



OPEN

DATA DESCRIPTOR

District-scale surface temperatures generated from high-resolution longitudinal thermal infrared images

Subin Lin¹, Vasantha Ramani¹, Miguel Martin¹, Pandarasamy Arjunan^{1,2}, Adrian Chong³, Filip Biljecki^{4,5}, Marcel Ignatius⁴, Kameshwar Poolla⁶ & Clayton Miller³✉

This paper describes a dataset collected by infrared thermography, a non-contact, non-intrusive technique to acquire data and analyze the built environment in various aspects. While most studies focus on the city and building scales, an observatory installed on a rooftop provides high temporal and spatial resolution observations with dynamic interactions on the district scale. The rooftop infrared thermography observatory with a multi-modal platform capable of assessing a wide range of dynamic processes in urban systems was deployed in Singapore. It was placed on the top of two buildings that overlook the outdoor context of the National University of Singapore campus. The platform collects remote sensing data from tropical areas on a temporal scale, allowing users to determine the temperature trend of individual features such as buildings, roads, and vegetation. The dataset includes 1,365,921 thermal images collected on average at approximately 10-second intervals from two locations during ten months.

Background & Summary

Urban ecosystems are the biggest, most dynamic, and most complicated man-made systems, with millions of people interacting and hundreds of governing agencies^{1,2}. Urban modernization has encouraged a significant portion of the global population to move to urban regions. Consequently, the built environment has grown rapidly, and building energy consumption has increased, which augments the release of CO₂^{3,4}. Improving energy efficiency has been urged by various sectors⁵. The scientific community has exerted significant efforts to enhance its understanding of the built environment by leveraging a wide range of sensing technologies⁴. Furthermore, the advent of computational techniques has revolutionized data collection and analysis, enabling us to process vast quantities of information¹.

Infrared thermography has been widely used in the built environment for many research purposes, such as urban heat island^{6,7}, building diagnostics^{5,8,9}, and urban heat fluxes¹⁰. Infrared thermography can provide images showing the surface temperature of different elements in the built environment with a lower cost and effort compared to similar kinds of sensor networks¹¹. It detects infrared electromagnetic radiation emitted by the inspected objects and is commonly used for non-destructive testing and non-contact diagnostic technology^{5,8}. It is based on an infrared imaging system calibrated to measure surfaces' emissive power distribution at various temperature ranges. The infrared camera generates a series of two-dimensional, readable thermal images, with different colors and tones representing different temperatures^{5,8}. Infrared systems have been leveraged to analyze the built environment across various scales: Satellites, capable of wide-ranging data collection, are

¹Berkeley Education Alliance for Research in Singapore, CREATE Tower 1 Create 6 Way, 138602, Singapore, Singapore. ²Robert Bosch Centre for Cyber-physical Systems, Indian Institute of Science, Bengaluru, Karnataka, 560012, India. ³Department of the Built Environment, College of Design and Engineering, National University of Singapore, 4 Architecture Drive, 117566, Singapore, Singapore. ⁴Department of Architecture, College of Design and Engineering, National University of Singapore, 4 Architecture Drive, 117566, Singapore, Singapore. ⁵Department of Real Estate, Business School, National University of Singapore, 15 Kent Ridge Drive, 119245, Singapore, Singapore. ⁶Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, CA, USA. ✉e-mail: clayton@nus.edu.sg

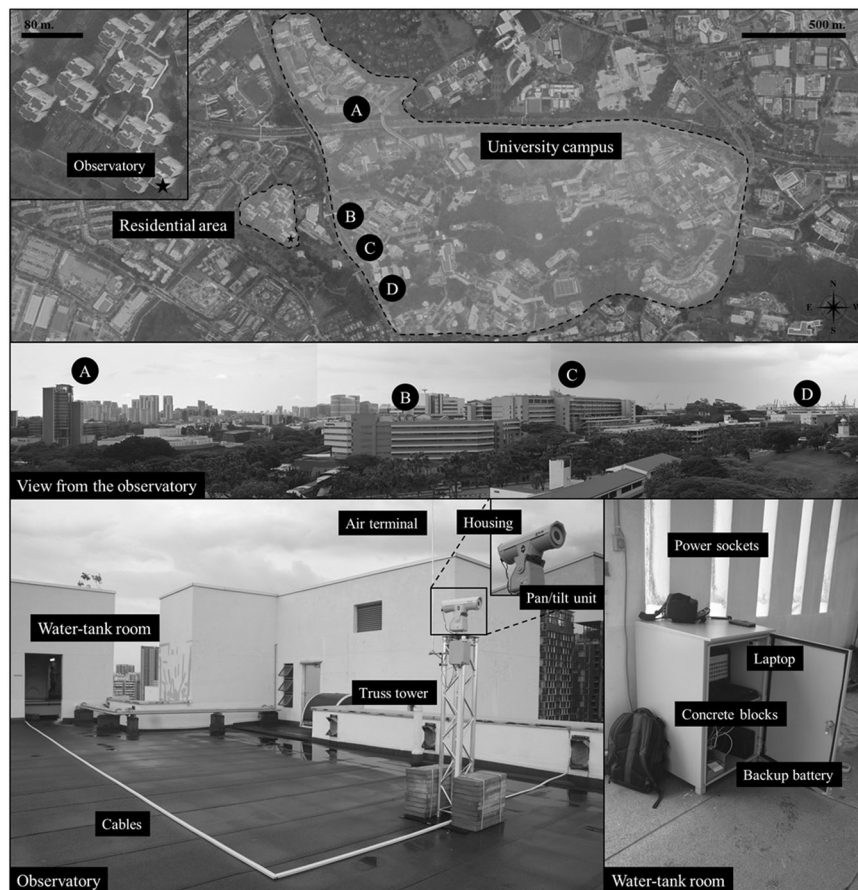


Fig. 1 Observatory installed on the rooftop of the residential area in Singapore and its captured information. Source of the imagery: Google Earth²⁰.

primarily used for city-scale analyses. This typically covers entire cities or large urban regions. Aerial vehicles, possessing a more focused reach, are frequently employed for investigations at both city and district scales. The term *district scale* indicates a specific urban district or neighborhood. It refers to a geographical area larger than a single building or block but smaller than an entire city or region. This level of analysis allows for considering the collective impact of multiple buildings and infrastructure elements within a defined urban district. Due to their unique vantage point, rooftop observatories are particularly effective for district-scale studies. These offer a detailed view of the urban microclimate within a specific district or neighborhood. Lastly, drones and handheld devices are used for detailed investigations at smaller scales. Their portability and close-range capabilities make them particularly suited for analyses at the building scale (focusing on individual buildings) and human scale (focusing on the immediate surroundings of individuals or small groups)¹¹. Many previous and current studies focus on the city and building scales and could not show the dynamic interactions on the district scale with high-resolution^{1,4}. Furthermore, urban climate studies have been focusing on the temperate climate zones, especially in North America and Europe, and there is a lack of such studies focusing on the tropical climate zones^{12–14}.

