POLITECNICO DI TORINO
Repository ISTITUZIONALE

Morphology-based classification and error-controlled reduction of urban weather files with Torino as a case study

Original
Morphology-based classification and error-controlled reduction of urban weather files with Torino as a case study / Jahanirahaei, A., Chiesa, G.. - In: BUILDING AND ENVIRONMENT. - ISSN 0360-1323. - ELETTRONICO. - 308:(2027). [10.1016/j.buildenv.2026.115528]

Availability:
This version is available at: 11583/3017696 since: 2026-11-09T10:46:35Z

Publisher:
Elsevier

Published
DOI:10.1016/j.buildenv.2026.115528

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)

21 November 2026

Building and Environment 308 (2027) 115528

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

Morphology-based classification and error-controlled reduction of urban weather files with Torino as a case study

Ali JahaniRahaei, Giacomo Chiesa *

Politecnico di Torino, Department of Architecture and Design, Viale Pier Andrea Mattioli 39, 10125 Torino, Italy

ARTICLE INFO

Keywords:
Urban weather generator
Urban morphology
k-medoids
Model reduction
Weather-file atlas
Decision support

ABSTRACT

Urban morphology creates intra-urban thermal variability that a single airport EnergyPlus Weather (EPW) file cannot represent. This study develops a reproducible GIS-to-Urban Weather Generator (UWG) framework for Torino, Italy, separating morphological classification from weather-product reduction. A preliminary building-seeded cell method was audited before adopting 100 m assignment cells characterised through overlapping 500 × 500 m neighbourhood supports. Support-size (250–1000 m) and 50 m grid-shift sensitivities were tested. Building fraction, area-weighted height, facade-to-site ratio, tree cover, and grass cover were extracted for 12,997 locations; 10,593 quality-screened supports were classified using robust scaling and CLARA-style k-medoids. Seven morphology classes balanced interpretability and spatial coherence, although four ranked best statistically, and classification robustness was moderate. Under a common UWG template, four representative EPWs were retained by curve inspection and minimax selection, with a 0.10°C maximum-error screen and a diminishing-return criterion. Maximum medoid hourly dry-bulb RMSE was 0.06°C with the delivered one-decimal EPWs and 0.04°C at diagnostic three-decimal serialisation. Across 105 stratified non-medoid supports, weighted mean and 95th-percentile RMSE were 0.05 and 0.10°C, but a 1.33°C industrial-tail error required a conservative post hoc restriction, not a validated applicability boundary. Synchronised processed 2022 station comparisons improved annual RMSE relative to Caselle only at Consolata. Humidity limitations preclude interpreting temperature fidelity as complete EPW validation. The atlas links coordinates to morphology classes, UWG configurations, and candidate EPWs. The restricted library distinguishes simulator-conditional reduction fidelity from observational accuracy.

1. Introduction

Hourly meteorological data serve as boundary conditions for urban climate assessment and built-environment analyses. Common simulation tools, including EnergyPlus and TRNSYS, ingest standardised weather files, and the EPW format is widely used to exchange hourly conditions. In practice, a city is often represented by a single typical meteorological file from an airport or another open site. This convention is convenient and supports consistent comparison, but the measurement context can differ substantially from the urban canopy in which most buildings and public spaces are located. Urban surfaces alter radiative exchange, heat storage, evapotranspiration, aerodynamic roughness, airflow, and anthropogenic heat release. Consequently, urban air temperature and humidity may differ from the nominal reference location, and these differences vary with urban form, land cover, season, and time of day [1,2,3].

The urban heat island (UHI) is therefore not one spatially uniform correction. The Local Climate Zone (LCZ) framework formalised the expectation that differences in surface structure, cover, fabric, and activity are relevant to urban temperature studies [3]. Observational studies have also demonstrated within-class and between-class variability, indicating that a categorical label does not remove the importance of local context [4,5]. At the same time, using a unique weather file for each building or block would create an unwieldy product and could imply a degree of spatial precision unsupported by either the forcing data or the urban canopy model.

This creates a representation problem between two limiting alternatives: either a single airport file for the entire city or thousands of independently generated files. The practical scientific question is not only whether urban weather differs from a reference station, but how much of the modelled spatial diversity must be retained in a compact and interpretable weather-file library.

* Corresponding author.
E-mail address: giacomo.chiesa@polito.it (G. Chiesa).

https://doi.org/10.1016/j.buildenv.2026.115528
Received 27 July 2026; Received in revised form 30 October 2026; Accepted 1 November 2026
Available online 2 November 2026
0360-1323/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

The Urban Weather Generator (UWG) was developed as a computationally efficient method for transforming rural or reference-station weather into neighbourhood-scale urban canopy conditions [6–8]. It couples an urban canopy energy balance with building and boundary-layer representations and can produce an EPW file suitable for downstream use. The method is substantially less computationally demanding than street-resolving computational fluid dynamics or three-dimensional microclimate simulation, but this efficiency relies on spatial aggregation. UWG represents characteristic neighbourhood conditions rather than the detailed flow around an individual facade.

Several studies have already linked urban morphology to the generation of representative weather. Palme et al. [9,10] used GIS-derived built area, green area, facade ratio, principal component analysis, and urban tissue categories to generate UWG weather files for four South American cities. Their released dataset included 17 urban files, with four urban categories for Lima, Antofagasta, and Guayaquil, and five for Valparaiso. Shen et al. [11] combined automated map capture, principal component analysis, Gaussian mixture clustering, Bayesian information criterion, calibration, and UWG for Shenzhen. Ma et al. [12] used street-network morpho-patches and localised weather in a multiscale workflow, while Hashemi et al. [13] coupled LCZ descriptions with UWG across seven cities in the United States. More recently, Liu et al. [14] demonstrated a scalable, citywide workflow that produced block-specific UWG files for 1,693 urban blocks.

These studies establish that morphology classification, representative urban forms, and category-specific UWG files are not new in themselves. A parallel body of work has organised urban form into archetypes for environmental analysis. Ratti et al. [15] illustrated archetype-based analysis of building form and environmental performance in an arid climate, while Chiesa and Li [16] linked morphology-based UWG weather generation to early-stage bioclimatic design, including applications in Torino. Joshi et al. [17] subsequently clustered 1007 blocks using 17 morphology variables, principal components, multiple cluster-number diagnostics, LCZ comparison, and land-surface-temperature evidence; their later work applied the archetypes to physical microclimate analysis [18]. Collectively, these studies indicate that internal cluster compactness is insufficient on its own: a classification should be physically interpretable, stable enough for its intended use, and tested against an external response. A medoid is an observed class member that can be mapped and passed directly to UWG; its centrality in morphology does not establish weather representativeness at class boundaries or extremes.

Within morphology-based classification, the clustering method is another source of variation. In this study, silhouette width and compactness indices were treated as complementary diagnostics rather than physical definitions of urban types [19]. Stability was evaluated using a spatial-block holdout, as nearby 500 m supports overlap, consistent with the broader principle that reproducible cluster structure should persist under perturbation [20]. K-medoids was selected for the Torino classifications because each representative is an observed support that can be mapped, inspected, and passed directly to UWG; the scalable CLARA-style (Clustering Large Applications) implementation follows the medoid framework reviewed by Schubert and Rousseeuw [21].

Spatial representation introduces a related choice concerning the area used to characterise an urban neighbourhood. Building blocks are intuitive spatial units because they follow streets and urban design logic, but they vary strongly in area and topology. A small pavilion in a park, a large industrial parcel, and a dense historic block can each become one polygon, even though their suitability as an urban climate support is very different. Conversely, a regular grid provides consistent spatial sampling but can cut through buildings or streets and can be sensitive to its origin. Urban meteorological metadata protocols distinguish local context from the instrument point itself, while morphology and sky-view studies show that urban descriptors have nontrivial and metric-specific spatial footprints [22,23]. Urban-climate model input reviews and multiscale morphology frameworks further show that the meaning of a descriptor depends on its spatial support [24,25].

The spatial unit must therefore be treated as a model choice and tested. It is useful to distinguish an assignment geometry, which tells a user or map where a weather category applies, from a climate-support geometry, which summarises the surrounding urban fabric used by UWG. This distinction also addresses the practical concern that a small plaza or isolated structure should not automatically create a unique neighbourhood climate category. Here, the 100 m cell locates the assignment, and the larger square supplies neighbourhood inputs; 100 m spacing is not 100 m meteorological resolution.

UWG outputs depend on both morphology and shared model assumptions. Sensitivity and calibration studies show that boundary-layer, reference-site, building-stock, vegetation, and anthropogenic parameters can materially influence results [26,27,28,29]. Holding non-morphological assumptions common across neighbourhoods isolates morphology in a controlled comparison; sensitivity tests are then needed to assess how uncertainty in shared inputs affects the resulting weather library.

Torino is a suitable test case because detailed municipal GIS information can be combined with a ten-year, seven-station hourly weather dataset covering 2014-2023 and simulation-ready representative EPWs [30]. Previous work has already examined reference-site selection and urban-rural classification in the metropolitan area [31], characterised Torino's urban heat and dry islands [32], and investigated the long-term UHI [33]. Borelli et al. [34] also used Torino blocks, morphology metrics, representative sampling, UWG, and degree-day correlations to map urban heat effects.

A remaining gap is the independent selection of weather-product resolution. Most studies either inherit the number of weather files from a predefined typology, select the number of morphology clusters before weather simulation, or generate a file for each spatial unit. Morphology categories and weather-product categories are thereby treated as if they must have the same cardinality. However, two morphologically distinct neighbourhoods may produce nearly identical hourly output under a fixed urban canopy model, while a small but unusual morphology may produce a distinct response. The minimum useful number of weather files should therefore be evaluated in weather-output space rather than assumed from input clustering.

Accordingly, Torino is used to test a weather-representation framework. The framework first derives morphology classes from consistent neighbourhood supports, then independently asks whether those classes require distinct annual EPWs. Measured stations are used for two separate evaluations: whether a reduced library reproduces an exact station-support UWG run, and whether either modelled file agrees with observed urban weather. Keeping these evaluations separate avoids treating model compression as observational validation.

This study aims to develop and critically evaluate a reproducible method for reducing citywide urban morphology into a compact library of UWG-generated EPWs for Torino, without assuming that the number of morphology classes equals the number of required weather files.

The study addresses four essential research questions:

1. How should urban morphology be spatially represented so that citywide inputs are reproducible, stable, and appropriate to UWG neighbourhood assumptions?
2. How many physically interpretable morphology classes are supported by the Torino GIS data when statistical quality, spatial stability, and real medoids are considered together?
3. How can a compact annual EPW library represent medoid and non-medoid UWG temperature responses under explicit error and diminishing-return criteria?
4. At measured urban stations, how well do the assigned library EPWs reproduce exact station-support UWG outputs, and how does this reduction fidelity differ from observational agreement?

The principal contribution is a two-stage representation method. Morphology is classified on standardised neighbourhood supports, after which weather outputs are reduced independently by exhaustive subset selection. This produces seven physically interpretable morphology classes but only four candidate EPWs, demonstrating that input and product resolutions need not have the same cardinality. The reproducibility package delivers the results through a public atlas: it includes the k=4, k=7, and k=8 spatial classifications and A-D assignment as reusable geospatial data, exposes the representative EPWs and exact UWG JSON configurations, and provides a coordinate lookup for architects, urban planners, and other built-environment users. The study also provides a transparent record of spatial-unit testing, input and output quality assessment, cluster sensitivity, model-parameter screening, and the limitations revealed by four urban stations. The weather-file count is conditional on the simulator, common configuration and temperature criterion; the library has an industrial applicability restriction enforced by the atlas lookup.

2. Materials and methods

2.1. Study design, area, and data

The study followed seven documented stages (Fig. 1):

1. assemble and harmonise municipal GIS and station weather data;
2. audit a preliminary irregular building-seeded context-cell method and diagnose its spatial limitations;
3. define and test a regular assignment grid with standardised morphology supports;
4. extract and validate citywide UWG-relevant inputs;
5. compare morphology cluster solutions and select real medoids;
6. run UWG under the final common template, reduce the medoid EPWs using hourly output error, and test non-medoid coverage; and
7. assign measured stations to classes and compare Caselle, exact-support UWG, measured urban EPWs, and library EPWs.

The two spatial representations were evaluated sequentially and were not combined in the final analysis. Method 1 was a preliminary diagnostic experiment used to test block-based context cells and expose their city-scale limitations. Method 2 supplied every morphology input used for clustering, UWG simulation, EPW reduction, mapping, and coordinate lookup.

The case study is the municipality of Torino in northwestern Italy. The adopted municipal boundary has an area of 130.004 km^2 and was represented in Monte Mario / Italy zone 1 (EPSG:3003) for metric geometry processing. The city contains dense historic fabric, nineteenth- and twentieth-century perimeter blocks, open residential districts, industrial and infrastructural areas, major parks, and heterogeneous peripheral development. This morphological range is useful for testing both common fabrics and small but physically distinctive classes.

The municipal GIS inputs, separate from the weather companion dataset, were:

- 127,392 building-volume polygons after boundary clipping, including building height attributes;
- 76,125 road-surface polygons, including vehicular, pedestrian, cycling, and mixed-secondary circulation;
- 79,715 mapped green-area polygons; and
- 116,858 tree records, of which 102,006 records classified as living trees were retained.

The weather dataset comprises hourly observations for seven stations from 2014 to 2023 and representative Average Meteorological Year (AvMY) EPWs [30]. The four urban stations used in the present evaluation were Consolata, Giardini Reali, Alenia, and Reiss Romoli. Caselle AvMY was used as the common UWG reference forcing because the related reference-station analysis identified it as the best available forcing for the tested Torino UWG configuration. Venaria La Mandria and Bauducchi remain part of the broader dataset but were not varied because sensitivity to non-urban reference stations was outside the scope of this study. The weather dataset and this study share authorship; reuse of the processed ARPA Piemonte records is not independent sensor-data preparation.

The supplied EPWs are representative-year profiles rather than

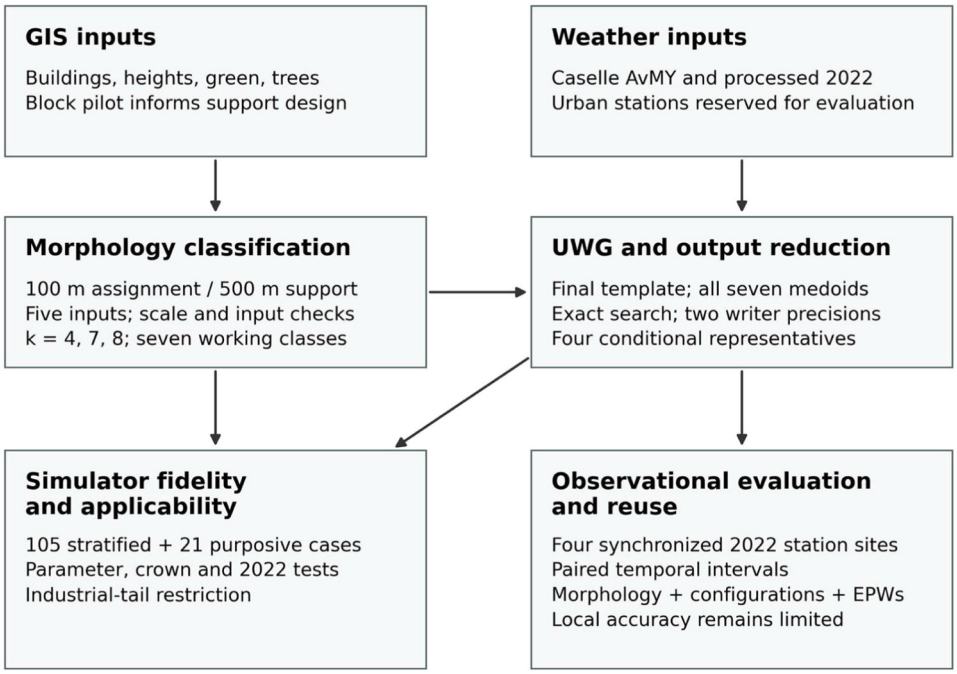


Fig. 1. Study workflow from GIS and meteorological inputs through spatial representation, morphology classification, UWG simulation, EPW reduction, and station-based evaluation. Morphology-class resolution and weather-product resolution are evaluated separately.

event-synchronous station records. Consequently, RMSE between two AvMY files describes differences between their representative hourly calendars and should not be interpreted as event-by-event forecast error. For event-aligned evaluation, processed 2022 annual EPWs from Caselle and all four urban stations were therefore used. The hottest complete year was selected after auditing 70 station-year files, without using station errors to fit clusters or to choose initial representatives. For 16 April–14 October, the four-urban-station mean in 2022 was 22.44°C, ranking first among 2014–2023 and exceeding the ten-year mean by 1.43°C.

