Discover Sustainability

https://doi.org/10.1007/s43621-026-04323-1

Article in Press

Geospatial machine learning maps urban heat inequality for climate resilient planning in Bologna

Received: 6 June 2026

Accepted: 3 August 2026

Published online: 12 August 2026

Cite this article as: Tabatabaei M., Saber A. & Pourzangbar A. Geospatial machine learning maps urban heat inequality for climate resilient planning in Bologna. *Discov Sustain* (2026). https://doi.org/10.1007/s43621-026-04323-1

Mahdiyeh Tabatabaei, Aniseh Saber & Ali Pourzangbar

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

ARTICLE IN PRESS

Geospatial machine learning maps urban heat inequality for climate resilient planning in Bologna

Mahdiyeh Tabatabaei [a]*, Aniseh Saber [a], Ali Pourzangbar [b]

[a] *Department of Architecture, Alma Mater Studiorum–University of Bologna, Bologna, Italy*

[b] *Institute for Water and Environment, Karlsruhe Institute of Technology, Karlsruhe, Germany*

** Corresponding Author: mahdiyeh.tabatabaei2@unibo.it*

Abstract

Sustainable, climate-resilient urban development requires evidence linking thermal exposure, social vulnerability, and the built environment. This study develops a geospatial machine-learning framework for Bologna, Italy, to identify areas where urban heat, social vulnerability, and green-infrastructure deficits overlap. The framework integrates land surface temperature (LST) with data on 51,976 building footprints, urban morphology, infrastructure, vegetation and land cover, meteorological conditions, and demographic and socioeconomic characteristics. A Random Forest (RF) model predicted building-level LST during summer 2023, while a Long Short-Term Memory (LSTM) model characterized hourly air-temperature dynamics using meteorological records from 2019–2023. With spatially tuned hyperparameters, the RF model achieved R^2 = 0.749, MAE = 1.16 °C, and RMSE = 1.50 °C under a conventional random split. Under five-fold spatially blocked cross-validation, pooled performance was R^2 = 0.579, MAE = 1.58 °C, and RMSE = 1.97 °C, indicating optimistic random validation while retaining predictive capability in spatially separated areas of Bologna. Local relief, built-up cover, tree cover, sky-view factor, shadow conditions, industrial and commercial land, and distance to parks were the most influential predictors. The tuned LSTM model captured one-hour-ahead air-temperature dynamics in 2023 with R^2 = 0.995, MAE = 0.43 °C, and RMSE = 0.59 °C, outperforming the persistence baseline RMSE of 1.07 °C. At the planning scale, a Heat Vulnerability Index (HVI) was developed for 90 statistical areas by integrating heat exposure with socioeconomic and demographic vulnerability. Sensitivity analysis showed that HVI rankings were robust to variations in component weights, with a median Spearman rank correlation of 0.992 and agreement in nine of ten areas between recalculated and baseline top ten rankings. CAAB, Michelino, Ducati–Villaggio INA, San Donnino, and Pilastro emerged as the highest-priority areas for targeted heat-mitigation and climate-adaptation interventions. By linking fine-scale heat patterns with local adaptive capacity, the framework supports prioritizing tree-canopy expansion, shading, cool surfaces, and improved access to green spaces. It contributes to the Sustainable Development Goals (SDGs) and can be adapted to other compact European cities after local calibration.

Keywords: Urban heat inequality; Sustainable urban development; Geospatial machine learning; Heat vulnerability index; Climate adaptation; Bologna

1. Introduction

ARTICLE IN PRESS

Cities are central arenas for climate action because they concentrate populations, infrastructure, economic activity, energy demand, and built assets. At the same time, urban form can amplify climate-related risks through dense impervious surfaces, limited vegetation cover, dense development, and anthropogenic heat emissions [1]. In Europe, this challenge is framed by the European Green Deal and the European Climate Law, which place climate neutrality by 2050 at the center of EU policy [2,3]. For urban planning, however, climate transition cannot focus only on reducing greenhouse gas emissions. It must also address rising temperatures, more frequent heatwaves, and unequal exposure to environmental hazards. This broader perspective aligns with disaster-risk-reduction frameworks that emphasize hazard, exposure, vulnerability, adaptive capacity, and risk-informed governance [4–6]. Climate-resilient planning, therefore, needs analytical approaches that connect physical heat exposure with social vulnerability and the spatial characteristics of the built environment.

From a sustainable-development perspective, urban heat adaptation lies at the intersection of public health, social equity, sustainable urban development, and climate action [5,10,57,60,79]. Heat-resilient planning can reduce heat-related health risks, address inequalities in the distribution of environmental burdens and cooling benefits and strengthen the resilience of cities and vulnerable communities [5,41–44,57,60]. These aims are closely aligned with Sustainable Development Goal (SDG) 3 on health and well-being, SDG 10 on reducing inequalities, SDG 11 on inclusive, safe, resilient, and sustainable cities, and SDG 13 on climate action [79]. Translating these global goals into local planning decisions, however, requires spatial evidence showing where thermal exposure, social vulnerability, and deficits in green and cooling infrastructure overlap [10,41–44,57–60].

Excessive heat is among the most urgent and spatially uneven urban climate hazards. The Urban Heat Island (UHI) effect causes urban areas to experience higher temperatures than surrounding rural or peri-urban areas [7,8]. It is driven by urban morphology, surface materials, land-cover composition, limited vegetation, reduced sky-view conditions, and anthropogenic heat emissions. Together, these factors influence Land Surface Temperature (LST), outdoor thermal comfort, cooling-energy demand, public health, and the overall livability of dense urban environments [8]. Yet heat exposure is not only a physical phenomenon. Its consequences are socially differentiated because people and neighborhoods vary in sensitivity, available resources, housing quality, mobility, and capacity to adapt [9,10]. Urban heat-risk assessment must therefore move beyond temperature mapping and examine how environmental exposure intersects with social vulnerability.

This need is particularly important in compact European cities, where historical urban fabrics, dense residential districts, limited open space, and competing land-use pressures constrain the range of possible interventions [11]. In these settings, heat adaptation cannot rely on generic greening or cooling strategies alone. Planning decisions must be spatially targeted, socially

ARTICLE IN PRESS

informed, and responsive to the city's specific morphology. Tree-canopy expansion, park accessibility improvement, changes in surface-material, shading strategies, and heat-sensitive zoning all require evidence about where heat accumulates, which built-environment characteristics contribute to it, and which communities are most vulnerable to its effects. Integrating these dimensions is therefore essential for translating urban-climate analysis into practical planning priorities.

Integrated assessments combining thermal exposure, built-environment characteristics, and social vulnerability are increasingly common in the urban-heat literature [10,45,57–63]. Nevertheless, existing studies differ considerably in spatial resolution, modelling strategy, and their connection to planning decisions. A practical challenge remains in compact European cities: linking interpretable building-scale surface-heat modelling with area-level vulnerability prioritization through datasets and methods that municipalities can reproduce and update. Rather than presenting integrated heat-vulnerability assessment as entirely new, this study addresses this specific planning-oriented gap. Recent advances in geospatial data science and machine learning make it possible to combine satellite-derived LST, building footprints, urban-morphology metrics, vegetation and land-cover indicators, park accessibility, hourly climate observations, and socioeconomic data within spatially explicit workflows. Machine-learning models can capture nonlinear relationships between heat patterns and multiple urban predictors, while index-based approaches can translate complex environmental and social information into planning-relevant priority maps. These methods, however, must be brought together within a coherent framework connecting fine-scale heat modelling, temporal climate dynamics, and area-level vulnerability assessment.

To address this gap, this study develops and tests an integrated geospatial machine-learning framework for Bologna that connects building-scale heat exposure, hourly air-temperature dynamics, green-infrastructure conditions and accessibility, and area-level socioeconomic vulnerability. It addresses three research questions: (1) how accurately can summer building-level LST be modelled using urban morphology, vegetation, land-cover, and park-proximity variables; (2) how well can a Long Short-Term Memory (LSTM) model represent one-hour-ahead air-temperature dynamics compared with a persistence baseline; and (3) which statistical areas should be prioritized when heat exposure and social vulnerability are assessed jointly? By answering these questions, the study seeks to support equitable climate adaptation, improve the allocation of urban cooling investments, and contribute to progress toward SDGs 3, 10, 11, and 13 [79].

The proposed framework integrates satellite-derived LST, 51,976 building footprints, urban-morphology indicators, vegetation and land-cover variables, park proximity, hourly climate records, and socio-economic and demographic data. It operates across three analytical scales. At the building scale, Random Forest (RF) regression is used to model summer LST and identify the

ARTICLE IN PRESS

spatial factors associated with building-level heat variation. At the temporal scale, LSTM modelling is applied to hourly climate data to examine air-temperature dynamics and compare deep-learning performance with a persistence baseline. At the planning scale, a Heat Vulnerability Index (HVI) is constructed for Bologna's statistical areas to identify priority zones where physical heat exposure overlaps with socio-economic and demographic vulnerability.

The novel contribution of this study lies in its multi-scale workflow connecting fine-scale heat modelling with temporal prediction and planning-oriented vulnerability assessment. At the building scale, it models summer LST across 51,976 building footprints and identifies interpretable spatial predictors using RF regression. At the temporal scale, it evaluates hourly air-temperature predictions against a persistence baseline. At the municipal planning scale, it constructs an HVI for 90 statistical areas. This design complements, rather than duplicates, the Bologna study by Saber et al. [45], which examined spatio-temporal heatwave risk using an optimized LSTM framework. The present study instead focuses on building-level surface-temperature variation, interpretable urban-form and green-infrastructure predictors, and statistical-area prioritization for climate-adaptation planning.

This study is organized as follows. Section 2 reviews the relevant literature on urban heat-risk assessment, geospatial machine learning, social vulnerability, and climate-resilient urban planning. Section 3 presents the study area, datasets, spatial-processing workflow, machine-learning models, and HVI construction. Section 4 reports and discusses the results, including building-level heat modelling, hourly climate dynamics, vulnerability mapping, and planning implications. Section 5 concludes by summarizing the main findings, discussing limitations, and outlining directions for future research.

ARTICLE IN PRESS

2. Literature review

From a sustainable-development perspective, urban heat is not solely an environmental or technical problem. It is also connected to public health, spatial inequality, access to environmental benefits, urban resilience, and the equitable allocation of adaptation resources [5,9,10,57,60,79]. Heat-mitigation measures such as tree planting, park development, shading, and cool-surface interventions can generate environmental, health, and social benefits, but these benefits may remain unevenly distributed when planning decisions do not account for social vulnerability and unequal adaptive capacity [38–44,57–61]. A sustainability-oriented heat assessment should therefore examine not only where temperatures are high, but also whether vulnerable populations have access to the resources, infrastructure, and environmental conditions required to cope with extreme heat [10,42–44,57–60]. This integrated perspective connects urban heat assessment with SDGs 3, 10, 11, and 13 [79].

Urban heat research has moved from describing the general UHI effect toward explaining why heat varies strongly within the same city and why its impacts are unevenly distributed across

ARTICLE IN PRESS

social groups. Oke's early work established that urban warming is rooted in changes to the surface energy balance, including heat storage in built materials, reduced evapotranspiration, modified radiative exchange, and anthropogenic heat emissions [7,14]. Later research has shown that these processes are shaped by land cover, building density, vegetation, surface materials, street geometry, and the configuration of open and shaded spaces [15–18]. For planning, urban heat should therefore be studied not only as a climatic phenomenon, but also as a spatial outcome of urban form and land-use decisions [19].

Thermal remote sensing has been central to this shift because it makes intra-urban heat differences visible at the scale of the whole city. Voogt and Oke [20] showed that thermal satellite data can support urban-climate analysis by identifying surface-temperature patterns and linking them to land-cover and morphological conditions. Weng [21] similarly emphasized the value of thermal infrared remote sensing for examining the relationship between LST, vegetation, imperviousness, and urban development. These studies have made LST one of the most widely used indicators for mapping surface UHI [22,23]. However, LST should not be treated as identical to air temperature or pedestrian-level heat stress [24]. Schwarz et al. [25] demonstrated that the relationship between land surface and air temperature is significant but context dependent, while recent research has warned that satellite-derived LST may not fully represent street-level heat stress in dense urban fabrics where shade, radiation, ventilation, and humidity strongly affect outdoor comfort [26]. Therefore, LST is most useful when interpreted together with urban morphology, climate data, and socio-spatial indicators [27].

Urban morphology provides an important bridge between surface-temperature patterns and planning interpretation. Stewart and Oke [28] introduced the Local Climate Zone (LCZ) framework to standardize the classification of urban and natural landscapes according to surface structure, cover, materials, and human activity. Bechtel et al. [29] later advanced LCZ mapping as a basis for global urban-climate databases. Nevertheless, LCZ classes can hide important local differences, particularly in compact European cities where building heights, street widths, courtyards, tree cover, and surface materials change over short distances [30,31]. For this reason, recent studies increasingly combine categorical urban classifications with continuous spatial indicators such as sky-view factor, building density, surface roughness, imperviousness, elevation, shadow conditions, vegetation cover, and distance to green spaces [32–34]. Urban canyon geometry, orientation, and surface reflectance also affect seasonal solar exposure, reinforcing the importance of three-dimensional urban form in the interpretation of urban thermal patterns [81]. In Milan, Puche et al. [35] showed that both land cover and urban morphology influence land surface and air temperatures, confirming that urban heat patterns are produced by interacting physical and spatial variables rather than by land cover alone.

Green infrastructure is one of the most consistent cooling factors identified in the literature, although its effect depends on scale, form, vegetation structure, and accessibility [36–38]. Trees

ARTICLE IN PRESS

and vegetated surfaces reduce heat through shading and evapotranspiration, while parks can create localized cooling effects that extend into surrounding areas [39]. Ziter et al. [38] found that tree-canopy cover can reduce daytime urban heat, particularly when canopy and impervious surfaces are considered together at appropriate spatial scales. Aram et al. [40] showed that park cooling intensity varies according to size, vegetation type, shape, surrounding density, and local climate. More recently, Li et al. [41] demonstrated that green spaces provide substantial but unequal cooling benefits globally. This finding is directly relevant to heat-resilient planning: green infrastructure is not only an environmental intervention, but also an issue of spatial equity [42,43]. Areas with high heat exposure and poor access to trees, parks, or shaded public spaces may require priority attention even if their absolute temperature values are not the highest in the city [44].

