Integrating Multi-Source Temperature Data and Explainable Deep Learning for Urban Microclimate Analysis

Shiqi Tu [12] , Qingming Zhan [*12] and Yinghui Xiao[12] and Liu Minghao[1] and Ruihan Qiu [12] and Zhihua Liu [1]
1 School of Urban Design, Wuhan University, Wuhan 430072, China
2 Research Center for Digital City, Wuhan University, Wuhan 430072, China
***Corresponding author:** Qingming Zhan, School of Urban Design, Wuhan University, Wuhan 430072, China
Email: qmzhan@whu.edu.cn
Tel: +86-139-9566-8639

Keywords: Urban microclimate; Land surface temperature; Air temperature; IoT sensor network; Explainable deep learning; Grad-CAM.

Abstract

Understanding the relationship between land surface temperature (LST) and near-surface air temperature is essential for fine-scale urban heat assessment. This study examines the spatial and temporal coupling between Landsat-8-derived LST and in-situ air temperature measured by a dense IoT sensor network across 19 sites on a university campus during June-August 2024. The campus includes varied building forms, surface materials, vegetation, and water bodies, providing a heterogeneous setting for evaluating surface–air thermal relationships. Rather than treating LST as a direct proxy for air temperature, the analysis compares spatial rankings, diurnal variations, and surface–air temperature differences to identify consistent and divergent thermal patterns. A convolutional neural network combined with Gradient-weighted Class Activation Mapping (Grad-CAM) was further used to test whether spatially reweighted LST information better corresponds to observed air temperature variability. Results show that LST presents stronger spatial differentiation than near-surface air temperature, whereas air temperature displays smoother spatial patterns and clear nighttime convergence. Surface–air temperature differences vary systematically across site environments, indicating heterogeneous coupling rather than random mismatch. The CAM-assisted regression analysis shows that emphasizing thermally relevant surface regions improves the correspondence between satellite-derived thermal information and ground observations. This study provides an interpretable framework for micro-scale analysis of surface–air temperature relationships and supports more reliable characterization of urban thermal environments through integrated satellite and sensor-based observations.

1. INTRODUCTION

Urban thermal environments have become a central concern in urban climate research due to their direct implications for heat exposure, thermal comfort, public health, and climate-sensitive urban design (Qiu et al., 2025). With accelerating urbanization, heterogeneous surface materials, complex building morphologies, and uneven vegetation distributions have intensified spatial variability in urban heat patterns, particularly during summer periods. Accurately characterizing this variability at fine spatial scales remains a key challenge for both scientific understanding and practical intervention (Schmidt et al., 2022).

Satellite-derived land surface temperature (LST) has been widely used to map urban thermal environments because of its spatial continuity and broad coverage (Liu et al., 2023). Numerous studies have demonstrated the effectiveness of LST in capturing surface thermal contrasts associated with impervious materials, vegetation cover, and solar exposure. However, LST represents radiative surface temperature rather than near-surface air temperature, which is more directly relevant to human thermal exposure. The physical distinction between surface and air layers, combined with atmospheric mixing and local ventilation, often leads to discrepancies between LST patterns and actual near-surface thermal conditions (Xie et al., 2025).

Existing research has frequently assessed the relationship between LST and air temperature through correlation analysis or regression-based validation using sparse meteorological stations. While such approaches provide useful aggregate statistics, they tend to obscure spatial structure and local heterogeneity. In particular, strong correlations at the city scale do not necessarily imply consistent spatial correspondence at neighborhood or micro-scale levels (Chao et al., 2020). As a result, the use of LST as a proxy for near-surface air temperature remains methodologically convenient but conceptually ambiguous (Wang and Shu, 2020).

Recent advances in urban microclimate monitoring have enabled the deployment of dense sensor networks that capture near-surface air temperature at high temporal resolution. These in-situ observations provide valuable reference information for examining fine-scale thermal variability that cannot be resolved by conventional meteorological stations. However, integrating point-based air temperature measurements with gridded satellite thermal data poses challenges related to scale mismatch, spatial representativeness, and interpretation (Zheng et al., 2023). Simply treating in-situ observations as ground truth for pixel-level validation risks conflating fundamentally different thermal quantities.

At the same time, machine learning techniques have been increasingly applied to urban climate studies to model complex, nonlinear relationships between surface characteristics and thermal conditions. Despite their predictive potential, many deep learning approaches function as black boxes, offering limited interpretability and weak physical grounding (Ke et al., 2021). This lack of transparency restricts their usefulness for scientific explanation and planning-oriented decision support, where understanding the spatial drivers of thermal patterns is as important as predictive accuracy (Su et al., 2024).

Against this background, there is a clear need for analytical frameworks that move beyond direct validation or prediction and instead focus on spatial consistency, relative structure, and interpretable relationships between surface and air temperature. Emphasizing spatial ranking, difference structure, and coupling behavior can provide deeper insight into how surface thermal signals translate into near-surface atmospheric conditions under heterogeneous urban settings (Shao et al., 2023).

This study addresses these challenges by integrating high-resolution satellite-derived LST with a dense IoT-based air temperature monitoring network deployed within a heterogeneous university campus environment. Rather than attempting to directly retrieve or predict absolute air temperature from LST, the analysis focuses on comparing spatial patterns, diurnal dynamics, and surface–air temperature differences across monitoring sites. Furthermore, an explainable deep learning framework is introduced, in which a convolutional neural network combined with Gradient-weighted Class Activation Mapping (Grad-CAM) is used to reweight spatial thermal information and assess its correspondence with observed air temperature variability.

The objectives of this study are threefold. First, to examine the spatial consistency and divergence between LST and near-surface air temperature by analyzing relative rankings and surface–air temperature differences across diverse microenvironments. Second, to investigate diurnal thermal response modes using high-frequency in-situ observations, revealing time-dependent patterns of spatial heterogeneity. Third, to evaluate the feasibility and interpretability of a CAM-assisted regression mapping approach for characterizing near-surface air temperature based on spatially reweighted LST information.

By explicitly distinguishing surface and air thermal signals and emphasizing interpretability over black-box prediction, this study contributes a refined methodological perspective on urban thermal analysis. The proposed framework offers a transferable approach for integrating satellite thermal imagery, dense sensor observations, and explainable machine learning, with implications for fine-scale heat assessment and climate-sensitive urban design (Figure 1).

