Citizen and machine learning-aided high-resolution mapping of urban heat exposure and stress

ENVIRONMENTAL RESEARCH
INFRASTRUCTURE AND SUSTAINABILITY

OPEN ACCESS

RECEIVED
7 December 2022

REVISED
24 July 2023

ACCEPTED FOR PUBLICATION
11 August 2023

PUBLISHED
25 August 2023

PAPER

Citizen and machine learning-aided high-resolution mapping of urban heat exposure and stress

Xuewei Wang[1,2,3], **Angel Hsu**[1,2,3,*] **and TC Chakraborty**[4]

[1] University of North Carolina at Chapel Hill, Data-Driven EnviroLab, Chapel Hill, NC, United States of America
[2] Department of Public Policy, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States of America
[3] University of North Carolina at Chapel Hill, Institute for Environment, Chapel Hill, NC, United States of America
[4] Atmospheric Sciences & Global Change Division, Pacific Northwest National Laboratory, Richland, WA, United States of America
[*] Author to whom any correspondence should be addressed.

E-mail: angel.hsu@unc.edu

Keywords: urban heat island, citizen science, heat stress, machine learning

Supplementary material for this article is available online

Abstract

Through conversion of land cover to more built-up, impervious surfaces, cities create hotter environments than their surroundings for urban residents, with large differences expected between different parts of the city. Existing measurements of ambient air temperature and heat stress, however, are often insufficient to capture the intra-urban variability in heat exposure. This study provides a replicable method for modeling air temperature, humidity, and moist heat stress over the urban area of Chapel Hill while engaging citizens to collect high-temporal and spatially-resolved air temperature and humidity measurements. We use low-cost, consumer-grade sensors combined with satellite remote sensing data and machine learning to map urban air temperature and relative humidity over various land-cover classes to understand intra-urban spatial variability of ambient heat exposure at a relatively high resolution (10 m). Our findings show that individuals may be exposed to higher levels of air temperature and moist heat stress than weather station data suggest, and that the ambient heat exposure varies according to land cover type, with tree-covered land the coolest and built-up areas the warmest, and time of day, with higher air temperatures observed during the early afternoon. Combining our resulting dataset with sociodemographic data, policymakers and urban planners in Chapel Hill can use data output from this method to identify areas exposed to high temperature and moist heat stress as a first step to design effective mitigation measures.

1. Introduction

During 2018–2020, a total of 3066 heat-related deaths occurred in the United States (CDC, 2022). Climate change is increasing the probability of intense, prolonged heat waves in many parts of the world (IPCC 2021, Vargas Zeppetello *et al* 2022). Extreme heat exposure in urban areas is of particular concern for public health. The urban heat island (UHI) effect, the phenomenon of higher temperatures in urban areas compared to their rural surroundings (Oke 1982), exacerbates the negative health effects of heat waves and extreme temperatures within cities (Iungman *et al* 2023). Exposure to excessive heat can kill people, and it can also cause general discomfort, respiratory problems, heat cramps and exhaustion, non-fatal heat strokes, and dehydration (Zander *et al* 2015, Heal and Park 2016, Park *et al* 2020). At the population level, extreme heat can lead to a loss of labor productivity and decreased learning (Tan *et al* 2010). These cumulative impacts are expected to increase in frequency and severity as the portion of the human population living in cities is projected to grow over the next few decades and climate change worsens (Ritchie and Roser 2018). The IPCC Sixth Assessment Report on Climate Change anticipates increases in the frequency and intensity of heat waves, with pernicious effects that interact dynamically with urban heat exposure (Krayenhoff *et al* 2018). Future urbanization is also expected to worsen average summer daytime and nighttime wet bulb globe

temperature (a more physiologically complete measure of heat stress than air temperature alone) of 0.5 °C–0.7 °C and up to ∼3 °C in some locations (Huang *et al* 2019).

Within cities, prior research has shown that temperatures (either air or land surface) vary considerably according to multiple factors. The presence (or absence) of urban green space, amount (size of city), form (building height, ratio of building height to width of street canyon, etc), and color (influencing reflectivity of solar radiation) of built-up structures, and intensity of human activity (through increased anthropogenic heat flux) all lead to intra-urban variability of land surface temperatures (LSTs) and the surface UHI effect (Chakraborty *et al* 2019, Benz and Burney 2021). Since urban landscapes are diverse and thermally complex, heat is unevenly distributed (Shi *et al* 2021, Chakraborty *et al* 2022), and research has shown differential access to green space and tree cover along with exposure to denser, built-up urban areas translates into varying potential for heat exposure (Chakraborty *et al* 2019, 2020, Ziter *et al* 2019). This intra-urban variability may also lead to different heat exposure by sociodemographic group, with people of color and poorer urban residents living in neighborhoods with higher land surface and air temperatures than non-Hispanic white and wealthy counterparts in nearly all major U.S. cities (Eliasson and Svensson 2003, Erell and Williamson 2007, Fenner *et al* 2017, Skarbit *et al* 2017, Bechtel *et al* 2019, Benz and Burney 2021, Hsu *et al* 2021, Qian *et al* 2022, Chakraborty *et al* 2023).

Despite the increasing severity and probability of heat waves and extreme heat waves disproportionately impacting urban areas, high-resolution measurements attuned to measure intra-urban heat variability are lacking for most cities (Muller *et al* 2013, Ziter *et al* 2019). Air temperature, feels-like temperature, and other weather data for an entire city are often based on one or a few meteorological stations at best, which are fixed either at a city's center or near an airport, which means they are not necessarily characteristic of the entire city's weather conditions; and cannot provide estimates of variability by neighborhood (van Hove *et al* 2015). Other types of information to contextualize weather data that are relevant include the type of land cover or local climate zone where a station is located. Therefore, most studies on heat-related mortality and morbidity utilizing these single-point weather station measurements make a critical limiting assumption that all people living in a geographic area are all exposed similarly (Rajagopalan *et al* 2020).

To address this limitation, researchers have turned to collecting ground-based data (i.e. air temperature, humidity, etc), which can be time and resource consuming, or using satellite remote sensing data. Numerical weather models of the urban surface and atmosphere are also increasingly being used as alternatives (Krayenhoff *et al* 2021, Wang *et al* 2023). But these models can be computationally expensive when run at fine scales and have their own uncertainties (Krayenhoff *et al* 2021). Satellite measurements have the advantage of inexpensively providing data using consistent methods at high frequency and at policy-relevant geographic scales (Manoli *et al* 2019, 2020, Chakraborty *et al* 2020, Hoffman *et al* 2020, Benz and Burney 2021, Hsu *et al* 2021). Of the factors that pertain to urban heat and are of interest from a public health perspective, however, satellites only measure LST—a significant limitation since this measurement is difficult to apply to pedestrian-level heat exposure (Stewart *et al* 2021, Venter *et al* 2021). In addition, satellite data are limited to certain times of the day (i.e. overpass times) and strictly clear-sky conditions (for optical and thermal imagery), and represent a directional view of land surface temperature dependent on sensor view angle (Shi *et al* 2021). A previous study found weak relationships between satellite-derived LST and ambient heat stress across scales during daytime (Chakraborty *et al* 2022).

To address these data limitations, citizen science methods, which engage local communities in collecting relevant data such as air temperature and humidity at a high temporal and spatial resolution, and private weather stations, have been proposed as possible solutions to provide more detailed data to improve estimates of intra-urban air temperature variation. Since 2017, the National Oceanic and Atmospheric Administration (NOAA) has partnered with more than 60 cities in the U.S. to design heat mapping campaigns that utilize mobile sensors attached to vehicles to map intra-urban heat variability throughout a typically hot summer day (NOAA n.d.). These data have been used to develop city-wide heat maps (Shandas *et al* 2019) and inform local policy regarding heat mitigation measures. Although these campaigns provide baseline measures of intra-urban heat variability for participating cities, Shi *et al* (2021), compared the NOAA citizen heat mapping data collected in Baltimore in 2019 to fixed sensor data and found several limitations, including the need for air temperatures above land surfaces other than roads collected through either bicycle or on-foot traverses to reduce sampling biases.

Recent advances in machine learning (ML) and other non-linear, non-parametric statistical modeling approaches provide potential to combine earth observation data along with ground-based measurements to develop more comprehensive and higher-resolution, continuous maps of heat exposure and stress across an urban area. Several previous studies have combined satellite remote imagery along with ground-based measurements, either citizen-collected data, defined as data collected from hand-held (Schwarz *et al* 2012) or mobile sensors attached to vehicles (Hart and Sailor 2009, Shandas *et al* 2019), weather stations (Ho *et al* 2014, Venter *et al* 2020, Vulova *et al* 2020, Shi *et al* 2021, Zumwald *et al* 2021, Burger *et al* 2022), using

non-parametric ML techniques for classification and regression modeling to predict and evaluate ambient air temperature. None of these studies, however, predict humidity, which is an important contributor to moist heat stress (Steeneveld *et al* 2011).

In this study, we provide a replicable, scalable approach to engaging citizens in utilizing low-cost, consumer-grade, hand-held sensors to map varied land-cover traverses of intra-urban heat, humidity, and heat stress variability. Combined with satellite remote sensing data on land cover, normalized difference vegetation index (NDVI; a measure of greenness), normalized difference built up index (NDBI; a proxy for urban structures), and LST, we develop a ML model to predict air temperature, humidity, and thus moist heat stress for the U.S. town of Chapel Hill, NC. The resulting dataset allows for examination of heat, humidity, and heat stress trends across the entire urban extent, which can aid policymakers in determining key areas to prioritize for heat mitigation measures such as albedo management and shade increase.

2. Materials and methods

2.1. Study area

Chapel Hill, NC is a town with a 2021 population of 61128 (US Census Bureau 2021) located within the Raleigh-Durham-Cary metropolitan statistical area, spanning the counties of Durham, Orange, and Chatham. It is the second largest metropolitan area in the state of North Carolina, with a 2019 population of 2079687. Its total area is 21.75 square miles (56.32 km^2), with a population density of 2871.2 people per square mile or 1107.75 people per square kilometer (US Census Bureau 2021). Chapel Hill is in a humid subtropical climate (Cfa), according to the Koppen-Geiger climate zone classification (Rubel and Kottek 2010), with typical air temperatures averaging between 32 °F and 88 °F, or about 0 °C and 31 °C.