2.2. Spatial discretisation and morphology extraction

Method 1 was a preliminary building-seeded context-cell experiment. It tested whether one climate unit could comprise a building block, its internal open space, and approximately half of each surrounding street. Because the available streets were road-surface polygons and intersections were separate geometries, circulation was dissolved, building-containing non-road cores were used as seeds, and minor isolated structures were suppressed before nearest-boundary Voronoi allocation of roads, intersections, plazas, green areas, and residual open fragments. Buildings, official green polygons, and living trees were then overlaid. Individual footprints were retained for area-weighted height, while touching footprints were dissolved into compounds, excluding party walls from facade-to-site calculations. Pilot and citywide QA combined geometry checks, robust outlier rules, and labelled manual-review maps. This method was not used to fit the reported clusters or generate the EPW library, although its limitations are given in the Discussion section.

The method was first tested in two central/western subsets and then applied to the full municipal boundary. A quality-assurance (QA) funnel combined hard geometry checks, robust outlier rules, and manual-review prioritisation. Indicators included cell area, built fraction, road fraction, green fraction, building count, seed-to-final area ratio, compactness, multipart geometry, height coverage, facade-to-site ratio, and edge clipping. Cells with high QA scores were labelled by cell identifier on review maps so that numerical records could be traced to geometry.

The method produced meaningful block-like cells in conventional street fabrics. Still, the full-city audit revealed that non-closed road surfaces and forced allocation of remote open land could create very large or multipart contexts. Building assignment by representative point also allowed some complete footprints to cross a cell boundary. These findings were treated as evidence about spatial support design rather than corrected through undocumented manual edits.

Method 2 was the adopted spatial representation and assigned distinct functions to two geometries. A deterministic 100 × 100 m grid identifies the map location and the recommended EPW category at that location, while an overlapping square centred on the grid location summarises neighbourhood morphology for UWG. This separation prevents an unusually small block, plaza, large park allocation, or industrial parcel from determining the spatial scale of the climate input.

The assignment grid was generated in the same projected coordinate system. A location was retained when the centre of its 100 m cell was covered by the exact municipal boundary. For the selected configuration, each centre received a 500 × 500 m support with a fixed area of 250,000 m^2. Fractions used this fixed denominator. Because no compatible external GIS halo was added, the supports intersecting the municipal edge retained the full denominator and were flagged as incomplete rather than normalised to a smaller clipped support.

Square supports with side lengths of 250, 500, and 1000 m were centred on each of the four urban stations. For each support, building fraction, area-weighted height, facade-to-site ratio, green cover, tree canopy, UWG tree and grass cover, and road fraction were recalculated directly from the source GIS. The support side was also used as the UWG *characteristic length (JSON key charlength*).

Each of the 12 station supports was simulated for January and July using AvMY Caselle and a preliminary Zone 1A template with archetype-specific floor heights, producing 24 successful UWG cases. This scale screen varied morphology and characteristic length together; it was separate from the annual Zone 4A experiments. Temperature was compared with the corresponding measured station, AvMY, for all hours, 07:00-18:00 daytime and 19:00-06:00 nighttime. Alenia was excluded from the scale selection because only 92.6%, 75.6%, and 65.8% of its 250 m, 500 m, and 1000 m supports, respectively, were covered by the municipal GIS.

The 500 m support was selected as a methodological compromise, not as a universal physical optimum. It contains multiple blocks, dampens the influence of isolated local exceptions, produces essentially the same aggregate station performance as 250 and 1000 m, and retains more local differentiation than 1000 m.

A matched-morphology test subsequently extracted all three support sizes across 385 centres. Of 336 probability draws across 28 strata in the 9,164-centre full-1000 m coverage frame, 304 passed quality checks at every scale. The remaining 32 lacked buildings or adequate height at 250 m. Weighted height and facade/site variances use those 304 centres; targeted examples are not entered into the estimator. This quantifies smoothing, not a measured atmospheric footprint or an independent optimisation of class separation (Supplementary Section S2).

To test whether the surface morphology depended strongly on the arbitrary grid origin, a 5 × 5 km central-Torino pilot containing 2,500 baseline locations was regenerated after shifting the entire 100 m grid by 50 m east and 50 m north. The shifted 500 m-support morphology surfaces were interpolated back to 2,401 common baseline coordinates, allowing comparison at the same physical locations rather than at equal row and column indices displaced by 70.7 m.

For building density, height, facade-to-site ratio, green cover, canopy cover, and road fraction, the prespecified acceptance criteria were Spearman rho ≥ 0.95, normalised RMSE ≤ 0.20 baseline standard deviations, and at least 95% agreement within the same or an adjacent baseline quintile.

The adopted grid was applied to the full Torino boundary. Primary variables were computed directly from source geometries rather than averaged from irregular cells – see equations from 1 to 3:

$$f_b = \frac{A_b}{A_s} \quad \text{(eq. 1)}$$

$$H = \frac{\sum_{i=1}^{n} A_i H_i}{\sum_{i=1}^{n} A_i} \quad \text{(eq. 2)}$$

$$F \Big/ S = \frac{P_{exposed} H}{A_s} \quad \text{(eq. 3)}$$

where A_s=250,000 m^2 is the fixed support area, f_b is building fraction, H is the valid-height footprint-area-weighted mean building height, and $P_{exposed}$ is the exterior perimeter of the global dissolved footprint union falling within the support, excluding shared walls and courtyard rings. Building, green, and road geometries were unioned within each support before their intersected areas were divided by A_s, preventing overlaps within a source layer from being counted more than once. A_b denotes the unioned building-footprint area; A_i and h_i are each the intersected valid-height footprint area and its height, and n is the number of valid-height footprints intersecting the support.

The tree canopy was estimated by buffering living-tree points using the recorded crown diameter. Missing crown diameters were filled sequentially from species medians where sample support was sufficient, from a height-based estimate, and finally from the citywide median. Overlapping crowns were merged before area calculation. UWG tree and grass cover were constrained by available non-building ground so that building and vegetation fractions did not exceed one. Tree height and density were retained as descriptive variables but were not direct inputs

to the clustering. Grass is the nonnegative mapped green remainder after allocating tree cover, capped by the remaining non-building area; it is not an independently observed grass layer. Individual tree heights and explicit ray-traced shadows are not passed to this UWG configuration.

Every row retained support-boundary overlap, edge status, valid-height footprint coverage, missing-input indicators, QA score, and QA reasons. The separate in-house validator reconstructed grid identifiers, geometry, formulas, fractions, QA logic, CRS, table joins, and file hashes. Four analysis populations were retained: all mapped locations; complete supports; complete supports with usable UWG geometry; and strict *zero-QA-score* supports.

Crown uncertainty was evaluated by five-fold spatially blocked masking, followed by 23 fixed-classifier scenarios: the central hierarchy, lower/upper conditional quartiles and 20 stochastic imputations. Fourteen annual medoid runs test the quartile canopy states. A separate extraction correction retains polygon components in intersection GeometryCollections and is assessed using frozen medoids, rather than silently substituted by a complete clustering refit (Supplementary Sections S1–S2).

2.3. Morphological classification

Five primary variables were used to cluster the strictly fitting population:

1. building fraction;
2. area-weighted average building height;
3. facade-to-site ratio;
4. estimated UWG tree-cover fraction; and
5. residual UWG grass-cover fraction.

These variables were selected because they map directly to the neighbourhood differences passed to UWG and represent plan density, vertical scale, exposed facade, canopy, and ground vegetation. Road fraction, total green percentage, tree count, tree density, and related metrics were retained for interpretation but excluded from the primary distance calculation. Adding them would give the same physical family repeated weight or introduce a variable without an independent mapping to the active UWG configuration.

The input audit included missingness, range checks, skewness, pairwise Spearman correlation, robust-scale profiles, principal components, and alternate feature treatments. Correlation was expected rather than interpreted as a data error: height and facade-to-site ratio had a Spearman's rho of 0.795; building fraction and facade-to-site ratio, 0.792; building fraction and grass cover, -0.697; and building fraction and height, 0.673. The primary solution retained the five direct inputs and applied median/IQR-robust scaling. A group-balanced sensitivity analysis reduced the repeated influence from correlated morphology and vegetation families. No single cluster count was assumed. K-means was used as a fast diagnostic for every integer k from 2 to 12. The lower bound tests whether the city is explained by a coarse open-versus-dense division. The upper bound is sufficiently large to expose secondary fabrics while remaining compatible with interpretable mapping, minimum class-size review, medoid simulation, and a compact public product. Values above 12 were not considered because the observed trend already showed a declining neighbour agreement, increasing fragmentation, and classes below approximately 3% of the fitting population for k=9-12. For each k, the following were calculated: inertia, silhouette width, Davies-Bouldin index, Calinski-Harabasz index, elbow curvature, minimum and maximum class size, neighbour-label agreement, connected-component statistics, and spatial stability. Stability used 20 repeats, with 20% of 1 km spatial blocks held out. Spatial blocks were used instead of random rows because the nearby 500 m supports overlap and are not independent; this initial screen did not eliminate overlap across the holdout boundary. The statistical ranking was restricted to k=4-8 after the full scan. The k=2 and k=3 solutions were stable coarse partitions but underrepresented the diversity of physically interpretable fabric. At the other end, k=9-12 produced progressively smaller classes, lower neighbour agreement, more components, and limited additional inertia reduction. The k=4-8 range, therefore, represented the useful balance between coarse and fragmented solutions rather than a claim that values outside the range were mathematically invalid. The composite rank uses weights of 0.25 for silhouette, 0.20 for Davies–Bouldin, 0.25 for spatial ARI, and 0.10 each for Calinski–Harabasz, neighbour agreement, and elbow curvature, with a two-point penalty for classes below the 1% minimum class size.

The three shortlisted counts, k=4, 7, and 8, were refitted with a CLARA-style scalable medoid procedure using Euclidean distance in robust-scaled feature space. Six trials were run for each candidate, with 2,000 fitting supports sampled per trial. After each trial, all 10,593 fitting supports were assigned to the sample medoids, and the full-population objective was evaluated. The best full-population solution was retained. Labels were reordered by increasing the median building fraction for consistent interpretation. The primary labels were compared with K-means on the same data, Manhattan-distance CLARA, Yeo-Johnson-transformed K-means, group-balanced K-means, and Ward hierarchical clustering, all on a fixed 2,500-support sample. Adjusted Rand index (ARI) was used to quantify the agreement. A conservative isolated-component reassignment was generated as a QA proposal but not applied to the scientific baseline, because a small spatial island can represent a genuine park, campus, civic site, or industrial fabric. Within each sample, ten-start K-means initialises real representatives, followed by within-class minimum-distance medoid updates for up to 40 iterations; this is not a globally optimal PAM. Additional fixed-algorithm feature tests and a 12-setting HDBSCAN sweep assess representation and noise. Ten paired spatial repeats withhold 20% of 2 km blocks, with training-only scaling and a 500 m Chebyshev buffer preventing positive support overlap across the split. Their four 1,500-support trials are compared with an equally budgeted reference (Supplementary Section S2).

2.4. UWG simulations and EPW reduction

UWG version 5.3.4 was run using AvMY Caselle as the common source. Five values varied by support and characteristic length was fixed at 500 m (Table 1):

All urban supports shared the same non-morphology assumptions. The complete 50-key JSON template was parsed and round-tripped through UWG to verify syntactic consistency. A preliminary one-at-a-time sensitivity screen used a Zone 1A template with archetype-specific floor heights and one 500 m Consolata support, 14-day January and July runs with seven settling days, and 82 successful cases. The screen was used to identify risky assumptions, not to calibrate the model to one station. Boundary/reference parameters such as observation height, nighttime urban boundary-layer height, reference height, minimum wind, and exchange coefficient produced the largest short-period effects. They were retained as common citywide assumptions because varying them by morphology class would confound atmospheric setup with neighbourhood form. The annual experiments used DOE zone 4A and an explicit citywide floor height of 3.05 m. The common template specified a canyon HVAC heat-release fraction of 1, a

Table 1
The six considered UWG parameters and the correlated GIS supporting sources.

UWG parameter	GIS/support source
bldheight	area-weighted average building height
blddensity	building fraction
vertohor	facade-to-site ratio
treecover	constrained tree-canopy fraction
grasscover	constrained residual green fraction
charlength	fixed 500 m support side

non-building sensible heat amplitude of 20 W/m^2, daytime and nighttime urban boundary-layer heights of 1000 and 80 m, and a reference height of 150 m; rural reference inputs, material properties, schedules, and building-stock fractions were also shared. The zone label is a reference-building-library choice rather than a direct equivalence to the Italian climate zone. Floor height was not varied by neighbourhood because no consistent citywide storey dataset was available. The common stock comprises 40% large office and 60% midrise apartment buildings, with post-1980 reference constructions, not district-specific buildings surveyed. All seven medoids were simulated with this common template to calculate the reduction curve.

The AvMY Caselle pressure field was supplied in hPa. A derivative forcing file was created in which only EPW column 10 (zero-based index 9) was multiplied by 100; the source file was left unchanged, and hashes for both files were recorded. Annual simulations began on 1 January and ran for 365 days with 300 s simulation steps and 3600 s weather intervals. Output checks required eight headers, 8,760 hourly rows and 35 weather fields; a duplicated 36th field was removed only after verifying equality with the last field. Humidity and dew-point violations were recorded separately and restricted the 2022 analysis to dry-bulb temperature. Annual UWG simulations were completed for all k=4 and k=7 medoids under the final common template. The k=7 set was then treated as the full medoid-level candidate library. For each nonempty subset of the seven EPWs, each original morphology class was assigned to its nearest retained EPW according to annual hourly dry-bulb RMSE, where $N = 8{,}760$ and t indexes matched hours:

$$RMSE_{ij} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}\left(T_{i,t} - T_{j,t}\right)^2} \qquad \text{(eq. 4)}$$

All 127 nonempty subsets were evaluated. For each library size, the selected subset minimised the maximum represented-class RMSE and then the cluster-share- weighted mean RMSE. This minimax-first rule prevents a small but distinct class from being sacrificed merely because it has a low share of fitting support. Weights are fitting-support shares, not population or city area; the secondary criterion improves typical representation among worst-error ties. The subset-search runtime was measured over 50 repetitions, excluding file loading and simulation.

Four files were retained as a compact working count after inspection of the complete reduction curve, not as a uniquely required minimum or universal climate tolerance. The 0.1°C serialisation is implementation-specific: in UWG v5.3.4, the EPW writer used here sets epw_precision=1 for dry-bulb temperature [35]; one decimal place is not an EPW-format requirement. Matched three-decimal writer diagnostics distinguish serialisation from model response, without implying greater physical accuracy. The final count was subsequently evaluated on non-medoid supports using prespecified screens: weighted 95th-percentile annual RMSE ≤0.20°C, purposive maximum ≤0.50°C, and added 95th-percentile error over the seven-medoid baseline ≤0.05°C.

