

DEVELOPMENT OF A FORECASTING TECHNIQUE FOR THE CHARLOTTE URBAN
HEAT ISLAND INTENSITY

by

Anna Stuck

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Earth Science

Charlotte

2020

Approved by:

Dr. Matthew Eastin

Dr. Casey Davenport

Mr. Terry Shirley

©2020
Anna Stuck
ALL RIGHTS RESERVED

ABSTRACT

ANNA STUCK. Development of a Forecasting Technique for the Charlotte Urban Heat Island Intensity (Under the direction of DR. MATTHEW EASTIN)

The urban heat island (UHI) of Charlotte (North Carolina) - a rapidly expanding subtropical city - was studied through statistical analysis and short-term predictive modeling. Our study used hourly surface observations taken from 12 weather stations over a 5-year period. One station was identified as the least urban and used as the rural reference. Two stations were deemed the most urban and were used for UHI analysis and development of a UHI predictive. Observations from the two urban stations were combined to account for dissimilar parameters observed by each station.

Previous work has shown that the daily UHI maximum (4-10 K) often occurs during nocturnal hours when optimal weather conditions (i.e., clear skies, light winds, low humidity, and strong static stability) are present. This study examines whether such nocturnal UHI maximum can be predicted from daytime weather parameters observed 6-9 hours prior. Using daytime weather parameters, two statistical model types (generalized linear models and bootstrap random forests) were evaluated for predicting the nocturnal UHI magnitude. The bootstrap random forest models were found to out-perform the generalized linear models. Stratification of the data by season and day of the week further improved random forest models.

DEDICATION

This thesis is dedicated to my mom, Ann Stuck. While you weren't able to see me complete my research, you never stopped motivating me and inspiring me along the way.

ACKNOWLEDGEMENTS

I would like to thank Dr. Matthew Eastin for guiding me through my research and helping me throughout my time here at UNC Charlotte. This project could not have been done without Dr. Eastin's patience and motivation. I would also like to give a special thanks to my committee, Dr. Casey Davenport and Mr. Terry Shirley for being incredibly flexible during these unprecedented times when scheduling around our crazy schedules and providing valuable feedback at my proposal defense. To my friends in the meteorology office, thank you from the bottom of my heart for being my support system, my best friends, and for making me a better meteorologist and person. To my mom, thank you for everything you did for me and for motivating me every day even when you are not here. Finally, I would like to thank my family and friends for always supporting me and believing in me.

TABLE OF CONTENTS

LIST OF TABLES.....	vii
LIST OF FIGURES	viii
CHAPTER I: INTRODUCTION AND PROBLEM STATEMENT.....	1
CHAPTER II: BACKGROUND	3
CHAPTER III: RESEARCH QUESTIONS AND GOALS.....	8
CHAPTER IV: DATA AND METHODS.....	9
CHAPTER V: RESULTS	14
CHAPTER VI: CONCLUSIONS AND FUTURE WORK.....	20
REFERENCES	38

LIST OF TABLES

TABLE 1: Surface weather station metadata, including latitude, longitude, elevation, local topographic relief, land use/cover characteristics, mean daily temperature (T_{DAY}), mean daily temperature range (ΔT_{DAY}), sky view factor (SVF), and the local climate zone (LCZ). This table is from Eastin et al. (2018).

TABLE 2: The correlation coefficients between each 15 LST meteorological parameters and the 21-23 average UHI intensity.

TABLE 3: Results of initial Bootstrap Random Forest model

TABLE 4: Results of bootstrap random forest model using the GRM optimal predictors.

TABLE 5: Results of bootstrap random forest model using the 5 predictors highest correlated to the mean UHI Intensity.

TABLE 6: Statistical Results of all model run performance.

TABLE 7: Cross Correlation analysis of wind speed with other predictors.

LIST OF FIGURES

FIGURE 1: Surface weather station locations superimposed on the 2011 USGS LULC product across the CMR. The blue circle denotes the rural reference site (WAD), the yellow circles denote the most urbanized sites (GAR and CLT) discussed in this study, and the plus sign denotes the central business district of the urban core. See Table 1 for station metadata and local LULC characteristics.

FIGURE 2: Wind rose of hourly observations at CLT for all good dates between June 2010 and May 2015.

FIGURE 3: Distribution of the Adjusted Temperature between June 2010 and May 2015.

FIGURE 4: Distribution of the Precipitation between June 2010 and May 2015.

FIGURE 5: Distribution of the Relative Humidity between June 2010 and May 2015.

FIGURE 6: Distribution of the Vapor Pressure between June 2010 and May 2015.

FIGURE 7: Distribution of the Station Pressure between June 2010 and May 2015.

FIGURE 8: Distribution of the Wind Speed between June 2010 and May 2015.

FIGURE 9: Distribution of the Wind Direction between June 2010 and May 2015.

FIGURE 10: Distribution of the Solar Radiation between June 2010 and May 2015.

FIGURE 11: Distribution of the Sky Cover between June 2010 and May 2015.

FIGURE 12: Distribution of the Cloud Height between June 2010 and May 2015.

FIGURE 13: Distribution of the Weather Factor between June 2010 and May 2015.

FIGURE 14: Distribution of the Turner Stability Index between June 2010 and May 2015.

FIGURE 15: Distribution of the Urban Heat Island Intensity between June 2010 and May 2015.

FIGURE 16: Scatterplots of observed (actual) UHI intensity vs the predicted UHI intensity from the initial BRF model fit for the training, validation, and testing subsets.

FIGURE 17: Actual UHI plotted with Predicted UHI for All Data.

FIGURE 18: Actual UHI plotted with Predicted UHI for spring.

FIGURE 19: Actual UHI plotted with Predicted UHI for summer.

FIGURE 20: Actual UHI plotted with Predicted UHI for fall.

FIGURE 21: Actual UHI plotted with Predicted UHI for winter.

FIGURE 22: Actual UHI plotted with Predicted UHI for weekend.

FIGURE 23: Actual UHI plotted with Predicted UHI for weekday.

CHAPTER I: INTRODUCTION AND PROBLEM STATEMENT

The global population is becoming increasingly urbanized; as of 2007 more than 50% of the population lives within urban areas, and by 2030 over 5 billion people will live within an urban area (Mills 2007). As the world becomes more urbanized, contributions to local climate change by Urban Heat Islands (UHIs) are expected to increase. While the UHI is one of the most widely studied urban climate phenomena, our understanding of the full range of forcing factors responsible for UHIs remains incomplete. For example, while it is well recognized that a UHI is primarily a nocturnal phenomenon that arises from differential heating/cooling rates between urban regions containing impervious built surfaces (buildings, roads, etc.) and rural regions containing natural vegetated surfaces (forest, grass, crops, etc.), our ability to predict UHI intensity remains a challenge. On any given night, UHIs result from a complex balance among multiple local factors that control the nocturnal cooling rates, including meteorological conditions, urban form, and human activity. Given that such factors are also influenced by climate, season, and geography, the governing UHI controls, and their relative importance, are often unique to each city (Grimmond 2006). Previous studies have found a diverse set of influencing factors on the UHI in relation to meteorological controls, atmospheric controls, and controls due to urban form. Significant meteorological controls such as wind speed and cloud cover have been found to be the most significant controlling factors in previous studies (Morris et al., 2001; Runnalls & Oke, 2000). Significant atmospheric controls found in previous studies include humidity (Hoffmann et al. 2012), stability (Holmer et al. 2013), and solar radiation (Steenefeld et al. 2011).

In this study, we developed a series of statistical models that provide short-term predictions of the nocturnal UHI intensity for Charlotte, North Carolina. The goal was to use afternoon observations to predict that evening's UHI intensity. Using previously acquired and quality-controlled meteorological data from standard surface weather stations, a series of statistical evaluations involving lagged cross-correlation analysis between various potential

afternoon parameters and the corresponding nocturnal UHI intensity were used to identify a reduced set of more optimal afternoon predictors. Then, a variety of model types were explored, including generalized linear models (GLMs) and nonlinear bootstrap random forests (BRFs), using the reduced set of optimal predictors.

CHAPTER II: BACKGROUND

Several studies have analyzed daily UHI variability based on surface weather observations. In an early study, Eliasson (1996) used observations from one urban and three rural meteorological stations near Goteborg, Sweden to develop statistical models that predict the intra-urban air temperature differences. Hourly observations over a 3-year period were separated into three categories depending on cloud cover. Then, a multiple regression model was developed for each category using a predictor pool that included total cloud cover, cloud base, air temperature, absolute air humidity, wind speed, wind direction, and a seasonal factor. The optimal regression models showed modest success in predicting the intra-urban temperature variations with wind speed and cloud cover being the most valuable predictors. These results showed that low winds speeds and limited cloud cover favor strong nocturnal UHIs.

Kim and Baik (2002) studied the daily maximum UHI intensity for Seoul, South Korea. Using hourly data from two surface weather stations (one urban; one rural) over a 24-year period, the UHI was found to be more intense during the cooler seasons and weaker in the warm seasons, while the most intense UHI are often observed during calm, cloudless nights. The authors developed a neural network to predict UHI intensity using four predictors, including the maximum UHI from the previous day, wind speed, cloud cover, and relative humidity. Of those predictors, the previous day UHI was positively correlated, while the others are negatively correlated, to the maximum UHI. Overall, the neural network (which explained 52.4% of the total variance) outperformed a simple multiple linear regression model using the same four predictors (which explained 46.1% of the total variance). These results suggest that nonlinear data-mining techniques (such as neural networks, regression trees, and random forests) may provide better UHI predictions than linear regression models.

Bottyan and Unger (2003) studied the effects of urban surface factors, such as the built-up ratio, the water surface ratio, the sky view factor, and building height on UHI intensity in Szeged, Hungary. Using mobile temperature observations between March 1999 and February 2000, the UHI intensity was measured once a week over a one-year period. A clear connection between the spatial distribution of UHI intensity and the land use parameters was observed. Also, in the cool season, a UHI intensity greater than 2 K covered about 37% of the investigated area, while only 2% of the area was similarly covered during the warm season. The authors constructed a stepwise multiple regression model for each season using the mean UHI intensity and the mean urban surface parameters. Overall, the sky view factor and building height proved to be the most important predictors. These results suggest that urban form and land cover play non-negligible roles in modulating UHI spatial distribution and maximum intensity.

