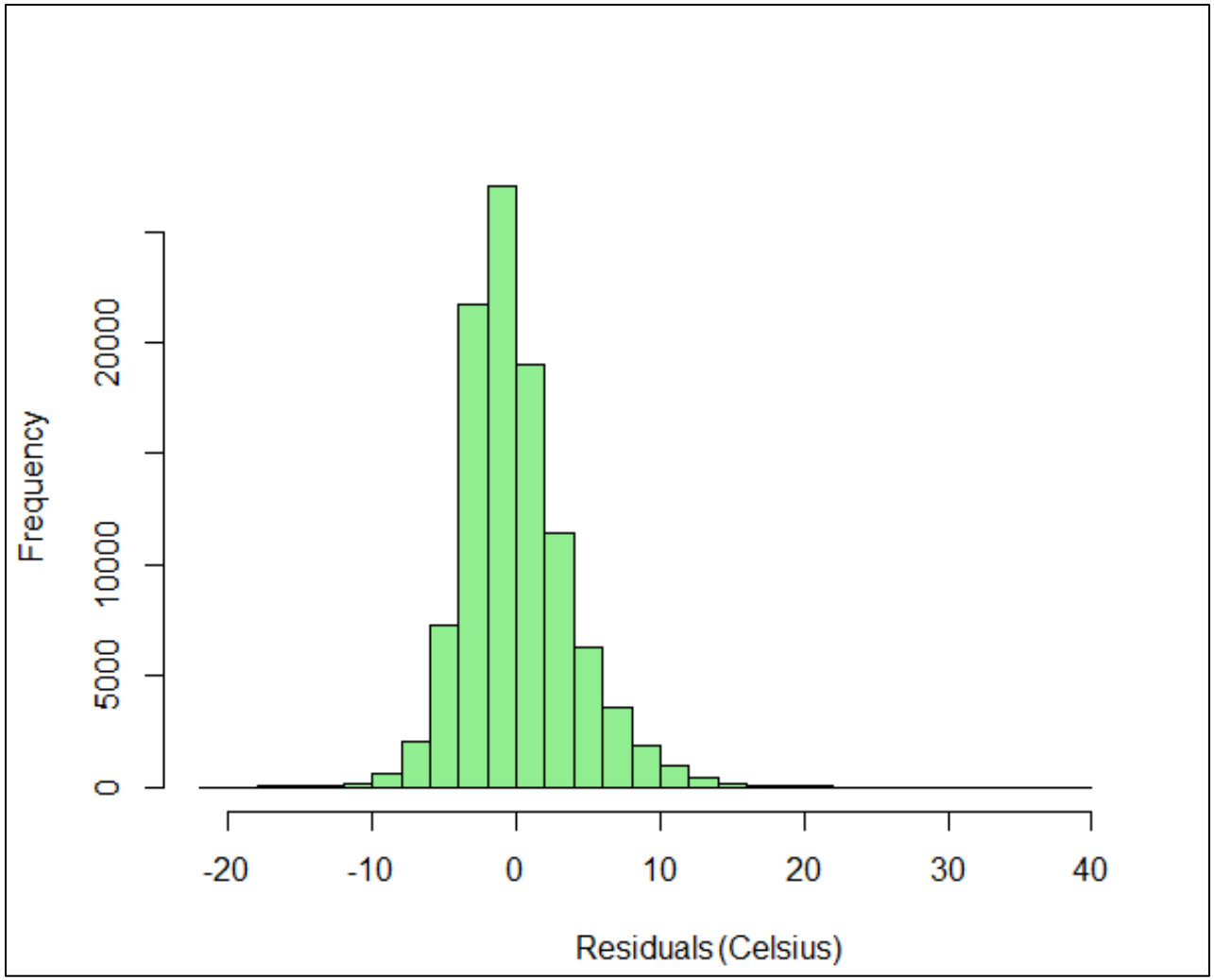

Figure 11. Histogram of residuals of the multiple linear regression LST prediction.

in a histogram. Almost all of the residuals are between -10 and 10 degrees Celsius. Areas that were predicted accurately are dispersed throughout the study area. In the northern half of the valley, the model tended to predict areas to be slightly cooler than they actually were for that specific ASTER image. At the southern end of the valley, temperatures were predicted to be far warmer. There appear to be some local phenomena where residuals are highest. For example, areas on the map that were largely composed of the NLV/SOIL class were predicted much cooler than they should have been. The southern portion of the map contained a large percentage of ILV in the form of crops. One possible reason why residuals were high around this area was because half of crops were classified as NLV/SOIL rather than ILV because the crops were sparse or recently harvested. It is also possible that air temperature in that area could be lower due to air flow from Utah Valley to the south. Analysis of additional ASTER scenes would help reveal whether these residuals are consistent through time.

