Doctoral Dissertation
Doctoral Program in Energy Engineering (38th Cycle)

Energy Performance Assessment of Building Stocks

Enhancing Quality Data Flows for Urban Energy Analysis

Matteo Piro

* * * * * *

Supervisors:
Prof. Ilaria Ballarini, Supervisor
Prof. Vincenzo Corrado, Co-Supervisor

Doctoral Examination Committee:

Prof. Pamela Fennell, University College London
Dr. Jacopo Vivian, University of Padua
Prof. Cristina Becchio, Politecnico di Torino

Politecnico di Torino
2026

I hereby declare that the contents and organisation of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

…………………………………......
Matteo Piro
2026

Abstract

The National Building Renovation Plan established under the newly approved Energy Performance of Buildings Directive (EU 2024/1275) can be effectively supported by combining Urban Building Energy Modelling (UBEM) with building archetyping. However, the limited availability and accuracy of city-wide energy consumption data, together with the high computational cost of large-scale simulations, hinder the validation and calibration of most UBEM applications.

To address these challenges, this Ph.D. research develops a comprehensive methodology to assess and improve the reliability of the entire UBEM workflow, from input data to model outputs.

The research first establishes a theoretical comparison between Building Energy Models (BEMs) and UBEMs, illustrating how simplifications inherited from single-building modelling affect geometry, envelope characterisation, zoning, internal gains, and technical building systems definition. It then examines the limitations of input data, focusing on a very common data source in UBEM as the Energy Performance Certificates (EPCs). Despite their widespread use, data quality issues arise. A dedicated data quality-checking procedure was developed and applied to regional EPC datasets, demonstrating the benefit of automated rules to detect unreliable or non-representative entries.

Reliable EPCs were subsequently used to generate representative buildings according to key segmentation criteria, emphasising the importance of non-geometric attributes in archetype-based UBEM approaches. The work also introduces a shift from representative buildings to representative urban blocks, which were identified and analysed under different climate scenarios.

Finally, the methodology integrates validation and calibration procedures tailored to data availability and modelling goals. This includes a workflow for assessing archetype reliability and an evolutionary optimisation approach for

calibrating in an UBEM environment single-building and multi-building energy demands.

Overall, this Ph.D. delivers an integrated framework that strengthens the credibility and practical applicability of UBEMs. By combining data quality assessment, representative building modelling, and scalable validation and calibration techniques, the work is aimed at supporting energy planners and policymakers in capturing the actual energy performance of the building stock and in designing effective renovation strategies and climate mitigation actions.

Acknowledgment

My sincere gratitude goes to my supervisors, professors Ballarini and Corrado, for providing guidance and feedback throughout the entire process of my research.

I would also like to acknowledge Dr. Kämpf for his vision, support, and for hosting me at the Idiap Research Institute in Martigny; working together has been a truly inspiring experience.

To my grandparents

Contents

List of Tables

List of Figures

Nomenclature

Quantities

A	area	m^2
ABD	average building distance	m
ABH	average building height	m
AWD	ambient warmness degree	°C
c	specific heat capacity	$J \cdot kg^{-1} \cdot K^{-1}$
CDD	cooling degree days	°C·d
COP	coefficient of performance	–
CR	compactness ratio	m^{-1}
CV(RMSE)	coefficient of variation of the root-mean-square error	–
d	thickness	m
E	electrical energy	kWh
EER	energy efficiency ratio	–
EP	energy performance	$kWh \cdot m^{-2}$
F or *f*	factor, fraction	–
FAR	floor area ratio	–
g	total solar energy transmittance	–
GR	green ratio	–
H	building height	m
HDD	heating degree days	°C·d
K	CO_2 emission coefficient	$kg \cdot kWh^{-1}$
L	length	m
LST	land surface temperature	°C
MBE	mean bias error	–
n	air change rate	1/h
NMBE	normalised mean bias error	–
PRH	percentage of representativeness hours	–
Q	areal or volumetric amount of heat	$kWh \cdot m^{-2}$ $kWh \cdot m^{-3}$
q	ventilation volumetric flow rate	$l \cdot s^{-1}$
R	thermal resistance	$m^2 \cdot K \cdot W^{-1}$
REC	relative compactness	–
RER	renewable energy ratio	–

SC	surface coverage	–
SCOP	seasonal coefficient of performance	–
SEER	seasonal energy efficiency ratio	–
SF	shape factor	m^{-1}
SVF	sky view factor	–
t	time	h d
U	thermal transmittance	$W \cdot m^{-2} \cdot K^{-1}$
UHII	urban heat island intensity	°C
V	volume	m^3
VAR	volume to area ratio	m
VRH	validity of representativeness hours	–
VtH	vertical to horizontal ratio	–
W	power	kW
WWR	window-to-wall ratio	–

Greek symbols

a	absorptance	–
α	azimuth angle	°
β	tilt angle	°
ε	long-wave emissivity of the surface	–
η	efficiency	–
ϑ	temperature	°C
λ	latitude	°
λ	thermal conductivity	$W \cdot m^{-1} \cdot K^{-1}$
μ	vapour resistance	–
ϕ	heat flow rate, heat load, power	W
κ	areal heat capacity	$J \cdot m^{-2} \cdot K^{-1}$
ρ	density	$kg \cdot m^{-3}$
ρ	reflectance	–
τ	transmittance	–
ψ	longitude	°

Subscripts

A	appliances	m	mean
a	air	n	net
bldg	building	nd	need
C	cooling	ngen	without generation
c	control	nren	non-renewable
cv	convective	obst	obstacles
CO2e	CO_2 emission equivalent	oc	occupants
coll	solar collector	op	opaque
d	distribution	pers.	occupant

e	emission	PV	photovoltaic
es	solar energy	Q_1, Q_3	first and third quartile
env	envelope	ren	renewable
fl	floor	s	seasonal
fr	frame	se	surface external
g	gross	sens	sensible
gl	glazing, glazed element	set	set-point
gl	overall	sh	shading
gn	generation	si	surface internal
H	heating	sol	solar
in	input, inlet	sto	storage
int	internal or indoor	tot	total
IQR	interquartile range	up	upper
L	lighting	use	useful
lat	latent	v	visible
ld	load	W	domestic hot water
lim	limit	w	window
lw	lower	wl	wall
m	mass related conductance or capacitance		

Acronyms

AB	Apartment Block
BA	Building Archetype
BEM	Building Energy Model (Modelling)
BEPS	Building Energy Performance Simulation
BIM	Building Information Model (Modelling)
BU	Building Unit
CEN	European Committee for Standardization
CFD	Computational Fluid Dynamics
CMA-ES	Covariance Matrix Adaptation Evolution Strategy
CP	Construction Period
CTI	Italian Thermotechnical Committee
DHW	Domestic Hot Water
EA	Evolutionary Algorithm
EPBD	Energy Performance of Buildings Directive
EPC	Energy Performance Certificate
EUI	Energy Use Intensity
HDE	Hybrid Differential Evolution
ISO	International Standardization Organization
LoD	Level of Detail
LW	Long-wave
KPI	Key Performance Indicator
ML	Machine Learning
NBRP	National Building Renovation Plan
REN	Recommended Energy Efficiency Measure

RWS	Rural Weather Station
SFH	Single-family house
SW	Short-wave
TBS	Technical Building System
TMY	Typical Meteorological Year
UBEM	Urban Building Energy Model (Modelling)
UCL	Urban Canopy Layer
UES	Urban Energy System
UHI	Urban Heat Island
UNI	Italian Organisation for Standardisation
USEM	Urban-Scale Energy Model
UWS	Urban Weather Station

Chapter 1

Introduction

The global population currently stands at approximately 7.8 billion and is projected to reach 10 billion in 2050. At present, 58 % of the world's population lives in urban areas, while 42 % resides in rural regions. However, the share of the urban population is expected to increase to 68 % within the next three decades. In Europe, despite a projected 3.7 % decline in total population between 2020 and 2050, the proportion of people living in urban areas is forecast to rise from 75 % in 2020 to 84 % in 2050 (Ritchie et al., 2024). Increasing urbanisation leads to a corresponding intensification of human activities and associated energy consumption. Consequently, effective urban energy planning becomes essential, requiring approaches that integrate sustainable development strategies and quantify the impacts of related policies.

To meet the European Union's 2030 and 2050 greenhouse gas emission reduction targets (European Commission, 2019), mitigation actions increasingly need to be planned and assessed beyond the single-building scale. This has strengthened the role of urban energy planning, which requires consistent, district-scale estimates of energy demand and the potential impact of retrofit and policy measures. In this context, Urban Building Energy Models (UBEMs) have emerged as a useful framework to analyse groups of buildings within a unified modelling approach. As a result, UBEM has become a prominent topic within the scientific community. City-scale energy modelling represents a critical tool for urban planners, public administrations, energy agencies, and validation bodies, as it supports the design, assessment, and prioritisation of energy efficiency actions for the building stock.

The National Building Renovation Plans (NBRPs), introduced under the revised Energy Performance of Buildings Directive (EPBD) (EU 2024/1275), provide a strategic framework for Member States to decarbonise the building stock. These plans define long-term pathways, policies, and measures aimed at improving energy efficiency, reducing greenhouse gas emissions, and supporting the transition towards a zero-emission building sector.

The credibility of NBRPs is compromised when large-scale energy models remain unvalidated and uncalibrated. However, the availability, accessibility, and quality of urban-scale data pose significant challenges to the accurate digital representation of building stock energy performance. In particular, the lack of city-wide building energy consumption data with sufficient spatial and temporal resolution constitutes a major technical barrier to the validation and calibration of large-scale model outputs. Consequently, without reliable data and modelling frameworks, the accuracy, robustness, and practical usefulness of policy and planning decisions based on UBEM results cannot be guaranteed.

This Ph.D. research aims to analyse the evolution of UBEMs, with particular focus on the increase of input data reliability, methodological approaches, and validation and calibration strategies, in order to identify methods, practices, and tools capable of supporting a more informed and reliable urban energy planning process. In this context, the research addresses the following research questions:

RQ1) What methodologies are currently adopted for the classification and development of building stock energy models, and what are the main challenges hindering their reliable implementation in urban planning?

RQ2) What simplifications and assumptions characterise UBEMs compared to traditional Building Energy Models (BEMs), and how do they affect the accuracy of energy simulations?

RQ3) How can the quality of city-scale energy data be improved to support aware urban energy planning?

RQ4) How can urban energy data effectively and representatively inform UBEMs?

RQ5) How do validation and calibration procedures contribute to reducing deviations between simulated outputs and measured energy consumption at the urban scale?

To address these research questions, this Ph.D. research aims to enhance the reliability of building stock energy performance assessment across the entire UBEM workflow. First, a critical literature review is conducted to classify the main methodological approaches adopted in building stock energy modelling and to identify the key challenges that currently hinder their reliable implementation in urban energy planning. Then, a systematic comparison of the assumptions and simplifications commonly adopted in UBEM tools is conducted against those used in detailed single-building energy models. Subsequently, a data quality-checking procedure applied to energy performance certificates is employed to generate *reliable* representative buildings, which serve as fundamental elements of an archetype-based UBEM approach. In this context, building archetypes are used as a trade-off strategy to mitigate input uncertainty while maintaining computational feasibility. Furthermore, depending on the objectives of the analysis, the availability of monitoring data, and the number of simulation iterations, both validation and calibration procedures are proposed to improve the accuracy of large-scale energy model outputs.

Overall, the interrelated research topics addressed within this Ph.D. framework are synthesised and visually presented in Figure 1.

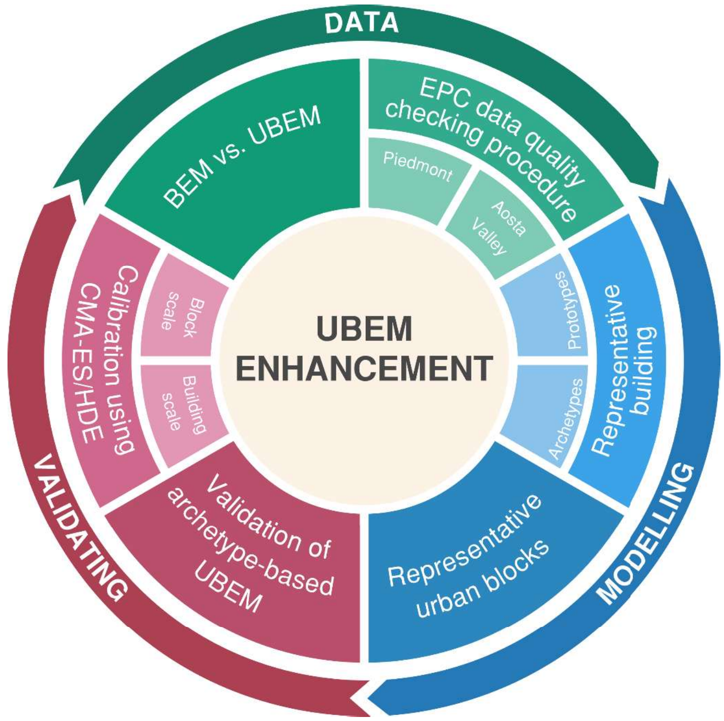


Figure 1. *Methodology workflow of the Ph.D. research*

Chapter 2 presents a literature review of the key factor influencing Urban Building Energy Modelling. It discusses the advantages and limitations of different building stock energy modelling approaches and examines city-scale uncertainties, highlighting the importance of collecting multi-domain urban data. The chapter then introduces the concept of representative buildings as a strategy to reduce the complexity of large-scale energy modelling while maintaining accuracy. Subsequently, it reviews studies assessing the reliability of UBEM outcomes through validation and calibration procedures. Finally, additional challenges and future research directions in UBEM are discussed.

Chapter 3 provides a theoretical comparison between BEMs and UBEMs, outlining the common simplifications and assumptions adopted in urban energy modelling.

Chapter 4 presents the data quality-checking procedure developed to assess and rank the reliability of energy certificates.

Chapter 5 introduces the evolution from the concept of a representative building to that of a representative urban block.

Chapter 6 describes a validation and calibration procedures aimed at enhancing the credibility of large-scale energy models.

Chapter 7 summarises the main findings of this Ph.D. research.

Chapter 2

State of the art

2.1 Overview

Part of the work described in this chapter has been previously published in an international peer-reviewed journal (Piro, Kämpf, et al., 2026), a book chapter (Ballarini et al., 2021), and a project deliverable (Ballarini et al., 2022). In Ballarini et al. (2021), the Ph.D. research focused on reviewing existing approaches for building stock energy modelling. Moreover, the contents of Ballarini et al. (2022) were revised and further elaborated to highlights the gaps, limitations, and technical barrier associated with current energy performance certificates. For the journal publication (Piro, Kämpf, et al., 2026), the specific contributions of the author are detailed in the corresponding CRediT authorship statement.

The literature review presented addresses key elements supporting the development of NBRPs, as outlined in the EPBD IV recast (EU 2024/1275). A fundamental component of these plans is the assessment of the energy performance of the existing building stock, while accounting for the diversity of building typologies.

Specifically, this chapter presents the approaches adopted for large-scale energy analysis and outlines the research challenges associated with UBEM, where uncertainties and simplifications in describing the phenomena arising within the urban environment increase significantly. In particular, special attention is devoted to the technical and social barriers associated with collecting large amounts of data for the development of UBEMs. Such data can be channelled into the creation of representative buildings, intended as a tool for reducing uncertainty in model inputs. Furthermore, the implementation of UBEMs is strongly hindered by the limited availability of validated and calibrated models, mainly due to the difficulties associated with collecting measured energy consumption data at the city scale.

To ensure the credibility and effectiveness of such models, this chapter examines several interrelated aspects: different building stock energy modelling approaches (Section 2.2), classification of uncertainties in UBEM (Section 2.3), the reliability and integration of multi-source urban data (Section 2.4), the use of representative buildings (Section 2.5), and strategies for large-scale validation and calibration procedures (Section 2.6), and further challenges in large-scale energy modelling. Finally, the last Section (2.8) is dedicated to final remarks.

2.2 Building stock energy modelling approaches

According to the scientific literature, building stock energy models are generally grouped into two methodological categories: *top-down* and *bottom-up* approaches (Abbasabadi & Mehdi Ashayeri, 2019; Johari et al., 2020; Kavgic et al., 2010; Swan & Ugursal, 2009). These approaches differ in terms of data aggregation, computational demand, level of detail, and intended application. More recently, Hong et al. (2025), through a ten-questions paper, highlighted several critical gaps building stock energy modelling research, emphasising the need for clearer methodological frameworks and more transparent modelling assumptions.

The choice between *top-down* and *bottom-up* approaches depends on the objectives of the analysis, the relationships among the considered variables, and the level of detail required for the input data. In general, *top-down* models are primarily employed to forecast future energy scenarios using historical time-series data at aggregated scales. *Bottom-up* models, instead, focus on characterising current conditions and exploring future intervention strategies, such as assessing or enhancing the energy performance of specific buildings. While *top-down* modelling offers a broad overview of the system with limited granularity, *bottom-up* modelling describes individual components, resulting in a more detailed understanding of the phenomenon under investigation. Figure 2 provides the binary classification schema of the building stock energy models.

However, recent advancements in big data accessibility, increased computing power, and the proliferation of modelling techniques have highlighted the limitations of this traditional binary classification. To address these gaps, Langevin et al. (2020), within the framework of IEA-EBC Annex 70, proposed a more up-to-date and nuanced framework that moves away from a hierarchical structure in favour of a flexible four-quadrant scheme, shown in Figure 3. Instead of adopting a purely hierarchical distinction, the framework organises building stock energy models within a four-quadrant scheme based on two main dimensions: the modelling design (*top-down* vs. *bottom-up*) and the degree of transparency (*black-box* vs. *white-box*).

A unique feature of this framework is its multi-layer organisation, which recognises that a building stock energy model must often represent multiple determinants of energy use. Beyond the primary “*Energy*” layer, the model should explicitly or implicitly define how it handles the occupant behaviour (“*People*”), physical characteristics ("*Building stock*"), and climatological and socio-economic context ("*Environment*") layers.

Furthermore, the framework emphasises four additional dimensions that are critical for defining the scope and interpretability of a study: *system boundaries* (spatial and temporal extent), *spatio-temporal resolution* (the unit of observation and time step), *dynamics* (how the stock and behaviours evolve over time), and *uncertainty*. Specifically, the temporal resolution (e.g., hourly vs. annual) determines the model's suitability for different use cases, such as investigating demand-side flexibility versus long-term renovation strategies.

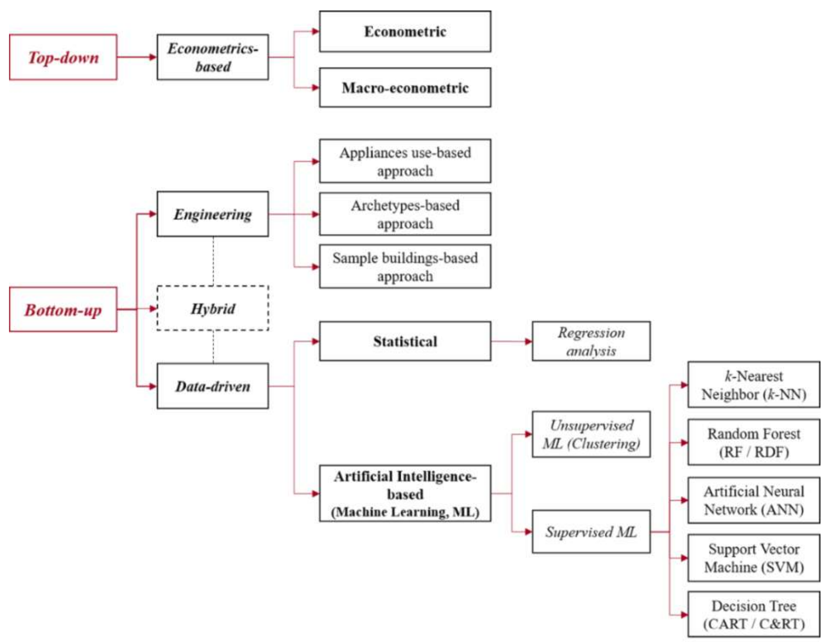


Figure 2. *Binary classification of the building stock energy models (Ballarini et al., 2021)*

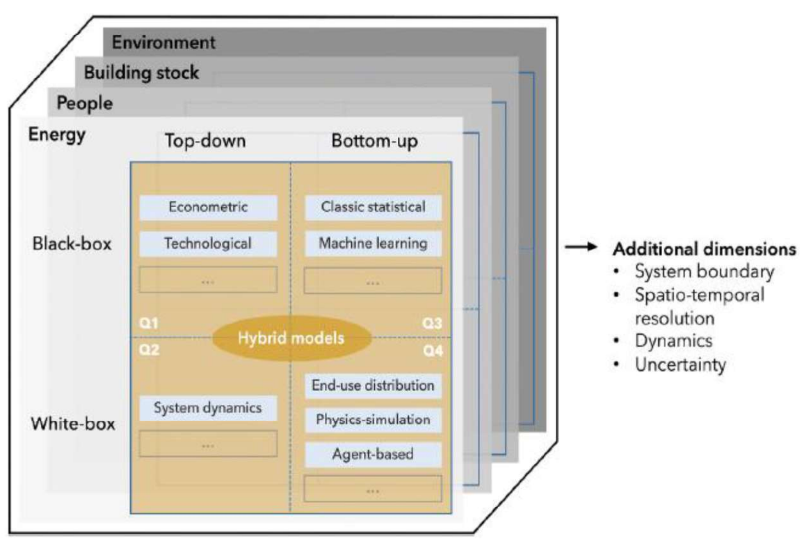


Figure 3. *Four-quadrant classification of building stock energy modelling (Langevin et al., 2020)*

In light of these classifications, the following sections provide an overview of *top-down* and *bottom-up* building stock energy models, discussing their main features, applications, strengths, and limitations as highlighted in the literature.

2.2.1 Top-down models

In general, the *top-down* approach addresses the analysis of complex systems by breaking their output into a finite set of components, allowing the identification of the interconnections among them. It relies on representing the system through a predefined number of “black boxes”, demonstrating mathematically that a given outcome can be reproduced without explicitly modelling the physical laws governing the underlying phenomena. The examination of the resulting output then enables the identification of missing resources or dependencies needed to fully describe the system.

Top-down models—typically grounded in historical time-series data—are designed to explore macro-level relationships across the energy, technological, and economic sectors. Building energy consumption is usually estimated as a function of economic variables (e.g., gross domestic product, unemployment rate, inflation, *per capita* income), energy prices, and/or climate variables. Because these models operate with highly aggregated data, they do not explicitly represent specific technological solutions or construction features.

A common example is the econometric model, developed to quantify the influence of economic variables on the outcomes of energy policies through statistical methods. However, its application is often challenged by the interdependence among energy and economic variables. The primary purpose of this approach—implicitly grounded in microeconomic theory—is to statistically predict the energy demand of a building stock, often using linear regression formulations with *per capita* income as a key explanatory variable.

The literature frequently distinguished between econometric and macro-econometric models, which differ mainly in their theoretical underpinnings: microeconomics for the former and macroeconomics for the latter. Macro-econometric models typically rely on national-level indicators, such as aggregate national income, rather than *per capita* metrics. Several examples of econometric *top-down* models can be found in literature. For instance, Bianco et al. (2009) developed a regression-based model capable of predicting urban energy use in Italy by correlating annual electricity demand with gross domestic product and population. Similarly, Haas & Schipper (1998) constructed multiple econometric models for OECD (Organisation for Economic Co-operation and Development) countries, relating residential energy demand to energy prices, household expenditure on energy, and degree-day values.

Due to issues such as multicollinearity, a major limitation of *top-down* building stock models is their inability to assess the impact of specific technological measures; they can only capture overall trends at an aggregate level.

The main strengths and limitations of *top-down* energy models are summarised in Table 1.

Table 1. *Strengths and limitations of top-down building stock energy models*

Strengths	Limitations
• Simplicity linked to reliance on aggregated historical data • Capability to extract relationships between factors that indirectly determine urban energy use • Ability to predict long-term energy consumption • Independence from detailed technological description	• Dependency on historical trends data • Generalisation of the existing conditions • Absence of determination of individual end-uses • Unsuitability to examine changes in technology

2.2.2 Bottom-up models

Unlike the *top-down* approach, the *bottom-up* model structures an ordered and sequential set of actions to achieve the final outcome, offering a detailed description of each subsystem and of its interactions with the others. *Bottom-up* building stock energy models rely on highly disaggregated and extensive datasets, which must be collected in advance through analytical, statistical, or empirical methods. These data are usually used to estimate the end-use energy consumption of individual buildings—or group of buildings—from which the total demand is then extrapolated to larger territorial scales. This enables the detailed assessment of potential energy retrofit strategies aimed at improving the performance of a defined building stock. Consequently, developers must be able to manage the complexity and uncertainty associated with numerous input parameters.

In the literature, *bottom-up* approaches are typically categorised into *engineering* (or *physical*), *data-driven*, and *hybrid* models.

Engineering (or *physical*) models rely on analytical calculations, generally involving the resolution of energy balance equations. This requires comprehensive knowledge of all input data, such as climatic conditions, geometric properties of buildings (surfaces, volumes, heights, window-to-wall ratios), thermo-physical characteristics of building envelope components, user behaviour, and technical system features. These data may derive from in-field measurements or empirical sources. *Engineering* approaches are the most widely used in the literature for building stock assessments and are commonly referred to as UBEMs. The next section of this chapter is dedicated to UBEMs and their related research developments.

Data-driven models start from large datasets and use statistical or artificial intelligence techniques to associate building characteristics and other factors—such as occupant behaviour—with energy use. Once identified, these mathematical relationships can be used to estimate end-use energy consumption (Abbasabadi & Mehdi Ashayeri, 2019). Statistical *data-driven* approaches typically apply regression techniques, such as linear or polynomial regression, which are simple and computationally efficient. However, their ability to capture complex real-world patterns is limited. This limitation can be overcome using

AI-based models, which employ machine learning (ML) algorithms that detect statistical patterns, fit the model, and infer functional relationships between energy use and building parameters. ML approaches include both *unsupervised* techniques (e.g., clustering) and *supervised* ones, such as neural networks. An example is offered by Pasichnyi, Wallin, & Kordas (2019) who developed a *data-driven* model of Stockholm's building stock to evaluate potential interventions; the optimal strategy reduced annual heating demand by 18 %.

A common issue in building stock modelling is the incomplete availability, accessibility, or accuracy of input data (Goy et al., 2020). *Data-driven* models often deliver higher accuracy, but their reliability depends strongly on the representativeness and granularity of the available dataset. *Engineering* models, by contrast, do not require historical energy-use records but are more sensitive to uncertainties arising from simplified building characterisation and user behaviour assumptions.

A comparative example is provided by Nageler et al. (2018), who developed both a dynamic, *physically* based model and a *data-driven* algorithm to estimate the thermal behaviour of a residential district comprising 34 buildings. Both approaches showed good agreement with measured data, and the authors discussed the suitable application domains of each.

Hybrid models aim to overcome the limitations of both *engineering* and *data-driven* approaches by combining their strengths. The stochastic nature of some parameters—especially those related to occupant behaviour—introduces uncertainty into *engineering* models. This challenge can be mitigated by integrating statistically derived inputs from historical datasets, thereby creating a more robust *bottom-up* methodology for determining building stock energy needs. These models typically extract information from multiple data sources to build a more complete and reliable representation.

Strengths and limits of the *engineering* and *data-driven bottom-up* energy models are summarised in Table 2 and Table 3, respectively.

Table 2. *Strengths and limitations of the engineering energy models*

***Engineering* models**	
Strengths	**Limitations**
• Independence from historical data • Easy assessment of energy efficiency measures • Adaptability with BIM tools	• Computationally time consuming • Loss of accuracy due to simplifications • Need of detailed input data

Table 3. *Strengths and limitations of the data-driven energy models*

***Data-driven* models**	
Strengths	**Limitations**
• Efficiency of processing time • Capability to include occupancy factors • Accurate prediction of energy use	• Relying on historical data • Limited adaptability in case of design variations (scalability) • Loss of accuracy in case of historical data unavailability

2.2.3 Urban Building Energy Models

Urban Building Energy Models (UBEMs), schematically represented in Figure 4, are generally classified as *bottom-up engineering* models, although the literature offers different interpretations. According to Sola et al. (2020) and Hong et al. (2020), the notion of UBEM is independent of the specific *bottom-up* methodology adopted—whether *engineering* or *data-driven*. In contrast, Reinhart & Cerezo Davila (2016) and Johari et al. (2020) associate UBEMs specifically with the *engineering bottom-up* models, which are developed to estimate the operational energy demand of a given building stock. This energy demand—related to energy services such as space heating, space cooling, domestic hot water (DHW) production, ventilation, artificial lighting—is calculated through the resolution of mass and energy conservation equations.

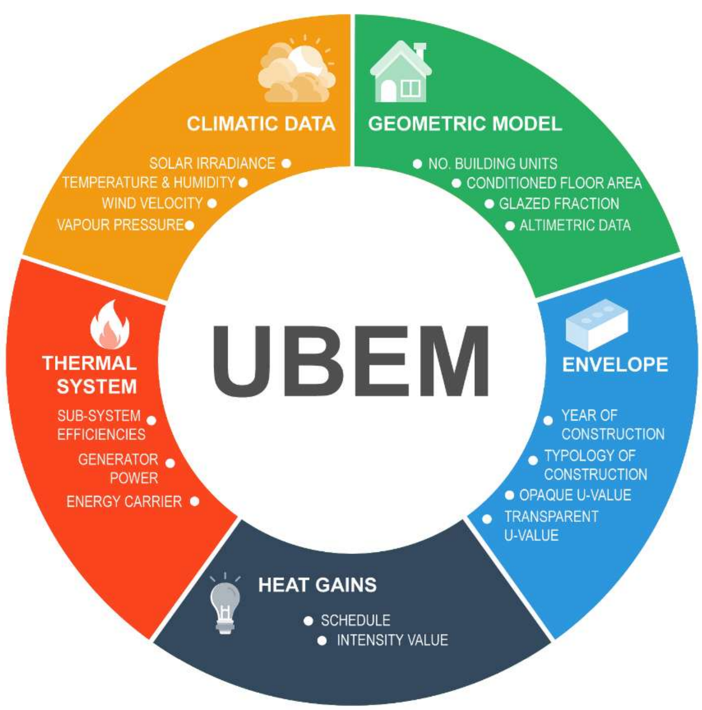


Figure 4. *Key elements of the Urban Building Energy Modelling*

Following this classification proposed by Swan & Ugursal (2009) for residential building stock energy models (Figure 2), *engineering* approaches—and thus UBEMs—can be further divided into three modelling strategies:

- *Appliance-use-based* approach, which estimate energy demand by analysing the appliances used within dwellings at regional or national scale, considering parameters such as type, use, efficiency, and consumption.
- *Archetype-based* models, which rely on representative building archetypes defined according to energy-relevant characteristics (e.g., construction period, building size, use category). The energy performance of each archetype is assessed and then scaled up to a broader territory.
- *Sample buildings-based* approach, which evaluate a large set of buildings for which detailed data are available. If this sample is representative of the broader regional or national stock, the total energy demand can be derived using weighted factors.

A schematic illustration of these three UBEM approaches is provided in Figure 5.

The first method—based on *appliance* use—is the least common, as it is applicable only to specific situations (e.g., estimating electrical demand when appliance type, power, and use profiles are known). Although it follows a *bottom-up* logic due to its reliance on disaggregated data, the most widely adopted UBEM approaches are the latter two, which differ mainly in the number of buildings represented. *Archetype-based* models typically involve a limited number of buildings constructed using large datasets, whereas *sample-based* models require the collection of detailed data for hundreds or thousands of buildings. An illustrative example is provided by Cerezo et al. (2017) who compared four archetype-based modelling methods on 336 residential buildings in Kuwait City. Two methods relied on deterministic modelling, one on probabilistic modelling, and the fourth introduced a Bayesian calibration approach, which achieved the closest agreement with measured energy use.

UBEMs serve as valuable tools for urban planners, designers, and public authorities aiming to assess the energy performance of urban environments and to support initiatives for improving the efficiency of the existing building stock. According to Reinhart & Cerezo Davila (2016), model accuracy tends to improve when moving from individual buildings to districts and regional scales. Across multiple-building UBEM applications, deviations between predicted and measured annual energy consumption for heating loads typically range from 7 % to 21 %. A notable example is the UBEM developed by Hedegaard et al. (2019), to estimate the energy consumption of a Danish neighbourhood consisting of 159 single-family houses served by district heating. The model was designed to predict daily consumption profiles that closely matched real data, thereby helping reduce peak network loads.

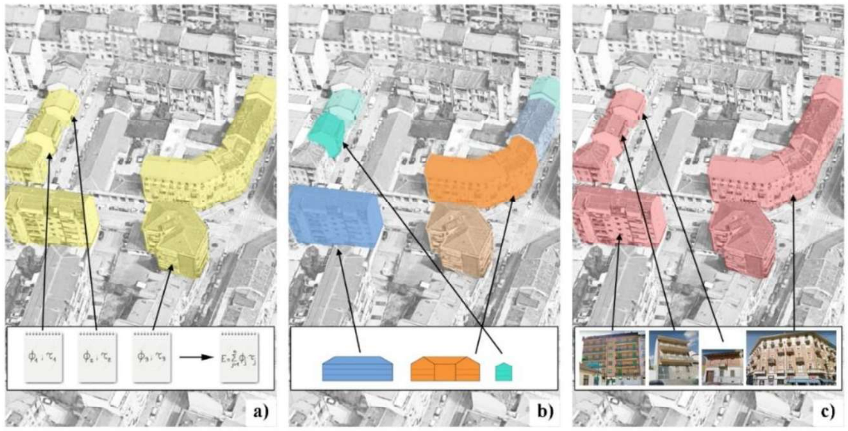


Figure 5. *Types of UBEMs: a) appliances-use based approach, b) archetypes-based approach, c) sample buildings-based approach (Ballarini et al., 2021)*

A wide range of UBEM development tools has emerged, most of them produced in academic settings. These tools commonly utilise Building Information Modelling (BIM) workflows to store, process, and visualise data. UBEMs frequently include a three-dimensional geometric representation, climate and topographic data, thermo-physical properties of envelope components, building-use information, and details of the technical systems.

To represent the urban geometry, the Open Geospatial Consortium (OGC) has developed the CityGML standard (Kolbe et al., 2021), which defines different Levels of Detail (LoD). The geometric and informational resolution of a UBEM depends on the purpose of the analysis. For example, LoD1—simple volumetric representations with planar surfaces—can be sufficient for building stock energy assessments. LoD2 incorporates sloped surfaces, enabling more accurate solar access studies. LoD3 and LoD4 further refine the model by adding external window frames and, potentially, internal building structures. Consequently, the appropriate level of detail varies depending on whether the simulation focuses on building-, district-, urban-, or regional-scale analyses.

2.2.4 Urban-Scale Energy Models

UBEMs, which focus exclusively on estimating building energy demand during the operational phase, are often combined with additional modelling tools to account for energy supply and distribution in urban areas (Allegrini et al., 2015). The urban environment can be viewed as a system in which multiple energy conversion processes take place across several sectors—buildings, industry, transport, and urban infrastructure. Understanding this complex urban thermodynamic system requires a holistic, integrated approach that captures the interactions and exchanges of different energy flows across multiple spatial and functional boundaries. In the literature, this broader concept is encompassed by the Urban Scale Energy Model (USEM), as defined by Sola et al. (2020).

A USEM is a virtual representation of urban energy flows with spatial and temporal resolution, designed to characterise a given urban energy system and explore strategies to enhance its performance. Through virtual modelling, a USEM can assess system losses, load peaks, unused energy fractions, and interactions between energy producers and consumers. As highlighted by Basu et al. (2019), a USEM is defined not only by physical components but also by the actors involved in energy distribution (e.g., suppliers, agents, consumers) and non-material factors such as user feedback, incentive schemes, and policy frameworks. Regarding the inclusion of urban energy components, models addressing building operational energy, transportation energy, and embodied energy in construction typically coexist but often function independently. Fully integrated models remain relatively rare, as also noted by Abbasabadi & Mehdi Ashayeri (2019). Within the USEM domain, and considering the various modelling approaches for both urban energy flows (generation, distribution, and demand) and energy-consuming sectors (buildings, transport, industry), a detailed classification of the principal existing tools can be established, as summarised in Table 4.

Table 4. *Classification of the main tools for urban energy use assessment*

Tool	**Reference**	**Model category**[(a)]	**Modelling approach**[(b)]	**Time step**[(c)]	**Modelling object**		
					Energy demand		Energy supply and network
					Buildings	Transport	
AutoBPS	Deng et al. (2023)	UBEM	EN	H	●		●
CEA	Fonseca et al. (2016)	USEM	EN	H	●	●	●
CityBES	Chen et al. (2017)	UBEM	EN	H	●		
CitySim	Robinson et al. (2009)	USEM	EN	H	●		●
EUReCA	Prataviera et al. (2021)	UBEM	EN	H	●		●
SimStadt	Nouvel et al. (2015)	UBEM	EN	M, H	●		●
TEASER	Remmen et al. (2018)	UBEM	EN	H	●		
UBEM.io	Ang et al. (2022)	UBEM	EN	H	●		
UMI	Reinhart et al. (2013)	UBEM	EN	H	●		
UrbanOpt	Kontar & Fleming (2020)	USEM	EN	H	●		

(a) UBEM = Urban Building Energy Model, USEM = Urban-Scale Energy Model
(b) EN = engineering
(c) M = monthly, H = hourly

2.2.5 Urban Energy Systems

Energy-efficient urban planning must be recognised that urban components—both energy consumers (e.g., buildings, transport) and energy suppliers—operate simultaneously across multiple scales and dimensions. This requires that models and tools used to analyse specific elements of the urban environment be integrated to reflect the inherent complexity of the urban system through a hierarchical approach. To address this need, the concept of the Urban Energy System (UES) was introduced, defined as "*the combined process of acquiring and using energy to satisfy the demands of a given urban area*" (Keirstead & Shah, 2013). As a dynamic system, a UES should be constructed to both represent the current state and simulate the future evolution of the urban energy performance, thereby supporting informed decision-making for future urban energy efficiency.

Compared with UBEMs or USEMs, a UES extends beyond the urban energy model itself. While UBEMs and USEMs focus mainly on characterising energy demand and flows, the UES encompasses the broader, multidisciplinary process required to conceptualise and develop such models dynamically. Although UES and USEM share an interdisciplinary perspective, the key distinction lies in the simultaneous consideration of all energy-related aspects that define the urban environment. UESs are often described metaphorically as metabolic systems (Keirstead & Shah, 2013), intended to assess social, economic, and environmental impacts. This analogy emphasises the analysis of all material and energy transformations occurring within the urban context and driven by external inputs. Thus, while USEMs aim to integrate multiple energy-related models wherever possible, a holistic integration is a fundamental requirement of UES. From this perspective, UBEMs and USEMs can be considered subsystems of the broader UES framework.

To capture the spatial and temporal dynamics of urban energy flows, UES modelling must acknowledge the involvement of multiple expertise domains. The development of an urban energy system therefore affects not only the territory's energy efficiency but also its economic, political, social, demographic, and cultural dimensions. These factors, in turn, influence urban energy flows, creating a circular and interdependent process.

2.3 Uncertainty in UBEM

Uncertainties hinder the reliable implementation of large-scale energy models. According to the ASHRAE Guideline-14 (ASHRAE, 2024), model uncertainty is defined as the "*range or interval of doubt surrounding a measured or calculated value within which the true value is expected to fall within some degree of confidence*". Based on Fennell et al. (2025) and Fennell et al. (2022), these deviations can be categorised into input-, model-, and output-related uncertainties (Figure 6).

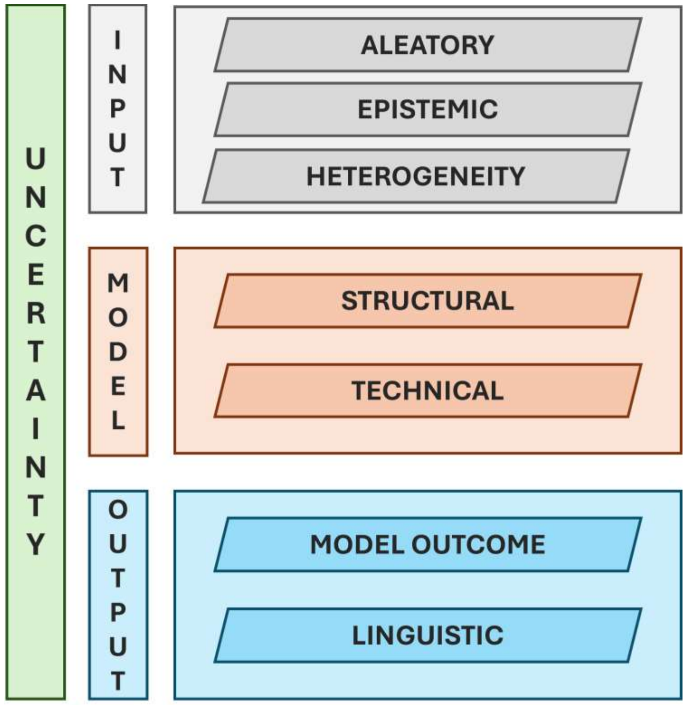


Figure 6. *Sources of uncertainties in UBEMs*

More specifically, they can be classified according to where they occur along the modelling chain:

- Input-related uncertainties
  - Aleatory (or stochastic): represents the inherent natural variation within a system that is fundamentally random. It is an intrinsic property of the system itself and cannot be reduced through additional knowledge.
  - Epistemic: arises from a lack of knowledge or imperfect information about a specific parameter. Unlike aleatory uncertainty, it can be reduced through additional data collection or improved research.
  - Heterogeneity: often classified under epistemic uncertainty in building stock models, this refers to the real-world variation among individual buildings that has been simplified into a single archetype cluster.
- Model-related uncertainties
  - Structural: a form of epistemic uncertainty resulting from an incomplete understanding of the real processes being modelled. It includes decisions regarding which physical phenomena to represent and which sub-models to adopt.
  - Technical: refers to errors generated by the software or hardware used for the simulation.

- Output-related uncertainties
    - Model outcome: the quantified range of results produced by the simulation itself.
    - Linguistic: arises from ambiguity or language-related issues when communicating model findings.

2.4 Data for building stock energy modelling

The main challenge faced by UBEMs developers is managing the high degree of epistemic uncertainty in determining certain energy-related data—particularly when dealing with existing building stocks. City-wide information, which is often inaccessible, unavailable, or inaccurate (Goy et al., 2020) is essential for informatising and evaluating the energy performance of the building stock. This raises critical issues related to data standardisation and harmonisation at the municipal or urban scale, as well as the need for interoperability between databases. Furthermore, due to privacy concerns, building-level consumption data is considered sensitive, which undermines the credibility of large-scale energy analyses and complicates the validation and calibration of UBEMs. As a result, UBEM developer must invest substantial effort in data collection (C. Wang et al., 2022) or employ multi-source data fusion techniques (Choi et al., 2025) to achieve highly accurate datasets.

In this context, the Energy Performance Certificate (EPC) is a crucial document for enhancing knowledge of the building stock's energy performance. The European Directive 2002/91/EC (European Commission, 2002) introduced the EPC as a mandatory requirement for constructed, sold, or rented out buildings. The primary objective of this document is to create a homogeneous framework for building energy-saving initiatives around EU Member States, aiming to achieve energy targets for the building stock. The certification process aims at providing a transparent tool in the building market to compare and assess the energy efficiency of the buildings. The EPC includes reference values such as current legal standards and benchmarks, enabling owners and tenants to evaluate the energy performance of a building or building unit.

Therefore, the EPC represents a core source of information that reflects the energy performance *status* of the building stocks, the fuel poverty of a specific country, and the monitoring of national or regional energy renovation programs (Hardy & Glew, 2019). The data quality of certificates influences the reliability, adequacy, plausibility, coherence, and accuracy—considered as the reduction of the uncertainty—of the information stored. The main aspects that reflect on the EPC data quality are related to the accuracy of the input data, the methodology and software applied, and the energy assessors' qualification (Li et al., 2019). Incorrect energy certificates influence the building economic assessment since they generate wrong energy performance labels: the more energy-efficient a building is, the more it is associated with a "premium price" (Hårsman et al., 2016; Khazal & Sønstebø, 2020). However, the knowledge and the experience of the technician in charge to carry out the EPC is the main factor that affects the data quality of the certificate. Tronchin & Fabbri (2012) have set up a round-robin test to evaluate the expertise of 162 technicians to estimate the energy performance of the same building. All the independent experts have had the same inputs and were beginners on the software to be used. The experiment revealed that approximately 70 % of energy assessors had correctly estimated the energy rating of the building assessed. Certainly, the aforementioned percentage can be influenced by the energy tool chosen by the independent experts. However, in Italy, this issue has been overcome as the software to calculate the energy performance of buildings is validated by the Italian Thermotechnical Committee (CTI) and presents a maximum deviation of ± 5 % compared to the corresponding energy indicators determined through technical standards (Italian Republic, 2015). So, the test of Tronchin & Fabbri (2012) clearly showed that the qualification of the technician in charge is essential for the quality of the EPC. In this context, the collection and the validation of the input data represent a basic skill to correctly conduct the subsequent energy certification phases.

European harmonisation of the minimum qualification level of an energy assessor, of the energy performance calculation methodology, and of the data stored in the certificate are fundamental requirements to homogenise and to align the EPC framework and content around European countries. High EPC data quality is a necessary condition to have a clear picture of the energy *status* of the building stock and to improve the knowledge and awareness on energy efficiency.

2.5 Representative buildings

To manage data uncertainty in large-scale energy analyses, collected information is often organised into building archetypes (BAs). This approach provides a balanced compromise, reducing modelling complexity while improving overall accuracy. Building archetypes describe the representative technological characteristics of specific portions of the building stock and consist of non-geometric datasets that define envelope properties, technical system features, and building-use patterns. BA datasets may be defined *deterministically*—using single representative values—or *probabilistically* (Prataviera et al., 2022)—using parameter ranges (Sokol et al., 2017). The archetype-based UBEM approach is one of the *engineering* methods identified by Swan & Ugursal (2009).

The BA concept is frequently confused with that of building prototypes. While prototypes refer to generic buildings with abstracted representative geometries (Carnieletto et al., 2021), building archetypes are not geometry-dependent and are instead assigned directly to buildings in the real urban context.

Although no harmonised methodology for BA generation currently exists (Borges et al., 2022), the *archetyping* process of the building stock typically involves two key steps: *segmentation* and *characterisation* (Pasichnyi, Wallin, & Kordas, 2019). *Segmentation* consists of classifying similar buildings according to criteria such as climate zone, use category, and construction period.

Characterisation then assigns representative parameters that generalise the energy behaviour of each segment.

EPCs are widely used data sources to develop BAs libraries (Johari et al., 2023). However, EPC data present several limitations, including reliance on static performance assessments, restricted availability of energy-related information, and varying levels of data quality (Piro et al., 2024).

For this reason, developing an archetype-based UBEM requires the use of recognised and validated libraries of representative buildings. These should include enough real and/or "virtual" residential and non-residential reference buildings to reflect the most common building categories within a given climatic zone. In Italy, the national building typology matrix originally developed under the TABULA project (Ballarini et al., 2014) has been updated by the results of the URBEM—*Urban Reference Buildings for Energy Modelling* project (Ferrando et al., 2025), which provides an extensive characterisation of building energy performance across the country. Similarly, the BUILD_ME project (BUILD_ME, 2025) offers comprehensive residential and non-residential building typologies for Lebanon, Egypt, and Jordan.

Currently, the focus is shifting from representative buildings toward representative urban blocks or districts. In this context, Zhang, Sun, Polly, et al. (2026) developed a set of representative district models classified according to urban density characteristics and building type distributions (e.g., "medium-density urban mixed-use district" or "suburban residential district"). These scorecards were used to perform a district-scale analyses aimed at evaluating multiple energy demand scenarios and the impact of distributed energy resources.

2.6 Large-scale validation and calibration procedures

Enhancing the credibility of National Building Renovation Plan (NBRP) is closely connected to reducing uncertainty in UBEM outputs. Depending on the aim of the analysis, this objective can be achieved through *validation* or *calibration* of large-scale energy models. *Validation* assesses whether the modelling framework and underlying assumptions are appropriate by comparing simulation results with benchmark data. This approach is typically adopted when only aggregated consumption data are available or when the computational cost of a full calibration is prohibitive.

Calibration, in contrast, involves adjusting a vector of input parameters so that simulated outputs match measured data as closely as possible, based on statistical indicators such as root-mean-square error (*RMSE*), coefficient of variation of *RMSE* (*CV-RMSE*), R^2, mean bias error (*MBE*), and mean absolute percentage error (*MAPE*). Given the increasing complexity of UBEMs with growing numbers of buildings, automated calibration approaches are generally favoured over manual tuning while still preserving engineering judgement.

Calibration of large-scale energy models may be conducted at various spatial levels—including individual buildings, districts, or entire building stocks.

According to the scientific literature (Fabrizio & Monetti, 2015; Sokol et al., 2017), automatic UBEM calibration methods can be grouped into three main categories, which can be used individually or in combination:

- *Bayesian* approaches, which integrate prior knowledge with observed data to iteratively update probability distributions of uncertain parameters.
- *Objective/penalty function* methods, which minimise discrepancies between measured and simulated data by systematically adjusting input parameters using deterministic or stochastic algorithms.
- *Meta-modelling* approaches, which construct simplified surrogate models that approximate the behaviour of computationally expensive simulations, thereby enabling faster calibration through reduced computational effort.

According to the proposed classification, objective/penalty function and Bayesian approaches are grouped within the optimisation-based techniques. Table 5 summarises the most relevant studies aimed at improving the reliability and credibility of UBEM outputs through validation or calibration, specifying the simulation engines employed and the utilisation of archetypes.

Table 5. *Large-scale validation and calibration approaches in literature*

Reference	UBEM tool (Simulation engine)	Archetype	Validation	Calibration		
				Optimisation techniques		
				Bayesian	Objective or penalty function	Meta-modelling
Cerezo Davila et al. (2016)	umi (EnergyPlus)	x	x			
Chen et al. (2020)	CityBES (EnergyPlus)					x
Dilsiz et al. (2023)	CitySim (CitySim solver)			x		
Faure et al. (2024)	MUBES (EnergyPlus)			x		
Gholami et al. (2021)	n/a (EnergyPlus)	x		x		
Haneef et al. (2021)	CitySim (CitySim solver)	x			x	
Heo et al. (2012)	n/a (EnergyPlus)	x		x		
Kim & Kim (2024)	umi (EnergyPlus)			x		x
Lundström & Akander	n/a (ISO14N model)			x		

Reference	UBEM tool (Simulation engine)	Archetype	Validation	Calibration		
				Optimisation techniques		
				Bayesian	Objective or penalty function	Meta-modelling
(2020)						
Mattsson et al. (2025)	CEA (CEA engine)	x			x	
Nageler et al. (2017)	IDA ICE (IDA ICE engine)		x			
Nagpal et al. (2019)	n/a (EnergyPlus)	x				x
Nouvel et al. (2017)	SimStadt (SimStadt engine)		x			
Sokol et al. (2017)	n/a (EnergyPlus)	x		x	x	x
Tognoli et al. (2023)	CitySim (CitySim solver)	x			x	
Vivian et al. (2025)	EUReCA (EUReCA engine)	x			x	
Zhang, Sun, Li, et al. (2026)	CityBES (EnergyPlus)				x	
C. K. Wang et al. (2020)	CitySim (CitySim solver)	x		x		

2.7 Further challenges in UBEM

Beyond the challenges related to city-scale data availability, quality, analysis, and model validation and calibration discussed in the previous sections, several additional factors influence the reliability of UBEM applications. These include occupant-related dynamics, urban microclimatic phenomena, and the simplified assumptions commonly adopted to solve the urban energy balance.

From this perspective, urban mobility—namely the activities and behavioural patterns of occupants across the analysed district or city—represents a major source of uncertainty. Occupant interactions within buildings influence internal heat gains and consequently affect energy performance predictions. Occupant behaviour can be represented through deterministic or stochastic approaches using either person-level or space-level data (Happle et al., 2018; Zhao et al., 2022). For example, Yang et al. (2025) employed mobile traffic data to investigate dynamic population distribution and improve urban space-use efficiency, while Abolhassani et al. (2022) relied on Wi-Fi sensing data to estimate occupant presence in two university buildings in Canada. In this context, Nejadshamsi et al. (2025) introduced the Transportation-Informed Building Occupancy approach to realistically assess citizen flows across different district in Montreal by combining heterogeneous datasets from travel survey, census data, and building-related information.

Although the Urban Heat Island (UHI) effect is widely recognised as a significant factor influencing building energy performance, only a limited number of studies have successfully quantified its impact with high accuracy. This limitation is mainly due to the challenges associated with collecting microscale climatic data through Urban Canopy Layer (UCL) analyses and modelling the complex drivers of the UHI while accounting for interactions between buildings and their surrounding environment within UBEMs. Among the available approaches, UCL energy balance models and computational fluid dynamics (CFD) techniques have demonstrated the highest reliability. From this perspective, Mun et al. (2026) coupled CitySim with a CFD model to assess the impact of cool coating strategies on pedestrian thermal comfort within an urban canyon. Similarly, G. Chen et al. (2020) employed a CitySim-CFD coupling approach to investigate whether non-uniform wall temperatures could be approximated by uniform wall temperatures in urban airflow analyses within an idealised 5x5 building array.

However, the application of such microscale models at district or city scale is often computationally prohibitive because of the large number of urban features and geometrical details involved. Therefore, simplified approaches aimed at avoiding CFD simulations while improving predictive accuracy are generally preferred. Using the Urban Weather Generator (Nakano et al., 2015), Palme et al. (2017) estimated increases in building energy demand ranging from 15 % to 200 % under UHI conditions. Similarly, Liu et al. (2026) applied the Urban Weather Generator to the city of Nanjing and showed that neglecting the urban heat island effect would likely lead to an underestimation of city-scale cooling demand by 11.4 % and an overestimation of heating demand. While most UHI-related correction methods, such as the Urban Weather Generator, primarily focus on modifying hourly outdoor air temperature and relative humidity, other studies have investigated the influence of urban morphology on local wind conditions. In this context, Salvati et al. (2020) proposed an algorithm for correcting hourly wind speed within urban canyons based on urban morphology indicators, including height-to-width and length-to-width ratios.

In addition to issues related to urban mobility and microclimate factors, systematic deviations in UBEM may also originate from the simplified assumptions adopted to solve the urban energy balance. Knowledge derived from conventional BEM has enabled the incorporation of assumptions and simplifications in UBEM heat-exchange calculations, as discussed by Salvalai et al. (2024). These simplifications concern geometric representation, thermal zoning discretisation, the characterisation of envelope components and technical systems, as well as differences in calculation procedures between BEM and UBEM frameworks. Within UBEM, urban geometry is typically represented according to the LoD classification, initially proposed by Kolbe (2009) and later updated by Kolbe et al. (2021). This classification ranges from LoD0, representing highly generalised geometries, to LoD3, corresponding to highly detailed urban models. However, the applicability of different LoDs to specific analytical purposes has been critically examined, demonstrating that increased geometric detail does not necessarily improve modelling performance, especially in urban-scale environmental and energy applications (Biljecki et al., 2016). In this context, Hwang et al. (2024) explored geometry-related biases using artificial neural networks and Bayesian calibration methods. Concerning thermal zoning subdivision, Johari et al. (2022) assessed acceptable simplification levels to achieve a balance between computational efficiency and modelling accuracy.

2.8 Remarks

The recently issued EPBD recast (EU 2024/1275) requires each Member State to establish "*a national building renovation plan to ensure the renovation of the national stock of residential and non-residential buildings, both public and private, into a highly energy efficient and decarbonised building stock by 2050, with the objective to transform existing buildings into zero-emission buildings*".

However, several challenges hinder effective implementation. These include the limited and low-quality informatisation of building stock energy performance, the lack of monitoring data required for validating and calibrating large-scale energy models, and the significant computational time needed to run multiple simulations with hundreds or thousands of buildings.

To address these challenges, NBRPs must be developed using a well-informed, archetype-based UBEM approach that ensures quality at every stage—from input data collection, cleaning, processing, to the verification of urban-scale outcomes.

Chapter 3

Building-Scale vs. Urban-Scale Energy Modelling

3.1 Overview

The work presented in this chapter has been previously presented at the 18th Building Simulation 2023 International Conference (Piro et al., 2023). In Piro et al. (2023), the Ph.D. activity focused on reviewing the simplifications introduced in UBEMs compared to BEMs, as well as on defining the methodology and its application for the numerical assessment of the differences between the two modelling approaches.

Given the growing adoption of UBEMs for analysing energy performance at district and city scale, it is first necessary to clarify the theoretical differences between traditional single-building approaches and urban-scale modelling frameworks.

The chapter begins with a theoretical overview of the key differences between BEM and UBEM, focusing on modelling assumptions related to: building geometry (Section 3.2.1), building envelope components (Section 3.2.2), zoning (Section 3.2.3), internal heat gains (Section 3.2.4), technical building systems (TBSs) (Section 3.2.5), connection with UBEM tool (Section 3.2.6), and calculation modules (Section 3.2.7). Subsequently, a quantitative assessment comparing BEM and UBEM approaches is presented (Section 3.3), followed by a discussion of the results (Section 3.4). The chapter concludes with a summary of the key findings and implications (Section 3.5).

Specifically, in Section 3.3 a case study application with BEM and UBEM is conducted using the same energy simulation tool (i.e., CitySim) for both cases. In this way, the peculiarities of the mathematical models, related to the differences between single-building and urban building energy models, were neutralised. The objective of this work is to emphasise the modelling characteristics between BEM and UBEM, providing a preliminary quantitative investigation. In particular, the case study studies and compares the shortwave (SW) and longwave (LW) exchange between apartment blocks in a BEM and an UBEM for a real city block located in northwest Italy.

3.2 The paradigm shift: Towards Urban Building Energy Modelling

Recently, a paradigm shift has emerged: the transition from BEMs, which focus on individual buildings, to UBEMs, which operate at the city-level scale. The experience gained from the development of Building Energy Performance Simulation (BEPS) tools (Crawley et al., 2008) has enabled the introduction of new assumptions and simplifications for modelling heat exchanges within urban environments.

Large-scale models make it possible to explore the complex, multi-domain energy flows that characterise urban systems. However, UBEMs lack a predefined spatial and temporal resolution, meaning that the level of geometric-informative detail generally decreases as the scale of analysis increases—a relationship clearly observable in the transition from BEMs to UBEMs.

Numerous studies have contributed to the advancement of both BEM and UBEM tools. Ferrando et al. (2020) and Abbasabadi & Mehdi Ashayeri (2019) provided comprehensive reviews of UBEM tools and applications aimed at evaluating the energy *status* of building stock, while Crawley et al. (2008) compared a wide range of energy simulation programs designed for single-building assessments.

A quantitative comparison between BEMs and UBEMs remains challenging and cannot be generalised when different BEPS tools are used. Two main groups of differences can be identified. The first concerns UBEM-specific simplifications in building modelling. For example, in large-scale energy analyses, transparent envelope components are typically parametrised through the window-to-wall ratio (*WWR*) and positioned at the façade's centre of gravity, whereas in BEMs, windows are explicitly spatially defined. The second group relates to the capacity of UBEMs to capture building-environment interactions. In UBEMs, surrounding buildings are modelled as dynamic entities, while in BEMs they are usually represented as static shading or reflective elements. Consequently, solar radiation reflected onto target building in a BEM may, through glazing components, influence the energy balance of adjacent buildings when assessed in UBEM context.

3.2.1 Geometrical model

In BEM, complex geometries are modelled within the constraints of the specific software employed. The shift from Computer-Aided Drawing (CAD) to BIM has revolutionised the Architectural Engineering and Construction (AEC) industry. Currently, both geometric and informational modelling are primarily conducted in environments distinct from energy simulation tools. Within this framework, data interoperability—the ability to exchange information between different simulation platforms, such as BIM and BEM—becomes a critical factor (Pezeshki et al., 2019). In this respect, the adoption of open data exchange formats such as Industry Foundation Classes (IFC) and Green Building XML (gbXML) has played a fundamental role.

In contrast, urban-scale representations typically rely on mass modelling, characterised by simplified 3D geometries generated through the vertical extrusion of building footprints. In UBEM, the level of geometric detail depends on the LoD adopted. To enable multiscale representation, Kolbe et al. (2021) updated the XML-based CityGML data model (Gröger et al., 2008), which defines four standardised LoDs (LoD0-3): LoD0 – "*highly generalized model*", LoD1 – "*block model/extrusion objects*", LoD2 – "*realistic, but still generalised model*", and LoD3 – "*highly detailed model*". In practice, UBEM applications most often employ LoD1 and LoD2. LoD1 corresponds to a simplified shoebox representation, while LoD2 allows for sloped or oblique surfaces, enabling more accurate solar analyses (e.g., shortwave irradiance on roofs, PV potential, or solar collector simulations). Notably, in v.3 of the CityGML standard (Kolbe et al., 2021), internal building elements—such as horizontal and vertical partitions—are included at all LoD levels. Despite these advancements, achieving full interoperability between BIM and UBEM environments remains an open research challenge.

3.2.2 Building envelope

In an UBEM, the specification of the building envelope components—vertical, upper, and lower horizontal closure—can typically be defined either by façade orientation or for the building as a whole. In the first case, if the stratigraphy varies along its vertical or horizontal axis, the developer is constrained in making detailed modifications. In the second case, only a single stratigraphic package can be assigned to each component (wall, floor, or roof). Both approaches inevitably introduce simplifications. By contrast, BEMs provide greater flexibility in modelling opaque envelope elements.

