scientific data

Check for updates

OPEN

DATA DESCRIPTOR

MORICHI: a Dataset to Study Urban Overheating during Extreme Heat in a Hot-Summer Humid Continental Climate

Miguel Martin[1], Clara Garcia-Sanchez[1], Jantien Stoter[1] & Mario Berges[2]✉

This paper describes a firsthand and open dataset comprising weather data collected from four street-level stations and microscale thermal images captured by a single infrared thermal camera during an extreme heat event in late August 2024 in Pittsburgh, United States. The weather data includes air temperature, relative humidity, wind speed and direction, and rainfall, measured at a height of 2 meters above the ground. From microscale thermal images, it is possible to assess the temperatures of built-up surfaces within a street canyon on a university campus, including a road, sidewalks, and building façades. Other factors that contribute to or mitigate urban overheating, such as waste heat emissions, traffic, and vegetation, can also be analyzed using the microscale thermal images. The weather data and microscale thermal images are publicly accessible in the 4TU.ResearchData repository under the CC BY 4.0 license. A Python library was developed to extract and process the data, particularly microscale thermal images, and is publicly available via the PIP package installer.

Background & Summary

Over the past two centuries, rural-to-urban migration has resulted in substantial global expansion of the built environment to accommodate urban residents. The built environment consists of buildings and roads that absorb heat from the sun during the day and release it at night, primarily as long-wave radiation. Due to the lack of heat dissipation from wind and evapotranspiration from trees, a considerable amount of heat remains trapped inside the street canyons of cities. As a result, cities have higher outdoor air temperatures than their suburban or rural surroundings, especially at night. This phenomenon is known as urban overheating or the heat island effect, and it has undesirable consequences for building energy consumption[1–4] and greenhouse gas emissions in cities around the world[5,6]. It also negatively impacts low-income populations[7–10] and the public health sector[11,12].

Unlike many other countries, urban overheating has not been regarded a significant climatic threat in the United States. Most states experience mild, cool, or very cold climates, where urban overheating may, on average, positively impact building energy consumption[13] and carbon emissions[14] over the year. Despite this, the country has started to take urban overheating more seriously than in the past. During extreme heat events, urban overheating threatens the building thermal resilience, which is buildings' ability to maintain a comfortable indoor environment for occupants during extreme heat events[15]. Since buildings in most states are not designed to withstand high temperatures, the risk of mortality from heat waves is significant[16].

The city of Pittsburgh, United States, is located in a region that could face major challenges in the public health sector due to urban overheating during extreme heat events. Its population is approximately 300,000, most of whom live in detached houses. According to the National Weather Services, temperatures can drop as low as -25 degrees Celsius in winter and rise as high as 35 degrees Celsius in summer[17]. As an industrial city, outdoor air quality used to be the primary environmental concern in Pittsburgh[18–21]. Nevertheless, a study by Ha *et al.*[22] shows that the risk of heatstroke in Allegheny County increases significantly if an extreme heat event lasts more than two days. Despite this alarming observation, few efforts have been made in Pittsburgh to understand the current or future consequences of urban overheating during heat waves.

[1]Delft University of Technology, Department of Urbanism, Delft, 2628 BL, Netherlands. [2]Carnegie Mellon University, Department of Civil and Environmental Engineering, Pittsburgh, PA 15213', USA. ✉e-mail: mberges@andrew.cmu.edu

Weather data collected during an extreme heat event in Pittsburgh can, apart from for Pittsburgh itself, be a valuable resource for other countries studying the magnitude of urban overheating. Pittsburgh has a hot-summer humid continental climate (Dfa in Köppen classification), similar to regions in the North-East United States, Eastern Europe, Central Asia, and East Asia[23]. A recent review by Wong *et al.*[24] shows that little weather data is available for urban areas in cities with a hot-summer humid continental climate, such as Pittsburgh. A key contribution in this area is the DCNet field experimental network, which has been operating since 2002 primarily in Washington, D.C., and New York City[25]. In New York City, two 10-meter-high meteorological towers were installed on building rooftops: one near Times Square and the other in a different Manhattan neighborhood. This kind of meteorological installation is suitable for assessing the impact of the heat emitted by the built environment on the urban boundary layer. However, it provides no information about weather conditions as experienced by pedestrians at the street level - a crucial consideration for studying the consequences of extreme heat events on the population. Therefore, street-level weather measurements from Pittsburgh would help other cities with similar climates understand both the magnitude of urban overheating during extreme heat events and the seriousness of the hazard this phenomenon poses to public safety.

Although weather data provide useful information about outdoor conditions during an extreme heat event, they are insufficient for identifying the specific built-environment elements that either contribute to or mitigate urban overheating. Analyzing the contributors and mitigators of urban overheating requires collecting weather data along with microscale thermal images of the built environment[26]. As reported in Martin *et al.*[27], microscale thermal images are captured by an infrared thermal camera installed on a rooftop observatory, drone, or mounting tripod. They observe the built environment within a range of one kilometer or less. For example, Dobler *et al.*[28] collected microscale thermal images of Midtown Manhattan, New York City, a location with a climate similar to that of Pittsburgh. Despite similar weather conditions, Midtown Manhattan's urban morphology is distinct from that of most areas in Pittsburgh, which is an important distinction when studying the contributors and mitigators of urban overheating. On the one hand, Midtown Manhattan is a business district characterized by high-rise buildings averaging 300 meters in height and avenues typically 30 meters wide[29,30]. On the other hand, Pittsburgh is largely covered by residential areas with buildings up to 12 meters high and streets at least 8 meters wide[31]. This results in a height-to-width ratio of about 10 for Midtown Manhattan, compared to a maximum of 1.5 in most Pittsburgh locations. Some U.S. cities, including Cincinnati[32,33] and Milwaukee[34,35], are morphologically closer to Pittsburgh and experience a similar climate. According to a review by Martin *et al.*[27], no thermal images have been captured in these cities.

Beyond regions with a hot-summer humid continental climate like Pittsburgh, urban overheating during extreme heat is a threat to public health in many cities around the world, particularly in an era of climate change. This explains why studies like the one published by Yan *et al.*[36] urge the scientific community to put effort into collecting weather data on extreme heat events at a global scale. Such a collective effort is instrumental in enhancing the analysis of building thermal resilience. Due to a scarcity of weather data from extreme heat events, little is known about the ability of buildings to adapt to such conditions, both now and in the future[37]. Recently, some efforts have been made to integrate extreme heat conditions into typical meteorological data[38]. However, these efforts fall short of capturing the impact of urban overheating on outdoor conditions within cities.

In the collective effort to augment data availability to study the influence of urban overheating on outdoor conditions during extreme heat events, particularly in a hot-summer humid continental climate, a dataset was created from measurements taken by weather stations and an infrared thermal camera on a university campus in Pittsburgh. The data were collected from July 26 to September 20, 2024, a period during which an extreme heat event was recorded with temperatures reaching 36 degrees Celsius during the day and 24 degrees Celsius at night. The outdoor air temperature and other weather parameters-such as relative humidity, wind speed, wind direction, and rainfall-were measured at 2 meters above ground every 5 minutes. Alongside weather data, microscale thermal images were captured with a resolution of 640 x 480 pixels and a frequency of 1 image per second. The absolute surface temperature obtained by microscale thermal images was calibrated against ground-truth measurements below a Root Mean Square Error (RMSE) of 1.7 degrees Celsius. Together, the weather data and microscale thermal images constitute a firsthand dataset of more than 1 terabyte, which is openly available to the scientific community in a repository administered by the Delft University of Technology (TU Delft). They can be shared and adapted without restriction on the condition that appropriate credit is given to the authors, a link to the CC BY 4.0 license is provided, and any changes made are indicated. A Python library was developed to facilitate the extraction and processing of data stored in the dataset.

Despite being primarily created to enable the scientific community to assess building thermal resilience in an area affected by urban overheating, the dataset can be used for various applications in urban microclimate modeling. One of them is the validation of physics-based urban microclimate models. Physics-based urban microclimate models, particularly those using computational fluid dynamics, are typically validated against measurements of outdoor conditions[39,40]. The agreement between estimates and measurements of outdoor conditions heavily depends on the assigned temperature of built-up surfaces, a key boundary condition in physics-based models. Therefore, the models' validity can be further demonstrated by comparing assigned boundary conditions and surface temperature measurements from microscale thermal images. As an alternative to the validation of physics-based models, both weather data and microscale thermal images can be used to train data-driven approaches. Traditionally, data-driven urban microclimate models have primarily been trained on typical weather data, which limits their generalizability[41]. This limitation can be overcome using training data collected during extreme events. Whether models are physics-based or data-driven, their reliability will greatly improve if they properly capture urban dynamics such as building operation and traffic. Urban dynamics can be captured using microscale thermal images collected at high frequency as shown in various studies[28,42].

