## CLIMATOLOGY

# Crowdsourced air temperatures contrast satellite measures of the urban heat island and its mechanisms

Zander S. Venter[1]*, Tirthankar Chakraborty[2], Xuhui Lee[2]

The ubiquitous nature of satellite data has led to an explosion of studies on the surface urban heat island (SUHI). Relatively few have simultaneously used air temperature measurements to compare SUHI with the canopy UHI (CUHI), which is more relevant to public health. Using crowdsourced citizen weather stations (>50,000) and satellite data over Europe, we estimate the CUHI and SUHI intensity in 342 urban clusters during the 2019 heat wave. Satellites produce a sixfold overestimate of UHI relative to station measurements (mean SUHI 1.45°C; CUHI 0.26°C), with SUHI exceeding CUHI in 96% of cities during daytime and in 80% at night. Using empirical evidence, we confirm the control of aerodynamic roughness on UHI intensity, but find evaporative cooling to have a stronger overall impact during this time period. Our results support urban greening as an effective UHI mitigation strategy and caution against relying on satellite data for urban heat risk assessments.

## INTRODUCTION

The urban heat island (UHI) effect is the most well-known impact of urbanization on local climate (1) and can aggravate heat stress, increase cooling load, modify frequency of extreme rain events, and influence air pollution (2–5). While the founding studies on UHI were based on observed air temperature ($T_{air}$) within the urban canopy [the canopy UHI (CUHI)] (6, 7), the proliferation of satellite-based land surface temperature (LST) measurements has led to the definition of the surface UHI (SUHI) (8). The explosion of SUHI studies is partly because observing LST has distinct advantages over ground-based measurements. Measurements from a single satellite usually use the same sensor, have complete global coverage, and eliminate the issues with setting up representative sites, which has been a major criticism of the CUHI literature (9). Consequently, almost all large-scale observational estimates of UHI are based on satellite data (10–12). However, it is expected that given the dependence of the coupling between LST and $T_{air}$, on, among other things, vegetation and aerodynamic roughness, CUHI and SUHI are not identical quantities (13–17). In particular, for public health considerations, CUHI is much more relevant than SUHI, because heat stress is a function of $T_{air}$, not necessarily LST (17, 18).

A handful of studies have performed simultaneous measurements of CUHI and SUHI and suggest that they may have different diurnal and seasonal trends (3, 19, 20, 17, 21, 10). Because of the influence of urban areas on local meteorology, it is not standard practice to measure $T_{air}$ in cities. Meteorological stations near urban areas are usually at airports, which are not representative of the urban core. Unfortunately, because of the scarcity of urban weather stations, multicity comparisons of CUHI and SUHI are practically nonexistent.

The dearth of urban weather stations has also precluded our ability to derive empirical evidence for determinants of UHI intensity at regional scales. A multitude of studies have argued or implicitly assumed that the UHI is primarily the result of a reduction in evaporative cooling in cities due to the lack of vegetation (12, 19, 22, 23). A recent study challenged this generally accepted paradigm by combining land surface modeling with an offline temperature attribution

method for a subset of cities in North America (24). The study demonstrated that the difference in aerodynamic roughness between a city and its surrounding rural reference is a stronger determinant of the daytime SUHI than the difference in evapotranspiration, thereby explaining the trend in SUHI intensity across climate zones for a sample of cities in North America. The model used in this study has a coarse implementation of urban land cover (25, 26). In the few Earth system models with explicit urban representation, the urban land is considered to be a completely built-up subgrid tile in a larger grid, with the forcing held constant. In reality, urban areas may contain more than built-up structures, including green space, snowpack, and bare soils. Moreover, the atmospheric forcing may not necessarily be equal between the city and its rural reference, due to the differences in local-scale aerosol concentrations and cloud cover. Alternative urban climate models do account for intra-urban land cover variations and are valuable for scenario testing and forecasting; however, they are only validated against a handful of station measurements (27).

We aimed to use the recent proliferation of crowdsourced citizen weather stations (CWSs) (28) to address gaps in our knowledge of the relative magnitude and distribution of SUHI and CUHI and empirically test whether aerodynamic roughness is an important determinant of UHI intensity. We do this using CWS $T_{air}$ measurements from >50,000 private weather stations throughout Europe during the month of July 2019. The time period chosen coincides with the heat wave that exceeded long-term averages by 5°C for more than three consecutive days (fig. S1). This was comparable to the 2003 heat wave, which was estimated to have killed more than 70,000 people (29). We find that the mean CUHI is lower than the mean SUHI during daytime in 96% of urban clusters, when the adverse effects of heat stress would be maximum. To further examine the coupling between the CUHI and the SUHI, we use a random forest (RF) regression to determine the controlling factors of each. Last, we extend this methodology to a subset of cities with building height measurements to determine if roughness is an important factor for the spatial variability of UHI (fig. S2).

## RESULTS

### CUHI versus SUHI

We found that city-scale UHI was overestimated by 1.19 ± 0.02°C (mean ± SE, $n$ = 342) if LST was used instead of $T_{air}$ (mean SUHI

[1]Terrestrial Ecology Section, Norwegian Institute for Nature Research—NINA, 0349 Oslo, Norway. [2]School of the Environment, Yale University, New Haven, CT, USA.
*Corresponding author. Email: zander.venter@nina.no

1.45°C versus mean CUHI 0.26°C; Figs. 1 and 2A). For the cities considered during July 2019, this reflects a sixfold overestimation in the urban contribution to local temperature if satellite observations are used instead of ambient air temperature measurements. The upward bias was present during both day and night with an average SUHI and CUHI of 2.00 ± 0.05°C and 0.05 ± 0.03°C during the day, respectively, and an average SUHI and CUHI of 0.90 ± 0.02°C and 0.47 ± 0.03°C during the night, respectively. Note that CUHI may be underestimated by up to 0.75°C in some cases because of an upward bias in rural CWS measured relative to established meteorological station measurements (fig. S12). Nevertheless, even after correcting for the positive bias, SUHI still exceeds CUHI by 40% at minimum. Satellite-based SUHI estimates are higher than CUHI in 96% of the cities during daytime and 80% of cities at night. This difference is consistent over the July 2019 period and the associated heat wave event (fig. S3); however, at a finer temporal resolution, it is apparent that the difference is maximized during the day and minimized at night (Fig. 2B). The temporal variance in SUHI is not well correlated to that of CUHI particularly during the night (fig. S4). Predominant daily weather explains part of the temporal variance in CUHI-SUHI differences (fig. S5). Specifically, SUHI exceeds CUHI more on days with relatively low cloud cover, wind speeds, and high diurnal air temperature ranges. While the CUHI is positive in 81% of the cities, the magnitude during daytime is less than 0.05 in 52% cities. A further analysis of the individual station data confirms that this is partly due to the location of the CWS. Weather stations in vegetated parts of the city show a lower CUHI (0.09 ± 0.02°C) than those situated on predominantly built-up surfaces (0.3 ± 0.02°C). Even when only the stations over built-up

surfaces are considered, satellite overestimates the UHI intensity by five times.

