Computers, Environment and Urban Systems 84 (2020) 101538

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Urban ambient air temperature estimation using hyperlocal data from smart vehicle-borne sensors

Yanzhe Yin[a,*], Navid Hashemi Tonekaboni[b], Andrew Grundstein[a], Deepak R. Mishra[a], Lakshmish Ramaswamy[b], John Dowd[c]

[a] Department of Geography, University of Georgia, Athens, GA 30602, USA
[b] Department of Computer Science, University of Georgia, Athens, GA 30602, USA
[c] Department of Geology, University of Georgia, Athens, GA 30602, USA

ARTICLE INFO

Keywords:
Human thermal comfort
Mobile sensing
Machine learning
Urban heat hazards
Smart and connected communities

ABSTRACT

High-quality temperature data at a finer spatio-temporal scale is critical for analyzing the risk of heat exposure and hazards in urban environments. The variability of urban landscapes makes cities a challenging environment for quantifying heat exposure. Most of the existing heat hazard studies have inherent limitations on two fronts; first, the spatio-temporal granularities are too coarse, and second, the inability to track the ambient air temperature (AAT) instead of land surface temperature (LST). Overcoming these limitations requires developing models for mapping the variability in heat exposure in urban environments. We investigated an integrated approach for mapping urban heat hazards by harnessing a diverse set of high-resolution measurements, including both ground-based and satellite-based temperature data. We mounted vehicle-borne mobile sensors on city buses to collect high-frequency temperature data throughout 2018 and 2019. Our research also incorporated key biophysical parameters and Landsat 8 LST data into Random Forest regression modeling to map the hyperlocal variability of heat hazard over areas not covered by the buses. The vehicle-borne temperature sensor data showed large temperature differences within the city, with the largest variations of up to 10 °C and morning-afternoon diurnal changes at a magnitude around 20 °C. Random Forest modeling on noontime (11:30 am – 12:30 pm) data to predict AAT produced accurate results with a mean absolute error of 0.29 °C and successfully showcased the enhanced granularity in urban heat hazard mapping. These maps revealed well-defined hyperlocal variabilities in AAT, which were not evident with other research approaches. Urban core and dense residential areas revealed larger than 5 °C AAT differences from their nearby green spaces. The sensing framework developed in this study can be easily implemented in other urban areas, and findings from this study will be beneficial in understanding the heat vulnerabilities of individual communities. It can be used by the local government to devise targeted hazard mitigation efforts such as increasing green space, developing better heat-safety policies, and exposure warning for workers.

1. Introduction

Extreme heat events associated with Urban Heat Islands (UHI) pose a serious public health risk in many urban areas in the U.S.(Harlan, Declet-Barreto, Stefanov, & Petitti, 2013). Climate change is likely to further exacerbate the dangers posed by UHI in cities across the U.S. and around the world, particularly in poorly planned communities (Kovats R. Sari & Koppe, 2006; Brooke Anderson, Bell, Brooke Anderson, & Bell, 2011). Simply put, the UHI is the excessive generation and retention of heat by the built environments within urban areas. Variables contributing to a city's UHI include the increased surface area of buildings, less greenspace and surface moisture, higher heat capacities of building materials, and increased exhaust from structures and automobiles (Ihara, Kikegawa, Asahi, Genchi, & Kondo, 2008; Kuttler, 2008). Buildings absorb more heat and retain more longwave radiation within the urban canyons. Values from 2009 from the U.S. Census Bureau identified 940 metropolitan and micropolitan statistical areas, and with continued warming of the climate, large numbers of communities and individuals in these towns and cities will be exposed to frequent extreme heat events. To protect these communities from extreme heat-related hazards, it is imperative that we comprehensively track and understand the within-UHI heat exposure variabilities of

* Corresponding author.
E-mail address: yanzhe.yin@uga.edu (Y. Yin).

https://doi.org/10.1016/j.compenvurbsys.2020.101538
Received 16 March 2020; Received in revised form 24 July 2020; Accepted 31 July 2020

various at-risk communities.

The largest and most broadly used resource is heat advisory and warnings from the National Weather Service (NWS) and the National Oceanic and Atmospheric Administration (NOAA). These services cover all the states and cities across the U.S., with hourly weather updates from ground-based observations and model output to each municipal city and county. The only limitation is their minimum unit for a report is city-level or county-level, which does not represent individual communities (Maier et al., 2014). Though all communities are and will continue being under heat risk, they do not share the same living environment, and thus, city/county-level reports do not adequately delineate specific heat hazard for each community.

Distinct from weather station data, satellite land surface temperature (LST) products are useful observations of intra-city variations in urban heat. LST is traditionally used as a proxy to delineate UHIs. With the development of more accurate and well-calibrated spaceborne thermal sensors such as NASA's ECOSTRESS, to provide LST images with better quality, the spatial resolution of LST images could reach 70 m without resampling, and the frequency is one to two images a week. Such spatial resolution is sufficient to provide details about which section of the city is hot and what is the intensity of the UHIs.

To date, urban temperature mapping has mostly focused on determining UHIs and environmental exposure through coarse resolution heat maps of cities using satellite (Harlan et al., 2013; Johnson & Wilson, 2009) or weather station data (Maier et al., 2014). However, within a known UHI, people are exposed to widely varying temperatures due to different surrounding environments (Harlan, Brazel, Prashad, Stefanov, & Larsen, 2006). Such variations in the environment bring many complexities to heat study at a granularity smaller than an urban block. The above studies suffer from a lack of high-resolution information needed to differentiate the effects of buildings and vegetation on the temperatures within and between neighborhoods, which ultimately influence hyperlocal urban ambient air temperature (AAT). Also, satellites lack the spatial and temporal resolution to track the actual Urban Heat Exposure (UHE) of individuals and communities over time. Most satellites only visit the same geographic location once every few days, and such a dataset cannot describe diurnal and daily variations.

Furthermore, satellites do not directly provide AAT, the heat conditions that people experience on the ground, which most closely tied to human thermal comfort. Satellite LST products represent surface temperature instead of near-surface AAT. Although LSTs have been used as a proxy for air temperatures, they differ in magnitude that LSTs are hotter in the day but relatively cooler at night (Oke, 1988). Also, LST and AAT are not perfectly correlated in various conditions (Ho, Knudby, Xu, Hodul, & Aminipouri, 2016; Kloog, Chudnovsky, Koutrakis, & Schwartz, 2012).

A significant advancement in human heat hazard monitoring will be to develop a product that captures the complex AAT variability in urban environments, identified as known-UHIs, at a highly granular scale. For that, we need temperature data with high spatio-temporal resolution and coverage. An exciting development in collecting such data is the increasing use of mobile sensors (e.g., vehicle-borne, human-borne), which can gather high resolution spatiotemporal data on ambient atmospheric conditions (Anjomshoaa et al., 2018; Eisenman, Lane, & Miluzzo, 2006; Honicky, Brewer, Paulos, & White, 2008). Indeed, many European cities have attempted to harness mobile sensing networks to improve AAT measurements (Erman, Van Hoesel, Havinga, & Jian, 2008; Hasenfratz et al., 2015; Overeem et al., 2013). Unfortunately, the widespread use of these sensors is limited due to the cost and the labor-intensive nature of data collection and processing requirements (Reis et al., 2015). Also, such sensor networks still have a limitation on spatial coverage because the data are mainly point or line data, unlike the satellite LST datasets, which can cover a large region.

To leverage the spatial continuity of satellite-derived products, some studies have established simple statistical relationships between LST and AAT (Nichol & Wong, 2008; Vancutsem, Ceccato, Dinku, & Connor, 2010; Zhu, Lu, & Jia, 2013). These relationships have only moderate explanatory power, in part because the near-surface urban area is a complex environment with interfering variables that alter the relationship between AAT and LST. These studies do not incorporate land cover and canopy conditions and shading among other biophysical parameters that may improve the relationship between LST and AAT (Frey, Rigo, & Parlow, 2007; Li et al., 2011; Yuan & Bauer, 2007). These environmental factors have rarely been used in an integrated model to estimate AAT. Further, these studies have not addressed the high-resolution AAT variabilities within known UHIs.

