



Regional Urban Heat Island Risk Mapping Using PCA and Fuzzy Clustering

Darniati¹, Muhammad Rizal Haris^{2*}, Muhammad Faisal³, Lisa Fitriani Ishak⁴, Titik Khawa Abdul Rahman⁵

^{1,2,3,4}Department of Informatics, Universitas Muhammadiyah Makassar, Makassar, Indonesia

⁵School of Science and Technology, Asia e University, Selangor, Malaysia

*105841103223@student.unismuh.ac.id

Abstract

Urban Heat Island (UHI) effects pose increasing challenges in rapidly urbanizing tropical regions, where land-cover transformation and population concentration intensify spatial thermal risk. This study proposes an integrated Principal Component Analysis-Fuzzy C-Means spatial framework to classify and map regional UHI risk across Sulawesi Island, Indonesia, using multi-source geospatial data. Land surface temperature, vegetation condition, built-up intensity, and population density were harmonized on a 1 km grid for the 2024-2025 period. Dimensionality reduction was performed to extract dominant spatial variance, followed by fuzzy clustering to represent gradual transitions in UHI risk. The results indicate a strong dominance of the first principal component, explaining more than 99.9% of total variance in both years, reflecting a stable integrated urban-environmental gradient. Three UHI risk classes were delineated, representing low-risk cool areas, moderate-risk transitional zones, and high-risk densely built urban regions. Although a two-cluster solution achieved the lowest numerical validity, a three-cluster configuration was adopted to better capture transitional urban conditions relevant to spatial planning. The resulting maps reveal concentrated UHI risk in major urban centers, particularly Makassar and expanding peri-urban corridors. The proposed framework provides a robust, interpretable, and policy-relevant approach for regional-scale UHI risk assessment in data-limited tropical environments.

Keywords: fuzzy clustering; principal component analysis; remote sensing; spatial risk mapping; urban heat island

How to Cite: D. Darniati, M. R. . Haris, M. Faisal, and L. F. . Ishak, "Regional Urban Heat Island Risk Mapping Using PCA and Fuzzy Clustering", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 10, no. 5, pp. 1538–1549, Oct. 2026, doi: [10.29207/resti.v10i5.7420](https://doi.org/10.29207/resti.v10i5.7420).
Permalink/DOI: <https://doi.org/10.29207/resti.v10i5.7420>

Received: December 26, 2025

Accepted: May 24, 2026

Available Online: October 18, 2026

This is an open-access article under the [CC BY 4.0 License](https://creativecommons.org/licenses/by/4.0/)

Published by [Ikatan Ahli Informatika Indonesia](https://www.iaii.or.id/)

1. Introduction

Urban Heat Island (UHI) refers to the systematic elevation of land surface or near-surface air temperatures within built-up areas relative to surrounding non-urban environments, primarily driven by accelerated urban expansion and anthropogenic land transformation [1]. In rapidly urbanizing tropical regions, impervious surface growth associated with transport networks, dense housing, and commercial development modifies surface biophysical properties, suppresses evapotranspiration, and enhances heat storage and radiative trapping, thereby intensifying thermal stress through surface energy-balance disruption [2].

Prolonged exposure to elevated urban temperatures has been associated with increasing environmental stress and adverse public-health outcomes, particularly in densely populated coastal and low-elevation settings

where high thermal intensity spatially overlaps with population concentration [3], [4]. Under persistently warm and humid tropical climates, UHI patterns tend to be cumulative and spatially stable rather than episodic, emphasizing the need for spatially explicit and regionally adaptive analytical approaches [5]. Remote sensing provides a scalable basis for regional UHI assessment by enabling spatially continuous and temporally consistent observations, especially in data-sparse regions. Long-term thermal patterns can be characterized using MODIS MOD11A2 land surface temperature (LST) products with stable radiometric performance and consistent temporal coverage [6].

Vegetation-mediated cooling can be represented using satellite-derived vegetation indices (e.g., NDVI), calculated from the normalized difference between red and near-infrared (NIR) reflectance, to discriminate vegetated and highly urbanized surfaces [7]. Urban structural intensity and imperviousness can be

quantified using the Global Human Settlement Layer (GHSL) built-up dataset [8], while gridded population datasets (e.g., WorldPop) enable explicit representation of human exposure by identifying the spatial intersection between thermal conditions and population density [9]. Beyond single-product analyses, recent studies emphasize data-centric and scalable urban mapping frameworks that integrate heterogeneous geospatial products into unified analytical structures to improve robustness, transferability, and interpretability across regions [10].

Consequently, UHI is increasingly conceptualized not merely as a thermal anomaly but as a coupled environmental and socio-spatial process governed by interactions among thermal conditions, vegetation dynamics, urban structure, and exposure [11], [12]. Complementing LST-based approaches, machine learning has also been used to enhance spatiotemporal mapping of urban temperature patterns beyond static retrievals [13], while empirical evidence consistently shows that urban vegetation exerts spatially heterogeneous cooling effects that regulate thermal environments across cities [14]. Meanwhile, advances in high-resolution land surface satellite products highlight continuing uncertainties related to atmospheric correction and emissivity, reinforcing the importance of careful product selection and harmonization in regional studies [15]. Empirical findings further indicate that UHI intensity and spatial extent are strongly controlled by land-cover composition, environmental factors, and urban morphology rather than temperature variability alone, consistent across diverse geographic settings [16], [17].

Importantly, exposure cannot be inferred from thermal magnitude alone: population density and compact urban form can amplify cumulative thermal stress through higher exposure and anthropogenic heat contributions [18], and long-term assessments increasingly integrate thermal patterns with population distribution to derive risk-relevant indicators [19]. However, regional-scale UHI assessment commonly involves multivariate datasets with strong inter-variable correlations, spatial heterogeneity, and uncertainty. To address these challenges, dimensionality reduction and soft clustering have been increasingly adopted. Feature-weighted fuzzy clustering developments support more robust clustering under heterogeneous features [20], and improved Fuzzy C-Means variants (including optimization-enhanced formulations) further strengthen clustering stability and partition quality [21].

In parallel, recent hybrid clustering applications in other domains reinforce the value of combining structure extraction and clustering for interpretability and robustness in complex, unlabeled data contexts [22]. Within heat-risk and mitigation studies, mapping frameworks have increasingly shifted toward planning-oriented outputs, including prioritization of UHI mitigation interventions via risk indexing and spatial decision support [23]. For fuzzy partition evaluation,

robust fuzzy cluster validity indices have been proposed to improve reliability in selecting cluster configurations under spatial complexity [24]. At broader scales, evidence also shows that urban sprawl form and growth dynamics shape thermal environments and modulate heat patterns during hot seasons [25], [26], while hourly analyses confirm that urban features influence land surface temperature distributions and reveal strong diurnal variability that static approaches may under-represent [27].

Recent advances in deep supervised and self-supervised learning for remote sensing feature extraction indicate continuing opportunities to improve representation learning in high-dimensional satellite data environments [28]. In exposure research, multi-source, high-temporal-resolution mapping demonstrates that heat exposure varies across population groups and can remain high even under moderate thermal intensity when population concentration is considered [29]. Broader Earth observation literature continues to highlight persistent challenges in scale dependency, data integration, and generalizability across heterogeneous urban systems, reinforcing the need for methods that remain interpretable and transferable across regions [30]. In Indonesia, UHI studies remain relatively fragmented and often city-centric, limiting regional-scale continuity across coastal, inland, and mountainous environments; evidence from Jakarta further suggests that volumetric built-up growth is closely tied to urban environmental impacts [5].

Despite growing advances in UHI research, several limitations remain. Many existing studies focus primarily on individual cities, limiting understanding of spatial heat-risk continuity across broader regional systems where urban, peri-urban, and rural environments interact. Moreover, numerous assessments rely mainly on thermal indicators without sufficiently integrating ecological and socio-demographic variables that shape heat exposure. In addition, relatively few studies employ interpretable hybrid analytical frameworks capable of simultaneously addressing multicollinearity among environmental indicators while capturing gradual spatial transitions in UHI risk across heterogeneous landscapes.