## UHI attribution

Using a combination of land cover, terrain, and city morphology variables in an RF regression model, we were able to explain more of the variance in local-scale SUHI [$R^2 = 0.49$, root mean square error (RMSE) = 1.22°C] than CUHI ($R^2 = 0.23$, RMSE = 1.05°C; table S1). Here, local-scale UHI refers to the urban-rural temperature differentials for each $T_{air}$ station or LST pixel. Land cover variables included evapotranspiration [using the normalized difference vegetation index (NDVI) as a proxy], surface albedo, and the fraction of impervious surface area (ISA). Terrain variables included elevation above sea level, distance to coastline, and a number of indices describing the diversity [topographic diversity index (TDIV)], relative position [topographic position index (TPI)], and solar heat load [continuous heat-insolation load index (CHILI)] of the surrounding terrain (see Materials and Methods for details). City morphology was characterized by the standard deviation in the heights of surrounding buildings. Multiple linear regression (MLR) models with the same explanatory data corroborate this finding with more variance explained in local-scale SUHI ($R^2 = 0.49$) than CUHI ($R^2 = 0.12$; table S2).

The urban-rural differentials in surface evapotranspiration and elevation contributed the most to explaining the variance in local-scale SUHI and CUHI, respectively, for all 342 clusters (fig. S6). When considering the subset of 30 urban clusters with building height data (fig. S15), evapotranspiration was the second most important variable for both CUHI and SUHI, whereas aerodynamic
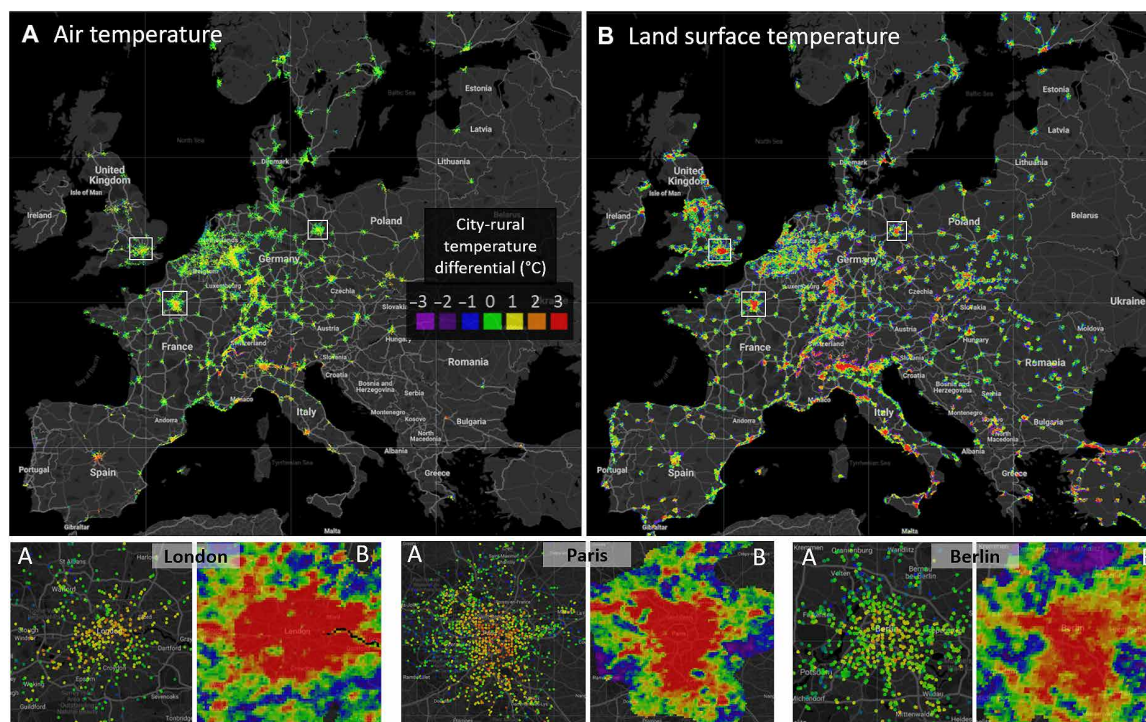


**Fig. 1. Distribution of city-rural temperature differentials for July 2019 in Europe.** (**A**) Private weather station ($n = 59,810$) air temperature and (**B**) satellite-derived LST are relativized to mean rural temperatures for 342 urban clusters. Temperature differentials give an indication of UHI intensity illustrated by the zoomed-extent windows for London, Paris, and Berlin.
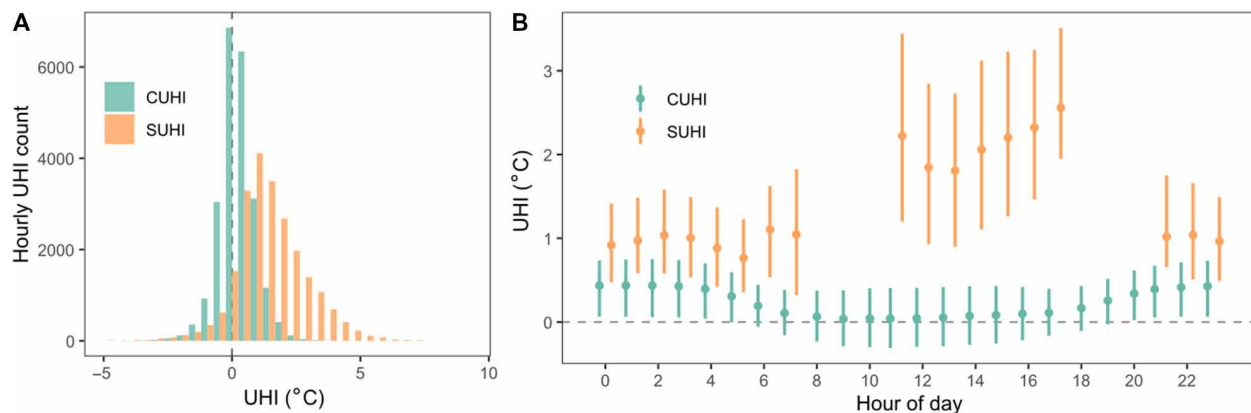
**Fig. 2. SUHI and CUHI intensities.** (**A**) Frequency distribution of all available hourly (local time) UHI intensities for European urban clusters ($n = 342$) during July 2019. (**B**) Hourly mean (point) and interquartile range (vertical line) for UHI intensities over each hour of the day.

roughness was less important in explaining variations in SUHI than CUHI (Fig. 3). For the cities considered, results from both RF (Fig. 3) and MLR models (table S2) reveal that evapotranspiration is a stronger determinant of the variability in CUHI than roughness during the day, although this difference fades at night (Fig. 3). Roughness is an important determinant of CUHI at night, where increases in city-rural roughness differentials contribute to larger CUHI intensities (Fig. 4). The partial dependence of local-scale UHI on evapotranspiration reveals that when evapotranspiration is lower inside relative to outside cities, both SUHI and CUHI have greater magnitudes, whereas city areas with higher evapotranspiration than rural areas result in negative SUHI and CUHI (Fig. 4). These partial dependencies of CUHI on roughness and evapotranspiration are consistent over multiple spatial scales (fig. S7).

Apart from considering the variation in local-scale UHI, we also explored the city-scale UHI by aggregating station measures up to city means. Using MLR of intercity variations in UHI resulted in similar patterns to those produced from the analysis of local-scale UHI (fig. S8). Mean city nighttime CUHI and SUHI were significantly ($P < 0.05$) lower in cities with higher urban-rural evapotranspiration differentials, lending support to the local-scale results in Figs. 3 and 4. However, city-rural aerodynamic roughness differentials had no significant ($P > 0.05$) effect on mean city CUHI and SUHI (fig. S8).