Schatz and Kucharik (2014) used hourly observations from 135 surface weather stations to document the spatiotemporal variability of UHI intensity across Madison, Wisconsin during an 18-month period (March 2012 – October 2013). From a seasonal perspective, the primary UHI peak occurred during the summer with a secondary peak during winter. The winter UHI was modulated by differential snow cover across the area, while the summer UHI was primarily modulated by wind speed and cloud cover (with the most intense UHIs occurring on nights with weak winds and clear skies). Overall, higher rural plant biomass in the summer promoted a more intense UHI due to the greater differences between the urban and rural cooling rates after sunset. These results suggest that metropolitan areas with substantial urban tree cover (such as Charlotte) may experience less intense UHIs during the warm season.

Recently, Eastin et al. (2018) used hourly observations from 12 surface weather stations (two urban, nine suburban, and one rural) to document the temporal variability of the UHI in Charlotte, North Carolina over a 5 year period (June 2010 to May 2015). Over 70% of dates exhibited prominent nocturnal UHIs (greater than 1 K). The most intense UHIs occurred on winter

nights with light winds, clear skies, low humidity, strong low-level stability, and no precipitation or frontal passage. The maximum hourly UHI typically occurred one to two hours before midnight. This was attributed to a positive sensible heat flux from urban surfaces that was maintained for 3-4 hours after sunset, causing urban cooling to lag rural cooling. Daily variations in nocturnal mean UHI intensity exhibited significant positive correlations with cloud-base height, atmospheric stability, NO₂ concentration, and total solar radiation. Daily variations in nocturnal mean UHI intensity exhibited significant negative correlations with relative humidity, wind speed, and cloud cover. When optimal weather for UHI development was present, UHIs were more intense on weekdays (greater air pollution) than on weekends (less air pollution). These results further suggest that local weather conditions and air pollution play substantial roles in modulating UHI intensity. The database used by Eastin et al. (2018) is used in this study – additional details are provided below.

Historically, the UHI has been studied primarily through observations, but high-resolution mesoscale models have increasingly been used to study the physical factors responsible for intense UHI events. However, an important hurdle that must be overcome when simulating any UHI event is a fair representation of urban factors known to influence UHI intensity (land use/cover fractions, sky view factor, building heights). Such detail requires substantial computing resources that have only recently become available. As a result, a limited number of mesoscale models have been capable of simulating UHIs (Best 2006).

In an early attempt, Johnson and Oke (1991) created two simple numerical models to simulate the surface-level UHI under ideal nocturnal conditions (weak winds and clear skies). One model used a full system of differential equations while the other used a “force-restore” approach (Johnson et al. 1991). Both models used computationally expensive parameterizations for building height, sky-view factor, internal building temperature, ground/soil temperature, surface thermal admittance, surface emissivity, and down-welling longwave radiation. Results from both

models were compared and tested against field data under similar ideal conditions. Overall, the force-restore model, also called the Surface Heat Island Model (SHIM), proved more effective at representing the cooling in both urban and rural surfaces when compared to field observations (Johnson et al. 1991).

In a follow-up study, Oke et al. (1991) used the SHIM to study the relative importance of numerous factors in developing moderate and intense UHIs. A series of sensitivity simulations revealed that UHIs are most impacted (in rank order) by (1) sky-view factor, (2) surface thermal admittance (surface type and/or materials), (3) wind speed, (4) cloud cover/height, (5) humidity, (6) urban pollution, and (7) anthropogenic heat release. These results further suggest that urban form and local weather conditions play a large role in modulating UHI intensity on a day-to-day basis, but also showed a substantial need for additional simulations and more sophisticated numerical models.

Lemonsu and Masson (2002) used a full-physics mesoscale numerical model to simulate diurnal variations in UHI intensity for Paris, France during two consecutive summer nights when strong anticyclonic conditions (weak winds, clear skies, low humidity, and strong low-level stability) prevailed. The spatiotemporal distribution of simulated surface temperature was well correlated with the available hourly observations during the period of interest, suggesting that critical governing factors were captured by the sophisticated surface parameterizations used in the model. An analysis of the surface energy budgets at several urban and rural locations revealed the following: (a) net radiation was slightly larger in the city due to a reduced sky-view that trapped the radiant energy; (b) sensible heat fluxes were much larger in the city due to limited vegetation cover, and they continued to be positive for a few hours after sunset; (c) latent heat fluxes were much larger in the rural area due to large vegetation cover; (d) storage heat fluxes were much larger in the city, and represented an atmospheric heat source at night. These results

reaffirm that urban factors play a non-negligible role in modulating UHI intensity when the local weather conditions are ideal for UHI formation.

Yang et al. (2015) used a full-physics mesoscale numerical model coupled with an urban-canopy model to examine UHI sensitivity to various aspects of urban form, including (a) building height and density, (b) building materials, (c) anthropogenic heat sources, and (d) land use/cover during ideal weather conditions (weak winds, clear skies, low humidity, and strong low-level stability). Results showed that the UHI was more intense when high-density low-rise buildings were more common (leading to less surface shade), when exterior building materials were less reflective (leading to more heat storage), when urban trees and parks are less common (leading to larger sensible heat fluxes and less latent heat fluxes), and when anthropogenic heat is large (during the summer when air-conditioner use is high and during the winter when buildings are heated). Since many urban areas (including Charlotte) consist of high-density low-rise buildings made with high heat-capacity materials and non-negligible anthropogenic heating, prominent UHIs should be more frequent and more intense when weather conditions are ideal for UHI formation. These results further reaffirm the idea that urban factors play an important secondary role in modulating UHI intensity.

Overall, the literature suggests that local weather conditions play a leading role in modulating the daily UHI intensity. The most intense UHIs occurred on nights with light winds, clear skies, low humidity, strong low-level stability, and no precipitation or frontal passage. If the weather conditions are favorable, then various measures of urban form dictate the spatiotemporal distribution of UHI intensity. These results suggest that predicting UHI intensity using local weather parameters is an important first step.

CHAPTER III: RESEARCH QUESTIONS AND GOALS

This project used the Eastin et al. (2018) database to evaluate a hierarchy of statistical models that provide short-term predictions of nocturnal UHI intensity for Charlotte, North Carolina. The goal was to use afternoon observations to predict the nocturnal UHI intensity 6-12 hours later. A variety of model types were explored, including generalized linear models (GLMs) and bootstrap random forests (BRFs), using a predictor pool consisting of meteorological parameters easily derived from standard surface weather observations. This thesis addresses the following research questions: Which afternoon meteorological parameters exhibit the greatest skill (i.e., lag correlation) at predicting the nocturnal UHI intensity? Which statistical models provide acceptable UHI predictions? Which model provides the best UHI prediction? We are expecting the daytime predictors with the biggest influence on the nocturnal UHI to include cloud cover, wind speed, humidity, and/or atmospheric stability, consistent with previous research. Regarding model type, we expected the non-linear data mining approaches will outperform the linear regression approaches.

CHAPTER IV: DATA AND METHODS

The data used was collected by 12 automated surface weather and air-quality stations across the Charlotte-Mecklenburg region. Of those stations, four were ASOS stations located at the Charlotte Airport (CLT), Gastonia (GAS), Monroe (MON), and Rock Hill (RKH). Seven stations are located near county airports in Chester (CHR), Concord (CON), Lancaster (LAN), Lincolnton (LIN), Salisbury (SAL), Statesville (STV), and Wadesboro (WAD) with one station being operated by Mecklenburg County Air Quality at Garinger High School (GAR). Each station was evaluated based on the land use land cover (LULC) percentages and placed into a category: rural, urban, or suburban (see Figure 1; Table 1). Rural sites were located greater than 40 km from the urban core, were surrounded by farmland and trees, and had the least amount of development. Of the sites classified as rural, the Wadesboro station was deemed the optimal rural site as it was located in an open grassy field, was far from any airport infrastructure, rarely downwind from the urban core, and had the smallest population. For the urban sites, both Garinger and Charlotte were the most developed and located less than 10 km from the urban core. While neither location was in the urban core, both GAR and CLT were often downwind of the urban core based on prevailing climatological synoptic patterns. Figure 2 shows a wind rose detailing long-term wind patterns for the Charlotte urban core (Eastin 2018). During the study period, wind primarily came from the southwest and the northeast. This allows for both locations to be under the influence of the urban core. Further details about the surface stations can be found in Eastin et al. (2018).

Using only the GAR, CLT, and WAD stations, hourly observations of station pressure, temperature, relative humidity, precipitation, wind speed, wind direction, cloud cover, cloud base, and net solar radiation were analyzed between June 2010 and May 2015. Due to latitudinal and elevation differences among the stations, all hourly observations were adjusted to a common elevation of 229 m using the dry adiabatic lapse rate and a common central latitude of 35.23°N

using -0.9°C per degree of latitude (the common elevation and latitude correspond to central Charlotte) (Lindén et al. 2015; Peterson 2003; Winkler et al. 1981). The hourly UHI intensity was calculated by subtracting the adjusted temperature of each urban site from the adjusted temperature of the rural site. Only dates with no missing hourly data were used to ensure a common dataset for each location. This criterion resulted in 1558 total dates and 404 optimal dates (where the weather conditions were favorable for the development of an intense UHI). Optimal weather conditions included clear to scattered skies, light winds, and low humidity with no precipitation or frontal passages (Eastin et al. 2018).

Two additional parameters were computed from the observations. First, the weather factor (WXF; Runnalls and Oke 2000) combines cloud cover and cloud type with surface winds and was computed using $\mathbf{WXF} = \mathbf{WS}^{-0.5} (1-kn^2)$ where WS is the wind speed, k is a nondimensional coefficient that accounts for decreasing cloud temperature with height and n is the cloud cover fraction. Determination of k requires unavailable hourly information on cloud type, depth, and optical thickness (Morgan et al. 1971). Thus, the k values used by Runnalls and Oke (2000) for low (0.88), medium (0.73), and high (0.24) clouds were assigned cloud base heights of 2, 5, and 8 km, respectively. Hourly estimates of k were computed through linear interpolation using the observed cloud base heights and assuming $k = 1$ at the surface and $k = 0$ for either clear skies or cloud base heights above 10km.

Second, to better understand how atmospheric stability influences UHI intensity, an hourly Pasquill-Gilford-Turner stability index (STB; Turner 1964) was computed at each site location. The STB scheme uses surface wind speed, cloud cover, and solar altitude to classify the stability of the atmosphere into seven different categories: 1) extremely unstable, 2) unstable, 3) slightly unstable, 4) neutral, 5) slightly stable, 6) stable, and 7) extremely stable. While a sounding-based temperature profile would be more ideal for estimating stability, no rawinsondes are launched

regularly in the Charlotte region. Thus, STB provides an alternative rough estimate at an hourly interval.