In UBEMs, glazed components are generally positioned at the centre of gravity of each façade or building floor. Windows openings are defined as a function of the *WWR*—either constant across the entire building or variable by façade. Consequently, users cannot directly control the exact spatial position or size of individual windows. This simplification can lead to uncertainties in estimating solar gains within thermal zones and in assessing shading effects from nearby obstructions. Moreover, when multiple glazing systems with differing optical or thermal properties exist on the same façade, their behaviour is often represented through averaged parameters. However, UBEM tools that use EnergyPlus as their simulation engine (e.g., UMI, CityBES, URBANopt) include a more detailed set of parameters, particularly for transparent envelope components. In contrast, BEMs define both opaque and transparent building envelope components explicitly, in terms of their thermophysical and spatial characteristics.

The representation of solar shading devices in UBEMs is also simplified—often modelled via a solar irradiance set-point. When only the solar energy transmittance of the window is imported, a weighting factor based on the operating schedule (opening/closing hours) of the shading system may be applied. The operational rules for such devices are outlined in the EN ISO 52016-1 technical standard (CEN, 2017a). In single-building energy models, instead, shading devices are explicitly defined with respect to both their spatial location and thermophysical properties.

Thermal bridges, typically neglected in large-scale models, can be approximated by incorporating total heat losses—previously estimated as point or linear thermal bridges—into the overall energy balance. Conversely, in BEMs, linear thermal transmittance can be calculated through numerical finite element methods or by referencing thermal bridge atlases, depending on the geometric or structural discontinuities under assessment.

3.2.3 Zoning

In an urban building energy model, the subdivision of internal space through vertical and/or horizontal partitions is generally limited. The representation of internal vertical partitions—often introduced through corrective factors—is typically intended only at calculating the thermal capacity of the building environment. In contrast, the horizontal subdivision offers the user greater control, allowing specification of the total building height, floor-to-floor height, or inter-storey height to define zones by floor. In some modelling environments, however, individual building elements—particularly vertical partitions—cannot be explicitly specified.

Consequently, the definition of thermal zones is strongly influenced by the implicit or absent modelling of internal partitions between different spaces. In most UBEM applications, a single thermal zone corresponds to the entire building. This simplification inevitably affects the representation of unconditioned spaces, the evaluation of the technical building system, and the estimation of internal heat gains, as the latter two are typically assigned to the predefined conditioned floor area. In this context, AutoBPS (Deng et al., 2023) automatically subdivides buildings into floors, assigning a corresponding thermal zone to each floor.

An implementation of these approaches is presented in the Appendix G of ANSI/ASHRAE/IES Standard 90.1 (ASHRAE, 2025), which provides the core-and-perimeter thermal block zoning method. This approach distinguishes between core spaces—located more than 4.5 m from an exterior wall—and perimeter spaces—located within 4.5 m of the building envelope. Building upon this framework, and from a stock-scale perspective, Dogan et al. (2014) present an algorithm integrated into UMI (Reinhart et al., 2013) that automatically divides each building floor into perimeter and core thermal zones. However, particularly for large and complex buildings, this method may fail to capture the energy dynamics of internal spaces adequately. To address this limitation, Zhou et al. (2025) introduced a more advanced percentage-based thermal zoning algorithm based on the discretisation of floor areas through Voronoi polygons and the aggregation of the resulting surfaces into different thermal zones.

By contrast, in a BEM, thermal zones are defined according to usage conditions, building complexity, the characteristics of the technical building systems, or specific regulatory requirements—for example, when assessing the energy performance of an individual building unit within an apartment block.

3.2.4 Internal heat gains

In both BEM and UBEM, internal heat gains are typically attributed to occupants, electrical appliances, and lighting equipment. For each category, depending on the simulation timestep, both the internal heat gains schedules and the corresponding intensity must be defined. According to Happle et al. (2018), occupant behaviour in energy models can be represented either deterministically or stochastically, with evaluations carried out on a per-person or per-space basis.

In UBEM applications, Zhao et al. (2022) categorise occupant-modelling methodologies into the following approaches: *schedule*-based, *standard*-based, *appliance*-based, and *statistic*-based. In UBEM, where uncertainty is inherently high, occupants' behaviour is most modelled deterministically using a schedule-based approach, in which occupants are assumed to follow predefined daily routine. In contrast, CitySim (Robinson et al., 2009) and City Energy Analyst (Fonseca et al., 2016) allow occupant activity to be represented through a stochastic schedule-based approach.

Because a UBEM thermal zone generally represents an entire building, the uncertainty in estimating internal heat gains increases when multiple use categories coexist within the same structure. For example, in an apartment block that includes commercial units on the ground floor, a prevailing building use and an average occupant behaviour must be assumed, as the same input data are applied to the whole building. Consequently, internal heat gains for each thermal zone are estimated parametrically, rather than through precise spatial definitions (e.g., considering the contribution of each individual light source). In this case, the thermal load from electrical equipment may be expressed as a function of the conditioned floor area (e.g., W/m^2).

In contrast, BEMs offer greater flexibility: the modeller can adopt either deterministic or stochastic approaches, with point-by-point or parametric definitions, depending on the model's complexity and the availability of input data.

3.2.5 Technical building systems

In a BEM, the technical building systems are specified by analytically or ideally modelling the performance of all subsystems—generation, storage, distribution, control, and emission. Each component is assessed in relation to its interaction with the surrounding environment, and the calculations are iteratively refined to ensure that heat losses are balanced according to the thermal zone's demand. This level of detail and accuracy, however, is generally unfeasible in UBEM.

Consequently, UBEM analyses typically focus on the type of energy system (e.g., boiler, heat pump, district heating, photovoltaic panels, solar collectors) and the key characteristics of the generator, such as thermal capacity, efficiency, energy carrier, and inlet/outlet temperatures. In some UBEM tools, simplified correction factors allow for the approximate consideration of heat losses within specific subsystems. Nevertheless, UBEM environments generally offer limited flexibility in incorporating national or local regulations—such as prescribed on/off periods for system operation (e.g., space heating schedules).

3.2.6 UBEM tool

This section presents the degrees of freedom of large-scale energy modelling tools, presented in Table 4, according to their characteristics in terms of geometry, building envelope components, zoning, internal heat gains, and TBS. Table 6 links these specifications to the corresponding software tools.

Table 6. *Characteristics of UBEM tools*

Tool	LoD	Envelope		Zoning	Gains[b]	TBS[c]
		Opaque[a]	Transparent[a]			
AutoBPS	LoD 1	surf-level	surf-level	floor-based	sched-based (D)	sub-sys-level
CEA	LoD1	bldg-level	bldg-level	single	sched-based (D), sched-based (S)	sub-sys-level
CityBES	LoD 1	surf-level	surf-level	single, core-and-perimeter	sched-based (D)	gen-level
CitySim	LoD 1, LoD 2	surf-level	surf-level	single, floor-based	sched-based (D), sched-based (S)	gen-level
EUReCA	LoD 1	bldg-level	bldg-level	single	sched-based (D)	sub-sys-level
SimStadt	LoD 1, LoD 2	bldg-level	bldg-level	single	sched-based (D)	sub-sys-level
TEASER	LoD 1, LoD 2	surf-level	surf-level	core-and-perimeter	sched-based (D)	gen-level
UBEM.io	LoD 1	bldg-level	bldg-level	core-and-perimeter	sched-based (D)	gen-level
UMI	LoD 1	bldg-level	bldg-level	core-and-perimeter	sched-based (D)	gen-level
UrbanOpt	LoD 1	surf-level	surf-level	core-and-perimeter	sched-based (D)	gen-level

[a] bldg-level = building-level, surf-level = surface-level
[b] sched-based (D) = schedule-based (deterministic), sched-based (S) = schedule-based (stochastic)
[c] gen-level = generation-level, sub-sys-level = sub-system-level

Specifically, the second column describes the complexity associated with representing urban geometry through different LoDs. The building envelope category refers to the possibility of assigning the thermal characteristics of opaque and transparent building elements either at surface level or at whole-building level. Regarding zoning, the approaches are classified into single-zone, floor-based zoning, and core-and-perimeter zoning configurations. Internal heat gains are categorised according to deterministic or stochastic approaches, combined with schedule-based modelling. Finally, the TBS category refers to the possibility of assessing system efficiency either at generator level or at sub-system level.

3.2.7 Calculation module

In both BEM and large-scale analyses, three main calculation approaches are commonly employed: (i) the monthly quasi-steady-state method, (ii) the simple hourly dynamic method, and (iii) the detailed hourly dynamic simulation. However, following the evolution of regulations on building energy performance assessment, hourly resolution became is the most widespread.

In BEPS tools, heat conduction is typically modelled using one of several formulations: (i) finite difference methods, (ii) lumped parameters approaches, (iii) conduction transfer functions, or (iv) response factor method (De Luca, 2022). Within UBEMs, the building envelope is often represented by a resistive-capacitive (R-C) network, where the degree of discretisation varies among models. Some adopt simplified schemes inspired by the 5R1C model of EN ISO 13790 (CEN, 2008) technical standard, while others employ 2R1C configurations (Kämpf & Robinson, 2007). UBEM tools such as UMI, URBANopt, and CityBES implement the EnergyPlus conduction model, providing a consistent physical basis across applications.

Convective heat transfer in both BEM and in UBEM simulations follows Newton’s law. Mainly, the different models are focused on the convective heat transfer coefficient calculation. External heat transfer phenomena are described by (i) buoyancy-driven heat convection, (ii) forced convection (i.e., wind-driven), or (iii) mixed heat convection (i.e., wind and buoyancy-driven) (De Luca, 2022). Depending on the modelling approach, the coefficient may be treated as constant or dynamically linked to weather conditions. In large-scale models, coupling UBEMs with CFD tools enhances the representation of convective phenomena, offering advantages over single-building simulations.

The solar irradiance distribution within UBEM scenes is evaluated using various algorithms. Some treat buildings as isolated entities, neglecting mutual interactions; others account for shading effects from surrounding structures but omit reflections. More advanced methods incorporate both reflections and shadowing, while often neglecting the absorbed and reflected components of glazing. In detailed algorithms, the sky vault is subdivided into discrete patches, each with defined radiative properties, enabling calculation of the partial or total sky view factor for each surface. In large-scale models, LW radiation exchange depends on the sky temperature, foreground surface temperature, and computed surface temperatures of the urban environment.

In BEMs, sky models vary according to the software implementation but typically include both isotropic and anisotropic formulations, with the latter being preferred for greater accuracy. External longwave radiation is exchanged among (i) the sky vault, (ii) the ground, and (iii) surrounding outdoor objects. In many single-building models, the surface temperature of these external objects is assumed equal to the outdoor air temperature, meaning that LW exchange occurs primarily between the building surfaces, the sky vault, and the ambient environment. In more detailed approaches, the Mean Radiant Temperature (*MRT*) of the external environment is explicitly computed to capture radiative effects more accurately.

3.3 Application

To neutralise any difference related to the mathematical model, the same tool (i.e., CitySim) was used, to analyse both the BEM and the UBEM scenarios. Four scenarios with progressive model complexity have been generated to investigate the mutual energy interactions between buildings.

3.3.1 Case study

The case study concerned a real urban block located in the northwest of Italy, within the municipality of Turin (239 m a.s.l.). The city block consisted of eleven similar apartment blocks (A_1, C_1, B_2, A_3, C_3, B_4, A_5, C_5, B_6, A_7, and C_7; see Figure 7), all constructed after 1960. Each building includes five occupied storeys above ground and is entirely residential, with no commercial activities. The main geometrical characteristics are reported in Table 7.

Table 7. *Geometrical characteristics of the buildings*

Quantity	**Symbol**	**Unit**	**Value**
Building footprint area	$A_{\text{footprint}}$	m^2	397
Overall thermally conditioned floor area	$A_{\text{fl;building}}$	m^2	1166
Overall thermal envelope area	A_{env}	m^2	2348
Gross volume	V	m^3	7387
Compactness ratio	$A_{\text{env}} \cdot V^{-1}$	m^{-1}	0.32

A homogeneous construction typology is adopted across the apartment blocks, consisting of a reinforced concrete skeleton with reinforced concrete and hollow-tile mixed floor, and hollow brick infill wall, as vertical closures. Figure 8 illustrates the *WWR* for each building and façade orientation.

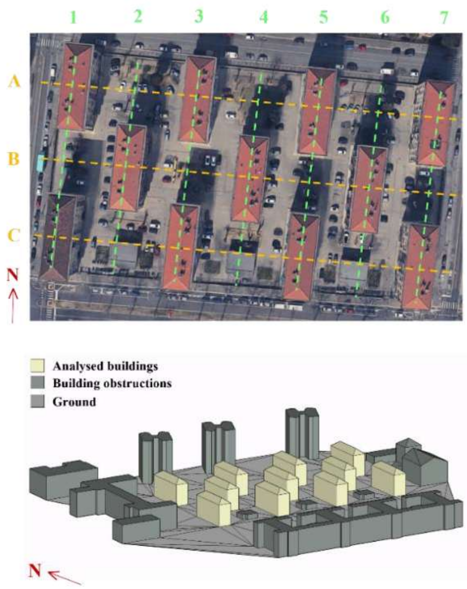


Figure 7. *Satellite view of the city block and 3D geometrical representation (Piro et al., 2023)*

The constructive characteristics of the opaque and transparent building envelope components were derived from the European project TABULA (*Typology Approach for Building Stock Energy Assessment*) (Ballarini et al., 2014). These features were identified based on the climatic zone, construction period, and building size and shape. The thermophysical properties of the materials were extracted from the technical report UNI/TR 11552 (UNI, 2014), which provided an *abacus* of opaque building envelope components. The windows and the French windows of the dwellings composing the apartment blocks were assumed to be single-glazed with wooden frame.

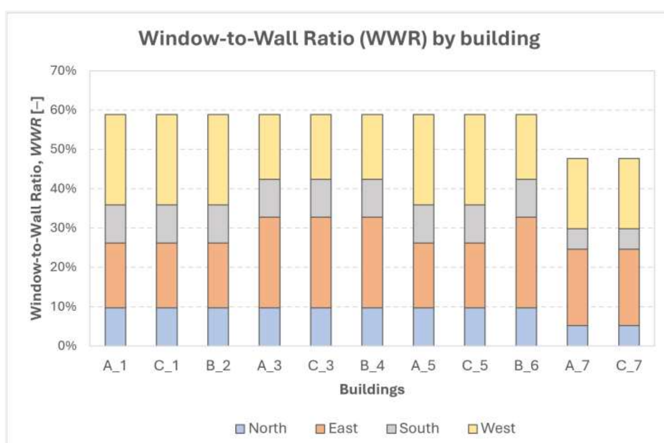


Figure 8. *WWR per façade orientation (Piro et al., 2023)*

The schedules and intensities of internal heat gains—including occupants, electrical appliances, and lighting equipment—were derived from the Italian National Annex of EN 16798-1 (UNI, 2025) and supplemented, where necessary, with data from the EN 16798-1 standard (CEN, 2019). Due to the absence of tailored occupancy data, a deterministic schedule-based approach representative of residential apartment blocks was therefore adopted.

The SW ground reflectance was set to 0.10 (Thevenard & Haddad, 2006). The reflectance coefficient of the analysed buildings depended on the solar absorption coefficient, which was assumed equal to 0.30 for light-coloured buildings (A_1, C_1, A_3, C_3, A_5, C_5, A_7, and C_7), and 0.60 for medium-coloured ones (B_2, B_4, and B_6). The reflectance coefficient for the nearby obstructed buildings was set to 0.20.

3.3.2 Energy models characteristics

This work presents a preliminary investigation into the quantitative energy interactions between buildings, aiming to highlight the complexity of energy exchange within an urbanised environment. Typically, even in detailed BEMs, the analysis focused primarily on the energy performance of a single building, treating the surroundings as a secondary element. In single-building simulations, neighbouring structures were generally modelled as obstructions, sometimes assigned a solar reflection coefficient. However, adjacent buildings could play a significant role—for example, by contributing to solar reflections within the urban context or influencing LW radiative exchanges through surfaces at different temperatures.

Two groups of simulations were carried out in CitySim, focusing on the apartment block located at the centre of the study area (building B_4, see Figure 7). From this perspective, four scenarios with increasing levels of complexity were developed (see Figure 9):

- In the first scenario (BEM_1), only building B_4 was simulated, without any obstructions or surrounding buildings.
- In the second scenario (BEM_2), neighbouring buildings were added as shading elements with a zero-reflection coefficient.
- The third scenario (BEM_3) included the same obstructions but with non-zero reflection coefficients.
- Finally, the fourth scenario (UBEM_1) represented a complete UBEM, where all eleven apartment blocks were explicitly modelled, allowing buildings to both reflect and absorb incident radiation.

The first group of simulations investigated the incident SW solar irradiance on a representative façade of building B_4 across the four models (BEM_1-UBEM_1), thereby analysing the solar irradiation exchanges between buildings.

The second group of simulations examined the variation of LW radiation on the same façade of building B_4, focusing on the BEM_3 and UBEM_1 configurations.

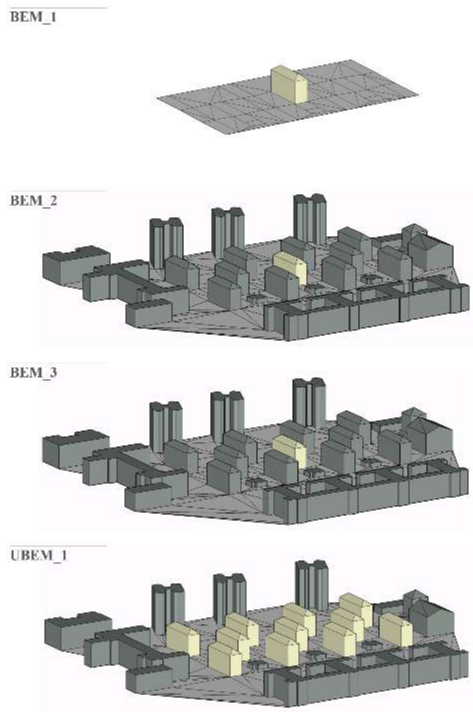


Figure 9. *Calculation scenarios (Piro et al., 2023)*

3.4 Results and discussion

3.4.1 Shortwave radiation

Figure 10 illustrated the mean monthly incident solar radiation on the north façade of apartment block B_4 for the four simulated scenarios. The results of this simulation set were strongly influenced by the shading effects of surrounding buildings and by the reflectance coefficients of the opaque surface elements defining each scenario.

The deviation between the first two configurations (BEM_1 vs. BEM_2) represented the impact of building shading, with variations ranging between – 33 % and – 38 %. The difference between BEM_2 and BEM_3, on the other hand, captured the effect of solar reflections within the urban context caused by neighbouring buildings and obstructions, resulting in an average increase of + 41 %. Finally, for each month, the comparison between BEM_3 and UBEM_1 highlighted the variation introduced by the modelling of transparent components across the entire urban scene.

According to the *Simplified Radiosity Algorithm* (SRA) implemented in CitySim, the reflected SW radiation incident on transparent envelope components was neglected. Consequently, the UBEM_1 scenario exhibited a constant monthly decrease in incident radiation between – 5 % and – 7 % compared to BEM_3. This deviation underlined the importance of a comprehensive representation of the surrounding context. In other words, the solar radiation reflected onto the analysed building in a BEM configuration (i.e., BEM_3) could instead be redirected, through glazing interactions, to influence the energy balance of adjacent buildings in the UBEM_1 scenario, thereby reducing the solar heat gains entering the thermal zone.

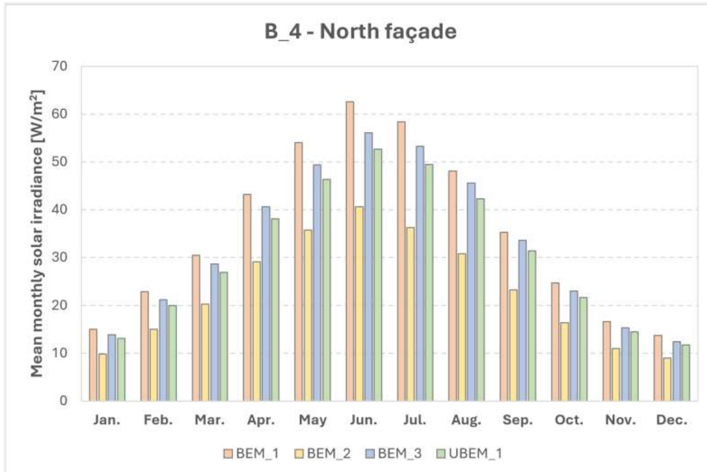


Figure 10. *Mean monthly solar irradiance on the North façade of B_4 (Piro et al., 2023)*

3.4.2 Longwave radiation

As described previously, in BEM the obstruction's surface temperature is set equal to the external air temperature or is determined through the *MRT*. In Figure 11, the mean monthly longwave irradiance exchanged on the West façade of the apartment block B_4 is presented.

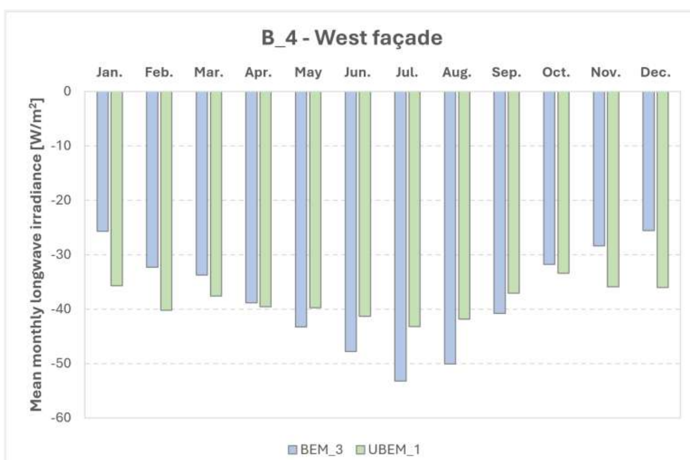


Figure 11. *Mean monthly longwave irradiance on the West façade of B_4 (Piro et al., 2023)*

The SRA integrated into CitySim calculated the mean environmental radiant temperature, which was influenced by several parameters: the sky temperature, the surface temperatures of the urban scene, the view factor between the analysed surface and each sky vault patch, the angle of incidence, and the solid angle (Robinson & Stone, 2005). In the BEM_3 scenario, the surface temperature of the obstructions was assumed to be equal to the default value of 288.15 K. Conversely, in the UBEM_1 scenario, the surface temperatures of neighbouring apartment blocks were computed in detail according to the thermophysical properties of their envelope materials.

As a result, the mean environmental radiant temperature in BEM_3 was higher than in UBEM_1 during the colder months (January-April and October-December), while the trend reversed between May and September. In the coldest months (January and December), the net longwave radiation balance between UBEM_1 and BEM_3 showed the greatest decrease, ranging from – 28 % to – 29 %. In contrast, in July, due to lower environmental radiant temperatures, the single-building model produced values 23 % higher than the large-scale simulation.

In conclusion, the surface temperature of the urban scene played a crucial role in accurately estimating the longwave radiation exchange and the surface temperature of the building under analysis.

3.5 Remarks

This chapter investigated the theoretical differences between UBEM and BEM tools. Simplifications and assumptions related to building geometry, envelope components, zoning, occupant behaviour and internal heat gains, and technical building systems were analysed for both urban-scale and single-building energy models. Furthermore, the main calculation modules were examined with respect to the three heat transfer mechanisms.

Using a city block located in the municipality of Turin (Italy) as a case study, this work provided a preliminary investigation into the SW and LW radiation deviations between UBEM and single-building energy performance simulations. The most significant differences between the two modelling approaches were attributed to the mutual energy interactions occurring among buildings, which influenced district-scale, building-oriented analyses. These interactions also had the potential to modify the local microclimate within the urban area.

In conclusion, it was considered essential to develop appropriate guidelines for both novice and experienced UBEM modellers, to support the exploration of modelling approaches and assumptions and to enhance the accuracy of urban building energy simulations.

Chapter 4

Improving large-scale data quality

4.1 Overview

Part of the work presented in this chapter has been previously published in Ballarini et al. (2023). In Ballarini et al. (2023), the Ph.D. research focused on defining a methodology to identify rules and scores for ranking the quality of energy certificates, as well as on its application to the EPC databases.

As mentioned in Section 2.4, "*Data for building stock energy modelling*", the EPC is a crucial and widely used document for capturing the energy performance of building stock. Energy audits, however, are not widely available, and since these documents are not required to be filed with a local or regional database, collecting them is challenging.

Nevertheless, due to the limitations of the current EPC configuration—such as insufficient energy-related data, static energy performance assessments, and low-quality parameters (Piro et al., 2024)—a methodology for scoring and ranking EPC data quality has been developed to evaluate the reliability of these documents.

This chapter is organised as follows: Section 4.2 illustrates the current system of checks integrated into different Italian regional EPC databases; Section 4.3 presents the EPC data quality-checking procedure developed within the framework of the European TIMEPAC project; Sections 4.4.1 and 4.4.2 describe its application to two Italian regional EPC databases, Piedmont and Aosta Valley, respectively; and Section 4.5 reports conclusion and general remarks, outlining the future characteristics of the next-generation EPC.

4.2 Background analysis

In Italy, the ownership for managing energy performance certificates is delegated to regions. The regions are also responsible for monitoring the quality of EPCs. Specifically, they carry out inspections to verify both the formal compliance and the technical accuracy of the certificates.

To harmonise the structure of the EPCs across regions, the Italian Thermotechnical Committee (CTI), in collaboration with several regional authorities, has developed a proposal for an XML standard for the exchange of EPCs input/output data. This XML (Extensible Markup Language) schema is intended for submitting EPCs to the information registries managed by the region or autonomous provinces. Two versions of the XML file are available: a "*reduced*" format (known as *ridotto* in Italian), which contain only the data included in the energy certificate (national format), and an "*extended*" format (known as *esteso* in Italian), which also includes a set of input data (e.g., building envelope characteristics) and output data (e.g., intermediate and final calculation results).

As an example, Figure 12 shows the data type (integer) and the admissible range for the construction period ("*typeAnno*") of the assessed building, which must fall between 1000 and 2100. However, such a check is insufficient to ensure the completeness and reliability of the data. For instance, historical building fall outside the scope of energy performance assessment, meaning that the lower limit could be raised. Similarly, the upper limit could be restricted by comparing the building's construction period with the certificate's issuing year.

Therefore, the controls embedded in the XML-format are very limited and should be completed by a more advance system of validation checks.

```
<xs:simpleType name="typeAnno">
    <xs:restriction base="xs:unsignedInt">
        <xs:minInclusive value="1000"/>
        <xs:maxInclusive value="2100"/>
    </xs:restriction>
```

Figure 12. *Example of upper and lower limits for building construction period (reduced XML)*

Typically, a syntactic check is followed by one or more semantic assessments aimed at evaluating the quality of the EPC data. The final technical control may then be followed by an on-site building inspection to confirm the plausibility of the information collected and used in the certification process. The syntactic check verifies the formal validity of the document, while the semantic assessment ensures the coherence, compatibility, and the physical admissibility of the certificate.

The scope and complexity of the quantitative checks as well as the data included in the verification procedure, are defined by each region. However, every year the regions are required to verify the quality of a representative sample to at least 2 % of the EPCs issued annually. The selection criteria for this sample vary by region. It may consist of EPCs randomly selected from the previous year or recently issued certificates. For example, the sample may be oriented toward certificates with the highest energy class, EPCs issued for new constructions, or those prepared by assessors with a high annual output of certificates.

The typical steps followed by Italian regions to establish the reliability and quality of EPCs are described below:

- <u>formal checks</u> are intended to verify the syntactic coherence of the data in the XML format. Typically, this coincides with the XSD (XML Scheme Definition) validation performed by the regions when they receive the imported reduced and/or extended XML file. However, the validity of the XSD does not imply the validity of the EPC. No cross-consistency checks are performed, nor are admissibility ranges included, this stage essentially verifies the XML-structure. The parameters checked during this stage may include:
    - the correct version of the file (v.12 for the reduced XML, v.5 for the extended XML),
    - the personal details of the energy assessor,
    - the corresponding ID codes,
    - compliance of the transmitted EPC with the standard (e.g., it must not contain more pages than required),
    - correctness of the relevant municipality codes (region code, postal code, cadastral code, etc.), and more.
- <u>preliminary checks</u> are aimed at carrying out a documentary assessment and verifying compliance with the procedures set out in the National Guidelines (Italian Republic, 2015). The parameters checked during this stage may include:
    - absence of mandatory energy services (e.g., space heating and domestic hot water for residential buildings, or artificial lighting for non-residential uses),
    - absence of at least one recommended energy efficiency measure,
    - use of non-compliant software for the assessed object under certification,
    - missing mandatory on-site inspection and/or incorrect inspection date,
    - likely presence of multiple building units in a single EPC,
    - energy class of newly constructed buildings lower than class E, and more.
- <u>in-depth checks</u> are intended to assess the severity of inconsistencies of one or more EPC data. The parameters subjected to compliance checks are verified against physical or experience-based ranges (see Table 8) and/or compared with representative statistical values (see Table 9).

  Table 8 summaries the admissibility ranges for relevant EPC data in the following Italian regions: Abruzzo (Abruzzo Region, 2023), Aosta Valley (Aosta Valley Region, 2021), and Liguria (Liguria Region, 2018). Most of these include acceptability ranges for geometrical parameters. For example, the compactness ratio (*CR*) and the summer effective solar area per unit of floor area ($A_{sol;est} \cdot A_{use}^{-1}$) illustrate how regions define different lower and upper thresholds for admissibility. In the case of the overall non-renewable energy performance ($EP_{gl;nren}$), the upper limit is established on an experience basis.

Table 8. *Confidence interval of some EPC data for Abruzzo, Aosta Valley, and Liguria regions*

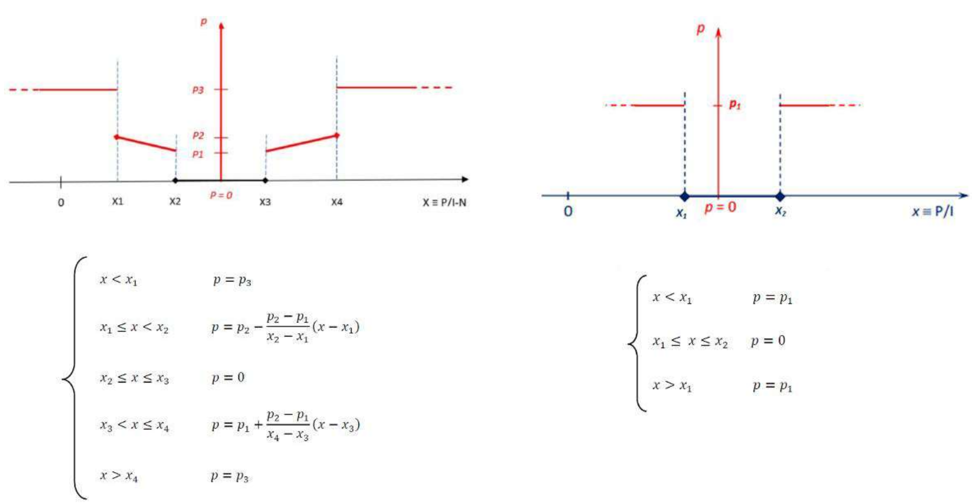
Quantity	Symbol	Unit	Italian region		
			Abruzzo	**Aosta Valley**	**Liguria**
Compactness ratio	CR	m^{-1}	[0.25; 2.0]	[0.2; 1.5]	n/a
Summer effective solar area per unit of floor area	$A_{sol;est} \cdot A_{use}^{-1}$	–	[0.01; 0.8]	n/a	[0.005; 0.8]
Theoretical height for residential building	$V_g \cdot A_{use}^{-1}$	m	[1.8; 3.5]	n/a	n/a
Theoretical height for non-residential building	$V_g \cdot A_{use}^{-1}$	m	[2.5; 7.0]	n/a	n/a
Useful floor area for building unit	A_{use}	m^2	n/a	(0; 300]	n/a
Air change rate (natural ventilation)	n	h^{-1}	n/a	n/a	[0.25; 0.35]
Overall non-renewable energy performance	$EP_{gl;nren}$	$kWh \cdot m^{-2}$	n/a	[0; 1500]	n/a

The Abruzzo and Liguria regions (Abruzzo Region, 2023; Liguria Region, 2018) assign a score to the parameters involved in the analysis by comparing them with their statistical reference values. The assigned score is equal to zero if the parameter complies with the reference, and greater than zero if it does not. Specifically, Table 9 shows the parameters (P), extracted from the EPC, that are compared with their representative statistical values (I), calculated for the specific cluster in Abruzzo and Liguria regions. For the selected energy certificates, the parameters are normalised by their corresponding weighted statistical average value (I), allowing the deviation to be assessed against the tolerance range. If the ratio $P \cdot I^{-1}$ falls within the admissibility range, the score (p) is zero; otherwise, p is calculated using the formulas reported in Figure 13. Accordingly, for some parameters the assigned score p is constant, while for others it is determined using the proposed formulas.

- inspection represents the final step for evaluating the adequacy and consistency of the collected data and the methodology applied in developing the EPC for a building or building unit. The building inspection is a mandatory part of the certification process. An innovation introduced in Annex 3 of the Guideline accompanying EPBD IV (EU 2024/1275) (EPBD, 2025) allows the independent expert to perform a virtual building visit, provided that an appropriate level of quality can be ensured; otherwise, a physical on-site visit is required.

Table 9. *EPC data checked against their statistical representative values*

Quantity	Symbol	Unit	Italian region	
			Abruzzo	**Liguria**
Mean overall heat transfer coefficient by thermal transmission	U_{op}	$\mathrm{W{\cdot}m^{-2}{\cdot}K^{-1}}$	x	x
Mean thermal transmittance of transparent building envelope	U_{w}	$\mathrm{W{\cdot}m^{-2}{\cdot}K^{-1}}$	x	x
Energy need for space heating	$EP_{\mathrm{H;nd}}$	$\mathrm{kWh{\cdot}m^{-2}}$		x
Seasonal space heating energy efficiency	$\eta_{\mathrm{s;H}}$	–	x	
Seasonal space cooling energy efficiency	$\eta_{\mathrm{s;C}}$	–	x	
Seasonal domestic hot water energy efficiency	$\eta_{\mathrm{s;W}}$	–	x	
Non-renewable energy performance per space heating	$EP_{\mathrm{H;nren}}$	$\mathrm{kWh{\cdot}m^{-2}}$	x	x
Non-renewable energy performance per space cooling	$EP_{\mathrm{C;nren}}$	$\mathrm{kWh{\cdot}m^{-2}}$	x	
Non-renewable energy performance per domestic hot water	$EP_{\mathrm{W;nren}}$	$\mathrm{kWh{\cdot}m^{-2}}$	x	x
Non-renewable energy performance per artificial lighting	$EP_{\mathrm{L;nren}}$	$\mathrm{kWh{\cdot}m^{-2}}$	x	
Overall non-renewable energy performance	$EP_{\mathrm{gl;nren}}$	$\mathrm{kWh{\cdot}m^{-2}}$	x	x



(a) **(b)**

Figure 13. *Trend of weight p as a function of the Parameter/Indicator: linear and constant (a), and constant (b) (Abruzzo Region, 2023)*

> The virtual visit must cover all relevant parts of the building or building unit (e.g., cellar, heating system, roof, at least one flat, and a clear view of windows and doors). In practice, a virtual visit involves the building owner or their representative (such as the building manager or construction supervisor) connecting remotely with the independent energy expert, for instance via a video-call platform.

In the event of a sanction, regional regulations provide for the application of a fine, direct payment or by means of an administrative warning, in accordance with Italian legislation (Italian Republic, 2005).

4.3 Data quality-checking procedure

An EPC data quality-checking procedure was initially developed within the framework of the EU H2020 TIMEPAC project and subsequently updated. This procedure involves scoring the EPC against a maximum error threshold, beyond which the certificate is deemed *unreliable*. Energy certificates that received an acceptable quality label have been used to create the representative dataset included in the building archetype schema.

4.3.1 The TIMEPAC project

TIMEPAC (*Towards Innovative Methods for Energy Performance Assessment and Certification of Buildings*, 2021-2024) (TIMEPAC, 2025) is an EU-funded project within the Next Generation EPC Horizon2020 cluster—a working group dedicated to advancing and improving the current energy certification framework.

Within the framework of Next Generation EPC Horizon2020 cluster (European Commission, 2024b), the EU funded several projects aimed at stimulating the evolution and large-scale deployment of future certification schemes. This group of projects focuses on enhancing the current energy certificate to make it more reliable, user-friendly, cost-effective, comparable across regions, and compliant with EU legislation. Projects such as X-tendo (Zuhaib et al., 2022), D^{2}EPC (Seduikyte et al., 2022), EUB SuperHub (Malinovec Puček et al., 2023), U-CERT (Carnero et al., 2022), and ePANACEA (Verheyen et al., 2022) contribute to enriching the current EPC framework with new Key Performance Indicators (KPIs), improving the usability and functionality of next-generation EPCs.

In particular, experience gained through the EUB SuperHub project led to the establishment of a CEN Workshop Agreement (CEN, 2024), which defines a harmonised core set of 21 transnational indicators for comparing building performance across European regions. The E-DYCE (Larsen et al., 2023) and ePANACEA (Verheyen et al., 2022) projects further explored the integration of standard energy performance assessments with actual building use. Meanwhile, X-tendo (Zuhaib et al., 2022) developed an automated scoring algorithm for validating energy certificates, thereby improving data quality and consistency.

TIMEPAC proposes a new generation of EPCs conceived as holistic, flexible, and through-life certification process. The TIMEPAC approach promotes a continuous flow of data (Sučić et al., 2022) across all stages of the certification process—generation, storage, analysis, and exploitation (Figure 14)—thus enabling a more efficient, transparent, and reliable EPC.

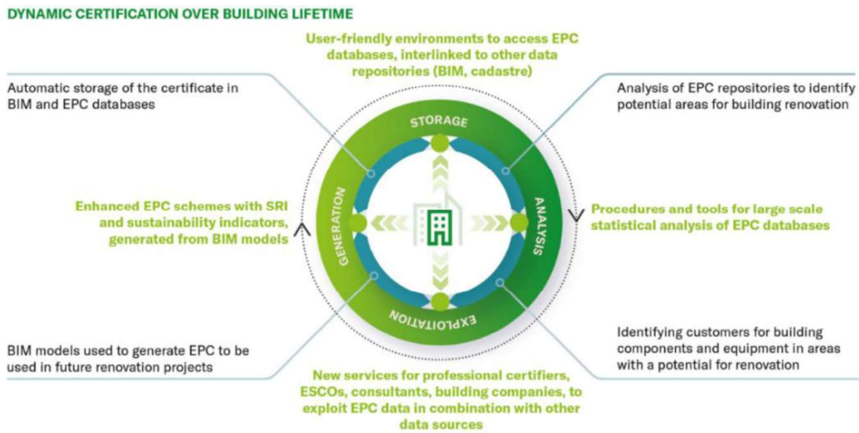


Figure 14. *TIMEPAC project flow (Sučić et al., 2022)*

The Transversal Deployment Scenarios (TDSs) constitute the core of the TIMEPAC project, focusing on the development of standardised procedures and digital tools across the six partner countries—Austria, Croatia, Cyprus, Italy, Spain, and Slovenia. The TDSs are interrelated (Figure 15) and address specific stages of the EPC process—generation, storage, analysis, and exploitation—while pursuing distinct objectives and engaging various target groups such as energy certifiers, energy agencies, auditors, architects, engineers, real estate agencies, construction companies, public authorities.

The project defines five TDSs:

i. Generating enhanced EPCs using BIM data, establishing a direct link between the generation of energy certificates and BIM models. This connection enables data enrichment through improved interoperability between BIM and the BEM environments,
ii. Enhancing EPC schemas through operational data integration, aimed at improving the accuracy and reliability of EPCs by incorporating real energy consumption data and other relevant information domains, such as IEQ, economic, and sustainability indicators,
iii. Creating Building Renovation Passport (BRP) from data repositories, to evaluate track building performance over time by comparing energy performance levels and the cost-effectiveness of different energy efficiency measures,
iv. Integration the Smart Readiness Indicator (SRI) and sustainability indicators into EPCs, assessing a building's smartness and its ability to

maintain indoor comfort, while incorporating new and existing indicators derived from the Level(s) framework, and finally

v. Conducting large-scale statistical analyses of EPC databases, examining regional and national datasets to leverage EPC information for supporting deep renovation strategies across the building stock. This TDS also addresses the enhancement of EPC data quality.

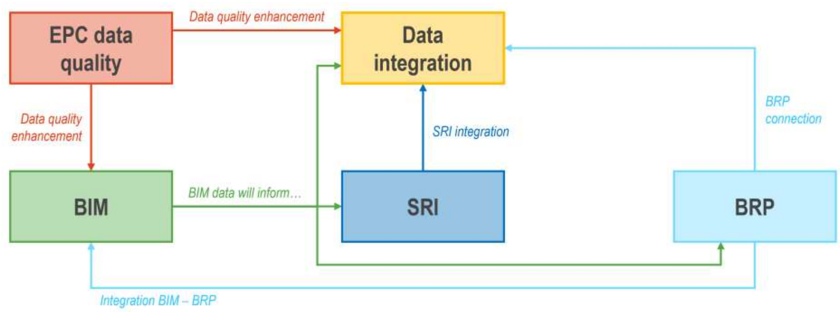


Figure 15. *TDSs connection in the TIMEPAC project*

4.3.2 Methodology

The EPC data quality-checking procedure was originally developed within the framework of the fifth TDS, "*Large scale statistical analysis of EPC databases*". The main objectives of this TDS are to support EPC data analysis and exploitation within the holistic EPC enhancement approach proposed by the TIMEPAC project. Specifically, this TDS aims to:

i. check and improve the quality of EPC data,
ii. exploit EPCs to perform energy balance analyses of the building stock through the definition of building archetypes, and
iii. provide relevant stakeholders with a methodology to conduct reliable refurbishment scenario assessments of their building portfolios.

The overall methodology, illustrated in Figure 16, originates from the analysis of regional or national EPC databases of the TIMEPAC partner countries. Once a common dataset is established, the subsequent work includes:

- clustering energy performance certificates, performing the EPC quality-checking, and conducting statistical analyses. The statistical analysis then serves as the cornerstone for:
    a. defining representative buildings (i.e., building stock archetypes) to benchmark the current energy status and develop renovation scenarios for urban areas, and
    b. developing a regression model and estimating confidence intervals for selected EPC input parameters.

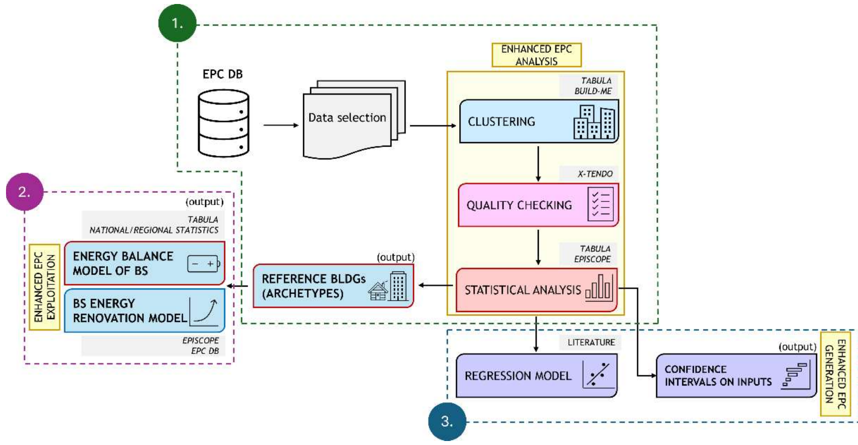


Figure 16. *Methodology flowchart of the TDS "Large scale statistical analysis of EPC databases"*

The clustering approach, which has been derived from the outcomes of the TABULA and EPISCOPE projects (IEE Project, 2013), plays an important role in grouping buildings with similar thermo-physical properties of the opaque and transparent building envelope, as well as comparable technical building system characteristics. In this work, the clustering approach refers to the manual grouping of certificates belonging to the same cluster. Once the clustering is performed, the EPC data quality-checking procedure is carried out. The procedure is based on a score attribution system that ensures the reliability of EPCs with a limited overall score error. This methodology draws inspiration from the X-tendo sister project, especially for the EPC data score attribution. The subsequent statistical analysis is carried out including only the EPCs with an acceptable scoring.

The statistical analysis is carried out for two important following phases that are:

- the generation of the building archetypes (section 1. represented in Figure 16),
- the generation of the semi-empirical models that deliver confidence intervals and distribution functions of the main input data of the EPC (section 3. represented in Figure 16).

In turn, the archetypes are adopted in the building stock energy model (section 2. represented in Figure 16). This model encompasses the current energy balance of the building stock and the renovation scenarios performed by means of a dedicated tool.

In the following sections, all the steps for generating reference buildings will be discussed.

4.3.2.1 EPC data selection

The aim of the EPC data selection is to identify the set of input and output metrics to be incorporated into the building archetype schema. These parameters were derived from regional or national energy certificates to establish a shared, harmonised, and replicable methodology for the subsequent operative phases.

Table 10 provides an overview of all the parameters considered in the analysis, along with their corresponding data types. Mostly of the data types fall into three main categories: string, numeric, and Boolean values. In cases where certain metrics listed in Table 10 are not available within a regional or national energy certificate, appropriate assumptions must be introduced, or—where possible—data should be retrieved from alternative sources or complementary databases.

Table 10. *EPC data types considered in the analysis*

Data name	Symbol	Unit	Data type
EPC ID code			string
Building use category			string
Assessed object			string
Building city			string
Number of building units	n_{unit}	–	number (integer)
Climatic region			string
Heating degree days	HDD	°C·day	number (integer)
Year of construction		–	number (integer)
Year of last renovation		–	number (integer)
Building typology			string
Building property			string
Building constructive typology			string
Thermally heated/cooled gross/net volume	$V_{H/C;g/n}$	m^3	number (decimal)
Thermally heated/cooled gross/net area	$A_{H/C;g/use}$	m^2	number (decimal)
Thermal envelope area	A_{env}	m^2	number (decimal)
Opaque/transparent thermal envelope area	$A_{env;op/w}$	m^2	number (decimal)
Thermal envelope area of walls/roof/floor	$A_{wl/fl;up/fl;lw}$	m^2	number (decimal)
Compactness ratio	CR	m^{-1}	number (decimal)
Mean thermal transmittance of the total/opaque/transparent building envelope	$U_{m;op/w}$	$W \cdot m^{-2} \cdot K^{-1}$	number (decimal)
Mean thermal transmittance of wals/roof/floor	$U_{wl/fl;up/fl;lw}$	$W \cdot m^{-2} \cdot K^{-1}$	number (decimal)
Glazing/frame system typology			string

Data name	Symbol	Unit	Data type
Heating/cooling/domestic hot water/ventilation/artificial lighting/transportation energy service			Boolean value
Heating/cooling/domestic hot water typology			string
Energy carrier for heating/cooling/domestic hot water/ventilation/artificial lighting/transportation			string
Main TBS type of space heating generator			string and/or number (integer)
Overall mean seasonal efficiency of the heating/cooling/domestic hot water system	$\eta_{s;H/C/W}$	–	number (decimal)
Mean seasonal efficiency of the heating generation/distribution/control/emission sub-system	$\eta_{H;gn/d/c/e}$	–	number (decimal)
PV/thermal solar system typology			string
Installed PV/thermal solar system power	$W_{PV/sol}$	kW	number (decimal)
Area of solar collectors	A_{coll}	m^2	number (decimal)
Energy need for heating/cooling/domestic hot water	$EP_{H/C/W;nd}$	$kWh \cdot m^{-2}$	number (decimal)
Non-renewable energy performance per heating/cooling/domestic hot water	$EP_{H/C/W;nren}$	$kWh \cdot m^{-2}$	number (decimal)
Overall renewable energy performance	$EP_{gl;ren}$	$kWh \cdot m^{-2}$	number (decimal)
Overall non-renewable energy performance	$EP_{gl;nren}$	$kWh \cdot m^{-2}$	number (decimal)
Delivered natural gas/district heating/solid biomass/thermal solar/PV system/electricity	$Q_{X;Y;in}$ $E_{X;Y;in}$	Nm^3 Sm^3 kg kWh	number (decimal)

4.3.2.2 EPC data aggregation

The EPC data aggregation aims to group buildings with similar thermo-physical properties of the opaque and transparent building envelope, as well as comparable TBSs characteristics. This approach involves assigning to an urbanised area real or theoretical reference buildings, both residential and non-residential, with known characteristics and energy consumption profiles. These reference buildings are useful for assessing energy and environmental performance on a larger scale.

The methodology for building clustering is derived from the outcomes of the TABULA project, which aimed to harmonise the European building typology, and its follow-up project, EPISCOPE, which focused on building stock monitoring. As shown in Figure 17, the defined clusters are based on climatic zone, the building use category, the year of construction, and the building size and shape for residential buildings. Specifically, these criteria are related to the following aspects:

- Climatic data: climate significantly affects a building's energy demand. In each country, the minimum energy performance requirements may vary from colder regions, where insulation levels are crucial, to warmer climates, which rely more on cooling systems and solar control strategies.
- Building use category: different building types (e.g., residential, office, school, and hospital) have distinct operational schedules, occupancy patterns, and internal heat gains (from occupants, lighting, and equipment). Additionally, user expectations and comfort requirements vary depending on the building's function.
- Construction period: building codes, materials, and construction practices evolve over time, greatly influencing thermal performance and energy efficiency. The construction period is essential for estimating the presence and quality of insulation in building envelope components. This factor is closely linked to the implementation of minimum energy performance requirements and their subsequent updates.
- Building size and shape: the building's form factor (surface-to-volume ratio) affects heat transfer through the envelope. Furthermore, building size can influences operational schedules and the intensity of internal heat gains. For example, kindergartens, elementary schools, and high schools all fall under the category of educational buildings, but the indoor comfort expectations and needs of their users differ.

The clustering approach for each EPCs follows a sequential procedure:

1 grouping by climatic zone: buildings are first grouped according to their climatic zone (a single country or region may encompass multiple climatic zones).
2 grouping by building use category: within each climatic zone, buildings are further categorised by their intended use (residential and non-residential), covering a wide range of building types.
3 grouping by construction period or building size: within each building use cluster, buildings are grouped based on either the construction period or building size, depending on specific criteria (e.g., the share of existing buildings represented, etc.).
4 further subdivision: within the innermost cluster, additional groupings can be performed according to the specific analysis scope (e.g., based on *U*-value ranges, type of space heating system, etc.).

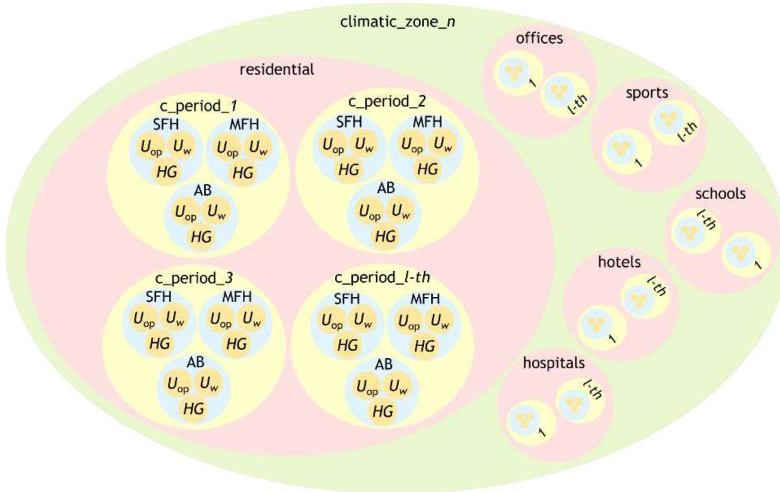


Figure 17. *EPC data aggregation methodology*

The classification of building use, reported in Table 11, have been derived from the technical standard EN ISO 52000-1 (CEN, 2017f), complemented by the outcomes of the TABULA and EPISCOPE projects. These building use categories are the most representative for summarising the building stock.

The proposed approach is primarily oriented towards single-use or predominantly single-use buildings, where one main activity determines the energy behaviour of the whole building. For this reason, the categories listed in Table 11 mainly refer to entire buildings (e.g., residential apartment blocks, offices, educational buildings, or hotels). This assumption facilitates the definition of representative archetypes and reduces the complexity associated with modelling mixed-use configurations at urban scale.

Nevertheless, the methodology does not preclude the development of mixed-use archetypes (e.g., residential buildings with offices or commercial spaces on the ground floor) or the inclusion of additional building uses, such as retail and commercial buildings. These configurations may be incorporated depending on the objectives of the analysis, the availability of input data, and the desired level of detail in the UBEM framework.

Regarding the third cluster step, the reference construction periods were derived from TABULA and are listed for Italy in Table 12. These periods reflect the main historical and regulatory phases that characterised the evolution of the Italian building stock: the years before and during the two World Wars, the post-war reconstruction period (1946-60), the years preceding the oil crisis (1961-75), the introduction of the first Italian energy efficiency regulation, Law 373/1976 (1976-90), the progressive implementation of energy performance regulations from Law 10/1991 (1991-05), to Legislative Decree 192/2005 and finally the more stringent energy performance requirements introduced by the Ministerial Decree of 26 June 2015 (post-2005).

Table 11. *EPC data types considered in the analysis*

Use category	Type	Description
Residential	*BLDNGCAT_RES_SINGLE*	Single-family houses of different types
	BLDNGCAT_RES_APPBLOCK	Apartment blocks (multi-family houses included)
Non-residential	*BLDNGCAT_OFF*	Offices
	BLDNGCAT_EDUC	Educational buildings
	BLDNGCAT_HOSP	Hospitals
	BLDNGCAT_HOTEL	Hotels and restaurants
	BLDNGCAT_SPORT	Sports facilities

Table 12. *Construction periods (TABULA-EPISCOPE)*

Country	Construction periods							
Italy	≤ 1900	1901-20	1921-45	1946-60	1961-75	1976-90	1991-05	> 2005

4.3.2.3 EPC data quality-checking

The EPC data quality-checking procedure is a fundamental phase to have reliable energy certificates free of inaccuracies for carrying out large-scale energy analyses. The deep renovation of building stock, in order to achieve the EU 2030 and 2050 GHG emission reduction goals, goes through the reliable creation of the urban building's energy models. In this regard, the energy certificate represents a crucial source of data to build urban energy analysis. Undoubtedly, the EPC information should be enhanced, real-time upgradeable, implemented, and made interoperable. However, it is crucial to establish the validity of the EPC data already contained in the energy certificate.

The aim of the EPC data quality procedure is to derive only the reliable energy certificates, whose information will be processed in the subsequent phase of the analysis. This methodology draws inspiration from the X-tendo project (Maia et al., 2022), especially for the EPC data score attribution. The proposed procedure is applicable to the energy certificates still uploaded into the reference local, regional, or national EPC databases. Moreover, it can be implemented in the EPC database system in the future, as a support to validation and control bodies in identifying the errors contained in the document, but also as a support to energy certifiers in providing warnings on input data in the EPC generation phase.

The EPC data quality-checking procedure provides the score attribution to parameters and values contained in the energy certificates. For each of the EPC data, a validity rule has been associated. For each rule, a score has been attributed to the non-respected rule. Therefore, the overall score, originating from summing the single score of each parameter, will be compared to the cut-off value. This value represents the acceptability threshold value for all energy certificate parameters. Whether the overall EPC data score is greater than the threshold value then all the EPC data will be neglected, and they will not appear in the subsequent phases (e.g., statistical analysis and reference buildings creation).

In Table A.1, in Appendix A some metrics have been defined as '*critical*' (C), whose validity is considered fundamental for statistical analysis and reference building generation, and '*non-critical*' (n-C) parameters. The non-compliance rules for critical parameters give a score greater than the threshold value and thus every EPC data will be neglected. The inaccuracies of these parameters drastically affect the reliability of the entire procedure, so they will not be considered. Moreover, the single-non-critical parameter with a non-null score (i.e., non-respected rule) will be discarded. The overall picture of EPC data involved in the energy certificate data quality-checking procedure, rules, and scores is shown in Table A.1, in Appendix A.

Three groups of rules are summarised as follows:

- (D) - Data types of checks are the simplest control and represent that set of rules which evaluate the data types (i.e., integer, string, Boolean value, etc.) of the data analysed. For instance, the compactness ratio is a decimal number and not a string, or the year of construction of the building must be an integer and not a Boolean value. These rules have been associated with both critical and non-critical parameters,
- (P) - Physical impossibility checks represent that set of rules which evaluate the order of magnitude of EPC data comparing them with the physical admissibility for that parameter. For example, the thermal energy need for space heating cannot be a negative value, or the thermal transmittance of the building envelope opaque component cannot be greater than a threshold value (e.g., 4.50 $W/(m^2 \cdot K)$). These rules have been associated with both critical and non-critical parameters, and
- (C) - Consistency checks represent that set of rules which determine the validity of a parameter compared to another one. For example, the energy carrier for the space cooling service should be invalid or null whether the energy service flag for space cooling is set to "NO" or "FALSE". The consistency checks could also refer to national regulations (e.g., the minimum average height of a building should not be less than reasonable values). These rules have been associated with both critical and non-critical parameters.

For each rule, scores begin at zero and increase as the validity rules are not met. The score depends on the number of critical and non-critical parameters. However, the magnitude of the non-respected rule score is different between critical and non-critical parameters.

Considering n as the total number of EPC data and m as the total amount of critical parameters selected within the set of EPC information (so that $n - m > 0$), the score for the non-critical parameter ($s_{\mathrm{n-C}}$) has been calculated as reported in Equation (1).

$$s_{\mathrm{n-C}} = \frac{1}{(n - m)} \tag{1}$$

Indeed, the invalidity score border, i.e., the threshold value beyond which an energy certificate, and all data contained therein, is considered rejected, is calculated as reported in Equation (2). The acceptability threshold value corresponds to the sum of the score of half of the EPC data considered in the analysis.

$$e = {}^{(n \cdot s_{\mathrm{n-C}})}\!/_{2} \tag{2}$$

The score for the non-respected rule for critical parameters (s_{C}) is set to 1, i.e., a value greater than the threshold one, as reported in Equation (3).

$$s_C = 1 \tag{3}$$

So, whether the overall score (s_{EPC}), calculated according to Equation (4), of the EPC is greater (>) than e then the EPC is rejected and will not be considered in the subsequent statistical analysis. Instead, as presented in Equation (5), if the overall score of the EPC is less or equal (≤) to e then the energy certificate is considered valid with reliable EPC data.

$$s_{\mathrm{EPC}} = \sum_{n-m} s_{\mathrm{n-C};n-m} + \sum_{m} s_{\mathrm{C};m} \tag{4}$$

$$\begin{cases} s_{\mathrm{EPC}} \leq e \rightarrow \text{accepted EPC} \\ s_{\mathrm{EPC}} > e \rightarrow \text{rejected EPC} \end{cases} \tag{5}$$

Obviously, the proposed procedure is a first check about the reliability of the EPC data, because energy certificates with respected rule outliers for some parameters may progress in the later stage of the work development. Mainly, the EPC data quality-check procedure sets lower limits because a priori it is difficult to set upper limits.

Table A.1 (Appendix A) reports the assessed EPC data, the level of criticality (i.e., critical or non-critical parameter), the typology and assigned rule. The highest level of criticality was assigned for the variables listed in Table 8 and Table 9.

As shown in Table A.1 (Appendix A), the following EPC data are deemed critical (C) with the corresponding reasons:

- EPC ID code: uniquely identify the energy certificate in the database,
- building use category, climatic region, and year of construction: essential data segmentation,
- thermally heated/cooled gross volume ($V_{H/C;g}$): correlated with useful floor area to approximate building height,
- thermally heated/cooled floor area ($A_{H/C;use}$): key geometric metric used to normalise energy performance indicators,
- thermal envelope area (A_{env}): required for accurately assessing heat losses through the building envelope components,
- mean thermal transmittance of opaque/transparent building envelope ($U_{m;op/w}$): directly correlated with thermal energy need for space heating,
- energy need for space heating ($EP_{H;nd}$): closely linked to heating energy demand, particularly relevant in winter-dominated climates such as Po Valley, alpine and sub-alpine areas (zones E and F, Italian legislation), and
- overall non-renewable energy performance indicator ($EP_{gl;nren}$): determines the energy class of the building or building unit.

As reported in Table 9, most of the proposed parameters coincide with those used in regional EPC data quality-checks against their statistical reference values.

4.4 Application

Regions extract different EPC parameters from the uploaded XML-format, either "*extended*" or "*reduced*" version. The EPCs of the Piedmont region are stored in the Information System for the Energy Performance of Buildings (SIPEE – *Sistema Informativo per la Prestazione Energetica degli Edifici*), while those of the Aosta Valley are managed within BEAUCLIMAT platforms. In 2017, following the transposition of the new Italian regulation on building energy performance calculation (M. D. 26 June 2015), the latter system was updated with a revised set of EPC data. However, thanks to the detailed description of the thermal building envelope components, both the pre- and post-2017 Aosta Valley EPC databases were considered in the analysis.

Table 13 compares the available data in Piedmont and Aosta Valley, distinguishing for the latter between *ante*- (*a-2017*) and *post*-2017 (*p-2017*). In particular, the *ante*-2017 EPC database includes *U*-values for different opaque envelope components (walls, roof, and floor), further differentiating their configuration (e.g., attic, pitched and flat roofs). For windows, the thermal transmittance is complemented by information on frame and glazing typologies. Such details enhance the usability and completeness of the building archetype schema, which in most other cases only provides mean U-values for opaque and transparent envelope components.

Sections 4.4.1 and 4.4.2 present the results of the EPC data quality-checking procedure, which are further discussed in the next chapter. Reported output include error frequencies (for both critical and non-critical parameters) and the overall EPC score for each database. Finally, significant quantitative and qualitative EPC data for residential buildings are presented across different climatic zones, while maintaining the same pre-/post-2017 distinction for the Aosta Valley dataset.