Additionally, built-up area and population density have been emphasized as essential societal variables for interpreting climate-hazard impacts, supporting the need to jointly incorporate imperviousness and exposure in risk frameworks [31]. Planning-oriented comparisons also show that priority zones for urban heat mitigation may differ depending on whether risk or shelter provision is emphasized, indicating that risk classification should be explicitly designed to support policy and spatial planning interpretation [32]. These gaps motivate a regional-scale hybrid framework for Sulawesi that integrates LST, vegetation, built-up intensity, and population exposure within a harmonized spatial structure, applies PCA to reduce

multicollinearity while preserving dominant spatial variability, and employs FCM to delineate low-, moderate-, and high-risk zones while representing gradual spatial transitions. By operationalizing UHI risk as a continuous, spatially heterogeneous process, the proposed PCA–FCM framework supports interpretable regional risk mapping and planning-relevant decision support in tropical archipelagic environments.

2. Methods

2.1 Research Framework

This study applies a hybrid PCA–FCM framework to classify regional-scale UHI risk across Sulawesi Island by integrating thermal conditions, vegetation dynamics, urban structural intensity, and population exposure into a unified analytical structure. The overall methodological workflow is illustrated in Figure 1,

outlining the sequential integration of multi-source datasets, dimensionality reduction, fuzzy clustering, cluster validation, and spatial reconstruction. In this study, dimensionality reduction is necessary to mitigate multicollinearity among thermal, vegetation, built-up, and population indicators while extracting the dominant latent structure that represents the integrated urbanization–exposure gradient underlying regional UHI risk patterns.

This framework is designed to preserve dominant spatial variance while representing gradual transitions in UHI risk, which are critical for regional planning applications. By structuring the analysis within a unified and reproducible workflow, the proposed framework ensures methodological consistency.

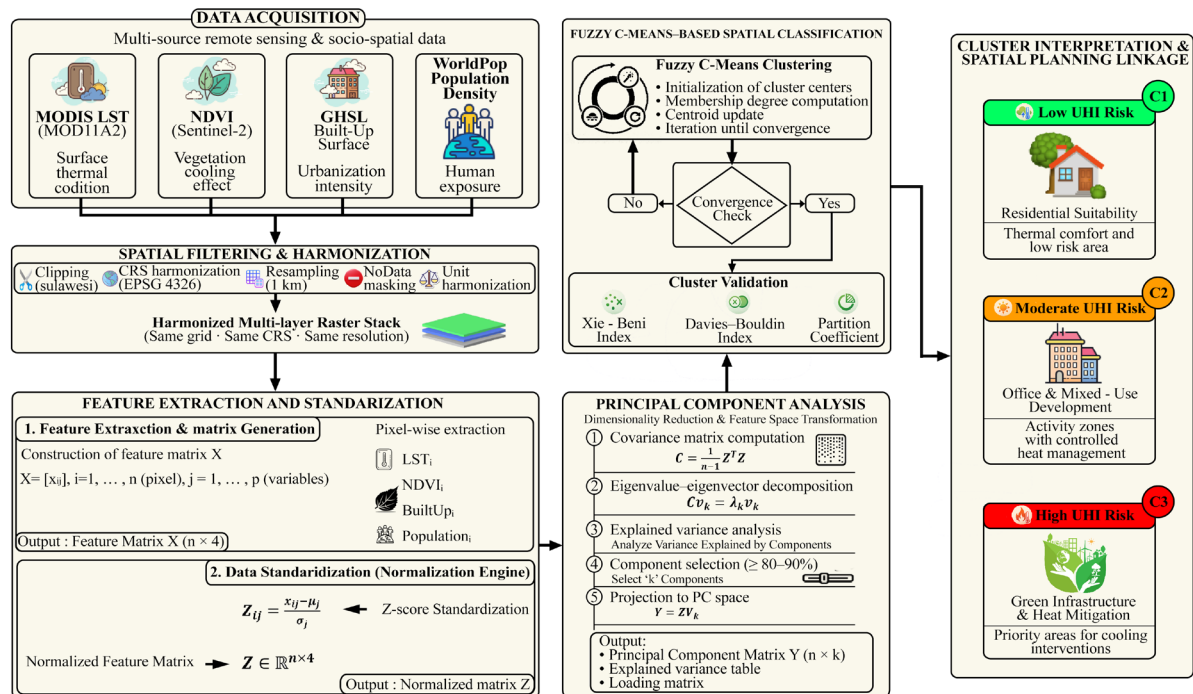


Figure 1. Proposed Integrated PCA–FCM Framework for Regional Urban Heat Island (UHI) Risk Assessment Integrating Multi-Source Geospatial Indicators.

Figure 1 illustrates the integrated workflow combining multi-source remote sensing indicators and population exposure data through spatial harmonization, PCA-based dimensionality reduction, and FCM clustering to delineate low-, moderate-, and high-UHI risk zones.

2.2 Dataset Acquisition

Thermal, ecological, structural, and exposure-related variables were integrated to represent the primary drivers of UHI risk, including land surface temperature, vegetation condition, built-up intensity, and population density [6-9]. A summary of the datasets and their spatial characteristics is provided in Table 1. All raster layers were clipped to the Sulawesi Island boundary, reprojected to a unified coordinate reference system, and resampled to a common 1 km spatial resolution to

ensure spatial consistency. Population density was incorporated only for grid cells with valid WorldPop observations, while pixels with missing population values primarily corresponding to sparsely inhabited or non-settlement areas were excluded from subsequent multivariate analysis. This treatment ensures that population exposure is represented consistently within urban and peri-urban contexts without biasing the PCA–FCM results. As shown in Table 1, the statistical distribution of each input variable is summarized for the 2024 dataset, the 2025 dataset, and the pooled dataset used for the integrated analysis. Q1 and Q3 denote the first and third quartiles representing the lower and upper bounds of the interquartile range, respectively, thereby characterizing data dispersion and variability prior to multivariate analysis.

Table 1. Statistical summary of input datasets (2024-2025)

Variable	N valid	Missing	Min	Q1	Median	Mean	Q3	Max	Std	Dataset
LST	212.279	0	12.5008	24.1129	26.2141	25.8944	28.0364	38.8763	3.0899	2024
LST	212.218	0	11.5901	23.2433	25.5194	25.0498	27.3260	36.8216	3.1131	2025
LST	424.497	0	11.5901	23.6614	25.8747	25.4721	27.6677	38.8763	3.1301	Pooled
NDVI	212.279	0	-0.1189	0.7673	0.8181	0.7932	0.8487	0.9193	0.0890	2024
NDVI	212.218	0	-0.1048	0.7678	0.8155	0.7917	0.8461	0.9195	0.0892	2025
NDVI	424.497	0	-0.1189	0.7675	0.8168	0.7925	0.8474	0.9195	0.0891	Pooled
BUILTUP	212.279	0	0	0	0	53.2560	2.8678	5053.3960	213.0145	2024
BUILTUP	212.218	0	0	0	0	53.2465	2.8593	5053.3960	213.0173	2025
BUILTUP	424.497	0	0	0	0	53.2512	2.8644	5053.3960	213.0157	Pooled
POPULATION	111.447	100.832	0	2.6331	15.0096	166.7828	117.0323	17506.2363	609.7069	2024
POPULATION	111.404	100.814	0	2.6377	15.0713	168.1338	117.8429	17645.0508	614.8622	2025
POPULATION	222.851	201.646	0	2.6352	15.0412	167.4581	117.4530	17645.0508	612.2885	Pooled

2.3 Feature Construction and Standardization

For each valid pixel, land surface temperature, vegetation index, built-up intensity, and population density values were combined into a feature vector defined in Equation 1.

$$x_i = [LST_i, NDVI_i, BuiltUp_i, Pop_i] \quad (1)$$

All pixel-level vectors were stacked into a feature matrix as expressed in Equation 2.

$$X = [x_1, x_2, \dots, x_n] \quad (2)$$

To eliminate scale effects among variables with different units, z-score standardization was applied according to Equation 3.

$$z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (3)$$

This standardization ensures numerical stability and equal contribution of all variables during dimensionality reduction and clustering.