To create a more cohesive urban dataset, the data from the GAR station was supplemented with data from the CLT station. Parameters used from the GAR station included adjusted temperature, precipitation, relative humidity, vapor pressure, station pressure, wind speed, wind direction, solar radiation, UHI intensity from the previous day, and UHI intensity (calculated as the adjusted temperature difference between GAR and WAD). Data included from CLT was cloud cover, cloud height, WXF, and STB. This data was combined to make one dataset including all desired variables.

To identify the optimal afternoon hour(s) for predictors and the optimal overnight hour(s) for UHI intensity, a series of lagged linear correlations were computed for all possible combinations of afternoon hours (12-16 LST) and overnight hours (19-23 LST) using the full set of 1558 good dates during the 5-year study period. The hours of 17-18 LST were excluded to account for sunset and to provide a true short-term prediction of UHI intensity. Morning hours (0-12 LST) were also excluded because the UHI is often most intense during the evening hours just before midnight (Eastin et al. 2018). The detailed correlation analysis showed that the 15 LST afternoon observations were best correlated with nighttime UHI intensities, but the correlation coefficients were roughly similar when the 15 LST observations were paired with either the 21 LST, 22 LST, or 23 LST estimates of UHI intensity. Thus, the 15 LST observations were chosen to represent afternoon predictors and the average UHI intensity between hours 21-23 LST was chosen to represent the nocturnal predictand. These selections will effectively allow for a 6-8 h forecast of UHI intensity.

The correlation coefficients between each 15 LST meteorological parameters and the 21-23 average UHI intensity are shown in Table 2. The predictors with the five largest correlation

coefficients are relative humidity, weather factor, cloud height, sky cover, and vapor pressure. It is expected that these predictors will be the most skillful predictors in model performance. The predictors with the 5 lowest correlation coefficients are adjusted temperature, wind direction, wind speed, precipitation, and the turner stability index. These predictors are expected to be the least skillful predictors. Station pressure and solar radiation were in the middle with correlation coefficients of just under 0.3. It is unknown how these predictors will perform in model development.

The next step was to develop and evaluate various statistical models using the various 15 LST observations as predictors of the 21-23 LST mean UHI intensity. The following statistical model types were evaluated using the JMP statistical software: generalized linear models (GLMs) and bootstrap random forests (BRFs).

A generalized linear model uses the following basic formula:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n + \varepsilon$$

where y is the response or dependent variable, x_n are the predictors, β_n are the estimators or weights for each predictor, β_0 is the intercept, n is the total number of predictors, and ε represents residual errors (which under the assumption of normalcy are assumed to be zero). The GLM generalizes regression by using a link function to relate the linear model to the response variable and using the variance magnitude of each measurement (i.e., estimators or weights) to be a function of the response. Assumptions of a GLM are that all predictors are independent (from one another), all predictors exhibit normal distributions, and the response exhibits a normal distribution (McCullagh and Nelder 1989).

A bootstrap random forest averages numerous decision trees to create an ensemble model. In lieu of using a mathematical formula, a decision tree uses a series of branches (equivalent to nested if/then statements in code) to guide the model through a series of decisions

to provide a prediction. Training data is randomly selected with replacement and used to create smaller datasets. These bootstrapped samples are then fed through multiple decision trees with the result being the averaged results from all trees. Each split in each tree included in the forest randomizes which predictors are included. Individually, the decision trees have a wide range of statistical significance. A bootstrap forest model averages the result from each tree to create a stronger model (Breiman 1996; Breiman et al. 1983).

CHAPTER V: RESULTS

The first model evaluated was a generalized linear model. Given that GLMs assume a normal distribution for all predictors and the predictand, the full dataset of adjusted temperature, vapor pressure, station pressure, sky cover, cloud height, weather factor, turner stability index, precipitation, relative humidity, wind speed, and mean nocturnal UHI intensity were checked for normalcy using the Kolmogorov-Smirnov test. Figures 3-14 show the distributions of each predictor with Figure 15 showing the distribution of the Urban Heat Island Intensity. A log transform was applied to all non-normal distributions (relative humidity, precipitation, wind direction, and wind speed).

Prior to any GLM fitting, the full dataset was split into independent modeling and testing subsets (often called developmental and validation subsets). The modeling subset consisted of the first three years (June 2010 – May 2013) or ~60% of the full dataset, while the testing subset consisted of the last two years (June 2013 – May 2015) or ~40% of the full dataset. A larger modeling subset was used to decrease overfitting and increase model robustness.

The initial GLM fit (using the 60% modeling subset) included all predictors and exhibited a corrected Akaike Information Criterion (AICc) near 260. More importantly, several individual predictors exhibited regression coefficients that were not statistically significant at the 5% level. To improve model fit, a stepwise process was invoked whereby the least skillful non-significant predictor was removed and a new GLM was fit. This process was repeated until all predictors exhibited statistically significant regression coefficients at the 5% level. The final “optimal” GLM included vapor pressure, weather factor, station pressure, solar radiation, relative humidity, and the UHI from the previous day as predictors in respective order of significance with an AICc over 4,300.

Validation statistics run on the testing years resulted in a mean absolute error in UHI intensity of 1.28 K, a root mean square error of 1.63 K, a mean percent error of 104% and a mean bias of 0.01 K. Given the dramatically increased AICc for all re-fit GLMs (including the optimal model) during the model development stage, combined with the rather large mean percent error during model validation, it was concluded that GLMs are not the most ideal statistical model type for predicting the Charlotte UHI intensity.

The next model evaluated was the Bootstrap Random Forest (BRF) model. For this model, no initial division of the full dataset into training and validation subsets was required since the JMP software automatically separates the data into training (60%), validation (20%), and testing (20%) subsets during model development. Training data is used to build the model and validation data is used *during model development* to provide an unbiased evaluation of the fit of the model on the training data while fine tuning hyper parameters. Finally, the testing data is the independent sample used for traditional model validation once an optimized model has developed. The separation process is random and the percentages are user-defined (the default values shown above were used for the initial evaluation). As with the GLMs, a log transform was applied to all non-normal distributions before any model development.

The initial BRF model fit included all predictors and exhibited a generalized R-square value of 0.71, suggesting BRFs may be a more ideal model type for predicting the Charlotte UHI intensity. Interestingly, however, model performance decreased with the validation and test subsets, exhibiting generalized R-square values of 0.391 and 0.334, respectively (see Table 3). Figure 16 shows the observed UHI intensity plotted with respect to the BRF predicted UHI intensity for the training, validation, and testing subsets; the training data clearly performed the best. The BRF was run using the optimal predictors from the final GLM which included vapor pressure, weather factor, station pressure, solar radiation, relative humidity, and the UHI from the previous day. This exhibited a training R-square of 0.29 with validation and test R-square values

of 0.26 and 0.25 respectively (see Table 4). BRF model was also run using the top 5 predictors highest correlated with mean UHI Intensity (see Table 2); relative humidity, weather factor, cloud height, sky cover, and vapor pressure. This run of the BRF performed better than the GLM optimal predictors with a training R-square of 0.361. Similar to the initial run, performance decreased with validation and test R-squares of 0.28 and 0.29 respectively (see Table 5.) Since model performance did not increase with limiting predictors included in model development, all predictors were included for the rest of the models.

Different user defined percentages for training, validation, and testing the BRF model were tested using the full dataset. The percentage breakdown with the highest overall performing testing R-square included 50% (853 dates) for model training, 40% (687 dates) for model validation, and 10% (167 dates) for model testing. The optimal settings or settings with the highest performing testing R-square were 175 trees with 10 predictors sampled per split and a minimum of 10 splits. This model run resulted in a training R-square of 0.795, a validation R-square of 0.362, and a testing R-square of 0.457. The top five predictors for this model run were vapor pressure, relative humidity, weather factor, temperature, and solar radiation. Figure 17 shows the actual UHI plotted against the predicted UHI which shows a loose correlation. This optimal model using all of the data was used as a baseline for the rest of model development on seasonal and weekend/weekday data (see below).

To better optimize the BRF, data was stratified by season (i.e., summer, fall, winter, and spring) and day of the week (i.e., weekday and weekend). Summer data included the months June, July, and August and fall data included September, October, and November. Winter data included December, January, and February with spring data including March, April, and May. Weekend data included Saturday's, Sunday's, and all National Holidays with weekday data including Monday-Friday with national holidays excluded.

For spring data, the optimal model used 50% (201 dates) for training, 40% (162 dates) for validation, and 10% (39 dates) for testing and resulted in a testing R-square of 0.415. This model run had 225 trees, 7 predictors sampled per split, and a minimum of 10 splits. The top five predictors were relative humidity, solar radiation, vapor pressure, weather factor, and temperature. Figure 18 shows the predicted UHI plotted with the actual UHI and displays a more scattered relationship than with all of the data.

The optimal BRF run on the summer data included 50% (224 dates) for training, 40% (183 dates) for validation, and 10% (46 dates) for testing. This model had 225 trees with 7 predictors sampled per tree, and a minimum of 10 splits per tree. The top five predictors were vapor pressure, the urban heat island from the previous day, solar radiation, relative humidity, and weather factor. Among all four seasons, the optimal summertime model performed the worst, with a testing R-square of 0.21. When looking at Figure 19, the relationship between the actual UHI and predicted UHI is relatively scattered with less correlation than was apparent with the optimal springtime model (Figure 18).

The fall BRF run included 50% (226 dates) for training, 40% (178 dates) for validation, and 10% (45 dates) for testing. This model had 100 trees with 3 predictors sampled per tree, and a minimum of 10 splits per tree. The testing R-square was 0.415, and the top five predictors were relative humidity, vapor pressure, solar radiation, temperature, and the urban heat island from the previous day. Figure 20 shows the actual UHI plotted with the previous UHI and shows a strong relationship with slight under prediction of the UHI compared to actual UHI.

The optimal winter BRF model included 204 dates in the training subset, 157 dates in the validation subset, and 42 dates in the testing subset, with 200 trees, 10 predictors sampled per split, and a minimum of 10 splits. The optimal model exhibited an R-square of 0.599, and the top five predictors were weather factor, cloud height, solar radiation, relative humidity, and the turner

stability index. Figure 21 shows a strong relationship between the observed and predicted UHI values for winter dates.

