

# Shifting Islands: How weather conditions and urban form shape the spatiotemporal character of Baltimore's urban heat island

Bianca Corpuz<sup>a,\*</sup>, Benjamin Zaitchik<sup>a</sup>, Darryn Waugh<sup>a</sup>, Anna Scott<sup>a</sup>, Tom Logan<sup>b</sup>

<sup>a</sup> Earth and Planetary Sciences, Johns Hopkins University, United States

<sup>b</sup> Civil and Natural Resources Engineering, University of Canterbury, New Zealand

## ARTICLE INFO

### Keywords:

Urban heat island  
Intra-urban air temperatures  
Generalized additive models (GAMs)

## ABSTRACT

The combined effects of climate change and urban expansion are elevating urban temperatures, amplifying health risks, and increasing summertime energy demand. Effective heat reduction and adaptation strategies rely on a thorough understanding of urban heat island (UHI) dynamics, which vary within cities and over time due to local urban form and meteorological conditions. While existing studies have examined the spatiotemporal structure of UHIs, they have not examined the space-time variability of air temperature at high spatial and temporal resolutions. This study addresses this gap by utilizing a ground-based sensor network of 53 iButton thermochrons, synoptic weather stations, and satellites to capture these spatiotemporal patterns. Using these data, we developed generalized additive models to predict daily, daytime and nighttime air temperature anomalies throughout Baltimore City. We found that 1) meteorological variables are more influential in characterizing the nighttime UHI than the daytime UHI; 2) spatiotemporal variability of nighttime UHIs is greatest in areas with sparse vegetation canopy cover; and 3) interactions between urban form and meteorology have a systematic but modest impact on the spatial structure of UHIs. The methodology and findings presented in this study offer tools for developing outdoor heat mitigation and adaptation strategies to reduce the UHI effect.

## 1. Introduction

With temperature increasing throughout the world, the frequency and intensity of heat events are expected to worsen (IPCC, 2022). Simultaneously, urbanization is rapidly advancing. In 2014, 54% of the global population were reported to be city dwellers with this percentage projected to increase to 66% by 2050 (United Nations DESA, 2019). Urbanization, coupled with the effects of climate change, lead to higher temperatures in urban environments which are intensified by urban heat islands (UHIs) (Tuholske et al., 2021). UHIs refer to the phenomenon where urbanized areas experience higher surface and air temperatures compared to surrounding regions (Oke, 1982; Sheridan and Allen, 2015). UHIs can pose significant health risks as they exacerbate adverse health effects such as asthma, chronic obstructive pulmonary disease, pregnancy risks, cardiovascular events, and mortality can be exacerbated by UHIs (Duan et al., 2020; Rodrigues et al., 2019; Zhang et al., 2017). Considering the impacts of UHIs on health outcomes and extreme heat exposure, it is critical to understand fine-scale spatiotemporal trends of air temperature within a city. This understanding can help facilitate the development of targeted, location-specific strategies for UHI mitigation and adaptation.

Appreciation of the magnitude and potential implications of the UHI has motivated efforts to mitigate elevated urban temperatures.

\* Corresponding author.

E-mail address: [bcorpuz1@jhu.edu](mailto:bcorpuz1@jhu.edu) (B. Corpuz).

When designing these strategies, it is important to recognize that the UHI is not uniform within a city, and the concept of the “urban heat archipelago” – i.e., fine spatial variability in the UHI—is generally recognized (Shepherd et al., 2014). That is, urban heat conditions are substantially stronger for some neighborhoods and land cover types than they are for others. This reality is clearly evident in the city of Baltimore, Maryland, which is the focus city for this study. In Baltimore, the UHI effect is a prevalent issue with studies reporting higher surface temperatures in some neighborhoods compared to others (e.g., Hoffman et al., 2020; Scott et al., 2017). Studies which have observed the distribution of urban heat in Baltimore have captured patterns of “the black butterfly,” a shape where the wings represent economically excluded, majority Black neighborhoods of Baltimore (Brown, 2021; G. Huang et al., 2011; Scott et al., 2017). While these studies examined the spatiotemporal structure of UHIs, they did not capture temperature variability for the entire city at high spatial and temporal resolutions. Our study further examines the space-time variability of air temperature in Baltimore and the impacts that urban form and meteorological conditions may have on the UHI. This study presents innovative approaches by investigating how the spatial and temporal variability of the UHI affect neighborhood-scale temperatures, which in turn can have implications for individual heat exposure. Additionally, it utilizes a nonlinear model capable of incorporating both time-varying and static variables. Furthermore, this study conducts a fine-scale variability analysis of the UHI focusing on air temperature rather than satellite derived temperature variables.

When characterizing the UHI one must be clear about the “heat” variable being studied. Here we focus on near-surface air temperature, as recorded by a distributed in situ network. This is in contrast to the many neighborhood-scale UHI studies that have used Land Surface Temperature (LST) measurements derived from satellite imagery when studying spatial variability patterns (e.g., B. Huang et al., 2013; Logan et al., 2020). These satellite observations are powerful, but they present space-time resolution challenges. Relying on LST estimates with high spatial resolution, like those derived from Landsat at 100 m or ECOSTRESS at 70 m, can compromise a study’s ability to resolve temporal variability due to low revisit frequency. Meanwhile, LST measurements with high temporal resolution, like those derived from VIIRS or MODIS or from microwave imagers, have limited spatial resolution (750 m to 1 km, and larger for microwave), which do not fully characterize surface temperature at finer, neighborhood scales (Anderson et al., 2021; Hulley et al., 2019). Perhaps more importantly, while LST is an important variable in its own right, it is not the same as air temperature in its physical meaning or its relevance to human comfort (Voogt and Oke, 2003). And though LST does correlate with air temperature at larger spatial scales, the correlation between LST and air temperature weakens at intra-urban scales (Good, 2016), so one cannot use spatiotemporal variability in LST as a proxy for variability in air temperature.

Obtaining reliable in situ air temperature measurements at high spatial resolution in urban environments, however, can be challenging. In recent years, a number of campaigns have taken on this challenge. This includes work with low-cost stationary temperature monitors distributed across the city (e.g., Gubler et al., 2021; Saverino et al., 2021; Scott et al., 2017) and mobile transit data from vehicle-mounted sensors (Shandas et al., 2019). There are merits to both approaches, as vehicle-mounted sensors can yield wide spatial coverage at limited cost, while stationary sensors offer constant monitoring at selected locations and are not limited to roadway environments (Shi et al., 2021; WMO, 2023). For the present study, we are concerned with space-time variability in the UHI at neighborhood scale, and for this purpose we find that a stationary network has advantages: it offers time-series data through the diurnal cycle and over the course of different weather conditions.

Diurnally, it is understood that the physical drivers of the city-scale UHI and its observed magnitude differ systematically between day and night. During the day, solar radiation is absorbed by urban materials such as buildings, asphalt, and concrete, and at night, the urban fabric releases the stored heat that was absorbed throughout the day (Kleereker et al., 2012; Oke, 1982). As a result, the intensity of city-scale UHIs are generally found to be larger and more pronounced at night compared to daytime conditions (Oke, 1982). However, when it comes to quantifying spatial patterns of UHIs using air temperature measurements, daytime UHIs have been found to be more complicated to characterize and exhibit more structure compared to nighttime UHIs, with hotspots within the urban matrix. Daytime UHIs have been found to be most influenced by green spaces, windspeed, and cloud cover, while the structure of nighttime UHIs is significantly affected by advection (Azevedo et al., 2016). Detailed study of these diurnal differences in the structure of the urban heat archipelago can advance our understanding of scale-dependency of UHI dynamics and inform the design of UHI mitigation strategies.

On the timescale of days to weeks, research has found that the mesoscale UHI is sensitive to weather conditions, as solar heating, wind speed, humidity, and other factors have a strong influence on the surface energy balance and atmospheric boundary layer processes relevant to the development of a UHI (Kong et al., 2021). For example, some researchers have found the UHI to be stronger under heat wave conditions, while others have found the opposite to be true (Li et al., 2019; Scott et al., 2018), and still others have emphasized the sensitivity of these results to how heat waves are defined with respect to the UHI (Fenner et al., 2019). However, it is unknown whether neighborhood scale variability within the UHI changes as a function of background weather conditions. Insomuch as fine scale structure in the UHI is relevant to health, energy use, and other socially relevant impacts of heat, it is important to account for space-time variability in the archipelago when mapping heat hazards. We hypothesize that such fine-scale influences of weather on UHI would appear in a statistical analysis in the form of interaction terms between static, physical properties and time-varying, synoptic meteorology variables.

In this study, datasets derived from satellite imagery, weather stations, and a ground-based sensor network of iButton thermochrons were utilized to study the influence of physical properties and meteorological conditions on the space-time variability of the UHI in Baltimore City. We included static, time-varying, and interaction terms between static and time-varying variables in our models to predict the spatial, temporal, and spatiotemporal variabilities in daytime and nighttime features of the urban heat archipelago. This work is concerned with predicting maximum and minimum air temperature anomalies at high spatial and temporal resolution. The objectives of this study include (1) estimating daily maximum and minimum air temperature anomalies throughout Baltimore, (2) quantifying spatial and temporal variance in local urban temperatures, (3) determining which static and time-varying variables are

useful in predicting daytime and nighttime UHIs, and (4) using interaction terms between static and time-varying variables to understand how meteorology influences the spatial structure of the urban heat archipelago.

## 2. Materials and methods

### 2.1. Study area

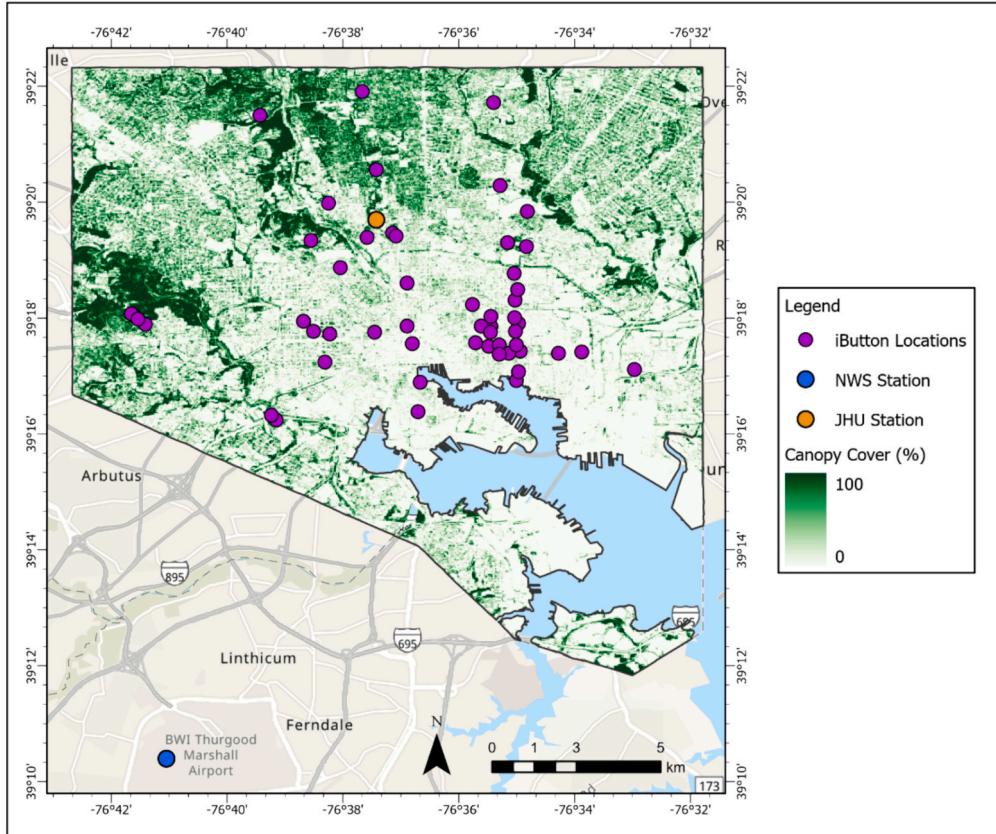
This study focuses on Baltimore, Maryland, a mid-sized city in the eastern United States bounded by 39.37° and 39.20° latitude and – 76.71° and – 76.53° longitude (Fig. 1). Baltimore is the most populous city in the state of Maryland and is located where the Patapsco River empties into the Chesapeake Bay and experiences a humid subtropical climate. The monthly average normals for our study period of July to September 2016 are 24.6 °C, 24.5 °C, and 20.7 °C. Throughout our study period, a total of 12 heat advisory days where temperatures exceeded 105 °F were issued by the NWS. According to the Maryland Department of Health and Mental Hygiene, there were >900 heat-related illness emergency department visits concerning hyperthermia, dehydration, heat exhaustion, heat stroke, sunburn, and 5 heat-related deaths reported during our study period ([Maryland DHMH, 2016](#)).