## DISCUSSION

While the impact of UHI on heat stress is only one of its many impacts, it has been the focus of a large number of studies, especially those looking at interactions between larger-scale phenomena like regional heat stress or global warming and urbanization (30–32). Studies on both individual (17, 33) and multiple cities (16) have shown that the SUHI is usually higher than the CUHI, especially during the daytime, which we confirm by scaling this comparison up to over 342 European cities during a heat wave period. The notably higher magnitude of SUHI has important implications for understanding and attributing the source of heat stress in urban areas. Because air temperature is the major constituent of heat stress, urban-rural air temperature differentials can help isolate the contribution of urbanization to local-scale heat exposure. Epidemiological studies have relied on near-surface air (ambient) temperature measured at
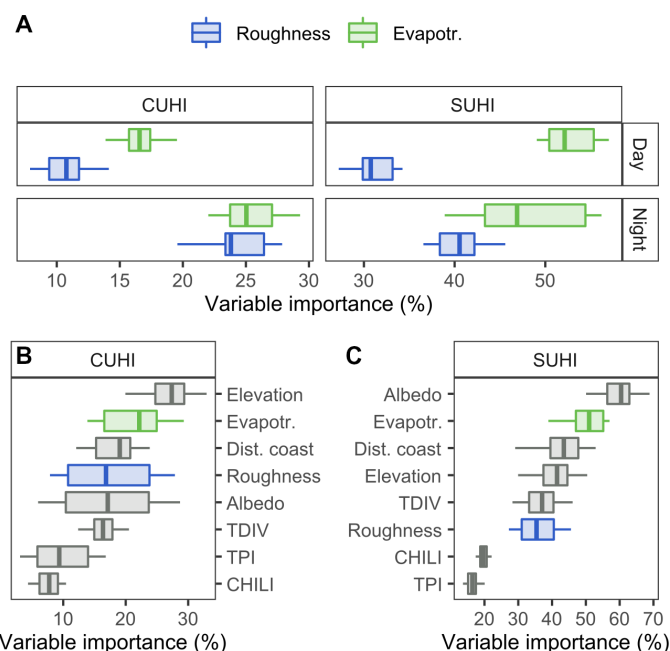


**Fig. 3. Determinants of SUHI and CUHI intensities.** (**A**) Predictor variable importance scores are plotted for RF models explaining the variance in local-scale ($n = 5703$ stations; 47,652 pixels) SUHI and CUHI over a subset of urban clusters (30) with building height data. Models are stratified into nighttime and daytime mean UHI for roughness and evapotranspiration. Mean importance values are plotted for the total predictor variable suite in (**B**) and (**C**).

fixed sites to quantify thermal exposure (34). However, spatially explicit LST data from satellites are gaining popularity and are being increasingly applied, either directly or for predicting air temperature, to public health risk studies (35, 36). In doing so, researchers run the risk of overestimating the magnitude of temperature exposure as well as the contribution of urbanization to the local temperature. This is particularly concerning given that the upward bias of satellite-based measures of UHI was greatest during conditions that are most relevant to heat risk epidemiology (daytime, low cloud cover, high diurnal temperature ranges, and low wind speeds). Proper
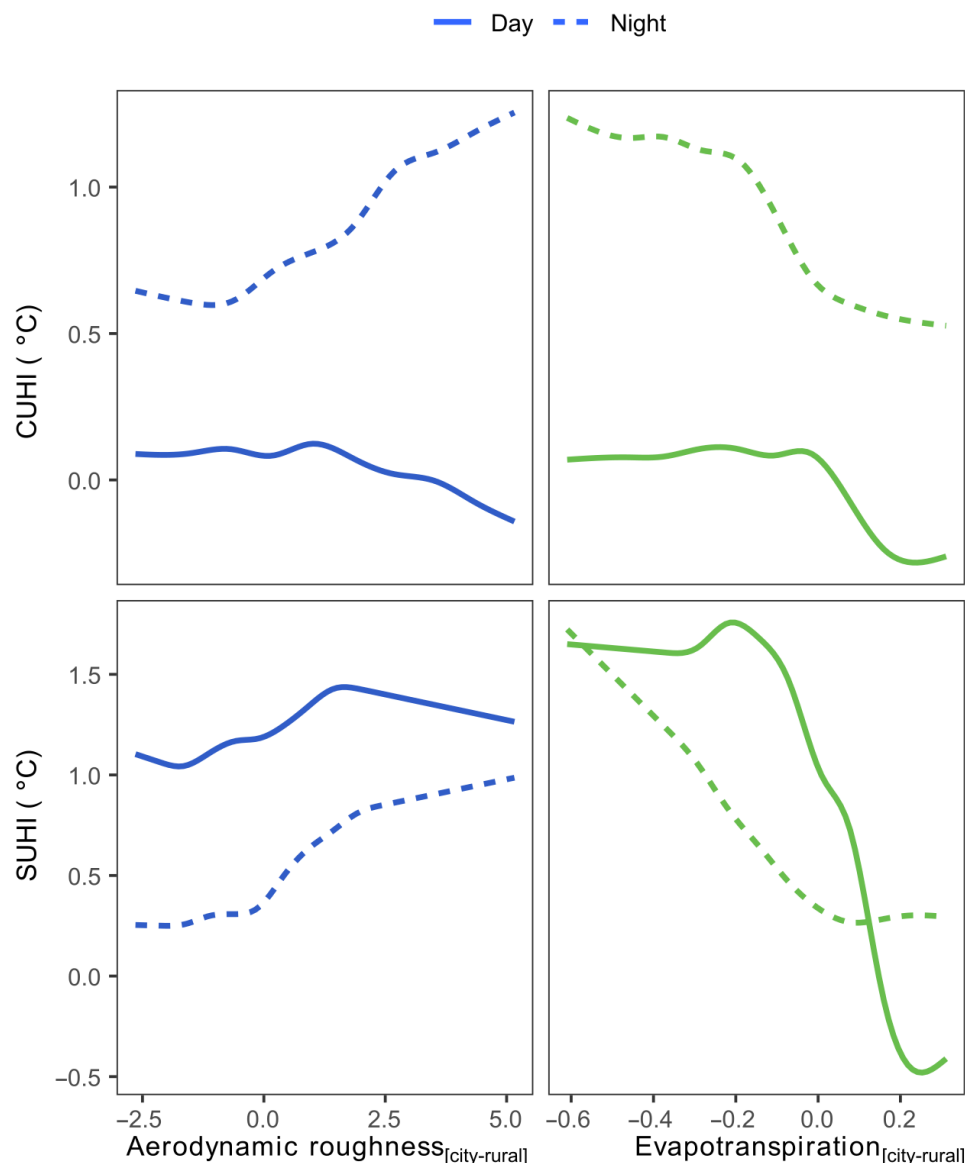
**Fig. 4. Partial dependence of UHI on roughness and evapotranspiration.** Lines represent the smoothed effect of roughness and evapotranspiration on local-scale ($n$ = 5703 stations; 47,652 pixels) SUHI and CUHI intensity after controlling for the effect of all other explanatory variables. RF models were used for daytime and nighttime daily UHI separately.

quantification of the public health implications of UHI requires dense monitoring networks in cities, and in the absence of these monitoring networks, crowdsourced air temperature measurements can be a reasonable alternative, as we demonstrate in the present study.

We also find that evapotranspiration and roughness were important determinants of both SUHI and CUHI intensity during this heat wave period (fig. S6). In particular, for the goal of reducing the effect of UHI on human health, urban green spaces remain an important mitigation option given that evapotranspiration was the dominant driver of UHI during the day (Fig. 3). However, urban greening may lead to elevated humidity levels, which is another contributor to human heat stress and can thereby counteract the benefit gained from evaporative cooling (37). In cities where urban greening is constrained by water limitations (e.g., in arid environments), an alternative may involve changing building structure, height, and spacing

so as to reduce intracity aerodynamic roughness. This heat mitigation strategy is equally supported by our results, showing that CUHI is exacerbated by greater city-rural roughness differentials, particularly at night. This alternative, while challenging as it requires fundamental changes to the city built environment morphology (24), can be considered for future urban planning.