2.4 Principal Component Analysis

PCA was employed to reduce dimensionality and mitigate multicollinearity among standardized variables [12], as is shown in Equation 4.

$$C = \frac{1}{n-1} Z^T Z \quad (4)$$

Eigenvalue decomposition of the covariance matrix yielded the principal components Equation 5.

$$C v_k = \lambda_k v_k \quad (5)$$

The standardized data were projected onto the reduced component space using Equation 6.

$$Y = Z V_k \quad (6)$$

Components were retained based on cumulative explained variance to preserve dominant spatial variability while ensuring compact representation.

2.5 Fuzzy C-Means Clustering

FCM clustering was applied to the PCA-transformed feature space to classify UHI risk while accounting for gradual spatial transitions [33], [24]. The FCM objective function minimized during clustering is defined in Equation 7.

$$J_m = \sum_{i=1}^n \sum_{k=1}^c u_{ik}^m \|y_i - v_k\|^2 \quad (7)$$

Cluster centroids and membership degrees were updated iteratively using Equation 8.

$$u_{ik} = \left(\sum_{j=1}^c \left(\frac{\|y_i - v_k\|}{\|y_i - v_j\|} \right)^{\frac{2}{m-1}} \right)^{-1} \quad (8)$$

2.6 Cluster Validation

Cluster robustness was evaluated using the Xie-Beni index, defined in Equation 9, which measures the balance between intra-cluster compactness and inter-cluster separation [24]:

$$XB = \frac{\sum_{i=1}^n \sum_{k=1}^c u_{ik}^m \|y_i - v_k\|^2}{n \cdot \min_{k \neq j} \|v_k - v_j\|^2} \quad (9)$$

Lower XB values indicate improved clustering quality. These spatial patterns were interpreted as indicators of low-, moderate-, and high-UHI risk to support climate-adaptive spatial planning at the regional scale.

3. Results and Discussions

3.1 PCA Structure and Dominant Variance Patterns

PCA was applied to the standardized multi-source dataset to extract dominant latent structures governing UHI risk across Sulawesi Island. The explained variance and cumulative variance of the retained principal components for the 2024 and 2025 datasets are reported separately in Tables 2 and 3, respectively.

The corresponding variance distribution patterns are further examined using comparative scree plots presented in Figure 2. To assess the robustness of dimensionality reduction, interrelationships among the original PCA input variables were examined using Pearson correlation analysis, as summarized independently for 2024 in Table 4 and 2025 in Table 5, thereby supporting the interpretability and statistical stability of the extracted principal components.

As shown in Table 2, the PCA results for 2024 indicate a pronounced dominance of the first principal component (PC1), which explains more than 99.9% of the total variance. A nearly identical variance structure is observed for the 2025 dataset Table 3, confirming

strong temporal stability of the extracted latent pattern. This interannual consistency suggests that regional UHI variability is primarily governed by persistent land-

cover composition, urban structural intensity, and population distribution rather than short-term thermal fluctuations.

Table 2. Explained Variance of Principal Components 2024

PC	Eigenvalue	Variance Explained	Variance Explained (%)	Cumulative Variance	Cumulative Variance (%)
PC1	371744.106526	0.9999951	99.999512	0.999995	99.999512
PC2	1.078826	0.000002902053	0.0002902053	0.999998	99.999802
PC3	0.408568	0.000001099051	0.0001099051	0.999999	99.999912
PC4	0.325875	0.0000008766057	0.00008766057	1.000000	100.000000

Table 3. Explained Variance of Principal Components 2025

PC	Eigenvalue	Variance Explained	Variance Explained (%)	Cumulative Variance	Cumulative Variance (%)
PC1	378057.0934	0.999995286	99.99952858	0.999995286	99.99952858
PC2	1.014994351	0.00000268475	0.000268475	0.999997971	99.99979706
PC3	0.406617219	0.00000107554	0.000107554	0.999999046	99.99990461
PC4	0.360614999	0.000000954	0.0000954	1.000000000	100.00000000

Examination of the PCA loading structure indicates that PC1 is characterized by strong positive contributions from land surface temperature, built-up intensity, and population density, accompanied by a negative contribution from NDVI. This loading configuration confirms that PC1 represents an integrated urbanization–exposure gradient, reflecting the combined influence of impervious surface dominance, vegetation suppression, and human concentration on surface thermal accumulation. The exceptionally high variance captured by PC1 >99.9% reflects strong interdependence among the input variables, reinforced by regional-scale spatial aggregation and the temporal stability of urban structural patterns during 2024–2025, while higher-order components capture only marginal residual variability.

Consistently, the Pearson correlation matrices for 2024 Table 4 and 2025 Table 5 reveal strong correlations among land surface temperature, built-up intensity, vegetation condition, and population density, confirming a high degree of multicollinearity in the original feature space prior to dimensionality reduction and justifying the subsequent application of PCA-based feature compression. Although the number of input variables is relatively limited, these variables exhibit strong interdependence and represent different dimensions of the same urbanization–exposure system; therefore, PCA is applied to consolidate correlated indicators into a single dominant latent structure that improves interpretability and numerical stability for subsequent clustering analysis.

Table 4. Pearson Correlation Matrix of PCA Input Variables (2024)

PC	Eigenvalue	Variance Explained	Variance Explained (%)	Cumulative Variance	Cumulative Variance (%)
PC1	371744.106526	0.9999951	99.999512	0.999995	99.999512
PC2	1.078826	0.000002902053	0.0002902053	0.999998	99.999802
PC3	0.408568	0.000001099051	0.0001099051	0.999999	99.999912
PC4	0.325875	0.0000008766057	0.00008766057	1.000000	100.000000

Table 5. Pearson Correlation Matrix of PCA Input Variables (2025)

PC	Eigenvalue	Variance Explained	Variance Explained (%)	Cumulative Variance	Cumulative Variance (%)
PC1	371744.106526	0.9999951	99.999512	0.999995	99.999512
PC2	1.078826	0.000002902053	0.0002902053	0.999998	99.999802
PC3	0.408568	0.000001099051	0.0001099051	0.999999	99.999912
PC4	0.325875	0.0000008766057	0.00008766057	1.000000	100.000000

A comparable correlation structure is observed for the 2025 dataset in Table 5, further confirming the temporal stability of inter-variable relationships and empirically justifying the application of PCA to consolidate redundant information into a reduced set of orthogonal components [16]. To further substantiate the dominance of PC1 and the consistency of variance structure across years, Figure 2 presents a comparative PCA scree plot analysis for the 2024 and 2025 datasets.

The dominance of PC1 is further confirmed by the scree and cumulative variance patterns illustrated in Figure 2, where a sharp inflection occurs immediately after the first component for both years. Higher-order components contribute only marginal residual variance,

indicating limited additional explanatory power at the regional scale [23]. To examine the spatial configuration and separability of observations in the reduced feature space, Figure 3 illustrates the distribution of PCA scores projected onto the first two principal components (PC1–PC2) for the integrated 2024–2025 dataset.

The structural configuration of the transformed data is illustrated in Figure 3, which shows an elongated distribution of observations along the PC1 axis within the PC1–PC2 feature space. This pattern confirms that the primary spatial contrast underlying UHI variability is captured by PC1, while higher-order components play a secondary role in structuring regional thermal

heterogeneity. The dominant variance patterns reported in Tables 2 and 3, the inter-variable relationships documented in Tables 4 and 5, and the variance and feature-space structures illustrated in Figures 2 and 3 collectively demonstrate that UHI risk across Sulawesi

is governed by a single, stable latent gradient. This robust PCA structure provides a statistically sound and physically interpretable basis for subsequent fuzzy clustering analysis.