The initial run of the BRF for weekends was the most optimal with 50% (266 dates) for training, 40% (214 dates) for training, and 10% (54 dates) for testing. With 100 trees, 3 terms sampled per split, and a minimum of 10 splits per tree, the optimal BRF for weekends has an R-square of 0.277. The top five predictors are relative humidity, weather factor, vapor pressure, cloud height, and solar radiation. Figure 22 shows the poor relationship between the predicted UHI and the actual UHI. The optimal run of the BRF for weekdays had user defined percentages of 40% (586 dates) for training, 30% (351 dates) for validation, and 30% (351 dates) for testing. With 225 trees, 7 splits per tree, and a minimum of 10 splits, the optimal model for weekdays had a testing R-square of 0.381. The top five predictors are relative humidity, weather factor, vapor pressure, turner stability index, and solar radiation. The weekday model has the turner stability index as one of the top predictors. This is important due to the influence of an inversion and pollution caused from individuals commuting to and from work. Compared to Figure 22, Figure 23 shows a slightly closer relationship between the predicted UHI when compared to the actual UHI.

Statistical validation was completed for each optimal BRF model and is shown in Table 6. To compute each validation metric, the entire dataset (i.e., the full dataset along with the seasonal and day-of-week subsets) was used (i.e., not just the 10% testing subsets). The average UHI for each model subset was also computed. Among the four seasons, spring exhibits the largest UHI and fall exhibits the smallest. For the mean absolute error (MAE), summer exhibits the smallest MAE and spring exhibits the largest. The MAE tells us how accurate the model is when predicting the UHI with a value of 0 being a perfect model performance. Of the four seasons winter and spring exhibit a positive bias indicating an over prediction of the UHI by the model. Fall exhibits virtually no bias at all and summer exhibits a slight negative bias indicating a slight under prediction of the UHI. Bias are a good indicator of how the model prediction trends and can be

easily subtracted or added to the model output to create a more accurate model prediction. Spring exhibits the largest percent error with 146.03% and winter exhibited the smallest percent error with 21.33%. Percent error is another way to measure how accurate the model prediction is.

For the weekend and weekday data, the weekdays on average have a slightly larger UHI by just 0.03 degrees Celsius. The weekend data has both a larger mean absolute error and a larger percent error than the weekday data meaning the model performance is not as accurate for weekends. This could be due to having a smaller sample size. Both the weekend and weekday data have a positive bias with weekends having a very small bias at just 0.03.

CHAPTER VI: CONCLUSIONS AND FUTURE WORK

While seasonality is not an effective predictor, knowing which season it is can greatly influence the effectiveness of general model performance and model predictors. For example, weather factor is a top predictor for winter, but does not perform as well in the other seasons. As mentioned before, weather factor combines surface winds and cloud cover and cloud type (Runnalls and Oke 2000). Winter is typically windier and days with low lying clouds are more common. High cloud cover and high winds would both inhibit the development of a large UHI. Solar radiation and relative humidity are the only common effective predictors for all four seasons. Solar radiation is an effective predictor for all seasons due the relationship between increased solar radiation and daytime urban heat storage. Greater amounts of solar radiation lead to greater daytime urban heat storage which transitions to a large UHI source after sunset. Relative humidity is an effective predictor for all seasons because of the “greenhouse effect” influence that atmospheric moisture has on surface air temperatures. Vapor Pressure is a top predictor for all models except for winter. Similar to relative humidity, vapor pressure is a measure of the moisture in the atmosphere. With the small day to day variations in atmospheric moisture during winter, vapor pressure would not be an effective predictor. Temperature is a top predictor for all of the data, spring, and fall models. This could be due to the large day to day temperature variations during the spring and fall when compared to the small variations during summer and winter.

Knowing if it is a weekend or weekday is also not an effective predictor, but model performance varies depending on the day of the week. The larger UHI on weekdays could be due to increased urban pollution caused by people commuting to and from work. An increase in urban pollution could lead to an increase in stored heat energy near the surface from an enhanced local greenhouse effect. This is backed up by the presence of the turner stability index as an effective predictor for the weekdays. As mentioned earlier, the turner stability index is a measure of stability in the atmosphere. A more stable atmosphere would be conducive for a larger UHI on weekdays

due to limited vertical mixing and air pollution dispersal. With limited vertical mixing, air pollution would enhance a local greenhouse effect. In addition to the turner stability index, relative humidity is an effective predictor for weekdays as well as weekends. A high relative humidity percentage would also be conducive for an enhanced local greenhouse effect. Knowing these details can improve the effectiveness of predicting the urban heat island.

Despite being an important predictor in previous studies, wind speed did not prove to be an effective predictor of the UHI for Charlotte, NC. Table 7 shows a cross correlation of wind speed with the other predictors used in this study. While wind speed is not a top five predictor for any BRF model, wind speed is highest correlated with other top predictors. Predictors highest correlated with wind speed include temperature, vapor pressure, weather factor, turner stability index, and solar radiation. The strong performance of these predictors in model performance could be the reason wind speed was not chosen as a top five predictor.

Future work would include expanding the dataset for model development and testing. For seasonality and weekends, model performance and accuracy would improve with a larger testing dataset. Model performance could be tested using predictors from two days prior instead of the previous day to test if that would result in an effective model.

TABLES

Table 1. Surface weather station metadata, including latitude, longitude, elevation, local topographic relief, land use/cover characteristics, mean daily temperature (T_{DAY}), mean daily temperature range (ΔT_{DAY}), sky view factor (SVF), and the local climate zone (LCZ). This table is from Eastin et al. (2018).

Station			Lat (°N)	Lon (°W)	Elev (m)	Relief ^a (m)	Developed ^b total (%)	Famland ^b total (%)	Forested ^b total (%)	Water ^b total (%)	Mean T_{DAY} (°C)	Mean ΔT_{DAY} (°C)	SVF ^c	LCZ ^d
Name	Identifier	Type												
Garinger	GAR	MCAQ	35.2401	-80.7857	232.0	8.2	95.1	1.3	3.1	0.5	16.30	10.56	0.882	6B
Charlotte	CLT	ASOS	35.2241	-80.9543	221.9	3.9	82.4	7.7	8.9	1.0	15.94	11.05	0.927	8C
Rock Hill	RKH	ASOS	34.9836	-81.0562	203.9	4.7	69.0	9.5	21.4	0.1	15.62	11.18	0.960	9D
Gastonia	GAS	ASOS	35.1968	-81.1555	242.9	5.8	51.4	10.9	36.6	1.1	15.54	11.32	0.873	9B
Salisbury	SAL	AWOS-IHP	35.6442	-80.5235	235.6	9.6	45.6	12.2	40.1	2.1	15.34	11.84	0.905	9B
Concord	CON	AWOS-IHP	35.3896	-80.7074	214.9	8.0	43.9	32.8	21.4	1.9	15.70	11.21	0.937	D9
Monroe	MON	ASOS	35.0147	-80.6242	207.0	6.7	36.4	38.3	24.7	0.6	15.79	11.10	0.926	D9
Statesville	STV	AWOS-IHP	35.7635	-80.9472	294.1	12.9	27.1	40.9	29.7	2.3	15.28	12.13	0.923	D9
Chester	CHR	AWOS-IHP	34.7904	-81.2007	200.3	5.4	20.1	26.1	52.4	1.4	15.18	11.72	0.892	B9
Lincolnton	LIN	AWOS-IHP	35.4862	-81.1565	266.7	-2.4	19.7	31.5	48.2	0.6	14.81	12.70	0.992	B9
Lancaster	LAN	AWOS-IHP	34.7266	-80.8531	148.1	-5.6	18.5	10.9	68.2	2.4	14.58	12.80	0.987	B9
Wadesboro	WAD	AWOS-IHP	35.0194	-80.0783	91.1	3.8	14.8	33.8	50.1	1.3	14.73	13.04	0.996	B9

^a Topographic relief is the difference between station elevation and mean elevation within a 1.5-km radius.

^b LULC percentages were determined within a 1.5-km radius of each station.

^c SVFs were determined within a 500-m radius of each station following L. Chen et al. (2012).

^d LCZs were determined following Stewart and Oke (2012)—the classification type listed first is primary.

Table 2. The correlation coefficients between each 15 LST meteorological parameters and the 21-23 average UHI intensity.

Predictor	UHI Mean
ATEMP	-0.0803
PRVAP	-0.357
PRSTN	0.286381
SOLRAD	0.290942
SKYCOV	-0.40454
CLDHGT	0.411627
WXFRAC	0.424822
TRNSTAB	-0.2704
Log[RAIN]	-0.14203
Log[RELHM]	-0.50654
Log[WINDDIR]	-0.09272
Log[WINDSPD]	-0.09272

Table 3. Results of initial Bootstrap Random Forest model.

Overall Statistics			
Individual Trees	RASE		
In Bag	0.856045		
Out of Bag	1.916325		
	RSquare	RASE	N
Training	0.713	1.0473665	1027
Validation	0.392	1.582881	342
Test	0.334	1.4776361	338

Table 4. Results of bootstrap random forest model using the GRM optimal predictors.

Overall Statistics			
Individual Trees	RASE		
In Bag	1.347839		
Out of Bag	1.807635		
	RSquare	RASE	N
Training	0.287	1.649475	1027
Validation	0.258	1.7482466	342
Test	0.246	1.5721165	338

Table 5. Results of bootstrap random forest model using the 5 predictors highest correlated to the mean UHI Intensity.

Overall Statistics			
Individual Trees	RASE		
In Bag	1.298898		
Out of Bag	1.738422		
	RSquare	RASE	N
Training	0.361	1.5621519	1027
Validation	0.284	1.7176523	342
Test	0.294	1.5212802	338

Table 6: Statistical Results of all model run performance.

Model	Mean Absolute Error	Root Mean Squared Error	Mean Percent Error	Mean Bias	Average UHI Intensity	R-Squared
All Data	0.99	1.30	56.21	0.14	0.35	0.46
Winter	1.09	1.43	21.33	0.14	0.31	0.6
Spring	1.28	1.67	146.03	0.55	0.47	0.42
Summer	0.82	1.07	41.42	-0.07	0.42	0.21
Fall	1.1	1.45	48.07	0	0.22	0.42
Weekend	1.13	1.47	55.26	0.03	0.33	0.28
Weekday	1.02	1.33	8.64	0.12	0.36	0.38

Table 7: Cross correlation analysis of wind speed with the other predictors.