The multiple linear regression model had an R^2 of 0.56, meaning a little more than half the variation in the data was explained using the variables provided. Evidently, there are many more variables that need to be explored in order to provide a more accurate prediction of LST. Examples of these variables might include configuration metrics used in other studies (Connors et al., 2013; Du et al., 2016; Zhou et al., 2011). These variables examine the landscape from an alternative perspective. Examples of these include patch density, edge density, shape index, and fractal dimension. The mentioned studies also seem to indicate that scale plays a large role in determining which variables are more valuable in predicting LST. Land cover configuration and scale should be researched to uncover more information of urban vegetation's ability to cool the environment.

Limitations

This study had its limitations. First, the Salt Lake Valley was classified using only one algorithm. It would be interesting to learn how well other algorithms such as support vector machines or object-based classification would classify an urban area using the provided datasets. This study did not deeply explore each individual variable's interaction with other individual variables. For example, LST in a pixel may significantly change if 30% of it is deciduous and the remaining makeup is 70% impervious surface or 70% building. The multiple linear regression model does not quantify these relationships. Other statistical models that account for these relationships may be able to predict LST using land cover composition variables more accurately.

The analysis was performed using one ASTER LST image. There may have been hourly, daily, monthly variations in LST that will go unaccounted for. Human-made features are known to appear hotter than surrounding areas at night time because of the urban heat island effect (Buyantuyev & Wu, 2010). Also, the landscape changes throughout the year. For example, deciduous trees drop their leaves in the fall, resulting in far less evapotranspiration. Thus, it is impossible to generalize the findings of this study for the entire year. It would be useful to repeat the study for different times of the year. Finally, urban vegetation's relationship with LST was only examined at the 90-meter scale. It would be useful to look at this relationship at coarser levels to potentially uncover any variation in regards to scale.

Finally, there was still strong spatial autocorrelation in the variables even after stratifying the observations. There was an obvious temperature trend from the northwest portion of the study area to the southeast as shown in both Figures 2 and 10. An

interesting future study would be to break the study area into smaller scale areas or neighborhoods and perform the analysis on each smaller unit as performed by Connors et al. (2013). Possible neighborhood types would include commercial, industrial, old residential, new residential, and other types. Not only would this help in accounting for spatial autocorrelation, but it would reveal if urban vegetation cover impacts LST the same way, in all types of neighborhoods.

Future studies could also address spatial autocorrelation by using a spatial filtering technique before creating the multiple linear regression model. Spatial filtering is performed by separating the spatial effects of a variable from its total effects (Getis & Griffith, 2002). The Getis Filtering approach is one such technique (Getis, 1990). It is performed by finding the distance within which areal units are spatially dependent (identified using either the G_i spatial statistic or a variogram) and examining each observation to discover its contribution to the spatial dependence in the dataset. Observations that contribute the most are filtered out into a separate dataset. In Getis and Griffith's (2002) demonstration of the technique, they found that adding the nonautocorrelated variables as predictors in their regression model caused the Adjusted R^2 to remain high (0.75). In addition, the spread of the residuals decreased and all of the coefficients became highly significant (Getis & Griffith, 2002). The Griffith Eigenfunction Decomposition Approach would also be a suitable spatial filtering technique for this study (Griffith, 2000). In this procedure, carefully selected eigenvectors for each variable are used to identify orthogonal patterns in the data. The eigenvectors act as surrogates for missing data when used in regression analyses. In Getis and Griffin's (2002) demonstration of the technique, they found that including eigenvectors as

predictors in their model increased the Adjusted R^2 value from 0.61 to 0.80 (Getis & Griffith, 2002). Eigenvectors have seen little use as predictor variables in similar urban heat island studies and would greatly enhance a regression model's ability to predict LST.

CONCLUSIONS

Urban vegetation plays an important role in providing comfort to humans in semiarid environments. This vegetation absorbs solar radiation, cools the environment through transpiration, and also provides shade. The Salt Lake Valley is situated in a semiarid region and relies on its urban vegetation to comfort its inhabitants during the warmer, summer months. Salt Lake City places a high importance on its urban forestry. In 2007, the mayor of Salt Lake City announced an initiative to plant one million trees in the county by 2017. The Salt Lake Urban Forestry program (www.slcgov.com/forestry) has made it their goal to facilitate a healthy and sustainable environment by protecting, maintaining, and planting trees. Smaller organizations like TreeUtah (treeutah.org) also make it their mission to educate people on the environmental and social benefits that trees provide as well as plant new trees. The results of this study could potentially be useful in urban planning. The land cover map generated by random forest could help identify areas in the Salt Lake Valley where urban vegetation is less abundant (Figure 7). Organizations would be able to plant trees or provide other forms of urban vegetation in those areas to help moderate temperatures and beautify these areas.