### 2.2. Datasets and data pre-processing

The data used in this study were obtained from a ground-based network of Maxim Integrated Products, Inc. “iButton” Model DS1923 Hygrochron thermochrons, a Davis Instruments Vantage Pro2 weather station (JHU Station), a National Weather Service (NWS) station, and satellite remote sensing products (Table 1).

#### 2.2.1. iButton air temperature data

We used air temperature data collected from a network of 53 fixed iButton sensors placed throughout Baltimore from July 18, 2016 to September 15, 2016 (Fig. 1). We obtained daily minimum and maximum temperatures for each iButton. Each iButton sensor was installed into a radiation shield, designed by researchers at the Maryland Institute College of Art, and attached to trees 2-m above the ground with full or partial shade while facing north (Fig. 2). All iButtons installed in this manner passed quality control checks for



**Fig. 1.** Distribution of observational network (53 iButton thermochrons) stationed in Baltimore from July to September 2016. The National Weather Service station utilized in this study is located within the Baltimore-Washington International airport grounds.

**Table 1**

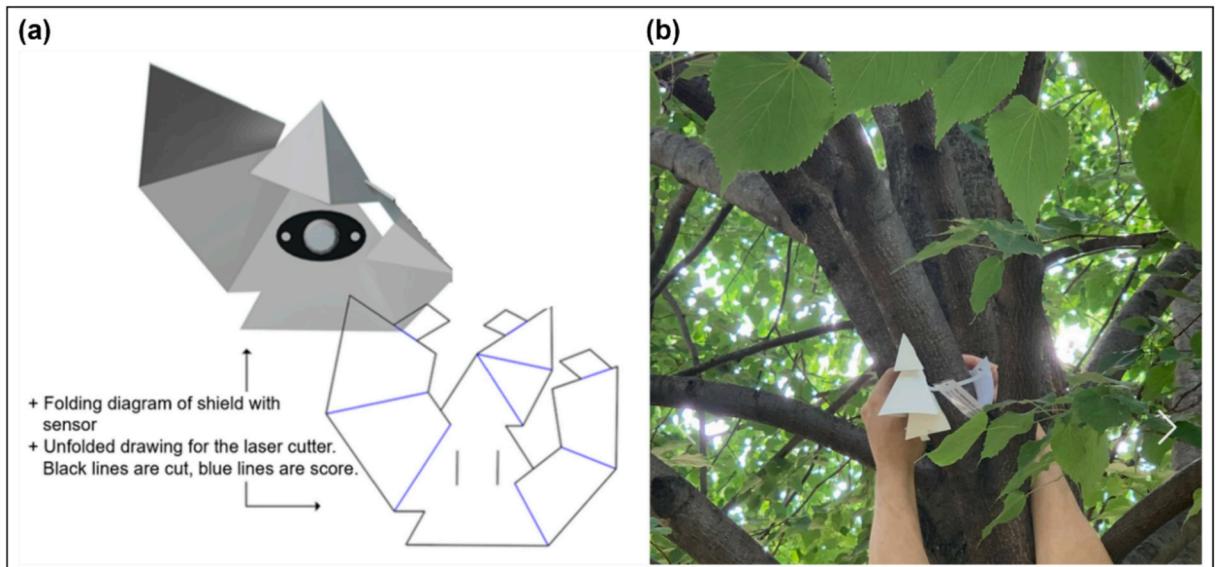
Datasets used in this study.

| Variable                                      | Dataset                             | Spatial Resolution | Temporal Resolution in this analysis |
|---|-------------------------------------|--------------------|--------------------------------------|
| Land Surface Temperature (LST)                | Landsat-8                           | 100 m              | Static (16-day composite)            |
| Albedo  | Landsat-8                           | 30 m               | Static (16-day composite)            |
| Normalized Difference Vegetation Index (NDVI) | USDA NAIP                           | 1 m                | Static (single snapshot)             |
| Sky View Factor (SVF)                         | <a href="#">Logan et al. (2020)</a> | 2 m                | Static                               |
| Land Use                                      | Chesapeake Conservatory             | 1 m                | Static (annual product)              |
| Park  | City of Baltimore                   | 1 m                | Static                               |
| Elevation                                     | Maryland Lidar Server               | 1 m                | Static (single snapshot)             |
| Canopy Cover (CC)                             | Maryland Lidar Server               | 30 m               | Static (single snapshot)             |
| Minimum Air Temperature (TMIN)                | NWS Station                         | Point measurement  | Daily                                |
| Maximum Air Temperature (TMAX)                | NWS Station                         | Point measurement  | Daily                                |
| Average Relative Humidity (RH)                | NWS Station                         | Point measurement  | Daily                                |
| Weather Type (WT)                             | NWS Station                         | Point measurement  | Daily                                |
| Average Daily Wind Speed (ADWS)               | NWS Station                         | Point measurement  | Daily                                |
| Average Solar Radiation (SR)                  | JHU Station                         | Point measurement  | Hourly                               |
| Air Temperature                               | iButton Network                     | Point measurement  | Hourly                               |
| Sensor Number (sensor)                        | iButton Network                     | Point measurement  | Static                               |

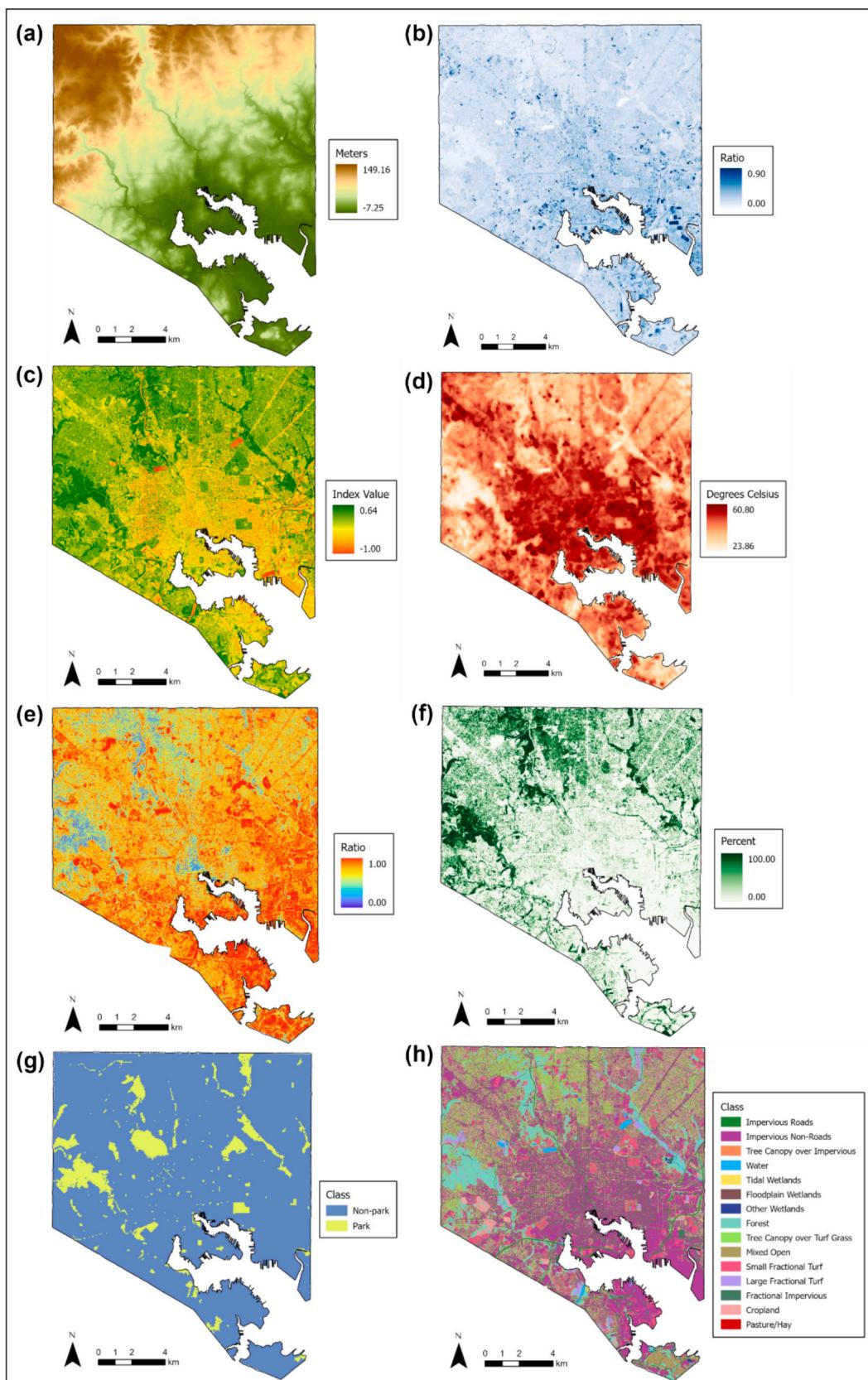
consistent representation of the diurnal cycle and an absence of breakpoints or unique peaks or troughs in the data. These 53 north-facing, tree-affixed iButtons represent a subset of an initial deployment of 79 iButtons, and they passed quality control checks of recording a smooth diurnal cycle in air temperature and lacking breakpoints that could result from tampering or other damage. Notably, all data from iButtons affixed to posts rather than trees was discarded, as it was found that those iButtons were prone to rapid daytime temperature rises that were inconsistent with data from weather stations or nearby iButtons. Any underestimate of air temperature measurements due to transpiration is mitigated by the consistent environmental conditions shared by iButtons when affixed beneath the canopy. That said, while the iButton sensors used in this study were placed in trees which provided consistent shading throughout our study period and protected against artifacts found to exist when measurements are taken in open sun or too close in proximity to impervious surfaces, there exists variability between tree types and that could affect the collected measurements ([Scott et al., 2017](#)). The iButton sensors have a reported accuracy of 0.5 °C and were programmed to record air temperature at an hourly temporal resolution ([Scott et al., 2017](#)). A thorough description of the iButton sensor network installation and methods is presented by [Scott et al., 2017](#). From the hourly measurements, we extracted daily minimum (nighttime) and maximum (daytime) air temperatures for our study period.

### 2.2.2. Meteorological data

In addition to our iButton sensor network, daily meteorological data included in this study were sourced from the National Weather Service (NWS) station at Baltimore-Washington International (BWI) Airport. This weather station is situated in an open field of grass within the BWI airport grounds. The following variables were included in our analysis: relative humidity (%), maximum air



**Fig. 2.** Images of radiation shield and sensor: (a) graphic illustrating how radiation shields, which store iButton thermochrons, are cut and folded and (b) an example of a constructed radiation shield and sensor deployed in a tree with adequate shade.