This work introduces the first thermal observatory deployed in Singapore, a tropical city, with a multi-modal platform with flexibility and high resolution to assess a wide range of dynamic processes in urban systems. The thermal observatory was set up on the roof of buildings that overlooked several educational buildings on the National University of Singapore campus. The platform gathers remote sensing data from a tropical area at a high temporal resolution at approximately 10-second intervals, providing the temperature trends of specific urban elements such as buildings, roads, and vegetation. Demonstration codes were given with data preprocessing, such as segmentation, in order to handle and analyze the generated raw data and allow users to use the data with high flexibility, catering to their research needs.

Methods

Thermal images data collection. Thermal images are taken from a FLIR A300 (9 Hz) thermal camera atop an urban-scale IR observatory, as shown in Fig. 1. The detailed specifications of the thermal camera are listed in Table 1. The thermal camera is protected by a housing against rains with IP67 protection. It was attached to a pan/tilt device that can spin 360 along its horizontal axis to record thermal images in multiple locations. These structures (thermal camera, housing, pan/tilt) were placed on a 2-m-high truss tower to avoid impediments while taking thermal images. Concrete blocks serve as a stabilizing support, and an air terminal serves as lightning

Image resolution	16 bit, 320 × 240 pixels
Sensor	Uncooled Microbolometer FPA
Thermal sensitivity	50 mK @ 30 °C
Spectral range	7.5 - 13 μ m
Field of view (FOV)	25° (H) and 18.8° (V)
Accuracy	\pm 2° or 2% of reading
Camera Weight	0.7 kg
Power supply	110/220 V AC
Size	170 mm × 70 mm × 70 mm

Table 1. Information about the FLIR A300 thermal camera.

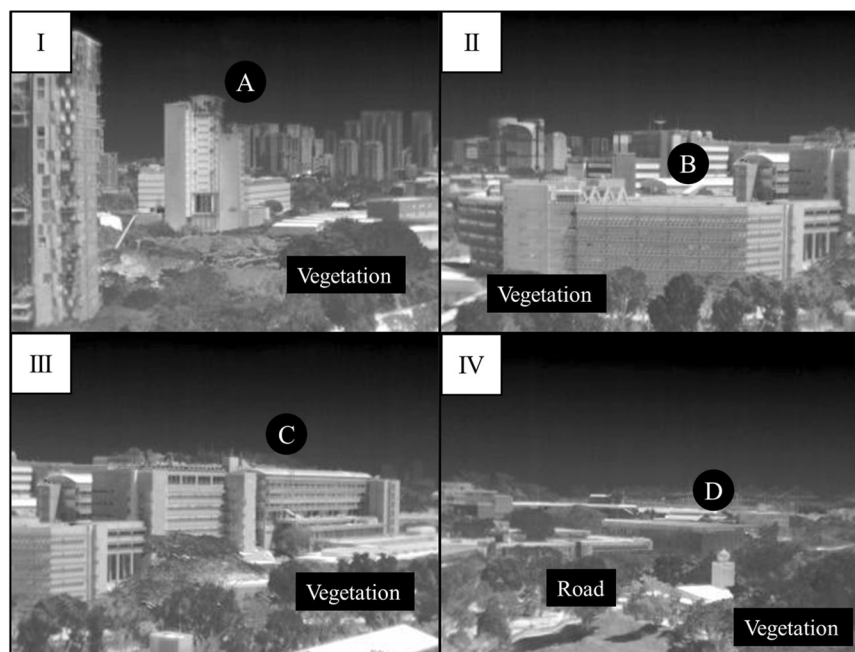


Fig. 2 Positions where the observatory capture thermal images on Kent Vale²⁰.

protection for the truss tower. Two plugs were added to the truss tower to allow a backup battery in the water tank room to power both the thermal camera and the motor of the pan/tilt device. The backup battery is continually charged from the building's electrical supply to enable the pan/tilt unit and thermal camera to be operational for up to two hours during a power outage. A laptop with a weatherproof casing in the water tank room was also linked to the pan/tilt device and thermal camera for customizing and reviewing the collection of thermal images.

The thermal images are collected from two separate locations, namely Kent Vale and S16. The Kent Vale observatory is located on the rooftop of a 42-m-tall building in a residential area, which is located in front of a university campus consisting of office and educational buildings, in order to access the district-scale analysis. The pan/tilt unit was configured from two separate software to allow the thermal camera to take images from four positions (I, II, III, and IV), as shown in Fig. 2. The first software was installed in a video encoder to control the positions of the pan/tilt unit to take a thermal image. The second software developed by NAX Instruments Pte Ltd was installed in the laptop to control the moment to take a thermal image, which could then be saved either in JPEG or FFF file format. The observatory can capture thermal images of four buildings, as shown in Fig. 1. Building A, known as CREATE, is centrally located at Position I and at a height of 68 meters. Curtain walls cover a significant portion of its facade. Buildings B (E1A) and C (EA) are each centered at positions II and III, respectively. Both buildings are roughly 27 meters tall, characterized by their concrete walls and single-pane windows. Building D, or SDE4, can be observed at Position IV. It is structured with metal grids mounted on a concrete frame and is designed to be a net-zero building. It stands at a height of approximately 24 meters. Surface temperatures of trees are also observed from the above positions. A road with traffic exists in front of Buildings B, C, and D, and the road can be observed in thermal images taken from Position IV. The observatory moved from Position I to IV sequentially and recorded thermal images at various intervals. The detailed time for the image taken is shown in its file name. Then, using a 4G Internet connection fitted on the laptop, the captured thermal images are stored on the Google Drive repository.

The S16 observatory is located on the rooftop of a nine-story building in an educational area on a university campus. The pan/tilt unit was configured to capture images from three distinct targets, denoted as Positions 1,

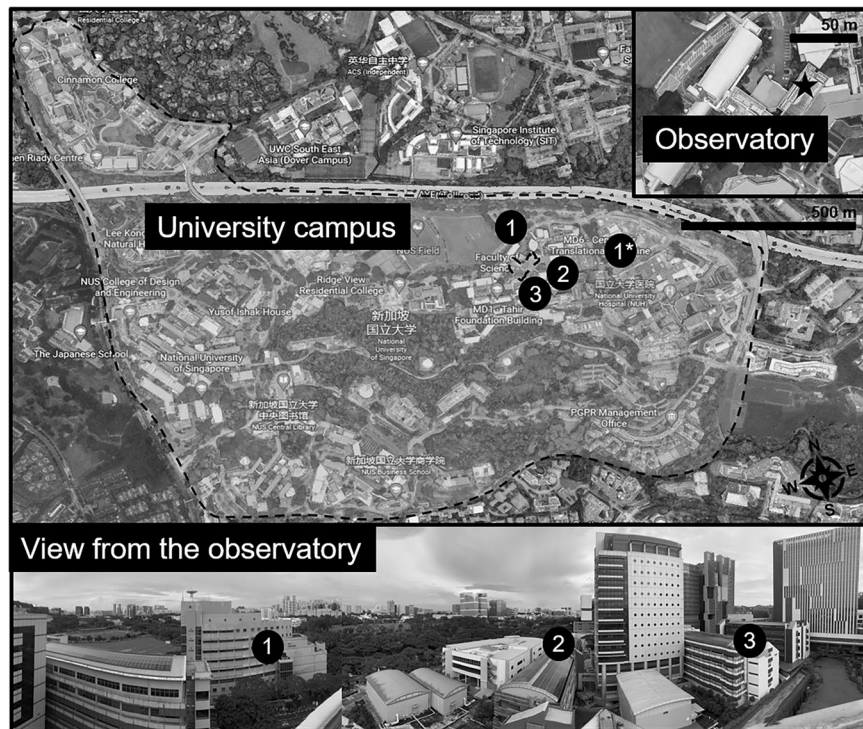


Fig. 3 Observatory installed on the rooftop of the university campus in Singapore and its captured information. Source of the imagery: Google Earth.