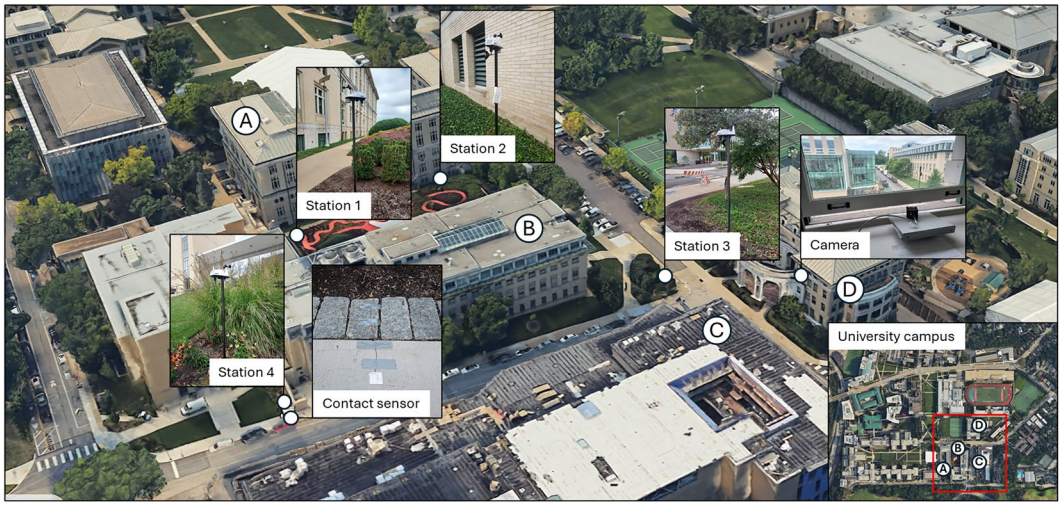


Fig. 1 Field experimental network in a university campus in Pittsburgh.

This paper serves as a data descriptor for the dataset openly available at the 4TU.ResearchData repository[43]. The first section describes the procedures used to create the dataset, particularly the collection and processing of weather data and microscale thermal images on a university campus in Pittsburgh. The second section details the content of the dataset and the repository where it is accessible. The third section gives a short overview of the analysis that can be performed using the dataset. The fourth section is dedicated to the technical validation of the dataset. Finally, the fifth section demonstrates how to easily access and process the data in the dataset using Python.

Methods

Experimental design. The dataset described in this paper was created from measurements made by four weather stations and one infrared thermal camera on a university campus in Pittsburgh, United States. Pittsburgh is home to over 145,000 students, who attend undergraduate and graduate programs in one of the six universities in the city[44]. About 11% of these students attend Carnegie Mellon University (CMU)[45]. The CMU campus (Lat. 40.4400517 and Long. -79.9394578) comprises buildings, labs, and residence halls that range from 3 to 9 stories in height. It also has green spaces and sports fields for accommodating various student activities.

The four weather stations and the infrared thermal camera used to create the dataset operated between buildings A, B, C, and D on the CMU campus from July 26 to September 20, 2024 (see Fig. 1). The buildings respective heights are approximately 33, 29, 26, and 31 meters[46]. Buildings A, B, and C consist of offices, classrooms, and laboratories. Building D is a newly constructed sports center. Buildings A and B are separated by a 35-meter-wide garden with walking trails. Cars can circulate and stop between buildings B, C and D on an asphalted road 10 meters wide. The central axes of the garden and the road are oriented north-south. Sidewalks are located next to the road to provide access for pedestrians. The sidewalks near Building B are shaded by trees and bordered by grass. A chiller plant is installed on the roof of Building B.

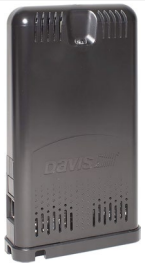
The dataset consists of weather data and microscale thermal images that were mainly collected in the street canyon delimited by buildings B and C. Its height-to-width ratio is around 1.0, a value commonly found in most urban areas of Pittsburgh except the central business district[31]. The composition of the street is also typical, comprising an asphalt road, concrete sidewalks, and grass with trees. The traditional brick construction of buildings in Pittsburgh is similar to that of building B but different from that of building C. Building C is a recently built sports center with curtain walls. Besides this small divergence, the street canyon between buildings B and C has a morphology and composition common in Pittsburgh. Therefore, the urban overheating experienced in this street canyon during extreme heat is likely representative of that in many other parts of Pittsburgh.

Weather measurements. Weather data were measured from four Davis Vantage Vue weather stations. Two of them (1 and 2) were placed in the garden between buildings A and B. At the same time, the other two (3 and 4) were located in the street canyon between buildings B and C. Based on the recommendations of Davis Instrument[47], weather stations were installed 2 meters above ground, mounted on a metallic pipe secured by a drilled hole. They were powered by solar panels and backup batteries, which allowed them to operate autonomously. Data collection was managed collectively through a cloud platform, which was connected to all weather stations via a WeatherLink Live gateway. At five-minute intervals, the cloud platform stored the average, minimum, and maximum values of 12 weather parameters measured by the stations. The range, accuracy, and resolution with which weather stations measured these parameters are described in Table 1. Some of these parameters, such as dry-bulb temperature, relative humidity, wind speed, wind direction, and rainfall, are measured from sensors on the weather stations. Others are derived from measured parameters to indicate outdoor thermal comfort. Table 3 shows how the weather stations calculated some of these parameters.

The WeatherLink Live gateway was installed inside Building D, less than 100 meters from the weather stations to minimize communication interference. As shown in Table 2, the gateway consists of internal sensors

Parameter	Range	Accuracy	Resolution
Dry-bulb temperature	−40 ... +150 °F	±1°F	0.1 °F
Relative humidity	0 ... 100 %	±3%	0.1 %
Wind speed	2 ... 180 mph	±2 mph	0.1 mph
Wind direction	0 ... 360°	±3°	0.1°
Rainfall	0 ... 199.99"	4%	0.01"
Dew point temperature (*)	−105 ... +130 °F	±3 °F	0.1 °F
Wind chill (*)	−110 ... +135 °F	±2°F	0.1 °F
Heat index (*)	−40 ... +165 °F	±3°F	0.1 °F
Wet-bulb temperature (*)	N/A	N/A	N/A
THW index (*)	N/A	N/A	N/A
Heating degree days (*)	N/A	N/A	N/A
Cooling degree days (*)	N/A	N/A	N/A

Table 1. Specification of measured and calculated (*) weather parameters by Davis Vantage Vue stations.



Parameter	Range	Accuracy	Resolution
Dry-bulb temperature	+32 ... +140 °F	±0. 5°F	0.1 °F
Relative humidity	1 ... 100 %	±2 %	0.1 %
Barometric Pressure	16" ... 32.50" Hg	N/A	0.01" Hg
Dew point temperature (*)	N/A	N/A	N/A
Heat index (*)	N/A	N/A	N/A
Absolute pressure	N/A	N/A	N/A

Table 2. Specification of measured and calculated (*) parameters by the WeatherLink Live gateway.

that can be used to measure indoor conditions. Like the weather stations, the minimum, average, and maximum values of each parameter measured by the WeatherLink Live gateway were recorded at 5-minute intervals. Throughout the measurements, the indoor space where the WeatherLink Live gateway was installed was minimally occupied, with little use of air conditioning.