2.2. Air temperature and humidity data collection

Working with local government officials in the Town of Chapel Hill, we identified five areas from the Town of Chapel Hill's Future Land Use Map and Extreme Heat Resiliency Assessment (Town of Chapel Hill 2020) for citizen scientists to map. These five neighborhoods—(1) Franklin St., (2) Meadowmont; (3) Southern Village; (4) University Place; and (5) Chapel Hill North were areas in Chapel Hill with relatively less tree cover, a higher density of built infrastructure (i.e. buildings, roads) and population than the rest of Chapel Hill (Town of Chapel Hill 2020). Each were identified as ranking 'high' on the Town's Extreme Heat Vulnerability rating, meaning the town has gauged these areas with the highest number of sensitive populations and highest percentage of developed land cover (>85%), with a low amount of tree canopy coverage (<33%) (Town of Chapel Hill 2020).

Two to three mile pedestrian routes were determined for each neighborhood to include a variety of land-cover types, from sidewalk, shaded/unshaded, pavement, and grassy areas (see figure S1). We planned routes first using Google maps to determine approximately 2–3 mile walking routes (or roughly 1 h) that avoided potentially dangerous areas (i.e. highways or road crossings without designated crosswalks). Our team then visited each neighborhood and tested each route and made adjustments to ensure traverse paths were safe for pedestrians and walkable. We developed two routes per neighborhood to maximize mapping coverage in the event that multiple volunteers showed for each time period. Citizen volunteers were recruited using email lists sent by the Town of Chapel Hill and Museum of Life and Science, as well as from the UNC Chapel Hill community. Since our Institutional Review Board exemption only covered adults over the age of 18, our volunteers had to be over the age of 18 to participate, even though no sensitive or personal data were collected. Participants were emailed walking routes prior to the study and were asked to assess their own ability to walk outside on a hot day while carrying a smartphone and Pocketlab sensor for at least an hour and a half.

We worked with NOAA to identify a hot day in August when rain was not forecasted to conduct the heat mapping campaign, which we identified as Saturday, 28 August 2021. On the mapping campaign day, we set up five hubs in each neighborhood as meeting points for volunteers to receive the sensors and mapping route instructions. We provided instructions to each volunteer to ensure that they held the sensors in their hands and that the sensor ports were unobstructed. In total around 40 citizen volunteers participated in two mapping sessions: from 2–3 pm EDT, UTC −05:00 and 5–6 pm to capture two time periods' air temperature and humidity measurements (see table 1). Thirteen volunteers participated from 2–3 pm, and 15 volunteers participated from 5–6 pm. For some neighborhoods, like the Franklin Street neighborhood, we had some instances of multiple volunteers mapping the same route during a session, but for most routes, we split volunteers so that they were not walking the same route. When more than two volunteers showed up to map a given route during a session, we had them start a route at different ends of the routes, to take into account any temporal fluctuations that might have occurred during the hour or so it took for them to map a route. Participants were able to select which side of roads and walkways they mapped. Since no participants

Table 1. Descriptive statistics of air temperature and RH data collected by citizens in all five neighborhoods.

Session	No. of volunteers	Variable	Count	Min	Mean	Median	Max	St.dev.
2–3 pm	12	Air temperature (°C)	20 551	33.3	37.8	37.7	42.6	2.0
2–3 pm	12	Relative humidity (%)	20 471	30.0	40.4	39.5	58.9	5.0
5–6 pm	12	Air temperature (°C)	12 996	29.9	34.7	35.1	39.2	1.6
5–6 pm	12	Relative humidity (%)	12 552	33.8	45.2	45.0	58.2	4.0

Table 2. PocketLab sensor specifications (from PocketLab Weather Specifications n.d.).

	Relative humidity	Air temperature
Range	0%–100%	−40 °C–85 °C (−40 °F–185 °F)
Resolution	0.02%	0.01 °C (0.02 °F)
Absolute accuracy	3%	0.5 °C (0.9 °F)

volunteered for Southern Village's 5–6 pm time period, we only recorded measurements for the 2–3 pm session for this neighborhood. table S1 shows the type and percent of each land cover that participants mapped during the two sessions.

2.3. Pocketlab sensors

For data collection, we utilized PocketLab(™) Weather sensors (PocketLab Weather Specifications n.d.) containing a three-in-one BME280 module that embeds sensors to gauge air temperature, relative humidity (RH), and barometric pressure. In our study, we set the PocketLab sensors to collect geographic location (latitude and longitude), ambient air temperature and RH data every 1 second (table 2). To evaluate the overall accuracy and validity of the PocketLab sensors to record air temperature and humidity, we conducted a comparison study with the PocketLab sensors and the NC State Climate Office's research-grade local weather station in Chapel Hill, located at the Old Horace Williams Airport (North Carolina State Climate Office, NC State University n.d.), which is located slightly north of the UNC-Chapel Hill campus (35°56′02.9″N 79°03′43.0″W). The weather station instruments for air temperature and RH are at 2 m height (further details on the sensor manufacturer and models can be found at the North Carolina State Climate Office, NC State University (n.d.) website for the Old Horace Williams Airport weather station). We were only able to conduct the comparison study on the afternoon of 6 February 2023, which had an average air temperature of 48.6 °F, min air temperature of 37.8 °F and max air temperature of 61.2 °F. We compared PocketLab sensors for each of the following conditions: handheld (∼1 m in height; 3 total sensors) in non-shade conditions and non-handheld (<1 m in height; 2 total sensors) on a table in both shade and non-shade conditions. For the sensors that were handheld, we minimized walking around the weather station since the purpose of the comparison was to see how well the sensors aligned with the weather station data.

To compare the measurements, we examined correlation coefficients between the PocketLab and weather station data, and conducted Analysis of Variance (ANOVA) statistical tests to compare the mean measurements of the PocketLab sensors and the weather station for the validation study (tables S2 and S3). Overall, we found that the Pocketlab sensors were responsive to changes in air temperature, aligned well with the weather station's absolute air temperature measurements ($r = 0.48$ comparing handheld sensor and weather station; see figure S2; table S2), and the way in which they were carried did not significantly ($p > 0.1$) impact their consistency s (−0.0006 °C; se: 0.045; table S3). During our comparison, a large drop in air temperature was observed after 15:30 from all sensors, which may have been due to clearing of previous cloudy conditions. The RH measurements did not align as closely with the weather station measurements ($r = 0.05$ comparing handheld sensor and weather station data; table S2) and showed higher significant differences compared to the handheld sensors (2.870%; se: 0.060). These differences in the RH values are to be expected, since RH is more difficult to measure, and the weather station sensor for RH is shaded, compared to our handheld sensors, which were gauging RH in sunlight. Regardless, since RH is strongly affected by air temperature, which we show in supplementary figure S3, we determined that our method using the PocketLab sensors can still provide measures of overall heat stress even if there might be underestimations in RH.

Participants were asked to place the PocketLab sensors on the palm of their hands (∼1 m above ground) so that the air temperature and humidity ports, which are located on the sides of the sensor casing, were not obstructed. During our preliminary testing of the devices, we determined that body temperature does not warm the device or lead to an increase in air temperature measurement. We asked participants to allow the

Table 3. Overview of spectral data collected.

Date type	Spatial resolution	Temporal resolution	Data source	Data acquired time	Data layer calculation
Land surface temperature (LST)	30 m (resampled from native resolution of 100 m)	16 d	Landsat 8 (USGS 2021)	24 August 2021; Average LST of August and September, 2021	—
Land use/land cover (LULC)	10 m	—	ESA Worldcover 2020 (Zanaga *et al* 2021)	2020	—
Satellite-derived indices	10–20 m	8 d	Sentinel-2 (Zanaga *et al* 2021)	on 5 September 2021	NDVI = (NIR − Red)/(NIR + Red) NDBI = (SWIR − NIR)/(SWIR + NIR)
Aerial imagery	1 m	2–3 years	The National Agriculture Imagery Program (NAIP) (Earth Resources Observation And Science (EROS) Center 2017)	2018	—

sensors to acclimate for a few minutes before recording. A csv data file was saved after each participant finished a mapping session.

After data were recorded and compiled, we observed some data quality issues related to some sensor, smartphone, or user errors. For example, we noticed that the first and last two minutes of a participant's data collection exhibited more fluctuations as participants were initializing their smartphone app and pairing with the PocketLab sensor. We therefore removed these data points. We also removed outliers greater than three standard deviations around the mean of each data file following the study on UHI by Chapman *et al* (2017) using crowdsourcing data. Since some volunteers mapped the same route during the same time of day, we averaged geographically overlapping readings (data with the same latitude and longitude value) to avoid data duplication. Table S4 documents the total number of participants whose data were used for the study after all quality control.

2.4. Spectral and remote sensing data collection

To develop city-wide maps of air temperature, humidity, and estimates of moist heat stress, we collected multi-spectral satellite remote sensing imagery, including LST, and land use/land cover (LULC). Since local temperature is related to surface characteristics (Chen *et al* 2006, Chakraborty *et al* 2019), we combined some of this multi-spectral data to calculate common indices for green (Chen *et al* 2014), and built-up area, NBDI (Zha *et al* 2003) (table 3). We also collected high-resolution aerial imagery from the USGS National Agriculture Imagery Program (Earth Resources Observation And Science (EROS) Center 2017).

2.4.1. Landsat LST data

Studies have found positive correlation between satellite-derived LST and air temperature (Mildrexler *et al* 2011, Zhang *et al* 2014), particularly during nighttime (Chakraborty *et al* 2022), and LST data have been widely used as a proxy for urban heat exposure in the past (Chakraborty *et al* 2019, Manoli *et al* 2019, Hoffman *et al* 2020, Hsu *et al* 2021, McDonald *et al* 2021, Mentaschi *et al* 2022). In this study, the LST data from satellite sensors Landsat 8 and MODIS were examined, since both data sources have their advantages and trade-offs. After comparing the spatial and temporal coverage of both sources of data, we decided to use the LST data from Landsat 8 for its higher spatial resolution (100 m) and temporal coverage for the time period we sampled. One cloud-free scene closest to the citizen-collected data on 24 August 2021 at 15:53 UTC from Landsat 8 was selected. Based on the National Weather Service (NWS), the selected Landsat 8 scene date (24 August 2021) and campaign data collection date (28 August 2021) have similar daytime air temperature ranges, where the weather station reported air temperatures for Chapel Hill range from 19.4 °C to 32.8 °C vs. 22.2 °C to 33.9 °C respectively for the two dates. To reduce the dependency on a single date

LST data and add robustness to the model, averaged LST data for August and September 2021 was also added as an independent variable. We filtered out any scene with 10% of cloud coverage or more during this time window.