Because urban heat is shaped by multiple interacting variables, machine-learning methods have become increasingly important in LST and UHI studies. Linear models can identify general associations, but they often struggle to represent non-linear relationships, threshold effects, and interactions among vegetation, morphology, land cover, and topography [45]. Previous integrated work in Bologna by Saber et al. [45] combined climatic, environmental, infrastructure, and demographic indicators within a spatio-temporal artificial-intelligence framework for heat-wave risk assessment. The present study, however, differs in analytical emphasis and spatial unit. It uses Random Forest regression to model summer LST at the level of individual building footprints, evaluates the predictive relevance of morphology, vegetation, land-cover, and park-proximity variables, and constructs an area-level HVI across Bologna's 90 statistical areas. The LSTM component is included as a complementary city-level analysis of air-temperature dynamics and is explicitly benchmarked against a persistence model. The contribution therefore lies in creating a reproducible, planning-oriented connection between interpretable building-level surface-heat modelling and municipal vulnerability prioritization, rather than in introducing the first integration of heat exposure and social vulnerability.

Accordingly, assessing the combined and non-linear effects of different factors on LST and urban heat more broadly requires methods that go beyond simple linear models. RF regression is well-suited to this task because it performs effectively with heterogeneous spatial datasets, captures complex interactions among environmental, morphological, and spatial variables, and provides interpretable estimates of predictor importance [46]. Wang et al.[47] used RF to examine how two- and three-dimensional urban characteristics affect LST, while Tanoori et al.[48] highlighted the growing role of machine-learning models in urban LST prediction. Recent machine-learning research has also integrated physical and socio-demographic determinants to support zone-specific and equity-oriented urban heat-mitigation policies [82]. However, predictive accuracy alone is not sufficient. Conventional random train–test splits can overestimate model performance because nearby observations often share similar spatial characteristics. Roberts et

ARTICLE IN PRESS

al.[49] therefore recommend validation strategies that account for spatial and temporal dependence. For planning-oriented research, machine-learning models should not only predict heat accurately but also produce interpretable evidence that can inform where and how interventions should be prioritized.

A second methodological strand concerns the temporal dynamics of urban heat. Many LST studies rely on one or a few satellite images, which are useful for mapping spatial contrasts but cannot fully represent hourly, daily, or seasonal temperature variation [50]. Urban heat risk changes over time because air temperature, radiation, wind, humidity, and heat storage vary across the day and across seasons [51]. LSTM models are increasingly used for time-series problems because they can learn temporal dependencies in sequential data. Yu et al. [52] applied LSTM models with urban sensor data to forecast localized temperature, and Wang et al. [53] examined LSTM forecasting of pedestrian-level urban air temperature. More recently, Saber et al. [45] developed an LSTM-based spatio-temporal framework for heat-wave risk assessment in Bologna, integrating climatic, environmental, infrastructure, and demographic indicators to capture the interaction between hazard, exposure, and vulnerability. Their results showed that temperature, vegetation condition, local climate zone, and access to urban services are important factors in shaping heat-wave risk, highlighting the value of LSTM models for planning-oriented urban heat analysis. This literature suggests that LSTM models can be useful for understanding temporal temperature dynamics, but they should be evaluated against simple baselines.

The social dimension of heat risk has developed through vulnerability and environmental justice research. Cutter et al. [9] defined social vulnerability as the social, economic, demographic, and institutional conditions that shape the capacity of people and communities to prepare for, respond to, and recover from hazards. Applied to urban heat, this means that risk cannot be explained by temperature alone. Older residents, low-income households, socially isolated individuals, migrants, people with limited mobility, and residents in poor-quality housing may face higher heat-related impacts even when exposed to similar thermal conditions [54,55]. Reid et al. [10] developed an influential approach to mapping heat vulnerability by combining demographic, socio-economic, health, and environmental indicators. Johnson et al. [56] later proposed an applied extreme heat vulnerability index using socio-economic and remotely sensed variables. Ellena et al. [57], in a systematic review of the heat-health nexus, showed that heat-related health outcomes depend on the interaction between hazard intensity, socio-economic vulnerability, and built-environment characteristics. These studies demonstrate that heat vulnerability is multidimensional and spatially uneven. Recent research in Johannesburg similarly showed that combining environmental, socioeconomic, and health indicators can reveal clustered patterns of intra-urban heat vulnerability and support targeted intervention priorities [80].

ARTICLE IN PRESS

Recent studies have refined this discussion by emphasizing both the usefulness and the limitations of heat vulnerability indices. Li et al. [12] found that HVI methods vary widely in terms of indicator selection, weighting, aggregation, and validation. This variation does not undermine the value of HVIs, but it does mean that they must be carefully adapted to local conditions and data availability. A generic index may miss locally important conditions such as housing quality, green-space accessibility, income structure, population age, or neighborhood morphology. Vulnerability mapping is therefore most useful when it is not treated as a separate social layer, but as part of an integrated spatial diagnosis of heat exposure and adaptive capacity [58,59].

Environmental justice studies further show that heat exposure often follows broader patterns of socio-spatial inequality. Ellena et al. [60] demonstrated in Turin that micro-scale urban heat risk reflects the interaction of thermal exposure, socioeconomic vulnerability, and built-environment characteristics. The broader implication is that urban heat is not randomly distributed and that heat-adaptation priorities should account for both exposure and social vulnerability [60,61]. For European cities, this concern is particularly important because many urban areas combine compact historical fabrics, aging populations, limited open space, and uneven access to green infrastructure [62]. In Bologna, Maccabiani et al. [61] examined social vulnerability to extreme heat and climate shelters, while Saber et al. [45] developed a spatio-temporal AI framework showing the importance of vegetation and green-space access in heat-wave risk mitigation. Together, these studies confirm the relevance of integrating heat exposure, social fragility, and access to cooling resources in local adaptation planning.

The available literature provides strong foundations for mapping urban heat, modelling its spatial drivers, analyzing temporal temperature dynamics, and assessing social vulnerability. Yet these research areas are still often developed separately. Remote-sensing studies may map heat without fully integrating socio-economic vulnerability. Vulnerability studies may identify sensitive populations without detailed modelling of building-scale heat drivers. Machine-learning studies may achieve high predictive performance but remain weakly connected to planning decisions [62,63]. Temporal models may forecast air temperature without being linked to spatial vulnerability or intervention priorities. The gap, therefore, is not the absence of methods, but the limited integration of these methods into a coherent and interpretable planning workflow.

This study responds to that gap by linking building-scale heat modelling, hourly climate-dynamics analysis, and statistical-area vulnerability assessment for Bologna. By combining satellite-derived LST, building morphology, vegetation, and land-cover indicators, proximity to green spaces, climate time series, and socio-economic variables, the framework connects three analytical questions that are commonly addressed at different spatial and analytical scales: where heat is concentrated, which spatial factors explain its distribution, and where physical exposure overlaps with social vulnerability. In doing so, it contributes to more targeted and equitable approaches to heat-resilient urban planning in compact European cities.

ARTICLE IN PRESS

3. Materials and methods

3.1 Study area and climatic context

This study focuses on Bologna (44.498955° N; 11.327591° E), a historic city in the Emilia-Romagna region of northern Italy. The municipality covers approximately 141 km² and had 392,791 residents as of 31 December 2024, making it one of Italy's most populous municipalities [64,65]. Bologna has a heterogeneous urban structure shaped by its compact historic center, dense residential neighborhoods, industrial and commercial areas, transport corridors, peri-urban fringes, and a distributed network of public parks and gardens. This diversity makes the city a suitable case study for examining how urban morphology, land-cover composition, green infrastructure, and socio-economic conditions influence heat exposure.

The city is also relevant from a climate-resilient planning perspective. Bologna is one of the European cities selected under the EU Mission for Climate-Neutral and Smart Cities, which aims to accelerate urban decarbonization and support climate-transition pathways [66]. In this context, Bologna provides an appropriate setting for testing an integrated framework that links urban heat-risk assessment with social-vulnerability mapping in an established European urban environment. The administrative structure of the municipality further supports this type of spatial analysis. Bologna is organized into six districts, 18 zones, and 90 statistical areas, allowing heat exposure and vulnerability patterns to be examined at multiple spatial scales.

Bologna's urban climate is strongly affected by seasonal variability, with summer conditions being particularly important for heat-risk assessment. Local thermal conditions are influenced by several interacting factors, including air temperature, wind speed, solar radiation, LST, vegetation cover, and urban geometry. These factors affect how heat is absorbed, stored, and released across different parts of the built environment.

To capture these interactions, this study integrates satellite-derived LST, hourly climate data, building morphology indicators, proximity to green infrastructure, and socio-demographic variables within a single analytical framework. The spatial context of the study area, including the municipal boundary, statistical areas, building footprints, parks and gardens, and the summer 2023 LST raster, is shown in **Fig. 1**.

ARTICLE IN PRESS

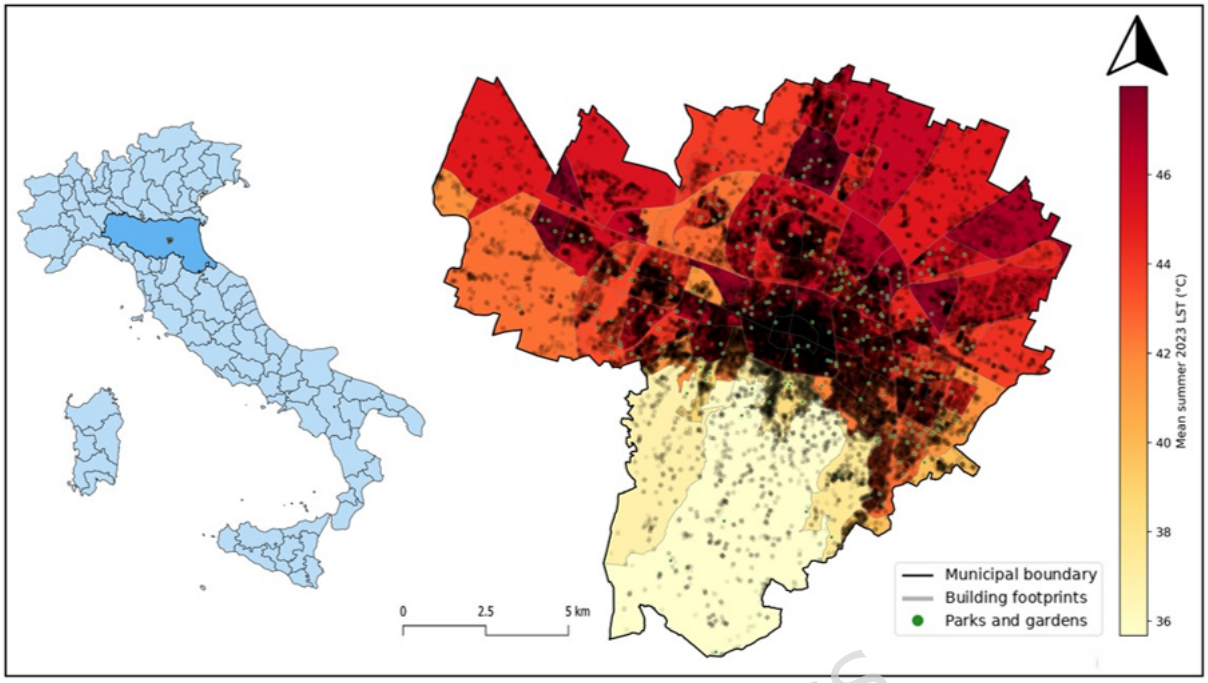


Fig. 1 Spatial context of Bologna, Italy, including the municipal boundaries, statistical areas, building footprints, green spaces, and Landsat-derived mean summer 2023 land surface temperature

3.2 Methodological approach

The methodological framework links building-level heat modelling, hourly climate-dynamics modelling, and heat-vulnerability assessment within a single urban-planning workflow. As shown in **Fig. 2**, the process begins with the collection of multi-source datasets, including Landsat-derived LST, hourly climate data, building-footprint information, land-cover and vegetation indicators, green-infrastructure data, and socio-economic and demographic statistics.

These datasets were then processed and harmonized in QGIS to create a consistent spatial database for the subsequent analysis. This step included the preparation of building-level attributes, statistical-area indicators, land-cover variables, and proximity measures related to parks and gardens. The resulting dataset formed the basis for both the machine-learning models and the vulnerability-index calculation.

In the modelling stage, two complementary approaches were applied. A random forest regression model was developed to estimate building-level summer LST and examine the influence of selected urban and environmental predictors. In parallel, an LSTM model was used to capture hourly air-temperature dynamics and produce one-hour-ahead temperature predictions. Both models were evaluated using R^2, Mean Absolute Error (MAE), and Root Mean Square Error (RMSE). The RF analysis used random and spatially blocked validation of the building-level dataset, whereas the LSTM model was trained on data from 2019–2022 and tested

ARTICLE IN PRESS

on the independent 2023 period. Feature-importance analysis was used to identify the most influential predictors of building-level thermal exposure.

The vulnerability component combined physical heat exposure with socio-economic and demographic indicators to construct an HVI for Bologna's statistical areas. The index was weighted according to three main dimensions: heat exposure, socio-economic conditions, and demographic sensitivity. The resulting HVI maps were used to highlight spatial differences in vulnerability and to identify areas with the greatest need for heat-related planning interventions.

In the final stage, the modelling and vulnerability outputs were translated into planning-oriented evidence. Feature-ranking results were used to interpret the main spatial drivers of heat exposure, hourly prediction curves were used to represent temporal heat dynamics, and high-HVI areas were delineated as priority zones. These outputs provide a decision-support basis for spatially targeted urban heat-resilience planning.