Cross Correlation with Wind Speed (m/s)	
Temperature (C)	-0.31
Precipitation (mm)	0
Relative Humidity (%)	0.01
Vapor Pressure (mb)	-0.28
Station Pressure (mb)	-0.04
Solar Radiation (W/m ²)	-0.18
Sky Cover (Frac)	0.05
Cloud Height (m)	-0.08
Weather Factor (frac)	-0.28
Turner Stability Index	0.34
Wind Direction (deg)	-0.03

FIGURES

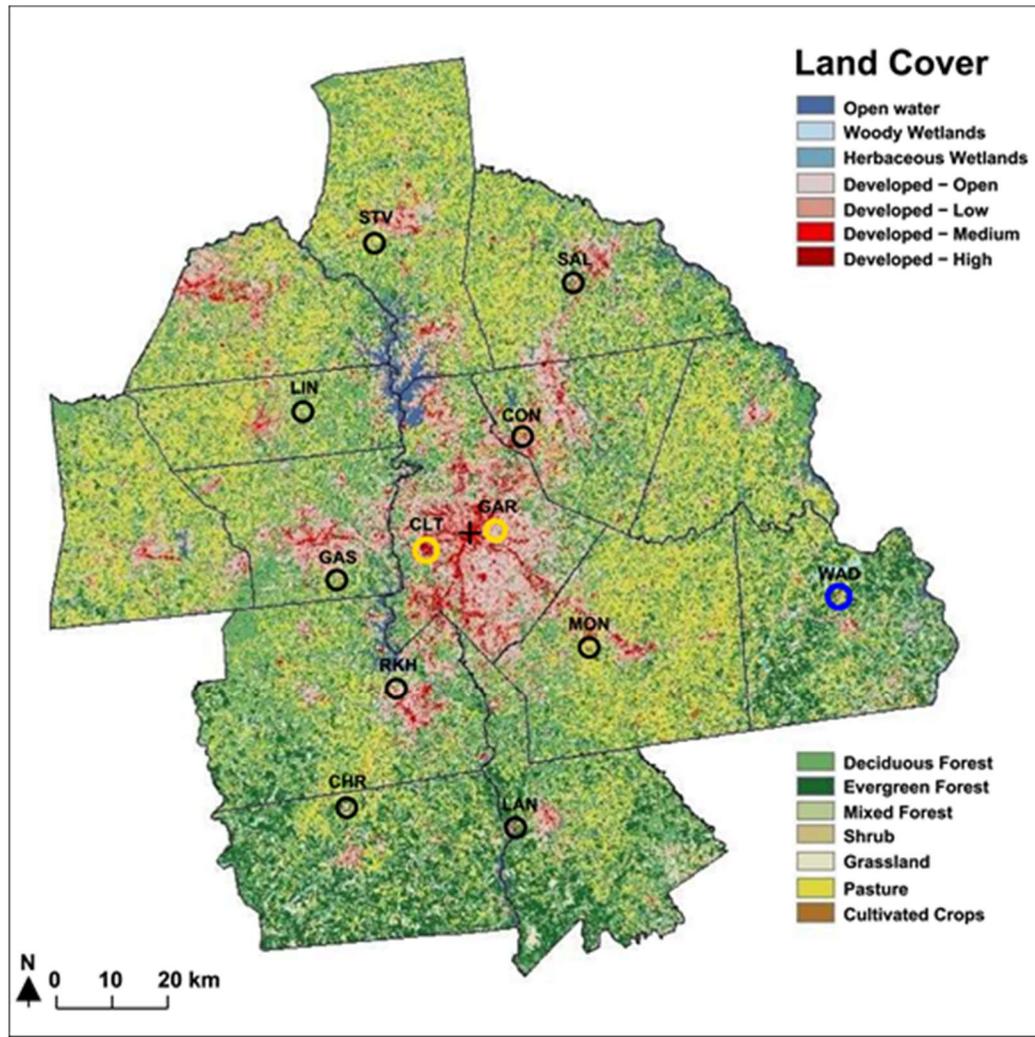


Figure 1. Surface weather station locations superimposed on the 2011 USGS LULC product across the CMR. The blue circle denotes the rural reference site (WAD), the yellow circles denote the most urbanized sites (GAR and CLT) discussed in this study, and the plus sign denotes the central business district of the urban core. See Table 1 for station metadata and local LULC characteristics.

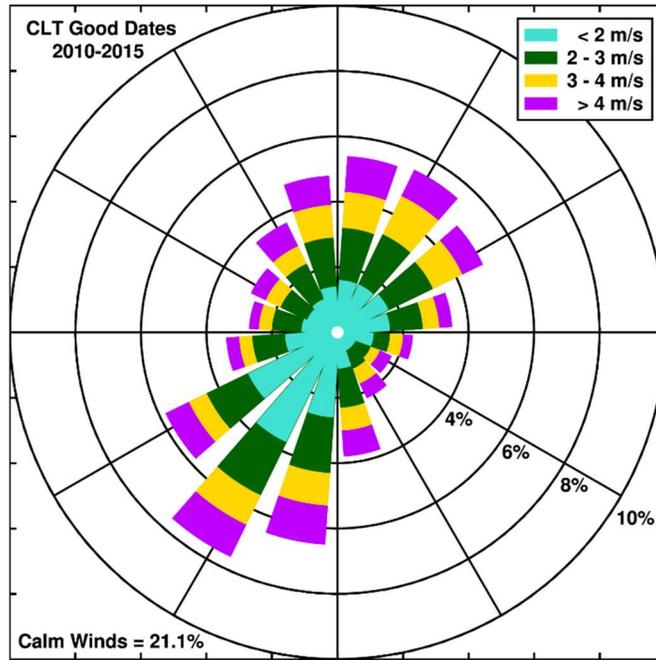


Figure 2: Wind rose of hourly observations at CLT for all good dates between June 2010 and May 2015.

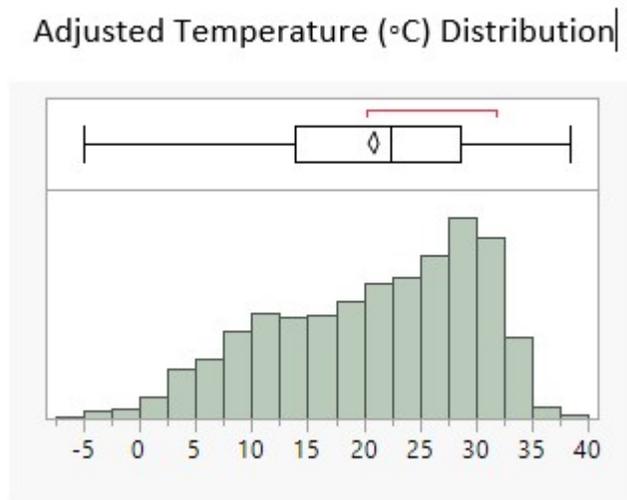


Figure 3: Distribution of the Adjusted Temperature (°C) between June 2010 and May 2015.

Precipitation (mm) Distribution

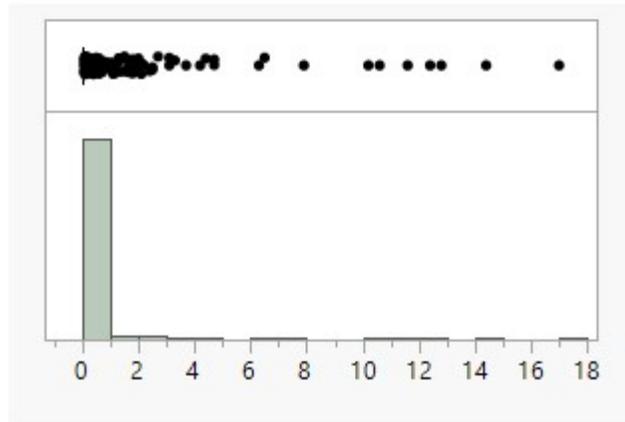


Figure 4: Distribution of the Precipitation (mm) between June 2010 and May 2015.

Relative Humidity (%) Distribution

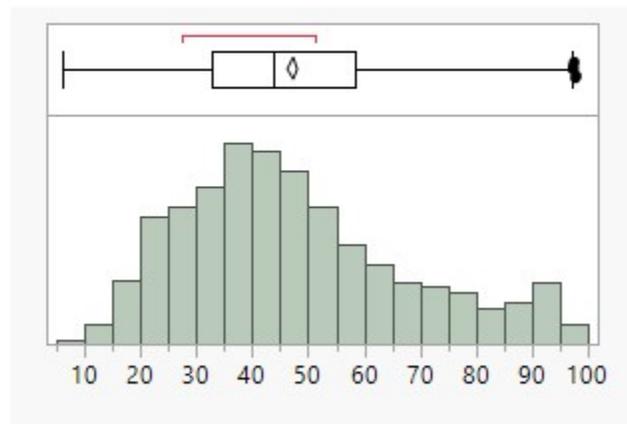


Figure 5: Distribution of the Relative Humidity (%) between June 2010 and May 2015.

Vapor Pressure (mb) Distribution

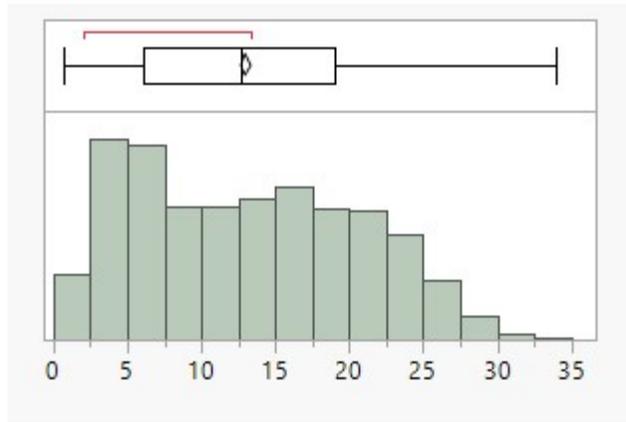


Figure 6: Distribution of the Vapor Pressure (mb) between June 2010 and May 2015.

Station Pressure (mb) Distribution

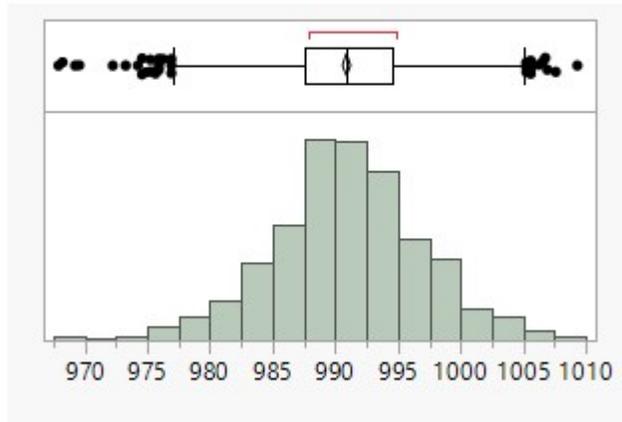


Figure 7: Distribution of the Station Pressure (mb) between June 2010 and May 2015.