2.4.2. LULC

Studies have found that different LULC types directly result in air temperature variance at and near the earth's surface because of different biophysical effects, which include land-surface energy exchanges (Mahmood *et al* 2013, Li and Gallagher 2019). Since air temperature data from our campaign is measured in 2021 and our collected land cover (i.e. built-up, grassland, tree, barren/sparse, and cropland) data is from 2020, we assume that patterns in near surface air temperature and its anomalies are consistent temporally. We detail the number and percentage of land-cover mapped in study in table S1.

2.4.3. Satellite-derived indices

Although we make the assumption that LULC types have a consistent impact on the near surface air temperature and the anomalies between the years 2020 and 2021, there might be seasonal variance in air temperature and potential LULC change. We incorporate NDVI and NBDI—two commonly-used satellite derived indices that have established relationships with LST (Chakraborty *et al* 2019) to aid in prediction for air temperature and RH. NDVI indicates vegetation coverage or 'greenness' and can distinguish between different LULC types and aid in understanding the relationship between vegetation and local climate (Rouse *et al* 1974, Ziter *et al* 2019). NDBI, developed by (Zha *et al* 2003), indicates built-up area as well as captures any new urban development of the study area between 2020 and 2021. Both indices are calculated with the satellite image from Sentinel-2 on 5 September 2021—the latest available and cloud free scene since data collection (table 3).

2.4.4. The national agriculture imagery program (NAIP)

The National Agriculture Imagery Program acquires aerial imagery at 1 m resolution during the 'leaf-on' season in the continental U.S starting from 2003. NAIP imagery is available every 2–3 years for each state. This high-resolution multispectral product can provide the details of land cover/land use, building structure, etc. for urban areas. All four bands of NAIP (Red, Green, Blue, Near-infrared) imagery from year of 2018 were evaluated as input variables.

2.5. Model selection and data preprocessing

We evaluated several ML models to predict air temperature, RH and thus moist heat stress across the entire spatial extent of our study area: multilinear regression (MLR), support vector regression (SVR), random forest (RF), and XGBoost. The MLR based model is one of the most popular multidimensional interpolation algorithms in urban studies for its ability to capture the correlation between local temperature and UHI with the intra-urban environment variances (Voogt and Oke 2003, Kousis *et al* 2021). We used MLR as the baseline model to compare the performance of other models because it is easy to implement. SVR with a radial kernel is an extension of the classifier support vector machine (SVM), which has also been used to study urban environments (Lai *et al* 2021). Studies on at or near surface air temperature modeling have found this estimator results in a lower root mean square error (RMSE) compared with other more complex algorithms (Chevalier *et al* 2011, Paniagua-Tineo *et al* 2011). RF regression is a nonparametric ML algorithm and known for picking up the non-linear relationship between the input variables (Breiman 2001). RF has been applied to multiple studies to produce high-resolution LST from relatively coarse data and shown solid performance or predicting spatially continuous air temperature when combined with satellite-derived information (Hutengs and Vohland 2016, Yang *et al* 2017, Venter *et al* 2020). Finally, XGBoost is one of the state-of-art regression models based on the gradient boosting decision tree model. We used a standard 80/20 split to create training and testing subsets of our data, meaning we used 80% of our data to train our model and held out 20% of the original data to test model performance. We used RMSE and r^2 as the model comparison metrics to evaluate how selected models performed on both training and test dataset. For instance, a low RMSE and high r^2 value indicates the model trained using 80% of the data was able to predict the 'unseen' 20% of the data.

We implemented the ML algorithms in the R Statistical Computing Environment (version 4.2). MLR is included within R version 4.2; SVM and RF are from R package caret version 6.0-92 (Kuhn 2008); and XGBoost from XGBoost R package version 1.6.0.1 (Chen and Guestrin 2016). All selected models, except MLR, are controlled by a set of hyperparameters, which constrain the learning process (Yang and Shami 2020). The hyperparameters in SVM are the penalty parameter for the error term, which controls the margin

of the decision boundary and sigma, which defines the distance between support vectors and the decision boundary. In RF, we tune the model with the number of trees, number of variables randomly sampled as candidates at each split, the minimum number of observations in a terminal node, and the matrix to use to split the data into two branches. In XGBoost, hyperparameters include the maximum depth of the tree, the learning rate, the minimum sum of weight in a node, and minimum loss reduction. We tuned each model with five-fold cross validation on a range of parameters to achieve the best parameters resulting in lowest RMSE and highest r^2 on the data's test set.

Two data preprocessing approaches were used in our study. First, two sets of variables were prepared and evaluated with all four ML algorithms for modeling air temperature. The first set with all variables is described in table 3, while the second set removed all NAIP bands. Although the NAIP data gives a high level of detail of the study area, the latest available data is from 2018 and the vegetation and urban extent information may have changed in the past 3 years. Meanwhile, in the input variables, Sentinel-2 NDVI and NDBI provide similar information about vegetation and built-up density and are more temporally consistent. Second, we scaled and centered the continuous variables with a preprocessing function (*prePrecoss*) from R package caret version 6.0-92. This preprocessing function resulted in a new dataset with rescaled and normalized input variables, which was achieved by subtracting the mean from each input variable and dividing by its standard deviation, so that each variable has a mean of zero and a standard deviation of 1. This approach is applied to the dataset for SVM model training only. SVM models are sensitive to feature scaling, as SVM takes input variables to find the margins of hyperplanes to best separate the data. The distances between data points with and without scaling are different. Thus, the SVM can be biased if trained with un-scaled data.

2.6. Model evaluation and final selection

Supplementary tables S5 and S6 shows the results of air temperature prediction from each model with the best parameters. The RF and XGBoost model have similar test RMSE and r^2 for both sessions, but outperform the other two models in terms of RMSE and r^2 For computation time, RF model training is about seven times faster than XGBoost training on an 8-core Windows PC. Comparing the results from different sets of variables with the same model, the set without NAIP yielded better results in both sessions (table S6). Based on the results from both sets of variables with all four selected models, we decided to use the RF model and the dataset without NAIP for its overall predictive performance and fast computation for our dataset. The best performing air temperature model results in a RMSE $= 0.76\ ^\circ$C, $r^2 = 0.86$ on the test data for the 2–3 pm session and RMSE $= 0.48\ ^\circ$C, $r^2 = 0.91$ on test data for the late afternoon (5–6 pm).

We used the same input variables as the air temperature model to build the RH models. Table S7 shows the results of the RH predictions from all four ML algorithms with their best parameters. The XGBoost is the best performing humidity model, which results in a RMSE $= 1.46\%$, $r^2 = 0.92$ on the test dataset for 2–3 pm session and RMSE $= 1.29\%$, $r^2 = 0.89$ on the test dataset for the 5–6 pm session.

To better understand the key predictors for air temperature and RH, we ran Monte Carlo cross-validation (Xu and Liang 2001, Chakraborty *et al* 2021) with the model run 50 times with random training and validation splits and evaluated the input features with permutation feature importance scores in the air temperature model and 'Gain' scores in the humidity model.

2.7. Heat stress calculation—Humidex

To get an estimate of the human physiological response to heat, we use the Humidex, an operational metric of moist heat stress (Masterton and Richardson 1979). Humidex is calculated for each pixel using the estimates of air temperature (T) and RH from the ML models. First, we calculate dew point temperature (T_d) from the following equation:

$$T_\mathrm{d} = \frac{243.04 \times \left\{\ln\left(\frac{\mathrm{RH}}{100}\right) + \frac{17.625 \times T}{243.04+T}\right\}}{17.625 - \left\{\ln\left(\frac{\mathrm{RH}}{100}\right) + \frac{17.625 \times T}{243.04+T}\right\}}. \tag{1}$$

Then Humidex is computed from:

$$\mathrm{Humidex} = T_\mathrm{d} + 0.5555 \times \left(6.11 \times \mathrm{e}^{5417.753 \times \left(\frac{1}{273.16} - \frac{1}{273.15+T_\mathrm{d}}\right)} - 10\right). \tag{2}$$

Note that Humidex is a unitless proxy for how hot it feels under shade, and results should be interpreted keeping this caveat in mind.

2.8. Air temperature, humidity, and Humidex prediction

After model evaluation, we re-trained the model with all of the citizen-collected data to generate the final model. The original input variable layers have different spatial resolutions. NDVI and NDBI from Sentinel-2, in addition to the land cover data, are 10 m spatial resolution, while the LST layers are 30 m spatial resolution (though these are resampled from the native 100 m resolution). To produce the predicted air temperature and humidity layers for the whole study area and retain the information from the LULC dataset, we chose 10 m as our target resolution. We resampled LST layers from 30 m to 10 m using the *rasterio (version 1.0.21)* package in Python. We then used the final model with the input variables that cover the whole study area at 10 m spatial resolution as inputs to predict the air temperature and RH. We also used the predicted air temperature and RH to calculate the Humidex as a proxy to simulate pedestrians' thermal comfort under shade. Finally, the predicted air temperature, RH, and Humidex results were converted into GeoTIFFs with the *rasterio (version 1.0.21)* package in Python.

2.9. Sociodemographic analysis

To explore disparities by sociodemographic group in Chapel Hill, replicating similar analyses in previous studies (Chakraborty *et al* 2019, Benz and Burney 2021, Hsu *et al* 2021) that used primarily LST, we extracted census tract and block group demographic data for Chapel Hill from the 2020 ACS 5 year Data Profile (US Census Bureau 2020). Since some areas of Chapel Hill's jurisdictional limits cover census tracts in other counties, we included additional census tracts in this analysis that overlap, which explains why our urban extent slightly differs from Chapel Hill's own administrative boundaries (supplementary figure S4). Median household income and race, according to the U.S. Census's categorizations of White, Black, Asian, Pacific Islander, Native American, and Other were extracted. For examining racial disparities, we use census tracts, the smallest level of aggregation for which the race data are available. To examine income-based disparities, we use census block groups.

The observed LST and predicted air temperature, humidity, and Humidex are aggregated to the census tracts and block groups. Associations between income and the heat exposure metrics are examined using Pearson's correlations in R language. To examine race-based disparity, we calculated population-weighted heat exposures for each census group, following the method in Hsu *et al* (2021).

