sustainability

MDPI

Article

Reducing the Sensing Burden: A Sensor-Light Machine Learning Framework for Thermal Comfort Assessment

Christos Mountzouris [1], Grigorios Protopsaltis [1], Nikos Andriopoulos [1], Dimitrios Koukiasas [1] and John Gialelis [2,*]

[1] Department of Electrical and Computer Engineering, University of Patras, 26504 Patras, Greece; mountzou@ece.upatras.gr (C.M.); g.protopsaltis@ac.upatras.gr (G.P.); nandriopoulos@ece.upatras.gr (N.A.)
[2] Industrial Systems Institute (ISI), Patras Science Park Building, Platani, 26504 Patras, Greece
* Correspondence: gialelis@isi.gr

Abstract

Thermal comfort shapes occupant health, well-being, and productivity and influences the sustainability and energy efficiency of the built environment. The Predicted Mean Vote (PMV) is the most widely used thermal comfort index, yet four of its six input parameters—globe temperature, clothing insulation, metabolic rate, and air velocity—require specialized, costly equipment or occupant self-reporting, which has long limited its practical large-scale application. This study introduces a machine learning (ML) framework aimed at estimating these four parameters using indoor and outdoor air temperature and relative humidity as its only sensor inputs, complemented by readily available contextual information and individual activity profiles. It exploits the climatic coupling of globe temperature and air velocity to the indoor–outdoor environment and the temperature- and activity-driven behavioral patterns that govern clothing insulation and metabolic rate. The proposed framework achieved strong predictive performance, explaining 85% of the variance in actual PMV values (R^2 = 0.85), with a near-zero mean residual (−0.041) and a residual standard deviation of 0.286. Approximately 91% of absolute errors fell below 0.5 PMV units—a deviation unlikely to shift the assigned thermal comfort category. Mapped to thermal comfort categories, predictions reached 80% accuracy, with a macro-averaged precision of 0.81 and recall of 0.80, exhibiting the highest performance for neutral and warm conditions while performing less accurately for cool discomfort. These results suggest that standard temperature and humidity sensors, combined with basic contextual information and individual activity profiles, could support reliable PMV-based thermal comfort assessment, advancing scalable, sensor-light comfort monitoring for energy-efficient, sustainable buildings.

Keywords: thermal comfort; predicted mean vote (PMV); machine learning; indoor environmental quality; building energy efficiency

Academic Editor: Tea Zakula

Received: 31 May 2026
Revised: 8 July 2026
Accepted: 14 July 2026
Published: 14 July 2026

1. Introduction

With people spending up to 90% of their time indoors, maintaining comfortable thermal environments is essential for health, well-being, and productivity. Thermal comfort is defined as that condition of mind reflecting satisfaction with the prevailing thermal conditions and is primarily shaped by climatic and physiological factors [1]. Beyond these, additional influences on thermal perception have been identified, including behavioral, cultural and socio-economic factors, as well as demographic characteristics such as gender and age [2–5]. Given these multifaceted influences and the inherent subjectivity of human

 https://doi.org/10.3390/su18147202

perception, thermal comfort assessment poses significant practical challenges. In response, two principal research approaches have emerged: the first emphasizes simplified, practical models that rely primarily on readily measurable climatic factors, inevitably involving assumptions and compromises regarding human factors; the second prioritizes personalization, developing models tailored to individuals' behavioural patterns, traits, and preferences. These two approaches represent a fundamental trade-off: complex, precise, resource-intensive models versus practical, scalable, cost-effective ones.

Thermal discomfort imposes costs that are often overlooked. Even short-term exposure can trigger acute physiological responses such as elevated heart rate and respiration rate [6,7]. Sustained discomfort has also been associated with mental health burdens, including stress, anxiety, and depression [8,9]. In workplace settings, these effects translate into reduced productivity, increased fatigue, and lower job satisfaction [10,11]. Although this cumulative burden is substantial, it remains difficult to quantify in practice, because thermal comfort cannot be measured reliably outside controlled environments.

Beyond these individual and organizational stakes, thermal comfort is also central to the European Union's energy and sustainability agenda [12]. With nearly 50% of total EU energy demand attributable to heating and cooling systems—rising to around 63% in the residential sector—thermal comfort is directly implicated in meeting the EU's sustainability and energy efficiency targets [13,14]. This agenda has two main dimensions: reducing energy waste and operational costs of these systems through intelligent, proactive management; and ensuring equitable access to comfortable indoor environments, thereby addressing energy poverty and supporting a fair energy transition.

The Predicted Mean Vote (PMV) is the most widely adopted thermal comfort index that estimates the mean thermal sensation of a large population exposed to the same indoor conditions [15]. This index incorporates six parameters spanning climatic and physiological factors: air temperature, relative humidity, mean radiant temperature, air velocity, metabolic rate, and clothing insulation. ASHRAE Standard 55, which governs thermal environmental conditions for human occupancy, defines a 7-point thermal sensation scale ranging from -3 to +3 that maps PMV values to thermal comfort categories. Negative values correspond to cool discomfort, positive values to warm discomfort, and 0 to thermal neutrality. The intervals between adjacent categories are considered grey zones; therefore, many studies adopt tolerance margins around each category to account for boundary ambiguity [16]. Although PMV is the foundational index for international standards such as ASHRAE Standard 55, its assumption of uniform thermal exposure is a critical limitation [17,18].

The present study addresses two main challenges in PMV-based thermal comfort assessment. The first challenge lies in measuring globe temperature and air velocity, which require specialized, costly equipment and dense spatial sampling due to their significant spatial variability indoors. The second challenge concerns clothing insulation and metabolic rate—two crucial parameters capturing individuals' physiological responses that must be self-reported in practice, posing a substantial practical barrier. To address these challenges, this study proposes an ML framework that jointly estimates the four hard-to-measure PMV parameters using readily available climatic inputs of indoor and outdoor air temperature and relative humidity as its only sensor inputs, complemented by contextual information and individual activity profiles. The primary contribution of this study lies in this integrated estimation under a reduced-sensing regime rather than the learning algorithms themselves.