Our study seeks to blend the strengths of remotely sensed data with its contiguous spatial coverage and the advantages of data collected from vehicle-borne mobile sensors with its capability of capturing highly granular AAT data through machine learning-based modeling to produce the first satellite-derived maps of AAT or exposure temperature. We present an innovative approach towards building AAT products with the ability to create a more robust conversion from LST to AAT by linking with biophysical and meteorological parameters. Human heat hazards, exacerbated by climate change, often disproportionally affect vulnerable urban communities. The first step towards resolving heat hazards is to understand heat distribution. Therefore, developing a framework to generate AAT or heat exposure maps would be extremely valuable in reducing heat hazards and enhancing urban resiliency.

2. Data and method

2.1. Vehicle borne sensor implementation

Vehicle-borne sensors in this study were designed and assembled to provide high temporal resolution AAT data along with location and time (Fig. 1). These sensor packages were assembled from scratch with microprocessor components. The cost of assembling one unit is about $100 USD. Each sensor includes an AdFruit M0 Feather microprocessor, a real-time clock, a GPS unit, a DS18B20 waterproof temperature sensor, a battery set, a micro S.D. card, and boards and wires that connect the components. IP67 waterproof Polycase boxes were used as the enclosure for the sensor assembly and provided an easy mounting opportunity on different kinds of surfaces on the vehicle. The sensor head is the only thing outside the box, and it is shielded by a white hollow plastic ball that ensures sufficient airflow but allows no direct solar radiation.

The sensor units ran Arduino ‘software serial’ libraries to control and communicate between the central processor and different components, the temperature sensor, real-time clock, GPS unit, data storage, and warning report. The GPS unit in the sensors was used for time adjustment and route identification. The real-time GPS data were retrieved from bus transit providers and their provided APIs which are operated by private businesses contracted through the University of Georgia (UGA). We have already developed architecture to filter and then upload the data to a cloud-based database as well to maintain the integrity and quality of the data. More details about the sensor component, assembly procedure, and software architecture are given in (N. H. Tonekaboni et al., 2018) and (Tonekaboni, 2019).

The vehicle-borne sensors shown in Fig. 1 were mounted on thirty UGA and Athens Transit buses. We visited the bus transit parking lots every month to replace the batteries, check the sensors, and access the data from the logger. The total cost for maintenance per month is less than $1 USD per unit. The buses traverse the Athens area from morning through evening recording temperature every five seconds.

2.2. Vehicle-borne sensor data

In the data collection process, we used the vehicle-borne mobile temperature sensors described in the above section to collect high-

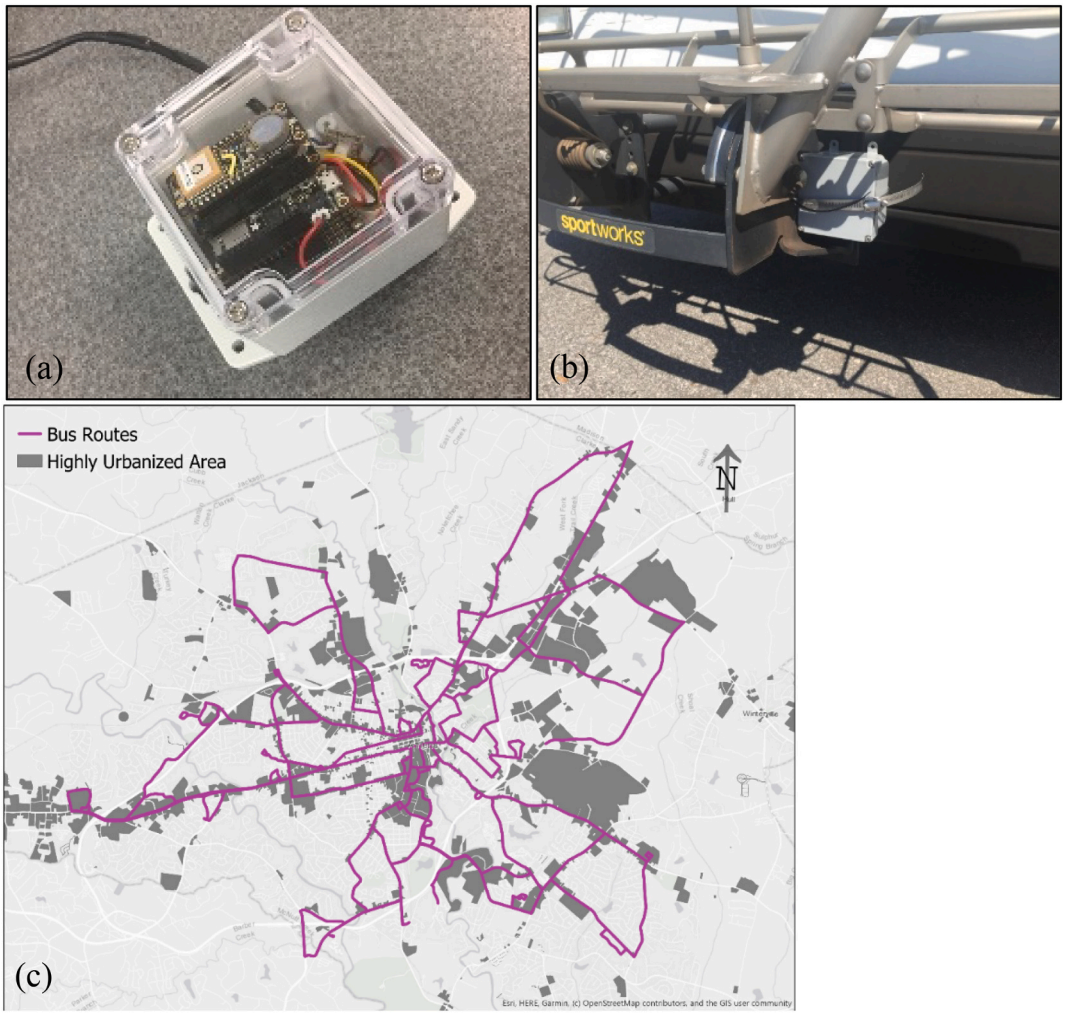

Fig. 1. Vehicle-borne sensor system, (a) Vehicle-borne sensor, (b) sensor mounted on bus, (c) bus routes in Athens, GA with highly urbanized areas such as malls, shoping plazas, urban core are combinmedly highlighted in dark gray.

resolution data within the greater Athens area. Athens, Georgia, is a town with a population of over 100,000. It is located in a humid subtropical climate in the southeastern U.S., characterized by long and hot summers transitioning into short and cold winters. Average monthly temperatures range from 6.4 °C in January to 27.0 °C in July; the maximum temperature on average reaches 32 °C or higher on 58 days annually (Arguez et al., 2010).

Sensors were mounted on more than thirty buses, which traversed the Athens area from 6 am to 9 pm, and recorded temperature every five seconds. We used these vehicle-borne sensors to collect AAT for six months (June–November) in 2018 and 6 months (May–October) in 2019. Data from the year of 2018 were used to calibrate the AAT prediction model, which also used LST and other influential biophysical parameters. Data from 2019 were used for model validation. Each year, within the 6-month deployment period, 90 million rows of raw AAT data (roughly 10 gigabytes) were collected from these vehicle-borne sensors before data cleaning and filtering. A quality-aware framework was used to filter erroneous data before being uploaded to the cloud database (see Tonekaboni et al., 2018 for details).

2.3. Remote sensing data

Several different satellite-based products were considered in this study. Landsat 8 Operational Land Imager (OLI) thermal infrared (TIR) data (100-m spatial resolution resampled to 30-m pixel size) for Athens, Georgia, were collected and processed for 2018 and 2019 from the USGS EarthExplorer website (https://earthexplorer.usgs.gov/). Multispectral PlanetScope products (Planet Team, 2017) at 3-m spatial resolution were acquired for 2018–2019 for developing high resolution Land Use Land Cover (LULC) maps. We performed supervised classification on four cloud-free PlanetScope images, two images from August 2018, and the other two from July, 2019. LiDAR data for Athens-Clarke County (ACC) collected by NOAA in 2015 were downloaded and preprocessed with the permission from ACC Geographic Information Service (https://www.athensclarkecounty.com/7632/Geospatial-Information-Office). LiDAR data were used to retrieve high-resolution canopy cover map for the study domain. Each dataset and associated products are explained in detail in the following subsections.