The medoid-level checks use the minimum attainable maximum hourly RMSE at each retained count: four-file error ≤0.10°C; the three-to-four reduction at least twice the four-to-five reduction; and a maximum-error penalty ≤0.05°C for the baseline representative identities rerun under a scenario, relative to that scenario's optimal four-file subset. The factor of two is an engineering diminishing-return heuristic: adding a fifth file should yield at most half the preceding gain. It is not a statistical confidence bound, a physical threshold, or a global optimisation of file count. These checks assess the curve-based working choice; the temperature tolerance alone can admit fewer files. Here, error-controlled denotes explicitly evaluated representation errors, not observed-climate accuracy (Supplementary Sections S4 and S7.2).

Non-medoid evaluation sampled 15 supports per class, with three in each of five medoid-distance bands (0–25, 25–50, 50–75, 75–95 and 95–100 percentiles): 105 cases in total. Spatial selection favoured different 1 km blocks and at least 500 m separation, with recorded relaxations. Class/distance post-stratification weights are therefore not exact simple-random inclusion weights. Another 21 purposive cases test boundary assignments, morphology tails and proposed cleanup changes, with zero weight in stratified estimates. All are compared with frozen class representatives and category assignments (Supplementary Section S4).

Seven shared-parameter alternatives were simulated for each medoid, yielding 49 annual runs: nighttime boundary-layer height of 40/120 m, exchange coefficient of 0.5/2.0, minimum wind speed of 2 m/s, road heat capacity of 2.2 MJ/(m^3 K), and older residential stock. Anthropogenic heat, mixing fraction and rural/reference terms remain fixed. Reoptimised and fixed-subset tests, engineering screens, crown experiments and the industrial-tail follow-up are specified in Supplementary Sections S3–S4; they do not reconstruct real citywide parameter heterogeneity.

An additional experiment used 84 annual integrations to test non-building heat magnitude, diurnal timing, HVAC waste-heat allocation and spatial heterogeneity. All seven medoids were evaluated with non-building sensible heat amplitudes of 10 or 40 W/m^2, flat or four-hour-delayed traffic profiles preserving annual prescribed heat, canyon HVAC heat-release fractions of 0 or 0.5, and two opposing road-fraction-ranked heat distributions bounded by 10 and 30 W/m^2. The baseline non-building sensible heat input of 20 W/m^2 is a schedule multiplier amplitude, corresponding to 10.01 W/m^2 annually; the canyon fraction partitions HVAC heat between canyon and roof release. The spatial scenarios preserve the fitting-support mean amplitude but do not estimate measured traffic. Both were also run for two GIS-selected road-share extremes per class, without probability weighting. All other inputs remained fixed. Complete subset searches distinguished unchanged baseline files from scenario-specific outputs. These AvMY tests used annual, December-January-February (DJF) / June-July-August (JJA) and the same clock day/night bins (Supplementary Section S7); Table 9 uses October-March and April-September half-years.

2.5. Station-based evaluation and reproducibility

The station layer was applied after the morphology classes and four-file reduction had been defined; measured temperatures did not determine the clusters or select the EPW representatives. The coordinates of Consolata, Giardini Reali, Alenia, and Reiss Romoli were overlaid on the 100 m assignment grid. Each coordinate matched one grid cell and its corresponding 500 m support, raw k=7 class, and reduced EPW category. Four additional annual UWG simulations used squares centred on the stations, rather than on the containing grid cells, and the accepted common template. The 2022 test comprises these four exact supports and all seven medoids under synchronised Caselle forcing.

Four synchronised 2022 series were compared for each station:

- processed Caselle 2022 weather;
- processed urban-station 2022 weather;
- the Caselle-forced station-specific 500 m-support UWG EPW; and
- the assigned representative simulated with the same 2022 forcing.

Annual hourly bias and RMSE were calculated after alignment by year, month, day, and EPW hour in fixed UTC+1. Annual RMSE and its improvement over Caselle were visualised. Bias is the mean (model minus reference), so positive station bias denotes a warmer model. RMSE is the square root of the mean squared hourly difference over the selected period. UHI was defined relative to Caselle as

$$\Delta T_{UHI} = T_{urban} - T_{Caselle} \qquad \text{(eq. 5)}$$

For 2022, annual, December–February (winter), June–August (summer), 07:00–18:00 (daytime), and 19:00–06:00 (nighttime) subsets were calculated using EPW hour labels; these are clock bins, not solar masks. AvMY source-relative results retain October–March and April–September seasons. Evaluation was explicitly nested: comparison

of station-support UWG with the assigned library EPW measured reduction fidelity, whereas comparison of either modelled EPW with the processed urban series assessed observational agreement. Paired circular moving-block bootstrap intervals use 5,000 seven-day resamples, with 1-, 3- and 14-day sensitivity checks. Uniform source flags do not distinguish measured from imputed hours; a separate flatline screen tests suspicious records (Supplementary Section S5).

Every major workflow stage used a separate generation/analysis script and a separate in-house verification script. Manifests recorded source paths, software versions, parameters, row counts, run time, output hashes, configuration hashes, and source-to-output lineage. Validation reconstructed GIS formulas, clustering assignments and metrics, EPW dimensions and physical constraints, temporal subsets, station matches, pairwise matrix identities, and selected KPIs. Warnings were retained rather than converted into silent exclusions. These scripts were developed within the same project, not by an independent external team.

3. Results

3.1. Spatial representation and citywide morphology dataset

The preliminary context-cell experiment produced intuitive block-like units in regular street fabrics and correctly retained green areas as cell attributes. In the 3.68 km^2 pilot, 55 minor isolated seeds were suppressed; at the city scale, 5,232 essentially non-overlapping cells covered more than 99.99% of the municipal area. These aggregate successes, however, were misleading: ranked QA identified oversized and multipart contexts, footprint-crossing assignments, and a strong divergence between cell-weighted and area-weighted morphology statistics (Table 2).

Exploratory clustering reinforced the diagnosis. At k=4 and k=5, small geometry- scale classes contained many spatially isolated cells and medoids with QA flags. The first method was therefore retained as a reproducible urban-block characterisation experiment but rejected as the final UWG support. The decisive issue was not invalid geometry; it was the attempt to force street allocation, mapping, and climatic neighbourhood support into the same variable-area polygon. The 250-, 500-, and 1000-m station-support tests produced very similar aggregate temperature results. Across Consolata, Giardini Reali, and Reiss Romoli, the mean RMSE over January/July and all-hour/day/night subsets was 1.57, 1.59, and 1.59°C, respectively. The complete scale-dependent spread was only 0.03°C. The 250 m support had the numerical minimum, but all three scales improved the mean January all-hour comparison and worsened the July comparison, while nighttime warming remained excessive. No scale solved the common-template transfer error.

The 500 m support was selected because its performance was effectively indistinguishable from the numerical minimum, it included multiple urban blocks, and it was less likely than the 250 m support to let a single plaza or small structure dominate the input. The 1000 m support offered no performance advantage and risked smoothing distinct fabrics.

All six morphology surfaces passed the 50 m grid-origin-shift criteria (Table 3). At common physical locations, Spearman correlations ranged from 0.994 to 0.998, normalised RMSE ranged from 0.054 to 0.121 baseline standard deviations, and every location remained within the same or an adjacent baseline quintile. Mean movement in the six-variable standardised space was 0.179, and its 95th percentile was 0.336.

In the matched scale test, weighted total height variance was 25.83, 19.27, and 12.77 m^2 at 250, 500, and 1000 m; total facade/site variance was 0.125, 0.103, and 0.075. Relative to 250 m, 500 m reduced these total variances by 25.4% and 18.1%, whereas 1000 m reduced them by 50.5% and 40.3%. Thus 500 m preserves more local contrast than 1000 m without establishing a unique physical optimum (Supplementary Table S8).

The within- and between-class decomposition clarifies this tradeoff. Both this decomposition and the total variances use the same 304 quality-matched centres and the same design weights; total variance equals within-class plus between-class variance. With the 500 m class labels held fixed, increasing support from 250 to 500 m reduced within-class height variance from 13.06 to 6.42 m^2 (50.8%), while between-class variance remained approximately 12.8 m^2. Within-class facade/site variance fell by 43.3%, with between-class variance changing from 0.084 to 0.079. At 1000 m, within-class variability decreased further, but between-class height and facade/site variances fell by approximately one third relative to 500 m. These results favour 500 m as a compromise between averaging and retained contrast. Because the class labels were fitted at 500 m and held fixed, the decomposition is reference-conditioned rather than an independent optimisation of support size (Supplementary Table S8).

These results support the regular-grid surface as stable for citywide characterisation, while the scale test supports 500 m as a transparent compromise rather than an optimised physical footprint. The adopted method generated 12,997 in-boundary assignment locations. Of these, 11,017 had complete 500 m supports, 1,980 were municipal-edge supports, and 10,593 met the strict *zero-QA-score* fitting rule. The run completed in 14.11 min with a peak process memory of 2.07 GB. The centre-in-boundary sampling rule represented 129.146 km^2, or 99.340%, of the exact polygon. The omitted 0.858 km^2 consisted of narrow boundary fragments whose 100 m cell centres fell outside the city. There were 135 locations without usable average-height or facade inputs: 122 contained no mapped buildings, and 13 had insufficient valid height. These rows remained visible and unassigned. Of the 102,006 living tree records, 78.48% had valid height and 69.73% had valid crown diameter; the canopy must consequently be interpreted as an estimated rather than remotely observed input (see Table 4 for the 500 m-support variable resume).

The in-house extraction audit passed 22 of 28 checks and issued six documented warnings, with no failures. All formulas, ranges, grid identifiers, geometries, CRS, table counts, QA reconstruction, and file hashes were verified. The warnings concerned edge support, missing or incomplete height, and imputed tree attributes. Crown masking gave diameter MAE 1.99 m and RMSE 2.67 m. Of 30,880 missing-crown trees,

Table 2
Main irregular-cell audit findings.

Finding	Result	Interpretation
Cells larger than 100,000 m^2	154	2.9% of cells but 50.2% of the mapped area
Cells larger than 1 km^2	8	24.1% of the mapped area
Largest cell	21.11 km^2	33 parts and 7,638 building polygons
Cells with *QA* score at least 3	749	46.2% of the mapped area
Assigned cells with >1% footprint outside	189	Building/cell topology inconsistency
Maximum reported building fraction	1.440	Physically impossible input

Table 3
Same-location morphology sensitivity after shifting the 100 m grid origin by 50 m east and north.

Variable and MAE unit	MAE	RMSE / SD (–)	Spearman (–)	Same/adjacent (%)
Building fraction	0.006	0.059	0.998	100.0
Height (m)	0.239	0.121	0.995	100.0
Facade-to-site ratio	0.015	0.077	0.997	100.0
Green (percentage points)	0.516	0.054	0.997	100.0
Canopy (percentage points)	0.235	0.102	0.995	100.0
Road fraction	0.005	0.086	0.994	100.0

Table 4
Citywide distributions of selected 500 m-support variables.

Variable and unit	Mean	Median	5th–95th percentile
Building fraction (fraction)	0.185	0.165	0.007–0.431
Mean building height (m)	11.39	11.22	4.86–18.90
Facade / site (ratio)	0.387	0.300	0.017–0.992
Mapped green (%)	19.33	15.42	2.29–47.85
Estimated tree canopy (%)	2.85	1.90	0.00–8.88
Road fraction (fraction)	0.208	0.216	0.036–0.374

21,926 also lack height, a subgroup with only 23 observed-crown analogues. Across propagation scenarios, k=7 and A–D assignment agreement remained at least 96.86% and 97.42%. The intersection correction affected 75 supports materially but no frozen k=7 labels; two k=4 and three k=8 labels changed, with all 19 medoid input differences below 10^{-7} (Supplementary Sections S1–S2).

3.2. Morphological classification

The k=2 solution yielded the highest silhouette (0.405) and spatial stability (mean ARI 0.977), but primarily represented a coarse open-versus-dense division. Increasing k exposed intermediate and special fabrics. Beyond k=8, the smallest K-means class fell below 3% of the fitting population, neighbour agreement declined from 0.831 at k=8 to 0.772 at k=12, and the number of spatial components increased from 282 to 458. The k=4-8 range, therefore, captured the useful transition between under-description and fragmentation.

Within that range, the diagnostic ranking favoured k=4, k=7, and k=8 (Table 5). The k=4 solution provided the strongest balance and spatial coherence. The k=7 solution improved Davies-Bouldin performance and revealed a distinct high-plan- density, low-facade class. The k=8 solution retained that fabric and separated an additional open, tree-rich class, but relied more strongly on the partly imputed canopy variable and produced a more fragmented map.

Final CLARA mean medoid distance decreased from 0.786 at k=4 to 0.665 at k=7 and 0.637 at k=8. The corresponding 95th-percentile distances were 1.506, 1.154, and 1.126. This diminishing improvement came with minimum class sizes of 19.4%, 3.68%, and 3.33%, respectively, confirming that compactness alone could not determine k.

The sensitivity agreement was moderate rather than universal. For k=7, ARI was 0.819 against the same-data K-means, 0.668 against Manhattan CLARA, 0.618 after Yeo-Johnson transformation, 0.700 for group-balanced features, and 0.544 against Ward clustering on the sample. The broad morphology structure recurred, but exact class boundaries depended on preprocessing and distance. The in-house clustering validator passed 38 checks with five methodological warnings and no errors. The k=4 solution produced a clear sequence from open low-rise to dense tall fabric, but merged large-roof industrial or campus forms into broader classes. The k=7 solution preserved this distinct combination as C6 (see Table 6): median building fraction 0.535, height 12.21 m, facade-to-site ratio 0.185, tree cover 0.005, and grass cover 0.027. High plan coverage combined with a low exposed facade is not a simple continuation of the dense perimeter-block class C5, whose median facade-to- site ratio was 0.834. Map inspection confirmed that C6 corresponded to coherent, large-footprint fabrics rather than to high statistical distance alone. Fixed-algorithm feature tests yielded ARI 0.496–0.749. Buffered spatial tests gave mean ARI 0.700, 0.653 and 0.599 for k=4, 7 and 8, under a different design from Table 5. HDBSCAN found two to seven groups but labelled 49.5–59.3% of supports as noise; its silhouette among assigned points cannot directly rank a full-coverage classification (Supplementary Table S3 for feature tests; repository HDBSCAN and buffered-stability evidence, Appendix A).

The k=8 solution added a small open, tree-rich class of 353 supports but offered only a small further decrease in medoid distance, had lower neighbour agreement, and was more sensitive to Manhattan and hierarchical alternatives. Because the tree canopy was the least complete primary input, k=8 was retained as a high-resolution sensitivity case. K=7 was retained retrospectively as a purpose-dependent working resolution, while k=4 remained the compact statistical baseline. C6 contains 3.68% of fitting supports, not 3.68% of the city area. Official geometry and land-use checks confirm the presence of industrial halls in 20 overlapping tail windows at Mirafiori and IVECO; these are two sites, not 20 independent districts (Supplementary Figures S1–S2).

No spatial cleanup was applied. The proposed k=7 cleanup would have affected 85 locations across 59 small components, but those exceptions had not been physically approved and could represent actual special fabrics. Of 13 successfully simulated boundary/cleanup cases with assignment switches, only three switches improved annual weather fidelity (Figs. 2-5).