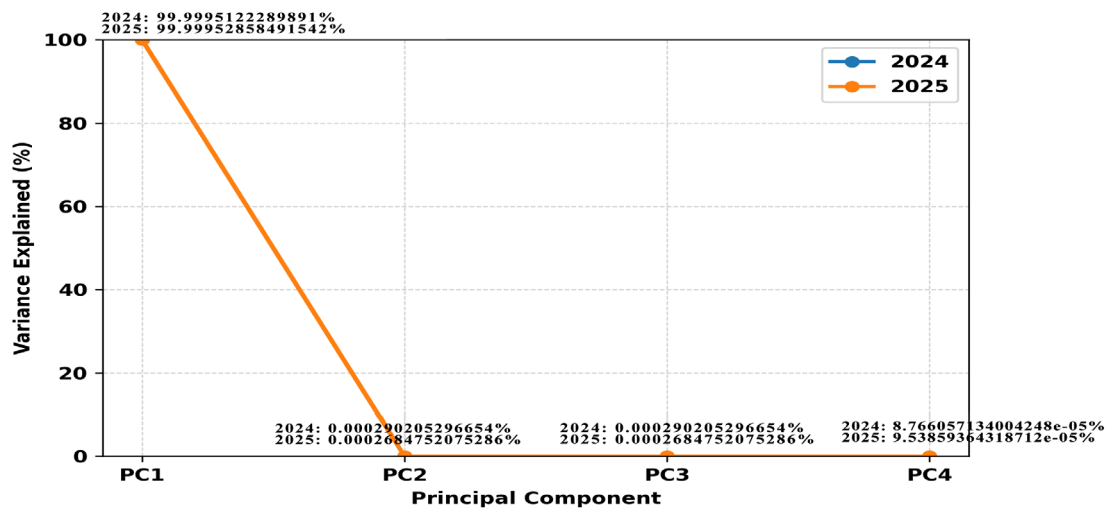


Figure 2.

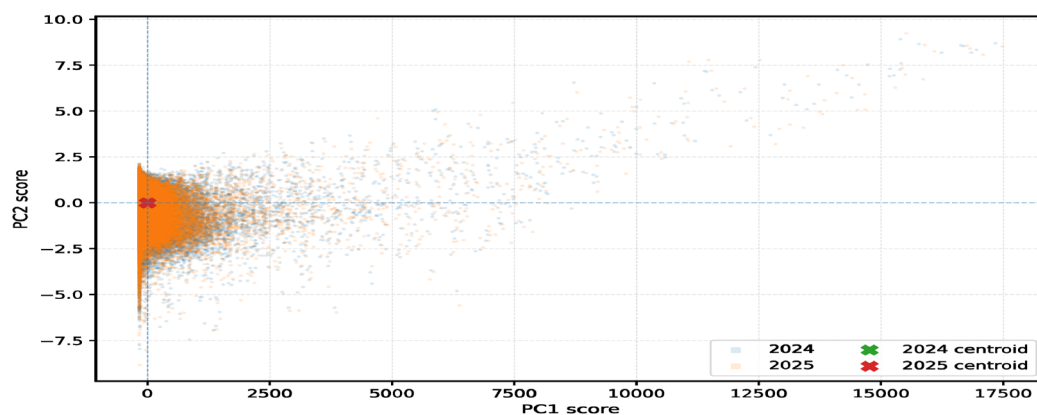


Figure 3. PCA feature-space structure based on PC1 and PC2 scores for the integrated 2024-2025 dataset

3.2 Fuzzy C-Means Clustering and UHI Risk Classification

Building upon the dominant variance structure identified through PCA, fuzzy clustering was applied to the reduced feature space to classify Urban Heat Island risk while explicitly accounting for gradual spatial transitions and uncertainty in cluster membership.

The PCA-transformed feature space was clustered using the Fuzzy C-Means (FCM) algorithm, which assigns each spatial unit degrees of membership to multiple clusters rather than enforcing hard boundaries. This property is particularly suitable for UHI analysis, where thermal risk evolves continuously across urban-peri-urban gradients rather than exhibiting abrupt spatial transitions [33]. To determine the optimal number of clusters, candidate solutions ranging from $c = 2$ to $c = 6$ were evaluated using the Xie-Beni (XB) validity index, which jointly assesses cluster compactness and separation. The resulting XB values for each candidate cluster configuration are

summarized in Table 6, providing an objective basis for subsequent cluster selection.

As shown in Table 6, the minimum XB value corresponds to the two-cluster solution, indicating the most compact mathematical partition. However, a three-cluster configuration was selected for subsequent analysis to explicitly represent transitional thermal environments that are critical for spatial planning, rather than collapsing intermediate-risk areas into binary classes [34].

Table 6. Xie-Beni (XB) Index for Candidate Numbers of Fuzzy Clusters ($c = 2-6$)

C	XB	iters	last diff
2	0.00378261	59	0.00000076
3	0.05336307	83	0.00000091
4	0.08298399	141	0.00000097
5	0.07235019	171	0.00000096
6	0.09100003	167	0.00000092

The three fuzzy clusters were interpreted as low-, moderate-, and high-UHI risk zones, with their

statistical characteristics derived from the pooled 2024-2025 dataset summarized in Table 7. High-risk clusters exhibit elevated land surface temperature, high built-up

intensity, and dense population concentration, while low-risk clusters are associated with stronger vegetation influence and lower anthropogenic pressure.

Table 7. Statistical Characteristics of UHI Risk Clusters Based on Pooled Data (2024-2025)

Cluster	UHI Risk Level	N	Mean LST	Mean NDVI	Mean BuiltUpDensity	Mean PopDensity
C1	Low	208931	26.993	0.782	53.608	71.188
C2	High	946	32.293	0.479	2772.239	7417.768
C3	Moderate	12974	29.233	0.679	675.240	1189.116

To further examine classification uncertainty, the distributions of fuzzy membership values were analyzed. Descriptive statistics for 2024 and 2025 are reported separately in Tables 8 and 9, indicating

stronger membership dominance in urban cores and more distributed membership patterns in peri-urban and mixed-use zones.

Table 8. Descriptive Statistics of Fuzzy Membership Values for UHI Risk Clusters (2024)

Cluster	Mean Membership	Std Membership	Min Membership	Max Membership
C1	0.33333313	0.469698161	0.000051	0.997585475
C2	0.33333343	0.469237506	0.000066	0.996933162
C3	0.33333343	0.469434381	0.000060	0.997212052

Table 9. Descriptive Statistics of Fuzzy Membership Values for UHI Risk Clusters (2025)

Cluster	Mean Membership	Std Membership	Min Membership	Max Membership
C1	0.33333343	0.469700515	0.000052	0.997588813
C2	0.33333313	0.469241083	0.000066	0.996938288
C3	0.33333343	0.469437927	0.000060	0.997217059

Concentrated high-UHI risk in major urban centers and expanding moderate-risk zones along peri-urban corridors. This spatial configuration is consistent with findings linking urban densification, population exposure, and built-up morphology to intensified thermal risk [34-37]. To translate the PCA-FCM clustering results into spatially explicit representations, the derived UHI risk classes were mapped across Sulawesi to examine their regional and local spatial configurations.

Figure 4 illustrates the spatial distribution of low-, moderate-, and high-UHI risk clusters, highlighting both island-scale patterns and city-level contrasts within major urban centers.

The FCM-based classification translates the dominant PCA-derived latent gradient into spatially explicit and policy-relevant UHI risk zones, providing a robust basis for interpreting regional thermal patterns and supporting targeted mitigation strategies.

