Smart and
Sustainable Built
Environment

Deep learning analysis on energy efficiency impacts on sustainable built environments: a Re-eVAM approach

Aurora Greta Ruggeri
Dipartimento di Architettura, Università degli Studi di Ferrara, Ferrara, Italy
Rubina Canesi and Giuliano Marella
ICEA – Dipartimento di Ingegneria Civile Edile e Ambientale,
Universita degli Studi di Padova, Padova, Italy
Laura Gabrielli
Dipartimento di Architettura, Università degli Studi di Ferrara, Ferrara, Italy, and
Massimiliano Scarpa
Faculty of Architecture, Iuav University of Venice, Venice, Italy

Received 29 October 2025
Revised 3 March 2026
22 April 2026
Accepted 20 May 2026

Abstract

Purpose – This study investigates the impact of energy efficiency (EE) improvements on residential property values in Italy, aligning with the European Union’s energy performance buildings directive IV (EPBD IV). By leveraging advanced AI and deep learning techniques, the research aims to quantify the relationship between energy performance certificates and market values, contributing to the development of smart and sustainable built environments. The study provides crucial insights for professionals, investors, and policymakers, supporting the transition towards more sustainable urban development and climate-responsive architecture.

Design/methodology/approach – The research employs a novel four-phase methodology, integrating big data analytics with deep learning techniques to address the complex challenges of sustainable built environments. It begins with comprehensive data mining of Italian Real Estate listings, followed by statistical analysis. A Residential building energy-efficient Valuation Model (Re-eVAM) is developed using deep artificial neural networks to forecast property values based on energy performances. Finally, scenario modelling evaluates potential market impacts under various EPBD IV implementation strategies, contributing to data-driven optimization for sustainable regulations and governance.

Findings – The study reveals significant potential impacts of EE improvements on RE values in the Italian market. In the worst-case scenario, insufficient implementation of energy retrofits could potentially devalue up to 74% of the overall RE market. Conversely, comprehensive EE improvements could enhance market value by up to 27% in the best-case scenario, demonstrating the economic viability of sustainable building practices. These findings underscore the critical role of well-designed policies in creating smart and healthy living environments, and the potential for significant value creation through EE improvements in the built environment.

Originality/value – This research provides unprecedented insights into the Italian residential RE market’s response to EE measures, utilizing a large, previously unexplored database. The innovative Re-eVAM approach offers a new tool for valuing EE in properties, showcasing the potential of machine learning in creating more sustainable and resilient urban environments. By quantifying the potential economic impacts of various EPBD IV implementation scenarios, this study contributes to the development of smart cities and guides sustainable

---

JEL Classification — SDG 7 (Affordable and clean energy), SDG 11 (Sustainable cities and communities)

Funding: This study was funded by the Department of Civil, Environmental and Architectural Engineering (ICEA), University of Padua, through SID-2025 funding (2025ICEA1SIDPROGETTI-00237) under the project SEnergy: Stochastic Life Cycle Cost Assessment for Energy Retrofitting Strategies in Residential Stock. The views and opinions expressed are solely those of the authors.

Smart and Sustainable Built Environment
Emerald Publishing Limited
e-ISSN: 2046-6102
p-ISSN: 2046-6099
DOI 10.1108/SASBE-10-2025-0681

urban investment strategies. The findings demonstrate the importance of data-driven approaches in achieving the UN Sustainable Development Goals in the built environment sector.

Keywords Energy efficiency, Property valuation, Building renovation, Sustainable environment, Deep artificial neural networks (DANN)

Paper type Research article

1. Introduction

The European Union's (EU) ambitious climate and energy targets have positioned the decarbonization of the building sector as a critical objective in realizing future energy-positive cities. This goal is to be achieved through the intelligent interconnection of increasingly energy efficient buildings. Energy Performance Certificates (EPCs) and the emerging Smart Readiness Indicators (SRIs) are expected to play a pivotal role in influencing consumer choices towards sustainable buildings. However, despite the mandatory implementation of EPCs across EU Member States, their impact on local RE markets remains uncertain. The 2024 Energy Performance of Buildings Directive (EPBD IV [1]) introduced some relevant measures. The approved EPBD IV differs significantly from its initial draft and proposal, presenting moderated requirements. This moderation was largely in response to concerns raised by some countries, including Italy, about the economic implications of the directive on Real Estate (RE) market values and retrofit costs. The final directive represents a compromise between the European Commission's ambitious environmental goals and the economic realities faced by member states. This revised approach aims to balance the urgent need for Energy Efficiency (EE) improvements in buildings with the diverse economic circumstances across EU countries.

The Energy Performance of Buildings Directive IV (EPBD IV) introduces comprehensive measures to improve EE in buildings across the European Union. Key targets include increasing the average energy performance of national residential building stock by 16% by 2030 and 20–22% by 2035, compared to 2020 levels. For non-residential buildings, it mandates refurbishment of 16% of the poorest performing structures by 2030, increasing to 26% by 2033 (European Commission, 2024). The directive allows flexibility for member states in implementation, particularly for private buildings. Countries can define their own standards and requirements, with options to target older, larger, or the most energy-intensive buildings. Moderation strategies include counting renovations completed since 2020 towards targets and rewarding proactive member states. Potential exemptions are allowed for certain categories, such as historic and agricultural buildings. EPBD IV also focuses on enhancing long-term renovation strategies and improving the reliability, quality, and digitalization of Energy Performance Certificates (EPCs), based on common criteria. Specific technology measures include postponing the phase-out of fossil-fuelled boilers to 2040, while ceasing subsidies for stand-alone units by 2025. Renewable energy incentives will continue for systems with significant renewable components.

Member States are expected to transpose the Directive into national legislation by 2025. These updated regulations are anticipated to significantly impact building stocks in terms of both EE and market appreciation, reflecting a balanced approach between ambitious goals and practical implementation across diverse national contexts.

Despite moderated targets, the EPBD IV's requirements remain challenging. The low efficiency and age of existing European buildings, along with urban planning constraints, pose significant renovation obstacles. At current rates, renovating the entire EU building stock would take over a century, even with new buildings mandated to be nearly-zero energy (European Commission, 2019). The European Commission estimates that achieving the proposed 55% climate target by 2030 requires an additional annual investment of €275 billion in building renovation, representing one of the largest investment gaps in the EU (European Commission, 2020). This financial burden is particularly significant given the composition of the European building stock (Figure 1), where 51% of the residential sector consists of buildings constructed before 1969, with 31% dating from before 1945 (Gevorgian *et al.*, 2021). These older buildings typically have the highest final energy consumption, with pre-1945 structures consuming an

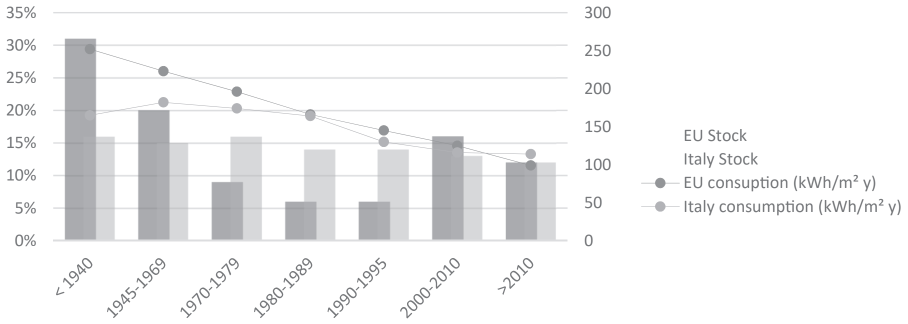

Figure 1. Covered floor area in residential sector per construction periods [%] (EU + UK and Italy), and residential energy consumption for space heating and domestic hot water [kWh/m^2 y] (EU + UK and Italy). Relaboration by the Authors from Eurach Research Report's Data (Gevorgian *et al*., 2021)

average of 252 kWh/m^2 annually for space heating and domestic hot water. In Italy, the situation mirrors the European trend, with 51% of residential buildings constructed before 1969. This ageing building stock presents unique challenges for renovation. The gradual decrease in energy consumption observed in newer buildings, down to 99 kWh/m^2 annually for post-2010 constructions, underscores the potential impact of widespread renovation but also highlights the scale of the challenge ahead. These factors collectively underscore the complexity of implementing the Green Homes Directive across diverse national contexts, each with its own unique building stock characteristics and renovation challenges.

Given the complexity and significant economic implications of EE measures, our study aims to provide a comprehensive, data-driven analysis of how EE could impact RE stock value in Italy, in anticipation of the future implementation of the Energy Performance of Buildings Directive IV (EPBD IV). The study employs a multi-step approach, beginning with the creation of a comprehensive Italian RE database. This is followed by the development of Re-eVAM (Residential building energy-efficient VAluation Model), a Deep Artificial Neural Network designed to forecast property values based on energy performance and other relevant characteristics. The model is applied to evaluate potential market value changes across five distinct scenarios, each representing different EE strategies and implementation levels. This methodology allows for a quantitative assessment of how various energy performance improvements could affect overall RE market values and potentially impact the valuation of buildings with lower Energy Efficiency.

This research contributes to academic discourse while offering actionable insights for stakeholders in the Energy Efficient housing sector. This study contributes to providing the first national-scale predictive assessment of energy performance differentials in the Italian residential market, overcoming the strong geographic fragmentation of previous city-level studies. Secondly, it integrates deep learning–based mass appraisal with regulatory scenario simulation, allowing the estimation of stock-level value reallocations under alternative EPBD IV enforcement intensities. Finally, it frames the model explicitly as a policy stress-testing tool rather than a purely predictive exercise, thereby linking AI-based valuation methods with forward-looking sustainability governance. The findings have, in fact, practical applications in investment decisions, property valuation, and sustainable urban development, aligning with broader energy and climate objectives. In this sense, Re-eVAM is positioned as an operational decision-support instrument for current sustainable built environment governance. Given the evolving landscape of EE in buildings and the lack of comprehensive Italian studies, this timely research provides a crucial understanding of energy performance's economic implications in the Italian residential market. By analyzing potential scenarios and their impacts, we aim to support informed decision-making in enhancing building Energy

Efficiency. Therefore, the originality of this study lies not only in its adoption of a deep learning-based evaluation framework, but also in its combination of country-level market data, nonlinear predictive models, and scenario-based regulatory stress tests, which provide policy-relevant evidence not available in previous studies.

The paper is structured as follows: following this introduction, an exhaustive literature review is presented, focussing on RE premiums related to EE at an international level. The third section describes the adopted methodology, detailing the various phases of the study and the data sampling process. Section 4 presents and discusses the results of our analysis. Finally, the conclusion summarizes the key findings, acknowledges the limitations of the study, and outlines potential future developments in this field of research.

2. Literature review

The intersection of Energy Efficiency, building renovation, and Real Estate economics has become a critical area of study, as it is a determining factor for all stakeholders involved in the decision-making process: policymakers who must design and implement regulations and incentives, and households who need to balance the investment costs of EE strategies against the potential energy bill savings, improved quality and thermal comfort, and the RE value premium they might achieve. However, while the policy debate has increasingly shifted toward large-scale decarbonization of national building stocks, much of the empirical literature remains fragmented and geographically localized, limiting its ability to provide policy-relevant evidence at systemic scale. Recent comprehensive reviews have emphasized that proper performance assessment of residential building renovation requires a holistic approach, integrating not only EE and sustainability metrics but also considerations of thermal comfort, life cycle impacts, and crucially, the evaluation of policy options and financial incentives necessary to overcome the financial constraints that impede renovation decisions (Chen and Lai, 2025).

Numerous strategies (Bandelow *et al.*, 2023; Canesi and Marella, 2023; Kaufmann *et al.*, 2023; Nguyen *et al.*, 2023; Streimikiene *et al.*, 2024), directives (von Malmborg *et al.*, 2023), laws, and incentives (Karytsas and Theodoropoulou, 2023; Liu *et al.*, 2023; Sarcina and Canesi, 2023) are in place to address the renovation of existing building stocks in Europe (Maduta *et al.*, 2023). The multifaceted objectives include mitigating technological obsolescence, enhancing the aesthetic aspects of large urban areas (Ascione *et al.*, 2011; Canesi, 2022; Martínez-Molina *et al.*, 2016; Mazzarella, 2015), and, most importantly, minimizing the climate impact that RE stocks have on the environment in terms of energy consumption and pollutant emissions (Ahmad *et al.*, 2018; Li *et al.*, 2017; Swan and Ugursal, 2009). Despite growing research attention, significant gaps remain in understanding the complex interplay between these factors, particularly regarding the evaluation of policy options, financial incentives, and holistic decision-making models that incorporate residents' intentions, safety, health, and wellbeing alongside traditional economic and environmental indicators (Chen and Lai, 2025). Yet, despite this regulatory proliferation, empirical evidence on how energy performance standards translate into measurable capital value adjustments remains heterogeneous and often inconsistent across markets. This inconsistency complicates the calibration of effective subsidy schemes and undermines the comparability of national transition pathways.

Assessing the EE of a building is a multifaceted procedure that significantly depends on the intended objectives and the relative weight assigned to each goal, requiring balance estimation between EE, cost savings, and overall lifecycle costs. One approach is based on calculating the energy performance of buildings under standardized boundary conditions. In this case, human behaviour (e.g. temperature set-point, internal heat gains intensity and periods), weather conditions and calculation procedures are set by laws, and the result is the so-called Energy Performance Certificate (EPC), which is aimed at comparing different buildings in the same location and with the same intended use. Alternatively, the energy performance can be calculated under custom boundary conditions. In this case, human behaviour, weather conditions and calculation procedures can vary at a large extent. The result is named energy audit, and the aim usually consists in the building's energy retrofit. This can be achieved through stepwise retrofitting, a phased approach that offers flexibility and overcomes financial barriers by implementing improvements in structured stages rather than a single intervention (Hulathdoowage *et al.*, 2025). Regardless of the EE calculation method used, whether standardized for EPCs or customized for energy audits, the majority of methodologies for assessing the feasibility of an energy retrofit campaign involve a comparison between the pre-retrofit and post-retrofit energy consumption. This approach has been extensively investigated by the authors in multiple previous research studies (Gabrielli and Ruggeri, 2022; Pittarello *et al.*, 2021; Ruggeri *et al.*, 2020). Implementing this approach requires that the benefits of energy savings outweigh the costs associated with renovations and other adjustments. In line with this, the EPBD recast and the European Union recommends the Life Cycle Cost (LCC) method for the economic analysis of investments in building energy retrofitting (BER). This method helps identify cost-optimal building energy retrofit projects (BERPs) while ensuring minimum energy performance standards are met (European Union, 2012). This approach underscores the importance of considering not just immediate costs and savings, but also long-term economic implications in EE assessments.

