sustainability

MDPI

Article

Multi-Objective Screening of Bio-Based Phase Change Materials for Building Envelopes Using Surrogate Models Across Italian Climates

Maria Grazia Insinga [1,2,*], Alessandro Muratore [1], Filippo Carollo [1] and Giuseppe Aiello [1,*]

[1] Department of Engineering, University of Palermo, 90128 Palermo, Italy; alessandro.muratore01@unipa.it (A.M.); filippo.carollo@unipa.it (F.C.)
[2] Istituto Euro-Mediterraneo di Scienza e Tecnologia, 90143 Palermo, Italy
* Correspondence: mariagrazia.insinga@unipa.it (M.G.I.); giuseppe.aiello03@unipa.it (G.A.)

Abstract

Bio-based phase change materials (PCMs) can increase transient heat storage in lightweight building envelopes, but their performance depends on the climate, transition properties, layer design, and assumptions used to translate thermal loads into carbon and cost indicators. Although PCM optimization, machine learning surrogates, and lifecycle assessment have each been studied extensively, comparatively few studies combine them while explicitly separating simulation-derived thermal outputs from scenario-dependent environmental and economic post-processing and benchmarking bio-based candidates against paraffin on a common wall area basis. This study develops a simulation-based screening framework for a south-facing office wall model using 18,000 EnergyPlus cases, climate-specific machine learning surrogates, TreeSHAP interpretation, NSGA-II optimization, and scenario-based lifecycle carbon and cost accounting. XGBoost achieved pooled held-out R^2 values of 0.974 for occupied discomfort degree-hours and 0.978 for total annual thermal demand. For Palermo, the directly re-simulated balanced configuration (T_m = 24.8 °C, L_h = 178 kJ/kg, 22 mm thickness, intermediate position) reduced occupant discomfort by 42.6% and the modeled single-zone total thermal demand by 6.0%. Under the central all-electric scenario (SCOP = SEER = 3.0, grid factor = 0.233 kg CO_2eq/kWh, 25 years), scenario-based net lifecycle carbon was −22.1 kg CO_2eq/m^2 for the analyzed south wall with an 8.0-year environmental payback, compared with −7.4 kg CO_2eq/m^2 and 19.2 years for RT28 paraffin. Energy savings did not recover the additional investment; the incremental lifecycle cost was +24.1 EUR/m^2. The theoretical contribution is a transparent, climate-dependent screening logic that couples surrogate interpretation with explicit evidence boundaries; the applied outcome is a palmitic–capric target-property region prioritized for laboratory validation rather than a deployment-ready product. Only directly re-simulated configurations are used for quantitative applied thermal claims; surrogate-only Pareto points are retained as exploratory screening candidates and are not interpreted as validated optima. The numerical results are specific to the modeled south-wall, single-zone boundary; whole-building and cross-regional application requires local recalibration and direct validation. By linking passive comfort, carbon accounting, material innovation, and responsible pre-experimental selection, the workflow is relevant to the decarbonization objectives represented by SDGs 7, 9, 11, 12, and 13.

Keywords: bio-based phase change materials; decarbonized building envelope; EnergyPlus; machine learning surrogate; multi-objective optimization; lifecycle carbon; lifecycle cost

Check for updates

Academic Editors: JeongWook Son and Jinyoung Kim

Received: 30 July 2026
Revised: 21 August 2026
Accepted: 31 August 2026
Published: 3 September 2026

1. Introduction

1.1. Building Envelopes, Thermal Mass, and Phase Change Materials

Buildings account for a substantial share of final energy use and energy-related emissions in Europe [1], while space heating and cooling remain the dominant operational end uses [2,3]. Improving envelope performance, particularly in lightweight and retrofitted buildings, is therefore central to the objectives of the recast Energy Performance of Buildings Directive [4].

Conventional insulation limits steady-state conductive heat flux but cannot attenuate indoor temperature oscillations in climates with large diurnal swings or intense solar irradiance. Thermal mass (dense concrete or brick) provides sensible heat storage that damps these oscillations [5]; however, adding sufficient mass to lightweight or refurbished constructions is often structurally impractical. Phase change materials (PCMs) offer a thermodynamically compact alternative: during melting, a PCM with latent heat L_h = 180 kJ/kg stores 180 kJ per kilogram, one to two orders of magnitude more than the sensible heat of an equivalent mass of brick or concrete over a comparable temperature interval [6–9]. Incorporated into a building envelope, a PCM layer absorbs heat gains during the day—remaining near its melting temperature T_m and limiting indoor temperature increases—then releases the stored heat during cooler nights, regenerating its capacity for the following cycle [10,11]. Insulation reduces the steady-state heat flux; the PCM reduces peak transient loads and improves free-running comfort. The two strategies are complementary [12,13].

Bio-based PCMs derived from fatty acids and their eutectics are a renewable-origin alternative to petroleum-derived paraffins [14]. Binary mixtures based on capric, lauric, palmitic, and stearic acids can be formulated within the temperature range relevant to indoor thermal comfort [15,16]. The recent evidence base includes state-of-the-art PCM envelope reviews [17], analyses of fatty acid limitations and development needs [18], advanced fatty-acid composite formulations [19], bio-derived composite and fiber insulation integrations [20,21], and comparative bio-based/paraffin envelope studies [22,23]. Collectively, these studies show that renewable origin alone does not guarantee superior building performance: conductivity, leakage control, support or encapsulation, cycling stability, hysteresis, and climate-specific activation remain formulation-dependent constraints [17–19].

The thermal performance of a PCM-enhanced envelope is a non-separable function of multiple interacting design variables. Melting temperature governs whether the material repeatedly enters and leaves the phase transition interval; a T_m that is too low or too high relative to the operating temperature range can leave the layer predominantly liquid or solid and reduce latent storage utilization. Latent heat governs storage capacity, but additional capacity is useful only when the layer can charge and regenerate within the daily thermal cycle. Thermal conductivity controls the rate of heat transfer into and out of the layer; thickness and the effective PCM content factor φ govern latent storage capacity and material demand, and wall position determines which thermal boundary the PCM interacts with most directly [24,25]. These interactions are non-linear, climate-dependent, and non-convex, motivating surrogate-assisted exploration of the sampled design space rather than exhaustive simulation of every candidate combination.

1.2. Artificial Intelligence and Multi-Objective Optimization for PCM Selection

Machine learning is increasingly used as a surrogate for computationally expensive building performance simulations [26,27]. Comparative studies show that ensemble tree methods can provide accurate and computationally efficient predictions of building loads, while Gaussian process and neural network models can also perform well when sampling is carefully designed [28–31]. Recent reviews of surrogate-assisted sustainable building optimization underline both the computational benefits and the need for stronger reporting of validation, model selection, and reproducibility practices [32–34]. Explainable AI techniques, particularly SHAP (SHapley Additive exPlanations), can extend surrogate modeling beyond prediction by revealing fitted non-linear thresholds and interactions [35]. In this study, such patterns are interpreted as properties of the trained numerical surrogate, not as direct evidence of physical causality.

Evolutionary multi-objective algorithms, including NSGA-II, are widely used for non-convex building design problems because they do not require differentiable objective functions [36]. Previous PCM studies have explored climate-dependent trade-offs between thermal performance and cost [37], and recent frameworks have combined surrogate models with multi-objective envelope optimization or energy–economic–environmental assessment [38–40]. Cross-climate studies of bio-based and conventional PCMs further underline the importance of evaluating operational and embodied impacts together [41–43]. The methodological gap is therefore narrower than a simple checklist suggests: the distinctive contribution of the present study is the coupling of bio-based PCM screening, free-running comfort, conditioned thermal demand, surrogate explainability, multi-objective optimization, and explicitly separated carbon/cost post-processing on a common south-wall functional unit. Table 1 provides an analytical map of selected contributions, while Section 5 extends the comparison to the recent PCM envelope and surrogate optimization literature.

Table 1. Analytical positioning of the present study relative to selected literature contributions). The comparison identifies methodological scope and remaining gaps rather than using a binary novelty checklist.

Study	Methodological Scope	What the Study Establishes	Gap Relative to the Present Screening Question
Ali et al. [30]	Machine learning surrogate for building energy simulation	Shows that data-driven surrogates can reproduce expensive building performance simulations with high efficiency	Not PCM-specific; no combined comfort, lifecycle carbon, LCC, or Pareto material screening
Mehrotra and Yi [31]	Adaptive sampling and machine learning emulators for architectural energy modeling	Demonstrates the importance of sampling strategy and emulator selection for simulation replacement	Architectural-massing focus; no PCM formulation, lifecycle accounting, or multi-objective material selection
Soares et al. [37]	Multi-dimensional optimization of PCM drywalls across climates	Establishes climate-dependent PCM design and thermal/economic trade-offs	No bio-based/paraffin common boundary comparison, surrogate XAI, or explicit lifecycle carbon accounting
Yang et al. [38]	Surrogate-assisted multi-objective PCM-enhanced building optimization	Demonstrates ML-assisted evolutionary optimization for PCM design	No bio-based formulation focus and no explicit separation of simulation outputs from scenario-based lifecycle assumptions
Idouanaou et al. [41]	Cross-climate evaluation of bio-based PCMs with operational and embodied carbon	Provides recent bio-based and lifecycle carbon evidence across multiple climates	No surrogate explainability, LCC, or Pareto screening of continuous PCM property targets
Present study	Bio-based PCM property space screening with EnergyPlus v24.1.0, XGBoost 3.4.1/TreeSHAP, NSGA-II, and scenario-based carbon/LCC	Integrates thermal screening, explainability, common wall area carbon/cost accounting, and explicit evidence boundaries	Screening-level manufacturability, three-climate scope, incomplete lifecycle boundary, and limited direct frontier validation

1.3. Research Gap, Questions, and Contributions

The scientific and applied problem addressed here is not the absence of PCM optimization studies, but the difficulty of translating a mathematically attractive optimum into a physically plausible and decision-relevant bio-based PCM target. Four gaps motivate the workflow. First, simulation-derived comfort and thermal loads are often combined with carbon and cost assumptions without a sufficiently visible boundary between numerical physics and scenario-dependent conversion. Second, bio-based candidates and paraffin benchmarks are not always compared on the same functional unit and accounting basis. Third, climate-dependent non-linear interactions between transition temperature, latent storage, layer geometry, and position require an efficient screening method but also careful interpretation of surrogate explanations. Fourth, accuracy on a random held-out test set does not by itself establish reliability at the Pareto boundary, where decision-critical solutions can lie. This study is structured around three research questions:

RQ1. Which PCM design variables most strongly influence free-running thermal discomfort and annual thermal demand across Mediterranean and continental Italian climates?

RQ2. Can machine learning surrogate models reproduce the EnergyPlus input–output mapping with sufficient held-out accuracy to support multi-objective screening?

RQ3. Which configurations provide the most defensible trade-offs between comfort, lifecycle carbon, and lifecycle cost under explicitly stated conversion and market assumptions?

This study therefore contributes: (i) an 18,000-case EnergyPlus parametric campaign covering six design variables and three Italian climates; (ii) comparison of five surrogate algorithms with held-out evaluation; (iii) TreeSHAP interpretation explicitly bounded to the fitted surrogate; (iv) NSGA-II screening of comfort, lifecycle carbon, and incremental lifecycle cost; (v) direct EnergyPlus verification of the balanced Palermo configuration; and (vi) bounded and threshold-based scenario checks that expose the dependence of environmental and economic conclusions on energy system, material, and integration assumptions. Scientifically, the objective is to identify activation and trade-off patterns without presenting a numerical optimum as a guaranteed manufacturable formulation. Practically, the objective is to narrow the experimental search toward candidate property regions that can subsequently be formulated, characterized, and validated.