3.3. EPW reduction and modelled urban heat island

All 11 annual simulations of the four k=4 and seven k=7 medoids succeeded under the final common template. Mean annual modelled UHI across k=7 remained approximately 1.69–1.95°C. Several outputs were nearly redundant, so k=7 retained a special morphology response without requiring all seven files in the weather product. The exhaustive subset curve showed a rapid decline in medoid representation error from 1 to 4 EPWs, followed by a smaller improvement from 4 to 5 (Fig. 6a; Supplementary Table S1). The four-file maximum and weighted mean annual hourly RMSE were 0.04 and 0.02°C at three-decimal serialisation. The exact 127-subset search took a median of 7.37 ms over 50 repetitions, excluding file loading and UWG simulation.

With the standard one-decimal writer, maximum and weighted mean medoid RMSE were 0.06 and 0.03°C. The improvement from four to five files was 0.003°C at one decimal and 0.009°C at three decimals. These rounded diagnostics describe serialisation effects, not physical precision; gains and screening decisions are calculated from unrounded values. Neither a significance test nor total input/model uncertainty establishes that the fifth-file gain is negligible.

The four-file minimax set merged C0+C1 and C3+C4+C5, while retaining C2 and C6. The selected representatives were real supports of C1, C2, C4, and C6, labelled A-D (Table 8). Across the strict fitting population, A represented 30.26%, B 9.21%, C 56.85%, and D 3.68%. The all-city lookup assigned 12,862 of 12,997 locations; the 135 missing-input locations remained blank. The all-city A–D counts are 4,697, 1,090, 6,662 and 413, including provisional and restricted supports.

Under the common Zone 4A template with a 3.05 m floor height, four representatives were selected from the seven annual medoid outputs. Annual mean modelled UHI was 1.69, 1.75, 1.81 and 1.95°C for A–D,

Table 5
K-means diagnostic comparison used to shortlist candidate counts.

k	Composite rank (–)	Silhouette (–)	Davies–Bouldin (–)	Screening ARI (–)	Neighbour agreement (–)	Smallest class (n)
4	2.30	0.298	1.217	0.925	0.877	2,193
7	2.45	0.305	1.056	0.841	0.848	498
8	2.80	0.299	1.108	0.907	0.831	382
5	2.95	0.286	1.131	0.902	0.865	1,525
6	4.50	0.278	1.147	0.676	0.841	961

Table 6
Physical profiles of the seven retained morphology classes. Values are medians over the strict fitting population.

Class	n	Built (–)	Height (m)	Facade/site (–)	Tree (–)	Grass (–)	Interpretation
C0	1,407	0.021	6.33	0.039	0.000	0.133	Very open, low-rise, mixed green
C1	1,798	0.049	7.33	0.114	0.000	0.392	Open, low-rise, green
C2	976	0.110	11.60	0.256	0.074	0.242	Open, mid-rise, tree-rich
C3	1,962	0.223	11.47	0.383	0.020	0.143	Moderate-density, mid-rise
C4	1,498	0.232	16.73	0.654	0.072	0.078	Moderate-density, taller, tree-influenced
C5	2,562	0.347	15.28	0.834	0.035	0.033	Dense, taller, low green
C6	390	0.535	12.21	0.185	0.005	0.027	Dense, large-footprint, low-facade

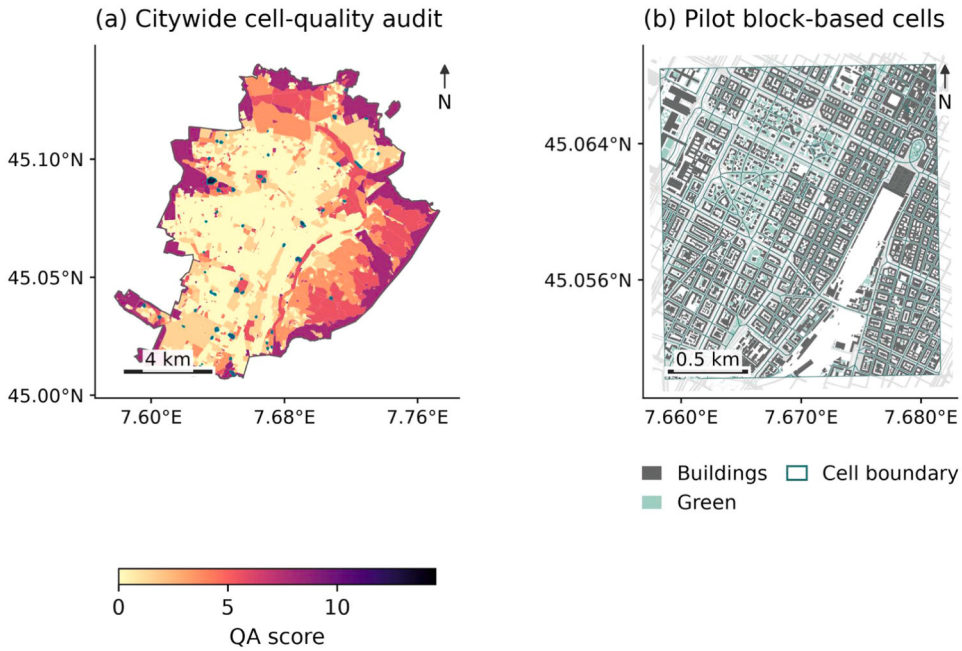


Fig. 2. Preliminary irregular-cell evidence: (a) citywide QA-score distribution, with outlines marking the 60 highest-ranked manual-review cases; and (b) building-seeded context cells in the central west pilot area, showing buildings, allocated road/open context, mapped green areas, and final boundaries.

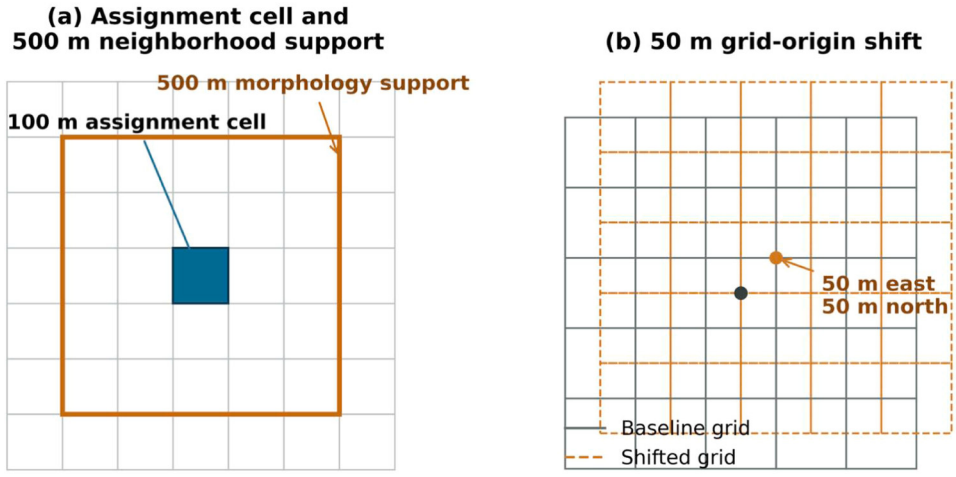


Fig. 3. Relationship between the 100 m assignment grid and the 500 m morphology support, including the 50 m grid-origin sensitivity test.

respectively (Table 9). Winter UHI exceeded summer UHI across all categories, and nighttime UHI was much larger than daytime UHI.

The accepted annual library spans 0.26°C in mean temperature. Although this range is modest, its rank order is stable across annual, seasonal, day, and night subsets. Absolute nighttime warming is discussed cautiously because station tests indicate common overprediction.

All 105 stratified and 20 of 21 purposive annual runs succeeded. The four-file weighted mean RMSE was 0.05°C, the weighted 95th percentile was 0.10°C, and the stratified maximum was 0.35°C, compared with 0.03°C and 0.08°C mean and 95th percentile for seven medoids (Table 7). Typical-case screens passed, but the 0.50°C stress screen failed at grid 1435, with a 1.33°C error already relative to its own C6 medoid

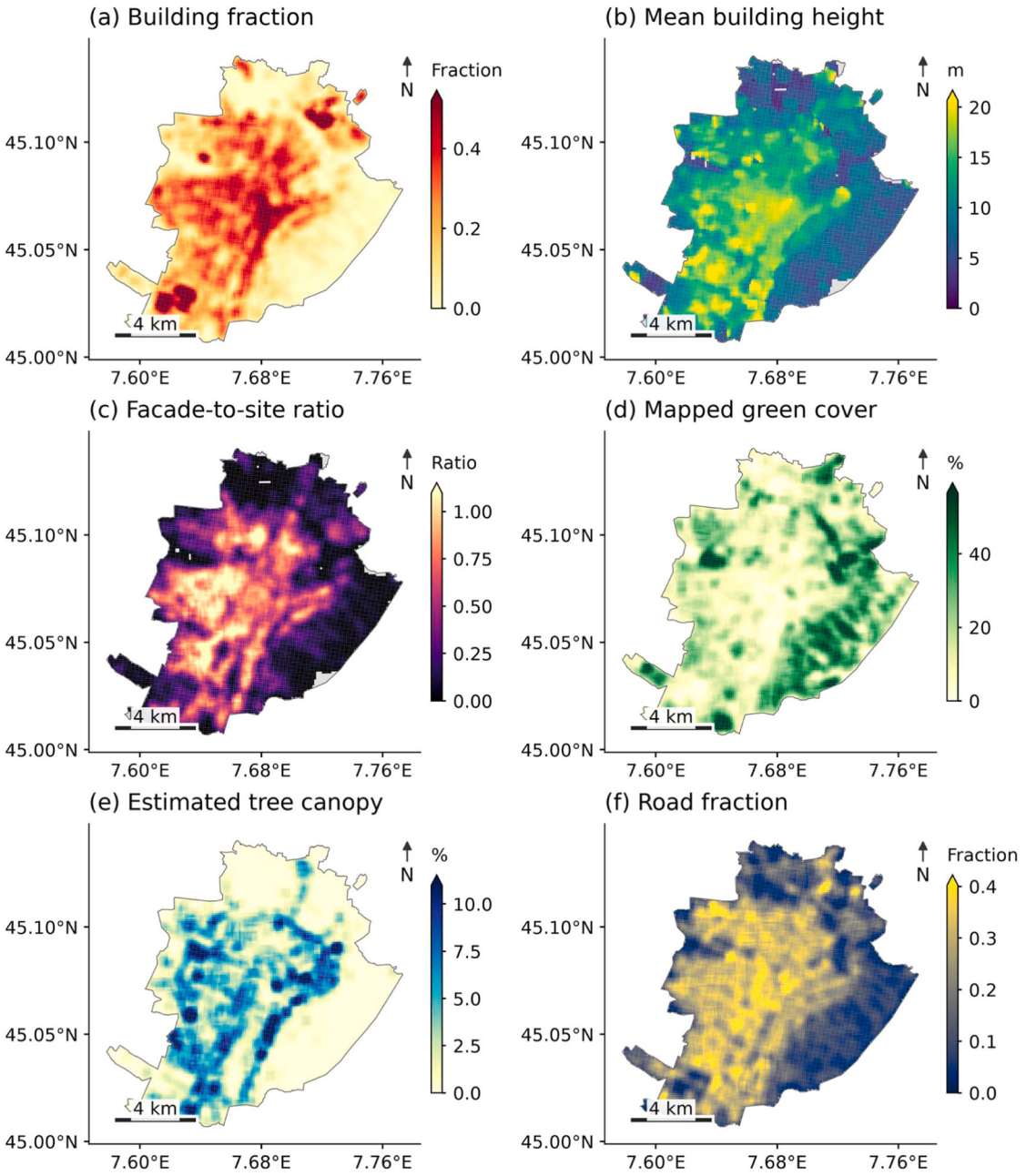


Fig. 4. Citywide 500 m-support morphology surfaces: (a) building fraction; (b) area-weighted average building height; (c) facade-to-site ratio; (d) mapped green cover; (e) corrected estimated tree-canopy cover; and (f) road-surface fraction. Color scales start at zero and use each panel's finite-value 98th percentile as the upper limit; extended colorbars denote values above that limit. Municipal outlines, north arrows, scale bars and WGS84 longitude and latitude ticks provide spatial reference.

(Fig. 8). Grid 8230 failed at 300 s and remains excluded; its 180 s diagnostic is not substituted. These sample maxima do not bound unsimulated supports.

A separate industrial-tail representative reduced the maximum across two prospective holdouts to 0.48°C, but the original triggering support remained at 0.56°C; replacing ordinary C6 globally worsened representation. A warehouse-only stock diagnostic changed annual response by up to 0.96°C RMSE without identifying actual local constructions. No fifth file was adopted. C6 supports at or above its 95th-percentile building fraction (approximately 0.795; the unrounded quantile is retained in the implementation) are outside the unqualified four-file recommendation domain. This conservative post hoc guardrail is not a physically validated applicability boundary. It does not retrospectively pass the original stress screen or re-estimate the stratified distribution (Supplementary Section S4).

The exact cutoff restricts 20 of 12,862 assigned locations; rounding its display to 0.795 changes no flags. Perturbing it by ±0.005 in building fraction changes the restricted count to 19–22, illustrating sensitivity rather than uncertainty bounds. A further 19 C6 supports between the strict-fitting 90th percentile (approximately 0.755) and the cutoff receive a near-threshold caution flag. Both percentiles use the same 390 strict-fitting C6 supports; all 413 assigned C6 locations are screened. The guard leaves 12,842 assigned locations unrestricted, including the 19 caution-flagged supports; other quality flags still apply. Locations below the restriction are not thereby validated.

All eight shared-parameter states, including baseline, passed the four-file medoid screens. The fixed set reached at most 0.07°C annual maximum RMSE, although the absolute same-medoid response changed

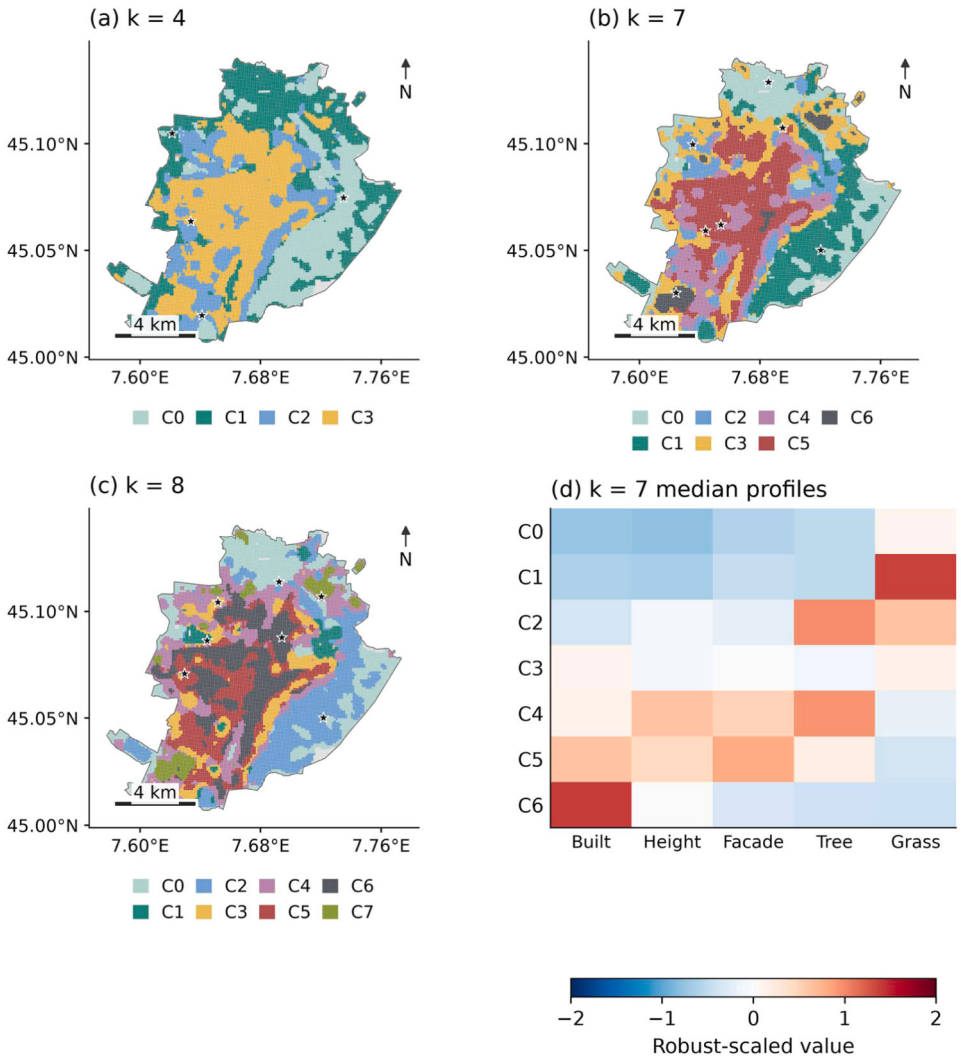


Fig. 5. Candidate morphology solutions: (a) k=4 compact baseline; (b) k=7 retained solution; (c) k=8 high-resolution sensitivity case; and (d) robust-scaled median input profiles for k=7. Stars identify medoid supports; municipal-edge assignments remain provisional. Corrected-input checks change two k=4 and three k=8 assignments; k=7 is unchanged.

by up to 0.92°C RMSE, and some secondary-optimal representatives changed. Crown quartile scenarios changed the same-medoid response by at most 0.005°C RMSE and retained low fixed-set error despite a changed upper-quartile optimum. With 2022 forcing, the AvMY-selected set gave a maximum medoid RMSE of 0.050°C, versus 0.048°C after reoptimisation. Detailed numerical diagnostics are in the supplementary material; neither these scenarios nor the crown test quantify total physical uncertainty.