However, traditional LCC and Cost–Benefit approaches predominantly focus on direct financial and environmental metrics, often overlooking capitalization effects in RE markets (Araújo *et al.*, 2016; Liu *et al.*, 2018; Sajid *et al.*, 2026; Zangheri *et al.*, 2022). This omission is particularly critical in contexts where market value appreciation may represent a substantial share of total retrofit benefits. Without incorporating value capitalization, economic feasibility assessments risk underestimating the true return on energy investments and misguiding both private and public decision-making. While environmental benefits (such as CO2 reduction), decreased energy costs, and improvements in inhabitants' quality of life are frequently considered, a crucial factor is often neglected: the increase in RE value resulting from these interventions. This oversight is significant, as the increased market value of a property, typically reflected in an improved Energy Performance Certificate (EPC), represents a substantial economic benefit for property owners (Bottero *et al.*, 2017; Gołąbeska, 2019; Kholodilin *et al.*, 2017; Mecca *et al.*, 2020; Moretti *et al.*, 2019; Surmann *et al.*, 2015; Tagliabue *et al.*, 2019). This value appreciation is not just a theoretical concept but an intangible externality that can significantly influence the overall economic assessment of EE measures. Failing to account for this increased market value can lead to underestimating the total benefits of energy retrofits, potentially skewing decision-making processes for both individual property owners and policymakers. The increase in market value is going to be strongly amplified by the recast of the EPBD IV.

According to recent literature, the relationship between EE and property values has been studied internationally, revealing a generally positive correlation. Within the European context, studies have demonstrated the potential economic benefits of EE in residential properties. Multi-criteria decision analysis demonstrates that EE measures positively affect life cycle costs, carbon emissions, property values, and indoor comfort (Dell'Anna *et al.*, 2020), while net-zero carbon retrofitting research highlights economic and environmental co-benefits (Weerasinghe *et al.*, 2025). De Ayala *et al*. report a 9.8% premium in Spain (De Ayala *et al.*, 2016). However, Galvin documents the prebound effect in Germany, showing discrepancies between EPC ratings and actual energy consumption (Galvin, 2023). Popescu *et al*. find that premiums vary significantly across markets and building types (Popescu *et al.*, 2012). These studies, while collectively supporting the existence of an "energy premium," reveal substantial variability in magnitude, direction, and statistical significance. Reported premiums range from negligible to double-digit percentages, suggesting that contextual, spatial, and methodological differences play a decisive role. Such heterogeneity raises a critical methodological question: are inconsistencies driven by true market variation, or by model specification limits, small samples, and linear econometric assumptions that may fail to capture complex non-linear interactions between location, structural attributes, and energy performance? Research specific to the Italian market remains limited. Studies conducted in

Bolzano (Bisello *et al.*, 2020), Turin (Bottero *et al.*, 2018), and Bari (Manganelli *et al.*, 2019; Morano *et al.*, 2020) confirm positive correlations between Energy Class and price premiums. In the Bolzano housing market, higher Energy Classes were associated with price premiums ranging from approximately 3%–6.5%, depending on property characteristics and model specification. Similar positive effects were observed in Turin and Bari, although the magnitude of the premium varied according to local market structure and spatial context. These findings are broadly consistent with international evidence (De Ayala *et al.*, 2016).

Nevertheless, despite confirming the existence of an energy-related capitalization effect, these contributions remain geographically circumscribed, typically based on relatively small samples and predominantly estimated through linear hedonic regressions applied to single metropolitan areas. The reported variability in premium magnitudes, ranging from modest single-digit percentages to nearly double-digit effects, suggests that local context, model specification, and spatial structure significantly influence results. Therefore, such findings cannot be directly generalized to the national building stock, nor can they adequately inform national-level policy instruments, particularly in the context of evolving European energy directives that require systemic and spatially comprehensive evidence for effective calibration. The absence of large-scale, spatially explicit, and methodologically advanced national analyses represents a major gap in the literature. Energy policy design, particularly in the context of progressive minimum performance standards, requires evidence that reflects systemic market behaviour rather than localized case studies. A national-scale approach is essential to capture spatial heterogeneity, regional income differentials, climatic diversity, and structural variation in building typologies. Without such a perspective, policy prescriptions risk being either inefficiently calibrated or socially regressive. From a methodological standpoint, most existing studies rely on traditional linear or semi-log hedonic models. While these approaches offer interpretability, they implicitly assume linearity, additivity, and stable marginal effects, which may not hold in RE markets characterized by spatial dependence, threshold effects, and non-linear capitalization dynamics. Deep Learning (DL) and Artificial Neural Networks (ANNs) provide a flexible modelling framework capable of capturing high-dimensional, non-linear interactions without imposing restrictive functional forms. However, while machine learning models offer superior predictive performance over traditional hedonic approaches, they face critical challenges related to the "black box" problem, data bias, and algorithmic transparency (Topraklı, 2024), necessitating dedicated testing frameworks to ensure responsible implementation and regulatory compliance (Wan and Lindenthal, 2023). Transparency is increasingly viewed not merely as a technical requirement but as a governance condition for responsible integration of AI systems in property taxation, mortgage underwriting, and sustainable policy evaluation.

Recognizing these methodological tensions, recent advancements in automated valuation models (AVMs) provide structured benchmarks and performance standards for large-scale property valuation systems. Recent contributions offer structured benchmarks and methodological reflections that frame the evolution of automated valuation models (AVMs). Comparative studies evaluating diverse AVM techniques, including stacked generalization combining tree-based algorithms, linear hedonic components, and advanced machine learning models, report out-of-sample predictive errors in the range of MdAPE ≈ 5–6% for optimized ensemble approaches applied to large housing datasets (Bjørgve *et al.*, 2026; Moreno-Foronda *et al.*, 2025). These findings indicate that ensemble and non-linear learning architectures can outperform traditional hedonic regressions in predictive accuracy, particularly when spatial features and validation rigour are properly addressed. In parallel, spatially enhanced AVMs demonstrate that explicit incorporation of geographic information and spatial cross-validation protocols is necessary to prevent overly optimistic accuracy estimates arising from non-spatial resampling strategies (Deppner and Cajias, 2024; Nyanda *et al.*, 2024). Complementing these empirical benchmarks, the recent systematic review synthesizes the AVM literature and identifies machine learning and hybrid ensemble systems as dominant methodological trends, while emphasizing persistent gaps related to spatial dependence treatment, interpretability, benchmarking transparency, and robustness diagnostics (El Jaouhari *et al.*, 2024). Collectively, this body of work underscores the importance of clearly articulated validation frameworks and standardized performance reporting in large-scale valuation systems. Responding directly to these requirements, the present study adopts a predictive, not causal, framing using ensemble Deep Artificial Neural Networks (DANN) with k-fold validation and transparent performance reporting. This approach addresses documented limitations by capturing complex non-linear interactions across heterogeneous predictors, demonstrating superior performance in high-dimensional spatial contexts while enabling robust policy-oriented scenario analysis.

In this sense, the primary contribution of this paper does not lie solely in the scale of the dataset, although its national coverage significantly strengthens external validity. Rather, the contribution lies in the integration of a large-scale spatial dataset with a non-linear deep learning architecture, enabling the estimation of energy-related capitalization effects under multiple policy scenarios at national level. This approach addresses both empirical inconsistencies observed in prior localized studies and methodological constraints associated with traditional linear hedonic models. These limitations underscore the need for a more comprehensive, up-to-date, and methodologically robust analysis of the Italian RE market in relation to Energy Efficiency. By combining national-scale data, advanced spatial controls, and a deep learning modelling framework, the present study seeks to provide policy-relevant evidence capable of supporting large-scale regulatory design rather than isolated urban case studies.

3. Materials and method

To examine how various attributes, particularly energy performance, affect residential property prices, this study employs the hedonic price method. The hedonic price method, founded on Lancaster’s consumer theory (Lancaster, 1966; Rosen, 1974), presumes that the utility of a good is derived from its attributes, rather than the good itself. This means that the value of a property is equivalent to the summation of the marginal costs of its various features. The hedonic price method allows the detection and quantification of the influence of specific attributes, such as energy performance, on the overall value of the property. This approach is particularly well-suited to this study since it allows for disentangling the complex bundle of factors affecting RE prices, with a specific focus on EE. The current study combines econometric hedonic pricing techniques with machine learning (Kalliola *et al.*, 2021; Yakub *et al.*, 2021). This choice stems from the fact that the combination of these two approaches allows for handling large volumes of data and numerous parameters with the highest predictive precision. This hybrid approach has been a very successful mass appraisal method in RE price assessments (IAAO, 2013).

The study was conducted in four distinct phases, as illustrated in Figure 2. Phase I: Data Mining and Database Construction involved creating a comprehensive database of residential properties in the Italian market by collecting and organizing data from various online sources. This process incorporated relevant variables identified through an extensive literature review,

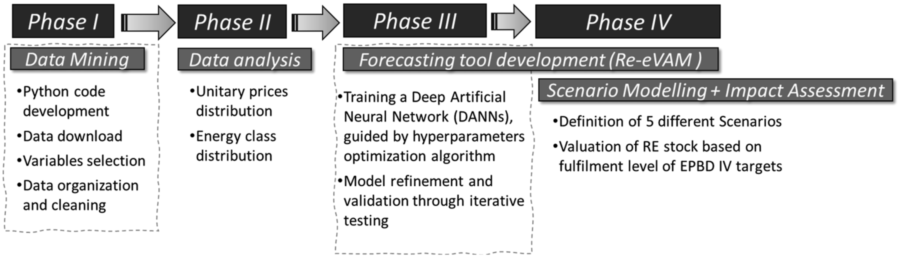

Figure 2. Study’s phases

focussing on factors influencing property values and EE indicators. Phase II: Data Analysis, Refinement, and Preliminary Exploration, involved thorough data cleaning and descriptive statistical analyses, focussing on energy performance certificates (EPCs) and their correlations with other property characteristics. Phase III developed Re-eVAM, a DANN designed to forecast property values based on energy performance and other relevant factors. Phase IV: Scenario Modelling and Impact Assessment applied Re-eVAM to evaluate EPBD IV's potential impact on market values across five distinct scenarios, each representing different EE strategies and implementation levels. This methodology allows for a quantitative assessment of how various EPBD IV implementation strategies could affect overall RE market values and potentially lead to depreciation of non-compliant building stock. In the following sections, each phase will be discussed in detail.

3.1 The case study

The study focuses on the largest Italian cities, encompassing approximately half of the country's population, to ensure a significant and representative sample of the national RE market. This approach provides a more robust dataset for machine learning algorithm training, enabling better estimation of marginal prices for various property features. Additionally, concentrating on major cities allows for the creation of large, homogeneous data clusters, minimizing potential distortions that could arise from including smaller towns with their distinct market dynamics, particularly regarding holiday homes.

Italy represents a particularly relevant case study for this research, as it can be considered a challenging context for the application of the EPBD IV, although a homogeneous comparison among Member States of the EU is not possible. Across EU, EPCs result from different choices in the transposition of the same directives. As such, the framework of EPC application is unharmonized and the comparison of the distribution of the building stock over the EPC scale in different countries is far from straightforward. For instance, differences in the labels of Energy Classes take place (Kamenders *et al.*, 2022). In Germany, EC ranges from A+ to H, while in Italy they range from A (including A4, A3, A2 and A1) to G. Moreover, the indicators contributing to the scores are different (Zinzi *et al.*, 2017) and the limits bounding the EC differ as well (Kamenders *et al.*, 2022). Italy presents a compelling case study for EPBD IV application due to its high proportion of old, energy-inefficient buildings (51.8% in classes F and G) (ENEA - Agenzia Nazionale Efficienza Energetica, 2026), recent experience with large-scale energy renovation incentives (SuperEcoBonus), resulting in over 116 billion euros in total costs and approximately 500,000 interventions from October 2020 to May 2024 (Camera dei Deputati, 2024), and high rate of homeownership (74%) (Eurostat, 2022). Beyond these unique factors, there is a notable lack of comprehensive studies on EPBD IV's potential impact on the Italian RE market, making this research particularly timely and relevant.

3.2 Phase I – data mining

An automated web-crawler was developed in Python to collect data from various Italian RE online marketplaces. The process involved identifying comprehensive websites, assessing available information and formats, defining common parameters, and developing specific crawler branches for each website. To balance efficiency and server load, the crawler implemented pauses between downloads and distributed activity across multiple websites for different municipalities. This approach ensured data consistency while minimizing server impact and managing download times effectively.

From the available online RE listings [2], a wide range of characteristics was collected. This comprehensive data structure is consistent with the hedonic pricing framework, according to which the market value of a RE unit can be interpreted as the result of the combined and interrelated marginal contributions of its attributes. Although the primary interest of this paper is to estimate the marginal effect of Energy Class (EC) on property value, it was therefore necessary to control for all major determinants that contribute to overall market pricing. Based on existing literature and appropriate units of measurement, the key explanatory variables were identified and operationalized for the empirical analysis. These variables are reported in Table 1, which reports their description, variable type, units of measurement (UM), and admitted values. The dataset includes a structured and heterogeneous set of predictors consistent with a hedonic pricing framework, in which the market value of a property is explained as a function of its structural, qualitative, technological, energy-related, and locational attributes. Each observation is uniquely identified by a UUID (advertisement code) and associated with a $\mathrm{Date}_{\mathrm{Creation}}$ variable, treated as a continuous temporal variable to allow tracking of listing dynamics. Continuous variables comprise Floor Area ($\mathrm{A}_{\mathrm{Floor}}$), geographic coordinates, Latitude (Lat) and Longitude (Long), and Market Value (M_{V}). Spatial location was explicitly incorporated through precise geographic coordinates, treated as continuous predictors within the neural network architecture, enabling the model to learn spatial price gradients through non-linear interactions with other property attributes. Unlike spatial econometric models that explicitly model spatial dependence structures (e.g. spatial lags or error terms), the DANN approach treats coordinates as standard continuous features alongside structural and qualitative characteristics. This specification assumes that spatial patterns in unitary prices—to the extent they exist at the national scale—are captured through learnt coordinate-based functions rather than through explicit neighbourhood dependence mechanisms. The nationwide geographical dispersion of training and testing observations, combined with the heterogeneity of urban contexts represented in the dataset, reduces the risk that model performance is driven predominantly by localized clustering or proximity effects specific to individual cities. Structural characteristics include the number of rooms ($\mathrm{n}_{\mathrm{Rooms}}$), bedrooms ($\mathrm{n}_{\mathrm{Bedrooms}}$), bathrooms ($\mathrm{n}_{\mathrm{Bathrooms}}$), and the maximum storey of the building ($\mathrm{n}_{\mathrm{Storey,Max}}$), all treated as discrete variables. A binary indicator identifies whether the unit is located on the top storey ($\mathrm{is}_{\mathrm{On\ top}}$ storey). Additional physical and functional amenities are captured through binary (dummy) variables such as Lift, Private Garden ($\mathrm{Garden}_{\mathrm{Private}}$), Shared Garden ($\mathrm{Garden}_{\mathrm{Shared}}$), Parking or private garage (Par), Cellar or attic (Cel), and Terrace (Ter). Qualitative and condition-related attributes are represented by ordinal variables, including Status of Maintenance (SoM), ranging from properties in need of renovation to new or under-construction units, and Energy Class (EC). The EC variable represents an ordinal scale ranging from 1 (class G) to 10 (class A4) [3]. Multi-level categorical variables include Furniture (Furn) and Concierge services (Con), both coded according to increasing qualitative intensity. Further technological and comfort-related features are incorporated through binary variables such as Alarm (A), Optical Fibre (OF), Building Automation (BA), Central Heating (CH), Fireplace, Air Conditioning (AC), Mechanical Ventilation (MecV), and the presence of a Photovoltaic System (PhS). Each variable was carefully defined with specific admitted values to ensure data consistency, reduce noise, and facilitate subsequent modelling procedures. The dependent variable, unitary Market Value (MV), was filtered on unitary price basis (€/m^2) within a predefined range to exclude extreme outliers and ensure a realistic representation of observed market conditions [4]. This comprehensive dataset forms the foundation for the subsequent econometric and machine learning analysis, allowing us to estimate the marginal contribution of Energy Class while appropriately controlling for the full set of structural, qualitative, technological, and locational determinants traditionally recognized in RE valuation models.