Table 13. *EPC data in Piedmont and Aosta Valley databases*

Field	EPC data	Symbol	Unit	Piedmont	Aosta Valley	
					***a*-2017** (*)	***p*-2017** (*)
General data	EPC ID code			x	x	x
	Building use category			x	x	x
	Assessed object			x	x	x
Identifying data	Building city			x	x	x
	No. building units	n_{unit}	–	x	x	x
	Climatic region			x	x	x
	Heating degree days	HDD	°C·d		x	x
	Year of construction		–	x	x	x
	Year of last renovation		–	x	x	x
	Building typology			x	x	x
	Building property			x	x	x
	Building constructive typology			x	x	x
Geometrical data	Thermally heated gross volume	$V_{H;g}$	m^3	x		x
	Thermally cooled gross volume	$V_{C;g}$	m^3	x		x
	Gross volume	V_g	m^3		x	
	Net volume	V_n	m^3		x	
	Thermally heated floor area	$A_{H;use}$	m^2	x		x
	Thermally cooled floor area	$A_{C;use}$	m^2	x		x
	Useful floor area	A_{use}	m^2		x	
	Gross floor area	$A_{fl;g}$	m^2	x	x	
	Thermal envelope area	A_{env}	m^2	x	x	x
	Opaque thermal envelope area	$A_{env;op}$	m^2	x	x	x
	Thermal envelope area of walls	A_{wl}	m^2		x	
	Thermal envelope area of roof	$A_{fl,up}$	m^2		x	

Field	EPC data	Symbol	Unit	Piedmont	Aosta Valley *a*-2017 (*)	Aosta Valley *p*-2017 (*)
	Thermal envelope area of floor	$A_{fl,lw}$	m^2		x	
	Transparent thermal envelope area	$A_{env;w}$	m^2	x	x	x
	Compactness ratio	CR	m^{-1}	x	x	x
Envelope	Mean thermal transmittance of opaque building envelope	$U_{m;op}$	$W \cdot m^{-2} \cdot K^{-1}$	x	x	x
	Mean thermal transmittance of walls	U_{wl}	$W \cdot m^{-2} \cdot K^{-1}$		x	
	Mean thermal transmittance of roof	$U_{fl,up}$	$W \cdot m^{-2} \cdot K^{-1}$		x	
	Mean thermal transmittance of floor	$U_{fl,lw}$	$W \cdot m^{-2} \cdot K^{-1}$		x	
	Glazing system typology				x	
	Frame system typology				x	
	Mean thermal transmittance of transparent building envelope	$U_{m;w}$	$W \cdot m^{-2} \cdot K^{-1}$	x	x	x
Technical building system	Space heating energy service			x	x	x
	Space cooling energy service			x	x	x
	Domestic hot water energy service			x	x	x
	Ventilation energy service			x		
	Artificial lighting energy service			x		
	Transportation energy service			x		
	Space heating system typology				x	x
	Space cooling system typology				x	x
	Space domestic hot water system typology				x	x
	Energy carrier (k) for space heating			x	x	x
	Energy carrier (k) for space cooling			x	x	x
	Energy carrier (k) for domestic hot water			x	x	x

Field	EPC data	Symbol	Unit	Piedmont	Aosta Valley *a*-2017 (*)	Aosta Valley *p*-2017 (*)
	Energy carrier (k) for ventilation			x		
	Energy carrier (k) for artificial lighting			x		
	Energy carrier (k) for transportation			x		
	Autonomous/centralised space heating system				x	x
	Autonomous/centralised domestic hot water system				x	x
	Overall mean seasonal efficiency of the space heating system	$\eta_{s;H}$	–	x		x
	Overall mean seasonal efficiency of the domestic hot water system	$\eta_{s;W}$	–	x		x
	Overall mean seasonal efficiency of the space cooling system	$\eta_{s;C}$	–	x		x
	Mean seasonal efficiency of the heating generation sub-system	$\eta_{H;gn}$	–	x		x
	Mean seasonal efficiency of the heating distribution sub-system	$\eta_{H;d}$	–	x		x
	Mean seasonal efficiency of the heating control sub-system	$\eta_{H;c}$	–	x		x
	Mean seasonal efficiency of the heating emission sub-system	$\eta_{H;e}$	–	x		x
	PV system typology					x
	Installed PV system power	W_{PV}	kW	x		x
	Thermal solar system typology					x
	Area of solar collectors	A_{coll}	m^2			x
	Installed thermal solar system power	W_{sol}	kW	x		x
Energy data	Energy need for space heating	$EP_{H;nd}$	$kWh \cdot m^{-2}$	x	x	x
	Energy need for domestic hot water	$EP_{W;nd}$	$kWh \cdot m^{-2}$	x		x

Field	EPC data	Symbol	Unit	Piedmont	Aosta Valley	
					a-2017 (*)	*p*-2017 (*)
	Energy need for space cooling	$EP_{\mathrm{C;nd}}$	kWh·m^{-2}	x	x	x
	Non-renewable energy performance per space heating	$EP_{\mathrm{H;nren}}$	kWh·m^{-2}	x		x
	Non-renewable energy performance per domestic hot water	$EP_{\mathrm{W;nren}}$	kWh·m^{-2}	x		x
	Non-renewable energy performance per space cooling	$EP_{\mathrm{C;nren}}$	kWh·m^{-2}	x		x
	Overall renewable energy performance indicator	$EP_{\mathrm{gl;ren}}$	kWh·m^{-2}	x		x
	Overall non-renewable energy performance indicator	$EP_{\mathrm{gl;nren}}$	kWh·m^{-2}	x		x
Delivered and produced energy	Delivered natural gas	$Q_{\mathrm{X;Y;in}}$	Nm^3 Sm^3	x		
	Delivered electricity	$E_{\mathrm{X;Y;in}}$	kWh	x		
	Delivered district heating	$Q_{\mathrm{X;Y;in}}$	kWh	x		
	Delivered solid biomass	$Q_{\mathrm{X;Y;in}}$	kg	x		
	Electrical energy produced by PV system	$E_{\mathrm{X;Y;in}}$	kWh	x		x
	Thermal energy produced by solar system	$Q_{\mathrm{X;Y;in}}$	kWh	x		x

(*) *a*-2017 = *ante*-2017; *p*-2017 = *post*-2017

Due to data availability constraints in the most recent EPC databases of the Piedmont and Aosta Valley regions, the conditioned net volume is not available. This limitation hinders the assessment of building height reliability and its comparison with the minimum height requirements established by Italian health and hygiene regulations. To address this issue, the conditioned useful floor area ($A_{\mathrm{H/C;use}}$) is converted into conditioned gross floor area ($A_{\mathrm{H/C;g}}$) using the simplified conversion factor introduced in UNI/TS 11300-1 (UNI, 2008), as follows:

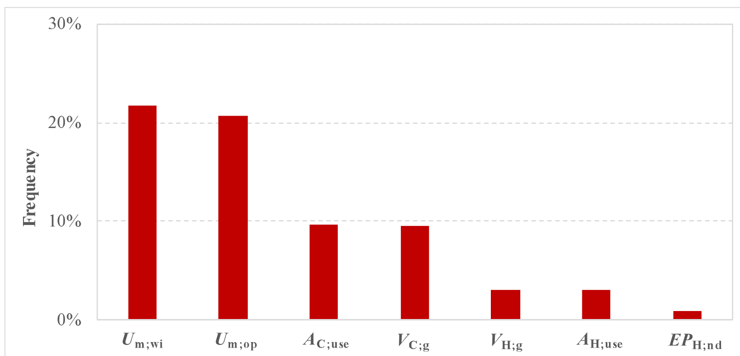
- $f_{\mathrm{n}} = 0.9761 - 0.3055 \cdot d_{\mathrm{m}}$
- where d_{m} represents the average thickness of the vertical enclosures, assumed equal to 0.45 m

4.4.1 Piedmont Region (Italy) energy certificate database

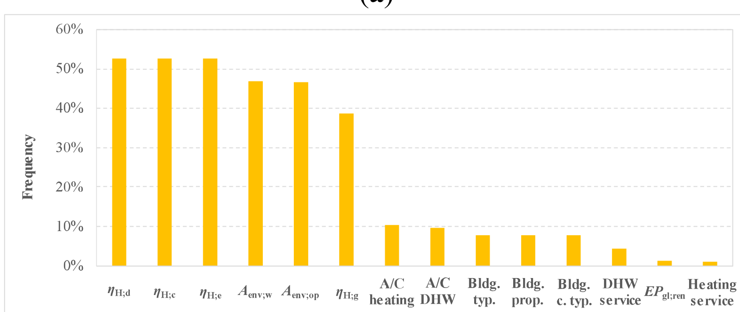
This section reports the outcomes of the EPC data quality-checking procedure applied to energy certificates from the Piedmont regional database.

Figure 18 shows the frequency of invalid rules for both critical and non-critical parameters, while Figure 19 presents the overall EPC score distribution, highlighting the shares of accepted and rejected certificates. In this latter case, green bars correspond to EPCs that passed the quality checks, while red bars indicate those rejected due to missing or inconsistent data.

For the *reliable* EPCs, the results are then discussed in terms of building constructive typologies, thermal properties of opaque and transparent envelope components, energy carriers, and TBS performance. All findings are reported separately for climatic zones E (from Figure 18 to Figure 23) and F (from Figure 24 to Figure 27).



(**a**)



(**b**)

Figure 18. *Error frequencies in Piedmont Region for critical (**a**) and non-critical parameters (**b**)*

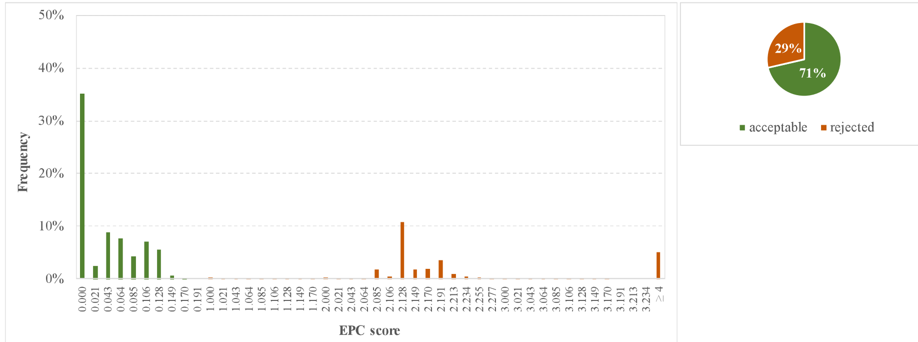


Figure 19. *Frequency distribution of EPC scores in the Piedmont Region: the green bar represents acceptable EPCs, while the red bars indicate non-acceptable EPCs*

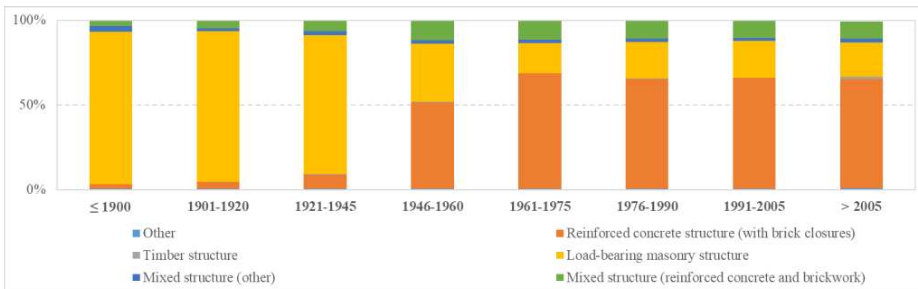
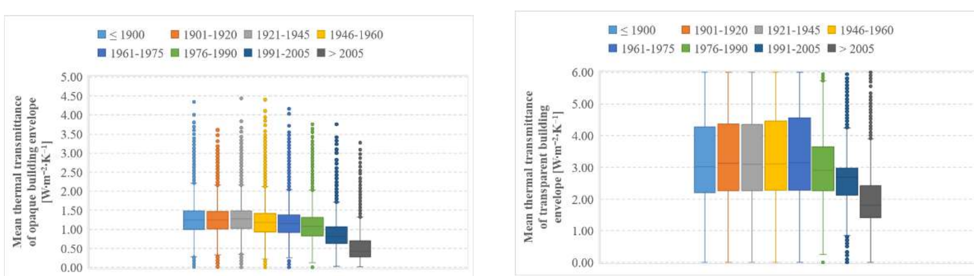


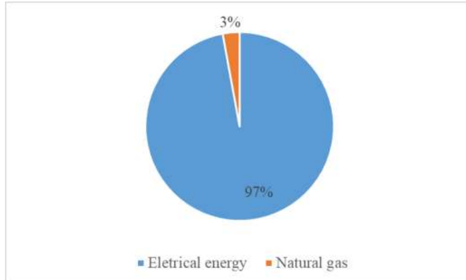
Figure 20. *Building constructive typologies of residential buildings by construction period in the Piedmont Region (climatic zone E)*



(**a**) (**b**)

Figure 21. *Mean thermal transmittance of opaque (**a**) and transparent (**b**) building envelope components in residential buildings by construction period in the Piedmont Region (climatic zone E)*

(**a**) (**b**)



(**c**)

Figure 22. *Energy carriers consumed in residential buildings in the Piedmont Region (climatic zone E): space heating (**a**), DHW (**b**), and space cooling (**c**)*

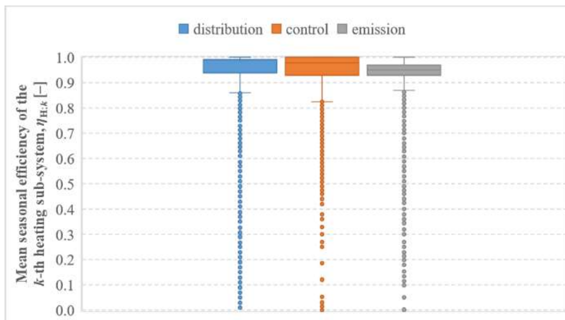


Figure 23. *Space heating distribution, control, and emission sub-systems in residential buildings in the Piedmont Region (climatic zone E)*

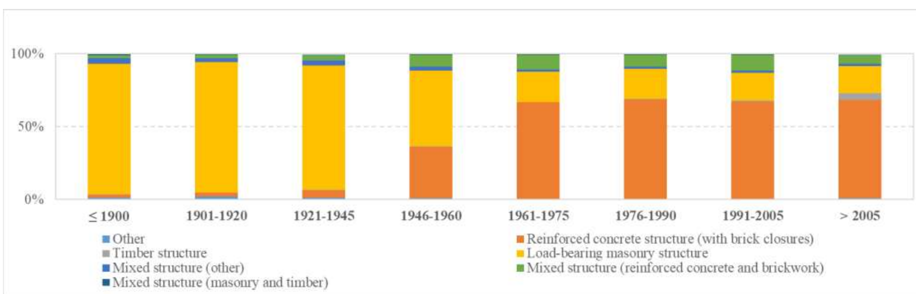
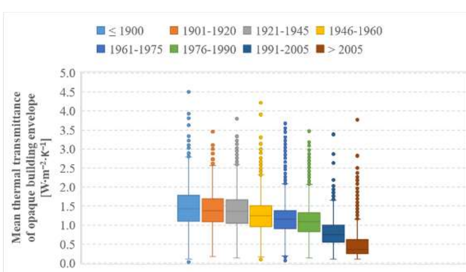
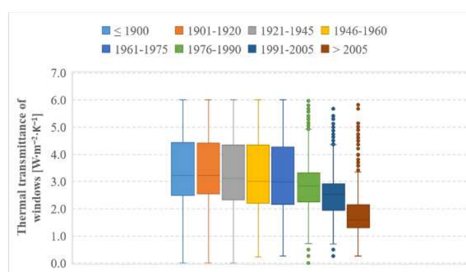


Figure 24. *Building constructive typologies of residential buildings by construction period in the Piedmont Region (climatic zone F)*

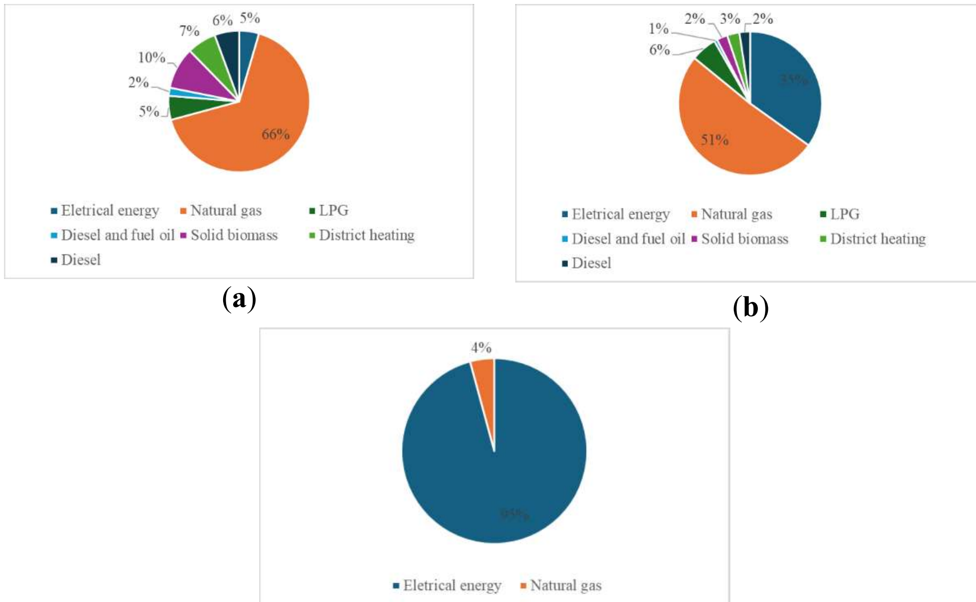


(**a**)



(**b**)

Figure 25. *Mean thermal transmittance of opaque (**a**) and transparent (**b**) building envelope components in residential buildings by construction period in the Piedmont Region (climatic zone F)*



(**a**) (**b**)

(**c**)

Figure 26. *Energy carriers consumed in residential buildings in the Piedmont Region (climatic zone F): space heating (**a**), DHW (**b**), and space cooling (**c**)*

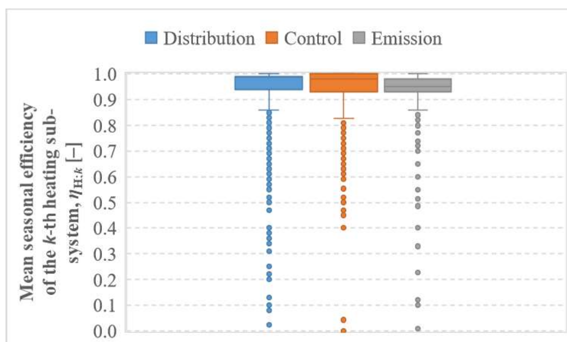


Figure 27. *Space heating distribution, control, and emission sub-systems in residential buildings in the Piedmont Region (climatic zone F)*

4.4.2 Aosta Valley Region (Italy) energy certificate database

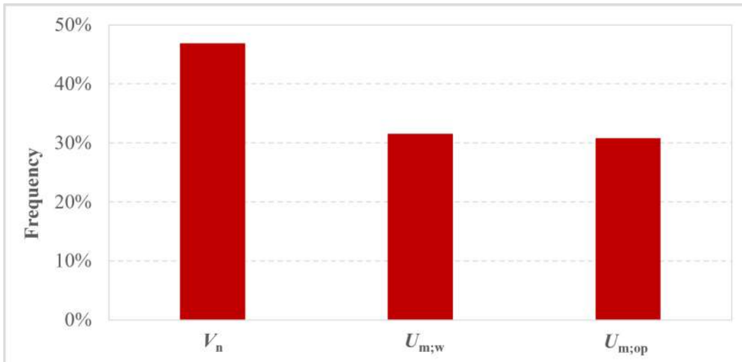
For the Aosta Valley EPC database, as noted in Section 4.4, two distinct datasets were analysed and scored. Therefore, the procedure differs slightly in the definition of critical and non-critical parameters. In this context, the results are presented separately for EPCs issued before 2017 and those issued from 2017 onward.

4.4.2.1 EPC database ante-2017

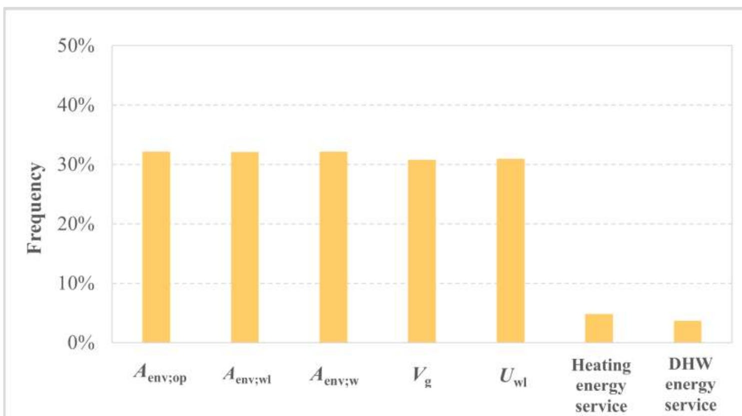
This section reports the outcomes of the EPC data quality-checking procedure applied to energy certificates issued before 2017 and extracted from the Aosta Valley regional database.

Figure 28 shows the frequency of invalid rules for both critical and non-critical parameters, while Figure 29 presents the overall EPC score distribution, highlighting the shares of accepted and rejected certificates. In this latter case, green bars correspond to EPCs that passed the quality checks, while red bars indicate those rejected due to missing or inconsistent data.

For the *reliable* EPCs, the results are then discussed in terms of glazing and frame types, as well as the thermal properties of opaque (distinguished by building element) and transparent envelope components. All findings are reported separately for climatic zones E (from Figure 30 to Figure 31) and E-F (from Figure 32 to Figure 33).



(**a**)



(**b**)

Figure 28. *Error frequencies in Aosta Valley Region for critical (**a**) and non-critical parameters (**b**)*

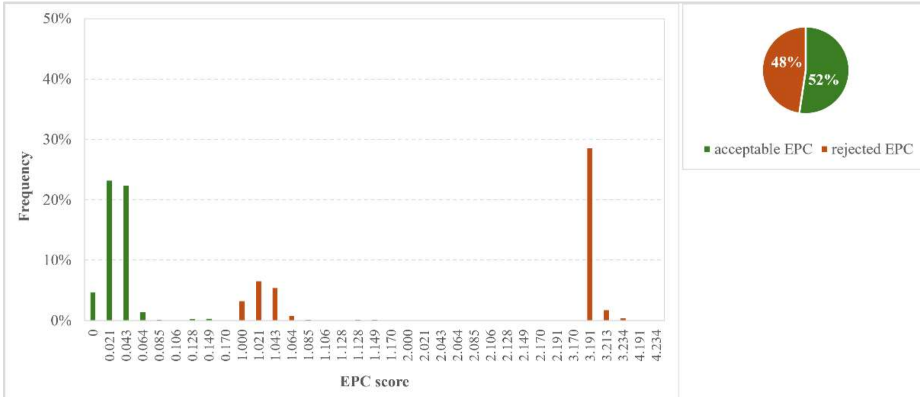
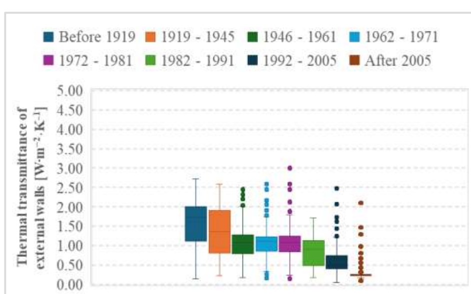


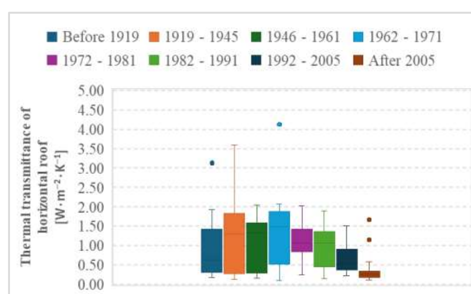
Figure 29. *Frequency distribution of EPC scores in the Aosta Valley Region: the green bar represents acceptable EPCs, while the red bars indicate non-acceptable EPCs*

Glazing type	Frame type	Before 1919	1919 - 1945	1946 - 1961	1962 - 1971	1972 - 1981	1982 - 1991	1992 - 2005	After 2005
Single glazing	Wooden frame	11%	14%	19%	26%	19%	2%	1%	0%
Single glazing	Metal frame without thermal break	1%	0%	1%	2%	2%	1%	0%	0%
Double glazing	Wooden frame	45%	42%	33%	30%	33%	60%	34%	6%
Double glazing	Metal frame without thermal break	0%	1%	1%	3%	2%	2%	14%	0%
Double glazing	Metal frame with thermal break	0%	0%	1%	4%	1%	1%	7%	0%
Double glazing	PVC frame	3%	6%	16%	14%	23%	8%	12%	2%
Double glazing with gas	Wooden frame	13%	19%	7%	3%	4%	18%	13%	5%
Double glazing with gas	Metal frame with thermal break	0%	0%	0%	0%	0%	0%	3%	1%
Double glazing with gas	PVC frame	0%	0%	5%	5%	3%	2%	3%	0%
Low-emissivity double glazing with gas	Wooden frame	18%	12%	5%	5%	4%	2%	7%	52%
Low-emissivity double glazing with gas	Metal frame with thermal break	0%	0%	0%	1%	0%	0%	2%	16%
Low-emissivity double glazing with gas	PVC frame	1%	3%	8%	5%	6%	3%	4%	12%
Triple glazing with gas	Wooden frame	2%	0%	0%	0%	0%	0%	0%	0%
Low-emissivity triple glazing with gas	Wooden frame	5%	0%	2%	0%	1%	0%	1%	1%
Low-emissivity triple glazing with gas	Metal frame with thermal break	0%	0%	0%	0%	0%	0%	0%	3%
Low-emissivity triple glazing with gas	PVC frame	0%	0%	2%	1%	0%	0%	1%	1%

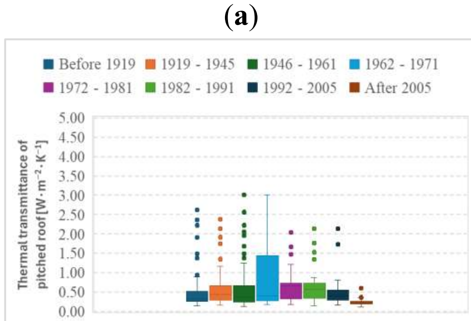
Figure 30. *Glazing–frame type combinations in residential buildings by construction period in the Aosta Valley Region (climatic zone E)*



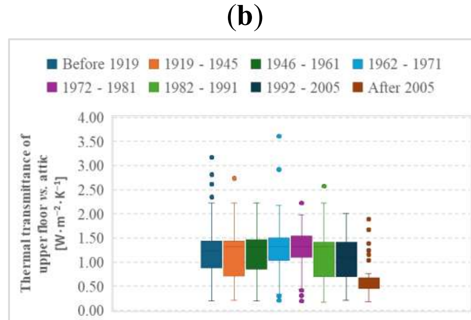
(a)



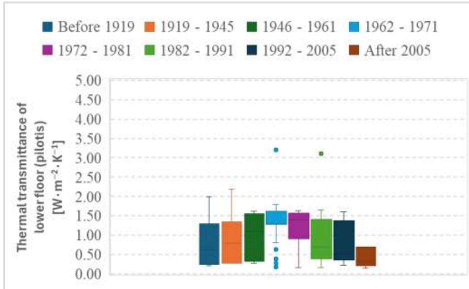
(b)



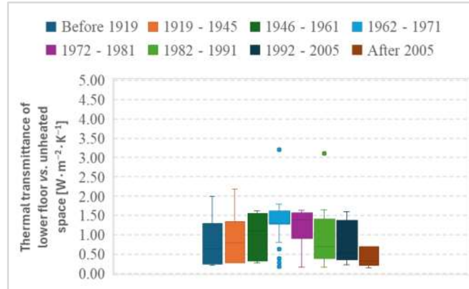
(c)



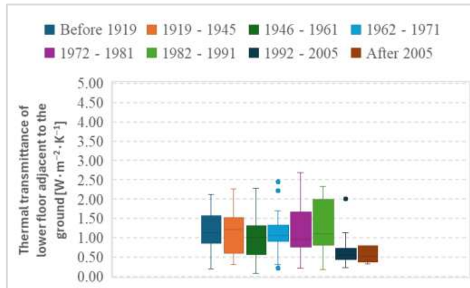
(d)



(**e**)



(**f**)

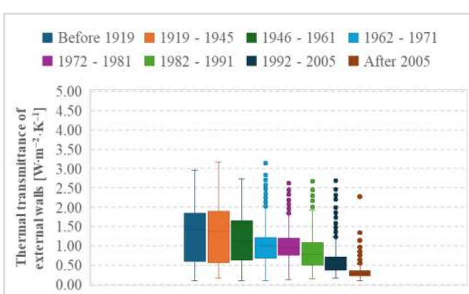


(**g**)

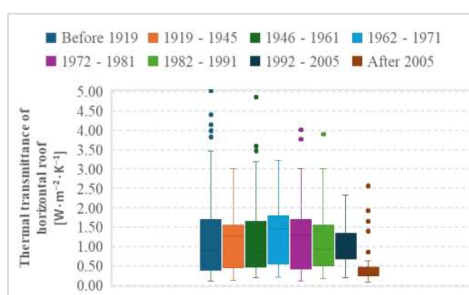
Figure 31. *Thermal transmittance of opaque building envelope components in residential buildings by construction period in the Aosta Valley Region (climatic zone E): external walls (**a**), horizontal roofs (**b**), pitched roofs (**c**), upper floors vs. attics (**d**), lower floors (pilotis) (**e**), lower floor above unheated space (**f**), and lower floor in contact with the ground (**g**)*

Glazing type	Frame type	Before 1919	1919 - 1945	1946 - 1961	1962 - 1971	1972 - 1981	1982 - 1991	1992 - 2005	After 2005
Single glazing	Wooden frame	16%	13%	18%	22%	17%	4%	0%	0%
Single glazing	Metal frame without thermal break	1%	0%	0%	0%	1%	0%	0%	0%
Double glazing	Wooden frame	47%	45%	40%	49%	62%	73%	63%	16%
Double glazing	Metal frame without thermal break	0%	0%	1%	1%	1%	1%	1%	0%
Double glazing	Metal frame with thermal break	0%	1%	2%	1%	1%	1%	0%	0%
Double glazing	PVC frame	2%	2%	5%	4%	3%	6%	2%	1%
Double glazing with gas	Wooden frame	11%	11%	11%	6%	4%	6%	18%	17%
Double glazing with gas	Metal frame with thermal break	0%	0%	0%	0%	0%	0%	2%	0%
Double glazing with gas	PVC frame	1%	1%	1%	2%	1%	1%	2%	2%
Low-emissivity double glazing with gas	Wooden frame	18%	21%	14%	10%	4%	5%	9%	47%
Low-emissivity double glazing with gas	Metal frame with thermal break	0%	1%	0%	1%	0%	0%	0%	3%
Low-emissivity double glazing with gas	PVC frame	2%	2%	4%	4%	3%	1%	1%	2%
Triple glazing with gas	Wooden frame	0%	0%	2%	0%	0%	0%	0%	0%
Low-emissivity triple glazing with gas	Wooden frame	2%	2%	2%	0%	1%	1%	0%	9%
Low-emissivity triple glazing with gas	PVC frame	0%	0%	1%	0%	1%	1%	0%	1%

Figure 32. *Glazing–frame type combinations in residential buildings by construction period in the Aosta Valley Region (climatic zone E-F)*



(**a**)



(**b**)

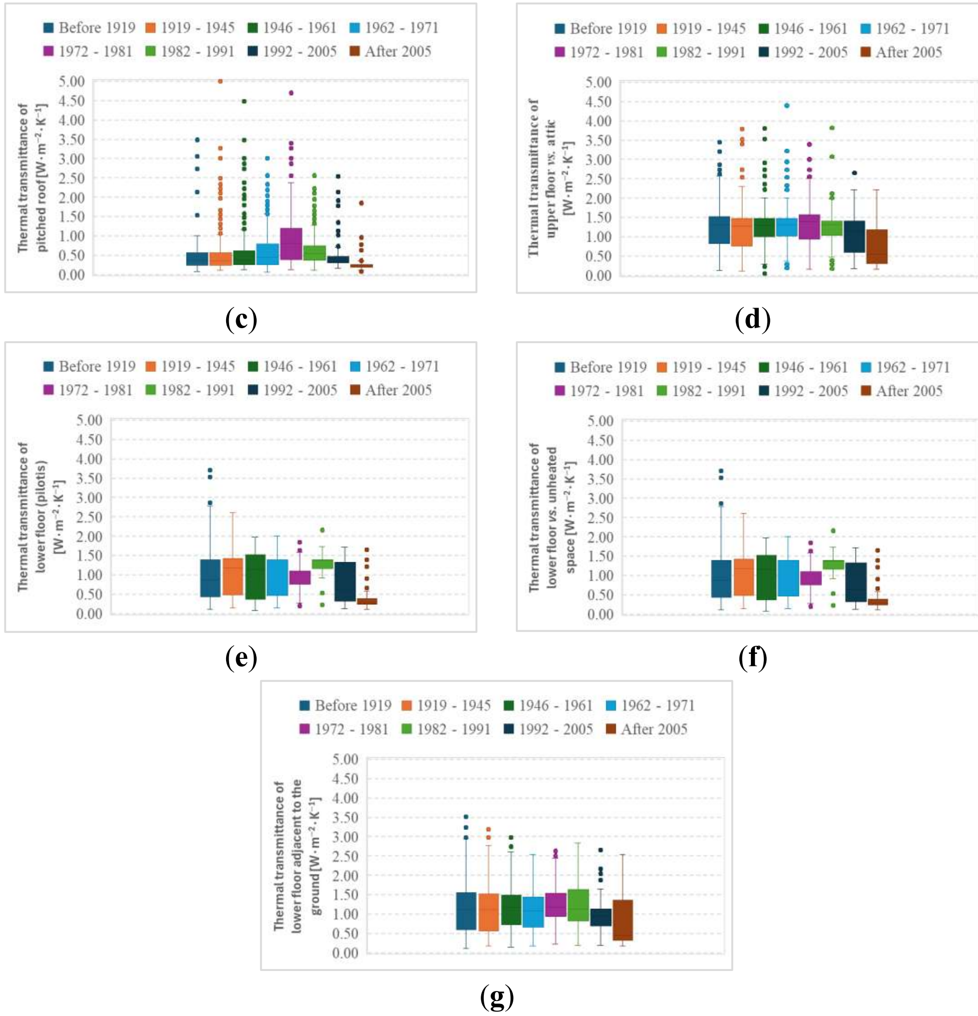


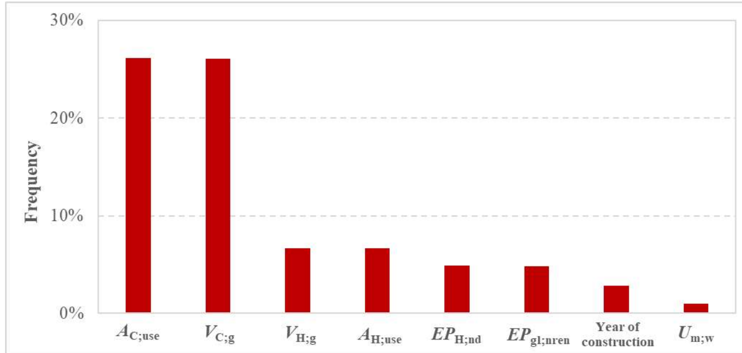
Figure 33. *Thermal transmittance of opaque building envelope components in residential buildings by construction period in the Aosta Valley Region (climatic zone E-F): external walls (**a**), horizontal roofs (**b**), pitched roofs (**c**), upper floors vs. attics (**d**), lower floors (pilotis) (**e**), lower floor above unheated space (**f**), and lower floor in contact with the ground (**g**)*

4.4.2.2 EPC database post-2017

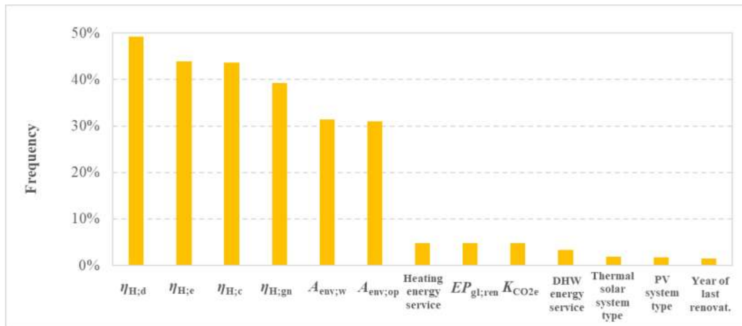
This section reports the outcomes of the EPC data quality-checking procedure applied to energy certificates issued after 2017 and extracted from the Aosta Valley regional database.

Figure 34 shows the frequency of invalid rules for both critical and non-critical parameters, while Figure 35 presents the overall EPC score distribution, highlighting the shares of accepted and rejected certificates. In this latter case, green bars correspond to EPCs that passed the quality checks, while red bars indicate those rejected due to missing or inconsistent data.

For the *reliable* EPCs, the results are then discussed in terms of building constructive typologies, thermal properties of transparent components, and TBS performance. All findings are reported separately for climatic zones E (from Figure 36 to Figure 39) and E-F (from Figure 40 to Figure 43).



(**a**)



(**b**)

Figure 34. *Error frequencies in Aosta Valley Region for critical (**a**) and non-critical parameters (**b**)*

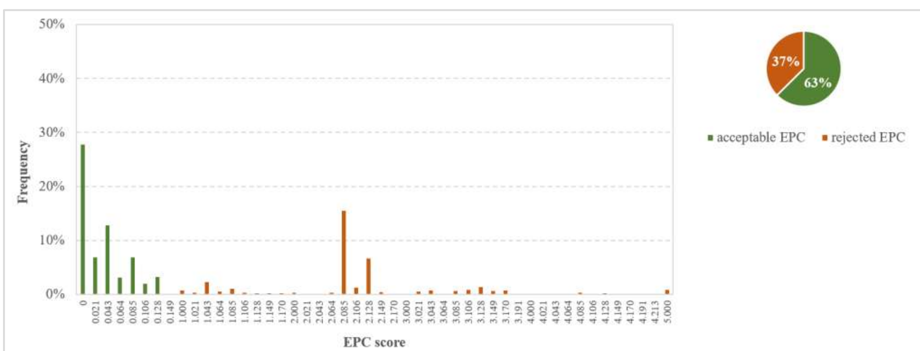


Figure 35. *Frequency distribution of EPC scores in the Aosta Valley Region: the green bar represents acceptable EPCs, while the red bars indicate non-acceptable EPCs*

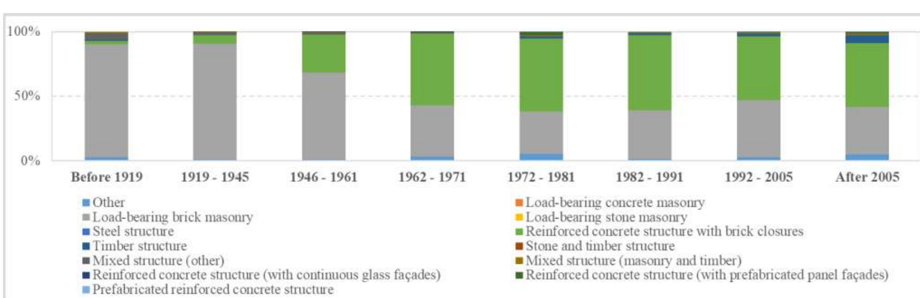


Figure 36. *Building constructive typologies of residential buildings by construction period in the Aosta Valley Region (climatic zone E)*

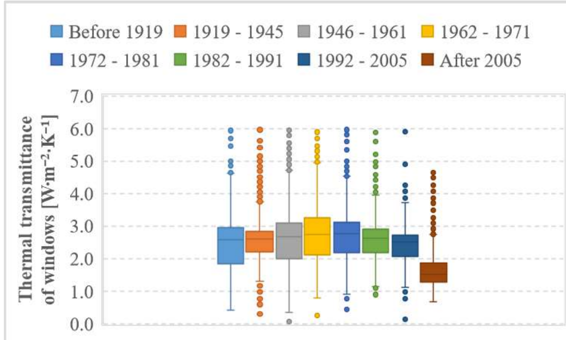


Figure 37. *Thermal transmittance of windows in residential buildings by construction period in the Aosta Valley Region (climatic zone E)*

Table 14. *Generation sub-systems for heating, cooling, and DHW in residential buildings in the Aosta Valley Region (climatic zone E). The three most representative space heating generators (condensing boiler, standard boiler, and district heating) are shown in bold*

Generation sub-system	Space heating system typology	Space cooling system typology	DHW system typology
Other	1 %	0 %	3 %
Electric boiler	0 %	0 %	33 %
Condensing boiler	**30 %**	11 %	21 %
Standard boiler	**40 %**	0 %	30 %
Air-to-air absorption heat pump	0 %	2 %	0 %
Water-to-water electric heat pump	1 %	20 %	1 %
Water-to-air electric heat pump	0 %	1 %	0 %
Air-to-water electric heat pump	2 %	24 %	3 %
Air-to-air electric heat pump	0 %	41 %	0 %
Stove or fireplace	2 %	0 %	0 %
District heating	**23 %**	0 %	9 %

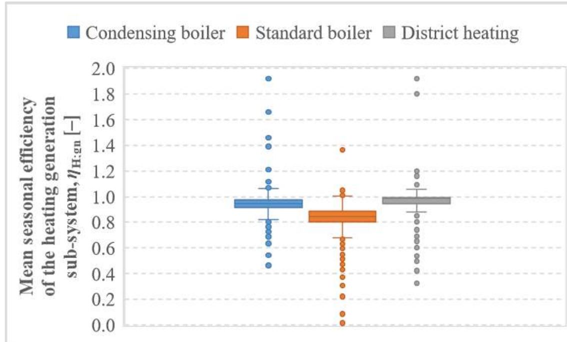


Figure 38. *Three most representative space heating generation sub-systems in residential buildings in the Aosta Valley Region (climatic zone E)*

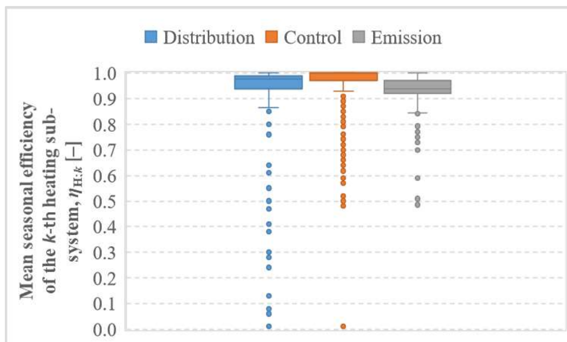


Figure 39. *Space heating distribution, control, and emission sub-systems in residential buildings in the Aosta Valley Region (climatic zone E)*

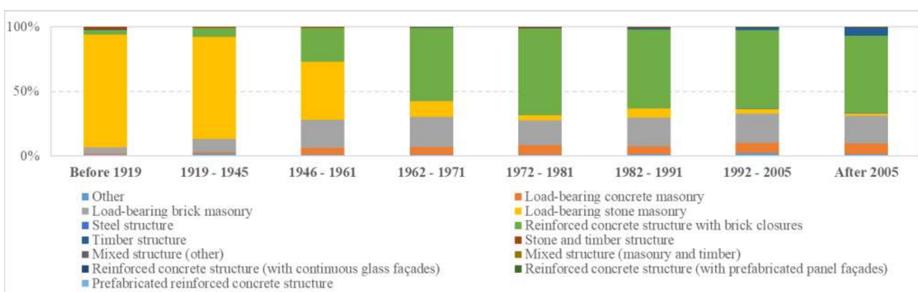


Figure 40. *Building constructive typologies of residential buildings by construction period in the Aosta Valley Region (climatic zone E-F)*

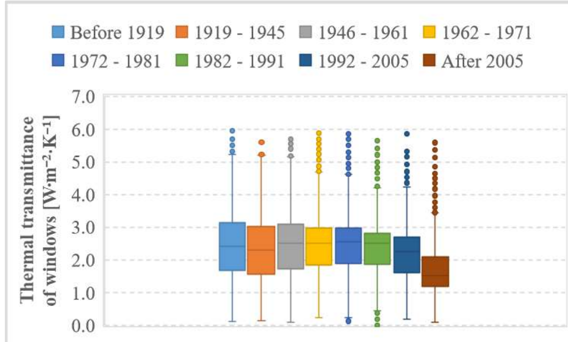


Figure 41. *Thermal transmittance of windows in residential buildings by construction period in the Aosta Valley Region (climatic zone E-F)*

Table 15. *Generation sub-systems for heating, cooling, and DHW in residential buildings in the Aosta Valley Region (climatic zone E-F). The three most representative space heating generators (condensing boiler, standard boiler, and stove or fireplace) are shown in bold*

Generation sub-system	**Space heating system typology**	**Space cooling system typology**	**DHW system typology**
Other	4 %	8 %	6 %
Electric boiler	0 %	0 %	25 %
Condensing boiler	**23 %**	0 %	20 %
Standard boiler	**52 %**	0 %	38 %
Water-to-water electric heat pump	0 %	4 %	0 %
Air-to-water electric heat pump	4 %	33 %	5 %
Air-to-air electric heat pump	0 %	55 %	0 %
Solar thermal system	0 %	0 %	1 %
Electric heating	1 %	0 %	0 %
Stove or fireplace	**10 %**	0 %	2 %
District heating	5 %	0 %	3 %

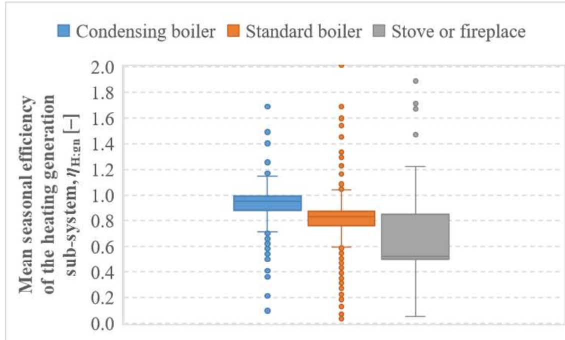


Figure 42. *Three most representative space heating generation sub-systems in residential buildings in the Aosta Valley Region (climatic zone E-F)*

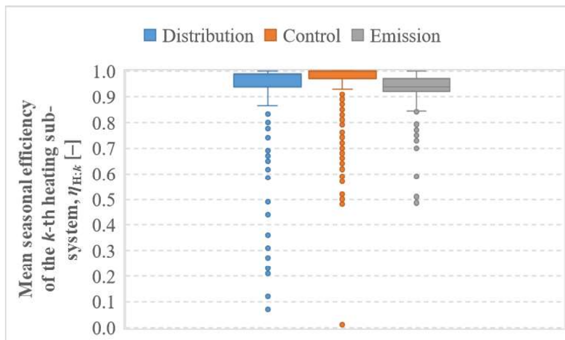


Figure 43. *Space heating distribution, control, and emission sub-systems in residential buildings in the Aosta Valley Region (climatic zone E-F)*

4.4.2.3 Trends in EPC unreliability for residential buildings

This section presents a further analysis of the residential building stock subset derived from the Aosta Valley EPC database post-2017. Specifically, it focuses on certificates that did not pass the data quality-checking procedure, with the aim of identifying systematic critical errors, particularly those associated with construction periods.

Figure 44 shows, for each construction period, the share of reliable and unreliable EPCs. This analysis is complemented by Figure 45, which reports the frequency of invalid critical parameters for each construction period. Interestingly, both the oldest and the most recent construction periods exhibit the highest share of invalid EPCs (Figure 44), reaching approximately 50 %. For the oldest buildings, this result is mainly attributable to invalid construction year values (see Figure 46) together with inconsistencies in cooled volumes and floor areas. For the most recent buildings, the high proportion of invalid certificates is primarily associated with discrepancies between the declared presence of the space-cooling energy service and the corresponding conditioned area or volume.

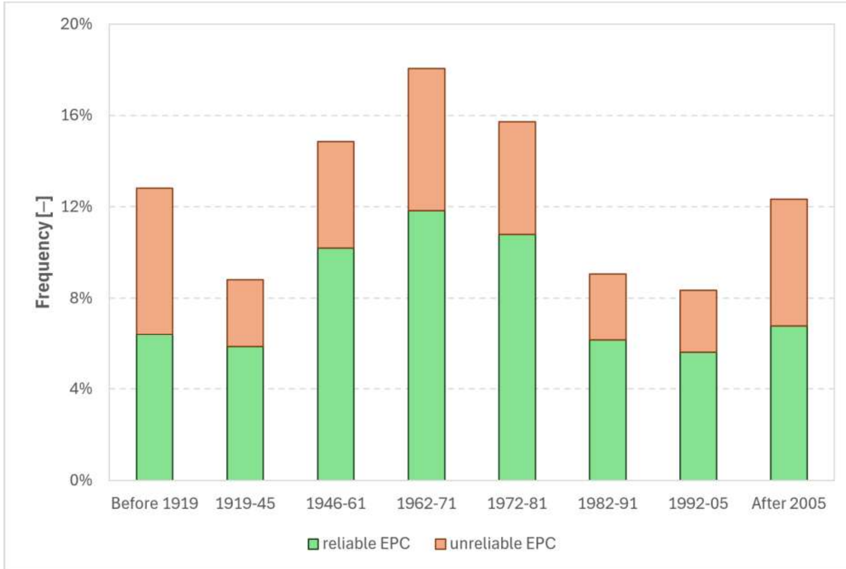


Figure 44. *Share of reliable and unreliable EPCs by construction period*

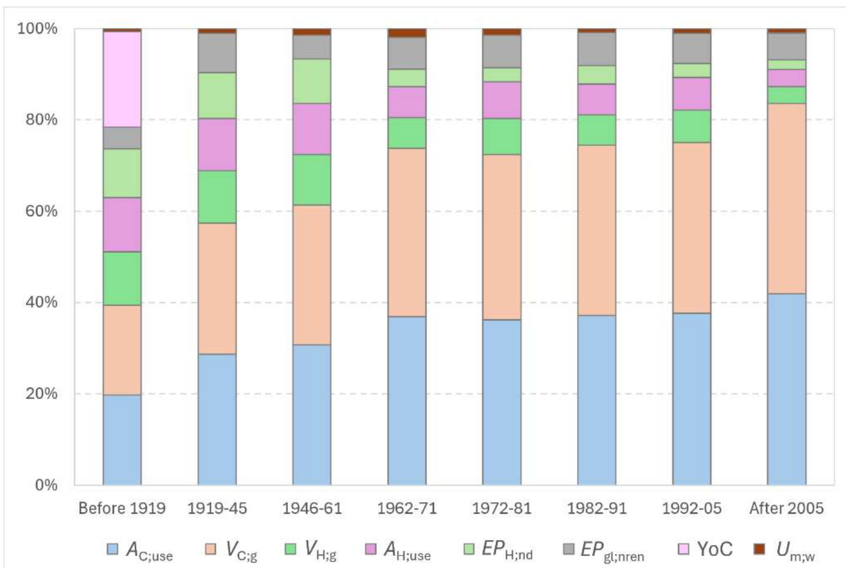


Figure 45. *Share of EPCs with invalid critical parameters by age*

Figure 46 reports the invalid construction period ranges declared in the EPCs stored in the BEAUCLIMAT portal of the Aosta Valley region for residential building stock. More than 50 documents reported a construction period between 1000 and 1100. This result is likely attributable to the XML schema, which imposes a lower bound of 1000 for this parameter (see Figure 12). In addition, a significant number of buildings were classified within the 1600-1700 period, while more than 140 buildings were assigned to the 1700-1800 period.

Although the validity of the most recent historical construction period cannot be excluded, a threshold year of 1800 was adopted in the EPC data quality-checking procedure. This choice was made to ensure consistency within the analysed sample and to limit the influence of potential outliers associated with historical buildings. Consequently, only buildings constructed after 1800 were retained for further analysis.

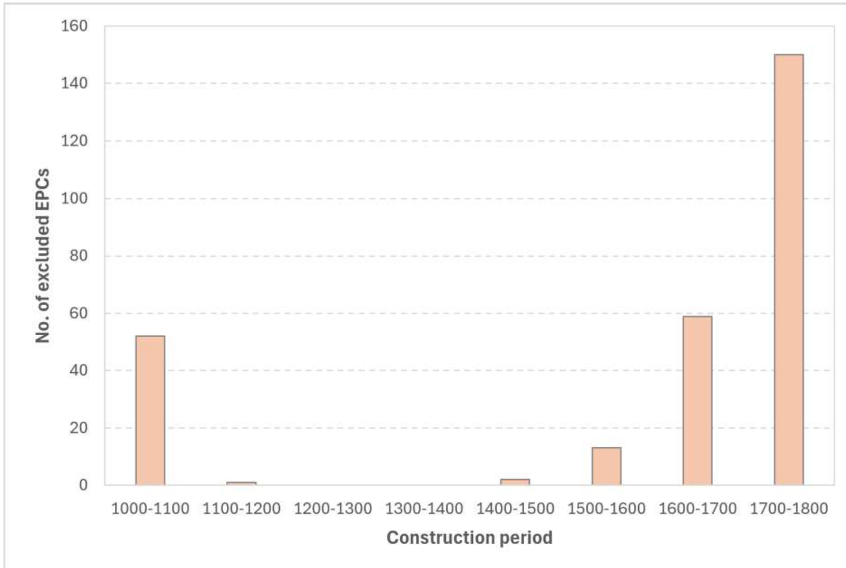


Figure 46. *Excluded construction period ranges*

According to Italian health and hygiene regulations, the minimum net internal height of habitable rooms is generally 2.70 m, reduced to 2.55 m in municipalities located above 1000 m a.s.l., with a further reduction of up to 0.10 m permitted in existing buildings undergoing energy retrofit interventions. Conservatively, a 2.30 m acceptable threshold limit was assumed. Figure 47 depicts the box-plot distribution of EPCs discarded due to invalid building height values.

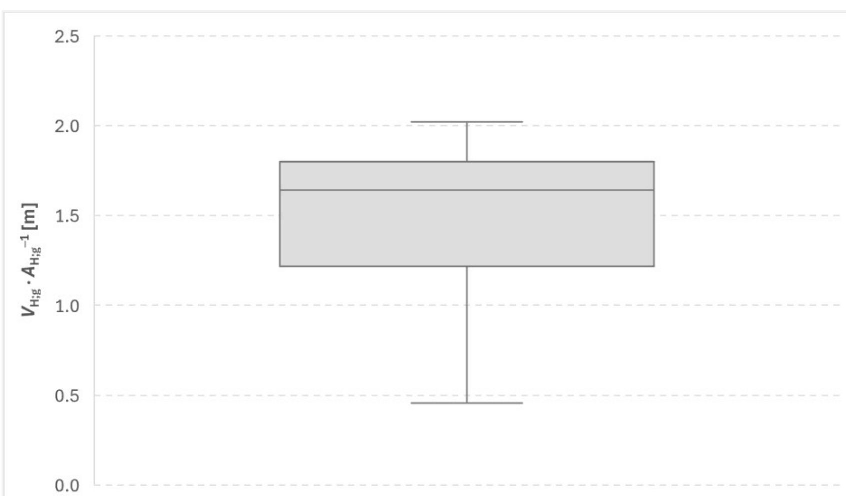


Figure 47. *Box-plot distribution of null heights*

It should be noted that, owing to data availability constraints, the thermally conditioned useful floor area was converted into gross floor area, as described in

Section 4.4. However, most EPCs were discarded because of inconsistencies in the reported sample, namely the absence of the space-heating energy service despite non-null values of both conditioned volume and floor area.

Figure 48 reports the ranges of *U*-values for transparent elements excluded during the EPC data-quality checking process. A total of 100 certificates were discarded due to null thermal transmittance values. Furthermore, Figure 48 shows threshold *U*-values calculated considering only the internal and external surface thermal resistances for horizontal and upwards heat flows. These surface resistances were derived from EN ISO 6946 (CEN, 2007a). For the subsequent statistical analysis, a threshold value of 6.0 W/(m^2·K) was adopted to standardise the heat-flow direction. *U*-values between 5.88 and 7.14 W/(m^2·K) may correspond to attic spaces equipped exclusively with horizontal single-glazed windows, representing a highly specific condition that was treated as an outlier. Conversely, *U*-values exceeding 7.14 W/(m^2·K) are physically implausible according to the surface thermal resistance limits defined in EN ISO 6946 (CEN, 2007a).

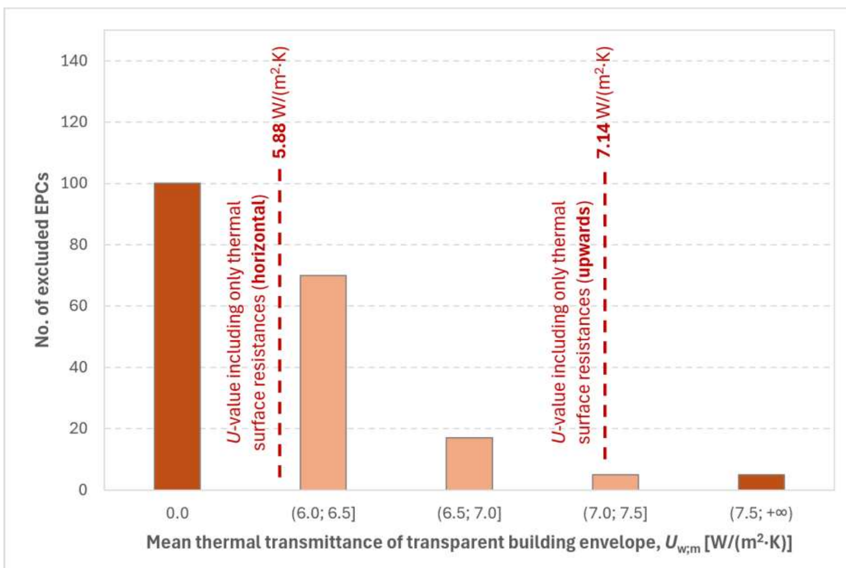


Figure 48. *Excluded U-value ranges for windows*

4.5 Remarks

The EPC data quality-checking procedure was applied to two distinct Italian databases. The structure and completeness of these datasets depend on which information is extracted from the *reduced* or *extended* formats of the uploaded certificates. Moreover, the non-critical parameter score ($s_{\mathrm{n-C}}$) is influenced by the number of both critical and non-critical parameters included in the analysis. Nevertheless, the proposed procedure harmonises the set of controls, ensuring its adaptability across different Italian regional contexts.

The proportion of discarded EPCs is closely linked to the stringency of the critical parameter checks, which directly affect the validity of the entire document. Across the analysed databases, between 29 % and 48 % of EPCs were rejected due to at least one unfulfilled rule associated with a critical parameter. Table 16, accompanied by the results reported in Section 4.4.2.3, summarises the most common reasons for the exclusion of EPC parameters according to their level of criticality.

In the Piedmont Region (Figure 19), approximately 71 % of EPCs were classified as acceptable, showing a predominance of valid certificates with relatively low EPC scores. In contrast, the Aosta Valley Region (Figure 29 and Figure 35) presents a more balanced distribution, with around 52-63 % of acceptable EPCs depending on the dataset considered. The proportion of rejected EPCs is significantly higher than in Piedmont, reflecting a greater incidence of data inconsistencies or missing information.

Overall, the results demonstrate that while both regional EPC databases contain a substantial share of reliable records, systematic data validation and control procedures remain essential to ensure the robustness and comparability of EPC-derived analyses at larger spatial scales.

Table 16. *Most common reasons for the exclusion of EPC parameters*

Level of criticality	Parameter(s)	Most common reason for parameter exclusion
Critical	YearofConstruction	Not compiled
	$V_{C;g}$ $A_{C;use}$	Inconsistency with the presence of the space cooling energy service
	$V_{H;g}$ $A_{H;use}$	Inconsistency with the calculation of the mean building height (2.30 m)
	A_{env}	Null or inconsistencies in the summation of opaque and transparent envelope areas
	$U_{m;op}$ $U_{m;w}$	Not compiled or out of range
	$EP_{H;nd}$	= 0 or null
	$EP_{gl;nren}$	Not compiled
Non-critical	$A_{env;op}$ $A_{env;w}$	Inconsistency with the total thermal envelope area
	$\eta_{H;k}$	Not compiled

4.5.1 Next-gen. EPCs: Future perspectives

Currently, the EPC is perceived by end-users mainly as an administrative burden required for newly constructed, sold, or leased buildings or building units. However, the next-generation EPC, as envisioned in the new EPBD recast (EPBD, 2024a), aims to evolve into a holistic instrument that goes beyond the sole energy domain. Buildings should be assessed from multiple perspectives—economic, environmental, and social—in line with the ISO 37100 technical standard (ISO, 2016), which highlights the interdependence among these dimensions in its definition of sustainability. Accordingly, future EPC datasets should be enriched with new KPIs, complementing the standard energy performance with information on the actual use of buildings.

A further critical issue of the current energy certification concerns the generally low quality of EPC data. In most cases, a systematic validation procedure should be implemented—either within building energy performance simulation tools or directly in EPC databases—to ensure data reliability. Existing deficiencies in EPC data quality hinder the effective use of certificates for both single-building assessments and large-scale analyses. In particular at the city scale, EPCs represent a key source of information for evaluating and ranking the overall energy performance of the building stock.