Although Zhao *et al.* (24) showed that the aerodynamic roughness of the city is an important determinant of the geographic distribution of UHI intensities in North America, it has been difficult to empirically test this hypothesis. For one, aerodynamic roughness is hard to determine from observations, especially for heterogeneous terrain like cities. Moreover, although it is possible to theoretically separate the effect of vegetation and roughness differentials on the UHI intensity, their effects are hard to isolate in practice, because the presence of vegetation also changes aerodynamic roughness and

cities are more heterogeneous than their simplified representations in climate models. Here, we attempt to do so statistically using proxies for evapotranspiration (NDVI) and aerodynamic roughness (standard deviation of building height; fig. S2). We find that, for the cities considered, evapotranspiration is a stronger determinant of the spatial variability in both SUHI and CUHI during the day. However, note that roughness still remains an important determinant of the magnitude of the CUHI magnitude, particularly at night. Consistent with the results of Zhao *et al.* (*24*), during daytime, as the urban-rural differential in aerodynamic roughness increases, the CUHI decreases, implying higher convective dissipation of heat from the urban area. For the urban-rural differential in NDVI, negative values imply less evapotranspiration in urban areas and, thus, higher UHI intensity, while positive values (possibly due to the presence of urban green space) lead to a reduction in local-scale UHI intensity due to local-scale evaporative cooling. For aerodynamic roughness, this pattern reverses at night for both CUHI and SUHI, as the rougher urban areas bring down warmer air from aloft compared to smoother rural landscape.

The overall dominance of evapotranspiration over roughness for determining variability of UHI intensity in our study is consistent with a recent study on a smaller subset of cities in Europe (*38*). Nevertheless, the generalizability of our results is limited to the temporal (summer) and geographical (Europe) extent of our analysis. The RF modeling framework used is by definition limited to the state space dimensions defined by the input dataset. Therefore, making inference to different climates and making predictions of how UHI will respond under climate change scenarios would require broadening our sampling over space and time. Despite this, our statistical modeling framework, driven by empirical data, is a unique and important complement to process-based numerical models and analytical attribution frameworks. Attribution frameworks, as used by Zhao *et al.* (*24*) and Li *et al.* (*39*), have the added advantage of being able to separate not only order of importance of variables but also the degree of relative importance at large scales. However, several such frameworks (*40*, *41*) do not necessarily agree with each other. On the other hand, process-based numerical models are restricted by both the quality of input data and the large potential variability in model parameters in urban areas (*42*, *43*). Our statistical approach relies on a priori assumptions about mechanistic links between predictor and response data, with the linear models assuming linear dependence between response and predictor, and the RFs also including nonlinear interactions. However, these results and methods remain useful tools for both hypothesis testing and hypothesis generation.

Crowdsourced weather data are now widely available (*44*), with dense CWS networks in major urban centers globally (https://weathermap.netatmo.com). Although data quality is a concern, it does not preclude the use of these data in urban climatology studies. Data quality control procedures used in the present study (fig. S9) are open source (*28*, *45*) and significantly reduce risk of including statistically implausible $T_{air}$ due to misplacement of sensors, solar exposition (radiative errors), inconsistent metadata, and device malfunctions. These quality control procedures are based on outlier detection and spatial correlation checks within the crowdsourced dataset and the option to cross-calibrate with reference meteorological stations. However, we did find evidence that quality control procedures do not completely remove solar contamination biases (fig. S12), particularly for detecting UHI. Netatmo CWS sensors have a large thermal mass compared to standard temperature sensors used at

weather stations, and many are likely placed in positions exposed to solar radiation, which results in an upward temperature bias. For studying urban climate, our results mirror the issues raised by Stewart (*9*) on the general UHI literature, with specifics of individual sensor placement possibly leading to a large degree of control on the CUHI variability (fig. S6). Ideally, station manufacturers would require customers to record more detailed metadata on station placement so that these data can be more readily screened for quality issues by researchers. At our end, we have tried to account for these variabilities through extensive quality assurance (fig S9), evaluating the station data against existing products both in bulk and under varying sky conditions (figs. S11 and S12), and by quantifying how the calculated CUHI intensity varies with both intra-urban land cover variability and weather conditions (figs. S5 and S14). Despite quality issues, dense $T_{air}$ station networks offer an unprecedented advantage over satellite imagery to derive hourly UHI estimates, irrespective of cloud cover, and provide temperature measures that are more relevant to human thermal comfort than satellite-derived LST. The extent to which crowdsourced weather stations can provide more reliable insights relative to established meteorological stations requires further investigation. Although denser networks of Netatmo CWS data are available within cities, established meteorological stations, which are rarely present in cities, have more stringent quality control to prioritize optimum station location and accurate data collection.

On the basis of our findings, we conclude that incorporating satellite-based measures (LST) of UHI in epidemiological research should be done with great caution, particularly on clear-sky days when the discrepancy with CUHI is at its maximum. Similarly, ground-based measures of UHI need to be adopted in epidemiological research with caution because the public health impacts of thermal exposure are also linked to humidity, wind speed, and behavioral differences and comorbidities that contribute to health outcomes. Assuming that epidemiologists do want to calculate CUHI, the relationship between SUHI and CUHI is not easily explained by a linear regression model (fig. S4), and therefore, it is difficult to come up with regression coefficients to convert SUHI to CUHI for incorporation in epidemiological studies. We suggest that crowdsourced and established meteorological weather stations should be used in parallel to calculate CUHI directly, after adequate cross-calibration and quality assurance of CWS. In doing so, caution should be taken when calculating UHI from stations in very different topographic contexts. The large importance scores for elevation and topographic diversity on CUHI suggest that the placement of the weather stations has a strong impact on the local-scale CUHI, as calculated in this study. Therefore, they can confound the UHI effect, which is by definition the effect of urban land cover and not elevation or topography on ambient temperature.

In contrasting SUHI and CUHI, we are not implying that one is better or worse than the other. Satellites and weather stations measure distinct and complementary components of the UHI effect, and SUHI remains useful for applications like weather prediction, because LST provides lower boundary conditions to models. Our observation that SUHI and CUHI differs serves as a point of departure for further discussion and hypothesis testing. Last, our results show that the UHI attribution is slightly different for SUHI and CUHI. In the context of public health risk and heat stress, evapotranspiration appears to be a stronger driver of CUHI during the day, although roughness is also important at night. Together, our results support the use of green space and building structures that reduce canopy

roughness as possible UHI mitigation options, particularly for ameliorating the impacts of future heat waves on the CUHI and, thus, public health.

## MATERIALS AND METHODS

### Air temperature data

The recent proliferation of CWS in urban areas allows dense meteorological observations to be crowdsourced (28, 45, 46). We downloaded hourly air temperature ($T_{air}$) data during July 2019 for all available CWS provided by Netatmo (https://netatmo.com/) in Europe. CWS are known to have quality issues caused by sensor misplacement, solar exposure, inconsistent metadata, and device or internet malfunctions (45). We followed the quality control method developed by (28) to clean the raw $T_{air}$ data for statistically implausible readings using the "Crowd-QC" package in R (47). The procedure flags stations with equivalent latitude and longitude coordinates and then applies a modified $z$ score approach for the detection of statistical outliers from the hourly $T_{air}$ distribution (fig. S9). The final step excludes stations with $T_{air}$ measures that, when correlated against the spatial median of hourly $T_{air}$ over a month, produce a Pearson's correlation coefficient of less than 0.9. The data cleaning procedure reduced the number of available stations from 113,215 to 95,084 stations.