The remainder of this study is organized as follows. Section 2 reviews related studies that applied ML to PMV estimation. Section 3 introduces the dataset and details the ML models developed to estimate the four hard-to-measure PMV parameters: globe temperature, clothing insulation, metabolic rate, and air velocity. Section 4 evaluates the predictive performance

https://doi.org/10.3390/su18147202

of the proposed framework. Section 5 discusses the results, addresses the limitations of this study, and outlines future research directions.

2. Related Works

ML-based PMV prediction from reduced feature sets has been investigated in several studies that address the practical difficulty of measuring or self-reporting some of the index's parameters, with air temperature consistently identified as a dominant predictor. Rahmanparast et al. [19] identified air temperature, mean radiant temperature, and metabolic rate as the most influential PMV predictors, and reported a strong linear correlation between air temperature and mean radiant temperature. Among multiple ML regression models trained on air temperature, metabolic rate, and clothing insulation as inputs, ensemble methods outperformed other models, with XGBoost (MAE = 0.125, R^2 = 0.93) and Gradient Boosting (MAE = 0.130, R^2 = 0.93) both reaching 95% accuracy under a $\pm$0.5 PMV unit tolerance. Similarly, Park and Woo [20] identified air temperature, globe temperature, clothing insulation, and metabolic rate as the strongest PMV predictors. Using globe temperature, metabolic rate, and clothing insulation as input features, the authors compared Random Forest, Artificial Neural Network (ANN), and Long Short-Term Memory (LSTM) models, with mean absolute percentage error (MAPE) between 5.45% and 6.72%.

Another line of work relies on the most readily measurable climatic parameters for PMV estimation, enabling broader practical use. Buratti et al. [21] proposed a simplified linear model for cost-effective, non-invasive, rapid evaluation of HVAC performance under near-neutral conditions, using air temperature, relative humidity, and the derived water vapor pressure as input features. For three predefined clothing insulation levels, their model achieved mean prediction errors close to zero with a standard deviation of $\pm$0.22. Palladino et al. [22] addressed several hard-to-measure PMV parameters using an ANN with a reduced feature set of air temperature, relative humidity, and clothing insulation. The remaining parameters were assumed fixed: a uniform mean radiant temperature, negligible air velocity, and a constant metabolic rate. Using field measurements from 18 residential settings in Italy, the proposed ANN achieved MAE < 0.3 for 83% of residences, rising to 93% when air temperature fell between 21 °C and 28 °C and relative humidity between 35% and 75%.

Atthajariyakul and Leephakpreeda [23] developed an ANN that estimated PMV using six climatic inputs—air temperature, globe temperature, wet-bulb temperature, relative humidity, mean radiant temperature, and air velocity—while assuming predefined values for clothing insulation and metabolic rate, achieving MSE = 0.11. A different approach was taken by Guenther and Sawodny [24], who proposed a Gaussian Process (GP) model in place of PMV, using engineering signals rather than thermal comfort parameters to address the uncertainty associated with self-reporting clothing insulation and metabolic rate. Using indoor and outdoor air temperatures, global solar radiation, and HVAC operation and fan status as inputs, the model achieved RMSE = 0.68—a 74% improvement in prediction error over the baseline PMV calculation. Dyvia and Arif [25] performed a sensitivity analysis identifying air temperature and mean radiant temperature as the most influential PMV predictors, with metabolic rate and clothing insulation showing only weak to moderate influence outside high-intensity indoor activities. Then, an ANN was developed for PMV estimation, achieving RMSE = 0.73 and R^2 = 0.66 with only the most influential climatic predictors as inputs, improving to RMSE = 0.14 and R^2 = 0.99 when all PMV parameters were included. Boutahri and Tilioua [26] examined the linear correlations between PMV and climatic factors, with air temperature and mean radiant temperature both showing moderate-to-strong linear correlation ($r \approx 0.65$). They also evaluated ensemble models for

PMV prediction, with XGBoost achieving MAE = MSE = 0.01 and R^2 = 0.97, and Random Forest achieving MAE = 0.02, MSE = 0.005, and R^2 = 0.97. Across these studies, ML-based PMV estimation has been pursued either by retaining hard-to-measure parameters as inputs, replacing them with non-PMV signals, or assuming fixed values—leaving the underlying measurement challenge unresolved.

Three observations emerge from this review. First, ML can effectively estimate PMV from reduced feature sets, provided that at least some hard-to-measure parameters are available as inputs [19–23,25,26]. Second, air temperature is consistently identified as a dominant PMV predictor, with mean radiant or globe temperature also contributing strongly when available [19,20,25,26]. Third, no prior study has estimated all four hard-to-measure PMV parameters—mean radiant temperature, air velocity, metabolic rate, and clothing insulation—from only the most readily available climatic inputs. Existing approaches either retain some hard-to-measure parameters as inputs, fix them at assumed values, or replace PMV entirely with alternative comfort models [19–26].

The present study addresses this gap by proposing an ML framework that jointly estimates the four hard-to-measure PMV parameters using indoor and outdoor air temperature and relative humidity as its only sensor inputs, complemented by contextual information and individual activity profiles. This reduced-sensing approach enables PMV-based thermal comfort assessment in settings equipped with only standard temperature and humidity sensors.