According to the literature, improving the accuracy and reliability of EPC information requires coordinated actions across several key areas:

Sustainability indicators:

- The EPC is primarily an energy-focused document based on the assessment of the EP indicator. However, future improvements should better address social, environmental, and economic aspects. Tsatsakis et al. (2017) proposed transforming the EPC from a static document into a dynamic tool by integrating real-consumption data. Caceres (2018) highlighted the need to include cost-effective and technically feasible energy efficiency measures within the EPC to guide end-users in renovation decisions. Although Parkinson et al. (2013) explored how a building's energy rating relates to occupant satisfaction, EPCs still lack sufficient information about indoor environmental quality. Zuhaib et al. (2022) recommended integrating new indicators into next-generation EPCs, such as the Smart Readiness Indicator—which reflects the capacity of building systems to adapt to occupants' needs—alongside indexes for indoor and outdoor air quality, real consumption information, and indicators informing users about the benefits of connecting to district heating or cooling networks. Moreover, Anagnostopoulos et al. (2015) suggested including fiscal incentive programs within EPCs to increase occupant awareness and encourage renovation activities.

Data interoperability:

- Interoperability is a cornerstone of the FAIR data principles—findability, accessibility, interoperability, and reusability (Wilkinson et al., 2016). Although data may exist in multiple formats, it should be managed as a unified entity linked across different databases (e.g., EPC, cadastral, statistical, or geographical databases). To enhance data quality, ownership and management of information should be centralised in a single system capable of communicating effectively with other repositories.

Database control activities:

- Each EPC database should incorporate analytical techniques to detect anomalies and ensure data consistency (Hardy & Glew, 2019; Prieler et al., 2017). Pasichnyi, Wallin, Levihn, et al. (2019) proposed a six-level framework for validating EPC data quality. A best-practice example is provided by the EPC database of the Emilia-Romagna region, which applies a two-stage control process (Marinosci & Morini, 2014). The first level (*Type A*) involves online preventive checks that guide assessors in submitting EPCs with plausible values and without technical inconsistencies, thereby improving data reliability and reducing typing errors.

Cost adaptation:

- EPC data quality is also influenced by market regulations. Several countries—such as Croatia, Denmark, Hungary, and Slovenia—have adopted standardised pricing frameworks for issuing energy certificates. According to Anagnostopoulos et al. (2015), the cost range for EPCs in single-family houses varies across twelve European countries.

Chapter 5

Representative buildings and urban blocks

5.1 Overview

Part of this work was previously published in an international peer-reviewed journal (Piro, Ballarini, et al., 2026) and conference proceedings (Ballarini et al., 2026; Piro, Ballarini, et al., 2025b, 2025a). In particular, Ballarini et al. (2026) report the first results of the CRiStAll project, within which the Ph.D. activity focused on the development and simulation of the UBEM model. In the conference proceedings by Piro, Ballarini, et al. (2025b, 2025a), the Ph.D. research focused on the development of the methodology and its application to different archetype-based UBEMs. For the journal publication (Piro, Ballarini, et al., 2026), the specific contributions of the author are detailed in the corresponding CRediT authorship statements.

This chapter addresses a key aspect of large-scale energy analysis, namely the reduction of uncertainty in the collection of crucial information through representative buildings. As discussed in Section 2.5, the *archetyping* process aimed at generating building prototypes and archetypes consists of *segmentation* and *characterisation* phases, whose application depends on the final planning or modelling purpose. From this perspective, prototype- and archetype-based approaches are adopted for the development of NBRPs and UBEMs, respectively. Furthermore, this chapter describes the evolution from the concept of representative building to that of the representative urban block. In the latter case, representative city blocks are thermally characterised using building archetypes. The overall methodology is depicted in Figure 49.

Moreover, the complete digital library of building prototype and archetype schemas is available in Ballarini et al. (2023) and on the URBEM web-site (URBEM, 2025), respectively. In addition, the methodology for shifting from representative buildings to representative urban blocks is part of the PRIN2022-PNRR CRiStAll research project.

This chapter is organised as follows: Section 5.2 presents the role of representative buildings in mitigating the uncertainty in modelling urban scale energy performance; Section 5.3 reports the paradigm shift from representative buildings to representative districts; and finally, Section 5.4 provides concluding remarks.

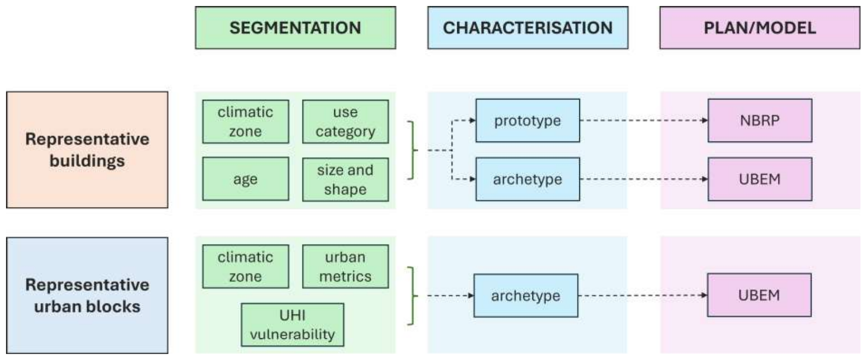


Figure 49. *Representative buildings and urban blocks development*

5.2 Representative buildings

Representative buildings help reduce the uncertainty associated with estimating the overall or partial balance of the building stock. These virtual objects are classified according to different categories or criteria, reflecting the average energy performance characteristic of specific building stock segments. They may include geometrical information and must contain thermophysical properties of the building envelope, technical building system characteristics, and user-related parameters.

Both building prototypes and building archetypes fall within the broader category of representative buildings, although their use depends on the purpose of the analysis. Building prototypes are geometry-dependant, and the median values of energy performance indicators are typically employed to statistically define baseline/*status quo* of building stock energy footprint, as well as to develop renovation scenarios. NBRP represents a prominent example of the use of building prototypes, as they rely on representative typologies to estimate renovation potential and energy-saving targets at large scales.

Conversely, building archetypes define a set of thermo-physical parameters assigned to reference urban geometries for energy simulations, forming the basis of archetype-based UBEMs. In this case, the result of the analysis will be the output of the UBEM tool.

5.2.1 Methodology

Both building prototypes and archetypes are defined according to the same criteria for segmenting the building stock based on energy performance. These criteria include the climatic zone, construction period, building use category, and, for residential buildings, the size and shape. These virtual objects were developed for single-family houses, apartment blocks, offices, and educational buildings.

Prototypes refer to representative buildings located in Piedmont region within climatic zone E, whereas archetypes encompass both climatic zones E and F for Piedmont and Aosta Valley regions. In both cases, the energy certificate database was used, following a quality and reliability assessment of the available certificates (see Sections 4.3 and 4.4).

5.2.1.1 Building prototypes

The NBRPs adopt a *top-down* approach to assess local or regional energy intensity based on representative energy performance indicators and their statistical distribution within the building stock. As mentioned earlier, this macroscopic approach requires the use of representative energy metrics for different building types, combined with their geometry share. Once the baseline energy footprint of the building stock is assessed and a fixed or variable renovation rate is defined, the analysis aims to outline a renovation pathway through the application of different scenarios for reducing overall energy demand.

The building protypes were developed within the framework of the H2020-TIMEPAC project for several EU countries, including Austria, Croatia, Cyprus, Italy (Piedmont region), Spain (Catalonia region), and Slovenia. These virtual objects were derived from reliable energy certificates, selected following the application of the EPC data quality-checking procedure. A common methodology, described in Section 4.3.2, was adopted across project partners and includes data selection, segmentation, and quality assessment.

Each reference building is represented by the statistical value of its main characterising parameter, grouping into the following categories: building geometry, thermal characteristics of the building envelope, technical building system properties, and energy indicators.

5.2.1.2 Building archetypes

The methodology to create an Italian digital library of building archetypes is aligned with the approach of the *Urban Reference Buildings for Energy Modelling* (URBEM, 2025) project. URBEM—an Italian national research initiative funded under the PRIN-2020 Programme—aims to develop a comprehensive library of representative Italian buildings for use in UBEM tools (CitySim, UMI, EUReCA, CEA, etc.).

The primary objective of URBEM is to develop a comprehensive database of archetype buildings that faithfully represents the national building stock. This database aims to reduce the uncertainties inherent in UBEM simulations, thereby serving as a reliable reference and analytical tool for public authorities and building stock managers in the effective management and preservation of their assets. Moreover, it will support the formulation of policies and incentive schemes that foster sustainable construction and renovation practices. Ultimately, URBEM seeks to enhance building energy efficiency and contribute to the long-term sustainability of the built environment.

Figure 50 shows a map of Italy divided into five coloured regions, each representing a different area of competence within the project. This territorial grouping is used to define the building archetypes.

Figure 50. *Territorial area of competence URBEM project*

The methodology (Figure 51) for generating building archetypes consist of three main phases:

i. UBEM data classification, which establishes a standardised framework for identifying the data required to develop large-scale energy models,
ii. database cataloguing, involving the collection and classification of recognised Italian databases at the local, regional, and national levels, and
iii. definition of a probabilistic building archetype dataset.

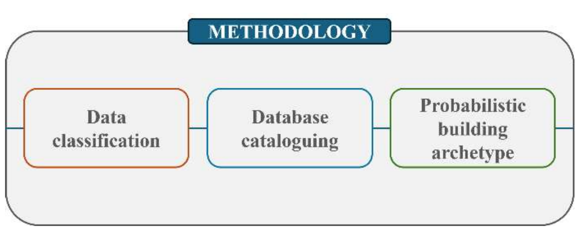


Figure 51. *URBEM methodology*

UBEM data classification

The classification of UBEM data aims to comprehensively identify all input parameters required to evaluate the energy performance of the building stock, independent of the simulation tool employed. This step is essential for pinpointing the relevant local, regional, and national databases necessary for the subsequent stages of the methodology. UBEM data can be broadly categorised into five groups: geometrical information (e.g., *WWR*); characteristics of transparent and of opaque building envelope components, occupancy data, Ferrando et al. (2020) and Abbasabadi & Mehdi Ashayeri (2019) presented an extensive review of UBEM tools developed to evaluate building stock performance. Ferrando et al. (2020) further classified physics-based tools into the following categories:

- reduced-order resistor-capacitor (R-C) models, and
- detailed dynamic thermal models.

According to the research conducted by Kamel (2022), the three most widely used UBEM tools are UMI, CityBES, and CitySim. The first two tools, UMI and CityBES, employ EnergyPlus as their simulation engine, whereas CitySim uses the CitySim Solver—a dedicated thermal model that discretises the building envelope components into an R-C network with a finite number of nodes.

Therefore, in the UBEM data classification the input needed for the most common reduced-order resistor-capacitor (CitySim) and detailed dynamic thermal models (UMI) were presented. The collected parameters were organised in different tables (from Table B.1 to Table B.11, in Appendix B), providing their nomenclature, symbols, and units of measure provided by the European Committee for Standardisation (CEN) and the International Organisation for Standardisation (ISO) technical standards.

Most of the parameters listed in the following tables (from Table B.1 to Table B.11, in Appendix B) serve as a starting point for understanding the input data requirements of UBEM tools. At the urban scale, some of these parameters are less likely to be available in tradition database, such as the visual properties of glazing. Nevertheless, identifying and systematising these parameters remains essential to ensure the usability of building archetype data in multiple UBEM applications.

Database cataloguing

Once the UBEM data classification is defined, it is necessary to review, list, and organise the existing Italian information systems to identify the relevant data for generating the BAs. Within the URBEM project, the databases were grouped into the following categories: (i) source type, indicating the origin of the information; (ii) accessibility, referring to the level of constraints associated with the information systems; (iii) database digitalisation, describing whether the data are available in digital (online or offline) or paper-based formats; and (iv) data granularity, measuring the level of detail of the information. Figure 52 illustrates the criteria adopted to classify local, regional, and national databases according to these categories. Furthermore, to address the uncertainty gap in UBEM, additional data from technical standards, statistical analyses, and multi-scale databases were identified and examined.

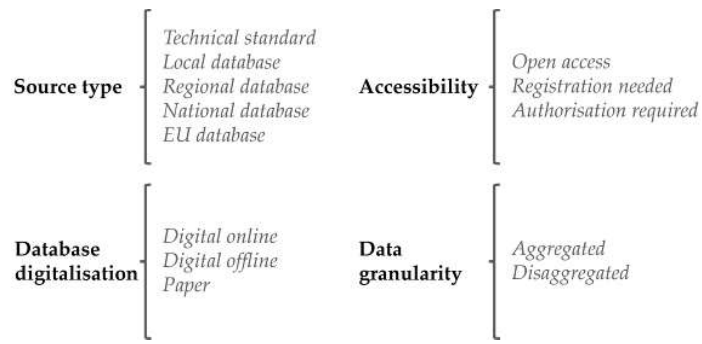


Figure 52. *Database cataloguing criteria adopted in the URBEM project (Piro, Ballarini, et al., 2025b)*

Probabilistic archetype schema

The criteria adopted to segment the building stock's energy performance include the climatic zone, building use category, and construction period. For residential buildings, an additional distinction was introduced based on building size and shape, differentiating between single-family houses and apartment blocks.

The proposed BA schema does not rely on data clustering techniques in the *unsupervised* machine-learning sense. Instead, the segmentation criteria were defined a priori based on literature review and on the structure commonly adopted in building stock characterisation methodologies. Consequently, the building stock was subdivided according to predefined categories, avoiding the generation of physically or statistically inconsistent combinations of building attributes. However, for some non-residential building categories, the number of available records within a single construction period was not always sufficient to ensure statistical representativeness. In these cases, adjacent construction periods were aggregated in order to guarantee a minimum threshold of five independent information sources for each archetype definition.

The BA schema is designed to represent datasets in a probabilistic manner, providing a range of values rather than a single deterministic number. This approach allows the statistical distribution of parameters to guide UBEM developers in selecting appropriate input values. Once the probabilistic archetypes are defined, the corresponding representative dataset can be associated with real building geometries according to the segmentation criteria previously described.

The variables included in the BA schema are classified into numerical and categorical parameters. For numerical variables, the mean, standard deviation, and the first, second, and third quartiles are provided. For the categorical variables—such as the building constructive typology or the glazing type—the percentage distribution of the different recurring items is shown.

5.2.2 Development of representative buildings

5.2.2.1 Building prototypes

The building prototypes were developed for the Piedmont region, within climatic zone E. Eight construction periods defined by the TABULA project were considered, along with four building use categories: single-family houses (SFHs), apartment blocks (BU(AB)), offices (OFF), and educational buildings (EDUC). This resulted in a total of 32 prototype schemes, represented in the building typology matrix reported in Table 17. An example of a building prototype schema for a building unit in apartment block constructed between 1961 and 1970 in climatic zones E of Piedmont (E_RES_BU(AB)_CP5) is provided in Appendix C. All datasheets are available in Ballarini et al. (2023). The criteria adopted for defining the building construction periods are described in Section 4.3.2.2.

Table 17. *Building typology matrix for residential buildings in Piedmont region, climatic zone E*

E_PIE	**CP1**	**CP2**	**CP3**	**CP4**	**CP5**	**CP6**	**CP7**	**CP8**
Cat.	≤ 1900	1901-20	1921-45	1946-60	1961-75	1976-90	1991-05	> 2005
SFH	x	x	x	x	x	x	x	x
BU(AB)	x	x	x	x	x	x	x	x
OFF	x	x	x	x	x	x	x	x
EDUC	x	x	x	x	x	x	x	x

Each building prototype schema comprises four categories: geometry, envelope, technical building system, and energy indicators. For each assessed parameter, the median value and interquartile ranges ($Q_3 - Q_2$ and $Q_2 - Q_1$) are provided. The parameters related to the first three categories, illustrated in Table 18, are described below:

- Geometry: The considered parameters include the compactness ratio (CR), thermally heated gross volume ($V_{H;g}$), thermally heated floor area ($A_{H;use}$), and the ratio of transparent envelope area to the total envelope surface (A_w/A_{env}). It should be note that the latter does not represent the WWR, because the opaque thermal envelope area is not distinct in wall, roof, or floor surfaces.
- Envelope: The mean thermal transmittance of opaque (U_{op}) and transparent (U_w) building envelope components is included. The EPC database for the Piedmont Region does not distinguish between different opaque building elements, which represents a limitation of the available data.
- Technical building system: The dataset includes information on the energy carriers predominantly used for space heating, space cooling, and domestic hot water production. Moreover, the mean seasonal

efficiency of the heating generation subsystem ($\eta_{H;gen}$) is presented for the two energy carriers with the highest occurrence in the sample. The overall energy efficiency ($\eta_{H;ngen}$), which also accounts for emission, control, and distribution subsystems, is subsequently reported.

Table 18. *Prototype schema – geometry, envelope, and TBS sections*

	Data	Symbol	Unit of measure	Median	$(Q_3 - Q_2)$	$(Q_2 - Q_1)$
Geometry	Compactness ratio	CR	m^{-1}	0.433	0.235	0.125
	Thermally heated gross volume	$V_{H;g}$	m^3	278	72	57
	Thermally heated floor area	$A_{H;use}$	m^2	72	19	15
	Transparent thermal envelope area on thermal envelope area	A_w/A_{env}	–	11 %	5 %	4 %
Envelope	Mean thermal transmittance of opaque building envelope	U_{op}	W/(m²·K)	1.157	0.215	0.228
	Mean thermal transmittance of transparent building envelope	U_w	W/(m²·K)	3.157	1.416	0.849
Technical building system	Energy carrier per space heating	Natural gas = 78 %; district heating = 15 %; others = 7 % (of the analysed sample)				
	Energy carrier per space cooling	Electricity = 99 %; natural gas = 1 % (of the analysed sample)				
	Energy carrier per domestic hot water	Natural gas = 74 %; electricity = 21 %; others = 5 % (of the analysed sample)				
	Mean seasonal efficiency of the heating generation sub-system (natural gas)	$\eta_{H;gen}$	–	0.926	0.126	0.136
	Mean seasonal efficiency of the heating generation sub-system (district heating)	$\eta_{H;gen}$	–	0.960	0.040	0.263
	Overall energy efficiency (without generation)	$\eta_{H;ngen}$	–	0.878	0.046	0.050

Table 19. *Prototype schema – energy indicators section*

	Data	Symbol	Unit of measure	Median	$(Q_3 - Q_2)$	$(Q_2 - Q_1)$
Energy indicators	Energy need for space heating	$EP_{H;nd;ztc}$	kWh/m²	95.8	51.2	33.2
	Energy need for space cooling	$EP_{C;nd;ztc}$	kWh/m²	16.0	10.8	8.4
	Energy need for domestic hot water	$EP_{W;nd;ztc}$	kWh/m²	18.5	1.3	1.2
	Seasonal space heating energy efficiency	$\eta_{s;H}$	–	0.730	0.070	0.120
	Seasonal space cooling energy efficiency	$\eta_{s;C}$	–	1.000	0.340	0.340
	Seasonal domestic hot water energy efficiency	$\eta_{s;W}$	–	0.630	0.110	0.270
	Non-renewable energy performance for space heating	$EP_{H;nren}$	kWh/m²	134.5	69.3	45.6
	Non-renewable energy performance for space cooling	$EP_{C;nren}$	kWh/m²	13.5	11.6	7.2
	Non-renewable energy performance for domestic hot water	$EP_{W;nren}$	kWh/m²	29.1	18.4	5.2
	Overall non-renewable energy performance	$EP_{gl;nren}$	kWh/m²	171.8	69.7	47.1
	Overall renewable energy performance	$EP_{gl;ren}$	kWh/m²	1.3	10.3	1.1
	Renewable Energy Ratio	RER	–	1 %	4 %	1 %

The second part of the prototype schema, shown in Table 19, reports energy indicators derived statistically from the analysis of the energy performance certificates. KPIs are presented for space heating, space cooling, and domestic hot water. Specifically, the energy need ($EP_{H/C/W;nd}$), seasonal energy efficiency ($\eta_{s,H/C/W}$), and non-renewable energy performance ($EP_{H/C/W;nren}$) are provided. Finally, the overall non-renewable ($EP_{gl;nren}$) and renewable ($EP_{gl;ren}$) energy performance, along with the renewable energy ratio (RER), are reported.

5.2.2.2 Building archetypes

Several building archetype schemes for the Aosta Valley and Piedmont regions were developed by the unit of the research of Politecnico di Torino. The BAs were derived from the energy performance certificates deemed reliable according to the EPC data quality-checking procedure, described in Sections 4.3 and 4.4. For both Italian regions, archetypes were created for residential and non-residential uses. Specifically, single-family houses (RES_SINGLE) and apartment blocks (RES_APPBLOCK) were defined for residential sector, while offices (OFF) and educational buildings (EDUC) were defined for the non-residential sector.

The archetypes for residential apartment blocks were developed starting from individual building units. This choice was motivated by the standard configuration of the Italian EPC system, in which energy certificates are generally issued for single dwelling units, whereas certification for the entire building is required only in specific cases, such as applications for fiscal incentive schemes. Nevertheless, the passive performance of the building fabric of the building unit reflects the characteristics of the whole building. Variations mainly arise in the representation of internal heat gains and TBS.

Regarding internal heat gains, when a single thermal zone is modelled, a predominant building use must be assigned. However, in the absence of detailed occupancy-related information, the URBEM scorecard refers to conventional schedule-based internal gains derived from the technical standard UNI EN 16798-1 (UNI, 2025).

Concerning the performance of technical building systems, only limited information was available from the EPC database. In particular, the dataset included the energy carrier associated with each energy service (space heating, space cooling, and DHW), together with the nominal power of the autonomous generation system. The efficiency of the generator has to be derived from appropriate technical datasheets, since, even within the same climatic zone, different local external conditions, corresponding to different municipalities, can affect system performance. Consequently, performing a statistical analysis on the aggregated cluster would have led to misleading results.

To increase the representativeness of the non-residential samples, construction periods were aggregated to ensure a minimum of five EPCs per group. The following assumptions were made for the two regions:

- Piedmont: Climatic zones E (Po Valley area) and F (alpine and sub-alpine areas) were considered separately. Building construction periods were defined in 10-year intervals starting from 1930.
- Aosta Valley: The regional territory includes both climate zones E and F. However, the city of Aosta and the municipalities closest to the capital (Quart, Saint-Christophe, and Sarre) were grouped within zone E, since the others—although technically classified as zone E—are more comparable to mountain buildings (zone F). For this reason,

archetypes have been defined with a combined “E-F” climate zone. Regarding the construction periods, these differ from those of Piedmont because the certifier could either enter an estimated year (selectable from predefined construction period ranges established by the database manager) or a specific year. To preserve this information, the construction period ranges defined by Aosta Valley were maintained, with exact years assigned to the corresponding intervals.

Table 20 to Table 26 present building typology matrices, indicating in each cell the number of EPCs considered in the analysis. An example of building archetype schema for an apartment block constructed between 1962 and 1971 in climatic zones E-F of the Aosta Valley region (RES_APPBLOCK_1962-1971_E-F_VAL) is provided in Appendix D. All datasheets are available on the project web-site: https://www.urbem.polimi.it/en/.

Table 20. *Building typology matrix for residential buildings in Aosta Valley region, climatic zone E*

E_VAL	**CP1**	**CP2**	**CP3**	**CP4**	**CP5**	**CP6**	**CP7**	**CP8**
Cat. (*)	-1919	1919-1945	1946-1961	1962-1971	1972-1981	1982-1991	1992-2005	2006-
R_S	19	31	39	41	38	42	42	86
R_A	657	849	1613	2152	1184	645	534	494

(*) R_S = RES_SINGLE; R_A = RES_APPBLOCK

Table 21. *Building typology matrix for non-residential buildings in Aosta Valley region, climatic zone E*

E_VAL	**CP1**	**CP2**	**CP3**
Cat. (*)	-1945	1946-1981	1982-2005
O	12	12	7
E	5	14	14

(*) O = OFF; E = EDUC

Table 22. *Building typology matrix for residential buildings in Aosta Valley region, climatic zones E-F*

E-F_VAL	**CP1**	**CP2**	**CP3**	**CP4**	**CP5**	**CP6**	**CP7**	**CP8**
Cat.	-1919	1919-1945	1946-1961	1962-1971	1972-1981	1982-1991	1992-2005	2006-
R_S	341	255	170	173	185	124	160	319
R_A	2278	1486	1933	2450	3190	1540	1176	1280

Table 23. *Building typology matrix for non-residential buildings in Aosta Valley region, climatic zones E-F*

E-F_VAL	CP1	CP2	CP3	CP4
Cat.	-1945	1946-1981	1982-2005	2006-
O	24	28	28	10
E	12	15	10	

Table 24. *Building typology matrix for residential and non-residential buildings in Piedmont region, climatic zone E*

E_PIE	CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8	CP9	CP10
Cat.	-1930	1931-1940	1941-1950	1951-1960	1961-1970	1971-1980	1981-1990	1991-2000	2001-2010	2011-
R_S	12807	3017	3106	3174	4181	3385	2141	2379	2491	2852
R_A	58853	14591	16087	48723	76504	33754	14645	12567	18517	13162
O	254	32	57	78	137	132	122	133	156	159
E	114	20	26	50	106	112	52	32	18	35

Table 25. *Building typology matrix for residential buildings in Piedmont region, climatic zone F*

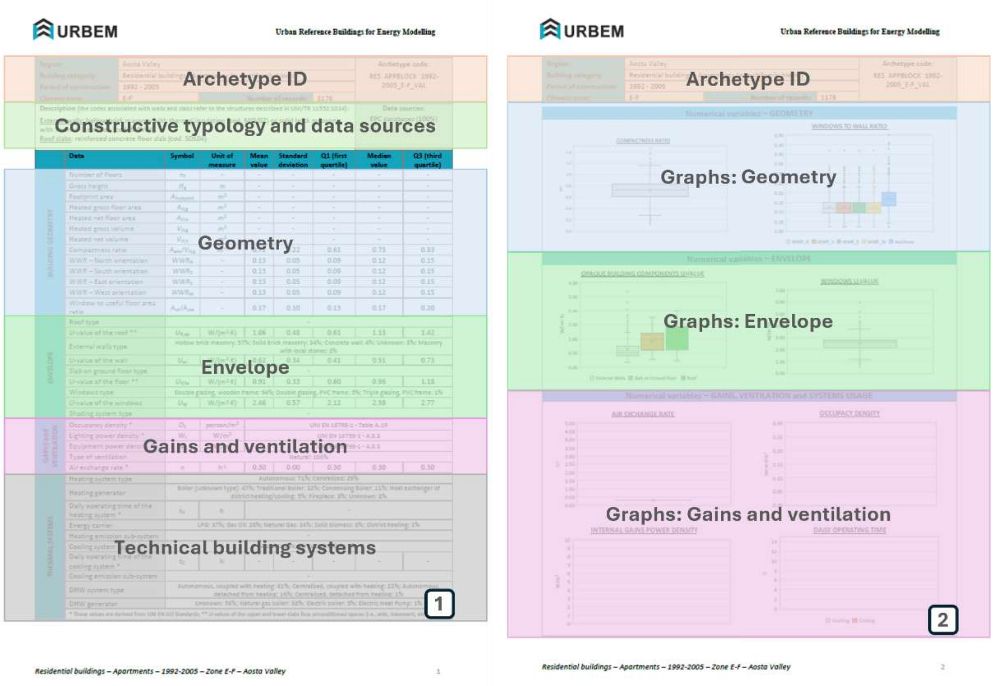
F_PIE	CP1	CP2	CP3	CP4	CP5	CP6	CP7	CP8	CP9	CP10
Cat.	-1930	1931-1940	1941-1950	1951-1960	1961-1970	1971-1980	1981-1990	1991-2000	2001-2010	2011-
R_S	3475	452	474	442	694	678	386	322	298	478
R_A	7739	1185	1516	2550	5946	5747	2101	1267	1301	1165

Table 26. *Building typology matrix for non-residential buildings in Piedmont region, climatic zone F*

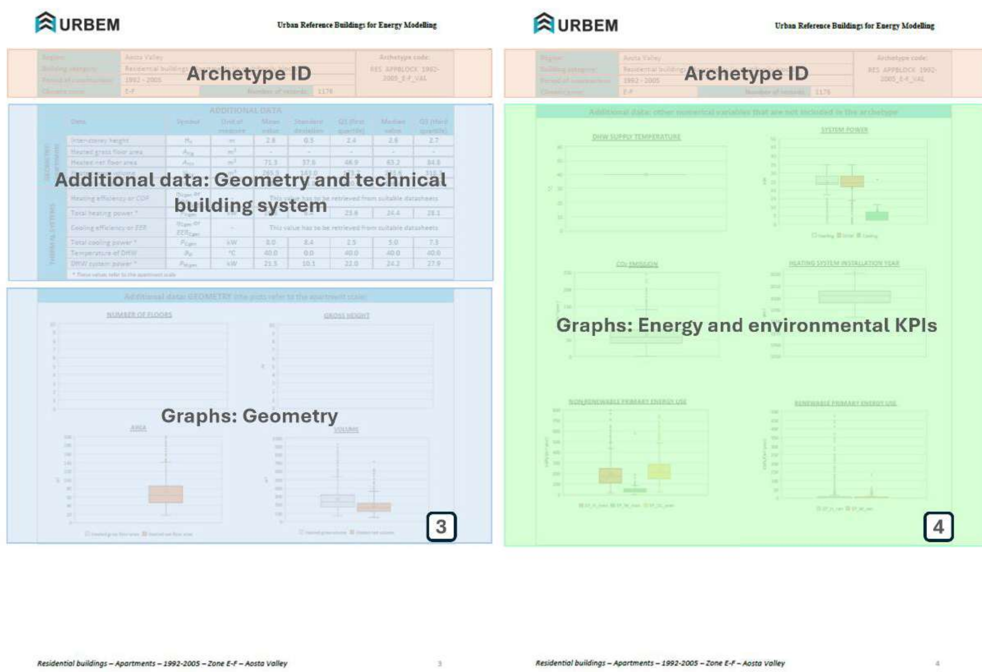
F_PIE	CP1	CP2	CP3	CP4	CP5
Cat.	-1950	1951-1980	1981-2000	2001-2010	2011-
O	61	26	16	9	10
E	39	30	6		

The URBEM archetype schema is a four-page document, illustrated in Figure 53, consisting of several sections, each with a specific layout and purpose. A slight variation in the graphical presentation is reported depending on whether the dataset represents a building or a building unit. The upper section of every page is identical and includes the identification information, such as the Italian region, building use category, construction period, climatic zone, number of records, and archetype code. Below the structure of the archetype schema for a representative building unit is described:

- **Page I –** General description and numerical-categorical data.
  This first page presents the numerical data of the archetype. Below the identification header, the constructive typologies of walls, roofs, and floors are described, referring to the alphanumeric codes of the UNI/TR 11552 (UNI, 2014) technical standard, which defines the sequence and thermal properties of the layers. The sources of data and their percentage shares are also reported (e.g., EPC databases, technical documentation, thermal plant registers).
  Categorical and numerical variables are grouped into four main sections: geometry, envelope, gains and ventilation, and technical building systems. Each variable is represented by its mean, standard deviation and first, second, and third quartiles. The geometry section mainly reports data on building height, useful floor area, volume, and *WWR* by façade orientation. The envelope section includes thermal transmittances for different building elements, while the window data specify the predominant frame and glazing types. Parameters related to internal heat gains—such as occupancy density (O_c), lighting power density (W_L), equipment power density (W_A)—are linked to their corresponding technical standard, serving as guidance for beginner UBEM developers. The last section characterises the space heating and DHW system data.
- **Page II –** Graphical visualisation of data: geometry, envelope, gains, ventilation and system usage.
  This page presents boxplot visualisations showing the distribution of geometrical and envelope data. For the envelope, one graph illustrates the *U*-values for different opaque elements, an another focuses on transparent components.
- **Page III –** Additional numerical section and graphical visualisation of data: gains, ventilation and system usage.
  The third page provides further numerical data and graphical representations related to internal gains, ventilation parameters, and building system operation.
- **Page IV –** Graphical visualisation of data: energy and environmental KPIs.
  The final page reports mainly KPIs, such as CO_2 emissions, renewable non-renewable primary energy use for heating and DHW ($EP_{H/W;ren/nren}$), and the overall non-renewable performance indicator ($EP_{gl;nren}$).



(**a**)



(**b**)

Figure 53. *URBEM archetype schema for representative building unit: pages 1-2 (**a**), pages 3-4 (**b**)*

5.2.3 Application at large-scale

5.2.3.1 Building prototypes

This section presents the key elements to develop a regional NBRP, aimed at assessing the current energy performance *status* and identifying potential renovation pathways to enhance energy efficiency at a large-scale.

The baseline energy balance can be established correlating the energy performance indicators included in the prototype schema with the geometric prevalence (number of buildings or total useful floor area) of the corresponding building types. This latter information can be obtained from the Italian National Institute of Statistics (ISTAT).

Regarding the energy refurbishment scenarios, the analysis relied solely on the information available in the energy performance certificates. Consequently, only the statistical variability of the overall non-renewable energy performance indicator ($EP_{gl;nren}$) was considered. This choice stems from the fact that, within the recommendation section of the Italian EPC, the effectiveness of an energy efficiency measure is assessed exclusively through the variation of $EP_{gl;nren}$.

It should be noted that the analysed renovation scenarios are derived directly from the EPC recommendations and do not include additional constraints or feasibility checks on the applicability of retrofit measures. Therefore, the study does not aim to identify optimal renovation packages, but rather to statistically evaluate the prevalence and potential impact of the most recurrent EPC-recommended actions.

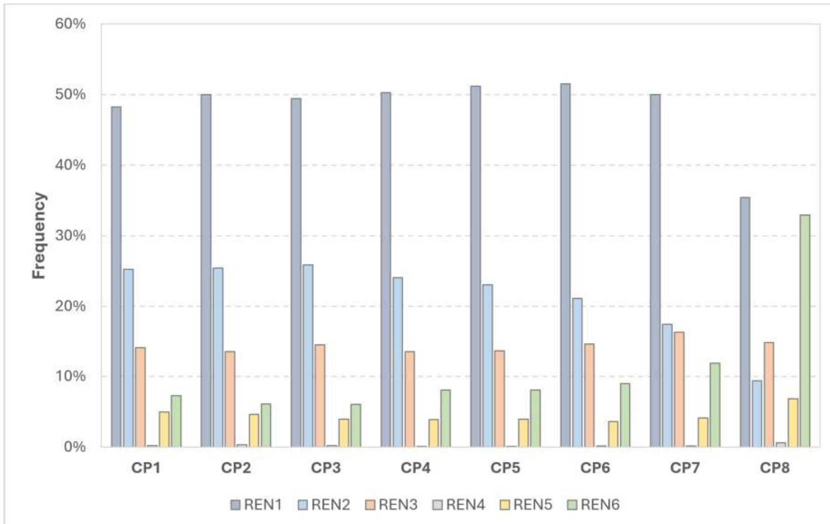
According to the EPC framework, the recommended energy efficiency measures (RENs) are grouped into six categories: REN1, REN2, REN3, REN4, REN5, and REN6. Each retrofit measure corresponds to specific interventions (see Table 27), addressing either the building fabric (REN1 and REN2) or the technical building systems (from REN3 to REN6).

Table 27. *Energy efficiency recommendations in the EPC schema*

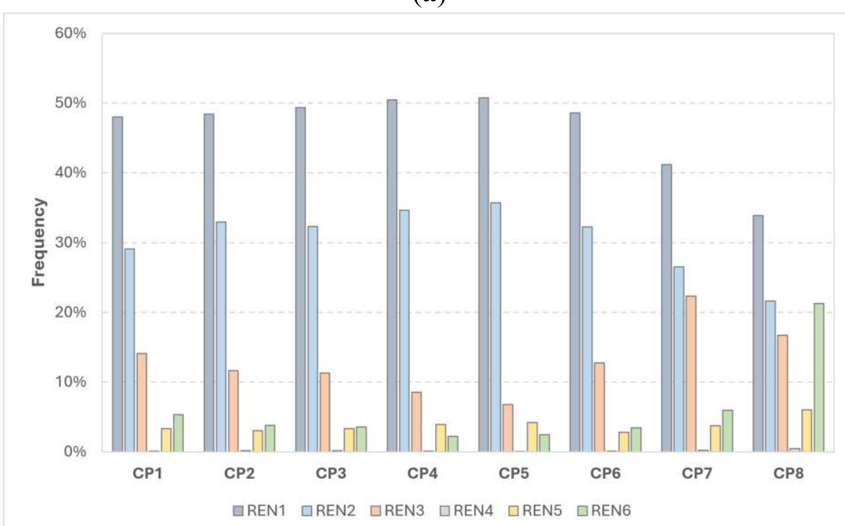
System	REN ID	Energy efficiency measure
Building fabric	REN1	Opaque building envelope
	REN2	Transparent building envelope
Technical building system	REN3	Space heating system
	REN4	Space cooling system
	REN5	Other systems
	REN6	Renewable energy plants

For each recommended energy efficiency measure recorded in the EPC database, the prevalence of each REN category was analysed by construction period and building type—SFH and BU(AB)—as shown in Figure 54a-b.

The most predominant renovation action is the refurbishment of opaque building elements (REN1), followed by the replacement of windows with more energy-efficient units (REN2). The third most frequent measure involves the replacement of technical systems for space heating. Moreover, the influence of policies promoting renewable energy sources can be observed in the more recent CPs, particularly for SFHs (Figure 54a).



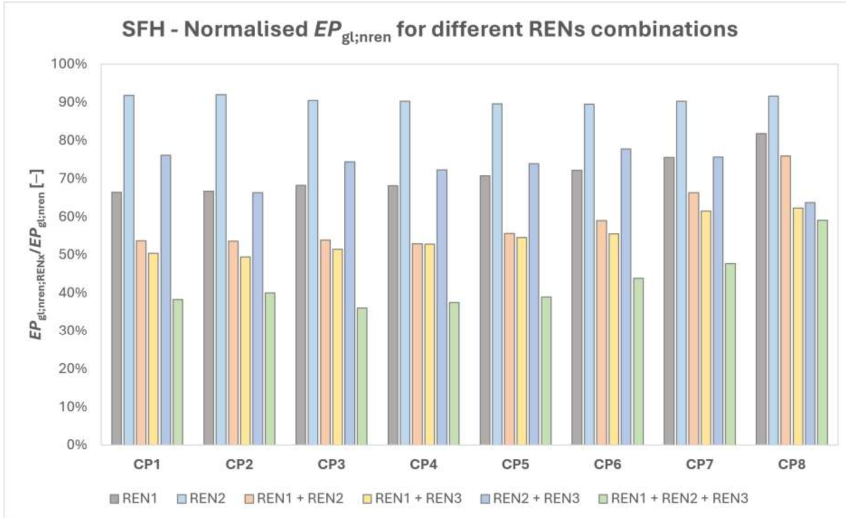
(a)



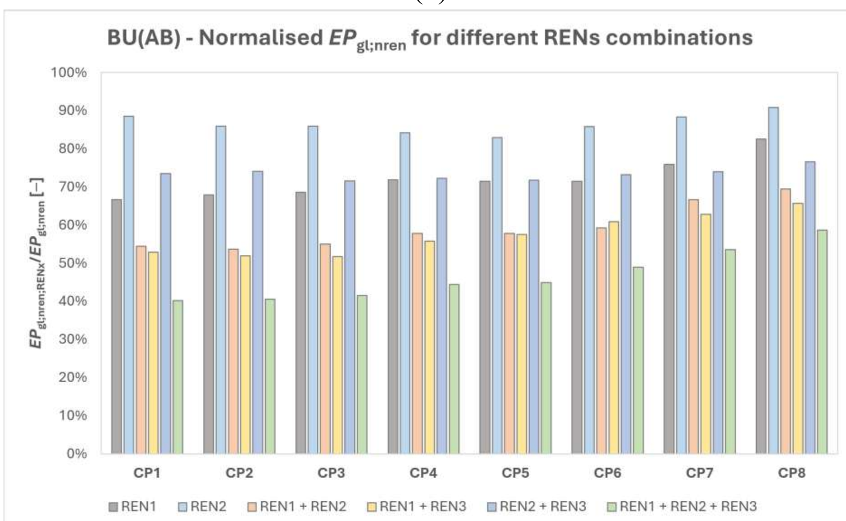
(b)

Figure 54. *Frequency for energy efficiency recommendations by CP for single-family houses (**a**) and building units in apartment blocks (**b**)*

The impact in the overall non-renewable energy performance indicator for the most widespread RENs, as well as the synergic effect obtained by combining differently REN1, REN2, and REN3 is illustrated in Figure 55a-b for SFHs and BU(AB), respectively.



(**a**)



(**b**)

Figure 55. *Percentage decrease in the $EP_{gl;nren}$ for individual and combined renovation measures in single-family house (**a**) and building unit in apartment block (**b**) for different construction periods*

5.2.3.2 Building archetypes

This section presents a study that applies the developed building archetypes to the municipality of Aosta, providing guidance for UBEM developers in optimising key BA parameters during model implementation.

Developing a large-scale energy model, even when using archetypes, is a time-consuming process that greatly benefits from clear recommendations and structured guidelines. Using a one-at-a-time local sensitivity analysis, the study aims to identify the most influential metrics within the BA schema, mainly derived from the EPCs of the Aosta Valley Region, used as a case study. The resulting data sheets assist UBEM developers in selecting appropriate values to characterise building energy performance and understanding how these choices affect simulation outcomes. To assess this influence, a discrete uniform distribution was applied over three intervals (first, second, and third quartiles), testing all possible combinations of BA parameters in CitySim to evaluate their impact on the thermal energy need for space heating.

The proposed methodology aims to evaluate the influence of key parameters within the BAs on urban energy modelling through an archetype-based UBEM approach. The procedure consists of the following steps:

1. extracting the relevant input data from the BA schema,
2. defining the variation range and probability distribution for each selected independent variable,
3. applying a statistical or data-driven method to quantify the relative influence of each input parameter on the model outcome, and
4. performing scenario analyses using a UBEM tool with varying inputs to assess the resulting output variability.

Multivariate regression analysis is commonly employed to evaluate the sensitivity of model outputs to different input parameters. The general formulation of a multivariate linear regression is presented in Equation (6), where Y_i denotes the output of the large-scale energy model, β_0 represents the intercept, β_j denotes the regression coefficients for the independent variables x_{ip}, and ε_i is the residual error for the i-th observation.

$$Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_j x_{ip} + \varepsilon_i \qquad i = 1, \dots, n \quad (6)$$

The relative importance of each predictor can be expressed through a weighting factor w_j, defined as the ratio between an individual coefficient and the sum of all coefficients, as reported in Equation (7). This normalisation is fundamental for comparing the influence of variables with different unit of measures.

$$w_j = \frac{\beta_j}{\sum_{k=1}^{p} \beta_k} \qquad j = 1, \dots, n \quad (7)$$

The case study examines a real residential block located in the municipality of Aosta (583 m a.s.l., heating degree days: 2876 °C·day), consisting of 18 buildings. The construction periods were identified following the methodology proposed by D'Alonzo et al. (2020). Figure 56 provides an overview of the analysed residential building stock, illustrating building codes, construction period ranges, and compactness ratios. Three construction age ranges were considered—1919-45 (BA_1), 1946-60 (BA_2), and 1961-70 (BA_3)—and the _APPBLOCK BA schema was assigned accordingly.

Figure 56. *Satellite view of assessed urban block (Piro, Ballarini, et al., 2025a)*

The LoD of the urban geometry model was progressively refined through successive steps. Initially, the building footprints (LoD0) were extracted from the OpenStreetMap database. These were then extruded to generate simplified shoebox models (LoD1), and roof slopes were subsequently incorporated (LoD2). Further refinements were implemented to improve model accuracy, such as converting erroneous building volumes into shading elements (see Figure 57). Nevertheless, despite the geometric refinements, each building volume was modelled as a single thermal zone.

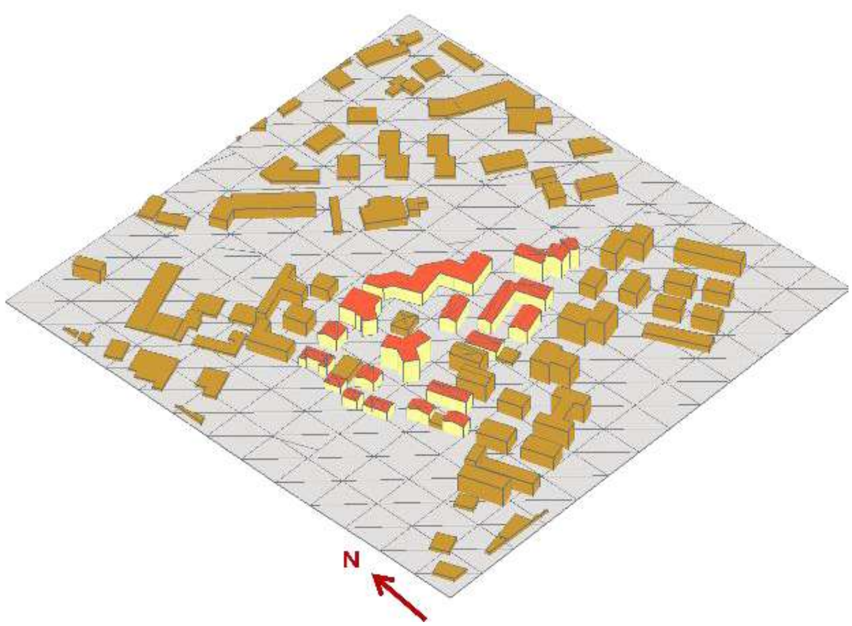


Figure 57. *Urban scene imported in CitySim Pro (Piro, Ballarini, et al., 2025a)*

The first (Q_1), second (Q_2), and third (Q_3) quartiles of key parameters—window-to-wall ratio (WWR), thermal transmittance of walls (U_{wl}), roof ($U_{fl;up}$), floor ($U_{fl;lw}$), and windows (U_w)—are summarised in Table 28. These values were statistically derived from the EPCs of the Aosta Valley Region within the UBEM research framework. In addition, two independent variables were included across all BA groups: the space-heating temperature set-point ($\theta_{H;set}$) and ventilation air flow rate (n). According to technical standards EN 12831-1 (CEN, 2017d) and UNI EN 16798-1 (UNI, 2025), these variables were assigned the following discrete intervals: n = {0.30, 0.40, 0.50} and $\theta_{H;set}$ = {18.0, 20.0, 21.0}, respectively. A discrete uniform distribution was applied to all inputs.

Table 28. *BA characteristics per construction period*

	BA_1 (1919-45)			BA_2 (1946-60)			BA_3 (1961-70)		
	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3	Q_1	Q_2	Q_3
WWR –	8 %	11 %	14 %	10 %	13 %	16 %	11 %	14 %	16 %
U_{wl} W·m^{-2}·K^{-1}	0.87	1.40	1.91	0.82	1.13	1.27	0.89	1.12	1.23
$U_{fl;up}$ W·m^{-2}·K^{-1}	0.88	1.32	1.45	1.09	1.34	1.45	1.10	1.32	1.42
$U_{fl;lw}$ W·m^{-2}·K^{-1}	1.07	1.15	1.36	1.10	1.15	1.20	1.11	1.14	1.19
U_w W·m^{-2}·K^{-1}	2.28	2.61	2.85	2.09	2.69	3.08	2.16	2.77	3.32

Simulations employed the Typical Meteorological Year (TMY) file developed by the Italian Thermotechnical Committee. Internal heat gain profiles and intensities—accounting for occupants, appliances, and lighting—were derived from the Italian National Annex of the UNI EN 16798-1 (UNI, 2025). Due to the absence of tailored occupancy data, a deterministic schedule-based approach representative of residential apartment blocks was therefore adopted.

The influence of seven independent parameters on the thermal energy need for space heating was evaluated for each BA group using a three-step interval and a discrete uniform distribution. As previously described, the three discrete values correspond to the first, second, and third quartiles of the BA schema. Using a one-at-a-time local sensitivity analysis, each input parameter was varied individually while keeping all others constant. This approach highlights the considerable influence of key UBEM input parameters on the overall energy performance of the building stock (see Figure 58). The algorithm used to generate and test all possible input combinations within CitySim is available in Kämpf (2024).

Three simulation sets were carried out for each BA schema representing residential buildings (_APPBLOCK) constructed in the periods 1919-45, 1946-60, and 1961-70. Each set included 3^7 = 2187 runs, resulting in a total of 3^8 = 6561 simulations.

Within each simulation group (e.g., BA_1), the thermal properties of the remaining representative building groups (BA_2 and BA_3) were kept constant, while the parameters of the target archetype were varied. To accurately account for short-wave and long-wave radiative exchanges within the urban environment, CitySim simulations of the other two BA groups were performed without treating them as shading objects.

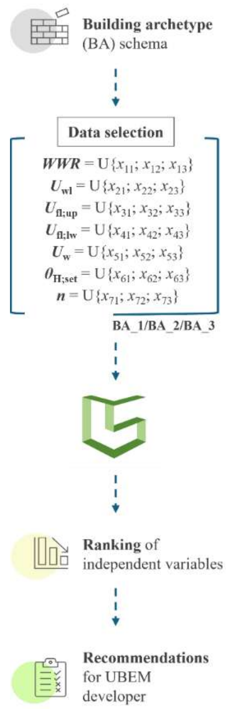


Figure 58. *Workflow diagram (Piro, Ballarini, et al., 2025a)*

Finally, the physical parameter values were normalised into logical values {0, 1, 2} to enable consistent, comparable regression analyses across all buildings and to calculate the weighting factor w_j, as defined in Equation (7).

The municipality of Aosta is characterised by a climate in which space heating dominates total energy demand. Accordingly, the analysis focused exclusively on variations in space heating energy need across different building configurations.

For each simulated building, a multivariate regression model was developed to correlate the seven independent variables with the thermal energy need for space heating. Figure 59 presents the simulation results, expressed as normalised regression coefficients for the 18 residential buildings within the analysed urban block.

The residential apartment blocks were classified into three construction periods: 1919-45 (buildings A3-F5), 1946-60 (buildings B3-G5), and 1961-70 (buildings A4, A5, and B2). The space-heating temperature set-point ($\theta_{H;set}$) emerged as the most influential variable across all construction periods, particularly for the more recent buildings, where it explains up to 42 % of the variation. The thermal transmittance of walls (U_{wl}) plays a major role in older buildings but progressively decreases in relevance from the 1946-60 to the 1961-70 periods. Under winter conditions, parameters related to the transparent envelope—namely, the *WWR* and window thermal transmittance (U_w)—show a relatively minor influence.

Among the opaque components, the thermal transmittance of the roof ($U_{fl;up}$) determines a greater impact than that of the floor ($U_{fl;lw}$). Moreover, ventilation rates show an increasing influence on overall energy performance.

These results confirm the usefulness of segmenting the building stock through building archetypes, as evidenced by the differing behaviours across representative buildings, the consistent trends within each age group, and the varying influence of independent variables. In this specific climatic context, parameters requiring less optimisation during large-scale energy model development are those associated with transparent envelope components and floor thermal transmittance.

Overall, the proposed methodology provides practical guidance for UBEM developers engaged in assessing the energy performance of building stocks through the use of archetypes.

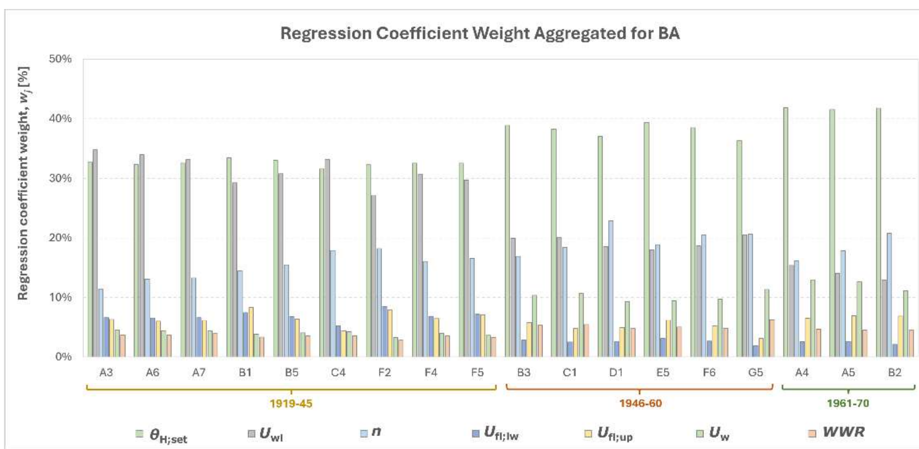


Figure 59. *Regression coefficient weight aggregated for BA (Piro, Ballarini, et al., 2025a)*

5.2.4 Discussion

For north-western Italy, a digital library was developed to systematise and digitalise information on building energy performance. Within this framework, probabilistic building prototypes and archetypes—providing representative ranges of possible values for key variables—were defined. These constitute a fundamental tool for multiple stakeholders aiming to characterise the energy *status* of both residential and non-residential building stocks.

Although the input parameters are intrinsically continuous, their variability was represented through three discrete levels corresponding to the quartiles of the archetype library. Consequently, the sensitivity analysis captures parameter influences over a discretised subset of the input space rather than over the full continuous domain.

For policymakers and public administrators, representative buildings serve as essential instruments to assess inefficiency levels and to guide deep renovation strategies in the most vulnerable urban or regional areas. From the UBEM developer's perspective, these reference buildings are useful and crucial for reducing modelling effort and mitigating uncertainty, thus enabling more robust and reliable urban-scale energy analyses.

However, the completeness of the archetype schemes depends strongly on the territorial data coverage and the degree of harmonisation among local databases. In the absence of specific information, e.g., occupant-related data and technical building system performance, the representative datasheet provides guidelines and references to relevant technical standards, although a basic technical knowledge is required for correct application.

Moreover, most of the energy certificates stored in the regional databases refer to individual units within multi-family buildings or apartment blocks. This introduces challenges in reconstructing the overall building geometry and estimating the capacity of shared technical building systems. From an archetype development perspective, working at the whole-building scale would therefore be preferable to ensure geometric and system representativeness.

The segmentation criteria of building size and shape are consistent with the approach adopted in DOE scorecards, which, for instance, provide typical characteristics for small, medium, and large offices. These distinctions influence occupancy-related aspects such as thermal comfort and indoor air quality requirements, internal gain intensity, and technical building system configurations. The current Italian EPC framework does not make this kind of distinction. For example, the educational building category includes all school levels (kindergartens, primary schools, universities, etc.), thereby limiting the granularity and accuracy of archetype development.

5.3 Representative urban blocks

The evolution from representative buildings to representative urban blocks plays a crucial role in assessing citizens' vulnerability to extreme heat waves. This transition is necessary to investigate the urban heat island effect at the micro-scale, where urban morphology—particularly the street-canyon configuration—becomes the main framework for analysis. This chapter describes a methodology for selecting representative parts of the city and assessing the short-, mid-, and long-term building stock energy performance.

5.3.1 Methodology

This chapter presents a methodology developed within the framework of the PRIN2022-PNRR CRiStAll project aimed at localising typical urban blocks, considered geometrical representative and vulnerable to UHI intensity, and assessing their energy performance in short-, medium-, and long-term. The exposure to microclimate variation of urbanised areas is studied through land surface temperature satellite images. Selected the representative city blocks, the approach foreseen the development of archetype-based UBEM, assigning typical non-geometric set of properties. This methodology was applied across different Italian climate zones with heterogeneous informative systems: Bari (climatic zone C), Rome (climatic zone D), and Turin (climatic zone E).

5.3.1.1 The CRiStAll project

The literature review indicates that quantifying UHIs over large territorial scales remains challenging due to the limited availability of microscale climate data—both historical observations and future projections—and the difficulty of extending results spatially across heterogeneous urban forms. Moreover, archetype-based UBEM studies often adopt predefined, idealised, or parametrically varied configurations, which constrains their ability to derive representative urban blocks directly from the full urban building stock.

These gaps can be addressed by meeting the following concurrent needs:

- incorporating local microclimatic conditions into urban-form characterisation,
- integrating urban morphology into archetype-based modelling, and
- coupling climate-change projections with urban forms within climate datasets.

The research was performed within the *Climate Resilient Strategies by Archetype-based Urban Energy Modelling* (CRiStAll) project (CRiStAll, 2025), funded by the European Union – Next Generation EU within the PRIN 2022 PNRR program, addresses three interconnected research lines:

A. the development of urban climate models that integrate both short-, medium-, and long-term climate change effects (future weather data) and UHI impacts at the micro-scale,
B. the implementation of an archetype-based UBEM representing typical urban configurations, and
C. the evaluation of climate-resilient and UHI mitigation strategies in urban environments.

An overview of the CRiStAll methodology is presented in Figure 60. The interrelations among these research lines can be summarised as follows:

- (A-B): the micro-scale urban climate model is informed by the morphology, surface materials, and building geometry characterising the typical urban configurations,

- (B-C): the UBEM, based on representative buildings within representative urban contexts, is expanded to include climate-resilient and UHI mitigation measures,
- (C-A): the effects of these strategies are then integrated back into the urban climate model.

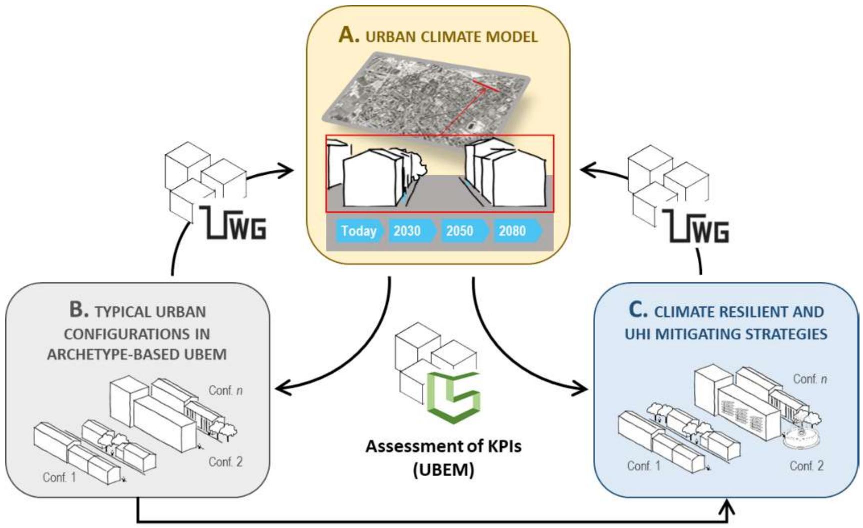


Figure 60. *CRiStAll project methodology*

5.3.1.2 Geometrical representativeness and thermal vulnerability

The selection of the urban block is based on connecting two sources of information: the geometry representativeness of city blocks, and the land surface temperature from NASA ECOSTRESS. This methodology was detailed expressed in Borelli et al. (2025) and Ballarini et al. (2026), and was developed by colleagues at the *Free University of Bozen-Bolzano (Italy)*.

Common urban metrics from literature, such as Floor Area Ratio (*FAR*), Surface Coverage (*SC*), Volume Area Ratio (*VAR*), Green Ratio (*GR*), Relative Compactness (*REC*), Average Building Height (*ABH*), Shape Factor (*SF*), Vertical to Horizontal Ratio (*VtH*), Average Building Distance (*ABD*), and Sky View Factor (*SVF*), were calculated for every urban blocks. Then a Spearman correlation test was performed for choosing the more independent metrics in order to facilitate the selection of the most geometrical representative archetypes: *GR*, *SC*, *ABH*, and *VtH*.

The second phase of the procedure employs satellite data to identify urban blocks most affected by the UHI phenomenon. In particular, ECOSTRESS satellite imagery of Land Surface Temperature (*LST*) for the target city is analysed.

Finally, the results of both first and second analyses are combined, identifying among the representative urban blocks, those with the largest average *LST*.

5.3.1.3 Key Performance Indicators

This section defines the KPIs adopted to quantify the intensity, frequency, and sensitivity of the urban warming effect under different climate change scenarios. Table 29 summarises the selected metrics, encompassing both building energy performance and climate-related indicators, along with their symbology, terminology, units, and corresponding references.

Table 29. *List of assessed KPIs*

Quantity	**Symbol**	**Unit**	**Ref.**
Energy Performance			
Energy need for space heating per unit conditioned floor area	$EP_{\mathrm{H;nd}}$	kWh·m^{-2}	CEN (2017f, 2017a)
Energy need for space cooling per unit conditioned floor area	$EP_{\mathrm{C;nd}}$	kWh·m^{-2}	
Peak heating load per unit conditioned floor area	$\phi_{\mathrm{H;ld}}$	W·m^{-2}	
Peak cooling load per unit conditioned floor area	$\phi_{\mathrm{C;ld}}$	W·m^{-2}	
Climate			
Ambient Warmness Degree	*AWD*	°C	Hamdy et al. (2017)
Heating Degree Days	*HDD*	°C·d	CEN (2007d)
Cooling Degree Days	*CDD*	°C·d	
Urban Heat Island Intensity	*UHII*	°C	Oke (1973)

The thermal energy need for space heating and cooling ($EP_{\mathrm{H/C;nd}}$) are selected to evaluate the energy performance of the building stock under different climate scenarios. In addition, the hourly heating and cooling peak loads ($\phi_{\mathrm{H/C;ld}}$) throughout the calculation period are essential for appropriately sizing technical building systems. These indicators can be directly obtained from any UBEM tool.