The remaining Netatmo stations included some data gaps, which can be caused by drops in wifi signal or battery failures (fig. S10A). Therefore, we also excluded stations that had more than 20% of their time series missing during July 2019. We performed a sensitivity analysis to show that stations with 80% of data will give monthly means that are within 0.1°C of the actual mean (fig. S10B). To do this, we selected 500 random stations with full data coverage and then iteratively removed a random subset of $T_{air}$ measurements at proportions ranging from 0 to 0.9. We calculated monthly $T_{air}$ means from the subsets and measured the absolute deviation from the actual monthly means. After excluding stations with <80% data coverage, we were left with 75,293 stations.

To validate the quality-controlled $T_{air}$ data, we used data from the Global Historical Climatology Network Daily (GHCND) dataset and the ERA5 fifth generation European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalysis of the global climate (48). We first calculated the mean value for daily $T_{air}$ maximum readings for each GHCND station over July 2019 within Europe. After doing the same for the Netatmo stations, we performed a spatial join with the GHCND stations by averaging Netatmo $T_{air}$ values from stations within a 2-km buffer of each GHCND (fig. S11C). Using the ERA5 gridded dataset, we calculated the July average for the daily 2-m maximum $T_{air}$ and performed a spatial join with the Netatmo data by averaging Netatmo $T_{air}$ values within each ERA5 grid cell. Given that the land module of ERA5 does not consider urban land cover, we only included Netatmo stations that were defined as rural according to the Global Human Settlement Layer (GHSL) produced by the European Joint Research Centre (49). The linear regression gives an RMSE of 0.95° and 1.31°C for the GHCND and ERA5 data, respectively (fig. S11). The Netatmo data appear to have a slight positive bias in $T_{air}$Max. Even within a 2 km buffer or 20 km grid cell, there is likely to be large variation in $T_{air}$ that is sensitive to microscale topographic effects, which, depending on the distribution of Netatmo stations, will skew the Netatmo readings. In addition, Netatmo stations do not always have radiation shields, which could also explain the

positive bias. We tested this by assessing how the bias changes on cloud cover days, when one would expect the effect of radiation to be minimal. Cloud cover does reduce the positive station bias; however, this bias was systematic and displayed little difference between urban and rural stations. Thus, this bias does not compromise the calculation of UHI (fig. S12).

### LST data

All raster and satellite data were extracted and processed within the Google Earth Engine cloud-computing platform (50). We collected daily LST data over Europe during July 2019 from the moderate-resolution imaging spectroradiometer (MODIS) satellites TERRA (MOD11A1) and AQUA (MYD11A1). LST data are preprocessed to clear-sky pixels with average LST error of less than or equal to 3 and are made available at 1 km × 1 km resolution. MODIS satellites collect data daily at approximately 1:30, 10:30, 13:30, and 22:30 local time and thus provide information on daytime and nighttime temperature. To obtain hourly LST values, we first converted the MODIS overpass times from local solar time to universal time coordinated (UTC) local time according to the MODIS user guide by subtracting the MODIS grid's longitude in degrees divided by 15. We then accounted for the time zone of each urban cluster in combination with daylight savings time to offset the UTC for each city. This resulted in LST readings with hour-specific time stamps that could be synced with the hourly Netatmo dataset. In the components of our analysis that aggregated data to daytime and nighttime UHI, we relied on the definitions of "day" and "night" implicit in the MOD11A1 and MYD11A1 raster datasets. These were LST values defined by the raster bands "LST_Day_1km" and "LST_Night_1km," respectively.

### Calculating UHIs

To calculate UHI magnitudes, we followed a rural buffer zone (also known as ring-based buffer) approach commonly used in UHI studies (20, 36, 51). In our analysis, we calculated both city-scale UHI and local-scale UHI. City-scale UHI is defined as the temperature difference between mean city and mean rural weather station $T_{air}$ or LST. Local-scale UHI is defined at the station/pixel level as the differential between each city weather station or satellite pixel and the rural mean. It thus provides a measure of intracity UHI variation. To define urban clusters, we used the GHSL "degree of urbanization" raster product, which labels 1 km × 1 km grid cells as uninhabited, rural, and low- and high-density urban (49). These classes are derived from both the GHSL built-up areas and GHSL population grid datasets. We isolated the low- and high-density urban pixels and used a connectivity function to define urban clusters using the same principles outlined in the City Clustering Algorithm (52). This involves iteratively aggregating neighboring urban pixels into a connected object defining the bounds of a unique urban cluster (fig. S13). We excluded urban clusters that were smaller than the 50th percentile of the urban cluster size class distribution (40 km²), because they have very few Netatmo stations. Even after this filtering, the sample size was substantial at 931 clusters.

To define the rural buffer zones, we buffered each urban cluster by 10 km by applying focal mode smoothing function so that there is no overlap in rural zones of neighboring clusters (fig. S13B). Although there are many ways of defining the rural reference zone (22), we adopted a 10-km buffer zone based on a precedent set by other studies and findings from Clinton and Gong (11) that UHI intensity does not change dramatically with varying buffer zone sizes.

Where there are two adjacent urban areas, the focal mode function will find the buffer zone border between the two that is equidistant to their original borders. According to best practice for UHI calculation (9), before calculating city-rural temperature differentials, we accounted for terrain effects by only including stations and LST pixels within 50 m of the mean urban cluster elevation. We used the Global Multi-resolution Terrain Elevation Data (GMTED) 2010 (53) at 7 arc sec resolution to perform this correction. We also used the MODIS land cover dataset (54) (MCD12Q1) to mask any pixels covered by water and any rural zone pixels containing urban land cover. To obtain more local-scale information on Netatmo weather station locations, we calculated the mode land cover class from the Copernicus Global Land Cover dataset (55) within a 300-m buffer of each station (fig. S14). LST-derived SUHI and Netatmo station–derived CUHI were calculated for daytime and nighttime temperatures for hours coincident with MODIS overpass. Netatmo weather stations intersecting the urban clusters plus rural buffers of interest amounted to 59,810. UHI was calculated for clusters with at least five Netatmo stations in both urban and rural zones to reduce potential bias from station error. This resulted in 342 urban clusters available for UHI analysis.

### RF modeling

Recent advances in machine learning and statistics have allowed researchers to extend the scope of application beyond prediction and classification alone. Regression tree algorithms such as the RF (56) are increasingly being used as tools for exploratory analysis and causal inference (57–59). Supervised RF models can give insight into the causal relationships between explanatory/predictor variables and the response variable in question. We use RF modeling as a tool to explore the marginal effects of aerodynamic roughness and evapotranspiration on local-scale UHI (station- or pixel-specific city-rural differentials).

City-rural temperature differentials were calculated from July 2019 averages of daily daytime and nighttime mean temperature time series for each Netatmo station and LST pixel. The temperature value for each Netatmo station or LST pixel was subtracted from the rural cluster average. This resulted in four response variable datasets stratified by temperature source (LST and $T_{air}$) and type (day, night). Separate RF models were built for each response variable.