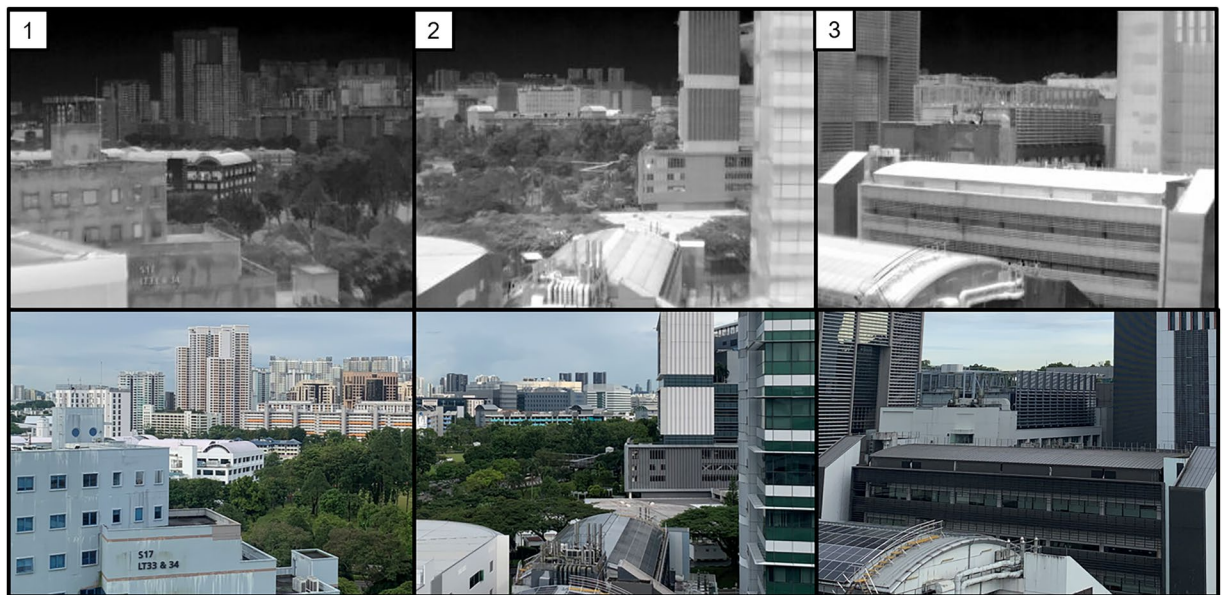


Fig. 4 Positions where the observatory capture thermal images on S16.

2, and 3 in Fig. 3, sequentially at different time intervals. The observatory can capture information from three different buildings with thermal images (Fig. 4). The observatory moved from Position 1 to 3 sequentially and recorded thermal images at various time intervals per view. The thermal images obtained in this study elucidate a range of elements, including air conditioning units, glass materials, vegetation, and solar panels. These elements can be systematically analyzed to attain an in-depth understanding of their thermal behavior and consequential impact on the ambient environment. Specifically, this research is poised to investigate the influence of these elements on local temperature variations, energy consumption patterns, and the phenomena of urban heat islands. These insights carry significant implications for urban planning and the development of efficient energy management strategies.

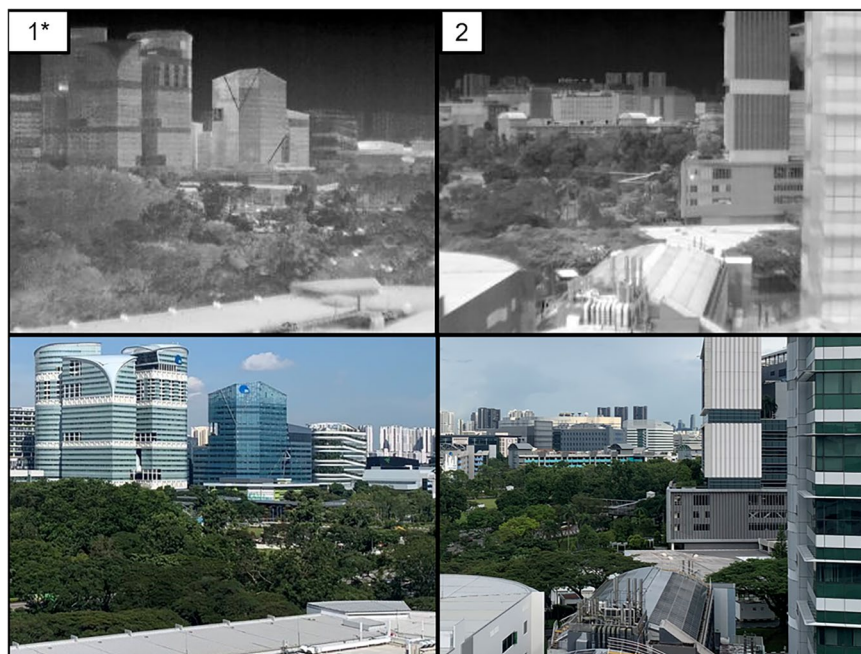


Fig. 5 Positions where the observatory capture thermal images on S16 after adjustment.

Due to the necessary adjustments to the second observatory's positioning system, thermal images at the S16 site were not consistently captured from the exact same locations—slight variations to the left or right were noted. Consequently, automated segmentation tools are strongly recommended to account for these minor positional differences. It is acknowledged that these variations may impact the interpretation of temporal patterns in the data. Future enhancements to the observatory system are intended to improve the stability of image capture, leading to more consistent data collection. Furthermore, the positions from which images were captured were altered due to network issues with the observatory system on both Kent Vale and S16. In an effort to achieve greater stability in the image capture locations on S16, measures were taken to ensure these positions were closely aligned with each other. Figure 5 depicts this more stable configuration.

The thermal camera comprises an optical system that directs radiation from the scene onto a microbolometer¹⁵. Long-wave infrared radiation changes the resistance of the detector, which is translated to apparent temperature (T_{obj}) readings as follows¹⁵:

$$T_{obj} = \frac{B}{\ln\left(\frac{R_1}{R_2(U_{obj} + O)} + f\right)} \quad (1)$$

where U_{obj} is the total signal response to the incident long-wave infrared radiation on the detector. This response is influenced by the material's emissivity, denoted by ε , and the current environmental conditions. For a comprehensive computation of the influence factors for U_{obj} , please refer to Eq. 6. The other terms, B , R_1 , R_2 , O , and f , are calibration constants determined by the camera manufacturer in a controlled environment and are stored in the thermal image metadata. However, these parameters usually need to be re-calibrated in an outdoor environment where a significant discrepancy can be observed between U_{obj} and T_{obj} . The re-calibrated parameters are presented in Table 2.