2. Materials and Methods

Research program. Figure 1 summarizes the complete numerical workflow and the evidence hierarchy used in this study. Candidate material descriptors, layer variables, climate, and reference wall conditions define the design space; Latin hypercube sampling then supplies the EnergyPlus v24.1.0 campaign. Free-running simulations provide occupied discomfort, whereas conditioned simulations provide heating and cooling loads. Climate-specific surrogates emulate the thermal response for candidate screening and are interpreted with TreeSHAP. During NSGA-II evaluation, thermal surrogate outputs are combined with the explicit analytical carbon and cost equations and scenario coefficients in Section 2.3; lifecycle indicators are therefore not learned outputs of EnergyPlus v24.1.0. Figure 1 also shows the modeled south-wall assembly, the three candidate PCM positions, and the palmitic–capric P4/S4 property target proposed for experimental follow-up. These are technical schematics of the numerical model and target formulation, not photographs of a physical specimen. The workflow deliberately separates direct simulation, surrogate prediction, and scenario-based post-processing so that the evidentiary status of every claim remains visible.

https://doi.org/10.3390/su18179041

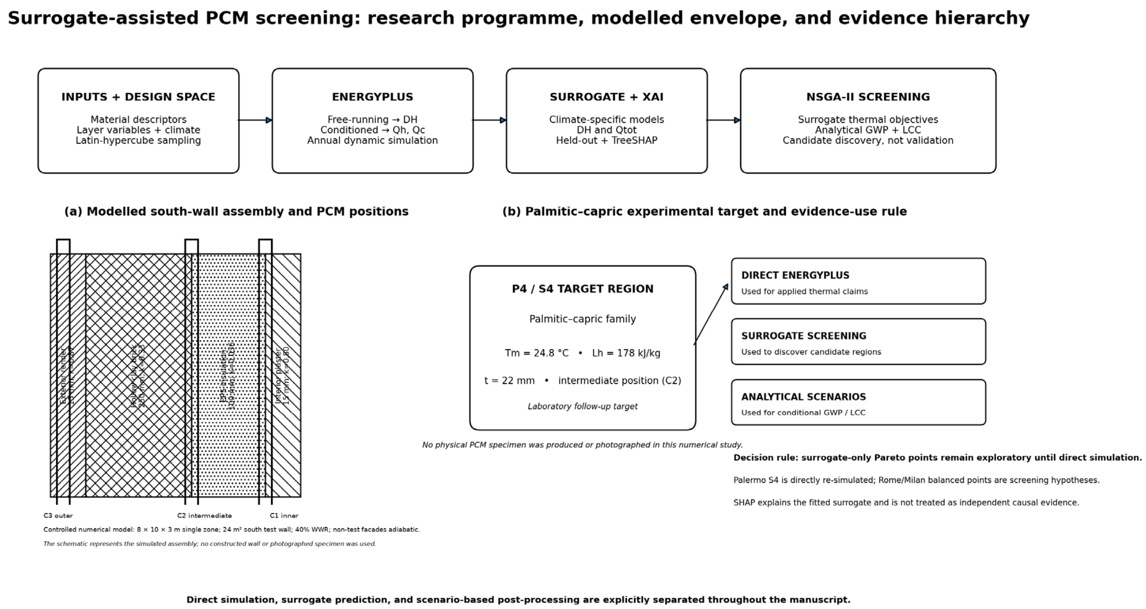

Figure 1. Research program, modeled south-wall assembly, palmitic–capric experimental target, and evidence hierarchy for surrogate-assisted PCM screening. Panel (**a**) is a technical schematic of the simulated wall and candidate PCM positions; panel (**b**) is a conceptual representation of the P4/S4 target property region and the evidence use rule. Neither panel represents a photographed experimental specimen. Direct EnergyPlus outputs, surrogate screening predictions, and analytical lifecycle results are distinguished throughout the manuscript.

2.1. Simulated Materials, Phase Change Representation, and Reference Building

This is a numerical screening study and did not use physical specimens, laboratory samples, or an experimentally monitored building. The "materials" are numerical descriptions of three literature-reported fatty-acid eutectic families and the commercial paraffin RT28, represented in EnergyPlus v24.1.0 (Table 2). The continuous design space is therefore interpreted as a target-property space for pre-experimental screening, not as evidence that every independently sampled combination of T_m, L_h, k_s, thickness, and effective PCM content can be manufactured. Representative numerical candidates are related to the closest plausible material family only to define experimental follow-up targets; manufacturability is never inferred from proximity in the property space. Because no physical specimen or case study building was tested, a photographic image would imply experimental evidence that does not exist. Figure 1 therefore provides explicit technical schematics of the simulated envelope and the palmitic–capric target region while preserving the numerical nature of the study.

Table 2. Nominal thermophysical properties of bio-based PCM candidates and paraffin benchmark [6,7,15,16,20,21,44–46].

Property	FA-1 (Capric–Lauric)	FA-2 (Palmitic–Capric)	FA-3 (Stearic–Palmitic)	RT28 (Paraffin)
Melting onset T_m (°C)	18.9	23.6	26.7	28.0
Latent heat L_h (kJ/kg)	143	171	183	244
Solid conductivity k_s (W/mK)	0.151	0.162	0.174	0.190
Hysteresis ΔT_h (°C)	0.6	0.7	0.9	0.9
Cyclic degradation D_c (%/1000 cycles)	2.1	1.7	1.4	0.8
Origin	Coconut/palm kernel oil	Palm/coconut oil	Animal/palm tallow	Petroleum refining

PCM heat storage was represented with the EnergyPlus v24.1.0 ConductionFinite-Difference heat balance algorithm and a piecewise linear enthalpy–temperature relation. The enthalpy curve was discretized at 0.2 °C, the finite-difference spatial discretization was 3 mm, and the zone timestep was 600 s [47]. The transition interval was represented symmetrically about the sampled melting temperature. Detailed supercooling, separate melting/freezing branches, and rate-dependent hysteresis were not resolved. This simplification is important because recent EnergyPlus v24.1.0 and annual building studies show that hysteresis representation can affect partial transition and annual PCM predictions [48,49]. The reported hysteresis values in Table 2 are therefore screening descriptors rather than fully resolved constitutive parameters and require experimental calibration.

Numerical resolution and validation context. The same 0.2 °C enthalpy discretization, nominal 3 mm spatial discretization, and 600 s zone timestep were used across all 18,000 cases to preserve within-campaign comparability. The 600 s timestep corresponds to six timesteps per hour, a temporal resolution also used in recent annual EnergyPlus v24.1.0 PCM studies [48]; this similarity is cited only as methodological context and is not evidence of convergence for the present model. At the solver level, the EnergyPlus PCM/ConductionFiniteDifference formulation has undergone analytical verification, comparative testing, and empirical validation for opaque wall assemblies [50], while subsequent work has shown that hysteresis treatment can materially affect partial transition predictions [49]. These publications support the solver lineage but do not validate the present wall discretization. With a nominal 3 mm spatial step, the sampled 5–40 mm PCM layer range is represented by approximately 2–13 intervals, making the thinnest cases particularly sensitive to discretization choice. Because a case-specific grid/timestep independence study was not preserved for this campaign, absolute performance values are interpreted as screening-level predictions; convergence testing remains necessary before product-level or code compliance claims.

For numerical implementation, the liquid-fraction function F(T) increases linearly from 0 to 1 across the prescribed transition interval, and PCM enthalpy is expressed as $H(T) = Hsens(T) + L_hF(T)$. The sampled factor φ is a homogenized effective PCM-content factor used to scale latent storage capacity and PCM areal mass; it is not a resolved microstructural volume fraction of a specified carrier. When $\varphi < 1$, the remaining fraction $(1 - \varphi)$ represents an unspecified support/encapsulation phase of the type commonly required in PCM composites [51]. Because consistent carrier-specific density, heat capacity, and conductivity data were unavailable for all candidate families, no constituent-level mixing rule is claimed. Instead, k_s is sampled directly as the effective conductivity of the PCM-containing layer, while φ scales only the latent-storage term and PCM mass used in lifecycle post-processing. This reduced-order definition is appropriate for property space screening but must be replaced by measured composite properties before experimental or product-level validation.

The reference model is a single-zone office measuring 8 m × 10 m × 3 m (80 m^2 floor area). The analyzed south-facing facade contains a 24 m^2 test wall area and a 40% window-to-wall ratio. The retained model specification reports, from outside to inside, 15 mm exterior render (k = 0.87 W/mK), 250 mm hollow clay brick (k = 0.33 W/mK), 100 mm expanded-polystyrene insulation (k = 0.036 W/mK), and 15 mm interior plaster (k = 0.80 W/mK), giving the reported overall U-value of 0.29 W/m^2K. When the PCM is introduced, the EPS thickness is adjusted to preserve the reference thermal resistance, thereby isolating latent storage effects from changes in steady-state U-value. Three PCM positions are considered: inner (C1), intermediate (C2), and outer (C3). The other vertical facades are adiabatic. Consequently, the model is designed for comparative south-wall screening within a controlled single-zone boundary and must not be interpreted as a whole-building energy model.

The reference geometry, envelope data, schedules, boundary conditions, and climate source used in the numerical model are described in Sections 2.1 and 2.2 and summarized in Tables 2–4.

Table 3. Climatic characterization of the three study locations. EPW files are from the IGDG dataset (EnergyPlus repository).

Parameter	Palermo	Rome	Milan
Climate Zone (ASHRAE)	3A	3B	5A
Köppen	Csa	Csa	Cfa
Mean annual temp. (°C)	18.5	15.7	13.0
HDD_{18} (°C·day)	601	1084	2404
CDD_{10} (°C·day)	1842	1287	512
Mean July temp. (°C)	27.8	25.6	23.4

Table 4. Design variables and sampling ranges (uniform distributions).

Variable	Symbol	Range	Unit
Melting temperature	T_m	18–32	°C
Latent heat	L_h	100–250	kJ/kg
Solid thermal conductivity	k_s	0.14–0.60	W/mK
Layer thickness	t	5–40	Mm
Effective PCM content factor	φ	0.20–0.80	–
Wall position	pos	{C1, C2, C3}	–

Occupancy follows an office schedule of 09:00–18:00 on weekdays, consistent with EN ISO 17772-1 [52]. The retained reference model specification reports an occupant density of 0.07 persons/m^2 and combined equipment-plus-lighting internal gains of 14 W/m^2 during occupied hours. Two complementary annual simulations are performed for every design: (a) a free-running simulation, with HVAC disabled, to quantify passive thermal buffering through occupied discomfort degree-hours; and (b) a conditioned simulation using EnergyPlus Ideal Loads Air System control, with occupied setpoints of 20 °C for heating and 26 °C for cooling, to calculate ideal heating and cooling loads. The space is free-floating outside occupied hours. This separation prevents active conditioning from masking the passive comfort effect while retaining a consistent estimate of energy demand changes.

2.2. Simulation Software, Parametric Design, and Performance Indicators

EnergyPlus v24.1.0 was used as the dynamic simulation engine [53]. For each climate, 6000 combinations were generated by Latin hypercube sampling implemented with SALib [54], with 2000 cases assigned to each wall position for 18,000 annual simulations. The material families in Table 2 were selected to span fatty-acid transition temperatures relevant to occupied-building conditions and to provide a paraffin benchmark; Table 4 then broadens these properties into a screening space to test sensitivity and interaction effects. Climate-specific surrogate models were trained separately so that fitted response surfaces were not pooled across thermally distinct locations. XGBoost 3.4.1 was selected for the final surrogate layer, TreeSHAP for interpretation, and pymoo for NSGA-II optimization. Software roles and reproducibility details are analyzed in Section 3.

Physical feasibility interpretation of the sampling space. The ranges in Table 4 intentionally exceed the nominal properties of the four benchmark materials because the optimization is used to identify desirable property regions, including values potentially achievable through eutectic reformulation, conductive fillers, form stabilization, or composite design [18,19]. The variables are sampled independently as a numerical design assumption; real fatty-acid formulations exhibit correlated thermophysical properties. Therefore, Pareto solutions are not automatically treated as manufacturable materials. A post-optimization feasibility gate is required: transition temperature, latent heat, conductivity, PCM content, encapsulation method, cycling stability, leakage resistance, and fire performance must all be jointly verified before a numerical solution is promoted to an experimental prototype.

Occupied discomfort degree-hours were calculated from the free-running operative temperature as DH = Σocc{max[Top(t) − 26, 0] + max[20 − Top(t), 0]}Δt, where the summation includes occupied timesteps only, and Δt is expressed in hours [55,56]. The conditioned model supplied annual heating load Qh, cooling load Qc, total thermal demand Qtot = Qh + Qc, and peak loads. Equivalent full cycles were calculated as Nfull = Elat,abs/(2 mPCM Lh), where Elat,abs is the annual absolute latent energy exchanged; the active fraction was defined as the fraction of occupied timesteps during which the layer-average PCM temperature lay within the prescribed transition interval. Annual latent energy exchanged was obtained by integrating the absolute latent heat flow over the annual simulation. Candidate benchmarks were S0 (no PCM), S1 (RT28, 25 mm, inner), S2 (nominal FA-2, 25 mm, inner), S3 (FA-2 with optimized transition properties, 25 mm, inner), and S4 (direct EnergyPlus simulation of the balanced Palermo design: T_m = 24.8 °C, L_h = 178 kJ/kg, 22 mm, intermediate position).

The indicator definitions above formalize the post-processing logic used in the manuscript and ensure consistent calculations of comfort, energy demand, activation, and cycling indicators across the simulation campaign.

2.3. Lifecycle Carbon Accounting and Lifecycle Cost

The functional unit is 1 m^2 of the analyzed south-facing wall over a 25-year reference period, following lifecycle assessment principles [57–59]. The comparative carbon boundary includes cradle-to-gate production of the PCM mass incorporated in the test wall and avoided operational electricity emissions relative to S0. Encapsulation/packaging, transport, installation, maintenance, demolition, and end-of-life processes are excluded because consistent candidate-specific inventories were unavailable. Their omission can bias the apparent environmental advantage of any PCM and is therefore treated explicitly as a boundary limitation rather than assumed to be negligible. The assessment is described as scenario-based lifecycle carbon accounting, not as a complete multi-indicator or product-level LCA. In addition to the central calculation, an algebraic omitted-process burden GWP_omitted is used only as a threshold stress test, such that GWP_net,extended = GWP_net + GWP_omitted. The bounded stress-test range of 0–30 kg CO_2eq/m^2 is not an inventory estimate or probability distribution; it identifies how much unmodeled embodied burden would be required to reverse the sign of a reported net carbon benefit.

Annual equivalent electricity savings are calculated separately for heating and cooling as ΔEel = ΔQheat/SCOP + ΔQcool/SEER. Net GWP is then calculated as GWPnet = GWPPCM,production − ΔEel × EFgrid × N, where ΔQheat and ΔQcool are annual ideal-load reductions per square meter, SCOP and SEER are the seasonal heating and cooling efficiency factors, EFgrid is the electricity carbon factor, and N is service life.