2.3.1. Land surface temperature (LST)

We calculated Landsat 8 LST from OLI TIR bands using the single-channel algorithm (Jiménez-Muñoz & Sobrino, 2010). There are two common types of algorithms to calculate LST from satellite thermal infrared images, Split-Window (S.W.) Algorithm (Qin, Dall'Olmo, Karnieli, & Berliner, 2001), which uses data from two thermal bands, and Single-Channel (S.C.) Algorithm, which uses only one thermal band data. Due to the reported data quality issue of Landsat 8 TIR band 11 (https://www.usgs.gov/land-resources/nli/landsat/landsat-8-oli-and-tirs-calibration-notices), we use the S.C. algorithm on Band 10 and avoided LST computation and comparison between the two algorithms. The basic theory of the S.C. algorithm is given in Eq. 1:

$$B(LST) = \frac{L_{sen} - L_{up} - \tau\cdot(1-\varepsilon)\cdot\frac{L_{down}}{\pi}}{\tau\cdot\varepsilon} \quad (1)$$

where B is the Planck function, L_{sen} is the radiance-at-the-sensor, L_{up} is the thermal path radiance, L_{down} is the downwelling irradiance, ε is the surface emissivity, and τ is the atmospheric transmissivity. Between June 2018 and November 2018, there were only three usable Landsat 8 cloud-free or minimally cloud-covered images that were used in LST estimation which are July 29th, August 14th, and August 30th.

2.3.2. Canopy cover

Canopy cover was derived from the LiDAR data by extracting the difference between the first and last returns of laser beams and co-analyzing the plant heights with LiDAR backscattered intensity data (MacFaden, O'Neil-Dunne, Royar, Lu, & Rundle, 2012). If an emitted laser beam hits a tree canopy, it is designed to produce multiple notable return pulses showing diffused reflection. This multiple-return characteristic of trees was combined with the extracted first-last return differences to identify the canopy and shading information. We imported 30 gigabytes of raw LiDAR data into ArcGIS and processed them into designated canopy cover with ArcGIS LAStools Toolbox (ESRI, Redlands, CA). The final output was a 3-m raster image with 0 and 1 values, where 0 means not covered by a canopy and 1 means covered by the canopy. The overall classification accuracy is 88.7% with a Kappa coefficient of 0.775. Next, we resampled the image to 30-m to match the pixel size of Landsat 8 LST products, with values range between 0 and 1 indicating the percentage of area in a 30-m pixel covered by the canopy.

2.3.3. LULC composition

We used a maximum-likelihood-logistic-regression-based supervised classification method (Treitz, Howarth, & Gong, 1992) to classify the 3-m resolution PlanetScope multispectral data into three main land cover types, vegetation, building/road/soil, and water. The overall classification accuracy is 92.6% and Kappa coefficient is 0.858 (as shown in Table 1.) Next, we resampled the 3-m LULC data from PlanetScope into 30-m to match the pixel size of Landsat 8 LST products. We did not use Landsat 8's multispectral bands for LULC classification to avoid excessive mixed pixel issues in urban areas. Mixed pixels are especially problematic for urban areas because of the nature of scattering from buildings and trees in cities. Each land cover composition in the newly resampled 30-m product was the ratio between the area of individual land cover type and the area of the whole 30-m pixel. This way, each mixed pixel was decomposed to three land cover types to explain the detailed LULC variability in the mixed pixels.

2.4. Weather station data

We used weather observations (air temperature and solar radiation) from a rooftop station operated by the Climatology Research Laboratory (CRL; http://weather. ggy.uga.edu/data/daily/) at UGA. This is the closest weather station to most places along the bus routes, and this station is located near a heavily urbanized part of the city of Athens. The CRL weather station takes observations every two minutes and saves the data in an open-access online FTP. Due to the different frequency of data collection between the weather station and satellite, we chose the closest weather station reading to the satellite time stamp. Other weather stations located in Athens-Clarke county have deficiencies. Several WeatherSTEM stations collect the necessary data but are not always located over a natural surface, which is common practice for representative weather station measurements. A National Weather Service/Federal Aviation Administration operates a weather station at Ben Epps Airport. This station is remote from the main bus routes and does not measure solar radiation. Even though the other weather stations were not used in model training and tuning, we use their data for AAT model validation.

Table 1
Producer, User and Overall Accuracy of the LULC supervised classification.

Classification	Impervious	Water	Vegetation	Truth	User accuracy
Impervious	**8900**	95	1699	10,694	0.832
Water	4	**1549**	3	1556	0.996
Vegetation	213	117	**16,221**	16,551	0.98
Total	9117	1761	17,923	**28,801**	**Overall Accuracy = 0.926**
Producer accuracy	0.976	0.88	0.905		**Kappa = 0.858**

2.5. Random forest modeling

We used the Random Forest model (Liaw, Wiener, et al., 2002), a commonly used algorithm in classification and regression, to predict AAT using multiple parameters. It is among the types of models that could be used where the input data for regression analysis have high spatial and temporal variabilities. In this algorithm, hundreds of decision trees are randomly generated, and all trees are independent of each other. A final classification or regression model is the most accurate one made by aggregating multiple layers of the decision trees. The trees are randomly generated, and the training, validation, and test set for the whole process are also all randomly drawn from the entire dataset to avoid the common overfitting and biases from statistical models.

The final formula for this study was:

$$T_{air} \sim f\{T_{LST}, LC_{veg}, LC_{man}, LC_{water}, CC, SR, T_{wea}\} \quad (2)$$

where T_{air} is AAT, T_{LST} is satellite-derived LST, LC_{veg}, LC_{man}, LC_{water} are the composition of the land cover types of vegetation, building/road/soil, and water, CC is canopy cover, S.R. is solar radiation, T_{wea} is weather station temperature.

Because the Random Forest Model result can be different based on the choices of randomly selected training and test sets, the best model could only be generated after trying different combinations of model parameters and training-test sets. However, with a large volume of data and almost infinite possibilities of model parameters, it is too computationally intensive to try all different combinations. Thus, an alternative solution is to use 'Hyperparameter' and K-Fold Cross-Validation. Hyperparameter is an alternative method where one can randomly change the parameter in the random forest regressor. Also, this method is a combination of two or three repeated parameterizations to reduce the time cost. We pre-define the choices for the number of estimators as 200, 400, 600, 800, and 1000 and found 1000 to be the best choice. Then we chose several possible numbers for each parameter, for instance, 500, 1000, and 1500, to find if the result from a previous parameterization is the global best. And we implied K-Fold Cross-Validation in the hyperparameter process with K = 3. Therefore, we iterated the validation three times with two-thirds of the total dataset as the training set and one-third as the testing set. The combination of Hyperparameter and 3-Fold Cross-Validation ensures the output model is the best possible result with the input data and parameters. Then we evaluated the model with the test dataset via different ways to calculate model errors.

AAT model evaluation was performed by computing the absolute and standardized differences between predictions and actual values for the test dataset. We also calculated the mean absolute percentage error (MAPE) via dividing the error by the actual values in the test dataset. Here, the test and validation dataset contained a portion of the data excluded from K-Fold Cross Validation, and the data especially left for model validation, which was the 2019 dataset in our study.

2.6. Other statistical modeling

We compared results from the Random Forest algorithm with other

commonly used approaches, including Multilinear Regression and Gradient Boost Model, to evaluate relative accuracy.

Multilinear Regression Model attempts to fit a linear equation between the targeted parameter and multiple explanatory variables (Kutner, Nachtsheim, Neter, Li, et al., 2005). The general formula looks like this:

$$y = \beta 0 + \beta 1x1 + \beta 2x2 + \cdots \beta kxk + \epsilon \quad (3)$$

where y is the dependent variable, $x1$, $x2$, …, xk are explanatory variables, ϵ are the residual terms, $\beta 0$, $\beta 1$, $\beta 2$, …, βk are correlation coefficients in this model.

Gradient Boost Model is a type of decision tree model constructed from multiple weak learners to produce better results (Friedman, 2002). However, this is different from Random Forest Regression and other decision trees because all tree regressors are predictions of the residuals between the constructed trees and the real value.

3. Results and discussion

3.1. Vehicle-borne AAT and hot spots

We first present a sample map of vehicle-borne sensor temperatures on August 14th, 2018 (Landsat 8 overpass day) for the morning (30 min before and after 9 am), noon (30 min before and after 12 pm), afternoon (30 min before and after 3 pm), and early evening (30 min before and after 6 pm) (Fig. 2). The bus routes cover the most populated parts of Athens-Clarke County comprehensively. However, less populated areas with fewer businesses such as in the southwest are not well covered.