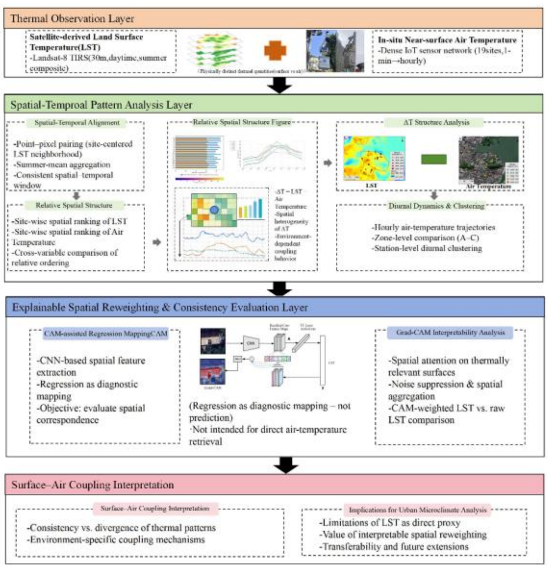


Figure 1. Research framework

2. MATERIALS AND METHODS

2.1 Study area

The study area is located within a large university campus in central China and is characterized by a heterogeneous hill–lake–urban composite landscape. The terrain combines gentle hills, open spaces, dense academic building clusters, and lakeside environments. A major water body is located along the eastern edge of the campus, while compact urban development borders the western side, creating pronounced contrasts in surface materials, vegetation cover, and ventilation conditions.

To facilitate spatial comparison, the campus was subdivided into three functional zones (Zones A–C) according to dominant land use, building morphology, and underlying surface characteristics rather than administrative boundaries. Zone A, located in the southwest, is dominated by teaching facilities, student residences, and laboratory clusters, with relatively high building density and extensive impervious surfaces. Zone B occupies the northern part of the campus and mainly contains engineering research complexes and academic laboratories, characterized by compact built forms interspersed with localized green spaces. Zone C is situated in the central and eastern lakeside area and includes administrative buildings, library facilities, and humanities-oriented teaching spaces, featuring higher vegetation coverage and proximity to water bodies.

This zoning scheme was intended to represent contrasts in microclimatic settings—such as shading conditions, surface thermal properties, and ventilation potential—at the campus scale, rather than to delineate formal climatic zones. The overall study area covers approximately 3 km². The spatial extent of the campus and the locations of the three zones are illustrated in Figure 2. It should be noted that the zoning does not assume predefined microclimatic differences, but reflects contrasts in morphological and surface characteristics that are known to influence local air temperature. Table 1 summarizes the 19 monitoring sites and their major microenvironmental characteristics.

Site (Zone)	Location Type & Microenvironmental Features
S1 (A)	Student residential area — Medium density; partial shade; mixed paving
S2 (A)	Campus entrance — High pedestrian flow; impervious ground; exposed surface
S3 (A)	Teaching facility cluster — Dense built-up; exposed roofs; limited vegetation
S4 (A)	Canteen vicinity — Impervious pavement; moderate shade; local heat sources
S5 (A)	Main plaza — Asphalt pavement; open area; minimal vegetation
S6 (B)	Campus gate — Compact built forms; shaded edges; pedestrian corridor
S7 (B)	Mixed-use junction — Residential cluster; mixed materials; scattered trees
S8 (B)	Open lawn area — Large vegetated surface; strong evapotranspiration
S9 (B)	Dormitory block — Concrete surfaces; medium density; partial shade
S10 (B)	Minor entrance — Walkway corridor; mixed ground materials
S11 (B)	Secondary gate — Asphalt pavement; partial shading from nearby trees
S12 (B)	Sports facility surroundings — Semi-open landscape; mixed impervious–vegetated cover
S13 (C)	Academic building vicinity — Vegetation shade; moderate heat accumulation
S14 (C)	Administrative complex — Reflective façade materials; moderate airflow

https://doi.org/10.5194/isprs-archives-XLIX-B3-2026-439-2026 | © Author(s) 2026. CC BY 4.0 License.

S15 (C)	Central open square — Paving stones; high solar exposure; minimal shade
S16 (C)	Rooftop platform — Full-sun exposure; strong radiant heating; high albedo contrast
S17 (C)	Roadside segment — Tree-lined corridor; moderated airflow; partial shading
S18 (C)	Lakeside walkway — Cooling from nearby water body; dense vegetation
S19 (C)	Dining facility region — Impervious pavement; human activity; moderate shading

Table 1. Overview of the 19 sampling sites (S1–S19) and their microenvironmental characteristics within the study area

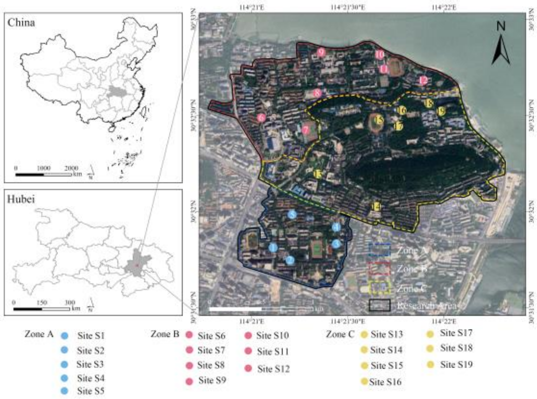


Figure 2. Research area

2.2 Data

2.2.1 Monitoring Site Selection and Deployment Strategy

The 19 monitoring sites were not selected based on the presence of pre-existing or permanently installed meteorological stations. Instead, all sites were deliberately chosen and instrumented for this study according to surrounding land-cover characteristics, functional use, and local morphological conditions.

The site selection aimed to capture a wide range of microclimatic environments at the campus scale, including areas near student dormitories, dining facilities, teaching clusters, plazas, roads, mixed-use activity nodes, rooftops, and lakeside corridors. This design prioritizes microclimatic diversity rather than statistical representativeness at the city scale.

2.2.2 In-Situ Air Temperature Measurements (CityGrid IoT Sensors)

Near-surface air temperature was measured using third-generation CityGrid urban microclimate monitoring stations (G3.5, Beijing Xiangsu Technology Co., Ltd.). The stations provide fixed-point outdoor observations of key meteorological variables, including air temperature, relative humidity, pressure, wind speed, wind direction, and illuminance. Air temperature was recorded at one-minute intervals within a measurement range of −20 to 60 °C and was aggregated into hourly means for analysis. Data from 19 monitoring sites across the campus were used in this study. Station coordinates were obtained from the CityGrid system and cross-checked using Google Earth Pro, with an estimated horizontal accuracy of ±0.0001° and vertical accuracy of ±0.5 m. The in-situ observations were used as reference measurements of near-surface atmospheric conditions rather than direct equivalents of satellite-derived LST (Table 2).