3.3 Phase II – data analysis

In this phase, we conducted a comprehensive data refinement and preliminary analysis process. Initially, the gathered data underwent a thorough cleaning and filtering procedure. The collected data were filtered and cleaned, standardizing the scales of the features downloaded to ensure homogeneity across the diverse websites used as data sources and removing duplicates and advertisements including outliers or unreliable/void values. This meticulous process ensured data consistency and reliability for subsequent analyses.

Table 1. Collected variables in the database, including variable types and Units of Measurement (UM)

Parameter	Description	Variable type	UM	Admitted values
UUID	Advertisement Code	Qualitative (Nominal)	String	Any
$\text{Date}_{\text{Creation}}$	Creation date	Continuous (Temporal)	Date	Any
Lat	Latitude	Continuous	°	$36 \leq x \leq 47$
Long	Longitude	Continuous	°	$7 \leq x \leq 19$
Type	Type	Qualitative (Nominal)	String	$x \in$ “Apartment”
A_{Floor}	Floor area	Continuous	m^2	$28 \leq x \leq 3000$
n_{Rooms}	Number of rooms	Discrete	–	$1 \leq x \leq 20$
n_{Bedrooms}	Number of bedrooms	Discrete	–	$1 \leq x \leq 13$
$n_{\text{Bathrooms}}$	Number of bathrooms	Discrete	–	$x \geq 1$
$n_{\text{Storey,Max}}$	Maximum storey	Discrete	–	$-1 \leq x \leq 22$
$\text{is}_{\text{On top storey}}$	The RE unit is on the top storey	Binary (Dummy)	–	$x \in \{0, 1\}$
Lift	Lift	Binary (Dummy)	–	$x \in \{0, 1\}$
$\text{Garden}_{\text{Private}}$	The RE unit has a private garden	Binary (Dummy)	–	$x \in \{0, 1\}$
$\text{Garden}_{\text{Shared}}$	The RE unit has a shared garden	Binary (Dummy)	–	$x \in \{0, 1\}$
Par	Parking space or Private garage	Binary (Dummy)	–	$x \in \{0, 1\}$
Cel	Cellar or attic	Binary (Dummy)	–	$x \geq 0$
Ter	Terrace	Binary (Dummy)	–	$x \in \{0, 1\}$
SoM	Status of Maintenance: 1: In need of renovation; 2: Average/Good; 3: Very good/Renovated; 4: New/Under construction.	Qualitative Ordinal	–	$x \in \{1; 2; 3; 4\}$
Furn	Furniture provided with the RE unit: 0: no furniture; 0.3: only kitchen; 0.5: partially furnished; 1.0: fully furnished.	Qualitative Ordinal	–	$x \in \{0.0, 0.3, 0.5, 1.0\}$
Con	Concierge: 0.0: No concierge; 0.3: Reception; 0.5: Concierge during working hours; 1.0: Concierge along the whole day.	Qualitative Ordinal	–	$x \in \{0.0, 0.3, 0.5, 1.0\}$
A	Alarm	Binary (Dummy)	–	$x \in \{0, 1\}$
OF	Optical fibre	Binary (Dummy)	–	$x \in \{0, 1\}$
BA	Building automation	Binary (Dummy)	–	$x \in \{0, 1\}$
EC	Energy Class: 1: G (the least efficient EC); 2: F; 3: E; 4: D; 5: C; 6: B; 7: A or A+ or A1; 8: A2; 9: A3; A4 (the most efficient EC): 10	Qualitative Ordinal	–	$x \in \{1, 2, 3; 4; 5; 6; 7; 8; 9; 10\}$
CH	Central heating	Binary (Dummy)	–	$x \in \{0, 1\}$

(*continued*)

Table 1. Continued

Parameter	Description	Variable type	UM	Admitted values
Fireplace	Fireplace	Binary (Dummy)	–	x Є {0, 1}
AC	Air-conditioning	Binary (Dummy)	–	x Є {0, 1}
MecV	Mechanical ventilation	Binary (Dummy)	–	x Є {0, 1}
PhS	Photovoltaic system	Binary (Dummy)	–	x Є {0, 1}
M_V	Market value	Continuous	€	Value filtered based on unitary price [€/m^2]: $400 \leq x \leq 20{,}000$

Following the data cleaning, we performed extensive descriptive statistical analyses to gain initial insights into our dataset. These analyses focused on several key aspects. We began by examining the frequency distribution of records based on two primary factors:

(1) Unitary Price (€/m^2), $UP = \frac{M_V}{A_{Floor}}$

(2) State of Maintenance (SoM) and Energy Class (EC) analyzed jointly.

The unitary price analysis provided insights into the price range and distribution of properties in our dataset, while the SoM and EC distribution investigation revealed the condition and Energy Efficiency profile of the properties. We then calculated and analyzed the mean unitary price, stratified by SoM and EC. This analysis allowed us to understand how property condition and EE relate to pricing, potentially revealing the impact of these factors on property values. The numerical encoding of ordinal variables (such as EC) does not impose linearity assumptions within the modelling framework. Moreover, the conversion of the Energy Class into a linear variable preserves its natural ordering (from G to A4), thus reflecting the monotonic ranking embedded in the classification system (where higher classes correspond to better energy performance) and makes it possible to use regression also for values of the Energy Class which have a low number of occurrences (hence, of training data). Finally, the ensemble of Deep Artificial Neural Networks (DANNs), through their non-linear activation functions, is capable of learning complex and non-linear relationships between input features and the output variable. Therefore, no linear marginal contribution of Energy Class to Market Value is assumed *a priori*.

To add a geographical dimension to our analysis, we utilized GeoDa, a specialized software for spatial data analysis. This enabled us to conduct a geographical analysis of our data, focussing on the spatial distribution of unitary prices and Energy Classes across the territory. We mapped the spatial distribution of properties' EC, allowing us to identify potential geographical patterns or clusters of high/low-value properties. Similarly, we visualized the spatial distribution of Energy Classes among properties, providing insights into the geographical patterns of EEE in the building stock. These spatial analyses were crucial in understanding the geographical context of our data, potentially revealing local trends, clusters, or disparities in property values and EE that might not be apparent from non-spatial statistical analyses.

3.4 Phase III – forecasting tool development

Re-eVAM utilizes DANNs for RE valuation. Figure 3 illustrates the structure of a typical DANN, which employs multiple stacked layers of interconnected computational units

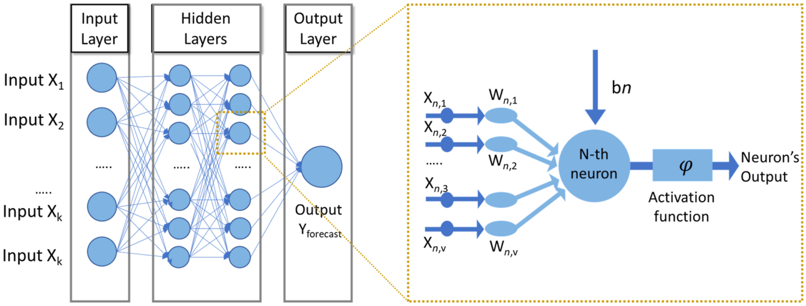

Figure 3. Structure of a deep Artificial neural network

(neurons), mimicking the structure of a brain. The model processes input parameters through several hidden layers, extracting complex correlations between property characteristics and market value (Pittarello *et al.*, 2021). Re-eVAM's structure includes an input layer for property data, multiple hidden layers for data processing, and an output layer that generates the predicted market value. This architecture allows Re-eVAM to capture intricate relationships in Real Estate data, potentially improving valuation accuracy, especially for Energy Efficient properties.

The ensemble of input parameters entering the DANN is represented by a column vector $[X_k]$, where $1 \leq k \leq K$, with K being the total number of inputs to the Artificial Neural Network (ANN), while a corresponding output $Y_{forecast}$ results from the network. In this study, the input parameters correspond to the features and characteristics of the buildings (Including Energy Class), whereas $Y_{forecast}$ is the Market Value (M_V). Figure 3 also provides a detailed view at the single-neuron level: multiple input values enter the single neuron, and one output is produced. Assuming that the layer hosting this single neuron consists of N neurons and the previous layer consists of V neurons, each nth neuron, with $1 \leq n \leq N$, receives information from the V neurons hosted in the previous layer, thus each piece of information is named $x_{n,v}$, with $1 \leq v \leq V$. The overall information is processed inside the neuron to provide output Y_n, through Eq. 1. Here, $W_{n,v}$ is the individual weight and b_n is named bias, while ϕ is the activation function giving the neuron's output value.

$$\forall\, n^{th} \text{ neuron}, Y_n = \varphi\left(\sum_{v=1}^{V} \left[\,(W_{n,v} * x_{n,v}) + b_n\,\right]\right) \qquad \text{Eq. (1)}$$

Given the number of hidden layers and the number of neurons per hidden layer, the DANN is trained to assess the market value of the corresponding RE unit (Du and Wang, 2023; Maselli, 2022; Mostofi *et al.*, 2022; Nikolov, 2023) based on the given values of the input parameters. It is important to clarify that the objective of the DANN framework is predictive valuation rather than causal inference: unlike hedonic regression models, the neural network does not provide directly interpretable coefficients associated with individual covariates, and the model is therefore employed as a forecasting tool for scenario simulation rather than as an econometric instrument to quantify marginal causal effects. This distinction is methodologically relevant: while the hedonic approach adopted in the companion spatial analysis enables identification and interpretation of energy class premiums controlling for confounders, the DANN prioritizes predictive accuracy and generalization capacity for aggregate stock valuation under alternative policy scenarios.

As illustrated in Figure 4, the modelling framework consists of three main components: the neural network architecture, the optimized training procedure, and the training database. The neural network architecture is composed of an input layer receiving the full set of explanatory

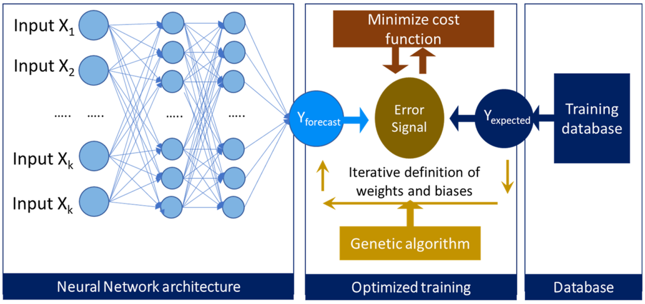

Figure 4. Optimized training process

variables (structural, qualitative, technological, energy-related, and spatial attributes), multiple fully connected hidden layers, and a single output node producing the predicted market value ($Y_{forecast}$). Through nonlinear activation functions, the hidden layers enable the model to capture complex and non-linear interactions among predictors, including potential threshold and interaction effects between EC, SoM, and locational characteristics. The training process consists of determining the values of weights and biases that minimize the difference between forecasted and actual market values (i.e. M_V), through an iterative optimization procedure. The predicted output ($Y_{forecast}$) is compared with the observed value ($Y_{expected}$) extracted from the training database, generating an error signal. This error is aggregated over the full set of training observations through a cost function, which summarizes the overall prediction discrepancy across the dataset. The optimization objective is to minimize this cost function, thereby progressively improving predictive accuracy. The optimal number of hidden layers and neurons per layer, known as hyperparameters, also require calibration. These hyperparameters define the network architecture and significantly affect model performance and generalization capacity. As shown in Figure 4, their selection was performed through an iterative optimization process based on a genetic algorithm. The genetic algorithm explores alternative architectural configurations by simulating evolutionary mechanisms (selection, crossover, and mutation), allowing the identification of the hyperparameter combination that yields the lowest prediction error. This hybrid approach enhances convergence robustness and reduces the risk of local minima, strengthening the reliability of the final DANN specification.

To train and identify the best DANN, the final database of 92,719 records was split into three parts: (1) 60% of the records were randomly selected to build up the training dataset, i.e. the records used to tune neurons' weights and biases; (2) 20% of the records were randomly selected to create the test dataset, i.e. the records which are used to evaluate the accuracy of the single DANN and determine whether further iterations are needed to improve the accuracy; (3) 20% of the records build up the validation dataset. The validation dataset is used to select the best DANN architecture. This new dataset contains records not included in the training set or in the testing set in order to evaluate the accuracy of the best DANN based on an independent set of records.

The model employs ensembles of DANNs rather than single DANNs to enhance reliability. The optimization algorithm selects the best ensemble based on regression accuracy estimation in the validation dataset.

3.5 Phase IV – case studies application

This study aims to investigate the potential influence of the EPBD IV Directive on the overall value of the national RE market. The research seeks to assess changes in market value resulting from different energy EE strategies, developed across different scenarios, considering the inherent uncertainties in future directive adoption. The objective is to support the fields of application beyond academic interpretation, including policy design under EPBD IV, calibration of retrofit incentives, prioritization of building-stock intervention strategies, and scenario-based support for valuers, investors, and public authorities. In this sense, the methodological framework is used not only to estimate aggregate market effects, but also to provide an operational basis for comparing alternative implementation pathways and their associated value redistributions.

Uncertainties in EPBD IV implementation stem from several factors. The ambitious 2050 zero-emissions goal for buildings presents a significant challenge, considering current conditions and the need to retrofit existing structures. Additionally, EU Member States have flexibility in adopting the Directive, with many implementation details yet to be defined. For instance, Italy has not yet adopted EPBD IV, leaving many specific details unspecified. Market values may also undergo substantial variations over time due to unpredictable social and economic factors, necessitating the use of current market values as a basis for projections in this study.

The generated five scenarios (S_0; $S_{A.1}$; $S_{A.2}$; S_B; S_C) were generated as combinations of different assumptions regarding how the directive could be implemented. These assumptions include whether non-compliant buildings in the existing stock will be retrofitted or not (ER Intervention), the target year for achieving zero emissions (Regulation framework), and the Market Value of the Non-Compliant Properties (M_V of NCPs). The interplay of these factors defines our five distinct scenarios (Table 2). To minimize arbitrariness, the scenario design was structured as a stress-testing exercise, rather than a precise forecast of future market developments. Each scenario combines transparent assumptions derived from the current EPBD IV regulatory framework, observed retrofit practices, and the residual value logic commonly used in valuation analysis. The goal is not to identify a scenario that will occur exactly as modelled, but to identify a plausible range of market responses under alternative regulatory and restructuring conditions.

In order to classify the RE units based on the need for retrofitting and to define the M_V of NCPs, we introduced the Overall Status parameter (O_S), as calculated as showed in Eq. 2.

$$O_S = EC + \frac{SoM}{10}; \; [-] \qquad \text{Eq. (2)}$$

where:

(1) *EC* is the value of "Energy Class" [1 to 10];

(2) *SoM* is the value of "State of Maintenance" [1 to 4].