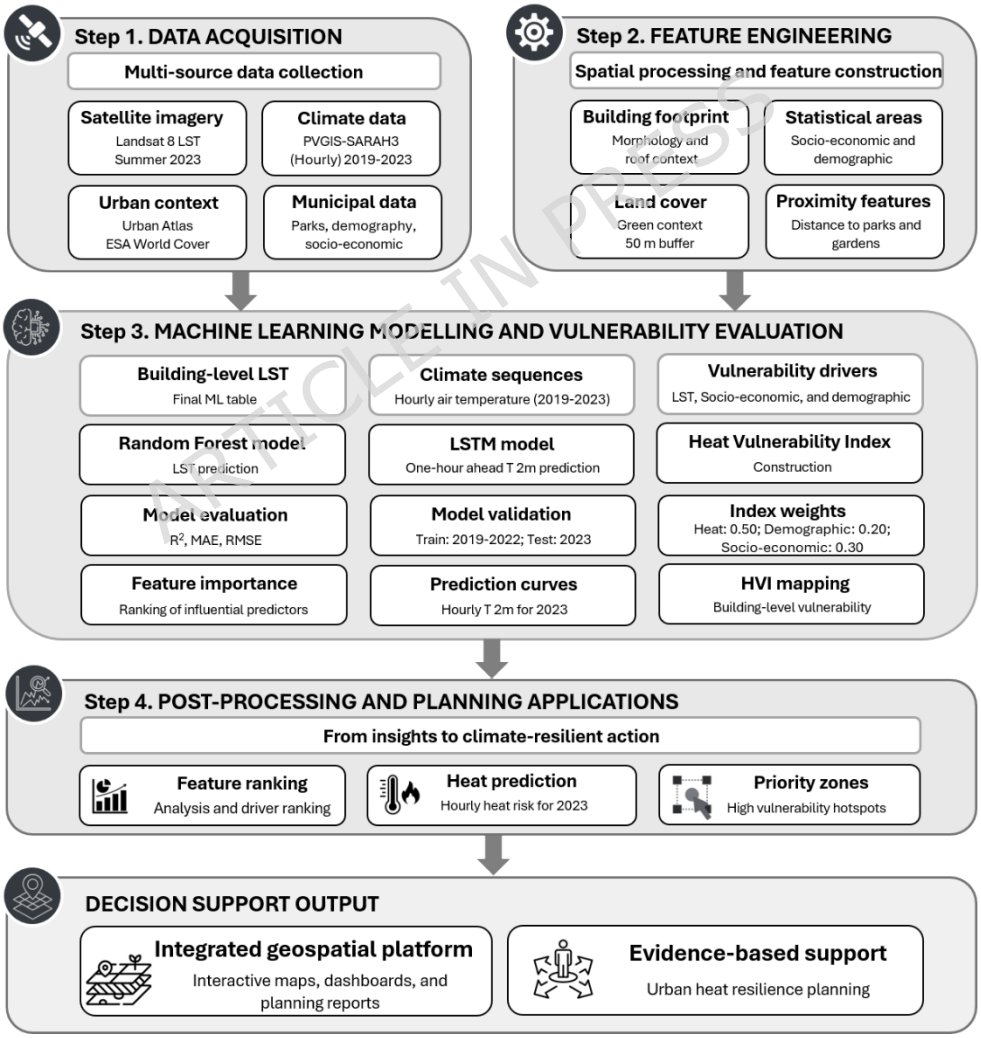


Fig. 2 Overall research workflow for building-level urban heat vulnerability assessment and climate-resilient planning

ARTICLE IN PRESS

3.3 Data sources and preprocessing

The analysis was based on six broad groups of data: hourly climate time series, satellite-derived LST, building-footprint and DEM-derived exposure indicators, land-cover information, parks and gardens, and socioeconomic and demographic datasets. . Table 1 summarizes the sources, spatial or temporal unit, main preprocessing steps, and analytical role of each dataset.

Hourly climate data were extracted from the PVGIS-SARAH3 database for a representative point in central Bologna at 44.494° N and 11.343° E. The PVGIS hourly tool provides hourly solar radiation and meteorological variables over multi-year periods [67,68]. The extracted period covered 2019-2023 and included air temperature at 2 m, wind speed at 10 m, beam irradiance, diffuse irradiance, reflected irradiance, and solar elevation. After removing metadata rows and invalid records, 43,824 valid hourly observations were retained for LSTM modelling.

Table 1. Data sources, preprocessing steps, and the analytical roles of input parameters

Dataset	**Source**	**Spatial/temporal unit**	**Main processing**	**Role in analysis**
Hourly meteorological data	PVGIS-SARAH3 [67]	Hourly time series, 2019–2023	Cleaning, removal of metadata and invalid records; time-feature extraction; sequence generation	LSTM modelling of hourly air-temperature dynamics
Land surface temperature	Landsat 8 Collection 2 Level 2 through Google Earth Engine and USGS documentation [69,70]	30 m Level 2 raster grid, summer 2023	QA/cloud masking; summer mean composite; Kelvin-to-Celsius conversion; reprojection to EPSG:25832	RF target variable and HVI heat-exposure indicator
Building footprint and DEM-derived exposure proxies	OSM building footprints; Copernicus DEM GLO-30; SAGA GIS/QGIS processing [71]	Building polygons; DEM-derived 30 m elevation context	Footprint-area calculation; DEM reprojection; SVF, shadow-index, and local-relief derivation; building-level zonal statistics	RF predictors
Land-cover context	Copernicus Urban Atlas 2018 [72]	Land-use/land-cover classes	Buffer-based land-cover percentage extraction around buildings	RF predictors
Built-up and tree-cover context	ESA WorldCover context [73]	10 m land-cover raster	Built-up-cover class 50 and tree-cover class 10 percentages extracted within 50 m building buffer	RF predictors

ARTICLE IN PRESS

Statistical areas	Comune di Bologna, Aree statistiche [74]	90 statistical-area polygons	Spatial-unit coordination and topology checking	HVI spatial unit
Income	Comune di Bologna, Redditi per area statistica [75]	Statistical area	Filtering by reference year and attribute join to statistical areas	Socio-economic vulnerability
Employment and education	Comune di Bologna, Occupazione e istruzion [77]	Statistical area	Aggregation, harmonization, and QGIS join	Socio-economic vulnerability
Demography	Comune di Bologna population historical dataset [76,77]	Quartiere, Zona hierarchy	Filtering by reference year, aggregation, and key-based join to statistical areas	Demographic vulnerability
Parks and gardens	Comune di Bologna, Elenco parchi e giardini [78]	Point layer	Reprojection and nearest-neighbor distance calculation	RF predictor

LST was derived from Landsat 8 OLI/TIRS Collection 2 Level 2 imagery accessed through Google Earth Engine [69]. Images intersecting the Bologna municipal boundary and acquired between 1 June and 31 August 2023 were initially selected to represent peak summer conditions. The collection was filtered using a scene-level cloud-cover threshold of less than 8%, resulting in four retained Landsat 8 scenes acquired on 20 June, 6 July, 7 August, and 23 August 2023. Thus, the scene-level cloud cover of every image included in the final collection was below 8%. Cloud- and cloud-shadow-contaminated pixels were additionally removed using the Landsat Collection 2 quality-assurance band. The remaining surface-temperature observations were averaged to produce the mean summer LST composite. Surface-temperature values were converted to degrees Celsius using the official scale factor and offset for the Landsat Collection 2 surface-temperature product [70]. The resulting raster was reprojected to EPSG:25832 to ensure spatial compatibility with the urban vector layers.

Building-footprint and terrain-exposure predictors were derived from OpenStreetMap building polygons and the Copernicus DEM GLO-30 product [71]. The building geometries were corrected, converted to single-part polygons, clipped to the Bologna municipal boundary, and reprojected to EPSG:25832. Building footprint area was calculated in square meters. The Copernicus elevation raster was reprojected to EPSG:25832 using bilinear resampling at a 30 m resolution and clipped to the study boundary. Because this elevation product does not provide reliable building-scale three-dimensional geometry, the derived sky-view, shadow, and relief variables were treated as DEM-derived exposure proxies rather than as complete representations of individual buildings or street-canyon geometry.

Land-cover and vegetation conditions were represented using Copernicus Urban Atlas 2018 and ESA WorldCover. Urban Atlas provides high-resolution land-use and land-cover information for European functional urban areas, while ESA WorldCover provides global land-cover products at

ARTICLE IN PRESS

10 m resolution based on Sentinel-1 and Sentinel-2 data [72,73]. These datasets were used to derive built-up-cover, tree-cover, and detailed land-cover-composition indicators around buildings.

Municipal open data from Comune di Bologna were used for statistical areas, income, employment, education, population, citizenship, birthplace, and park locations. The statistical area layer was selected as the main spatial unit for vulnerability mapping because it provides a finer spatial subdivision than the six traditional districts [74]. Income, employment, education, and demographic indicators were integrated into this spatial unit to construct the vulnerability layer, while the parks and gardens dataset was used to calculate each building's distance to the nearest green space [78].

3.4 Spatial harmonization and feature construction

Spatial harmonization and feature construction were carried out in QGIS 3.x using EPSG:25832 as the common coordinate reference system. This step converted raster, point, polygon, and tabular datasets into two main analytical units: building footprints for LST modelling and statistical areas for vulnerability mapping.

For the building-level dataset, mean summer LST values were extracted from the summer Landsat composite using building footprints as zonal units. Zonal statistics transferred raster-based thermal information to building polygons and produced the target variable for the RF model. This procedure transformed the satellite-derived thermal raster into a building-level modelling table.

DEM-derived exposure predictors were calculated in QGIS 3.38.2 using GDAL and SAGA GIS processing tools. Sky-view factor was calculated using the SAGA Terrain Analysis Lighting tool with a maximum search radius of 500 m, 32 directional sectors, and a multiscale factor of 1.25. A shadow-exposure index was constructed from four GDAL hillshade rasters generated with a solar altitude of 45° and azimuth angles of 45°, 135°, 225°, and 315°. The four hillshade rasters were averaged and normalized to a 0–1 scale, with lower values representing greater terrain-related shading.

Local relief was calculated as the difference between the maximum and minimum elevation within an approximately 500 m circular neighborhood using SAGA focal statistics. Mean sky-view factor, shadow-index, and local-relief values were extracted for each building footprint using zonal statistics and joined to the building dataset through a unique building identifier. Land-cover and vegetation indicators were calculated separately from Copernicus Urban Atlas and ESA WorldCover, using buffer-based summaries to represent ESA WorldCover built-up and tree-cover percentages and the surrounding Urban Atlas land-cover composition. Green-infrastructure proximity was represented through a nearest-neighbor spatial join. Parks and gardens were first reprojected to the common coordinate system, and the nearest green-space point was identified

ARTICLE IN PRESS

for each building. The resulting distance was retained as a predictor describing proximity to potential local cooling resources.

The socio-economic and demographic data required a separate area-level workflow. The statistical area layer was adopted as the spatial unit for vulnerability mapping because it provides a fine subdivision of Bologna. Income data were joined directly to the statistical area polygons. Employment and education records were aggregated to the same spatial level, while demographic variables were harmonized through the available district-zone keys and then linked to the statistical areas. This produced an area-level vulnerability dataset containing income, employment, education, population, gender, citizenship, birthplace, and age-group indicators.

Finally, LST statistics were also calculated for each statistical area. Mean, minimum, maximum, and standard deviation values were extracted from the reprojected summer Landsat raster. The mean area-level LST was used as the heat-exposure component of the HVI, while the socio-economic and demographic variables represented the vulnerability dimensions.

3.5 Building-level LST modelling with RF

RF regression was used to model the relationship between building-level LST and the spatial predictors derived in the previous step. Urban heat patterns are shaped by several interacting factors, including urban morphology, vegetation, land-cover composition, shading conditions, and proximity to green infrastructure. Because these relationships are often non-linear and may involve interaction effects, a tree-based ensemble model was considered suitable for the analysis. RF combines multiple decision trees and reduces model instability through ensemble averaging. It can also provide feature-importance measures, which are useful for interpreting the relative contribution of different predictor groups in a planning context [46].

ARTICLE IN PRESS

The response variable was the mean summer LST calculated for each building footprint. The 29 predictors were organized into four groups: building-footprint and DEM-derived exposure context, ESA WorldCover built-up and tree-cover context, Urban Atlas land-cover composition, and green-infrastructure proximity. **Table 2** summarizes these predictor groups, while Supplementary **Table S1** lists all predictors individually together with their definitions, source datasets, units, and calculation methods.

Table 2. Predictor groups and their descriptions used in the RF model

Predictor group	**Description**	**Role in RF model**
Building footprint and terrain-exposure context	Building footprint area and DEM-derived sky-view, local-relief, and shadow proxies	Represents building size and terrain-scale surface-exposure conditions
Built-up and tree-cover context	Green-cover and tree-cover percentages are within 50 m of each building	Represents local built-up intensity and tree-cover context
Land-cover composition	Urban Atlas-based surrounding land-cover percentages	Represents built and natural surface composition

ARTICLE IN PRESS

Green-infrastructure proximity	Represents access to nearby cooling resources	Represents proximity to nearby cooling resources

Potential dependence among the 29 predictors was assessed using pairwise Spearman rank correlations. Absolute correlation coefficients of 0.70 or greater were treated as strong associations. Because the land-cover percentages are compositional and several predictors describe overlapping surface conditions, the correlation analysis was used to qualify the interpretation of feature importance rather than to remove variables automatically. The complete correlation matrix is provided in Supplementary **Fig. S2**. RF hyperparameters were tuned using spatially grouped cross-validation, with mean RMSE as the selection criterion. A structured random search evaluated 24 configurations drawn from a predefined parameter space comprising 150 or 300 trees, square-root or 0.30 feature sampling, minimum leaf sizes of 1, 2, 5, or 10, maximum tree depths of 25, 40, or unrestricted, and minimum split sizes of 2 or 5. Initial tuning was conducted on a spatially stratified sample of 15,002 buildings using 3 km spatial groups and a 300 m exclusion buffer. The leading configurations were subsequently confirmed on the complete dataset using three-fold spatially grouped cross-validation. The selected configuration used 300 trees, square-root feature sampling, a maximum depth of 40, a minimum leaf size of 10, and a minimum split size of 5. The parameter search and selected settings are reported in Supplementary **Table S2**. The selected configuration was then used in the subsequent RF analyses. The same 3 km spatial-block grid and 300 m exclusion-buffer rule were used for hyperparameter tuning and final spatial validation; tuning used three grouped folds for configuration confirmation, whereas final performance was reported from five spatial folds.

Model performance was assessed using the coefficient of determination, MAE, and RMSE, with MAE and RMSE calculated as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | y_i - \hat{y}_i |$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$

where y_i is the observed building-level LST, $\hat{y}_i$ is the predicted LST, and n is the number of observations. Feature importance was then used to examine which spatial predictors contributed most strongly to the modeled variation in building-level LST.

Model performance was evaluated using two complementary validation strategies. First, a conventional 80:20 random train–test split with a random state of 42 was retained for comparison. Second, five-fold spatially blocked cross-validation was conducted using building-

ARTICLE IN PRESS

centroid coordinates. An empirical semivariogram of building-level LST indicated that spatial dependence weakened substantially at a separation distance of approximately 2.8 km (Supplementary **Fig. S1**). Therefore, the study area was divided into 3 km spatial blocks, and all buildings within the same block were assigned to the same validation fold. To further reduce spatial leakage, training buildings located within 300 m of any test building were excluded in each fold. Median imputation was fitted separately within each training fold. Performance was reported using pooled out-of-fold predictions and the mean and standard deviation of the fold-specific metrics.

3.6 LSTM tuning and benchmarking for one-hour-ahead air temperature prediction

The RF model described the spatial variation of building-level LST, while the LSTM model was used to represent the temporal component of heat dynamics. LSTM networks are well-suited to time-series modelling because they can learn dependencies across sequential observations. In this study, the LSTM model was developed to predict air temperature one hour ahead using the previous 24 hours of climate observations.