The additional heat tests retained small medoid compression errors: the best four-file maximum annual RMSE was 0.035-0.043°C across nine states including baseline. All three-decimal screens passed. The weighted secondary criterion preferred C1/C3/C5/C6 in three alternatives, but the baseline C1/C2/C4/C6 identities, simulated under each scenario with nearest assignments, remained minimax-equivalent throughout. With one-decimal output, the reverse road-ranked scenario failed the curve-improvement screen, although four-file maximum RMSE remained 0.059°C (Supplementary Section S7.2 and Table S11).

Traffic magnitude and timing changed the same-medoid annual series by at most 0.047°C RMSE. Setting the canyon HVAC heat-release fraction to zero produced a larger maximum of 0.119°C and lowered the C6 summer-night mean temperature by 0.24°C; unchanged baseline files then reached 0.260°C summer-night RMSE. Across the 14 selected non-medoids, the maximum error relative to their scenario-class medoid was 0.347/0.363°C for the two spatial patterns, versus 0.349°C at baseline. These purposive checks do not estimate citywide error percentiles (Supplementary Table S12 and Figures S3-S6).

3.4. Station-based evaluation

Each urban station was uniquely located on the assignment grid (Table 10). Consolata and Giardini Reali were not assigned to the same raw class: they belonged to C5 and C4, respectively. Their strongly different building density, facade-to-site ratio, and vegetation were therefore retained on the morphology map. They shared library EPW C only because the C4 and C5 medoid outputs were weather-redundant, under the fixed-template temperature criterion. Alenia's C3 assignment maps to C, while Reiss Romoli is assigned to D. No urban station evaluates A or B.

Only 85.72% of the Alenia support was covered by municipal GIS, leaving 35,698.7 m^2 outside the extraction. Its class and modelled weather were retained as lower-confidence evidence. Absolute station agreement in the synchronised processed 2022 data was mixed (Table 11). At Consolata, station-support UWG reduced annual RMSE from 2.10°C for Caselle to 1.37°C, and library C gave 1.32°C. At Giardini Reali, Alenia, and Reiss Romoli, unchanged Caselle remained closer to processed urban temperature than either local-support UWG or the

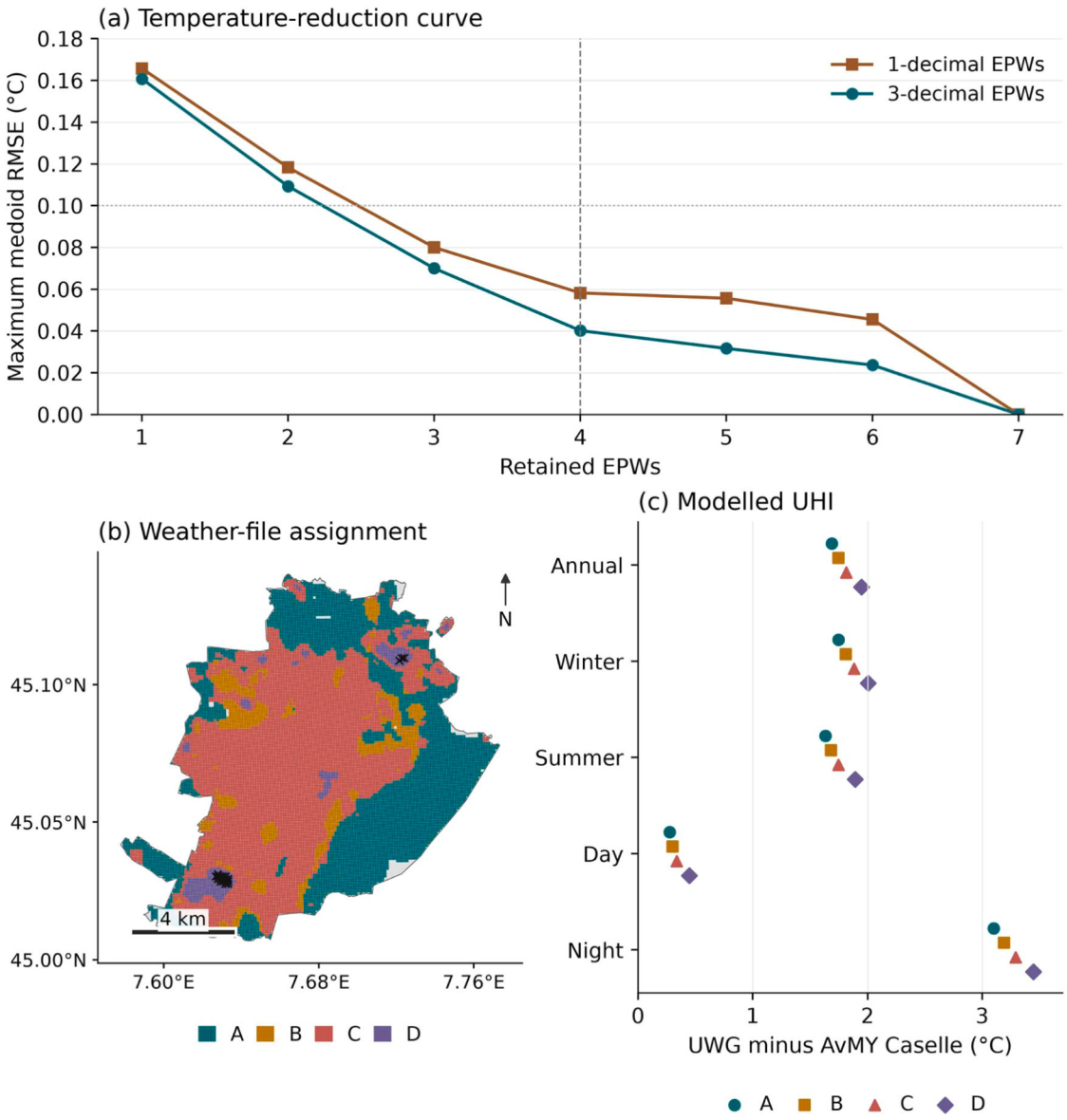


Fig. 6. Weather-product reduction and delivery: (a) final-template minimax error at both writer precisions as the seven k=7 medoid outputs are reduced; (b) citywide A-D EPW assignment with restricted high-density C6 supports marked; and (c) annual, seasonal, daytime, and nighttime UHI of the accepted four-file candidate library relative to Caselle.

Table 7
Non-medoid fidelity to exact-support UWG temperature. Main statistics use the 105 stratified supports and post-stratification weights; the purposive maximum is a separate, unweighted stress result. All RMSE values are in°C; three-decimal serialisation is primary.

Statistic	Seven medoids (°C)	Four files (°C)	Population / scope
Weighted mean	0.03	0.05	105 stratified
Weighted median	0.03	0.04	105 stratified
Weighted 95th percentile	0.08	0.10	105 stratified
Observed stratified maximum	0.35	0.35	105 stratified
One-decimal weighted mean	0.05	0.06	Writer sensitivity
One-decimal weighted P95	0.09	0.11	Writer sensitivity
Purposive maximum	1.33	1.33	20 successful stress cases

Table 8
Accepted four-file candidate library and representative support inputs. Grid IDs identify 100 m assignment-cell centres; simulations use the associated 500 m squares (coordinates in Supplementary Table S2).

EPW	Raw classes	Centre grid ID	Height (m)	Built (–)	Facade/site (–)	Tree (–)	Grass (–)
A	C0, C1	3788	7.35	0.047	0.111	0.003	0.412
B	C2	10026	11.55	0.128	0.230	0.076	0.244
C	C3, C4, C5	5210	17.27	0.230	0.699	0.076	0.104
D	C6	1524	13.27	0.577	0.222	0.003	0.014

assigned library file. Consolata's RMSE improvement was 0.73°C (95% paired seven-day block interval 0.59–0.87°C) for exact-support UWG and 0.78°C (0.64–0.92°C) for the library (Fig. 7). These intervals quantify temporal sampling, not total sensor or model uncertainty.

In the AvMY comparison, measured urban annual means spanned 0.85°C versus approximately 0.17°C for exact-support UWG; the AvMY

Table 9
Mean temperature difference between accepted library EPWs and AvMY Caselle. Values are in°C; winter/summer are October–March/April–September, and day/night use EPW labels 07–18/19–06. Three-decimal writer diagnostics are summarised to two decimals.

EPW	Annual (°C)	Winter (°C)	Summer (°C)	Day (°C)	Night (°C)
A	1.69	1.75	1.63	0.28	3.10
B	1.75	1.81	1.68	0.30	3.19
C	1.81	1.88	1.75	0.34	3.29
D	1.95	2.01	1.89	0.45	3.45

library spans 0.26°C. The shared UWG template, therefore, compresses the observed inter-station thermal diversity. Category C illustrates this directly: both exact support UWG files are close to the same library file, but measured Consolata and Giardini Reali have substantially different relationships with Caselle.

Seasonal and diurnal results explain part of the error. For example, the AvMY Consolata-support UWG had daytime bias -0.23°C and nighttime bias +1.15°C; at Giardini Reali, the corresponding values were -0.44 and +2.83°C. Similar day-night asymmetry occurred at Alenia and Reiss Romoli. The model can be slightly cool or close to observations during the day while being too warm at night, so an annual bias alone understates the structural mismatch. For 2022, the statement that only Consolata improves refers specifically to annual dry-bulb RMSE, not every season or time of day. The annual direction persists with 1-, 3-, 7- and 14-day bootstrap blocks. Excluding 62 suspicious Reiss Romoli flatline hours changes annual RMSE by at most 0.03°C (Supplementary Tables S5–S6).

Table 12 summarises the principal robustness tests; detailed designs and results are provided in the supplementary material.

4. Discussion

At the present level of evidence, seven working morphology classes were retained to distinguish the principal physical fabrics identified in the selected Torino feature space, and four EPWs reproduce the seven medoid-level UWG temperature responses within a maximum annual hourly RMSE of 0.06°C with the delivered one-decimal EPWs (0.04°C at diagnostic three-decimal serialisation). The distinction is central: seven describes urban form, while four describes a reduced weather product under one model, source EPW, and error definition. The result gives built-environment users a manageable alternative to both an unmodified airport file and thousands of nominally local files. This is a temperature-reduction result under shared assumptions, not evidence that Torino has four observed climates or universally accurate local weather files.

The four-file selection is supported by three pieces of evidence. First, it is the exact minimax subset for four retained medoids. Second, the maximum error improves by 0.009°C at three-decimal serialisation when moving from four to five files, whereas the improvement from three to four is 0.030°C. Third, the 105 stratified non-medoid tests give a weighted 95th-percentile error of 0.10°C. The result is therefore more than an arbitrary merging of adjacent classes. Nevertheless, the answer is conditional. The exhaustive reduction used one medoid per morphology class, followed by a stratified evaluation of within-class supports. A medoid is real and central in input space, but it cannot reveal the weather spread of distant, boundary, or atypical members. The four-file library is consequently restricted and simulator-conditional. The industrial-tail failure limits applicability even though typical-case errors are small; the subsequent restriction does not retroactively turn that failed screen into a pass.

The fifth-file gain must be interpreted alongside serialisation and input sensitivity. UWG v5.3.4 writes dry-bulb temperature to one decimal by default [35]: the four-to-five reduction in maximum medoid RMSE is 0.003°C at that precision and 0.009°C with three-decimal output. At three decimals, the largest annual same-medoid response to a conditional crown-quartile change is 0.005°C; this measures input sensitivity, not compression error. Recomputed four-file maximum RMSE is approximately 0.04°C across the crown states and does not exceed 0.07°C across the seven shared-parameter alternatives (Supplementary Sections S3–S4). These bounded tests support compact representation under the tested assumptions, but their response magnitudes are not confidence intervals and cannot be combined into total height, crown and model uncertainty. They therefore do not establish that the fifth-file benefit is statistically insignificant.

Existing morphology-to-weather studies commonly derive the number of weather files from input categories [13,9,10,11] or generate localised files for every selected patch or block [14,12]. The Torino results show why a second reduction stage is useful. Consolata and Giardini Reali differ strongly in density, facade exposure, and vegetation and therefore appropriately retain different morphology classes. Under the fixed UWG configuration, however, their class-medoid weather outputs are nearly identical and can share a single public file without claiming that their urban forms are identical. Conversely, the C6 large-footprint class accounts for only 3.68% of strict fitting supports but should not be removed. Its low facade-to-site ratio at very high plan density is physically unlike dense perimeter blocks, and its UWG response is the most distinct in the k=7 set. A product designed only to minimise average error could discard this small class. The minimax-first objective preserves it. This two-stage logic is the study's main methodological contribution. Input-space clustering supports physical interpretation and spatial assignment; output-space reduction supports product economy. Neither stage is required to perform the other function. For practical use, the companion Torino Urban Morphology and EPW Atlas packages the study's spatial classifications and candidate weather files. The browser and command-line interfaces accept latitude and longitude, identify the containing 100 m assignment cell, report its k=4, k=7, and

Table 10
Urban-station morphology and EPW assignments. Descriptors are for station-centred squares; class labels follow the containing grid cell.

Station	Grid ID	Class	EPW	Built (–)	Height (m)	Facade/site (–)	Tree (–)	Grass (–)	Status
Consolata	7096	C5	C	0.506	16.98	0.810	0.029	0.000	Fit member
Giardini Reali	6455	C4	C	0.269	17.96	0.428	0.118	0.124	Fit member
Alenia	7925	C3	C	0.190	10.15	0.242	0.011	0.097	Edge provisional
Reiss Romoli	11464	C6	D	0.400	9.04	0.302	0.000	0.083	Fit member

Table 11
Annual dry-bulb comparison with synchronised processed urban-station weather in 2022. All values are in°C; positive bias is model minus station.