Climate-related KPIs are directly dependant on external air temperature. The Ambient Warmness Degree (*AWD*), as defined in Hamdy et al. (2017), quantifies both the intensity and duration of elevated outdoor temperatures relative to a reference base temperature—typically 18 °C, which represents the lower comfort threshold during the heating season. As expressed in Equation (8), only positive deviations between i-th hourly outdoor air temperature ($\theta_{\mathrm{e};i}$, in °C) and the base temperature (θ_{b}, in °C) are considered in the summation and normalised by the total number of occupied hours (N) for the building.

$$AWD = \frac{\sum_{i=1}^{N} \left(\theta_{\mathrm{e};i} - \theta_{\mathrm{b}}\right)^{+} t_i}{\sum_{i=1}^{N} t_i} \tag{8}$$

The Heating Degree Days (*HDD*) and Cooling Degree Days (*CDD*) are synthetic indicators used to characterise how warm or cold a given area is over a specified period. The Urban Heat Island Intensity (*UHII*) quantifies the temperature difference between an urban weather station (UWS) and a rural weather station (RWS). This indicator can be derived either from the average monthly temperature difference between UWS and RWS (Oke, 1973), or the monthly maximum and minimum outdoor air temperature at the two weather stations (Battista et al., 2023). The first method is illustrated in Equation (9), where the $UHII_m$ denotes the value calculated for the m-th month. In this equation, $\bar{\theta}_{\mathrm{e,UWS;h};m}$ (in °C) represents the average monthly external air temperature measured at the UWS, while $\bar{\theta}_{\mathrm{e,RWS;h};m}$ (in °C) refers to that recorded at the RWS.

$$UHII_m = \bar{\theta}_{\mathrm{e,UWS;h};m} - \bar{\theta}_{\mathrm{e,RWS;h};m} \tag{9}$$

5.3.2 Application

5.3.2.1 Case studies

The methodology, developed as part of the Ph.D. research activity, was applied to three Italian municipalities belonging to different climate zones. The representative urban blocks are illustrated in Figure 61, through the visualisation in CitySim Pro environment. UBEM scenes are composed by buildings, shadings, and ground meshes, where the 3D objects are represented with a LoD1. Moreover, an example of scorecard for a representative urban block located in the municipality of Bari is provided in Appendix E.

However, in this section the urban blocks A, B, and C for Turin will be analysed. Table 30 presents the main general and geometric urban metrics for blocks A, B, and C. The selected KPIs include the number of thermally assessed buildings, surface coverage (*SC*), average building height (*ABH*), and compactness ratio ($A_{\mathrm{env}} \cdot V^{-1}$). Among the three, urban block C shows the highest median compactness ratio, which is expected to influence the heat transfer through the building envelope. In contrast, urban block B is characterised by taller buildings and the highest degree of urbanisation, as reflected by its *SC* value.

Table 30. *Relevant urban metrics for urban blocks A, B, and C in Turin*

Parameter	Unit	Urban block		
		A	B	C
No. of assessed buildings	–	17	8	9
Surface Coverage, *SC*	–	0.322	0.388	0.323
Average Building Height, *ABH*	m	18.8	20.5	17.4
Compactness Ratio (median), $A_{\mathrm{env}} \cdot V^{-1}$	m^{-1}	0.300	0.251	0.316

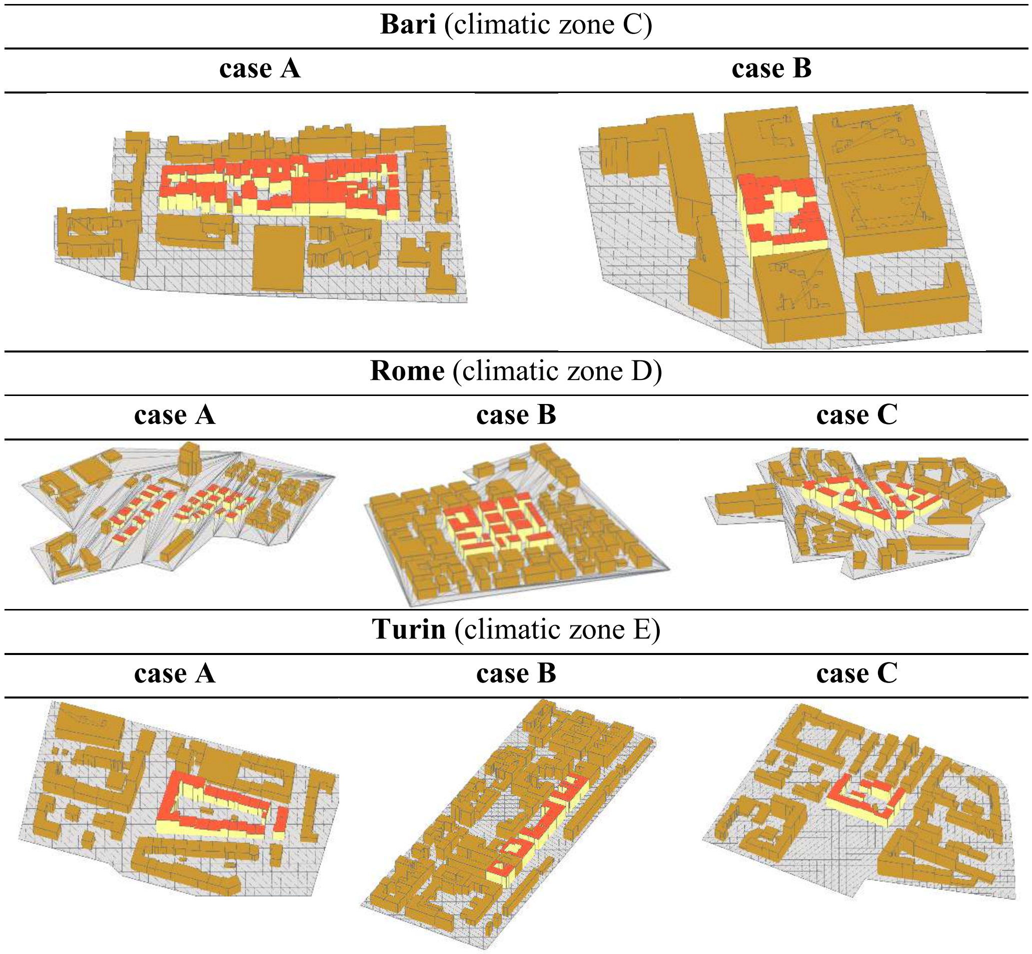


Figure 61. *Representative urban blocks for Bari, Rome, and Turin*

Figure 62, Figure 63, and Figure 64 report the 3D geometric model of urban block A, B, and C highlighting the associated building codes. From this perspective, Table 31, Table 32, and Table 33 report both the geometrical and thermophysical characteristics of the assessed buildings, for the analysed urban blocks. These parameters include building floor area (A_{fl}), total thermal envelope area (A_{env}), volume (V), window-to-wall ratio (WWR), compactness ratio ($A_{env}{\cdot}V^{-1}$), and mean thermal transmittance for opaque (U_{op}) and transparent (U_{wi}) components.

Figure 62. *Urban block A (Turin) visualised in CitySim Pro*

Table 31. *Geometrical and mean thermal characteristics for opaque and transparent components for urban block A (Turin)*

Bldg. ID	Constr. period	A_{fl} m^2	A_{env} m^2	V m^3	WWR –	$A_{env}\cdot V^{-1}$ m^{-1}	U_{op} W/(m^2·K)	U_w W/(m^2·K)
A_1		524	648	1917		0.338	1.45	
A_2		776	1357	2654		0.511	1.42	
A_3		371	671	1508		0.445	1.47	
A_4		748	1045	2860		0.365	1.47	
A_5		639	743	2364		0.314	1.46	
A_6		13200	11126	45540		0.244	1.43	
A_7		1387	1315	5086		0.258	1.46	
A_8	1921-45	877	954	3244	14 %	0.294	1.47	3.09
A_9		849	880	2943		0.299	1.45	
A_10		980	949	3409		0.278	1.48	
A_11		875	937	3227		0.290	1.47	
A_12		689	707	2480		0.285	1.48	
A_13		775	868	2964		0.293	1.49	
A_14		665	713	2380		0.300	1.47	
A_15		351	480	1321		0.363	1.51	
A_16	1946-60	587	661	2085	13 %	0.317	1.28	3.12
A_17		1401	1680	5044		0.333	1.27	

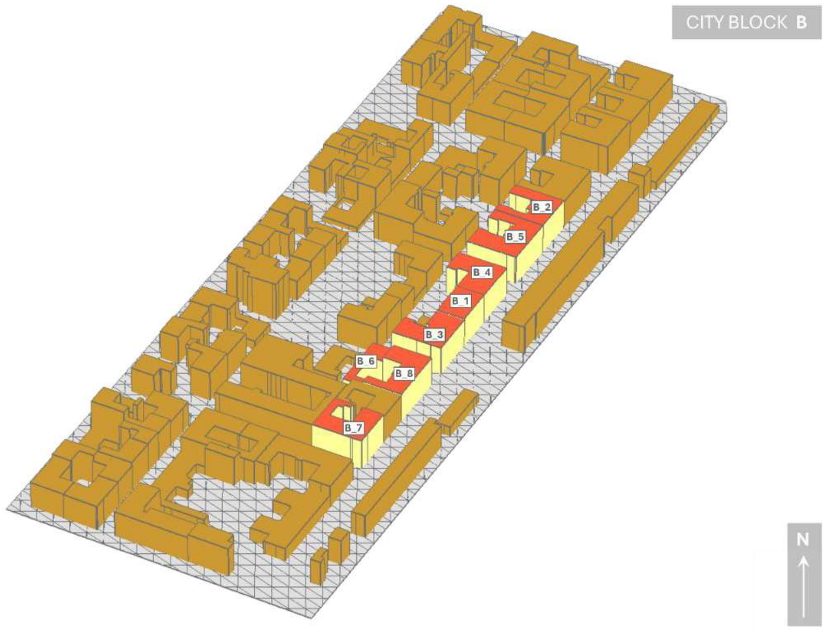


Figure 63. *Urban block B (Turin) visualised in CitySim Pro*

Table 32. *Geometrical and mean thermal characteristics for opaque and transparent components for urban block B (Turin)*

Bldg. ID	**Constr. period**	A_{fl} m^2	A_{env} m^2	V m^3	WWR –	$A_{env} \cdot V^{-1}$ m^{-1}	U_{op} W/(m^2·K)	U_w W/(m^2·K)
B_1		2909	2103	10179		0.207	1.50	
B_2		4018	3558	14197		0.251	1.47	
B_3	1901-20	4698	3969	15816	16 %	0.251	1.48	3.10
B_4		4701	3922	16298		0.241	1.48	
B_5		5738	5050	19988		0.253	1.47	
B_6		1804	2370	6351		0.373	1.45	
B_7	1921-45	4434	4036	15963	14 %	0.253	1.47	3.09
B_8		4569	3968	15838		0.251	1.47	

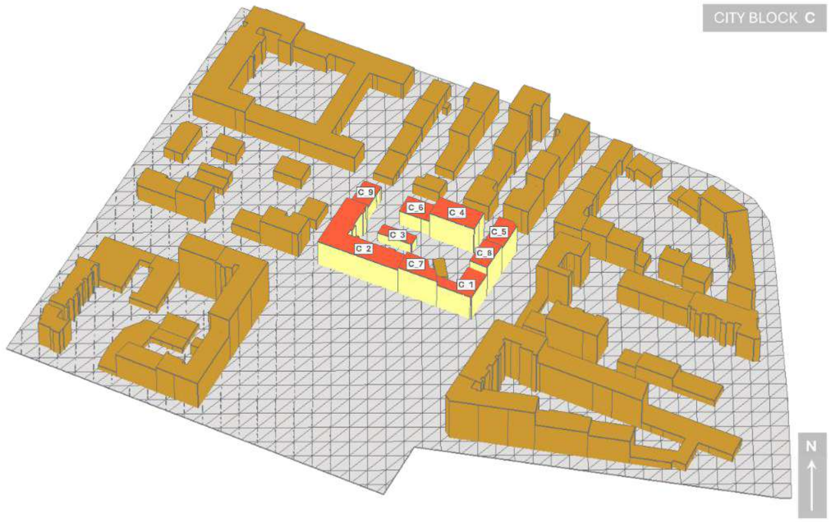


Figure 64. *Urban block C (Turin) visualised in CitySim Pro*

Table 33. *Geometrical and mean thermal characteristics for opaque and transparent components for urban block C (Turin)*

Bldg. ID	**Constr. period**	A_{fl} m^2	A_{env} m^2	V m^3	*WWR* –	$A_{env}\cdot V^{-1}$ m^{-1}	U_{op} W/(m^2·K)	U_w W/(m^2·K)
C_1		1440	1819	5328		0.341	1.45	
C_2	1901-20	5310	4854	17081	16 %	0.284	1.47	3.10
C_3		278	682	1027		0.664	1.48	
C_4		2103	2197	7149		0.307	1.30	
C_5		1373	1584	5013		0.316	1.28	
C_6	1946-60	678	1043	2645	13 %	0.394	1.36	3.12
C_7		1355	1521	4854		0.313	1.28	
C_8		1312	1327	4593		0.289	1.27	
C_9	1961-75	681	1063	2589	10 %	0.411	1.25	3.16

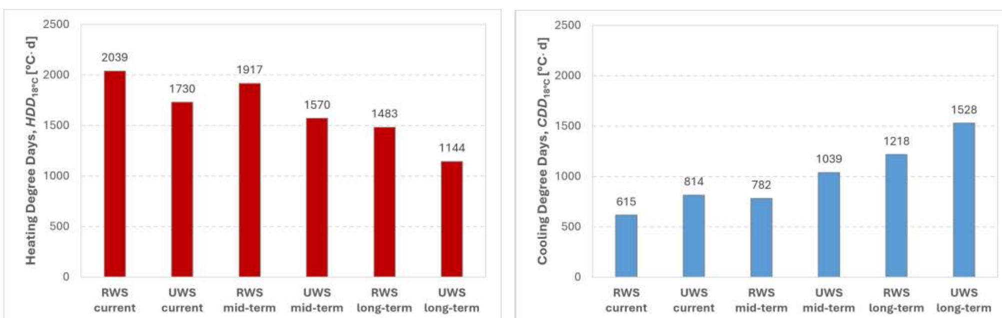
The use category and construction period served as the main criteria for assigning the corresponding building archetypes. Specifically, the thermal properties of opaque and transparent envelope components were derived from archetype schemas developed within the PRIN2020 URBEM and the TIMEPAC projects. Moreover, the internal heat gain schedules and intensities—including occupants, appliances, and lighting—were taken from the Italian National Annex of UNI EN 16798-1 (UNI, 2025). Due to the absence of tailored occupancy data, a deterministic schedule-based approach representative of residential apartment blocks was therefore adopted.

5.3.2.2 Rural and urban weather data under current, mid-, and long-term scenarios

The rural climatic data were obtained from the Turin Bauducchi weather station, located in an open-field area (latitude 44.96° N, longitude 7.70° E, altitude 226 m a.s.l.). This station was selected as the nearest rural reference site offering ten years of continuous measurements (1994-2003). The dataset, provided by ARPA Piemonte (Regional Environmental Protection Agency) (ARPA, 2025), was used to generate the TMY representative of current climate conditions.

Subsequently, mid- and long-term future weather files were produced following the methodology of the IEA-EBC Annex 80 Weather Data Task Group (Machard et al., 2024). The creation of current, mid-, and long-term urban weather files was carried out using the Urban Weather Generator (UWG) (Bueno et al., 2013; Nakano et al., 2015), based on 3D models of the investigated urban blocks. These models incorporated both geometric and physical parameters representative of the urban morphology and the corresponding building archetypes.

Figure 65 presents the Heating and Cooling Degree Days—calculated for a base temperature of 18 °C (HDD_{18} and CDD_{18}, respectively)—in accordance with UNI 10349-3 (UNI, 2016), for the UWS under short-, mid-, and long-term scenarios for the urban block A. The results show a 15-23 % decrease in *HDD*s and a 25-33 % increase in *CDD*s for urban areas compared to rural ones, highlighting a stronger and more persistent urban effect during the summer period.



(**a**) (**b**)

Figure 65. *HDDs (**a**) and CDDs (**b**) based on RWS and UWS data for the current, mid-, and long-term periods in urban block A (Piro, Ballarini, et al., 2026)*

Figure 66 illustrates the Ambient Warmness Degree, calculated with a reference temperature of 18 °C (AWD_{18}) under different climate conditions for urban block A, as defined by Equations (8). The AWD_{18} was computed for the summer period (June-August).

Under current conditions, AWD_{18} reaches 5.4 °C at the rural site and 7.0 °C in the urban area, confirming the presence of a warmer urban microclimate. In the mid-term scenario, both stations show a further increase, reaching 6.5 °C (RWS) and 8.4 °C (UWS). The long-term scenario indicates a marked intensification of warm conditions, with AWD_{18} values of 9.3 °C for the RWS and 11.2 °C for the UWS. The latter suggests that, on average, the outdoor air temperature exceeds the 18 °C threshold by more than 11 °C during the analysed warm period. Slight differences on the order of tenths of a degree were observed for AWD_{18} in urban blocks B and C.

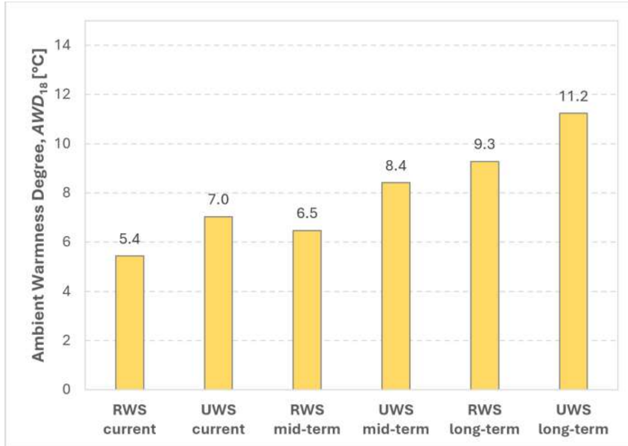
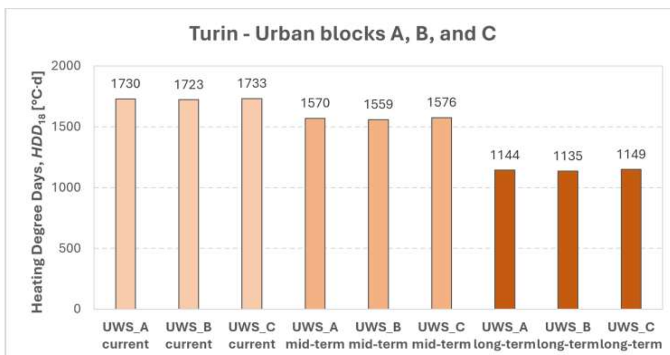
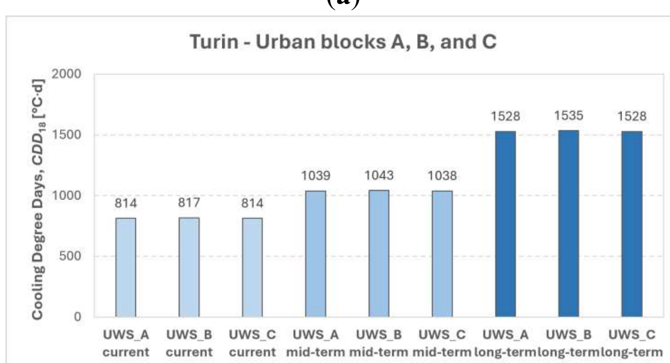


Figure 66. *AWD_{18} based on RWS and UWS data for the current, mid-, and long-term periods in urban block A (Piro, Ballarini, et al., 2026)*

Additionally, Figure 67 presents the aggregated values of HDD_{18} and CDD_{18} and Figure 68 illustrates the *UHII*, calculated according to Equations (9), for the three city blocks, based on UWS data. Specifically, monthly *UHII* values are based on UWS data measured at long-term period.



(**a**)



(**b**)

Figure 67. *HDD_{18} (**a**) and CDD_{18} (**b**) for urban blocks in Turin*

Slight variations can be observed for urban block B compared to A and C. In both cases, these minor deviations are attributable to differences in urban morphology, since the street-level anthropogenic sensible heat was kept constant across simulations. Indeed, the most urbanised area, block B, characterised by the highest *SC* (Table 30), exhibits the largest temperature differences between UWS and RWS data. This results in a stronger UHI effect, leading to a reduction in *HDD*s and an increase in both *CDD*s and monthly *UHII* values.

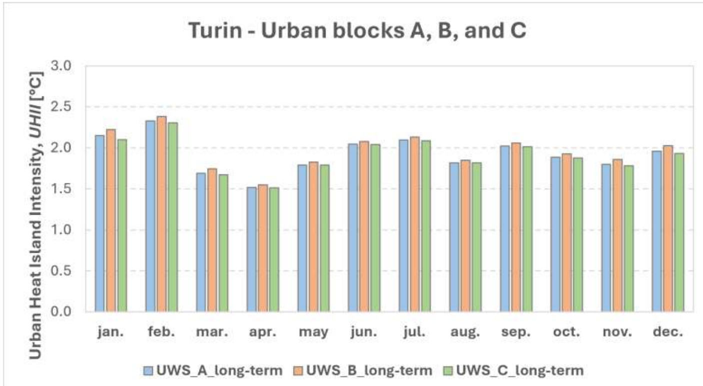


Figure 68. *Long-term scenario UHII at UWS for the urban blocks*

5.3.3 Results and discussion

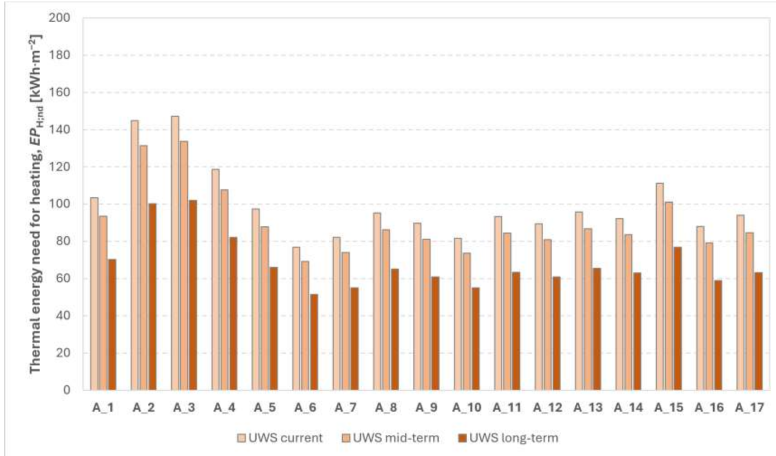
5.3.3.1 Disaggregated for urban block

The results in terms of thermal energy need and peak loads for space heating and cooling, disaggregated for the buildings within the representative urban blocks A, B, and C in the municipality of Turin are presented. Specifically, the trends of $EP_{\mathrm{H/C;nd}}$ and $\phi_{\mathrm{H/C;ld}}$ under short-, medium- and long-term climate scenarios, based on UWS data, are illustrated.

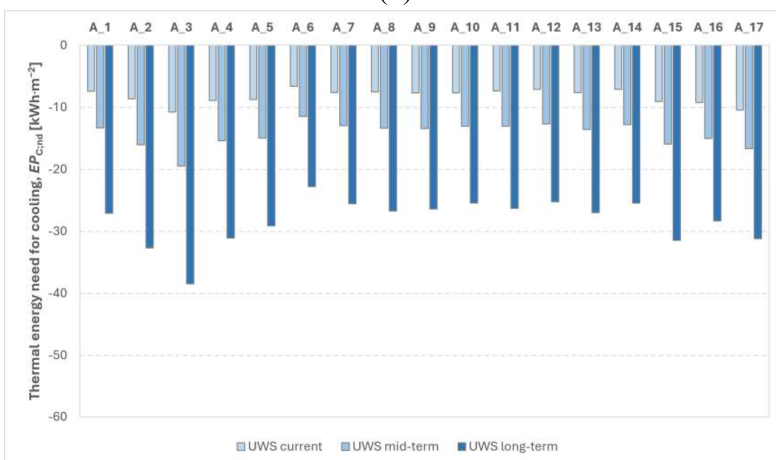
The fluctuations in building energy need ($EP_{\mathrm{H/C;nd}}$), which are strongly influenced by variations in external air temperature, are more pronounced than those in peak heating/cooling loads ($\phi_{\mathrm{H/C;ld}}$). The latter depend not only on outdoor air temperature but also on solar irradiance and the thermal inertia of the buildings.

Figure 69 and Figure 70 present the results for portfolio A in terms of $EP_{\mathrm{H/C;nd}}$ and $\phi_{\mathrm{H/C;ld}}$, respectively. Figure 71 and Figure 72 report the same indicators for the representative urban blocks B and C.

Overall, the results reveal a consistent pattern: climate change markedly reduces the thermal energy need for space heating while substantially increases cooling requirements, with effects that intensify toward the long-term horizon. The energy intensity at the city-block scale is influenced by multiple factors, including building geometry, construction-related thermal properties, and the morphological configuration of the urban fabric. Among these, geometry and morphology emerged as the dominant factors driving variations in thermal energy needs for both heating and cooling—and consequently in peak heating and cooling loads.

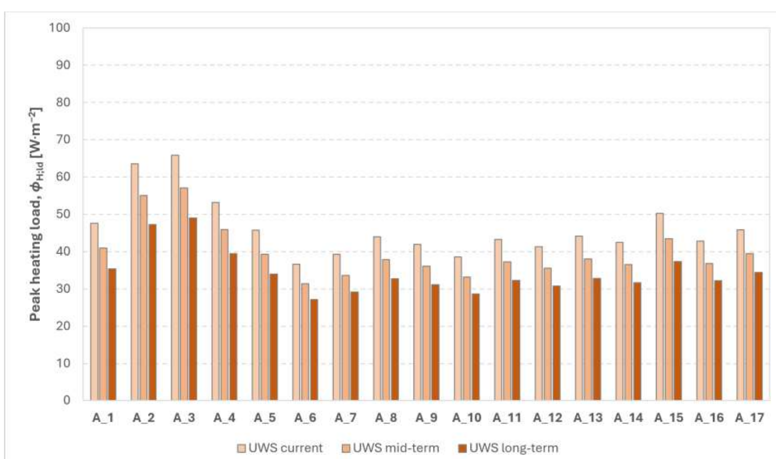


(**a**)

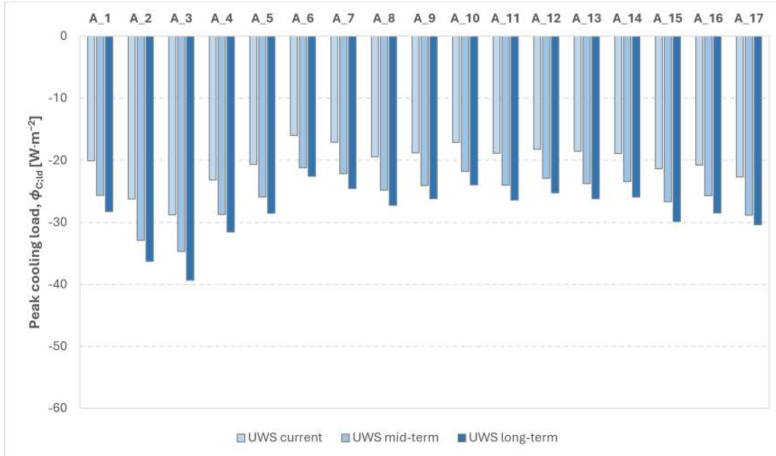


(**b**)

Figure 69. *Thermal energy need for space heating (**a**) and cooling (**b**) using UWS data for the short-, mid-, and long-term periods for urban block A (Piro, Ballarini, et al., 2026)*

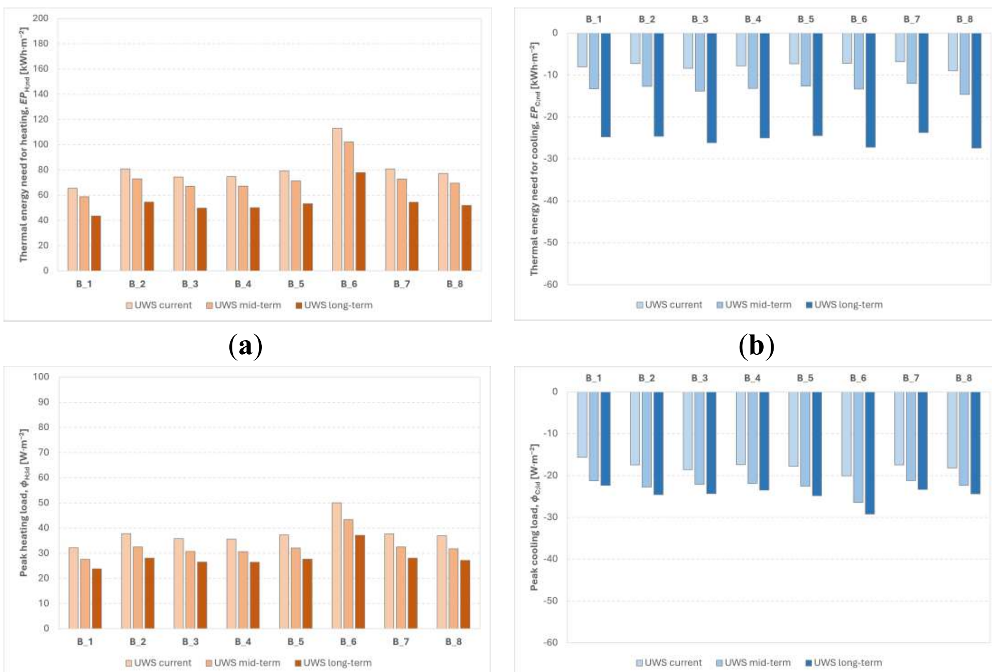


(**a**)



(**b**)

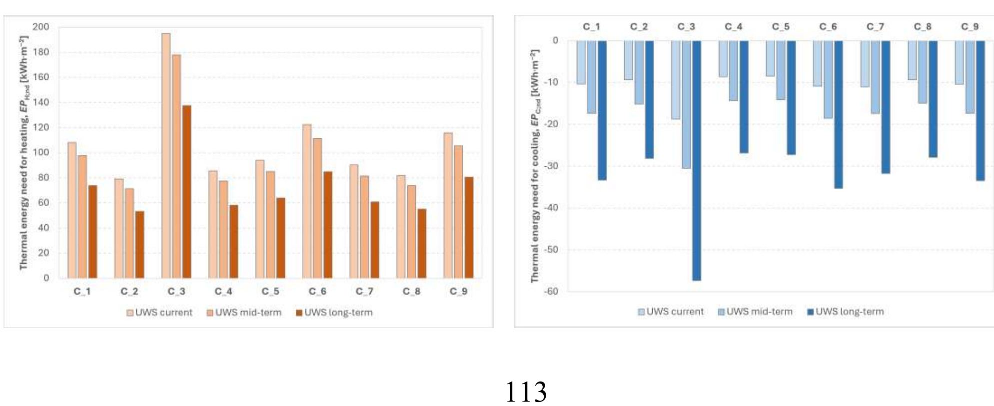
Figure 70. *Peak heating (**a**) and cooling (**b**) loads using UWS data for the short-, mid-, and long-term periods for urban block A (Piro, Ballarini, et al., 2026)*



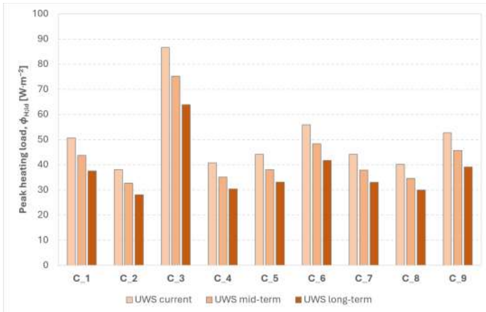
(**a**) (**b**)

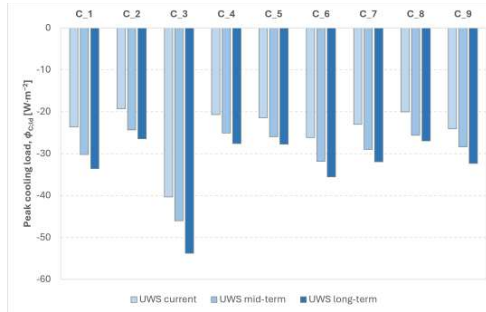
(**c**) (**d**)

Figure 71. *Thermal energy need for space heating (**a**) and cooling (**b**), and peak heating (**c**) and cooling (**d**) loads using UWS data for the short-, mid-, and long-term periods for urban block B*



(**a**) (**b**)



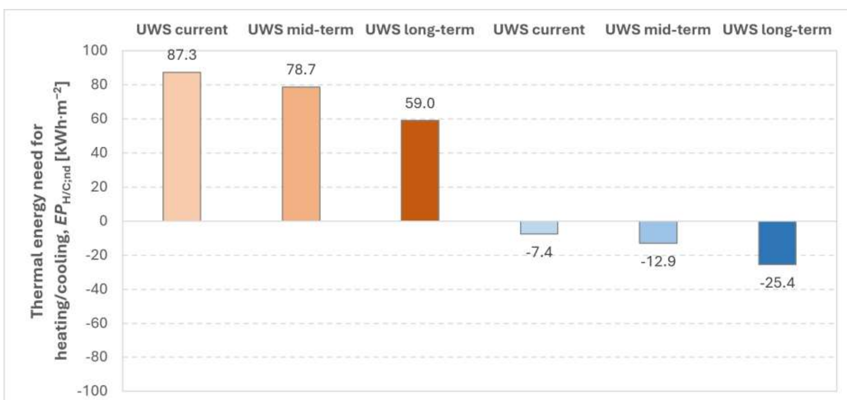


(**c**) (**d**)

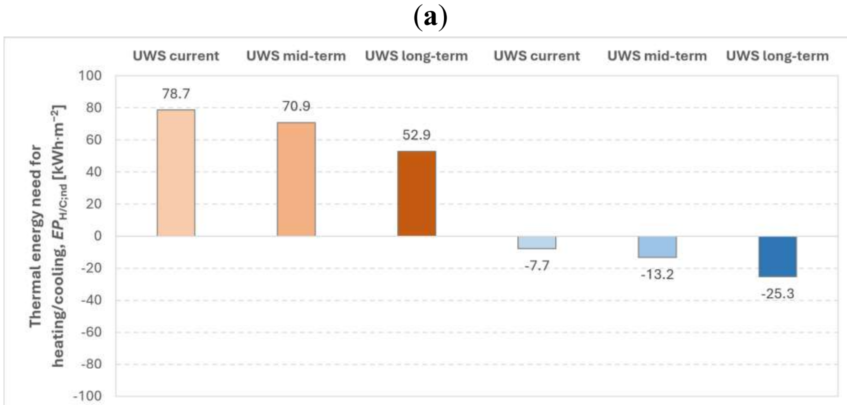
Figure 72. *Thermal energy need for space heating (**a**) and cooling (**b**), and peak heating (**c**) and cooling (**d**) loads using UWS data for the short-, mid-, and long-term periods for urban block C*

5.3.3.2 Aggregated for representative city blocks

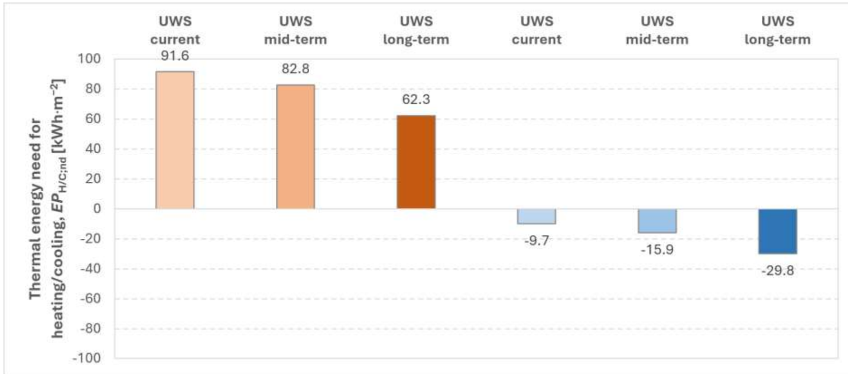
The aggregated results for the three representative urban blocks of the municipality of Turin, emphasising the differences in energy performance within the same urban context, are presented. Specifically, Figure 73 illustrates the thermal energy needs for space heating and cooling—calculated as the net floor area-weighted average of the buildings within urban blocks A, B, and C, respectively—using UWS climate data for three distinct climate periods.



(**a**)



(**b**)



(**c**)

Figure 73. *$EP_{H/C;nd}$ for urban blocks A (**a**), B (**b**), and C (**c**) using UWS data for the current, mid-, and long-term periods (Piro, Ballarini, et al., 2026)*

Several key factors influence the energy intensity at the city-block scale. First, the geometry of the buildings within the UBEM domain determines the extent of heat transfer through both opaque and transparent envelope components. Second, the non-geometric characteristics defined in the building archetype schema—particularly those related to construction period—affect the building typology and the thermal properties of envelope elements. Finally, the morphological configuration of each block, including shading patterns and façade orientation, influences the contribution of solar heat gains.

Overall, the geometric and morphological attributes of the urban blocks emerge as the dominant factors governing variations in thermal energy needs for both heating and cooling. The compactness ratio ($A_{env}{\cdot}V^{-1}$) has a direct impact on the thermal energy demand for space heating. As shown in Table 30, urban block C exhibits the highest $A_{env}{\cdot}V^{-1}$, followed by blocks A and B. A similar trend is observed for $EP_{H;nd}$ in the long-term scenario, with values of 62.3 kWh/m² for block C, 59.0 kWh/m² for block A, and 52.9 kWh/m² for block B.

Conversely, the influence on space cooling is less straightforward. A higher compactness ratio implies a larger surface area exposed to solar radiation, which can increase cooling needs. In the long-term scenario, urban block C records the highest $EP_{C;nd}$ at – 29.8 kWh/m², while blocks A and B show comparable values of about – 25 kWh/m².

Table 34 presents the overall variation in urban block $EP_{H/C;nd}$ relative to the current scenario for case studies A, B, and C. In the mid-term, the reduction in heating need ranges between – 9.8 % and – 10.0 %, while the cooling need rises by 63.4 % to 73.3 %. The long-term scenario amplifies these trends, with the heating need decreasing by more than 30 % and the cooling need more than doubling (+ 206.7 % to + 242.2 %). This substantial rise in cooling energy need highlights the growing impact of warmer outdoor conditions and the increasing importance of summer thermal comfort in future climate contexts.

Table 34 *Overall variation in urban block $EP_{H/C;nd}$ relative to the current scenario for the assessed case studies A, B, and C*

Scenarios	$\Delta EP_{H;nd}$ kWh·m^{-2}	$\Delta EP_{C;nd}$ kWh·m^{-2}	$\Delta EP_{H;nd}$ –	$\Delta EP_{C;nd}$ –
Urban block A				
Mid-term - current	– 8.6	+ 5.4	– 9.8 %	+ 73.3 %
Long-term - current	– 28.3	+ 18.0	– 32.4 %	+ 242.2 %
Urban block B				
Mid-term - current	– 7.8	+ 5.4	– 10.0 %	+ 70.1 %
Long-term - current	– 25.8	+ 17.5	– 32.7 %	+ 226.3 %
Urban block C				
Mid-term - current	– 8.7	+ 6.2	– 9.6 %	+ 63.4 %
Long-term - current	– 29.3	+ 20.1	– 32.0 %	+ 206.7 %

5.4 Remarks

In this chapter, a methodology for creating both representative buildings and urban blocks was developed and presented. The first approach, which derives from the output of the quality-checking procedure, was applied to the territories of the Piedmont and Aosta Valley regions. The second procedure focuses on selecting representative urban blocks within the typical Italian climate zones of Turin, Bari, and Rome.

Reference buildings—classified as prototypes and archetypes—play a crucial role in both *top-down* or *bottom-up* urban planning approaches aimed at assessing the energy and environmental performance of the building stock. In both cases, reference buildings contribute to increasing the digitalisation and knowledge base of various building types. Specifically, prototypes serve as a useful tool for developing NBRPs that support large-scale interventions. Conversely, datasets describing non-geometrical properties constitute the foundation for UBEM archetype-based approaches, which can be employed for multiple purposes. These virtual representative entities are essential resources for a range of stakeholders, including public administrations, energy agencies, and research institutions.

Finally, urban blocks identified as both geometrically representative and vulnerable to urban heat island effects were analysed to assess the current and future energy footprints of the building stock. This approach highlights the importance of selecting representative urban blocks while considering building morphology within urbanised areas—an aspect further emphasised in the results part.

Chapter 6

Enhancing the credibility of large-scale energy models

6.1 Overview

The work presented in this chapter was carried out at the Energy Informatics Group of the Idiap Research Institute in Martigny (Switzerland), under the supervision of Dr. Jérôme H. Kämpf.

Part of the work described in this chapter has been previously published in an international peer-reviewed journal (Piro, Kämpf, et al., 2026) and conference proceeding (Piro, Kämpf, et al., 2025). In particular, in Piro, Kämpf, et al. (2025), the Ph.D. activity focused on defining a methodology and its application for identifying the most representative subset of values within the probabilistic building archetype schema. For the journal publication (Piro, Kämpf, et al., 2026), the specific contributions of the author are detailed in the corresponding CRediT authorship statements.

This chapter addresses a key challenge in Urban Building Energy Modelling: the verification of model outputs. Depending on the purpose of the analysis, the availability of data, and computational constraints, either validation or calibration procedures can be applied to enhance the reliability of large-scale energy assessments. Given the heterogeneous informatisation context, this chapter proposes a validation and calibration framework at the UBEM scale.

This chapter is organised as follows: Section 6.2 introduces a validation procedure aimed at evaluating the reliability of the information embedded within the building archetypes; Section 6.3 presents a calibration methodology based on an evolutionary optimisation algorithm and applied at different spatial granularities; finally, Section 6.4 provides general remarks and conclusions.

6.2 Validation

The credibility of a NBRP relies on the accuracy of large-scale energy models, which in turn depends on minimising the gap between simulated and actual energy performance of the building stock. BAs, often derived from uncertain and incomplete EPCs, play a key role in reducing this discrepancy. Therefore, a validation framework is required to assess their actual representativeness.

This research introduces two novel metrics—the Validity of Representativeness Hours (*VRH*) and the Percentage of Representativeness Hours (*PRH*)—to evaluate, at each simulation step, how effectively a probabilistic BA schema captures real operational building conditions. The methodology was implemented in CitySim and applied to a residential district in the municipality of Turin, comprising apartment blocks classified into three construction periods. The results demonstrate that BA inputs can accurately reproduce the real energy performance of the building stock. By iteratively varying sensitive input parameters within their statistical ranges, the corresponding *PRH* values were compared to assess the degree of representativeness.

The study provides valuable insights for two main audiences. For UBEM developers, it offers guidance on optimising BA data by identifying the most representative input values. For policymakers, it enhances understanding of the actual energy performance of the building stock, thereby supporting more informed and effective decision-making in urban energy planning.

6.2.1 Methodology

Validation of an energy model is a common approach for assessing the reliability of estimated energy outcomes by comparing simulated results with observed benchmark data. This approach is typically adopted when only aggregated consumption data are available or when the computational cost of a full calibration procedure is prohibitive. According to the literature, the statistical indicators most commonly used for this purpose include the root-mean-square error (*RMSE*), coefficient of variation of *RMSE* (*CV(RMSE)*), coefficient of determination (R^2), mean bias error (*MBE*), and mean absolute percentage error (*MAPE*).

The methodology proposed here introduces two interrelated KPIs—the Validity of Representativeness Hours (*VRH*) and the Percentage of Representativeness Hours (*PRH*)—to evaluate the extent to which the simulated energy demand matches the measured energy consumption within a predefined tolerance range, based on a given set of inputs derived from the probabilistic building archetype schema. These KPIs are inspired by the concept of percentage hours of discomfort used in thermal comfort analysis, where, at each assessed timestep, it is verified whether the estimated operative temperature falls within the acceptable comfort range, and the compliant hours are subsequently aggregated over the calculation period.

BAs are deterministic or probabilistic datasets of non-geometrical properties designed to characterise building energy performance by capturing various technological features and usage patterns. The proposed methodology aims to assess the effectiveness of BA inputs in accurately modelling the actual energy behaviour of the building stock, based on their degree of representativeness.

The approach consists of the following steps:

1. identifying and extracting the relevant input parameters from the BA schema,
2. defining the probability distribution and variation range for each selected variable,
3. using a UBEM tool to perform scenario analyses with varying input parameters, and
4. computing, for each simulation, the *VRH* and *PRH* metrics.

For different input sets, the *VRH*, defined in Equation (10), determines whether the simulated delivered thermal energy to the generation sub-system (gen;in;sim), calculated using a UBEM tool, at the *i*-th timestep (sub-hourly, hourly, or daily), falls within the predefined lower (lw) and upper (up) consumption limits (lim) for the assessed building (*bldg*). Depending on the availability and reliability of measured data, this formulation can be applied to various energy services—including space heating (H), space cooling (C), and/or domestic hot water (W)—either individually or in combination (e.g., total delivered energy as the sum of space heating and DHW demands). The upper and lower consumption limits may be statistically derived, incorporating a tolerance margin based on measurement uncertainty, or defined as the maximum and minimum observed values in the building consumption dataset.

$$VRH_{\mathrm{H/C/W};bldg;i} = \begin{cases} 1, & Q_{\mathrm{H/C/W;gen;in;lw;lim};i} \le Q_{\mathrm{H/C/W;gen;in;sim};bldg;i} \le Q_{\mathrm{H/C/W;gen;in;up;lim};i} \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

Subsequently, the *PRH* quantifies the ability of the simulated BA inputs to capture the building's operational representativeness. As expressed in Equation (11), *PRH*—calculated as the ratio of *VRH* to the total number of timesteps (N) over the simulation period—provides an aggregated indicator of how effectively the tested input combination, drawn from the probabilistic BA schema, reproduces real operational conditions.

$$PRH_{\mathrm{H/C/W};bldg} = \frac{\sum_{i}^{N} VRH_{\mathrm{H/C/W};bldg;i}}{N} \cdot 100 \tag{11}$$

This approach provides a high degree of flexibility, allowing its application across diverse contexts and datasets with varying spatial and temporal resolutions. Spatial flexibility enables the method to be applied to energy consumption data at both disaggregated (building-level) and aggregated (block- or district-level) scales. Likewise, temporal flexibility allows the simulated building energy consumption to be aggregated on a sub-hourly, hourly, or daily basis, depending on the resolution of the available measurements.

6.2.2 Application

The proposed methodology was applied to a case study in the municipality of Turin, consisting primarily of apartment blocks modelled in CitySim. To evaluate the reliability of the archetype on a building-by-building basis, the thermal properties of opaque and transparent envelope components, along with the space heating set-point temperatures, were varied in five incremental steps within their respective statistical ranges. The simulated space heating energy use was then compared against real upper and lower limits, derived from anonymised hourly district heating consumption data collected from over 100 substations serving apartment blocks in Turin. For each simulation and building, the *PRH* values were calculated over the space heating period defined by Italian legislation.

6.2.2.1 Case study

The case study focuses on a real residential block located in Turin (*HDD*: 2162 °C·d), comprising 357 buildings categorised into three construction periods: 1946-60 (BA_1), 1961-70 (BA_2), and 1971-80 (BA_3). Building age and height data were obtained from the open database of the Turin Geoportal (Geoportal of the Metropolitan city of Turin, 2025). The LoD adopted for the analysed urban block corresponds to LoD1, where buildings are represented as simplified shoebox geometries. Figure 74 shows the segmentation of buildings according to construction period and the corresponding 3D urban scene modelled in CitySim. Non-residential buildings and residential buildings that are outside the selected construction periods were included as shading objects. This area of Turin is connected to the district heating network.

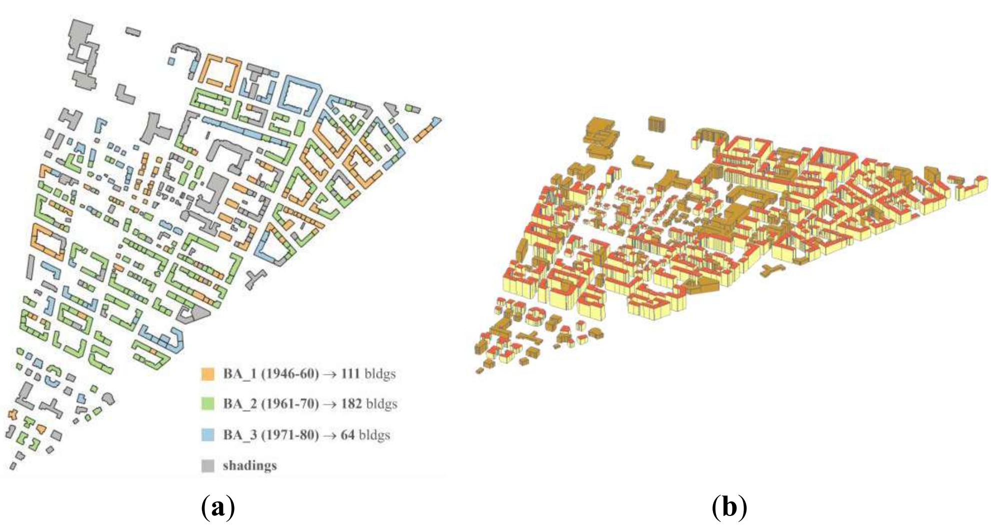


Figure 74. *Case study segmented per BAs (**a**); CitySim model (**b**) (Piro, Kämpf, et al., 2025)*

The procedure entails varying four parameters to evaluate how different input combinations affect the calculation of *PRH*. The BA data were obtained from the

Italian research project URBEM. Table 35 reports the thermal transmittance values of walls (U_{wl}), roofs ($U_{\mathrm{fl;up}}$), and windows (U_{w}), as well as the heating set-point temperature ($\theta_{\mathrm{H;set}}$) for each BA group. Specifically, the thermal characteristics of the building envelope were sourced from the URBEM BA schema, where statistical ranges were subdivided into five discrete steps, while $\theta_{\mathrm{H;set}}$ values were derived from the EN 16798-1 standard (CEN, 2019).

Table 35. *BA characteristics per construction period*

	BA_1 (1946-60)					BA_2 (1961-70)					BA_3 (1971-80)				
U_{wl} W/(m²·K)	1.27	1.16	1.05	0.93	0.82	1.23	1.15	1.06	0.98	0.89	1.24	1.16	1.09	1.01	0.93
$U_{\mathrm{fl;up}}$ W/(m²·K)	1.45	1.28	1.10	0.92	0.75	1.42	1.29	1.16	1.04	0.91	1.42	1.25	1.09	0.93	0.77
U_{w} W/(m²·K)	4.51	3.97	3.42	2.88	2.33	4.58	4.02	3.45	2.89	2.32	4.49	3.94	3.39	2.83	2.28
$\theta_{\mathrm{H;set}}$ °C	21.0	20.5	20.0	19.0	18.0	21.0	20.5	20.0	19.0	18.0	21.0	20.5	20.0	19.0	18.0

6.2.2.2 Configuration steps

This section illustrates the implementation of the proposed methodology. For each i-th hour, the lower and upper bounds—computed using Equation (12)—were statistically derived from anonymised 2021 measurements collected at 105 substations serving apartment blocks in Turin. A 10 % tolerance, following ASHRAE Guideline-14 (ASHRAE, 2023), was applied when defining the measurement limits based on the first and third quartiles (Q_1 and Q_3) and the interquartile range (IQR). For ensuring an apple-to-apple comparison, the building energy consumption were normalised over the conditioned volumes. For privacy reason, the district heating provider obscured the addresses of the connected buildings, determining limitations in identifying building envelope energy efficiency. Figure 75 presents the hourly aggregation of space heating energy demand measured at 105 district heating substations for the first week of January 2021. Red bars represent upper and lower whiskers including the tolerance 10 % of margin. This procedure was applied to the entire activation period of the heating systems: from mid-October to mid-April (Figure 76 and Figure 77). In Figure 76 are evident some missing days in measurements due to technical issues. For these five days hourly recordings of the closest day were assumed. The space heating temperature set-point was aligned with the operation of the district heating network substations: free-floating conditions during night-time and fixed at a constant during the remaining hours.

$$\begin{cases} Q_{\mathrm{H;gen;in;lw;lim};i} = \max\left[\min\left(Q_{\mathrm{H;gen;in};i}\right);\ Q_{\mathrm{H;gen;in};Q_1;i} - (1.5+0.1)\cdot Q_{\mathrm{H;gen;in};IQR;i}\right] \\ Q_{\mathrm{H;gen;in;up;lim};i} = \min\left[\max\left(Q_{\mathrm{H;gen;in};i}\right);\ Q_{\mathrm{H;gen;in};Q_3;i} + (1.5+0.1)\cdot Q_{\mathrm{H;gen;in};IQR;i}\right] \end{cases} \tag{12}$$

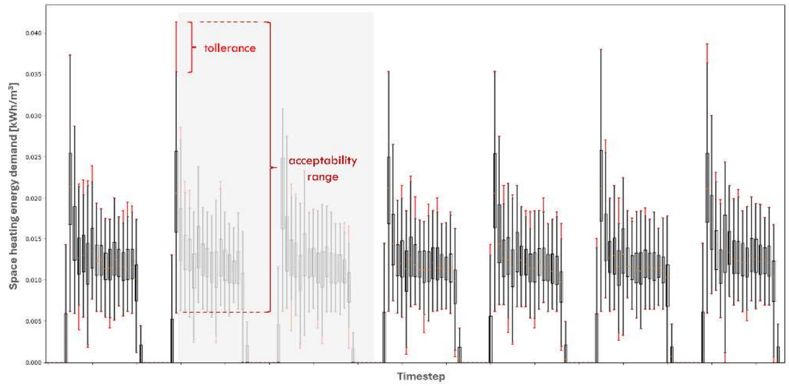


Figure 75. *Hourly aggregation of space heating energy demand for the first week of January*

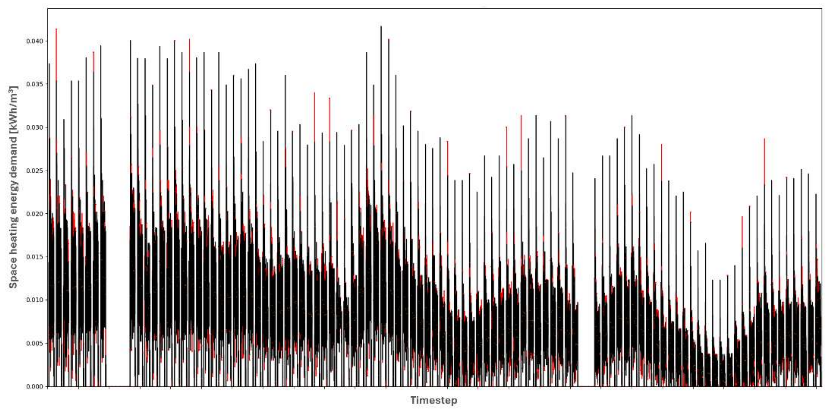


Figure 76. *Hourly aggregation of space heating energy demand for the first part of the year (from 1st of January to 15th of April)*

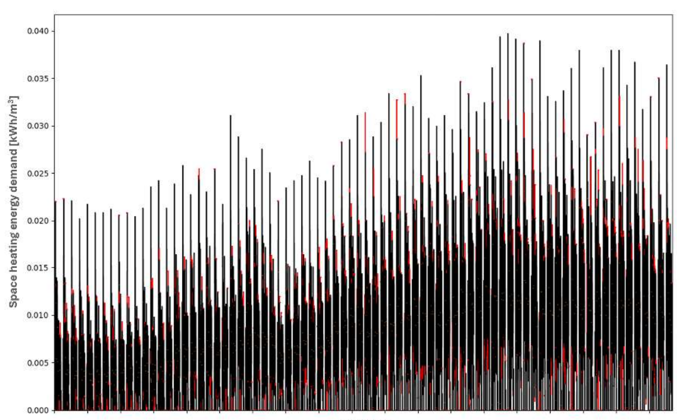


Figure 77. *Hourly aggregation of space heating energy demand for the second part of the year (from 15th of October to 31st of December)*

In accordance with the case study, Equations (13) and (14) were adapted to determine *VRH* and *PRH*, considering only the space heating energy service. This adjustment was required to match the scope of the available data. The summation for *PRH* calculations was extended over the heating period established.

$$VRH_{\mathrm{H};bldg;i} = \begin{cases} 1, & Q_{\mathrm{H;gen;in;lw;lim};i} \leq Q_{\mathrm{H;gen;in;sim};bldg;i} \leq Q_{\mathrm{H;gen;in;up;lim};i} \\ 0, & \text{otherwise} \end{cases} \tag{13}$$

$$PRH_{\mathrm{H};bldg} = \frac{\sum_{i}^{N} VRH_{\mathrm{H};bldg;i}}{N} \cdot 100 = \frac{\sum_{j=1}^{M=2520} VRH_{\mathrm{H};bldg;j} + \sum_{k=6889}^{L=8760} VRH_{\mathrm{H};bldg;k}}{M + (L - k)} \cdot 100 \tag{14}$$

The algorithm used to test all possible input combinations in conjunction with CitySim is available online (Kämpf, 2024). A discrete uniform distribution was assigned to the four independent variables listed in Table 35. Accordingly, for each archetype group, $5^4 = 625$ simulations were performed, evaluating *PRH* at every simulation phase and for each analysed building. Within each simulation set, the target BA cluster was thermally simulated, while the remaining two archetype groups acted as shading elements.

6.2.3 Results and discussion

The results of the validation procedure are presented in the following sections both in disaggregated and aggregated forms, in each case considering the categorisation into building archetypes.

6.2.3.1 Maximisation of PRH

This section presents the findings related to the maximisation of *PRH*, categorised by building archetype. Within each archetype group, the results are reported in a disaggregated manner for individual buildings.

Urban morphology and solar gains influence differently the space heating energy demands of buildings; consequently, buildings respond differently to variations in the input parameters. Figure 78, Figure 79, and Figure 80 show the maximum *PRH* values for BA_1, BA_2, and BA_3, respectively.

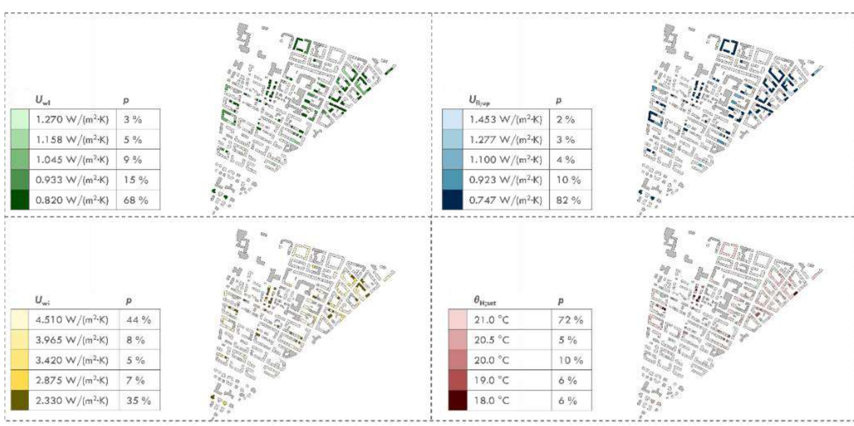


Figure 78. *Maximisation of PRH for BA_1 (Piro, Kämpf, et al., 2025)*

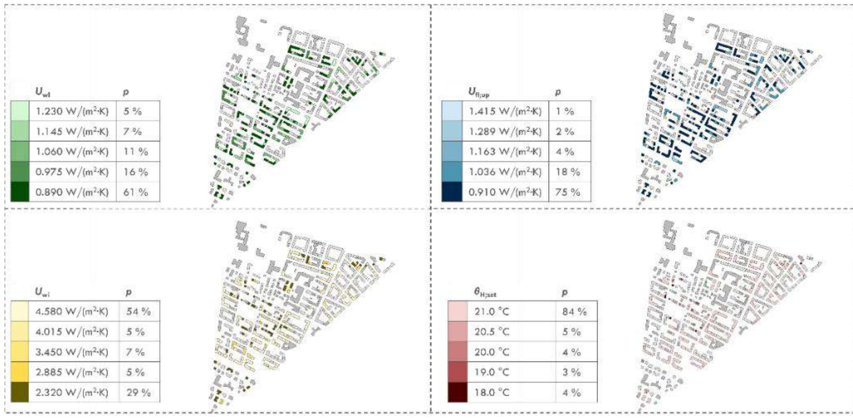


Figure 79. *Maximisation of PRH for BA_2*

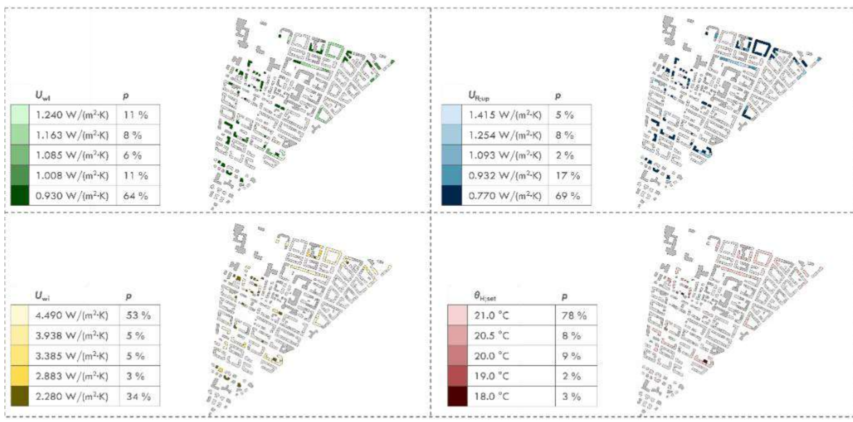


Figure 80. *Maximisation of PRH for BA_3*

Overall, an insulated opaque envelope with a heating set-point of 21 °C results in a better match between simulated and actual space heating energy demand. For BA_1, insulated walls (0.820 W/(m^{2}K)) and roof (0.747 W/(m^{2}K)) with a heating set-point of 21 °C have a higher probability of achieving elevated *PRH* values. A similar trend is observed for BA_2 and BA_3, where U_{wl} and $U_{\mathrm{fl;up}}$ are equal to 0.890 W/(m^{2}K) and 0.930 W/(m^{2}K), and 0.910 W/(m^{2}K) and 0.770 W/(m^{2}K), respectively. Greater uncertainty is associated with the thermal performance of transparent elements, for which older windows are preferred.