The O_S parameter is introduced as an operational ranking tool aimed at identifying the segment of the building stock subject to regulatory intervention under different EPBD IV scenarios, accounting also for market behaviour. It is not conceived as a theoretically grounded composite valuation index. The formulation reflects the regulatory primacy of Energy Class in determining compliance, while the State of Maintenance is incorporated as a secondary

Table 2. Scenarios identified to depict how EPBD IV may affect RE market value.(Residual Value = RV)

Scenarios	ER interventions	Regulation framework	M_V of NCPs
S_0	No (Ec_n)	Current	100%
$S_{A.1}$	No (Ec_n)	EPBD IV @ 2040	0%
$S_{A.2}$	No (Ec_n)	EPBD IV @ 2040	Max (0,RV)
S_B	Yes (up to EC_{n+2})	EPBD IV @ 2040	100%
S_C	Yes (up to EC_{10})	EPBD IV @ 2050	100%

refinement factor. By scaling SoM by one order of magnitude, the ordinal hierarchy of Energy Class is preserved, ensuring that differences between energy classes remain dominant in the ranking structure. The O_S parameter ranges from 1.1 to 10.4. Lower O_S values indicate higher primary energy consumption per unit area and therefore identify assets with lower energy performance that would benefit most from energy retrofitting interventions, as presented in Table 3.

The O_S was calculated for each property in the database, and its frequency distribution was determined. This parameter plays a crucial role in identifying which properties require retrofitting in each of the developed scenarios. Each scenario is described in detail below, outlining the specific assumptions and parameters used in the modelling process.

3.5.1 Scenario S_0. This scenario represents the current status of buildings and regulations. As such, the records collected from the RE online marketplaces are directly used. However, as a reference, the overall value of the stock [5] was also calculated using the developed forecast tool to verify its accuracy. This scenario considers 100% of the current $\mathrm{M_V}$ of the properties as they are.

3.5.2 Scenario $S_{A.1}$. This scenario considers buildings in their current status, but within a regulatory framework complying with EPBD IV. In such a context, the market value of RE units non-compliant with the laws ($\mathrm{M_V}$ of NCPs) will suffer a decrease in value. No detail in this regard is available yet, but, in previous draughts, EPBD IV stated that by 2030 and 2033, the worst performing assets (in particular, those worse than Energy Class E and D respectively) should be retrofitted. It is worth noting that the adoption of EPBD IV will establish a new energy classification system, which is currently undefined. Hence, given the lack of information, this study considers the existing energy classification system and the values of EC given in the advertisements collected. Furthermore, it was established that NCP units should be considered unavailable for sale or rental until retrofitted.

Therefore, $\mathrm{S_{A.1}}$ considers that NCP units would be removed from the market until retrofitted, as a worst-case scenario. The $\mathrm{M_V}$ of NCPs is considered null. The study considers the EPBD IV requirement for member states to achieve at least 55% of the decrease in average primary energy use through the renovation of the 43% worst-performing residential buildings, with reference to targets set for 2040. Based on this directive, and assuming a linear decrease in primary energy use, the research determines that 78.2% of the collected RE units should be considered for retrofitting, where 78.2% is assessed as 0.43/0.55 [6]. This percentage is then applied in the scenarios, with scenarios $\mathrm{S_B}$ and $\mathrm{S_C}$ treating these units as retrofitted, while $\mathrm{S_{A.1}}$ and $\mathrm{S_{A.2}}$ consider them as not retrofitted. Categories from 1.1 to 3.4 account for approximately 74.16% of the records. To reach the 78.2% worst-performing residential buildings, 4.04% of the occurrences must be selected. Since category 3.3 accounts for approximately 5.17% of the occurrences, approximately 80% of these RE units in category 3.3 was randomly selected.

Table 3. Overall Status parameter (O_S) computation

SoM		Worst			Best
EC		1	2	3	4
Worst	1	1.1	1.2	1.3	1.4
	2	2.1	2.2	2.3	2.4
	3	3.1	3.2	3.3	3.4
	4	4.1	4.2	4.3	4.4
	5	5.1	5.2	5.3	5.4
	6	6.1	6.2	6.3	6.4
	7	7.1	7.2	7.3	7.4
	8	8.1	8.2	8.3	8.4
	9	9.1	9.2	9.3	9.4
Best	10	10.1	10.2	10.3	10.4

In summary, this scenario presents a worst-case analysis by assigning a null market value to the 78.2% of properties identified as the worst-performing buildings in the database. While this assumption is admittedly extreme, it serves to illustrate the potential magnitude of suspended market value that could result from inadequate policies promoting energy retrofit interventions. This approach provides a baseline for assessing the economic risks associated with insufficient support for EE improvements in the RE sector.

3.5.3 Scenario $S_{A.2}$. This scenario is similar to $S_{A.1}$, but the MV of NPCs is not considered null. Instead, it is calculated based on an average cost of retrofitting derived from recent data available from the Italian tax credit known as SuperEcoBonus and property statistics from ISTAT (Italian National Institute of Statistics), as follows:

1. Total cost of SuperEcoBonus: approximately 116 B€;
2. Total number of residential buildings retrofitted via SuperEcoBonus tax credit: approximately 0.5 M.
3. Total number of residential buildings in Italy: approximately 12.2 M.
4. Total number of residential RE units in Italy: approximately 31.2 M.
5. Average floor area of RE units in Italy: approximately 87 m^2.

Based on these figures, the average cost of adequate energy retrofit interventions is estimated at about 1,039 €/m^2 for an improvement of two Energy Classes. In $S_{A.2}$, the unitary MV of NPCs is calculated as the maximum value between zero and the Residual Value (RV), where RV is the difference between the MV of the building after retrofitting and the retrofitting costs, set at 1,039 €/m^2.

3.5.4 Scenario S_B. In S_B the NCPs are appropriately retrofitted to comply with EPBD IV, with particular reference to the targets of 2040. Therefore, the values of *SoM* and *EC* of 78.2% of the records were updated as follows:

1. *EC* is increased by two classes, as expressed in Eq. 3. This is assumption is based on the fact that the Italian "SuperEcoBonus" tax reduction scheme required an improvement of at least two EC, and that the draft of EPBD IV considered EC_3 and EC_4 as minimum requirements for 2030 and 2033, respectively.
2. *SoM* was increased to 3 (Very good/Renovated) or maintained at 4 (New/Under construction) if already at that level, as expressed in Eq. 4. This reflects the assumption that energy retrofitting would significantly improve the overall condition of the property.

Finally, the M_V of the NPCs was updated, based on $EC^{Updated}$ and $SoM^{Updated}$, as follows:

$$EC^{Updated} = EC + 2\,[-] \qquad \text{Eq. (3)}$$

where:

1. $EC^{Updated}$ is the value of *EC*, updated after the energy retrofit intervention $[-]$;
2. *EC* is the value of "Energy Class" $[-]$ with no energy retrofit intervention.

$$SoM^{Updated} = max(SoM, 3)\,[-] \qquad \text{Eq. (4)}$$

where:

1. $SoM^{Updated}$ is the value of *SoM*, updated after the energy retrofit intervention $[-]$;
2. *SoM* is the value of "State of Maintenance" $[-]$ with no energy retrofit intervention.

Scenario S_C

This is the best-case scenario, where the RE units are retrofitted in order to achieve zero-emission targets (EC_{10}), which is the target for 2050. This could seem a tough achievement, but, on the other hand, it could be much more difficult to achieve this target by iterations. As a matter of fact, the overall cost of interventions would be probably lower if the zero-emission target is achieved in just one unique energy retrofit intervention, instead of two sequential interventions (one complying with 2040 targets and one complying with 2050 targets). Therefore, the values of *SoM* and *EC* of the RE units selected for energy retrofit (i.e. 78.2% of the records) were updated as expressed in Eq. 5 and Eq. 4.

$$EC^{Updated} = 10\,[-] \qquad \text{Eq. (5)}$$

4. Results and discussion per phases

4.1 Phase I - data mining

The study focused on the 451 largest municipalities in Italy (Table A1 in Annex), collectively representing approximately half of the country's population. This approach ensures a significant and representative sample of the national RE market, allowing for more generalizable findings. Larger municipalities were chosen for their active and diverse RE markets, substantial online advertisements, and tendency to lead in implementing EE measures. These cities represent a significant portion of Italy's economic activity and housing stock, crucial for understanding EPBD IV's potential impact. The selection includes municipalities from various regions, capturing regional variations in RE markets and EE practices. Table 4 provides a detailed breakdown of the total number of municipalities, overall population, and proportion covered by the selected municipalities for each region.

Figure 5 illustrates the population distribution among the selected cities, displaying the cumulative population in descending order. This visualization demonstrates the concentration in larger urban areas and the gradual inclusion of smaller municipalities. The population-based selection criterion aims to create a model that accurately represents Italy's diverse RE market, focussing on areas where EPBD IV implementation is likely to have the most significant impact.

The data mining process yielded 325,485 advertisements, containing parameters as described in Table 1. A rigorous data cleaning process was then implemented, involving the removal of duplicates, outliers, and entries with unreliable or missing values. The dataset was refined to focus exclusively on apartments, ensuring a more homogeneous sample and reducing variability in property characteristics. This decision was made due to the underrepresentation of villas and single-family houses, which could lead to less accurate evaluations. Certain variables were excluded from the regression model, including those reflecting subjective buyer preferences (Furn, $n_{Storey,Max}$), features more relevant to single-family homes ($Garden_{Private}$, $Garden_{Shared}$) and attributes with limited data availability (Fireplace, MecV, OF, Con, BA). This comprehensive filtering approach resulted in a high-quality, consistent dataset focused on the most relevant and consistently available variables for apartment properties. To avoid redundancy, each advertisement was compared against all the others to find duplicates, based on the simultaneous fulfilment of the following conditions:

1. The values of latitude differ by no more than 0.0005°;
2. The values of longitude differ by no more than 0.0005°;
3. The market values differ by no more than 5%;
4. The values of floor area differ by no more than 5%;
5. No difference in the status of maintenance takes place;

Table 4. Distribution of municipalities considered among the Italian regions, including location and existing population

	Main data about the region					Data coverage			
		Municipalities		Population		Municipalities		Population	
Region	Location	Number [−]	%	Number [p]	%	Number [−]	% of regional	Number [p]	% of regional
Abruzzo	South	305	3.8%	1.31 M	2.2%	13	4.3%	0.58 M	44.0%
Basilicata	South	131	1.6%	0.58 M	1.0%	2	1.5%	0.13 M	21.9%
Calabria	South	404	5.1%	1.96 M	3.3%	9	2.2%	0.63 M	32.3%
Campania	South	550	6.9%	5.77 M	9.7%	59	10.7%	3.48 M	60.3%
Emilia-Romagna	North-East	331	4.2%	4.34 M	7.3%	32	9.7%	2.37 M	54.6%
Friuli-Venezia Giulia	North-East	215	2.7%	1.22 M	2.1%	5	2.3%	0.41 M	33.9%
Lazio	Centre	378	4.8%	5.50 M	9.3%	32	8.5%	4.03 M	73.2%
Liguria	North-West	234	2.9%	1.57 M	2.6%	9	3.8%	0.94 M	59.8%
Lombardia	North-West	1,516	19.1%	9.70 M	16.3%	61	4.0%	3.86 M	39.8%
Marche	Centre	229	2.9%	1.54 M	2.6%	14	6.1%	0.68 M	43.8%
Molise	South	136	1.7%	0.31 M	0.5%	2	1.5%	0.08 M	26.0%
Piemonte	North-West	1,197	15.0%	4.36 M	7.3%	28	2.3%	2.00 M	45.8%
Puglia	South	258	3.2%	4.05 M	6.8%	43	16.7%	2.40 M	59.1%
Sardegna	Islands	377	4.7%	1.64 M	2.8%	13	3.4%	0.66 M	40.4%
Sicilia	Islands	390	4.9%	5.00 M	8.4%	53	13.6%	3.18 M	63.5%
Toscana	Centre	274	3.4%	3.67 M	6.2%	34	12.4%	2.10 M	57.1%
Trentino-Alto Adige	North-East	292	3.7%	1.03 M	1.7%	4	1.4%	0.29 M	28.4%
Umbria	Centre	92	1.2%	0.88 M	1.5%	7	7.6%	0.47 M	52.7%
Valle d’Aosta	North-West	74	0.9%	0.13 M	0.2%	1	1.4%	0.03 M	26.9%
Veneto	North-East	571	7.2%	4.86 M	8.2%	30	5.3%	1.69 M	34.9%
TOTAL		7,954	100%	59.40 M	100%	451	5.7%	30.00 M	50.5%

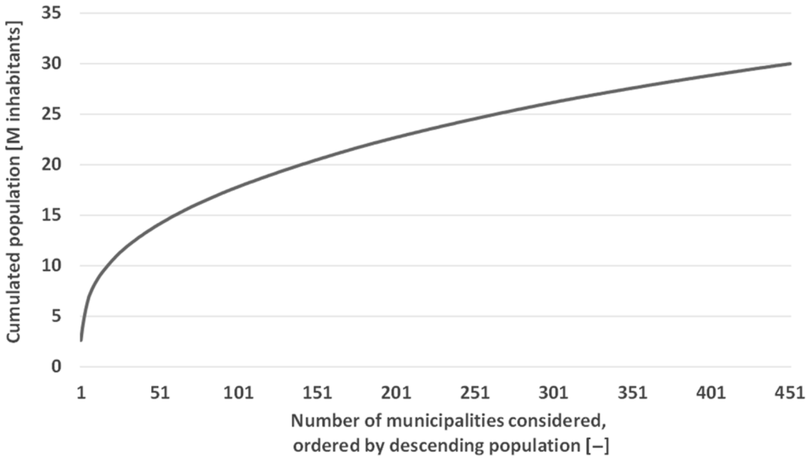

Figure 5. Cumulated population of the municipalities considered, ordered by population

(6) No difference in the number of rooms takes place;

(7) No difference in the number of bathrooms takes place;

(8) No difference in the value of the Energy Class takes place.

Among the duplicates identified, the most recent advertisement was retained. At the end of the data cleaning process, 92,719 advertisements remained.

4.2 Phase II – data analysis

After the data cleaning process, the database was explored. Figure 6 illustrates the distribution of unitary prices (€/m2) within our dataset. The distribution appears right-skewed, with a concentration of properties in the lower to middle price ranges and a long tail extending towards higher prices. This pattern likely reflects the diverse nature of Italy's RE market, encompassing both affordable housing in smaller towns and high-value properties in downtown areas and tourist destinations. The highest frequency of occurrences is observed in the range of 1400–2000 €/m^2, suggesting that this represents the typical price range for a significant portion of Italian properties. The dataset's wide price range underscores the significant variability in Italy's RE market. The cumulative frequency line provides further insights, showing a steep rise in the lower to middle price ranges before plateauing around the 6000 €/m^2 mark. This suggests that while Italy has a sizeable affordable and mid-range housing stock, there's also a distinct luxury segment. The fact that approximately 80% of properties are priced below 4000 €/m^2 indicates that a significant portion of Italy's housing stock remains relatively affordable, which could be crucial when considering the implementation of EE measures across the country. The implementation of the EPBD IV directive could have varying effects across these price ranges, potentially leading to more significant value increases in the mid-range market where EE improvements might be more noticeable and impactful on overall property value. We will therefore attempt to better understand the possible impacts through scenario analysis.