(caption on next page)

**Fig. 3.** Physical characteristics of Baltimore urban area depicted by satellite-derived variables: (a) elevation; (b) albedo; (c) normalized difference vegetation index; (d) land surface temperature; (e) sky view factor; (f) canopy cover; (g) park; and (h) land use.

temperature (°C), minimum air temperature (°C), weather type, and average wind speed (m/s). The scheme for weather type is that defined by the Global Historical Climatology Network Daily Database. The daily, maximum and minimum air temperature measurements obtained from the NWS station represent the highest and lowest hourly temperature, respectively, which is not exactly analogous to the method of maximum and minimum calculation for iButtons which is based on maximum or minimum one-hour average temperature. Hourly solar radiation (kJ/m<sup>2</sup>) data were obtained from a Davis Instruments Vantage Pro2 weather station installed on the roof of Johns Hopkins University's Olin Hall, which we refer to as the "JHU Station", and were processed to daily averages.

### 2.2.3. Satellite remote sensing data

Landsat-8 Operational Land Imager/Thermal Infrared Sensor (OLI/TIRS) Collection 2 Level 2 data were accessed to obtain Land Surface Temperature (LST; 100 m resolution) and albedo data (30 m resolution) for our iButton location sites (U.S. Geological Survey, 2016) (Fig. 3). The Landsat-8 image collected on 18 July 2016 was used for data extraction as this file had the least cloud cover during our study period. The weather conditions preceding this date were relatively humid but had no significant rainfall, which is indicative of the radiative heating we would expect for rain-free summer conditions in this coastal setting in the mid-Atlantic region of the United States. A scale factor of 0.00341802 and offset of 149.0 per pixel were applied to band-10 digital-number data to transform LST data into a usable format (U.S. Geological Survey, 2023). The resulting LST product in Kelvin was converted to degrees Celsius.

To calculate for albedo, Landsat-8 OLI data were first converted to top of atmosphere (TOA) reflectance according to Eq. (1) where  $M_p$  represents band-specific multiplicative re-scaling factor of 0.0000275,  $A_p$  represents band-specific additive re-scaling factor of -0.2,  $Q_{cal}$  represents the digital number,  $\theta_{SE}$  represents the local sun elevation angle of 57.64038290:

$$\rho_\lambda = \frac{M_p * Q_{cal} + A_p}{\sin(\theta_{SE})} \quad (1)$$

Once the TOA reflectances with a correction for the sun angle were calculated, albedo was estimated using the formula presented in Liang, 2001 and Smith, 2010 based on reflectances ( $\rho$ ) in Landsat bands 1, 3, 4, 5, and 7:

$$\alpha_{shortwave} = \frac{0.356 \rho_1 + 0.356 \rho_3 + 0.356 \rho_4 + 0.356 \rho_5 + 0.356 \rho_7 - 0.0018}{0.356 + 0.130 + 0.373 + 0.085 + 0.072} \quad (2)$$

Normalized Difference Vegetation Index (NDVI) was extracted from USDA's National Agriculture Imagery Program (NAIP) imagery, accessed through the NOAA Office of Coastal Management's Data Access Viewer (NOAA, 2015). NAIP imagery includes 4 bands (red, green, blue, and near-infrared) and has a spatial resolution of 1 m. NDVI was extracted according to Eq. (3) where  $\rho$  represents reflectance in the red and near-infrared bands:

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (3)$$

Maryland State's Lidar Topography Server was used to obtain elevation and tree canopy cover data for Baltimore City at 1-m resolution. Terrain elevation (DEM) and tree canopy cover data were retrieved from Maryland's Mapping and GIS Data Portal (Maryland iMAP Portal, 2013, 2015). Land use data with a spatial resolution of 1 m was obtained from the Chesapeake Bay Program Land Use/Land Cover Data Project. These data categorized land use into 15 separate classes. Lastly, Sky View Factor (SVF) data was obtained with 2-m resolution. The SVF extraction methodology can be found in (Logan et al., 2020). A thematic map of park areas was obtained from the City of Baltimore Department of Planning and was also sampled to 1-m gridded resolution.

We examined 30-m and 100-m radius buffers and 1-m grid cells for each iButton sensor location to characterize the site environment. We found the explanatory variables to have similar measurements for each analytical unit and perform similarly when used in our models. For each explanatory variable, we performed a one-way ANOVA test comparing their measurements derived from a 1-m grid cell, 30-m radius buffer, and 100-m radius buffer. No statistically significant differences were detected at the 0.05 significance level. For these reasons, all remote sensing products used in this study were resampled to a 1-m cell size using the nearest neighbor resampling method to maintain a consistent spatial resolution.

### 2.3. Generalized additive modeling

Generalized Additive Models (GAMs), provide a method to fit linear and non-linear variables within a model (Hastie and Tibshirani, 1986). GAMs were used to predict daily minimum and maximum air temperature anomalies throughout the entire city of Baltimore with the R package "mgcv" (Wood, 2020). Our explanatory variables consisted of static, satellite-derived and cadastral variables, time-varying, weather station-derived variables, and interaction terms between static and time-varying variables (full set of individual variables seen in Table 1). The sensor number variable was treated as a random effect term in both models to account for non-independence of the daily observations. For our minimum model, we defined our response variable (MinTA) as the daily difference in minimum temperature between the NWS station and each iButton. For our maximum model, we defined our response variable (MaxTA) as the daily difference in maximum temperature between the NWS station and each iButton. We refer to these

response variables as “anomalies,” and they should be interpreted as the spatial anomaly each day relative to temperature recorded at the NWS airport weather station.

To create our GAMs, we first calculated Pearson correlation coefficients and variance-inflation factors to avoid multicollinearity between explanatory variables in our models. Our datasets were then randomly split into training and testing subsets which excluded both dates from our study period and iButton location data. Our training and testing subsets comprised 70% and 30% of the data, respectively. Our modeling process began with a full-covariate model which consisted of a comprehensive suite of interaction terms between static and time-varying variables to explore those which exhibit a statistically significant response. The reduced-covariate models in our study removed interaction terms that were not statistically significant ( $\alpha > 0.05$ ). The reduced-covariate models contained all static, time-varying, and interaction terms as seen in Eq. (4), where  $T$  is TMIN or TMAX, gg is the identity link function, *Reduced* is the response variable for the reduced-covariate models,  $\beta_0$  is the intercept term;  $s$  is the smoothing function, *Sensor* is the random effect term, and  $\epsilon$  is the error term:

$$g(\text{Reduced}) = \beta_0 + s(LST) + s(Albedo) + s(NDVI) + s(SV) + s(Elevation) + s(CC) + s(T) + s(RH) + s(ADWS) + s(SR) + s(Sensor) + Land\ Use + Park + WT + s(CC^*RH) + s(CC^*T) + s(SV^*ADWS) + s(SV^*T) + s(ALBEDO^*T) + \epsilon \quad (4)$$

Using our training and testing subsets, we performed supervised backwards subset-selection. Monte Carlo Cross Validation, an iterative resampling method where data are randomly split into modeling and evaluation groups, was performed with 10 resamples to obtain adjusted  $R^2$  and RMSE values to evaluate the quality of each fitted GAM (Picard and Cook, 1984). The final, reduced-covariate models were then used to map the spatiotemporal variability of daytime and nighttime UHIs throughout Baltimore.

#### 2.4. Partial dependence and variable importance plots

GAMs were also implemented in this study to observe the relationships between response variables and both linear and non-linear explanatory variables. The influence of predictor variables was examined using Partial Dependence Plots (PDPs) and Variable Importance Plots (VIPs) (Hastie et al., 2009). PDPs displayed temperature anomaly predictions across the range of values for each continuous variable and can be interpreted as showing the marginal change in the prediction for a small change in the plotted variable; thus, the slope of a PDP shows the direction of a predictor’s influence on the prediction over a given range of predictor values (Badr et al., 2023). PDPs are used to explore and visualize the relationships between explanatory variables and air temperature anomalies. VIPs displayed the ranking of each continuous and categorical predictor by the computed mean absolute accumulated local effect (ALE) values. ALE values were used in this study as non-zero collinearity between these explanatory variables exist despite controlling for multicollinearity through statistical procedures. Further, ALEs calculate the marginal effect based on realistic model predictions (Molnar, 2022).

#### 2.5. Spatial and temporal variability

We examine the magnitude of UHI variability through two lenses: temporal variability and spatial variability. First, we calculate the temporal standard deviation in model-estimated daily MinTA and MaxTA for each grid cell. This yields a gridded map of temporal standard deviation ( $\sigma_{temporal,grid}$ ), calculated as:

$$\sigma_{temporal,grid} = \sqrt{\frac{\sum_{k=1}^K (x_{ijk} - \mu_{ij})^2}{K}}, \quad (5)$$

where  $x_{ijk}$  represents values within grid cells,  $i$  represents the value within the  $i$ th grid cell along the x-axis of a single daily map,  $j$  represents the value within the  $j$ th grid cell along the y-axis of a single daily map,  $k$  represents the  $k$ th daily map out of  $K$  total maps, and  $\mu_{ij}$  represents the mean value within grid cells across  $K$  total maps. Similarly, we calculate the temporal standard deviation across all iButton locations. This yields a map of temporal standard deviation ( $\sigma_{temporal,iButton}$ ) at iButton locations, calculated as:

$$\sigma_{temporal,iButton} = \sqrt{\frac{\sum_{k=1}^K (x_k - \mu)^2}{K}}, \quad (6)$$

where  $k$  and  $K$  have the same definition as in Eq. (5),  $x_k$  represents values at an iButton location and  $\mu$  represents the mean value at an iButton location.

We emphasize that our response variables (MinTA and MaxTA) are defined in terms of differences from the synoptic weather station, such that a large  $\sigma_{temporal,grid}$  at a given location indicates large variability in the difference between that location and the NWS station, and not necessarily large variability in absolute temperature at that location—for example, if the NWS station has large temperature variability over a given period while a location in the city has small temperature variability,  $\sigma_{temporal,grid}$  will be large for that point on account of the difference in behavior between that location and the weather station location.

Second, we calculate the standard deviation across all iButton locations on each day in the study period. This yields a timeseries of the spatial standard deviation values for iButtons ( $\sigma_{spatial,iButton}$ ), calculated as:

$$\sigma_{\text{spatial},i\text{Button}} = \sqrt{\frac{\sum_{n=1}^N (x_n - \mu)^2}{N}}, \quad (7)$$

where  $\sigma_{\text{spatial},i\text{Button}}$  represents spatial standard deviation,  $x_n$  represents values at iButton sensor locations,  $n$  represents the  $n$ th iButton sensor,  $N$  represents the total number of iButton sensors, and  $\mu$  represents the mean value across  $N$  total iButton sensors. Similarly, we calculate the spatial standard deviation for grid cells at iButton locations to assess whether our model can capture observed variability. This yields a timeseries of the spatial standard deviation values for grid cells ( $\sigma_{\text{spatial},\text{grid}}$ ), calculated as:

$$\sigma_{\text{spatial},\text{grid}} = \sqrt{\frac{\sum_{n=1}^N (x_{jn} - \mu)^2}{N}}, \quad (8)$$

where  $i$  and  $j$  have the same definition as in Eq. (5),  $n$  and  $N$  have the same definition as in Eq. (7),  $x_{jn}$  represents values within grid cells that encompass iButton sensors, and  $\mu$  represents the mean value across grid cells of the  $N$  total iButton sensors.

### 3. Results and discussion

#### 3.1. Performance metrics of nighttime and daytime UHI models

**Table 2** summarizes our nighttime (MinTA) and daytime (MaxTA) UHI model comparisons when using 30% holdout data. All covariates in the models were statistically significant at the  $\alpha < 0.05$  level. The RMSE values output by performing Monte Carlo cross-validation were used to compare each model, with lower RMSE values indicating a better fit to the data. The full-covariate and reduced-covariate are as described in Section 2.3. The static terms model only included satellite-derived variables and the static and time-varying terms model included both satellite-derived and meteorological variables.