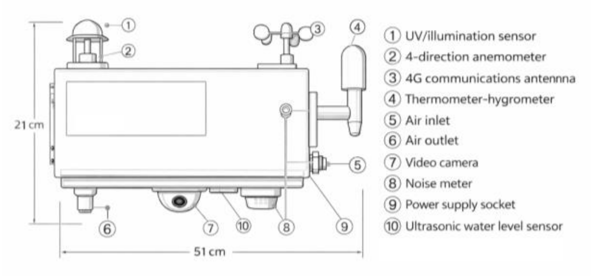


Figure 3. Equipment diagram

NO.	Range	Interval
1	Pressure	300-1100hPa
2	Wind velocity	0-45m/s
3	Wind direction	0-360°
4	Illumination	0-200000Lux
5	Air temperature	-20-60℃
6	Air humidity	0-100%RH

Table 2. Measurement ranges of key environmental variables recorded by the CityGrid in-situ monitoring sensors.

2.2.3 Satellite-Derived LST Data (Landsat-8 TIRS)

Satellite-derived land surface temperature (LST) was obtained from the Landsat-8 Collection 2 Level-2 Science Product through Google Earth Engine. The LST data were retrieved from the Thermal Infrared Sensor at a spatial resolution of 30 m. All clear-sky daytime scenes acquired from June to August 2024 were selected, and cloud-contaminated pixels were removed using the Quality Assessment band. After filtering, eight valid scenes were retained, all acquired at approximately 10:30 local solar time. A summer LST composite was generated using pixel-wise median aggregation to reduce residual noise and scene-specific effects. The LST data were used to characterize spatially continuous surface thermal patterns. They were not calibrated by the in-situ air temperature observations, because LST and near-surface air temperature represent different thermal quantities. Instead, the two datasets were independently processed and compared to examine spatial consistency and divergence in surface–air temperature coupling.

2.3 Methodology

2.3.1 In-situ Air Temperature Monitoring Network and Deployment Strategy

Field air-temperature measurements were collected using an IoT-based sensor network deployed across representative campus environments (Figure 4). Monitoring sites were selected according to building density, vegetation coverage, surface materials, and proximity to water bodies, covering teaching areas, dormitory zones, open paved spaces, and lakeside corridors (Figure 4a). Sensors were mounted on existing utility poles at approximately 2.5–3.0 m above ground to ensure stable installation and reduce direct ground-surface interference (Figure 4c). Each sensor unit included an integrated temperature probe housed within a protective enclosure for continuous outdoor monitoring (Figure 4b). The network recorded air temperature at one-minute intervals from June to August 2024. All sensors were calibrated before deployment, with measurement uncertainty controlled within ±0.3 °C.

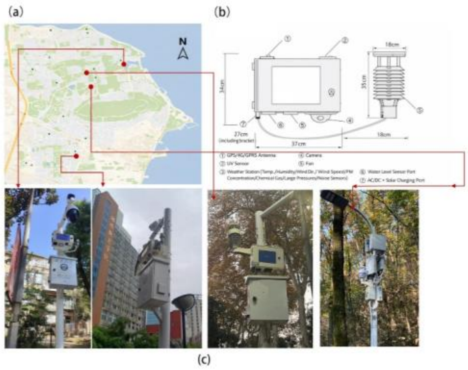


Figure 4. Configuration and deployment of the in-situ air temperature observation system. (a) Spatial distribution of the sensor network across the study area. (b) Structural schematic of the IoT-based air temperature sensor, showing key components and installation configuration. (c) Field deployment of sensors mounted on utility poles in representative microclimatic environments.

2.3.2 Satellite-derived Land Surface Temperature Data

Satellite-derived LST was obtained from clear-sky Landsat-8 observations from June to August 2024. Cloud-contaminated and low-quality pixels were removed using quality assessment information and spatial consistency checks. For each monitoring site located at (x_i, y_i), LST values were extracted from the corresponding satellite pixel and, where needed, from a local neighborhood to reduce point–pixel mismatch. The neighborhood-based LST was calculated as:

$$LST_i^{(r)} = \frac{1}{|\Omega_r|} \sum_{(x,y)\in\Omega_r} LST(x,y) \qquad (1)$$

where Ω_r denotes the set of pixels within a predefined neighborhood centered on monitoring site i.

The summer-averaged LST and air temperature at site i were calculated as:

$$L\bar{S}T_i = \frac{1}{N_i}\sum_{t=1}^{N_i} LST_{i,t}\,, \bar{T}_i = \frac{1}{N_i}\sum_{t=1}^{N_i} T_{i,t} \qquad (2)$$

where N_i represents the number of valid observations during the study period.

The surface–air temperature difference was then defined as:

$$\Delta T_i = L\bar{S}T_i - \bar{T}_i \qquad (3)$$

This metric was used to evaluate systematic deviations and spatial heterogeneity in surface–air temperature coupling.

2.3.3 Temporal Pairing and Difference Analysis

To ensure comparability, LST and in-situ air temperature were paired at the same monitoring sites and over the same summer observation period. This point-wise pairing established a one-to-one correspondence between surface thermal information and near-surface air temperature observations.

Based on Eq. (3), ΔT_i was treated as a diagnostic metric to quantify the magnitude and direction of deviation between LST and air temperature. Spatial variations in ΔT_i were analyzed to identify systematic differences in surface–air temperature coupling across local campus environments.

2.3.4 CAM-assisted Regression Mapping

A CAM-assisted regression mapping framework was used to evaluate whether LST-derived spatial information could better correspond to near-surface air temperature. In this framework, the in-situ measured air temperature T_i served as the response variable, while the predictor information X_i was derived from spatial LST patterns around each monitoring site.

The regression relationship was expressed as:

$$\hat{T}_i = f(X_i) \qquad (4)$$

where $\hat{T}_i$ denotes the model-estimated air temperature and $f(\cdot)$ represents the regression function learned by the model. A convolutional neural network was used to extract spatial thermal features from the LST input. Model performance was evaluated using RMSE, MAE, Pearson correlation coefficient, and regression slope (Figure 5).