Nearly one-third of the study area was classified by random forest as some form of urban vegetation. Because urban vegetation in a semiarid climate requires a lot of irrigated water to keep it healthy, some may question whether or not it is worth devoting valuable water resources to it. The results of this study indicate that increasing

percentages of urban vegetation types in this study area do lead to cooler daytime summer temperatures. Results also indicate that irrigated vegetation is more strongly correlated with temperature than native vegetation of the study area, and that tree cover is associated with a larger decrease in LST than ILV cover. Typically, human-made features such as buildings are strongly associated with higher LST. That relationship was not as strong in this study area. Since the multiple linear regression model only explained approximately half of the variation in LST, it is recommended that further research include land cover configuration variables. Knowledge of those variables would further explain and validate the benefits of the Salt Lake Valley's urban forest.

APPENDIX

Table 6: Confusion matrix generated internally using OOB sample.

	Reference Data								
Class	CONIF	DECID	ILV	IMPER	BUILDING	NLV/SOIL	WATER	TOTAL	**UA(%)**
CONIF	5518	666	32	17	62	5	0	6300	0.876
DECID	279	9147	189	16	19	18	3	9671	0.946
ILV	12	156	23892	56	5	6	0	24127	0.99
IMPER	3	25	10	24404	292	381	4	25119	0.972
BUILDING	19	14	1	248	47358	58	9	47707	0.993
NLV/SOIL	5	14	11	461	38	41366	1	41896	0.987
WATER	0	1	0	107	4	6	18022	18140	0.993
TOTAL	5836	10023	24135	25309	47778	41840	18039	172960	-
PA(%)	0.946	0.913	0.99	0.964	0.991	0.989	0.999	-	-

Overall Accuracy = 98%

Kappa Coefficient = 0.98

REFERENCES

Akbari, H. (2009). Cooling our communities. A guidebook on tree planting and light-colored surfacing. Washington, DC: *Lawrence Berkeley National Laboratory*.

Akbari, H., Rosenfeld, A., Taha, H., & Gartland, L. (1996). Mitigation of summer urban heat islands to save electricity and smog. In *76th Annual Meteorological Society Meeting, Atlanta, GA* (Jan 28 – Feb 2 1996).

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5-32.

Buyantuyev, A., & Wu, J. (2010). Urban heat islands and landscape heterogeneity: Linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. *Landscape Ecology*, *25*(1), 17-33.

Chehata, N., Guo, L., & Mallet, C. (2009). Airborne lidar feature selection for urban classification using random forests. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *39*(Part 3/W8), 207-212.

Chen, X. L., Zhao, H. M., Li, P. X., & 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*(2), 133-146.

Congalton, R. G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, *37*(1), 35-46.

Connors, J. P., Galletti, C. S., & 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*(2), 271-283.

Dimoudi, A., & Nikolopoulou, M. (2003). Vegetation in the urban environment: Microclimatic analysis and benefits. *Energy and Buildings*, *35*(1), 69-76.

Du, S., Xiong, Z., Wang, Y. C., & Guo, L. (2016). Quantifying the multilevel effects of landscape composition and configuration on land surface temperature. *Remote Sensing of Environment*, *178*, 84-92.

Dwyer, J. F., McPherson, E. G., Schroeder, H. W., & Rowntree, R. A. (1992). Assessing the benefits and costs of the urban forest. *Journal of Arboriculture*, *18*, 227-227.

Gatziolis, D., & Andersen, H. E. (2008). *A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest* (pp. 1-32). US Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Getis, A. (1990). Screening for spatial dependence in regression analysis. *Papers in Regional Science*, *69*(1), 69-81.

Getis, A., & Griffith, D. A. (2002). Comparative spatial filtering in regression analysis. *Geographical Analysis*, *34*(2), 130-140.

Gillespie, A., Rokugawa, S., Matsunaga, T., Cothern, J. S., Hook, S., & Kahle, A. B. (1998). A temperature and emissivity separation algorithm for Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images. *Geoscience and Remote Sensing, IEEE Transactions on*, *36*(4), 1113-1126.

Gluch, R., Quattrochi, D. A., & Luvall, J. C. (2006). A multi-scale approach to urban thermal analysis. *Remote Sensing of Environment*, *104*(2), 123-132.

Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random forests for land cover classification. *Pattern Recognition Letters*, *27*(4), 294-300.

Guan, H., Yub, J., Lia, J., & Luoc, L. (2012). Random forests-based feature selection for land-use classification using LIDAR data and orthoimagery. *International Archives of the Photogrammetry. Remote Sensing and Spatial Information Sciences*, *39*, B7.