The RMSE values ( $^{\circ}\text{C}$ ) decreased as additional covariates were added to the model with values ranging from 1.62 (static terms model) to 0.74 (full-covariate model) for the MinTA model and from 1.02 (static terms model) to 0.72 (full-covariate model) for the MaxTA model. In addition to RMSE value comparisons, predictive skill was assessed by comparing the adjusted  $R^2$  values, with a higher adjusted  $R^2$  value indicating a better fit of the data. The adjusted  $R^2$  values increased as additional covariates were added to the model with values ranging from 0.26 (static terms model) to 0.84 (full-covariate model) for the MinTA model and from 0.64 (static terms model) to 0.81 (full-covariate model) to for the MaxTA model.

While the full-covariate models predicting MinTA and MaxTA yield lower RMSE and adjusted  $R^2$  values, these models were not selected as our final models. Models such as these which contain many covariates can exhibit low bias and high variance, and they also complicate interpretation of model results. For this reason, we performed supervised backwards subset-selection to obtain reduced-variable models which could capture the spatial structure of the nighttime and daytime UHIs, albeit with higher RMSE and reduced adjusted  $R^2$  values. Our final models confirm strong predictive skill and acceptable model performance. The final, reduced-covariate models for MinTA and MaxTA predictions are seen in Eqs. (9) and (10) where  $g$ ,  $\beta_0$ ,  $s$ , *Sensor*, and  $\varepsilon$  have the same definition as in Eq. (4). Each GAM used a Gaussian distribution and restricted maximum likelihood (REML) for parameter smoothing:

$$g(\text{MinTA}) = \beta_0 + s(LST) + s(CC) + s(Elevation) + s(TMIN) + s(ADWS) + s(SR) + s(RH) + s(Sensor) + WT + s(CC*RH) + s(CC*TMIN) + \varepsilon \quad (9)$$

$$g(\text{MaxTA}) = \beta_0 + s(LST) + s(CC) + s(NDVI) + s(TMAX) + s(ADWS) + s(SR) + s(RH) + s(Sensor) + WT + s(CC*RH) + s(CC*TMAX) + \varepsilon \quad (10)$$

The final, reduced-covariate models for nighttime and daytime air temperature anomalies differ in the static variables included, with the MinTA model including elevation and not NDVI, and the reverse for the MaxTA model. These findings suggest NDVI played a larger role in characterizing daytime UHIs, which is consistent with results of [Chun and Guhathakurta, 2017](#) who found vegetation to be the dominant factor for daytime UHIs, and [Qiao et al., 2013](#) who found temperatures in urban environments to exhibit the strongest relationship with NDVI during the summer season and daytime conditions. Our nighttime model suggests that elevation plays a larger role in characterizing nighttime UHIs, which is consistent with findings from a study conducted by [Khandelwal et al., 2018](#), who found

**Table 2**  
Model comparisons based on predictive skill.

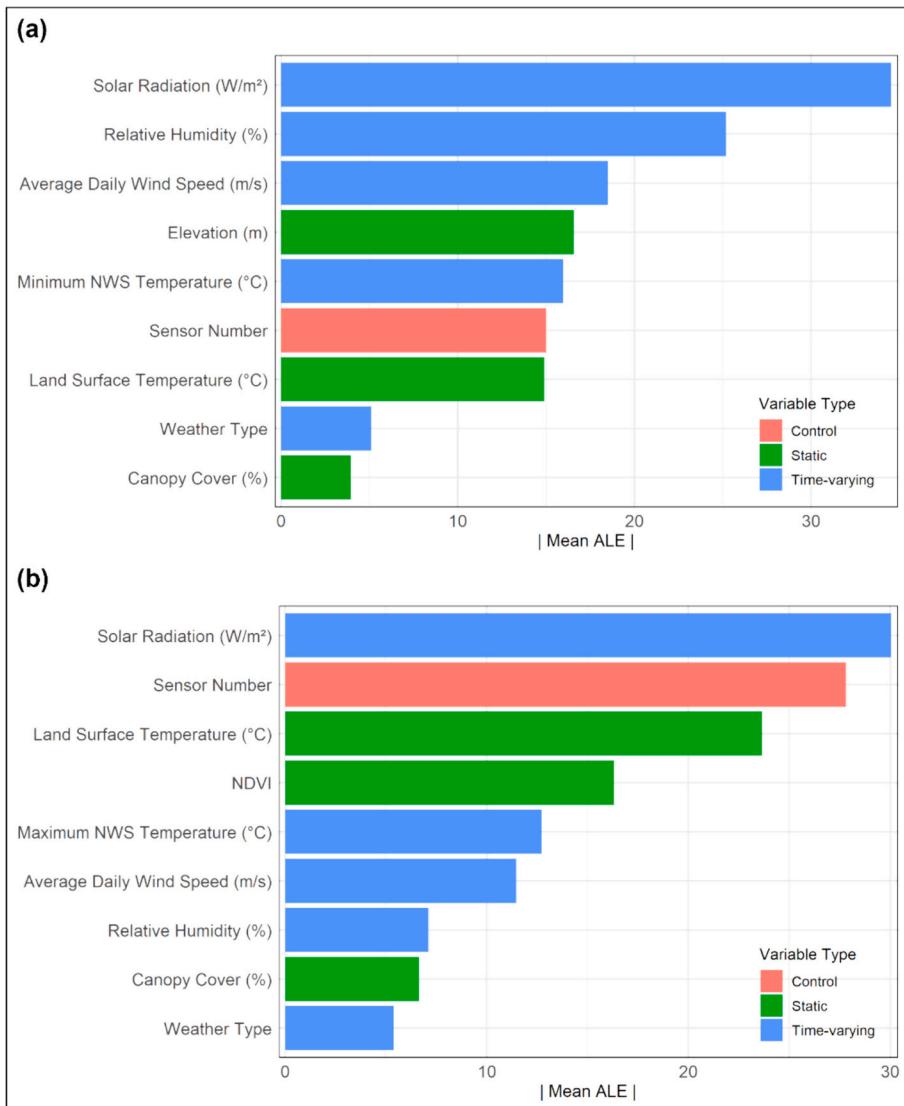
| Model                         | MinTA Model |                         | MaxTA Model |                         |
|-------------------------------|-------------|-------------------------|-------------|-------------------------|
|                               | RMSE        | Adjusted R <sup>2</sup> | RMSE        | Adjusted R <sup>2</sup> |
| Static Terms                  | 1.62        | 0.26                    | 1.02        | 0.64                    |
| Static and Time-Varying Terms | 0.76        | 0.83                    | 0.73        | 0.80                    |
| Full-Covariate Model          | 0.74        | 0.84                    | 0.72        | 0.81                    |
| Reduced-Covariate Model       | 0.75        | 0.84                    | 0.72        | 0.81                    |

that when using night-time images specifically, elevation exhibited negative relationships in all seasons.

The present study finds that static physical landscape characteristics alone do not sufficiently predict MinTA or MaxTA, as they exhibited the lowest adjusted  $R^2$  values of 0.26 and 0.64 and the highest RMSE values of 1.62 and 1.02, for the MinTA and MaxTA models, respectively. By including both static and time-varying variables in our models, we found significant improvements in our models' predictive skills, with decreased RMSE values of 0.76 and 0.73 and increased adjusted  $R^2$  values of 0.83 and 0.80 for the MinTA and MaxTA models, respectively. These model improvements predictive skill are consistent with the literature, as meteorological conditions have been found to be major determinants of UHI intensity (Santamouris, 2015). This underscores the need to include meteorological variables when characterizing summer nighttime and daytime UHIs.

### 3.2. Impact of temporal and spatial controls on spatial structure

The relative importance of the variables included in our final, reduced-covariate models is shown in Fig. 4. Our VIPs display the ranking of each predictor by their computed mean absolute ALE values. For our MinTA model, solar radiation ranked highest in importance (ALE = 34.56) followed by relative humidity (ALE = 25.20), average daily wind speed (ALE = 18.50), and elevation (ALE = 16.58) (Fig. 4a). For our MaxTA model, solar radiation again ranked highest (ALE = 30.08). This was followed by sensor number, a random effect variable that was included in both models to account for non-independence in our daily air temperature measurements. The importance of this variable in the MaxTA model suggests that sensor-specific or location-specific characteristics that are not



**Fig. 4.** Variable importance plots of our (a) MinTA model terms and (b) MaxTA model terms using mean absolute accumulated local effects (ALEs). Higher ALE values for a specific term indicate that the removal of that term will result in a larger reduction in the accuracy of the model's prediction.

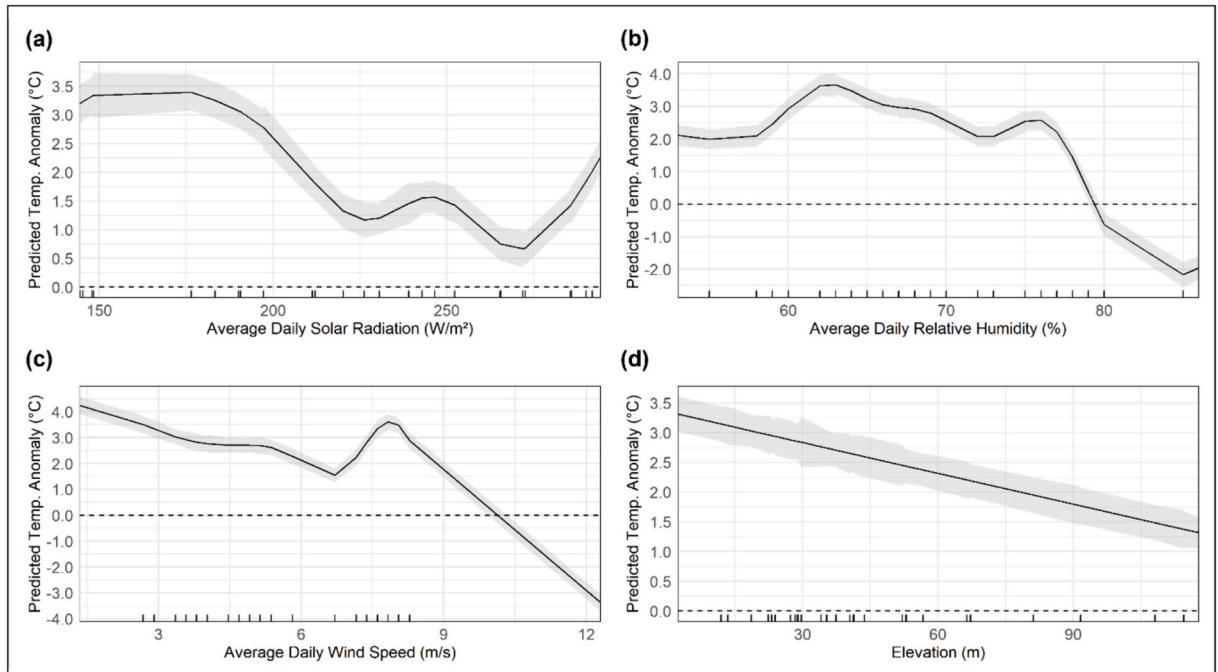
included in our set of predictors can have a significant impact on daytime temperatures. The meteorological and physical variables which follow solar radiation are land surface temperature (ALE = 23.67), NDVI (ALE = 16.31), and maximum NWS temperature (ALE = 12.73) (Fig. 4b). The three highest ranked variables in the MinTA model were NWS station-derived meteorological variables which exhibit temporal variability with the fourth highest being a physical characteristic which varies in space. However, while solar radiation also ranked highest in the MaxTA model, it did not demonstrate the same influence of meteorological conditions in predicting air temperature anomalies as observed in the MinTA model. Though there are some similarities in the list of important variables for predicting MinTA and MaxTA, the character of these relationships can vary significantly. These differences can be effectively illustrated through PDPs.

As seen in Fig. 5, MinTA has a u-shaped relationship with solar radiation (Fig. 5a) and negative associations with relative humidity, wind speed, and elevation (Fig. 5b-5d). In Fig. 6, we find that MaxTA exhibits a consistently positive relationship with solar radiation (Fig. 6a) and a linear, positive association with LST (Fig. 6b). In contrast, MaxTA has negative associations with NDVI and maximum NWS temperature (Fig. 6c-6d).