Infrared thermography. Microscale thermal images were captured by a FLIR A50 infrared thermal camera, whose specifications are outlined in Table 4. The infrared thermal camera was installed on the fourth floor of building D, which corresponds to a height of 30 meters above ground. It was connected to a laptop using an Ethernet cable. The laptop was powered by the building's electrical supply, and the FLIR Research Studio software was used to manage data collection. Using the software, microscale thermal images were captured at a frequency of 1 Hertz and stored in a video file every 5-minute intervals. To minimize disturbances while capturing microscale thermal images, the camera was positioned behind an open window. With this installation setup, the radiometric values $\hat{U}_{ij}$ captured by the infrared thermal camera at pixels ij were primarily influenced by reflected sky radiation and outdoor air transmission, that is:

$$\hat{U}_{ij} = \frac{1}{\varepsilon_{ij}\tau_{out}} U_{ij} - \frac{1-\varepsilon_{ij}}{\varepsilon_{ij}} U_{sky} - \frac{1-\tau_{out}}{\varepsilon_{ij}\tau_{out}} U_{out} \quad (1)$$

where U_{ij} (in Volts) is the true radiometric value of the observed object at pixel ij, ε_{ij} (0-1) the emissivity of the object, U_{sky} (in Volts) and U_{out} (in Volts) the radiometric value associated to the sky temperature T_{sky} (in K) and outdoor air temperature T_{out} (in K), respectively, and τ_{out} (0-1) the outdoor air transmissivity. Similar to the infrared thermal camera used by Martin *et al.*[26], the surface temperature (T_{ij}) of the observed object at pixels ij can be expressed as:

$$T_{ij} = b \cdot \ln\left[\frac{r_1}{r_2 \cdot \left(U_{ij} + O\right)} + f\right]^{-1} \quad (2)$$

where b, r_1, r_2, O, and f are parameters to be calibrated against ground truth measurements. It is important to note that calibrating these parameters is only necessary when the microscale thermal images are used for quantitative studies that rely on absolute surface temperature (T_{ij}). Qualitative studies such as waste heat and traffic detection can easily be conducted based on radiometric values (U_{ij}).

To calibrate parameters b, r_1, r_2, O, and f for the Pittsburgh's conditions, a thermocouple of type J was placed on the sidewalk next to weather station 4. This type of thermocouple can measure the surface temperature within a

Parameter	Formula
Cooling degree days - T_{CDD} (unitless)	$T_{CDD} = \begin{cases} \left(\max_{\mathbb{T}} T - \min_{\mathbb{T}} T\right)/2 - 65 & \text{if } \left(\max_{\mathbb{T}} T - \min_{\mathbb{T}} T\right)/2 \geq 65 \\ 0 & \text{otherwise} \end{cases}$
	where: • T - Dry bulb temperature (in ^{o}F) • $\mathbb{T}$ - Time period
Dew point - T_d (in ^{o}C)	$T_d = T - ((100 - RH)/5)$
	where: • T - Dry bulb temperature (in ^{o}C) • RH - Relative humidity (in %)
Heating degree days - T_{HDD} (unitless)	$T_{HDD} = \begin{cases} 65 - \left(\max_{\mathbb{T}} T - \min_{\mathbb{T}} T\right)/2 & \text{if } \left(\max_{\mathbb{T}} T - \min_{\mathbb{T}} T\right)/2 < 65 \\ 0 & \text{otherwise} \end{cases}$
	where: • T - Dry bulb temperature (in ^{o}F) • $\mathbb{T}$ - Time period
Heat index - T_{HI} (in ^{o}F)	$T_{HI} = \begin{cases} \frac{1}{2}[T + 61.0 + 1.2(T - 68.0) + 0.094RH] & \text{if } T_{HI} \leq 80^o \text{ F} \\ -42.379 + 2.04901523T + 10.14333127RH & \\ -.22475541TRH - .00683783T^2 & \\ -.05481717RH^2 + .00122874T^2RH & \\ +.00085282TRH^2 - .00000199T^2RH^2 & \text{otherwise} \end{cases}$
	where: • T - Dry bulb temperature (in ^{o}F) • RH - Relative humidity (in %)
THW index - T_{THW} (in ^{o}F)	$T_{THW} = T_{HI} - 1.072V$
	where: • T_{HI} - Heat index (in ^{o}F) • V - Wind speed (in mph)
Wet bulb temperature - T_{wb} (in ^{o}C)	$T_{wb} = T\arctan[0.151977(RH + 8.313658)^{1/2}] + \arctan(T + RH)$
	$\arctan(RH - 1.676331) + 0.00391838RH^{3/2}\arctan(0.023101RH) - 4.686035$
	where: • T - Dry bulb temperature (in ^{o}C) • RH - Relative humidity (in %)
Wind chill - T_{wc} (in ^{o}F)	$T_{wc} = 35.74 + 0.6215T - 35.75(V^{0.16}) + 0.4275T(V^{0.16})$
	where: • T - Dry bulb temperature (in ^{o}F) • V - Wind speed (in mph)

Table 3. Calculation of some weather parameters available in the Davis Vantage Vue stations.

	Feature	Value
	Field of view	95^o (W) × 74^o (H)
	Image resolution	464 (W) × 348 (H)
	Accuracy	±2°C or ±2% of reading
		(for ambient temperature 15°C-35°C)
	Operating temperature range	−20 to 50 °C
	Storage temperature range	−20 to 175 °C
	Spectral range	7.5 - 14 μm
	Image format	16-bit

Table 4. Specifications of the FLIR A50 infrared thermal camera.

range of -210 and 760 degrees Celsius, with an accuracy of ±0.6 degrees Celsius and a resolution of 0.03 degrees Celsius. The thermocouple was protected against solar radiation using an aluminum foil. It was connected to a HOBO UX0100-014M data logger, which was used to measure the surface temperature of the sidewalk every minute. The installation of the ground truth follows the procedure described by Martin *et al.*[26]. This study shows that FLIR infrared thermal cameras can accurately measure the absolute surface temperature (T_{ij}) on different types of surfaces, provided that the parameters b, r_1, r_2, O, and f are manually or automatically calibrated at a single point to an RMSE below 1.7 degrees Celsius. The discrepancy between the surface temperature assessed by the camera and that measured by the ground truth is strongly influenced by the parameters O and r_2. In accordance with Martin *et al.*[26], these two parameters were adjusted until the discrepancy fell below an RMSE of 1.7 degrees Celsius Fig. 2.

Data collection. From July 26, 2024, the four Davis Vantage Vue weather stations, the FLIR A50 camera, and the HOBO thermocouple collected weather data, microscale thermal images, and ground truth measurements, respectively. As shown in Fig. 3, data collection for microscale thermal images was interrupted on the following dates: the last week of July, the first week of August, August 25, and the second week of September. Between the last week of July and the first week of August, interruptions were due to ongoing system configuration. Indeed, infrared thermal cameras are complex systems that require more time to be properly configured than weather stations or contact sensors. Other interruptions were caused by unexpected operating system updates. Ground truth measurements were stopped at the end of the first week of September. While weather data and microscale thermal images were collected using devices constantly connected to the Internet, an autonomous HOBO data logger collected the ground truth measurements. To avoid data loss and resynchronize its clock, the data logger was stopped and relaunched at least once a week throughout the data collection period.

Fig. 2 Structure of the MORICHI dataset.

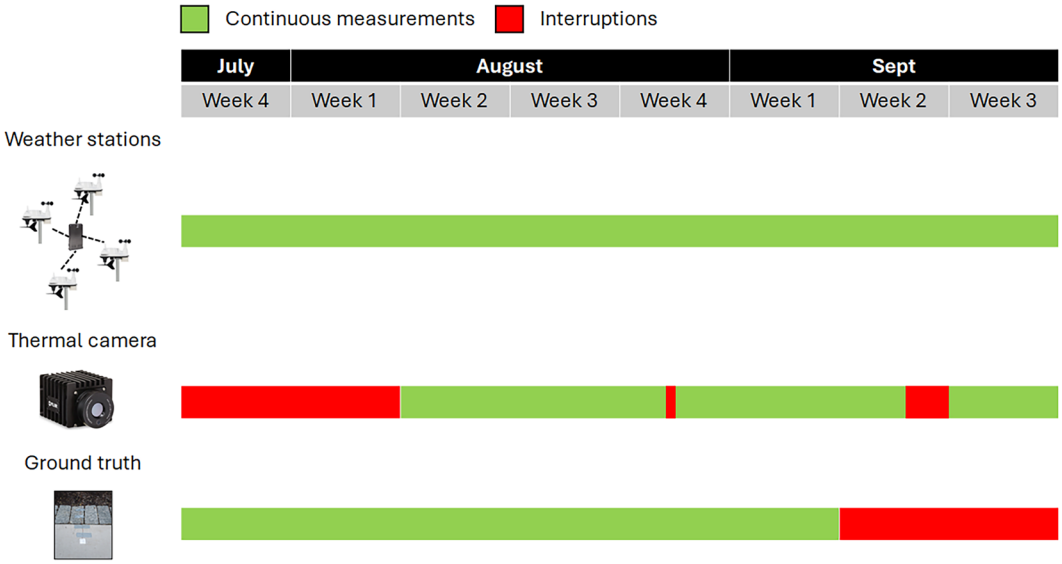


Fig. 3 Data collection of weather conditions, microscale thermal images, and ground truth.