Figure 7 displays the spatial distribution of properties across the 451 analyzed municipalities, categorized by EPC clusters. This geographical representation divides the Italian building stock into three distinct EE clusters, showing a widespread distribution across all examined municipalities. The visualization highlights the predominance of lower efficiency buildings, the increasing presence of mid-range efficiency structures, and the

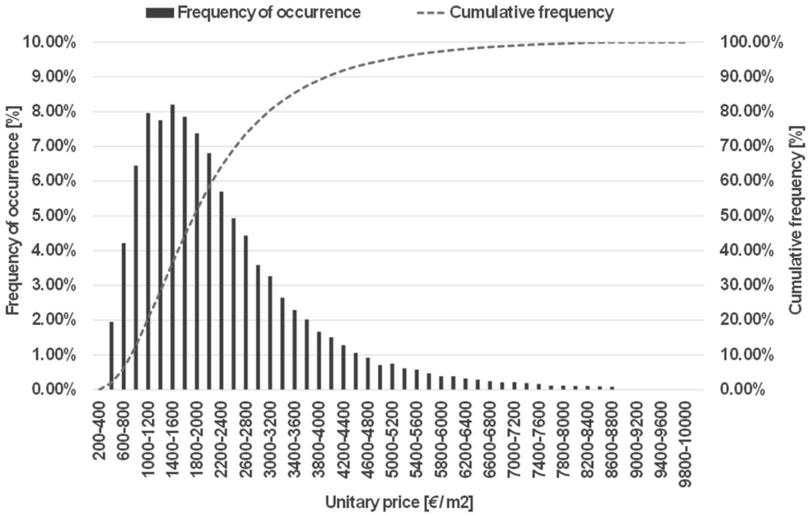

Figure 6. Frequency distribution of the records downloaded per unitary price (€/m^2)

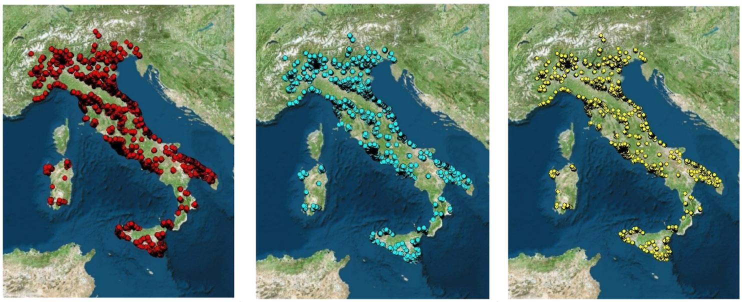
Energy classes: E, F, G | Energy classes: B, C, D | Energy classes: A1, A2, A3, A4

Figure 7. Distribution of the observations clusterized by EPCs, through GeoDa (Anselin *et al.*, 2010)

emerging, albeit limited, adoption of high-efficiency standards. This distribution pattern underscores the significant potential for EE improvements nationwide, particularly in upgrading the numerous E, F, and G class buildings to more efficient categories.

Table 5 and Table 6 offer key insights into the distribution (%) and pricing dynamics (€/m^2) of Italy's residential RE market, focussing on SoM and EC. A majority (52.35%) of properties fall into the lowest EE category (EC$_1$), highlighting significant potential for energy retrofitting. Most properties (47.03%) are in Average/Good condition (SoM$_2$), suggesting that structural maintenance has been prioritized over EE improvements. The data shows an inverse relationship between Energy Efficiency and frequency, with only 2.00% in the highest efficiency category (EC$_{10}$). High-efficiency (EC$_7$-EC$_{10}$) and excellent maintenance (SoM$_4$) properties are rare (4.17% of the dataset), potentially commanding premium prices due to their scarcity.

Table 6 supports these findings by illustrating the impact of SoM and EC on average unitary prices [7]. Although the neural network does not produce directly interpretable coefficients, descriptive evidence reported in Table 6 confirms that EC exhibits a systematic association

Table 5. Frequency distribution of values of SoM and EC over the database

SoM	Worst			Best	Total
EC	1	2	3	4	
Worst 1	9.79%	28.76%	13.16%	0.64%	52.35%
2	1.50%	8.39%	5.95%	0.20%	16.04%
3	0.62%	5.15%	5.17%	0.20%	11.14%
4	0.23%	2.64%	3.98%	0.24%	7.09%
5	0.08%	1.04%	2.27%	0.35%	3.74%
6	0.03%	0.40%	1.61%	0.56%	2.60%
7	0.03%	0.32%	1.43%	1.78%	3.55%
8	0.02%	0.13%	0.28%	0.33%	0.76%
9	0.01%	0.09%	0.19%	0.42%	0.71%
Best 10	0.01%	0.09%	0.26%	1.64%	2.00%
Total	12.32%	47.03%	34.29%	6.36%	

Table 6. Average unitary price (€/m^2) of buildings depending on SoM and EC

SoM	Worst			Best	Mean
EC	1	2	3	4	
Worst 1	1,922	1,983	2,572	2,563	2,260
2	1,984	2,009	2,557	2,575	2,281
3	2,084	2,039	2,590	2,685	2,349
4	1,981	2,116	2,600	2,803	2,375
5	2,080	2,098	2,699	2,535	2,353
6	1,996	2,316	2,893	2,844	2,512
7	1,287	2,073	3,050	3,272	2,421
8	2,486	1,996	2,924	3,097	2,625
9	1,177	1,819	3,035	3,359	2,348
Best 10	1,763	1,956	3,180	3,318	2,554
Total	1,876	2,040	2,810	2,905	

with unitary prices, even if its impact appears secondary to the SoM. This observed pricing structure supports the inclusion of Energy Class as a relevant predictor within the valuation framework. State of Maintenance emerges as a more significant factor in determining property values compared to Energy Class. The price difference between the worst (SoM_1) and best (SoM_4) maintained properties is substantial, approximately 55% (1,876 €/m^2 vs 2,905 €/m^2), and this trend holds across all ECs. This underscores the importance of overall property condition in the Italian market. The estimated association between EC and pricing, while present, shows a more nuanced pattern [8]. The highest EC categories show higher mean unitary prices by about 13%. Compared with prior empirical evidence reported in the Literature Review Section, the premiums emerging from our national dataset appear comparatively stronger. A plausible explanation lies in the temporal dimension: most existing studies rely on data collected between 2016 and 2020, whereas our analysis is based on substantially more recent observations. Over the last few years, the progressive tightening of European energy directives and the growing public awareness of sustainability issues may have intensified the market capitalization of energy performance, thereby amplifying price differentials across EC levels. There is a general upward trend in prices as EE improves, but this relationship is not strictly linear. This non-linearity also helps clarify part of the heterogeneity documented in earlier localized studies, which often employ linear hedonic specifications and may therefore underrepresent threshold or interaction effects across higher EC categories. Interestingly, the highest average price (2,625 €/m^2) is observed for EC_8, rather than the highest Energy Class (EC_{10}), suggesting that factors beyond EE play a role in determining top-tier property values. Properties combining high EE with excellent maintenance command the highest prices in the market. Those in EC_9-EC_{10} with SoM_4 average over 3,300 €/m^2, corresponding to an observed price differential exceeding 75% compared to the lowest category (1.1). This combined effect suggests that the capitalization of energy efficiency is increasingly embedded within a broader quality premium, rather than operating as an isolated determinant. Such interaction effects are only partially explored in previous Italian contributions, which are typically geographically circumscribed and based on smaller samples. This significant price difference highlights the market's valuation of both EE and overall property quality. While EE does influence property values, the relative predictive contribution appears secondary to the state of maintenance in the Italian RE market. By leveraging a recent, large-scale national dataset, our findings provide updated evidence that complements and extends earlier localized analyses, suggesting that the economic relevance of EC may have evolved over time in response to regulatory pressure, financial incentives, and heightened environmental awareness. These predictive patterns suggest potential implications for both policymakers and property owners. It suggests that while EE improvements are crucial for meeting climate goals, they should be implemented alongside overall property maintenance to maximize both environmental and economic benefits. The data also indicate a substantial predicted potential for value differentiation through energy retrofitting, particularly for the large stock of well-maintained but energy-inefficient properties that dominate the Italian housing market. In this sense, the present results contribute to bridging the gap between earlier city-level studies and the need for nationally representative evidence capable of informing contemporary policy design under the evolving European energy framework.

4.3 Phase III – forecasting tool development

The study employs ensembles of DANNs rather than single DANNs. Preliminary tests revealed that while some individual DANNs performed well, others produced forecasts with low accuracy. In contrast, DANN ensembles generally demonstrated more reliable forecasts [9]. Specifically, sets of 8 DANNs were utilized, with training and testing sets extracted via k-fold grouping. Each DANN within the ensemble is initialized with a distinct fixed random seed (ranging from 1234 to 1241) to ensure full reproducibility: under this specification, the training procedure is deterministic and produces stable aggregate results, while the diversity of initial weight configurations enables the ensemble to capture complementary patterns in the data and reduce prediction variance through averaging. The optimization algorithm trained these DANNs using adjustable hyperparameters within specified ranges, as outlined in Table 7. The target variable predicted by the DANN ensemble is the unitary market value expressed in euros per square metre (€/m^2), obtained by dividing the total asking price by the Floor Area (A_{Floor}). This normalization was adopted to enhance comparability across heterogeneous dwellings and to mitigate heteroskedasticity associated with size effects. During training, both input and output variables were scaled using a MinMax normalization procedure, and prediction outputs were subsequently rescaled to their original monetary unit (€/m^2) to ensure direct economic interpretability. Table 7 reports the fixed and adjustable hyperparameters defining the architecture and training configuration of the DANN ensembles. Fixed hyperparameters include the scaling method (MinMax), optimization algorithm (Adam), activation functions, loss function, dropout rate, and training control parameters such as maximum epochs and early stopping patience. Adjustable hyperparameters concern the architectural structure of the network, namely the number of hidden layers and the number of nodes per layer, which were optimized within predefined ranges. The choice of Mean Squared Logarithmic Error (MSLE) as loss function was motivated by its robustness in handling

Table 7. Fixed and Adjustable hyperparameters of trained DANNs

Fixed hyperparameter	Choices
Input scaler [−]	Min-Max
Output scaler [−]	Min-Max
Optimization algorithm [−]	Adam
Activation function of the hidden layers [−]	Rectified Linear Unit
Activation function of the output layer [−]	Linear
Loss function [−]	MSLE
Dropout [%]	25%
Maximum number of epochs [epoch]	1,500
Patience for early stopping [epoch]	100
Batch size [−]	32

Adjustable hyperparameter	Minimum value	Maximum value
Number of hidden layers [−]	1	3
Number of nodes per layer [−]	8	100

proportional deviations in price prediction, particularly in the presence of heterogeneous price distributions across different market segments.

Optimization library HyperOpt was used, which assessed the best performing ensemble by measuring the performance of each ensemble on the validation dataset. The best ensemble identified by HyperOpt consisted of DANNs with 1 hidden layer and 68 nodes. The hidden layer employs a Rectified Linear Unit (ReLU) activation function, enabling the model to capture non-linear relationships while mitigating vanishing gradient issues, while the output layer uses a linear activation to produce continuous price predictions. The model is trained using the Adam optimizer, a computationally efficient stochastic gradient descent variant with adaptive learning rates, and the Mean Squared Logarithmic Error (MSLE) loss function, which penalizes proportional rather than absolute deviations and is therefore well-suited to heterogeneous price distributions. Early stopping is applied after 100 epochs without improvement on the validation set (maximum 1500 epochs total) to prevent overfitting and ensure generalization capacity.

The predictive performance of this optimized ensemble architecture is reported in Table 8, which presents accuracy metrics computed on the held-out test dataset. The table presents the predictive performance metrics of the selected ensemble, computed on the test dataset. Error indicators (Mean Absolute Error - MAE -, Mean Squared Error – MSE -, Root Mean Squared Error - RMSE) are expressed in euros per square metre (€/m^2), ensuring direct interpretability of prediction deviations relative to observed market prices. The Adjusted R-Squared (R^2) measures the proportion of variance explained by the model, adjusted for model complexity.

The MAE of 202.91 €/m^2 and the RMSE of 334.28 €/m^2 should be interpreted in relation to the distribution of observed unitary prices in the dataset. Considering that the average

Table 8. Figures of accuracy performance of the chosen ensemble of DANNs

Figure	Value
Mean Absolute Error (MAE)	202.91 €/m^2
Mean Squared Error (MSE)	111,743.96 €/m^2
Root Mean Squared Error (RMSE)	334.28 €/m^2
Mean Absolute Percentage Error (MAPE)	0.10
Adjusted R-Squared (R^2)	0.94

unitary market value in the sample exceeds 2,000 €/m^2 (as reported in the descriptive statistics section), the RMSE corresponds to approximately 15–17% of the mean value, indicating a contained prediction dispersion relative to overall market variability. Equivalently, the Mean Absolute Percentage Error (MAPE) of 0.10 confirms that, on average, prediction errors remain around 10% of observed values. The high Adjusted R^2 of 0.94 indicates that the model captures 94% of the structural price variation across the national residential sample. These performance levels are consistent with large-scale hedonic and machine learning applications in RE modelling, where intrinsic spatial and qualitative heterogeneity typically prevents near-perfect prediction accuracy. These metrics are further supported by international benchmarking. Recent international studies, reported in the Literature Review Section, applying ensemble learning methods show that the observed MAPE of 10% is consistent with the performance range reported for heterogeneous national-scale datasets, though positioned toward its upper bound (Bjørgve *et al.*, 2026; Moreno-Foronda *et al.*, 2025). From a professional appraisal perspective, the International Association of Assessing Officers recommends a Coefficient of Dispersion below 15% for residential properties, with which the achieved MAPE is broadly aligned (IAAO, 2013). Similarly, the Adjusted R^2 of 0.94 is consistent with published ensemble AVM studies, where R^2 values typically range between 0.85 and 0.95. These comparisons confirm that the proposed deep learning ensemble satisfies both academic benchmarking standards and professional mass appraisal expectations.

4.4 Phase IV – application to the RE over the considered territory

The resulting ensemble of DANNs [10] has been used to develop forecasts [11] about the market value of the RE stock because of EPBD IV. The results of the five selected Scenarios are depicted in Table 9, based on the assumptions previously described, offering valuable insights into the potential economic impacts of EE policies on the Italian RE market. It is important to clarify that the projected variations reported in Table 9 do not represent marginal energy-efficiency price premiums at the individual property level, as commonly estimated in hedonic techniques. Instead, they reflect aggregate market value changes under systemic regulatory scenarios affecting a large share of the national residential building stock. In particular, Scenario SA.1 represents an intentionally extreme stress-test assumption in which non-compliant properties are temporarily excluded from the market, while Scenario SC reflects a structural transition scenario in which the same portion of the stock is upgraded to the highest energy standard, together with the consequent improved maintenance conditions. Therefore, the reported percentage variations should be interpreted as macro-level revaluation effects under rather different regulatory transition assumptions (stress test), in order to define the minimum and maximum baselines in the overall evaluation of the considered building stock. Accordingly, the scenario results should be interpreted comparatively, i.e. as relative policy-sensitive exposure ranges, rather than as exact predictions of future transaction outcomes.