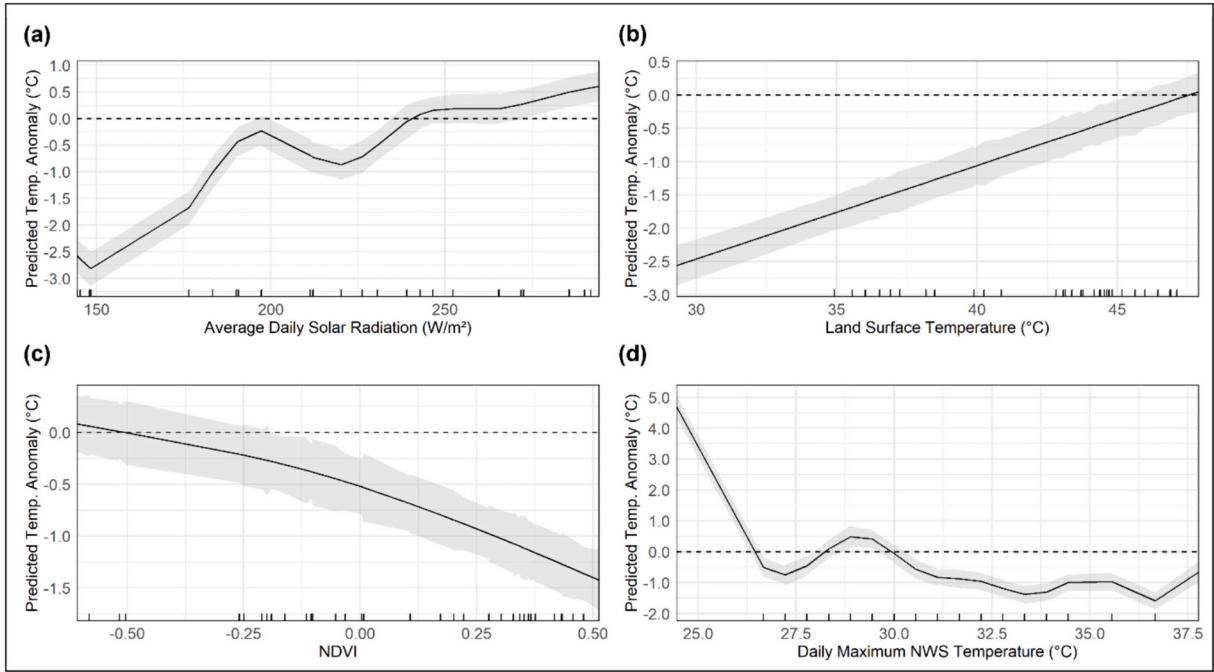
The relationship solar radiation has with MinTA and MaxTA illustrates how our models are able to capture diverse, non-linear relationships between the response and predictor variables. The positive association between MaxTA and solar radiation is likely a direct result of higher incoming solar radiation absorbed by the relatively dry urban surface, leading to local heating through sensible heat flux and atmospheric absorption of surface-emitted thermal infrared radiation (Oke, 1982). Some contribution of measurement bias also cannot be ruled out: though the iButtons were shielded from direct solar radiation, sunnier conditions could result in more radiative heating of the metal iButton casings in reflective or low shade environments (Maclean et al., 2021; Terando et al., 2017). The u-shaped MinTA relationship, however, requires further investigation, and might result from the fact that (daytime) solar radiation is an imperfect indicator of nighttime conditions; a direct estimate of overnight cloud cover may be more useful.

Focusing on the relationships illustrated in Fig. 5, we find the negative relationship between relative humidity and MinTA (Fig. 5b) to be consistent with the understanding that higher humidity conditions impede nighttime radiative cooling, through the atmospheric absorption and reemission of thermal infrared radiation. This effect is expected to be larger in open areas that normally have a wide sky view and cool relatively efficiently at night (e.g., the airport location of the NWS station or open grassy spaces in the city) than in densely urbanized areas in which urban form impedes nighttime radiative cooling even under low humidity conditions (Scott et al., 2018). The negative relationship between wind speed and MinTA (Fig. 5c) is intuitive, as higher wind speed would be expected to lead to greater mixing and less local structure in the UHI (e.g., Kim and Baik, 2002), though we caution that the dramatic negative slope at high wind speeds seen in Fig. 5c is the product of a small number of high wind days.

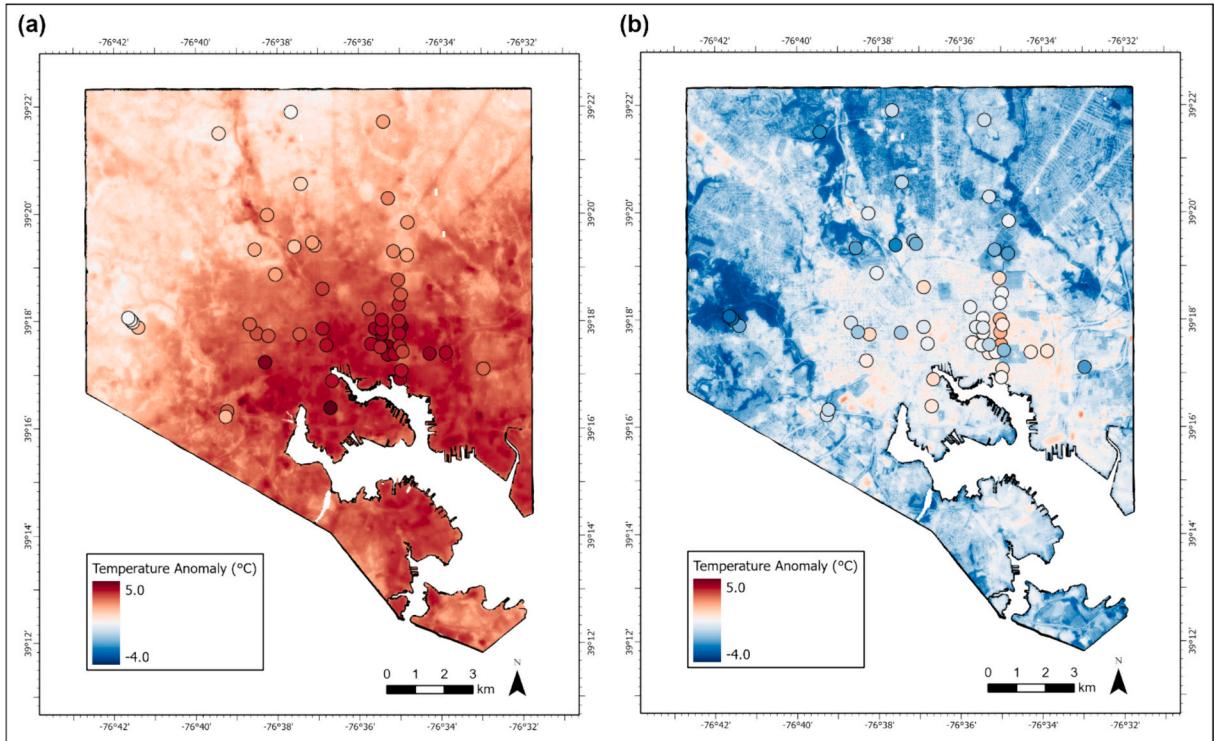
The negative association between elevation and MinTA (Fig. 5d) reflects the fact that higher elevations tend to be cooler, and thus more similar to the extra-urban conditions captured by the NWS station which is situated in an open field of grass within airport grounds (Khandelwal et al., 2018). For MaxTA, we see that LST is a strong predictor alongside solar radiation (Fig. 6b), with a



**Fig. 5.** Partial dependence plots of the four highest ranked covariates in our MinTA model based on computed mean absolute accumulated local effects: (a) solar radiation; (b) relative humidity; (c) average daily wind speed; and (d) elevation. Gray shaded bands denote the 95% confidence interval.



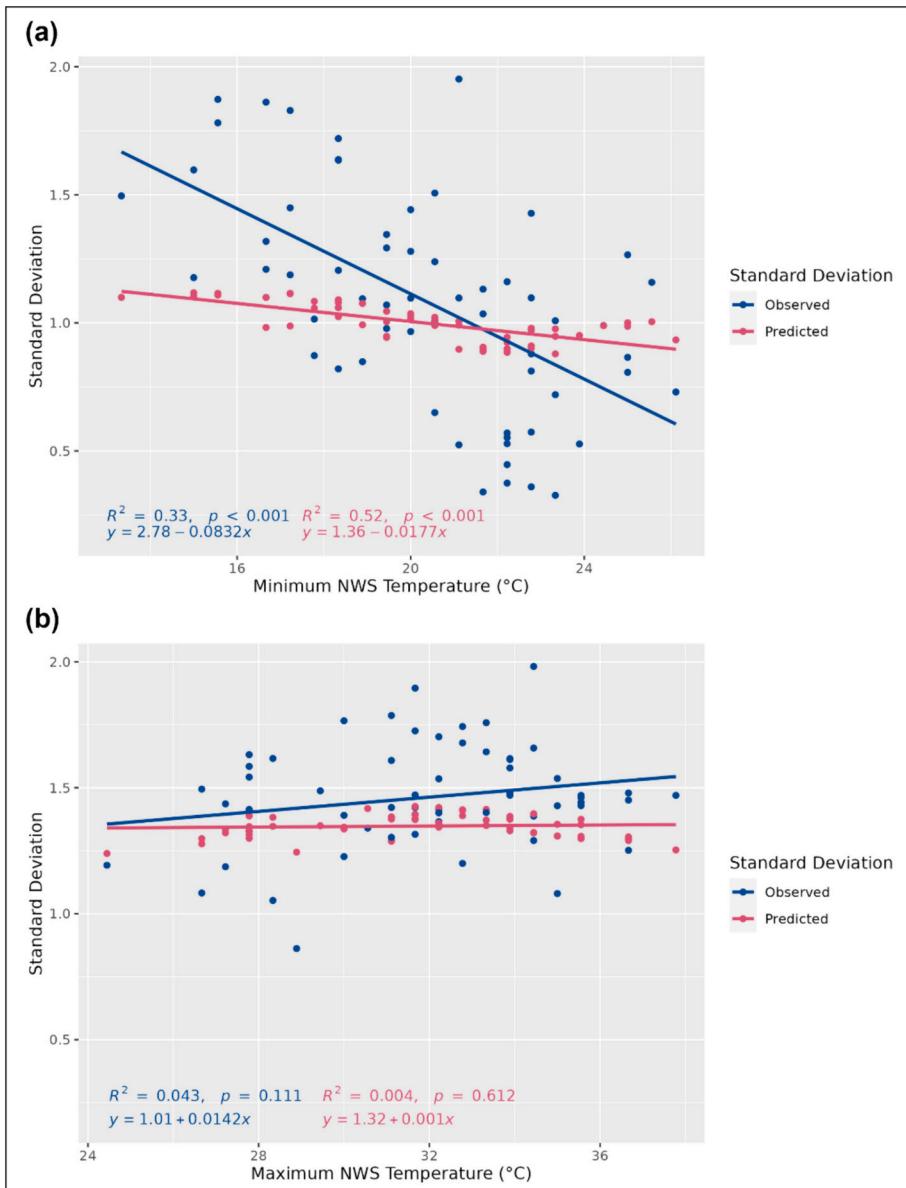
**Fig. 6.** Partial dependence plots of the four highest ranked covariates in our MaxTA model based on computed mean accumulated local effects: (a) solar radiation; (b) land surface temperature; (c) normalized difference vegetation index; and (d) maximum NWS temperature. Gray shaded bands denote the 95% confidence interval.