Weather data, microscale thermal images, and ground truth measurements were collected in the original time zones of their respective data loggers. While this requires post-processing to align the data for analysis, it preserves data integrity by preventing gaps or overlaps caused by time zone changes. It also prevents configuration errors, such as the inversion of days and months, which is a typical problem when adapting a data collection system from European to US standards.

Among the collected data, microscale thermal images require the most ethical consideration. When researchers use an infrared thermal camera to observe the outdoor built environment, pedestrians might appear in the thermal images. To preserve pedestrians' anonymity, the infrared thermal camera was installed far enough away so that they could not be identified. Additionally, data managers checked all ethical issues related to the collected data before publishing it in the 4TU.ResearchData repository[48].

Data Records

Weather data and microscale thermal images collected from July 26 to September 20, 2024, on the CMU campus are stored in the 4TU.ResearchData repository under the title *Meteorological Observations and Radiometric Images for Cold Climate Heat Wave Investigations* (MORICHI). The 4TU.ResearchData repository is hosted and managed by TU Delft in accordance with its legal statements. It implements the Findable, Accessible, Interoperable, and Reusable (FAIR) framework by curating, sharing, and preserving data for long-term access. Only members of

Name	Size (in megabytes)	Description
MO.zip	13	Contains weather data collected by each station
GT.zip	0.4	Contains measurements of the contact sensor used to validate microscale thermal images
TIR_YYYY_MM_DD.zip	4,000–18,000	Contains microscale thermal images collected over a day

Table 5. Compressed files of the MORICHI dataset stored in the 4TU.ResearchData repository.

Instrument	Format	Time zone
Weather stations	MM/DD/YY hh:mm [AM|PM]	Pittsburgh time (GMT-5)
Infrared thermal camera	YYYY_MM_DD_hh_mm_ss_[AM|PM]	London time (GMT-0)
Contact sensor	MM/DD/YY hh:mm:ss [AM|PM]	New York time (GMT-4)

Table 6. Format and time zone considered by the internal clock of each instrument for data storage.

the 4TU.ResearchData Consortium can create and modify datasets in the 4TU.ResearchData repository. While data in the 4TU.ResearchData repository is publicly available by default, access can be embargoed or restricted if there is a reason to do this. Various open data access licenses are available, including CC0, CC BY, CC BY-SA, CC BY-ND, CC BY-NC, CC BY-NC-SA, CC BY-NC-ND, and ODbL. The MORICHI dataset, for instance, was stored under the CC BY 4.0 license. It means the dataset can be freely distributed, remixed, and augmented for any purpose, including commercial use, with proper attribution to the creating authors and institutions.

Within the 4TU.ResearchData repository, the MORICHI dataset is accompanied by metadata that includes a list of authors, their affiliations, funding agencies, the publisher, a brief description, data formats, categories, keywords, geolocation, and time coverage. An MD5 checksum is generated for each file stored in the repository. Most of the stored files are compressed folders containing either weather data or microscale thermal images. Their sizes are shown in Table 5. Apart from compressed files, it is possible to find a list of Jupyter notebooks explaining how to use the MORICHI dataset.

Figure 2 illustrates the structure with which the MORICHI dataset can be downloaded in a local machine or remote server when running the script **install.py**. The data structure consists of two folders: the Meteorological Observations (MO) and the Radiometric Images (RI). The MO folder includes the weather data collected by each of the four Davis Vantage Vue weather stations and the WeatherLink Live gateway. Each file in this folder is in Comma-Separated Values (CSV) format and specifies the station identifier, along with the start and end dates of data collection. The other folder, RI, has several subfolders, each corresponding to a specific day on which the infrared thermal camera captured microscale thermal images. Each subfolder contains microscale thermal image recordings in a format readable only by the FLIR Research Studio software or the Python SDK. The names of the recordings indicate the period during which the FLIR A50 camera captured the microscale thermal images. The RI folder also includes a subfolder, Ground Truth (GT), where the contact surface sensor measurements are stored. The structure and usage of the MORICHI dataset are further detailed in the accompanying README file table 6.

Weather data collected by the Davis Vantage Vue weather stations and the WeatherLink Live gateway are stored in CSV tables consisting of 29 and 13 columns, respectively, as shown in Tables 7 and 8. The first column indicates the date and time. Each subsequent column corresponds to the 5-minute average, minimum, or maximum value of a measured or assessed weather parameter, with most of these values expressed as real numbers with one decimal place. Measured weather parameters include dry-bulb temperature (in degrees Fahrenheit), relative humidity (in percent), wind speed (in miles per hour), wind direction (in cardinal coordinates), rainfall (in inches), and barometric pressure (in inches of Mercury). Assessed weather parameters mainly provide indications of thermal sensation at the pedestrian level, where the Davis Vantage Vue weather stations were installed. Some of them, such as dew point temperature, wet bulb temperature, and the heat index (all in degrees Fahrenheit), are thermal sensation metrics that account for the combined influence of temperature and humidity on thermal comfort. Others like wind chill (in degrees Fahrenheit) consider temperature and wind speed. The THW index (in degrees Fahrenheit) is a comprehensive indicator that builds upon the heat index by also incorporating wind speed to assess thermal sensation. Apart from thermal sensation, other indicators such as heating and cooling degree days relate outdoor conditions to building energy consumption. For each measured and assessed weather parameter, a total of 16,275 measurements were made from July 26 to September 20, 2024. This period is reflected in the file names where the weather data are stored. Files corresponding to data collected by weather stations are prefixed by their identifier. The file sizes are 3.26 megabytes for data created from the weather stations and 1.70 megabytes for that collected by the WeatherLink Live gateway.

Microscale thermal images captured by the FLIR A50 camera are stored in sequence within files that are considerably larger than those storing weather data. Each file is approximately 95 megabytes in size and contains 300 microscale thermal images. Their names are prefixed by TIR and indicate the 5-minute interval during which microscale thermal images were captured by the FLIR A50 camera. The camera uses a semi-proprietary format, SEQuence (SEQ), to store a sequence of 348-by-464 matrices of 16-bit integers, representing radiometric values ($\hat{U}_{ij}$). The sequence is associated with metadata that provide information about the camera, images, calibration parameters, geolocation, and timestamps. SEQ files can be read using the FLIR Research Studio or Software Development Kit. They can be converted to other video file formats, such as MP4 or AVI, using the **ffdshow** tool developed by FLIR.

Label	Unit	Data type	Description
Temp	°F	one-decimal real number	Average dry-bulb temperature recorded over a 5-minute period
High Temp	°F	one-decimal real number	Highest dry-bulb temperature recorded over a 5-minute period
Low Temp	°F	one-decimal real number	Lowest dry-bulb temperature recorded over a 5-minute period
Hum	%	one-decimal real number	Average relative humidity recorded over a 5-minute period
High Hum	%	one-decimal real number	Highest relative humidity recorded over a 5-minute period
Low Hum	%	one-decimal real number	Lowest relative humidity recorded over a 5-minute period
Dew Point	°F	one-decimal real number	Average dew point temperature recorded over a 5-minute period
High Dew Point	°F	one-decimal real number	Highest dew point temperature recorded over a 5-minute period
Low Dew Point	°F	one-decimal real number	Lowest dew point temperature recorded over a 5-minute period
Wet Bulb	°F	one-decimal real number	Average wet-bulb temperature recorded over a 5-minute period
High Wet Bulb	°F	one-decimal real number	Highest wet-bulb temperature recorded over a 5-minute period
Low Wet Bulb	°F	one-decimal real number	Lowest wet-bulb temperature recorded over a 5-minute period
Avg Wind Speed	mph	one-decimal real number	Average wind speed recorded over a 5-minute period
High Wind Speed	mph	one-decimal real number	Highest wind speed recorded over a 5-minute period
Prevailing Wind Direction	cardinal direction	string	Average wind direction recorded over a 5-minute period
High Wind Direction	cardinal direction	string	Most frequent wind direction recorded over a 5-minute period
Wind Chill	°F	one-decimal real number	Average wind chill recorded over a 5-minute period
Low Wind Chill	°F	one-decimal real number	Lowest wind chill recorded over a 5-minute period
Heat Index	°F	one-decimal real number	Average heat index recorded over a 5-minute period
High Heat Index	°F	one-decimal real number	Highest heat index recorded over a 5-minute period
THW Index	°F	one-decimal real number	Average temperature-humidity-wind index recorded over a 5-minute period
High THW Index	°F	one-decimal real number	Highest temperature-humidity-wind index recorded over a 5-minute period
Low THW Index	°F	one-decimal real number	Lowest temperature-humidity-wind index recorded over a 5-minute period
Wind Run	miles	one-decimal real number	Total wind run recorded over a 5-minute period
Rain	inches	two-decimal real number	Total rainfall recorded over a 5-minute period
High Rain Rate	inches per hour	two-decimal real number	Highest rainfall rate recorded over a 5-minute period
Cooling Degree Days	-	three-decimal real number	Total cooling degree days recorded over a 5-minute period
Heating Degree Days	-	three-decimal real number	Total heating degree days recorded over a 5-minute period

Table 7. Specification of weather parameters stored in CSV files created from measurements of the four Davis Vantage Vue weather stations.