The target variable was hourly air temperature at 2 m. The input sequence included air temperature, wind speed at 10 m, solar-radiation components, solar elevation, and cyclic time features representing the hour of day and day of year. The lookback window was set to 24 hours, and the forecast horizon was one hour. Data from 2019–2022 were used for training, and the year 2023 was reserved for testing. This temporal split was used to evaluate the model on an unseen year rather than on randomly selected hourly records.

LSTM hyperparameters were selected using three rolling-origin temporal-validation folds: training on 2019 and validating on 2020, training on 2019–2020 and validating on 2021, and training on 2019–2021 and validating on 2022. Six candidate configurations were evaluated, spanning lookback windows of 24, 48, and 72 hours; 24, 32, and 64 hidden units; one or two LSTM layers; dropout values of 0 and 0.20; learning rates of 0.001 and 0.002; and batch sizes of 512 and 1024. During standardized configuration screening, each fold used 5,000 evenly spaced training sequences and a maximum of eight epochs, and mean validation RMSE was used to rank the configurations. These screening metrics were used only for configuration selection.

The best-performing architecture used a 24-hour lookback, one LSTM layer with 64 hidden units, no dropout, a learning rate of 0.001, and a batch size of 512. Training duration was subsequently refined using the complete 2019–2021 period for training and 2022 for validation, resulting in 15 epochs. The selected model was then retrained using all observations from 2019–2022 and evaluated on the untouched 2023 test year. The candidate configurations and selected settings are reported in Supplementary **Table S2**. The final architecture therefore consisted of one LSTM layer with 64 hidden units followed by a fully connected output layer. It was trained for 15 epochs using the Adam optimizer and mean squared error loss, with a batch size of 512.

ARTICLE IN PRESS

The LSTM component was therefore treated as a benchmark test of whether a sequence-based neural network could improve on persistence at the one-hour forecast horizon, rather than assuming that greater model complexity would necessarily produce better predictions. A persistence baseline was also calculated to provide a simple reference model. This baseline assumes that the air temperature in the next hour is equal to the current-hour temperature. Comparing the LSTM with this baseline allowed the evaluation to test whether the neural network captured useful temporal structure beyond the strong short-term continuity normally present in hourly temperature data.

3.7 Heat vulnerability index construction

An HVI was constructed to identify statistical areas where high thermal exposure coincides with socio-economic and demographic vulnerability. Land surface temperature alone does not fully describe vulnerability, since areas with similar heat exposure may differ substantially in income, employment, education, age structure, citizenship, and demographic composition. For this reason, the HVI was designed as a composite index that combines physical exposure with indicators of social vulnerability.

The index was calculated for the 90 statistical areas of Bologna. As shown in **Fig. 3**, three components were included. The heat-exposure component was represented by the mean summer LST calculated for each statistical area. The socio-economic component included income, employment, and education, which were used to describe differences in adaptive capacity. The demographic component included the share of residents aged 65 years and older, the female population share, the foreign population share, and the share of residents born abroad.

ARTICLE IN PRESS

ARTICLE IN PRESS

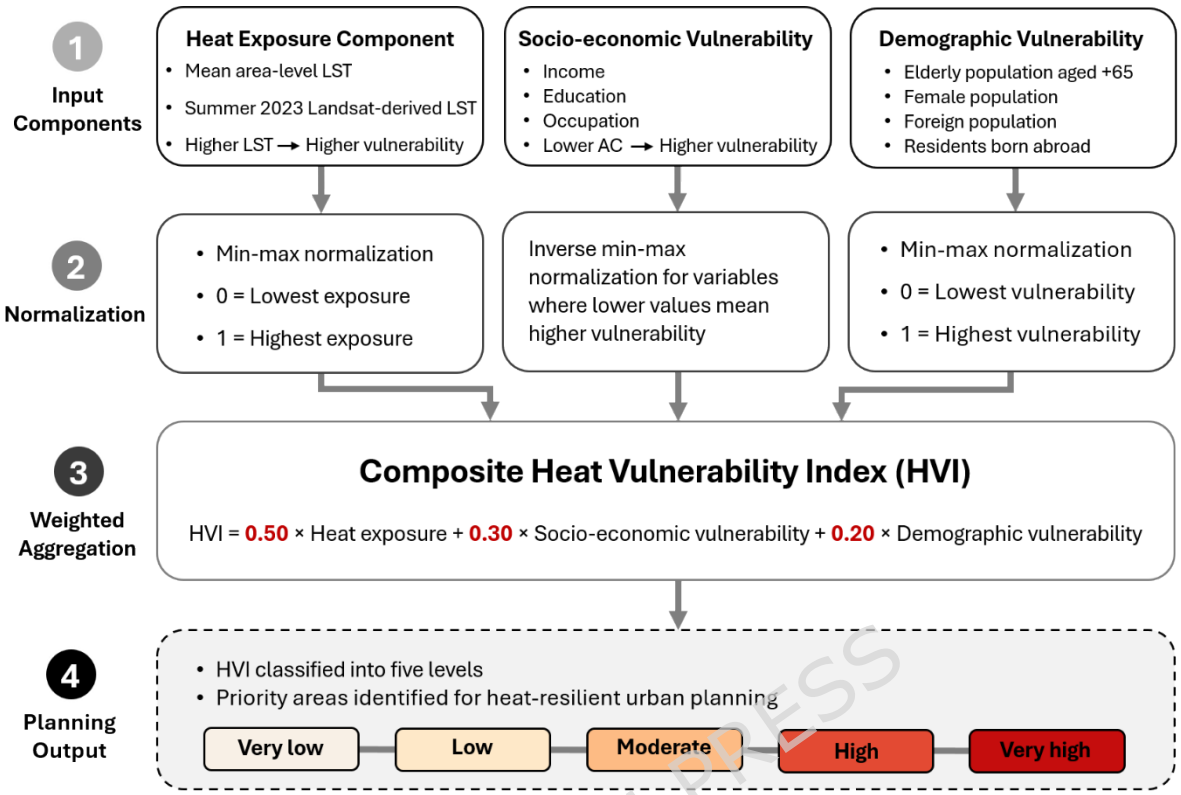


Fig. 3 Construction of the HVI combining heat exposure, socio-economic, and demographic vulnerability

All indicators were normalized to a common 0–1 scale before aggregation. For indicators where higher values indicate higher vulnerability, min–max normalization was applied as follows:

$$S_i = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

where S_i is the normalized score, X_i is the original value for statistical area *i*, and x_{min} and x_{max} are the minimum and maximum values across all statistical areas.

For indicators where lower values indicate higher vulnerability, such as income and employment rate, the normalized score was inverted:

$$S_i = 1 - \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

After normalization, the indicators within each vulnerability dimension were combined to produce component-level scores. The final HVI was then calculated using a weighted aggregation:

$$HVI = 0.50 \times HeatExposure + 0.30 \times SocioEconomicVulnerability + 0.20 \times DemographicVulnerability$$

The highest weight was assigned to heat exposure because the index was intended to identify priority areas under heat-risk conditions. Socio-economic vulnerability received the second-highest weight, reflecting the role of income, employment, and education in shaping adaptive

ARTICLE IN PRESS

capacity. Demographic vulnerability represented population sensitivity. The final HVI scores were classified into five levels: very low, low, moderate, high, and very high. These classes were used to map priority areas for heat-resilience urban planning.

3.8 HVI weight-sensitivity analysis

The robustness of the HVI ranking was evaluated by varying the weights assigned to heat exposure, socioeconomic vulnerability, and demographic vulnerability. In addition to the baseline weights of 0.50, 0.30, and 0.20, respectively, a local sensitivity analysis was conducted using 91 valid weight combinations. Heat-exposure weights ranged from 0.40 to 0.60, socioeconomic weights from 0.20 to 0.40, and demographic weights from 0.10 to 0.30, while the three weights were constrained to sum to one. A broader robustness analysis evaluated 111 combinations in which each component weight ranged from 0.10 to 0.70. For each combination, HVI scores and rankings were recalculated. Robustness was evaluated using the Spearman rank correlation with the baseline ranking and the overlap with the baseline top ten areas.

4. Results and discussion

4.1 Building-level LST and thermal exposure

The Landsat-derived summer 2023 LST composite revealed substantial variation in thermal exposure among buildings across Bologna. Valid mean LST values were obtained for 51,976 building footprints. As summarized in **Table 3**, mean summer LST ranged from 30.47 °C to 57.97 °C, with an overall mean of 44.70 °C and a standard deviation of 3.03 °C. The interquartile range extended from 43.26 °C to 46.71 °C, indicating that most buildings fell within a relatively warm thermal band, while a smaller number experienced distinctly cooler or hotter surface conditions.

ARTICLE IN PRESS

Table 3. Descriptive statistics of building-level mean summer LST in Bologna

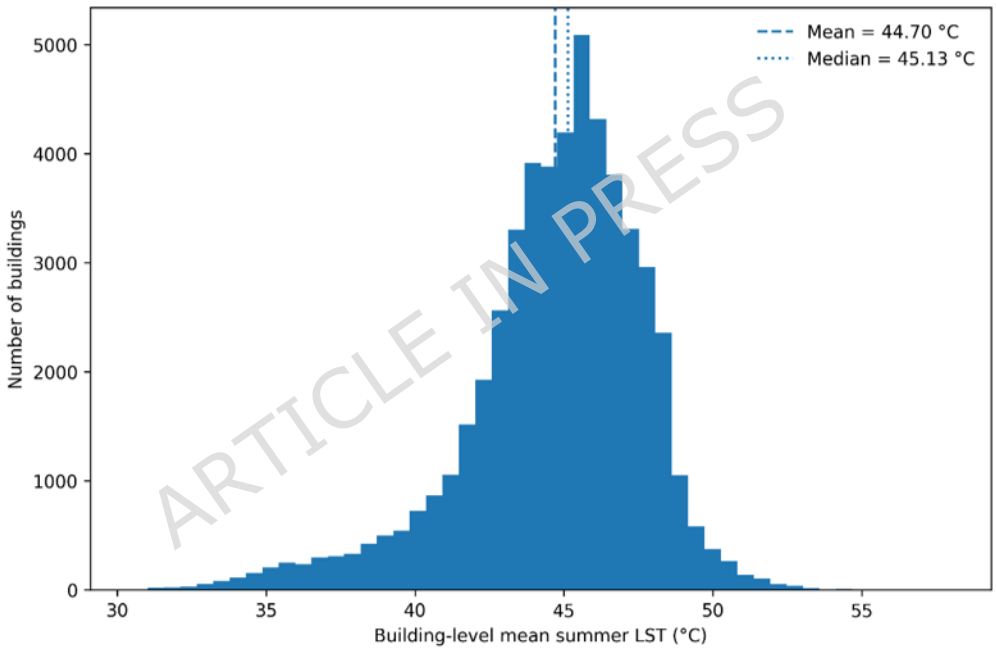
Indicator	**Value**
Buildings with valid LST values	51,976
Mean of LST values	44.70 °C
Standard deviation of LST values	3.03 °C
Minimum of LST values	30.47 °C
25th percentile of LST values	43.26 °C
Median of LST values	45.13 °C
75th percentile of LST values	46.71 °C
Maximum of LST values	57.97 °C

ARTICLE IN PRESS

The distribution shown in **Fig. 4** is centred around 45 °C, with the median slightly higher than the mean. This pattern suggests that, although most buildings fall within a narrow range of high summer surface temperatures, local conditions create clear deviations from the central tendency. These differences are important for the subsequent modelling stage, as they indicate that building-level heat exposure is not uniform across the city.

These spatial differences provided the basis for the RF analysis. Rather than treating Bologna as a thermally homogeneous urban area, the building-level dataset allowed LST to be examined in relation to urban morphology, land-cover composition, vegetation, and proximity to green infrastructure. This interpretation is consistent with evidence that intra-urban heat patterns reflect both regional climate and local differences in surface characteristics, vegetation, and built form [17,19,41].

Fig. 4 Distribution of building-level mean summer LST in Bologna, derived from the summer 2023 Landsat 8 LST composite



ARTICLE IN PRESS

4.2 Random forest modelling of building-level LST

Using the spatially tuned RF configuration, the conventional 80:20 random split produced R^2 = 0.749, MAE = 1.16 °C, and RMSE = 1.50 °C. Under five-fold spatially blocked cross-validation with 3 km blocks and a 300 m exclusion buffer, pooled out-of-fold performance decreased to R^2 = 0.579, MAE = 1.58 °C, and RMSE = 1.97 °C. Mean fold-level performance was R^2 = 0.525 ± 0.175, MAE = 1.60 ± 0.22 °C, and RMSE = 1.98 ± 0.27 °C. **Table 4** summarizes the model performance under the random split and the two spatial cross-validation reporting approaches, including pooled out-of-fold predictions and the mean and standard deviation across the five folds. The decline under spatial validation shows that the random split was optimistic because nearby buildings share similar thermal and environmental conditions. Even so, the positive pooled

ARTICLE IN PRESS

spatial R^2 indicates that the model retained meaningful predictive capability in spatially separated parts of Bologna. The analysis included 51,976 buildings and 29 predictors. Fold-level spatial cross-validation values are reported as mean ± standard deviation across five folds.

Table 4. Performance of the spatially tuned Random Forest model under random and spatial validation

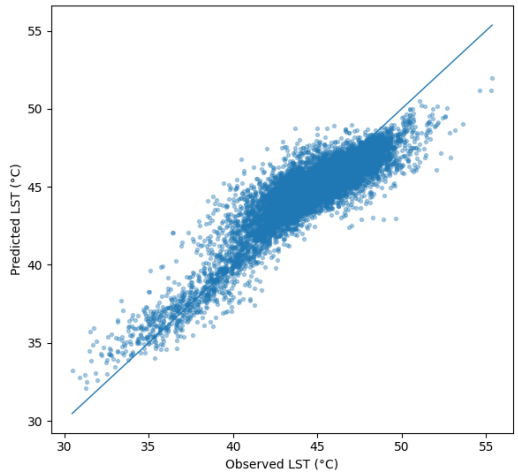
Validation strategy	**R^2**	**MAE**	**RMSE**
Random 80:20 split	0.749	1.16 °C	1.50 °C
Spatial CV pooled predictions	0.579	1.58 °C	1.97 °C
Spatial CV mean ± SD	0.525 ± 0.175	1.60 ± 0.22 °C	1.98 ± 0.27 °C

Panel (a) of **Fig. 5** shows the reduction in R^2 from random to spatial validation, while panel (b) shows that the pooled observed and predicted LST values remained broadly aligned with the 1:1 line under spatial cross-validation. Errors were larger at the coolest and hottest ends of the LST distribution, indicating some regression toward the mean in spatially separated areas. These deviations may reflect local surface materials, microscale shading, mixed Landsat pixels, or urban characteristics that were not fully represented by the available predictors.