6.2.3.2 Representative combinations

Progressive acceptable thresholds of *PRH* were defined and tailored for the different building archetype groups. These thresholds were established to highlight the most representative combinations that simultaneously maximise the *PRH* values for all buildings within each cluster. In practical terms, this approach identifies the most representative set of archetype parameters that best match the real building stock's energy consumption. The analysis was based on the following assumptions:

- the baseline acceptability threshold limit was set at 60 % (corresponding to the number of heated hours defined for climatic zone E); thus, only combinations with *PRH* values greater than 60 % were considered acceptable,
- a dynamic threshold was introduced for each archetype group to reduce the number of possible combinations and to maximise the highest achievable *PRH*s, and
- a combination was considered valid only if the *PRH* value of every building exceeded the specified threshold. If even one building did not meet this condition, the combination was discarded.

Figure 81 shows the number of valid occurrences for each building within the corresponding archetype cluster. When adopting a 75 % threshold, the initial assumptions prove more appropriate for the third BA, as the number of valid occurrences reaches nearly 25 % of the tested solutions.

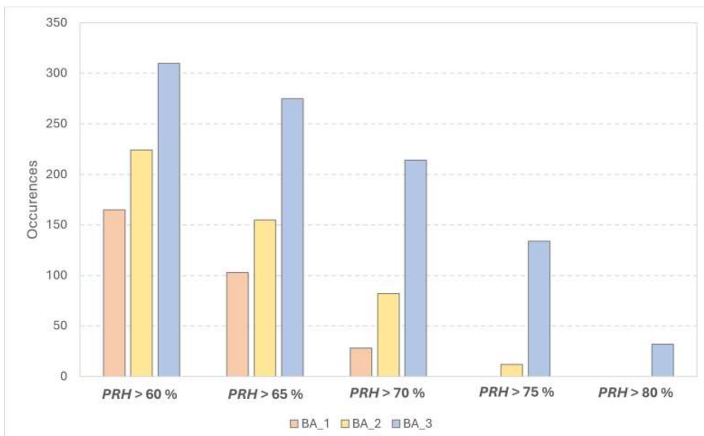
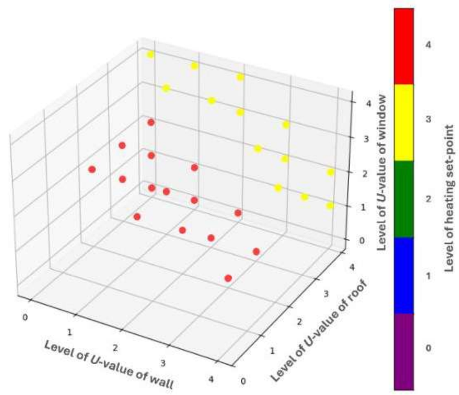


Figure 81. *No. of combinations with highest PRH values for BAs*

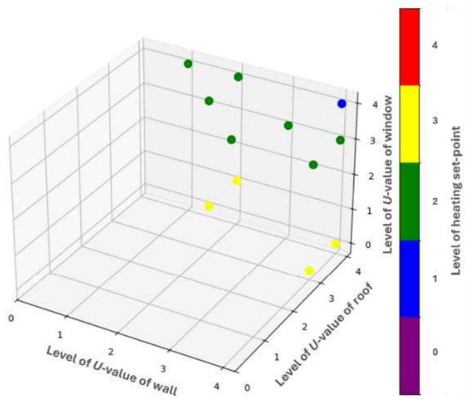
Figure 82, Figure 83, and Figure 84 present 4D scatter plots illustrating the best parameter combinations that yield the highest *PRH* values.

For buildings constructed between 1951 and 1960 (BA_1), a *PRH* exceeding 70 % can be achieved by combining roof insulation with a space heating set-point of either 18 °C or 19 °C (Figure 82). In the case of buildings built between 1961 and 1970 (BA_2), attaining a *PRH* greater than 75 % requires the combined application of insulation on both opaque and transparent envelope components, together with a heating set-point of 20 °C (Figure 83). Finally, for the most recent construction period (BA_3, 1971-80), achieving a *PRH* above 80 % can be accomplished by combining roof insulation with a space heating set-point of 20.5 °C (Figure 84).



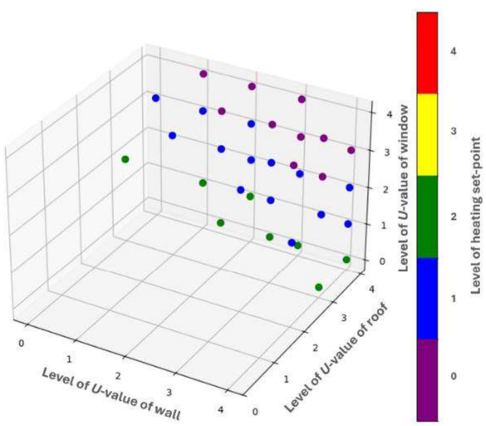
BA_1 (1946-60)	0	1	2	3	4
U_{wl} W/(m^2·K)	1.27	1.16	1.05	0.93	0.82
$U_{fl;up}$ W/(m^2·K)	1.45	1.28	1.10	0.92	0.75
U_w W/(m^2·K)	4.51	3.97	3.42	2.88	2.33
$\theta_{H;set}$ °C	21.0	20.5	20.0	19.0	18.0

Figure 82. *Best combinations for BA_1 with PRH > 70 %*



BA_2 (1961-70)	0	1	2	3	4
U_{wl} W/(m^2·K)	1.23	1.15	1.06	0.98	0.89
$U_{fl;up}$ W/(m^2·K)	1.42	1.29	1.16	1.04	0.91
U_w W/(m^2·K)	4.58	4.02	3.45	2.89	2.32
$\theta_{H;set}$ °C	21.0	20.5	20.0	19.0	18.0

Figure 83. *Best combinations for BA_2 with PRH > 75 %*



BA_3 (1971-80)	0	1	2	3	4
U_{wl} W/(m^2·K)	1.24	1.16	1.09	1.01	0.93
$U_{fl;up}$ W/(m^2·K)	1.42	1.25	1.09	0.93	0.77
U_w W/(m^2·K)	4.49	3.94	3.39	2.83	2.28
$\theta_{H;set}$ °C	21.0	20.5	20.0	19.0	18.0

Figure 84. *Best combinations for BA_3 with PRH > 80 %*

In the following tables (from Table 36 to Table 38) the best combinations, obtained increasing dynamically the *PRH* thresholds, for each assessed BA group, are reported.

Table 36. *Best combinations for BA_1 (PRH > 71.7 %)*

	U_{wl}		$U_{fl;up}$		U_w		$\theta_{H;set}$	
combination	level –	value W/(m^2·K)	level –	value W/(m^2·K)	level –	value W/(m^2·K)	level –	value °C
comb_1	4	0.82	4	0.75	2	3.42	3	19.0
comb_2	1	1.16	4	0.75	4	2.33	3	19.0
comb_3	1	1.16	4	0.75	0	4.51	4	18.0
comb_4	2	1.05	4	0.75	0	4.51	4	18.0
comb_5	0	1.27	4	0.75	1	3.97	4	18.0

Table 37. *Best combinations for BA_2 (PRH > 75.6 %)*

	U_{wl}		$U_{fl;up}$		U_w		$\theta_{H;set}$	
combination	level –	value W/(m^2·K)	level –	value W/(m^2·K)	level –	value W/(m^2·K)	level –	value °C
comb_1	4	0.89	4	0.91	4	2.32	1	20.5
comb_2	4	0.89	4	0.91	3	2.89	2	20.0
comb_3	2	1.06	4	0.91	4	2.32	2	20.0
comb_4	4	0.89	4	0.91	0	4.58	3	19.0
comb_5	2	1.06	4	0.91	1	4.02	3	19.0

Table 38. *Best combinations for BA_3 (PRH > 81.2 %)*

	U_{wl}		$U_{fl;up}$		U_w		$\theta_{H;set}$	
combination	level –	value W/(m^2·K)	level –	value W/(m^2·K)	level –	value W/(m^2·K)	level –	value °C
comb_1	4	0.93	4	0.77	3	2.83	0	21.0
comb_2	3	1.01	3	0.93	4	2.28	0	21.0
comb_3	1	1.16	4	0.77	4	2.28	0	21.0
comb_4	2	1.09	4	0.77	4	2.28	0	21.0
comb_5	4	0.93	3	0.93	2	3.39	1	20.5
comb_6	3	1.01	4	0.77	2	3.39	1	20.5

6.2.3.3 Discussion

The presented results depend on the purpose of the analysis—namely, whether the objective is to match the actual energy consumption of a single building or that of the entire building stock. Building energy balances are strongly influenced by morphology and solar gains (orientation and available surface for short-wave radiation); therefore, buildings respond differently to variations in the input parameters.

The results of the validation procedure are consistent in both approaches, except for BA_1, where the aggregated analysis suggests a lower heating set-point (18-19 °C). Overall, for the assessed building archetype groups, in order to evaluate the representativeness of the archetype data with respect to actual building-stock energy use, the energy consumption trends are better matched when assuming an insulated opaque envelope and increasing the heating set-point to 20-21 °C.

6.3 Calibration

This chapter introduces a novel calibration procedure that employs the CMA-ES/HDE evolutionary algorithm to minimise the *CV(RMSE)* objective function. The main innovation lies in preserving a fully scalable framework by performing calibration directly at the UBEM scale, rather than using the results to train inverse models. A further contribution of this work is the calibration of large-scale energy models at multiple spatial resolutions—specifically, both at the single-building level and at the urban-block level.

The methodology was applied to hundreds of buildings in the municipality of Monthey (Switzerland), in the Canton of Valais, using an archetype-based UBEM approach simulated in the CitySim environment. The analysed urban blocks incorporated varying degrees of spatial data granularity, including both aggregated and disaggregated energy consumption records. The study highlights the importance of homogeneous urban data as a prerequisite for adopting consistent modelling methodologies across different municipalities and national contexts.

The automatic calibration algorithm minimises *CV(RMSE)* by comparing simulated and measured daily energy demand for space heating and domestic hot water. Four independent parameters are adjusted during the process: air change rate, occupied area per person, daily DHW volume per person, and space-heating set-point temperature. Given their strong influence on model outputs, these variables are widely adopted in large-scale calibration studies (Sokol et al., 2017; C. K. Wang et al., 2020). The optimisation algorithm delivers calibration results for both individual buildings and urban blocks.

6.3.1 Methodology

6.3.1.1 General workflow

This section presents the theoretical framework of the automatic calibration procedure developed for UBEM. Increasing the accuracy of large-scale energy models is essential for reducing UBEM output uncertainty and, in turn, strengthening the credibility of NBRPs for their intended users. The applicability of the procedure depends on the purpose of the analysis, the computational effort required by the optimisation algorithm, and the availability of measurement data—either at the individual-building scale or for groups of buildings. The workflow, illustrated in Figure 85, consists of the following steps:

- Identifying and extracting the relevant input parameters from the BA schema.

- Defining variation ranges and probability distributions for each selected parameter.
- Formulating the optimisation objective function to be minimised or maximised and/or specifying stopping criteria such as a maximum iteration count or a convergence threshold.
- Running the UBEM tool to quantify output variability across multiple input combinations.
- Computing the performance metric required to evaluate the objective function at each simulation step.

The calibration relies on a fully CMA-ES/HDE (covariance matrix adaptation evolution strategy and hybrid differential evolution) evolutionary algorithm to adjust model parameters and refine estimates of space heating and domestic hot water energy demand at the building-stock level. This approach constitutes an innovative contribution to UBEM, enabling a closer alignment between simulated and measured energy use. Given the different spatial resolutions at which energy consumption data may be available, the procedure supports calibration at both the single-building and urban-block scales, aiming to minimise the coefficient of variation of the root-mean-square error, *CV(RMSE)*. To ensure a feasible runtime, a predefined maximum number of simulations is set according to the model's complexity at each scale.

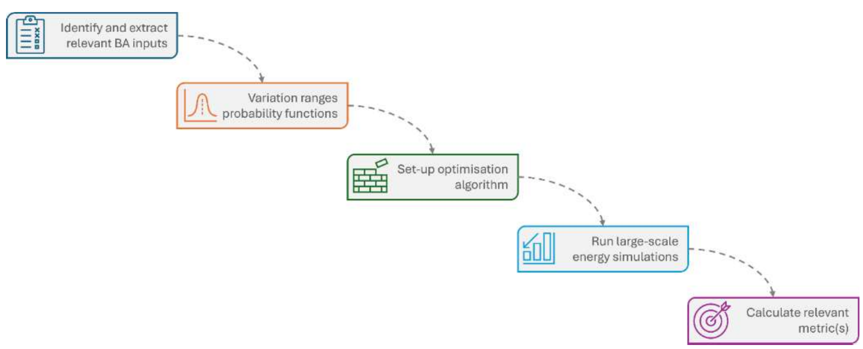


Figure 85. *Methodology for calibrating large-scale energy models (Piro, Kämpf, et al., 2026)*

CV(RMSE) expresses the agreement between simulated and observed data and is reported as a percentage. It is calculated using Equation (15), where y_i represents the daily measured energy demand from the district heating network, and $\hat{y}_i$ is the corresponding daily value predicted by CitySim. N denotes the number of available daily observations, and $\bar{y}$ is the mean of all measured values y_i. Since the temporal resolution of energy consumption data differs across buildings, *CV(RMSE)* is computed only for buildings for which valid daily measurements are available.

$$CV(RMSE) = \frac{\sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{N}}}{\bar{y}} \cdot 100 \tag{15}$$

For multi-building calibration, $\hat{y}_i$ represents the sum of the i-th daily delivered energy for space heating and domestic hot water across all buildings considered. This aggregated simulated value is then compared with the measured thermal energy demand of the urban block, y_i ASHRAE Guideline-14 (ASHRAE, 2023) provides reference *CV(RMSE)* thresholds for assessing whether a model can be considered as calibrated, depending on the temporal resolution of the available measurements: 30 % for hourly data and 15 % for monthly data. Since no established benchmark exists for daily observations, this study adopts a *CV(RMSE)* threshold of 25 %, positioned midway between the two ASHRAE reference values.

6.3.1.2 Calibration through CMA-ES/HDE algorithm

Evolutionary algorithms (EAs) are optimisation methods inspired by the principles of natural evolution and are widely used across engineering applications. They typically operate by initialising a population of candidate solutions, selecting a subset at random, and generating new individuals through crossover and mutation. The most promising individuals from both the parent and offspring groups are then retained to form the next generation. This iterative evolutionary cycle continues until a predefined stopping condition is reached, such as a maximum number of iterations or a convergence threshold.

The CMA-ES/HDE optimisation scheme belongs (Figure 86) to this family of EAs and integrates two complementary approaches: the covariance matrix adaptation evolution strategy (CMA-ES) and hybrid differential evolution (HDE) algorithms (Kämpf et al., 2010; Kämpf & Robinson, 2009). The hybrid approach exploits the global search ability of differential evolution alongside the rapid local convergence of CMA-ES, providing robust and efficient optimisation in complex, high-dimensional, and multimodal search spaces.

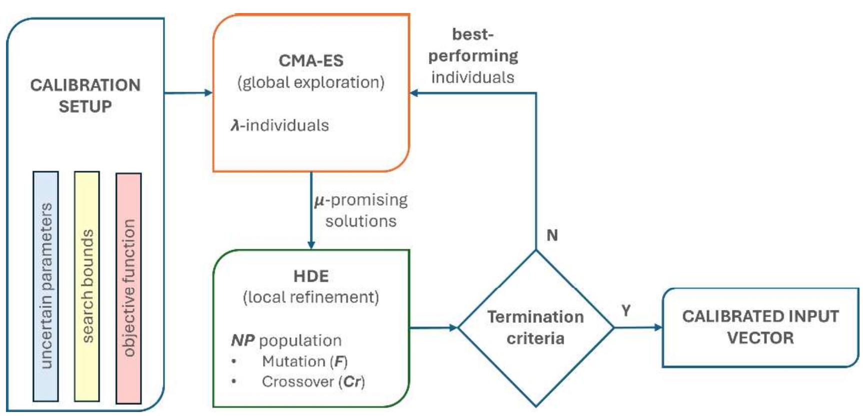


Figure 86. *CMA-ES/HDE calibration framework used for UBEM parameter optimisation (Piro, Kämpf, et al., 2026)*

Table 39 summarises the main parameters adopted for the hybrid CMA-ES/HDE algorithm. The hybrid framework consists of two interacting populations that evolve cooperatively. The CMA-ES component operates with a population size of λ individuals, defined as $\lambda = 4 + \lfloor 3\ln(n) \rfloor$, where *n* represents the number of continuous decision variables, and uses an initial step size of σ = 0.2. During each CMA-ES iteration, candidate solutions are evaluated and ranked, and the best-performing $\mu = \lambda/2$ individuals are selected to update the mean vector and covariance matrix of the sampling distribution.

The selected μ solutions are subsequently transferred to the HDE component, which employs a population size of $NP = 30$. Instead of relying on randomly generated candidates, the HDE population is initialised and evolved around the injected CMA-ES solutions through differential mutation and binomial crossover operators, using a mutation factor $F = 0.3$ and a crossover probability $Cr = 0.1$. In this way, the HDE stage performs a local refinement of the search space by producing improved variants of the solutions generated by CMA-ES.

Table 39. *Parameters used for the hybrid CMA-ES/HDE algorithm*

Algorithm	Parameters		
CMA-ES	$\lambda = 4 + \lfloor 3\ln(n) \rfloor$	$\mu = \lambda/2$	$\sigma = 0.2$
HDE	$NP = 30$	$F = 0.3$	$Cr = 0.1$

The procedure starts by initialising two separate populations: a group of μ-individuals handled by CMA-ES, and a group of ($NP - nt$) individuals governed by HDE. At each iteration, CMA-ES evaluates its population and selects the top-performing *nt* individuals, which are transferred to the HDE algorithm. HDE then randomly generates ($NP - nt$) new candidates, integrating the injected *nt* solutions into its mutation and recombination operators to enhance exploration of the search space. After evaluation, HDE returns its best individuals to the CMA-ES population, contributing to the update of its sampling distribution. This reciprocal exchange of individuals continues until the termination criterion is satisfied.

6.3.1.2 Statistical analysis

To quantitatively evaluate the influence of key parameters on simulated building energy demand, a quadratic multivariate regression model was developed based on the simulation results obtained during the iterations of the automatic calibration procedure.

Quadratic multivariate regression is a commonly adopted approach for representing non-linear physical relationships. Its general formulation is presented in Equation (16), where Y_i denotes the simulated building energy consumption, β_0 represents the intercept term, and β_i, β_{ii}, and β_{ij} correspond to the coefficients associated with the linear, quadratic, and interaction terms of the independent variables x_i, x_i^2, and $x_i x_j$, respectively.

$$Y_i = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \beta_{ij} x_i x_j \tag{16}$$

To evaluate the relative contribution of the linear, quadratic, and interaction effects, the weights w_{x_i}, $w_{x_i^2}$, and $w_{x_i x_j}$ were computed separately. These weights are defined as the ratio between the absolute value of each coefficient ($|\beta_i|$, $|\beta_{ii}|$, and $|\beta_{ij}|$) and the sum of the absolute values of all regression coefficients, excluding the intercept term.

6.3.2 Application

An automatic calibration methodology based on the CMA-ES/HDE evolutionary algorithm was applied to a Swiss case study located in the municipality of Monthey. The approach was implemented on approximately 200 residential buildings simulated in CitySim through an archetype-based UBEM framework. This modelling strategy reduces uncertainty in the representation of buildings' passive energy performance while enabling the analysis of energy-use patterns related to occupant behaviour and interactions. Archetypes were assigned according to the construction period, and non-geometrical attributes were mapped onto the actual building geometries.

The calibration procedure aimed to minimise the *CV(RMSE)* between simulated daily space-heating and DHW demands and the measured district heating consumption data for the year 2020. Calibration was conducted at both the single-building and multi-building scales, depending on the spatial resolution of the available measured data. Four independent variables were adjusted within predefined ranges: air change rate, occupied floor area per person, daily DHW consumption per person, and indoor space-heating set-point temperature.

6.3.2.1 Case study

The case study is located in the municipality of Monthey, in the Canton of Valais, Switzerland (430 m a.s.l.). Monthey has a population of roughly 16,000 inhabitants and contains over 2100 buildings. Figure 87 highlights the thermally simulated buildings in yellow with red roofs and indicates shading objects in brown.

Residential buildings make up 84 % of the analysed stock, consisting of 80 % multi-family houses and 4 % single-family houses. The remaining 16 % are non-residential buildings, including 8 % with unknown use, 3 % hospitals, 2 % research facilities, 1 % offices, and other minor categories.

Geometric information was extracted from the SwissBUILDINGS3D (SwissBUILDINGS3D, 2025) open geographic database. These data were converted into a GeoPackage format that distinguishes thermally simulated zones from surrounding shading objects. The resulting files were then imported into CitySim Pro to perform urban-scale energy simulations. The level of detail adopted for the urban block is LoD1, meaning buildings are represented using simplified shoebox geometries. Ground surfaces were omitted during calibration to reduce computational load.

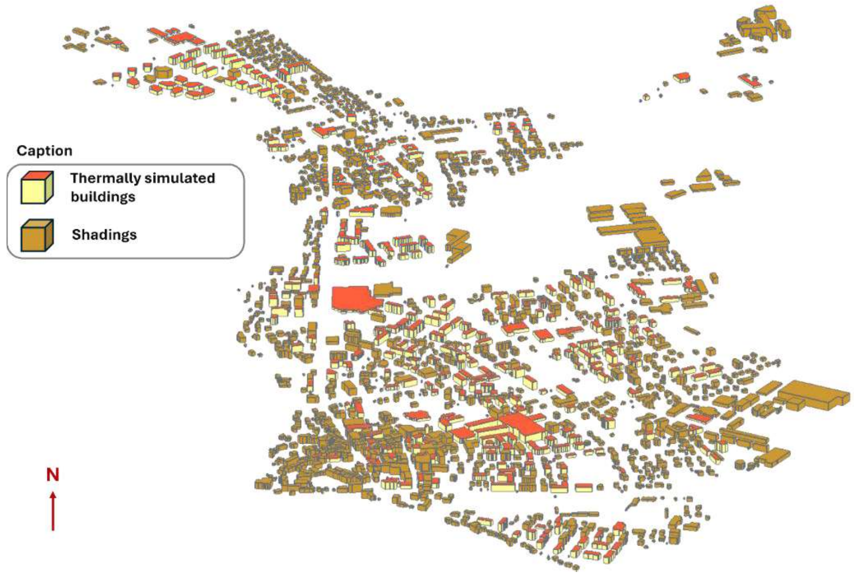


Figure 87. *Whole UBEM for the city of Monthey (Piro, Kämpf, et al., 2026)*

The thermal properties of both transparent and opaque envelope components were derived from the work of Perez (2014). These characteristics are classified by construction period and summarised in Table 40. For opaque elements, the BA schema includes the thermal transmittance of walls (U_{wl}), roofs ($U_{fl;up}$), and floors ($U_{fl;lw}$). Transparent components are described by the window-to-wall ratio (*WWR*), window thermal transmittance (U_w), and glazing total solar energy transmittance (*g*-value). Standard internal heat gains—from occupants, lighting, and electrical appliances—are assigned according to building use category.

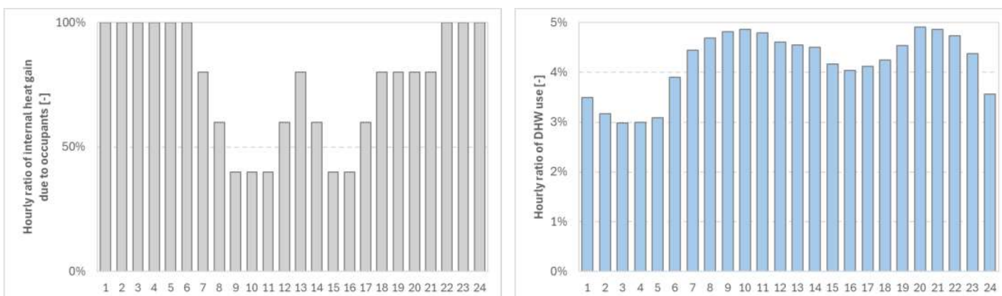
Table 40. *BA thermal and geometrical envelope characteristics*

	Opaque envelope			Transparent envelope		
year	U_{wl} W/(m²·K)	$U_{fl;up}$ W/(m²·K)	$U_{fl;lw}$ W/(m²·K)	***WWR*** −	U_w W/(m²·K)	***g*-value** −
1900-45	1.22	0.70	1.60	25 %	2.30	0.47
1946-60	1.53	0.70	1.50			
1961-70	1.02	0.65	1.30			
1971-80	0.55	0.60	1.10			
1981-90	0.89	0.43	0.68			
1991-00	0.69	0.31	0.49			
post-2000	0.51	0.22	0.25	35 %	1.70	0.49

Standard internal heat gains, generated by occupants, lighting systems, and electrical appliances, were differentiated according to building use category. Figure 88 illustrates the hourly profiles of internal heat gains associated with occupants (Figure 88a) and DHW consumption (Figure 88b) for residential buildings; in Figure 88b, the DHW hourly profile is expressed as a percentage distribution. In the simulations, both the space-heating and DHW systems were assumed to operate continuously, without intermittent schedules. Furthermore, occupant-related profiles were considered constant throughout the year: daily DHW consumption was assumed identical for every day, while hourly occupancy-related internal heat gains followed the same fixed distribution over the entire simulation period.

Space heating and domestic hot water demand are met through connections to the district heating network. Each building is equipped with a substation and a storage unit, dimensioned to meet its specific thermal requirements. Delivered heat is metered on the client side, meaning the substations operate with an equivalent system efficiency of 1.0.

Climatic data were obtained from the weather station in Aigle (46.33° N, 6.91° E, 381 m a.s.l.), located approximately 10 km from Monthey, which is the nearest station with complete weather records.



(**a**) (**b**)

Figure 88. *Usage schedules for occupants (**a**) and DHW use (**b**) (Piro, Kämpf, et al., 2026)*

During the iterative EA-based automatic calibration, four key parameters were varied to capture different physical processes influencing the building energy balance: the air change rate n, which governs ventilation-related heat losses; the occupied area per person A_{pers}, which affects internal gains from occupants; the daily domestic hot water volume per person V_{W}; and the heating set-point temperature $\theta_{\mathrm{H;set}}$, which influences the building's space heating demand. Owing to their strong influence on model outputs, these variables are widely used in large-scale calibration studies (Sokol et al., 2017; C. K. Wang et al., 2020).

A continuous range was defined for each parameter, with lower and upper bounds determined from Swiss and European technical standards. Table 41 summarises these four parameters along with their corresponding variation ranges and references. The hygiene-related air change rate (n) reflects infiltration and the natural ventilation flow rate. The upper bound of the occupied area per person also accounts for cases in which buildings may be classified as inhabited but have no actual occupants.

Overall, the proposed methodology is based on an archetype-driven UBEM approach, in which the passive thermal properties of opaque and transparent envelope components, together with the WWR, are defined at the archetype level and are therefore considered less uncertain than operational parameters, for which only standardised or conventional assumptions are generally available at the urban scale.

Table 41. *Continuous parameter ranges for automatic calibration*

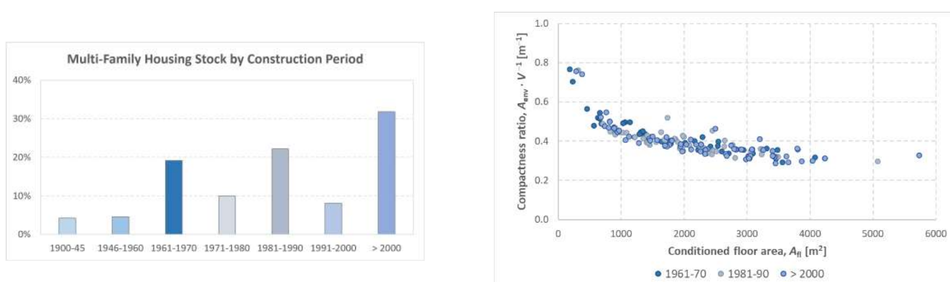
Parameter	Symbol	Unit	Range	Ref.
Air change rate	n	h^{-1}	[0.5; 1.2]	CEN (2007e)
Occupied area per person	A_{pers}	m^2/pers.	[28.3; 15,000]	CEN (2019)
Daily volume of DHW per person	V_W	l/(d·pers.)	[10.0; 50.0]	SIA (2006)
Heating set-point temperature	$\theta_{H;set}$	°C	[17.0; 25.0]	CEN (2019)

Single-building block calibration

The calibration procedure was applied to approximately 200 individual buildings. The remaining cases were calibrated at the urban-block scale due to the high level of aggregation in their energy consumption data.

Figure 89a shows the distribution of construction periods for multi-unit residential buildings, with the largest shares built in 1961-70, 1981-90, and after 2000. Among the individually calibrated buildings, 94 % are residential, and 96 % of these are apartment blocks.

In this section and in the following one (6.3.3.1), the analysis focuses on the 1961-70 archetype group, which represent a substantial portion of the calibrated stock. Figure 89b depicts the relationship between the conditioned floor area (A_{fl}) and the compactness ratio ($A_{env} \cdot V^{-1}$) for buildings constructed during this period. The compactness is defined as the ratio between the total thermal envelope area and the building volume.



(a) **(b)**

Figure 89. *Multi-family housing stock by construction period (**a**); relationship between conditioned floor area and building compactness ratio for residential houses built between 1961–1970 (**b**) (Piro, Kämpf, et al., 2026)*

The non-residential share of the stock—representing 6 %—includes hospitals, commercial buildings, offices, hotels, and buildings with unknown functions. For computational efficiency, each case was modelled using a customised large-scale energy model that incorporates neighbouring buildings within a 100-metre radius. Figure 90 illustrates the full UBEM scenario, subdivided into single-building sub-models for large-scale energy analysis, enabling individual assessments of buildings M3547, M3548, and M3549 as examples.

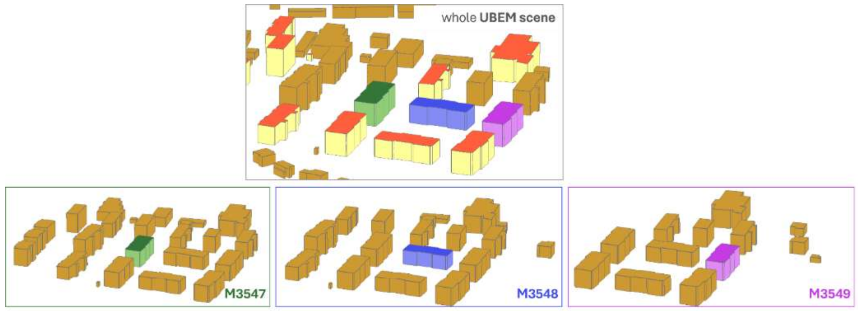
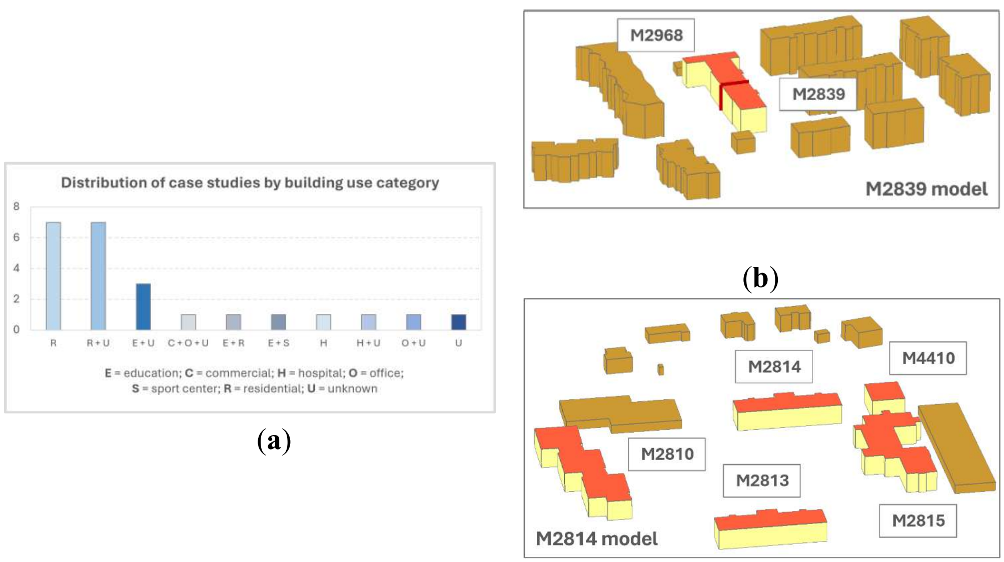


Figure 90. *Large-scale energy model discretisation: from the whole UBEM scene to M3547, M3548, and M3549 models with their surroundings (Piro, Kämpf, et al., 2026)*

Urban block calibration

In 24 urban blocks—comprising a total of 65 buildings—the district heating substation's energy meter serves multiple constructions. Each block contains between two and five buildings, which are jointly modelled within a single UBEM scenario. Most of these cases consist exclusively of residential buildings or a combination of residential and unknown-use buildings, representing 14 out of the 24 urban blocks. The remaining blocks are more heterogeneous, featuring mixed functions such as educational, hospital, commercial, and office uses, each occurring only once or in a small number of cases (Figure 91a).



(**a**)

(**b**)

(**c**)

Figure 91. *Case studies by building use category (**a**); model M2839 composed of two residential buildings (**b**); model M2814 composed of five hospitals (**c**) (Piro, Kämpf, et al., 2026)*

Figure 91b-c present two example case studies undergoing the iterative automatic calibration procedure. The first is a fully residential two-building case (M2839), consisting of apartment blocks M2839 and M2968 (Figure 91b). The second is a fully non-residential case (M2814), which includes buildings M2814, M2810, M2815, M4410, and M2813 (Figure 91c). Although these buildings are officially classified as hospitals in the geoportal, they appear to operate mainly as residential care facilities for elderly and disabled individuals.

In both examples, the sub-UBEM is discretised in the same manner: all assessed buildings are modelled together, along with surrounding shading objects located within a 100-metre radius.

Table 42 and Table 43 summarise the general, geometric, and thermophysical properties of the buildings comprising the M2839 and M2814 urban blocks, respectively. Specifically, it reports each building's construction period, use type, conditioned floor area (A_{fl}), total thermal envelope area (A_{env}), gross volume (V), window-to-wall ratio (WWR), compactness ratio ($A_{env} \cdot V^{-1}$), and the mean thermal transmittance of opaque (U_{op}) and transparent (U_w) elements.

Table 42. *General, geometrical, and mean thermal characteristics for opaque and transparent building envelope components for urban block M2839*

Bldg. code	**Age**	**Use** (*)	A_{fl} m²	A_{env} m²	V m³	WWR –	$A_{env} \cdot V^{-1}$ m⁻¹	U_{op} W/(m²K)	U_w W/(m²K)
M2839	1961-70	R	1422	1576	4265	0.25	0.37	1.00	2.30
M2968	1971-80		2027	2341	6081		0.38	0.70	
(*) R = residential									

Table 43. *General, geometrical, and mean thermal characteristics for opaque and transparent building envelope components for urban block and M2814*

Bldg. code	**Age**	**Use** (*)	A_{fl} m²	A_{env} m²	V m³	WWR –	$A_{env} \cdot V^{-1}$ m⁻¹	U_{op} W/(m²K)	U_w W/(m²K)
M2814	1961-70	E	1899	2462	5696	0.25	0.43	0.99	2.30
M2810			3947	4659	11841		0.39	0.99	
M2815			3351	4201	10052		0.42	0.99	
M4410			799	1122	2396		0.47	1.00	
M2813	1971-80		1847	2408	5542		0.43	0.72	
(*) E = homes for elderly and disabled people									

6.3.2.2 Building energy consumption

The consumption dataset represents the delivered energy demand for space heating and domestic hot water supplied by the district heating network. It comprises 314 energy meters in the municipality of Monthey: 290 correspond to individual buildings, while the remaining 24 energy meters provide aggregated readings for urban blocks consisting of two to five buildings within the same UBEM scenario.

To ensure adequate representation of the heating season, only meters with data coverage exceeding two-thirds of the year were retained. This threshold helps capture a sufficient number of space-heating days and enhances the reliability of consumption patterns by limiting the influence of incomplete records. Calibration was therefore performed only for days with valid measurements. To remain consistent with the UBEM output timestep, consumption recorded on the leap day was removed.

The dataset varies in both temporal and spatial detail. In terms of temporal resolution, some meters report values every three hours, whereas others provide daily measurements. Owing to this heterogeneity—and to account for domestic hot water storage dynamics—all records were aggregated to a daily timestep.

With respect to spatial resolution, certain meters correspond to single buildings, while others serve multiple buildings. For meters covering multiple buildings, calibration was conducted at the level of the entire urban block.

6.3.2.3 Configuration steps

The calibration analysis starts with the construction of UBEM sub-models. For each building under evaluation, the optimisation algorithm performs a one-year calibration using the combined space heating and domestic hot water energy consumption recorded in 2020. These data, provided by the energy distributor, were aggregated to daily values to compensate for missing hourly entries and to better capture DHW storage tank dynamics. As explained in Section 6.3.2.2, only buildings with at least two-thirds daily data coverage were included in the calibration process.

The EA CMA-ES/HDE algorithm (EA CMA-ES/HDE, 2025) begins by initialising a population and randomly selecting individuals to generate offspring through crossover and mutation. The most promising candidates from the parent and offspring groups are retained to form the next generation, and this evolutionary cycle is repeated until the termination criterion—defined as a maximum number of iterations—is met. When additional calibration parameters are introduced, such as through the inclusion of more buildings, the solution space expands. This increases the difficulty of efficiently locating the optimum and leads to longer computation times, as more simulations are required to sufficiently explore the enlarged parameter space. Consequently, the number of simulations needed for convergence depends on the degrees of freedom associated with each sub-UBEM, which vary according to how many buildings are thermally modelled. Table 44 reports the maximum number of individuals (simulations) executed for case studies involving one to five buildings.

Table 44. *Tested individuals in CMA-ES/HDE per different scenario*

	No. of individuals
One-building case studies	2000
Two-buildings case studies	3000
Three-buildings case studies	4000
Four-buildings case studies	5000
Five-buildings case studies	6000

6.3.3 Results and discussion

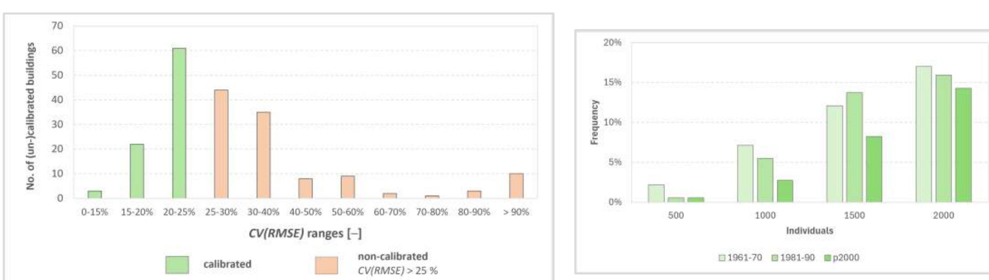
6.3.3.1 Calibration at building-scale

<u>Residential buildings: calibration results</u>

The calibration results for residential buildings constructed in 1961-70, 1981-90, and after 2000 at the building scale are presented in Figure 92a. Overall, the calibration procedure was successful in 43 % of the analysed cases, while 12 % of the tested buildings achieved a *CV(RMSE)* lower than 20 %. By adopting the more permissive threshold of 30 % for hourly data recommended by ASHRAE Guideline-14 (ASHRAE, 2023), 66 % of the sample would be considered acceptable. This comparison indicates that changes in the selected *CV(RMSE)* threshold mainly affect buildings with values close to the limit, whereas the overall calibration trends remain substantially consistent.

Figure 92b illustrates the proportion of case studies satisfying the calibration criterion, i.e., *CV(RMSE)* $\leq$ 25 %, for each individual class and for the three analysed construction periods (1961-70, 1981-90, and post-2000). A clear increasing trend can be observed across the classes: the proportion of successfully calibrated cases is very limited for the smallest classes (approximately 500 individuals), increases for intermediate classes (around 1000 individuals), and reaches the highest values for the largest classes (approximately 1500-2000 individuals), where the acceptance rate ranges between about 8 % and 16 %, depending on the construction period.

Across all classes, the 1961-70 and 1981-90 cohorts consistently exhibit higher calibration success rates than the post-2000 buildings, whose performance remains comparatively lower, particularly in the intermediate-size classes. Overall, the results suggest that larger individual classes are associated with a greater likelihood of meeting the *CV(RMSE)* threshold. Furthermore, the analysis highlights systematic differences among construction periods, with newer buildings proving more difficult to calibrate within this dataset, potentially due to the passive thermal characteristics of their envelope components.



(a) **(b)**

Figure 92. *Distribution of CV(RMSE) ranges across the (un-)calibrated buildings (**a**); frequency of case studies with CV(RMSE) ≤ 25 % across individual classes, by construction period (**b**) (Piro, Kämpf, et al., 2026)*

Residential buildings constructed in the periods 1961-70, 1981-90, and post-2000 represent the most populated clusters (Figure 89a). Figure 93 presents the boxplots of the calibrated input parameters, namely the ventilation air change rate (n), occupied area per person (A_{pers}), daily DHW volume per person (V_W), and heating set-point temperature ($\theta_{H;set}$).

The air change rate includes both infiltration and natural ventilation required for hygiene purposes. As shown in Figure 93a, most buildings converge towards $n = 0.5$ h^{-1}, with limited dispersion and only a few outliers. This indicates that the ventilation air change rate stabilises around a representative value and is not systematically exploited as a compensating parameter during the calibration process. In contrast, the occupied area per person (Figure 93b) exhibits a relatively uniform distribution, with a median value of 28.9 m²/person, close to the conventional reference value recommended by EN 16798-1 (CEN, 2019).

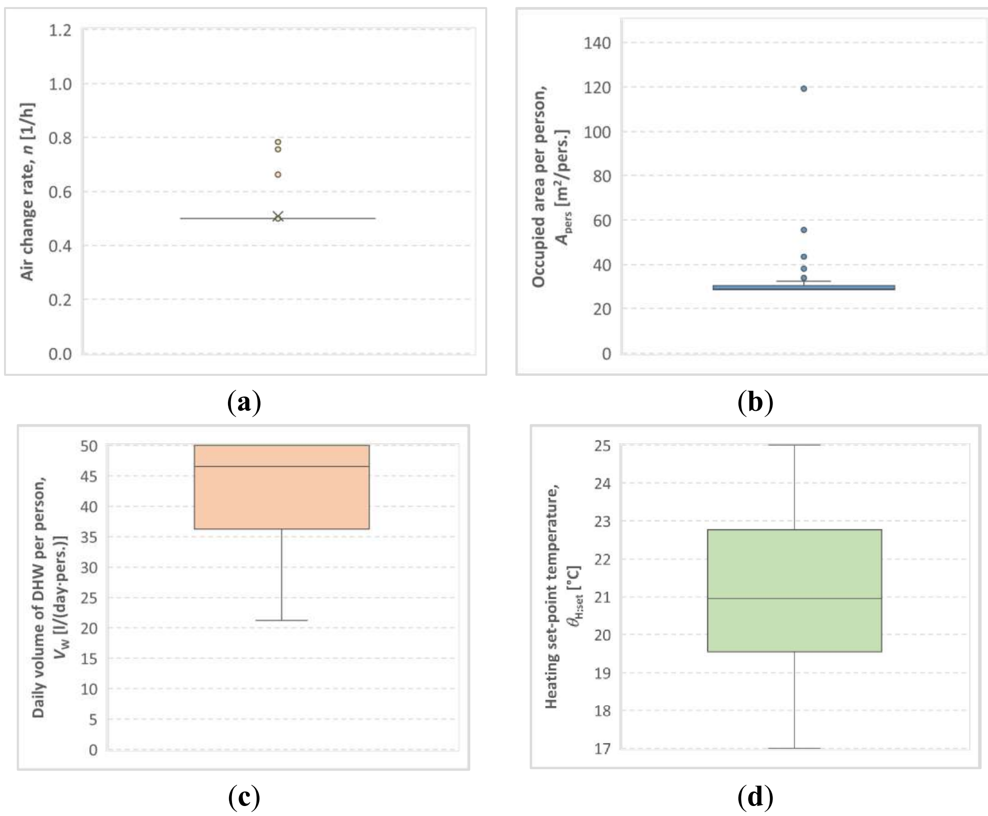


Figure 93. *Boxplots of input variables after calibration process for residential buildings: air change rate (**a**), occupied area per person (**b**), daily volume of DHW per person (**c**), and heating set-point temperature (**d**) (Piro, Kämpf, et al., 2026)*

The calibrated daily DHW volume per person (Figure 93c) lies within a relatively narrow yet elevated range, with a median of 46.6 l/(d·pers.), which is consistent with typical residential consumption patterns. Finally, the heating set-point temperature (Figure 93d) shows a median value of 21 °C and an interquartile range of 3.2 °C, highlighting moderate variability across the analysed sample. Overall, $\theta_{H,set}$ and V_W display the greatest dispersion among the calibrated parameters.

The resulting set of calibrated parameters provides a realistic representation of building-stock operation, confirming a commonly adopted heating set-point temperature around 21 °C, a typical ventilation air change rate of 0.5 h^{-1}, an occupancy density aligned with EN 16798-1 (CEN, 2019), and DHW consumption values within expected residential ranges. These findings support the physical plausibility and robustness of the proposed automatic calibration procedure.

Figure 94 presents the monthly distribution of the Normalised Mean Bias Error ($NMBE$) for the analysed dataset, which includes calibrated residential buildings constructed in 1961-70. The $NMBE$ was calculated as $(S - M)/M$, where S denotes the simulated monthly energy consumption and M the measured value.

In the simulations, the hourly diversity profiles for occupancy and DHW consumption (Figure 88a-b) were assumed to remain constant throughout the year, while the thermal zone was considered continuously occupied, i.e., without intermittent operation of space-heating and DHW systems associated with actual occupancy patterns. Overall, the model shows good performance during the winter months—particularly in January, November, and December, as well as February—where $NMBE$ values are concentrated around zero with limited dispersion. This indicates a strong agreement between simulated and measured energy use under conditions dominated by space-heating demand.

In contrast, the largest variability occurs during the mid-season months, especially September and October, when heating system operation is more irregular and often absent in real conditions. Under these circumstances, the model tends to overestimate energy consumption, resulting in higher positive biases and greater dispersion. A distinct pattern is also observed in July, plausibly associated with reduced occupancy during summer holiday periods. Since the model assumes constant internal gains and DHW demand throughout the year, this simplification may contribute to the discrepancy with measured consumption. In this respect, the implementation of intermittent system operation could represent a relevant topic for future investigation.

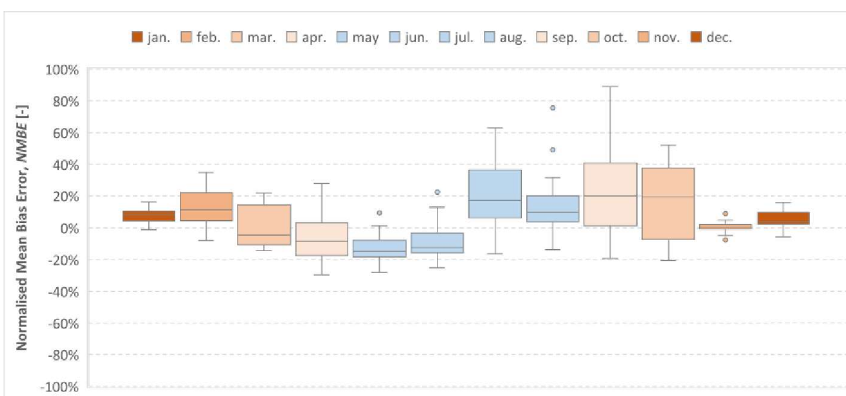
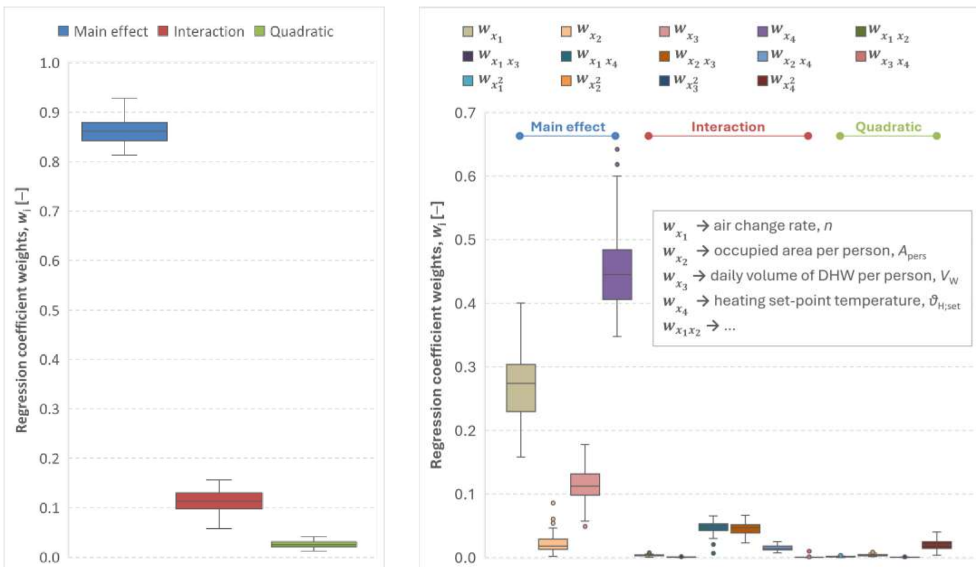


Figure 94. *Monthly NMBE for calibrated residential buildings built in 1961-70 (Piro, Kämpf, et al., 2026)*

Multivariate quadratic regression analysis

This section presents the results of the multivariate quadratic regression model developed for residential apartment blocks constructed between 1961 and 1970. The analysed sample comprises more than 100,000 simulations performed across 54 buildings. Figure 95a illustrates the regression coefficient weights aggregated into main, interaction, and quadratic effects, whereas Figure 95b reports the corresponding disaggregated regression terms.

In this analysis, the simulated space-heating and DHW energy demand computed with CitySim (Y_i) is correlated with the variation of four input parameters considered during the calibration procedure: air change rate (x_1), occupied area per person (x_2), daily DHW volume per person (x_3), and heating set-point temperature (x_4). Prior to fitting the regression model, all independent variables were standardised using a z-score normalisation approach.



(a) **(b)**

Figure 95. *Weights of regression coefficients for residential buildings constructed in 1961-70: aggregated by main, interaction, and quadratic effects (**a**), and disaggregated by independent variables (**b**) (Piro, Kämpf, et al., 2026)*

Figure 95a highlights the predominant influence of the main effects of the selected variables on the simulated energy demand. The median contribution of the first-order terms is approximately 86 %, whereas interaction effects account for a median contribution of about 11 %, reflecting the combined influence of pairs of variables. In contrast, the quadratic terms, which describe the non-linear curvature of the relationship between the predictors and Y_i, provide a comparatively limited contribution.

Figure 95b shows that variations in the heating set-point temperature, followed by the ventilation air change rate, exert the strongest influence on the overall building energy demand. Within the analysed sample, the median weight associated with w_{x_4} is approximately 45 %, while w_{x_1} is around 27 %. Concerning the interaction effects, the combination of heating set-point temperature and ventilation air change rate ($w_{x_1 x_4}$) predominantly governs the sensitivity of space-heating demand, whereas the interaction between the number of occupants in the thermal zone and the daily DHW volume ($w_{x_2 x_3}$) has the greatest influence on DHW demand variations. Among the quadratic terms, the largest contribution is associated with the non-linear effect of the heating set-point temperature ($\theta_{H;set}$).

Overall, the quadratic regression model provides an accurate representation of the local input-output relationships, achieving an excellent fit across all analysed buildings (R^2 = 0.99). The analysis therefore identifies the most influential variables and interactions governing building energy demand within the sample considered.

6.3.3.2 Calibration at block-scale

The complexity of the calibration procedure increases with the number of buildings considered simultaneously. In the case of a single building, the combined space-heating and domestic hot water demand is governed by four parameters. However, extending the same methodology to multi-building calibrations substantially increases the number of unknowns, rising from eight parameters for two buildings to twenty parameters for five buildings.

To mitigate imbalances during urban-block calibration—such as unrealistically high heating set-point temperatures in some buildings and excessively low values in others—and to reduce the solution space explored by the optimiser, the interquartile ranges of the calibrated inputs obtained at the single-building scale (Figure 93) were adopted as constraints for the urban-block calibration process. Specifically, the admissible ranges were restricted to A_{pers} = [28.3; 33.3] m²/person, V_W = [35.0; 50.0] l/(day·person), and $\theta_{H;set}$ = [19.5; 23.0] °C, instead of the broader initial ranges of A_{pers} = [28.3; 15,000] m²/person, V_W = [10.0; 50.0] l/(day·person), and $\theta_{H;set}$ = [17.0; 25.0] °C, respectively. The ventilation air change rate was instead fixed at 0.5 h^{-1}.

This section presents the results obtained from the automatic calibration procedure based on the CMA-ES/HDE optimisation approach. In particular, the successful calibration outcomes achieved for the urban-block case studies M2839 and M2814 are discussed. The results are reported in terms of *CV(RMSE)* for the entire analysed urban blocks, together with the calibrated values of occupied area per person (A_{pers}), daily DHW consumption per occupant (V_W), and space-heating set-point temperature ($\theta_{H;set}$).

Two-building case study

Table 45 reports the results of the calibration process for the residential block M2839, which consists of two apartment buildings. By achieving a *CV(RMSE)* of 16.4 %, the optimiser maintains nearly constant values for both the occupied area per person and the daily DHW consumption per occupant, while primarily adjusting the space-heating set-point temperature.

The results highlight a compensation effect within the simulated case study, namely the tendency of the model to modify one or more parameters in order to reproduce the observed energy consumption. In this specific case, the optimisation algorithm increases the heating set-point temperature of the building with the largest conditioned volume (M2968), while compensating by reducing the $\theta_{H;set}$ of building M2839 to 19.6 °C.

Although M2968 is a relatively recent building characterised by a lower average opaque thermal transmittance (Table 42), modifying its winter heating set-point temperature proves more effective because of its larger total thermal envelope surface area (A_{env}). As a result, the overall heat transfer coefficient of building M2968 exceeds that of building M2839, making it the dominant contributor in regulating the aggregated heating demand of the urban block.

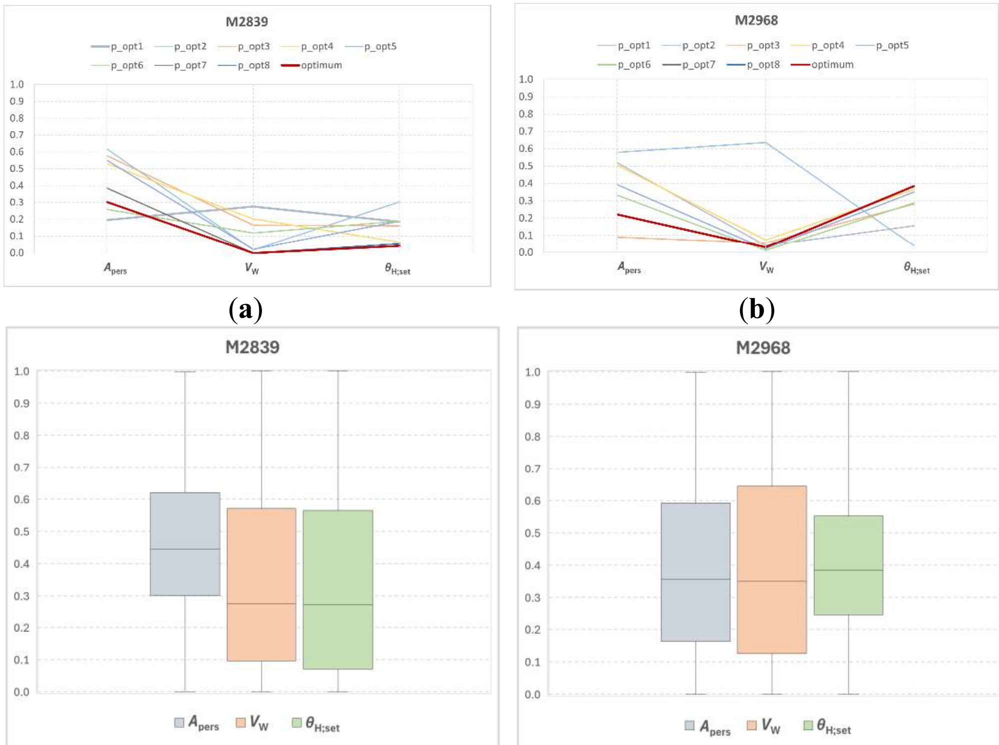
Table 45. *Calibration results for urban block M2839*

Block ID	**Bldg. code**	***CV(RMSE)*** –	$\boldsymbol{A_{pers}}$ m²/pers.	$\boldsymbol{V_W}$ l/(d·pers.)	$\boldsymbol{\theta_{H;set}}$ °C
M2839	M2839	16.4 %	32	35.0	19.6
	M2968		32	35.5	20.8

This compensation effect is illustrated in Figure 96a-b, where the coloured piecewise linear curves represent the partial optimal solutions identified by the hybrid differential evolution algorithm, i.e., the best-performing individuals, while the red curve denotes the optimal configuration achieving a *CV(RMSE)* of 16.4 % for buildings M2839 and M2968, respectively. The partial optima correspond to very similar performance levels, with *CV(RMSE)* values ranging from 16.6 % to 16.4 %, thereby highlighting the existence of multiple near-equivalent solutions.

Although all input parameters were normalised within the interval [0; 1], distinct behaviours can be observed among the variables. In both buildings, the occupied area per person exhibits the greatest variability, indicating its role as a compensating parameter during the optimisation process. Nevertheless, because the admissible range was intentionally restricted, this apparent variability corresponds only to limited absolute changes. In contrast, the space-heating set-point temperature of building M2839 displays substantially lower variability than the other inputs, suggesting that it acts as a constraining parameter in maintaining the overall energy balance of the urban block.

This behaviour is further confirmed by Figure 96c-d, which show the complete distribution of the tested input parameters explored by the optimiser using the same normalisation approach. The concentration of solutions within narrow intervals for certain variables, combined with broader dispersions for others, indicates that the optimisation process relies on compensatory trade-offs among the inputs. In particular, the heating set-point temperature of building M2968 appears to play a dominant role in regulating the equilibrium of the system, while the remaining variables adjust around this balance.



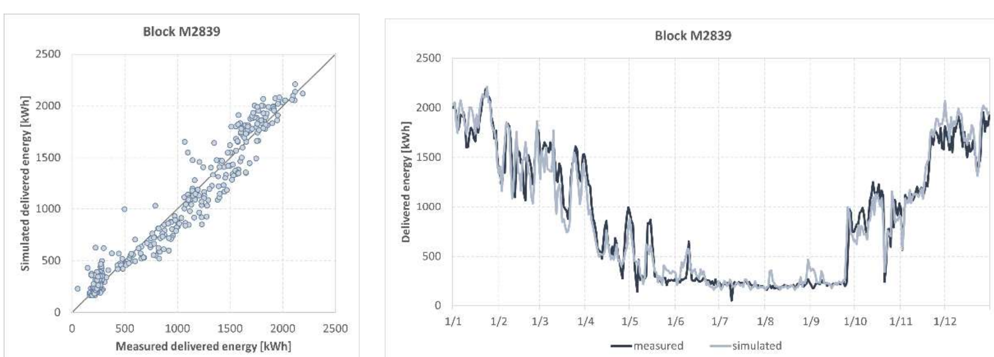
(**a**) (**b**)

(**c**) (**d**)

Figure 96. *Urban block M2839: partial and global optima and full distribution of simulated inputs, for M2839 (**a-c**) and M2968 (**b-d**) multi-family houses, respectively (Piro, Kämpf, et al., 2026)*

Figure 97a-b compares the daily measured and simulated delivered thermal energy for space heating and domestic hot water for block M2839. As shown in Figure 97a, the model demonstrates a strong agreement with the observations ($R^2 = 0.95$), with the data points closely distributed around the bisector, indicating consistent performance across the full range of daily energy loads.

The time-series comparison presented in Figure 97b further confirms that the simulation successfully reproduces both the seasonal trend and the timing of the main variations in energy demand.



(**a**) (**b**)

Figure 97. *Daily measured vs simulated delivered energy for urban block M2839: measured-simulated (**a**) and time-series comparison (**b**) (Piro, Kämpf, et al., 2026)*

The remaining discrepancies can be partially attributed to the absence of an explicit representation of intermittent system operation and occupant-driven usage patterns, which were assumed to remain continuous throughout the year. This modelling simplification may therefore limit the capability of the simulation to accurately capture short-term demand fluctuations.

Five-building case study

For the residential urban block M2814, consisting of five buildings, the optimiser achieved a *CV(RMSE)* of 21.6 % after generating and evaluating approximately 6000 individuals. The corresponding calibration results are reported in Table 46. In this case, the input vector minimising the objective function converged towards constant values across all five buildings for both the daily DHW consumption per occupant, V_W = 35.0 l/(d·pers.), and the space-heating set-point temperature, $\theta_{H;set}$ = 19.5 °C.

As a consequence, the automatic calibration procedure primarily adjusted the occupied area per person, with calibrated values ranging between 29 and 31 m²/person. Within this context, lower occupant densities are generally associated with smaller buildings, whereas higher densities tend to correspond to larger buildings. Although these results remain physically plausible, the optimiser mainly modifies a restricted subset of parameters while maintaining others unchanged, thereby indicating limited parameter identifiability at the urban-block scale.