Table 9. Overall value of the RE units collected in the five Scenarios

Scenario	Overall M_V [B€]	% Compared to scenario 0	Notes
S_0	21.499	–	Calculated via records collected
	21.586	–	Calculated via the ensemble of DANNs
$S_{A.1}$	5.662	26.3%	Calculated via records collected
	5.667	26.2%	Calculated via the ensemble of DANNs
$S_{A.2}$	14.129	65.7%	Calculated via records collected
	14.104	65.3%	Calculated via the ensemble of DANNs
S_B	24.008	111.2%	Calculated via the ensemble of DANNs
S_C	27.361	126.8%	Calculated via the ensemble of DANNs

In the worst-case scenario ($S_{A.1}$), where energy retrofit subsidies are inadequate, approximately 74% of the overall value could be frozen. This dramatic reduction highlights the critical importance of well-designed subsidy programs in facilitating the transition to more Energy Efficient buildings. This scenario underscores the potential for significant market disruption if property owners are unable to meet compliance requirements due to financial constraints. $S_{A.2}$ presents a slightly less severe but still concerning outcome. In this case, if energy retrofit costs are discounted from the market value of non-compliant Real Estate units, the overall value would decrease by approximately 35%. This scenario reflects a more nuanced market response, where the costs of necessary upgrades are factored into property valuations. More optimistically, the model suggests that adequate energy retrofit subsidies could increase the overall value by approximately 11% (S_B). This finding emphasizes the potential for well-designed policy interventions to not only achieve environmental goals but also to stimulate economic value in the RE sector. The best-case scenario (S_C) presents an even more optimistic picture. If energy retrofits aimed at achieving zero emissions are implemented for 78.2% of the worst-performing RE units, the overall value could increase by about 27%. This scenario illustrates the significant potential for value creation through comprehensive EE improvements.

A particularly noteworthy observation is the substantial difference of approximately 70% between $S_{A.2}$ and S_B. This large gap underscores the critical role that policy and market responses play in determining outcomes. It's important to consider that $S_{A.2}$, while representing a worst-case scenario, assumes market values are adequately discounted based on energy retrofit costs. However, as compliance deadlines approach, a "panic-for-compliance" phenomenon could further depress the market values of non-compliant units, potentially pushing the overall value towards the more severe outcome of $S_{A.1}$. Conversely, S_B may be conservative in its projections. To achieve full compliance by 2050, even more comprehensive retrofitting might be necessary, potentially driving results closer to those of S_C, especially if adequate subsidies are in place. The table also provides a comparison between values calculated from collected records and those derived from the ensemble of DANNs, showing close alignment and thus validating the model's accuracy. This consistency lends credibility to the forecasts for S_B and S_C, which are based solely on the DANN ensemble.

Given the critical role of retrofit costs in determining scenario outcomes, particularly the 70% gap between $S_{A.2}$ and S_B, and recognizing that cost assumptions may vary depending on technological choices, market conditions, and implementation contexts, a formal sensitivity analysis was performed to assess the robustness of the modelling results. To evaluate the robustness of the modelling results, a one-parameter sensitivity analysis was performed on the retrofit cost assumption adopted in Scenario SA.2. The baseline unitary cost of 1,039 €/m^2 was varied by a maximum of $\pm$20%, capturing plausible fluctuations in construction costs, technological choices, and implementation conditions. For each cost configuration, the overall market value was recalculated under the same scenario structure as in Table 10.

In scenario $S_{A.2}$, at the baseline cost estimate of 1,039 €/m^2, the overall market value is projected to decrease by approximately 34.4% relative to the current market conditions (Scenario S_0). Scenario $S_{A.2}$ represents the most realistic configuration under assumptions of low renovation rates and market-driven residual value calculations, making it a particularly relevant reference for policy evaluation. Since $S_{A.2}$ outcomes are strictly dependent on renovation costs, the sensitivity analysis reveals that the projected market value decline of 34.4% varies within a relatively narrow range: $+6.3\%$ points under lower costs (-20% scenario) and -5.7% points under higher costs ($+20\%$ scenario). Consequently, while retrofit cost uncertainty introduces some variability, the aggregate market impact remains structurally stable, with the central estimate of approximately one-third market value loss holding across plausible cost configurations. These findings suggest that the main conclusions of the study are not driven by a single point estimate of retrofit cost but remain structurally robust within a realistic range of cost variability.

Table 10. Sensitivity analysis on retrofit costs

Scenario	Costs variation [%]	Costs [€/m^2]	Overall M_V [B€]	% Compared to S_0	% decrease in total MV
$S_{A.2}$	−20%	831	15.54	72.0%	−28.0%
	−15%	883	15.18	70.3%	−29.7%
	−10%	935	14.84	68.7%	−31.3%
	−5%	987	14.50	67.2%	−32.8%
	Best estimate	1,039	14.17	65.6%	−34.4%
	+5%	1,091	13.84	64.1%	−35.9%
	+10%	1,143	13.53	62.7%	−37.3%
	+15%	1,195	13.23	61.3%	−38.7%
	+20%	1,247	12.93	59.9%	−40.1%

Having established the robustness of the quantitative estimates to cost assumptions, it is essential to contextualize the broader interpretative framework within which these results should be understood. The implications discussed in this study should be interpreted in light of the scenario-based and assumption-driven nature of the modelling framework. The estimated variations in overall market value do not represent predictive forecasts, but rather conditional outcomes derived from specific regulatory and market configurations. The magnitude of the projected effects depends on assumptions regarding compliance thresholds, retrofit costs, and market adjustment mechanisms. While the direction of the results highlights the structural relevance of regulatory design, further research incorporating formal sensitivity analysis on additional parameters, including compliance timelines, subsidy availability, and behavioural response elasticities, and alternative modelling strategies would be necessary to support direct policy prescriptions. In this sense, the present work should be understood as an exploratory quantitative assessment aimed at informing the debate on energy regulation and market dynamics, rather than as a definitive evaluation of future market trajectories.

In conclusion, these results highlight the profound impact that EE policies and their implementation can have on the RE market. They underscore the importance of well-designed subsidy programs and the potential for significant value creation through EE improvements. However, they also warn of the risks of inadequate policy responses, which could lead to substantial market value losses. The wide range of outcomes documented across scenarios, from 26% market value retention ($S_{A.1}$) to 127% appreciation (S_C), illustrates that regulatory design, enforcement intensity, and financial support mechanisms are not merely technical details but structural determinants of market trajectories. The relative stability of scenario $S_{A.2}$ outcomes across cost variations further suggests that policy uncertainty regarding implementation timelines and compliance thresholds may be more consequential than cost fluctuations in shaping market expectations. These findings provide crucial insights for policymakers, investors, and property owners as they navigate the transition to a more Energy Efficient building stock.

5. Conclusions

This study investigated the potential economic consequences of enhanced energy performance standards on the market value of the Italian residential Real Estate stock. The analysis was conducted within the implementation framework of the Energy Performance of Buildings Directive (EPBD IV), thereby explicitly linking valuation dynamics to the current European regulatory transition. The study developed a comprehensive database of Italian RE and employed a Residential building energy efficient VAluation Model (Re-eVAM) using Deep Artificial Neural Networks (DANNs). Five potential scenarios of EE implementation were considered, each revealing varying impacts on RE stock values. Rather than providing point-estimate forecasts of individual property values, the Re-eVAM framework is designed as a macro-level analytical instrument to simulate structural value reallocation under regulatory transition. More precisely, the scenario results should be interpreted as deterministic stress-test configurations rather than probabilistic forecasts: the objective is to explore the potential upper and lower bounds of market exposure under alternative regulatory implementations of EPBD IV. The reported variations therefore represent structural outcome ranges conditional on specific assumptions, particularly the share of buildings considered non-compliant, assumed retrofit costs, and market reaction mechanisms (temporary exclusion, value discounting, or full upgrading), rather than predicted market trajectories. Adjusting any of these parameters would change the magnitude of estimated effects, but not their overall direction. The purpose is therefore to illustrate how different regulatory designs can systematically expand or contract aggregate market value, with the structural conclusion remaining stable even if specific numerical values vary under alternative assumptions. Its contribution lies in quantifying potential distributional and systemic implications of energy performance standards at national scale. Beyond the Italian case study, the broader implication of this work lies in the definition of a transferable research framework that combines large-scale RE data mining, deep learning-based valuation, and regulatory scenario simulation. In this sense, Re-eVAM can support future investigations on comparative housing markets, cross-country EPBD IV scenario analysis, and the development of AI-based decision-support models for sustainable built environment transitions.

The results highlighted a significant difference between the worst-case and best-case scenarios, underscoring the substantial economic implications of EE improvements. The magnitude and heterogeneity of these effects suggest that EPBD IV implementation may produce differentiated wealth impacts across building typologies, territorial contexts, and ownership structures, potentially reshaping relative price equilibria within the residential market. These findings clarify that energy retrofit strategies can be assessed not only as technical upgrading measures, but also as valuation-sensitive interventions relevant to investment screening, market-risk assessment, subsidy calibration, and territorial policy design. Energy retrofitting should therefore not be interpreted solely as a regulatory compliance cost, but as a structural factor influencing capital allocation decisions within the housing stock. Such consideration would allow for a more equitable distribution of resources across a larger public and prevent over-emphasis of certain measures. From the perspective of property investors, the results indicate that retrofit interventions must be evaluated within a risk-adjusted investment framework, where expected value premiums are assessed against intervention costs, holding periods, and local market liquidity conditions. The evidence shows that value gains are scenario-dependent and not automatically sufficient to offset capital expenditures, thus requiring integrated financial feasibility analysis. For urban policymakers, the findings highlight the importance of anticipating spatial redistribution effects. Uniform regulatory enforcement without calibrated support mechanisms may disproportionately affect lower-value or peripheral markets, amplifying territorial disparities. The integration of energy transition policies with urban regeneration and social housing strategies becomes therefore essential to ensure balanced implementation. At the national level, the significant variation observed across scenarios underscores the need for differentiated and fiscally sustainable subsidy schemes. Ex -ante simulation tools such as Re-eVAM can support governments in stress-testing alternative regulatory pathways and in calibrating incentive mechanisms according to building characteristics, transition gaps, and market absorption capacity. In this sense, the model provides an operational decision-support framework for the design of effective retrofit programmes.

Some limitations should be acknowledged. The study relies on advertised asking prices rather than actual transaction data, and focuses exclusively on major Italian cities, which may limit generalizability to other contexts. Additionally, the analysis uses the current EPC classification system, while EPBD IV will introduce a new framework once finalized. The scenarios are necessarily based on assumptions regarding retrofit implementation, costs, and subsidy mechanisms, given the ongoing regulatory transition. While we do not provide formal probabilistic uncertainty bounds, it is important to clarify that the magnitude of the projected variations depends directly on this limited set of modelling assumptions. For instance, lower retrofit costs or gradual compliance mechanisms would reduce the severity of negative outcomes, while more partial upgrading strategies would moderate the positive scenarios. Finally, the model's predictive capacity is calibrated to historical data and may require recalibration as market and regulatory conditions evolve. While this relies on numerous assumptions due to the wide range of options available in improving building energy performance and RE market trends, it provides valuable insights for stakeholders in preparing and fine-tuning energy retrofit strategies. Beyond the Italian case, the methodological architecture of Re-eVAM is adaptable to other EU Member States, provided that country-specific datasets on building characteristics, transaction prices, energy classifications, and retrofit cost structures are available. This adaptability also opens further research applications in comparative policy evaluation, territorial benchmarking of energy-transition risks, and multi-country assessment of capitalization effects associated with minimum energy performance standards. Given the structural heterogeneity of European housing systems, the broader implementation of EPBD IV is likely to generate asymmetric impacts across national building stocks; the proposed framework enables comparative scenario analysis and supports a more coordinated European transition strategy.

Future research should incorporate formal uncertainty quantification and sensitivity analyses to assess how alternative retrofit costs, compliance rates, and market adjustment mechanisms may influence the magnitude of the projected effects. Additionally, future work should focus on analyzing the RE premium in relation to the costs of EE interventions, and the energy savings achieved over the property's lifetime. This analysis would help determine whether immediate sale to capitalize on the value increase or long-term use to accumulate energy savings is more beneficial, considering the initial retrofit costs. Additionally, further studies should examine the Life Cycle Cost of these scenarios, evaluating the economic and financial sustainability of energy retrofitting interventions against their potential value increase. Future work should also i) address the current study's limitations by incorporating actual transaction data, extending geographical coverage to non-urban areas, and adapting the model once the new EPC classification system under EPBD IV is finalized, and ii) incorporate credit market responses, collateral valuation adjustments, and dynamic modelling of phased policy implementation, in order to capture transitional adjustment processes rather than static scenario endpoints.

Despite its limitations, this study represents a significant step towards understanding the complex interplay between EE improvements and RE market dynamics in Italy. Overall, the findings indicate that EPBD IV implementation constitutes not merely an environmental compliance requirement, but a structural economic transformation of residential capital values, with implications not only for investors, policymakers, and national governments, but also for future research on AI-based valuation, regulatory stress testing, and sustainable built environment transitions.

Ethics statement

This research did not involve human participants or animals. The study relies exclusively on publicly available Real Estate listing data. No personal data, identifiable information, or interactions with human subjects were involved in this research. Therefore, ethical approval was not required for this study.

AI assistance disclosure

The authors confirm that AI-based tools (Grammarly and Lucrez-IA) were used solely for proofreading and identifying typographical errors in this manuscript. No AI tools were used for draughting, content generation, data analysis, interpretation of results, or any substantive aspects of the research. All intellectual contributions, including research design, methodology, analysis, and scientific writing, were conducted entirely by the authors.