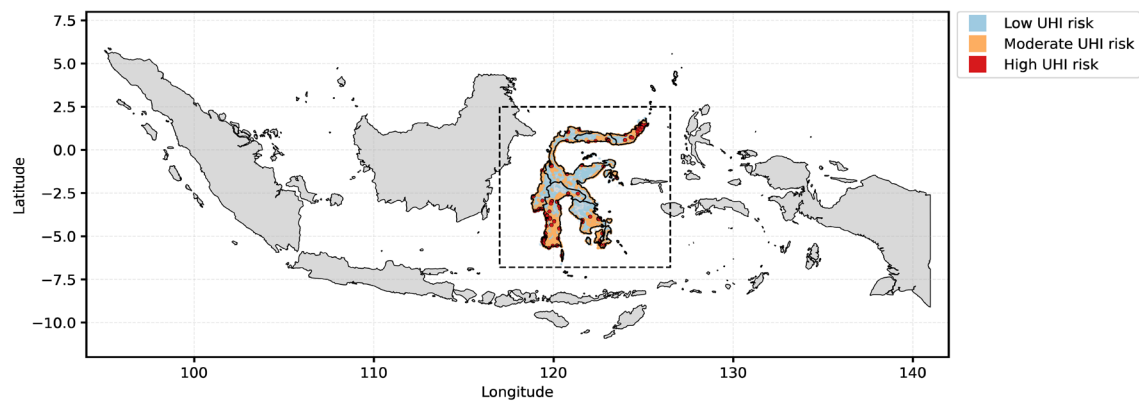
3.3 Spatial Distribution of UHI Risk and Planning Implications

The spatial distribution of UHI risk clusters derived from the PCA-FCM framework is illustrated in Figure 4, revealing clear regional differentiation across Sulawesi. High-UHI risk zones are consistently concentrated in major urban centers, including Makassar, Manado, Kendari, Palu, and Gorontalo, where dense built-up morphology, limited vegetation cover, and high population concentration coincide. This spatial coherence with known urban cores supports the physical plausibility of the classification and reflects the combined effects of urban densification, volumetric growth, and intensified anthropogenic heat accumulation in tropical coastal cities [38], [39].

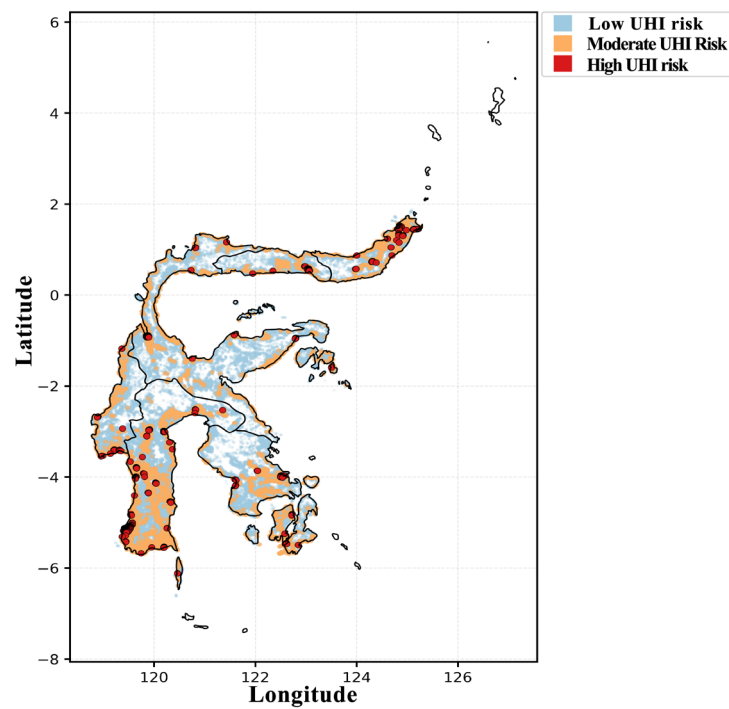
Moderate-UHI risk clusters form continuous peri-urban belts surrounding urban cores and extending along transportation corridors. These transitional zones exhibit mixed land-cover composition and intermediate population density, indicating increasing exposure as urban expansion progresses, consistent with reported urban-peri-urban thermal gradients in rapidly urbanizing regions [40]. Low-UHI risk clusters are primarily associated with vegetated, less urbanized, or topographically elevated areas, underscoring the cooling role of green infrastructure and land-cover heterogeneity in moderating surface thermal conditions [41].

Similar cooling effects of vegetation-dominated landscapes have been widely reported across tropical and subtropical urban contexts [42], [43]. From a planning perspective, the spatially explicit UHI risk classification provides a practical basis for differentiated heat mitigation strategies. High-risk urban cores require immediate interventions focused on urban greening, surface material modification, and heat-resilient urban design, while moderate-risk peri-urban zones offer strategic opportunities for preventive planning through controlled development and green corridor integration [44], [45].

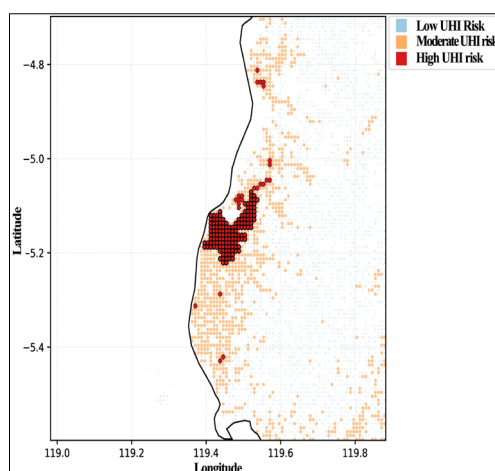
The fuzzy nature of the classification further supports adaptive planning by explicitly capturing gradual spatial transitions rather than enforcing rigid zoning boundaries. Collectively, these spatial patterns highlight the progressive outward diffusion of UHI risk from urban cores toward surrounding transitional landscapes and adjacent peri-urban zones.



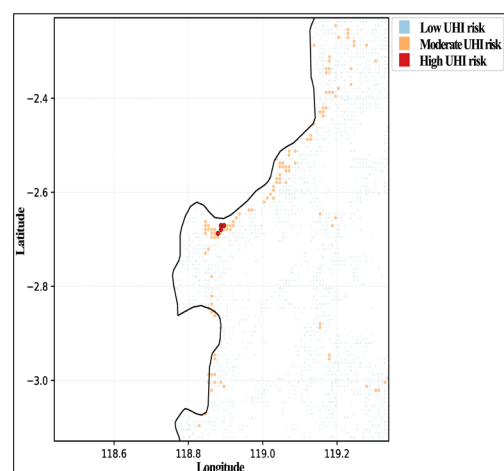
(a)



(b)



(c)



(d)

Figure 4. Spatial Patterns of Urban Heat Island (UHI) Risk Clusters Across Sulawesi (2024-2025): (a) Geographic context of the study area showing the location of Sulawesi within the Indonesian archipelago, (b) Local Spatial Pattern of UHI Risk Clusters in South Sulawesi, (c) UHI Risk Clusters in Makassar (South Sulawesi), (d) UHI Risk Clusters in Mamuju (West Sulawesi), (e) UHI Risk Clusters in Palu (Central Sulawesi), (f) UHI Risk Clusters in Manado (North Sulawesi), (g) UHI Risk Clusters in Kendari (Southeast Sulawesi), (h) UHI Risk Clusters in Gorontalo (Gorontalo)