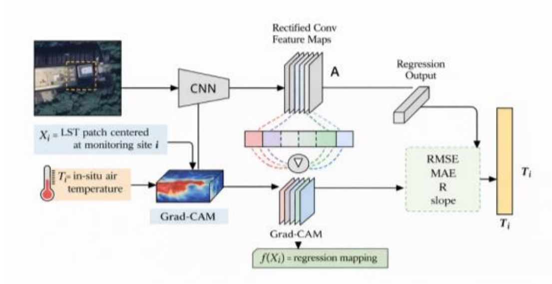


Figure 5. CAM-assisted regression mapping framework for near-surface air temperature estimation based on LST spatial information.

2.3.5 Grad-CAM-based Interpretability and Consistency Evaluation

Gradient-weighted Class Activation Mapping was applied to interpret the trained regression model. Grad-CAM identifies spatial regions within the LST input that contribute most strongly to the regression output (Liu et al., 2025).

The activation map for site i was expressed as:

$$CAM_i(x,y) = \mathrm{ReLU}\,(\sum_k w_k A_k(x,y)) \qquad (5)$$

where $A_k(x,y)$ denotes the activation of the k-th feature map at spatial location (x,y), and w_k represents the corresponding gradient-derived weight. CAM-weighted LST representations were then compared with in-situ air temperature observations to evaluate whether spatial reweighting improved surface–air temperature consistency (Figure 6).

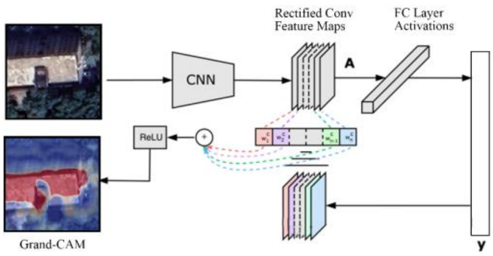


Figure 6. Grad-CAM–guided feature reweighting framework for LST retrieval.

3. RESULTS

3.1 Comparison of Spatial Patterns and Surface–Air Temperature Coupling

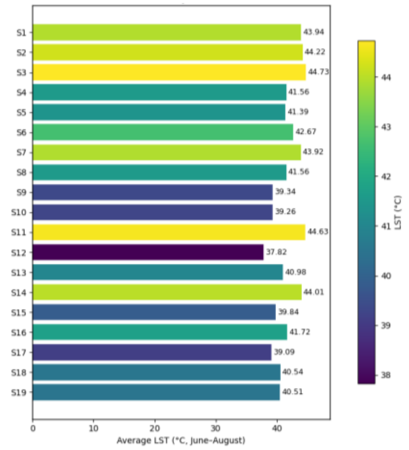
Figure 7 compares the spatial patterns of land surface temperature (LST), measured air temperature, and their difference (ΔT = LST − Air T) across monitoring sites during June-August 2024. The results reveal clear differences in spatial ranking and structural variability between surface and air temperature fields.

The spatial ranking of LST shows strong differentiation among monitoring sites (Figure 7a). Sites associated with impervious surfaces and exposed built environments consistently exhibit higher LST values, whereas vegetated or shaded environments remain at the lower end of the ranking. The ordering of sites indicates a pronounced spatial contrast in surface thermal conditions across the campus.

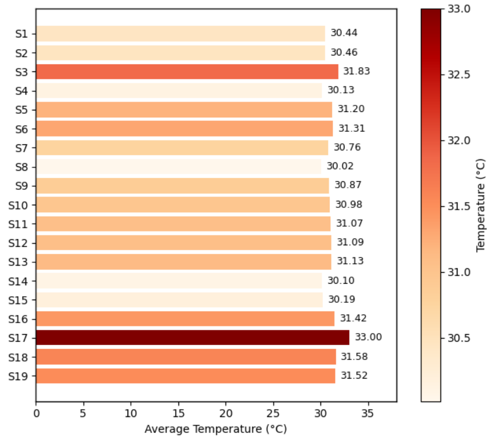
By comparison, the spatial ranking of measured air temperature is notably more uniform (Figure 7b). Although higher air temperatures are still observed at several built-up locations, the overall spread across sites is substantially reduced relative to LST. This suggests weaker spatial differentiation in near-surface air temperature despite pronounced surface thermal contrasts.

The structure of surface–air temperature differences further highlights systematic spatial patterns (Figure 7c). ΔT values vary markedly across environments, with impervious and highly exposed sites showing larger positive ΔT, while vegetated and shaded sites display smaller differences. The spatial distribution of ΔT does not simply mirror either LST or air temperature alone, but reflects distinct surface–air thermal relationships across different site types.

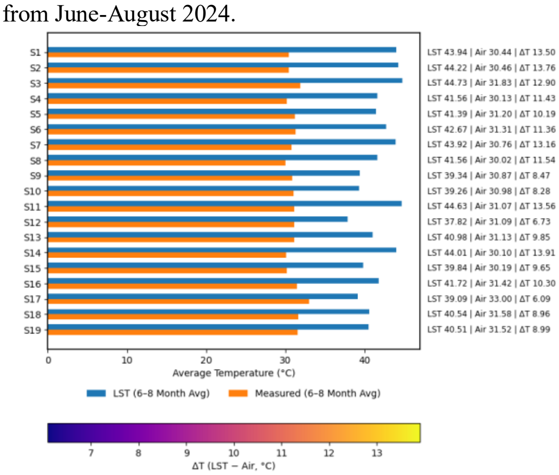
Taken together, these results indicate that while LST and air temperature share broadly consistent spatial tendencies, their relative rankings and difference structures diverge across environments. The magnitude and spatial organization of ΔT reveal clear differences in surface–air temperature coupling strength among urban forms, underscoring that relative spatial patterns and ΔT structure provide more informative evidence than absolute temperature agreement alone.



(a) Average land surface temperature (LST) across monitoring sites from June-August 2024.



(b) Average measured air temperature across monitoring sites from June-August 2024.



(c) Variations in surface–air temperature difference (ΔT = LST − Air T) across different environments.

Figure 7. Comparison between land surface temperature (LST) and field-measured air temperature across monitoring sites of the study campus from June-August 2024.