The central scenario uses SCOP = 3.0 and SEER = 3.0, EF_grid = 0.233 kg CO_2eq/kWh, and N = 25 years [60]. Direct EnergyPlus benchmark calculations retain Q_h and Q_c separately and convert them through SCOP and SEER before applying the grid factor or tariff. The optimization layer, however, uses a surrogate for Q_tot = Q_h + Q_c. Accordingly, Q_tot is sufficient for the bounded optimization scenarios reported here only because SCOP and SEER are assigned equal values in every low-, central-, and high-benefit combination. The surrogate Pareto set does not support independent variation of SCOP and SEER; such an analysis would require separate climate-specific surrogate models for Q_h and Q_c. No claim is therefore made about Pareto rankings under asymmetric heating and cooling conversion efficiencies.

The incremental lifecycle cost is calculated as LCCinc = $C_0 - \Sigma$t[(ΔEel $\times$ celectricity)/(1 + r)ˆt]. The initial cost is decomposed as C_0 = ρPCM tφ cPCM + Cint(pos), where cPCM is the PCM unit price, and Cint is an author-defined screening allowance for encapsulation, handling, and installation. Central allowances of 7.5 EUR/m^2 for inner placement (C1), 20.0 EUR/m^2 for intermediate placement (C2), and 25.0 EUR/m^2 for outer placement (C3) encode relative integration complexity rather than market quotations. Because this assumption can represent a large share of initial costs, the bounded Cint range is expanded to 0–60 EUR/m^2 as a stress test, and critical break-even values are reported in the Results. The range is not presented as a market distribution. The central electricity tariff is 0.28 EUR/kWh, and the real discount rate is 3% [61]. A positive LCCinc indicates that the additional investment is not recovered by discounted energy savings during the reference period.

2.4. *Machine Learning Surrogates and Multi-Objective Optimization*

Five algorithms were compared for the two simulation-derived targets used in optimization: multiple linear regression (MLR), random forest (RF), Gaussian process regression (GPR), multi-layer perceptron (MLP), and XGBoost. Within each climate, cases were stratified by wall position and divided into training, validation, and held-out test sets (70/15/15); one climate-specific model was then trained with wall position as a categorical predictor. Available workflow records document five-fold cross-validation on the training partition. Continuous inputs were standardized to zero mean and unit variance for GPR and MLP, whereas RF and XGBoost were trained without feature scaling. Grid search cross-validation was used for MLR and GPR; Bayesian optimization with Optuna was used for RF, XGBoost, and MLP. The documented RF search space comprised 100–500 trees, a maximum depth of 5–20, and a minimum sample of 1–10 per leaf. The documented XGBoost search space comprised 100–600 estimators, a learning rate of 0.01–0.30, a maximum depth of 3–8, a subsample of 0.60–1.00, L1 regularization of 0–1, and L2 regularization of 0–5. For the MLP, 2–5 hidden layers and 64–512 neurons per layer were explored. Model development used R^2, RMSE, and NRMSE, with NRMSE defined as the RMSE divided by the observed held-out target range. This disclosure supports protocol-level reproducibility of model selection; exact bitwise replication is not claimed because final fitted hyperparameter vectors, split indices, and random seeds were not preserved. The manuscript therefore does not reconstruct those values by assumption, and no conclusion is made to depend on a fabricated fitted state.

SHAP values were calculated for the selected XGBoost models with TreeSHAP [35]. Global importance was quantified using mean absolute SHAP values on held-out instances, while dependence and pair-interaction analyses were used to identify fitted non-linear activation windows and variable interactions. These quantities explain the fitted surrogate; they do not establish causal thermophysical mechanisms and can reflect model error or sampling density. Section 4.2 therefore reports the two direct EnergyPlus consistency checks that can be supported by the retained benchmark outputs: the P4/S4 frontier-point residual and a matched-geometry S2 $\rightarrow$ S3 transition property comparison across all three climates.

Because T_m and L_h change jointly between S2 and S3, that comparison is explicitly treated as a directional consistency check rather than a one-factor local derivative.

NSGA-II was implemented in pymoo [62] with a population of 200 and a fixed termination at 500 generations. Available optimization records specify simulated-binary crossover (SBX, $\eta c = 20$), polynomial mutation ($\eta m = 20$), and a crossover probability of 0.9; optimization was run independently for each climate. The three minimized objectives were DH, net GWP, and incremental LCC. DH and Q_tot were predicted by climate-specific XGBoost surrogates, while GWP and LCC were calculated analytically from predicted thermal demand and the scenario coefficients in Table 5. Four representative solutions were retained for screening: P1 (comfort-focused), P2 (low-material carbon–cost compromise), P3 (lowest LCC), and P4 (balanced knee candidate). Equal-weight TOPSIS together with a knee-point criterion was used to identify the balanced candidate. The balanced Palermo property set was re-simulated directly in EnergyPlus as S4, and only this directly checked configuration is used for quantitative applied thermal conclusions. The remaining Palermo Pareto representatives and the Rome/Milan balanced configurations are explicitly retained as exploratory surrogate-screening candidates rather than validated optima. Exact optimization seed, mutation probability, and repeat-run convergence traces were not preserved; therefore, no stochastic stability claim is made beyond the reported archived run.

Table 5. Central lifecycle carbon/LCC assumptions and bounded scenario ranges.

Parameter	Central Value	Bounded Range	Use
PCM density	850 kg/m^3	800–900 kg/m^3	Areal mass; literature-informed screening assumption [15,16,63–65]
Bio-based embodied GWP	1.4 kg CO_2eq/kg	1.0–2.0	Cradle-to-gate PCM production; ecoinvent/secondary literature [63,66,67]
Paraffin embodied GWP	2.9 kg CO_2eq/kg	2.2–3.6	Cradle-to-gate PCM production; ecoinvent/secondary literature [44,45,66]
PCM unit price	4.2 EUR/kg	2.8–8.1 EUR/kg	Author-defined screening price; supplier quotation required
Electricity tariff	0.28 EUR/kWh	0.20–0.36	Screening tariff assumption; bounded range contextualized with Eurostat electricity price statistics [61]
Grid carbon factor	0.233 kg CO_2eq/kWh	0.160–0.306	Italian screening factor contextualized with ISPRA electricity sector reporting [60]
Heating SCOP	3.0	2.0–4.0	Author-defined all-electric heating scenario
Cooling SEER	3.0	2.0–4.0	Author-defined all-electric cooling scenario
Discount rate/life	3%/25 years	1–5%/20–30 years	Author-defined financial horizon and sensitivity bounds
Integration allowance C_int	7.5/20.0/25.0 EUR/m^2 (C1/C2/C3)	0–60 EUR/m^2	Author-defined stress-test range; critical break-even value reported
Additional omitted-process GWP burden	0 kg CO_2eq/m^2	0–30 kg CO_2eq/m^2	Threshold stress test for excluded encapsulation/packaging, transport, installation, maintenance, and end-of-life burdens; not an inventory estimate

Evidence use rule for Pareto results. Surrogate optimization is used for candidate discovery, not as a substitute for direct dynamic validation. A surrogate-selected Pareto point is considered eligible for an applied performance claim only after a matched direct EnergyPlus check. Under this rule, Palermo S4 is decision-qualified within the present numerical model, whereas P1–P3 and the Rome/Milan balanced points remain exploratory hypotheses. This rule is applied consistently in the Results, the Discussion, and the Conclusions: unverified frontier points may illustrate trade-offs or guide follow-up simulation, but they are not used as validated evidence of realized building performance, product superiority, or deployment readiness.

2.5. Scenario Analysis and Internal Consistency Checks

The robustness of lifecycle conclusions was examined with bounded low-, central-, and high-benefit parameter combinations rather than interpreted probabilistically. "Benefit" denotes avoided carbon or costs attributed to the PCM relative to S0, not the sustainability of the wider energy system. The analysis varies PCM embodied-carbon factors, the electricity tariff, the grid carbon factor, equal heating/cooling conversion efficiency, the discount rate, service life, PCM costs, and the integration allowance within Table 5. The separate GWP_omitted stress test evaluates the additional unmodeled embodied burden required to erase the central carbon advantage. The purpose is to identify changes in sign and ranking under explicit boundary conditions, not to estimate occurrence probabilities.

Assumption provenance and status. The PCM density is a representative fatty-acid screening value within ranges reported in the PCM literature [15,16,18,63–65]. The embodied GWP coefficients are comparative cradle-to-gate screening factors informed by ecoinvent 3.10 and secondary sources [44,45,63,66,67], not environmental product declarations. The central electricity carbon factor is an explicit Italian screening value contextualized with ISPRA electricity sector reporting [60], while the central tariff and its bounded range are modeling assumptions contextualized with Eurostat electricity price statistics [61]; neither is presented as a universal market parameter. SCOP, SEER, service life, discount rate, PCM price, integration allowances, and the omitted-process GWP stress-test range are likewise explicit modeling assumptions. The last two are intentionally varied over broad ranges to expose sign changes or break-even thresholds and should be replaced by supplier quotations and primary lifecycle inventories before design or procurement decisions. Recent lifecycle envelope optimization also shows that grid decarbonization can materially change operational carbon conclusions [40]; the static grid-factor treatment used here is therefore scenario simplification rather than a forecast.

Dimensional consistency was checked throughout the analytical post-processing. EnergyPlus ideal heating and cooling loads were converted separately to equivalent electricity using SCOP and SEER before application of the grid factor or tariff. Carbon and cost calculations use the same annual-energy basis, functional unit, service life, and discount convention. This separation prevents unit inconsistencies and makes the distinction between simulated thermal outputs and scenario-dependent conversion assumptions explicit.

3. Modeling Procedure

3.1. Numerical Simulation, Sampling, and Software Environment

This study is based exclusively on numerical modeling; no physical specimen, experimental measurement series, or observational building dataset was analyzed. The 18,000 observations used for surrogate modeling are outputs of the parametric EnergyPlus campaign. The numerical workflow used EnergyPlus v24.1.0 for dynamic building simulation, SALib 1.5.2 for Latin hypercube sampling, XGBoost 3.4.1 for the selected surrogate models, TreeSHAP for explainability, and pymoo for NSGA-II optimization. Lifecycle carbon and cost indicators were calculated from the analytical equations and assumptions reported in Sections 2.3 and 2.5 rather than predicted directly by EnergyPlus (See Table 6). Throughout the Results, evidence is labeled according to source: direct EnergyPlus benchmark output, surrogate screening prediction, or analytical scenario post-processing. This hierarchy is maintained to avoid conflating model acceleration with independent physical validation.

Table 6. Software components and their role in the numerical workflow.

Software Component	Role in the Study
EnergyPlus v24.1.0	Dynamic annual simulation of the reference building, free-running comfort assessment, Ideal Loads heating and cooling demand, and PCM modeling with ConductionFiniteDifference.
SALib 1.5.2 [54]	Latin hypercube sampling of the six-dimensional PCM design space for the 18,000-case numerical campaign.
XGBoost 3.4.1	Climate-specific surrogate modeling of discomfort degree-hours and total annual thermal demand.
TreeSHAP [35]	Global feature importance, dependence, and interaction analysis for interpretation of the selected XGBoost surrogate models.
pymoo [62]	Implementation of NSGA-II multi-objective optimization with 200 individuals and 500 generations.
ecoinvent 3.10 [66]	Source of background cradle-to-gate coefficients used in the scenario-based lifecycle carbon accounting.

The modeling procedure is organized as a staged pipeline rather than as a set of interchangeable software tools. EnergyPlus provides the annual free-running and conditioned thermal response; SALib supplies the Latin hypercube design points; climate-specific XGBoost models emulate the two thermal targets used for screening; TreeSHAP interprets those fitted response surfaces; pymoo executes the NSGA-II search; and ecoinvent 3.10 is used only as a source of background cradle-to-gate coefficients in the analytical lifecycle post-processing. This separation is important because carbon and cost indicators are not EnergyPlus outputs, and surrogate accuracy is not treated as independent physical validation.

3.2. Surrogate Training, Validation, and Explainability

Reproducibility status. Available workflow records support protocol-level reconstruction of feature scaling, five-fold cross-validation, tuning approach, principal RF/XGBoost search ranges, MLP architecture bounds, and the main NSGA-II genetic operators. Tables 7 and 8 distinguish protocol-reproducible elements from details that are unavailable for exact bitwise replication, principally final fitted parameter vectors, split indices, random seeds, and some optimization trace settings. The inference policy adopted in the Results is deliberately independent of reconstructing those missing values: only direct EnergyPlus results are used for applied thermal claims, while surrogate-only outputs remain screening candidates. No unavailable parameter is retrospectively invented.

Table 7. Reproducibility and validation information available from the archived workflow.

Workflow Element	Information Available	Qualification/Inference Control
Simulation campaign	EnergyPlus v24.1.0; 18,000 annual cases; 6000 per climate; six sampled design variables	No public model/code archive; study claims are restricted to the documented numerical campaign and reported evidence.