Emissivity (ε) measures a material's ability to emit infrared energy. Different materials have different emissivities, which can significantly impact the surface temperature measurements. Therefore, accounting for the material's emissivity under observation is crucial when processing thermal images. The dataset is presented in a radiometric format to preserve the raw thermal data, facilitating necessary corrections based on specific emissivity values. The total signal response U_{obj} is also affected by ambient conditions, such as air temperature and humidity. By incorporating local weather data into the calculations, these influences can be corrected, and the accuracy of the surface temperature measurements can be improved. Please refer to the Usage Notes section for further information.

Data preprocessing. Data preprocessing was done before analyzing the images for the Kent Vale dataset. First, data filtering was conducted to manually remove images unsuitable for analysis. Thermal images captured during the rainy period were disturbed by rain droplets and needed removal. Then, a convolution neural network (CNN) classification model was applied to exclude blurred images taken while the pan/tilt unit was moving.

Constant	Original parameters	Calibrated parameters
B	1396.6000	1425.0000
f	1.0000	1.0000
O	-6303.0000	-9550.0000
R1	14911.1846	14911.1850
R2	0.0108	0.0120
Position	RMSE (in °C)	MBE (in °C)
A	1.24	-0.21
B	1.69	0.91
C	1.00	0.55

Table 2. Calibration parameters of the measurement[20].

The CNN model is also used to classify the building type for the first location. The model was trained on 4316 pre-classified images and then used to classify thermal image datasets of Kent Vale.

Weather station data collection. The weather station locations are shown in Fig. 6. The weather stations measured meteorological data in multiple locations on the university campus around 2 meters above the ground, except for Locations 1 and 7 on rooftops. The six recorded parameters are air temperature (°C), relative humidity (RH, %), dew point (°C), wind speed (m/s) and direction (degree), gust speed (m/s), and solar radiation (W/m²) at a 1-minute interval. For location 9, only the first three parameters were recorded. The weather stations were calibrated before installation. The weather station specifications are listed in Table 3. The weather station installment details are available in Chen *et al.*¹⁶, and Yu *et al.*¹⁷.

Data Records

The thermal image and weather station data for the same periods are published in Zenodo with open access¹⁸. The thermal images dataset contains two locations: Kent Vale and S16. A set of 483,915 thermal images were taken from Kent Vale between November 8, 2021, and March 8, 2022, while 882,006 thermal images were taken from S16 between August 3, 2022, and December 14, 2022. The images taken time are indicated in their file names. The data were converted and stored in JPEG format. The thermal images were stored in folders according to locations and views and listed sequentially. The detailed hierarchy of the data folders is listed in the data/README file on GitHub (<https://github.com/buds-lab/project-iris-dataset>) and Fig. 7. Missing values are intended, as the preprocessing process filters unsuitable images for analysis. The CNN model classified the images taken at Kent Vale into five different groups: CREATE (I), E1A (II), EA (III), SDE4 (IV), and Corrupted. The Corrupted group collected images that were not classified into any of the above four views, which means that these collected images were of low quality and unable to be distinguished; hence, they needed to be excluded. The images taken at S16 were not entirely classified; while some were divided into three different viewpoints (view_1, view_2, and view_3), some were not, requiring further classification. The summary of the provided dataset is shown in Table 4. In Zenodo, the Kent Vale dataset is stored in four independent zip files in sequential order, while the S16 dataset is stored in six independent zip files in sequential order. In each subcategory, the data is zipped in days. The code structure demonstration is also available in the Readme file in the GitHub repository. Examples of the data folder hierarchy include: 2022-08-03/view_1/smap-2022-08-03T13-48-18.19.jpeg; 20220306/create/20220306000001.jpeg. The code demonstration contains three subcategories: data, notebook, and src. The data directory shows all the original and converted/processed data; The notebook directory includes three notebooks containing three preprocess steps; and the src directory runs preprocessing scripts. A more detailed description of the code file organization and explanation is available in the Code Availability section. The weather station data is in Excel format and stored in Zenodo as WS_data.zip. It contains some missing values; therefore, prior data preprocessing is needed.

Technical Validation

Sensitivity analysis. Before evaluating the infrared receptor (T_{ij}) from the observatory, a sensitivity analysis on factors that might significantly impact its variance was carried out. First-order Sobol index¹⁹ was used to estimate the contribution of a variable to the variance of T_{ij} . The total Sobol index was also computed to consider the variables' interactions.

During the sensitivity analysis, the variables and their corresponding constraints are measured and shown in Table 5²⁰. The Satteli sampler²¹ on thermal images taken at the four positions during a sunny day in Singapore was used to calculate the first order and total Sobol indices. T_{ij} was determined for each sample at various times of the day and thermal camera placements. The sensitivity of each variable at different periods was given using the mean indices across the four positions.

Parameters calibration. The thermal images' surface temperature estimates were calibrated in accordance with the information gathered by contact surface sensors, as shown in Fig. 8 at three different positions. Between December 2021 and March 2022, a heat flux (HF) sensor and a temperature probe were installed at Position *a* on the surface of a window in Building B. The HF sensor and the temperature probe were connected to a Hioki data logger and measured the surface temperature every 3 seconds. Another type of data logger, the UbiBot WS1 indoor monitoring sensor, was used to connect temperature probes at position *b* and *c*, respectively.