3.2 Diurnal Air Temperature Variations across Campus Zone

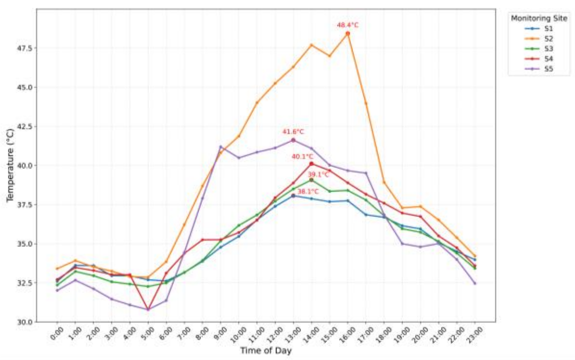
Figure 8 illustrates the hourly air temperature variations across three campus zones during June-August 2024, revealing both common diurnal characteristics and clear zone-specific differences. In all three zones, air temperature follows a pronounced unimodal diurnal cycle, with relatively stable conditions during the early morning, a rapid increase after sunrise, peak temperatures occurring between late morning and early afternoon, and a gradual decline toward evening.

In Zone A (Figure 8a), temperature curves across monitoring sites are closely aligned throughout the day. Daytime warming is moderate, and the timing of peak temperature is relatively consistent among sites. The diurnal amplitude remains limited, and post-peak cooling proceeds smoothly with minimal divergence between curves. This indicates low intra-zone variability in both peak magnitude and cooling rate within Zone A.

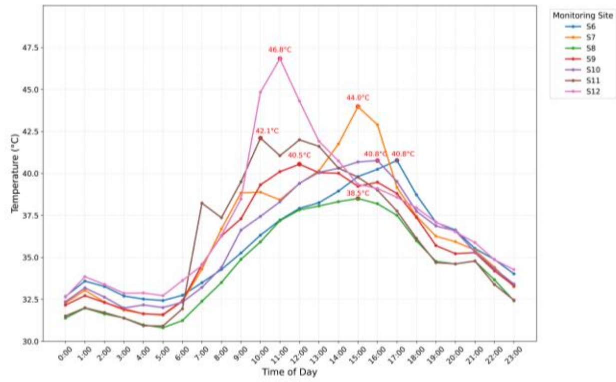
Zone B (Figure 8b) exhibits greater heterogeneity in diurnal temperature behavior. While all sites display a similar timing of daytime warming, peak temperatures vary substantially among stations, resulting in a wider spread around midday. Several sites show sharper morning temperature increases and higher daytime maxima compared to others, followed by more variable afternoon cooling trajectories. This leads to the largest intra-zone dispersion during daytime hours among the three zones.

Zone C (Figure 8c) shows the most pronounced diurnal contrasts. Certain sites reach substantially higher daytime maxima than the rest, with elevated temperatures persisting into the early afternoon. Although cooling occurs toward evening across all sites, the rate of temperature decline differs noticeably, maintaining clear separation among curves for an extended period. Compared with Zones A and B, Zone C exhibits both larger diurnal amplitudes and stronger persistence of daytime heat.

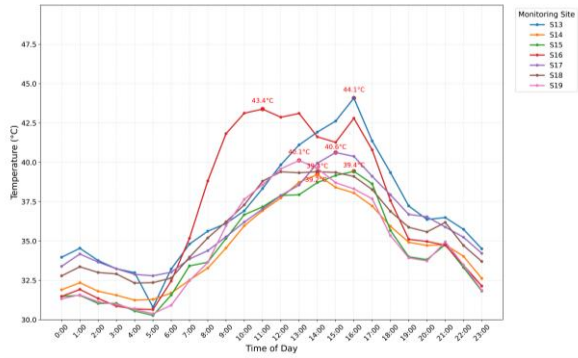
Across all zones, nighttime temperatures converge to a relatively narrow range, whereas daytime temperatures diverge markedly, particularly around peak heating hours. This pattern indicates that spatial differentiation of air temperature within the campus is most evident during periods of daytime warming, while nocturnal conditions remain comparatively homogeneous.



(a) Hourly temperature variation in the Zone A from June-August 2024.



(b) Hourly temperature variation in the Zone B from June-August 2024.



(c) Hourly temperature variation in the Zone C from June-August 2024.

Figure 8. Hourly temperature variations across different campus zones of the study area from June to August 2024: (a) Zone A; (b) Zone B; (c) Zone C.

Figure 9 presents the station-level diurnal air temperature patterns grouped by time-series clustering (k = 5), revealing clear and systematic differences in temporal structure among clusters. The clustering separates monitoring sites primarily according to diurnal amplitude, timing of peak temperature, and persistence of elevated temperatures. Distinct clusters emerge with coherent internal patterns, indicating that stations within the same cluster share highly similar hourly temperature trajectories over the diurnal cycle.

Cluster 0 is characterized by a sharp and pronounced midday peak, with rapid temperature increases during the morning hours and delayed cooling in the afternoon. Stations in this cluster exhibit the largest daytime temperature maxima relative to early morning conditions, resulting in a steep diurnal rise and an asymmetric temperature curve. Cluster 1 displays muted daytime peaks and relatively low nocturnal temperatures, with reduced diurnal amplitude and rapid evening cooling. The temperature curves in this cluster remain comparatively flat during daytime hours, indicating limited daytime amplification. Cluster 2 shows a sustained daytime thermal plateau, where temperatures rise steadily and remain elevated over an extended period before gradually decreasing after sunset. This cluster exhibits slower post-peak cooling and a prolonged high-temperature phase compared with other groups.

Cluster 3 represents an intermediate pattern, with moderate diurnal amplitude and balanced heating and cooling rates. Temperature curves in this cluster display smoother transitions between daytime and nighttime conditions, without extreme peaks or prolonged plateaus. Cluster 4 consists of a single station with an extreme diurnal profile, exhibiting the highest daytime maximum and the strongest persistence of elevated temperatures among all stations. This isolated cluster highlights the presence of highly localized thermal behavior that deviates substantially from the dominant campus-wide patterns.

As summarized in Table 3, the five clusters collectively capture a wide spectrum of diurnal temperature behaviors, ranging from sharply peaked and highly persistent profiles to damped and rapidly cooling patterns. Stations belonging to the same cluster are not confined to a single spatial zone, indicating that diurnal thermal behavior varies independently of broad zonal classification. These results demonstrate that station-level thermal variability is structured by distinct diurnal response modes, which cannot be adequately represented by daily mean temperatures or single-time-point observations alone.