3. Results

3.1. Local-scale air temperature and humidity

Although recorded weather station data from the NC State Climate Office Chapel Hill weather monitoring station showed measured air temperature maximum of 33 °C and a minimum air temperature of 30 °C for 28 August 2021 (North Carolina State Climate Office, NC State University n.d.) our citizen volunteers mapped higher air temperatures in all five neighborhoods, ranging from 33.3 °C to 42.6 °C in the 2–3 pm session, and 29.9 °C to 39.2 °C in the 5–6 pm session (figure S5). The highest air temperatures were measured during the early afternoon session in the Franklin St. (38.7 ± 2.0 °C) and University Place (39.1 ± 1.4 °C) neighborhoods, which feature a high proportion of asphalt parking lots and pavement, compared to surrounding areas and other neighborhoods. Other weather conditions for that day include average 79% humidity and average 3.4 mph wind speed (North Carolina State Climate Office, NC State University n.d.).

Volunteer-collected RH ranged from 30% to 58.9% in the 2–3 pm session and 33.8%–58.2% for the 5–6 pm session (figure S5). Although the RH ranges are similar between the two time periods, we still observed differences in the humidity distribution between hubs. Franklin St. (38.6 ± 3.4%) and University Place (38.5 ± 4.4%) are neighborhoods with the lowest recorded RH. The Meadowmont neighborhood, located in the southeast area of Chapel Hill, has relatively high humidity in both time periods (48.7 ± 2.3% for the 2–3 pm session; and 44.2 ± 2.3% in the 5–6 pm session), which may be related to high vegetation and tree coverage in the neighborhood. Chapel Hill North showed the largest differences in humidity between the 2–3 pm session (39.37 ± 3.23%) and 5–6 pm session (47.25 ± 6.09%), but this difference is likely due to the fact that not enough volunteers showed up to map both routes (see figure S1), and so different routes were mapped. Volunteers mapped the built-up and parking lots during the 2–3 pm session, while they mapped more residential areas with high tree coverage during the 5–6 pm session (see figures S1(c) and (d)). We further discuss the limitations of this difference in the Discussion.

3.2. Predictors of air temperature and RH

Figure 1 shows the importance scores of all features in the air temperature and RH models for the 2–3 pm session. The permutation feature importance indicates how much the model depends on each input feature determined by how much the model score decreases when a certain feature is removed. The gain value represents the fractional contribution of each input feature to the model calculated by taking each feature's

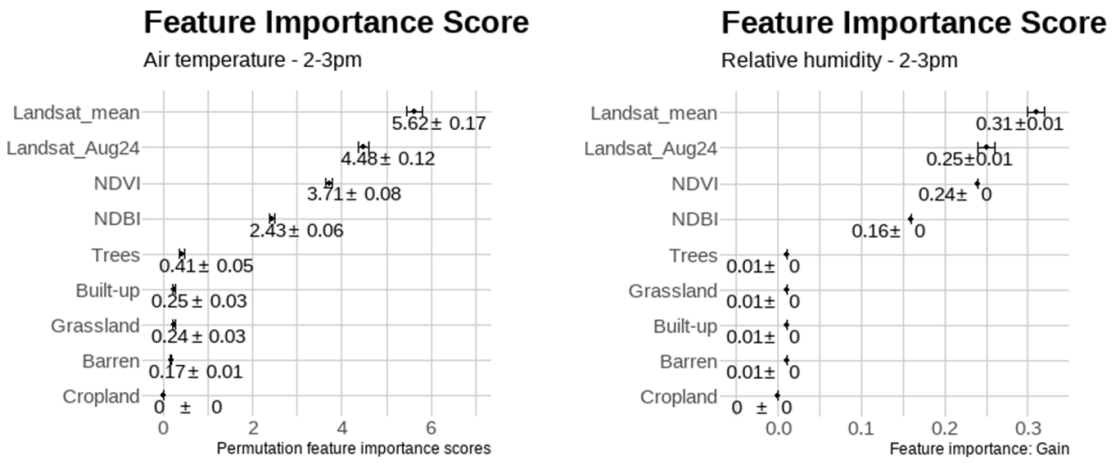

Figure 1. Permutation feature importance scores for air temperature (left) and for RH (right) models (based on the 50 random training and validation splits).

contribution (improvement in accuracy) for each tree. A higher value in both permutation feature importance or the gain value indicates a higher importance of this feature for generating a prediction. The Landsat 8 two-month averaged LST is the top ranking predictor for both the air temperature and RH models. The next most important features include the Landsat 8 LST data closest to the citizen-science data collection date. NDVI and NDBI also have high contributions to both the air temperature and humidity models. Among all five land cover/ land use dummy variables, whether an area is classified as the 'tree' land cover class has the most predictive effect in the air temperature models.

Our recorded RH and air temperature show strong negative correlation (figure S3), which is consistent with expectations. We provide measured RH, absolute humidity (calculated from our measured air temperature and RH), and dew point temperature (Tdew) in supplementary tables S7–S9, which show that RH is indeed determined by air temperature, while absolute humidity is less dependent on air temperature.

3.3. Modeled air temperature, humidity, and Humidex

Figure 2 shows predicted air temperature, RH, and Humidex for the 2–3 pm and 5–6 pm sessions in our study area. Overall, we found the 2–3 pm session is hotter and drier than the 5–6 pm session, with the predicted mean air temperature as 35.8 ± 0.7 °C for 2–3 pm sessions and 31.9 ± 1.9 °C for 5–6 pm sessions; the predicted mean RH as 38.6 ± 1.5% for 2–3 pm sessions and 56.7 ± 1% for 5–6 pm sessions. Also we found different patterns of the heat distribution between the 2–3 pm and 5–6 pm sessions. The 2–3 pm session tends to be more uniformly hot over the study area, where 90% of the area's predicted air temperature ranges from 35.2 °C (5th percentile) to 37.3 °C (95th percentile) and has a standard deviation of 0.7 °C. In the 5–6 pm session, the predicted air temperature varies more across the study area, with 90% of the area ranging from 30.4 °C (5th percentile) to 35.5 °C (95th percentile) with a 1.9 °C standard deviation. As a function of air temperature and RH, the Humidex shows similar patterns as predicted air temperature and RH on both sessions. In the 2–3 pm session, the Humidex ranges from 41.7 (5th percentile) to 45.2 (95th percentile) for 90% of the study area, with an average value of 42.9 ± 1.1. In the 5–6 pm session, the Humidex varies from 38.8 (5th percentile) to 48.5 (95th percentile) for 90% of the study area, with values on average 41.6 ± 3.5. This summary analysis was applied to the land cover types of interest only (see figure 3(b)).

3.4. Land cover analysis

Based on the ESA Worldcover 2020 data, which may not necessarily be precisely attuned to capture all of the variation of land cover within an urban area, our study area has seven land cover types: 82.8% of the total area is classified as Trees, followed by 9.8% Built-up area and 6.2% Grassland. The remaining 1.2% of the area is classified as Barren or sparse vegetation, Open water, Cropland, and Herbaceous wetland, which our volunteers did not sample and we therefore did not include in our final analysis (see figure 3(b) and table 4). In table 4, we determined predicted 2–3 pm air temperature by land cover classes, and calculated each land cover class's air temperature difference from the Tree cover class, since it has the coolest predicted air temperature compared to other land cover classes. On average the Built-up area is 1.2 °C hotter than tree-covered areas and Grassland is 0.7 °C hotter. Figure 3(a) shows the mean air temperature differences from the tree class. Half of the study area is hotter than the mean air temperature of the Tree class, with 31.5% of the total area less than 1 °C hotter; 13.3% area 1 °C–2 °C hotter; and only 2% area 2 °C hotter.

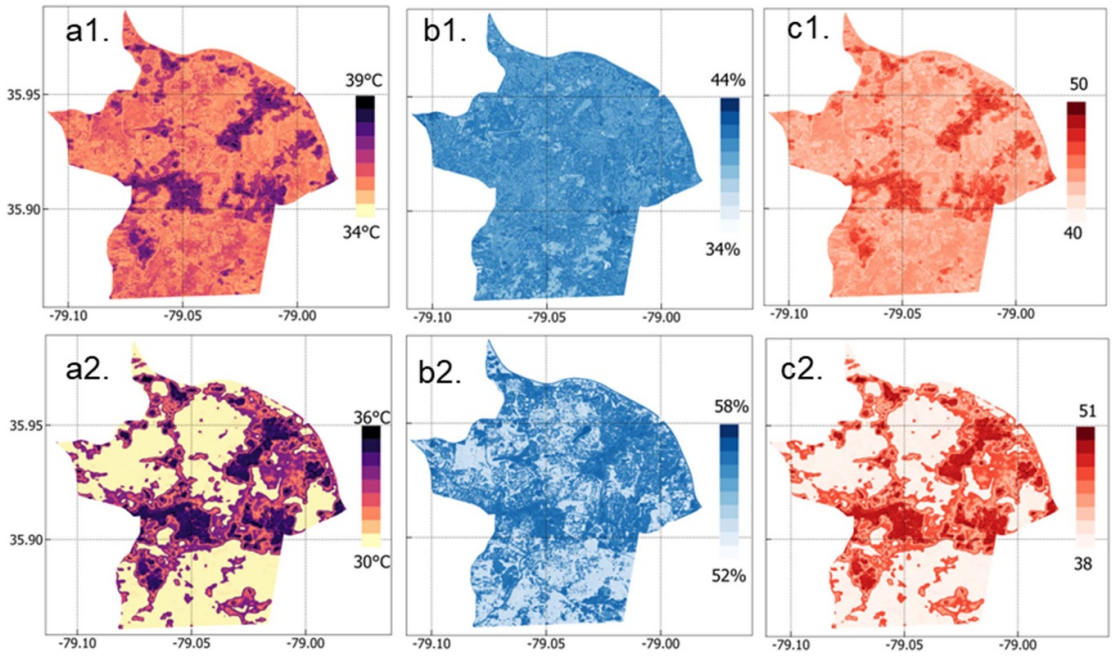

Figure 2. Maps of predicted air temperature, RH, and Humidex of the study area (10 m spatial resolution). (a1) predicted air temperature of 2–3 pm session; (b1) predicted RH of 2–3 pm session; (c1) Humidex of 2–3 pm session; (a2) predicted air temperature of 5–6 pm session; (b2) predicted RH of 5–6 pm session; (c2) Humidex of 5–6 pm session.