From these data points, both spatial and temporal variations are evident. Spatially "hot spots" evident by red color-coding is located across the city at places with a higher level of urbanization. Downtown Athens, the shopping plazas in the west corridor, and several other suburban plazas all stand out in the image as hot spots. At noon and afternoon times, the hot spots reached temperatures over 40 °C while at the same time, some suburban and places well shaded by buildings or canopies were below 30 °C. The maximum difference was 14 °C during noontime. LST product shows a similar noontime spatial pattern that urban core areas reached 52 °C, and rural areas were around 30 °C. The maximum difference was 22 °C in the Landsat 8 LST image on that day. Diurnal variability is also evident across the area, as shown in the temperature range of each image. On that day, the AAT increased from around 18 °C in the early morning to around 40 °C in the late afternoon.

These maps show evident diversity in temperatures across the city, but the dataset is limited in spatial coverage to areas with bus routes, which hinders spatial extrapolation and applications. The LST satellite products provide the spatial contiguity, although at a coarser spatial resolution and with quite different values from vehicle-borne AAT (Fig. 3). In the clipped grid, AAT recorded by the vehicle-borne sensors ranged from 29.69 to 32.75 °C, while the Landsat 8 LST value in the same grid was 39 °C, which is significantly higher than AAT. Nevertheless, Landsat 8 LST images cover the whole urban area, with many cells without any bus AAT data. To better facilitate the bus AAT data to explain the urban heat phenomenon beyond the bus routes, we trained the machine learning models such as random forest and others using the overlapping bus AAT and Landsat 8 LST along with spatially decomposed LULC and other biophysical parameters. The machine learning models were aimed at predicting AAT at places not covered by the vehicle-borne sensors. After a robust training and validation process, the models can be implemented every time Landsat 8 passes over the study site producing a novel Landsat-8 resolution or a higher resolution spatial AAT product. These AAT products can then be used to study community exposure risk or heat hazard risk in an urban environment. That is the premise behind the machine learning modeling.

3.2. Random forest model training and tuning

According to previous research about biophysical parameters

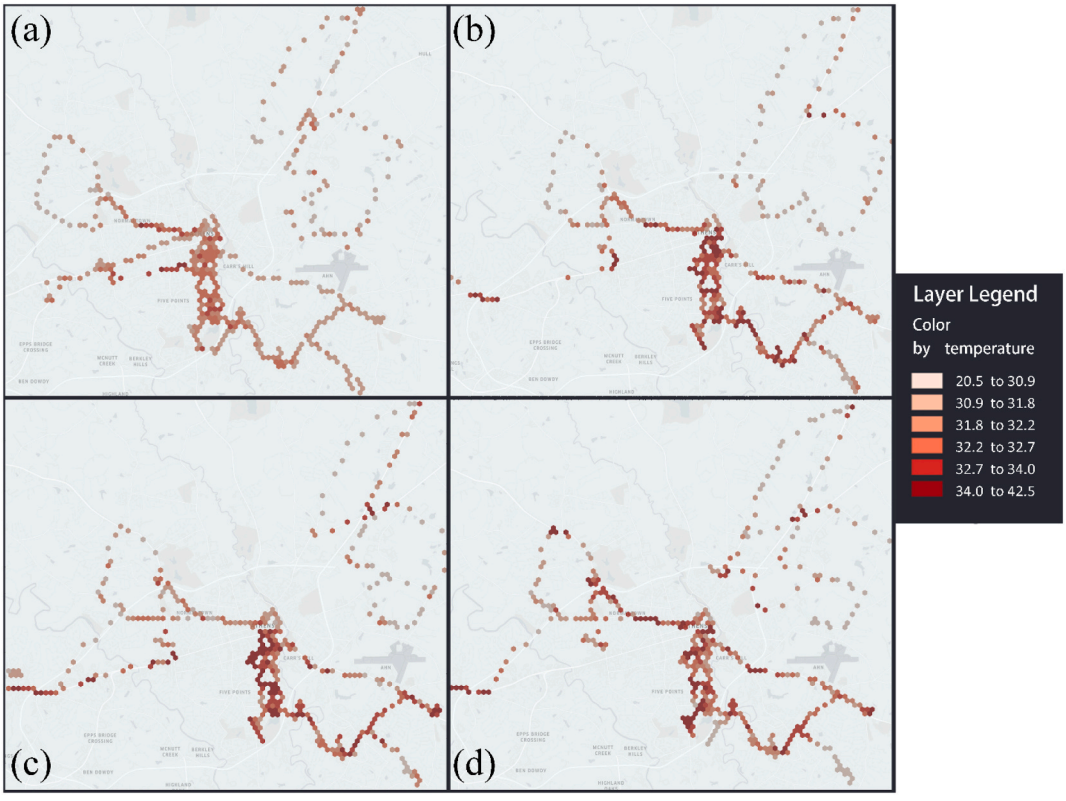

Fig. 2. Morning (a), noon (b), afternoon (c), and evening (d) time snapshots of AAT data from bus sensors(Data classified using natural breaks).

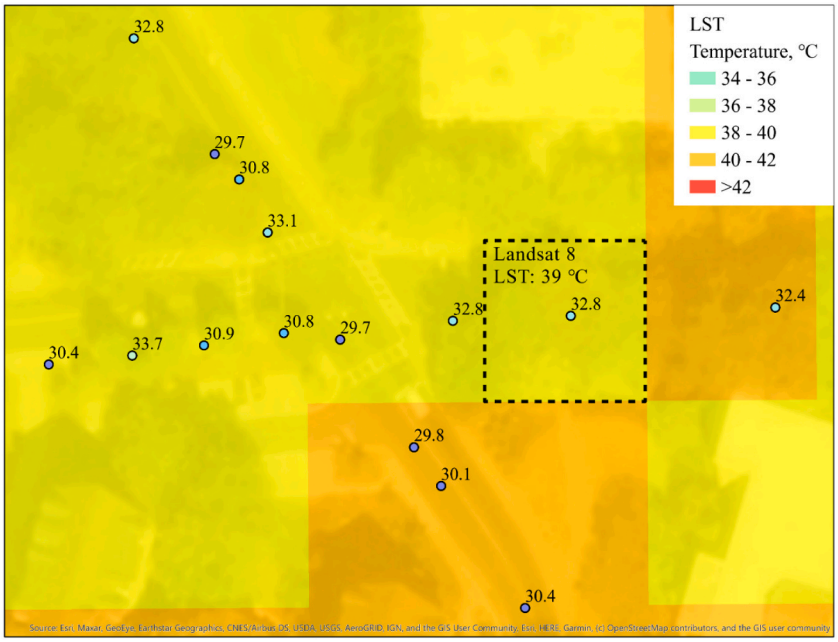

Fig. 3. Comparison between bus AAT data and Landsat 8 LST on spatial resolution and contiguity (each dot is a bus AAT reading; each grid is a Landsat 8 30-m pixel; data classified using equal intervals).

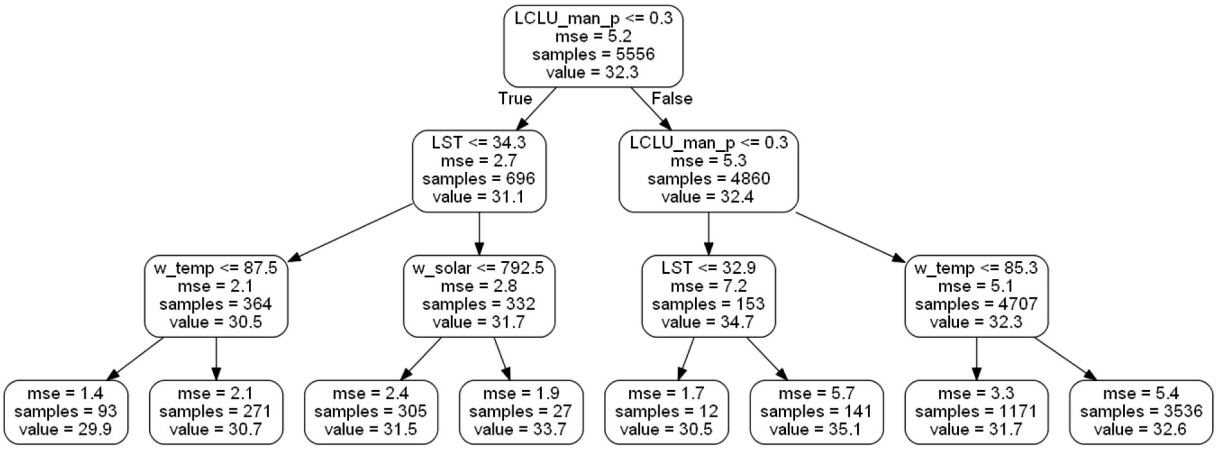

Fig. 4. A demo three-layers random forest model dendrogram used in this study (mse: mean standard error; samples: number of samples within each branch; value: the predicted value for the samples in this branch).