Label	Unit	Data type	Description
Inside Temp	°F	one-decimal real number	Average inside dry-bulb temperature recorded over a 5-minute period
High Inside Temp	°F	one-decimal real number	Highest inside dry-bulb temperature recorded over a 5-minute period
Low Inside Temp	°F	one-decimal real number	Lowest inside dry-bulb temperature recorded over a 5-minute period
Inside Hum	%	one-decimal real number	Average inside relative humidity recorded over a 5-minute period
High Inside Hum	%	one-decimal real number	Highest inside relative humidity recorded over a 5-minute period
Low Inside Hum	%	one-decimal real number	Lowest inside relative humidity recorded over a 5-minute period
Inside Dew Point	°F	one-decimal real number	Average inside dew point temperature recorded over a 5-minute period
Inside Heat Index	°F	one-decimal real number	Average inside heat index recorded over a 5-minute period
Barometer	inches Hg	three-decimal real number	Average inside barometric pressure recorded over a 5-minute period
High Bar	inches Hg	three-decimal real number	Highest inside barometric pressure recorded over a 5-minute period
Low Bar	inches Hg	three-decimal real number	Average inside barometric pressure recorded over a 5-minute period
Absolute Pressure	inches Hg	three-decimal real number	Average inside absolute pressure recorded over a 5-minute period

Table 8. Specification of parameters stored in CSV files created from measurements of the WeatherLink Live gateway.

The date of collected data appears in two locations: in the first column for weather data and ground truth measurements, and in the filenames for microscale thermal images. The first column for weather data and ground truth measurements includes both the date and time. In contrast, the filenames for microscale thermal images contain only the date. A Python module called **irim** was developed to extract the date and time of each microscale thermal image captured by the FLIR A50 camera. As described in Table 6, the date and time of weather data, microscale thermal images, and ground truth measurements are expressed with a different format and time zone. As for the time zone, it was preferred to keep the original format that the device used to timestamp the collected data. This prevents errors related to misconfiguration. Python libraries used for data analytics, such as Pandas, facilitate timestamp standardization.

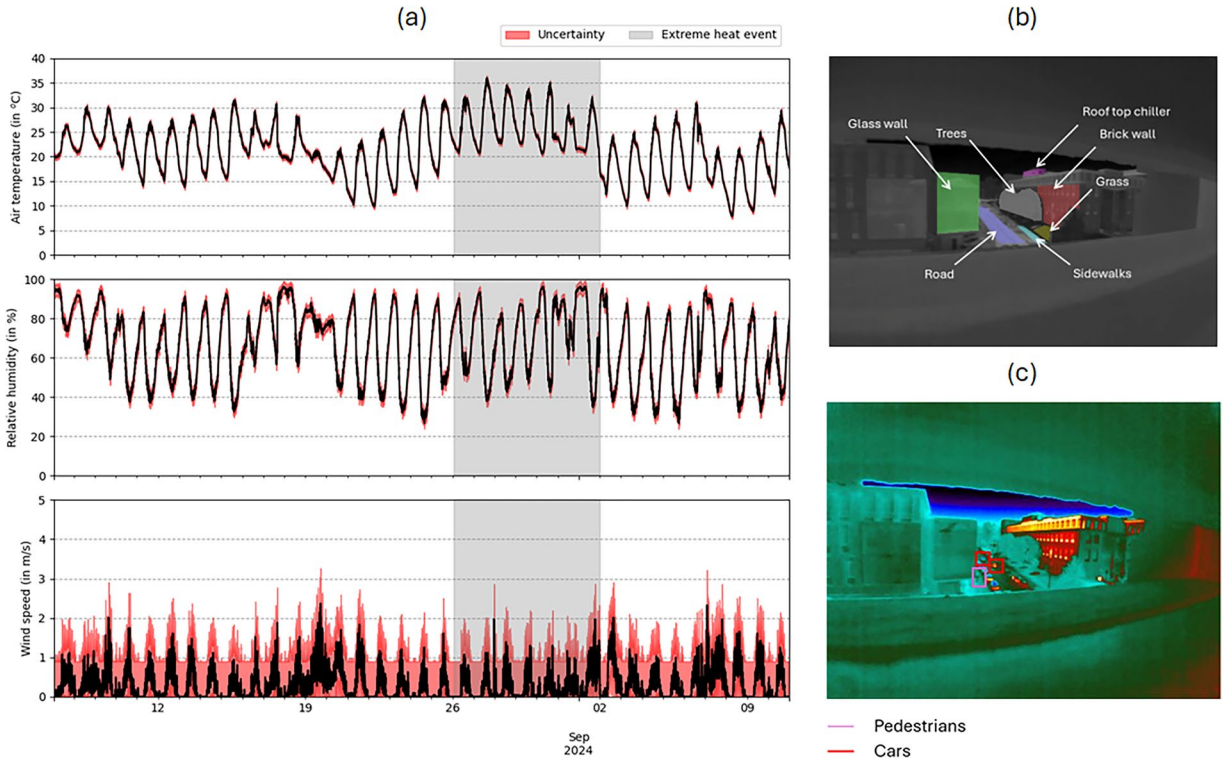


Fig. 4 Overview of the MORICHI dataset, including the outdoor conditions (**a**), static elements that can be monitored from microscale thermal images (**b**), and dynamic elements that can be detected from microscale thermal image (**c**).

Data Overview

Figure 4 illustrates the outdoor conditions from the weather data and the static/dynamic elements of the built environment observed in the microscale thermal images. This overview indicates various analyses that can be performed using the MORICHI dataset:

- Firstly, it demonstrates that outdoor thermal comfort can be monitored during an extreme heat event in a hot-summer humid continental climate. Based on outdoor air temperature measurements, an extreme heat event was recorded between August 26 and September 1, 2024. During this extreme heat event, air temperature, relative humidity, and wind speed were measured within a street canyon (i.e., weather stations 3 and 4) and on top of a garden (i.e., weather stations 1 and 2). This means the MORICHI dataset can potentially be used to study the ability of greenery to improve outdoor thermal comfort during an extreme heat event, a key factor in building more heat-resilient urban spaces.
- Secondly, it proves that contributors to and mitigators of urban overheating can be tracked during the extreme heat event. Potential contributors include a road, a sidewalk, two building walls, and one rooftop chiller. Trees and grass are the potential mitigators captured by microscale thermal images during the extreme heat event. As highlighted by Martin *et al.*[26], one can assess the heat fluxes of contributors and mitigators using both their surface temperatures from microscale thermal images and the outdoor conditions from weather stations. Urban heat fluxes are relevant information for the scientific community to understand the causes of urban overheating and potential strategies to minimize its magnitude, particularly during an extreme heat event.
- Thirdly, it shows that dynamic elements such as pedestrians and traffic can be detected from microscale thermal images. These elements, mainly traffic, are sources of anthropogenic heat, which can further contribute to urban overheating. However, this is not the most important aspect that can be studied using pedestrian and traffic detection. It can be used to analyze how urban overheating during an extreme heat event impacts human behavior. This analysis could result in valuable guidance to authorities on how to protect the population from extreme heat in cities with a hot-summer humid continental climate.

These examples are three of the many possible analyses that can be performed using the MORICHI dataset.