Annex

Table A1. List of the municipalities considered by the web crawler, ordered by region and population

Region	Municipalities considered
Abruzzo	Pescara, L'Aquila, Teramo, Chieti, Montesilvano, Avezzano, Vasto, Lanciano, Roseto degli Abruzzi, Sulmona, Francavilla al Mare, Ortona, Giulianova
Basilicata	Potenza, Matera
Calabria	Reggio di Calabria, Catanzaro, Corigliano-Rossano, Lamezia Terme, Cosenza, Crotone, Rende, Vibo Valentia, Castrovillari
Campania	Napoli, Salerno, Giugliano in Campania, Torre del Greco, Pozzuoli, Casoria, Caserta, Castellammare di Stabia, Afragola, Benevento, Marano di Napoli, Acerra, Portici, Avellino, Cava de' Tirreni, Ercolano, Aversa, Battipaglia, Scafati, Casalnuovo di Napoli, Nocera Inferiore, San Giorgio a Cremano, Torre Annunziata, Marcianise, Pomigliano d'Arco, Maddaloni, Quarto, Eboli, Caivano, Melito di Napoli, Arzano, Pagani, Somma Vesuviana, Mugnano di Napoli, Sant'Antimo, Nola, Angri, Santa Maria Capua Vetere, Sarno, Marigliano, Frattamaggiore, Villaricca, Gragnano, San Giuseppe Vesuviano, Boscoreale, Sant'Anastasia, Mondragone, Bacoli, Pompei, Pontecagnano Faiano, Orta di Atella, Qualiano, Nocera Superiore, Ottaviano, Volla, Castel Volturno, Ariano Irpino, Cardito, Sessa Aurunca
Emilia-Romagna	Bologna, Modena, Parma, Reggio nell'Emilia, Ravenna, Rimini, Ferrara, Forlì, Piacenza, Cesena, Imola, Carpi, Faenza, Sassuolo, Casalecchio di Reno, Cento, Riccione, Formigine, Lugo, Castelfranco Emilia, San Lazzaro di Savena, Valsamoggia, Cervia, San Giovanni in Persiceto, Fidenza, Cesenatico, Correggio, Scandiano, Vignola, Mirandola, Comacchio, Argenta
Friuli-Venezia Giulia	Trieste, Udine, Pordenone, Gorizia, Monfalcone
Lazio	Roma, Latina, Guidonia Montecelio, Fiumicino, Aprilia, Viterbo, Pomezia, Tivoli, Velletri, Civitavecchia, Anzio, Frosinone, Rieti, Nettuno, Terracina, Ardea, Monterotondo, Albano Laziale, Marino, Ladispoli, Ciampino, Fondi, Formia, Cisterna di Latina, Cerveteri, Cassino, Fonte Nuova, Alatri, Sora, Sezze, Genzano di Roma, Ceccano
Liguria	Genova, La Spezia, Savona, Sanremo, Imperia, Rapallo, Chiavari, Ventimiglia, Albenga
Lombardia	Milano, Brescia, Monza, Bergamo, Como, Varese, Busto Arsizio, Sesto San Giovanni, Cinisello Balsamo, Cremona, Pavia, Vigevano, Legnano, Gallarate, Rho, Lecco, Mantova, Paderno Dugnano, Cologno Monzese, Lodi, Seregno, Lissone, Desio, Rozzano, Cantù, Saronno, Voghera, Cesano Maderno, San Giuliano Milanese, Bollate, Pioltello, Limbiate, Corsico, Segrate, Brugherio, Crema, Abbiategrasso, San Donato Milanese, Cernusco sul Naviglio, Treviglio, Desenzano del Garda, Parabiago, Buccinasco, Garbagnate Milanese, Bresso, Vimercate, Lainate, Giussano, Seriate, Montichiari, Mariano Comense, Cesano Boscone, Lumezzane, Muggiò, Meda, Dalmine, Magenta, Seveso, Nova Milanese, Peschiera Borromeo, Castiglione delle Stiviere
Marche	Ancona, Pesaro, Fano, Ascoli Piceno, San Benedetto del Tronto, Senigallia, Macerata, Jesi, Civitanova Marche, Fermo, Osimo, Fabriano, Falconara Marittima, Porto Sant'Elpidio
Molise	Campobasso, Termoli
Piemonte	Torino, Novara, Alessandria, Asti, Moncalieri, Cuneo, Collegno, Rivoli, Nichelino, Settimo Torinese, Vercelli, Biella, Grugliasco, Chieri, Pinerolo, Casale Monferrato, Venaria Reale, Alba, Verbania, Bra, Carmagnola, Novi Ligure, Tortona, Chivasso, Fossano, Ivrea, Orbassano, Mondovì
Puglia	Bari, Taranto, Foggia, Andria, Barletta, Lecce, Brindisi, Altamura, Molfetta, Cerignola, Bitonto, Manfredonia, Trani, San Severo, Bisceglie, Martina Franca, Monopoli, Corato, Gravina in Puglia, Fasano, Modugno, Francavilla Fontana, Lucera, Grottaglie, Massafra, Ostuni, Nardò, Manduria, Canosa di Puglia, Gioia del Colle, Mesagne, San Giovanni Rotondo, Galatina, Putignano, Triggiano, Terlizzi, Santeramo in Colle, Noicattaro, Conversano, Ruvo di Puglia, Mola di Bari, Copertino, Ginosa

(*continued*)

Table A1. Continued

Region	Municipalities considered
Sardegna	Cagliari, Sassari, Quartu Sant'Elena, Olbia, Alghero, Nuoro, Oristano, Carbonia, Selargius, Iglesias, Assemini, Capoterra, Porto Torres
Sicilia	Palermo, Catania, Messina, Siracusa, Marsala, Gela, Ragusa, Trapani, Caltanissetta, Vittoria, Agrigento, Bagheria, Modica, Acireale, Mazara del Vallo, Paternò, Misterbianco, Alcamo, Barcellona Pozzo di Gotto, Sciacca, Licata, Caltagirone, Monreale, Augusta, Carini, Adrano, Canicattì, Favara, Milazzo, Castelvetrano, Partinico, Avola, Mascalucia, Comiso, Aci Catena, Giarre, Erice, Niscemi, Enna, Misilmeri, Gravina di Catania, Belpasso, Termini Imerese, Scicli, Lentini, Noto, Biancavilla, Palma di Montechiaro, San Cataldo, Floridia, Piazza Armerina, Pachino, San Giovanni la Punta
Toscana	Firenze, Prato, Livorno, Arezzo, Pistoia, Lucca, Pisa, Grosseto, Massa, Carrara, Viareggio, Siena, Scandicci, Sesto Fiorentino, Empoli, Capannori, Cascina, Campi Bisenzio, Piombino, Camaiore, Rosignano Marittimo, San Giuliano Terme, Poggibonsi, Pontedera, Cecina, San Miniato, Bagno a Ripoli, Quarrata, Pietrasanta, Montevarchi, Figline e Incisa Valdarno, Fucecchio, Cortona, Massarosa
Trentino-Alto Adige	Trento, Bolzano, Rovereto, Merano
Umbria	Perugia, Terni, Foligno, Città di Castello, Spoleto, Gubbio, Assisi
Valle d'Aosta	Aosta
Veneto	Venezia, Verona, Padova, Vicenza, Treviso, Rovigo, Chioggia, Bassano del Grappa, San Donà di Piave, Schio, Mira, Belluno, Conegliano, Castelfranco Veneto, Villafranca di Verona, Montebelluna, Vittorio Veneto, Mogliano Veneto, Spinea, Valdagno, Mirano, Arzignano, Portogruaro, Legnago, Jesolo, San Giovanni Lupatoto, Albignasego, Montecchio Maggiore, Thiene, Selvazzano Dentro

Notes

1. This Directive will undergo formal approval by the Council, and the implementation is expected in 2026 upon publication in the EU Official Journal.

2. A limitation of the study concerns the use of online listing prices rather than finalized transaction values. Asking prices may deviate from sale prices due to negotiation dynamics, strategic overpricing, and local demand–supply conditions, generating potential measurement error. In prime micro-locations, transactions may occur at a premium relative to listings, whereas in weaker or peripheral markets discounts are more frequent. This may introduce both systematic bias (if deviations are structurally correlated with specific market segments) and random noise. However, given the large-scale dataset and the focus on comparative scenario analysis rather than precise transaction price estimation, such distortions are unlikely to compromise the robustness of the results. Random measurement error tends to attenuate estimated coefficients rather than generate spurious relationships, while opposing premium and discount effects across locations may partially offset in aggregate modelling. Listing data are widely adopted in large-scale hedonic and machine learning real estate modelling due to their availability, granularity, and ability to capture market expectations in real time. In this perspective, they remain particularly suitable for analyzing regulatory transition scenarios such as those related to evolving energy performance standards.

3. Although numerically encoded for modelling purposes, EC does not represent a continuous variable with equal interval distances in terms of primary energy consumption. The numerical encoding was adopted to preserve the natural ordering of the classification system, without implying linear marginal effects between adjacent classes.

4. In the modelling framework, the target variable predicted by the DANN is the unitary market value expressed in euros per square metre (€/m^2), rather than the total listing price. The total asking price was therefore normalized by the Floor Area (A_{Floor}) to reduce heteroskedasticity effects associated with property size and to enhance comparability across heterogeneous dwellings. No logarithmic transformation was applied to the dependent variable, as preliminary tests showed stable variance behaviour and satisfactory model performance in level terms. All continuous predictors were

standardized prior to training to improve convergence stability, while the predicted output was subsequently rescaled to its original monetary unit (€/m^2) for interpretability.

5. In this study, the term “overall value” refers to the aggregated asking prices of the sampled listings rather than finalized transaction prices. Although asking prices may deviate from effective transaction values due to negotiation dynamics, strategic pricing behaviour, and local demand–supply conditions, such deviations may occur in both positive (premium) and negative (discount) directions. Given the large sample size and the focus on comparative and scenario-based analysis rather than precise transaction-level appraisal, aggregated asking prices are considered a reasonable proxy for market value dynamics. At scale, systematic price formation patterns tend to prevail over level negotiation noise, allowing the estimation of relative value differentials and structural effects with acceptable approximation reliability.
6. It is important to distinguish between the provisions explicitly stated in the EPBD IV Directive and the modelling assumptions adopted in this study. While the Directive establishes minimum energy performance targets and progressive renovation obligations, it does not specify market value adjustments nor quantify the exact share of buildings requiring retrofit within a defined timeframe. The translation of regulatory objectives into quantitative market scenarios therefore necessarily involves modelling assumptions. In this framework, the 78.2% threshold is not arbitrarily imposed but derives from the empirical distribution of Energy Classes in the national dataset, corresponding to the cumulative share of units below the assumed minimum performance level. The scenarios should thus be interpreted as analytical stress-testing configurations based on the observed structure of the Italian housing stock, rather than as deterministic forecasts mandated by the Directive.
7. Although deep neural networks are often characterized as “black-box” models, as specified in the Literature Section, interpretability in this study is achieved through structured post-estimation analysis of predicted price distributions across Energy Classes (EC), maintenance levels (SoM), and scenario configurations. By comparing conditional predicted means and price gradients across categories, the model enables the identification of systematic valuation patterns embedded in the data. The objective is therefore not to extract individual parameter coefficients, but to capture non-linear valuation structures relevant for large-scale policy analysis.
8. Although the DANN operates as a non-linear predictive model, descriptive post-estimation analysis of predicted price distributions across Energy Classes and maintenance levels was conducted to improve interpretability. These comparisons allow the identification of systematic pricing gradients across categories, without implying direct causal transmission mechanisms.
9. To ensure methodological robustness, model development followed a structured validation protocol including k-fold cross-validation, hyperparameter optimization via HyperOpt, and early stopping criteria to prevent overfitting. Ensemble averaging was adopted to reduce variance and improve generalization performance relative to single-network specifications. This design choice is consistent with bias–variance trade-off principles in machine learning and enhances model stability when applied to heterogeneous national-scale housing data.
10. The modelling framework adopted in this study is predictive rather than causal. The DANN ensemble estimates conditional associations between property characteristics and observed market prices within the available dataset. Therefore, references to “impacts”, “effects”, or “marginal contributions” should be interpreted as predictive marginal contributions within the model structure, and not as causal treatment effects in a strict econometric sense. The analysis does not rely on exogenous variation or identification strategies designed to isolate causal mechanisms; rather, it aims to capture structural price patterns embedded in observed market data. Consequently, the results should be interpreted as scenario-based predictive simulations of potential market adjustments under alternative regulatory configurations.
11. The reliability of the scenario projections depends on the predictive accuracy of the valuation model. As reported in Table 8, the optimized DANNs ensemble achieves high explanatory power (Adjusted R^2 = 0.94) and low prediction error, supporting its use for scenario simulation purposes.

References

Ahmad, T., Chen, H., Guo, Y. and Wang, J. (2018), "A comprehensive overview on the data driven and large scale based approaches for forecasting of building energy demand: a review", *Energy and Buildings*, Vol. 165, pp. 301-320, doi: 10.1016/j.enbuild.2018.01.017.

Anselin, L., Syabri, I. and Kho, Y. (2010), in Fischer, M.M. and Getis, A. (Eds), *GeoDa: An Introduction to Spatial Data Analysis BT - Handbook of Applied Spatial Analysis: Software Tools, Methods and Applications*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 73-89, doi: 10.1007/978-3-642-03647-7_5.

Araújo, C., Almeida, M., Bragança, L. and Barbosa, J.A. (2016), "Cost–benefit analysis method for building solutions", *Applied Energy*, Vol. 173, pp. 124-133, doi: 10.1016/j.apenergy.2016.04.005.

Ascione, F., De Rossi, F. and Vanoli, G.P. (2011), "Energy retrofit of historical buildings: theoretical and experimental investigations for the modelling of reliable performance scenarios", *Energy and Buildings*, Vol. 43 No. 8, pp. 1925-1936, doi: 10.1016/j.enbuild.2011.03.040.

Bandelow, N.C., Hornung, J., Sager, F. and Schröder, I. (2023), "Energy efficiency, housing, and economic policy", *European Policy Analysis*, Vol. 9 No. 3, pp. 196-199, doi: 10.1002/epa2.1185.

Bisello, A., Antoniucci, V. and Marella, G. (2020), "Measuring the price premium of energy efficiency: a two-step analysis in the Italian housing market", *Energy and Buildings*, Vol. 208, 109670, doi: 10.1016/j.enbuild.2019.109670.

Bjørgve, E., Oust, A., Pollestad, A.J., Sandnes, C. and Sønstebø, O.J. (2026), "Comparing housing valuation techniques and stacked generalization: exploiting explainable AI", *The Journal of Real Estate Finance and Economics*, Vol. 72 No. 3, pp. 640-665, doi: 10.1007/s11146-025-10030-x.

Bottero, M., Bravi, M., Mondini, G. and Talarico, A. (2017), "Buildings energy performance and real estate market value: an application of the spatial auto regressive (SAR) model", *Green Energy and Technology*, Vol. 0, pp. 221-230, doi: 10.1007/978-3-319-49676-4_16.

Bottero, M., Bravi, M., Dell'Anna, F. and Mondini, G. (2018), "Valuing buildings energy efficiency through Hedonic Prices Method: are spatial effects relevant?", *Valori e Valutazioni*, Vol. 2018 No. 21, pp. 27-39.

Camera dei Deputati (2024), "La dimensione economica del superbonus", 29 May, available at: https://temi.camera.it/leg19/post/la-dimensione-economica-del-superbonus.html#:~:text=In%20base%20agli%20ultimi%20dati,di%20investimenti%20ammessi%20a%20detrazione (accessed 6 December 2024).

Canesi, R. (2022), "Urban policy sustainability through a value-added densification tool: the case of the South Boston area", *Sustainability (Switzerland)*, Vol. 14 No. 14, p. 8762, doi: 10.3390/su14148762.

Canesi, R. and Marella, G. (2023), "Towards European transitions: indicators for the development of marginal urban regions", *Land*, Vol. 12 No. 1, p. 27, doi: 10.3390/land12010027.

Chen, W. and Lai, J. (2025), "Performance assessment of residential building renovation: a scientometric analysis and qualitative review of literature", *Smart and Sustainable Built Environment*, Vol. 14 No. 3, pp. 625-648, doi: 10.1108/SASBE-09-2023-0276.

De Ayala, A., Galarraga, I. and Spadaro, J.V. (2016), "The price of energy efficiency in the Spanish housing market", *Energy Policy*", Vol. 94, pp. 16-24, doi: 10.1016/j.enpol.2016.03.032.

Dell'Anna, F., Bottero, M., Becchio, C., Corgnati, S.P. and Mondini, G. (2020), "Designing a decision support system to evaluate the environmental and extra-economic performances of a nearly zero-energy building", *Smart and Sustainable Built Environment*, Vol. 9 No. 4, pp. 413-442, doi: 10.1108/SASBE-09-2019-0121.