Station	Station minus Caselle (°C)	Caselle RMSE (°C)	Support bias (°C)	Support RMSE (°C)	Library RMSE (°C)	Support–library RMSE (°C)
Consolata	1.44	2.10	0.44	1.37	1.32	0.14
Giardini Reali	0.80	1.38	0.95	2.57	2.57	0.01
Alenia	0.65	1.59	1.03	1.85	1.91	0.08
Reiss Romoli	0.89	1.46	0.87	1.93	2.03	0.17

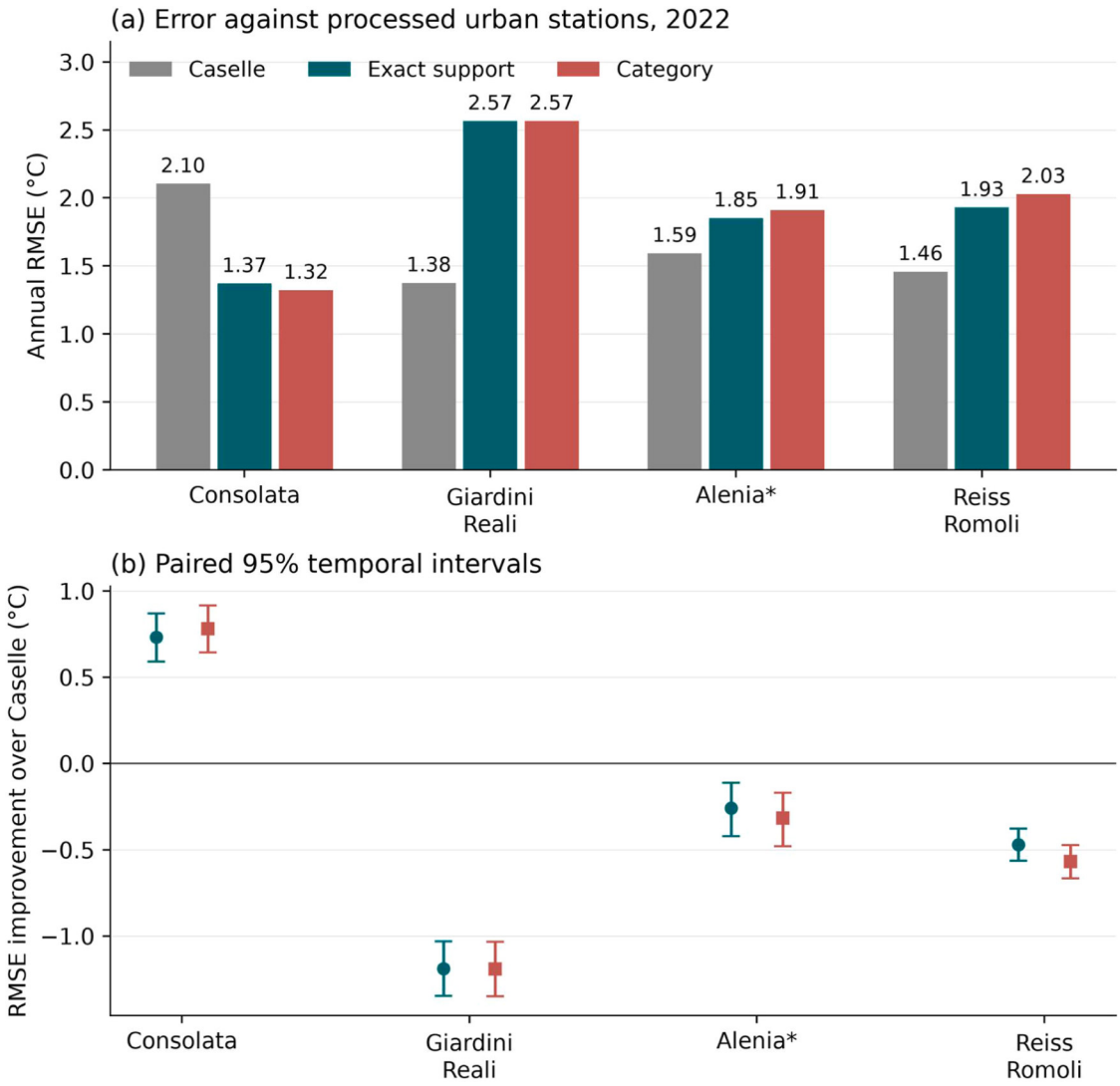


Fig. 7. Synchronised 2022 station evaluation: (a) annual dry-bulb RMSE of Caselle, exact station-support UWG and assigned-category UWG against processed urban weather; and (b) RMSE improvement over Caselle with paired seven-day moving-block 95% intervals (5,000 replicates). Positive improvement means a lower RMSE. Three-decimal UWG serialisation is used, and Alenia is provisional because of incomplete GIS coverage.

k=8 classes, 500 m-support morphology descriptors, and quality flags, and return the assigned A-D candidate weather file where the applicability criteria permit a recommendation. The browser provides direct downloads of both the EPW and the exact UWG JSON configuration used to generate it. A canonical GeoPackage contains the complete grid assignments, cluster medoids, municipal boundary, morphology variables, class labels, and EPW categories; equivalent GeoJSON, shapefile, and tabular distributions support reuse beyond the lookup interface. Architects and urban planners can therefore select a documented candidate for the urban weather file without rerunning GIS extraction, clustering, or UWG, while researchers can reuse the classifications as spatial strata for other urban-microclimate studies. The lookup is a transparent decision aid: its displayed UHI values are UWG-modelled differences from Caselle, not measured local corrections or guarantees of point-scale meteorological accuracy. The 0.26°C annual mean range across accepted library files is smaller than might be expected from the visual and statistical diversity of Torino morphology. This result quantitatively indicates that, under the common UWG assumptions, much of the morphology-induced annual temperature response is redundant. The practical consequence is that Torino may require fewer modelled urban files than morphology classes. The annual mean also conceals larger temporal structure. The A–D nighttime UHI range is 0.35°C. The ordering of A-D is stable across annual, seasonal, day, and night subsets. A small annual mean spread, therefore, should not be interpreted as identical hourly behaviour.

This model-specific bridge complements existing morphology products (Table 13). The global LCZ map uses 100 m categorical pixels [36]; UT-GLOBUS aggregates canopy parameters on 1 km^2 supports at a 300 m sliding step [37]; GLAMOUR provides continuous height and footprint estimates [38]. Their variable definitions require reconciliation before reuse: UT-GLOBUS includes roofs and walls in its surface ratio, whereas the present facade/site approximation excludes roofs, party walls and courtyard rings. GLAMOUR estimates building fraction and area-weighted height, but additional facade and vegetation inputs are needed for this UWG workflow. The contribution is the explicit link between local inputs, morphology classes and tested weather representatives; an aligned Torino benchmark against these products was not performed.

The station comparison assesses observational agreement separately from reduction fidelity. The exact station-support simulations and assigned library files are close to each other, establishing that the reduction preserves the behaviour of the fixed model. They are not consistently close to measurements. Only Consolata improves annual dry-bulb RMSE over Caselle in the synchronised 2022 test; Giardini

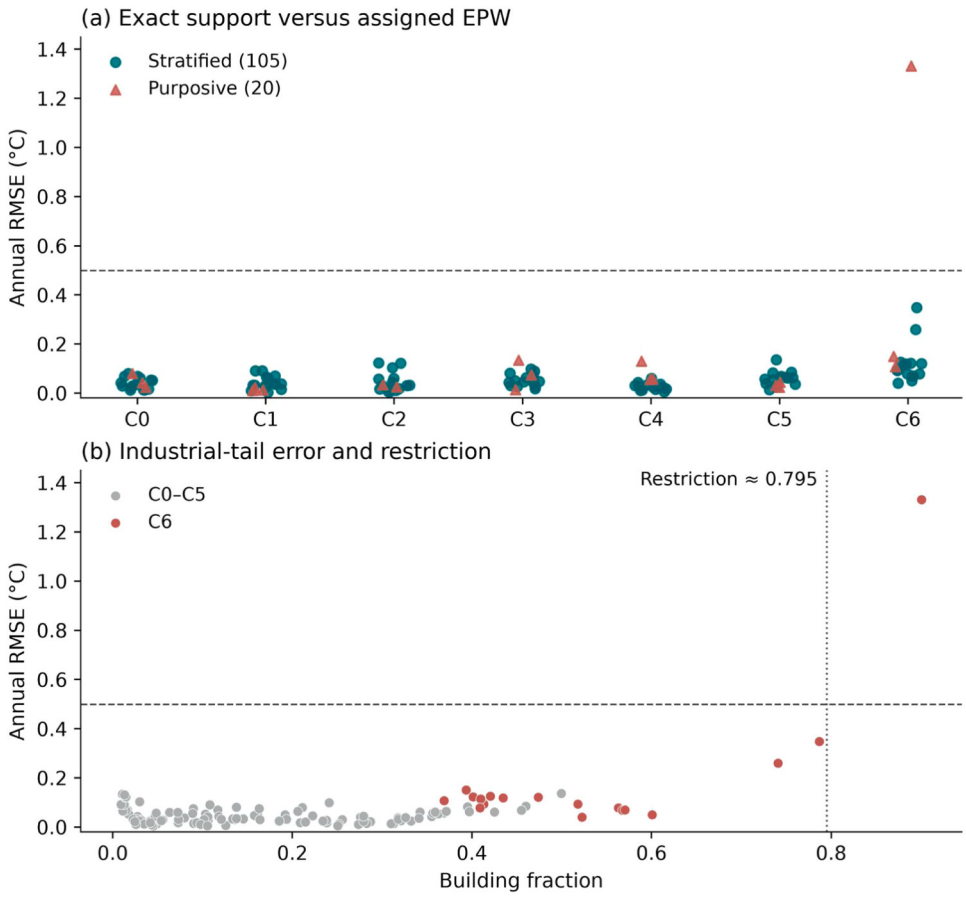


Fig. 8. Non-medoid applicability: (a) annual exact-support versus assigned four-file RMSE by class for 105 stratified and 20 successful purposive cases; (b) the same errors against building fraction, with C6 distinguished. Points are individual, potentially overlapping supports, not independent neighbourhoods. The 0.50°C line is the frozen stress screen; the vertical line marks the subsequently investigated C6 density-tail threshold. Weighted typical-case summaries are in Table 7. The failed 300 s case is excluded, not assigned zero error.

Table 12
Scientific robustness tests and their limits. These are separate analyses within this project, not verification by an independent external team.

Test	Main finding	Scope / limitation
Matched support size	500 m: weighted total height/facade variance −25.4%/−18.1% versus 250 m	Same 304 centres and design weights as the within/between decomposition; not a physical optimum
Feature/spatial clustering	Feature ARI 0.496–0.749; buffered k7 mean 0.653	Moderate reference agreement, not true-class accuracy
Crown missingness	96.86% minimum k7 agreement; 0.005°C largest primary response	Conditional imputation; missing-height subgroup weakly represented
Non-medoid response	Weighted mean/P95: 0.05/ 0.10°C	105 stratified supports; original population, not a filtered-domain estimate
Industrial tail	1.33°C trigger; alternative still 0.56°C at trigger	20 restricted and 19 caution-flagged overlapping C6 supports; post hoc safeguards, no fifth file
Common parameters	Largest fixed-set maximum 0.066°C; absolute response up to 0.92°C	Seven alternatives; anthropogenic heat and canyon HVAC fraction fixed
Traffic and HVAC heat	Best-four maximum 0.035-0.043°C; HVAC response up to 0.119°C	84 annual runs; hypothetical spatial heat; one-decimal curve exception
2022 medoid reduction	Fixed-set maximum 0.050°C	Cooling-season rank 1/10; not all non-medoids resimulated
2022 observations	Annual improvement only at Consolata	Processed data; A/B untested; humidity outputs excluded

Reali, Alenia, and Reiss Romoli are modelled too warm, especially at night. This difference between fidelity and accuracy is important for model-reduction research. A reduced model or library can emulate its parent simulator extremely well while inheriting the parent's structural bias. In Torino, the measured urban AvMY means span more than three times the accepted library range. The model's common boundary-layer, reference-site, building-stock, material, schedule, and anthropogenic assumptions appear to suppress part of the site-to-site diversity. The parameter screen showed that atmospheric/reference settings can have much larger temperature effects than the tested zone and floor-height choices, but adjusting them independently for each neighbourhood would undermine the controlled morphology comparison. The strongest nighttime mismatch is consistent with recent scrutiny of UWG nighttime boundary-layer height [39] and with sensitivity evidence from Toulouse [26]. These references support further physical investigation, not post hoc tuning to force agreement at one station. Multi-station calibration would require simultaneous raw observations, defensible parameter ranges, and an explicit uncertainty framework such as that advocated by Mao et al. [28]. The annual parameter alternatives support compact medoid representation despite changes in absolute warming, but do not map real spatial variation in stock and activity or uniquely attribute nighttime bias to boundary-layer exchange, heat storage or anthropogenic schedules.

These targeted heat tests strengthen the conditional reduction evidence, but the 0.10°C medoid tolerance alone permits three files in seven states and two in the two HVAC-allocation states at three-decimal precision. Four is therefore a fidelity-complexity compromise, not a uniquely required minimum, and the one-decimal curve exception warrants retaining the precision qualification. Traffic magnitude and

Table 13
Scope and reuse of selected morphology products. Pixel spacing, aggregation support and meteorological resolution are distinct; this is a conceptual comparison, not an aligned accuracy benchmark.

Product	Spatial basis	Descriptors and purpose	Weather linkage and limits
Global LCZ map Demuzere et al. [36]	100 m categorical pixels; neighbourhood typology.	Standard classes of structure, cover, fabric and activity; generic canopy-parameter ranges for comparable urban types.	Inputs for urban-climate models, not a prescribed EPW count. Local within-class variation remains.
UT-GLOBUS Kamath et al. [37]	Individual building footprints/heights; canopy parameters over 1 km^2 at a 300 m step.	Plan fraction, roof-plus-wall surface ratio, area-weighted height and height distribution; continuous model inputs.	WRF-Urban and street-scale applications. Footprint/height uncertainty and definitions need checking before UWG reuse.
GLAMOUR Li et al. [38]	0.0009° pixels, approximately 100 m at the equator; global urban coverage.	Satellite/deep-learning estimates of building fraction and area-weighted height; continuous hydroclimate inputs.	No weather-file selection stage. Retrieval uncertainty; facade and vegetation require other data.
Present framework	100 m assignment cells; overlapping 500 × 500 m GIS supports in Torino.	Building fraction, area-weighted height, facade/site, tree and residual grass cover; local k-medoids classes.	Class-to-UWG-to-EPW link with dry-bulb error checks. Shared template, edge gaps and industrial-tail restriction.

timing perturbations changed the same-medoid annual temperature by at most 0.047°C RMSE, whereas routing all HVAC waste heat from canyon to roof produced a larger response of 0.119°C. For C6, that redistribution reduced summer-night mean temperature by 0.24°C relative to baseline (Supplementary Section S7). This difference is consistent with the direct contribution of canyon-released HVAC heat to the modelled canopy energy balance. There is no evidence that real traffic is unimportant or that HVAC routing uniquely explains the station bias. The baseline already includes a diurnal traffic profile, and road-share scenarios are diagnostic proxies rather than calibrated traffic data. Geometry, heat storage, atmospheric exchange, and parameter interactions remain alternative or interacting explanations that require observation-constrained testing.

The irregular-cell analysis provided methodological evidence concerning the urban design features that a block-based unit can represent and clarified the difference between geometric enclosure and climatic support. In regular street fabrics, the method provides intuitive units with surrounding road space, internal green areas, and facade metrics. Limitations arise when incomplete road topology or extensive open areas require allocation over distances beyond a neighbourhood scale. Statistical clustering then treats a 500 m^2 cell and a 21 km^2 cell as equivalent observations, although their derived ratios have substantially different support meanings.