3. Materials and Methods

3.1. Dataset Description

The open-access ASHRAE DB II serves as the primary dataset for this study, which comprises more than 100,000 recordings from 70 thermal comfort field studies conducted worldwide between 1982 and 2021 [27]. This dataset encompasses indoor and outdoor climatic conditions, participants' metabolic rate and clothing insulation, subjective thermal sensation and comfort responses, PMV and Predicted Percentage Dissatisfied (PPD) indices, and participant demographics, as well as region, climate zone, season, cooling and heating strategy, and building characteristics. Nevertheless, a notable limitation is the considerable number of incomplete recordings due to differences in research protocols and objectives and limited access to specialized sensing equipment across studies.

3.2. Data Cleaning

The ASHRAE DB II dataset comprises n = 109,033 records and d = 53 features, of which $d_n = 41$ are numerical and $d_c = 12$ are categorical. Following a data cleaning procedure, the original feature space was transformed into a lower-dimensional space, retaining the most informative and non-redundant features for the subsequent modeling. Let $\mathbf{X}$ represent the original feature space, partitioned into a numerical subspace $\mathbf{X_n}$ and a categorical subspace $\mathbf{X_c}$, where $\mathbb{D}$ denotes the set of admissible categorical values in $\mathbf{X_c}$. Likewise, let $\mathbf{X}'$ represent the reduced feature space, partitioned analogously into a numerical $\mathbf{X}'_\mathbf{n}$ and a categorical $\mathbf{X}'_\mathbf{c}$ subspace, retaining the same sample size n and $d' < d$ features. The structure of the two subspaces is defined in Equations (1) and (2), where ${d_n}' < d_n$ and ${d_c}' < d_c$ denote the number of numerical and categorical features in $\mathbf{X}'$.

$$\mathbf{X} = \mathbf{X_n} \times \mathbf{X_c}, \mathbf{X_n} \in \mathbb{R}^{n\times 41}, \mathbf{X_c} \in \mathbb{D}^{n\times 12} \tag{1}$$

$$\mathbf{X}' = \mathbf{X}'_\mathbf{n} \times \mathbf{X}'_\mathbf{c},\ \mathbf{X}'_\mathbf{n} \in \mathbb{R}^{n\times d'_n},\ \mathbf{X}'_\mathbf{c} \in \mathbb{D}^{n\times d'_c} \tag{2}$$

First, non-predictive features were removed from $\mathbf{X}$, including research metadata, timestamps, and recording and building identifiers. The subject identifier was preserved to support subject-aware partitioning, with the individual participant serving as the unit of partitioning rather than the individual record. All observations associated with a given participant were assigned exclusively to a single partition, ensuring that no participant's repeated measurements were shared between the training and test sets, thereby precluding subject-level information leakage. Next, incomplete features—specifically those exceeding 75% missing values—were removed from $\mathbf{X}$, including structural characteristics of the occupied spaces (doors, windows, blinds) and specific HVAC equipment indicators (heater, fan). Subsequently, redundant features capturing repeated measurements of the same quantity were also removed from $\mathbf{X}$, retaining only a single representative variable in each case: height-specific recordings of air, globe, and mean radiant temperatures, and air velocity (at 0.1, 0.6, and 1.1 m above the floor) were replaced by their corresponding occupied-zone averages, whereas time-averaged metabolic rates over the previous 10, 20, 30, and 60 min were omitted in favor of the current metabolic rate. Finally, participants' subjective responses for thermal sensation, preference, and thermal and air-movement acceptability were excluded from $\mathbf{X}$, as modeling subjective thermal responses was beyond the scope of this study.

Following data cleaning, the reduced feature space $\mathbf{X}'$ comprised d′ = 13 features, including $d_n{}' = 11$ numerical and $d_c{}' = 2$ categorical features (Table 1), thereby forming the $\mathcal{D}_0$ dataset. This dataset preserved all n = 109,033 recordings from the original ASHRAE DB II, with data completeness handled separately at the sub-model level.

Table 1. The features included in the reduced space $\mathbf{X}'$.

Feature	Type	Space	Symbol	Units
Indoor air temperature	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{ta}	°C
Indoor relative humidity	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{rh}	%
Indoor globe temperature	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{tg}	°C
Indoor air velocity	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{av}	m/s
Metabolic rate	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{mr}	met
Clothing insulation	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{clo}	clo
Outdoor air temperature (Current)	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{taoc}	°C
Outdoor air temperature (Monthly mean)	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{taom}	°C
Outdoor relative humidity (Current)	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{rhoc}	%
Season	Categorical	$\mathbf{X}'_\mathbf{c}$	x_{sn}	-
Building cooling mode	Categorical	$\mathbf{X}'_\mathbf{c}$	x_{cm}	-
PMV	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{pmv}	-
Subject ID	Numerical	$\mathbf{X}'_\mathbf{n}$	x_{sid}	-

From the $\mathcal{D}_0$ dataset, the subset of records with complete information for the PMV input parameters and all sub-model predictors was extracted to enable ground-truth comparison during the framework's evaluation. This subset was partitioned using an 85/15 split, yielding a held-out test set $\mathcal{D}_{test}$ ($|\mathcal{D}_{test}| = 3197$), with the $\mathcal{D}_{test}$ set being mutually exclusive from the sub-model training and testing. The split was stratified on the two categorical features of the reduced space, season and building cooling mode, to preserve their marginal proportions across partitions. Subject disjointness was maintained throughout, such that no participant was represented in both the held-out set and any sub-model training or test data.