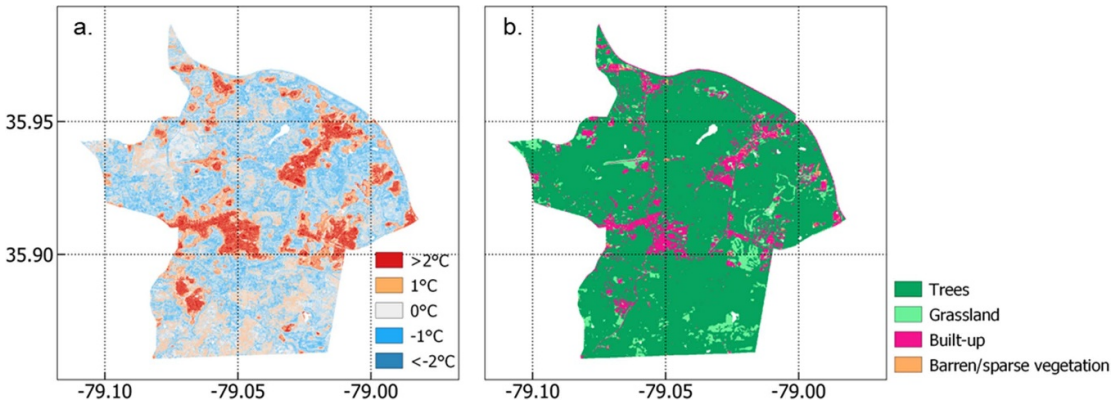

Figure 3. Maps of predicted air temperature differences (°C) and land cover (10 m spatial resolution). (a) Air temperature differences (°C) from mean air temperature of the area covered by trees; (b) land cover types from ESA Worldcover 2020.

Table 4. Descriptive statistics of predicted air temperature by land cover and the mean differences from the tree land cover class.

Land cover	Count	Area %	Min (°C)	Mean (°C)	Median (°C)	Max (°C)	St. dev. (°C)	Difference (°C)
Trees	724 863	82.8%	34.3	35.8	35.6	38.9	0.6	0.0
Built-up	85 920	9.8%	34.5	37.0	37.1	38.9	0.8	1.2
Grassland	53 978	6.2%	35.0	36.5	36.3	38.9	0.6	0.7
Barren/sparse vegetation	6585	0.8%	34.4	37.1	37.1	39.0	1.0	1.3

Note: all land cover data is from ESA Worldcover 2020 (Zanaga *et al* 2022).

3.5. Sociodemographic analysis by census tracts

While some obvious spatial patterns in heat, greenness, race, and income can be observed by census tract (figure 5), we did not find substantial differences in the exposure heat metrics between these different racial demographic groups (supplementary table 11), except a minor difference in LST between White and Black (32.4 °C versus 33.2 °C). The hottest census tracts (in figure 5, those with the highest average Humidex) are found in the downtown area of Chapel Hill, which also has the lowest NDVI (a proxy for tree cover and greenness). These census tracts correspond with the University of Chapel Hill and Franklin Street neighborhoods, which have high student populations and explains why these areas have the lowest income

per capita (Krizek 1995). According to the 2020 ACS Census, 76% of Chapel Hill residents identify as White, 13% as Asian, 10% as Black, and less than 2% as Native American or Other (see table S12).

4. Discussion

This study provides a replicable method for modeling spatially-resolved air temperature, humidity, and moist heat stress over an urban area and engages citizens to gather air temperature and humidity data, which are used to train the models. Since high-resolution, individual-scale air temperature, humidity and heat stress data are difficult and costly to monitor, this study sought to develop an approach using low-cost, consumer-grade sensors combined with satellite remote sensing data and ML to map urban heat over various land-cover classes to understand intra-urban spatial variability of heat at a relatively high resolution (10 m). We anticipate that our approach can be replicated with reasonable confidence by others, as the Pocketlab sensors we employed are consumer-grade and accessible for purchase, and the satellite-remote sensing data and land-cover data used as predictive variables in our model are publicly available free of charge. Compared to other methods of local or 'hyperlocal' heat and weather monitoring, such as the use of crowdsourced private weather stations (Venter *et al* 2020), the method we employed can potentially capture more intra-urban variability, since measurement routes can be customized and designed, whereas the use of private monitoring station data is subject to individual users' locations and potential sampling biases and are generally fixed. Private, crowdsourced monitoring stations, however, have the ability to take repeated measurements over a longer time period with minimal human intervention.

This study contributes to an emerging area of scholarship on estimating intra-urban variability in microclimate using ML (Venter *et al* 2020, Hanoon *et al* 2021, Chen *et al* 2023). Moreover, it is also one of the first to use these methods to predict the spatial variability of RH, and thus estimates of ambient moist heat stress. We find that, when combined with other ancillary information, satellite-derived LST can be a strong predictor of ambient air temperature, even though using LST directly as a proxy for urban heat exposure may be misleading (Chakraborty *et al* 2022, Turner *et al* 2022). Overall, since our ML model identified LST from both the Landsat two-month average and the closest time period as being the variables contributing the most predictive power to our model, this means that the air temperature variability is embedded within LST variability, as reflected in the feature importance scores (figure 2). The resulting datasets and maps can be utilized within various decision making contexts, from individuals determining where to live or urban planners and policymakers developing urban heat mitigation measures to protect citizens. We discuss some of our key findings and their implications for policy as well as some of the study's limitations.

4.1. Local-scale heat exposure and stress

Our findings here show that individuals may be exposed to higher levels of heat and heat stress than what weather station data provide, and this heat varies according to land cover type and throughout the day. With our volunteer-collected data, we found citizen-measured air temperatures ranged from 33.3 °C to 42.6 °C—on average higher than what local weather monitoring stations recorded for the day of data collection (32.2 °C–32.4 °C from 2–3 pm; and 32.5 °C–33.1 °C from 5–6 pm) (North Carolina State Climate Office, NC State University n.d.). According to weather.gov, the highest recorded air temperature in Chapel Hill, NC is 105 °F or 40.5 °C. We observed multiple instances of citizen-measured air temperature readings higher than 40 °C or 104 °F, which the NWS classifies as 'extreme caution' (32 °C–41 °C), and several readings that fall into the NWS's 'danger' zone (41 °C–54 °C). As illustrated in figure 4, which displays differences in predicted air temperature and humidity in our five mapped neighborhoods, the hottest areas, Franklin St. (average predicted air temperature 36.1 $\pm$ 0.79 °C during the 2–3 pm session) and University Place (36.4 $\pm$ 1.0 °C during the 2–3 pm session), tended to intersect with census tracts that had the lowest greenness or NDVI (mean 0.54 for Franklin Street and 0.46 for University Place, compared to 0.61 for Southern Village) and the least amount of classified tree land cover class (less than 0.5 for Franklin Street and 0.65 for University Place, compared to 0.88 for Chapel Hill North and 0.85 for Southern Village). The Franklin Street area is also located within census tracts that have the highest average built-up area (average 0.51, compared to 0.09 in Chapel Hill North). The Franklin Street area corresponds with the census tracts with the highest average Humidex values out of all neighborhoods (figure 5), likely due to its relative lack of tree cover and greater built extent compared to other areas of Chapel Hill.

4.2. Sociodemographic patterns

We did not find substantial variation in air temperature, humidity or heat stress by racial demographic, as other studies have found for different samples (Harlan *et al* 2006, Wong *et al* 2016, Voelkel *et al* 2018, Chakraborty *et al* 2019, Benz and Burney 2021, Hsu *et al* 2021), probably because the majority of census tracts in Chapel Hill are predominantly White (figure S6). In our analysis, the lowest income areas also

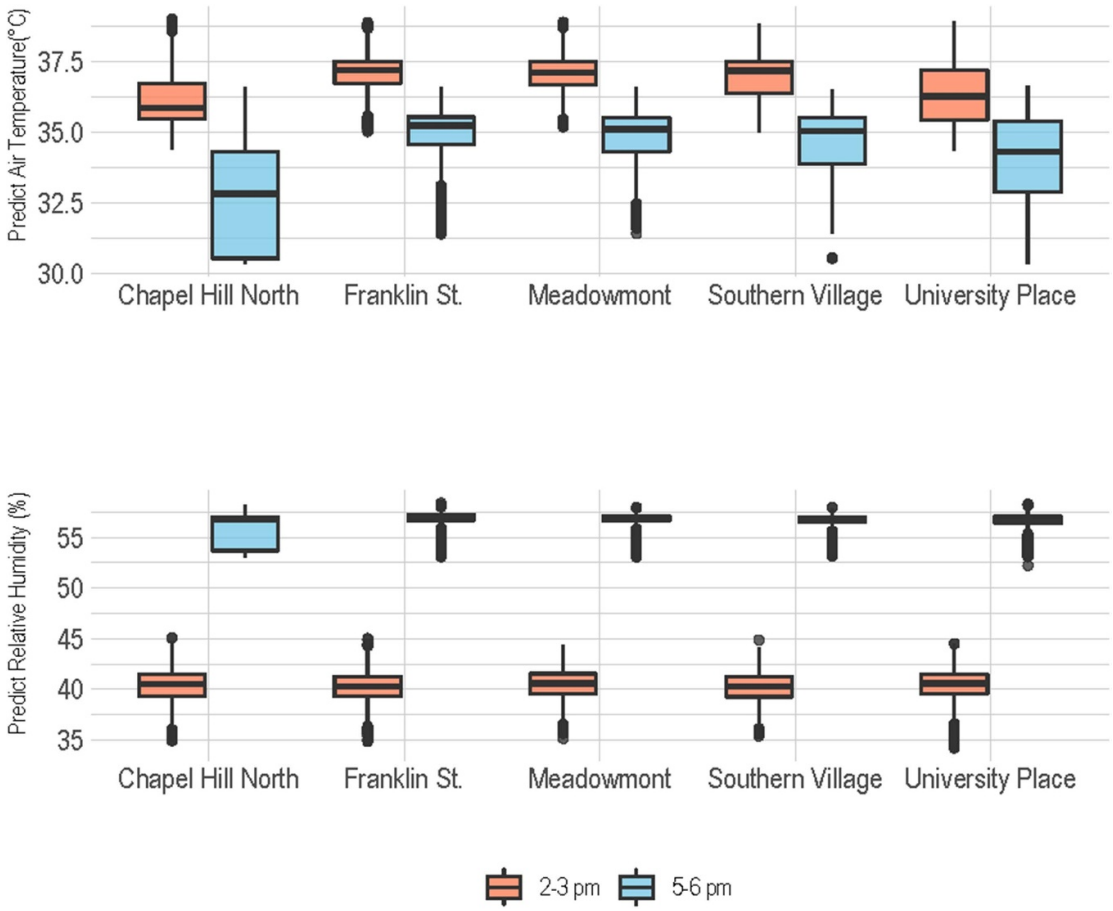

Figure 4. Distributions of predicted air temperature (top panel) and RH (bottom panel) for five neighborhoods mapped in Chapel Hill.