**Fig. 7.** Spatial distribution of air temperature anomalies from our (a) MinTA model and (b) MaxTA model averaged across our study period of July to September 2016. Overlayed circles represent (a) nighttime air temperature anomalies and (b) daytime air temperature anomalies from our iButton observations averaged across our study period. Air temperature anomalies are defined relative to the NWS station and are displayed at 1-m resolution.

relatively linear, positive association. This is not a surprise, given the general correlation that exists between satellite-derived LST and air temperature, but the fact that the association is more predictive for MaxTA than it is for MinTA reflects the more defined local structure of the air temperature UHI during the day and is consistent with the understood importance of local radiative balance to air temperature under daytime conditions. We emphasize that LST in this study is a static variable, drawn from a single Landsat image, so the relationship observed here should be interpreted as an indication of the spatial character of the UHI and not a reflection of temporal correlation.

NDVI is another static (in this study) variable that shows strong association with MaxTA (Fig. 6c): more highly vegetated areas show lower MaxTA values, as green areas in the city have reduced air temperature and are more similar to conditions outside of the city during the day. This is consistent with findings from Chun and Guhathakurta, 2017 where NDVI was found to be significant in estimating daytime UHIs rather than nighttime UHIs. The relationship between maximum NWS temperature and MaxTA tends to be negative across the range of NWS temperature data, but the association is dominated by a strong decline in predicted MaxTA with NWS



**Fig. 8.** The relationships between (a) spatial standard deviation and nighttime, minimum NWS air temperature and (b) spatial standard deviation and daytime, maximum NWS air temperature. Pink dots represent the spatial standard deviations from our (a) MinTA model and (b) MaxTA model. Blue dots represent the spatial standard deviations from our (a) minimum, nighttime air temperature anomalies and (b) maximum, daytime air temperature anomalies observed in iButton measurements. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

temperature held down by cooler days with NWS temperatures  $<27^{\circ}\text{C}$  (Fig. 6d). The negative character of this relationship is consistent with studies that have found a reduction in UHI under warmer conditions (e.g., Scott et al., 2018), but the dramatic nonlinear effect at low temperatures requires further investigation before any specific mechanistic conclusions can be drawn.

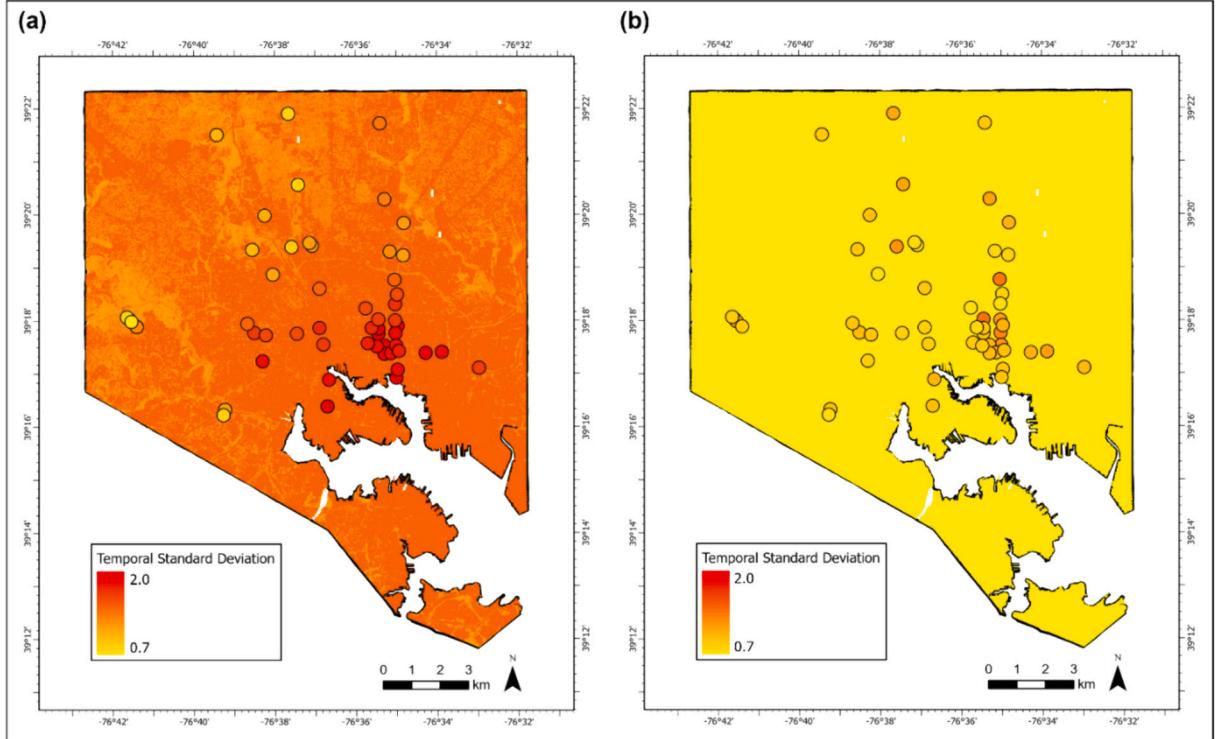
### 3.3. Spatial variability of daytime and nighttime UHIs

Our final MinTA and MaxTA models were used to map spatial patterns of nighttime and daytime UHIs. The spatial pattern of time-averaged MinTA and MaxTA for our study period are seen in Fig. 7. At night, we find air temperatures throughout the city to be warmer than the NWS station, on average (Fig. 7a), as the UHI exhibits a smooth, city-wide spatial pattern, with a gradient in UHI magnitude between the downtown area and the city outskirts. However, we find a broader distribution of air temperatures during the day, with the center of the city warmer than the weather station and areas outside of the city center cooler (Fig. 7b). These cooler in-city areas are associated with parks, urban tree stands, and highly tree-covered neighborhoods that are more forest-like than the NWS station, which is located in an open airport field. Despite these differences, both the nighttime and daytime UHI effect are seen to be concentrated in the highly urbanized, Baltimore downtown area in the center of the city. Our findings are in agreement with a number of studies which have also demonstrated industrial, urban centers being warmer than their surrounding regions (e.g., Chun and Guhathakurta, 2017; Qiao et al., 2013; Saverino et al., 2021).

Fig. 7 also presents varying degrees of UHI structure in nighttime and daytime UHIs. We find the nighttime UHI to have less structure than the daytime UHI. In both the observations and models, differences in air temperature anomalies between parks and non-parks were found. The average MaxTA in parks is  $-2.09^{\circ}\text{C}$  for both the model and observations, while the non-park average MaxTA is  $-0.72^{\circ}\text{C}$  in the model and  $-0.10^{\circ}\text{C}$  in the observations. The average MinTA in parks is  $1.91^{\circ}\text{C}$  in the model and  $1.86^{\circ}\text{C}$  in the observations, while the non-park average MinTA is  $2.57^{\circ}\text{C}$  in the model and  $2.78^{\circ}\text{C}$  in the observations. While parks are seen to be cooler during both nighttime and daytime conditions with respect to their surroundings, their cooling effect is more pronounced during the daytime. At night, these areas with high tree canopy cover are more similar to their surroundings. A study conducted by Azevedo et al., 2016 similarly found nighttime UHIs to be less defined and suggested advection as the major influence.

### 3.4. Spatiotemporal variability of daytime and nighttime UHIs

We now turn to the question of variability in MinTA and MaxTA. First, we consider temporal variability by looking at the spatial standard deviation in iButton temperature anomalies calculated for each day. We are concerned with both observed variability and our

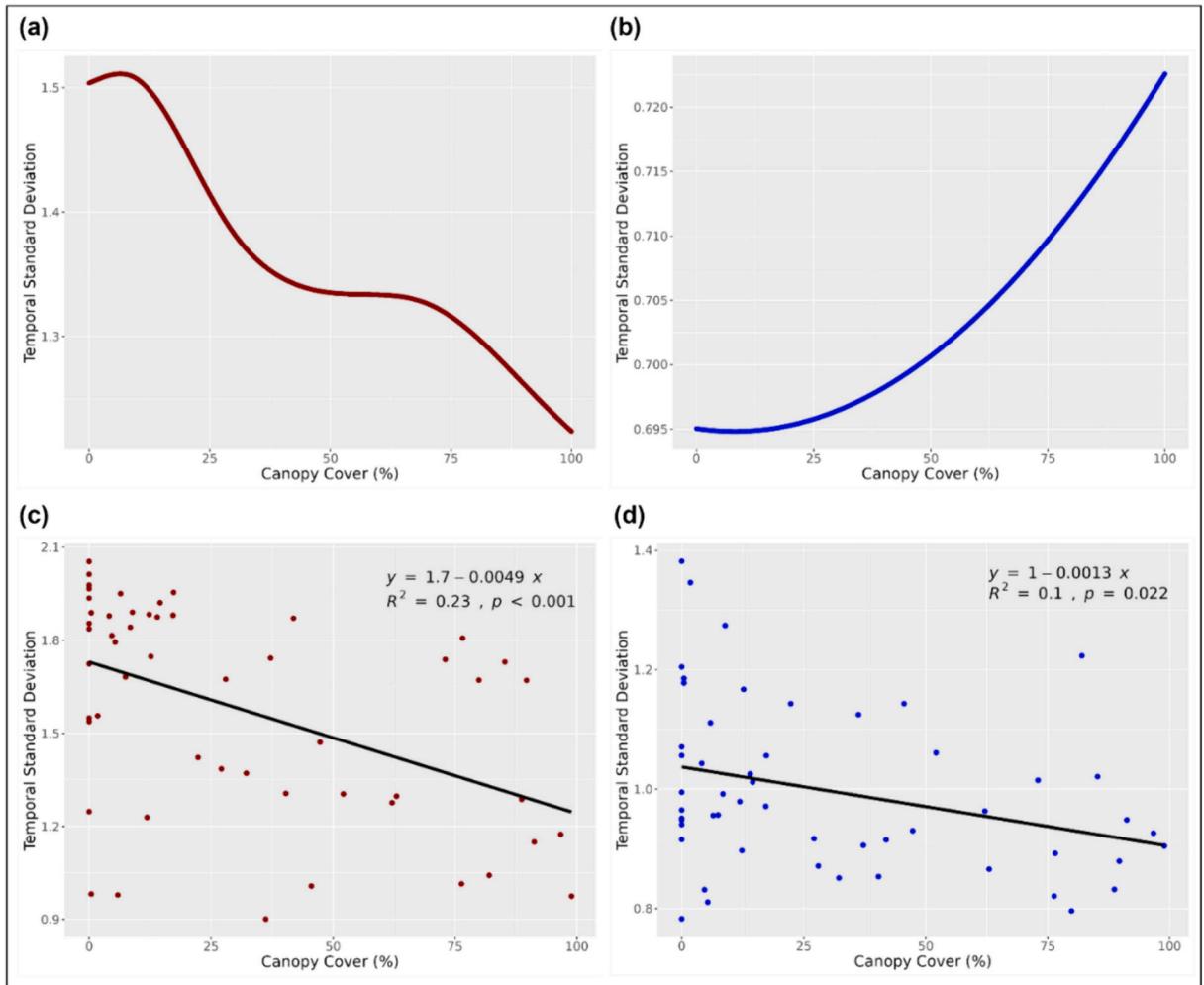


**Fig. 9.** Spatial patterns of temporal variability (temporal standard deviations, SD) from our (a) MinTA model and (b) MaxTA model. Overlaid circles represent (a) temporal SD of minimum, nighttime air temperature anomalies and (b) temporal SD of maximum, daytime air temperature anomalies observed in iButton measurements.

models' ability to capture it, noting that NWS air temperature is a significant but not a leading predictor in our MinTA and MaxTA models. Results for both observation and model are shown in Fig. 8. In Fig. 8a, we see that as the minimum NWS temperature increases, both the observed and predicted spatial standard deviations in MinTA decrease. At night, both observed and predicted spatial standard deviations exhibit a negative linear trend, with slopes of  $-0.083$  and  $-0.017$ , respectively. These trends are accompanied by statistically significant  $R^2$  values of  $0.33$  and  $0.52$ , respectively. In Fig. 8b, we see that as the maximum NWS temperature increases, the predicted and observed MaxTA standard deviations are flat or show a slight increase. During the day, the observed and predicted spatial standard deviations have slopes of  $0.014$  and  $0.001$ , respectively. However, their corresponding  $R^2$  values of  $0.043$  and  $0.004$  are not statistically significant.