Numerical transformations were applied to the two categorical features retained in the reduced space $\mathbf{X}'$. Building cooling mode, comprising air-conditioned, naturally ventilated, and mixed-mode categories, was one-hot encoded. Season was cyclically encoded to capture its inherent periodicity, ensuring that calendar-adjacent seasons remain close in the feature space. Each season was assigned to an integer index ($x_{sn} = 0$ for winter, $x_{sn} = 1$ for spring, $x_{sn} = 2$ for summer, and $x_{sn} = 3$ for autumn), which was then mapped onto the unit circle, as described in Equations (3) and (4).

$$x_{sn,s} = \sin\left(\frac{2\pi}{4} \times x_{sn}\right) \tag{3}$$

$$x_{sn,c} = \cos\left(\frac{2\pi}{4} \times x_{sn}\right) \tag{4}$$

3.3. Estimation of Globe Temperature Using OLS Linear Regression

Globe temperature represents the equilibrium response of a globe sensor to simultaneous convective and radiative heat exchange with its surroundings, integrating the effects of air temperature, mean radiant temperature, and air velocity into a single thermal metric. Because air temperature directly influences the convective component of this equilibrium, a strong correlation between air and globe temperature naturally emerges. This relationship was exploited in the present study to estimate globe temperature from the more readily available air temperature. However, the strength of this relationship is not uniform, as the relative contributions of convective and radiative heat transfer vary with both season and building cooling mode. Accordingly, the air–globe temperature relationship was examined separately for each season–building cooling mode combination.

Within $\mathcal{D}_0$, a total of 22,686 measurements with complete records for air temperature, globe temperature, season, and building cooling mode were included, forming the dataset $\mathcal{D}_1$. Recordings corresponding to autumn and spring seasons were excluded, as they accounted for less than 5% of the $\mathcal{D}_1$ dataset, yielding $|\mathcal{D}_1| = 22,108$. For each season–building cooling mode combination $g = (x_{sn}, x_{cm})$, a stratified subset $\mathcal{D}_1^{(g)}$ was created, and subsequently symmetrically trimmed at its own 0.02 and 99.98 percentiles to mitigate the influence of extreme globe temperature outliers. Let ${\beta_0}^{(g)}$ and ${\beta_1}^{(g)}$ denote the intercept and the slope estimated for subset $\mathcal{D}_1^{(g)}$, respectively. The Ordinary Least Squares (OLS) model regressing globe temperature x_{tg} on air temperature x_{ta} is described by Equation (5).

$$\hat{x}_{tg} = {\beta_0}^{(g)} + {\beta_1}^{(g)} \times x_{ta} \tag{5}$$

For each $\mathcal{D}_1^{(g)}$ subset, the OLS assumptions for linearity, homoscedasticity, and normality of residuals were assessed through residual diagnostics. As presented in Figure 1, residuals were randomly scattered around zero with no curvilinear trends in any $\mathcal{D}_1^{(g)}$ subset, supporting the linearity assumption. Breusch–Pagan test results indicated pervasive heteroscedasticity across $\mathcal{D}_1^{(g)}$ subsets $(p < 0.01)$, with non-constant residual variance bypassed only in the case of the summer naturally ventilated subset $(p = 0.64)$. The direction of this heteroscedasticity varied by subset: residual spread widened toward higher fitted temperatures in the winter air-conditioned and winter naturally ventilated subsets, whereas the summer air-conditioned subset showed the opposite pattern, with spread narrowing as fitted globe temperature increased. The mixed-mode subsets exhibited only mild, though statistically detectable, variance changes. To ensure valid inference despite these violations, heteroscedasticity-consistent (HC3) standard errors were used in the subsequent OLS analysis.

Quantile–quantile (Q–Q) plots of the residuals are presented in Figure 2. Across all $\mathcal{D}_1^{(g)}$ subsets, observed and theoretical quantiles were closely aligned through the central range (approximately ± 2 standard deviations), with departures confined to the tails. These

https://doi.org/10.3390/su18147202

were most pronounced in the summer air-conditioned and winter naturally ventilated subsets, reflecting both heavy tails and asymmetry, consistent with the elevated kurtosis (up to 12.83) and non-negligible skewness (up to 1.50) reported in Table 2. The summer mixed-mode subset was closest to normality, with the remaining subsets showing intermediate departures. Although Jarque–Bera tests rejected normality in every subset, this reflects the test's high sensitivity at large sample sizes, where even minor departures from normality reach statistical significance, rather than a substantive violation. Crucially, the validity of the reported models does not rest on residual normality. OLS coefficient estimates remain unbiased under non-normal errors, and the models are employed for prediction rather than small-sample hypothesis testing, so the efficiency and exact-inference arguments that motivate normality do not apply. To verify that the tail behavior did not materially affect the fitted relationships, response transformations (log and Box–Cox) were additionally examined for the most affected subsets; these did not meaningfully alter the estimated slope β_1 or improve predictive error, while degrading the direct interpretability of the slope as the change in globe temperature per unit change in air temperature, and were therefore not adopted.