The standardised support method reduces variable-support artefacts and provides a single, reproducible denominator. It also makes isolated exceptions contextual: a plaza or small park structure contributes to, but does not define, a 500 m neighbourhood. The grid-origin test indicates that the resulting surfaces are not materially controlled by a 50 m shift. This design is transferable because a new location can be characterised with the same support rule without first solving a citywide street-allocation topology. The tradeoff is that a square support is not a meteorological source-area model. It is isotropic, does not vary with wind direction or atmospheric stability, and can cross barriers such as rail corridors or rivers. The 500 m support balances within-class smoothing and retained between-class contrast in the matched morphology test, while the preliminary station experiment showed little aggregate temperature-error sensitivity to support size. Neither test identifies a universal atmospheric footprint. The study advances beyond the closest precedents in a specific way. Palme et al. [9,10] demonstrated morphology categories and a small urban weather-file family; Shen et al. [11] demonstrated automated morphology clustering and UWG; Joshi et al. [17] provided a rigorous morphology-archetype method; Ma et al. [12] localised microclimate with morpho-patches; Hashemi et al. [13] linked LCZs and UWG; and Liu et al. [14] demonstrated scalable block-specific UWG generation. Borelli et al. [34] already connected Torino morphology, UWG, and citywide heat mapping. The methodological contribution is the explicit decoupling of morphology resolution from weather-product resolution, quantified through an error-versus-library-size curve and tested separately for model fidelity and observational accuracy. The public atlas operationalises this distinction by preserving multiple morphological resolutions while returning a single candidate A-D EPW for a queried coordinate. The results also reveal an important limitation of success-oriented reduction workflows: a library can accurately compress modelled differences, while the parent simulator underrepresents measured spatial diversity.

Transfer to another city requires compatible boundaries, footprints, heights and vegetation, local input quality checks, defensible forcing and urban assumptions, and repeated clustering, non-medoid tests and station evaluation. Torino's categories and four-file count do not transfer unchanged. Extraction cost depends on geometry volume and complexity; subset enumeration grows as 2^n-1. For more than 20 candidates, greedy/local-search or optimisation methods may be appropriate, but were not benchmarked here.

The main limitations are as follows.

First, the four-file reduction is based on seven medoid simulations. Within-class sampling found small typical errors but failed in an industrial tail; sample maxima are not citywide error bounds. Second, municipal GIS ends at the administrative boundary. Edge supports were flagged and excluded from fitting, but their published assignments remain provisional. Alenia illustrates this limitation. Third, canopy cover is partly imputed from incomplete tree attributes. The k=8 tree-rich class was therefore not adopted as the primary solution, but tree uncertainty still affects k=7 inputs. Fourth, the facade-to-site ratio is a GIS approximation based on exposed footprint perimeter and average height. It omits courtyard rings and does not explicitly represent directional frontal area, height variability, sky-view factor, or three-dimensional canyon geometry. Learned features [40] were not tested and would require a defensible mapping to UWG inputs. Fifth, all neighbourhoods share one UWG template. This is necessary for a controlled comparison of morphology but omits spatial variation in building use, schedules, anthropogenic heat, materials, traffic, and atmospheric exposure. The accepted Zone 4A and 3.05 m floor height are transparent common choices, not a complete Torino calibration. Sixth, temperature is the only physical basis of the reduction. Standard UWG 5.3.4 does not generate a morphology-dependent humidity ratio in this workflow, and wind, radiation, and precipitation were not used to distinguish EPWs. A weather library intended for other applications may require a multivariable distance. Seventh, the synchronised 2022 comparison uses processed files whose flags do not recover measured versus imputed hours. It is not raw-sensor-only validation or a multi-year uncertainty assessment. Finally, no measured urban station belongs to library categories A or B, and only four urban station contexts are evaluated. The station network, therefore, provides a valuable but incomplete external layer.

Further evaluation should address the failed industrial tail, poorly observed tree subgroup, local building stock and multi-year station behaviour before broadening the applicability claim. An interpretable surrogate can subsequently predict annual and period-specific UWG temperature KPIs, nearest EPW, and expected approximation error from

morphology. Spatial-block cross-validation and out-of-domain warnings are necessary because nearby supports overlap. Such a surrogate would emulate UWG rather than observations, unless station residuals were modelled separately. Additional urban-climate extensions include: (i) testing spatially heterogeneous building stock and activity; (ii) obtaining a compatible external GIS halo; (iii) adding directional frontal area or sky-view factor; (iv) comparing the local classes with LCZs; (v) using raw simultaneous station years for multi-site evaluation; and (vi) investigating nighttime boundary-layer representation without class-specific tuning.

4.1. Humidity and weather-file scope

Physical morphology-dependent urban dry-island behaviour could not be evaluated using the active UWG 5.3.4 workflow. The model preserved the rural forcing humidity ratio in the canyon at each time step. Reconstructed annual Caselle-minus-urban mixing-ratio differences were below 0.001 g/kg and arose from EPW temperature and relative-humidity serialisation. These small residuals are not evidence of resolved moisture removal. Relative humidity decreased in the warmer modelled files, but this was a temperature-driven thermodynamic effect rather than modelled moisture removal. UDI was therefore excluded from cluster selection and EPW reduction. This does not contradict measured UDI in Torino [32]; it identifies a model scope limitation. The EPW field audit found only dry bulb and relative humidity varying between medoids; dew point and wind were modified but shared, and 31 other fields were copied from forcing. The 2022 audit also found relative humidity above 100% and dew point above dry bulb, restricting that experiment to temperature. Temperature compression does not validate solar gains, ventilation, comfort or complete EPW equivalence.

5. Conclusions

This study asked how a morphologically heterogeneous city can be represented by a small set of urban weather files without assuming one file per block or one file per morphology class. The answer required separating three decisions: the unit used to map a location, the area used to characterise neighbourhood morphology, and the number of weather products retained after simulation. An initial irregular block/street allocation method created intuitive cells but also generated oversized and topologically inconsistent contexts at the city scale. The adopted method, therefore, used a deterministic 100 m assignment grid with overlapping 500 m morphology supports. The support was selected after 250/500/1000 m station tests and a 50 m grid-origin shift, and the resulting 12,997-location dataset passed in-house checks for geometry, formulas, range, and provenance. Robust-scaled clustering of 10,593 clean supports identified k=4 as the compact baseline, k=7 as a purpose-dependent, detailed solution with moderate classification robustness, and k=8 as a tree-sensitive, high-resolution alternative. K=7 was retained because it preserved a dense, large-footprint, low-facade fabric absent from k=4. The annual UWG simulation then showed that seven morphology classes did not require seven distinct products. Four representative EPWs, retained as a curve-based fidelity–complexity compromise and checked against error and diminishing-return criteria, reproduced the seven medoid outputs with 0.06°C maximum annual hourly RMSE and 0.03°C weighted mean RMSE at the delivered one-decimal precision. The four urban stations supported the internal reduction but not blanket model validation. The 105 stratified supports gave a weighted 95th-percentile error of 0.10°C, but a 1.33°C industrial-tail failure required a conservative post hoc restriction rather than establishing a validated applicability boundary. Only Consolata improved the annual RMSE over Caselle against the synchronised processed 2022 temperature. The fixed UWG configuration compressed measured inter-station diversity and produced excessive nighttime warming at several sites.

The defensible present conclusion is therefore that Torino can be represented by a four-file candidate library under the tested simulator assumptions and applicability restriction while retaining seven classes for urban-form mapping and interpretation. The accompanying public atlas provides access: a latitude-longitude query returns the local morphology classes, A-D category, representative support characteristics, and candidate EPW where applicable. The k=4, k=7, and k=8 layers also remain available for other urban-microclimate analyses that require a coarser or finer morphology partition. Unconditional citywide sufficiency is not established by the sampling; the atlas withholds location-specific EPW recommendations for the restricted industrial supports and locations with missing inputs. More broadly, the study shows that weather-file compression and meteorological accuracy are separate properties and should be evaluated separately whenever urban weather libraries are derived from a common simulator.

Data and code availability

The scientific morphology and weather-file outputs are archived in the public GitHub repository as release v0.2.0 (commit 451c87db536cc4fa8922668cf32fb290f88e23c7). The atlas interface and geospatial distributions are provided in in release v0.3.0 (https://github.com/alijahanirahaei/torino-urban-epw-atlas/releases/tag/v0.3.0). The four representatives EPWs and exact UWG configurations retain their v0.2.0 filenames and hashes. It contains the citywide k=4, k=7, and k=8 assignments; the retained k=7 classification; the reduced A-D weather assignment; a canonical EPSG:3003 GeoPackage and convenience GeoJSON, shapefile, and CSV distributions; four annual candidate EPWs; the exact UWG JSON configurations used to generate them; cluster diagnostics and figures; checksums, provenance, and validation manifests; and browser and command-line coordinate lookup tools. The measured weather dataset is documented by JahaniRahaei et al. [30]. The repository route is: https://github.com/alijahanirahaei/torino-urban-epw-atlas. It also includes an interactive atlas: https://alijahanirahaei.github.io/torino-urban-epw-atlas/. Repository software is released under the MIT License; the derived spatial data, tables, figures, UWG configurations, and generated EPWs are released under the Creative Commons Attribution 4.0 International License (CC BY 4.0). Redistribution of source municipal GIS and measured station files remains subject to the rights stated by their original providers. Appendix A briefly summarises the repository contents. The release uses C1/C2/C4/C6 as A-D representatives and includes seven named GeoPackage layers: grid_assignments, cluster_medoids, city_boundary, morphology_k4, morphology_k7, morphology_k8 and weather_categories. The repository and atlas downloads are accessible without authentication. Fixed-classifier reconstruction, seven-medoid reduction, and annual UWG execution are supported by the supplied features, configurations, scripts, and provenance.

Supplementary methods and results (Sections S1–S7, Tables S1–S13 and Figures S1–S6), model configurations and heat-sensitivity reproduction files are available in a separately versioned supplementary release v1.1.0: https://github.com/alijahanirahaei/torino-urban-epw-atlas/releases/tag/supplement-v1.1.0. These materials document execution, provenance and interoperability limits.

Declaration of generative AI and AI-assisted technologies in the manuscript preparation process

The authors used ChatGPT and Codex (OpenAI) to assist with literature synthesis, manuscript organization, drafting, and language editing. All AI-assisted material was critically reviewed and revised by the authors, who checked the cited sources and scientific claims and remain fully accountable for the article.

CRediT authorship contribution statement

Ali JahaniRahaei: Writing – review & editing, Writing – original

draft, Software, Methodology, Data curation. **Giacomo Chiesa:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition.

Declaration of competing interest

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

Acknowledgment

The authors thank the Regional Agency for the Protection of the Environment of Piedmont Region (ARPA Piemonte) for the weather station data used in this paper.

This research has received funding under the “P2022AEN9Y URBAN GENERATION project – A design approach for climate change adaptation and mitigation” – (PRIN PNRR 2022 program - CUP D53D23020090001), funded by European Union - NextGeneration EU "MUR – Ministero dell’Università e della Ricerca “Bando PRIN 2022 PNRR” - D.D. n. 1409 del 14-9-2022".

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.buildenv.2026.115528.

Appendix A. Public atlas and reusable research package

The public atlas operationalises the paper's two output levels (Figs. A1 and A2). A user may select a point on the map or enter latitude and longitude. The interface returns the 100 m assignment identifier; k=4, k=7, and k=8 morphology classes; the assigned A-D weather category; six principal 500 m-support descriptors; and edge or outlier quality flags. For eligible locations, the selected annual EPW and its exact UWG JSON configuration can be downloaded directly, and a result-specific link can be shared. The displayed UHI values carry an explicit asterisk because they are UWG-simulated differences from AvMY Caselle rather than measured local UHI. The lookup withholds EPW and configuration recommendations for 20 high-density C6 supports among 12,862 assigned locations; a further 135 locations have missing inputs. This is a conservative post hoc guardrail motivated by the industrial-tail failure, not a physically validated domain boundary. An additional 19 C6 supports between the strict-fitting 90th and 95th building-fraction percentiles carry a near-threshold caution flag while retaining candidate downloads. Membership below the restriction does not establish validated coverage. Categories A and B have no urban evaluation station, so their observational accuracy remains untested. Municipal-edge and other quality-flagged assignments remain provisional. Figs. A1 and A2 show release v0.3.0 with its morphology layers and candidate-library downloads.

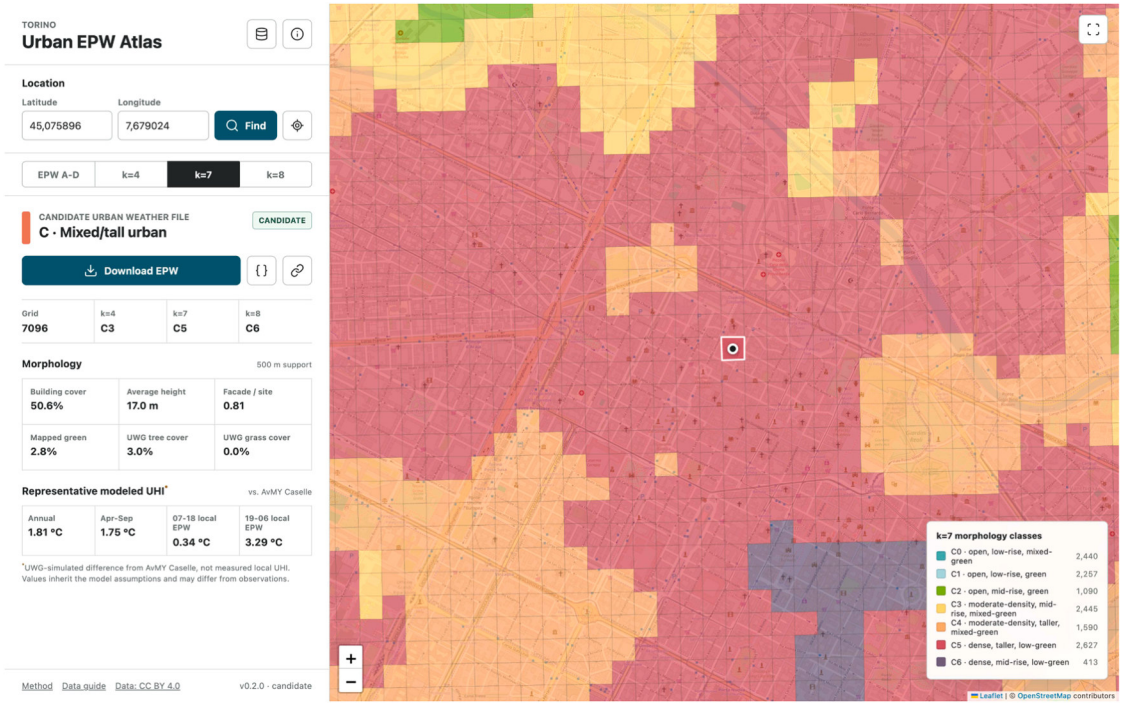


Fig. A1. Atlas desktop interface (v0.3.0) after a coordinate query at Consolata. The panel reports the candidate weather category, multiple morphology resolutions, 500 m support descriptors, quality status, direct EPW and UWG configuration downloads, and UWG modelled UHI relative to Caselle. The asterisk distinguishes modelled values from measured local UHI.

For GIS and reproducibility use, the repository distributes a canonical EPSG:3003 GeoPackage containing 12,997 grid assignments, the 19 medoids across k=4, k=7, and k=8, and the municipal study boundary. WGS84 GeoJSON, EPSG:3003 shapefiles, non-spatial CSV tables, cluster profiles and selection diagnostics, figures, checksums, build scripts, and validation tests are also provided. The morphology layers can therefore support studies that do not require an EPW, although classifications optimised for UWG-relevant thermal morphology may not be optimal for every wind, pollution, radiation, hydrological, or outdoor-comfort question. Code and research outputs are licensed under the dual MIT and CC BY 4.0 licenses stated above.