Wind Speed (m/s) Distribution

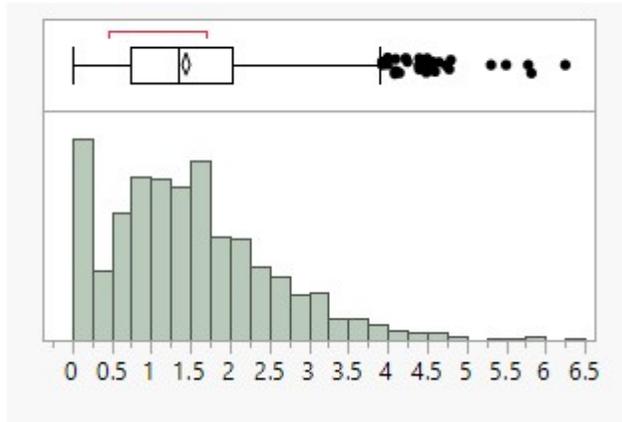


Figure 8: Distribution of the Wind Speed (m/s) between June 2010 and May 2015.

Wind Direction (deg) Distribution

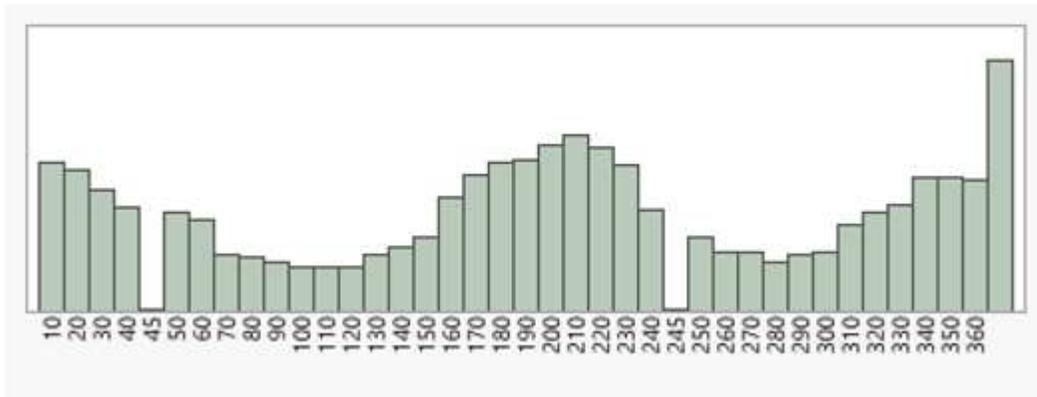


Figure 9: Distribution of the Wind Direction (deg) between June 2010 and May 2015.

Solar Radiation (W/m2) Distribution

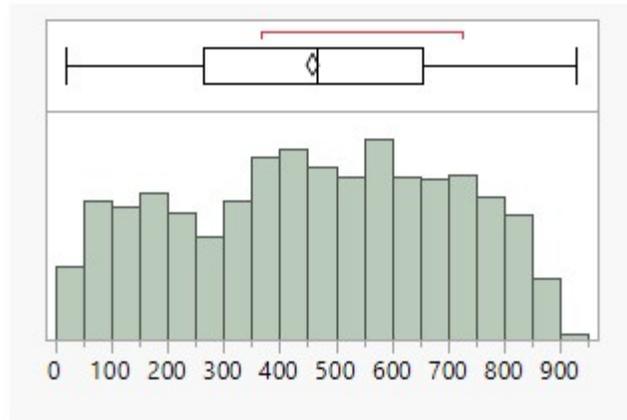


Figure 10: Distribution of the Solar Radiation (W/m2) between June 2010 and May 2015.

Sky Cover (frac) Distribution

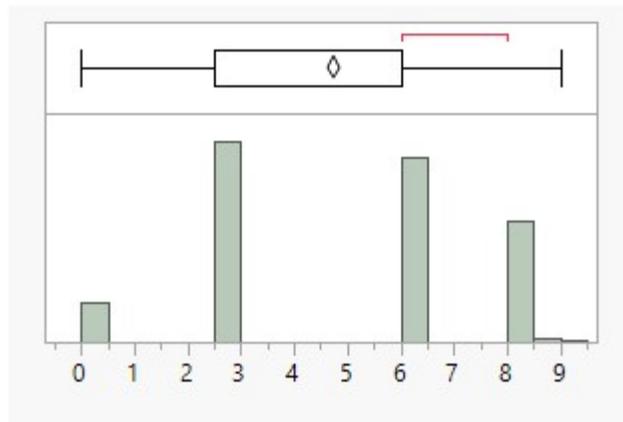


Figure 11: Distribution of the Sky Cover (frac) between June 2010 and May 2015.

Cloud Height (m) Distribution

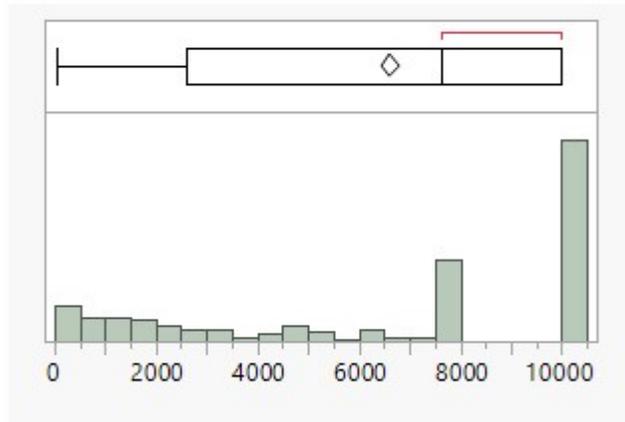


Figure 12: Distribution of the Cloud Height (m) between June 2010 and May 2015.

Weather Factor (frac) Distribution

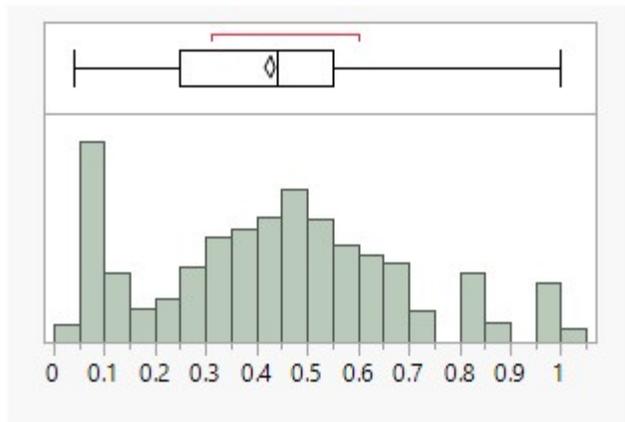


Figure 13: Distribution of the Weather Factor between June 2010 and May 2015.

Turner Stability Index Distribution

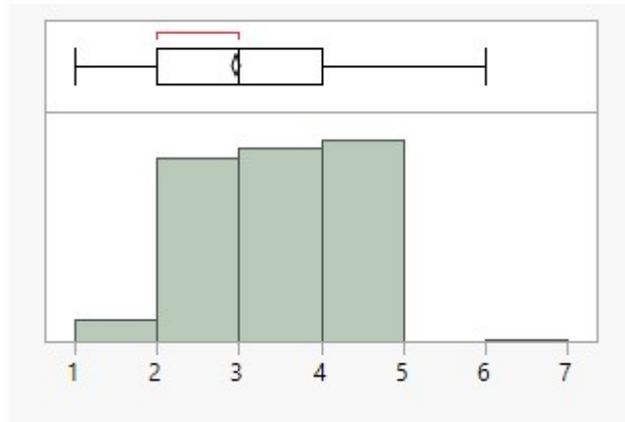


Figure 14: Distribution of the Turner Stability Index between June 2010 and May 2015.

UHI (°C) Distribution

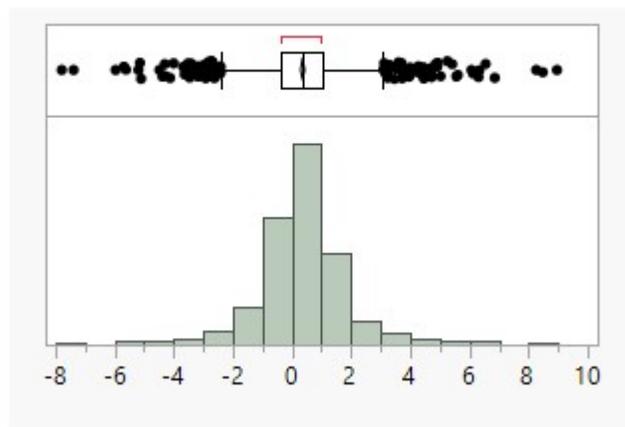


Figure 15: Distribution of the Urban Heat Island Intensity (°C) between June 2010 and May 2015.

Actual by Predicted Plot

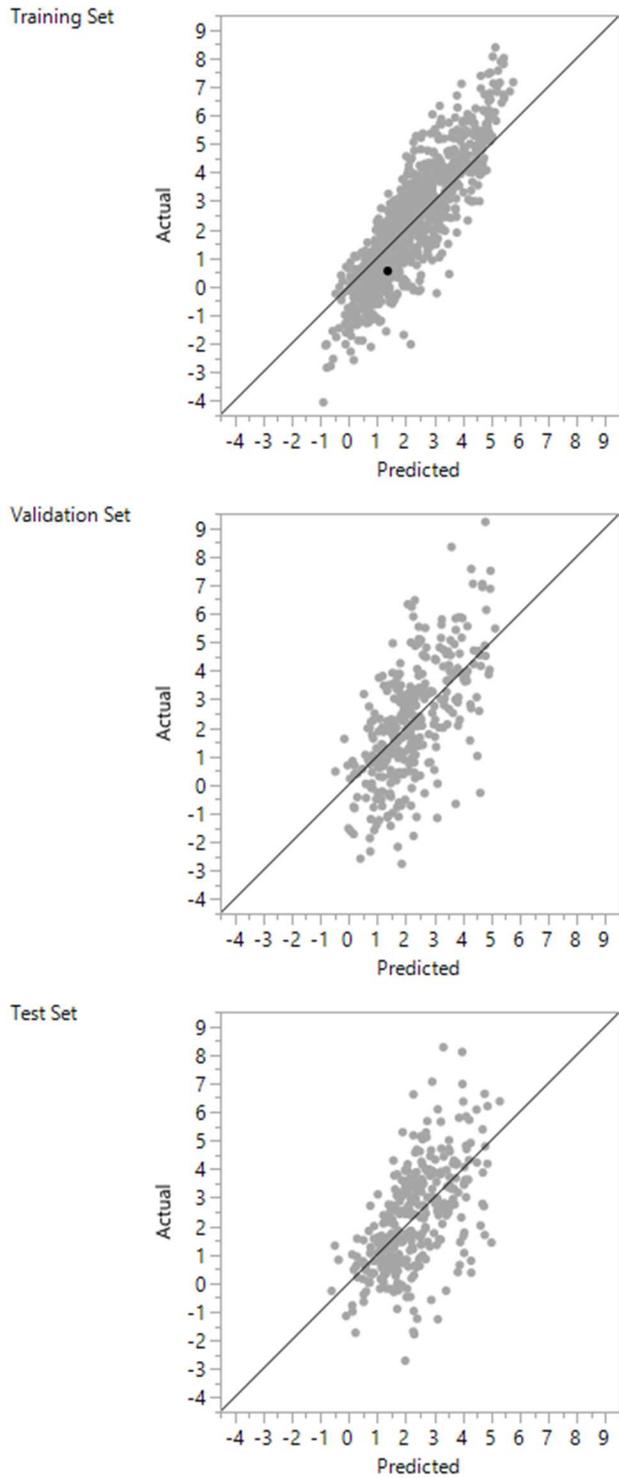


Figure 16. Scatterplots of observed (actual) UHI intensity vs the predicted UHI intensity from the initial BRF model fit for the training, validation, and testing subsets.