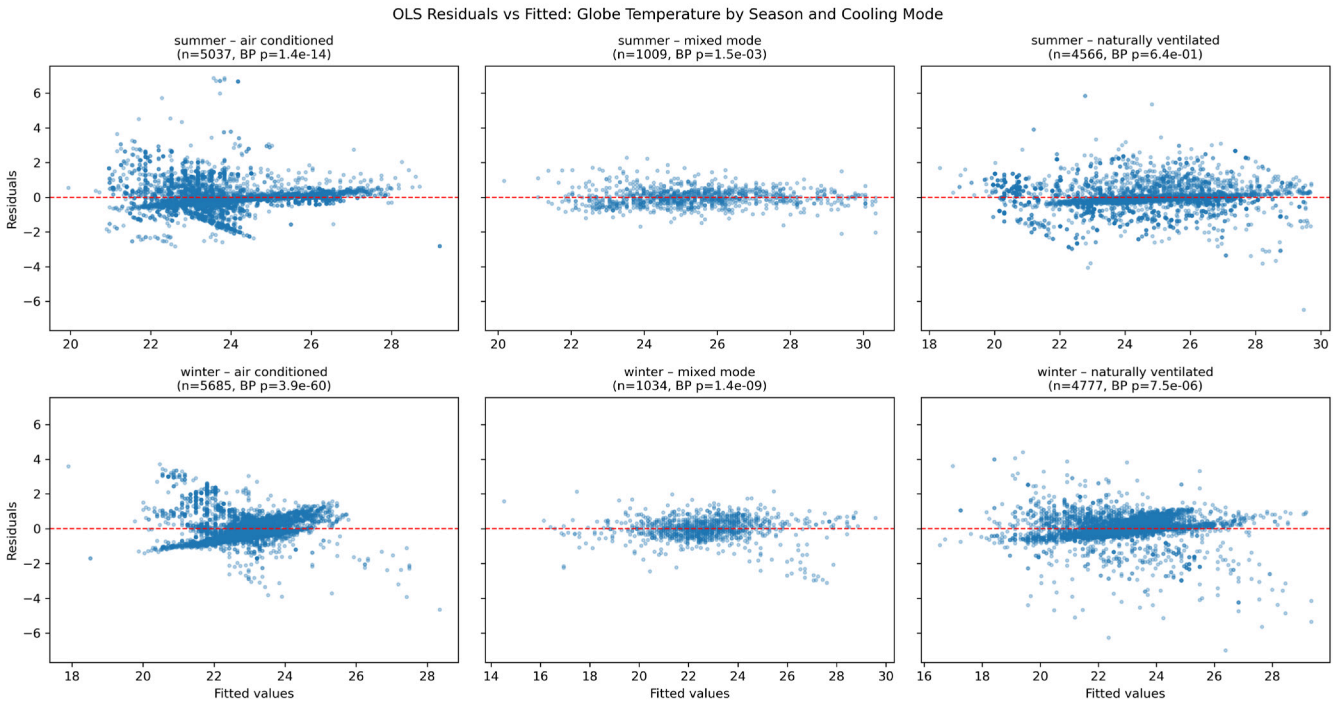

Figure 1. Residuals plots for fitted globe temperatures during winter and summer months.

The six fitted OLS models for globe temperature estimation are presented in Figure 3. Across all subsets, the estimated slopes lay close to unity (0.78–0.93), indicating an approximately one-to-one relationship between air and globe temperature. The relationship was steepest in summer mixed-mode spaces, where globe temperature increased by about 0.93 °C per 1 °C rise in air temperature, and shallowest in winter air-conditioned spaces, at about 0.78 °C. Slopes were consistently lowest in the air-conditioned subsets and higher in the mixed-mode and naturally ventilated ones, consistent with the partial decoupling of radiant from air temperature under active conditioning. The explanatory power of the models followed the same pattern: air temperature explained the most variance in globe temperature in the mixed-mode and naturally ventilated subsets (R^2 up to 0.93) and the least in the air-conditioned subsets ($R^2 = 0.71$ in summer and $R^2 = 0.61$ in winter). This lower fit reflects the influence of conditioned surfaces on mean radiant temperature, which weakens the air–globe relationship the models exploit.

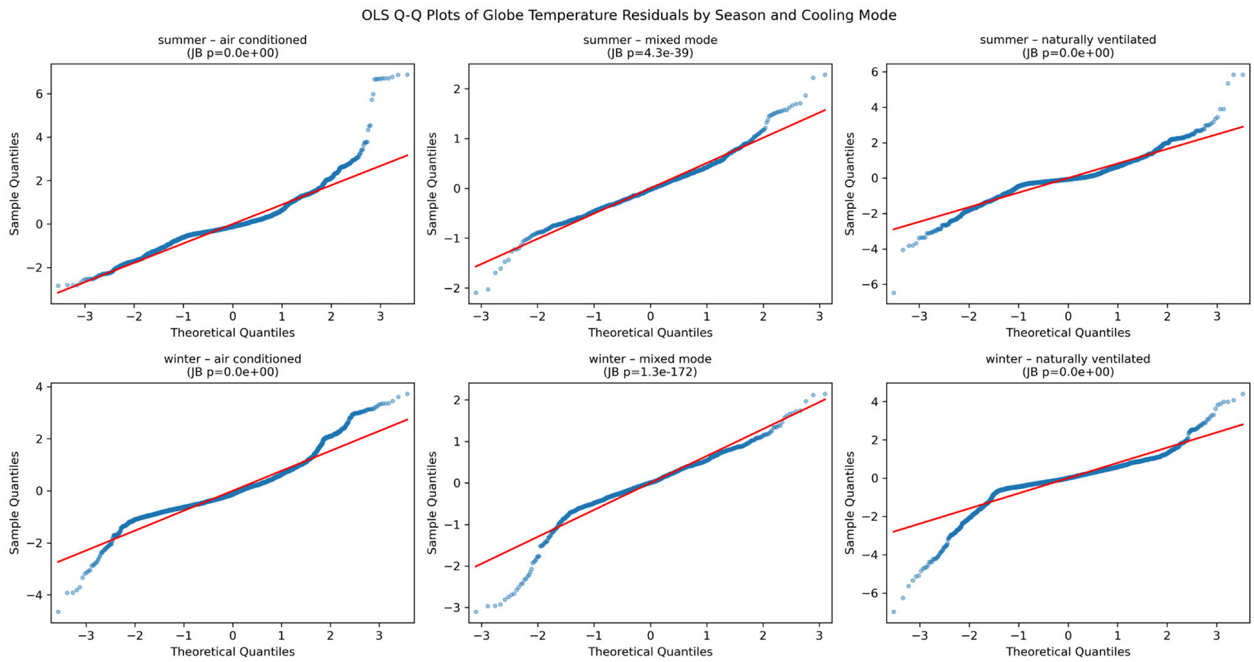

Figure 2. Q-Q plots for globe temperature residuals for summer and winter months.