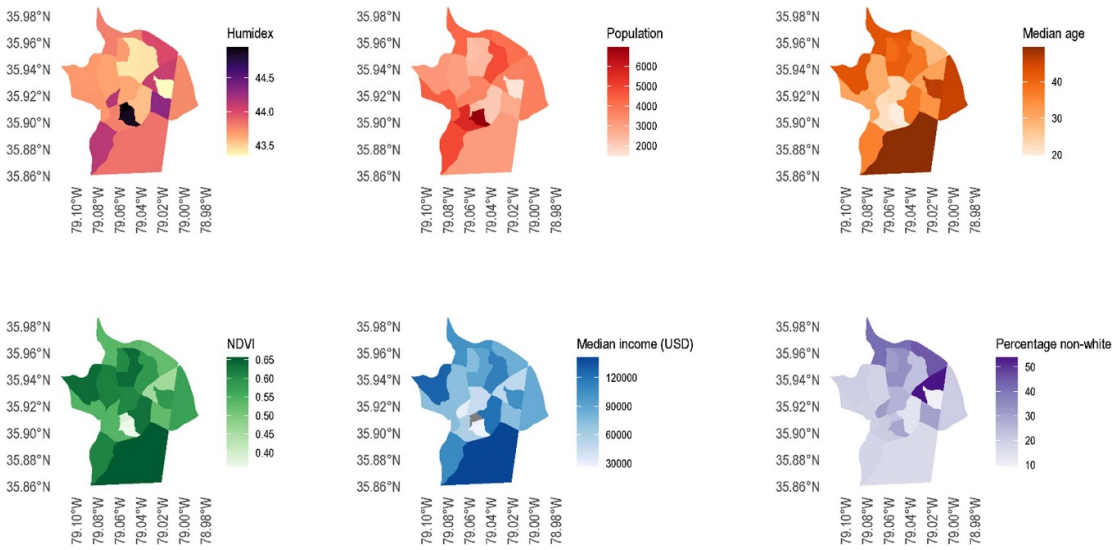

Figure 5. Maps of heat (Humidex), population, median age, greenness (NDVI), median income, and race by census tract in Chapel Hill for the 2–3 pm session. The Humidex and NDVI are both unitless.

correspond to census tracts with a significant proportion of the population between ages 20–24: the Franklin Street census tracts have more than 43% of the population in this demographic, compared to census tracts in the Chapel Hill North neighborhood, with only around 5% of the population in this age bracket (figure 5). Since the University of North Carolina at Chapel Hill surrounds the Franklin Street neighborhood, the high

percentage of people aged 20–24 living in these census tracts makes sense and explains the relatively lower income compared to neighboring census tracts.

While RH shows an inverse relationship with air temperature (figure S3), which is expected within urban areas due to urbanization-induced drying (Oke *et al* 2017, Chakraborty *et al* 2022), relative importance of humidity for human heat risk is still an active area of research (Sherwood 2018). Finally, heat stress also depends on wind speed, solar radiation, etc., which can be modified by urban morphology, shading structures, etc. (Fitria *et al* 2019, Kleerekoper *et al* 2012, Li 2021, Steeneveld *et al* 2011, Yuan *et al* 2020). These factors should be considered in future studies to better understand the heterogeneity of heat stress within urban environments.

4.3. Policy implications for mitigating and managing urban heat

The resulting air temperature, humidity, and heat stress maps and datasets can be used to inform urban heat mitigation planning and policy. Identification of urban heat hotspots, combined with sociodemographic data, can help identify potential vulnerable communities and areas for targeted intervention. Our findings that tree-covered areas, aside from water bodies, are the coolest, with built-up urban areas and grasslands (e.g. parks or unshaded greenspaces) 1.2 °C and 0.7 °C warmer, respectively, is in line with previous studies (Zhang *et al* 2014, Aram *et al* 2019, Ziter *et al* 2019, Aboelata *et al* 2020) that suggest albedo management and tree planting or green space development could help mitigate heat. Since the hottest areas of Chapel Hill are coincident with the highest population density, lowest greenness, and lowest income, policymakers could use the data and maps we produced here to develop strategies to increase green space, tree cover, and shade in particularly hot areas. Cities are starting to incorporate insights derived from high-resolution heat mapping into urban planning decisions. For example, in the case of Raleigh, NC, data collected from a July 2021 NOAA heat mapping campaign led to a city council vote to reallocate $70 000 into a pavement rejuvenation project to coat more than 150 000 yards of roadway with titanium dioxide to make it more reflective (Retana 2022, The City Council of the City of Raleigh 2002). This project specifically aims to increase the reflectivity of existing pavements to cool the air above it and surrounding areas, although researchers note an increased risk for night-time drivers since titanium-coated pavements can be more reflective (Cheela *et al* 2021). Additionally, high-albedo surfaces like coated sidewalks could also lead to an increased heat load during daytime for pedestrians due to increased mean radiant air temperature, which strongly affects daytime heat stress (Schneider *et al* 2023).

4.4. Limitations

Several limitations need to be kept in mind when contextualizing the results of the study. The first set of limitations relate to the sensors themselves. The PocketLab sensors have their own sources of error. What each sensor measures is representative of a fetch or source area, the extent of which depends on several factors, including atmospheric stability and measurement height (Oke 2007). The height:fetch ratio can vary between 1:10 for extremely unstable conditions i.e. a sensor at 1 m height gets information from a 10 m region upwind of the sensor, to over 1:500 for very stable conditions. Since the combination of the summer daytime period and urban roughness elements would lead to more unstable conditions (Bowne and Ball 1970), we expect the height:fetch ratio to be between 1:50–1:100. The citizen measurements would therefore likely represent a fetch of between 75 to 150 m during the study period, with most of the collected information coming from regions closest to the sensor. Even so, here we estimate gridded air temperatures at 10 m to also take advantage of the 10 m land cover dataset, which is a feature in our ML models. Overall, the attribution of measurements from a larger fetch to a smaller area would reduce expected variability. Essentially, when considering only this error, our gridded dataset provides more conservative estimates of near-surface heterogeneity in air temperature and RH at 10 m resolution.

Second, the PocketLab sensors report absolute accuracy (3% for RH and 0.5 °C for air temperature), but the actual measurements in a non-laboratory setting (so not where these accuracies were derived) are likely subject to higher error or deviation, as we found in our comparison with a local weather station, although the setting for that station (i.e. abandoned airport field) is quite different from the urban areas we had volunteers assess. As we show in tables S2, S3 and figure S2, the PocketLab sensors' measurements for RH are more variable compared to other research grade sensors' data, but this difference should have less impact on our estimates of Humidex, which is more dependent on air temperature. Third, the sensors could also be faulty or have radiative errors, although we lack data on these specific errors. We attempted, however, to address these issues by having multiple volunteers map the same route during each session. Fourth, the sensors we used were only able to measure air temperature and humidity, although there are other relevant heat metrics, such as mean radiant air temperature, which measures the total heat load on the body, considering radiation, shade and humidity in addition to ambient air temperature. All these potential sources of errors mean that, although variability in the urban microclimate is expected, we should be cautious about the absolute

quantitative estimates from these measurements. As such, focusing on which neighborhoods are the warmest (rather than how much warmer a neighborhood is) may be a more reasonable use of these and other similar sensors. However, this is a limitation of the data source (also see next paragraph on human errors) and not our data-driven methodological approach. One could, if funding permits, use multiple research-grade sensors to get higher quality observations to train similar ML algorithms.

The second set of limitations are related to engaging citizens as a primary mode of data collection, since volunteers pose another source of error. Although we provided the volunteers training before collecting measurements, there could be some individual error introduced because most of them were first-time users of the PocketLab sensors. For instance, when evaluating the quality of data collected, we found one user's recorded humidity data was extremely low while air temperature data appeared within a normal range. These measurements were not used in our analyses. This error may be due to a temporary issue with the sensor or an individual's accidental blockage of the humidity sensor port. One way to evaluate the sensitivity of our data to these possible errors, future studies could expand data collection involving more volunteers and a larger study area. Volunteers can make mistakes due to their misinterpretation of directions or the study parameters. Although the intention was to assign enough volunteers for each neighborhood to be mapped twice during each session, the Chapel Hill North neighborhood was not adequately staffed, resulting in different routes being mapped during the 2–3 pm and 5–6 pm sessions. Since our model relies on the assumption that the input variables (e.g. land cover, LST, etc.) are related to the citizen-mapped air temperature, this difference in mapping routes for the 2–3 and 5–6 pm models should not have affected our model accuracy, since we also trained the models individually. Additionally, we were only able to conduct the campaign one summer afternoon, whereas being able to collect data on multiple days would identify any potential idiosyncrasies with the data our volunteers collected on our campaign day.

Another source of potential error is in the use of satellite data, which measures LST and surface reflectance at the top of the canopy or foliage. For instance, satellite-derived LST over trees would only reflect the radiative temperature of the top of the tree canopy and not the air temperature (or even LST) of the ground surface below the canopy (Cheung *et al* 2021). Due to the high correlation ($R^2 = 0.86$ for the 2–3 pm session and $R^2 = 0.91$ for the 5–6 pm session) between the predicted and measured air temperature using LST as one of the inputs, we are reasonably confident in the overall variability predicted in our air temperature estimates. Additionally, as illustrated in figure S7, when examined directly within the assumptions of a linear model, satellite-derived LST is only weakly positively correlated ($r = 0.28$) with our citizen-collected air temperature data, although its inclusion improved the predictive power of our RF model, which accounts for many non-linearities. Although our volunteers collected data over the major urban land-covers in our study area (see table S1), there are some areas that are relevant for heat exposure, such as people's backyards and water, that we could not directly observe.