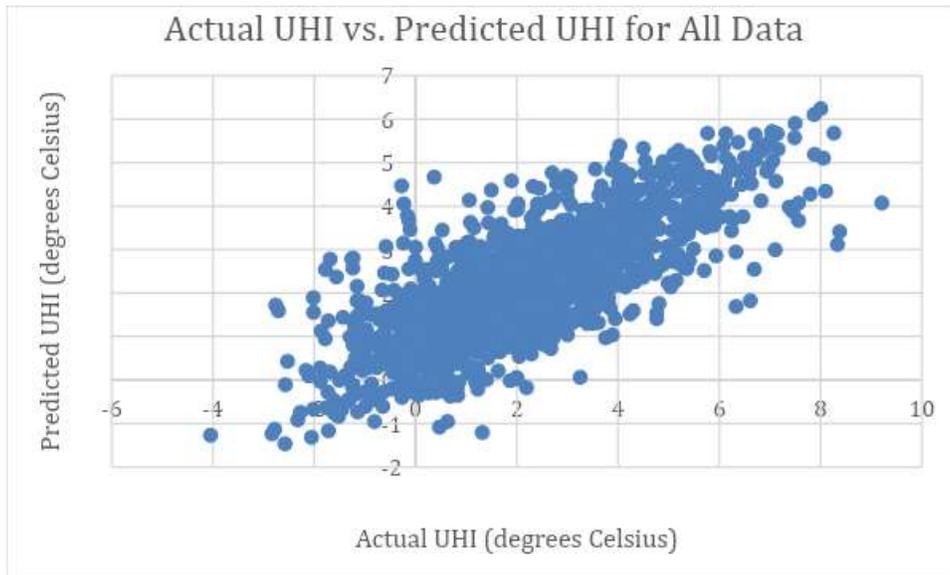


Figure 17: Actual UHI plotted with Predicted UHI for All Data.

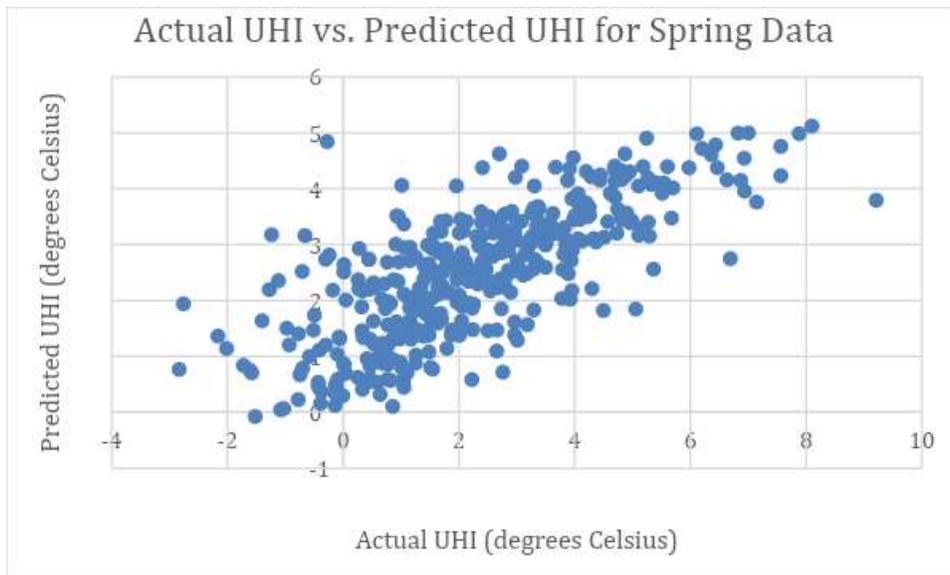


Figure 18: Actual UHI plotted with Predicted UHI for spring.

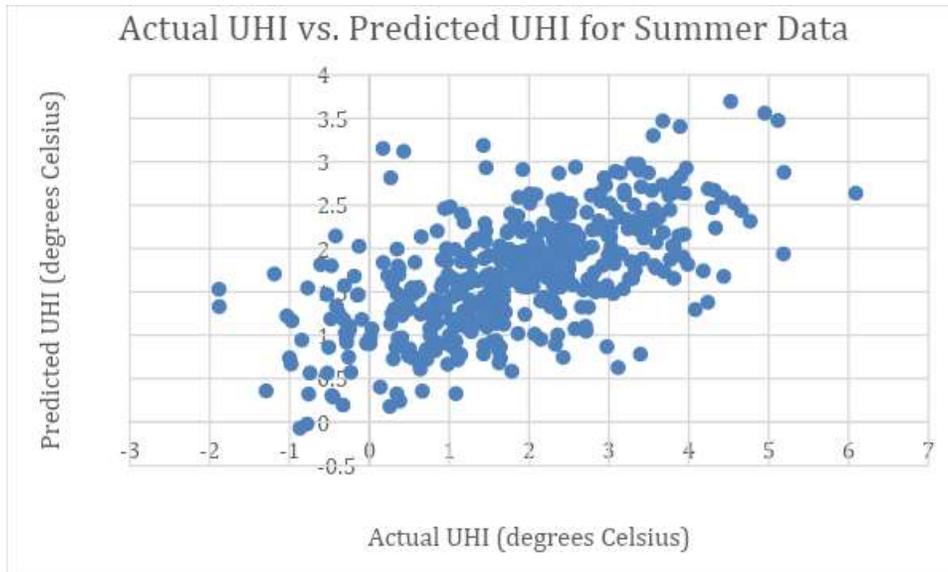


Figure 19: Actual UHI plotted with Predicted UHI for summer.

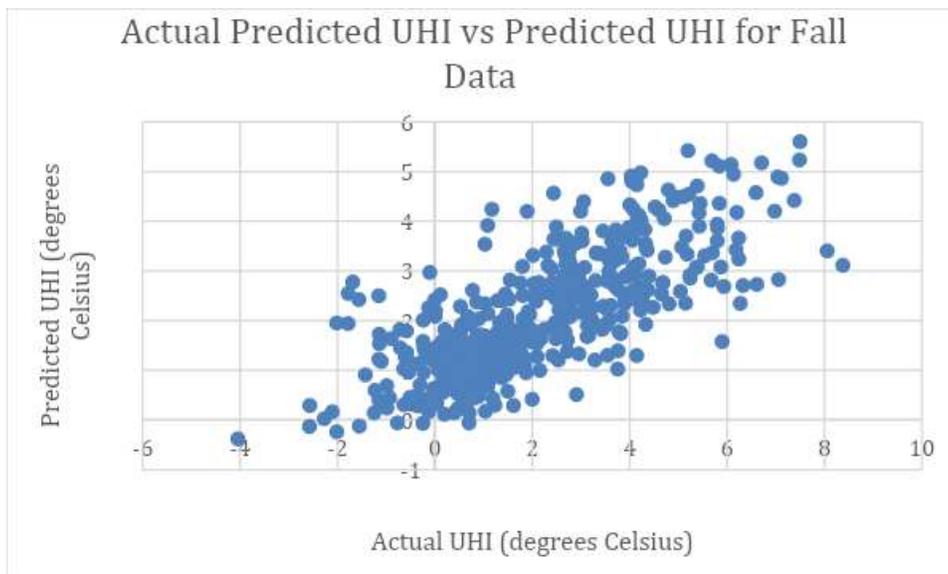


Figure 20: Actual UHI plotted with Predicted UHI for fall.

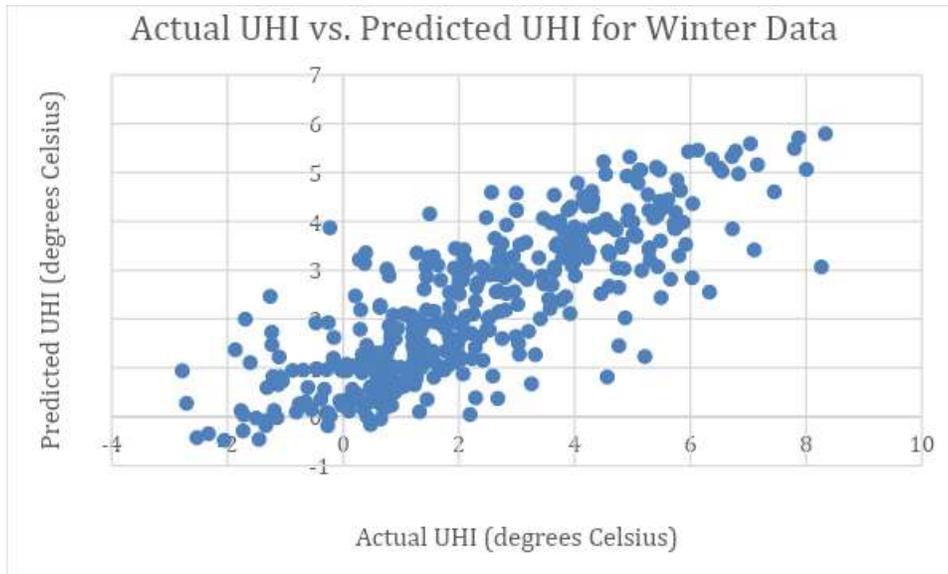


Figure 21: Actual UHI plotted with Predicted UHI for winter.