Technical Validation

Image noise. By default, microscale thermal images are matrices of size $M \times N$ consisting of radiometric values $\hat{U}_{i,j}[n]$ captured by the infrared thermal camera at time $t_n = t_0 + n \cdot \Delta t$. Noise can then affect their quality both spatially and temporally. The spatial influence of noise on microscale thermal images can be assessed using the method developed by Donoho and Johnstone[49]. This method computes the standard deviation σ_U over U,

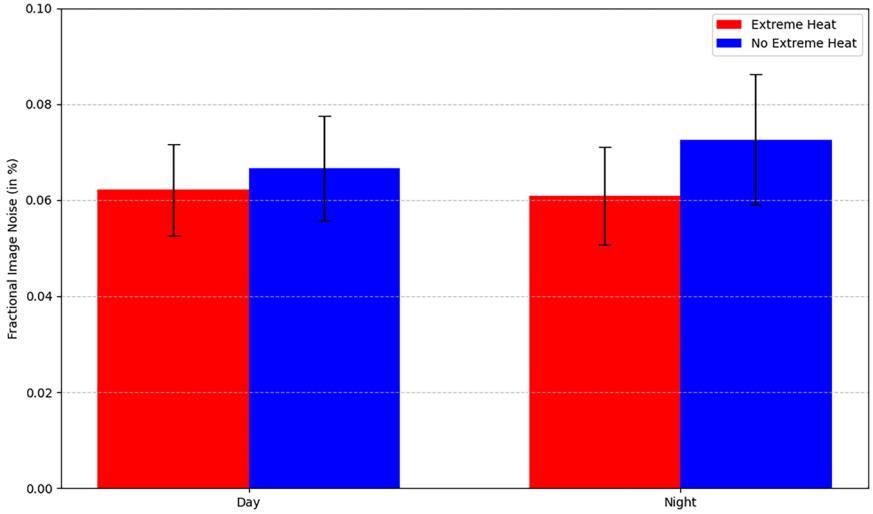


Fig. 5 Fraction of the noise contained in microscale thermal images collected both during the day and at night, and both during and outside of an extreme heat event.

under the assumption that noise is normally distributed. To compare its significance at different moments during data collection, σ_U was normalized with respect to the mean of radiometric values ($\overline{U}$):

$$FIN = 100 \times (\sigma_U / \overline{U}) \tag{3}$$

where *FIN* is the magnitude of noise contained in a microscale thermal image expressed as a percentage of the mean radiometric value ($\overline{U}$). Its magnitude was calculated for microscale thermal images collected both during the day and at night, and both during and outside of an extreme heat event.

Figure 5 shows the results from the spatial analysis of noise in microscale thermal images collected between July 26 and September 20, 2024. The magnitude of the noise was consistently less than 0.1% of the mean radiometric value, regardless of when the microscale thermal images were collected. This outcome dispels two concerns regarding the spatial quality of microscale thermal images. One concern was the high ambient temperatures during the extreme heat event, which sometimes reached 35 degrees Celsius. According to the specifications of the FLIR A50 infrared thermal camera, an accuracy of ± 2 degrees Celsius is guaranteed within an ambient temperature range of 15 to 35 degrees Celsius. The other concern, reported by Martin *et al.*[26], was the potential for this accuracy to degrade at night. While slight variations in noise magnitude are observed during the extreme heat event and at night, these do not constitute major inaccuracies in radiometric values collected by the infrared thermal camera.

Temporal stability. An Allan deviation analysis was performed to assess the temporal influence of noise and short-term drift on both microscale thermal images and weather data. To accomplish this, three variables-air temperature, relative humidity, and wind speed-were each extracted from the weather data as a separate time series ($X_{Temp}[n]$, $X_{Hum}[n]$, and $X_{WS}[n]$). The temporal influence of noise and short-term drift on microscale thermal images was calculated using the time series of radiometric values $X_{Rad}[n] = U_{i_{GT},j_{GT}}[n]$, collected at the ground truth position (i_{GT}, j_{GT}). To perform the Allan deviation analysis, $X_{Temp}[n]$, $X_{Hum}[n]$, $X_{WS}[n]$, and $X_{Rad}[n]$ were first decomposed into a trend $T[n]$, a seasonal component $S[n]$, and residuals $\varepsilon[n]$, such that $X[n] = T[n] + S[n] + \varepsilon[n]$. As the collection period was shorter than two months, the long-term drift was considered to have a negligible impact on the weather data and microscale thermal images. Therefore, the trend and seasonal components, $T[n] + S[n]$, can be assumed to contain only the natural or expected variations in the data, such as air temperature, relative humidity, wind speed, and radiometric values. To determine the influence of noise and short-term drift on both microscale thermal images and weather data, the Allan deviation $\sigma(\tau)$ was then calculated from the residuals $\varepsilon[n]$ over various averaging times τ. The magnitude of $\sigma(\tau)$ was normalized by the mean of each time series to compare the temporal stability between the weather data and the microscale thermal images:

$$FAD = \sigma(\tau) \& \#x00f7; \frac{1}{N}\sum_{n=1}^{N} X[n] \tag{4}$$

where *FAD* is the fractional Allan deviation.

The temporal stability of weather data and microscale thermal images, as assessed by the Allan deviation analysis, is shown in Fig. 6. Three main observations can be made:

- The stability of the weather data is consistent across all stations. A small difference exists between the wind speed collected by weather station 1 and that collected by the other stations.

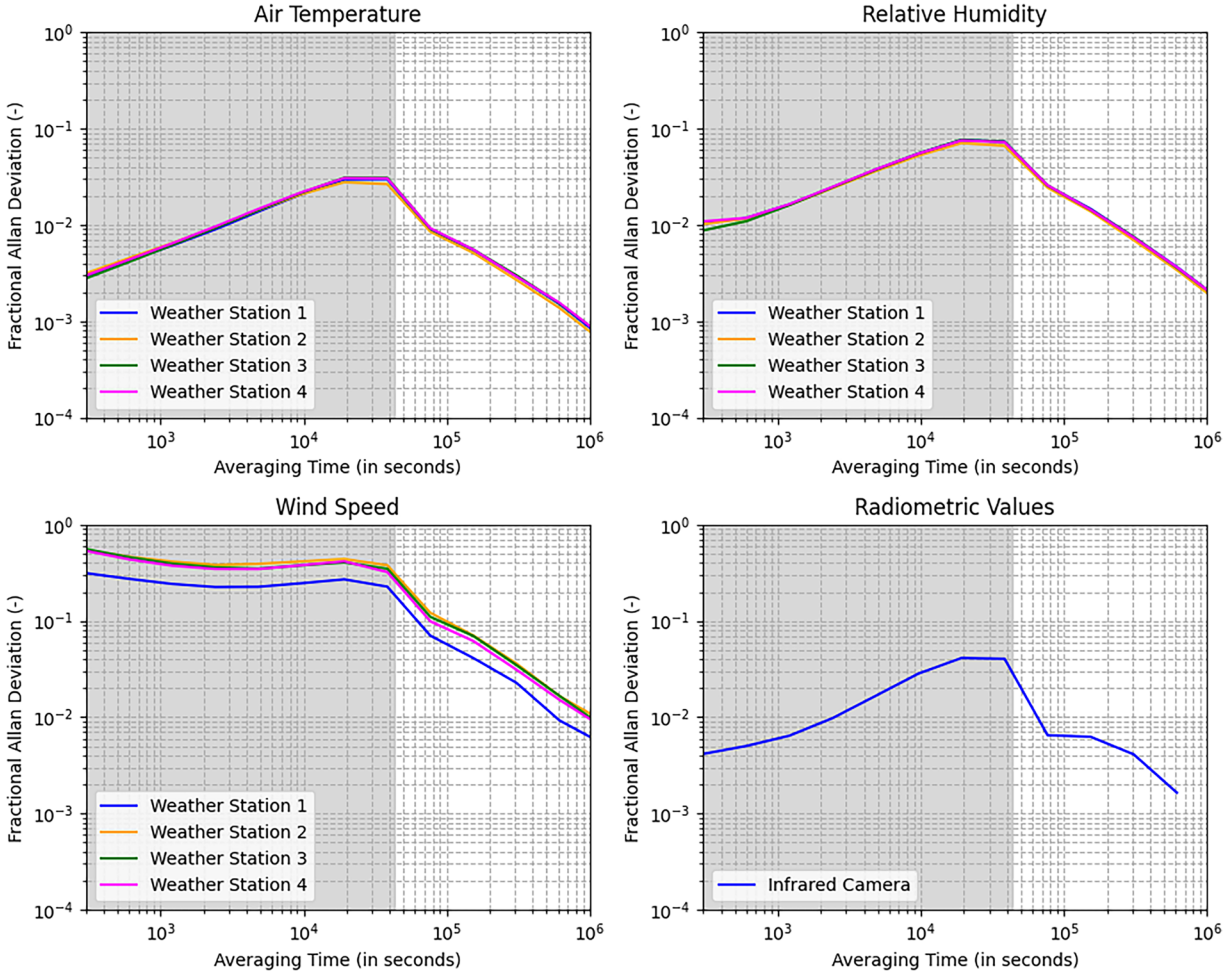


Fig. 6 Fraction Allan deviation calculated on residuals of weather data and microscale thermal images at the ground truth location. The gray region corresponds to the range of averaging time from 10 minutes to 12 hours, which is expected to be more dominated by the noise than the short-term drift.