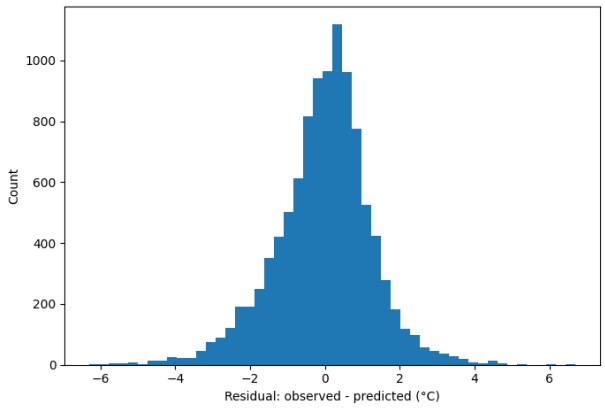
The feature-importance ranking indicates that DEM-derived terrain and surface-exposure proxies, built-up and tree-cover context, Urban Atlas land-cover composition, and park proximity contributed most strongly to the modelled LST variation (**Fig. 5c**). Local relief had the highest mean importance, followed by built-up-cover percentage, tree-cover percentage, the DEM-derived sky-view factor, and the shadow index. Industrial, commercial, and public land, distance to the nearest park or garden, continuous urban fabric, and forest cover also contributed to prediction. These importance values describe predictive relevance rather than the direction or causal effect of each predictor.

ARTICLE IN PRESS

(a)



(b)



ARTICLE IN PRESS

(c)

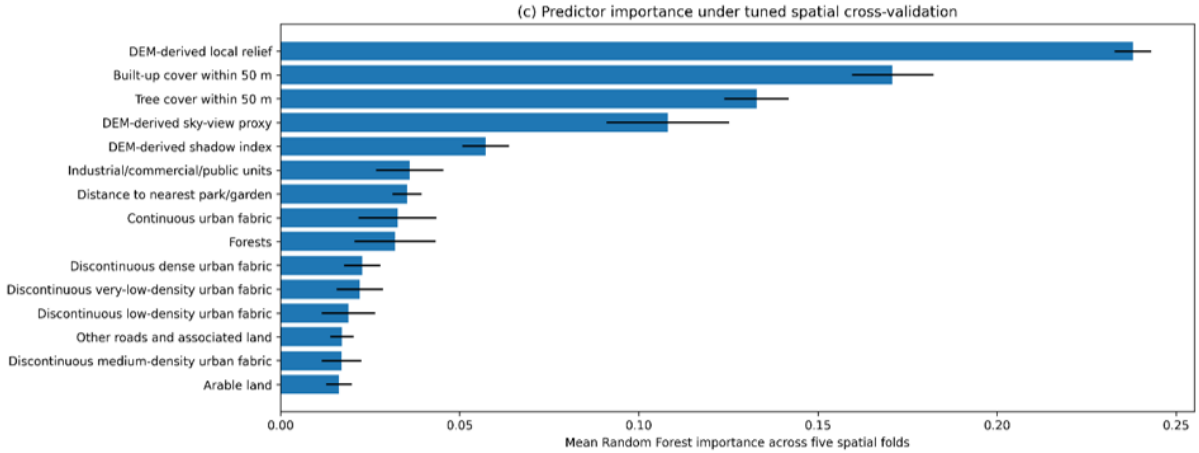


Fig. 5 Random Forest validation and interpretation results: (a) comparison of R^2 under random and spatial validation, (b) pooled observed-versus-predicted LST under five-fold spatial block cross-validation, and (c) mean predictor importance across the five spatial folds.

4.3 Spatial predictors of building-level heat exposure

The feature-importance ranking provides insight into the predictor groups associated with building-level LST variation, although individual values should be interpreted cautiously because correlated predictors can share or dilute importance. The Spearman analysis identified two strongly correlated predictor pairs: the DEM-derived sky-view and shadow proxies ($\rho = 0.721$) and the built-up and tree-cover percentages ($\rho = -0.717$). No other pair had $|\rho| \geq 0.70$ (Supplementary Fig. S2). As shown in Table 5 and Fig. 5c, local relief was the highest-ranked predictor, followed by built-up-cover percentage, tree-cover percentage, the DEM-derived sky-view factor, and the shadow index. Because building-scale three-dimensional geometry was not directly modelled, the terrain-derived variables should be understood as exposure proxies rather than detailed representations of individual buildings or street canyons.

Built-up and tree-cover percentages were both among the strongest contributors. Their negative correlation shows that they represent opposing but overlapping land-cover conditions, so their values should not be interpreted as independent effects. The high ranking of tree cover is consistent with its potential cooling role, whereas the importance of built-up cover indicates predictive relevance without establishing the direction of its relationship with LST. Similarly, the correlated sky-view and shadow proxies are best interpreted jointly as broader indicators of terrain and surface exposure. Distance to the nearest park or garden also remained among the leading predictors, supporting the relevance of green-infrastructure proximity to local thermal conditions.

The land-cover results further show that the surrounding urban fabric contributes to building-level thermal exposure. Industrial, commercial, and public urban land, continuous urban fabric, forest cover, and dense discontinuous urban fabric all appeared among the ten leading

ARTICLE IN PRESS

predictors. This pattern is consistent with previous UHI research identifying vegetation, impervious surfaces, urban geometry, and surface exposure as important controls of intra-urban temperature variation. The rankings nevertheless represent model-based associations rather than direct causal effects. Their main value is to identify plausible spatial drivers that can guide further analysis and support planning hypotheses.

Table 5. Top ten predictors from the tuned RF model averaged across five spatial folds

Rank	Indicator	Mean importance
1	Local relief / terrain-context indicator	0.238
2	Built-up-cover percentage within 50 m	0.171
3	Tree-cover percentage within 50 m	0.133
4	Mean DEM-derived sky-view factor proxy	0.108
5	Mean DEM-derived shadow index	0.057
6	Industrial, commercial, and public urban land within 50 m	0.036
7	Distance to the nearest park or garden	0.035
8	Continuous urban fabric within 50 m	0.033
9	Forest cover within 50 m	0.032
10	Discontinuous dense urban fabric within 50 m	0.023

ARTICLE IN PRESS

4.4 Tuned LSTM benchmarking against the persistence baseline

The temporal component was evaluated using PVGIS-SARAH3 hourly climate data. The LSTM model was trained on observations from 2019–2022 and tested on the independent 2023 period. A 24-hour lookback window was used to predict air temperature one hour ahead.

Hyperparameter tuning selected a 24-hour lookback, 64 hidden units, one LSTM layer, no dropout, a learning rate of 0.001, and a batch size of 512. After retraining on the complete 2019–2022 period for 15 epochs, the model achieved R^2 = 0.995, MAE = 0.43 °C, and RMSE = 0.59 °C on the independent 2023 test year (**Table 6**). The annual and July comparisons show that the model captured both seasonal variation and short-term daily cycles, including most temperature peaks (**Figs. 6a and 6b**). Observed and predicted values were closely aligned with the 1:1 line (**Fig. 6c**), while the residuals were centred near zero, with a mean residual of 0.13 °C (**Fig. 6d**). Training loss declined consistently over the 15 epochs and approached convergence after the initial rapid decrease (**Fig. 6e**).

Table 6. Performance of the tuned LSTM model and persistence baseline for one-hour-ahead air-temperature prediction

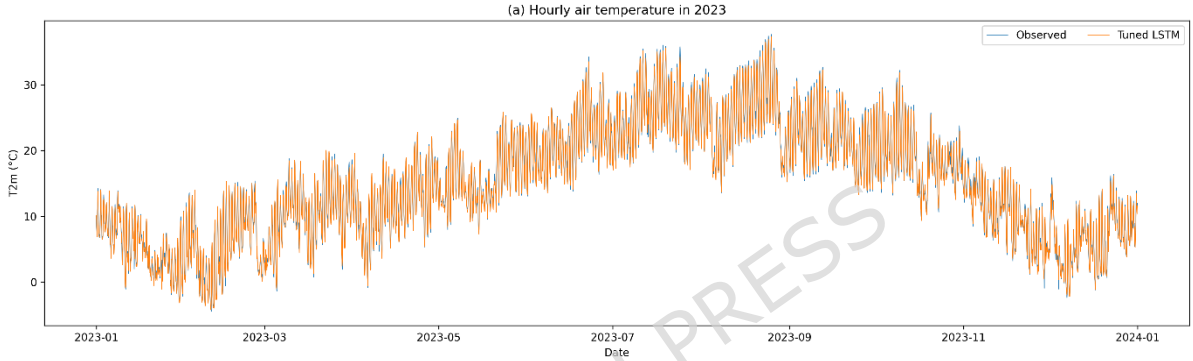
Model	R^2	MAE	RMSE
Tuned LSTM	0.995	0.43 °C	0.59°C

ARTICLE IN PRESS

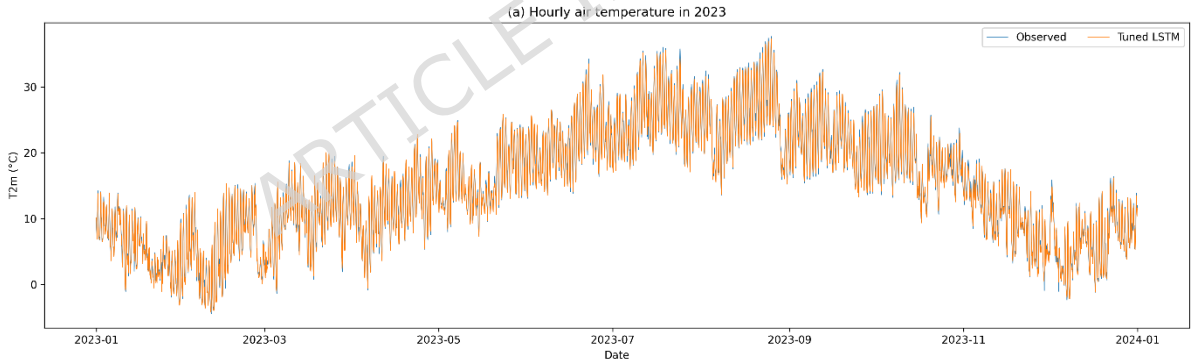
Persistence baseline	0.984	0.78 °C	1.07 °C

The persistence baseline performed strongly because hourly air temperature is highly autocorrelated and the next-hour value is often close to the current observation. Nevertheless, the tuned LSTM reduced RMSE by approximately 45% relative to persistence at the one-hour horizon. This improvement indicates that the sequence model captured useful temporal structure beyond simple short-term continuity. Retaining the persistence benchmark remains important, however, because it provides a demanding and transparent reference for evaluating models of highly autocorrelated temperature series.

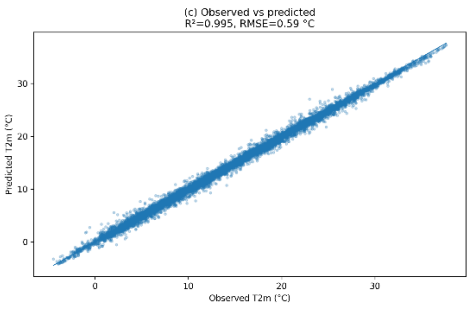
(a)



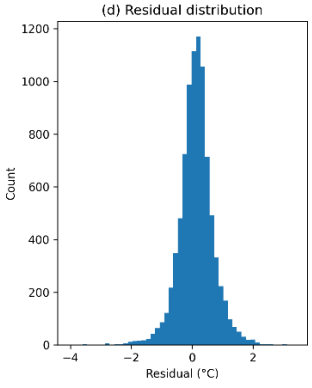
(b)



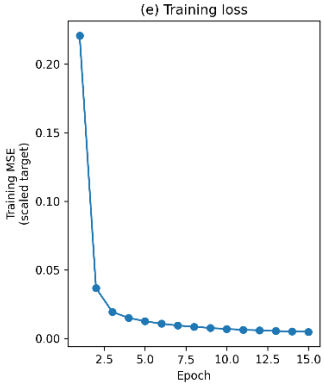
(c)



(d)



(e)



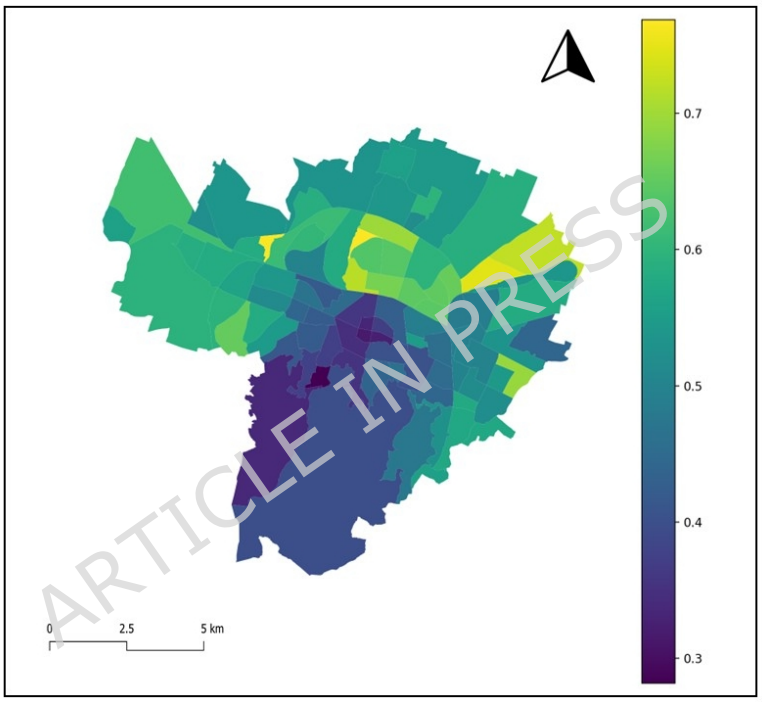
ARTICLE IN PRESS

Fig. 6 Tuned LSTM results for one-hour-ahead air temperature prediction in 2023: (a) annual observed and predicted air temperature, (b) July 2023 detail, (c) observed-versus-predicted values, (d) residual distribution, and (e) training loss.