Table 2. Fitted OLS globe-temperature models and residual diagnostics by season and cooling mode.

Season	Cooling Mode	n	β_0	β_1	SE(β_1)	R^2	BP p *	Skew	Kurtosis
Summer	Air-conditioned	5037	3.57	0.857	0.007	0.708	<0.001	1.50	11.24
Summer	Mixed-mode	1009	2.34	0.933	0.009	0.931	0.002	0.44	4.85
Summer	Naturally ventilated	4566	2.03	0.922	0.006	0.868	0.64	0.11	7.70
Winter	Air-conditioned	5685	5.42	0.780	0.014	0.608	<0.001	0.86	6.38
Winter	Mixed-mode	1034	2.88	0.896	0.013	0.905	<0.001	−1.01	6.78
Winter	Naturally ventilated	4777	2.56	0.896	0.010	0.806	<0.001	−1.42	12.83

* BP p denotes the *p*-value of the Breusch-Pagan test for heteroskedasticity.

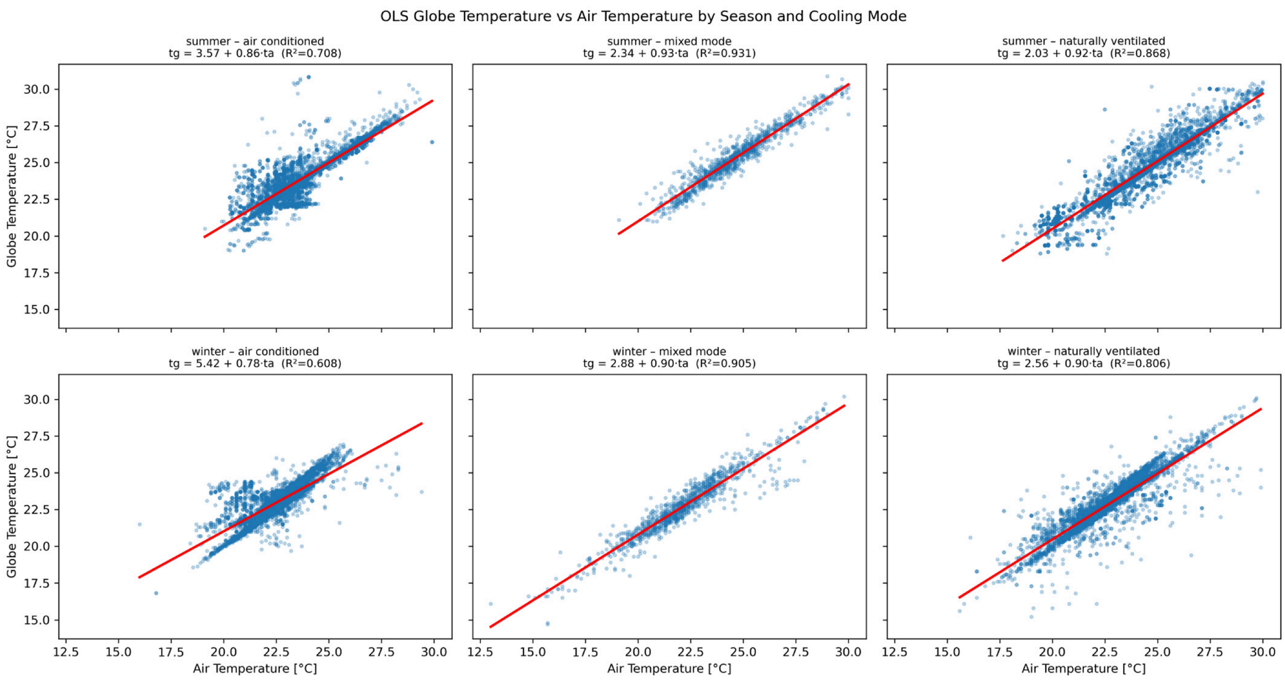

Figure 3. OLS linear regression models for globe temperature estimation during summer and winter months.

3.4. Estimation of Clothing Insulation Using Random Forest Regressor

Clothing insulation acts as a thermal barrier between the body and its environment, regulating heat transfer and supporting the body's thermal equilibrium. Although it is a critical parameter for thermal comfort assessment, it must be self-reported in practice, limiting the practical applicability of PMV-based models. A non-intrusive estimation of clothing insulation is therefore particularly valuable for integration into dynamic thermal management systems—motivating the approach proposed in this section.

Pearson's and Spearman's coefficients were computed between clothing insulation and candidate climatic predictors to assess linear and monotonic dependencies (Figure 4), using complete records from $\mathcal{D}_0$ and excluding the held-out $\mathcal{D}_{\text{test}}$. The linear correlations were moderate to strong and negative across all temperature-related variables, suggesting systematic thermoregulatory adjustment in clothing behavior: insulation levels decreased as thermal conditions became warmer and increased under cooler conditions. Indoor and outdoor temperature variables exhibited linear dependencies of comparable magnitude, with indoor air and globe temperatures both yielding $r = -0.70$, followed by current outdoor air temperature ($r = -0.67$) and mean monthly outdoor air temperature ($r = -0.66$). Spearman's correlations further confirmed strong negative monotonic relationships between clothing insulation and all temperature variables ($|\rho| > 0.71$), demonstrating that clothing insulation decreased consistently with increasing temperature irrespective of the precise functional form of the association. Indoor relative humidity showed weak associations with clothing insulation, yielding $r = -0.19$ and $\rho = -0.22$.