Table 7. *Cont.*

Workflow Element	Information Available	Qualification/Inference Control
Data partition	Within-climate split stratified by wall position; 70/15/15 training/validation/test; five-fold CV on training partition	Protocol-level split/CV structure is documented; exact bitwise split replication is not claimed because indices/seed are unavailable.
Surrogate comparison	MLR, RF, GPR, MLP, and XGBoost; scaling/tuning protocol and principal RF/XGBoost/MLP search ranges recovered (Table 8)	Search protocol is documented; exact final fitted state is not reconstructed, and no claim depends on invented hyperparameters.
Evaluation	R^2, RMSE, and NRMSE; aggregate held-out results; direct P4 frontier residual and S2 $\rightarrow$ S3 cross-climate checks	Pooled interpolation metrics are reported as such. Climate-specific residual distributions are not invented; decision region inference is governed by the evidence hierarchy reported in Section 4.2.
Explainability	TreeSHAP global/dependence/interaction interpretation plus direct EnergyPlus consistency checks reported in Section 4.2	SHAP is model interpretation only. Direct checks take precedence; no one-at-a-time causal sensitivity claim is made.
Optimization	pymoo NSGA-II; population 200; 500 generations; SBX $\eta c = 20$; polynomial mutation $\eta m = 20$; crossover probability 0.9; equal-weight TOPSIS/knee selection retained	The archived run defines screening candidates only; no repeat-run stochastic stability claim is made without the missing seed/trace.
Direct verification	Balanced Palermo P4 re-simulated as S4; pointwise residual quantified; direct S2 $\rightarrow$ S3 checks available for Palermo/Rome/Milan	Only Palermo S4 is directly decision-qualified. P1–P3 and Rome/Milan balanced candidates are explicitly exploratory and excluded from validated applied claims.

Table 8. Recovered surrogate training and optimization settings.

Model/Step	Recovered Protocol/Search Space	Unavailable Detail/Consequence
Preprocessing and validation	One climate-specific model per climate; 70/15/15 training/validation/test; wall-position stratification; five-fold CV on training set; GPR/MLP standardized; RF/XGBoost unscaled	Exact categorical encoding, split indices, and random seed
MLR/GPR	Grid search cross-validation; GPR uses standardized continuous inputs	MLR intercept/regularization details and exact GPR kernel/grid/final fitted settings
Random forest	Bayesian optimization (Optuna); 100–500 trees; max depth 5–20; min samples per leaf 1–10	max_features/min_samples_split details, final selected RF vector, and seed
XGBoost	Bayesian optimization (Optuna); 100–600 estimators; learning rate 0.01–0.30; max depth 3–8; subsample 0.60–1.00; L1 0–1; L2 0–5	colsample_bytree/min_child_weight/gamma/early stopping details, final selected XGBoost vector, and seed

https://doi.org/10.3390/su18179041

Table 8. *Cont.*

Model/Step	Recovered Protocol/Search Space	Unavailable Detail/Consequence
MLP	Bayesian optimization (Optuna); 2–5 hidden layers; 64–512 neurons per layer; standardized inputs	Activation/optimizer, learning-rate/dropout bounds, epochs, batch size, early stopping, final architecture, and seed
NSGA-II	Population 200; 500 generations; SBX $\eta c = 20$; polynomial mutation $\eta m = 20$; crossover probability 0.9; run independently by climate; balanced candidate selected using equal-weight TOPSIS plus knee criterion	Exact optimization seed, mutation probability, repeat-run convergence trace, and tie-breaking trace

The retained workflow therefore supports protocol-level validation of the modeling chain: the split structure, cross-validation logic, model-specific scaling, principal hyperparameter search spaces, held-out metrics, and the available direct EnergyPlus checks can be reconstructed. By contrast, exact split indices, random seeds, final fitted parameter vectors, and some optimization traces are not available; these missing items are explicitly treated as limits on bitwise replication rather than filled by assumption.

3.3. Surrogate Training and Optimization Settings

Table 8 consolidates the recoverable model training and optimization specifications used to reproduce the methodological protocol. The settings reported here are those retained in the archived workflow and are deliberately separated from unavailable fitted-state details.

Reproducibility tier. The workflow is reproducible at the protocol/specification level from the retained documentation, but not at the bitwise fitted-model level. This distinction is consequential rather than cosmetic: the manuscript does not use unrecoverable seeds or final parameter vectors to support any new numerical result, and all decision-critical applied claims are anchored to direct EnergyPlus outputs or explicit analytical equations.

4. Results

4.1. Thermal Activation and Energy Performance

All 18,000 EnergyPlus simulations were completed without convergence failure. Table 9 summarizes the simulation-derived thermal outputs pooled across climates. Lifecycle indicators are reported separately because they depend on the explicit post-processing assumptions in Table 5.

Table 9. Statistical summary of EnergyPlus thermal outputs across all 18,000 simulations.

Output Variable	Unit	Min	Mean	Max	Std. Dev.
Discomfort degree-hours (DH, free-running)	°C·h/year	41.3	213.7	412.8	78.4
Annual cooling demand Q_c (conditioned)	kWh/year	1847	4621	7304	1103
Annual heating demand Q_h (conditioned)	kWh/year	621	2184	5817	1041
Total energy demand Q_tot	kWh/year	3112	6805	11,437	1682
Annual PCM cycles N_full	cycles/year	0	94.3	213	56.8
PCM active fraction f_active	–	0	0.24	0.72	0.16

https://doi.org/10.3390/su18179041

DH is obtained from the free-running model; Q_c, Q_h, and Q_tot are ideal thermal loads from the conditioned model.

PCM activation exhibited a climate-dependent transition-temperature window. In Palermo, the greatest cycling occurred around 25–27 °C; the corresponding window shifted downward in Rome and Milan. Intermediate positioning generally increased the fraction of time spent within the transition interval, whereas the inner position produced the strongest direct reduction in indoor discomfort. Outer positioning was the least reliable because the layer was more frequently saturated or inactive. (See Table 10).

Table 10. Annual discomfort degree-hours (DH, free-running model) for benchmark scenarios.

Scenario	Palermo DH (°C·h/yr)	Palermo DH Δ%	Rome DH	Rome Δ%	Milan DH	Milan Δ%
S0 (no PCM)	326	–	248	–	187	–
S1 (RT28, 25 mm, inner)	198	−39.3%	160	−35.5%	134	−28.3%
S2 (FA-2 nominal, 25 mm, inner)	224	−31.3%	176	−29.0%	148	−20.9%
S3 (FA-2, T_m + L_h optim., 25 mm, inner)	203	−37.7%	158	−36.3%	129	−31.0%
Balanced [1]	187	−42.6%	162	−34.7%	134	−28.3%

[1] For Palermo, the balanced configuration is the direct EnergyPlus verification S4 of the P4 property set (T_m = 24.8 °C, L_h = 178 kJ/kg, t = 22 mm, intermediate position), and its DH value of 187 °C·h/year is used for headline reporting. The Rome and Milan values in the balanced row are surrogate-model predictions for the climate-specific configurations reported later in Section 4.4. S0–S3 are direct EnergyPlus benchmark simulations in all three climates.

The conditioned benchmark results show modest but consistent reductions in annual total thermal demand (Table 11). For the balanced configuration, the 6.0% Palermo reduction is from the direct EnergyPlus simulation S4, whereas the 3.3% Rome and 3.1% Milan reductions are surrogate model predictions for their climate-specific balanced configurations. (See Figure 2). Cooling reductions were partly offset by small increases in heating demand, illustrating why separate heating and cooling loads are retained in direct post-processing even though the optimization layer uses Q_tot and equal SCOP/SEER values in the reported scenarios.

Table 11. Annual energy demands from the conditioned model for benchmark scenarios.

Scenario	Palermo Q_c (kWh)	Palermo Q_h (kWh)	Palermo Q_tot (kWh)	Rome Q_tot (kWh)	Milan Q_tot (kWh)
S0	5512	1204	6716	6183	7941
S1 (RT28)	5082 (−7.8%)	1238 (+2.8%)	6320 (−5.9%)	5928 (−4.1%)	7683 (−3.3%)
S2 (FA-2 nom.)	5167 (−6.3%)	1227 (+1.9%)	6394 (−4.8%)	5998 (−3.0%)	7733 (−2.6%)
Balanced [1]	5082 (−7.8%)	1231 (+2.2%)	6313 (−6.0%)	5982 (−3.3%)	7694 (−3.1%)

[1] S0–S2 are direct EnergyPlus benchmark simulations in all climates. The balanced Palermo values are direct S4 EnergyPlus results; the Rome and Milan total demand values are surrogate model predictions for the climate-specific balanced configurations reported later in Section 4.4.

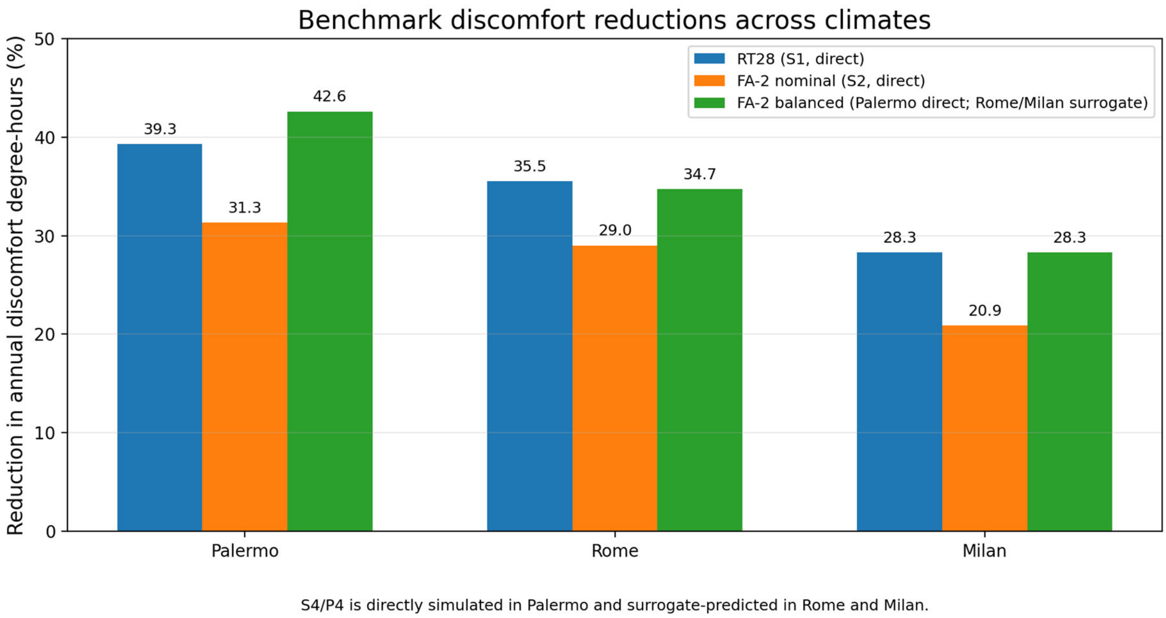

Figure 2. Benchmark reduction in annual occupied discomfort degree-hours. S1 and S2 are direct EnergyPlus simulations in all three climates. The balanced configuration is direct for Palermo (S4/P4 property set) and surrogate-predicted for the climate-specific Rome and Milan configurations.

4.2. Surrogate Performance, Frontier Diagnostics, and Evidence Qualification

Table 12 reports pooled held-out performance across the three climate-specific models (2700 test cases in total; 900 per climate). XGBoost achieved the highest available aggregate accuracy for both thermal targets, with $R^2 = 0.974$ for DH and 0.978 for Q_tot. These metrics quantify interpolation on stratified random held-out samples; they are not treated as a substitute for climate-specific residual distributions or frontier validation. Because the archived prediction vectors needed to reconstruct those distributions are unavailable, the revision adopts a stricter inference rule rather than inventing residual statistics: pooled metrics are used only to justify surrogate screening, while direct EnergyPlus evidence governs applied claims at decision-critical points. Training times are retained as descriptive archived values and are not presented as hardware-independent benchmarks.

Table 12. Pooled held-out predictive performance across three climate-specific surrogate models (2700 test cases total; 900 per climate). Training times are descriptive archived values, not hardware-normalized benchmarks.

Model	DH R^2	DH NRMSE	Q_tot R^2	Training Time
Multiple Linear Regression	0.803	0.094	0.821	<1 s
Random Forest	0.957	0.042	0.963	3.2 min
XGBoost	0.974	0.034	0.978	1.8 min
Gaussian Process Regression	0.958	0.041	0.961	12.4 min
Multi-Layer Perceptron	0.966	0.037	0.970	8.7 min

Within the fitted DH surrogate, melting temperature had the highest mean absolute SHAP contribution, followed by latent heat, layer thickness, and position. The fitted dependence profile was U-shaped, with higher predicted discomfort at transition temperatures below or above the climate-specific activation region. High latent heat produced favorable predictions only when the transition temperature enabled repeated cycling. For Q_tot, transition temperature and layer geometry also remained influential. These SHAP patterns are treated as hypotheses about the fitted EnergyPlus response surface rather than standalone proof of physical mechanisms. The direct diagnostics in Table 13 use only archived EnergyPlus outputs and therefore add an explicit check at the decision boundary and a cross-climate directional check of the transition property interpretation.

Table 13. Direct EnergyPlus consistency diagnostics supported by the retained benchmark outputs.