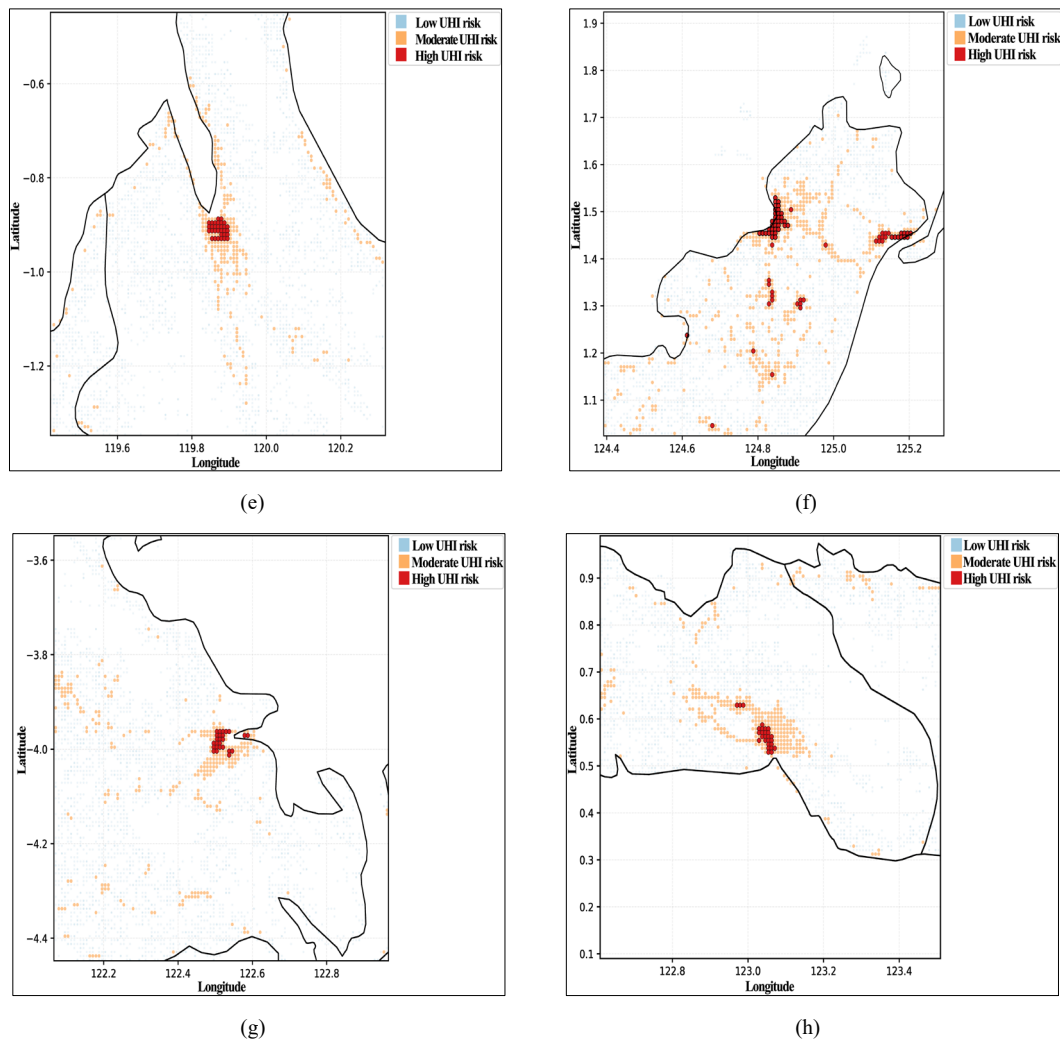


Figure 4. (continued)

3.4 Temporal Consistency and Interannual Stability of UHI Risk Patterns

To assess temporal robustness, UHI risk patterns derived from the 2024 and 2025 datasets were compared at both statistical and spatial levels. The results indicate strong interannual consistency, with negligible changes in the dominant variance structure, cluster configuration, and spatial distribution of risk zones.

The persistence of a strongly dominant first principal component across both years confirms that the integrated thermal-vegetation-built-up-population gradient remains structurally stable over time. Consistently, the fuzzy clustering results preserve clear separation among low-, moderate-, and high-risk classes, indicating that the identified patterns are not driven by short-term anomalies but reflect underlying urban form and land-cover characteristics. Spatially, high-UHI risk zones remain concentrated in major urban centers, while moderate-risk belts consistently occupy peri-urban and transitional areas. This stability aligns with prior evidence that urban densification, population exposure, and land-cover configuration

generate cumulative and temporally persistent thermal patterns, particularly in rapidly urbanizing tropical regions [39], [40]. The observed interannual coherence strengthens confidence in the reliability and generalizability of the PCA-FCM framework for regional-scale UHI risk assessment, supporting its application for medium-term monitoring and comparative analysis across time.

3.5 Methodological Strengths and Limitations

The proposed PCA-FCM framework exhibits several methodological strengths for regional-scale UHI risk assessment. First, the integration of PCA effectively mitigates multicollinearity among thermal, vegetation, built-up, and population variables while preserving the dominant variance structure, thereby enhancing interpretability and numerical stability. Second, the use of fuzzy clustering explicitly represents gradual spatial transitions and uncertainty in UHI risk classification, which is more consistent with the continuous nature of urban thermal processes than hard partitioning approaches. These strengths enable the framework to capture complex interactions between urban form, densification patterns, and environmental drivers that

shape spatial heat accumulation [34], [36]. Despite these advantages, certain limitations should be acknowledged.

The analysis relies on satellite-derived land surface temperature and aggregated ancillary datasets, which may not fully capture fine-scale microclimatic variability or short-term atmospheric dynamics. In addition, while the PCA-FCM approach effectively characterizes spatial structure, it does not explicitly model causal mechanisms linking urban morphology to thermal responses. As highlighted in recent studies, integrating explanatory or machine-learning-based models could further enhance process-level understanding of the interactions between urban form, air quality, and heat-related outcomes [37]. These limitations do not undermine the validity of the results but rather delineate the scope of applicability of the proposed framework. The PCA-FCM approach is best suited for comparative regional assessment, spatial screening, and planning-oriented UHI risk mapping, while more detailed process-based analyses can be incorporated in future extensions.

3.6 Implications for Heat Mitigation and Future Research

The spatially explicit UHI risk patterns derived from the PCA-FCM framework provide actionable insights for regional heat mitigation and climate-adaptive planning. High-risk urban cores identified in this study require priority interventions focused on expanding and optimizing urban green spaces, improving surface material reflectivity, and integrating heat-resilient urban design principles. Empirical evidence consistently demonstrates that well-distributed and context-sensitive green infrastructure can significantly reduce urban thermal stress and associated health risks, particularly under persistent heat conditions [41], [43]. Moderate-risk peri-urban zones represent strategic intervention areas where preventive planning can effectively limit future UHI intensification.

Controlled urban expansion, preservation of green corridors, and early integration of heat mitigation measures in developing areas can prevent these transitional zones from evolving into high-risk heat hotspots. The fuzzy nature of the classification further supports adaptive planning by highlighting areas with mixed risk characteristics, enabling flexible and phased mitigation strategies rather than rigid zoning. From a research perspective, the proposed framework can be extended by incorporating additional dimensions such as urban energy consumption, socio-economic vulnerability, and policy-driven land-use scenarios. Hybrid decision-support approaches that integrate multi-objective optimization and spatial decision-making techniques may further enhance the translation of UHI risk assessment into actionable urban policy instruments [46]. In addition, hybrid multi-objective decision-support approaches that combine spatial clustering outputs with policy-oriented evaluation frameworks can further strengthen the translation of

UHI risk mapping into actionable urban and environmental policy instruments [47]. Future studies may also explore finer spatial resolutions and longer temporal horizons to better capture microclimatic variability and long-term urban transformation processes. The PCA-FCM-based UHI risk mapping framework offers a scalable and interpretable foundation for supporting heat mitigation planning and guiding future research on urban climate resilience in rapidly urbanizing tropical regions.

4. Conclusions

This study demonstrates that the integrated PCA-FCM framework provides a robust and interpretable approach for regional-scale Urban Heat Island (UHI) risk assessment across Sulawesi by effectively reducing multicollinearity among thermal, vegetation, built-up, and population variables while capturing gradual spatial transitions between low-, moderate-, and high-risk zones. The results confirm that dominant UHI risk patterns are governed by persistent urban structural intensity, vegetation suppression, and population concentration rather than short-term thermal variability, enabling a spatially consistent characterization of UHI risk across heterogeneous tropical urban contexts. Methodologically, the proposed hybrid framework advances regional UHI analysis by combining dimensionality reduction and fuzzy clustering to preserve physically meaningful variance structures and represent transitional urban-peri-urban environments. Practically, the derived UHI risk typology provides policy-relevant spatial information that supports heat mitigation prioritization and climate-adaptive urban planning. Future research may enhance this framework by integrating temporal modeling and higher-resolution remote sensing products to better capture intra-urban dynamics and long-term thermal evolution, as well as by embedding PCA-FCM outputs into decision-support systems to strengthen data-driven urban climate governance in rapidly urbanizing tropical regions.