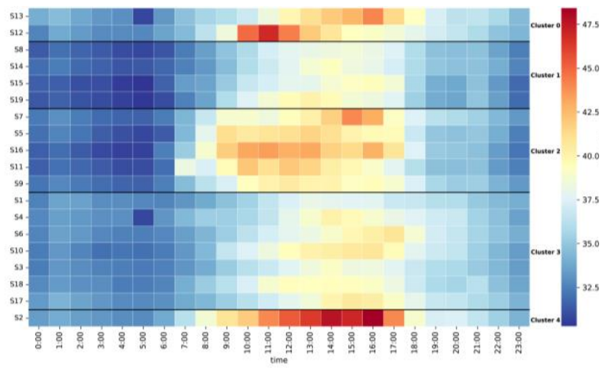


Figure 9. Station-level diurnal temperature patterns grouped by time-series clustering (k= 5).

Cluster	Stations	Diurnal behavior	Mechanism	Key attributes
Cluster 0	S12, S13	Midday peak; rapid heating; delayed cooling	Radiative dominance	Impervious surface; full exposure; little shading
Cluster 1	S8, S14, S15, S19	Low night temperature; weak daytime peak; fast cooling	Shading; evapotranspiration	Tree canopy; vegetation; courtyard shelter
Cluster 2	S5, S7, S9, S11, S16	Daytime plateau; slow nighttime cooling	Heat storage release	Dense buildings; high wall ratio; poor ventilation
Cluster 3	S1, S3, S4, S6, S10, S17, S18	Moderate amplitude; balanced cycle	Mixed coupling	Transitional form; partial shading; mixed surfaces
Cluster 4	S2	Extreme daytime maximum; strong persistence	Radiation trapping; heat storage	Exposed plaza; building enclosure; little vegetation

Table 3. Mechanism-oriented interpretation of time-series clustering results.

3.3 Interpretable CAM-Based Retrieval of Near-Surface Air Temperature

Figure 10 presents the Grad-CAM–based interpretability results of the CAM model for representative surface types, including vegetation, building roofs, asphalt pavement, and concrete pavement. The activation maps reveal clear spatial heterogeneity in model responses across different surface components. High activation intensities are consistently observed over surfaces with strong thermal relevance, while surrounding background regions exhibit weaker or suppressed responses. This pattern indicates that the CAM model does not treat the input scene uniformly, but instead assigns differentiated spatial weights to heterogeneous surface elements.

Notably, the Grad-CAM maps exhibit more spatially coherent activation patterns, with reduced high-frequency variability at the pixel level. Compared with the original imagery, the highlighted regions appear smoother and more aggregated, suggesting that the CAM model effectively filters out fine-scale noise and emphasizes thermally representative surface structures. These results demonstrate that Grad-CAM provides an interpretable view of how the model selectively focuses on surface components that are potentially more informative for near-surface temperature estimation.

The quantitative effect of this spatial reweighting is further evaluated through regression analysis, as shown in Figure 11. For the CAM-based retrieval, a strong linear relationship is observed between the retrieved temperature and the in-situ measured air temperature, with a regression slope of 2.00 and a high Pearson correlation coefficient (R = 0.884). The data points are tightly clustered around the fitted regression line, and the corresponding confidence interval remains relatively narrow across the full range of retrieved values, indicating stable numerical behavior and a consistent response to variations in measured air temperature.

In contrast, the LST-based retrieval exhibits a substantially weaker relationship with in-situ measurements. The regression slope is markedly lower (0.53), and the correlation coefficient is limited (R = 0.158). The scatter of data points is considerably larger, accompanied by a wide confidence interval, suggesting that LST-derived temperatures show greater variability and reduced consistency when compared with near-surface air temperature under the same conditions.

Taken together, Figures 10 and 11 indicate that the CAM-based approach achieves a pronounced improvement in both the strength and stability of the numerical relationship between satellite-derived temperature and in-situ observations. The enhanced correlation, steeper regression slope, and reduced dispersion collectively demonstrate that the CAM-retrieved temperature provides a more consistent and representative characterization of near-surface thermal conditions than the conventional LST-based retrieval.

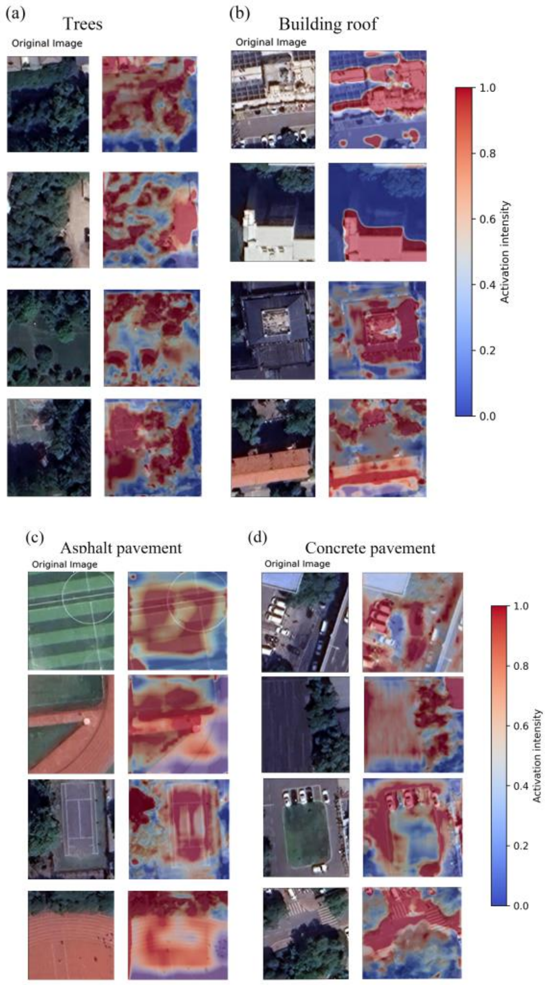


Figure 10. Attention activation maps for typical land-cover types. Panels (a)–(d) show trees, building roofs, asphalt pavement, and concrete pavement, respectively. For each sample, the original RGB image and the corresponding attention activation map are displayed side by side. Warmer colors indicate higher activation intensity.