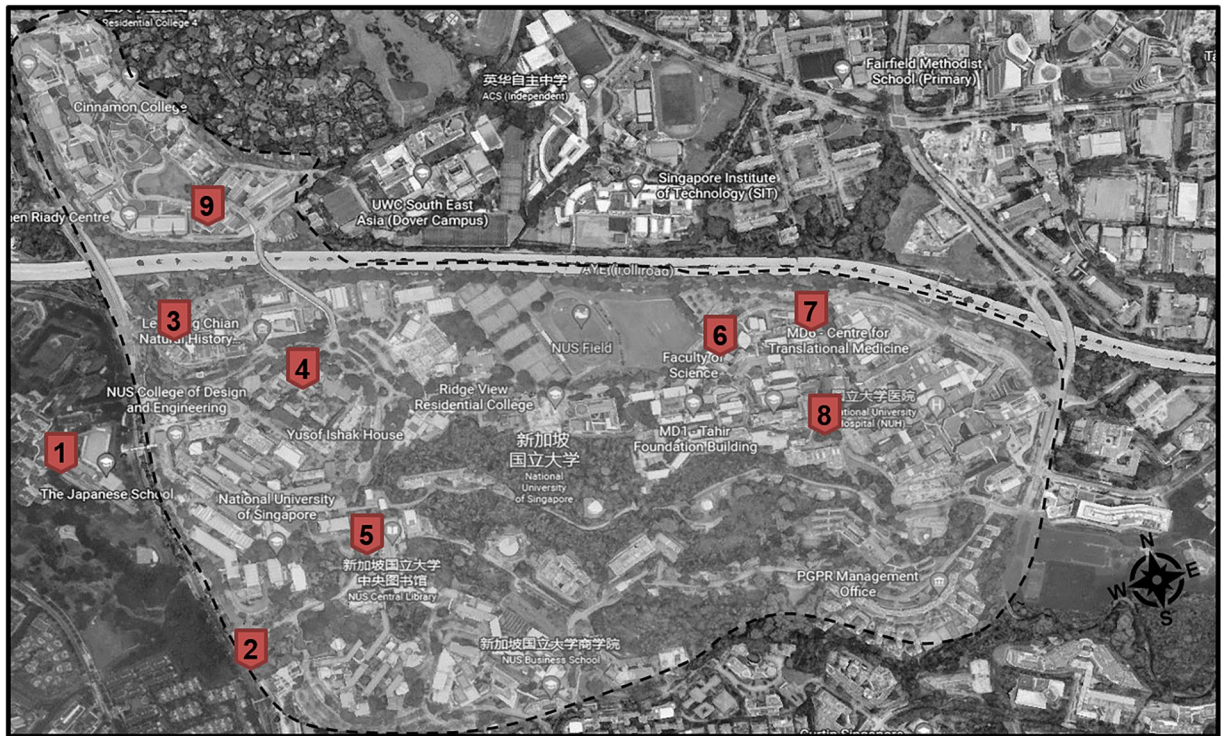


Fig. 6 Weather station locations^{16,17,43}.

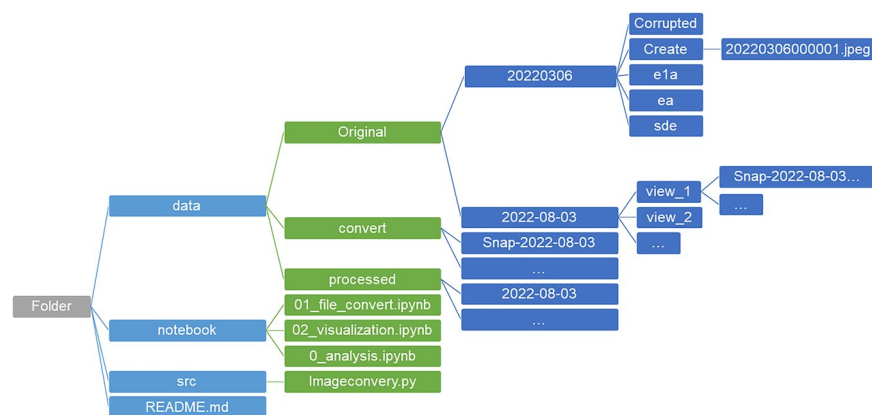


Fig. 7 File structure demonstration example for the dataset.

Duration	Parameter	Temperature	Solar Radiation	Wind Speed
01/11/2021-08/03/2022 01/08/2022-14/12/2022	Range	-40 °C-75 °C	0-1280 W/m ²	0-50 m/s
	Accuracy	±0.21 °C	±10 W/m ²	0.2 m/s
	Resolution	0.01 °C	1.25 W/m ²	0.2 m/s

Table 3. Specifications of weather station sensors¹⁷.

The output voltage recorded by the infrared receptor at position ij in the thermal image is represented by an array U_{ij}^r and a header in these files. Equation 2 can be used to transform U_{ij}^r into longwave radiations (L_{ij}^r) between 7.5 and 13 micrometers:

$$U_{ij}^r = cL_{ij}^r \tag{2}$$

where the term c is not truly constant but depends on the emissivity of the surface material being observed, with lower emissivities reflecting more radiation from other sources.

Duration	Resolution (sec)	Locations	Views	Examples	Corrupted size	Dataset size
08/11/2021-08/03/2022	vary	Kent Vale	I	72036	100758	483915
			II	96289		
			III	127506		
			IV	87326		
03/08/2022-14/12/2022	vary	S16	no classify	/	/	882006

Table 4. Summary of the provided datasets at each location: duration of each location, number of views per group, the number of examples per view, corrupted size per view and dataset size.

Variable	Minimum value	Maximum value
Thermal emissivity (–)	0.7	1.0
Sky temperature (°C)*	11.0	33.0
Air temperature (°C)**	19.0	37.0
Air relative humidity (%)**	30.0	100.0
Distance object (m)	100.0	1000.0
Window temperature (°C) ⁺	20.0	60.0
Window transmittance (–)	0.8	1.0

Table 5. Constrains of each variable considered during the sensitivity analysis of T_{ij} . *The sky temperature ($T^{sky} = g(U^{sky})$) was measured by Miguel *et al.*²⁰. **Weather data recorded by the Meteorological Service of Singapore³². ⁺Measured from a HOBO UX100-014M sensor

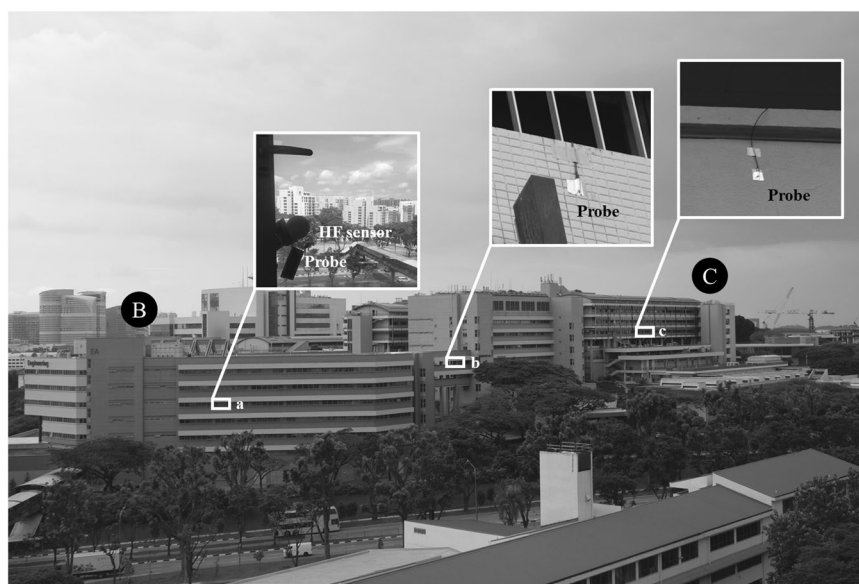


Fig. 8 Locations where sensors were placed in to calibrate thermal images and measure the surface temperature²⁰.

The FLIR A300 (9 Hz) thermal camera that was installed in the observatory was calibrated such that the following Eq. 3 holds true for the surface temperature of a target element detected by the infrared receptor (T_{ij}^r):

$$T_{ij}^r = g(U_{ij}^r) = b \ln \left[\frac{r_1}{r_2(U_{ij}^r + O)} + f \right]^{-1} \quad (3)$$

where b , r_1 , r_2 , O , and f are calibrated parameters.

However, the parameters in Eq. 3 need to be calibrated in order to minimize the discrepancy between T_{ij}^r and T_{ij} in an outdoor environment. In this case, these values are acquired in a controlled setting where the target element's actual surface temperature T_{ij} at location ij in the thermal picture is almost identical to T_{ij}^r .