Declarations

AI Usage Statement

The authors used a generative artificial intelligence tool, namely ChatGPT, solely to assist with English-language editing, grammatical refinement, text organization, and improving the manuscript's clarity during the writing process.

Author Contributions

Conceptualization, D. and M.R.H.; methodology, M.F., D. and M.R.H.; software, M.R.H., L.F.I.; validation, D., M.F., and T.K.A.R.; formal analysis, M.F. and T.K.A.R.; investigation, M.R.H., D. and T.K.A.R.; resources, D., M.F.; data curation, L.F.I.; writing original draft preparation, D.; writing review and editing, M.R.H., M.F.; visualization, L.F.I., T.K.A.R.; supervision, T.K.A.R.; project administration, M.F.; funding acquisition, M.F., L.F.I., and D. All authors

have read and agreed to the published version of the manuscript.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interest Statement

The authors declare no conflict of interest.

Ethical Clearance

Ethical clearance was not required for this study because it relied exclusively on secondary geospatial data, satellite-derived products, and aggregated population data, without direct involvement of human participants or animal subjects. The study was conducted in accordance with applicable ethical standards.

Data and Code Availability

The datasets used in this study are publicly available from their respective data providers, including NASA's MODIS products, the Global Human Settlement Layer (GHSL), and WorldPop. The code and processed datasets generated during this study are available from the corresponding author upon reasonable request.

References

- [1] L. Yao, S. Sun, C. Song, Y. Wang, and Y. Xu, "Recognizing surface urban heat 'island' effect and its urbanization association in terms of intensity, footprint, and capacity: A case study with multi-dimensional analysis in Northern China," *J. Clean. Prod.*, vol. 372, p. 133720, 2022, doi: <https://doi.org/10.1016/j.jclepro.2022.133720>.
- [2] R. M. Abou Samra, "Investigating and mapping day-night urban heat island and its driving factors using Sentinel/MODIS data and Google Earth Engine. Case study: Greater Cairo, Egypt," *Urban Clim.*, vol. 52, p. 101729, 2023, doi: <https://doi.org/10.1016/j.uclim.2023.101729>.
- [3] C. Zhang, Y. Yang, and L. Yu, "Assessing urban surface thermal environment and heat health risk in Chinese cities: A twenty-year study," *Urban Clim.*, vol. 59, p. 102304, 2025, doi: <https://doi.org/10.1016/j.uclim.2025.102304>.
- [4] R. Jia et al., "Population heat exposure risk from the perspective of urban heat island spatial expansion in China during 2005–2020," *Urban Clim.*, vol. 56, p. 101987, 2024, doi: <https://doi.org/10.1016/j.uclim.2024.101987>.
- [5] T. Sarker, P. Fan, J. P. Messina, N. Mujahid, E. Aldrian, and J. Chen, "Impact of Urban built-up volume on Urban environment: A Case of Jakarta," *Sustain. Cities Soc.*, vol. 105, p. 105346, 2024, doi: <https://doi.org/10.1016/j.scs.2024.105346>.
- [6] R. E. N. Rongcai, W. U. Guoxiong, and C. A. I. Ming, "MOD11A2 MODIS/Terra Suhu/Emisivitas Permukaan Tanah 8 Hari L3 Global 1km SIN Grid V006," 2015, doi: <https://doi.org/10.5067/MODIS/MOD11A2.006>.
- [7] R. E. N. Rongcai, W. U. Guoxiong, and C. A. I. Ming, "MODIS/Terra Vegetation Indices Monthly L3 Global 1km SIN Grid V061," 2021, doi: <https://doi.org/10.5067/MODIS/MOD13A3.061>.
- [8] E. C. J. R. Centre, *GHSL data package 2023*. Publications Office of the European Union, 2023. doi: 10.2760/098587.
- [9] Kholif, "Constrained estimates of 2015-2030 total number of people per grid square at a resolution of 3 arc (approximately 100m at the equator) R2025A version v1," p. 6, 2025, doi: <https://hub.worldpop.org/doi/10.5258/SOTON/WP00839>.
- [10] Y. Zhou and Q. Weng, "Building up a data engine for global urban mapping," *Remote Sens. Environ.*, vol. 311, p. 114242, 2024, doi: <https://doi.org/10.1016/j.rse.2024.114242>.
- [11] F. Hung, M. Aboelnour, D. Wood, and J. Zhang, "Urban heat risks reimagined: A framework for global physical hazard assessment," *Urban Clim.*, vol. 64, p. 102659, 2025, doi: <https://doi.org/10.1016/j.uclim.2025.102659>.
- [12] Z. Zhu, S. Qiu, and S. Ye, "Remote sensing of land change: A multifaceted perspective," *Remote Sens. Environ.*, vol. 282, p. 113266, 2022, doi: <https://doi.org/10.1016/j.rse.2022.113266>.
- [13] P. Shen, "Spatiotemporal mapping of urban air temperature and UHI under TMY condition: A reference station based machine learning approach," *Energy Build.*, vol. 343, p. 115923, 2025, doi: <https://doi.org/10.1016/j.enbuild.2025.115923>.
- [14] A. Wang, Y. Dai, M. Zhang, and E. Chen, "Exploring the cooling intensity of green cover on urban heat island: A case study of nine main urban districts in Chongqing," *Sustain. Cities Soc.*, vol. 124, p. 106299, 2025, doi: <https://doi.org/10.1016/j.scs.2025.106299>.
- [15] S. Liang et al., "Advancements in high-resolution land surface satellite products: A comprehensive review of inversion algorithms, products and challenges," *Sci. Remote Sens.*, vol. 10, p. 100152, 2024, doi: <https://doi.org/10.1016/j.srs.2024.100152>.
- [16] I. Onjel, V. Amihăse, D. Micu, A. Dumitrescu, and S. Cheval, "Influence of environmental factors on land surface temperature and surface urban heat island. A cross-country analysis in Romania," *Sustain. Cities Soc.*, vol. 128, p. 106454, 2025, doi: <https://doi.org/10.1016/j.scs.2025.106454>.
- [17] Z. Lin, H. Xu, X. Yao, C. Yang, and D. Ye, "How does urban thermal environmental factors impact diurnal cycle of land surface temperature? A multi-dimensional and multi-granularity perspective," *Sustain. Cities Soc.*, vol. 101, p. 105190, 2024, doi: <https://doi.org/10.1016/j.scs.2024.105190>.
- [18] D. Liu et al., "Spatio-temporal patterns and population exposure risks of urban heat island in megacity Shanghai, China," *Sustain. Cities Soc.*, vol. 108, p. 105500, 2024, doi: <https://doi.org/10.1016/j.scs.2024.105500>.
- [19] H. Xu, C. Li, Y. Hu, R. Kong, Q. Wang, and Y. Zhou, "Long-term temporal and spatial divergence patterns of urban heat risk in the Beijing–Tianjin–Hebei urban agglomeration," *Urban Clim.*, vol. 56, p. 102085, 2024, doi: <https://doi.org/10.1016/j.uclim.2024.102085>.
- [20] A. Golzari Oskouei et al., "Feature-weighted fuzzy clustering methods: An experimental review," *Neurocomputing*, vol. 619, p. 129176, 2025, doi: <https://doi.org/10.1016/j.neucom.2024.129176>.
- [21] H. Zhang and S.-L. Huang, "Improved fuzzy C-means clustering algorithm based on fuzzy particle swarm optimization for solving data clustering problems," *Math. Comput. Simul.*, vol. 233, pp. 311–329, 2025, doi: <https://doi.org/10.1016/j.matcom.2025.02.012>.
- [22] M. Thamrin et al., "Machine learning for global trade analysis: a hybrid clustering approach using DBSCAN, elbow, and SOM," *IAES Int. J. Artif. Intell.*, vol. 14, no. 4, p. 3033, Aug. 2025, doi: <https://doi.org/10.11591/ijai.v14.i4.pp3033-3046>.
- [23] E. Elmarakby and H. Elkadi, "Prioritising urban heat island mitigation interventions: Mapping a heat risk index," *Sci. Total Environ.*, vol. 948, p. 174927, 2024, doi: <https://doi.org/10.1016/j.scitotenv.2024.174927>.
- [24] B. Yan and Z. Xie, "A robust fuzzy cluster validity index based on local distances," *Expert Syst. Appl.*, vol. 295, Jan. 2026, doi: <https://doi.org/10.1016/j.eswa.2025.128883>.
- [25] Y. Ji, Z. Li, Y. Liu, H. Tang, H. Sun, and T. Feng, "How urban sprawl patterns shape the thermal environment during hot summers: An empirical analysis of 338 Chinese cities," *Environ. Impact Assess. Rev.*, vol. 117, p. 108251, 2026, doi: <https://doi.org/10.1016/j.eiar.2025.108251>.
- [26] G. K. Sharma and V. V. Ghuge, "Assessment of urban growth dynamics using landscape expansion index: A study of 514 metropolitan cities from 1975 to 2020," *Land use policy*, vol. 153, p. 107557, 2025, doi: <https://doi.org/10.1016/j.landusepol.2025.107557>.
- [27] Q. Wang, H. Wang, L. Ren, J. Chen, and X. Wang, "Hourly impact of urban features on the spatial distribution of land surface temperature: A study across 30 cities," *Sustain. Cities*