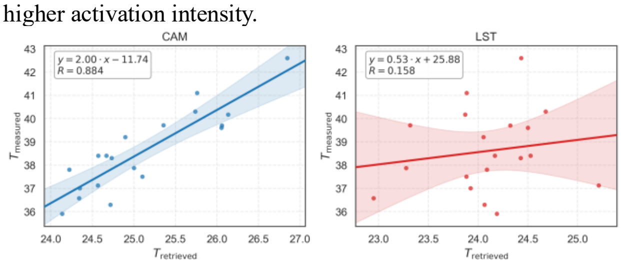


Figure 11. Linear regression between retrieved temperature and in-situ measured air temperature for the CAM-based retrieval (left) and the LST-based retrieval (right).

4. DISCUSSION

4.1 Spatial Consistency and Divergence Between Surface and Air Temperature in Relation to Existing Studies

The comparison between land surface temperature (LST) and near-surface air temperature reveals both consistency and divergence in their spatial patterns, which aligns with and extends findings reported in previous urban climate studies. Numerous studies have shown that LST captures strong spatial contrasts associated with surface materials, vegetation cover, and solar exposure, whereas near-surface air temperature tends to exhibit smoother spatial gradients due to atmospheric mixing and ventilation effects. The pronounced spatial differentiation of LST observed in this study is therefore consistent with established understanding of surface thermal behavior in heterogeneous urban environments (Chen et al., 2020).

However, the relatively uniform spatial ranking of air temperature across monitoring sites highlights a key limitation of using LST as a direct proxy for near-surface thermal conditions, a point that has been noted but not always explicitly quantified in prior research. By explicitly analyzing the spatial structure of surface–air temperature differences (ΔT), this study demonstrates that surface and air temperatures do not diverge randomly, but follow systematic patterns linked to local environmental settings (Wang et al., 2022). Compared with studies that rely primarily on correlation analysis between LST and air temperature, the present results provide a more nuanced characterization of surface–air temperature coupling by emphasizing relative spatial differences and difference structures rather than absolute agreement alone (Dai et al., 2018).

4.2 Diurnal Thermal Response Modes and Implications for Spatial Heterogeneity

The time-series analysis and clustering of station-level air temperature reveal that spatial heterogeneity in thermal conditions is strongly time-dependent, with the most pronounced differentiation occurring during daytime heating periods. While nighttime temperatures converge across the campus, daytime temperature trajectories diverge substantially, particularly around peak heating hours. This finding is consistent with previous observations that urban thermal heterogeneity is amplified under strong radiative forcing and attenuated under nocturnal conditions when surface–air coupling weakens (Yang et al., 2023).

The identification of distinct diurnal response modes through clustering further demonstrates that thermal behavior cannot be adequately represented by daily mean temperatures or single-time-point observations. Stations belonging to the same diurnal cluster are not confined to specific spatial zones, indicating that local thermal responses are shaped by site-specific surface and morphological characteristics rather than broad functional zoning alone. This result complements existing literature by providing empirical evidence that temporal response patterns constitute an independent dimension of urban thermal heterogeneity, reinforcing the need to integrate diurnal dynamics into the interpretation of urban heat patterns (Tu et al., 2025).

4.3 Feasibility and Effectiveness of CAM-Assisted Regression Mapping for Near-surface Temperature Characterization

The Grad-CAM–assisted regression analysis provides important insights into the feasibility of using spatially reweighted LST information to characterize near-surface air temperature. The activation maps demonstrate that the regression model selectively focuses on thermally relevant surface components, such as impervious pavements and exposed roofs, while suppressing less informative background regions. This spatial selectivity is consistent with physical expectations and supports the interpretability of the modeling framework.

More importantly, the quantitative comparison between CAM-based retrieval and conventional LST-based retrieval shows a substantial improvement in numerical consistency with in-situ air temperature observations. The stronger correlation, steeper regression slope, and reduced dispersion observed for the CAM-based results indicate that incorporating spatial attention mechanisms effectively enhances the surface–air temperature correspondence. These findings provide empirical validation for the feasibility of the proposed regression mapping approach, demonstrating that LST-derived spatial information, when appropriately reweighted, can serve as a reliable basis for near-surface temperature characterization (Sützl et al., 2024).

Compared with existing approaches that treat LST as a single-point predictor or rely on direct statistical correlations, the CAM-assisted framework offers a more robust and interpretable alternative. While the current implementation focuses on thermal information alone, the demonstrated improvement suggests strong potential for extending this framework to multi-source predictors in future studies, further enhancing predictive capability and physical interpretability.

4.4 Implications and Limitations

This study provides several methodological and practical implications for understanding surface–air temperature coupling in complex urban environments. By explicitly distinguishing spatial patterns of land surface temperature, near-surface air temperature, and their differences, the results demonstrate that relative spatial structure and surface–air temperature coupling offer more informative insights than absolute temperature agreement alone. The CAM-assisted regression framework further illustrates that spatially reweighted thermal information can substantially improve consistency between satellite-derived surface signals and in-situ air temperature observations, highlighting the potential of interpretable machine-learning approaches for urban thermal analysis.

Despite these contributions, several limitations should be acknowledged. First, the proposed regression framework is trained and evaluated within a relatively small and localized learning domain, namely a single university campus. Although the results demonstrate strong internal consistency and clear improvements over conventional LST-based comparisons, the generalizability of the findings to other urban contexts remains uncertain. Differences in urban form, climate background, and surface composition may affect the transferability of the learned surface–air temperature relationships. Future studies should therefore test the framework across multiple sites and cities to evaluate its scalability and robustness.

Second, the current analysis focuses exclusively on temperature-related information derived from LST, without incorporating additional urban morphological or environmental variables. While this design choice allows a clear examination of surface–air thermal coupling, it inevitably limits the predictive scope of the model. Urban air temperature is influenced by a wide range of factors, including building density, surface roughness, vegetation structure, impervious-surface fraction, and topography. Integrating such variables into a multi-source predictive framework represents an important direction for future research and may further enhance both accuracy and physical interpretability.

Third, the division of the study area into three campus zones (Zones A–C) was primarily intended to organize spatial analysis and facilitate comparison across contrasting local environments, rather than to represent formally defined climatic classes. Although the observed thermal differences among zones reflect variations in surface composition, shading, and exposure, the zoning scheme itself is not directly linked to established urban climate classification systems. Future work could adopt standardized frameworks such as Local Climate Zones (LCZ) to provide a more systematic and transferable basis for relating surface–air temperature behavior to urban form (Qiu et al., 2026).