The FLIR A300 camera was calibrated by manually tuning the parameters in Eq. 3 until there was a satisfactory agreement between surface temperature estimations and observations. The agreement was assessed in terms of the Mean Bias Error (MBE) and Root Mean Square Error (RMSE) as shown in Eqs. 4 and 5:

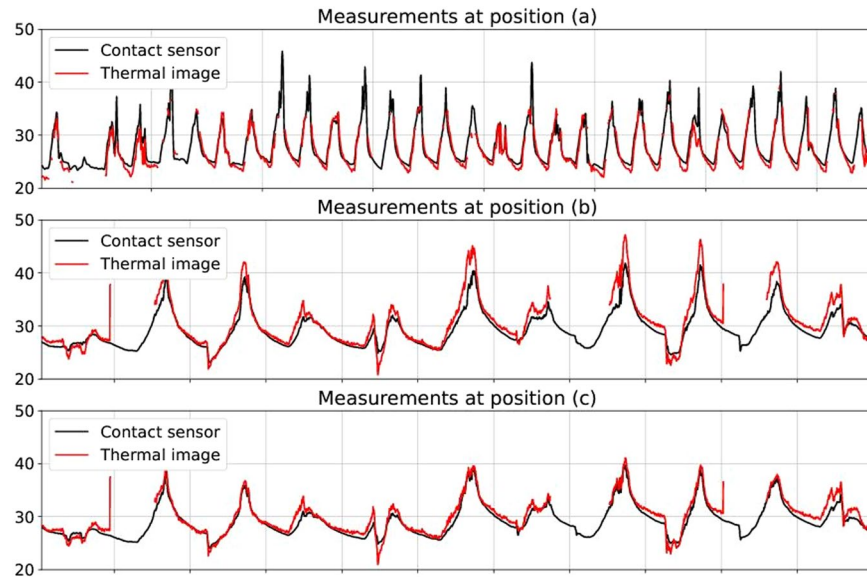


Fig. 9 Comparison of estimated and measured surface temperatures after calibration of the A300 infrared camera at the observatory²⁰.

$$MBE = \frac{1}{n} \sum_{i=1}^n (T_S^n - T_{S,m}^n)^2 \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (T_S^n - T_{S,m}^n)^2} \quad (5)$$

where $T_{S,m}^n$ is the surface temperature obtained by contact surface sensors, and T_S^n is the surface temperature calculated from thermal pictures captured by the FLIR A300 camera at time $t = t_0 + n \cdot \Delta t$. The calibration aims to achieve an MBE as close to 0 with the lowest RMSE. The RMSE and MBE were calculated with different Δt at three different positions to optimize the temporal resolution for each location (30 minutes at Position A and 5 minutes for Positions B and C). At Positions A-C, the RMSE and MBE obtained after calibration are shown in Table 2, achieving an RMSE below 2 degrees Celsius and an MBE below ± 1 degree Celsius. Furthermore, Fig. 9 illustrates a comparison between the estimated and measured surface temperatures after calibration at Positions A-C, offering additional information that enables the data to be freely accessed and tailored to suit specific needs and interests in further research.

Usage Notes

Factors consideration. When evaluating T_{ij} from the rooftop observatory, several factors must be considered. First, the buildings, streets, and trees that could be seen from the observatory are exposed to longwave radiations mainly coming from the skydome. Then, in tandem with the longwave radiation released by the element (L) being observed from the observatory, the longwave radiation from the sky (L^{sky}) is reflected into the air. Both L^{sky} and L pass through the atmosphere before arriving at the observatory, then together with the longwave radiation emitted by the atmosphere (L^{atm}), transmitting through the window of the housing before reaching the infrared receptor. Along with other types of radiation, the infrared receptor also captures longwave radiation from the window (L^{win}). The camera output voltage corresponding to U_{ij} combining the relationship defined in Eq. 3 and the aforementioned factors could be expressed as the following equation:

$$U_{ij} = \frac{1}{\varepsilon_{ij} \tau_{ij}^{atm} \tau^{win}} U_{ij}^r - \frac{1 - \varepsilon_{ij}}{\varepsilon_{ij}} U^{sky} - \frac{1 - \tau_{ij}^{atm}}{\varepsilon_{ij} \tau_{ij}^{atm}} U^{atm} - \frac{1 - \tau^{win}}{\varepsilon_{ij} \tau_{ij}^{atm} \tau^{win}} U^{win} \quad (6)$$

where ε_{ij} is the thermal emissivity from the element at position ij in the thermal image, τ_{ij}^{atm} is the transmissivity of the atmosphere between the element and the rooftop observatory, and τ^{win} is the transmissivity of the window.

T_{ij} could be calculated through Eq. 3 if U_{ij} is known. ε_{ij} and τ^{win} could be estimated from the material properties of the element and the window, respectively. τ_{ij}^{atm} can be calculated from weather data collected from a weather station.

Considering building material properties is paramount in analyzing heat fluxes, as various materials exhibit distinct thermal behaviors. Specifically, Building A's facade comprises two steel walls and one glass wall. In contrast, Buildings B and C predominantly feature concrete walls, while Building D's facade uniquely incorporates

Portion	Emissivity (0-1)	Density (kg/m ³)	Specific heat (J/kg-K)	Thickness (cm)	Thermal conductivity ((W/m-K))	Reference
Steel	0.88	7940	507	10	—	4,33
Glass	0.93	2500	836	2.5	0.974	34–37
Concrete	0.90	2400	1180	30	0.62	4,38–40
Portion	Emissivity (0-1)	leaf area index (m ² /m ²)				Reference
Vegetation (Tropical trees)	0.98	4				41,42

Table 6. Thermal properties of building facades and vegetation²⁰.

steel and concrete walls. The diversity in materials necessitates a detailed understanding of their emissive properties, which are critical in accurately calculating temperatures. Any oversight in this regard could potentially lead to misleading results. Therefore, the thermal properties, including the corresponding emissivities of the building facades and vegetation, have been meticulously defined. The thermal properties of the building facades and vegetation, including the corresponding emissivities, are defined and presented in Table 6, which has been constructed based on the relevant literature.

Weather stations data. The weather information obtained from the weather stations could be used together with the thermal images gathered by the rooftop observatory to estimate urban heat fluxes. Eight weather stations measured air temperature, relative humidity, dew point, wind speed and direction, gust speed, and solar radiation.

Weather stations near the observatory camera could be used to calculate τ_{ij}^{atm} in the following equation according to Waldemar and Klecha^{20,22}:

$$\tau_{ij}^{atm} = x \cdot \exp - \sqrt{d_{ij}} (\alpha_1 + \beta_1 \sqrt{\omega}) + (1 - x) \cdot \exp - \sqrt{d_{ij}} (\alpha_2 + \beta_2 \sqrt{\omega}) \quad (7)$$

where d_{ij} is the distance between the element and the observatory at position ij of the thermal image, ω is the atmosphere's water vapor content, α , β , and x are empirical coefficients. ω can be estimated from the air temperature (T_{air}) and relative humidity (ϕ) measured from the weather stations using the following equation:

$$\omega = \phi \cdot \exp [\gamma_0 + \gamma_1 T_{air} + \gamma_2 T_{air}^2 + \gamma_3 T_{air}^3] \quad (8)$$

where γ is another set of empirical coefficients described in Waldemar and Klecha^{20,22}.