Direct Check	EnergyPlus/Surrogate Values	Derived Diagnostic	Interpretation
P4/S4 Palermo frontier point	P4 surrogate DH = 204; S4 direct DH = 187 °C·h/year	+17 °C·h/year; 9.1% surrogate overprediction; 5.2 percentage difference in baseline reduction	Pointwise Pareto-region residual; direct decision-point check
S2 → S3 Palermo	224 → 203 °C·h/year	−9.4% DH	Matched geometry; transition properties changed jointly
S2 → S3 Rome	176 → 158 °C·h/year	−10.2% DH	Matched geometry; transition properties changed jointly
S2 → S3 Milan	148 → 129 °C·h/year	−12.8% DH	Matched geometry; transition properties changed jointly

The P4/S4 comparison quantifies the available surrogate-versus-direct error at a Pareto decision point: the surrogate overpredicts Palermo DH by 17 °C·h/year (9.1% relative to the direct value). This local error is larger than the 3 °C·h/year surrogate-predicted separation between P2 (207) and P4 (204), so their comfort ordering is explicitly classified as unresolved rather than validated. The S2 → S3 comparisons provide an additional direct EnergyPlus consistency check: with the FA-2 family, 25 mm thickness, and inner placement held constant, changing the transition-property parameterization reduces DH by 9.4%, 10.2%, and 12.8% in Palermo, Rome, and Milan, respectively. Because T_m and L_h change together, these values are directional checks and not one-at-a-time local sensitivities. Table 14 formalizes the resulting evidence hierarchy. A 10–20-point direct frontier campaign, climate-specific held-out residual distributions, and a one-at-a-time EnergyPlus sensitivity sweep are not reconstructed from absent raw files; instead, any conclusion that would require those missing analyses is withheld. This converts uncertainty at the frontier into an explicit decision boundary rather than an unqualified optimization claim.

Table 14. Evidence hierarchy and permitted interpretation of reviewer-critical results.

Reviewer-Critical Result	Evidence Actually Available	Status in This Manuscript	Permitted Interpretation
Palermo balanced thermal performance	Matched direct EnergyPlus S4 simulation of the P4 property set	Direct numerical evidence for the stated model boundary	May support quantitative Palermo south-wall/single-zone thermal claims
P4 surrogate prediction	XGBoost prediction plus matched S4 direct residual	Frontier diagnostic only	Used to quantify local surrogate error; not substituted for the direct S4 value

Table 14. *Cont.*

Reviewer-Critical Result	Evidence Actually Available	Status in This Manuscript	Permitted Interpretation
S2 → S3 transition property response	Direct EnergyPlus benchmarks in Palermo, Rome, and Milan	Directional physical consistency evidence	Supports relevance of transition property tuning, not a one-factor causal derivative
P1–P3 Palermo and balanced Rome/Milan points	Surrogate screening + analytical GWP/LCC post-processing	Exploratory candidates	May guide follow-up simulations; may not be presented as validated optima or realized performance
SHAP patterns	TreeSHAP on held-out surrogate predictions, triangulated with direct checks above	Model interpretation	Hypothesis generation only; no independent causal physics claim

Table 14 is the operative inference control rule for the revised manuscript. It prevents a high pooled R^2 value from being used to overstate unverified frontier accuracy and prevents SHAP from being promoted to physical validation. Accordingly, the absence of a reconstructable 10–20-point campaign does not generate surrogate-only applied claims: those claims are excluded by design until matched direct simulations are available. (See Figure 3).

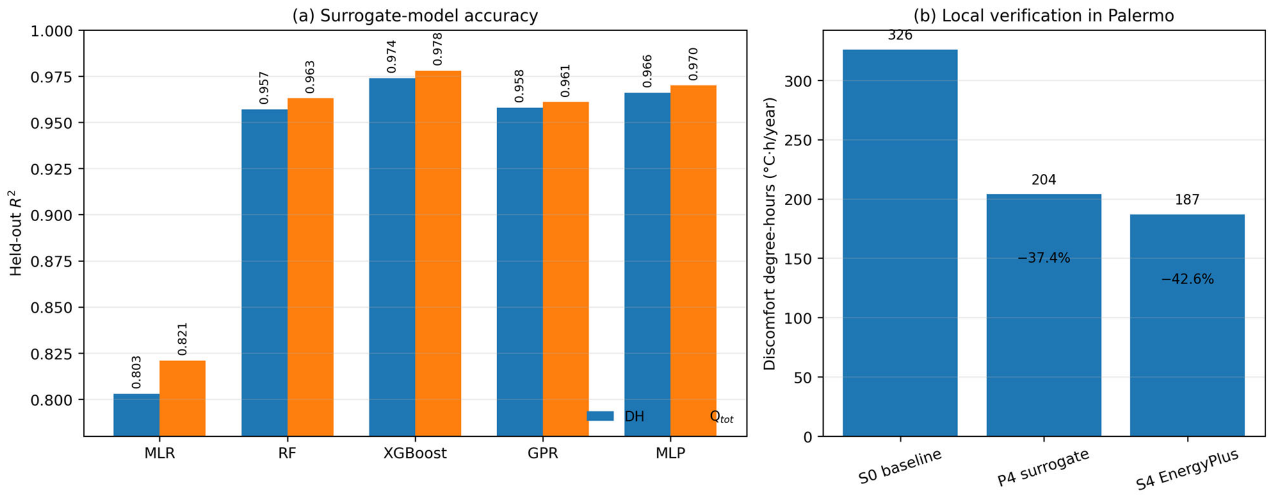

Figure 3. Surrogate performance and direct verification: (**a**) held-out R^2 for discomfort degree-hours (DH) and total thermal demand (Q_tot); (**b**) Palermo discomfort for the no-PCM baseline, the P4 surrogate prediction, and the directly simulated S4 configuration.

4.3. Environmental and Economic Performance

Table 15 presents the central-scenario carbon calculation for the direct Palermo benchmarks. Ideal heating- and cooling-load savings were converted separately to equivalent electricity using SCOP = 3.0 and SEER = 3.0 before applying the grid carbon factor. PCM areal mass was calculated from layer thickness, the effective PCM content factor φ, and the 850 kg/m^3 screening density; the comparative cradle-to-gate PCM production coefficients were 1.4 kg CO_2eq/kg for the bio-based PCM and 2.9 kg CO_2eq/kg for paraffin. Encapsulation/packaging, installation, transport, maintenance, demolition, and end-of-life impacts remain outside the central carbon boundary and are addressed only through the threshold stress test.

Table 15. Lifecycle carbon breakdown for benchmark scenarios (Palermo, 25-year period, per m^2 south wall).

GWP Component (kg CO_2eq/m^2)	S0	S1 (RT28)	S2 (FA-2 Nominal)	S4 (FA-2 Optimised)
PCM production (cradle-to-gate)	0	+24.7	+11.9	+10.5
Avoided operational emissions	0	−32.0	−26.1	−32.6
Net GWP	0	−7.4	−14.2	−22.1
Environmental payback (years)	-	19.2	11.4	8.0

Central scenario: SCOP = SEER = 3.0, EF_grid = 0.233 kg CO_2eq/kWh, and 25-year service life. Values are rounded and represent a comparative screening boundary limited to PCM production and avoided operational emissions; they are not material-specific environmental product declarations.

All three PCM benchmarks produce net carbon savings within the central Palermo accounting boundary, but the magnitude is lower than would be obtained by treating thermal loads as equivalent electricity at SCOP = SEER = 1.0. The optimized bio-based configuration S4 has net GWP of -22.1 kg CO_2eq/m^2 and an environmental payback of 8.0 years; RT28 yields -7.4 kg CO_2eq/m^2 and a 19.2-year payback, while S2 yields -14.2 kg CO_2eq/m^2. Because excluded processes are not assigned fabricated central inventories, their influence is expressed as a threshold: an additional omitted-process burden greater than 7.4, 14.2, or 22.1 kg CO_2eq/m^2 would reverse the central net-carbon sign of S1, S2, or S4, respectively. These values are algebraic break-even thresholds, not estimates of packaging, transport, installation, maintenance, or end-of-life impacts. The comparison therefore demonstrates why complete primary inventories are required before product-level environmental claims.

The initial cost is transparently decomposed into the PCM material cost and the position-dependent integration allowance defined in Section 2.3. Under the central all-electric assumptions, none of the representative designs recover their additional investment from energy savings alone within 25 years. P3 has the lowest incremental LCC (12.9 EUR/m^2), whereas P4 retains an incremental LCC of 24.1 EUR/m^2. The expanded integration cost stress test reveals the critical condition directly: with all other central assumptions fixed, the break-even integration allowance Cint* is 5.0 EUR/m^2 for P2 and 7.1 EUR/m^2 for P3. For P1 and P4, no non-negative integration allowance can produce energy-only break-even because material cost alone exceeds the present value of energy savings by 48.7 and 4.1 EUR/m^2, respectively. Thus, integration cost uncertainty can change the economics of low-material solutions, but central P4 profitability would additionally require a lower material price, higher tariff, different system efficiency, longer life, or monetized non-energy benefits. Its energy-only break-even electricity tariff remains approximately 0.53 EUR/kWh. (See Table 16).

Table 16. Central economic indicators for representative Palermo solutions (25 years, per m^2 of south wall).

Indicator	P1 (Comfort)	P2 (Low Material)	P3 (Lowest LCC)	P4 (Balanced)
PCM areal mass (kg/m^2)	18.8	4.1	2.3	7.5
PCM material cost (EUR/m^2)	78.9	17.1	9.8	31.4
Integration allowance Cint (EUR/m^2)	7.5	20.0	20.0	20.0
Initial additional cost C_0 (EUR/m^2)	86.4	37.1	29.8	51.4
Annual thermal load saving ($kWhth/m^2{\cdot}yr$)	18.6	13.6	10.4	16.8
Annual electricity cost saving ($EUR/m^2{\cdot}yr$)	1.73	1.27	0.97	1.57
PV of 25-year energy savings (EUR/m^2)	30.2	22.1	16.9	27.3
Incremental LCC (EUR/m^2)	+56.2	+15.0	+12.9	+24.1
Break-even electricity tariff (EUR/kWh)	0.80	0.47	0.49	0.53

4.4. Pareto Screening and Climate-Dependent Candidate Qualification

The Palermo surrogate-selected representative set exhibits clear screening trade-offs. P1 is comfort-focused and requires the greatest material quantity. P2 is a low-material carbon–cost compromise; P3 has the lowest incremental LCC; and P4 lies near the knee of the comfort–carbon–cost surface. These relationships describe the archived surrogate/analytical decision surface, not four independently validated EnergyPlus optima. Because only the P4 property set has a matched direct Palermo simulation, P1–P3 are retained solely to illustrate candidate diversity and to prioritize subsequent direct checks. No applied recommendation is based on their unverified thermal ordering. (See Table 17).

P4 is retained as the primary laboratory follow-up candidate because it is the only balanced Pareto property set with a matched direct Palermo EnergyPlus check and because its target properties are compatible with the palmitic–capric family. The direct S4 result, not the surrogate P4 DH value, supports the stated Palermo thermal reduction. Its positive incremental LCC of +24.1 EUR/m^2 is reported explicitly, so no claim of current energy-only profitability or deployment readiness is made.

Figure 4 visualizes the surrogate/analytical screening trade-offs among the four selected Palermo candidates. The bubble area represents incremental LCC. The figure is a candidate prioritization map rather than a validation plot: relative positions of P1–P3 are not interpreted as confirmed EnergyPlus rankings, whereas P4 is linked to its matched direct S4 check in Tables 7–9.

The surrogate screening suggests a downward shift in the balanced transition-temperature target from Palermo to Milan, consistent with cooler operating conditions. The Palermo value is linked to the directly checked S4 property set; the Rome and Milan values are hypotheses generated by the climate-specific surrogates. Therefore, the numerical Rome/Milan reductions, carbon values, paybacks, and LCC values in Table 18 are not presented as validated performance forecasts. Their role is to demonstrate how the workflow changes the candidate search region with climate and to define the configurations that should be re-simulated next.

Table 17. Surrogate-selected Palermo screening candidates. P4 has a matched direct EnergyPlus check (S4); P1–P3 remain exploratory until directly re-simulated.

Property	P1 (Comfort)	P2 (Low Material)	P3 (Lowest LCC)	P4 (Balanced)
T_m (°C)	25.8	24.1	23.4	24.8
L_h (kJ/kg)	213	148	132	178
k_s (W/mK)	0.17	0.16	0.18	0.17
Thickness t (mm)	34	16	11	22
Effective PCM content factor φ	0.65	0.30	0.25	0.40
Wall position	Inner (C1)	Intermediate (C2)	Intermediate (C2)	Intermediate (C2)
DH (°C·h/year) [1]	182 (−44.2%)	207 (−36.5%)	253 (−22.4%)	204 (−37.4%)
Net GWP (kg CO_2eq/m^2) [1]	−9.8	−20.6	−16.8	−22.1
Incremental LCC (EUR/m^2) [1]	+56.2	+15.0	+12.9	+24.1

[1] DH values in this table are surrogate predictions used during screening. Net GWP and incremental LCC are analytical post-processing results under the central assumptions in Table 5. The P4 property set was additionally simulated directly in Palermo as S4; direct S4 thermal values are reported in Tables 10 and 11 and used for applied thermal reporting. P1–P3 are not described as validated optima.