- Wind speed is the most unstable of the weather variables measured. Indeed, measuring wind speed at 2 meters is more challenging compared to measuring air temperature and relative humidity. More advanced wind sensors, such as ultrasonic sensors, provide more stable measurements of low wind speeds than 3-cup anemometers do. However, these kinds of sophisticated sensors can hardly be incorporated into an all-in-one weather station.
- Microscale thermal images appear to have a temporal stability equivalent to that of air temperature data. This result implies that a proper analysis of factors contributing to and mitigating urban overheating can be performed using weather data and microscale thermal images; this capability is one of the main motivations for the MORICHI dataset.

Overall, both weather data and microscale thermal images seem to be temporally stable enough for analysis.

Calibration of the surface temperature. The purpose of calibrating the surface temperature assessed from microscale thermal images (T_S^{TIR}) is to minimize the discrepancy with ground truth measurements (T_S^{GT}). The discrepancy was estimated using the Root Mean Square Error (RMSE) and the Mean Bias Error (MBE), that is:

$$RMSE = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(T_S^{TIR}[n] - T_S^{GT}[n])^2} \tag{5}$$

$$MBE = \frac{1}{N}\sum_{n=1}^{N}(T_S^{TIR}[n] - T_S^{GT}[n]) \tag{6}$$

where $t_n = t_0 + n \cdot \Delta t$ corresponds to timestamps at which the surface temperature was measured by the contact surface sensor and assessed by microscale thermal images, and N the total number of timestamps.

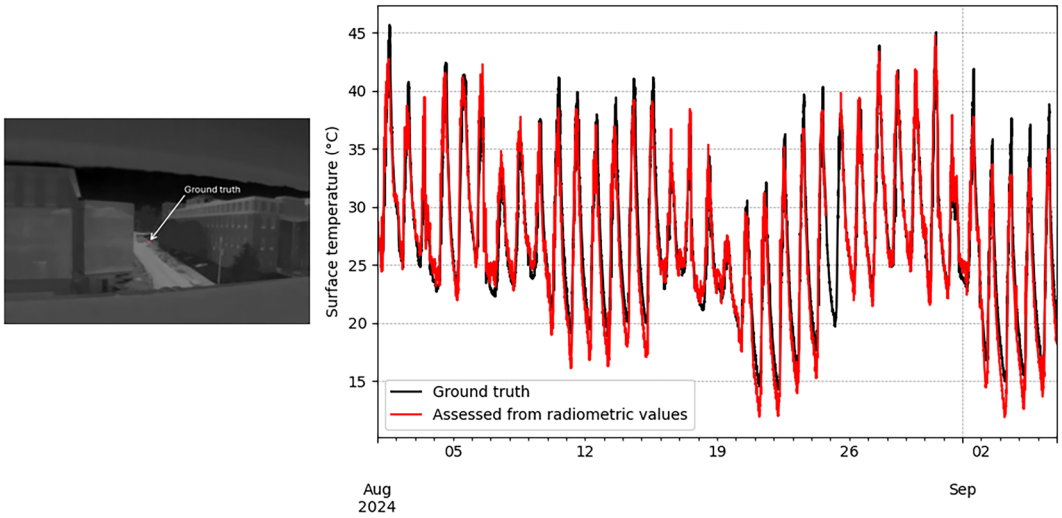


Fig. 7 Discrepancy between the surface temperature measured by the contact surface sensor (black) and this assessed from thermal images captured by the infrared camera (red).

Figure 7 shows the discrepancy between T_S^{GT} and T_S^{TIR} after calibration. The sidewalks where the contact sensor was placed are assumed to have an emissivity of 0.90, consistent with the value for concrete surfaces defined by Mandanici *et al.*[50]. Using Google Earth, the distance between the infrared thermal camera and the contact sensor was determined to be 115 meters. The outdoor temperature and humidity recorded by Weather Station 4 were used to calculate the air transmissivity between the infrared thermal camera and the contact sensor. After manually adjusting parameters b, f, O, r_1, and r_2, it was possible to achieve a RMSE of 1.66 degrees Celsius and a MBE of -0.42 degrees Celsius. This discrepancy was achieved under the following values:

$$b = 1396.6,\ f = 1.0,\ O = -1500.0,\ r_1 = 14911.1846,\ \text{and}\ r_2 = 0.0121 \tag{7}$$

Using these values for parameters b, f, O, r_1, and r_2, the discrepancy between the surface temperature from microscale thermal images and the ground truth at other positions is then expected to be similar to or lower than the RMSE of 1.7 degrees Celsius. As also observed by Martin *et al.*[26], the surface temperature derived from microscale thermal images is typically underestimated compared to the ground truth, particularly at night. Although the exact magnitude of the surface temperature may not be perfectly determined from microscale thermal images, their daily variation appears to be accurately captured over the measurement period.

Usage Notes

A detailed description on how to use the MORICHI dataset is accessible at https://github.com/tudelft3d/MORICHI. This public repository also contains a Python script, **install.py**, for installing the MORICHI dataset as illustrated in Figure 2. After installing the MORICHI dataset on a local computer or remote server, users can learn how to extract and process weather data and microscale thermal images with the provided Jupyter notebooks, such as **Extract_weather_data_using_python.ipynb** and **Analyze_thermal_images_using_python.ipynb**.

Data availability

The MORICHI dataset is available on the 4TU.ResearchData (https://data.4tu.nl), where it can be found by searching for “MORICHI". In this repository, the MORICHI dataset consists of a sequence of compressed files containing weather data and microscale thermal images.

Code availability

A Python module, **irim**, has been developed as part of the *sciencespy* library to facilitate the extraction and processing of microscale thermal images in the MORICHI dataset. The *sciencespy* Python library is publicly available under the GNU GPL license. It can be installed using the PIP package installer as:

Received: 2 April 2025; Accepted: 31 January 2026;
Published online: 10 February 2026

References

1. Santamouris, M., Cartalis, C., Synnefa, A. & Kolokotsa, D. On the impact of urban heat island and global warming on the power demand and electricity consumption of buildings-A review. *Energy and buildings* **98**, 119–124 (2015).
2. Li, X. *et al.* Urban heat island impacts on building energy consumption: A review of approaches and findings. *Energy* **174**, 407–419 (2019).
3. Singh, M. & Sharston, R. Quantifying the dualistic nature of Urban Heat Island effect (UHI) on building energy consumption. *Energy and Buildings* **255**, 111649 (2022).