Deppner, J. and Cajias, M. (2024), "Accounting for spatial autocorrelation in algorithm-driven hedonic models: a spatial cross-validation approach", *The Journal of Real Estate Finance and Economics*, Vol. 68 No. 2, pp. 235-273, doi: 10.1007/s11146-022-09915-y.

Du, B. and Wang, Y. (2023), "Real estate price evaluation system based on BP neural network algorithm", *Lecture Notes in Electrical Engineering*, Vol. 1031 LNEE, pp. 553-561, doi: 10.1007/978-981-99-1428-9_68.

El Jaouhari, A., Samadhiya, A., Kumar, A., Šešplaukis, A. and Raslanas, S. (2024), "Mapping the landscape: a systematic literature review on automated valuation models and strategic applications in real estate", *International Journal of Strategic Property Management*, Vol. 28 Nos 5 SE-Articles, pp. 286-301, doi: 10.3846/ijspm.2024.22251.

ENEA - Agenzia Nazionale Efficienza Energetica (2026), "Sistema Informativo sugli Attestati di Prestazione Energetica", *Sistema Informativo Sugli Attestati Di Prestazione Energetica (SIAPE)*, 24 February 2026.

European Commission (2019), "OVERVIEW | Decarbonising the non-residential building stock", *European Commission*, available at: https://build-up.ec.europa.eu/en/resources-and-tools/articles/overview-decarbonising-non-residential-building-stock (accessed 12 June 2025).

European Commission (2020), *Questions and Answers on the Renovation Wave*, European Commission, available at: https://ec.europa.eu/commission/presscorner/detail/en/qanda_20_1836 (accessed 12 June 2025).

European Commission (2024), "Energy performance of buildings directive", 2024, available at: https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en?etransnolive=1 (accessed 4 February 2025).

European Union (2012), "Commission Delegated Regulation (EU) No 244/2012 of 16 January 2012 supplementing Directive 2010/31/EU of the European Parliament and of the Council on the energy performance of buildings by establishing a comparative methodology framework for calculating", *Official Journal of the European Union*, European Union.

Eurostat (2022), "Owning or renting? What is the EU's housing situation?", *Eurostat*, available at: https://ec.europa.eu/eurostat/web/products-eurostat-news/-/wdn-20211230-1 (accessed 24 February 2026).

Gabrielli, L. and Ruggeri, A.G. (2022), "Sustainability and energy efficiency in twentieth-century Italian built heritage", in Naddeo, V., Choo, K.H. and Ksibi, M. (Eds), *Water-Energy-Nexus in the Ecological Transition, Advances in Science, Technology & Innovation*. Springer, Cham, doi: 10.1007/978-3-031-00808-5_2.

Galvin, R. (2023), "How prebound effects compromise the market premium for energy efficiency in German house sales", *Building Research and Information*, Vol. 51 No. 5, pp. 501-517, doi: 10.1080/09613218.2023.2176284.

Gevorgian, A.S., Pezzutto, S., Zambotti, S., Croce, S., Filippi Oberegger, U., Lollini, R., Kranzl, L. and Müller A. (2021), *European Building Stock Analysis*, in Research, E. (Ed.), Bolzano, Italy.

Gołąbeska, E. (2019), *The Impact of the Energy Efficiency of the Building to its Market Value*, Vol. 70, Ekonomia i Srodowisko, pp. 55-62, doi: 10.34659/2019/3/34.

Hulathdoowage, N.D., Karunasena, G., Udawatta, N. and Liu, C. (2025), "The applicability of stepwise retrofitting in the Australian residential sector: an exploratory qualitative study", *Smart and Sustainable Built Environment*, pp. 1-30, doi: 10.1108/SASBE-04-2025-0180.

IAAO - The International Association of Assessing Officers (2017), *Standard on Mass Appraisal of Real Property*, Kansas City, Missouri, ISBN 978-0-88329-248-8, available at: https://www.iaao.org/wp-content/uploads/StandardOnMassAppraisal.pdf (accessed 15 June 2026).

Kalliola, J., Kapočiūte-Dzikiene, J. and Damaševičius, R. (2021), "Neural network hyperparameter optimization for prediction of real estate prices in Helsinki", *PeerJ Computer Science*, Vol. 7, pp. 1-25, doi: 10.7717/peerj-cs.444.

Kamenders, A., Stivriņš, R. and Žogla, G. (2022), *Minimum Energy Performance Standards (MEPS) in the Residential Sector*, Riga Technical University, The European Economic and Social Committee (EESC), Riga, doi: 10.2864/304127.

Karytsas, S. and Theodoropoulou, E. (2023), "Awareness and utilization of incentive programs for household energy-saving renovations: empirical findings from Greece", *Sustainability (Switzerland)*, Vol. 15 No. 18, p. 13923, doi: 10.3390/su151813923.

Kaufmann, M., Veenman, S., Haarbosch, S. and Jansen, E. (2023), "How policy instruments reproduce energy vulnerability - a qualitative study of Dutch household energy efficiency measures", *Energy Research and Social Science*, Vol. 103, 103206, doi: 10.1016/j.erss.2023.103206.

Kholodilin, K.A., Mense, A. and Michelsen, C. (2017), “The market value of energy efficiency in buildings and the mode of tenure”, *Urban Studies*, Vol. 54 No. 14, pp. 3218-3238, doi: 10.1177/0042098016669464.

Lancaster, K.J. (1966), “A new approach to consumer theory”, *Journal of Political Economy*, Vol. 74 No. 2, pp. 132-157, doi: 10.1086/259131.

Li, W., Zhou, Y., Cetin, K., Eom, J., Wang, Y., Chen, G. and Zhang, X. (2017), “Modeling urban building energy use: a review of modeling approaches and procedures”, *Energy*, Vol. 141, pp. 2445-2457, doi: 10.1016/j.energy.2017.11.071.

Liu, Y., Liu, T., Ye, S. and Liu, Y. (2018), “Cost-benefit analysis for Energy Efficiency Retrofit of existing buildings: a case study in China”, *Journal of Cleaner Production*, Vol. 177, pp. 493-506, doi: 10.1016/j.jclepro.2017.12.225.

Liu, Z., Yu, C., Qian, Q.K., Huang, R., You, K., Visscher, H. and Zhang, G. (2023), “Incentive initiatives on energy-efficient renovation of existing buildings towards carbon–neutral blueprints in China: advancements, challenges and prospects”, *Energy and Buildings*, Vol. 296, 113343, doi: 10.1016/j.enbuild.2023.113343.

Maduta, C., D’Agostino, D., Tsemekidi-Tzeiranaki, S., Castellazzi, L., Melica, G. and Bertoldi, P. (2023), “Towards climate neutrality within the European union: assessment of the energy performance of buildings directive implementation in member states”, *Energy and Buildings*, Vol. 301, 113716, doi: 10.1016/j.enbuild.2023.113716.

Manganelli, B., Morano, P., Tajani, F. and Salvo, F. (2019), “Affordability assessment of energy-efficient building construction in Italy”, *Sustainability (Switzerland)*, Vol. 11 No. 1, p. 249, doi: 10.3390/su11010249.

Martínez-Molina, A., Tort-Ausina, I., Cho, S. and Vivancos, J.L. (2016), “Energy efficiency and thermal comfort in historic buildings: a review”, *Renewable and Sustainable Energy Reviews*, Vol. 61, pp. 70-85, doi: 10.1016/j.rser.2016.03.018.

Maselli, G. (2022), “Evaluating the impact of urban renewal on the residential real estate market: artificial neural networks versus multiple regression analysis”, *Lecture Notes in Networks and Systems*, Vol. 482 LNNS, pp. 702-712, doi: 10.1007/978-3-031-06825-6_66.

Mazzarella, L. (2015), “Energy retrofit of historic and existing buildings. The legislative and regulatory point of view”, *Energy and Buildings*, Vol. 95, pp. 23-31, doi: 10.1016/j.enbuild.2014.10.073.

Mecca, U., Moglia, G., Piantanida, P., Prizzon, F., Rebaudengo, M. and Vottari, A. (2020), “How energy retrofit maintenance affects residential buildings market value?”, *Sustainability (Switzerland)*, Vol. 12 No. 12, p. 5213, doi: 10.3390/su12125213.

Morano, P., Rosato, P., Tajani, F. and Di Liddo, F. (2020), *An Analysis of the Energy Efficiency Impacts on the Residential Property Prices in the City of Bari (Italy) BT - Values and Functions for Future Cities*, Springer International Publishing, Mondini, G., Oppio, A., Stanghellini, S., Bottero, M. and Abastante, F. (Eds), Cham, pp. 73-88, doi: 10.1007/978-3-030-23786-8_5.

Moreno-Foronda, I., Sánchez-Martínez, M.-T. and Pareja-Eastaway, M. (2025), *Comparative Analysis of Advanced Models for Predicting Housing Prices: A Review*, Urban Science, doi: 10.3390/urbansci9020032.

Moretti, N., Tagliabue, L.C., Dejaco, M.C. and Cecconi, F.R.E. (2019), “Location-based data driven model for real estate market value analysis based on energy performance certification”, *Journal of Physics: Conference Series*, Vol. 1343 No. 1, 012052, doi: 10.1088/1742-6596/1343/1/012052.

Mostofi, F., Toğan, V. and Başağa, H.B. (2022), “Real-estate price prediction with deep neural network and principal component analysis”, *Organization, Technology and Management in Construction*, Vol. 14 No. 1, pp. 2741-2759, doi: 10.2478/otmcj-2022-0016.

Nguyen, T.T., Nguyen, D.N. and Pham, H.T.L. (2023), “Survey data on energy-saving policies, energy price, crisis and household energy-saving behavior”, *Data in Brief*, Vol. 51, 109646, doi: 10.1016/j.dib.2023.109646.

Nikolov, V. (2023), “A neural network based approach for estimation of real estate prices”, *Lecture Notes in Networks and Systems*, Vol. 652 LNNS, pp. 474-481, doi: 10.1007/978-3-031-28073-3_34.

Nyanda, F., Muyingo, H. and Wilhelmsson, M. (2024), “Machine learning valuation in dual market dynamics: a case study of the formal and informal real estate market in dar es Salaam”, *Buildings*, Vol. 14 No. 10, p. 3172, doi: 10.3390/buildings14103172.

Pittarello, M., Scarpa, M., Ruggeri, A.G., Gabrielli, L. and Schibuola, L. (2021), “Artificial neural networks to optimize zero energy building (Zeb) projects from the early design stages”, *Applied Sciences (Switzerland)*, Vol. 11 No. 12, p. 5377, doi: 10.3390/app11125377.

Popescu, D., Bienert, S., Schützenhofer, C. and Boazu, R. (2012), “Impact of energy efficiency measures on the economic value of buildings”, *Applied Energy*, Vol. 89 No. 1, pp. 454-463, doi: 10.1016/j.apenergy.2011.08.015.

Rosen, S. (1974), “Hedonic prices and implicit markets: product differentiation in pure competition”, *Journal of Political Economy*, Vol. 82 No. 1, pp. 34-55, doi: 10.1086/260169.

Ruggeri, A.G., Calzolari, M., Scarpa, M., Gabrielli, L. and Davoli, P. (2020), “Planning energy retrofit on historic building stocks: a score-driven decision support system”, *Energy and Buildings*, Vol. 224, 110066, doi: 10.1016/j.enbuild.2020.110066.

Sajid, Z.W., Khan, S.A., Hussain, F., Ullah, F., Khushnood, R.A. and Soliman, N. (2026), “Assessing economic and environmental performance of infill materials through BIM: a life cycle approach”, *Smart and Sustainable Built Environment*, Vol. 15 No. 1, pp. 9-39, doi: 10.1108/SASBE-11-2023-0341.

Sarcina, A. and Canesi, R. (2023), “Renewable energy community: opportunities and threats towards green transition”, *Sustainability (Switzerland)*, Vol. 15 No. 18, 13860, doi: 10.3390/su151813860.

Streimikiene, D., Kyriakopoulos, G.L., Ślusarczyk, B. and Stankuniene, G. (2024), “Policies and measures for energy efficiency improvement at households: a bibliometric analysis”, *International Journal of Renewable Energy Development*, Vol. 13 No. 1, pp. 31-51, doi: 10.14710/ijred.2024.57769.

Surmann, M., Brunauer, W. and Bienert, S. (2015), “How does energy efficiency influence the Market Value of office buildings in Germany and does this effect increase over time?”, *Journal of European Real Estate Research*, Vol. 8 No. 3, pp. 243-266, doi: 10.1108/JERER-04-2015-0018.

Swan, L.G. and Ugursal, V.I. (2009), “Modeling of end-use energy consumption in the residential sector: a review of modeling techniques”, *Renewable and Sustainable Energy Reviews*, Vol. 13 No. 8, pp. 1819-1835, doi: 10.1016/j.rser.2008.09.033.

Tagliabue, L.C., Cecconi, F.R., Moretti, N. and Dejaco, M.C. (2019), “The influence of energy performance certification the market value of residential buildings”, *IOP Conference Series: Earth and Environmental Science*, Vol. 290 No. 1, 012062, doi: 10.1088/1755-1315/290/1/012062.

Topraklı, A.Y. (2024), “AI-driven valuation: a new era for real estate appraisal”, *Journal of European Real Estate Research*, Vol. 18 No. 1, pp. 105-120, doi: 10.1108/JERER-04-2024-0031.

von Malmborg, F., Björklund, M. and Nordensvärd, J. (2023), “Framing the benefits of European Union policy expansion on energy efficiency of buildings: a Swiss knife or a Trojan horse”, *European Policy Analysis*, Vol. 9 No. 3, pp. 219-243, doi: 10.1002/epa2.1184.

Wan, W.X. and Lindenthal, T. (2023), “Testing machine learning systems in real estate”, *Real Estate Economics*, Vol. 51 No. 3, pp. 754-778, doi: 10.1111/1540-6229.12416.

Weerasinghe, L.N.K., Darko, A. and Chan, A.P.C. (2025), “Impacts of and relationships between benefits of residential building net zero carbon retrofits”, *Smart and Sustainable Built Environment*, pp. 1-25, doi: 10.1108/SASBE-06-2025-0297.

Yakub, A.A., Hishamuddin, M.A., Kamalahasan, A., Abdul Jalil, R.B. and Salawu, A.O. (2021), “An integrated approach based on artificial intelligence using anfis and ANN for multiple criteria real estate price prediction", *Planning Malaysia*, Vol. 19 No. 3, pp. 270-282, doi: 10.21837/PM.V19I17.1005.

Zangheri, P., D'Agostino, D., Armani, R. and Bertoldi, P. (2022), "Review of the cost-optimal methodology implementation in member states in compliance with the energy performance of buildings directive", *Buildings*, Vol. 12 No. 9, p. 1482, doi: 10.3390/buildings12091482.

Zinzi, M., Pagliaro, F., Agnoli, S., Bisegna, F. and Iatauro, D. (2017), "Assessing the overheating risks in Italian existing school buildings renovated with nZEB targets", *Energy Procedia*, Vol. 142 No. February 2018, pp. 2517-2524, doi: 10.1016/j.egypro.2017.12.192.

Corresponding author
Rubina Canesi can be contacted at: rubina.canesi@unipd.it