- Soc., vol. 113, p. 105701, 2024, doi: <https://doi.org/10.1016/j.scs.2024.105701>.
- [28] M. P. da Silva et al., "Advancing agricultural remote sensing: A comprehensive review of deep supervised and Self-Supervised Learning for crop monitoring," *Comput. Graph.*, vol. 133, p. 104434, 2025, doi: <https://doi.org/10.1016/j.cag.2025.104434>.
- [29] J. Zhang et al., "Spatiotemporal dynamic mapping of heat exposure risk for different populations in city based on hourly multi-source data," *Sustain. Cities Soc.*, vol. 107, p. 105454, 2024, doi: <https://doi.org/10.1016/j.scs.2024.105454>.
- [30] A. Chauhan et al., "Chapter 10 - Earth observation applications for urban mapping and monitoring: research prospects, opportunities and challenges," in *Earth Observation in Urban Monitoring*, A. Kumar, P. K. Srivastava, P. Saikia, and R. K. Mall, Eds., Elsevier, 2024, pp. 197–229. doi: <https://doi.org/10.1016/B978-0-323-99164-3.00007-0>.
- [31] D. Ehrlich, T. Kemper, M. Pesaresi, and C. Corbane, "Built-up area and population density: Two Essential Societal Variables to address climate hazard impact," *Environ. Sci. Policy*, vol. 90, pp. 73–82, 2018, doi: <https://doi.org/10.1016/j.envsci.2018.10.001>.
- [32] W. Qian, F. Rao, X. Li, and D. Lai, "Mapping priority zones for urban heat mitigation in Shanghai: Heat risk vs. shelter provision," *Comput. Environ. Urban Syst.*, vol. 121, p. 102330, 2025, doi: <https://doi.org/10.1016/j.compenvurbsys.2025.102330>.
- [33] L. H. Son, "Enhancing clustering quality of geo-demographic analysis using context fuzzy clustering type-2 and particle swarm optimization," *Appl. Soft Comput.*, vol. 22, pp. 566–584, 2014, doi: <https://doi.org/10.1016/j.asoc.2014.04.025>.
- [34] Y. Zhou, H. Huang, J. Ma, Q. Liu, and Y. Guo, "Complex spatiotemporal thermal effects caused by urban densification: Patterns, characteristics, mechanisms and strategies," *Build. Environ.*, vol. 288, p. 114008, 2026, doi: <https://doi.org/10.1016/j.buildenv.2025.114008>.
- [35] A. Soltani, P. Azizi, A. Rahimioun, and M. Sedaghatfard, "Volumetric urban sprawl: Horizontal and vertical growth in two metropolitans," *J. Urban Manag.*, vol. 14, no. 4, pp. 1004–1021, 2025, doi: <https://doi.org/10.1016/j.jum.2025.03.006>.
- [36] A. Soltani, A. Dehghani, and P. Azizi, "Volumetric insights into urban growth analysis: Investigating vertical and horizontal patterns," *Sustain. Cities Soc.*, vol. 130, p. 106589, 2025, doi: <https://doi.org/10.1016/j.scs.2025.106589>.
- [37] M. Davtalab and S. Byčenkienė, "Synergistic effects of spatial urban form on PM2.5 concentration and urban heat islands: A multi-scale explanatory machine learning," *J. Clean. Prod.*, vol. 538, p. 147226, 2026, doi: <https://doi.org/10.1016/j.jclepro.2025.147226>.
- [38] T. Ye et al., "Billions of people exposed to increasing heat but decreasing greenness from 2000 to 2022," *Innov.*, vol. 6, no. 5, p. 100870, 2025, doi: <https://doi.org/10.1016/j.xinn.2025.100870>.
- [39] Z. Stalhandske et al., "Global assessment of population exposure to multiple climate-related hazards from 2003 to 2021: a retrospective analysis," *Lancet Planet. Heal.*, vol. 9, no. 8, p. 101295, 2025, doi: <https://doi.org/10.1016/j.lanplh.2025.101295>.
- [40] F. Zhang, G. Chen, J. Hua, and C. Ren, "Investigating the dynamic pattern of regional heat health risks: A case study of Guangdong Province, China," *Urban Clim.*, vol. 57, p. 102127, 2024, doi: <https://doi.org/10.1016/j.uclim.2024.102127>.
- [41] J. Yin, H. Wang, and J. Zhang, "Exploring the thermal effects and influencing factors of urban parks during diurnal variations," *Sustain. Cities Soc.*, vol. 129, p. 106499, 2025, doi: <https://doi.org/10.1016/j.scs.2025.106499>.
- [42] Y. Xiao et al., "Seasonal variability and multi-factor analysis of urban park cooling effects: A case study of Nanjing, China," *Sustain. Cities Soc.*, vol. 131, p. 106711, 2025, doi: <https://doi.org/10.1016/j.scs.2025.106711>.
- [43] M. Chen, X. Li, A. Dang, Y. Weng, and S. Qiu, "Adapting to heatwaves: Optimizing urban green spaces in Beijing to reduce heat health risks," *Sustain. Cities Soc.*, vol. 130, p. 106600, 2025, doi: <https://doi.org/10.1016/j.scs.2025.106600>.
- [44] T. Zhang et al., "A new framework for assessing and dealing with heat risk from an urban resilience perspective," *J. Clean. Prod.*, vol. 479, p. 144008, 2024, doi: <https://doi.org/10.1016/j.jclepro.2024.144008>.
- [45] M. Faisal et al., "A Hybrid MOO, MCGDM, and Sentiment Analysis Methodologies for Enhancing Regional Expansion Planning: A Case Study Luwu - Indonesia," *Int. J. Math. Eng. Manag. Sci.*, vol. 10, no. 1, pp. 163–188, Feb. 2025, doi: <https://doi.org/10.33889/IJMEMS.2025.10.1.010>.
- [46] M. Faisal et al., "A Hybrid MOO-MCDM-PSOM Approach for Mapping Urban Carbon Trade Policies," Feb. 12, 2025, doi: <https://doi.org/10.33889/IJMEMS.2025.10.6.101>.
- [47] X. Dong, X. Li, Y. Ye, D. Su, R. Yang, and A. Lausch, "Measuring urban thermal environment from accessibility-based perspective: A case study in a populous city," *Geogr. Sustain.*, vol. 5, no. 3, pp. 329–342, 2024, doi: <https://doi.org/10.1016/j.geosus.2024.02.004>.