Guo, L., Chehata, N., Mallet, C., & Boukir, S. (2011). Relevance of airborne lidar and multispectral image data for urban scene classification using Random Forests. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(1), 56-66.

Griffith, D. A. (2000). A linear regression solution to the spatial autocorrelation problem. *Journal of Geographical Systems*, *2*(2), 141-156.

Grimmond, C. S. B., & Oke, T. R. (1991). An evapotranspiration-interception model for urban areas. *Water Resources Research*, *27*(7), 1739-1755.

Huang, G., Zhou, W., & Cadenasso, M. L. (2011). Is everyone hot in the city? Spatial pattern of land surface temperatures, land cover and neighborhood socioeconomic characteristics in Baltimore, MD. *Journal of Environmental Management*, *92*(7), 1753-1759.

Kinney, P., Shindell, D., Chae, E., & Winston, B. (2001). Climate change and public health: Impact assessment for the NYC metropolitan region. Washington, DC: US

Global Change Research Program, US National Assessment of the Potential Consequences of Climate Variability and Change, Metropolitan East Coast Assessment.

Lowry Jr, J. H. (2010). Spatial analysis of urbanization in the Salt Lake Valley: An urban ecosystem perspective. *All Graduate Theses and Dissertations*, *746*. Retrieved from http://digitalcommons.usu.edu/etd/746

Mitchell, M. W. (2011). Bias of the Random Forest out-of-bag (OOB) error for certain input parameters. *Open Journal of Statistics*, *1*(03), 205.

Motohka, T., Nasahara, K. N., Oguma, H., & Tsuchida, S. (2010). Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sensing*, *2*(10), 2369-2387.

NASA LP DAAC. (2001). *ASTER Level 2 Surface Temperature Product*. Retrieved from https://doi.org/10.5067/ASTER/AST_08.003.

Nowak, D. J., & Dwyer, J. F. (2007). Understanding the benefits and costs of urban forest ecosystems. In J. E. Kuser (Ed.), *Handbook of urban and community forestry in the northeast* (pp. 11-25). New York, NY: Kluwer Academic.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673-690.

OpenTopography. (2014). *State of Utah Acquired Lidar Data - Wasatch Front*. Retrieved from http://opentopo.sdsc.edu/raster?opentopoID=OTSDEM.122014.26912.1.

Pal, M. (2005). Random forest classifier for remote sensing classification. *International Journal of Remote Sensing*, *26*(1), 217-222.

Population Reference Bureau (2013). *2013 World population datasheet*. Retrieved from http://www.prb.org/pdf13/2013-population-data-sheet_eng.pdf.

Poumadere, M., Mays, C., Le Mer, S., & Blong, R. (2005). The 2003 heat wave in France: Dangerous climate change here and now. *Risk Analysis*, *25*(6), 1483-1494.

Quattrochi, D. A., & Ridd, M. K. (1998). Analysis of vegetation within a semi-arid urban environment using high spatial resolution airborne thermal infrared remote sensing data. *Atmospheric Environment*, *32*(1), 19-33.

Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *67*, 93-104.

Rouse Jr, J., Haas, R. H., Schell, J. A., & Deering, D. W. (1974). Monitoring vegetation systems in the Great Plains with ERTS. *NASA special publication*, *351*, 309.

Schmid, K., Waters, K., Dingerson, L., Hadley, B., Mataosky, R., Carter, J., & Dare, J. (2008). Lidar 101: An introduction to lidar technology, data, and applications. *NOAA Coastal Services Center*, 76.

Smith, W. K., & Hinckley, T. M. (1995). *Resource physiology of conifers: Acquisition, allocation, and utilization*. San Diego, CA: Academic Press.

Turner, M. G. (2005). Landscape ecology: What is the state of the science? *Annual Review of Ecology, Evolution, and Systematics*, *36*, 319-344.

U.S. Census Bureau. (2015). *QuickFacts: Salt Lake County, Utah*. Retrieved from http://www.census.gov/quickfacts/table/PST045215/49035

Utah AGRC. (2015). *2014 Naip 1 Meter Orthophotography*. Retrieved from https://gis.utah.gov/data/aerial-photography/2014-naip-1-meter-orthophotography/

Weier, J., & Herring, D. (2011). Measuring Vegetation (NDVI & EVI). Retrieved from http://earthobservatory.nasa.gov/Features/MeasuringVegetation/

Weng, Q., Lu, D., & Schubring, J. (2004). Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. *Remote sensing of Environment*, *89*(4), 467-483.

Yuan, F., & 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.

Zhou, W., Huang, G., & 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*(1), 54-63.