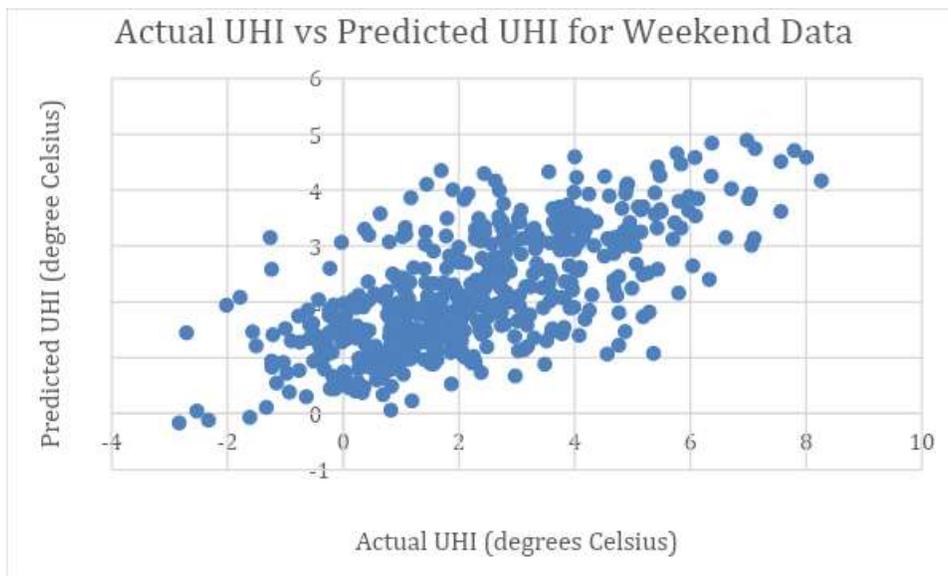


Figure 22: Actual UHI plotted with Predicted UHI for weekend.

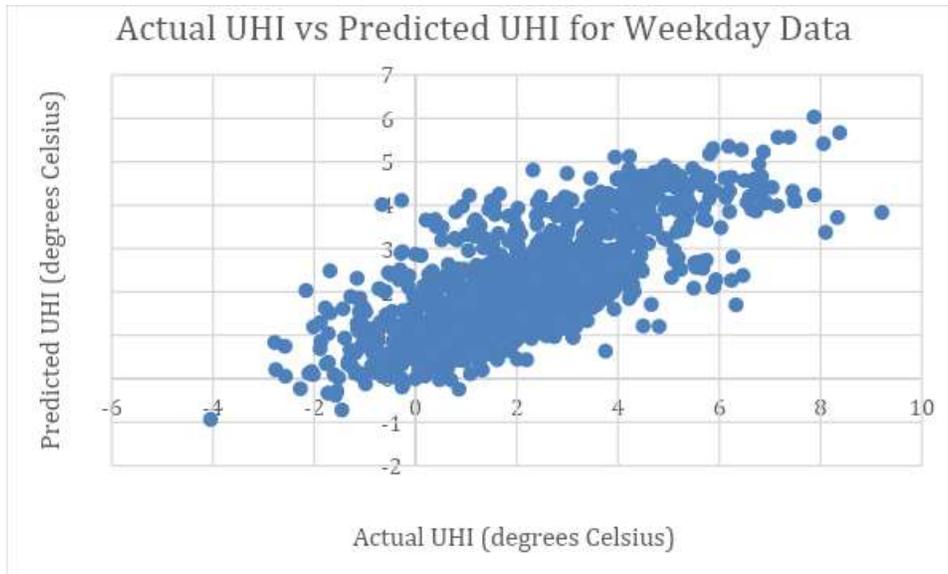


Figure 23: Actual UHI plotted with Predicted UHI for weekday.

REFERENCES

- Best, M. J. (2006). Progress towards better weather forecasts for city dwellers: From short range to climate change. *Theoretical and Applied Climatology*, 84(1), 47–55. <https://doi.org/10.1007/s00704-005-0143-2>
- Bottyán, Z., & Unger, J. (2003). A multiple linear statistical model for estimating the mean maximum urban heat island. *Theoretical and Applied Climatology*, 75(3–4), 233–243. <https://doi.org/10.1007/s00704-003-0735-7>
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123–140. <https://doi.org/10.1007/BF00058655>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1983). *Classification and Regression Trees*. <https://doi.org/10.2307/2530946>
- Eastin, M. D., Baber, M., Boucher, A., Di Bari, S., Hubler, R., Stimac-Spalding, B., & Winesett, T. (2018). Temporal Variability of the Charlotte (Sub)Urban Heat Island. *Journal of Applied Meteorology and Climatology*, 57(1), 81–102. <https://doi.org/10.1175/JAMC-D-17-0099.1>
- Eliasson, I. (1996). Intra-urban nocturnal temperature differences: A multivariate approach. *Climate Research*, 7, 21–30. <https://doi.org/10.3354/cr007021>
- Grimmond, C. S. B. (2006). Progress in measuring and observing the urban atmosphere. *Theoretical and Applied Climatology*, 84(1), 3–22. <https://doi.org/10.1007/s00704-005-0140-5>
- Hoffmann, P., Krueger, O., & Schlünzen, K. H. (2012). A statistical model for the urban heat island and its application to a climate change scenario. *International Journal of Climatology*, 32(8), 1238–1248. <https://doi.org/10.1002/joc.2348>
- Holmer, B., Thorsson, S., & Lindén, J. (2013). Evening evapotranspirative cooling in relation to vegetation and urban geometry in the city of Ouagadougou, Burkina Faso. *International Journal of Climatology*, 33(15), 3089–3105. <https://doi.org/10.1002/joc.3561>
- Johnson, G., Oke, T., Lyons, T., Steyn, D., Watson, I., & Voogt, J. (1991). Simulation of surface urban heat islands under 'IDEAL' conditions at night part 1: Theory and tests against field data. *Boundary-Layer Meteorology*, 56(3), 275–294. <https://doi.org/10.1007/BF00120424>
- Kim, Y., & Baik, J. (2002). Maximum Urban Heat Island Intensity in Seoul. *Journal of Applied Meteorology*, 41(6), 651–659. [https://doi.org/10.1175/1520-0450\(2002\)041<0651:MUHIII>2.0.CO;2](https://doi.org/10.1175/1520-0450(2002)041<0651:MUHIII>2.0.CO;2)
- Lemonsu, A., & Masson, V. (2002). Simulation of a Summer Urban Breeze Over Paris. *Boundary-Layer Meteorology*, 104(3), 463–490. <https://doi.org/10.1023/A:1016509614936>

- Lindén, J., Esper, J., & Holmer, B. (2015). Using Land Cover, Population, and Night Light Data for Assessing Local Temperature Differences in Mainz, Germany. *Journal of Applied Meteorology and Climatology*, 54(3), 658–670. <https://doi.org/10.1175/JAMC-D-14-0124.1>
- McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models* (2nd ed.). Springer US. <https://doi.org/10.1007/978-1-4899-3242-6>
- Mills, G. (2007). Cities as agents of global change. *International Journal of Climatology*, 27(14), 1849–1857. <https://doi.org/10.1002/joc.1604>
- Morris, C. J. G., Simmonds, I., & Plummer, N. (2001). Quantification of the Influences of Wind and Cloud on the Nocturnal Urban Heat Island of a Large City. *Journal of Applied Meteorology*, 40(2), 169–182. [https://doi.org/10.1175/1520-0450\(2001\)040<0169:QOTIOW>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0169:QOTIOW>2.0.CO;2)
- Oke, T. R., Johnson, G. T., Steyn, D. G., & Watson, I. D. (1991). Simulation of surface urban heat islands under 'ideal' conditions at night part 2: Diagnosis of causation. *Boundary-Layer Meteorology*, 56(4), 339–358. <https://doi.org/10.1007/BF00119211>
- Peterson, T. C. (2003). Assessment of Urban Versus Rural In Situ Surface Temperatures in the Contiguous United States: No Difference Found. *Journal of Climate*, 16(18), 2941–2959. [https://doi.org/10.1175/1520-0442\(2003\)016<2941:AOUVRI>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<2941:AOUVRI>2.0.CO;2)
- Runnalls, K. E., & Oke, T. R. (2000). Dynamics and Controls of the Near-Surface Heat Island of Vancouver, British Columbia. *Physical Geography*, 21(4), 283–304. <https://doi.org/10.1080/02723646.2000.10642711>
- Schatz, J., & Kucharik, C. (2014). Seasonality of the Urban Heat Island Effect in Madison, Wisconsin. *Journal of Applied Meteorology and Climatology*, 53(10), 2371–2386. <https://doi.org/10.1175/JAMC-D-14-0107.1>
- Sözen, A., Arcaklioğlu, E., & Özalp, M. (2004). Estimation of solar potential in Turkey by artificial neural networks using meteorological and geographical data. *Energy Conversion and Management*, 45(18), 3033–3052. <https://doi.org/10.1016/j.enconman.2003.12.020>
- Steenefeld, G. J., Koopmans, S., Heusinkveld, B. G., Hove, L. W. A. van, & Holtslag, A. a. M. (2011). Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in the Netherlands. *Journal of Geophysical Research: Atmospheres*, 116(D20). <https://doi.org/10.1029/2011JD015988>
- Turner, D. B. (1964). A Diffusion Model for an Urban Area. *Journal of Applied Meteorology*, 3(1), 83–91. [https://doi.org/10.1175/1520-0450\(1964\)003<0083:ADMFAU>2.0.CO;2](https://doi.org/10.1175/1520-0450(1964)003<0083:ADMFAU>2.0.CO;2)
- Winkler, J. A., Skaggs, R. H., & Baker, D. G. (1981). Effect of Temperature Adjustments on the Minneapolis-St. Paul Urban Heat Island. *Journal of Applied Meteorology*, 20(11), 1295–1300. [https://doi.org/10.1175/1520-0450\(1981\)020<1295:EOTAOT>2.0.CO;2](https://doi.org/10.1175/1520-0450(1981)020<1295:EOTAOT>2.0.CO;2)
- Yang, Z., Dominguez, F., Gupta, H., Zeng, X., & Norman, L. (2015). Urban Effects on Regional Climate: A Case Study in the Phoenix and Tucson “Sun Corridor.” *Earth Interactions*, 20(20), 1–25. <https://doi.org/10.1175/EI-D-15-0027.1>