Preprocessing. Following the classification of CNN, segmentation of the desired region and extraction of its radiometric data are required in order to evaluate the dataset. Flirextractor python package²³ is recommended, as it allows extracting temperature data from regions of interest in the thermal images and storing it in CSV (comma-separated values) format with a simple process. The temperature data from thermal images were converted using Eq. 1. The python package Labelme²⁴ is helpful in the segmentation of regions of interest for analysis. The relevant codes are provided in the Code Availability section.

Image information. The thermal image data are classified and stored based on their positions. The image datasets contain information including time, device positioning, and observatory locations (Kent Vale or S16). The building detail and surrounding environment could be seen based on the corresponding real images in this work. Analysis can be conducted in the time domain, frequency domain, or both, based on the nature of the time series and the degree of information to be retrieved¹⁵.

Potential applications. The thermal images provide surface temperatures of urban elements such as buildings, roads, and vegetation. Some example applications of surface temperature data are urban heat island analysis^{20,25}, urban energy monitoring²⁶, and building thermal performance monitoring^{8,15,27}. A detailed application of urban observatory is available in Dobler's work¹. Thermal images collected by the observatory can provide high-resolution data to study dynamic interactions in buildings at the district scale, much smaller than the city scale captured by satellite thermal image data, which are used frequently in relevant fields. Further studies of the thermal properties of non-residential buildings can be conducted and compared with those estimated at the city scale. The image data also lends itself to several digital image processing techniques, including edge detection, image enhancement, segmentation, and statistical analysis. For instance, edge detection could help discern different elements within the urban environment, such as buildings, vegetation, and transportation infrastructure. Image enhancement techniques could enhance the visibility of thermal patterns, and segmentation could be used to isolate areas of interest. The dataset provided here is well-suited for various applications, especially those involving urban heat island effects, building energy efficiency studies, and urban microclimate analysis. However, there remains potential for improving the accuracy of surface temperature measurements. Discussions about the data and innovative approaches to its interpretation and application are welcomed.

Privacy and safety control. These thermal images form the foundation for various research analyses related to urban living environments. For instance, they can be used to study temperature variations within an urban block, understand the effect of building materials on heat retention, or analyze the influence of vegetation on local microclimates. Researchers are encouraged to use various analytical techniques to probe the data, including computational statistics, machine learning, and sequential analysis. Machine learning, for instance, could be

used to predict future temperature patterns based on past data, while sequential analysis could help understand temporal changes in the thermal landscape. However, data misuse needs to be prevented^{1,28}. Since IoT cameras can be easily abused²⁹, the dataset collected here should include sufficient privacy protections for relevant individuals and cities^{1,30}. In such case, the images collected in this work do not contain personally identifiable information³¹ such as facial features, and the individual feature cannot be tracked. All taken images were strictly limited in pixel resolution so that interiors of buildings cannot be seen¹.

Code availability

Python 3 was used for data processing and analysis. The simple code demonstration suggested in this work can be found in the GitHub repository stated in the Data Record section. The code demonstrates the methodology for extraction and subsequent processing of the thermal images. Few images were selected from the dataset for illustration purposes. The code uses two specific GitHub packages: Flirextractor²³ and Labelme²⁴. Flirextractor is an efficient Python package for extracting temperature data from thermal images and converting it into an array, then saving it as a CSV file for further access. The link to the GitHub package describing the usage of Flirextractor in detail is as follows: <https://github.com/aloisklink/flirextractor>. Labelme is a graphical image annotation tool written in Python. It allows users to demarcate the desired area of any shape with simple mouse clicks. Detailed instructions for Labelme can be found at the following link: <https://github.com/wkentaro/labelme>. The GitHub code repository exhibits a meticulously structured hierarchy designed to enhance navigation and understandability. It incorporates three primary directories: 'data', 'notebook', and 'src'. The 'data' directory segregates original, converted, and processed files, whereas the 'notebook' directory encompasses Jupyter notebooks detailing file conversion, data visualization, and data analysis procedures. The 'src' directory hosts a Python script dedicated to image conversion. Furthermore, the 'data' folder manifests a systematic organization structured to efficiently manage different stages of data processing. It consists of three primary subdirectories: 'original', 'convert', and 'labelme'. The 'original' and 'processed' subdirectories mirror each other, containing dated folders representing different acquisition dates, with each date folder further divided into 'view_1', 'view_2', and 'view_3' subdirectories. These subdirectories contain the respective images captured from each view named "smap-YYYY-MM-DDTHH-MM-SS.MS.jpeg". During the data processing phase, corresponding 'json' files for each image are generated in the 'labelme' directory. In addition, for each image, a dedicated directory is established within the 'convert' subdirectory, encapsulating original, processed, and labeled images and a text file enumerating the detected labels. Detailed file structure tree plots, included within the respective README files, furnish comprehensive visual representations of the file and folder organization, thus facilitating an understanding of the hierarchical structure and interrelationships among different components of the codebase. Moreover, an example involving a representative image has been conducted to illustrate the efficacy of the employed image processing techniques; the results highlight the segmentation, further processing, and final presentation of the image augmented with enhanced colors and informative labels. Please contact the corresponding author directly if you have any particular requirements.

Received: 14 February 2023; Accepted: 14 November 2023;

Published online: 02 December 2023

References

- 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).
- Liu, P. & Biljecki, F. A review of spatially-explicit geospatial applications in urban geography. *International Journal of Applied Earth Observation and Geoinformation* **112**, 102936. <https://doi.org/10.1016/j.jag.2022.102936> (2022).
- Iea (2022), buildings, tracking report. *buildings* <https://www.iea.org/reports/buildings> (2022).
- Miguel, M. *et al.* A physically-based model of interactions between a building and its outdoor conditions at the urban microscale. *Energy and Buildings* **237**, 110788 (2021).
- Lucchi, E. Applications of the infrared thermography in the energy audit of buildings: A review. *Renewable and Sustainable Energy Reviews* **82**, 3077–3090 (2018).
- Ngie, A., Abutaleb, K., Ahmed, F., Darwish, A. & Ahmed, M. Assessment of urban heat island using satellite remotely sensed imagery: a review. *South African Geographical Journal = Suid-Afrikaanse Geografiese Tydskrif* **96**, 198–214 (2014).
- Almeida, C. R. D., Teodoro, A. C. & Gonçalves, A. Study of the urban heat island (uhi) using remote sensing data/techniques: A systematic review. *Environments* **8**, 105 (2021).
- Balaras, C. A. & Argiriou, A. Infrared thermography for building diagnostics. *Energy and buildings* **34**, 171–183 (2002).
- Arjunan, P. *et al.* Operational characteristics of residential air conditioners with temporally granular remote thermographic imaging. In *BuildSys '21: Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation*, Proceedings of the 8th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, 184–187. <https://doi.org/10.1145/3486611.3486659> (2021).
- Sham, J. F., Lo, T. Y. & Memon, S. A. Verification and application of continuous surface temperature monitoring technique for investigation of nocturnal sensible heat release characteristics by building fabrics. *Energy and Buildings* **53**, 108–116 (2012).
- 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).
- Chew, L. W., Liu, X., Li, X.-X. & Norford, L. K. Interaction between heat wave and urban heat island: A case study in a tropical coastal city, Singapore. *Atmospheric Research* **247**, 105134 (2021).
- Roth, M. Review of urban climate research in (sub) tropical regions. *International Journal of Climatology: A Journal of the Royal Meteorological Society* **27**, 1859–1873 (2007).
- Ignatius, M. *et al.* Local climate zones: Lessons from Singapore and potential improvement with street view imagery. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **X-4/W2-2022**, 121–128. <https://doi.org/10.5194/isprs-annals-X-4-W2-2022-121-2022> (2022).
- Ramani, V. *et al.* Longitudinal thermal imaging for scalable non-residential hvac and occupant behaviour characterization <https://doi.org/10.48550/ARXIV.2211.09288> (2022).
- Chen, S. *et al.* Atlas: Software for analysing the relationship between urban microclimate and urban morphology in a tropical city. *Building and Environment* **208**, 108591 (2022).