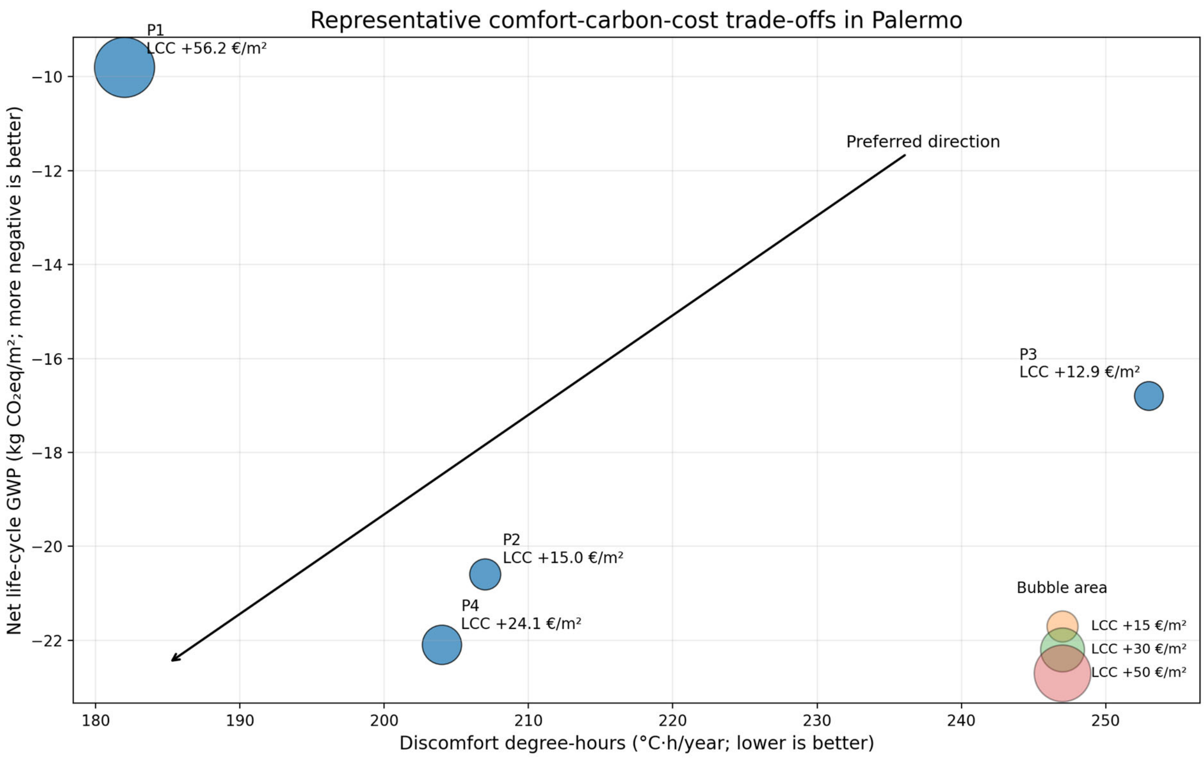

Figure 4. Representative Palermo trade-offs between discomfort, net lifecycle GWP, and incremental LCC. The bubble area is proportional to incremental LCC; lower discomfort and a more negative GWP are preferred.

Table 18. Climate-specific balanced screening candidates. Palermo's thermal performance is direct EnergyPlus S4 evidence; Rome and Milan values are exploratory surrogate estimates and are not used as validated applied-performance claims.

Property	Palermo	Rome	Milan
T_m (°C)	24.8	23.1	21.4
L_h (kJ/kg)	178	165	152
Thickness t (mm)	22	20	17
DH reduction (%) [1]	−42.6	−34.7	−28.3
Net GWP (kg CO_2eq/m^2)	−22.1	−5.8	−9.5
Environmental payback (years)	8.0	16.1	13.1
Incremental LCC (EUR/m^2)	+24.1	+38.7	+35.6

[1] Palermo DH is from the direct EnergyPlus simulation S4. Rome and Milan DH reductions are surrogate screening estimates for their respective balanced candidates and have not been directly re-simulated. Net GWP and incremental LCC are analytical post-processing results under the central assumptions in Table 5. Rome/Milan values are retained only to formulate climate-dependent hypotheses and follow-up simulation targets.

The bounded scenario check confirms that the carbon conclusion for P4 remains favorable across the tested combinations, although the margin narrows in the low-benefit case with a low-carbon grid, high SCOP/SEER, and shorter service life (See Table 19). Economic performance is more conditional: P4 remains cost-increasing in the low-benefit and central cases but becomes cost-saving in the high-benefit combination of lower conversion efficiency, high electricity price, long service life, and low discount rate. These are scenario envelopes, not occurrence probabilities. Figure 5 displays the corresponding change in sign and magnitude.

Table 19. Bounded scenario check for the balanced Palermo configuration P4.

Outcome	Low-Benefit	Central	High-Benefit	Interpretation
Net GWP P4 (kg CO_2eq/m^2)	−2.9	−22.1	−66.6	Carbon saving remains negative
Environmental payback (years)	14.6	8.0	3.8	Strong grid/efficiency dependence
Incremental LCC P4 (EUR/m^2)	+40.9	+24.1	−26.6	Economic sign is conditional
Discounted energy payback	>30 years	>25 years	19 years	No central energy-only payback

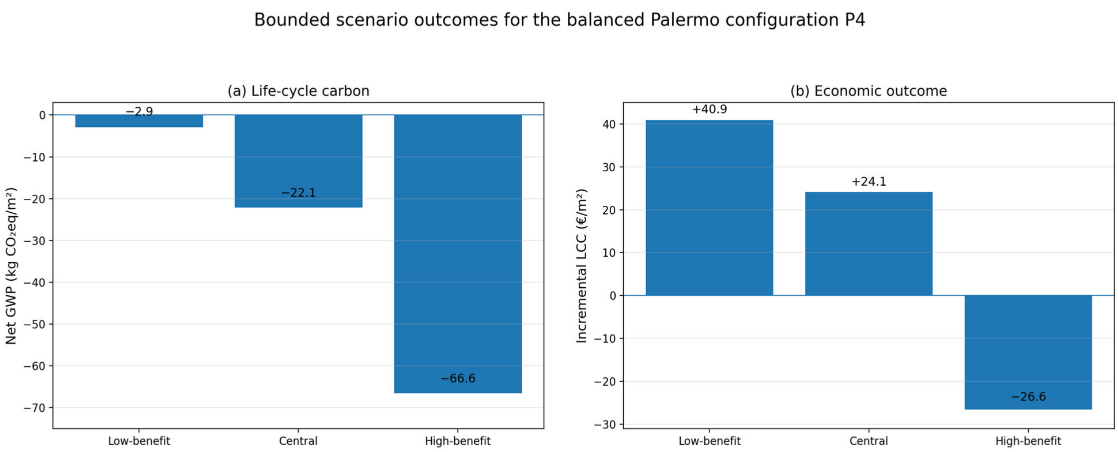

Figure 5. Bounded scenario outcomes for the balanced Palermo configuration P4: (**a**) net lifecycle GWP and (**b**) incremental LCC. Negative GWP denotes net carbon savings; negative LCC denotes cost recovery. Low-/high-benefit labels refer to the benefit attributed to the PCM relative to S0.

https://doi.org/10.3390/su18179041

Low-benefit carbon case: SCOP = SEER = 4.0, EF_grid = 0.160 kg CO_2eq/kWh, 20 years. High-benefit carbon case: SCOP = SEER = 2.0, EF_grid = 0.306 kg CO_2eq/kWh, 30 years. Low-benefit cost case: SCOP = SEER = 4.0, tariff 0.20 EUR/kWh, r = 5%, 20 years. High-benefit cost case: SCOP = SEER = 2.0, tariff 0.36 EUR/kWh, r = 1%, 30 years.

5. Discussion

5.1. Thermophysical Drivers and Climate Dependence

The surrogate assigns the largest mean absolute SHAP contribution to melting temperature for free-running discomfort. Its fitted U-shaped dependence is qualitatively consistent with the requirement that a PCM repeatedly traverse its transition interval during charging and regeneration. The evidence hierarchy in Table 14 governs interpretation: direct EnergyPlus results take precedence over surrogate explanations. The P4/S4 comparison provides a quantified decision-point residual, while the matched-geometry S2 → S3 benchmark reduces direct DH in all three climates (Table 13), supporting the relevance of transition property tuning. The S2 → S3 check remains directional because T_m and L_h change jointly. Published EnergyPlus PCM verification/validation and hysteresis studies support the solver lineage while also showing why partial-transition behavior requires care [48–50]. SHAP is therefore used only to generate and organize hypotheses about the fitted response surface; it is not presented as independent causal evidence.

The fitted surrogate also indicates diminishing thermal benefit from increasing latent heat when the transition window is poorly matched to the daily thermal cycle. A physically plausible interpretation is that storage capacity is useful only if sufficient charging and re-solidification occur between successive load cycles; otherwise, additional latent capacity can remain underutilized. Because the present study does not provide a one-at-a-time EnergyPlus sweep of latent heat at fixed transition temperature, this explanation is presented as a testable interpretation of the fitted response surface rather than as a directly validated causal law. Experimental formulation and direct sensitivity analysis should therefore target the combination of latent capacity and reliable regeneration rather than maximizing latent heat in isolation.

The screening trend toward a lower preferred T_m from Palermo (24.8 °C, directly verified for the balanced property set) to Milan (21.4 °C, surrogate prediction) supports the hypothesis that transition temperature should be calibrated to local operating conditions. A product optimized for Mediterranean performance may therefore be suboptimal in a cooler context, but the numerical values cannot be generalized beyond the three studied climates. Within the selected Palermo representatives, intermediate positioning (C2) occurs in P2–P4, whereas the comfort-focused P1 uses inner placement (C1). This is a study-specific screening pattern, not a universal placement rule; wall construction, boundary conditions, orientation, and climate can alter the preferred position.

5.2. Environmental and Economic Trade-Offs

Under the central scenario, bio-based configurations outperform RT28 in net GWP because their lower assumed cradle-to-gate impact compensates for similar or slightly lower operational savings. For Palermo, environmental payback decreases from 19.2 years for RT28 to 8.0 years for the optimized bio-based design. This conclusion remains conditional on the secondary inventory coefficients, the truncated lifecycle boundary, a static grid factor, and the all-electric SCOP/SEER assumptions. Recent PCM 3E analysis and broader envelope optimization likewise show that economic and carbon conclusions depend strongly on system boundaries, hysteresis representation, and grid assumptions [39,40]. Primary industrial carbon data and dynamic operational emission scenarios could shift absolute values and should be obtained before product-level claims.

The economic analysis is deliberately conservative. Once ideal thermal loads are converted through SCOP and SEER, the central 25-year energy savings do not recover the installed PCM cost. The lowest-LCC design carries an incremental cost of 12.9 EUR/m^2, and the balanced design's is 24.1 EUR/m^2. Accordingly, the near-term case for PCM deployment rests on a combination of lower installed costs, higher electricity prices, avoided cooling capacity investment, and monetized overheating resilience rather than on energy savings alone.

5.3. The AI Layer: Acceleration and Interpretation

The surrogate layer reduces the need for repeated annual EnergyPlus evaluations during optimization and therefore makes broad design-space screening computationally practical. A hardware-independent acceleration factor is not claimed here because complete compute environment and timing metadata were not preserved. The defensible computational conclusion is qualitative: once trained, the surrogate replaces repeated dynamic simulations during NSGA-II evaluation, while decision-critical candidates still require direct EnergyPlus verification before experimental selection.

The AI layer is therefore best understood as an acceleration and prioritization tool. It identifies climate-dependent regions of interest and fitted interactions that would be expensive to explore exhaustively with annual simulation, but its reliability is bounded by the sampled design space, available training records, and local surrogate error. Recent reviews of surrogate-based sustainable building design likewise emphasize that optimization quality depends on validation strategy and reproducibility rather than on global prediction metrics alone [34]. In this study, economics remains outside the learned thermal model and enters only through explicit scenario equations, while Pareto candidates remain provisional until directly re-simulated.

5.4. Sustainability Implications

The framework treats sustainability as a multi-criteria decision problem rather than equating lower thermal demand with a sustainability claim. Lifecycle carbon accounting links embodied production burden with avoided operational emissions; incremental LCC tests whether those benefits are economically recoverable within the chosen horizon, and occupied discomfort degree-hours quantify passive thermal resilience within the free-running numerical model. The representative solutions demonstrate that these dimensions do not improve simultaneously. This makes transparent multi-objective decision support preferable to selecting a PCM from a single thermal metric and clarifies which conclusions are simulation-derived, which are scenario-dependent, and which still require experimental evidence.

For decarbonized construction, the main contribution is methodological: climate-specific screening can reduce the risk of specifying a PCM that remains inactive, fails to regenerate, or shifts impacts from operation to material production. The approach supports SDG 7 through energy efficiency assessment, SDG 9 through materials and digital innovation, SDG 11 through climate-responsive and resilient buildings, SDG 12 through lifecycle and resource considerations, and SDG 13 through explicit carbon accounting [1,68,69]. These links do not establish product-level sustainability certification; they identify the evidence required before responsible scale-up and procurement.

The recent literature confirms that no single component of the present workflow is individually unprecedented. The contribution lies in how the components are combined and bounded for bio-based PCM pre-experimental screening. Table 20 therefore replaces an overly binary novelty claim with an analytical comparison of scope, strengths, and unresolved gaps across closely related recent studies.

Table 20. Analytical positioning relative to recent PCM envelope, surrogate optimization, and lifecycle studies. The comparison is qualitative and is intended to delimit methodological scope rather than rank study quality.