The GeoPackage contains assignment, medoid and boundary layers, plus separate k=4, k=7, k=8 and weather-category layers. The research directory supplies the twelve-setting HDBSCAN results, buffered spatial-stability summaries, feature tests and non-medoid sampling and fidelity tables. These files contain descriptors and assignments, not building geometries or a Dragonfly model; Python was the production UWG route, while Grasshopper supported preliminary exploration. From the supplied features and fixed classifier, users can reconstruct the exports and repeat the seven-medoid reduction; exact JSONs and a hash-checked runner support annual UWG reproduction with the documented source EPW. Reconstruction from raw municipal GIS, complete clustering refits and a Dragonfly import benchmark are outside the package's scope (Supplementary Section S6). The supplementary package is provided in the repository's research/supplementary/ directory and as a downloadable ZIP in release supplement-v1.1.0. It includes the supplementary document in PDF and Word formats, Tables S1–S13 as CSV files, Figures S1–S6, exact UWG configurations, heat-sensitivity results and a reusable simulation script. Tables S1, S2 and S10–S12 are also reproduced in the supplementary document; the remaining tables are provided as companion CSVs. A README and file index describe the contents, while checksums enable file-integrity verification.

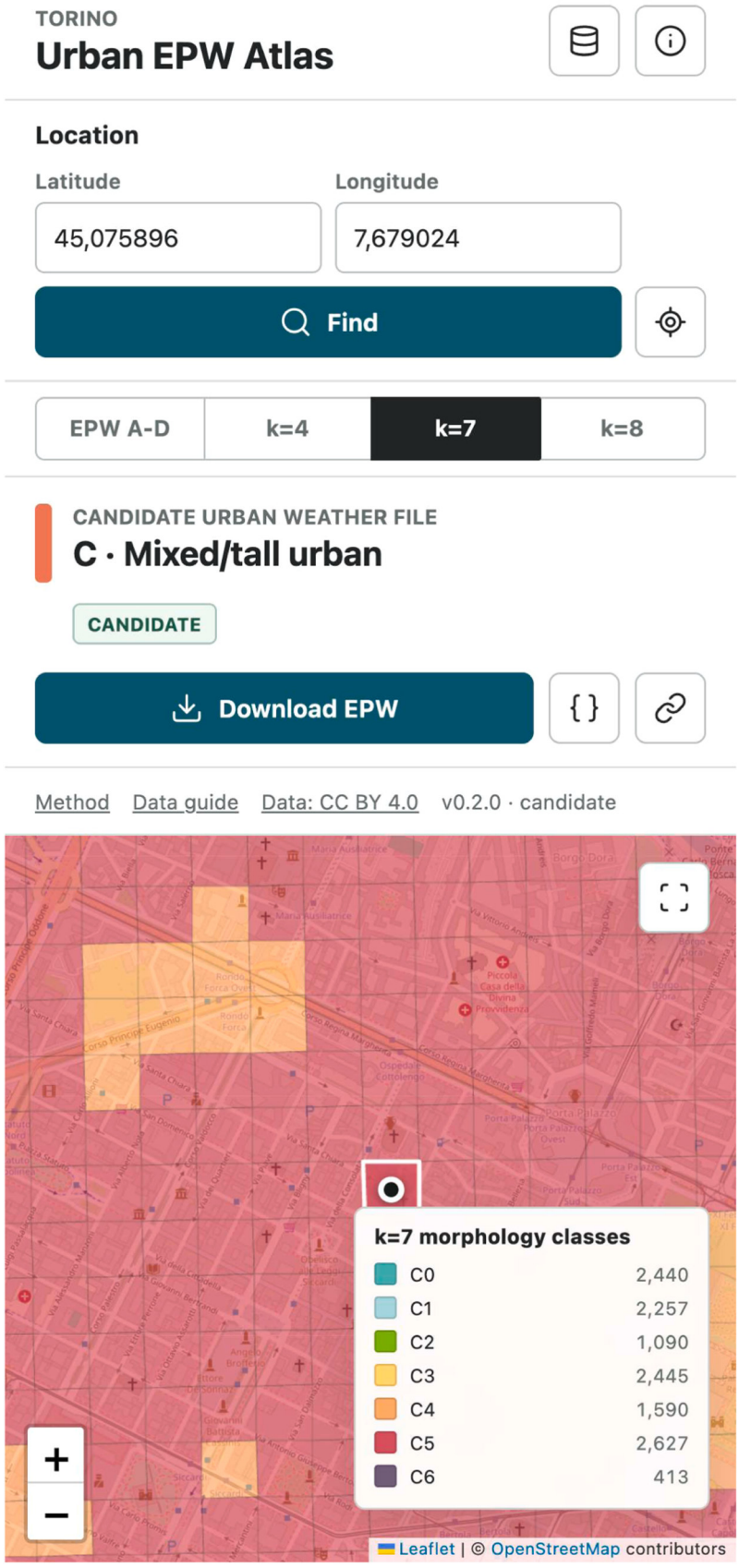


Fig. A2. Atlas responsive mobile interface (v0.3.0) showing coordinate lookup, selected category, EPW and UWG-configuration controls, and the k=7 morphology layer.

Data availability

Data will be made available on request.

References

[1] A.J. Arnfield, Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island, Int. J. Climatol. 23 (2003) 1–26, https://doi.org/10.1002/joc.859.

[2] T.R. Oke, The energetic basis of the urban heat island, Q. J. R. Meteorol. Soc. 108 (1982) 1–24, https://doi.org/10.1002/qj.49710845502.

[3] I.D. Stewart, T.R. Oke, Local Climate Zones for urban temperature studies, Bull. Am. Meteorol. Soc. 93 (12) (2012) 1879–1900, https://doi.org/10.1175/BAMS-D-11-00019.1.

[4] D. Fenner, F. Meier, B. Bechtel, M. Otto, D. Scherer, Intra and inter Local Climate Zone variability of air temperature as observed by crowdsourced citizen weather stations in Berlin, Germany, Meteorol. Z. 26 (5) (2017) 525–547, https://doi.org/10.1127/metz/2017/0861.

[5] M. Lehnert, J. Geletic, P. Dobrovolny, M. Jurek, Temperature differences among local climate zones established by mobile measurements in two Central European cities, Clim. Res. 75 (2018) 53–64, https://doi.org/10.3354/cr01508.

[6] B. Bueno, L. Norford, J. Hidalgo, G. Pigeon, The urban weather generator, J. Build. Perform. Simul. 6 (4) (2013) 269–281, https://doi.org/10.1080/19401493.2012.718797.

[7] B. Bueno, J. Hidalgo, G. Pigeon, L. Norford, V. Masson, Calculation of air temperatures above the urban canopy layer from measurements at a rural operational weather station, J. Appl. Meteorol. Climatol. 52 (2) (2013) 472–483, https://doi.org/10.1175/JAMC-D-12-083.1.

[8] B. Bueno, M. Roth, L. Norford, R. Li, Computationally efficient prediction of canopy level urban air temperature at the neighbourhood scale, Urban Clim. 9 (2014) 35–53, https://doi.org/10.1016/j.uclim.2014.05.005.

[9] M. Palme, L. Inostroza, G. Villacreses, A. Lobato-Cordero, C. Carrasco, From urban climate to energy consumption. Enhancing building performance simulation by including the urban heat island effect, Energy Build. 145 (2017) 107–120, https://doi.org/10.1016/j.enbuild.2017.03.069.

[10] M. Palme, L. Inostroza, G. Villacreses, A. Lobato, C. Carrasco, Urban weather data and building models for the inclusion of the urban heat island effect in building performance simulation, Data Br. 14 (2017) 671–675, https://doi.org/10.1016/j.dib.2017.08.035.

[11] P. Shen, J. Liu, M. Wang, Fast generation of microclimate weather data for building simulation under heat island using map capturing and clustering technique, Sustain. Cities Soc. 71 (2021) 102954, https://doi.org/10.1016/j.scs.2021.102954.

[12] R. Ma, T. Wang, Y. Wang, J. Chen, Tuning urban microclimate: a morpho-patch approach for multi-scale building group energy simulation, Sustain. Cities Soc. 76 (2022) 103516, https://doi.org/10.1016/j.scs.2021.103516.

[13] F. Hashemi, G. Mills, U. Poerschke, L.D. Iulo, G. Pavlak, L. Kalisperis, A novel parametric workflow for simulating urban heat island effects on residential building energy use: coupling local climate zones with the Urban Weather Generator, a case study of seven U.S. cities, Sustain. Cities Soc. 110 (2024) 105568, https://doi.org/10.1016/j.scs.2024.105568.

[14] K. Liu, X. Xu, D. Lin, R. Zhang, L. Zhao, A. Abuduwayiti, F. Causone, A scalable and efficient framework for city-scale building energy modeling with microclimate considerations, Sustain. Cities Soc. 138 (2026) 107187, https://doi.org/10.1016/j.scs.2026.107187.

[15] C. Ratti, D. Raydan, K. Steemers, Building form and environmental performance: archetypes, analysis and an arid climate, Energy Build. 35 (1) (2003) 49–59, https://doi.org/10.1016/S0378-7788(02)00079-8.

[16] G. Chiesa, Y. Li, Including Urban Heat Island in bioclimatic early-design phases: a simplified methodology and sample applications, Sustainability 13 (11) (2021) 5918, https://doi.org/10.3390/su13115918.

[17] M.Y. Joshi, A. Rodler, M. Musy, S. Guernouti, M. Cools, J. Teller, Identifying urban morphological archetypes for microclimate studies using a clustering approach, Build. Environ. 224 (2022) 109574, https://doi.org/10.1016/j.buildenv.2022.109574.

[18] M.Y. Joshi, A. Rodler, M. Musy, S. Guernouti, J. Teller, Influence of urban morphology on potential of green roofs in regulating local microclimate: a case study of Liege, Belgium, Urban Clim. 58 (2024) 102144, https://doi.org/10.1016/j.uclim.2024.102144.

[19] P.J. Rousseeuw, Silhouettes: a graphical aid to the interpretation and validation of cluster analysis, J. Comput. Appl. Math. 20 (1987) 53–65, https://doi.org/10.1016/0377-0427(87)90125-7.

[20] Ben-Hur, A., Elisseeff, A., & Guyon, I. (2002). A stability based method for discovering structure in clustered data. https://doi.org/10.1142/9789812799623_0002.

[21] E. Schubert, P.J. Rousseeuw, Fast and eager k-medoids clustering: O(k) runtime improvement of the PAM, CLARA, and CLARANS algorithms, Inf. Syst. 101 (2021) 101804, https://doi.org/10.1016/j.is.2021.101804.

[22] A. Middel, J. Lukasczyk, S. Zakrzewski, M. Arnold, R. Maciejewski, Sky view factor footprints for urban climate modeling, Urban Clim. 25 (2018) 120–134, https://doi.org/10.1016/j.uclim.2018.05.004.

[23] C.L. Muller, L. Chapman, C.S.B. Grimmond, D.T. Young, X. Cai, Toward a standardized metadata protocol for urban meteorological networks, Bull. Am. Meteorol. Soc. 94 (8) (2013) 1161–1185, https://doi.org/10.1175/BAMS-D-12-00096.1.

[24] F. Biljecki, Y.S. Chow, Global building morphology indicators, Comput., Environ. Urban Systems 95 (2022) 101809, https://doi.org/10.1016/j.compenvurbsys.2022.101809.

[25] V. Masson, W. Heldens, E. Bocher, M. Bonhomme, C. Burmeister, C. de Munck, et al., City-descriptive input data for urban climate models: model requirements, data sources and challenges, Urban Clim. 31 (2020) 100536, https://doi.org/10.1016/j.uclim.2019.100536.

[26] R. Hamdi, C. Lallemand, J. Le Bras, V. Masson, Evaluation of the Urban Weather Generator on the city of Toulouse (France), Appl. Sci. 14 (1) (2024) 185, https://doi.org/10.3390/app14010185.

[27] J. Mao, J.H. Yang, A. Afshari, L.K. Norford, Global sensitivity analysis of an urban microclimate system under uncertainty: design and case study, Build. Environ. 124 (2017) 153–170, https://doi.org/10.1016/j.buildenv.2017.08.011.

[28] J. Mao, Y. Fu, A. Afshari, P.R. Armstrong, L.K. Norford, Optimization-aided calibration of an urban microclimate model under uncertainty, Build. Environ. 143 (2018) 390–403, https://doi.org/10.1016/j.buildenv.2018.07.034.

[29] A. Salvati, M. Palme, L. Inostroza, Key parameters for urban heat island assessment in a Mediterranean context: a sensitivity analysis using the Urban Weather Generator model, IOP Conf. Ser.: Mater. Sci. Eng. 245 (2017) 082055, https://doi.org/10.1088/1757-899X/245/8/082055.

[30] A. JahaniRahaei, M. Milelli, G. Chiesa, Urban weather dataset for building energy simulations: data collection and EPW file generation for Torino, Italy (2014-2023), Data Br. 61 (2025) 111708, https://doi.org/10.1016/j.dib.2025.111708.

[31] F. Bassani, V. Garbero, D. Poggi, L. Ridolfi, J. von Hardenberg, M. Milelli, An innovative approach to select urban-rural sites for Urban Heat Island analysis: the case of Turin (Italy), Urban Clim. 42 (2022) 101099, https://doi.org/10.1016/j.uclim.2022.101099.

[32] M. Milelli, F. Bassani, V. Garbero, D. Poggi, J. von Hardenberg, L. Ridolfi, Characterization of the urban heat and dry island effects in the Turin metropolitan area, Urban Clim. 47 (2023) 101397, https://doi.org/10.1016/j.uclim.2022.101397.

[33] D. Garzena, F. Acquaotta, S. Fratianni, Analysis of the long-time climate data series for Turin and assessment of the city's urban heat island, Weather 74 (10) (2019) 353–359, https://doi.org/10.1002/wea.3292.

[34] Borelli, G., Ballarini, I., Corrado, V., Gasparella, A., & Pernigotto, G. (conference 2024; proceedings record 2025). Assessment and mapping of the urban heat island effect: a preliminary analysis on the impact on urban morphology for the City of Turin, Italy. *Proceedings of Building Simulation Applications BSA 2024*, 525-531. https://doi.org/10.13124/9788860462022_64.

[35] Ladybug Tools, Urban Weather Generator (UWG), version 5.3.4 [Computer software]. https://pypi.org/project/uwg/5.3.4/, 2020.

[36] M. Demuzere, J. Kittner, A. Martilli, G. Mills, C. Moede, I.D. Stewart, J. van Vliet, B. Bechtel, A global map of local climate zones to support earth system modelling and urban-scale environmental science, Earth Syst. Sci. Data 14 (2022) 3835–3873, https://doi.org/10.5194/essd-14-3835-2022.

[37] H.G. Kamath, M. Singh, N. Malviya, A. Martilli, L. He, D. Aliaga, C. He, F. Chen, L. A. Magruder, Z.-L. Yang, D. Niyogi, GLObal Building heights for Urban Studies (UT-GLOBUS) for city- and street-scale urban simulations: development and first applications, Sci. Data 11 (2024) 886, https://doi.org/10.1038/s41597-024-03719-w.

[38] R. Li, T. Sun, S. Ghaffarian, M. Tsamados, G. Ni, GLAMOUR: GLobAl building MOrphology dataset for URban hydroclimate modelling, Sci. Data 11 (2024) 618, https://doi.org/10.1038/s41597-024-03446-2.

[39] D. David, Nighttime urban boundary layer height in the Urban Weather Generator: analysis and dynamic modelling, Urban Clim. 65 (2026) 102797, https://doi.org/10.1016/j.uclim.2026.102797.

[40] P. Patel, R. Kalyanam, L. He, D. Aliaga, D. Niyogi, Deep learning-based urban morphology for city-scale environmental modeling, PNAS Nexus 2 (3) (2023) pgad027, https://doi.org/10.1093/pnasnexus/pgad027.