Table 2
Tuning result for the random forest regressor parameters.

Bootstrap	Max depth	Max features	Min samples leaf	Min samples split	Number of estimators
False	80	2	1	2	1000

influencing AAT (Frey et al., 2007; Li et al., 2011; Yuan & Bauer, 2007), there is an agreement that the local environment has a significant influence on air temperature. Human-made materials, vegetation, closeness to a waterbody, shadow of building or trees, and time of the day are important contributors to changes in air temperature. Based on previous findings and availability of datasets, we choose five parameters (i.e., LST, LULC, Canopy Cover, Weather Station Temperature, and Solar Radiation) to model AAT. Because LST from Landsat 8 is only available during noontime on certain days (temporal resolution: 16 days), we used the corresponding pairs of vehicle-borne AAT and Landsat 8 LST datasets during the 30-min window of Landsat 8 overpass for model training. We trained and validated the random forest model using hyperparameter and K-Fold Cross-Validation. A three-layer dendrogram is presented in Fig. 4, and the tuning results for the regressor parameters are shown in Table 2.

In this dendrogram, LCLU_man_p is the percentage of the land cover being human-made materials, w_temp and w_solar are the weather station temperature and solar radiation observations respectively, and LST is Landsat 8 land surface temperature. This dendrogram can be interpreted as a simplified regression tree if the tree only has three layers. At every node, each set of input data will make a Yes or No decision based on the criteria, and all input data will be categorized into two branches. If the tree stops there, each set of data will be signed with a value shown in the box. For example, if a sample has less than 30% of land cover with human-made materials and weather station temperature observations are smaller than 30.83 °C (87.5 °F), then the predicted value will be assigned as 29.9 °C together with other 93 samples and the predicted values for these samples will have an MSE of 1.4 °C.

Bootstrapping is a standard method for sampling data points (with or without replacement) in model tuning. The parameters used in

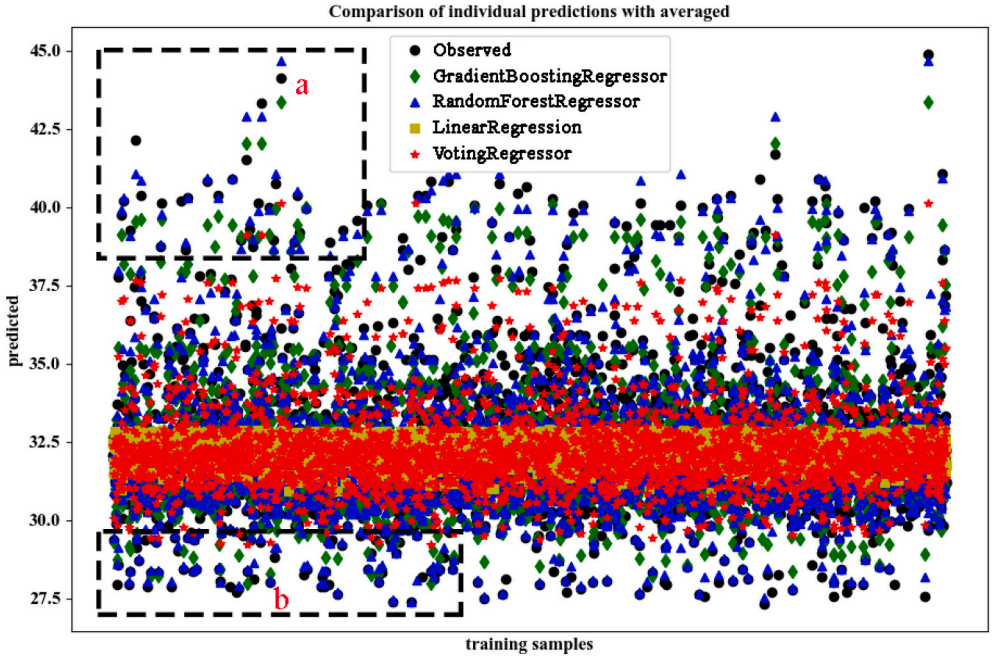

Fig. 5. Comparison between predicted values from Random Forest model and other regression models, and the actual observed values.

bootstrapping are, (1) Max depth - maximum number of levels in each decision tree, (2) Max features - maximum number of features considered for splitting a node, (3) Min samples leaf - minimum number of data points allowed in a leaf node, (4) Min samples split - minimum number of data points placed in a node before the node is split, and (5) Number of estimators - number of trees in the forest. For instance, the demo three-layer tree in Fig. 4 has (1) a Max depth of 3, (2) Max features of 1, (3) Min samples in a leaf of 12, (4) Min samples split of 153, and (5) the number of the estimator as 1 because this demo dendrogram is just one tree. Also, in the K-Fold Cross-Validation, the number K is three, which means every time we use one-third of the total data to test the newly generated model.

We compared the random forest model output to the output from the traditional linear regression model, Gradient boosting regressor model, to illustrate that those methods do not outperform the random forest model. We compared the observed values from the test dataset and the predicted values from all the above regression models (Fig. 5) and calculated the Mean Absolute Error (MAE) and Mean Standard Error (MSE) for each regression model (Table 3).

Sample points inside the two rectangles in Fig. 5 (a and b) were used to evaluate the prediction performance between different regression models. It is evident that the predicted AAT values from the Random Forest model are closer to the observed ones than the other models. MAE and MSE values for all regression models showed that the Random Forest model outperformed other methods (Table 3).

3.3. Model testing

We assessed the effectiveness of the hyperparameter and model tuning in the performance of the model using error parameters such as MAE, MSE, and MAPE (Table 4). These parameters were used to assess the differences between the original default Random Forest model and the tuned model. The tuned model has a MAE of 0.29 °C, a MSE of 0.57 °C, and a MAPE of 0.86% compared to a MAE of 0.36 °C, MSE of 0.71 °C and a MAPE of 1.08% before hyperparameter tuning and validation. Moreover, in the model, each parameter contributes a significant portion to explain the model variance ranging from LST contributing 35% to Percentage of Water contributing 5%.

Table 3
MAE and MSE comparison for different regression models used to predict AAT.

Model	MAE, °C	MSE, °C
Random Forest	0.29	0.57
Gradient Boost	0.74	1.06
Linear Regressor	1.57	2.24
Voting Regressor	0.81	1.13

Table 4
Contribution of variables in the tuned Random Forest Model.

Most important parameter	Contribution of the variables (explained variance)
LST	0.35
Weather station temperature	0.18
Percentage of Vegetation	0.13
Percentage of soil/building/road	0.13
Percent of canopy cover	0.09
Weather station solar radiation	0.07
Percentage of Water	0.05
Total	1

Due to the overrepresentation of data collected from places with impervious surfaces, we included weather station data to represent temperature over suburbs and other less urbanized regions. We randomly extracted 10,800 data points with predicted AAT and LST from Landsat 8. For the weather station temperature, we used noon time data from the UGA CRL Weather Station, Athens Epps Airport Weather Station and three WeatherSTEM Weather Station data. These weather stations represent both central urbanized city and suburbs less-populated Athens. According to the plot below (Fig. 6), the modeled AAT output is closer to weather station temperature than LST. And the about 6 °C variation in AAT is also identical with the spatial variation found from Result Part 1 which cannot be captured by Weather Station Data.

The observed versus predicted AAT of the tuned model revealed the error to be smaller than two standard deviations (2 S.D.) without any particular pattern of over or underestimation (Fig. 7). But we still identified a few outliers with differences greater than 5 °C.