We collected a range of predictor variables to explain the variance in city-rural temperature differentials (table S3) based on knowledge derived from previous reviews of important determinants of UHI (20, 36). The primary variables of interest to test our hypothesis were aerodynamic roughness and evapotranspiration (fig. S2). To include roughness in our model, we limited the analysis to a subset of 30 urban clusters (fig. S15), where building height data are made available through the Copernicus urban atlas for 2012 (https://land.copernicus.eu/local/urban-atlas). We calculated three morphological parameters that characterize aerodynamic roughness, which are often used in urban canopy models (60, 61). These included mean building height, building fraction (horizontal building surface divided by total surface), and the standard deviation of building height. We used the normalized difference vegetation index (NDVI) as a proxy for evapotranspiration by calculating the median NDVI from all available cloud-masked Landsat 8 Operational Land Imager surface reflectance scenes over Europe during July 2019. Global datasets for potential and actual evapotranspiration do not have coverage in urban areas (62). Given that NDVI is the primary input into the

evapotranspiration algorithms and is highly correlated with potential evapotranspiration (63–65), we argue that it is an adequate proxy to be used in our study.

Apart from the two primary predictor variables, we collected additional land cover and terrain morphology predictors (table S3). We used the Copernicus Imperviousness Density 2015 layer at 100-m resolution (66) to extract ISA. Distance to ocean was calculated using the distance to the nearest coast layer produced by NASA (https://oceancolor.gsfc.nasa.gov/docs/distfromcoast/). Median black sky albedo was calculated using the MODIS MCD43A3 product (67) for Europe over July 2019 for all available bands, and this was reduced into a single raster by averaging across bands. Elevation above sea level and terrain slope was calculated from the GMTED dataset at 30-m resolution. To characterize solar insolation and terrain relief, we also extracted the CHILI, TDIV, and TPI at 90-m resolution (68). These variables have been used as surrogate variables for terrain variability that may induce a variety of temperature and moisture conditions. The TDIV variable is calculated with the TPI and CHILI indices within a multiscale diversity framework to define the diversity of physiographic units within a moving window. See "multi-scale diversity" in (68).

To merge the predictor data with the LST data, we resampled all predictor variables up to the 1 km × 1 km resolution of LST data by calculating the mean values within each LST pixel. For the $T_{air}$ data, we extracted the mean predictor values within a range of circular buffer zones around each Netatmo weather station location. Buffer zone radii used included 30, 50, 100, 200, 400, 800, 1600, 3200, and 6400 m and were chosen based on precedents set by previous studies modeling urban $T_{air}$ using statistical approaches (17, 69). Separate RF models were built for each buffer size, and model results were averaged across all models.

We assessed model performance by withholding 30% of the dataset from the model training stage and thereafter testing model predictions against it. By regressing observed temperature on predicted temperature, we derived the RMSE and adjusted $R^2$ as measures of model accuracy and fit, respectively. This process was repeated 10 times, in a bootstrapping procedure, for each model to smooth over the random variation caused by the splitting of training and validation datasets.

The RF algorithm measures predictor variable importance by quantifying the increase in prediction errors when a predictor is permuted in the validation data. We also derived partial dependence curves describing the marginal effect of each predictor variable on temperature differentials when holding all other predictors constant. We collected variable importance scores and partial dependence curves for each iteration of the bootstrapping procedure. By using RF modeling for exploratory analysis of predictor importance and marginal effects on temperature differentials, we needed to ensure that predictor variables were not collinear, which would confound results. We therefore first tested for multicollinearity in the predictor set by calculating variance inflation factor (VIF) values for each predictor and only including those with a VIF of less than 5 (70). The VIF score is based on a cross-correlation analysis where pairwise correlations between each combination of predictor variables are performed. We found that building fraction, building height, and the standard deviation of building height were collinear so we dropped building fraction and building height from the predictor set. Therefore, building height standard deviation was the selected proxy for aerodynamic roughness.

As a further test of the effect of potential multicollinearity and the sensitivity of the model to the predictor variable set, we created a null model where only roughness and evapotranspiration were included as predictors and then iteratively added other predictors to the model to explore the effect on variable importance. We found that the difference in importance between roughness and evapotranspiration was relatively invariant with the addition of potentially collinear predictors (fig. S16). Overall variable importance scores were derived from the average of this iterative modeling procedure.

## Linear regression modeling

To supplement and corroborate the RF model estimates of variable importance for roughness and evapotranspiration as determinants of local-scale UHI, we also produced MLR models. The MLR analysis was limited to the subset of urban clusters ($n$ = 30) with building height data because the point was to corroborate the RF model variable importance scores, informing the core hypothesis about UHI attribution. Separate MLR models were made for daytime and nighttime LST and $T_{air}$ datasets. Temperature was regressed on all explanatory variables, and we then used the "relaimpo" package in R to calculate relative importance scores for roughness and evapotranspiration (71) using the "lmg" method (72). All variables were included in the model; however, the relative importance of roughness versus evapotranspiration was determined by decomposing the residual $R^2$ using the "always" parameter to define the variables that are held constant. In addition to examining local-scale UHI, we also explored city-scale UHI by examining the contribution of city mean roughness and evapotranspiration differentials to explaining the variation in UHI intensity between cities. We therefore performed a linear regression of daytime and nighttime CUHI and SUHI on mean city-rural roughness and evapotranspiration differentials.

Last, we also used linear regression to assess how daily weather influences the temporal variation in local-scale SUHI-CUHI differences. We regressed SUHI-CUHI differences on daily fractional cloud cover, diurnal temperature range, and wind speed. Daily cloud cover probabilities were derived from the PATMOS-x dataset (73), temperature data from the Netatmo stations, and wind speed from the ERA5 daily reanalysis dataset.

## SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/22/eabb9569/DC1

## REFERENCES AND NOTES

1. T. R. Oke, G. Mills, A. Christen, J. A. Voogt, Urban heat island, in *Urban Climate* (Cambridge Univ. Press, 2017).
2. J.-Y. Han, J.-J. Baik, H. Lee, Urban impacts on precipitation. *Asia Pac. J. Atmos. Sci.* **50**, 17–30 (2014).
3. A. M. Rizwan, L. Y. C. Dennis, C. Liu, A review on the generation, determination and mitigation of Urban Heat Island. *J. Environ. Sci.* **20**, 120–128 (2008).
4. M. Santamouris, On the energy impact of urban heat island and global warming on buildings. *Energ. Buildings* **82**, 100–113 (2014).
5. C. Sarrat, A. Lemonsu, V. Masson, D. Guedalia, Impact of urban heat island on regional atmospheric pollution. *Atmos. Environ.* **40**, 1743–1758 (2006).
6. L. Howard, *The Climate of London: Deduced from Meteorological Observations, Made at Different Places in the Neighbourhood of the Metropolis, Volume 1* (W. Phillips, George Yard, Lombard Street, sold also by J. and A. Arch, Cornhill; Baldwin, Cradock, and Joy, and W. Bent, Paternoster Row; and J. Hatchard, Picadilly, 1818).
7. H. E. Landsberg, Atmospheric changes in a growing community (the Columbia, Maryland experience). *Urban Ecol.* **4**, 53–81 (1979).
8. K. P. Gallo, J. D. Tarpley, A. L. McNab, T. R. Karl, Assessment of urban heat islands: A satellite perspective. *Atmos. Res.* **37**, 37–43 (1995).
9. I. D. Stewart, A systematic review and scientific critique of methodology in modern urban heat island literature. *Int. J. Climatol.* **31**, 200–217 (2011).
10. T. Chakraborty, C. Sarangi, S. N. Tripathi, Understanding diurnality and inter-seasonality of a sub-tropical urban heat island. *Bound.-Lay. Meteorol.* **163**, 287–309 (2017).
11. N. Clinton, P. Gong, MODIS detected surface urban heat islands and sinks: Global locations and controls. *Remote Sens. Environ.* **134**, 294–304 (2013).
12. S. Peng, S. Piao, P. Ciais, P. Friedlingstein, C. Ottle, F.-M. Bréon, H. Nan, L. Zhou, R. B. Myneni, Surface urban heat island across 419 global big cities. *Environ. Sci. Technol.* **46**, 696–703 (2012).
13. M. Jin, R. E. Dickinson, Land surface skin temperature climatology: Benefitting from the strengths of satellite observations. *Environ. Res. Lett.* **5**, 044004 (2010).
14. J. H. Lawrimore, M. J. Menne, B. E. Gleason, C. N. Williams, D. B. Wuertz, R. S. Vose, J. Rennie, An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *J. Geophys. Res. Atmos.* **116**, D19121 (2011).
15. D. J. Mildrexler, M. Zhao, S. W, A global comparison between station air temperatures and MODIS land surface temperatures reveals the cooling role of forests. *J. Geophys. Res. Biogeo.* **116**, G03025 (2011).
16. P. Zhang, L. Bounoua, M. L. Imhoff, R. E. Wolfe, K. Thome, Comparison of MODIS land surface temperature and air temperature over the continental USA meteorological stations. *Can. J. Remote Sens.* **40**, 110–122 (2014).
17. H. C. Ho, A. Knudby, Y. Xu, M. Hodul, M. Aminipouri, A comparison of urban heat islands mapped using skin temperature, air temperature, and apparent temperature (Humidex), for the greater Vancouver area. *Sci. Total Environ.* **544**, 929–938 (2016).
18. G. B. Anderson, M. L. Bell, R. D. Peng, Methods to calculate the heat index as an exposure metric in environmental health research. *Environ. Health Perspect.* **121**, 1111–1119 (2013).
19. Y. Y. Cui, B. de Foy, Seasonal variations of the urban heat island at the surface and the near-surface and reductions due to urban vegetation in Mexico City. *J. Appl. Meteorol. Climatol.* **51**, 855–868 (2012).
20. P. E. Phelan, K. Kaloush, M. Miner, J. Golden, B. Phelan, H. Silva, R. A. Taylor, Urban heat island: Mechanisms, implications, and possible remedies. *Annu. Rev. Environ. Resour.* **40**, 285–307 (2015).
21. K. Wang, S. Jiang, J. Wang, C. Zhou, X. Wang, X. Lee, Comparing the diurnal and seasonal variabilities of atmospheric and surface urban heat islands based on the Beijing urban meteorological network. *J. Geophys. Res. Atmos.* **122**, 2131–2154 (2017).
22. T. Chakraborty, X. Lee, A simplified urban-extent algorithm to characterize surface urban heat islands on a global scale and examine vegetation control on their spatiotemporal variability. *Int. J. Appl. Earth Obs. Geoinf.* **74**, 269–280 (2019).
23. Z. S. Venter, N. H. Krog, D. N. Barton, Linking green infrastructure to urban heat and human health risk mitigation in Oslo, Norway. *Sci. Total Environ.* **709**, 136193 (2020).
24. L. Zhao, X. Lee, R. B. Smith, K. Oleson, Strong contributions of local background climate to urban heat islands. *Nature* **511**, 216–219 (2014).
25. K. W. Oleson, G. B. Bonan, J. Feddema, M. Vertenstein, E. Kluzek, Technical description of an urban parameterization for the Community Land Model (CLMU) (No. NCAR/TN-480+STR, University Corporation for Atmospheric Research, 2010).
26. D. Li, S. Malyshev, E. Shevliakova, Exploring historical and future urban climate in the Earth System Modeling framework: 1. Model development and evaluation. *J. Adv. Model. Earth Syst.* **8**, 917–935 (2016).
27. K. De Ridder, D. Lauwaet, B. Maiheu, *UrbClim*—A fast urban boundary layer climate model. *Urban Clim.* **12**, 21–48 (2015).
28. A. Napoly, T. Grassmann, F. Meier, D. Fenner, Development and application of a statistically-based quality control for crowdsourced air temperature data. *Front. Earth Sci.* **6**, 118 (2018).
29. J.-M. Robine, S. L. K. Cheung, S. Le Roy, H. Van Oyen, C. Griffiths, J.-P. Michel, F. R. Herrmann, Death toll exceeded 70,000 in Europe during the summer of 2003. *C. R. Biol.* **331**, 171–178 (2008).
30. K. Huang, X. Li, X. Liu, K. C. Seto, Projecting global urban land expansion and heat island intensification through 2050. *Environ. Res. Lett.* **14**, 114037 (2019).
31. S. Jiang, X. Lee, J. Wang, K. Wang, Amplified urban heat islands during heat wave periods. *J. Geophys. Res. Atmos.* **124**, 7797–7812 (2019).
32. L. Zhao, M. Oppenheimer, Q. Zhu, J. W. Baldwin, K. L. Ebi, E. Bou-Zeid, K. Guan, X. Liu, Interactions between urban heat islands and heat waves. *Environ. Res. Lett.* **13**, 034003 (2018).
33. S. Bonafoni, R. Anniballe, M. Pichierri, in *2015 Joint Urban Remote Sensing Event (JURSE)* (IEEE, 2015), pp. 1–4.
34. B. Armstrong, M. L. Bell, M. de Sousa Zanotti Stagliorio Coelho, Y.-L. L. Guo, Y. Guo, P. Goodman, M. Hashizume, Y. Honda, H. Kim, E. Lavigne, P. Michelozzi, P. H. N. Saldiva, J. Schwartz, M. Scortichini, F. Sera, A. Tobias, S. Tong, C.-F. Wu, A. Zanobetti, A. Zeka, A. Gasparrini, Longer-term impact of high and low temperature on mortality: An international study to clarify length of mortality displacement. *Environ. Health Perspect.* **125**, 107009 (2017).

35. L. Shi, I. Kloog, A. Zanobetti, P. Liu, J. D. Schwartz, Impacts of temperature and its variability on mortality in New England. *Nat. Clim. Chang.* **5**, 988–991 (2015).

36. D. Zhou, J. Xiao, S. Bonafoni, C. Berger, K. Deilami, Y. Zhou, S. Frolking, R. Yao, Z. Qiao, J. A. Sobrino, Satellite remote sensing of surface urban heat islands: Progress, challenges, and perspectives. *Remote Sens.* **11**, 48 (2018).

37. A. L. Hass, K. N. Ellis, L. Reyes Mason, J. M. Hathaway, D. A. Howe, Heat and humidity in the city: Neighborhood heat index variability in a mid-sized city in the southeastern United States. *Int. J. Environ. Res. Public Health* **13**, 117 (2016).

38. G. Manoli, S. Fatichi, E. Bou-Zeid, G. G. Katul, Seasonal hysteresis of surface urban heat islands. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 7082–7089 (2020).

39. D. Li, W. Liao, A. J. Rigden, X. Liu, D. Wang, S. Malyshev, E. Shevliakova, Urban heat island: Aerodynamics or imperviousness? *Sci. Adv.* **5**, eaau4299 (2019).

40. X. Lee, M. L. Goulden, D. Y. Hollinger, A. Barr, T. A. Black, G. Bohrer, R. Bracho, B. Drake, A. Goldstein, L. Gu, G. Katul, T. Kolb, B. E. Law, H. Margolis, T. Meyers, R. Monson, W. Munger, R. Oren, K. T. Paw U, A. D. Richardson, H. P. Schmid, R. Staebler, S. Wofsy, L. Zhao, Observed increase in local cooling effect of deforestation at higher latitudes. *Nature* **479**, 384–387 (2011).