This behaviour, already observed during the single-building calibration process, becomes more pronounced at larger spatial scales, where the number of unknown variables in the mathematical system increases and compensation effects among parameters are more likely to emerge.

Table 46. *Calibration results for urban block M2814*

Block ID	Bldg. code	*CV(RMSE)* –	A_{pers} m^2/pers.	V_W l/(d·pers.)	$\theta_{H;set}$ °C
M2839	M2814	21.6 %	30	35.0	19.5
	M2810		29		
	M2815		29		
	M4410		31		
	M2813		31		

Figure 98 presents the monthly delivered energy intensity for the urban block M2814, normalised by the total conditioned floor area. The stacked columns represent the simulated energy demand of the individual buildings (M2814, M2810, M2815, M4410, and M2813), while the adjacent grey bars correspond to the measured monthly energy consumption. The results indicate that the overall energy demand of the urban block is predominantly governed by a limited subset of buildings. In particular, buildings M2810 and M2815 provide the largest contribution throughout the year, especially during the heating season (January-March and October-December), when they largely determine the magnitude of the winter demand peaks. Together, these two buildings account for more than 60 % of the total conditioned floor area of the urban block (Table 43).

The figure further highlights the complexity of the calibration problem, as the dominant influence of a few buildings increases the sensitivity of the optimisation process to a restricted number of parameters. Moreover, a systematic underestimation of the simulated combined space-heating and domestic hot water demand can be observed during the coldest months, particularly January and December. This discrepancy may be related to the assumption of continuous, non-intermittent operation of the space-heating and DHW technical building systems. The effect is further amplified by the representation of each building as a single thermal zone, which limits the model's capability to reproduce heterogeneous demand profiles among different dwelling units within the same building.

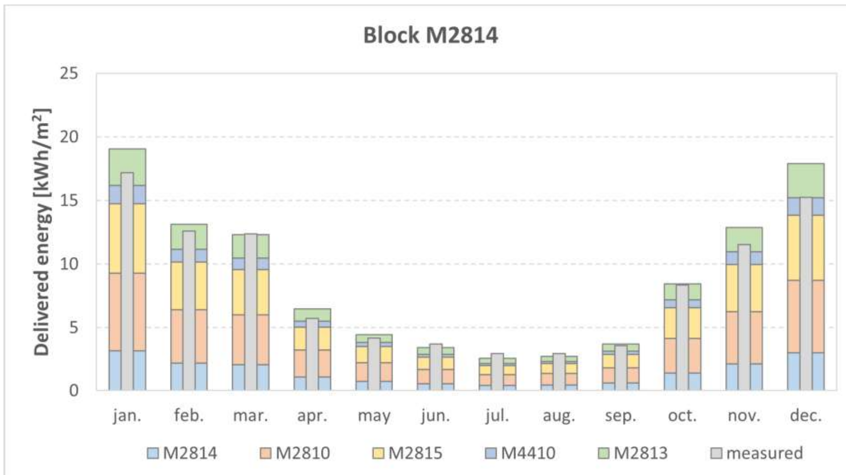
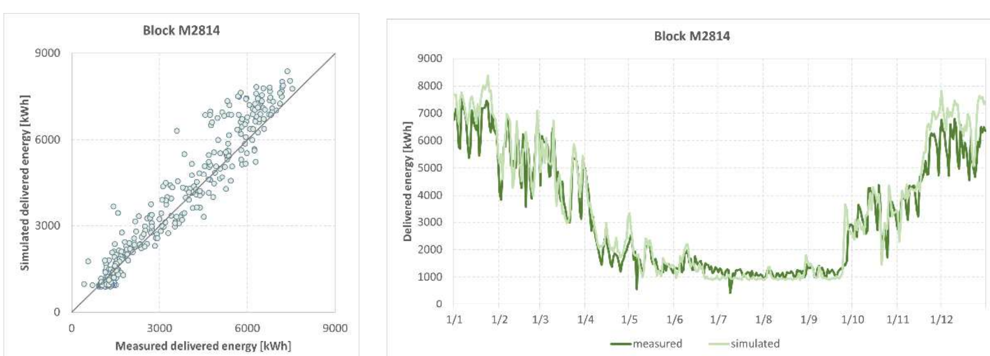


Figure 98. *Monthly delivered energy intensity for the urban block M2814, expressed in kWh/m². The stacked columns represent the simulated contributions of the individual buildings (M2814, M2810, M2815, M4410, and M2813), while the adjacent grey column indicates the monthly measured energy consumption (Piro, Kämpf, et al., 2026)*

At the daily temporal resolution, Figure 99a presents the comparison between measured and simulated values, while Figure 99b illustrates the temporal distribution of the delivered energy demand.



(**a**) (**b**)

Figure 99. *Daily measured vs simulated delivered energy for urban block M2814: measured-simulated (**a**) and time-series comparison (**b**) (Piro, Kämpf, et al., 2026)*

The measured-simulated comparison shown in Figure 99a reveals a strong correlation between simulated and observed values (R^2 = 0.94). Both figures further confirm the presence of discrepancies between simulated and measured energy demand during the winter months, consistent with the trends previously identified at the monthly scale.

<u>Discussion</u>

The calibration results highlight both methodological and modelling limitations inherent to multi-building calibration at the urban-block scale. When several buildings are calibrated simultaneously, the optimisation process exhibits a pronounced compensation effect, whereby discrepancies in one building are balanced through parameter adjustments in others. Although this phenomenon is also observed in single-building calibration, where compensatory parameter interactions may still occur, it becomes substantially more evident at the aggregated urban-block scale. In this context, larger buildings tend to dominate the overall energy demand because of their greater conditioned volume and envelope surface area. Consequently, the optimiser preferentially adjusts parameters associated with these buildings—such as the space-heating set-point temperature—in order to minimise the global error metric.

Within this framework, differences between the parameters obtained through single-building and group-level calibration are expected, since the two approaches pursue fundamentally different optimisation objectives. Single-building calibration identifies building-specific parameter sets that best reproduce the measured consumption of each individual building. Conversely, urban-block calibration seeks a shared set of parameters that minimises an aggregated error metric across multiple buildings simultaneously. The resulting solution therefore represents a compromise among buildings characterised by different thermal properties and demand profiles, which may lead to deviations from the optima identified at the individual-building scale.

These findings underline the importance of constraining the admissible variability range of the calibrated parameters. Broad parameter bounds increase the probability of compensatory solutions capable of achieving satisfactory statistical agreement with measured data while converging towards values that are only weakly identifiable or marginally physically plausible. Restricting the parameter search space is therefore essential to guide the optimisation towards solutions that are not only statistically robust, but also physically meaningful.

Beyond calibration-specific considerations, the results also reflect intrinsic limitations of the adopted modelling framework. The UBEM employed in this study is based on a simplified representation in which each building is modelled as a single thermal zone. This assumption limits the capability of the model to capture intra-building thermal heterogeneity and differences in operation among individual zones or dwelling units. Furthermore, the model does not explicitly account for intermittent system operation, either in terms of scheduling or intensity modulation. Heating operation and occupant-related variability are therefore assumed to remain continuous throughout the year (Figure 88), reducing the model’s ability to reproduce short-term fluctuations and peak demand dynamics.

Finally, the interpretation of calibration performance is further complicated by the use of aggregated measured energy data combining space-heating and domestic hot water consumption, as exemplified in Figure 98. Such aggregation obscures the relative contribution of individual end uses and limits the capability of the model to attribute discrepancies to specific demand components.

6.3.3.3 General considerations

Urban Building Energy Models build on principles established in building energy simulation tools, incorporating a series of assumptions and simplifications to represent heat exchanges at the urban scale. As the spatial scale of analysis increases, the level of geometric and informational detail typically decreases, thereby amplifying uncertainties associated with input data in large-scale energy assessments. This issue is compounded by the scarcity of city-scale energy consumption datasets with adequate spatial and temporal resolution, which constrains the calibration and validation of UBEMs. Moreover, due to the high computational burden of detailed models, grey-box or black-box approaches are often preferred. Combined with the limited availability of reliable consumption data, these constraints reduce the feasibility of iterative validation processes required to enhance model accuracy and reliability.

The UBEM domain still lacks standardised methodologies. Consequently, modelling assumptions are frequently drawn from technical standards, reports, and best-practice references developed for individual building simulations. For example, ASHRAE Guideline-14 (ASHRAE, 2023) specifies *MBE* and *CV(RMSE)* thresholds for model calibration in building energy modelling. However, given the greater simplifications, assumptions, and uncertainties inherent in UBEMs, it is unclear whether the same statistical acceptability criteria should apply; more lenient thresholds are likely more suitable for urban-scale applications.

UBEM development relies heavily on local or regional geospatial datasets that—at best—provide information on building use, construction period, and volumetric attributes. Errors in these datasets, such as misclassified construction periods, incorrect use categories, or geometric inaccuracies, directly propagate through the modelling chain. These errors affect the assignment of non-geometric properties (e.g., archetypes), the representation of occupant-related loads, and ultimately the estimation of building energy performance.

Ideally, UBEM calibration would involve modelling all thermally simulated buildings within a single unified scenario, computing total stock energy demand as the sum of individual building consumptions, and iteratively adjusting parameters until calibration criteria are satisfied. However, this approach faces two major constraints. First, the computational time required to simulate hundreds or thousands of buildings simultaneously—coupled with the multiple iterations typical of calibration procedures—is prohibitively high. Second, the dimensionality of the input parameter space expands exponentially as more buildings are included, substantially slowing convergence toward an optimal solution.

Additionally, conducting automatic calibration at the scale of aggregated urban blocks can introduce imbalances, as the procedure may result in unrealistically high energy demand estimates for buildings with larger conditioned volumes. To reduce such compensation effects, calibration at the city-block scale may be more appropriate than using broader district-level aggregations.

The main limitations affecting the applicability of the proposed methodology are summarised below:

- Limited availability of high-resolution data. Access to large-scale building energy consumption datasets with adequate spatial and temporal detail remains scarce, restricting the potential for robust calibration.
- High computational demand. The calibration process is computationally intensive. On average, running 2000 individuals (i.e., a single-building calibration) requires roughly two hours. All simulations were performed in CitySim on a workstation equipped with an Intel Core i9-14900K processor (24 cores, 32 threads, 3.2 GHz) and 32 GB RAM. Surrogate modelling could provide a more efficient alternative to mitigate this computational burden.
- Lack of engineering judgement in automated optimisation. Fully automated calibration may bias convergence toward adjustments of the most sensitive parameters, potentially suppressing variability in others. However, manual calibration at the city-block scale is impractical, making automated procedures an unavoidable compromise.

6.4 Remarks

The credibility of NBRPs is weakened when large-scale energy models are applied without prior validation or calibration procedures. A major obstacle is the limited availability of city-wide building energy consumption data with sufficient spatial and temporal resolution, which restricts the ability to verify or adjust model outputs. At the urban scale, validation is typically favoured over calibration because iterative simulations demand considerable computational resources. As a result, in the absence of robust data and well-validated models, the accuracy, reliability, and practical value of UBEM-based decision-making processes cannot be assured.

Depending on the completeness of the available building energy consumption data, this chapter introduces a novel validation and calibration procedure designed to assess the reliability of UBEM findings.

Chapter 7

Conclusion

A crucial role in fostering a reliable implementation of large-scale energy analyses lies in assessing every stage of the Urban Building Energy Modelling process. This need derives primarily from several shortcomings: (i) the lack of monitoring data at the city-scale, (ii) uncertainty in input data—particularly those related to local climate variability and building energy characteristics—and (iii) in physically based environments, the substantial computational time required to run simulations and perform multiple iterations to enhance output reliability.

Within this framework, the Ph.D. has developed methods and applications involving all phases of an Urban Building Energy Modelling process, from inputs to outputs, providing novel insights to enhance the reliability and reproducibility of UBEM workflow.

A critical literature review was conducted by comparing the traditional binary methodology adopted for the classification of urban-scale energy models with the more recent framework introduced by the IEA-EBC Annex 70, characterised by four quadrants based on different dimensions of interest. Furthermore, the distinction between UBEM, USEM, and UES was clarified. Finally, the current research gaps, including data availability, uncertainty, and quality, urban microclimate integration, and urban mobility, were identified and analysed as key aspects requiring further investigation within present and future urban energy modelling research.

The research then moved to a theoretical comparison between BEMs and UBEMs. It was highlighted that increasing the spatial scale inevitably leads to a reduction in both the level of detail and the representativeness of the energy model. Given the simplifications and assumptions typically adopted in large-scale energy models—often derived from experience gained in single-building energy modelling—this comparison analysed key modelling components: geometry, building envelope characterisation, zoning, internal heat gains, TBSs, and calculation modules. Furthermore, these simplifications were related to the corresponding UBEM tools most commonly adopted within the scientific community. Particular attention was devoted to short-wave and long-wave radiative exchanges, where differences between UBEM and BEM results of approximately 7 % and 30 %, respectively, were identified. These findings provide useful guidelines for researchers and developers involved in urban energy modelling.

Subsequently, the research addressed the limitations of input data, focusing on building energy performance certificates, which, although widely issued, present several shortcomings: (i) low data quality, (ii) static energy performance assessments, and (iii) a limited number of energy-related KPIs. To overcome these issues, a data quality-checking procedure was conceived and applied to the EPC databases from the Piedmont and Aosta Valley regions, for which EPC data are available at a different level of detail. It was demonstrated that, through the proposed EPC quality assessment procedure, approximately 30 % to 50 % of certificates may contain technical inconsistencies that undermine the reliability of their application in urban-scale analyses, including urban energy planning, energy vulnerability assessment, and urban energy efficiency evaluation. The research highlighted that the most critical parameters affecting EPC reliability are frequently geometrical variables, such as building volume and useful floor areas, consequently raising concerns regarding the reliability of non-geometrical parameters, which are generally of greater interest for energy analyses. Finally, the proposed EPC quality assessment procedure, modular and adaptable to different informational contexts, proved to be a practical methodology for identifying technically reliable certificates. The implementation of this procedure may benefit both database managers and local or regional authorities interested in obtaining a reliable overview of the energy performance level of the building stock.

After identifying the reliable EPCs, these were used to generate representative buildings according to several segmentation criteria. It was demonstrated that the use of representative buildings derived from urban energy datasets represents an effective approach for informing large-scale energy models in the absence of detailed single-building data. To this aim, one of the outcomes of this Ph.D. research was the development of building prototypes and archetypes representative of the construction typologies of north-west Italy. These tools provide information regarding the energy performance status of buildings and represent useful support for both NBRP implementation and UBEM development. However, the energy performance of representative buildings changes when they are analysed within the urban microclimate context. For this reason, a paradigm shift is required: from the concept of representative building to that of representative urban block, integrating not only geometrical aspects but also the effects generated by the UHI phenomenon. The proposed methodology proved to be adaptable to different climatic contexts and varying levels of urban data availability and harmonisation. The behaviour of these representative urban blocks was assessed under short-, medium-, and long-term climate scenarios.

Depending on the purpose of the analysis, the availability of data, and computational constraints, validation and calibration procedures were applied to enhance the reliability of large-scale energy assessments. They contribute to reducing discrepancies between simulated and measured energy consumption by identifying the parametric combinations that are most representative of actual operating conditions. The proposed validation procedure introduced two novel indicators aimed at evaluating actual operational representativeness. The methodology identified the most probable parameter combinations leading to improved agreement with real energy consumption profiles derived from hourly district heating network data. Furthermore, the validation approach showed that archetypes associated with more recent construction periods tend to better reproduce measured consumption profiles, probably due to the lower variability in construction typologies. Conversely, urban-scale calibration results demonstrated that, in approximately 45 % of the analysed cases, it was possible to achieve *CV(RMSE)* values below 25 %. In addition, the advantages and limitations of calibration at multiple-building scale were quantitatively evaluated in cases where a single meter recorded the consumption of multiple buildings. In this context, the importance of introducing constraints to limit the dominance of larger and more energy-intensive buildings within the calibration process was highlighted.

In conclusion, this Ph.D. provides a comprehensive framework contributing to the improvement of the reliability, credibility, and practical applicability of UBEMs. By integrating data quality assessment, representative-building approaches, and scalable validation and calibration techniques, the work advances the foundations for more trustworthy large-scale energy simulations. This contribution is particularly valuable for supporting energy planners and policymakers in the development of effective NBRPs and climate mitigation actions.

References

Abbasabadi, N., & Mehdi Ashayeri, J. K. (2019). Urban energy use modeling methods and tools: A review and an outlook. *Building and Environment*, *161*. https://doi.org/10.1016/j.buildenv.2019.106270.

Abolhassani, S. S., Amayri, M., Bouguila, N., & Eicker, U. (2022). A new workflow for detailed urban scale building energy modeling using spatial joining of attributes for archetype selection. *Journal of Building Engineering*, *46*. https://doi.org/10.1016/j.jobe.2021.103661.

Abruzzo Region. (2023). *Servizio di Controllo degli Attestati di Prestazione Energetica (DPR n. 75/2013 - L.R. n. 32/2015, modificata dalla L.R. n. 50/2017) - Disciplinare Prestazionale. In Italian.*

Agenzia Regionale per la Protezione Ambientale (ARPA). (2025). *ARPA Piemonte web-site*. Retrieved from https://www.arpa.piemonte.it/. Accessed April 30, 2025.

Allegrini, J., Orehounig, K., Mavromatidis, G., Ruesch, F., Dorer, V., & Evins, R. (2015). A review of modelling approaches and tools for the simulation of district-scale energy systems. *Renewable and Sustainable Energy Reviews*, *52*, 1391–1404. https://doi.org/10.1016/j.rser.2015.07.123.

American Society of Heating, R. and A.-C. E. (ASHRAE). (2023). *Measurement of Energy, Demand, and Water Savings - ASHRAE Guideline 14*. www.ashrae.org.

American Society of Heating, R. and A.-C. E. (ASHRAE). (2025). *Energy Standard for Sites and Buildings Except Low-Rise Residential Buildings - ANSI/ASHRAE/IES Standard 90.1*. www.ashrae.org.

Anagnostopoulos, F., Arcipowska, A., & Mariottini, F. (2015). *Energy Performance Certificates as tools to support and track renovation activities*.

Ang, Y. Q., Berzolla, Z. M., Letellier-Duchesne, S., Jusiega, V., & Reinhart, C. (2022). UBEM.io: A web-based framework to rapidly generate urban building energy models for carbon reduction technology pathways. *Sustainable Cities and Society*, *77*. https://doi.org/10.1016/j.scs.2021.103534.

Aosta Valley Region. (2021). *Caratteristiche e Modalità di Gestione del Sistema di Certificazione Energetica Regionale Beauclimat (DGR 1249-2021). In Italian.*

Ballarini, I., Bianco Mauthe Degerfeld, F., Corrado, V., De Luca, G., & Piro, M. (2022). *Deliverable 1.3 - Report on EPC data analysis (TIMEPAC)*. www.timepac.eu.

Ballarini, I., Corgnati, S. P., & Corrado, V. (2014). Use of reference buildings to assess the energy saving potentials of the residential building stock: The experience of TABULA project. *Energy Policy*, *68*, 273–284. https://doi.org/10.1016/j.enpol.2014.01.027.

Ballarini, I., Corrado, V., & Piro, M. (2021). Building Stock Energy Models and ICT Solutions for Urban Energy Systems. In IGI Global (Ed.), *Handbook of Research on Developing Smart Cities Based on Digital Twins* (pp. 490–514). IGI Global. https://doi.org/10.4018/978-1-7998-7091-3.ch022.

Ballarini, I., Piro, M., Corrado, V., Pernigotto, G., Borelli, G., & Gasparella, A. (2026). A Methodology for the Assessment of the Urban Heat Island Effect by Exploiting the Urban Archetype Approach. In C. Zilio, F. Busato, L. Mazzarella, & M. Noro (Eds.), *Proceedings of the 15th REHVA HVAC World Congress - CLIMA 2025* (Vol. 763, pp. 1179–1191). Springer. https://doi.org/10.1007/978-3-032-06810-1_118.

Ballarini, I., Piro, M., & Tootkaboni P, M. (2023). *Deliverable 2.5 - Procedures and services to undertake large-scale statistical analysis of EPCs databases - Transversal Deployment Scenario 5 (TIMEPAC)*. www.timepac.eu.

Basu, S., S. E. Bale, C., Wehnert, T., & Topp, K. (2019). A complexity approach to defining urban energy systems. *Cities*, *95*. https://doi.org/10.1016/j.cities.2019.05.027.

Battista, G., Evangelisti, L., Guattari, C., Roncone, M., & Balaras, C. A. (2023). Space-time estimation of the urban heat island in Rome (Italy): Overall assessment and effects on the energy performance of buildings. *Building and Environment*, *228*. https://doi.org/10.1016/j.buildenv.2022.109878.

Bianco, V., Manca, O., & Nardini, S. (2009). Electricity consumption forecasting in Italy using linear regression models. *Energy*, *34*(9), 1413–1421. https://doi.org/10.1016/j.energy.2009.06.034.

Biljecki, F., Ledoux, H., & Stoter, J. (2016). An improved LOD specification for 3D building models. *Computers, Environment and Urban Systems*, *59*, 25–37. https://doi.org/10.1016/j.compenvurbsys.2016.04.005.

Borelli, G., Ballarini, I., Corrado, V., Gasparella, A., & Pernigotto, G. (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: 6th IBPSA-Italy Conference Bozen-Bolzano*, 525–531. https://doi.org/10.13124/9788860462022.

Borges, P., Travesset-Baro, O., & Pages-Ramon, A. (2022). Hybrid approach to representative building archetypes development for urban models – A case study in Andorra. *Building and Environment*, *215*. https://doi.org/10.1016/j.buildenv.2022.108958.

Bueno, B., Norford, L., Hidalgo, J., & Pigeon, G. (2013). The urban weather generator. *Journal of Building Performance Simulation*, *6*(4), 269–281. https://doi.org/10.1080/19401493.2012.718797.

BUILD_ME. (2025). *BUILD_ME web-site*. Retrieved from https://www.buildings-mena.com/typologies. Accessed April 30, 2025.

Caceres, A. G. (2018). Shortcomings and suggestions to the EPC recommendation list of measures: In-depth interviews in six countries. *Energies*, *11*(10). https://doi.org/10.3390/en11102516.

Carnero, P., Van Dijk, D., Mignani, N., & Ana, G. (2022). A common European EPB Assessment and Certification scheme. U-CERT's proposal. *Proceedings of CLIMA 2022 Conference: 14th REHVA HVAC World Congress Rotterdam*. https://doi.org/10.34641/clima.2022.279.

Carnieletto, L., Ferrando, M., Teso, L., Sun, K., Zhang, W., Causone, F., Romagnoni, P., Zarrella, A., & Hong, T. (2021). Italian prototype building models for urban scale building performance simulation. *Building and Environment*, *192*. https://doi.org/10.1016/j.buildenv.2021.107590.

Cerezo, C., Sokol, J., AlKhaled, S., Reinhart, C., Al-Mumin, A., & Hajiah, A. (2017). Comparison of four building archetype characterization methods in urban building energy modeling (UBEM): A residential case study in Kuwait City. *Energy and Buildings*, *154*, 321–334. https://doi.org/10.1016/j.enbuild.2017.08.029.

Cerezo Davila, C., Reinhart, C. F., & Bemis, J. L. (2016). Modeling Boston: A workflow for the efficient generation and maintenance of urban building energy models from existing geospatial datasets. *Energy*, *117*, 237–250. https://doi.org/10.1016/j.energy.2016.10.057.

Chen, G., Rong, L., & Zhang, G. (2020). Comparison of urban airflow between solar-induced thermal wall and uniform wall temperature boundary conditions by coupling CitySim and CFD. *Building and Environment*, *172*. https://doi.org/10.1016/j.buildenv.2020.106732.

Chen, Y., Deng, Z., & Hong, T. (2020). Automatic and rapid calibration of urban building energy models by learning from energy performance database. *Applied Energy*, *277*. https://doi.org/10.1016/j.apenergy.2020.115584.

Chen, Y., Hong, T., & Piette, M. A. (2017). City-Scale Building Retrofit Analysis: A Case Study using CityBES. *Proceedings of Building Simulation 2017: 15th Conference of IBPSA San Francisco*, 259–266. https://doi.org/10.26868/25222708.2017.071.

Choi, S., Yi, D. H., Kim, D. W., & Yoon, S. (2025). Multi-source data fusion-driven urban building energy modeling. *Sustainable Cities and Society*, *123*. https://doi.org/10.1016/j.scs.2025.106283.

Climate Resilient Strategies by Archetype-based Urban Energy Modelling (CRiStAll). (2025). *CRiStAll web-site*. Retrieved from https://cristall.polito.it. Accessed April 30, 2025.

Crawley, D. B., Hand, J. W., Kummert, M., & Griffith, B. T. (2008). Contrasting the capabilities of building energy performance simulation programs. *Building and Environment*, *43*(4), 661–673. https://doi.org/10.1016/j.buildenv.2006.10.027.

D'Alonzo, V., Novelli, A., Vaccaro, R., Vettorato, D., Albatici, R., Diamantini, C., & Zambelli, P. (2020). A bottom-up spatially explicit methodology to estimate the space heating demand of the building stock at regional scale. *Energy and Buildings*, *206*. https://doi.org/10.1016/j.enbuild.2019.109581.

De Luca, G. (2022). *Energy performance assessment of buildings: a novel approach for model validation* [Doctoral dissertation, Politecnico di Torino]. https://iris.polito.it/retrieve/1558666e-2f80-4688-a206-afb6f33af8ac/conv_deluca_phd_thesis_r1_2022-09-14.pdf.

Deng, Z., Chen, Y., Yang, J., & Causone, F. (2023). AutoBPS: A tool for urban building energy modeling to support energy efficiency improvement at city-scale. *Energy and Buildings*, *282*. https://doi.org/10.1016/j.enbuild.2023.112794.

Dilsiz, A. D., Nweye, K. E., Wu, A. J., Kämpf, J. H., Biljecki, F., & Nagy, Z. (2023). How spatio-temporal resolution impacts urban energy calibration. *Energy and Buildings*, *292*. https://doi.org/10.1016/j.enbuild.2023.113175.

Dogan, T., Reinhart, C., & Michalatos, P. (2014). Automated multi-zone building energy model generation for schematic design and urban massing studies. *Proceedings of ESim 2014: 8th Conference of IBPSA-Canada*, 123–136.

EA CMA-ES/HDE. (2025). *EA CMA-ES/HDE GitHub web-site*. Retrieved from https://github.com/jkaempf/EA. Accessed April 30, 2025.

European Commission. (2002). *Directive 2002/91/EC of the European Parliament and of the Council of 16 December 2002 on the energy performance of buildings*.

European Commission. (2019). *The European Green Deal*. https://eur-lex.europa.eu/resource.html?uri=cellar:b828d165-1c22-11ea-8c1f-01aa75ed71a1.0002.02/DOC_1&format=PDF.

European Commission. (2024a). *Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast)*.

European Commission. (2024b). *Next-generation of Energy Performance Assessment and Certification*. https://cordis.europa.eu/programme/id/H2020_LC-SC3-B4E-4-2020/en.

European Commission. (2025). *Energy performance certificates (Articles 19-21, Annex V) and independent control systems (Article 27, Annex VI) - Annex 3*.

European Committee for Standardization (CEN). (2007a). *Building components and building elements - Thermal resistance and thermal transmittance - Calculation method (ISO 6946:2007).*

European Committee for Standardization (CEN). (2007b). *Building materials and products - Hygrothermal properties - Tabulated design values and procedures for determining declared and design thermal values Incorporating (EN ISO 10456).*

European Committee for Standardization (CEN). (2007c). *Heating systems in buildings - Method for calculation of system energy requirements and system efficiencies - Part 3-3: Domestic hot water systems, generation (EN 15316-3-3).*

European Committee for Standardization (CEN). (2007d). *Hygrothermal performance of buildings - Calculation and presentation of climatic data Part 6: Accumulated temperature differences (degree-days) (ISO 15927-6).*

European Committee for Standardization (CEN). (2007e). *Thermal performance of buildings - Transmission and ventilation heat transfer coefficient - Calculation method (ISO 13789:2007).*

European Committee for Standardization (CEN). (2008). *Energy performance of buildings - Calculation of energy use for space heating and cooling (EN ISO 13790).*

European Committee for Standardization (CEN). (2011a). *Glass in building - Determination of luminous and solar characteristics of glazing (EN 410).*

European Committee for Standardization (CEN). (2011b). *Glass in building — Determination of thermal transmittance (U value) — Calculation method (EN 673)*.

European Committee for Standardization (CEN). (2017a). *Energy performance of buildings - Energy needs for heating and cooling, internal temperatures and sensible and latent heat loads - Part 1: Calculation procedures (EN ISO 52016-1).*

European Committee for Standardization (CEN). (2017b). *Energy performance of buildings - Method for calculation of system energy requirements and system efficiencies - Part 4-1: Space heating and DHW generation systems, combustion systems (boilers, biomass), Module M3-8-1, M8-8-1 (EN 15316-4-1).*

European Committee for Standardization (CEN). (2017c). *Energy performance of buildings - Method for calculation of system energy requirements and system efficiencies - Part 4-2: Space heating generation systems, heat pump systems, Module M3-8-2, M8-8-2 (EN 15316-4-2).*

European Committee for Standardization (CEN). (2017d). *Energy performance of buildings - Method for calculation of the design heat load - Part 1: Space heating load, Module M3-3 (EN 12831-1).*

European Committee for Standardization (CEN). (2017e). *Energy performance of buildings - Method for calculation of the design heat load - Part 3: Domestic hot water systems heat load and characterisation of needs, Module M8-2, M8-3 (EN 12831-3).*

European Committee for Standardization (CEN). (2017f). *Energy performance of buildings - Overarching EPB assessment - Part 1: General framework and procedures (ISO 52000-1:2017)*.

European Committee for Standardization (CEN). (2017g). *Energy performance of buildings - Ventilation for buildings - Part 5-1: Calculation methods for energy requirements of ventilation and air conditioning systems (Modules M5-6, M5-8, M6-5, M6-8, M7-5, M7-8) — Method 1: Distribution and generation (EN 16798-5-1).*

European Committee for Standardization (CEN). (2017h). *Energy performance of buildings-Method for calculation of system energy requirements and system efficiencies - Part 1: General and Energy performance expression, Module M3-1, M3-4, M3-9, M8-1, M8-4 (EN 15316-1).*

European Committee for Standardization (CEN). (2017i). *Energy performance of buildings-Method for calculation of system energy requirements and system* *efficiencies - Part 4-3: Heat generation systems, thermal solar and photovoltaic systems, Module M3-8-3, M8-8-3, M11-8-3 (EN 15316-4-3).*

European Committee for Standardization (CEN). (2017j). *Energy performance of buildings-Thermal, solar and daylight properties of building components and elements - Part 1: Simplified calculation method of the solar and daylight characteristics for solar protection devices combined with glazing (EN ISO 52022-1).*

European Committee for Standardization (CEN). (2017k). *Energy performance of buildings-Ventilation for buildings - Part 7: Calculation methods for the determination of air flow rates in buildings including infiltration (Modules M5-5) (BS EN 16798-7).*

European Committee for Standardization (CEN). (2017l). *Thermal performance of buildings - Heat transfer via the ground - Calculation methods (EN ISO 13370).*

European Committee for Standardization (CEN). (2018a). *Air conditioners, liquid chilling packages and heat pumps, with electrically driven compressors, for space heating and cooling - Testing and rating at part load conditions and calculation of seasonal performance (EN 14825).*

European Committee for Standardization (CEN). (2018b). *Thermal performance of building components - Dynamic thermal characteristics - Calculation methods (EN ISO 13786).*

European Committee for Standardization (CEN). (2019). *Energy performance of buildings-Ventilation for buildings - Part 1: Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics - Module M1-6 (EN 16798-1).*

European Committee for Standardization (CEN). (2020a). *Photovoltaic devices - Part 1: Measurement of photovoltaic current-voltage characteristics (EN IEC 60904-1).*

European Committee for Standardization (CEN). (2020b). *Thermal performance of windows, doors and shutters: Calculation of thermal transmittance. Part 1, General (EN ISO 10077-1).* British Standards Institution.

European Committee for Standardization (CEN). (2024). *EUB SuperHub - A harmonization of KPIs for supporting the next generation of EPCs (CWA 18127)*.

Fabrizio, E., & Monetti, V. (2015). Methodologies and advancements in the calibration of building energy models. *Energies*, *8*(4), 2548–2574. https://doi.org/10.3390/en8042548.

Faure, X., Lebrun, R., & Pasichnyi, O. (2024). Impact of time resolution on estimation of energy savings using a copula-based calibration in UBEM. *Energy and Buildings*, *311*. https://doi.org/10.1016/j.enbuild.2024.114134.

Fennell, P. J., Ebrahimigharehbaghi, S., Mata, É., Kokogiannakis, G., Amrith, S., Ignatiadou, S., & Piamo, S. Lo. (2025). Urban building energy models: how can we improve the treatment of uncertainty for energy policy decision-making? *Environmental Research Communications*, *7*(2). https://doi.org/10.1088/2515-7620/ad9438.

Fennell, P. J., Van Hove, M., Weinberg, L., Bennett, G. J., Delghust, M., Forthuber, S., Jakob, M., Mata, É., Nägeli, C., Reyna, J. L., & Catenazzi, G. (2022). Challenges and Lessons Learned in Applying Sensitivity Analysis to Building Stock Energy Models. *Proceedings of Building Simulation 2021: 17th Conference of IBPSA Bruges*, 2203–2210. https://doi.org/10.26868/25222708.2021.30960.

Ferrando, M., Causone, F., Banfi, A., Corrado, V., Ballarini, I., Piro, M., Zarrella, A., Carnieletto, L., Borgato, N., Evola, G., Detommaso, M., Nicoletti, F., Vallati, A., & Fiorini, C. V. (2025). Developing an Italian Library of Reference Buildings for Urban Building Energy Modeling (UBEM): Lessons Learnt from the URBEM Project. *Energies*, *18*(22), 6026. https://doi.org/10.3390/en18226026.

Ferrando, M., Causone, F., Hong, T., & Chen, Y. (2020). Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches. *Sustainable Cities and Society*, *62*. https://doi.org/10.1016/j.scs.2020.102408.

Fonseca, J. A., Nguyen, T. A., Schlueter, A., & Marechal, F. (2016). City Energy Analyst (CEA): Integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts. *Energy and Buildings*, *113*, 202–226. https://doi.org/10.1016/j.enbuild.2015.11.055.

Geoportal of the Metropolitan City of Turin. (2025). *Geoportal of the Metropolitan City of Turin*. Retrieved from http://geoportale.comune.torino.it/web/. Accessed April 30, 2025.

Gholami, M., Torreggiani, D., Tassinari, P., & Barbaresi, A. (2021). Narrowing uncertainties in forecasting urban building energy demand through an optimal archetyping method. *Renewable and Sustainable Energy Reviews*, *148*. https://doi.org/10.1016/j.rser.2021.111312.

Goy, S., Maréchal, F., & Finn, D. (2020). Data for urban scale building energy modelling: Assessing impacts and overcoming availability challenges. *Energies*, *13*(6). https://doi.org/10.3390/en13164244.

Gröger, G., Kolbe, T. H., Czerwinski, A., & Nagel, C. (2008). *OpenGIS ® City Geography Markup Language (CityGML) Encoding Standard*. http://www.opengeospatial.org/legal/.

Haas, R., & Schipper, L. (1998). Residential energy demand in OECD-countries and the role of irreversible efficiency improvements. *Energy Economics*, *20*(4), 421–442. https://doi.org/10.1016/S0140-9883(98)00003-6.

Hamdy, M., Carlucci, S., Hoes, P. J., & Hensen, J. L. M. (2017). The impact of climate change on the overheating risk in dwellings—A Dutch case study. *Building and Environment*, *122*, 307–323. https://doi.org/10.1016/j.buildenv.2017.06.031.

Haneef, F., Pernigotto, G., Gasparella, A., & Kämpf, J. H. (2021). Application of urban scale energy modelling and multi-objective optimization techniques for building energy renovation at district scale. *Sustainability*, *13*(20). https://doi.org/10.3390/su132011554.

Happle, G., Fonseca, J. A., & Schlueter, A. (2018). A review on occupant behavior in urban building energy models. *Energy and Buildings*, *174*, 276–292. https://doi.org/10.1016/j.enbuild.2018.06.030.

Hardy, A., & Glew, D. (2019). An analysis of errors in the Energy Performance certificate database. *Energy Policy*, *129*, 1168–1178. https://doi.org/10.1016/j.enpol.2019.03.022.

Hårsman, B., Daghbashyan, Z., & Chaudhary, P. (2016). On the quality and impact of residential energy performance certificates. *Energy and Buildings*, *133*, 711–723. https://doi.org/10.1016/j.enbuild.2016.10.033.

Hedegaard, R. E., Kristensen, M. H., Pedersen, T. H., Brun, A., & Petersen, S. (2019). Bottom-up modelling methodology for urban-scale analysis of residential space heating demand response. *Applied Energy*, *242*, 181–204. https://doi.org/10.1016/j.apenergy.2019.03.063.

Heo, Y., Choudhary, R., & Augenbroe, G. A. (2012). Calibration of building energy models for retrofit analysis under uncertainty. *Energy and Buildings*, *47*, 550–560. https://doi.org/10.1016/j.enbuild.2011.12.029.

Hong, T., Chen, Y., Luo, X., Luo, N., & Lee, S. H. (2020). Ten questions on urban building energy modeling. *Building and Environment*, *168*. https://doi.org/10.1016/j.buildenv.2019.106508.

Hong, T., Malik, J., Langevin, J., Liu, Z., & Yan, D. (2025). Ten questions on building stock modeling to inform energy efficiency and sustainability. *Building and Environment*, *284*. https://doi.org/10.1016/j.buildenv.2025.113465.

Hwang, J., Lim, H., & Lim, J. (2024). Reducing uncertainty of building shape information in urban building energy modeling using Bayesian calibration. *Sustainable Cities and Society*, *116*. https://doi.org/10.1016/j.scs.2024.105895.

IEE Project. (2013). *IEE Project - TABULA and EPISCOPE web-site*. Retrieved from https://episcope.eu/. Accessed April 30, 2025.

International Organization for Standardization (ISO). (1997). *Geometric product specifications (GPS). Surface texture: profile method. Terms, definitions and surface texture parameters (ISO 4287)*.

International Organization for Standardization (ISO). (2016). *Sustainable cities and communities-Vocabulary (ISO 37100).*

Italian Organisation for Standardisation (UNI). (2008). *Energy performance of buildings - Part 1: Evaluation of energy need for space heating and cooling (UNI/TS 11300-1). In Italian.*

Italian Organisation for Standardisation (UNI). (2014). *Opaque envelope components - Thermo-physical parameters (UNI/TR 11552). In Italian.*

Italian Organisation for Standardisation (UNI). (2016). *Riscaldamento e raffrescamento degli edifici. Dati climatici. Differenze di temperatura cumulate (gradi giorno) ed altri indici sintetici (UNI 10349-3). In Italian.*

Italian Organisation for Standardisation (UNI). (2025). *Italian National Annex to EN 16798-1 (UNI EN 16798-1:2019/NA).*

Italian Republic. (2005). *Attuazione della direttiva 2002/91/CE relativa al rendimento energetico nell'edilizia (D.Lgs. 192/2005). In Italian.*

Italian Republic. (2015). *Decreto interministeriale 26 giugno 2015 - Applicazione delle metodologie di calcolo delle prestazioni energetiche e definizione delle prescrizioni e dei requisiti minimi degli edifici. In Italian.*

Johari, F., Munkhammar, J., Shadram, F., & Widén, J. (2022). Evaluation of simplified building energy models for urban-scale energy analysis of buildings. *Building and Environment*, *211*. https://doi.org/10.1016/j.buildenv.2021.108684.

Johari, F., Peronato, G., Sadeghian, P., Zhao, X., & Widén, J. (2020). Urban building energy modeling: State of the art and future prospects. *Renewable and Sustainable Energy Reviews*, *128*. https://doi.org/10.1016/j.rser.2020.109902.

Johari, F., Shadram, F., & Widén, J. (2023). Urban building energy modeling from geo-referenced energy performance certificate data: Development, calibration, and validation. *Sustainable Cities and Society*, *96*. https://doi.org/10.1016/j.scs.2023.104664.

Kamel, E. (2022). A Systematic Literature Review of Physics-Based Urban Building Energy Modeling (UBEM) Tools, Data Sources, and Challenges for Energy Conservation. *Energies*, *15*(22). https://doi.org/10.3390/en15228649.

Kämpf, J. H. (2024). *Evolutionary Algorithm - CMA-ES/HDE*. Retrieved from https://github.com/jkaempf/EA. Accessed April 30, 2025.

Kämpf, J. H., & Robinson, D. (2007). A simplified thermal model to support analysis of urban resource flows. *Energy and Buildings*, *39*(4), 445–453. https://doi.org/10.1016/j.enbuild.2006.09.002.

Kämpf, J. H., & Robinson, D. (2009). A hybrid CMA-ES and HDE optimisation algorithm with application to solar energy potential. *Applied Soft Computing Journal*, *9*(2), 738–745. https://doi.org/10.1016/j.asoc.2008.09.009.

Kämpf, J. H., Wetter, M., & Robinson, D. (2010). A comparison of global optimisation algorithms with standard benchmark functions and real-world applications using EnergyPlus. *Journal of Building Performance Simulation*, *3*, 103–120. https://doi.org/10.1080/19401490903494597.

Kavgic, M., Mavrogianni, A., Mumovic, D., Summerfield, A., Stevanovic, Z., & Djurovic-Petrovic, M. (2010). A review of bottom-up building stock models for energy consumption in the residential sector. *Building and Environment*, *45*(7), 1683–1697. https://doi.org/10.1016/j.buildenv.2010.01.021.

Keirstead, J., & Shah, N. (2013). *Urban Energy Systems - An Integrated Approach* (J. Keirstead & N. Shah, Eds.; 1st Edition). Routledge. https://doi.org/10.4324/9780203066782.

Khazal, A., & Sønstebø, O. J. (2020). Valuation of energy performance certificates in the rental market – Professionals vs. nonprofessionals. *Energy Policy*, *147*. https://doi.org/10.1016/j.enpol.2020.111830.

Kim, H. G., & Kim, S. S. (2024). UBEM calibration method by energy consumption disaggregation using a change-point model. *Journal of Building Performance Simulation*, 1–28. https://doi.org/10.1080/19401493.2024.2412135.

Kolbe, T. H. (2009). Representing and Exchanging 3D City Models with CityGML. In S. Lee Jiyeong and Zlatanova (Ed.), *3D Geo-Information Sciences* (pp. 15–31). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-87395-2_2.

Kolbe, T. H., Kutzner, T., Smyth, C. S., Nagel, C., Roensdorf, C., & Heazel, C. (2021). *OGC Document: 20-010 OGC City Geography Markup Language (CityGML) Part 1: Conceptual Model Standard*. http://portal.opengeospatial.org/public_ogc/change_request.php.

Kontar, R. El, & Fleming, K. (2020). URBANopt: An Open-Source Software Development Kit for Community and Urban District Energy Modeling. *Proceedings of 2020 Building Performance Analysis Conference and SimBuild Co-Organized by ASHRAE and IBPSA-USA*, 293–301.

Langevin, J., Reyna, J. L., Ebrahimigharehbaghi, S., Sandberg, N., Fennell, P., Nägeli, C., Laverge, J., Delghust, M., Mata, Van Hove, M., Webster, J., Federico, F., Jakob, M., & Camarasa, C. (2020). Developing a common approach for classifying building stock energy models. *Renewable and Sustainable Energy Reviews*, *133*. https://doi.org/10.1016/j.rser.2020.110276.

Larsen, O. K., Pomianowski, M. Z., Chiesa, G., Belias, E., de Kerchove d'Exaerde, T., Flourentzou, F., Fasano, F., & Grasso, P. (2023). E-DYCE - Dynamic approach to the dynamic energy certification of buildings. *Proceedings of International Conference CISBAT 2023*, *2600*(3), 1–6. https://doi.org/10.1088/1742-6596/2600/3/032015.

Li, Y., Kubicki, S., Guerriero, A., & Rezgui, Y. (2019). Review of building energy performance certification schemes towards future improvement. *Renewable and Sustainable Energy Reviews*, *113*. https://doi.org/10.1016/j.rser.2019.109244.

Liguria Region. (2018). *Regione Liguria - Regolamento Consiglio 1-2018 (Allegati). In Italian.*

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

Lundström, L., & Akander, J. (2020). Bayesian Calibration with Augmented Stochastic State-Space Models of District-Heated Multifamily Buildings. *Energies*, *13*(1). https://doi.org/10.3390/en13010076.

Machard, A., Salvati, A., P. Tootkaboni, M., Gaur, A., Zou, J., Wang, L. L., Baba, F., Ge, H., Bre, F., Bozonnet, E., Corrado, V., Luo, X., Levinson, R., Lee, S. H., Hong, T., Salles Olinger, M., Machado, R. M. e. S., da Guarda, E. L. A., Veiga, R. K., … Holzer, P. (2024). Typical and extreme weather datasets for studying the resilience of buildings to climate change and heatwaves. *Scientific Data* , *11*(1). https://doi.org/10.1038/s41597-024-03319-8.

Maia, I., Kranzl, L., Toth, Z., Volt J, Monteiro, C., & Fragoso, R. (2022). *D4.4 Description of methodologies and concepts for the technical implementation of each feature regarding improved handling and use of EPC data in selected implementing countries*. https://x-tendo.eu.

Malinovec Puček, M., Khoja, A., Bazzan, E., & Gyuris, P. (2023). A Data Structure for Digital Building Logbooks: Achieving Energy Efficiency, Sustainability, and Smartness in Buildings across the EU. *Buildings*, *13*(4). https://doi.org/10.3390/buildings13041082.

Marinosci, C., & Morini, G. L. (2014). Check-in and control activities on the energy performance certificates in Emilia-Romagna (Italy). *Energy Procedia*, *45*, 434–442. https://doi.org/10.1016/j.egypro.2014.01.047.

Mattsson, M., Danielski, I., Olofsson, T., & Nair, G. (2025). Archetypes-based calibration for urban building energy modelling. *Energy and Buildings*, *343*. https://doi.org/10.1016/j.enbuild.2025.115843.

Mun, D. D., Kämpf, J. H., & Kim, J. J. (2026). Effects of cool coatings on urban microclimate and outdoor thermal comfort: A CFD–CitySim pro coupled simulation study. *Energy and Buildings*, *359*. https://doi.org/10.1016/j.enbuild.2026.117245.

Nageler, P., Koch, A., Mauthner, F., Leusbrock, I., Mach, T., Hochenauer, C., & Heimrath, R. (2018). Comparison of dynamic urban building energy models (UBEM): Sigmoid energy signature and physical modelling approach. *Energy and Buildings*, *179*, 333–343. https://doi.org/10.1016/j.enbuild.2018.09.034.

Nageler, P., Zahrer, G., Heimrath, R., Mach, T., Mauthner, F., Leusbrock, I., Schranzhofer, H., & Hochenauer, C. (2017). Novel validated method for GIS based automated dynamic urban building energy simulations. *Energy*, *139*, 142–154. https://doi.org/10.1016/j.energy.2017.07.151.

Nagpal, S., Mueller, C., Aijazi, A., & Reinhart, C. F. (2019). A methodology for auto-calibrating urban building energy models using surrogate modeling techniques. *Journal of Building Performance Simulation*, *12*(1), 1–16. https://doi.org/10.1080/19401493.2018.1457722.

Nakano, A., Bueno, B., Norford, L., & Reinhart, C. F. (2015). Urban Weather Generator - A novel workflow for integrating Urban Heat Island effect within urban design process. *Proceedings of Building Simulation 2015: 14th Conference of IBPSA Hyderabad*. https://doi.org/10.26868/25222708.2015.2909.

Nejadshamsi, S., Eicker, U., Bentahar, J., & Wang, C. (2025). Improving urban-scale building occupancy and energy use estimation using a transportation-informed building occupancy estimation framework. *Energy and Buildings*, *333*. https://doi.org/10.1016/j.enbuild.2025.115468.

Nouvel, R., Brassel, K.-H., Bruse, M., Duminil, E., Coors, V., Eicker, U., & Robinson, D. (2015). SimStadt, a new workflow-driven urban energy simulation platform for CityGML city models. *Proceedings of International Conference CISBAT 2015*, 889–894. https://doi.org/10.5075/epfl-cisbat2015-889-894.

Nouvel, R., Zirak, M., Coors, V., & Eicker, U. (2017). The influence of data quality on urban heating demand modeling using 3D city models. *Computers, Environment and Urban Systems*, *64*, 68–80. https://doi.org/10.1016/j.compenvurbsys.2016.12.005.

Oke, T. R. (1973). City size and the Urban Heat Island. *Atmospheric Environment*, *7*, 769–779.

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

Parkinson, A., De Jong, R., Cooke, A., & Guthrie, P. (2013). Energy performance certification as a signal of workplace quality. *Energy Policy*, *62*, 1493–1505. https://doi.org/10.1016/j.enpol.2013.07.043.

Pasichnyi, O., Wallin, J., & Kordas, O. (2019). Data-driven building archetypes for urban building energy modelling. *Energy*, *181*, 360–377. https://doi.org/10.1016/j.energy.2019.04.197.

Pasichnyi, O., Wallin, J., Levihn, F., Shahrokni, H., & Kordas, O. (2019). Energy performance certificates — New opportunities for data-enabled urban energy policy instruments? *Energy Policy*, *127*, 486–499. https://doi.org/10.1016/j.enpol.2018.11.051.

Perez, D. (2014). *A framework to model and simulate the disaggregated energy flows supplying buildings in urban areas* [Doctoral dissertation, Ecole Polytechnique Fédérale de Lausanne (EPFL)]. https://infoscience.epfl.ch/handle/20.500.14299/101151.

Pezeshki, Z., Soleimani, A., & Darabi, A. (2019). Application of BEM and using BIM database for BEM: A review. *Journal of Building Engineering*, *23*, 1–17. https://doi.org/10.1016/j.jobe.2019.01.021.

Piro, M., Ballarini, I., & Corrado, V. (2023). From Building Energy Models (BEM) to Urban Building Energy Models (UBEM): input data and modelling approaches. *Proceedings of Building Simulation 2023: 18th Conference of IBPSA Shanghai*, *18*, 3475–3482. https://doi.org/10.26868/25222708.2023.1662.

Piro, M., Ballarini, I., & Corrado, V. (2025a). Assessing data uncertainty in large-scale energy model. *Proceedings of Building Simulation 2025: 19th Conference of IBPSA Brisbane*. https://doi.org/10.26868/25222708.2025.1706.

Piro, M., Ballarini, I., & Corrado, V. (2025b). Building Archetypes Supporting the National Building Renovation Plan. *Proceedings of Building Simulation Applications BSA 2024: 6th IBPSA-Italy Conference Bozen-Bolzano*, 183–190. https://doi.org/10.13124/9788860462022_24.

Piro, M., Ballarini, I., Tootkaboni, M. P., Corrado, V., Pernigotto, G., Borelli, G., & Gasparella, A. (2026). The Urban Heat Island Under Climate Change: Analysis of Representative Urban Blocks in Northwestern Italy. *Energies*, *19*(3), 660. https://doi.org/10.3390/en19030660.

Piro, M., Bianco Mauthe Degerfeld, F., Ballarini, I., & Corrado, V. (2024). The challenges for a holistic, flexible and through-life updated energy performance certificate. *Sustainable Energy Technologies and Assessments*, *69*. https://doi.org/10.1016/j.seta.2024.103922.

Piro, M., Kämpf, J. H., Ballarini, I., & Corrado, V. (2025). Assessing the reliability of archetype-based Urban Building Energy Simulations: A case study analysis in Turin (Italy). *Journal of Physics: Conference Series*, *3140*, 62028. https://doi.org/10.1088/1742-6596/3140/6/062028.

Piro, M., Kämpf, J. H., Ballarini, I., & Corrado, V. (2026). A scalable, automatic, and evolutionary algorithm for calibrating urban building energy models. *Sustainable Cities and Society*, *144*. https://doi.org/10.1016/j.scs.2026.107428.

Prataviera, E., Romano, P., Carnieletto, L., Pirotti, F., Vivian, J., & Zarrella, A. (2021). EUReCA: An open-source urban building energy modelling tool for the efficient evaluation of cities energy demand. *Renewable Energy*, *173*, 544–560. https://doi.org/10.1016/j.renene.2021.03.144.

Prataviera, E., Vivian, J., Lombardo, G., & Zarrella, A. (2022). Evaluation of the impact of input uncertainty on urban building energy simulations using uncertainty and sensitivity analysis. *Applied Energy*, *311*. https://doi.org/10.1016/j.apenergy.2022.118691.

Prieler, M., Leeb, M., & Reiter, T. (2017). Characteristics of a database for energy performance certificates. *Energy Procedia*, *132*, 1000–1005. https://doi.org/10.1016/j.egypro.2017.09.704.

Reinhart, C. F., & Cerezo Davila, C. (2016). Urban building energy modeling - A review of a nascent field. *Building and Environment*, *97*, 196–202. https://doi.org/10.1016/j.buildenv.2015.12.001.

Reinhart, C. F., Dogan, T., Jakubiec, A., Rakha, T., & Sang, A. (2013). UMI - An urban simulation environment for building energy use, daylighting and walkability. *Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association*, 476–483. https://doi.org/10.26868/25222708.2013.1404.

Remmen, P., Lauster, M., Mans, M., Fuchs, M., Osterhage, T., & Müller, D. (2018). TEASER: An open tool for urban energy modelling of building stocks. *Journal of Building Performance Simulation*, *11*(1), 84–98. https://doi.org/10.1080/19401493.2017.1283539.

Ritchie, H., Samborska, V., & Roser, M. (2024). *Urbanization*. Published online at OurWorldinData.org. Retrieved from: https://ourworldindata.org/urbanization.

Robinson, D., Haldi, F., Leroux, P., Perez, D., Rasheed, A., & Wilke, U. (2009). CitySim: Comprehensive Micro-Simulation of Resource Flows for Sustainable Urban Planning. *Proceedings of Building Simulation 2009: 11th International IBPSA Conference Glasgow*, 1083–1090. https://doi.org/10.26868/25222708.2009.1083-1090.

Robinson, D., & Stone, A. (2005). Holistic Radiation Modelling with a Fast Simplified Radiosity Algorithm. *Proceedings of Building Simulation 2005: 9th International IBPSA Conference Montréal*, 1043–1050.

Salvalai, G., Zhu, Y., & Sesana, M. M. (2024). From building energy modeling to urban building energy modeling: A review of recent research trend and simulation tools. *Energy and Buildings*, *319*. https://doi.org/10.1016/j.enbuild.2024.114500.

Salvati, A., Palme, M., Chiesa, G., & Kolokotroni, M. (2020). Built form, urban climate and building energy modelling: case-studies in Rome and Antofagasta. *Journal of Building Performance Simulation*, *13*(2), 209–225. https://doi.org/10.1080/19401493.2019.1707876.

Seduikyte, L., Georgali, P.-Z., Panteli, C., Koltsios, S., Morsink-Georgali, P.-Z., Chatzipanagiotidou, P., Stavros, K., Stasiulienė, L., Spūdys, P., Pupeikis, D., Jurelionis, A., & Fokaides, P. (2022). Next-Generation Energy Performance Certificates, What novel implementation do we need? *Proceedings of CLIMA 2022 Conference: 14th REHVA HVAC World Congress Rotterdam*. https://doi.org/10.34641/clima.2022.348.

Sokol, J., Cerezo Davila, C., & Reinhart, C. F. (2017). Validation of a Bayesian-based method for defining residential archetypes in urban building energy models. *Energy and Buildings*, *134*, 11–24. https://doi.org/10.1016/j.enbuild.2016.10.050.

Sola, A., Corchero, C., Salom, J., & Sanmarti, M. (2020). Multi-domain urban-scale energy modelling tools: A review. *Sustainable Cities and Society*, *54*. https://doi.org/10.1016/j.scs.2019.101872.

Sučić, B., Stegnar, G., Ballarini, I., Gorrino, A., Sicilia, Á., Charalambides, A., & Vryonidi, M. (2022). *Guidelines for the implementation of a continuous EPC workflow*. https://timepac.eu.

Swan, L. G., & Ugursal, V. I. (2009). Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. *Renewable and Sustainable Energy Reviews*, *13*(8), 1819–1835. https://doi.org/10.1016/j.rser.2008.09.033.

Swiss Society of Engineers and Architects (SIA). (2006). *Conditions d'utilisation standard pour l'energie et les installations du bâtiment (SIA 2024). In French.*

swissBUILDINGS3D. (2025). *swissBUILDINGS3D web-site*. Retrieved from https://www.swisstopo.admin.ch/en/landscape-model-swissbuildings3d-2-0. Accessed April 30, 2025.

Thevenard, D., & Haddad, K. (2006). Ground reflectivity in the context of building energy simulation. *Energy and Buildings*, *38*(8), 972–980. https://doi.org/10.1016/j.enbuild.2005.11.007.

Tognoli, M., Peronato, G., & Kämpf, J. H. (2023). Prediction of Building Hourly Heating Demand from CityGML Files: Training Workflow and Deployment as an API. *Building Simulation Conference Proceedings*, *18*, 2932–2939. https://doi.org/10.26868/25222708.2023.1570.

Towards Innovative Methods for Energy Performance Assessment and Certification (TIMEPAC). (2025). *TIMEPAC web-site*. Retrieved from https://www.timepac.eu. Accessed April 30, 2025.

Tronchin, L., & Fabbri, K. (2012). Energy Performance Certificate of building and confidence interval in assessment: An Italian case study. *Energy Policy*, *48*, 176–184. https://doi.org/10.1016/j.enpol.2012.05.011.

Tsatsakis, K., Tsitsanis, A., & Malavazos, C. (2017). Enhanced human-centric building energy performance rating: Operational rating and behavioral change in the scope of OrbEEt project. *Proceedings of Global Internet of Things Summit (GIoTS) 2017*, 1–6. https://doi.org/10.1109/GIOTS.2017.8016268.

Urban Reference Buildings for Energy Modelling (URBEM). (2025). *URBEM web-site*. Retrieved from https://www.urbem.polimi.it/. Accessed April 30, 2025.

Verheyen, J., Lambie, E., Boneta, M. F., Maia, I., Kranzl, L., Urbanz, T., & Frick, D. (2022). Next generation energy performance assessment. *Proceedings of CLIMA 2022 Conference: 14th REHVA HVAC World Congress Rotterdam*. https://doi.org/10.34641/clima.2022.260.

Vivian, J., Prataviera, E., Bano, G., & Zarrella, A. (2025). Development and Calibration of an Urban Building Energy Model for the City of Padua. *Building Simulation Applications*, 77–84. https://doi.org/10.13124/9788860462022.

Wang, C., Ferrando, M., Causone, F., Jin, X., Zhou, X., & Shi, X. (2022). Data acquisition for urban building energy modeling: A review. *Building and Environment*, *217*. https://doi.org/10.1016/j.buildenv.2022.109056.

Wang, C. K., Tindemans, S., Miller, C., Agugiaro, G., & Stoter, J. (2020). Bayesian calibration at the urban scale: a case study on a large residential heating demand application in Amsterdam. *Journal of Building Performance Simulation*, *13*(3), 347–361. https://doi.org/10.1080/19401493.2020.1729862.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). Comment: The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*. https://doi.org/10.1038/sdata.2016.18.

Yang, Y., Tan, Z., & Schläpfer, M. (2025). Assessing the space-use efficiency of French cities by coupling city volumes with mobile data traffic. *Sustainable Cities and Society*, *124*. https://doi.org/10.1016/j.scs.2025.106292.

Zhang, W., Sun, K., Li, H., Rodriguez-Garcia, L., Heleno, M., & Hong, T. (2026). Calibration of urban building energy model using smart meter data for district peak load prediction. *Applied Energy*, *407*. https://doi.org/10.1016/j.apenergy.2025.127348.

Zhang, W., Sun, K., Polly, B., Goldwasser, D., Fleming, K., & Hong, T. (2026). District-Scale Analysis of Electricity Load and Strategies to Improve Energy

Reliability Using Prototype District Models. *Sustainable Cities and Society*, *143*. https://doi.org/10.1016/j.scs.2026.107348.

Zhao, K., Fennell, P. J., & Tomei, J. (2022). Representations of people in Urban Building Energy Models. *Proceedings of Building Simulation 2021: 17th Conference of IBPSA Bruges*, 3601–3609. https://doi.org/10.26868/25222708.2021.30660.

Zhou, J., Fennell, P., Korolija, I., & Ruyssevelt, P. (2025). Percentage-based thermal zoning approach for enhanced stock-level building energy performance modelling. *Energy and Buildings*, *329*. https://doi.org/10.1016/j.enbuild.2024.115231.

Zuhaib, S., Schmatzberger, S., Volt, J., Toth, Z., Kranzl, L., Eugenio Noronha Maia, I., Verheyen, J., Borragán, G., Monteiro, C. S., Mateus, N., Fragoso, R., & Kwiatkowski, J. (2022). Next-generation energy performance certificates: End-user needs and expectations. *Energy Policy*, *161*. https://doi.org/10.1016/j.enpol.2021.112723.

Publications

Journal articles

Piro, M., Kämpf, J. H., Ballarini, I., & Corrado, V. (2026). A scalable, automatic, and evolutionary algorithm for calibrating urban building energy models. *Sustainable Cities and Society*, *144*. https://doi.org/10.1016/j.scs.2026.107428.

Piro, M., Ballarini, I., P.Tootkaboni, M., Corrado, V., Pernigotto, G., Borelli, G., & Gasparella, A. (2026). The Urban Heat Island Under Climate Change: Analysis of Representative Urban Blocks in Northwestern Italy. *Energies, 19*(3), 660. https://doi.org/10.3390/en19030660.