As we observed in the vehicle-borne AAT data (Fig. 3), typically, there are many data points within one Landsat 8 LST pixel. Any time

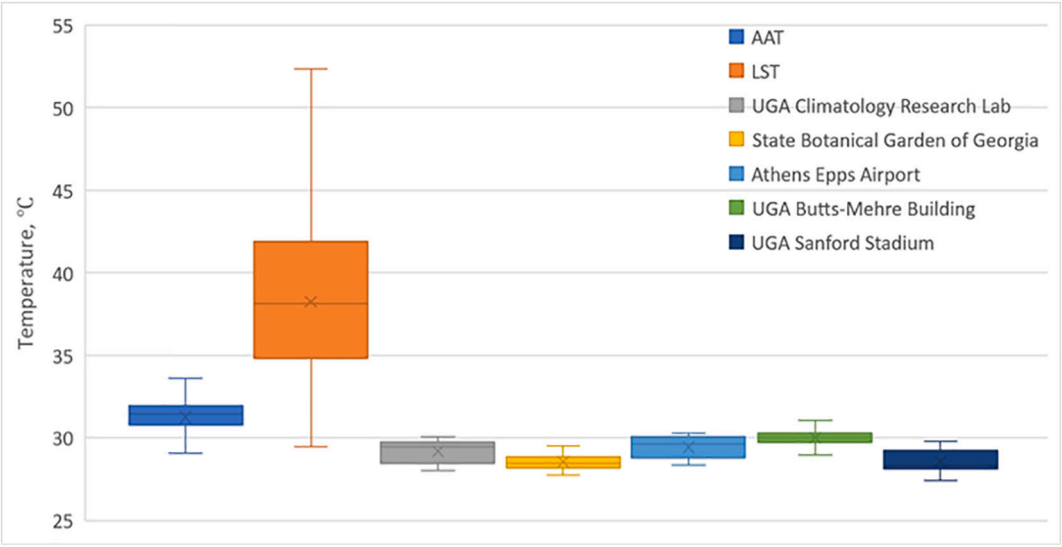

Fig. 6. Comparison between AAT, LST and Weather Station Temperature at noontime (box represent 25th and 75th percentile, the whiskers represent 5th and 95th percentiles, and the line in the box represents median).

there is an abrupt change in vehicle-borne AAT due to either a change in land cover type within a Landsat 8 pixel or being at the proximity of an unnatural heat source such as exhaust or engine of another vehicle, there will be an increased deviation between LST and AAT. In the Random Forest Regression process, the decision tree will balance between normal data and outliers instead of overfitting the pattern of outliers. Therefore, the model will not respond to the outliers caused by factors not included as model variables well.

3.4. Model output visualization and analysis

Finally, all important parameters, LST, canopy cover, land cover composition, weather station temperature, and solar radiation, were used in the Random Forest model for AAT predictions for the entire Athens-Clarke County for each cloudless Landsat image. Fig. 8 presents an example of the estimated 30-m resolution AAT and Landsat 8 LST distribution across Athens-Clarke County at noontime of August 14th, 2018 and August 1st 2019.

From this predicted AAT map, we can pick the hot spots across the city, showing both the highly urbanized downtown areas, business plazas, and the residential districts. The hot spots during noontime on August 14th, 2018 peaked around 34 °C at the downtown and shopping plazas in Athens, and there were areas found to be 5–6 °C cooler several meters away. The hot spots during noontime on August 1st 2019 peaked around 36 °C at the downtown and shopping plazas in Athens, and there were also cool spots around them. These cool spots near downtown and plazas are mostly missing from the LST maps because surrounding environments usually contaminate LST pixels being coarser resolution. Residential districts typically have a large amount of vegetation coverage, which subdues these hot spots on the LST maps. The similar but opposite phenomenon is that heated buildings and roads were 4–5 °C higher than the green space near those buildings or apartment complexes.

For a quantitative comparison, we selected all the urban core developed areas and plotted their corresponding AAT and LST for pixels with changing percentage of vegetation since % vegetation within a pixel plays a significant role in reducing LST (Fig. 9). Both graphs show the negative correlation between the percentage of vegetation and

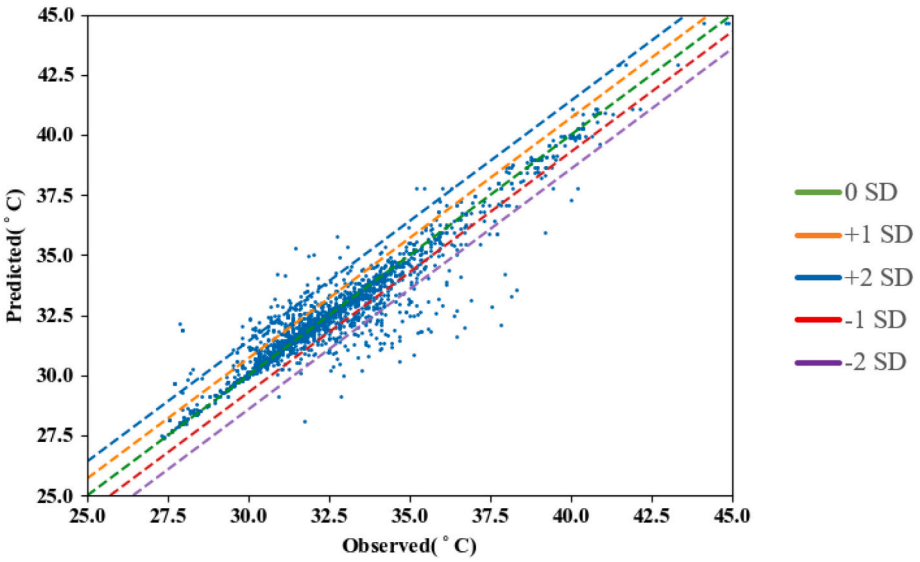

Fig. 7. Visualization of Random Forest model error: Actual vs. Predicted AAT (each line represents a certain standard deviation- SD from the mean).

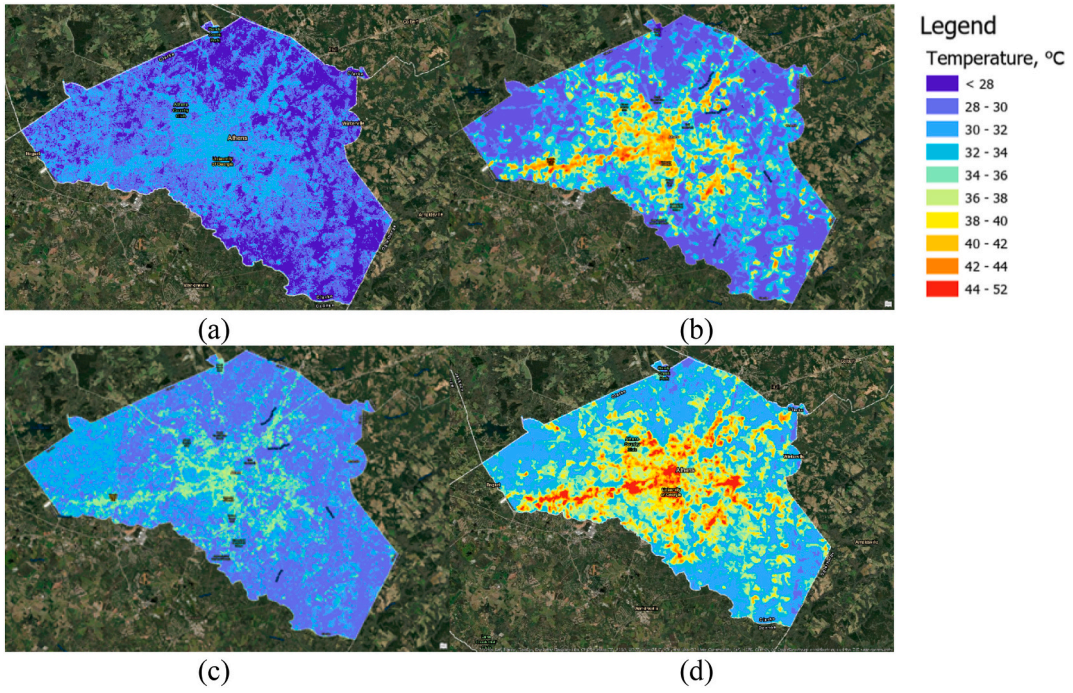

Fig. 8. Comparison between Random Forest predicted AAT map (a), and Landsat 8 LST map of Athens-Clarke County (b) (Noon, August 14th, 2018), predicted AAT map (c), and Landsat 8 LST map (d) (Noon, August 1st 2019). Data classified using equal intervals.