5. Conclusion

In this study, we developed a method for applying satellite remote sensing and citizen-collected air and humidity data to develop a high resolution (10 m) map of air temperature, humidity, and heat stress. We confirm previous studies that show air temperatures are hottest for impervious, built-up urban areas and coolest for forested, tree-covered areas. Compared to individual weather station measurements, we find individuals are exposed to greater variability of air temperatures and heat stress, and that this exposure is greatest during the daytime and cools off in the late afternoon, although the amount of this difference between daytime and nighttime differs. Ultimately, this method and approach can be replicated and scaled at a relatively low cost (or even for more expensive sensors if funding permits) and provide detailed information for decision-makers and urban planners seeking to mitigate urban heat and its human health effects.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://doi.org/10.15139/S3/ZI9ER6.

Acknowledgments

TC's contribution was supported by COMPASS-GLM, a multi-institutional project supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research as part of the Earth and Environmental Systems Modeling program. PNNL is operated for the Department of Energy by Battelle Memorial Institute under Contract DE-AC05-76RL01830. This work was also supported by grant to A H by the Samuel Family Foundation (Grant No. 5116545).

Ethical statement

Data collection by citizen volunteers was reviewed and exempted by UNC's Institutional Review Board Committee (Project ID: 21-1748).

ORCID iD

Angel Hsu https://orcid.org/0000-0003-4913-9479

References

Aboelata A and Sodoudi S 2020 Evaluating the effect of trees on UHI mitigation and reduction of energy usage in different built up areas in Cairo *Build. Environ.* **168** 106490
Bechtel B, Demuzere M, Mills G, Zhan W, Sismanidis P, Small C and Voogt J 2019 SUHI analysis using local climate zones—A comparison of 50 cities *Urban Clim.* **28** 100451
Benz S A and Burney J A 2021 Widespread race and class disparities in surface urban heat extremes across the United States *Earth's Future* **9** e2021EF002016
Bowne N E and Ball J T 1970 Observational comparison of rural and urban boundary layer turbulence *J. Appl. Meteorol.* **9** 862–73
Breiman L 2001 Random forests *Mach. Learn.* **45** 5–32
Burger M, Gubler M and Brönnimann S 2022 Modeling the intra-urban nocturnal summertime air temperature fields at a daily basis in a city with complex topography *PLoS Clim.* **1** e0000089
Centers for Disease Control and Prevention (CDC) 2022 QuickStats: percentage distribution of heat-related deaths, by age group—National Vital Statistics System, United States, 2018–2020 *MMWR Morb. Mortal. Wkly. Rep.* **71** 808
Chakraborty T C, Newman A, Qian Y, Hsu A and Sheriff G 2023 *Residential Segregation and Urban Heat Stress Disparities in the United States* (*OneEarth*) (https://doi.org/10.2139/ssrn.4231649)
Chakraborty T C, Sarangi C and Lee X 2021 Reduction in human activity can enhance the urban heat island: insights from the COVID-19 lockdown *Environ. Res. Lett.* **16** 054060
Chakraborty T, Hsu A, Manya D and Sheriff G 2019 Disproportionately higher exposure to urban heat in lower-income neighborhoods: a multi-city perspective *Environ. Res. Lett.* **14** 105003
Chakraborty T, Hsu A, Manya D and Sheriff G 2020 A spatially explicit surface urban heat island database for the United States: characterization, uncertainties, and possible applications *ISPRS J. Photogramm. Remote Sens.* **168** 74–88
Chakraborty T, Venter Z S, Qian Y and Lee X 2022 Lower urban humidity moderates outdoor heat stress *AGU Adv.* **3** e2022AV000729
Chapman L, Bell C and Bell S 2017 Can the crowdsourcing data paradigm take atmospheric science to a new level? A case study of the urban heat island of London quantified using Netatmo weather stations *Int. J. Climatol.* **37** 3597–605
Cheela V S, John M, Biswas W and Sarker P 2021 Combating urban heat island effect—A review of reflective pavements and tree shading strategies *Buildings* **11** 93
Chen G, Hua J, Shi Y and Ren C 2023 Constructing air temperature and relative humidity-based hourly thermal comfort dataset for a high-density city using machine learning *Urban Clim.* **47** 101400
Chen T, De Jeu R A M, Liu Y Y, Van der Werf G R and Dolman A J 2014 Using satellite based soil moisture to quantify the water driven variability in NDVI: a case study over mainland Australia *Remote Sens. Environ.* **140** 330–8
Chen T and Guestrin C 2016 XGBoost: a scalable tree boosting system *Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining* pp 785–94
Chen X-L, Zhao H-M, Li P-X and Yin Z-Y 2006 Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes *Remote Sens. Environ.* **104** 133–46
Cheung P K, Jim C Y and Hung P L 2021 Preliminary study on the temperature relationship at remotely-sensed tree canopy and below-canopy air and ground surface *Build. Environ.* **204** 108169
Chevalier R F, Hoogenboom G, McClendon R W and Paz J A 2011 Support vector regression with reduced training sets for air temperature prediction: a comparison with artificial neural networks *Neural Comput. Appl.* **20** 151–9
Earth Resources Observation And Science (EROS) Center 2017 *National Agriculture Imagery Program (NAIP)* (U.S. Geological Survey) (https://doi.org/10.5066/F7QN651G)
Eliasson I and Svensson M K 2003 Spatial air temperature variations and urban land use-a statistical approach *Meteorol. Appl.* **10** 135–49
Erell E and Williamson T 2007 Intra-urban differences in canopy layer air temperature at a mid-latitude city *Int. J. Climatol.* A **27** 1243–55
Fenner D, Meier F, Bechtel B, Otto M and Scherer D 2017 Intra and inter local climate zone variability of air temperature as observed by crowdsourced citizen weather stations in Berlin, Germany *Meteorol. Z.* **26** 525–47
Fitria R, Kim D, Baik J and Choi M 2019 Impact of biophysical mechanisms on urban heat island associated with climate variation and urban morphology *Sci. Rep.* **9** 19503
Hanoon M S, Ahmed A N, Zaini N, Razzaq A, Kumar P, Sherif M, Sefelnasr A and El-Shafie A 2021 Developing machine learning algorithms for meteorological temperature and humidity forecasting at Terengganu state in Malaysia *Sci. Rep.* **11** 18935
Harlan S L, Brazel A J, Prashad L, Stefanov W L and Larsen L 2006 Neighborhood microclimates and vulnerability to heat stress *Soc. Sci. Med.* **63** 2847–63
Hart M A and Sailor D J 2009 Quantifying the influence of land-use and surface characteristics on spatial variability in the urban heat island *Theor. Appl. Climatol.* **95** 397–406
Heal G and Park J 2016 Reflections—temperature stress and the direct impact of climate change: a review of an emerging literature *Rev. Environ. Econ. Policy* **10** 347–62
Ho H C, Knudby A, Sirovyak P, Xu Y, Hodul M and Henderson S B 2014 Mapping maximum urban air temperature on hot summer days *Remote Sens. Environ.* **154** 38–45
Hoffman J S, Shandas V and Pendleton N 2020 The effects of historical housing policies on resident exposure to intra-urban heat: a study of 108 US urban areas *Climate* **8** 12
Hsu A, Sheriff G, Chakraborty T and Manya D 2021 Disproportionate exposure to urban heat island intensity across major US cities *Nat. Commun.* **12** 1–11