5. CONCLUSIONS

This study examined the spatial and temporal relationships between satellite-derived land surface temperature (LST) and in-situ near-surface air temperature in a heterogeneous campus environment. Rather than treating LST as a direct proxy for air temperature, the analysis focused on spatial structure, relative ranking, diurnal dynamics, and surface–air temperature coupling. The results show that LST exhibits stronger spatial differentiation than near-surface air temperature, especially across sites with different surface materials, solar exposure, and vegetation conditions. Air temperature shows smoother spatial patterns, reflecting the effects of atmospheric mixing and local ventilation. The spatial variation in surface–air temperature differences further indicates that LST–air temperature divergence is systematic and environment-dependent rather than random.

The time-series analysis shows that air temperature heterogeneity is strongly time-dependent. Daytime heating produces clear divergence among monitoring sites, whereas nighttime temperatures tend to converge. The clustering results identify distinct diurnal response modes that are shaped more by site-specific surface and morphological conditions than by broad functional zones. The CAM-assisted regression framework further demonstrates that spatially reweighted LST information improves consistency with in-situ air temperature observations. Grad-CAM results show that the model emphasizes thermally relevant surface components, leading to stronger correlations and reduced dispersion compared with conventional LST-based retrievals. These findings suggest that micro-scale heat assessment should use LST as structured surface thermal information rather than as a direct substitute for near-surface air temperature. With further validation across different urban contexts, the proposed framework can support fine-scale urban climate assessment and climate-sensitive urban design.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (No. 52078389).

REFERENCES

Chao, Z., Wang, L., Che, M., Hou, S., 2020. Effects of Different Urbanization Levels on Land Surface Temperature Change: Taking Tokyo and Shanghai for Example. Remote Sensing 12, 2022. https://doi.org/10.3390/rs12122022

Chen, M., Zhou, Yuan, Hu, M., Zhou, Yaliu, 2020. Influence of Urban Scale and Urban Expansion on the Urban Heat Island Effect in Metropolitan Areas: Case Study of Beijing–Tianjin–Hebei Urban Agglomeration. Remote Sensing 12, 3491. https://doi.org/10.3390/rs12213491

Dai, Z., Guldmann, J.-M., Hu, Y., 2018. Spatial regression models of park and land-use impacts on the urban heat island in central Beijing. Science of The Total Environment 626, 1136–1147. https://doi.org/10.1016/j.scitotenv.2018.01.165

Ke, X., Men, H., Zhou, T., Li, Z., Zhu, F., 2021. Variance of the impact of urban green space on the urban heat island effect among different urban functional zones: A case study in Wuhan. Urban Forestry & Urban Greening 62, 127159. https://doi.org/10.1016/j.ufug.2021.127159

Liu, M., Gou, Z., Zhao, Q., 2025. Enhancing building performance evaluation through Google street View: A multimodal transfer learning framework. Energy and Buildings 344, 115968. https://doi.org/10.1016/j.enbuild.2025.115968

Liu, Z., Fu, L., Wu, C., Zhang, Zhenzhen, Zhang, Zhaoyang, Lin, X., Li, X., Hu, Y., Ge, H., 2023. Spatialized importance of key factors affecting park cooling intensity based on the park scale. Sustainable Cities and Society 99, 104952. https://doi.org/10.1016/j.scs.2023.104952

Qiu, R., Zhan, Q., Tu, S., 2025. Long-term global trends and influencing factors of surface urban cool and heat islands. International Journal of Digital Earth 18, 2489731. https://doi.org/10.1080/17538947.2025.2489731

Qiu, R., Zhan, Q., Tu, S., Li, C., 2026. Interannual variability of surface urban Heat Islands across global cities: Patterns and climatic drivers. Urban Climate 65, 102772. https://doi.org/10.1016/j.uclim.2026.102772

Schmidt, R., Chen, Z., Paz Soldan, V., 2022. Choice Posture, Architecture, and Infrastructure: Systemic Behavioral Design for Public Health Policy. She Ji: The Journal of Design, Economics, and Innovation, Design and Public Health 8, 504–525. https://doi.org/10.1016/j.sheji.2022.08.002

Shao, L., Liao, W., Li, P., Luo, M., Xiong, X., Liu, X., 2023. Drivers of global surface urban heat islands: Surface property, climate background, and 2D/3D urban morphologies. Building and Environment 242, 110581. https://doi.org/10.1016/j.buildenv.2023.110581

Su, R., Yang, C., Xu, Z., Luo, T., Yang, L., Liu, L., Wang, C., 2024. Comparing Characteristics of the Urban Thermal Environment Based on the Local Climate Zone in Three Chinese Metropolises. ISPRS International Journal of Geo-Information 13, 61. https://doi.org/10.3390/ijgi13020061

Sützl B.S., Strebel D.A., Rubin A., Wen J., Carmeliet J., 2024. Urban morphology clustering analysis to identify heat-prone neighbourhoods in cities. Sustainable Cities and Society 107, 105360. https://doi.org/10.1016/j.scs.2024.105360

Tu, S., Zhan, Q., Qiu, R., Li, C., 2025. Spatiotemporal Dynamics and Driving Factors of Urban Expansion in the Chongqing Metropolitan Area Based on Nighttime Light Remote Sensing. Buildings 15, 3306. https://doi.org/10.3390/buildings15183306

Wang, R., Voogt, J., Ren, C., Ng, E., 2022. Spatial-temporal variations of surface urban heat island: An application of local climate zone into large Chinese cities. Building and Environment 222, 109378. https://doi.org/10.1016/j.buildenv.2022.109378

Wang, W., Shu, J., 2020. Urban Renewal Can Mitigate Urban Heat Islands. Geophysical Research Letters 47, e2019GL085948. https://doi.org/10.1029/2019GL085948

Xie, X., Liu, M., Yang, Y., Gou, Z., 2025. Optimizing green exposure for summer outdoor thermal comfort in subtropical campus environments: A machine learning approach integrating climatic and physiological factors. Urban Climate 62, 102518. https://doi.org/10.1016/j.uclim.2025.102518

Yang, F., Yousefpour, R., Zhang, Y., Wang, H., 2023. The assessment of cooling capacity of blue-green spaces in rapidly developing cities: A case study of Tianjin’s central urban area. Sustainable Cities and Society 99, 104918. https://doi.org/10.1016/j.scs.2023.104918

Zheng, S., Chen, X., Liu, Y., 2023. Impact of urban renewal on urban heat island: Study of renewal processes and thermal effects. Sustainable Cities and Society 99, 104995. https://doi.org/10.1016/j.scs.2023.104995