4.5 Social vulnerability and heat vulnerability index

The social vulnerability index was calculated for Bologna's 90 statistical areas using indicators related to income, employment, education, age structure, gender, citizenship, and birthplace. Its spatial distribution and the 15 highest-scoring areas are shown in **Fig. 7**. CNR, Lungo Reno, Pilastro, CAAB, and Ex Mercato Ortofrutticolo recorded the highest values, while **Table 7** lists the ten leading areas.

(a)



(b)

ARTICLE IN PRESS

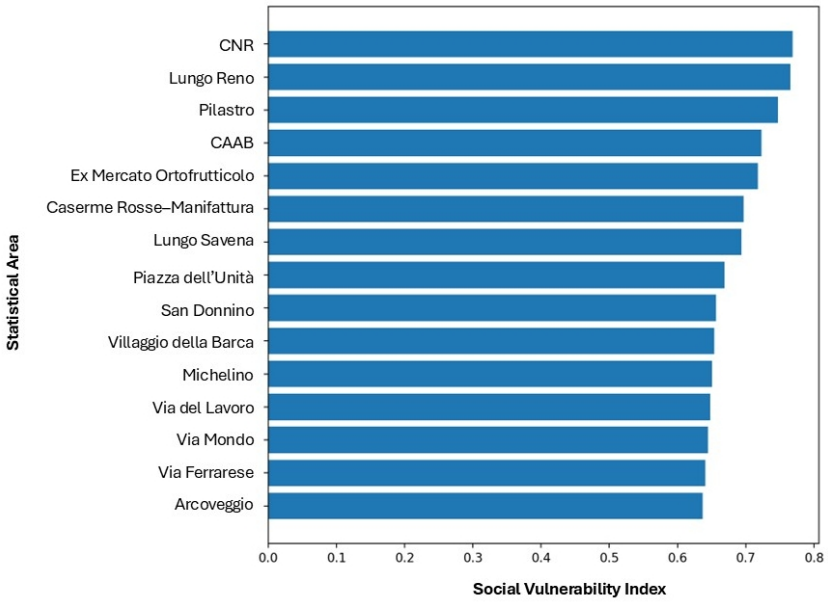


Fig. 7 Social vulnerability results for Bologna's statistical areas: (a) spatial distribution of the social vulnerability index and (b) the 15 statistical areas with the highest social vulnerability scores

Social vulnerability was not confined to the historic center. Higher values occurred mainly in the northern, north-eastern, and western parts of the municipality, particularly within Navile, San Donato–San Vitale, and Borgo Panigale–Reno. This pattern shows that vulnerable populations are also concentrated in peripheral and semi-peripheral areas, where heat exposure may compound broader socioeconomic disadvantages.

Table 7. Statistical areas with the highest social vulnerability scores

Rank	Statistical area	Quartiere	Social vulnerability index
1	CNR	Navile	0.769
2	Lungo Reno	Borgo Panigale–Reno	0.766
3	Pilastro	San Donato–San Vitale	0.747
4	CAAB	San Donato–San Vitale	0.723
5	Ex Mercato Ortofrutticolo	Navile	0.718
6	Caserme Rosse–Manifattura	Navile	0.697
7	Lungo Savena	Savena	0.694
8	Piazza dell'Unità	Navile	0.669
9	San Donnino	San Donato–San Vitale	0.656
10	Villaggio della Barca	Borgo Panigale–Reno	0.654

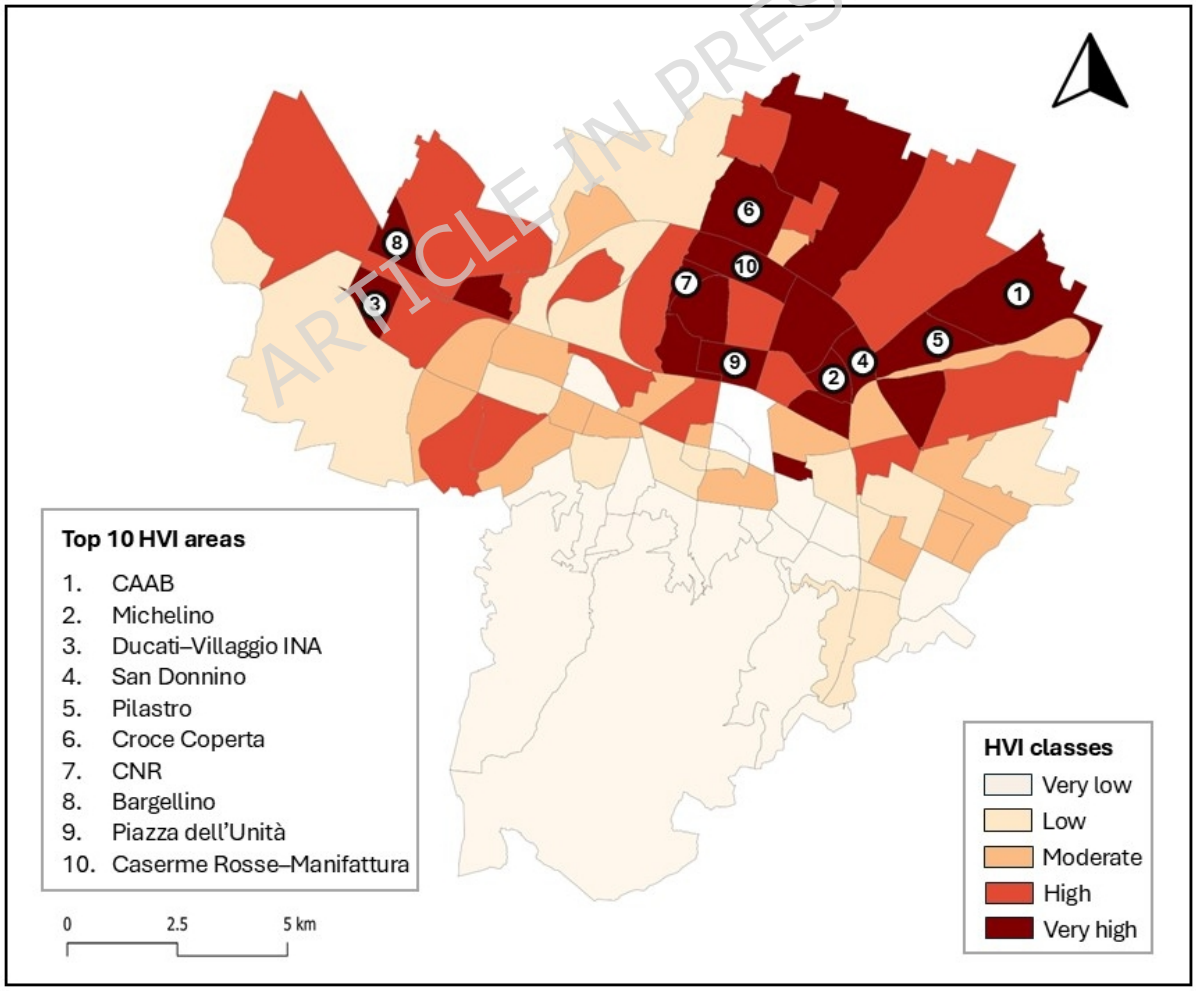
The final HVI was developed by combining heat exposure, socio-economic vulnerability, and demographic vulnerability. The three components were weighted as follows:

ARTICLE IN PRESS

$$HVI = \mathbf{0.50} \times HeatExposure + \mathbf{0.30} \times SocioeconomicVulnerability + \mathbf{0.20} \times DemographicVulnerability$$

The spatial distribution of the HVI is shown in **Fig. 8**, and the ten highest-priority statistical areas are listed in Table 8. Across the 90 areas, mean LST ranged from 35.68 °C to 47.98 °C, with an average of 44.01 °C. This area-level average differs from the building-level mean of 44.70 °C because the two values are based on different spatial units of aggregation. CAAB, in San Donato–San Vitale, recorded the highest HVI score (0.822), followed by Michelino, Ducati–Villaggio INA, San Donnino, Pilastro, Croce Coperta, CNR, Bargellino, Piazza dell'Unità, and Caserme Rosse–Manifattura.

Comparing the social vulnerability and final HVI rankings illustrates why neither social conditions nor temperature alone is sufficient for prioritization. CNR had the highest social vulnerability score but ranked seventh after heat exposure was incorporated. In contrast, CAAB ranked fourth for social vulnerability and moved to first place in the combined HVI. Priority therefore emerges where physical heat exposure, socioeconomic disadvantage, and demographic sensitivity overlap.



ARTICLE IN PRESS

Fig. 8 HVI across Bologna's 90 statistical areas. The index combines area-level heat exposure, socio-economic vulnerability, and demographic vulnerability, and is classified into five vulnerability levels from very low to very high

Table 8. Statistical areas with the highest HVI values

Rank	Statistical area	Quartiere	HVI	Main planning interpretation
1	CAAB	San Donato–San Vitale	0.822	Highest combined heat-vulnerability priority
2	Michelino	San Donato–San Vitale	0.790	High heat exposure and vulnerability overlap
3	Ducati–Villaggio INA	Borgo Panigale–Reno	0.780	Priority for heat-sensitive redevelopment
4	San Donnino	San Donato–San Vitale	0.777	Priority for green infrastructure improvement
5	Pilastro	San Donato–San Vitale	0.775	High social vulnerability and heat-risk relevance
6	Croce Coperta	Navile	0.775	Heat vulnerability priority area
7	CNR	Navile	0.752	Highest social vulnerability, also high HVI
8	Bargellino	Borgo Panigale–Reno	0.749	Priority western area
9	Piazza dell'Unità	Navile	0.740	Social vulnerability and heat-risk overlap
10	Caserme Rosse - Manifattura	Navile	0.739	Priority area in Navile

4.6 Sensitivity of the HVI ranking

The HVI ranking remained highly stable under plausible changes in component weights. Across the 91 local weight combinations, the median Spearman rank correlation with the baseline ranking was 0.992, and the minimum was 0.973. The median overlap with the baseline top ten was nine areas, while the minimum overlap was seven. CAAB remained the highest-ranked area under all 91 local combinations, and the first six baseline areas—CAAB, Michelino, Ducati–Villaggio INA, San Donnino, Pilastro, and Croce Coperta—remained in the top ten in every local configuration. Under the broader set of 111 combinations, the median Spearman correlation was 0.921 and the median top ten overlap was seven areas. The principal priority areas were therefore robust to plausible changes in weighting, although more extreme shifts in policy priorities affected the lower positions in the ranking.

Fig. 9 summarizes this stability under the two sensitivity settings. Panel (a) presents the rank ranges of the baseline top 15 areas across the 91 local weight combinations, while panel (b) presents the Spearman correlations with the baseline ranking across the 111 broader combinations.

ARTICLE IN PRESS

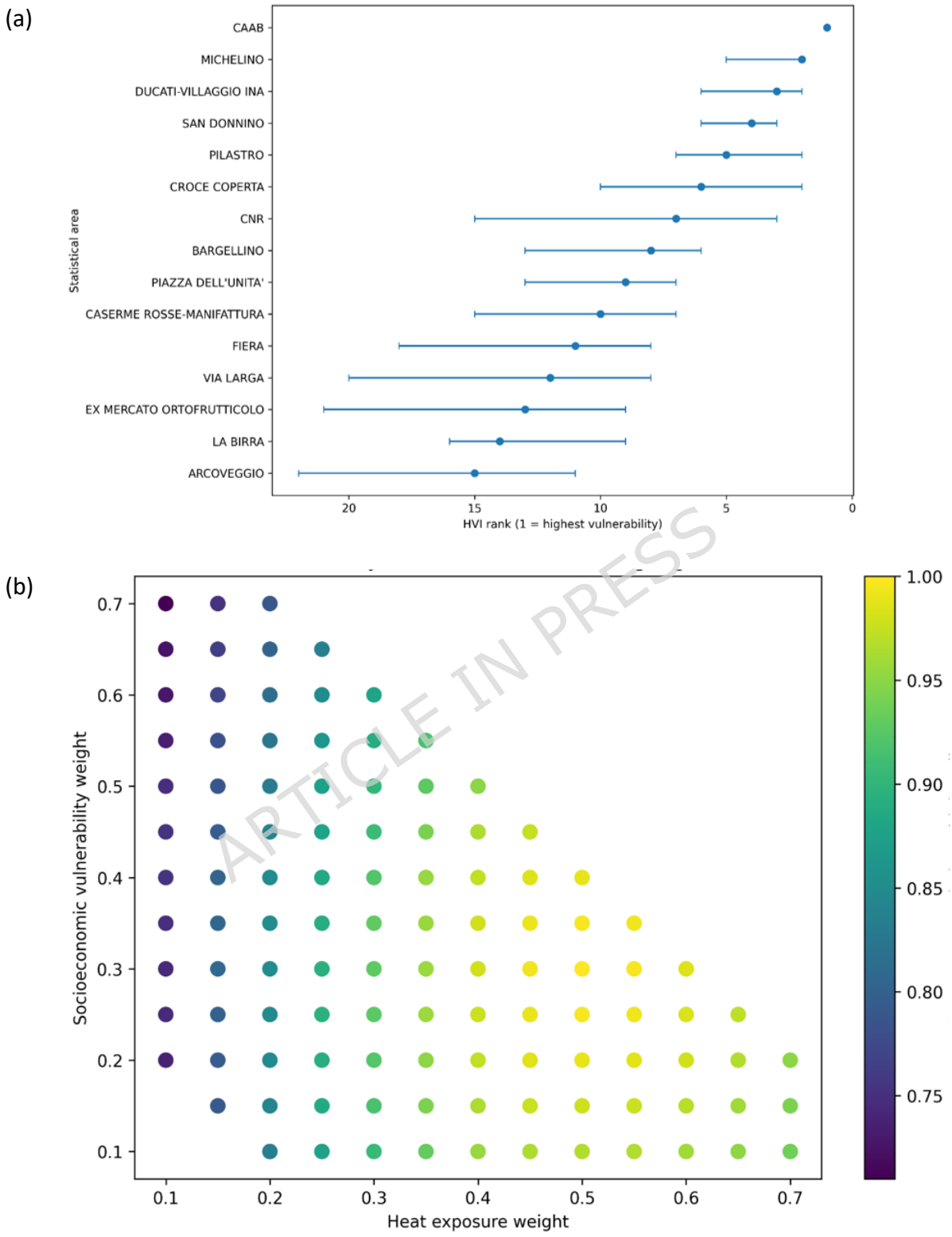


Fig. 9 HVI weight-sensitivity analysis showing (a) rank ranges of the baseline top 15 statistical areas across 91 local weight combinations and (b) Spearman rank correlation with the baseline ranking across 111 broader weight combinations.