Notably, our models underestimate the spatial structure of the nighttime and daytime UHIs. They do, however, capture the general direction of the relationships. The tendency towards greater nighttime UHI structure at lower temperature is consistent with previous studies, which have pointed to reduced radiative cooling in open areas under hot, humid conditions as the reason for reduced urban-rural contrast (e.g., Scott et al., 2018). Previous studies have also found that the relationship between UHI structure and synoptic air temperature is noisier for daytime conditions, as processes other than radiative cooling play more significant roles during the day.

Turning to the modeled spatial pattern of temporal variability (Fig. 9), we find that standard deviations in MinTA are largest in areas of little to no tree canopy cover and smallest in areas of high tree canopy cover, with a Pearson correlation coefficient of  $-0.97$  (Fig. 9a). This spatial pattern can also be seen in iButton observations. Fig. 10 further emphasized this point by plotting the relationship between the temporal variability and percent tree canopy over. Consistent with Fig. 9a, Fig. 10a shows that the temporal standard deviation in MinTA drops with increasing tree canopy cover in our model. The results for iButton observations are noisy, but the lowest variability is generally found at the highest tree canopy sites (Fig. 10c). For MaxTA, modeled patterns of temporal variability are muted (Fig. 9b). Similarly, the temporal variability of MaxTA calculated from our iButton observations show some spatial variability but no



**Fig. 10.** The relationships between canopy cover and temporal standard deviation from our: (a) MinTA model; (b) MaxTA model; (c) minimum, nighttime air temperature anomalies observed in iButton measurements; and (d) maximum, daytime air temperature anomalies observed in iButton measurements.

systematic spatial structure (Fig. 9b). The relationship between predicted and observed variability in MaxTA and percent tree canopy cover is small: the model shows a positive association with a Pearson correlation coefficient of 0.93, but over an extremely small range of variability (Fig. 10b), while iButton observations are dominated by noise (Fig. 10d).

The pattern seen in Fig. 9a reflects our definition of MinTA (and MaxTA): since we define anomalies relative to the NWS station, places that have temperature variability similar to the station will tend to have low MinTA standard deviation while those with different drivers or expressions of temperature variability will have relatively large temporal standard deviation. Drawing from the reasoning and findings of previous studies described above (e.g., Scott et al., 2018; Kim and Baik, 2002; Ścieżor, 2020), in the case of the nighttime UHI we understand that efficiency of radiative cooling underlies much of this difference in temporal variability. The NWS station and less urbanized areas of Baltimore City will have higher nighttime temperatures under humid conditions, when radiative cooling is compromised. On those nights the station and urban green areas become more similar to densely urbanized locations, while on nights with lower humidity (which are generally cooler at night) their air temperatures diverge from the elevated heat conditions of the urban core. Since MinTA is defined relative to the NWS station, this phenomenon maps as a higher temporal standard deviation in the urban core. While the modeled patterns of temporal variability for MaxTA are muted, this result must be interpreted in the context of a statistical model that fails to capture much of the spatial variance. Indeed, the iButton observations did not exhibit systematic spatial structure at city scale. The variability observed in our MaxTA observations may be due to differences in sun exposure of our iButton thermochrons at various locations and subject to noise due to measurement error—factors that could contribute to the large importance of the random effect variable in our model of MaxTA (Maclean et al., 2021; Terando et al., 2017)—as well as to true microclimate variability, as captured in the variance that the model is able to explain.

Finally, we are interested in the nature of interactions between temporally-varying synoptic variables (as measured at weather stations) and spatially-varying landscape variables. Our hypothesis in designing this study was that these interactions could be diagnosed in our statistical models and thus help us to explain the character of relationships between neighborhood-scale urban form and the urban heat archipelago. In a formal sense, this hypothesis was confirmed: we found statistically significant interaction terms in both our MinTA and MaxTA models (see Eqs. (9 and 10)). In each model, the significant interaction terms all involve canopy cover.

For MinTA, we find that interactions between canopy cover and NWS station (i.e., synoptic) relative humidity and minimum air temperature were significant. For relative humidity, we find that high humidity conditions are associated with reduced MinTA, consistent with the understanding that high background humidity reduces nighttime radiative cooling across all land cover types. The interaction term shows that this effect is slightly stronger for low canopy cover sites, as we would expect if the mechanism at fine scale mirrors that seen at city-scale: that the humidity “blanketing” effect is greater at sites that are relatively open and not obstructed by trees or other barriers to radiative cooling. The result for NWS minimum air temperature is in the same direction and is consistent with elevated dewpoint (i.e., higher humidity) conditions leading to lower MinTA.

For MaxTA, we find statistically significant interaction terms between canopy cover and NWS station relative humidity and maximum air temperature. The statistical significance of the interaction term involving NWS maximum air temperature is dominated by response at the low end of the air temperature range, and is not particularly relevant for understanding heat risks. The relative humidity response is noisy, suggesting strong local effects, but generally shows that MaxTA declines more dramatically for high canopy cover sites than for low canopy cover sites under low synoptic relative humidity conditions. In other words, under dry weather patterns, higher canopy cover sites see a reduction in MaxTA. This observation requires further investigation, but it is generally consistent with a mechanism under which elevated evaporative demand leads to increased transpiration from tree-covered sites, reducing local temperature relative to surrounding, lower canopy cover areas. As with all of our interaction term results, this effect is statistically significant but small in magnitude, and we hesitate to overinterpret the physical importance of the result based on this diagnostic study alone.

#### 4. Conclusions

Numerous studies have examined the spatiotemporal structure of UHIs, but they have not captured temperature variability across an entire city at both high spatial and temporal resolutions, as this study has accomplished (e.g., G. Huang et al., 2011; Scott et al., 2017). Key features of the current study include: (1) the use of a high-density urban air temperature monitoring network; (2) application of GAMs to capture nonlinear relationships between proposed predictors and local air temperature; (3) direct comparison of statistical models of daytime vs. nighttime urban temperature anomalies; and (4) a focus on temporal variability in spatial structure of the UHI, including the derivation of statistical models that predict daily changes in urban air temperature anomalies at fine spatial scales. None of these features are entirely unique to this study, but together they offer new insights on UHI processes and some practical takeaways for UHI monitoring to support outdoor heat mitigation strategies.

First, we find that while daytime and nighttime UHIs can be characterized by similar variables, the same meteorological and physical conditions may influence them differently. Among all variables studied, the most influential variable for predicting local nighttime and daytime urban temperature anomalies is solar radiation, but the shape of the relationship—and thus the physical reason for the association—is different between night and day. Beyond solar radiation, we find that nighttime local temperature anomalies were most influenced by daily meteorological conditions whereas daytime local temperature anomalies are influenced equally by quasi-static landscape conditions and time-varying factors. These results point to the need to consider the specific purpose of outdoor, extreme heat mitigation when implementing climate adaptation strategies. Strategies that target the daytime UHI, for example, may be less effective at reducing nighttime temperatures.

Second, we find that localized urban temperature anomalies—i.e., how much warmer or cooler it is in different locations in the city relative to a synoptic NWS station placed at a nearby airport—vary in time as well as in space. This indicates that a simple application

of a time-averaged urban heat island map may be inadequate when estimating heat risk across a city. Instead, local government officials concerned with heat response and mitigation might consider using an analysis that accounts for the fact that changing weather conditions such as weakening winds or increased humidity can change patterns of heat across the city. This could include differences in which neighborhoods fall above heat advisory thresholds.

Third, we find that there is meaningful information in a high density, stationary temperature monitoring network. The iButtons capture meaningful spatial and spatiotemporal variability patterns, and the presence of such networks is required to develop and evaluate models that predict urban heat as a function of weather conditions and physical characteristics. Localized differences in air temperatures throughout the city relative to a central weather service station may be important information for local government officials and city planners who rely on weather service station data to create heat adaptation and mitigation strategies or issue heat-related weather alerts.

As previously noted, our nighttime and daytime models underestimate the observed spatial variability of minimum and maximum daily air temperature anomalies. For MinTA, the model underestimates spatial variability but the magnitude of variability was modest for both observation and models, such that the error might not have major implications for heat exposure analyses. For daytime, site-specific variability in observations is relatively high. The model's inability to capture the magnitude of this variability could mean that there are site-specific factors not accounted for in the model and/or that variability in observations is strongly influenced by measurement error. These limitations could be addressed in future studies by ensuring instruments are better shaded and actively aspirated to reduce impacts of sun exposure. Our models also fail to capture the magnitude of temporal variability in fine-scaled UHI structure, especially for daytime conditions. Including interaction terms in the model does make it structurally possible to capture relationships that drive such space-time variability, and we do see some statistical significance in these terms, but further work with refined or extended datasets might be necessary to capture the full magnitude of these terms. While these models were not created to obtain the highest possible predictive power, we acknowledge these limitations and note our models can be improved in these ways. Despite these limitations, the presented models provide reasonable predictive performance and interpretability of temporal and spatial patterns. It is our hope that the methodology presented in this paper can be further developed to aid cities in their air temperature monitoring plans and practices.

#### CRediT authorship contribution statement

**Bianca Corpuz:** Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Benjamin Zaitchik:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Darryn Waugh:** Writing – review & editing, Data curation. **Anna Scott:** Writing – review & editing, Data curation. **Tom Logan:** Writing – review & editing, Data curation.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Data availability

Data will be made available on request.

#### Acknowledgements

This research was supported by the National Oceanic and Atmospheric Administration (NOAA) Mid-Atlantic Regional Integrated Sciences and Assessments (MARISA) and NOAA Climate Programs Office (CPO) Award (No. NA21OAR4310147) and by the US Department of Energy (DOE) Office of Science Award (No. DE-SC0023217).

#### References

- Anderson, M.C., Yang, Y., Xue, J., Knipper, K.R., Yang, Y., Gao, F., Hain, C.R., Kustas, W.P., Cawse-Nicholson, K., Hulley, G., Fisher, J.B., Alfieri, J.G., Meyers, T.P., Prueger, J., Baldocchi, D.D., Rey-Sanchez, C., 2021. Interoperability of ECOSTRESS and Landsat for mapping evapotranspiration time series at sub-field scales. *Remote Sens. Environ.* 252, 112189 <https://doi.org/10.1016/j.rse.2020.112189>.
- Azevedo, J., Chapman, L., Muller, C., 2016. Quantifying the daytime and nighttime urban Heat Island in Birmingham, UK: a comparison of satellite derived land surface temperature and high resolution air temperature observations. *Remote Sens.* 8 (2), 153. <https://doi.org/10.3390/rs8020153>.
- Badr, H.S., Colston, J.M., Nguyen, N.-L.H., Chen, Y.T., Burnett, E., Ali, S.A., Rayamajhi, A., Satter, S.M., Van Trang, N., Eibach, D., Krumkamp, R., May, J., Adegnika, A.A., Manouana, G.P., Kremsner, P.G., Chilengi, R., Hatyoka, L., Debes, A.K., Ateudjieu, J., Kosek, M.N., 2023. Spatiotemporal variation in risk of Shigella infection in childhood: a global risk mapping and prediction model using individual participant data. *Lancet Glob. Health* 11 (3), e373–e384. [https://doi.org/10.1016/S2214-109X\(22\)00549-6](https://doi.org/10.1016/S2214-109X(22)00549-6).
- Brown, L.T., 2021. *The Black Butterfly: The Harmful Politics of Race and Space in America*. Johns Hopkins University Press.
- Chun, B., Guhathakurta, S., 2017. Daytime and nighttime urban heat islands statistical models for Atlanta. *Environ. Plann. B: Urban Analyt. City Sci.* 44 (2), 308–327. <https://doi.org/10.1177/0265813515624685>.
- Duan, R.-R., Hao, K., Yang, T., 2020. Air pollution and chronic obstructive pulmonary disease. *Chron. Diseases. Translation. Med.* 6 (4), 260–269. <https://doi.org/10.1016/j.cdtm.2020.05.004>.