41. S. Luyssaert, M. Jammet, P. C. Stoy, S. Estel, J. Pongratz, E. Ceschia, G. Churkina, A. Don, K. Erb, M. Ferlicoq, B. Gielen, T. Grünwald, R. A. Houghton, K. Klumpp, A. Knohl, T. Kolb, T. Kuemmerle, T. Laurila, A. Lohila, D. Loustau, M. J. McGrath, P. Meyfroidt, E. J. Moors, K. Naudts, K. Novick, J. Otto, K. Pilegaard, C. A. Pio, S. Rambal, C. Rebmann, J. Ryder, A. E. Suyker, M. Varlagin, M. Wattenbach, A. J. Dolman, Land management and land-cover change have impacts of similar magnitude on surface temperature. *Nat. Clim. Chang.* **4**, 389–393 (2014).

42. E. S. Krayenhoff, A. M. Broadbent, L. Zhao, M. Georgescu, A. Middel, J. A. Voogt, A. Martilli, D. J. Sailor, E. Erell, Cooling hot cities: A systematic and critical review of the numerical modelling literature. *Environ. Res. Lett.* , (2021).

43. Y. T. Kwok, E. Y. Y. Ng, Trends, topics, and lessons learnt from real case studies using mesoscale atmospheric models for urban climate applications in 2000–2019. *Urban Clim.* **36**, 100785 (2021).

44. Z. S. Venter, O. Brousse, I. Esau, F. Meier, Hyperlocal mapping of urban air temperature using remote sensing and crowdsourced weather data. *Remote Sens. Environ.* **242**, 111791 (2020).

45. F. Meier, D. Fenner, T. Grassmann, M. Otto, D. Scherer, Crowdsourcing air temperature from citizen weather stations for urban climate research. *Urban Clim.* **19**, 170–191 (2017).

46. K. Hammerberg, O. Brousse, A. Martilli, A. Mahdavi, Implications of employing detailed urban canopy parameters for mesoscale climate modelling: A comparison between WUDAPT and GIS databases over Vienna, Austria. *Int. J. Climatol.* **38**, e1241–e1257 (2018).

47. T. Grassmann, A. Napoly, F. Meier, D. Fenner, *CrowdQC v1.2.0—Quality Control for Crowdsourced Data from CWS* (Technische Universität Berlin, 2018).

48. Copernicus Climate Change Service (C3S), ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate (2017); 10.24381/cds.adbb2d47.

49. M. Pesaresi, S. Freire, GHS Settlement grid, following the REGIO model 2014 in application to GHSL Landsat and CIESIN GPW v4-multitemporal (1975-1990-2000-2015). JRC Data Cat. (2016).

50. N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, R. Moore, Google Earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).

51. P. A. Mirzaei, F. Haghighat, Approaches to study Urban Heat Island – abilities and limitations. *Build. Environ.* **45**, 2192–2201 (2010).

52. H. D. Rozenfeld, D. Rybski, J. S. Andrade Jr., M. Batty, H. E. Stanley, H. A. Makse, Laws of population growth. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 18702–18707 (2008).

53. J. J. Danielson, D. B. Gesch, "Global multi-resolution terrain elevation data 2010 (GMTED2010)" (2331–1258, US Geological Survey, 2011).

54. M. A. Friedl, D. Sulla-Menashe, B. Tan, A. Schneider, N. Ramankutty, A. Sibley, X. Huang, MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ.* **114**, 168–182 (2010).

55. M. Buchhorn, M. Lesiv, N.-E. Tsendbazar, M. Herold, L. Bertels, B. Smets, Copernicus global land cover layers—Collection 2. *Remote Sens.* **108**, 1044 (2020).

56. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).

57. C. Crisci, B. Ghattas, G. Perera, A review of supervised machine learning algorithms and their applications to ecological data. *Ecol. Model.* **240**, 113–122 (2012).

58. S. Wager, S. Athey, Estimation and inference of heterogeneous treatment effects using random forests. *J. Am. Stat. Assoc.* **113**, 1228–1242 (2018).

59. Q. Zhao, T. Hastie, Causal interpretations of black-box models. *J. Bus. Econ. Stat.* **39**, 272–281 (2019).

60. Y. Choi, S. Lee, H. Moon, Urban physical environments and the duration of high air temperature: Focusing on solar radiation trapping effects. *Sustainability* **10**, 4837 (2018).

61. V. Masson, W. Heldens, E. Bocher, M. Bonhomme, B. Bucher, C. Burmeister, C. de Munck, T. Esch, J. Hidalgo, F. Kanani-Sühring, Y.-T. Kwok, A. Lemonsu, J.-P. Lévy, B. Maronga, D. Pavlik, G. Petit, L. See, R. Schoetter, N. Tornay, A. Votsis, J. Zeidler, City-descriptive input data for urban climate models: Model requirements, data sources and challenges. *Urban Clim.* **31**, 100536 (2020).

62. S. Running, Q. Mu, M. Zhao, MOD16A2 MODIS/Terra Net Evapotranspiration 8-Day L4 Global 500m SIN Grid V006. *NASA EOSDIS Land Process. DAAC* (2017).

63. C. Carter, S. Liang, Evaluation of ten machine learning methods for estimating terrestrial evapotranspiration from remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **78**, 86–92 (2019).

64. D. Courault, B. Seguin, A. Olioso, Review on estimation of evapotranspiration from remote sensing data: From empirical to numerical modeling approaches. *Irrig. Drain. Syst.* **19**, 223–249 (2005).

65. K. Zhang, J. S. Kimball, S. W. Running, A review of remote sensing based actual evapotranspiration estimation. *WIREs Water* **3**, 834–853 (2016).

66. A. Lefebvre, C. Sannier, T. Corpetti, Monitoring urban areas with Sentinel-2A data: Application to the update of the Copernicus high resolution layer imperviousness degree. *Remote Sens.* **8**, 606 (2016).

67. C. Schaaf, Z. Wang, MCD43A3-MODIS/Terra+Aqua BRDF/Albedo Daily L3 Global-500 m V006. *NASA EOSDIS Land Process. DAAC* **10** (2015).

68. D. M. Theobald, D. Harrison-Atlas, W. B. Monahan, C. M. Albano, Ecologically-relevant maps of landforms and physiographic diversity for climate adaptation planning. *PLOS ONE* **10**, e0143619 (2015).

69. L. Massetti, M. Petralli, G. Brandani, S. Orlandini, An approach to evaluate the intra-urban thermal variability in summer using an urban indicator. *Environ. Pollut.* **192**, 259–265 (2014).

70. A. F. Zuur, E. N. Ieno, C. S. Elphick, A protocol for data exploration to avoid common statistical problems. *Methods Ecol. Evol.* **1**, 3–14 (2010).

71. U. Groemping, Relative importance for linear regression in R: The package relaimpo. *J. Stat. Softw.* **17**, 1–27 (2006).

72. R. H. Lindeman, P. F. Merenda, R. Z. Gold, *Introduction to Bivariate and Multivariate Analysis* (Scott Foresman, 1980).

73. A. K. Heidinger, M. J. Foster, A. Walther, X. Zhao; NOAA CDR Program, NOAA Climate Data Record (CDR) of Cloud Properties from AVHRR Pathfinder Atmospheres - Extended (PATMOS-x), Version 5.3 (NOAA National Centers for Environmental Information, 2014); doi: 10.7289/V5348HCK.

**Citation:** Z. S. Venter, T. Chakraborty, X. Lee, Crowdsourced air temperatures contrast satellite measures of the urban heat island and its mechanisms. *Sci. Adv.* **7**, eabb9569 (2021).