ARTICLE IN PRESS

4.7 Implications for sustainable development and climate resilient urban planning

The findings have direct implications for sustainable development because they identify where environmental exposure, limited adaptive capacity, and deficits in cooling infrastructure intersect. Previous research has shown that access to green infrastructure and its cooling benefits can be spatially unequal, making the social targeting of adaptation measures an important part of sustainable urban planning [41–44,57,60]. In practice, the HVI can help municipal authorities direct limited resources toward areas where tree planting, shade provision, cool materials, improved park access, and heat-health measures are likely to produce both thermal and social benefits. These applications align with SDG 3 through the reduction of heat-related health risks, SDG 10 through attention to disadvantaged communities, SDG 11 through more inclusive and resilient urban development, and SDG 13 through locally targeted climate adaptation [79]. The HVI should nevertheless be understood as a decision-support and prioritization tool, not as a direct measure of SDG achievement.

The priority areas should not all receive the same package of interventions. Neighbourhoods with limited vegetation and high surface temperatures may benefit from tree-canopy expansion, shaded pedestrian routes, stronger green-space connectivity, and cool or permeable surface materials [36–44]. Where high exposure also overlaps with socioeconomic or demographic vulnerability, these physical measures may need to be complemented by accessible cooling spaces, targeted risk communication, and support for sensitive population groups [10,54–60]. Such a differentiated approach is more effective than uniform citywide action because it brings together environmental performance, social equity, and efficient use of public resources.

The workflow also has potential value beyond Bologna because it relies largely on open or widely available datasets and reproducible GIS operations. Landsat can provide summer LST; Copernicus DEM, Urban Atlas, and ESA WorldCover can describe terrain, land cover, vegetation, and urban context; OpenStreetMap or municipal cadastral sources can provide building footprints; and local statistical offices can supply socioeconomic and demographic data. Hourly observations for temporal modelling can similarly be obtained from meteorological stations or European climate services.

In other cities, the framework can be implemented as a modular sequence: harmonizing geospatial layers, extracting building-level LST and spatial predictors, training an interpretable model of heat exposure, analysing temporal climate dynamics, and constructing a vulnerability index at an appropriate local statistical scale. This structure is particularly relevant to compact European cities with historic centers, heterogeneous neighborhoods, limited land availability, and growing pressure to adapt the built environment to extreme heat.

Transferability, however, should not be confused with direct duplication. Indicator weights, HVI thresholds, vulnerability variables, and planning priorities need to be adapted to local climate,

ARTICLE IN PRESS

data availability, urban morphology, policy goals, heritage constraints, and community needs. Because the framework was evaluated only in Bologna, its performance in other settings has not yet been demonstrated empirically. Multi-city validation is therefore needed before broader generalization, particularly because model accuracy, predictor importance, and vulnerability rankings may change with local environmental and social conditions.

5. Summary and conclusions

This study developed an integrated geospatial machine-learning framework to assess urban heat exposure, temporal climate dynamics, and social vulnerability in Bologna, Italy. By combining Landsat-derived land surface temperature (LST), building footprints, urban-morphology indicators, vegetation and land-cover variables, park proximity, hourly meteorological data, and socioeconomic and demographic indicators, the framework provides planning-oriented evidence on where physical heat exposure and social vulnerability overlap. The main findings are as follows:

- The building-level analysis revealed substantial intra-urban variation in surface heat across Bologna. Summer 2023 LST values were obtained for 51,976 building footprints, ranging from 30.47 °C to 57.97 °C, with a mean of 44.70 °C. These results show that heat exposure is highly uneven across the city and cannot be adequately represented by citywide averages alone.
- Using spatially tuned hyperparameters, the Random Forest (RF) model achieved $R^2 = 0.749$, MAE = 1.16 °C, and RMSE = 1.50 °C under the conventional random split. Under five-fold spatially blocked cross-validation, pooled performance decreased to $R^2 = 0.579$, MAE = 1.58 °C, and RMSE = 1.97 °C. This comparison shows that random validation was optimistic, although the tuned model retained meaningful predictive capability in spatially separated locations. The most influential predictors included DEM-derived local relief, built-up cover, tree cover, sky-view and shadow proxies, industrial and commercial land, and proximity to parks.
- The tuned Long Short-Term Memory (LSTM) model captured hourly air-temperature dynamics in 2023 with $R^2 = 0.995$, MAE = 0.43 °C, and RMSE = 0.59 °C. It outperformed the persistence baseline, which achieved $R^2 = 0.984$, MAE = 0.78 °C, and RMSE = 1.07 °C. This finding highlights the value of systematic hyperparameter tuning and the need to compare complex forecasting models with simple baselines.
- The Heat Vulnerability Index (HVI) showed that heat-risk priority is not determined by temperature alone. By combining area-level LST with socioeconomic and demographic vulnerability across 90 statistical areas, the analysis identified CAAB as the highest-priority area, followed by Michelino, Ducati-Villaggio INA, San Donnino, Pilastro, Croce Coperta, CNR, Bargellino, Piazza dell'Unità, and Caserme Rosse-Manifattura. Comparing the social-

ARTICLE IN PRESS

vulnerability and final HVI rankings confirmed that the areas with the greatest social vulnerability are not always those with the highest combined heat-risk priority.

- The results provide practical guidance for climate-resilient urban planning in Bologna. High-HVI areas should be prioritized for tree-canopy expansion, shaded pedestrian routes, cool-surface strategies, improved park accessibility, and heat-sensitive zoning. These measures are particularly important in dense, historically developed areas where limited land availability requires targeted and spatially efficient adaptation. Sensitivity analysis also confirmed that the HVI ranking was highly robust to plausible changes in component weights, with a median Spearman correlation of 0.992 and median retention of nine of the baseline top 10 areas.
- The framework can potentially be adapted to other compact European cities after local calibration and external validation. Its indicator weights, vulnerability variables, classification thresholds, and planning interpretations should be adjusted to reflect local climate, urban morphology, data availability, social structure, policy objectives, and heritage constraints.

The proposed framework offers a reproducible, spatially detailed, and interpretable decision-support approach for jointly prioritizing heat mitigation, reductions in social vulnerability, and green-infrastructure improvements. It can help urban planners, public authorities, and policymakers design more targeted, equitable, and climate-resilient adaptation strategies for Bologna and comparable compact European cities.

5.1 Limitations and future directions

Several limitations should be considered when interpreting the results. First, the LST layer was derived from Landsat 8 imagery. Although Landsat provides consistent thermal information at the city scale, its spatial resolution cannot fully capture microscale variation in narrow streets, shaded pedestrian spaces, small courtyards, or individual building facades. Future studies could address this limitation by integrating higher-resolution thermal imagery, UAV-based surveys, mobile temperature measurements, or fixed in-situ sensors.

Second, the representation of three-dimensional urban form was limited by the available elevation data. The DEM-derived sky-view, shadow, and relief variables describe terrain-scale and surface-exposure conditions but do not reliably capture individual building heights, street-canyon height-to-width ratios, built volume, aerodynamic roughness, facade orientation, or mean radiant temperature. The analysis also did not include Local Climate Zone classifications or UMEP-based mean radiant temperature simulations. Future research should integrate a citywide digital surface model, building-height data, detailed street geometry, and microclimate modelling to assess the additional explanatory value of more complete three-dimensional urban form indicators.

Third, the RF model identifies statistical associations between urban predictors and building-level LST and should not be interpreted as a causal model. Although only two predictor pairs exceeded

ARTICLE IN PRESS

the predefined correlation threshold, correlated and compositional variables can still redistribute impurity-based feature importance among predictors. The ranking should therefore be interpreted primarily at the predictor-group level. Future studies could strengthen interpretation using SHAP values, partial-dependence plots, conditional or grouped permutation importance, and causal-inference approaches where appropriate data are available.

Fourth, the tuned LSTM was evaluated only for a one-hour-ahead forecast at a single representative PVGIS location. Although it outperformed persistence at this horizon, the result does not establish its performance for longer lead times, heatwave events, or spatially distributed urban conditions. Future studies should test multi-step forecasting, incorporate additional meteorological stations or urban sensor networks, and evaluate models designed specifically to predict extreme-temperature events.

Fifth, the HVI depends on the selected indicators, weighting scheme, and classification thresholds. The 0.50-0.30-0.20 weighting structure was designed to prioritize heat exposure while still accounting for socioeconomic and demographic vulnerability. The sensitivity analysis showed that the ranking was highly stable under plausible local changes, although broader reweighting produced larger shifts, particularly among lower-ranked priority areas. The sensitivity of the classification thresholds was not assessed and should be examined in future applications.

Finally, although the framework could be adapted to other European cities, it requires local calibration. Differences in climate, data availability, building morphology, social indicators, planning priorities, and heritage constraints may affect both model performance and policy interpretation. Future research should therefore apply the workflow across multiple years, different heatwave events, and other urban contexts. Integrating in-situ air-temperature or thermal-comfort measurements would also strengthen validation and improve the framework's usefulness for operational heat-risk planning.

ARTICLE IN PRESS

Funding

This research received no external funding.

Competing Interests

The authors declare that they have no competing interests or personal relationships that could have appeared to influence the work reported in this study.

Data availability

The datasets used in this study were obtained from publicly available sources, including PVGIS-SARAH3, Landsat 8 Collection 2 Level 2 imagery, Copernicus DEM, Copernicus Urban Atlas, ESA WorldCover, OpenStreetMap, and open datasets provided by Comune di Bologna. The processed

ARTICLE IN PRESS

datasets generated during this study are available from the corresponding author upon reasonable request.

Code availability

The code used for data preprocessing, machine-learning modelling, performance evaluation, and visualization is available from the corresponding author upon reasonable request.

Authors' contributions

Mahdiyeh Tabatabaei: Writing – original draft, conceptualization, research design, geospatial data processing, machine-learning modelling, interpretation of results, and manuscript revision. **Aniseh Saber:** Writing – original draft, conceptualization, data curation and preparation, formal analysis, interpretation of results, visualization, and manuscript revision. **Ali Pourzangbar:** Methodology, investigation, formal analysis, conceptualization. All authors reviewed and approved the final manuscript.

Declarations

Ethical approval

Not applicable. This study used publicly available and aggregated environmental, geospatial, socioeconomic, and demographic data and did not involve human participants, animals, interventions, or identifiable personal information.

Consent to participate

Not applicable.

Consent to publish

Not applicable.

ARTICLE IN PRESS

References

[1] Akinsanola AA, Singhai P, Taguela TN, Folorunsho AH, Adeyeri OE, Morakinyo TE, et al. A review of urban resilience to weather and climate extremes. City and Built Environment 2025;3. https://doi.org/10.1007/s44213-025-00063-6.

[2] European Commission. The European Green Deal. COM(2019) 640 final. Brussels: European Commission 2019.

[3] European Parliament and the Council. European Climate Law. vol. 2021. 2021.

[4] Sendai framework for disaster risk reduction 2015-2030. Australian Journal of Emergency Management 2015;30.

ARTICLE IN PRESS

[5] Dodman D, Hayward, B. M, Pelling V, Broto CW, Chow E, Chu R, et al. Cities, Settlements and Key Infrastructure. In: Climate Change 2022: Impacts, Adaptation and Vulnerability. Climate Change 2022 – Impacts, Adaptation and Vulnerability, 2022.

[6] IPCC. Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. 2014.

[7] Oke TR. The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society 1982;108. https://doi.org/10.1002/qj.49710845502.

[8] Santamouris M. Recent progress on urban overheating and heat island research. Integrated assessment of the energy, environmental, vulnerability and health impact. Synergies with the global climate change. Energy Build 2020;207. https://doi.org/10.1016/j.enbuild.2019.109482.

[9] Cutter SL, Boruff BJ, Shirley WL. Social vulnerability to environmental hazards. Soc Sci Q 2003;84. https://doi.org/10.1111/1540-6237.8402002.

[10] Reid CE, O'Neill MS, Gronlund CJ, Brines SJ, Brown DG, Diez-Roux A V., et al. Mapping community determinants of heat vulnerability. Environ Health Perspect 2009;117. https://doi.org/10.1289/ehp.0900683.

[11] Hincks S, Carter J, Connelly A. A new typology of climate change risk for European cities and regions: Principles and applications. Global Environmental Change 2023;83. https://doi.org/10.1016/j.gloenvcha.2023.102767.

[12] Li F, Yigitcanlar T, Nepal M, Thanh KN, Dur F. Understanding Urban Heat Vulnerability Assessment Methods: A PRISMA Review. Energies (Basel) 2022;15. https://doi.org/10.3390/en15196998.

[13] Tabatabaei M, Saber A. Equity- and Heritage-Informed Rooftop PV Screening for the Sustainable Development of Renewable Energy Communities in Bologna, Italy. Sustainability (Switzerland) 2026;18. https://doi.org/10.3390/su18104774.

[14] Tabatabaei M. Analysis of Traditional Iranian Vernacular Architecture in Response to Sunlight: An Approach to Sustainable Development. Urban Sustainability, vol. Part F4270, 2024. https://doi.org/10.1007/978-981-96-1116-4_14.

[15] Bao Y, Huang Z, Yin G, Ren S, Yan X, Qi J. Quantifying the impact of building material stock and green infrastructure on urban heat island intensity. Build Environ 2025;280. https://doi.org/10.1016/j.buildenv.2025.113068.

[16] Jonathan R, Lin T, Lun I, Widijatmoko SD, Tang Y-T. Assessing the Impact of Building Surface Materials on Local Thermal Environment Using Infrared Thermal Imagery and Microclimate Simulations. Buildings 2026;16. https://doi.org/10.3390/buildings16020334.

ARTICLE IN PRESS

[17] Li Y, Schubert S, Kropp JP, Rybski D. On the influence of density and morphology on the Urban Heat Island intensity. Nat Commun 2020;11. https://doi.org/10.1038/s41467-020-16461-9.

[18] Cheng X, Liu Y, Dong J, Corcoran J, Peng J. Opposite climate impacts on urban green spaces' cooling efficiency around their coverage change thresholds in major African cities. Sustain Cities Soc 2023;88. https://doi.org/10.1016/j.scs.2022.104254.

[19] Jang S, Jung J. Urban form and green space structure as drivers of urban heat mitigation. Sustain Cities Soc 2025;130. https://doi.org/10.1016/j.scs.2025.106597.

[20] Voogt JA, Oke TR. Thermal remote sensing of urban climates. Remote Sens Environ 2003;86. https://doi.org/10.1016/S0034-4257(03)00079-8.

[21] Weng Q. Thermal infrared remote sensing for urban climate and environmental studies: Methods, applications, and trends. ISPRS Journal of Photogrammetry and Remote Sensing 2009;64. https://doi.org/10.1016/j.isprsjprs.2009.03.007.