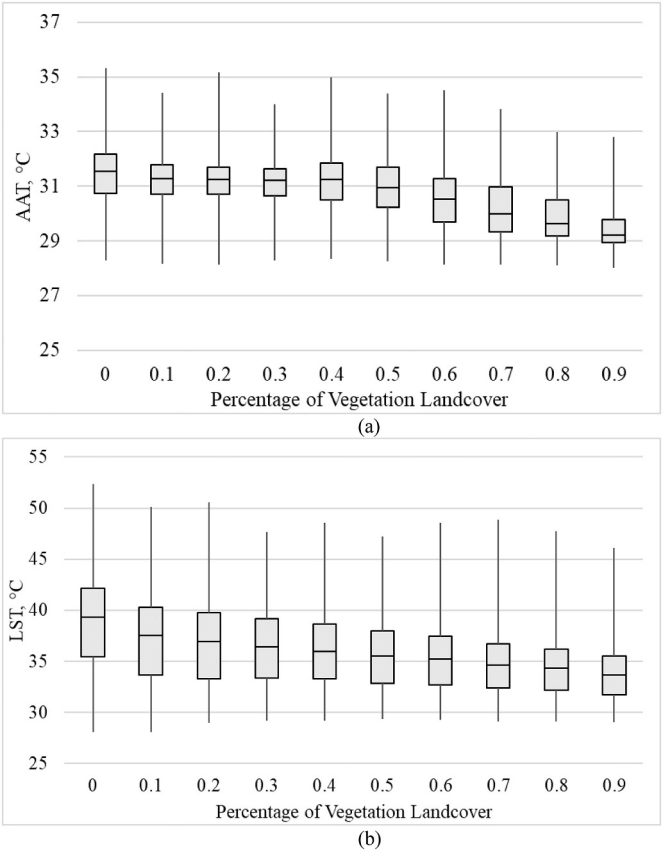

Fig. 9. Relationship between percentage of vegetation cover and (a) AAT, and (b) LST (box represent 25th and 75th percentile, the whiskers represent 5th and 95th percentiles, and the line in the box represents median).

temperature, which means an increase in vegetation cover reduces both AAT and LST. However, LST instantly decreases as the vegetation cover increases because more vegetation means less artificial materials like asphalt and pavement and higher evaporative cooling. AAT does not decrease when the vegetation cover increases from 0 to 50% because, in a mixed pixel, the near-surface temperature is strongly influenced by other land cover types within the pixel and around the pixel. In other words, AAT is not sensitive to increasing vegetation cover up to a certain threshold as opposed to LST, which shows an immediate reduction. The threshold of vegetation cover for AAT to decline is roughly 50% at 30-m resolution (this study). However, that threshold will most likely change with a change in spatial resolution. AAT Shows a significant decline when vegetation cover within a pixel exceeds 50%.

Previous research tried either to solely link LST with AAT (Nichol & Wong, 2008; Vancutsem et al., 2010; Zhu et al., 2013) or use coarse resolution data to estimate AAT (Harlan et al., 2013; Johnson & Wilson, 2009). A few studies incorporated different parameters into AAT modeling but did not have revisited data for model validation and scaling up (Makido, Shandas, Ferwati, & Sailor, 2016). Our modeling framework enables a broader application to other cities and countries. The results indicate that the sensing is scalable and replicable to other cities of different sizes, especially to those with well-connected city transit systems or other public utility vehicles (e.g., sanitation, maintenance, law enforcement) (Anjomshoaa et al., 2018). The public utility vehicles can also be used to collect AAT data and make them available as open-source and free to use datasets.

3.5. Limitations

There are a few limitations to be recognized in our study. First, in using vehicle-borne sensors, the AAT sensing was limited to roadways. Future developments should include AAT from human-borne sensors into the model. Human-borne sensors allow temperature monitoring in areas that are accessible only to pedestrians and allow tracking of individual heat exposure within known UHIs (De Nazelle et al., 2013; Tonekaboni et al., 2018). A second issue involves the raster format of our remotely sensed data. Modifiable Areal Unit Problem (MAUP) is an

inevitable problem in studies using raster data, irrespective of how we choose boundaries of pixels. This problem would be minimized in future studies when we have a much higher spatial resolution model but still cannot be eliminated completely. In future studies, we will improve the spatio-temporal resolution of AAT maps in several ways. Spatially, we will build models that have a spatial resolution of 3-m. Temporally, we will ensure that the model adapts to diurnal AAT changes, which will help in producing AAT maps every hour for an urban area and meeting the ultimate goal of generating dynamic AAT maps for public and urban planning.

4. Conclusions

In this study, we successfully developed and implemented a vehicle-borne drive-by thermal sensing system in a mid-sized city and used the Random Forest model to predict ambient air temperature (AAT) at places without such high-resolution AAT data. Our analysis shows that AAT, which is a better indicator of human heat exposure, has a complicated relationship with land cover. LST traditionally used in urban heat island studies may not be sufficient to indicate human thermal comfort as it does not have a strong adjacency effect and almost has a binary relationship with land cover types or shading. LST is not an accurate indicator of human thermal comfort. For instance, many places can have extremely high LST but people do not usually experience it as extreme thermal discomfort because of different radiative properties of the ground (absorptive vs reflective), whereas, AAT shows a higher variability as it is derived from a complex relationship between land cover types and other biophysical and meteorological parameters. This novel approach outlines a path to utilize public vehicles or transit vehicles as sensor carriers with minimum additional cost and environmental impact. The long lifespan and low energy requirement of mobile sensors also provide an excellent continuity in the AAT observations. Machine learning techniques such as Random Forest modeling can be valuable in scaling up the mobile sensor observations to a larger geographic region.

The findings from this study will be beneficial for understanding the heat vulnerabilities of individual communities. Better identification of communities at risk of heat stress will help local, state, and federal officials to devise more informed and targeted mitigation efforts, and therefore make communities and cities more livable, workable, and sustainable. The results will have transformative social benefits. For instance, this knowledge will help identify where structural changes can be made to known urban UHIs via increased greenspace, posting warning signs, aid employers in developing better heat-safety policies for their workers, and guide agencies such as the NWS in how to communicate public warnings on heat safety more effectively.

Author Statement

Yanzhe Yin: Conceptualization, Data curation, Software, Formal analysis, Visualization, Writing - original draft.

Navid Hashemi Tonekaboni: Data curation, Software, Visualization, Validation.

Andrew Grundstein: Conceptualization, Methodology, Investigation, Writing- Review and Editing.

Deepak R. Mishra: Conceptualization, Methodology, Funding acquisition, Project administration, Writing- Review and Editing.

Lakshmish Ramaswamy: Conceptualization, Methodology, Writing-Review and Editing.

John Dowd: Data curation, Software.

Acknowledgments

This research was funded by (NSF) S&CC: Smart & Connected Communities program (Grant # 1637277). Thanks are due to our colleagues Sujeet Kulkarni, Himanshu Pendyala, and Omid Setayeshfar, who contributed to this research by developing our mobile applications. We would also like to thank the University of Georgia's (UGA) Transportation Services and Athens-Clarke County (ACC) Unified Government's Transit Department for allowing us to mount the sensors on their buses and answer numerous questions throughout the projects. Special Thanks to Butch McDuffie, Director ACC Transit and Don Walter, Director UGA Transit for their support.

References

Anjomshoaa, A., Duarte, F., Rennings, D., Matarazzo, T. J., Desouza, P., Ratti, C., ... Desouza, P. (2018). City scanner: building and scheduling a mobile sensing platform for smart city services. *IEEE Internet of Things Journal, 5*(6), 4567–4579. https://doi.org/10.1109/jiot.2018.2839058.

Arguez, A., Durre I., Applequist S., Squires M., Vose R., Yin X., & Bilotta R. (2010). NOAA's U.S. Climate Normals (1981–2010). *NOAA National Centers for Environmental Information* 10: V5PN93JP.

Brooke Anderson, G., Bell, M. L., Brooke Anderson, G., & Bell, M. L. (2011). Heat waves in the United States: Mortality risk during heat waves and effect modification. *Environmental Health Perspectives, 119*(2), 210–218. https://doi.org/10.1289/relationships.temperature-mortality.

De Nazelle, A., Seto, E., Donaire-Gonzalez, D., Mendez, M., Matamala, J., Nieuwenhuijsen, M. J., & Jerrett, M. (2013). Improving estimates of air pollution exposure through ubiquitous sensing technologies. *Environmental Pollution, 176*(May), 92–99. https://doi.org/10.1016/j.envpol.2012.12.032.

Eisenman, S., Lane, N., & Miluzzo, E. (2006). MetroSense project: People-centric sensing at scale. WSW at SenSys.