- Fenner, D., Holtmann, A., Meier, F., Langer, I., Scherer, D., 2019. Contrasting changes of urban heat island intensity during hot weather episodes. Environ. Res. Lett. 14 (12), 124013 <https://doi.org/10.1088/1748-9326/ab506b>.
- Good, E.J., 2016. An in situ-based analysis of the relationship between land surface “skin” and screen-level air temperatures: land skin-air temperature relationship. J. Geophys. Res. Atmos. 121 (15), 8801–8819. <https://doi.org/10.1002/2016JD025318>.
- Gubler, M., Christen, A., Remund, J., Brönnimann, S., 2021. Evaluation and application of a low-cost measurement network to study intra-urban temperature differences during summer 2018 in Bern, Switzerland. Urban Clim. 37, 100817 <https://doi.org/10.1016/j.ul clim.2021.100817>.
- Hastie, T., Tibshirani, R., 1986. Generalized additive models. Stat. Sci. 1 (3) <https://doi.org/10.1214/ss/1177013604>.
- Hastie, T., Tibshirani, R., Friedman, J.H., 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer.
- Hoffman, J.S., Shandas, V., 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 (1), 12. <https://doi.org/10.3390/cl8010012>.
- Huang, G., Zhou, W., Cadenasso, M.L., 2011. Is everyone hot in the city? Spatial pattern of land surface temperatures, land cover and neighborhood socioeconomic characteristics in Baltimore, MD. J. Environ. Manag. 92 (7), 1753–1759. <https://doi.org/10.1016/j.jenvman.2011.02.006>.
- Huang, B., Wang, J., Song, H., Fu, D., Wong, K., 2013. Generating high spatiotemporal resolution land surface temperature for urban Heat Island monitoring. IEEE Geosci. Remote Sens. Lett. 10 (5), 1011–1015. <https://doi.org/10.1109/LGRS.2012.2227930>.
- Hulley, G., Shivers, S., Wetherley, E., Cudd, R., 2019. New ECOSTRESS and MODIS land surface temperature data reveal fine-scale heat vulnerability in cities: a case study for Los Angeles County, California. Remote Sens. 11 (18), 2136. <https://doi.org/10.3390/rs11182136>.
- IPCC, 2022. *Summary for Policymakers (Climate Change 2022): Impacts, Adaptation, and Vulnerability*. Cambridge University Press, pp. 3–33.
- Khandelwal, S., Goyal, R., Kaul, N., Mathew, A., 2018. Assessment of land surface temperature variation due to change in elevation of area surrounding Jaipur, India. Egypt. J. Remote Sens. Space Sci. 21 (1), 87–94. <https://doi.org/10.1016/j.ejrs.2017.01.005>.
- Kim, Y.-H., Baik, J.-J., 2002. Maximum urban Heat Island intensity in Seoul. J. Appl. Meteorol. 41 (6), 651–659. [https://doi.org/10.1175/1520-0450\(2002\)041<0651:MUHII>2.0.CO;2](https://doi.org/10.1175/1520-0450(2002)041<0651:MUHII>2.0.CO;2).
- Kleerekooper, L., van Esch, M., Salcedo, T.B., 2012. How to make a city climate-proof, addressing the urban heat island effect. Resour. Conserv. Recycl. 64, 30–38. <https://doi.org/10.1016/j.resconrec.2011.06.004>.
- Kong, J., Zhao, Y., Carmeliet, J., Lei, C., 2021. Urban Heat Island and its interaction with heatwaves: a review of studies on mesoscale. Sustainability 13 (19), 10923. <https://doi.org/10.3390/su131910923>.
- Li, D., Liao, W., Rigden, A.J., Liu, X., Wang, D., Malyshev, S., Shevliakova, E., 2019. Urban heat island: aerodynamics or imperviousness? Sci. Adv. 5 (4), eaau4299. <https://doi.org/10.1126/sciadv.aau4299>.
- Liang, S., 2001. Narrowband to broadband conversions of land surface albedo I. Remote Sens. Environ. 76 (2), 213–238. [https://doi.org/10.1016/S0034-4257\(00\)00205-4](https://doi.org/10.1016/S0034-4257(00)00205-4).
- Logan, T.M., Zaitchik, B., Guikema, S., Nisbet, A., 2020. Night and day: the influence and relative importance of urban characteristics on remotely sensed land surface temperature. Remote Sens. Environ. 247, 111861 <https://doi.org/10.1016/j.rse.2020.111861>.
- Maclean, I.M.D., Duffy, J.P., Haesen, S., Govaert, S., De Frenne, P., Vanneste, T., Lenoir, J., Lembrechts, J.J., Rhodes, M.W., Van Meerbeek, K., 2021. On the measurement of microclimate. Methods Ecol. Evol. 12 (8), 1397–1410. <https://doi.org/10.1111/2041-210X.13627>.
- Maryland DHMH, 2016. 2016 Heat-Related Illness Surveillance Report, pp. 1–6. <https://health.maryland.gov/preparedness/Documents/2016%20Summary%20Heat%20Report.pdf>.
- Maryland iMAP Portal, 2013. Maryland canopy cover [dataset]. REST Services Directory. <https://imap.maryland.gov/>.
- Maryland iMAP Portal, 2015. City of Baltimore LiDAR [dataset]. Maryland LiDAR Inventory. <https://imap.maryland.gov/>.
- Molnar, C., 2022. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*, Second Edition. Christoph Molnar.
- NOAA, 2015. 2015 USDA NAIP 4-Band 8 Bit Imagery: Maryland [Dataset]. Data Access Viewer, Digital Coast. <https://coast.noaa.gov/dataviewer/#/>.
- Oke, T.R., 1982. The energetic basis of the urban heat island. Q. J. R. Meteorol. Soc. 108 (455), 1–24. <https://doi.org/10.1002/qj.49710845502>.
- Picard, R.R., Cook, R.D., 1984. Cross-validation of regression models. J. Am. Stat. Assoc. 79 (387), 575–583. <https://doi.org/10.1080/01621459.1984.10478083>.
- Qiao, Z., Tian, G., Xiao, L., 2013. Diurnal and seasonal impacts of urbanization on the urban thermal environment: a case study of Beijing using MODIS data. ISPRS J. Photogramm. Remote Sens. 85, 93–101. <https://doi.org/10.1016/j.isprsjprs.2013.08.010>.
- Rodrigues, M., Santana, P., Rocha, A., 2019. Effects of extreme temperatures on cerebrovascular mortality in Lisbon: a distributed lag non-linear model. Int. J. Biometeorol. 63 (4), 549–559. <https://doi.org/10.1007/s00484-019-01685-2>.
- Santamouris, M., 2015. Analyzing the heat island magnitude and characteristics in one hundred Asian and Australian cities and regions. Sci. Total Environ. 512–513, 582–598. <https://doi.org/10.1016/j.scitotenv.2015.01.060>.
- Saverino, K.C., Routman, E., Lookingbill, T.R., Eanes, A.M., Hoffman, J.S., Bao, R., 2021. Thermal inequity in Richmond, VA: the effect of an unjust evolution of the urban landscape on urban Heat Islands. Sustainability 13 (3), 1511. <https://doi.org/10.3390/su13031511>.
- Ściążor, T., 2020. The impact of clouds on the brightness of the night sky. J. Quant. Spectrosc. Radiat. Transf. 247, 106962 <https://doi.org/10.1016/j.jqsrt.2020.106962>.
- Scott, A.A., Zaitchik, B., Waugh, D.W., O’Meara, K., 2017. Intraurban temperature variability in Baltimore. J. Appl. Meteorol. Climatol. 56 (1), 159–171. <https://doi.org/10.1175/JAMC-D-16-0232.1>.
- Scott, A.A., Waugh, D.W., Zaitchik, B.F., 2018. Reduced urban Heat Island intensity under warmer conditions. Environ. Res. Lett. 13 (6), 064003 <https://doi.org/10.1088/1748-9326/aabd6c>.
- Shandas, V., Voelkel, J., Williams, J., Hoffman, J., 2019. Integrating satellite and ground measurements for predicting locations of extreme urban heat. Climate 7 (1), 5. <https://doi.org/10.3390/cl7010005>.
- Shepherd, J.M., Andersen, T., Strother, C., Horst, A., Lahouari, B., Mitra, C., 2018. Urban climate archipelagos: a new framework for urban impacts on climate. Earthzine. <https://earthzine.org/urban-climate-archipelagos-a-new-framework-for-urban-impacts-on-climate/>.
- Sheridan, S.C., Allen, M.J., 2015. Changes in the frequency and intensity of extreme temperature events and human health concerns. Curr. Clim. Chang. Rep. 1 (3), 155–162. <https://doi.org/10.1007/s40641-015-0017-3>.
- Shi, R., Hobbs, B.F., Zaitchik, B.F., Waugh, D.W., Scott, A.A., Zhang, Y., 2021. Monitoring intra-urban temperature with dense sensor networks: fixed or mobile? An empirical study in Baltimore, MD. Urban Clim. 39, 100979 <https://doi.org/10.1016/j.ul clim.2021.100979>.
- Smith, R.B., 2010. The heat budget of the earth’s surface deduced from space. Yale University Center for Earth Observation Rep. [http://yceo.yale.edu/sites/default/files/files/Surface\\_Heat\\_Budget\\_From\\_Space.pdf](http://yceo.yale.edu/sites/default/files/files/Surface_Heat_Budget_From_Space.pdf).
- Terando, A.J., Youngsteadt, E., Meineke, E.K., Prado, S.G., 2017. Ad hoc instrumentation methods in ecological studies produce highly biased temperature measurements. Ecol. Evol. 7 (23), 9890–9904. <https://doi.org/10.1002/ece3.3499>.
- Tuholske, C., Caylor, K., Funk, C., Verdin, A., Sweeney, S., Grace, K., Peterson, P., Evans, T., 2021. Global urban population exposure to extreme heat. Proc. Natl. Acad. Sci. 118 (41), e2024792118 <https://doi.org/10.1073/pnas.2024792118>.
- U.S. Geological Survey, 2016. Landsat Collection 2-Level 2 [Dataset]. USGS Earth Explorer. <https://earthexplorer.usgs.gov/>.
- U.S. Geological Survey, 2023. How Do I Use a Scale Factor with Landsat Level-2 Science Products? [https://www.usgs.gov/faqs/how-do-i-use-scale-factor-landsat-level-2-science-products?qt-news\\_science\\_products=0#qt-news\\_science\\_products](https://www.usgs.gov/faqs/how-do-i-use-scale-factor-landsat-level-2-science-products?qt-news_science_products=0#qt-news_science_products).
- United Nations DESA, 2019. *World Urbanization Prospects: The 2014 Revision*.
- Voogt, J., Oke, T., 2003. Thermal remote sensing of urban climates. Remote Sens. Environ. 86 (3), 370–384. [https://doi.org/10.1016/S0034-4257\(03\)00079-8](https://doi.org/10.1016/S0034-4257(03)00079-8).
- WMO, 2023. *Guidance on Measuring, Modelling and Monitoring the Canopy Layer Urban Heat Island (CL-UHI)*, vol. No. 1292. WMO.
- Wood, S., 2020. Package “mgcv”. <https://cran.r-project.org/web/packages/mgcv/mgcv.pdf>.
- Zhang, Y., Yu, C., Wang, L., 2017. Temperature exposure during pregnancy and birth outcomes: an updated systematic review of epidemiological evidence. Environ. Pollut. 225, 700–712.