[22] Khan MS, Ullah S, Chen L. Comparison on land-use/land-cover indices in explaining land surface temperature variations in the city of beijing, china. Land (Basel) 2021;10. https://doi.org/10.3390/land10101018.

[23] Cao W, Zhou W, Yu W, Wu T. Combined effects of urban forests on land surface temperature and PM2.5 pollution in the winter and summer. Sustain Cities Soc 2024;104. https://doi.org/10.1016/j.scs.2024.105309.

[24] Zhang P, Zhang J, Liu Z, Liu Y, Chen Z. Relationship between land surface temperature and air quality in urban and suburban areas: Dynamic changes and interaction effects. Sustain Cities Soc 2025;118. https://doi.org/10.1016/j.scs.2024.106043.

[25] Schwarz N, Schlink U, Franck U, Großmann K. Relationship of land surface and air temperatures and its implications for quantifying urban heat island indicators - An application for the city of Leipzig (Germany). Ecol Indic 2012;18. https://doi.org/10.1016/j.ecolind.2012.01.001.

[26] Briegel F, Pinto JG, Christen A. Is satellite land surface temperature an appropriate proxy for intra-urban variability of daytime heat stress? Remote Sens Environ 2025;331. https://doi.org/10.1016/j.rse.2025.115045.

[27] Logan TM, Zaitchik B, Guikema S, Nisbet A. Night and day: The influence and relative importance of urban characteristics on remotely sensed land surface temperature. Remote Sens Environ 2020;247. https://doi.org/10.1016/j.rse.2020.111861.

[28] Stewart ID, Oke TR. Local climate zones for urban temperature studies. Bull Am Meteorol Soc 2012;93. https://doi.org/10.1175/BAMS-D-11-00019.1.

ARTICLE IN PRESS

[29] Bechtel B, Alexander PJ, Böhner J, Ching J, Conrad O, Feddema J, et al. Mapping local climate zones for a worldwide database of the form and function of cities. ISPRS Int J Geoinf 2015;4. https://doi.org/10.3390/ijgi4010199.

[30] Wang M, Wang R, Ouyang W, Tan Z. A multi-scale mapping approach of local climate zones: A case study in Hong Kong. Urban Clim 2025;61. https://doi.org/10.1016/j.uclim.2025.102446.

[31] Saber A. Effects of window-to-wall ratio on energy consumption: Application of numerical and ann approaches. Journal of Soft Computing in Civil Engineering 2021;5. https://doi.org/10.22115/SCCE.2021.281977.1299.

[32] Stumpe B, Stuhrmann N, Marschner B. Variability of cooling distances and intensities of urban green spaces over six hot summer days affects their relevance for UHI mitigation in Berlin, Germany. Sustain Cities Soc 2026;137. https://doi.org/10.1016/j.scs.2026.107139.

[33] Dirksen M, Ronda RJ, Theeuwes NE, Pagani GA. Sky view factor calculations and its application in urban heat island studies. Urban Clim 2019;30. https://doi.org/10.1016/j.uclim.2019.100498.

[34] Almashhour R, Alzaatreh A. Identifying building structure factors for urban heat mitigation: a hybrid methodology using Fuzzy Delphi Method and Confirmatory Factor Analysis. International Journal of Sustainable Energy 2025;44. https://doi.org/10.1080/14786451.2025.2457377.

[35] Puche M, Vavassori A, Brovelli MA. Insights into the Effect of Urban Morphology and Land Cover on Land Surface and Air Temperatures in the Metropolitan City of Milan (Italy) Using Satellite Imagery and In Situ Measurements. Remote Sens (Basel) 2023;15. https://doi.org/10.3390/rs15030733.

[36] Soltanifard H, Amani-Beni M. The cooling effect of urban green spaces as nature-based solutions for mitigating urban heat: insights from a decade-long systematic review. Clim Risk Manag 2025;49. https://doi.org/10.1016/j.crm.2025.100731.

[37] Richards DR, Fung TK, Belcher RN, Edwards PJ. Differential air temperature cooling performance of urban vegetation types in the tropics. Urban For Urban Green 2020;50. https://doi.org/10.1016/j.ufug.2020.126651.

[38] Ziter CD, Pedersen EJ, Kucharik CJ, Turner MG. Scale-dependent interactions between tree canopy cover and impervious surfaces reduce daytime urban heat during summer. Proc Natl Acad Sci U S A 2019;116. https://doi.org/10.1073/pnas.1817561116.

[39] Su Y, Wu J, Zhang C, Wu X, Li Q, Liu L, et al. Estimating the cooling effect magnitude of urban vegetation in different climate zones using multi-source remote sensing. Urban Clim 2022;43. https://doi.org/10.1016/j.uclim.2022.101155.

ARTICLE IN PRESS

[40] Aram F, Higueras García E, Solgi E, Mansournia S. Urban green space cooling effect in cities. Heliyon 2019;5. https://doi.org/10.1016/j.heliyon.2019.e01339.

[41] Li Y, Svenning JC, Zhou W, Zhu K, Abrams JF, Lenton TM, et al. Green spaces provide substantial but unequal urban cooling globally. Nature Communications 2024;15. https://doi.org/10.1038/s41467-024-51355-0.

[42] Wang J, Xu C, Pauleit S, Kindler A, Banzhaf E. Spatial patterns of urban green infrastructure for equity: A novel exploration. J Clean Prod 2019;238. https://doi.org/10.1016/j.jclepro.2019.117858.

[43] Xia G, He G, Zhang X. Assessing the Spatial Equity of Urban Park Green Space Layout from the Perspective of Resident Heterogeneity. Sustainability (Switzerland) 2024;16. https://doi.org/10.3390/su16135631.

[44] Li L, Kross A, Ziter CD, Eicker U. Analyzing spatial patterns of urban green infrastructure for urban cooling and social equity. Urban For Urban Green 2025;112. https://doi.org/10.1016/j.ufug.2025.128983.

[45] Saber A, De Luca C, Pourzangbar A, Tondelli S, Bell ML. Heat wave risk assessment in Bologna using Spatio-temporal artificial intelligence: Leveraging LSTM enhanced by the hippopotamus optimization algorithm. Sustain Cities Soc 2025;131. https://doi.org/10.1016/j.scs.2025.106671.

[46] Belgiu M, Drăgu L. Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing 2016;114. https://doi.org/10.1016/j.isprsjprs.2016.01.011.

[47] Wang Q, Wang X, Zhou Y, Liu D, Wang H. The dominant factors and influence of urban characteristics on land surface temperature using random forest algorithm. Sustain Cities Soc 2022;79. https://doi.org/10.1016/j.scs.2022.103722.

[48] Tanoori G, Soltani A, Modiri A. Machine Learning for Urban Heat Island (UHI) Analysis: Predicting Land Surface Temperature (LST) in Urban Environments. Urban Clim 2024;55. https://doi.org/10.1016/j.uclim.2024.101962.

[49] Roberts DR, Bahn V, Ciuti S, Boyce MS, Elith J, Guillera-Arroita G, et al. Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. Ecography 2017;40. https://doi.org/10.1111/ecog.02881.

[50] Tasan M, Dąbrowska J, Michałowska K, Uciechowska-Grakowicz A. Long-term dynamics of urban heat island and hot spots in Wrocław: A 25-year satellite-based analysis using machine learning. Sustain Cities Soc 2025;132. https://doi.org/10.1016/j.scs.2025.106797.

ARTICLE IN PRESS

[51] Cheval S, Amihăesei VA, Chitu Z, Dumitrescu A, Falcescu V, Irașoc A, et al. A systematic review of urban heat island and heat waves research (1991–2022). Clim Risk Manag 2024;44. https://doi.org/10.1016/j.crm.2024.100603.

[52] Yu M, Xu F, Hu W, Sun J, Cervone G. Using Long Short-Term Memory (LSTM) and Internet of Things (IoT) for Localized Surface Temperature Forecasting in an Urban Environment. IEEE Access 2021;9. https://doi.org/10.1109/ACCESS.2021.3116809.

[53] Wang H, Zhang J, Yang J. Time series forecasting of pedestrian-level urban air temperature by LSTM: Guidance for practitioners. Urban Clim 2024;56. https://doi.org/10.1016/j.uclim.2024.102063.

[54] Sahani J, Kumar P, Debele SE. Assessing demographic and socioeconomic susceptibilities to heatwaves in the Southeastern United Kingdom. Sustain Cities Soc 2024;117. https://doi.org/10.1016/j.scs.2024.105958.

[55] Olsen JR, Niedzwiedz C, Nicholls N, Wheeler BW, Ho FK, Pell JP. Sociodemographic and geographic inequalities in exposure to projected hot and extreme summer days in England: A nationwide socio-spatial analysis. Environ Int 2025;197. https://doi.org/10.1016/j.envint.2025.109351.

[56] Johnson DP, Stanforth A, Lulla V, Luber G. Developing an applied extreme heat vulnerability index utilizing socioeconomic and environmental data. Applied Geography 2012;35. https://doi.org/10.1016/j.apgeog.2012.04.006.

[57] Ellena M, Breil M, Soriani S. The heat-health nexus in the urban context: A systematic literature review exploring the socio-economic vulnerabilities and built environment characteristics. Urban Clim 2020;34. https://doi.org/10.1016/j.uclim.2020.100676.

[58] Rathi SK, Chakraborty S, Mishra SK, Dutta A, Nanda L. A heat vulnerability index: Spatial patterns of exposure, sensitivity and adaptive capacity for urbanites of four cities of india. Int J Environ Res Public Health 2022;19. https://doi.org/10.3390/ijerph19010283.

[59] Qian Y, Liu T. Heat vulnerability assessment: A systematic review of critical metrics. Hygiene and Environmental Health Advances 2025;15. https://doi.org/10.1016/j.heha.2025.100138.

[60] Ellena M, Melis G, Zengarini N, Gangi E Di, Ricciardi G, Mercogliano P, et al. Micro-scale UHI risk assessment on the heat-health nexus within cities by looking at socio-economic factors and built environment characteristics: The Turin case study (Italy). Urban Clim 2023;49. https://doi.org/10.1016/j.uclim.2023.101514.

[61] Maccabiani E, Usmani M, Nanni R, Napolitano M. Investigating Social Vulnerability to Extreme Heat: Heat Islands and Climate Shelters in Urban Contexts: The Case of Bologna. ISPRS Int J Geoinf 2025;14. https://doi.org/10.3390/ijgi14010017.

ARTICLE IN PRESS

[62] Koutra S, Ioakimidis CS. Unveiling the Potential of Machine Learning Applications in Urban Planning Challenges. Land (Basel) 2023;12. https://doi.org/10.3390/land12010083.

[63] Chaturvedi V, de Vries WT. Machine Learning Algorithms for Urban Land Use Planning: A Review. Urban Science 2021;5. https://doi.org/10.3390/urbansci5030068.

[64] Comune di Bologna. Le tendenze demografiche a Bologna nel 2024. Bologna: I numeri di Bologna Metropolitana; 2025. Available from: https://inumeridibolognametropolitana.it/notizie/le-tendenze-demografiche-bologna-nel-2024. n.d.

[65] Comune di Bologna. Territorio: 100 grafici. Bologna: I numeri di Bologna Metropolitana; 2025. n.d.

[66] European Commission. EU Mission: Climate-Neutral and Smart Cities. Brussels: European Commission 2026.

[67] European Commission Joint Research Centre. Photovoltaic Geographical Information System (PVGIS). Brussels: European Commission 2026.

[68] Tabatabaei M, Antonini E. Machine Learning for Optimizing Urban Photovoltaics: A Review of Static and Dynamic Factors. Sustainability (Switzerland) 2025;17. https://doi.org/10.3390/su17188308.

[69] Google Earth Engine. USGS Landsat 8 Level 2, Collection 2, Tier 1. Mountain View: Google Developers; 2026 2026.

[70] United States Geological Survey. Landsat Collection 2 Surface Temperature. Reston: USGS 2024.

[71] Copernicus Data Space Ecosystem. Global and European Digital Elevation Model: Copernicus DEM GLO-30. Brussels: Copernicus 2026.

[72] Copernicus Land Monitoring Service. Urban Atlas Land Cover/Land Use 2018 (vector), Europe, 6-yearly. Urban Atlas 2018.

[73] European Space Agency. ESA WorldCover 10 m 2021 global land-cover product. Paris: ESA; 2021 2021.

[74] Comune di Bologna. Aree statistiche di Bologna. Bologna: Open Data Bologna. Available from: https://opendata.comune.bologna.it/explore/dataset/aree-statistiche/ 2026.

[75] Comune di Bologna. Comune di Bologna. Redditi per area statistica. Bologna: Open Data Bologna; 2026. Available from: https://opendata.comune.bologna.it/explore/dataset/redditi-per-area-statistica/. n.d.

ARTICLE IN PRESS

ARTICLE IN PRESS

[76] Comune di Bologna. Popolazione per quartiere, zona, sesso, eta, cittadinanza e luogo di nascita: serie storica. Bologna: Open Data Bologna; 2026. Available from: https://opendata.comune.bologna.it/explore/dataset/popolazione-per-quartiere-zona-sesso-eta-cittadinanza-provincia-di-nascita-serie/. n.d.

[77] Comune di Bologna. Occupazione e istruzione a Bologna. Bologna: Open Data Bologna; 2026. Available from: https://opendata.comune.bologna.it/explore/dataset/occupazione_istruzione/. n.d.

[78] Comune di Bologna. Elenco parchi e giardini. Bologna: Open Data Bologna; 2026. Available from: https://opendata.comune.bologna.it/explore/dataset/carta-tecnica-comunale-toponimi-parchi-e-giardini/ n.d.

[79] United Nations. Transforming our world: the 2030 Agenda for Sustainable Development. Resolution adopted by the General Assembly on 25 September 2015, A/RES/70/1. New York: United Nations; 2015.

[80] Parker C, Mahlasi C, Govindasamy T, et al. Quantifying intra-urban socio-economic and environmental vulnerability to extreme heat events in Johannesburg, South Africa. International Journal of Biometeorology. 2025;69:2501–2517. https://doi.org/10.1007/s00484-025-02971-y.

[81] Vartholomaios A. Classification of the influence of urban canyon geometry and reflectance on seasonal solar irradiation in three European cities. Sustainable Cities and Society. 2021;75:103379. https://doi.org/10.1016/j.scs.2021.103379.

[82] Syeda AQ, Castillo-Villar KK, Alaeddini A. Sustainable Urban Heat Island Mitigation Through Machine Learning Integrating Physical and Social Determinants for Evidence-Based Urban Policy. Sustainability. 2025;17(15):7040. https://doi.org/10.3390/su17157040.

ARTICLE IN PRESS