Ferrando, M., Causone, F., Banfi, A., Corrado, V., Ballarini, I., Piro, M., Zarrella, A., Carnieletto, L., Borgato, N., Evola, G., et al. (2025). Developing an Italian Library of Reference Buildings for Urban Building Energy Modeling (UBEM): Lessons Learnt from the URBEM Project. *Energies, 18*, 6026. https://doi.org/10.3390/en18226026.

Bianco Mauthe Degerfeld, F., P. Tootkaboni, M., Piro, M., Ballarini, I., & Corrado, V. (2025). Toward Zero-Emission Buildings in Italy: A Holistic Approach to Identify Actions Under Current and Future Climates. *Energies, 18*(11), 2721. https://doi.org/10.3390/en18112721.

Piro, M., Bianco Mauthe Degerfeld, F., Ballarini, I., & Corrado, V. (2024). The challenges for a holistic, flexible and through-life updated energy performance certificate. *Sustainable Energy Technologies and Assessments, 69*. https://doi.org/10.1016/j.seta.2024.103922.

Bianco Mauthe Degerfeld, F., Piro, M., De Luca, G., Ballarini, I., & Corrado, V. (2023). The Application of EN ISO 52016-1 to Assess Building Cost-Optimal Energy Performance Levels in Italy. *Energy Reports, 10*, 1702–1717. https://doi.org/10.1016/j.egyr.2023.08.014.

Piro, M., Ballarini, I., & Corrado, V. (2020). Use of Energy Performance Certificates data repositories in Urban Building Energy Models. *AiCARR Journal, 64(5)*, https://doi.org/10.36164/AiCARRJ.64.05.03.

Book chapters

Corrado, V., Ballarini, I., Piro, M., & P.Tootkaboni, M. (2026). Sostenibilità ambientale e qualità dell’ambiente interno. In A. Osello & M. Cassis (Maggioli Editore), *Il Gemello Digitale della Forma Urbana - L’ex Manifattura Tabacchi di Torino*. In print. In Italian.

Ballarini, I., Bianco Mauthe Degerfeld, F., Corrado, V., Piro, M., Sicilia Gómez, Á., Sučić, B., et al. (2025). Tools and protocols for buildings sustainability assessment and certification. In F. Asdrubali & U. Berardi (Taylor & Francis), *Sustainability Certifications, Labels and Tools in the Built Environment - How to Evaluate, Certificate and Reduce the Energy and Environmental Impacts of Buildings* (pp. 79–207). Taylor & Francis. https://doi.org/10.1201/9781032705149-4.

Ballarini, I., Corrado, V., & Piro, M. (2021). Building Stock Energy Models and ICT Solutions for Urban Energy Systems. In M. Del Giudice & A. Osello (IGI Global), *Handbook of Research on Developing Smart Cities Based on Digital Twins* (pp. 490–514). IGI Global. https://doi.org/10.4018/978-1-7998-7091-3.ch022.

Papers in conference proceedings

Piro, M., Tootkaboni, M. P., Borelli, G., Ballarini, I., & Corrado, V. (2026). Climate-Resilient Strategies for Urban Heat Island Mitigation Using Representative Urban Blocks. In print in the *Proceedings of Building Simulation Applications (BSA): 7th Conference of IBPSA-Italy*, *Bozen-Bolzano (Italy)*.

Piro, M., Ballarini, I., & Corrado, V. (2025). Assessing data uncertainty in large-scale energy model. *Proceedings of Building Simulation (BS) 2025: 19th Conference of IBPSA*. https://doi.org/10.26868/25222708.2025.1706.

Piro, M., Kämpf, J. H., Ballarini, I., & Corrado, V. (2025). Assessing the reliability of archetype-based Urban Building Energy Simulations: A case study analysis in Turin (Italy). *Journal of Physics: Conference Series, 3140*, 62028. https://doi.org/10.1088/1742-6596/3140/6/062028.

Ballarini, I., Piro, M., Corrado, V., Pernigotto, G., Borelli, G., & Gasparella, A. (2026). A Methodology for the Assessment of the Urban Heat Island Effect by Exploiting the Urban Archetype Approach. In C. Zilio, F. Busato, L. Mazzarella, & M. Noro (Eds.), *Proceedings of the 15th REHVA HVAC World Congress - CLIMA 2025* (Vol. 763, pp. 1179–1191). Springer. https://doi.org/10.1007/978-3-032-06810-1_118.

Piro, M., Ballarini, I., & Corrado, V. (2024). Building archetypes supporting the national building renovation plan. *Proceedings of Building Simulation Applications (BSA): 6th Conference of IBPSA-Italy*, 183–190. https://doi.org/10.13124/9788860462022_24.

Piro, M., Ballarini, I., & Corrado, V. (2023). From Building Energy Models (BEM) to Urban Building Energy Models (UBEM): input data and modelling approaches. *Proceedings of Building Simulation (BS) 2023: 18th Conference of IBPSA*, *Shanghai (China)*, 3475–3482. https://doi.org/10.26868/25222708.2023.1662.

Piro, M., Bianco Mauthe Degerfeld, F., De Luca, G., Ballarini, I., & Corrado, V. (2022). Hourly-Simplified Calculation To Identify Cost-Optimal Energy Performance Requirements For The Italian Building Stock. *Proceedings of Building Simulation Applications (BSA): 7th Conference of IBPSA-Italy*, *Bozen-Bolzano (Italy)*, 153–160.

Abstracts in conference proceedings

Piro, M., Bianco Mauthe Degerfeld, F., Ballarini, I., & Corrado, V. (2024). The challenges for a holistic, flexible and through-life updated Energy Performance Certificate. *Proceedings of SDEWES Conference 2023: 18th SDEWES Conference*, *Dubrovnik (Croatia)*.

Bianco Mauthe Degerfeld, F., Piro, M., De Luca, G., Ballarini, I., & Corrado, V. (2023). The Application of EN ISO 52016-1 to Assess Building Cost-Optimal Energy Performance Levels in Italy. *Proceedings of SDEWES Conference 2022: 17th SDEWES Conference*, *Paphos (Cyprus)*.

Reports

Ballarini, I., Pernigotto, G., Tootkaboni, M. P., & Piro, M. (2026). Atlas of the typical urban context configuration: KPIs to assess the resilient strategies under future urban climate (D4.3).

Pernigotto, G., & Piro, M. (2025). Atlas of the typical urban context configuration: KPIs under future urban climate. CRiStAll project deliverable (D3.3).

Pernigotto, G., & Piro, M. (2025). Atlas of the typical urban context configuration: model features. CRiStAll project deliverable (D3.2).

Piro, M., Ballarini, I., & Pernigotto, G. (2025). Typical urban context configurations using archetypes. CRiStAll project deliverable (D3.1).

Gorrino, A., Soma, P., Madrazo, L., Sicilia, Á, Pečkaj, M., Čižman, J., Ballarini, I., Corrado, V., Bianco Mauthe Degerfeld, F., Piro, M., et al. (2023). Intermediate Exploitation Plan. TIMEPAC project deliverable (D5.11).

Ballarini, I., Piro, M., & P.Tootkaboni, M. (2023). Procedures and services to undertake large-scale statistical analysis of EPCs databases. TIMEPAC project deliverable (D2.5).

Corrado, V., Ballarini, I., Bianco Mauthe Degerfeld, F., De Luca, G., & Piro, M. (2022). Proceedings of the TIMEPAC-22 International Workshop. TIMEPAC project deliverable (D5.7).

Ballarini, I., Bianco Mauthe Degerfeld, F., Corrado, V., De Luca, G., & Piro, M. (2022). Report on EPC data analysis. TIMEPAC project deliverable (D1.3).

Corrado, V., Ballarini, I., Bianco Mauthe Degerfeld, F., & Piro, M. (2021). Aggiornamento della metodologia di calcolo dei livelli ottimali di

prestazione energetica in funzione dei costi (Report RdS/PTR2021/108 – ENEA). In Italian.

Research projects

Involvement in the PRIN-2022 PNRR CRiStAll project (https://cristall.polito.it/, 2023-26), with co-authorship of four project deliverables.

Involvement in the PRIN-2020 URBEM project (https://www.urbem.polimi.it/, 2022-25), with co-authorship of two project deliverables and regional scorecards for Piedmont and Aosta Valley.

Involvement in the H2020 TIMEPAC project (https://timepac.eu/, 2021-24), with co-authorship of three project deliverables, a factsheet, and a guideline.

Involvement in the Erasmus+ Energy Efficiency Expert project (https://eeexpert-project.eu, 2020-23), with co-authorship of training materials for energy efficiency experts.

Appendix A – EPC data quality checking procedure

Rules and scores of EPC data quality-checking procedure

Table A.1. *EPC data criticality, codification, and typology of rule(s)*

Field	Critical (C) or non-critical (n-C)	EPC data	Symbol	Unit	Data type	Typology of rule(s) (*)	Rule
General data	C	EPC ID code			string	D	not null
	C	Building use category			string	D	not null
	n-C	Assessed object			string	D	not null
Identifying data	n-C	Building city			string	D	not null
	n-C	No. building units	n_{unit}	–	number	D, P	$n_{unit} > 0$ (if field populated)
	C	Climatic region			string	D	" A" (**) " B" " C" " D" " E" " F"
	n-C	Heating degree days	HDD	°C·d	number	D, P, C	Climatic region = " A" $\rightarrow HDD \leq 600$ Climatic region = " B" $\rightarrow 601 \leq HDD \leq 900$ Climatic region = " C" $\rightarrow 901 \leq HDD \leq 1400$ Climatic region = " D" $\rightarrow 1401 \leq HDD \leq 2100$ Climatic region = " E" $\rightarrow 2101 \leq HDD \leq 3000$ Climatic region = " F" $\rightarrow HDD > 3000$
	C	Year of construction		–	number	D, P	1800 < Year of construction < 2023 (The upper limit is related to the issued EPC year)

Field	Critical (C) or non-critical (n-C)	EPC data	Symbol	Unit	Data type	Typology of rule(s) (*)	Rule
	C	Year of last renovation		–	number	D, P, C	1800 < Year of construction < 2023 (if field populated) AND Year of last renovation > Year of construction
	n-C	Building typology			string	D	not null
	n-C	Building property			string	D	not null
	n-C	Building constructive typology			string	D	not null
Geometrical data	C	Thermally heated gross volume	$V_{H;g}$	m^3	number	D, P, C	$V_{H;g} > 0$ Space heating energy service = "TRUE" $V_{H;g} \cdot A_{H;g;use}^{-1} \geq 2.30$ m
	C	Thermally cooled gross volume	$V_{C;g}$	m^3	number	D, P, C	$V_{C;g} > 0$ Space cooling energy service = "TRUE" $V_{C;g} \cdot A_{C;g;use}^{-1} \geq 2.30$ m
	n-C	Gross volume	V_g	m^3	number	D, P, C	$V_g > 0$ $V_g \cdot V_n^{-1} > 1.00$ (present in the old XML format)
	C	Net volume	V_n	m^3	number	D, P, C	$V_n > 0$ $V_n \cdot A_{H;use}^{-1} \geq 2.30$ m (present in the old XML format)
	C	Thermally heated floor area	$A_{H;use}$	m^2	number	D, P, C	$A_{H;use} > 0$ Space heating energy service = "TRUE" $V_{H;g} \cdot A_{H;g;use}^{-1} \geq 2.30$ m
	C	Thermally cooled floor area	$A_{C;use}$	m^2	number	D, P, C	$A_{C;use} > 0$ Space cooling energy service = "TRUE" $V_{C;g} \cdot A_{C;g;use}^{-1} \geq 2.30$ m

Field	Critical (C) or non-critical (n-C)	EPC data	Symbol	Unit	Data type	Typology of rule(s) (*)	Rule
	C	Thermally heated gross floor area	$A_{H;g;use}$	m^2	number	D, P, C	$A_{H;g;use} > 0$ Space heating energy service = " TRUE" $V_{H;g} \cdot A_{H;g;use}{}^{-1} \geq 2.30$ m
	C	Useful floor area	A_{use}	m^2	number	D, P	$A_{use} > 0$ (present in the old XML format)
	C	Gross floor area	$A_{fl;g}$	m^2	number	D, P	$A_{fl;g} > 0$ (present in the old XML format)
	C	Thermal envelope area	A_{env}	m^2	number	D, P	$A_{env} \geq 0$
	n-C	Opaque thermal envelope area	$A_{env;op}$	m^2	number	D, P, C	$0.75 \cdot A_{env} \leq A_{env;op} + A_{env;w} \leq A_{env}$
	n-C	Thermal envelope area of walls	A_{wl}	m^2	number	D, P, C	$0.75 \cdot A_{env} \leq A_{env;op} + A_{env;w} \leq A_{env}$
	n-C	Thermal envelope area of roof	$A_{fl,up}$	m^2	number	D, P, C	$0.75 \cdot A_{env} \leq A_{env;op} + A_{env;w} \leq A_{env}$
	n-C	Thermal envelope area of floor	$A_{fl,lw}$	m^2	number	D, P, C	$0.75 \cdot A_{env} \leq A_{env;op} + A_{env;w} \leq A_{env}$
	n-C	Transparent thermal envelope area	$A_{env;w}$	m^2	number	D, P, C	$0.75 \cdot A_{env} \leq A_{env;op} + A_{env;w} \leq A_{env}$
	n-C	Compactness ratio	CR	m^{-1}	number	D, P	$CR \geq 0$
Envelope	C	Mean thermal transmittance of opaque building envelope	$U_{m;op}$	$W \cdot m^{-2} \cdot K^{-1}$	number	D, P	$U_{op} = (0.0; 4.5]\ W \cdot m^{-2} \cdot K^{-1}$
	n-C	Mean thermal transmittance of walls	U_{wl}	$W \cdot m^{-2} \cdot K^{-1}$	number	D, P	$U_{wl} = (0.0; 4.5]\ W \cdot m^{-2} \cdot K^{-1}$

Field	Critical (C) or non-critical (n-C)	EPC data	Symbol	Unit	Data type	Typology of rule(s) (*)	Rule
	n-C	Mean thermal transmittance of roof	$U_{fl,up}$	$W \cdot m^{-2} \cdot K^{-1}$	number	D, P	$U_{op} = (0.0; 4.5]\ W \cdot m^{-2} \cdot K^{-1}$
	n-C	Mean thermal transmittance of floor	$U_{fl,lw}$	$W \cdot m^{-2} \cdot K^{-1}$	number	D, P	$U_{op} = (0.0; 4.5]\ W \cdot m^{-2} \cdot K^{-1}$
	n-C	Glazing system type			string	D	not null
	n-C	Frame system type			string	D	not null
	C	Mean thermal transmittance of transparent building envelope	$U_{m;w}$	$W \cdot m^{-2} \cdot K^{-1}$	number	D, P	$U_{w} = (0.0; 6.0]\ W \cdot m^{-2} \cdot K^{-1}$
Technical building system	n-C	Space heating energy service			boolean	D	" TRUE"
	n-C	Space cooling energy service			boolean	D	" TRUE"
	n-C	Domestic hot water energy service			boolean	D, C	"TRUE" if (Building use category = " E1.1" OR Building use category = " E1.2") (***) " FALSE"
	n-C	Ventilation energy service			boolean	D	" TRUE" " FALSE"
	n-C	Artificial lighting energy service			boolean	D	" TRUE" " FALSE"
	n-C	Transportation energy service			boolean	D	" TRUE" " FALSE"
	n-C	Space heating system typology			string	D, C	not null Generator nominal power ($P_{N;k}$) > 0

Field	Critical (C) or non-critical (n-C)	EPC data	Symbol	Unit	Data type	Typology of rule(s) (*)	Rule
	n-C	Space cooling system typology			string	D, P, C	not null $P_{N;k} > 0$
	n-C	Space domestic hot water system typology			string	D, P, C	not null $P_{N;k} > 0$
	n-C	Energy carrier (k) for space heating			string	D, P, C	not null $P_{N;k} > 0$
	n-C	Energy carrier (k) for space cooling			string	D, P, C	not null $P_{N;k} > 0$
	n-C	Energy carrier (k) for DHW			string	D, P, C	not null $P_{N;k} > 0$
	n-C	Energy carrier (k) for ventilation			string	D, P, C	not null $P_{N;k} > 0$
	n-C	Energy carrier (k) for artificial lighting			string	D, P, C	not null $P_{N;k} > 0$
	n-C	Energy carrier (k) for transportation			string	D, P, C	not null $P_{N;k} > 0$
	n-C	Autonomous/centralised space heating system			string	D, P, C	not null (if not specified autonomous if $P_{N;H} \leq 35$ kW)
	n-C	Autonomous/centralised domestic hot water system			string	D, P, C	not null (if not specified autonomous if $P_{N;W} \leq 35$ kW)
	n-C	Overall mean seasonal efficiency of the space heating system	$\eta_{s;H}$	–	string	D, P, C	Space heating energy service = " TRUE" $\eta_{s;H} > 0$
	n-C	Overall mean seasonal efficiency of the DHW system	$\eta_{s;W}$	–	string	D, P, C	Domestic hot water energy service = " TRUE" $\eta_{s;W} > 0$

Field	Critical (C) or non-critical (n-C)	EPC data	Symbol	Unit	Data type	Typology of rule(s) [(*)]	Rule
	n-C	Overall mean seasonal efficiency of the space cooling system	$\eta_{s;C}$	–	string	D, P, C	Space cooling energy service = “ TRUE” $\eta_{s;C} > 0$
	n-C	Mean seasonal efficiency of the heating generation sub-system	$\eta_{H;gn}$	–	string	D, P, C	Space heating energy service = “ TRUE” $\eta_{H;gn} > 0$
	n-C	Mean seasonal efficiency of the heating distribution sub-system	$\eta_{H;d}$	–	string	D, P, C	Space heating energy service = “ TRUE” $\eta_{H;d} = (0.0; 1.0]$
	n-C	Mean seasonal efficiency of the heating control sub-system	$\eta_{H;c}$	–	string	D, P, C	Space heating energy service = “ TRUE” $\eta_{H;c} = (0.0; 1.0]$
	n-C	Mean seasonal efficiency of the heating emission sub-system	$\eta_{H;e}$	–	string	D, P, C	Space heating energy service = “ TRUE” $\eta_{H;e} = (0.0; 1.0]$
	n-C	PV system typology			string	D, C	not null if $W_{PV} > 0$
	n-C	Installed PV system power	W_{PV}	kW	number	D, C	$W_{PV} > 0$ (if field populated)
	n-C	Thermal solar system typology			string	D, C	not null if $W_{sol} > 0$
	n-C	Area of solar collectors	A_{coll}	m^2	number	D, C	$A_{coll} \geq 0$ (if field populated)

Field	Critical (C) or non-critical (n-C)	EPC data	Symbol	Unit	Data type	Typology of rule(s) (*)	Rule
Energy data	C	Energy need for space heating	$EP_{H;nd}$	$kWh \cdot m^{-2}$	number	D, P, C	Space heating energy service = "TRUE" $EP_{H;nd} > 0$
	n-C	Energy need for domestic hot water	$EP_{W;nd}$	$kWh \cdot m^{-2}$	number	D, P, C	Domestic hot water energy service = "TRUE" $EP_{W;nd} > 0$
	n-C	Energy need for space cooling	$EP_{C;nd}$	$kWh \cdot m^{-2}$	number	D, P, C	Space cooling energy service = "TRUE" $EP_{C;nd} > 0$
	n-C	Non-renewable energy performance per space heating	$EP_{H;nren}$	$kWh \cdot m^{-2}$	number	D, P, C	Space heating energy service = "TRUE" $EP_{H;nren} > 0$
	n-C	Non-renewable energy performance per domestic hot water	$EP_{W;nren}$	$kWh \cdot m^{-2}$	number	D, P, C	Domestic hot water energy service = "TRUE" $EP_{W;nren} > 0$
	n-C	Non-renewable energy performance per space cooling	$EP_{C;nren}$	$kWh \cdot m^{-2}$	number	D, P, C	Space cooling energy service = "TRUE" $EP_{C;nren} > 0$
	n-C	Overall renewable energy performance indicator	$EP_{gl;ren}$	$kWh \cdot m^{-2}$	number	D, P	$EP_{gl;ren} \geq 0$
	C	Overall non-renewable energy performance indicator	$EP_{gl;nren}$	$kWh \cdot m^{-2}$	number	D, P	$EP_{gl;nren} \geq 0$
Delivered and produced energy	n-C	Delivered natural gas	$Q_{X;Y;in}$	Nm^3 Sm^3	number	D, P, C	$Q_{X;Y;in} > 0$ Energy carrier = "Natural gas" Unit = "Nm^3" OR "Sm^3"
	n-C	Delivered electricity	$E_{X;Y;in}$	kWh	number	D, P, C	Energy carrier = "Electricity" Unit = "kWh"

Field	Critical (C) or non-critical (n-C)	EPC data	Symbol	Unit	Data type	Typology of rule(s) [(*)]	Rule
	n-C	Delivered district heating	$Q_{X;Y;in}$	kWh	number	D, P, C	$Q_{X;Y;in} > 0$ Energy carrier = " District heating" Unit = " kWh"
	n-C	Delivered solid biomass	$Q_{X;Y;in}$	kg	number	D, P, C	$Q_{X;Y;in} > 0$ Energy carrier = " Solid biomass" Unit = " kg"
	n-C	Electrical energy produced by PV system	$E_{X;Y;in}$	kWh	number	D, P, C	$E_{X;Y;in} > 0$ Energy carrier = " Photovoltaic" Unit = " kWh"
	n-C	Thermal energy produced by solar system	$Q_{X;Y;in}$	kWh	number	D, P, C	$Q_{X;Y;in} > 0$ Energy carrier = " Solar thermal" Unit = " kWh"

(*) D = data types of checks; P = physical impossibility checks; C = consistency checks.
(**) The Italian territory is divided into climatic zones based on heating degree day (*HDD*): A, B, C, D, E, and F. Piedmont and Aosta Valley regions belong to climatic zones E and F.
(***) In residential buildings (" E1.1" and " E1.2"), the domestic hot water system must always be modelled, using either the actual system or, if absent, a fictious one

Appendix B – UBEM input data classification

UBEM input data classification

Table B.1. *General information*

Description	Symbol	Unit	Reference standard
Building height	H	m	-
Latitude of the building	ϕ	°	-
Longitude of the building	ψ	°	-
No. of building units	-	–	-
No. of storeys	-	–	-
Storey height	-	m	-

Table B.2. *Thermal zone characteristics*

Description	Symbol	Unit	Reference standard
Usage period for domestic hot water (in the calculation interval)	$t_{W;use}$	h *or* d	EN 15316-1 (CEN, 2017h)
Usage period for heating (in the calculation interval)	$t_{H;use}$	h *or* d	EN 15316-1 (CEN, 2017h)
Useful floor area per thermal zone	A_{use}	m^2	-
Value of areal thermal capacity of air and furniture, per thermally conditioned zone	$\kappa_{m;int}$	$J/(m^2 \cdot K)$	EN ISO 52016-1 (CEN, 2017a)

Table B.3. *Building envelope characteristics*

Description	Symbol	Unit	Reference standard
Thickness of the material layer in the component	d	m	-
Design thermal conductivity	λ	$W/(m \cdot K)$	EN ISO 6946 (CEN, 2007a)
External surface resistance	R_{se}	$(m^2 \cdot K)/W$	EN ISO 6946 (CEN, 2007a)
Internal surface resistance	R_{si}	$(m^2 \cdot K)/W$	EN ISO 6946 (CEN, 2007a)
Hemispherical emissivity of the surface	ε	–	EN ISO 6946 (CEN, 2007a)
Density of building component	ρ_m	kg/m^3	EN ISO 13786 (CEN, 2018b)
Specific heat capacity of building component	c_m	$J/(kg \cdot K)$	EN ISO 13786 (CEN, 2018b)
Solar absorption coefficient of external opaque surfaces	$a_{sol;k}$	–	EN ISO 52016-1 (CEN, 2017a)
Visible absorption coefficient of external opaque surfaces		–	-

Description	Symbol	Unit	Reference standard
Water vapour resistance factor	μ	–	EN ISO 10456 (CEN, 2007b)
Roughness of building component		–	ISO 4287 (ISO, 1997)
Thickness of the glass		m	-
Thermal conductivity of glass	λ	W/(m·K)	EN 673 (CEN, 2011b)
Thermal transmittance of window	U_{W}	$\mathrm{W/(m^2{\cdot}K)}$	EN ISO 10077-1 (CEN, 2020b)
Total solar energy transmittance of the transparent part of window *wi*	$g_{\mathrm{gl};wi}$	–	EN ISO 52016-1 (CEN, 2017a)
Solar direct reflectance of the side of the glazing facing the incident radiation	ρ_{es}	–	EN ISO 52022-1 (CEN, 2017j)
Solar direct reflectance of the side of the glazing facing away from the incident radiation	ρ'_{es}	–	EN ISO 52022-1 (CEN, 2017j)
Solar direct transmittance	τ_{es}	–	EN ISO 52022-1 (CEN, 2017j)
Light reflectance of the side of the glazing facing the incident radiation	ρ_v	–	EN 410 (CEN, 2011a)
Light reflectance of the side of the glazing facing away from incident radiation	ρ'_v	–	EN ISO 52022-1 (CEN, 2017j)
Area of window element *wi*	A_{wi}	$\mathrm{m^2}$	EN ISO 52016-1 (CEN, 2017a)
Glazed area of window element *wi*	$A_{\mathrm{gl};wi}$	$\mathrm{m^2}$	EN ISO 52016-1 (CEN, 2017a)
Frame area fraction of window *wi*, ratio of the projected frame area to the overall projected area of the glazed element of window *wi*	$F_{\mathrm{fr},wi}$	–	EN ISO 52016-1 (CEN, 2017a)
Thermal transmittance of the frame	U_{f}	$\mathrm{W/(m^2{\cdot}K)}$	EN ISO 10077-1 (CEN, 2020b)
Thermal transmittance of the glazing	U_{g}	$\mathrm{W/(m^2{\cdot}K)}$	EN ISO 10077-1 (CEN, 2020b)
Tilt angle of the window *wi* (from horizontal, measured upwards facing)	β_{wi}	°	EN ISO 52016-1 (CEN, 2017a)
Orientation angle of the window *wi*, obtained from the geometric data of the construction element, in degrees (expressed as the geographical azimuth angle of the horizontal projection of the inclined surface normal	g_{wi}	°	EN ISO 52016-1 (CEN, 2017a)
Dimensionless shading reduction factor for external obstacles	$F_{\mathrm{sh;obst}}$	–	EN ISO 52016-1 (CEN, 2017a)

Description	Symbol	Unit	Reference standard
Emissivity of the inside surface of the glass	ε_i	–	EN 410 (CEN, 2011a)
Emissivity of the external surface of the glass	ε_e	–	EN 410 (CEN, 2011a)
Solar irradiance limit for operation of solar shading devices		W/m^2	EN ISO 52016-1 (CEN, 2017a)
Total solar energy transmittance of the combination of glazing and shading, when the solar shading is in use	$g_{gl;sh;wi}$	–	EN ISO 52016-1 (CEN, 2017a)
Weighted fraction of the time with the solar shading in use, e.g. as a function of the intensity of incident solar radiation	$f_{sh;with}$	–	EN ISO 52016-1 (CEN, 2017a)

Table B.4. *Occupancy*

Description	Symbol	Unit	Reference standard
Convective heat released by a person	$Q_{sens,cv}$	W	EN 16798-1 (CEN, 2019)
Default occupancy per building use	O_c	$persons/m^2$	EN 16798-1 (CEN, 2019)
Diversity factors for occupants		–	EN 16798-1 (CEN, 2019)
Diversity factors for the appliances		–	EN 16798-1 (CEN, 2019)
Diversity factors for the lighting system		–	EN 16798-1 (CEN, 2019)
Latent heat released by a person	Q_{lat}	W	EN 16798-1 (CEN, 2019)
Radiative heat released by a person	$Q_{sens,rd}$	W	EN 16798-1 (CEN, 2019)
Solar shading devices schedule		–	-
Convective fraction of the internal gains	$f_{int,cv}$	–	EN ISO 52016-1 (CEN, 2017a)
Specific internal heat flow rate due to dissipated heat from appliances	$q_{int;A;t}$	W/m^2	EN ISO 52016-1 (CEN, 2017a)
Specific internal heat flow rate due to metabolic heat from occupants	$q_{int;oc;t}$	W/m^2	EN ISO 52016-1 (CEN, 2017a)
Specific internal heat flow rate due to recoverable losses from lighting	$q_{int;L;t}$	W/m^2	EN ISO 52016-1 (CEN, 2017a)

Table B.5. *Generators – Boiler*

Description	Symbol	Unit	Reference standard
Ambient temperature of the hot water tank	θ_a	°C	EN 12831-3 (CEN, 2017e)
Average hot water temperature in circulation system without operation	$\theta_{W,avg}$	°C	EN 15316-3-3 (CEN, 2007c)
Cold-water temperature (e.g., 10°C)	$\theta_{W;C}$	°C	EN 12831-3 (CEN, 2017e)
Daily DHW use in litre per square meter in a day	$v_{DHW,day}$	$l/(m^2 \cdot d)$	EN 16798-1 (CEN, 2019)
Density of water	ρ_W	kg/m^3	EN 12831-3 (CEN, 2017e)
DHW system efficiency	$\eta_{W;sys}$	–	EN 15316-1 (CEN, 2017h)
Energy loss of the hot water tank at the time *t* (minute)	$Q_{W;sto;t}$	kWh/min	EN 12831-3 (CEN, 2017e)
Generator efficiency at full load	$\eta_{gen;Pn}$	–	EN 15316-4-1 (CEN, 2017b)
Generator output at full load	P_n	kW	EN 15316-4-1 (CEN, 2017b)
Heating generator Fuel type *i*	*H_GEN_FUEL_i*	–	EN 15316-1 (CEN, 2017h)
Heating system efficiency	$\eta_{H,m,y}$	–	EN 15316-1 (CEN, 2017h)
Maximum temperature of the water stored in the hot water tank according to design specifications	$\theta_{W;sto;max}$	°C	EN 12831-3 (CEN, 2017e)
Minimum temperature of the mixed water drawn at the tap (needs temperature)	$\theta_{W;draw}$	°C	EN 12831-3 (CEN, 2017e)
Required hot water temperature	$\theta\_{W;em}$	°C	EN 15316-1 (CEN, 2017h)
Seasonal Space Heating Energy Efficiency	$\eta_{s,h}$	–	EN 14825 (CEN, 2018a)
Size of the hot water tank (internal volume)	V_{sto}	l	EN 12831-3 (CEN, 2017e)
Specific thermal capacity of water	c_W	$kJ/(kg \cdot K)$	EN 12831-3 (CEN, 2017e)
Standby loss of the hot water tank per day (specified by manufacturer)	$q_{sb;sto}$	kWh/d	EN 12831-3 (CEN, 2017e)
Maximum flow rate of the heating fluid		m^3/h	-

Table B.6. *Generators – Chiller*

Description	Symbol	Unit	Reference standard
Energy Efficiency Ratio	*EER*	kWh/kWh	EN 14825 (CEN, 2018a)
Fuel type	*HP_FUEL*	–	EN 15316-4-2 (CEN, 2017c)
Maximum flow rate of the cooling fluid		m^3/h	-
Seasonal Energy Efficiency Ratio	*SEER*	kWh/kWh	EN 14825 (CEN, 2018a)
Seasonal Space Cooling Energy Efficiency	$\eta_{s,c}$	–	EN 14825 (CEN, 2018a)

Table B.7. *Generators – District Heating*

Description	Symbol	Unit	Reference standard
Annual mean external temperature	$\theta_{e,m}$	°C	EN 12831-1 (CEN, 2017d)
Circulation pump efficiency	η	–	-
Depth of pipe from surface	z	m	EN 15316-3 (CEN, 2007c)
Distance between parallel pipes		m	-
Input temperature of the heating circuit	$\theta_{H,in}$	°C	EN 15316-3 (CEN, 2007c)
Length of the pipe in the zone	L	m	EN 15316-3 (CEN, 2007c)
Outer diameter of the pipe	d_a	m	EN 15316-3 (CEN, 2007c)
Thermal conductivity of insulation	λ_D	W/(m·K)	EN 15316-3 (CEN, 2007c)
Thermal conductivity of the ground	λ_g	W/(m·K)	EN ISO 13370 (CEN, 2017l)

Table B.8. *Generators – Heat Pump*

Description	Symbol	Unit	Reference standard
Coefficient of Performance	*COP*	kWh/kWh	EN 14825 (CEN, 2018a)
Energy Efficiency Ratio	*EER*	kWh/kWh	EN 14825 (CEN, 2018a)
Fuel type	*HP_FUEL*	–	EN 15316-4-2 (CEN, 2017c)
Maximum flow rate of the cooling fluid		m^3/h	-

Description	Symbol	Unit	Reference standard
Maximum flow rate of the heating fluid		m^3/h	-
Seasonal Coefficient of Performance	*SCOP*	kWh/kWh	EN 14825 (CEN, 2018a)
Seasonal Energy Efficiency Ratio	*SEER*	kWh/kWh	EN 14825 (CEN, 2018a)

Table B.9. *Generators – Photovoltaic*

Description	Symbol	Unit	Reference standard
Azimuth angle of PV modules	α	°	EN 15316-4-3 (CEN, 2017i)
Number of PV modules	N	–	EN 15316-4-3 (CEN, 2017i)
Open-circuit voltage	V_{0C}	V	EN IEC 60904-1 (CEN, 2020a)
Peak power	P_{pk}	kW	EN 15316-4-3 (CEN, 2017i)
Tilt angle of the PV modules	β	°	EN 15316-4-3 (CEN, 2017i)
Total area of PV modules (without frame)	A	m^2	EN 15316-4-3 (CEN, 2017i)

Table B.10. *Generators – Solar Thermal*

Description	Symbol	Unit	Reference standard
Collector azimuth of the collector	$\alpha_{sol;ori}$	°	EN 15316-4-3 (CEN, 2017i)
Collector heat loss coefficient	a_1	$W/(m^2 \cdot K)$	EN 15316-4-3 (CEN, 2017i)
Collector reference area	$A_{sol;mod}$	m^2	EN 15316-4-3 (CEN, 2017i)
Number of collector modules installed	N_{col}	–	EN 15316-4-3 (CEN, 2017i)
Peak collector efficiency	η_0	–	EN 15316-4-3 (CEN, 2017i)
Solar irradiance on the collector plane	I_{sol}	W/m^2	EN 15316-4-3 (CEN, 2017i)
Temperature dependent collector heat loss coefficient	a_2	$W/(m^2 \cdot K)$	EN 15316-4-3 (CEN, 2017i)
Tilt angle of the collector	$\alpha_{sol;tilt}$	°	EN 15316-4-3 (CEN, 2017i)

Table B.11. *Generators – Ventilation*

Description	Symbol	Unit	Reference standard
Heat recovery control options	*HEAT_REC_TYPE*	–	EN 16798-5-1 (CEN, 2017g)
Latent efficiency of heat recovery		–	-
Maximum external temperature for the natural ventilation system		°C	-
Maximum relative humidity for the natural ventilation system		–	EN 16798-1 (CEN, 2019)
Minimum external temperature for the natural ventilation system		°C	-
Hourly schedule of the mechanical ventilation system		–	-
Set-point temperature of air mechanically supplied into the thermal zone		°C	-
Nominal heat recovery temperature efficiency at design air velocity	$\eta_{\mathrm{hr;nom}}$	–	EN 16798-5-1 (CEN, 2017g)
Nominal humidity recovery efficiency at design air velocity	$\eta_{\mathrm{xr;nom}}$	–	EN 16798-5-1 (CEN, 2017g)
Sensible efficiency of heat recovery		–	-
Usage period for ventilation (in the calculation interval)	$t_{\mathrm{V;use}}$	h *or* d	EN 15316-4-1 (CEN, 2017b)
Ventilation flow rate for building materials	q_{B}	l/s·(m^2)	EN 16798-1 (CEN, 2019)
Ventilation flow rate for persons	q_{P}	l/s·(person)	EN 16798-1 (CEN, 2019)
Ventilation system operation	*VENT_SYS_OP*	–	EN 16798-7 (CEN, 2017k)

Appendix C – Building prototype schema

Building prototype
E_RES_BU(AB)_CP5

PIEDMONT REGION EPC DATABASE - E_RES_BU(AB)_CP5						
	Data	**Symbol**	**Unit of measure**	**Median**	**($Q_3 - Q_2$)**	**($Q_2 - Q_1$)**
Geometry	Compactness ratio	CR	m^{-1}	0.433	0.235	0.125
	Thermally heated gross volume	$V_{H;g}$	m^3	278	72	57
	Thermally heated floor area	$A_{H;use}$	m^2	72	19	15
	Transparent thermal envelope area on thermal envelope area	A_w/A_{env}	–	11 %	5 %	4 %
Envelope	Mean thermal transmittance of opaque building envelope	U_{op}	$W/(m^2{\cdot}K)$	1.157	0.215	0.228
	Mean thermal transmittance of transparent building envelope	U_w	$W/(m^2{\cdot}K)$	3.157	1.416	0.849
Technical building system	Energy carrier per space heating	Natural gas = 78 %; district heating = 15 %; others = 7 % (of the analysed sample)				
	Energy carrier per space cooling	Electricity = 99 %; natural gas = 1 % (of the analysed sample)				
	Energy carrier per domestic hot water	Natural gas = 74 %; electricity = 21 %; others = 5 % (of the analysed sample)				
	Mean seasonal efficiency of the heating generation sub-system (natural gas)	$\eta_{H;gen}$	–	0.926	0.126	0.136
	Mean seasonal efficiency of the heating generation sub-system (district heating)	$\eta_{H;gen}$	–	0.960	0.040	0.263
	Overall energy efficiency (without generation)	$\eta_{H;ngen}$	–	0.878	0.046	0.050
Energy indicators	Energy need for space heating	$EP_{H;nd;ztc}$	kWh/m^2	95.8	51.2	33.2
	Energy need for space cooling	$EP_{C;nd;ztc}$	kWh/m^2	16.0	10.8	8.4
	Energy need for domestic hot water	$EP_{W;nd;ztc}$	kWh/m^2	18.5	1.3	1.2
	Seasonal space heating energy efficiency	$\eta_{s;H}$	–	0.730	0.070	0.120
	Seasonal space cooling energy efficiency	$\eta_{s;C}$	–	1.000	0.340	0.340
	Seasonal domestic hot water energy efficiency	$\eta_{s;W}$	–	0.630	0.110	0.270
	Non-renewable energy performance for space heating	$EP_{H;nren}$	kWh/m^2	134.5	69.3	45.6
	Non-renewable energy performance for space cooling	$EP_{C;nren}$	kWh/m^2	13.5	11.6	7.2
	Non-renewable energy performance for domestic hot water	$EP_{W;nren}$	kWh/m^2	29.1	18.4	5.2
	Overall non-renewable energy performance	$EP_{gl;nren}$	kWh/m^2	171.8	69.7	47.1
	Overall renewable energy performance	$EP_{gl;ren}$	kWh/m^2	1.3	10.3	1.1
	Renewable Energy Ratio	RER	–	1 %	4 %	1 %

Appendix D – Building archetype schema

Building archetype
RES_APPBLOCK_1962-1971_E-F_VAL

Region:	Aosta Valley			**Archetype code:** RES_APPBLOCK_1962-1971_E-F_VAL
Building category:	Residential buildings - Apartments (in multifamily blocks)			
Period of construction:	1962 - 1971			
Climatic zone:	E-F	**Number of records:**	2450	

Description (the codes associated with walls and slabs refer to the structures described in UNI/TR 11552:2014):

External walls: hollow brick masonry with air gap (cod. MCV01).
Roof slabs: reinforced concrete floor slab (cod. SOL04).

Data sources: EPC databases (100%)

	Data	Symbol	Unit of measure	Mean value	Standard deviation	Q1 (first quartile)	Median value	Q3 (third quartile)
BUILDING GEOMETRY	Number of floors	n_f	–	-	-	-	-	-
	Gross height	H_g	m	-	-	-	-	-
	Footprint area	$A_{footprint}$	m²	-	-	-	-	-
	Heated gross floor area	$A_{H;g}$	m²	-	-	-	-	-
	Heated net floor area	$A_{H;n}$	m²	-	-	-	-	-
	Heated gross volume	$V_{H;g}$	m³	-	-	-	-	-
	Heated net volume	$V_{H;n}$	m³	-	-	-	-	-
	Compactness ratio	$A_{env}/V_{H;g}$	m^{-1}	0.63	0.24	0.43	0.66	0.79
	WWR – North orientation	WWR_N	–	0.13	0.06	0.10	0.12	0.16
	WWR – South orientation	WWR_S	–	0.13	0.06	0.10	0.12	0.16
	WWR – East orientation	WWR_E	–	0.13	0.06	0.10	0.12	0.16
	WWR – West orientation	WWR_W	–	0.13	0.06	0.10	0.12	0.16
	Window to useful floor area ratio	A_{wi}/A_{use}	–	0.18	0.07	0.14	0.17	0.21
ENVELOPE	Roof type	-						
	U-value of the roof **	$U_{fl;up}$	W/(m²·K)	1.26	0.47	1.05	1.32	1.46
	External walls type	Hollow brick masonry: 58 %; Solid Brick masonry: 29 %; Masonry with local stones: 6 %; Concrete wall: 4 %; Unknown: 3 %						
	U-value of the wall	U_{wl}	W/(m²·K)	0.99	0.44	0.73	1.01	1.21
	Slab on ground floor type	-						
	U-value of the floor **	$U_{fl;lw}$	W/(m²·K)	1.07	0.26	0.99	1.11	1.19
	Windows type	Double glazing, wooden frame: 67 %; Single glazing, wooden frame: 22 %; Double glazing, PVC frame: 9 %; Triple glazing, wooden frame: 1 %; Triple glazing, PVC frame: 1 %						
	U-value of the windows	U_W	W/(m²·K)	2.74	0.99	2.08	2.67	3.09
	Shading system type	-						
GAINS and VENTILATION	Occupancy density *	O_C	person/m²	UNI EN 16798-1 - Table A.19				
	Lighting power density *	W_L	W/m²	UNI EN 16798-1 - A.8.3				
	Equipment power density *	W_A	W/m²	UNI EN 16798-1 - A.8.3				
	Type of ventilation	Natural: 100 %						
	Air exchange rate *	n	h^{-1}	0.30	0.00	0.30	0.30	0.30
THERMAL SYSTEMS	Heating system type	Centralized: 54%; Autonomous: 46 %						
	Heating generator	Boiler (unknown type): 50 %; Traditional Boiler: 29 %; Condensing Boiler: 9 %; Heat exchanger of district heating/cooling: 5 %; Fireplace: 4 %; Unknown: 2 %; Air-source heat pump: 1 %						
	Daily operating time of the heating system *	t_H	h	-				
	Energy carrier	Gas Oil: 56 %; Natural Gas: 19 %; LPG: 13 %; Solid biomass: 8 %; District heating: 4 %						
	Heating emission sub-system	-						
	Cooling system type	Absent: 100 %						
	Daily operating time of the cooling system *	t_C	h	-	-	-	-	-
	Cooling emission sub-system	-						
	DHW system type	Autonomous, detached from heating: 49 %; Autonomous, coupled with heating: 31 %; Centralized, coupled with heating: 19 %; Centralized, detached from heating: 1 %						
	DHW generator	Unknown: 61 %; Natural gas boiler: 22 %; Electric boiler: 16 %; Electric Heat Pump: 1 %						

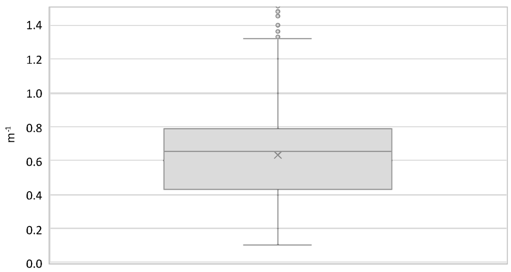
* These values are derived from UNI EN ISO Standards; ** *U*-values of the upper and lower slabs face unconditioned spaces (i.e., attic, basement, etc.)

Residential buildings – Apartments – 1962-1971 – Zone E-F – Aosta Valley

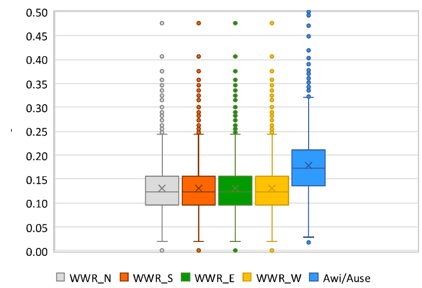
Region:	Aosta Valley			Archetype code:
Building category:	Residential buildings - Apartments (in multifamily blocks)			RES_APPBLOCK_1962-1971_E-F_VAL
Period of construction:	1962 - 1971			
Climatic zone:	E-F	Number of records:	2450	

Numerical variables – GEOMETRY

COMPACTNESS RATIO

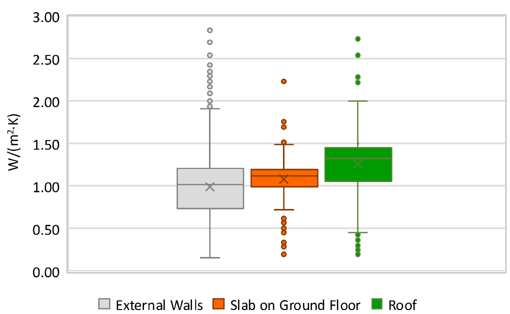


WINDOWS TO WALL RATIO

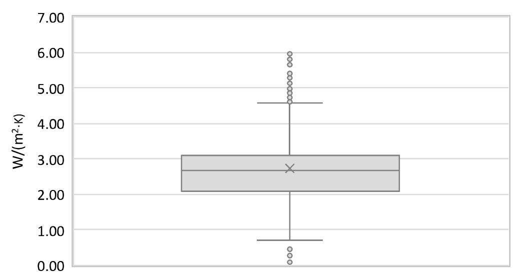


Numerical variables – ENVELOPE

OPAQUE BUILDING COMPONENTS U-VALUE

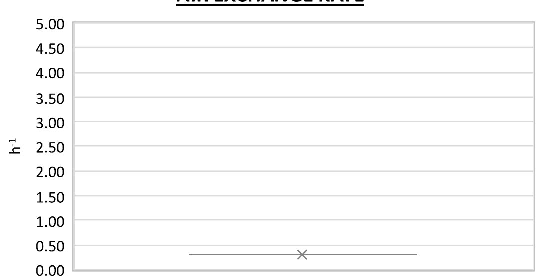


WINDOWS U-VALUE

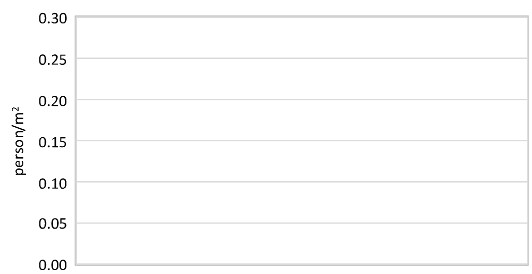


Numerical variables – GAINS, VENTILATION and SYSTEMS USAGE

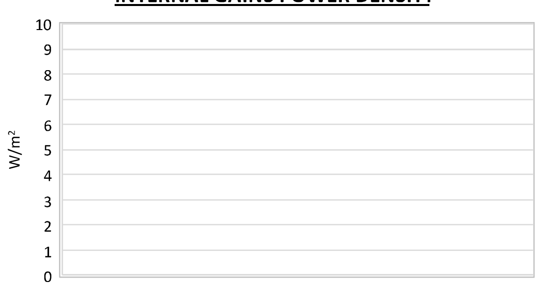
AIR EXCHANGE RATE



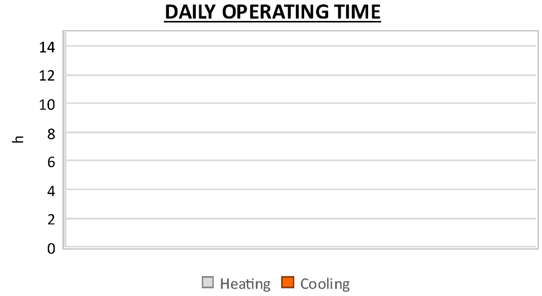
OCCUPANCY DENSITY



INTERNAL GAINS POWER DENSITY



DAILY OPERATING TIME



Residential buildings – Apartments – 1962-1971 – Zone E-F – Aosta Valley

Region:	Aosta Valley			**Archetype code:** RES_APPBLOCK_1962-1971_E-F_VAL
Building category:	Residential buildings - Apartments (in multifamily blocks)			
Period of construction:	1962 - 1971			
Climatic zone:	E-F	**Number of records:**	2450	

ADDITIONAL DATA

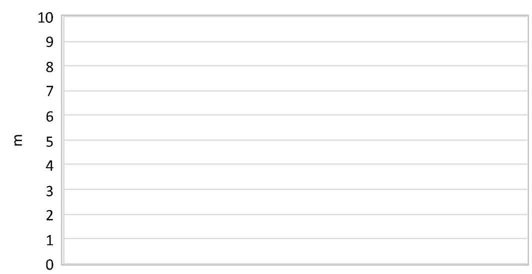
	Data	Symbol	Unit of measure	Mean value	Standard deviation	Q1 (first quartile)	Median value	Q3 (third quartile)
GEOMETRY: apartments	Inter-storey height	H_n	m	2.6	0.3	2.4	2.5	2.7
	Heated gross floor area	$A_{H;g}$	m²	-	-	-	-	-
	Heated net floor area	$A_{H;n}$	m²	64.8	30.3	44.0	60.5	79.0
	Heated gross volume	$V_{H;g}$	m³	237.2	115.5	160.0	218.2	289.2
	Heated net volume	$V_{H;n}$	m³	162.6	76.5	111.6	150.7	200.0
THERMAL SYSTEMS	Heating efficiency or *COP*	$\eta_{H;gen}$ or $COP_{H;gen}$	–	This value has to be retrieved from suitable datasheets				
	Total heating power *	$P_{H;gen}$	kW	22.8	8.5	19.5	24.1	29.0
	Cooling efficiency or *EER*	$\eta_{C;gen}$ or $EER_{C;gen}$	–	This value has to be retrieved from suitable datasheets				
	Total cooling power *	$P_{C;gen}$	kW	3.9	2.1	2.5	3.0	4.4
	Temperature of DHW	ϑ_W	°C	40.0	0.0	40.0	40.0	40.0
	DHW system power *	$P_{W;gen}$	kW	11.8	12.4	1.2	2.0	24.0

* These values refer to the apartment scale

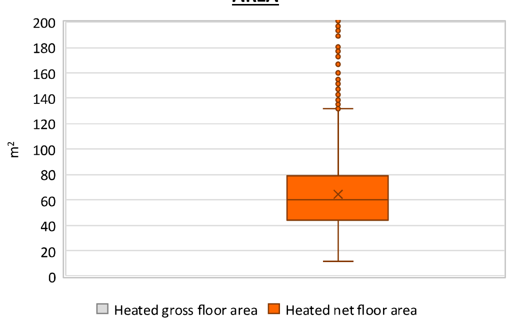
Additional data: GEOMETRY (the plots refer to the apartment scale)

NUMBER OF FLOORS

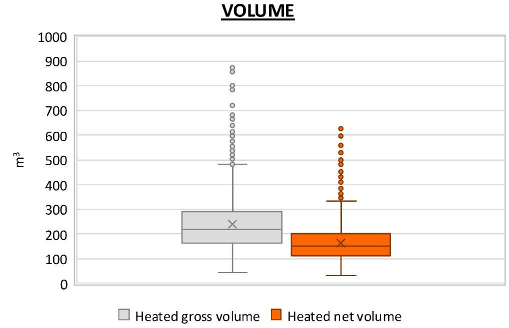
GROSS HEIGHT



AREA



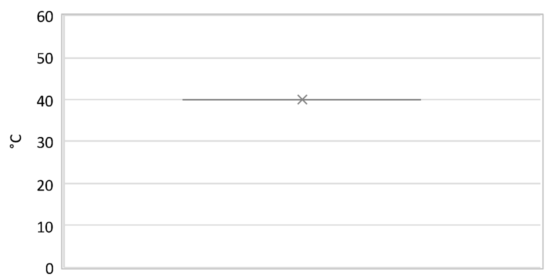
VOLUME



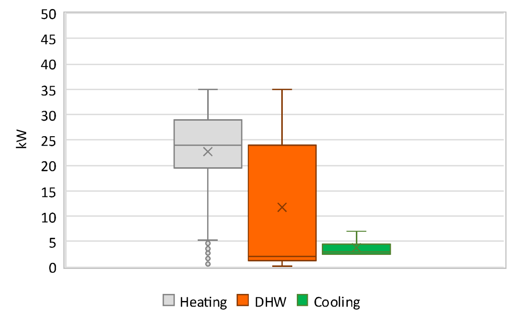
Region:	Aosta Valley			**Archetype code:** RES_APPBLOCK_1962-1971_E-F_VAL
Building category:	Residential buildings - Apartments (in multifamily blocks)			
Period of construction:	1962 - 1971			
Climatic zone:	E-F	**Number of records:**	2450	

Additional data: other numerical variables that are not included in the archetype

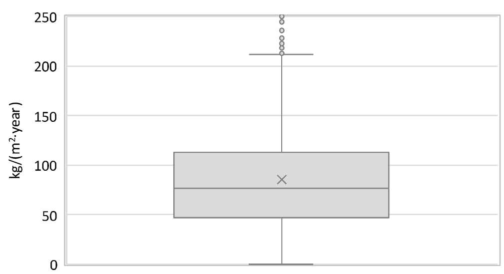
DHW SUPPLY TEMPERATURE



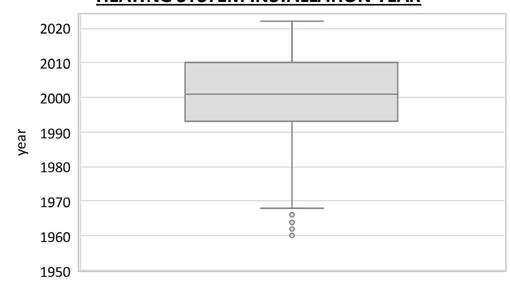
SYSTEM POWER



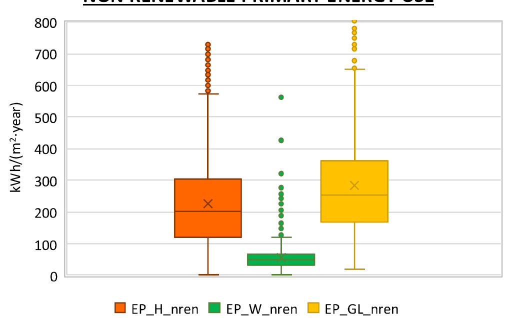
CO_2 EMISSION



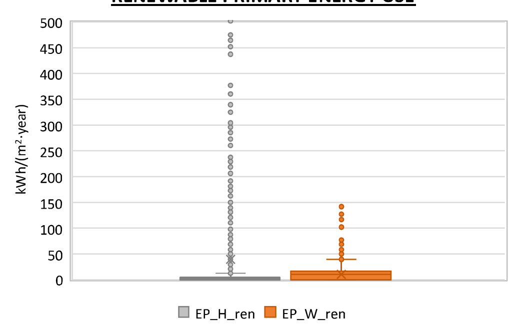
HEATING SYSTEM INSTALLATION YEAR



NON-RENEWABLE PRIMARY ENERGY USE



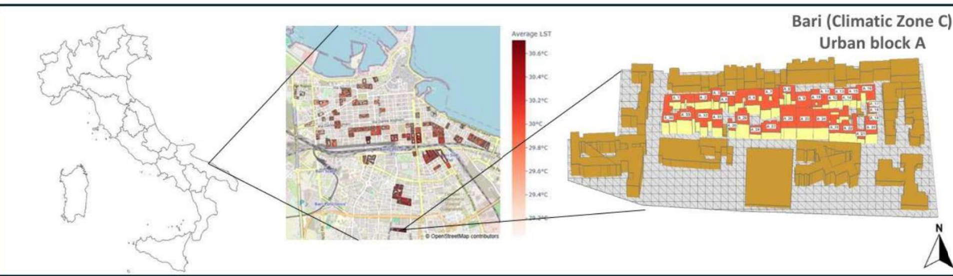
RENEWABLE PRIMARY ENERGY USE



Appendix E – Representative urban block schema

Representative urban block
Bari – Urban Block A

URBAN DATA



Main Urban Metrics for Block A		Unit
No. of assessed buildings in the block	34	–
Surface Coverage	0.424	–
Average Building Height	18.7	m
Compactness Ratio median	0.288	m^{-1}
Green Ratio	0.00	–

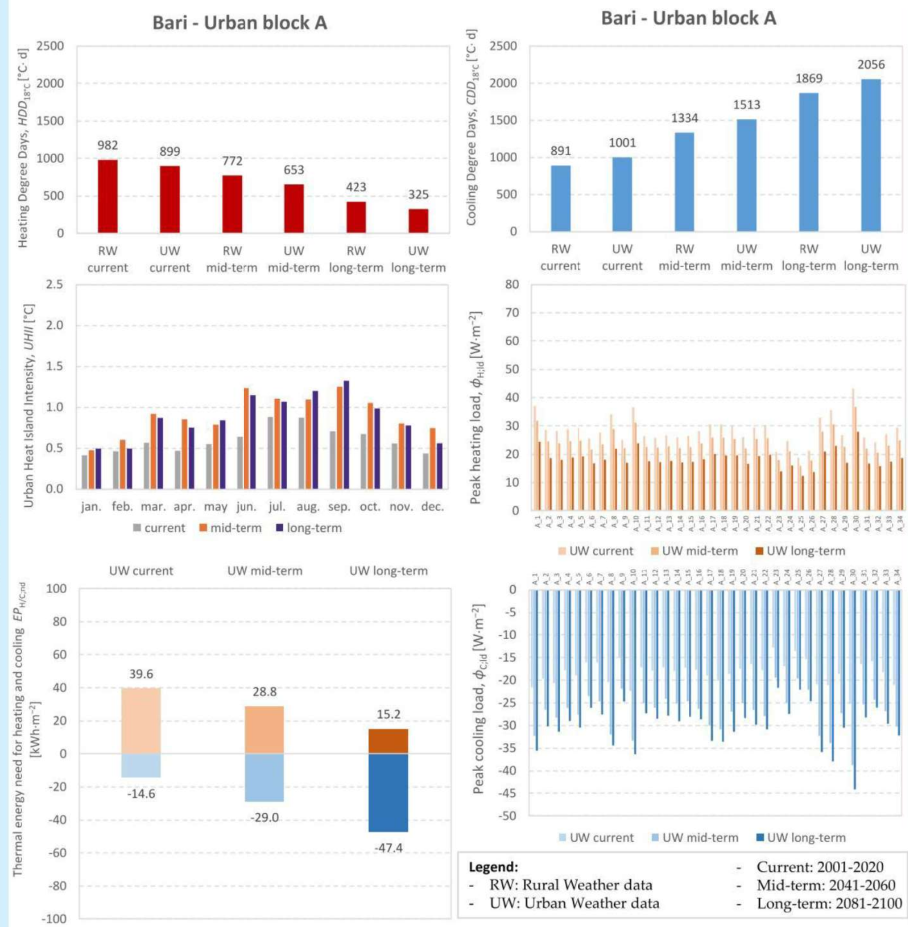
BUILDING DATA

Building data for Urban block A

Bldg. code	Constr. period	A_{fl} [m²]	A_{env} [m²]	V [m³]	WWR [–]	$A_{env} \cdot V^{-1}$ [m⁻¹]	U_{op}^{*} [W·m⁻²·K⁻¹]	U_{wi}^{*} [W·m⁻²·K⁻¹]
A_1		2379	3353	8922		0.376	0.93	
A_2		1396	1436	4655		0.308	0.92	
A_3		1511	1474	5395		0.273	0.92	
A_4		497	533	1865		0.286	0.94	
A_5		593	660	2225		0.297	0.93	
A_6		3622	3261	12075		0.270	0.93	
A_7		3801	3756	12670		0.296	0.92	
A_8		2178	2662	7778		0.342	0.91	
A_9		1417	1453	4722		0.308	0.95	
A_10		1221	1726	4580		0.377	0.93	
A_11		738	702	2768		0.254	0.94	
A_12		671	647	2515		0.257	0.94	
A_13		562	545	2107		0.258	0.94	
A_14		599	579	2247		0.258	0.94	
A_15		544	523	2039		0.257	0.94	
A_16		2519	2570	8398		0.306	0.92	
A_17	1976-90	402	472	1506	18 %	0.313	0.93	3.70
A_18		612	685	2040		0.336	0.92	
A_19		589	648	2208		0.293	0.94	
A_20		1645	1528	5485		0.279	0.93	
A_21		360	441	1200		0.367	0.94	
A_22		2389	2615	7964		0.328	0.92	
A_23		1731	1209	5771		0.209	0.94	
A_24		5061	4158	18074		0.230	0.93	
A_25		3938	2254	14065		0.160	0.94	
A_26		4338	2917	15493		0.188	0.93	
A_27		1428	1762	5355		0.329	0.93	
A_28		532	791	1774		0.446	0.93	
A_29		3024	2700	10081		0.268	0.92	
A_30		490	822	1836		0.448	0.92	
A_31		1966	1799	6554		0.274	0.93	
A_32		2446	2058	8153		0.252	0.93	
A_33		3039	2777	10131		0.274	0.92	
A_34		1564	1623	5586		0.291	0.92	

*Values were derived from URBEM database (https://www.urbem.polimi.it/databasedifici/)

CLIMATIC DATA & KPIs



Legend:
- RW: Rural Weather data
- UW: Urban Weather data
- Current: 2001-2020
- Mid-term: 2041-2060
- Long-term: 2081-2100