Huang K, Li X, Liu X and Seto K C 2019 Projecting global urban land expansion and heat island intensification through 2050 *Environ. Res. Lett.* **14** 114037
Hutengs C and Vohland M 2016 Downscaling land surface temperatures at regional scales with random forest regression *Remote Sens. Environ.* **178** 127–41
IPCC 2021 Summary for policymakers *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge University Press) pp 3–32
Iungman T *et al* 2023 Cooling cities through urban green infrastructure: a health impact assessment of European cities *Lancet* **401** 577–89
Kleerekoper L, Van Esch M and Salcedo T B 2012 How to make a city climate-proof, addressing the urban heat island effect *Resour. Conserv. Recycl.* **64** 30–38
Kousis I, Pigliautile I and Pisello A L 2021 Intra-urban microclimate investigation in urban heat island through a novel mobile monitoring system *Sci. Rep.* **11** 9732
Krayenhoff E S, Broadbent A M, Zhao L, Georgescu M, Middel A, Voogt J A, Martilli A, Sailor D J and Erell E 2021 Cooling hot cities: a systematic and critical review of the numerical modelling literature *Environ. Res. Lett.* **16** 053007
Krayenhoff E S, Moustaoui M, Broadbent A M, Gupta V and Georgescu M 2018 Diurnal interaction between urban expansion, climate change and adaptation in US cities *Nat. Clim. Change* **8** 1097–103
Krizek K J 1995 Patterns of use in Main Street activity: a case study of downtown Chapel Hill, NC *Carol. Plann.* **20** 62–70
Kuhn M 2008 Building predictive models in R using the caret package *J. Stat. Softw.* **28** 1–26
Lai J *et al* 2021 Meteorological controls on daily variations of nighttime surface urban heat islands *Remote Sens. Environ.* **253** 112198
Li X 2021 Investigating the spatial distribution of resident's outdoor heat exposure across neighborhoods of Philadelphia, Pennsylvania using urban microclimate modeling *Sustain. Cities Soc.* **72** 103066
Li Y and Gallagher K P 2019 The environmental impact of China-financed coal-fired power plants in South East Asia *Boston University, Global Development Policy Center, Global China Initiative Working Paper* 7
Mahmood R, Keeling T, Foster S A and Hubbard K G 2013 Did irrigation impact 20th century air temperature in the High Plains aquifer region? *Appl. Geogr.* **38** 11–21
Manoli G, Fatichi S, Bou-Zeid E and Katul G G 2020 Seasonal hysteresis of surface urban heat islands *Proc. Natl Acad. Sci.* **117** 7082–9
Manoli G, Fatichi S, Schläpfer M, Yu K, Crowther T W, Meili N, Burlando P, Katul G G and Bou-Zeid E 2019 Magnitude of urban heat islands largely explained by climate and population *Nature* **573** 55–60
Masterton J M and Richardson F A 1979 *Humidex: A Method of Quantifying Human Discomfort Due to Excessive Heat and Humidity* (Environment Canada, Atmospheric Environment)
McDonald R I, Biswas T, Sachar C, Housman I, Boucher T M, Balk D, Nowak D, Spotswood E, Stanley C K and Leyk S 2021 The tree cover and temperature disparity in US urbanized areas: quantifying the association with income across 5,723 communities *PLoS One* **16** e0249715
Mentaschi L, Duveiller G, Zulian G, Corbane C, Pesaresi M, Maes J, Stocchino A and Feyen L 2022 Global long-term mapping of surface temperature shows intensified intra-city urban heat island extremes *Glob. Environ. Change* **72** 102441
Mildrexler D J, Zhao M and Running S W 2011 A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests *J. Geophys. Res.* **116** G03025
Muller C L, Chapman L, Grimmond C S B, Young D T and Cai X 2013 Sensors and the city: a review of urban meteorological networks *Int. J. Climatol.* **33** 1585–600
NOAA n.d. Mapping campaigns *National Integrated Heat Health Information System* (available at: www.heat.gov/pages/mapping-campaigns) (Accessed 16 August 2023)
North Carolina State Climate Office, NC State University *Cardinal* (available at: https://products.climate.ncsu.edu/cardinal/request) (Accessed January 2023)
Oke T R 1982 The energetic basis of the urban heat island *Q. J. R. Meteorol. Soc.* **108** 1–24
Oke T R 2007 Siting and exposure of meteorological instruments at urban sites *Air Pollution Modeling and its Application XVII* **615** 31
Oke T R, Mills G, Christen A and Voogt J A 2017 *Urban Climates* (Cambridge University Press)
Paniagua-Tineo A, Salcedo-Sanz S, Casanova-Mateo C, Ortiz-García E G, Cony M A and Hernández-Martín E 2011 Prediction of daily maximum temperature using a support vector regression algorithm *Renew. Energy* **36** 3054–60
Park R J, Goodman J, Hurwitz M and Smith J 2020 Heat and learning *Am. Econ. J.: Econ. Policy* **12** 306–39
PocketLab *PocketLab Weather Specifications* (available at: https://support.thepocketlab.com/knowledge/pocketlab-weather-specifications) (Accessed 9 March 2023)
Qian Y, Chakraborty T C, Li J, Li D, He C, Sarangi C, Chen F, Yang X and Leung L R 2022 Urbanization impact on regional climate and extreme weather: current understanding, uncertainties, and future research directions *Adv. Atmos. Sci.* **39** 819–60
Rajagopalan P, Andamon M M and Paolini R 2020 Investigating thermal comfort and energy impact through microclimate monitoring-a citizen science approach *Energy Build.* **229** 110526
Ram F, García E H, Solgi E and Mansournia S 2019 Urban green space cooling effect in cities *Heliyon* **5**
Retana J 2022 Are Raleigh's roads making the city hotter? *CBS17.Com* (available at: www.cbs17.com/news/local-news/wake-county-news/are-raleighs-roads-making-the-city-hotter/)
Ritchie H and Roser M 2018 Urbanization *Our World in Data* (available at: https://ourworldindata.org/urbanization)
Rouse J W Jr, Haas R H, Schell J A and Deering D W 1974 Monitoring vegetation systems in the Great Plains with ERTS *Goddard Space Flight Center 3d ERTS-1 Symp.* (NASA) p 1
Rubel F and Kottek M 2010 Observed and projected climate shifts 1901-2100 depicted by world maps of the Köppen-Geiger climate classification *Meteorol. Z.* **19** 135
Schneider F A, Ortiz J C, Vanos J K, Sailor D J and Middel A 2023 Evidence-based guidance on reflective pavement for urban heat mitigation in Arizona *Nat. Commun.* **14** 1467
Schwarz N, Schlink U, Franck U and Großmann K 2012 Relationship of land surface and air temperatures and its implications for quantifying urban heat island indicators—an application for the city of Leipzig (Germany) *Ecol. Indic.* **18** 693–704
Shandas V, Voelkel J, Williams J and Hoffman J 2019 Integrating satellite and ground measurements for predicting locations of extreme urban heat *Climate* **7** 5
Sherwood S C 2018 How important is humidity in heat stress? *J. Geophys. Res.* **123** 11–808
Shi H, Xian G, Auch R, Gallo K and Zhou Q 2021 Urban heat island and its regional impacts using remotely sensed thermal data—a review of recent developments and methodology *Land* **10** 867

Skarbit N, Stewart I D, Unger J and Gál T 2017 Employing an urban meteorological network to monitor air temperature conditions in the 'local climate zones' of Szeged, Hungary *Int. J. Climatol.* **37** 582–96
Steeneveld G-J, Koopmans S, Heusinkveld B G, Van Hove L W A and Holtslag A A M 2011 Quantifying urban heat island effects and human comfort for cities of variable size and urban morphology in the Netherlands *J. Geophys. Res.* **116** D20129
Stewart I D, Krayenhoff E S, Voogt J A, Lachapelle J A, Allen M A and Broadbent A M 2021 Time evolution of the surface urban heat Island *Earth's Future* **9** e2021EF002178
Tan J *et al* 2010 The urban heat island and its impact on heat waves and human health in Shanghai *Int. J. Biometeorol.* **54** 75–84
The City Council of the City of Raleigh (2002) Regular Meeting—Third Tuesday—Work Session & Afternoon Session (available at: http://go.boarddocs.com/nc/raleigh/Board.nsf/goto?open&id=CELPSM660913)
Town of Chapel Hill 2020 Future land use map—update to chapel hill 2020 *FlippingBook* (available at: https://online.flippingbook.com/view/26191/72-73/?sharedOn=)
Turner V K, Rogers M L, Zhang Y, Middel A, Schneider F A, Ocón J P, Seeley M and Dialesandro J 2022 More than surface temperature: mitigating thermal exposure in hyper-local land system *J. Land Use Sci.* **17** 79–99
US Census Bureau 2020 *American Community Survey 5-Year Data (2009-2020)* (Census.Gov) (available at: www.census.gov/data/developers/data-sets/acs-5year.html)
US Census Bureau 2021 *U.S. Census Bureau QuickFacts: Chapel Hill town, North Carolina* (available at: www.census.gov/quickfacts/fact/table/chapelhilltownnorthcarolina/PST045221)
USGS 2021 Landsat 8 *U.S. Geological Survey* (available at: www.usgs.gov/landsat-missions/landsat-8)
van Hove L W A, Jacobs C M J, Heusinkveld B G, Elbers J A, van Driel B L and Holtslag A A M 2015 Temporal and spatial variability of urban heat island and thermal comfort within the Rotterdam agglomeration *Build. Environ.* **83** 91–103
Vargas Zeppetello L R, Raftery A E and Battisti D S 2022 Probabilistic projections of increased heat stress driven by climate change *Commun. Earth Environ.* **3** 1–7
Venter Z S, Brousse O, Esau I and Meier F 2020 Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data *Remote Sens. Environ.* **242** 111791
Venter Z S, Chakraborty T and Lee X 2021 Crowdsourced air temperatures contrast satellite measures of the urban heat island and its mechanisms *Sci. Adv.* **7** eabb9569
Voelkel J, Hellman D, Sakuma R and Shandas V 2018 Assessing vulnerability to urban heat: a study of disproportionate heat exposure and access to refuge by socio-demographic status in Portland, Oregon *Int. J. Environ. Res. Public Health* **15** 640
Voogt J A and Oke T R 2003 Thermal remote sensing of urban climates *Remote Sens. Environ.* **86** 370–84
Vulova S, Meier F, Fenner D, Nouri H and Kleinschmit B 2020 Summer nights in Berlin, Germany: modeling air temperature spatially with remote sensing, crowdsourced weather data, and machine learning *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **13** 5074–87
Wang J, Qian Y, Pringle W, Chakraborty T C, Hetland R, Yang Z and Xue P 2023 Contrasting effects of lake breeze and urbanization on heat stress in Chicago metropolitan area *Urban Clim.* **48** 101429
Wong M S, Peng F, Zou B, Shi W Z and Wilson G J 2016 Spatially analyzing the inequity of the Hong Kong urban heat island by socio-demographic characteristics *Int. J. Environ. Res. Public Health* **13** 317
Xu Q-S and Liang Y-Z 2001 Monte Carlo cross validation *Chemom. Intell. Lab. Syst.* **56** 1–11
Yang L and Shami A 2020 On hyperparameter optimization of machine learning algorithms: theory and practice *Neurocomputing* **415** 295–316
Yang Y, Cao C, Pan X, Li X and Zhu X 2017 Downscaling land surface temperature in an arid area by using multiple remote sensing indices with random forest regression *Remote Sens.* **9** 789
Yuan C, Adelia A S, Mei S, He W, Li -X-X and Norford L 2020 Mitigating intensity of urban heat island by better understanding on urban morphology and anthropogenic heat dispersion *Build. Environ.* **176** 106876
Zanaga D *et al* 2021 ESA WorldCover 10 m 2020 v100 Data set Zenodo (https://doi.org/10.5281/zenodo.5571936)
Zanaga D, Van De Kerchove R, Daems D, De Keersmaecker W, Brockmann C, Kirches G, Wevers J, Cartus O, Santoro M and Fritz S 2022 *ESA WorldCover 10 m 2021 v200* (https://doi.org/10.5281/zenodo.5571935)
Zander K K, Botzen W J, Oppermann E, Kjellstrom T and Garnett S T 2015 Heat stress causes substantial labour productivity loss in Australia *Nat. Clim. Change* **5** 647–51
Zha Y, Gao J and Ni S 2003 Use of normalized difference built-up index in automatically mapping urban areas from TM imagery *Int. J. Remote Sens.* **24** 583–94
Zhang P, Bounoua L, Imhoff M L, Wolfe R E and Thome K 2014 Comparison of MODIS land surface temperature and air temperature over the Continental USA Meteorological Stations *Can. J. Remote Sens.* **40** 110–22
Ziter C D, Pedersen E J, Kucharik C J and Turner M G 2019 Scale-dependent interactions between tree canopy cover and impervious surfaces reduce daytime urban heat during summer *Proc. Natl Acad. Sci.* **116** 7575–80
Zumwald M, Knüsel B, Bresch D N and Knutti R 2021 Mapping urban temperature using crowd-sensing data and machine learning *Urban Clim.* **35** 100739