Study	Main Scope	Relevant Strength	Remaining Gap Relative to the Present Screening Question
Jiao et al. [17]	State-of-the-art PCM envelope review	Broad synthesis of integration strategies and recent envelope research	Review scope; no integrated surrogate–multi-objective–lifecycle workflow for bio-based candidate screening
Herrera et al. [18]	Fatty-acid PCMs for building applications	Recent synthesis of bio-sources, thermophysical limitations, incorporation, and LCA needs	No climate-specific surrogate optimization or Pareto screening
Chan et al. [48]	Annual EnergyPlus PCM simulation with basic and hysteresis methods	Shows climate and hysteresis sensitivity in annual building simulation	Paraffin-focused; no XAI, lifecycle carbon/cost, or multi-objective screening
Feng et al. [49]	Enhanced PCM hysteresis modeling in EnergyPlus	Demonstrates importance of constitutive hysteresis representation	Model development focus rather than bio-based multi-criteria design screening
Elwy and Hagishima [34]	Systematic review of surrogate-based sustainable building optimization	Highlights surrogate efficiency, validation, and methodological diversity	Not PCM-specific and does not establish material formulation feasibility
Zhang et al. [39]	Energy–economic–environmental PCM wall analysis with hysteresis	Integrates operational performance, economics, carbon, and phase transition detail	No surrogate explainability or bio-based continuous multi-objective screening
Hepple et al. [42]; Lahoud et al. [43]	Climate- and typology-dependent PCM envelope applications	Demonstrates that PCM placement/performance is context-dependent	Different building types/climates; no unified surrogate + lifecycle decision workflow
Li et al. [40]	Surrogate-assisted multi-objective envelope optimization with dynamic grid factors	Combines comfort, carbon, cost, surrogate modeling, and NSGA-II-type optimization	Building envelope variables rather than PCM formulation; not specific to bio-based PCMs
Present study	18,000 EnergyPlus cases + XGBoost/TreeSHAP + NSGA-II + scenario-based carbon/LCC	Common south-wall functional unit, explicit separation of simulated thermal outputs from post-processing, and bio-based/paraffin comparison	Continuous-property manufacturability, three-climate scope, incomplete lifecycle boundary, and only one directly re-simulated Pareto solution

5.5. Limitations and Scope of Applicability

The revised manuscript applies explicit scope controls so that each known limitation has a corresponding restriction on inference. First, independently sampled PCM properties do not preserve all formulation correlations; Pareto points are therefore target-property combinations and pass through a manufacturability gate before any experimental recommendation [18,19,70,71]. Second, the adopted EnergyPlus enthalpy representation does not resolve detailed hysteresis/supercooling, and no case-specific grid/timestep convergence record is available. Published solver-level verification/validation [48–50] supports the numerical method but not the convergence of this exact wall; accordingly, results are screening-level model predictions, and no product-, code-, or experimentally validated claim is made. Third, the 24 m^2 south wall in a single-zone model is the functional comparison surface. For Palermo, S4 saves 403 kWh/year, equal to 16.8 kWhth/m^2 of the analyzed wall and 6.0% of the modeled zone's ideal thermal demand; this value is not extrapolated to a whole building. Fourth, only Palermo, Rome, and Milan are studied. Rome/Milan balanced configurations are surrogate hypotheses, and the numerical values are not generalized to Northern Europe, the Middle East, or other typologies [42,43]. Fifth, Q_tot is the optimization surrogate target; asymmetric SCOP/SEER rankings are therefore explicitly outside of the scope, while direct benchmark post-processing retains Q_h and Q_c separately. Sixth, the central carbon boundary excludes encapsulation/packaging, transport, installation, maintenance, demolition, and end-of-life; the omitted-process threshold stress test quantifies the extra burden required to reverse the sign but is not represented as a fabricated inventory. Seventh, integration allowances are screening assumptions rather than quotations, so economic conclusions are bounded by the 0–60 EUR/m^2 stress test and break-even conditions rather than presented as market forecasts. Eighth, frontier validation is limited to the direct P4/S4 residual and S2 $\rightarrow$ S3 cross-climate checks. Instead of treating unvalidated Pareto points as confirmed optima, Table 14 excludes them from applied performance claims until direct re-simulation. SHAP likewise remains non-causal until a one-at-a-time direct sensitivity campaign is available. Ninth, the archived workflow supports protocol-level but not bitwise computational reproducibility because final fitted parameter vectors, split indices, seeds, and some trace settings were not preserved. These missing values are not reconstructed by assumption; Section 3 specifies exactly which elements are available and what conclusions are permitted from them. These controls make the inferential boundary of this study explicit and ensure that no conclusion exceeds the evidence actually retained.

5.6. Future Research Directions

Future research should extend, rather than retroactively inflate, the evidence available in the present screening study. First, directly re-simulate at least 10–20 Pareto candidates per climate, including the Rome and Milan balanced candidates, and report objective-wise errors plus residual distributions in the frontier region. Second, perform one-at-a-time EnergyPlus perturbations around balanced configurations for T_m, L_h, thickness, effective PCM content factor, and position, and compare finite-difference responses with SHAP dependence patterns. Third, train separate Q_h and Q_c surrogates so that SCOP and SEER can vary independently. Fourth, perform a case-specific grid/timestep convergence study and validate the adopted PCM parameterization against an experimental or established numerical benchmark, including separate melting/freezing branches and hysteresis where supported by measured material data. Fifth, formulate and characterize a P4-equivalent palmitic–capric composite using DSC, T-history where appropriate, leakage testing, fire safety assessment, conductivity measurements, and 1000–3000 thermal cycles before climate chamber validation of a test wall specimen. Sixth, replace screening carbon factors and integration allowances with supplier-specific inventories, transport/installation/end-of-life data, quotations, and dynamic electricity–carbon scenarios. Seventh, extend the model to multi-zone and whole-building typologies with thermal bridges and explicit HVAC systems. Finally, apply the workflow to substantially different climates, including Northern Europe and the Middle East. The transferable contribution is the evidence-controlled screening methodology; numerical targets, costs, grid factors, and formulations must be recalibrated and directly validated locally.

6. Conclusions

6.1. Theoretical and Methodological Conclusions

1. The numerical screening identifies a climate-dependent activation region in which transition temperature has the largest fitted influence on free-running discomfort, followed by latent heat and layer geometry. The downward Palermo-to-Milan shift is reported as a surrogate-screening trend rather than a universal material constant. Direct evidence supports transition property relevance through the S2 $\rightarrow$ S3 benchmark in all three climates, while only the Palermo balanced property set is directly checked at the Pareto decision level.
2. XGBoost achieved pooled held-out R^2 values of 0.974 for discomfort and 0.978 for total thermal demand, demonstrating suitability for candidate discovery within the sampled design space. The P4/S4 residual shows why high global accuracy cannot certify a Pareto point. The methodological rule adopted here is therefore explicit: surrogate models locate candidates, direct EnergyPlus simulation qualifies candidates for applied thermal claims, and SHAP organizes hypotheses but does not prove causal physics.
3. Separating simulation-derived thermal outputs from analytical carbon and cost post-processing makes the sustainability argument auditable. Under the stated central assumptions, the directly simulated Palermo S4 configuration reduces occupied discomfort by 42.6% and the modeled single-zone total thermal demand by 6.0%; its scenario-based net carbon is favorable, whereas incremental LCC remains positive. The contribution is therefore an evidence-controlled, multi-criteria pre-experimental screening method, not a claim of a universally optimal, experimentally validated, or deployment-ready PCM product.

6.2. Practical Recommendations for Italian Applications

1. For the Palermo numerical case, the directly re-simulated balanced target (T_m = 24.8 °C, L_h = 178 kJ/kg, 22 mm thickness, intermediate position) is the only Pareto-derived configuration used for quantitative applied thermal reporting and is therefore the most defensible laboratory follow-up target. Its results are expressed per square meter of the analyzed south wall and as a percentage of the modeled single-zone ideal thermal demand, not as whole-building performance.
2. The palmitic–capric family should be treated as an experimental formulation target, not a deployment-ready product. Practical development should jointly verify transition behavior, composite conductivity, support/encapsulation, leakage resistance, fire safety, cyclic stability, manufacturability, and supplier-specific embodied impacts.
3. Current energy-only profitability is not supported by the central assumptions: P4 has an incremental LCC of 24.1 EUR/m^2, and its material cost alone exceeds the present value of central energy savings by 4.1 EUR/m^2 before any positive integration allowance. Italian implementation should therefore be conditional on lower installed/material costs, different energy price/system assumptions, or monetized resilience and capacity benefits.

6.3. Transferability Beyond Italy

1. The workflow can be transferred to other regions, but the numerical candidate values cannot. Rome and Milan values in this study remain exploratory surrogate outputs. Application to Northern European climates should re-estimate lower transition-

temperature regions, heating-dominated behavior, building typology, and winter regeneration; application to hot or arid Middle Eastern climates should explicitly test high night-time temperatures, incomplete solidification, cooling-dominated operation, solar gains, and ventilation strategies.

2. Every new region requires local weather files, comfort/set-point assumptions, envelope typology, orientation, HVAC conversion efficiencies, grid carbon trajectories, tariffs, material inventories, and direct Pareto validation. The contribution to sustainability practice is therefore a reproducible decision logic for identifying what must be recalibrated and validated before responsible scale-up, consistent with the energy efficiency, innovation, resilient-building, responsible production, and climate action dimensions represented by SDGs 7, 9, 11, 12, and 13 [1,68,69].

Author Contributions: Conceptualization: M.G.I., A.M., G.A., and F.C., Methodology: M.G.I., A.M., G.A., and F.C., Software M.G.I. and A.M., Validation M.G.I., A.M., and G.A., Formal Analysis: M.G.I. and A.M., Investigation: M.G.I. and A.M., Data Curation: M.G.I. and A.M., Writing—Original Draft: M.G.I. and A.M., Writing—Review and Editing: M.G.I. and A.M., Visualization: M.G.I., A.M., G.A., and F.C., Supervision: M.G.I., A.M., G.A., and F.C., Project Administration and Funding Acquisition: G.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was carried out within an industrial PhD program.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No experimental or observational dataset was generated or analyzed. The reported findings derive from the EnergyPlus numerical simulation campaign and the analytical post-processing described in the manuscript. No public model, code, or simulation output repository is available for this study; the resulting reproducibility constraints are stated explicitly in Section 3. EPW weather files are publicly available through the EnergyPlus weather repository. Ecoinvent 3.10 data are commercially licensed and cannot be redistributed. All numerical assumptions required to reproduce the analytical carbon and cost calculations reported in the manuscript are stated in Sections 2.3 and 2.5 and in Table 5.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. IPCC. Climate Change 2023: Synthesis Report. In *Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*; IPCC: Geneva, Switzerland, 2023.
2. European Commission. *Directive 2010/31/EU on the Energy Performance of Buildings*; Official Journal of the EU: Luxembourg, 2010.
3. International Energy Agency. *Global Status Report for Buildings and Construction*; IEA: Paris, France, 2023.
4. European Commission. *Directive 2024/1275 on the Energy Performance of Buildings (Recast)*; Official Journal of the EU: Luxembourg, 2024.
5. Navarro, L.; de Gracia, A.; Colclough, S.; Browne, M.; McCormack, S.J.; Griffiths, P.; Cabeza, L.F. Thermal energy storage in building integrated thermal systems: A review. Part 1. Active storage systems. *Renew. Energy* **2016**, *88*, 526–547. [CrossRef]
6. Zalba, B.; Marín, J.M.; Cabeza, L.F.; Mehling, H. Review on thermal energy storage with phase change: Materials, heat transfer analysis and applications. *Appl. Therm. Eng.* **2003**, *23*, 251–283. [CrossRef]
7. Baetens, R.; Jelle, B.P.; Gustavsen, A. Phase change materials for building applications: A state-of-the-art review. *Energy Build.* **2010**, *42*, 1361–1368. [CrossRef]
8. Kuznik, F.; David, D.; Johannes, K.; Roux, J.-J. A review on phase change materials integrated in building walls. *Renew. Sustain. Energy Rev.* **2011**, *15*, 379–391. [CrossRef]
9. Sharma, A.; Tyagi, V.V.; Chen, C.R.; Buddhi, D. Review on thermal energy storage with phase change materials and applications. *Renew. Sustain. Energy Rev.* **2009**, *13*, 318–345. [CrossRef]
10. Stritih, U.; Tyagi, V.V.; Stropnik, R.; Paksoy, H.; Haghighat, F.; Joybari, M.M. Integration of passive PCM technologies for net-zero energy buildings. *Sustain. Cities Soc.* **2018**, *41*, 286–295. [CrossRef]