17. Yu, Z. *et al.* Dependence between urban morphology and outdoor air temperature: A tropical campus study using random forests algorithm. *Sustainable Cities and Society* **61**, 102200 (2020).
18. Lin, S. *et al.* Infrared investigation in singapore (iris) dataset. *Zenodo* <https://zenodo.org/doi/10.5281/zenodo.7463995> (2022).
19. Sobol, I. M. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation* **55**, 271–280 (2001).
20. Martin, M., Ramani, V. & Miller, C. *Infrared investigation in singapore (iris) observatory: Urban heat island contributors and mitigators analysis using neighborhood-scale thermal imaging* <https://doi.org/10.48550/ARXIV.2210.11663> (2022).
21. Saltelli, A. *et al.* Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer physics communications* **181**, 259–270 (2010).
22. Waldemar, M. & Klecha, D. Modeling of atmospheric transmission coefficient in infrared for thermovision measurements. In *Proceedings of the Sensor*, 903–907 (2015).
23. aloisklink/flirextractor v1.0.2: An efficient gplv3-licensed python package for extracting temperature data from flir irt images. *flirextractor* <https://github.com/aloisklink/flirextractor> (2020).
24. Russell, B. C., Torralba, A., Murphy, K. P. & Freeman, W. T. Labelme: a database and web-based tool for image annotation. *International journal of computer vision* **77**, 157–173 (2008).
25. Weng, Q. & Quattrochi, D. A. *Urban remote sensing* (CRC press, 2018).
26. Yang, X. X. *Urban remote sensing: Monitoring, synthesis and modeling in the urban environment* (John Wiley & Sons, 2021).
27. Kylii, A., Fokaides, P. A., Christou, P. & Kalogirou, S. A. Infrared thermography (irt) applications for building diagnostics: A review. *Applied Energy* **134**, 531–549 (2014).
28. Floridi, L. *et al.* Ai4people—an ethical framework for a good ai society: Opportunities, risks, principles, and recommendations. *Minds and machines* **28**, 689–707 (2018).
29. Tom Yeh, M. *et al.* Designing a moral compass for the future of computer vision using speculative analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 64–73 (2017).
30. *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (Cambridge University Press, 2014).
31. McCallister, E. *Guide to protecting the confidentiality of personally identifiable information*, vol. 800 (Diane Publishing, 2010).
32. Historical daily records, 2022. *climate-historical-daily* <http://www.weather.gov.sg/climate-historical-daily> (2022).
33. Meneghetti, G., Ricotta, M. & Atzori, B. A synthesis of the push-pull fatigue behaviour of plain and notched stainless steel specimens by using the specific heat loss. *Fatigue & Fracture of Engineering Materials & Structures* **36**, 1306–1322 (2013).
34. Raman, R. & Thakur, A. Thermal emissivity of materials. *Applied Energy* **12**, 205–220 (1982).
35. Ritland, H. N. Density phenomena in the transformation range of a borosilicate crown glass. *Journal of the American Ceramic Society* **37**, 370–377 (1954).
36. Sharp, D. & Ginther, L. Effect of composition and temperature on the specific heat of glass. *Journal of the American Ceramic Society* **34**, 260–271 (1951).
37. Wang, X. *et al.* Optically transparent bamboo with high strength and low thermal conductivity. *ACS Applied Materials & Interfaces* **13**, 1662–1669 (2020).
38. Iffat, S. Relation between density and compressive strength of hardened concrete. *Concrete Research Letters* **6**, 182–189 (2015).
39. De Schutter, G. & Taerwe, L. Specific heat and thermal diffusivity of hardening concrete. *Magazine of Concrete research* **47**, 203–208 (1995).
40. Kim, K.-H., Jeon, S.-E., Kim, J.-K. & Yang, S. An experimental study on thermal conductivity of concrete. *Cement and concrete research* **33**, 363–371 (2003).
41. Olioso, A. Simulating the relationship between thermal emissivity and the normalized difference vegetation index. *International Journal of Remote Sensing* **16**, 3211–3216 (1995).
42. Ganguly, S. *et al.* Generating vegetation leaf area index earth system data record from multiple sensors. part 2: Implementation, analysis and validation. *Remote Sensing of Environment* **112**, 4318–4332 (2008).
43. 1°17'50"N, 103°46'37"E. **Google Earth**. march 05, 2016. november 01, 2022. <https://earth.google.com/web/@1.29728934,103.77702845,23.33038212a,2430.67786248d,35y,27.55033278h,0t,0r> (2016).

Acknowledgements

This research is supported by the National Research Foundation and the Prime Minister's Office of Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) program. It was funded through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST2) program. BEARS was established by the University of California, Berkeley, as a center for intellectual excellence in research and education in Singapore. This research is also supported by the Multi-scale Digital Twins for the Urban Environment: From Heartbeats to Cities project, funded by the Singapore Ministry of Education (MOE) Tier 1 Academic Research Fund, NUS Resilience and Growth Postdoctoral Fellowship - Smart Cities and Urban Analytics supported by National Research Foundation (NRF), and the Johnson Controls Open Blue Innovation Laboratory. Furthermore, we would like to acknowledge the contributions of the Development of Integrated Multi-scale and Multiphysics Urban Microclimate Model project, funded by the National University of Singapore and supported by the University Campus Infrastructure (UCI) and Office of the Deputy President (Research & Technology).

Author contributions

S.L. wrote the original draft and assisted in data collection and formal analysis, V.S. and M.M. designed and deployed the data collection platform, P.A. and M.I. supported the data collection process, A.C. and F.B. provided conceptual input for the data platform deployment and reviewed the manuscript, K.P. edited and reviewed the manuscript. and C.M. supervised the deployment, data collection, and development outlined in this article and edited and reviewed the manuscript. All authors have reviewed the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

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

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.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023