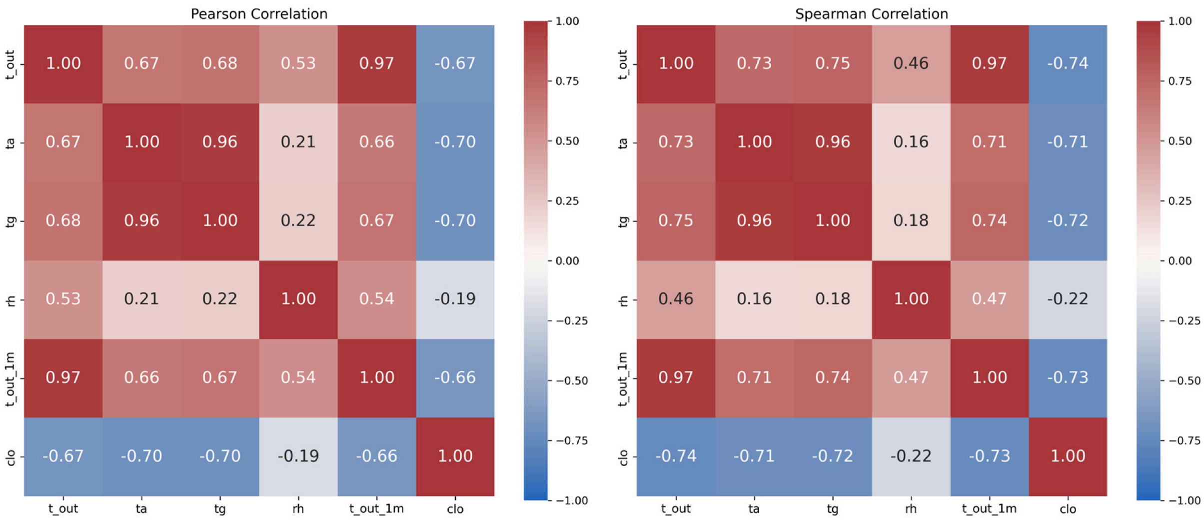

Figure 4. Pearson and Spearman correlation heatmaps between clothing insulation and climatic predictors.

Mutual Information (MI) was computed between clothing insulation and the candidate predictors to investigate non-linear dependencies. Current and mean monthly outdoor air temperatures yielded the highest values (MI $\approx$ 0.73). Indoor air and globe temperatures yielded MI = 0.51 and MI = 0.50, respectively, suggesting similar predictive value for clothing insulation estimation. Indoor relative humidity, by contrast, showed the weakest dependency (MI = 0.26), consistent with the weak associations identified earlier.

Given the coexistence of linear and non-linear dependencies identified above, a Random Forest (RF) regression model was adopted as a flexible mapping function $\mathcal{G}(\cdot)$ to estimate clothing insulation $\hat{x}_{\text{clo}}$, as described in Equation (6). The proposed model takes as input a vector $\mathbf{x}_{\text{RF}}$ comprising outdoor and indoor temperatures, indoor relative humidity, and cyclical seasonal components, formally defined in Equation (7). Although the clothing model is trained on measured globe temperature x_{tg}, at inference it receives the estimated globe temperature $\hat{x}_{\text{tg}}$ from the OLS regression model (Section 3.3), so that any error in the globe estimate propagates into the clothing insulation prediction.

$$\hat{x}_{\text{clo}} = \mathcal{G}(\mathbf{x}_{\text{RF}}) \tag{6}$$

$$\mathbf{x}_{\text{RF}} = \left[x_{\text{taoc}}, x_{\text{taom}}, x_{\text{ta}}, x_{\text{tg}}, x_{\text{rh}}, x_{\text{sn,s}}, x_{\text{sn,c}}\right]^{\text{T}} \tag{7}$$

A Tree-structured Parzen Estimator (TPE) was used to tune the hyperparameters of the RF model, with 5-fold cross-validation and R^2 as the objective function. The search space $\mathcal{H}$ was defined over integer-valued ranges for the number of estimators, maximum tree depth, minimum samples per leaf, and minimum samples to split, as specified in Equation (8). Within $\mathbf{X}'$, 5321 samples were identified with complete records for all predictors in $\mathbf{x}_{\text{RF}}$ and the clothing insulation x_{clo}, constituting $\mathcal{D}_2$ dataset. This dataset was partitioned into training and test subsets using an 85/15 ratio, reserving 799 samples as a held-out test set for final model evaluation. The optimal hyperparameters of $z_{\text{est}} = 256$, $z_{\text{depth}} = 20$, $z_{\text{leaf}} = 8$, and $z_{\text{split}} = 9$ yielded $R^2 = 0.71$, MAE = 0.108, and RMSE = 0.14 on the held-out test set. Given that clothing insulation values typically range between 0.4 and 1.2 clo, the MAE of 0.108 clo corresponds to an estimation error of approximately 12% of this range, indicating practically useful predictive accuracy.

$$\mathcal{H} = \left\{z_{\text{est}} \in [40, 500], z_{\text{depth}} \in [2, 22], z_{\text{leaf}} \in [1, 10], z_{\text{split}} \in [2, 10]\right\} \tag{8}$$

To evaluate the generalization performance of the RF model, regression-fit and residual diagnostics were examined on the held-out $\mathcal{D}_{\text{test}}$, as illustrated in Figure 5. Predictive accuracy was highest for mid-range clothing insulation values, reflecting their closer correspondence to commonly observed garment ensembles in built environments. Toward the extremes, however, the model exhibited a mild regression-to-the-mean tendency: it slightly overestimated low insulation values (≈0.2–0.4 clo) and increasingly underestimated higher values, with the largest deviations near the upper end of the retained range (≈1.2 clo). This compression toward central values is consistent with the averaging behavior of ensemble tree models and contributes to the overall $R^2 = 0.71$, as the residual variance concentrates at the extremes. The residual plot reflected this pattern, with predominantly positive residuals at low predicted values and a gradual shift toward negative residuals at higher predicted values; residuals were otherwise centered near zero, mostly within ±0.2 clo, with no pronounced funnel-shaped heteroscedasticity.