Erman, A. T., Van Hoesel, L., Havinga, P., & Jian, W. (2008). Enabling mobility in heterogeneous wireless sensor networks cooperating with UAVs for mission-critical management. *IEEE Wireless Communications, 15*(6), 38–46. https://doi.org/10.1109/MWC.2008.4749746.

Frey, C. M., Rigo, G., & Parlow, E. (2007). Urban radiation balance of two coastal cities in a hot and dry environment. *International Journal of Remote Sensing, 28*(12), 2695–2712. https://doi.org/10.1080/01431160600993389.

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics and Data Analysis, 38*(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2.

Harlan, S. L., Brazel, A. J., Prashad, L., Stefanov, W. L., & Larsen, L. (2006). Neighborhood microclimates and vulnerability to heat stress. *Social Science and Medicine, 63*(11), 2847–2863. https://doi.org/10.1016/j.socscimed.2006.07.030.

Harlan, S. L., Declet-Barreto, J. H., Stefanov, W. L., & Petitti, D. B. (2013). Neighborhood effects on heat deaths: Social and environmental predictors of vulnerability in Maricopa County, Arizona. *Environmental Health Perspectives*. https://doi.org/10.1289/ehp.1104625.

Hasenfratz, D., Saukh, O., Walser, C., Hueglin, C., Fierz, M., Arn, T., ... Thiele, L. (2015). Deriving high-resolution urban air pollution maps using mobile sensor nodes. *Pervasive and Mobile Computing, 16*, 268–285. https://doi.org/10.1016/j.pmcj.2014.11.008.

Ho, H. C., Knudby, A., Xu, Y., Hodul, M., & Aminipouri, M. (2016). A comparison of urban Heat Islands mapped using skin temperature, air temperature, and apparent temperature (humidex), for the greater Vancouver area. *Science of the Total Environment, 544*(February), 929–938. https://doi.org/10.1016/j.scitotenv.2015.12.021.

Honicky, R. J., Brewer, E. A., Paulos, E., & White, R. M. (2008). N-SMARTS: Networked suite of mobile atmospheric real-time sensors. *Proceedings of the ACM SIGCOMM 2008 conference on computer communications—2nd ACM SIGCOMM Workshop on Networked Systems for Developing Regions, NSDR'08* (pp. 25–29). New York, NY: ACM Press. https://doi.org/10.1145/1397705.1397713.

Ihara, T., Kikegawa, Y., Asahi, K., Genchi, Y., & Kondo, H. (2008). Changes in year-round air temperature and annual energy consumption in office building areas by urban Heat-Island countermeasures and energy-saving measures. *Applied Energy, 85*(1), 12–25. https://doi.org/10.1016/j.apenergy.2007.06.012.

Jiménez-Muñoz, J. C., & Sobrino, J. A. (2010). A Single-Channel algorithm for land-surface temperature retrieval from ASTER data. *IEEE Geoscience and Remote Sensing Letters, 7*(1), 176–179. https://doi.org/10.1109/LGRS.2009.2029534.

Johnson, D. P., & Wilson, J. S. (2009). The socio-spatial dynamics of extreme urban heat events: The case of heat-related deaths in Philadelphia. *Applied Geography, 29*(3), 419–434. https://doi.org/10.1016/j.apgeog.2008.11.004.

Kloog, I., Chudnovsky, A., Koutrakis, P., & Schwartz, J. (2012). Temporal and spatial assessments of minimum air temperature using satellite surface temperature measurements in Massachusetts, USA. *Science of the Total Environment, 432*(August), 85–92. https://doi.org/10.1016/j.scitotenv.2012.05.095.

Kutner, M. H., Nachtsheim, C. J., Neter, J., Li, W., et al. (2005). *Applied linear statistical models. 5*. Irwin, NY: McGraw-Hill.

Kuttler, W. (2008). The urban climate-basic and applied aspects. *Urban Ecology: An International Perspective on the Interaction Between Humans and Nature*https://doi.org/10.1007/978-0-387-73412-5_13.

Li, J., Song, C., Lu, C., Zhu, F., Meng, X., & Jianguo, W. (2011). Impacts of landscape structure on surface urban Heat Islands: A case study of Shanghai, China. *Remote Sensing of Environment, 115*(12), 3249–3263. https://doi.org/10.1016/j.rse.2011.07.008.

Liaw, A., Wiener, M., et al. (2002). Classification and regression by RandomForest. *R News, 2*(3), 18–22.

MacFaden, S. W., O'Neil-Dunne, J. P. M., Royar, A. R., Lu, J. W. T., & Rundle, A. G. (2012). High-resolution tree canopy mapping for new York City using LIDAR and

object-based image analysis. *Journal of Applied Remote Sensing, 6*(1), https://doi.org/10.1117/1.jrs.6.063567 63567–1.

Maier, G., Grundstein, A., Jang, W., Li, C., Naeher, L. P., & Shepherd, M. (2014). Assessing the performance of a vulnerability index during oppressive heat across Georgia, United States. *Weather, Climate, and Society, 6*(2), 253–263. https://doi.org/10.1175/WCAS-D-13-00037.1.

Makido, Y., Shandas, V., Ferwati, S., & Sailor, D. (2016). Daytime variation of urban Heat Islands: The case study of Doha, Qatar. *Climate, 4*(2), 32. https://doi.org/10.3390/cli4020032.

Nichol, J. E., & Wong, M. S. (2008). Spatial variability of air temperature and appropriate resolution for satellite-derived air temperature estimation. *International Journal of Remote Sensing, 29*(24), 7213–7223. https://doi.org/10.1080/01431160802192178.

Oke, T. R. (1988). The urban energy balance. *Progress in Physical Geography: Earth and Environment, 12*(4), 471–508. https://doi.org/10.1177/030913338801200401.

Overeem, A., Robinson, J. C. R., Leijnse, H., Steeneveld, G. J., Horn, B. K. P., & Uijlenhoet, R. (2013). Crowdsourcing urban air temperatures from smartphone battery temperatures. *Geophysical Research Letters, 40*(15), 4081–4085. https://doi.org/10.1002/grl.50786.

Planet Team. (2017). Planet application program Interface: In space for life on earth. San Francisco, CA, p. 40.

Qin, Z., Dall'Olmo, G., Karnieli, A., & Berliner, P. (2001). Derivation of Split window algorithm and its sensitivity analysis for retrieving land surface Temperature from NOAA-advanced very high resolution radiometer data. *Journal of Geophysical Research Atmospheres, 106*(D19), 22655–22670. https://doi.org/10.1029/2000JD900452.

Reis, S., Seto, E., Northcross, A., Quinn, N. W. T., Convertino, M., Jones, R. L., & Holger R. Maier, et al. (2015). Integrating modelling and smart sensors for environmental and human health. *Environmental Modelling and Software, 74*(December), 238–246. https://doi.org/10.1016/j.envsoft.2015.06.003.

Sari, K. R., & Koppe, C. (2006). Heatwaves: Past and future impacts on health. *Integration of public health with adaptation to climate change: Lessons learned and new directions* (pp. 136–160). CRC Press.

Tonekaboni, N. H., Grundstein, A., Ramaswamy, L., Kulkarni, S., Mishra, D., & Yin, Y. (2018). Scouts: A smart community centric urban heat monitoring framework. In *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities, ARIC 2018*. doi: https://doi.org/10.1145/3284566.3284570.

Tonekaboni, S. N. H. (2019). Towards effective harnessing of crowd sensing and drive-by sensing for high-resolution urban heat analysis: Challenges and approaches. University of Georgia.

Treitz, P. M., Howarth, P. J., & Gong, P. (1992). Application of satellite and GIS technologies for land-cover and land-use mapping at the rural-urban fringe: A case study. *Photogrammetric Engineering and Remote Sensing, 58*(4), 439–448.

Vancutsem, C., Ceccato, P., Dinku, T., & Connor, S. J. (2010). Evaluation of MODIS land surface temperature data to estimate air temperature in different ecosystems over Africa. *Remote Sensing of Environment, 114*(2), 449–465. https://doi.org/10.1016/j.rse.2009.10.002.

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. https://doi.org/10.1016/j.rse.2006.09.003.

Zhu, W., Lu, A., & Jia, S. (2013). Estimation of daily maximum and minimum air temperature using MODIS land surface temperature products. *Remote Sensing of Environment, 130*(March), 62–73. https://doi.org/10.1016/j.rse.2012.10.034.