4. Yang, M., Wang, H., Yu, C. W. & Cao, S.-J. A global challenge of accurately predicting building energy consumption under urban heat island effect (2023).
5. Phelan, P. E. *et al*. Urban heat island: mechanisms, implications, and possible remedies. *Annual Review of Environment and Resources* **40**, 285–307 (2015).
6. Chen, R. & You, X.-Y. Reduction of urban heat island and associated greenhouse gas emissions. *Mitigation and adaptation strategies for global change* **25**, 689–711 (2020).
7. Santamouris, M. Recent progress on urban overheating and heat island research. Integrated assessment of the energy, environmental, vulnerability and health impact. Synergies with the global climate change. *Energy and Buildings* **207**, 109482 (2020).
8. Hsu, A., Sheriff, G., Chakraborty, T. & Manya, D. Disproportionate exposure to urban heat island intensity across major US cities. *Nature communications* **12**, 2721 (2021).
9. Ramly, N. *et al*. Identifying vulnerable population in urban heat island: A literature review. *International Journal of Public Health Research***13** (2023).
10. Chen, S., Bruhn, S. & Seto, K. C. Trends in socioeconomic disparities in urban heat exposure and adaptation options in mid-sized US cities. *Remote Sensing Applications: Society and Environment* **36**, 101313 (2024).
11. Vargo, J., Stone, B., Habeeb, D., Liu, P. & Russell, A. The social and spatial distribution of temperature-related health impacts from urban heat island reduction policies. *Environmental Science & Policy* **66**, 366–374 (2016).
12. Heaviside, C., Macintyre, H. & Vardoulakis, S. The urban heat island: implications for health in a changing environment. *Current environmental health reports* **4**, 296–305 (2017).
13. Lowe, S. A. An energy and mortality impact assessment of the urban heat island in the US. *Environmental Impact Assessment Review* **56**, 139–144 (2016).
14. Roxon, J., Ulm, F.-J. & Pellenq, R.-M. Urban heat island impact on state residential energy cost and CO2 emissions in the United States. *Urban Climate* **31**, 100546 (2020).
15. Siu, C. Y. *et al*. Evaluating thermal resilience of building designs using building performance simulation–A review of existing practices. *Building and Environment* **234**, 110124 (2023).
16. Anderson, G. B. & Bell, M. L. Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 US communities. *Environmental health perspectives* **119**, 210–218 (2011).
17. Service, N. W. Nws pittsburgh climate page. https://www.weather.gov/pbz/climate. Accessed: 2025-03-07.
18. Tan, Y., Lipsky, E. M., Saleh, R., Robinson, A. L. & Presto, A. A. Characterizing the spatial variation of air pollutants and the contributions of high emitting vehicles in Pittsburgh, PA. *Environmental science & technology* **48**, 14186–14194 (2014).
19. Gentile, D. A., Morphew, T., Elliott, J., Presto, A. A. & Skoner, D. P. Asthma prevalence and control among schoolchildren residing near outdoor air pollution sites. *Journal of Asthma* **59**, 12–22 (2022).
20. Lange, C. L., Smith, V. A. & Kahler, D. M. Pittsburgh air pollution changes during the COVID-19 lockdown. *Environmental Advances* **7**, 100149 (2022).
21. Togami, C. & Murphy, M. W. The politics of airing grievances: an analysis of air quality knowledge and ignorance in Pittsburgh. *Local Environment* **29**, 446–459 (2024).
22. Ha, S., Talbott, E. O., Kan, H., Prins, C. A. & Xu, X. The effects of heat stress and its effect modifiers on stroke hospitalizations in Allegheny County, Pennsylvania. *International archives of occupational and environmental health* **87**, 557–565 (2014).
23. Cui, D., Liang, S. & Wang, D. Observed and projected changes in global climate zones based on köppen climate classification. *Wiley Interdisciplinary Reviews: Climate Change* **12**, e701 (2021).
24. Wong, C. H. M., Kwok, Y. T., He, Y. & Ng, E. Government-involved urban meteorological networks (umns): A global review. *Urban Climate* **61**, 102409 (2025).
25. Hicks, B. B., Pendergrass III, W. R., Vogel, C. A. & Artz, R. S. On the drag and heat of washington, dc, and new york city. *Journal of Applied Meteorology and Climatology* **53**, 1454–1470 (2014).
26. Martin, M., Ramani, V. & Miller, C. InfraRed Investigation in Singapore (IRIS) Observatory: Urban heat island contributors and mitigators analysis using neighborhood-scale thermal imaging. *Energy and Buildings* **307**, 113973 (2024).
27. Martin, M., Chong, A., Biljecki, F. & Miller, C. Infrared thermography in the built environment: A multi-scale review. *Renewable and sustainable energy reviews* **165**, 112540 (2022).
28. Dobler, G. *et al*. The urban observatory: a multi-modal imaging platform for the study of dynamics in complex urban systems. *Remote sensing* **13**, 1426 (2021).
29. Secret NYC. NYC Is Home To The World's Second Tallest Skyline. https://secretnyc.co/tallest-skyline-nyc/ [Online; accessed 2025-09-11] (2023).
30. Museum of the City of New York. Making The Plan - The Greatest Grid. https://thegreatestgrid.mcny.org/greatest-grid/making-the-plan/12. [Online; accessed 2025-09-11].
31. City of Pittsburgh. City of Pittsburgh, PA Mixed Use Zoning Districts. https://ecode360.com/45474194#45474224 [Online; accessed 2025-09-11] (2025).
32. City of Cincinnati. The Cincinnati Zoning Code. https://www.cincinnati-oh.gov/cincyzoning/zoning-code-text/ [Online; accessed 2025-09-11] (2025).
33. City of Sharonville, Ohio. Engineering Standards. https://www.sharonville.org/DocumentCenter/View/113/Engineering-Standards-PDF [Online; accessed 2025-09-11] (2025).
34. City of Milwaukee. Milwaukee Code of Ordinances, Chapter 295: Zoning. https://city.milwaukee.gov/DCD/Planning/Zoning/Code [Online; accessed 2025-09-11] (2025).
35. City of Milwaukee Department of Public Works. Standard Street Cross Sections. https://city.milwaukee.gov/mpw/divisions/infrastructure/Coordination/Standard-Street-Cross-Sections.pdf [Online; accessed 2025-09-11] (2025).
36. Yan, D., Wu, Y., Malik, J. & Hong, T. Ten questions on future and extreme weather data for building simulation and analysis in a changing climate. *Building and Environment* 112461 (2024).
37. Ren, Z. A literature review on the use of weather data for building thermal simulations. *Energies* **18**, 3653 (2025).
38. Machard, A. *et al*. Typical and extreme weather datasets for studying the resilience of buildings to climate change and heatwaves. *Scientific data* **11**, 531 (2024).
39. Toparlar, Y., Blocken, B., Maiheu, B. & van Heijst, G. J. F. A review on the cfd analysis of urban microclimate. *Renewable and Sustainable Energy Reviews* **80**, 1613–1640 (2017).
40. Mirzaei, P. A. Cfd modeling of micro and urban climates: Problems to be solved in the new decade. *Sustainable Cities and Society* **69**, 102839 (2021).
41. Wang, H., Yang, J., Chen, G., Ren, C. & Zhang, J. Machine learning applications on air temperature prediction in the urban canopy layer: A critical review of 2011–2022. *Urban Climate* **49**, 101499 (2023).
42. Wang, Z. *et al*. A review of vehicle detection techniques for intelligent vehicles. *IEEE Transactions on Neural Networks and Learning Systems* **34**, 3811–3831 (2022).
43. Martin Fehlmann, M., Berges, M., Stoter, J. & Garcia Sanchez, C. Meteorological observations and radiometric images for cold climate heatwave investigations (morichi), https://doi.org/10.4121/c5fb8062-af5d-491a-8faf-8be7fc390f48.v1 (2025).
44. Pittsburgh Perspectives. Profile of college students in the pittsburgh region (2023).
45. Carnegie Mellon University. Common data set 2024-2025. Tech. Rep., Carnegie Mellon University (2025).

46. Microsoft Buildings Footprint Training Data with Heights. https://www.arcgis.com/home/item.html?id=3b0b8cf27ffb49e2a2c8370f9806f267 Accessed: 2025-01-14. (2019).
47. Instrument, D. What is the best height at which to install my weather station? https://support.davisinstruments.com/article/gbuon361gs-what-is-the-best-height-at-which-to-install-my-weather-station. Accessed: 2025-01-14.
48. TU Delft. Manage confidential data: Personal data. https://www.tudelft.nl/en/library/data-management/research-data-management/manage-confidential-data-personal-data (2025). Accessed: 2025-10-02.
49. Donoho, D. L. & Johnstone, I. M. Ideal spatial adaptation by wavelet shrinkage. *biometrika* **81**, 425–455 (1994).
50. Mandanici, E., Conte, P. & Girelli, V. A. Integration of aerial thermal imagery, LiDAR data and ground surveys for surface temperature mapping in urban environments. *Remote Sensing* **8**, 880 (2016).

Acknowledgements

This research has received funding from the European Union's Horizon Research and Innovation programme under the Marie Sklodowska-Curie 2021 grant agreement No 101059484. We also wish to address our sincerest gratitude to Mr. Steven Guenther, Mr. Robert Penn, and Mr. Jason Sturges from the Facility Management team of Carnegie Mellon University, without whom the presented data could not have been collected. This work was supported, in part, by the National Aeronautics and Space Administration (NASA) as part of the Space Technology Research Institute (STRI) Habitats Optimized for Missions of Exploration (HOME) 'SmartHab' project [grant number 80NSSC19K1052].

Author contributions

M.M. conducted the field experimental campaign, formatted the data of MORICHI, and deployed the dataset in the 4TU.ResearchData repository implemented the *irim* module in the *sciencespy* Python library and wrote the original draft of the paper, C.G.S. supervised the field experiments, managed administrative aspects of the project, and revised the original draft of the paper, J.S. managed administrative aspects of the project and revised the original draft of the paper, M.B. provided the financial support to purchase instruments, supervised the field experiments, connected with Carnegie Mellon library to have the publication in open access, and revised the original draft of the paper.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.