11. Akeiber, H.; Nejat, P.; Majid, M.Z.A.; Wahid, M.A.; Jomehzadeh, F.; Famileh, I.Z.; Calautit, J.K.; Hughes, B.R.; Zaki, S.A. A review on phase change material (PCM) for sustainable passive cooling in building envelopes. *Renew. Sustain. Energy Rev.* **2016**, *60*, 1470–1497. [CrossRef]
12. Tyagi, V.V.; Buddhi, D. PCM thermal storage in buildings: A state of art. *Renew. Sustain. Energy Rev.* **2007**, *11*, 1146–1166. [CrossRef]
13. de Gracia, A.; Cabeza, L.F. Phase change materials and thermal energy storage for buildings. *Energy Build.* **2015**, *103*, 414–419. [CrossRef]
14. Pielichowska, K.; Pielichowski, K. Phase change materials for thermal energy storage. *Prog. Mater. Sci.* **2014**, *65*, 67–123. [CrossRef]
15. Kenisarin, M.; Mahkamov, K. Solar energy storage using phase change materials. *Renew. Sustain. Energy Rev.* **2007**, *11*, 1913–1965. [CrossRef]
16. Sarı, A. Thermal characteristics of a eutectic mixture of myristic and palmitic acids as phase change material. *Appl. Therm. Eng.* **2003**, *23*, 1205–1213. [CrossRef]
17. Jiao, K.; Lu, L.; Zhao, L.; Wang, G. Towards Passive Building Thermal Regulation: A State-of-the-Art Review on Recent Progress of PCM-Integrated Building Envelopes. *Sustainability* **2024**, *16*, 6482. [CrossRef]
18. Herrera, P.; De la Hoz Siegler, H.; Clarke, M. Fatty Acids as Phase Change Materials for Building Applications: Drawbacks and Future Developments. *Energies* **2024**, *17*, 4880. [CrossRef]
19. Sun, Y.; He, F.; Wang, J.; Wang, L.; Fu, B.; Tang, W. Preparation and characterization of fatty acid ternary eutectic mixture/Nano-SiO_2 composite phase change material for building applications. *Appl. Energy* **2024**, *376*, 124221. [CrossRef]
20. Grzybek, J.; Nazari, M.; Jebrane, M.; Terziev, N.; Tippner, J.; Petutschnigg, A.; Schnabel, T. Bio-based phase change material for enhanced building energy efficiency: A study of beech and thermally modified beech wood for wall structures. *Energy Storage* **2024**, *6*, e568. [CrossRef]
21. Kosny, J.; Kossecka, E.; Brzezinski, A.; Tleoubaev, A.; Yarbrough, D. Dynamic thermal performance analysis of fiber insulations containing bio-based phase change materials (PCMs). *Energy Build.* **2012**, *52*, 122–131. [CrossRef]
22. Li, A.; Yang, Y.; Ren, Y.; Wan, Y.; Wu, W.; Zhang, H. Bio-based composite phase change material utilizing wood fiber and coconut oil for thermal management in building envelopes. *J. Clean. Prod.* **2024**, *469*, 143177. [CrossRef]
23. Jahangir, M.H.; Alimohamadi, R. A comparative evaluation on energy consumption of a building using bio-based and paraffin-based phase change materials integrated to external building envelope. *Energy Rep.* **2024**, *11*, 3914–3930. [CrossRef]
24. Zhou, D.; Zhao, C.Y.; Tian, Y. Review on thermal energy storage with phase change materials (PCMs) in building applications. *Appl. Energy* **2012**, *92*, 593–605. [CrossRef]
25. Regin, A.F.; Solanki, S.C.; Saini, J.S. Heat transfer characteristics of thermal energy storage system using PCM capsules: A review. *Renew. Sustain. Energy Rev.* **2008**, *12*, 2438–2458. [CrossRef]
26. Ascione, F.; Bianco, N.; De Stasio, C.; Mauro, G.M.; Vanoli, G.P. Artificial neural networks to predict energy performance and retrofit scenarios for any member of a building category: A novel approach. *Energy* **2017**, *118*, 999–1017. [CrossRef]
27. Bourdeau, M.; Zhai, X.Q.; Nefzaoui, E.; Guo, X.; Chatellier, P. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustain. Cities Soc.* **2019**, *48*, 101533. [CrossRef]
28. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; Association for Computing Machinery (ACM): New York, NY, USA, 2016; pp. 785–794.
29. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning*; MIT Press: Cambridge, MA, USA, 2006.
30. Ali, A.; Jayaraman, R.; Mayyas, A.; Alaifan, B.; Azar, E. Machine learning as a surrogate to building performance simulation: Predicting energy consumption under different operational settings. *Energy Build.* **2023**, *286*, 112940. [CrossRef]
31. Mehrotra, A.; Yi, H. Effect of adaptive intelligent sampling and machine-learning emulators in surrogate energy modeling of architectural massing. *J. Build. Eng.* **2023**, *72*, 106614. [CrossRef]
32. Shirzadi, N.; Lau, D.; Stylianou, M. Surrogate modeling for building design: Energy and cost prediction compared to simulation-based methods. *Buildings* **2025**, *15*, 2361. [CrossRef]
33. Teixeira, B.; Carvalhais, L.; Pinto, T.; Vale, Z. Explainable AI framework for reliable and transparent automated energy management in buildings. *Energy Build.* **2025**, *347*, 116246. [CrossRef]
34. Elwy, I.; Hagishima, A. The artificial intelligence reformation of sustainable building design approach: A systematic review on building design optimization methods using surrogate models. *Energy Build.* **2024**, *323*, 114769. [CrossRef]
35. Lundberg, S.M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J.M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S.-I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2*, 56–67. [CrossRef]
36. Deb, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* **2002**, *6*, 182–197. [CrossRef]
37. Soares, N.; Gaspar, A.R.; Santos, P.; Costa, J.J. Multi-dimensional optimization of the incorporation of PCM-drywalls in lightweight steel-framed residential buildings in different climates. *Energy Build.* **2014**, *70*, 411–421. [CrossRef]

38. Yang, H.; Xu, Z.; Shi, Y.; Tang, W.; Liu, C.; Yunusa-Kaltungo, A.; Cui, H. Multi-objective optimization designs of phase change material-enhanced building using the integration of the Stacking model and NSGA-III algorithm. *J. Energy Storage* **2023**, *68*, 107807. [CrossRef]
39. Zhang, L.; Zhang, Q.; Jin, L.; Cui, X.; Yang, X. Energy, economic and environmental (3E) analysis of residential building walls enhanced with phase change materials. *J. Build. Eng.* **2024**, *84*, 108503. [CrossRef]
40. Li, C.; Pan, Y.; Liu, Z.; Liang, Y.; Yuan, X.; Huang, Z.; Zhou, N. Optimal design of building envelope towards life cycle performance: Impact of considering dynamic grid emission factors. *Energy Build.* **2024**, *323*, 114770. [CrossRef]
41. Idouanaou, A.; Malha, M.; Kardellass, S.; Bah, A.; Ansari, O.; El Attar, R.; Cherqi, O. Integrated evaluation of bio-based phase change materials to reduce operational and embodied carbon in service buildings across multiple climate zones. *Buildings* **2025**, *15*, 3720. [CrossRef]
42. Hepple, R.; Zhao, Y.; Yang, R.; Zhang, Q.; Yang, S. Investigating the Effects of PCM-Integrated Walls on Thermal Performance for UK Residential Buildings of Different Typologies. *Buildings* **2024**, *14*, 3382. [CrossRef]
43. Lahoud, C.; Chahwan, A.; Rishmany, J.; Yehia, C.; Daaboul, M. Enhancing Energy Efficiency in Mediterranean Coastal Buildings Through PCM Integration. *Buildings* **2024**, *14*, 4023. [CrossRef]
44. Sarı, A.; Karaipekli, A. Thermal conductivity and latent heat thermal energy storage characteristics of paraffin/expanded graphite composite. *Appl. Therm. Eng.* **2007**, *27*, 1271–1277. [CrossRef]
45. Rubitherm GmbH. *Technical Data Sheet RT28*; Rubitherm GmbH: Berlin, Germany, 2023. Available online: https://www.rubitherm.com/ (accessed on 30 August 2026).
46. Sarı, A.; Karaipekli, A. Preparation, thermal properties and thermal reliability of palmitic acid/expanded graphite composite as form-stable PCM for thermal energy storage. *Sol. Energy Mater. Sol. Cells* **2009**, *93*, 571–576. [CrossRef]
47. U.S. Department of Energy. *EnergyPlus Engineering Reference*; version 24.1.0; U.S. Department of Energy: Washington, DC, USA, 2024.
48. Chan, Y.; Hoke, T.; Meredith, K.; Chen, X. Annual Simulation of Phase Change Materials for Enhanced Energy Efficiency and Thermal Performance of Buildings in Southern California. *Energies* **2025**, *18*, 847. [CrossRef]
49. Feng, F.; Fu, Y.; Yang, Z.; O'Neill, Z. Enhancement of phase change material hysteresis model: A case study of modeling building envelope in EnergyPlus. *Energy Build.* **2022**, *276*, 112511. [CrossRef]
50. Tabares-Velasco, P.C.; Christensen, C.; Bianchi, M. Verification and validation of EnergyPlus phase change material model for opaque wall assemblies. *Build. Environ.* **2012**, *54*, 186–196. [CrossRef]
51. Jamekhorshid, A.; Sadrameli, S.M.; Farid, M. A review of microencapsulation methods of phase change materials (PCM) as a thermal energy storage (TES) medium. *Renew. Sustain. Energy Rev.* **2014**, *31*, 531–542. [CrossRef]
52. *ISO E.N. 17772-1:2017*; Energy Performance of Buildings—Indoor Environmental Quality—Part 1: Indoor Environmental Input Parameters for the Design and Assessment of Energy Performance of Buildings. European Committee for Standardization: Brussels, Belgium, 2017.
53. Crawley, D.B.; Lawrie, L.K.; Winkelmann, F.C.; Buhl, W.F.; Huang, Y.J.; Pedersen, C.O.; Strand, R.K.; Liesen, R.J.; Fisher, D.E.; Witte, M.J.; et al. EnergyPlus: Creating a new-generation building energy simulation program. *Energy Build.* **2001**, *33*, 319–331. [CrossRef]
54. Herman, J.; Usher, W. SALib: An open-source Python library for sensitivity analysis. *J. Open Source Softw.* **2017**, *2*, 97. [CrossRef]
55. Fanger, P.O. *Thermal Comfort: Analysis and Applications in Environmental Engineering*; McGraw-Hill: New York, NY, USA, 1970.
56. *ASHRAE Standard 55-2023*; Thermal Environmental Conditions for Human Occupancy, American Society of Heating. Refrigerating and Air-Conditioning Engineers: Atlanta, GA, USA, 2023.
57. *ISO 14040:2006*; Environmental Management—Life Cycle Assessment—Principles and Framework. International Organization for Standardization: Geneva, Switzerland, 2006.
58. *ISO 14044:2006*; Environmental Management—Life Cycle Assessment—Requirements and Guidelines. International Organization for Standardization: Geneva, Switzerland, 2006.
59. Hauschild, M.Z.; Rosenbaum, R.K.; Olsen, S.I. (Eds.) *Life Cycle Assessment: Theory and Practice*; Springer International Publishing: Cham, Switzerland, 2018.
60. ISPRA. *Le Emissioni di CO_2 nel Settore Elettrico Nazionale e Regionale*; Rapporto 413/2025; Istituto Superiore per la Protezione e la Ricerca Ambientale: Rome, Italy, 2025.
61. Eurostat. Electricity Prices for Non-Household Consumers. 2024. Available online: https://ec.europa.eu/eurostat/ (accessed on 30 August 2026).
62. Blank, J.; Deb, K. Pymoo: Multi-objective optimization in Python. *IEEE Access* **2020**, *8*, 89497–89509. [CrossRef]
63. Mekhilef, G.; Golriz, N.; Dorez, H.; Ferry, B.; Bourmaud, S.; Lopez-Cuesta, J.M. Development of biobased and biodegradable PCM nanocomposites for building applications. *Prog. Org. Coat.* **2017**, *110*, 175–183.
64. Kalnæs, S.E.; Jelle, B.P. Phase change materials and products for building applications: A state-of-the-art review and future research opportunities. *Energy Build.* **2015**, *94*, 150–176. [CrossRef]

65. Cabeza, L.F.; Castell, A.; Barreneche, C.; de Gracia, A.; Fernández, A.I. Materials used as PCM in thermal energy storage in buildings: A review. *Renew. Sustain. Energy Rev.* **2011**, *15*, 1675–1695. [CrossRef]
66. Ecoinvent Centre. *Ecoinvent Database v3.10*; Swiss Centre for Life Cycle Inventories: Zürich, Switzerland, 2023.
67. Feldman, D.; Banu, D.; Hawes, D.W. Development and application of organic phase change mixtures in thermal storage gypsum wallboard. *Sol. Energy Mater. Sol. Cells* **1995**, *36*, 147–157. [CrossRef]
68. United Nations. *Transforming Our World: The 2030 Agenda for Sustainable Development*; United Nations: New York, NY, USA, 2015.
69. IEA. *Net Zero by 2050: A Roadmap for the Global Energy Sector*; International Energy Agency: Paris, France, 2021.
70. Rathgeber, C.; Miró, L.; Cabeza, L.F.; Hiebler, S. Measurement of enthalpy curves of phase change materials via DSC and T-History: When are both methods needed to estimate the behaviour of the bulk material in applications? *Thermochim. Acta* **2014**, *596*, 79–88. [CrossRef]
71. Mandilaras, I.; Stamatiadou, M.; Katsourinis, D.; Zannis, G.; Founti, M. Experimental thermal characterization of a Mediterranean residential building with PCM gypsum board walls. *Build. Environ.* **2013**, *61*, 93–103. [CrossRef]

https://doi.org/10.3390/su18179041