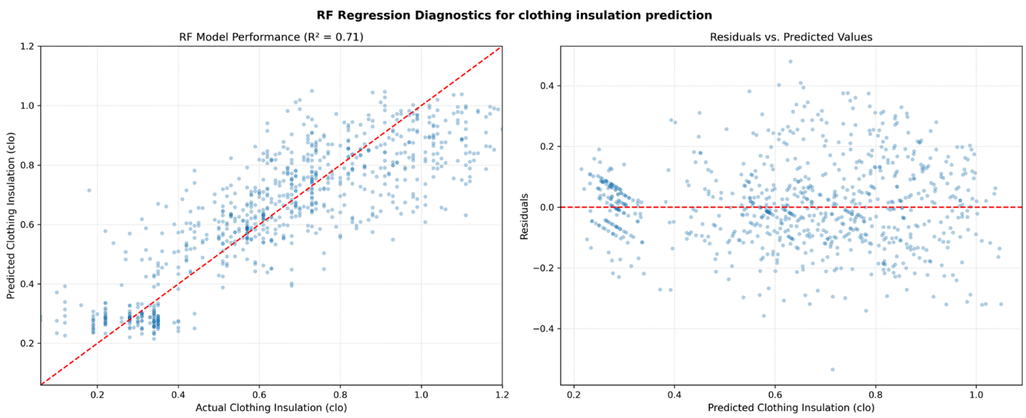

Figure 5. Regression fit and residual plots of the RF regression model for clothing insulation prediction.

3.5. Estimation of Metabolic Rate Using k-Means Clustering

Metabolic rate quantifies the rate at which the body converts stored energy into heat and mechanical work, governing internal heat production and serving as a critical determinant of an individual's thermal perception. Unlike globe temperature and clothing insulation, metabolic rate is an intrinsic physiological process governed by physical activity, body composition, health status, and individual lifestyle—largely independent of climatic conditions and therefore not directly estimable from them. In the absence of per-occupant metabolic sensing, which is rarely instrumented in operational indoor environments, metabolic rate can instead be approximated at the individual level by assigning a representative rate derived from each occupant's historical activity profile. To this end, k-means clustering was used to group individuals into metabolic rate profiles, assigning a representative rate to each cluster based on historical observations.

Within $\mathbf{X}'$, 38,635 samples were identified with complete metabolic rate records spanning j = 1197 individuals (approximately 32 samples per individual) forming the $\mathcal{D}_3$ dataset. To ensure that $\mathcal{D}_3$ reflected typical indoor activity patterns, upper trimming at the 98th percentile (met ≤ 1.88) was applied to exclude atypically high metabolic rates, preserving sedentary and moderate activity levels while targeting extreme high-intensity outliers. Per-individual feature vectors were then computed; individuals with a single observation, for which a standard deviation could not be defined, were excluded, leaving 1090 individuals for clustering. The elbow method was applied over cluster counts ranging from 2 to 10, identifying k = 2 as the optimal number of clusters based on within-cluster inertia.

Let $\mathbf{m}_j \in \mathbb{R}^5$ denote the feature vector corresponding to the j-th individual, containing the mean, minimum, maximum, median, and standard deviation of metabolic rate observations in $\mathcal{D}_3$. Let $c_j \in \{0, 1\}$ denote the assigned cluster label for the j-th individual and $\boldsymbol{\mu}_k \in \mathbb{R}^5$ the centroid of cluster k. Then, the assignment of each individual to a cluster is expressed as in Equation (9).

$$c_j = \underset{k \in \{0,1\}}{\operatorname{argmin}} \left\| \mathbf{m}_j - \boldsymbol{\mu}_k \right\|^2 \tag{9}$$

As illustrated in Figure 6, k-means clustering yielded two distinct metabolic rate profiles. Principal Component Analysis (PCA), applied for visualization, separated the clusters along the first principal component, with Cluster 0 (c_0) concentrated in the negative PC1 region and Cluster 1 (c_1) in the positive PC1 region, though the two distributions remained contiguous near the boundary. Cluster c_0 was characterized by metabolic rates between approximately 1.0 and 1.2 met with limited intra-cluster variance, representing light to sedentary indoor activities. In contrast, Cluster 1 (c_1) exhibited higher metabolic rates between 1.2 and 1.6 met with substantially greater intra-cluster variability, reflecting a broader spectrum of moderate to higher-intensity indoor activities. Specifically, Cluster 0 (c_0) showed a mean metabolic rate of 1.17 ± 0.05 met and Cluster 1 (c_1) a mean of 1.26 ± 0.07 met, the larger standard deviation of c_1 reflecting its broader range of activity levels.

To quantify the error introduced in PMV assessment by representing each individual's metabolic rate with a static cluster-level value rather than the actual observed measurement, PMV was computed under both conditions across the leak-free $\mathcal{D}_3$ records (n = 30,404). All other PMV inputs were held fixed so that the difference isolated the effect of substituting metabolic rate with a representative cluster-level value. Replacing the true per-sample metabolic rate with the assigned representative produced a near-zero systematic bias $(\Delta\text{PMV} = +0.05)$ and a mean absolute deviation of 0.28 PMV units, with the 95th percentile of $|\Delta\text{PMV}|$ at 0.86 and an RMSE of 0.44. At the level of the comfort verdict, the thermal-acceptability classification $(|\text{PMV}| \leq 0.5)$ was preserved in 79.4% of observations. These results indicate that the substitution introduces a bounded error with negligible systematic bias, under which the comfort classification remains consistent with that obtained from the observed metabolic rate in the majority of cases. The residual discrepancy is attributable to the pronounced within-individual variability of metabolic rate, which a single static profile cannot capture. This accuracy cost, together with the activity range over which the proxy remains applicable, is discussed further in Section 5.

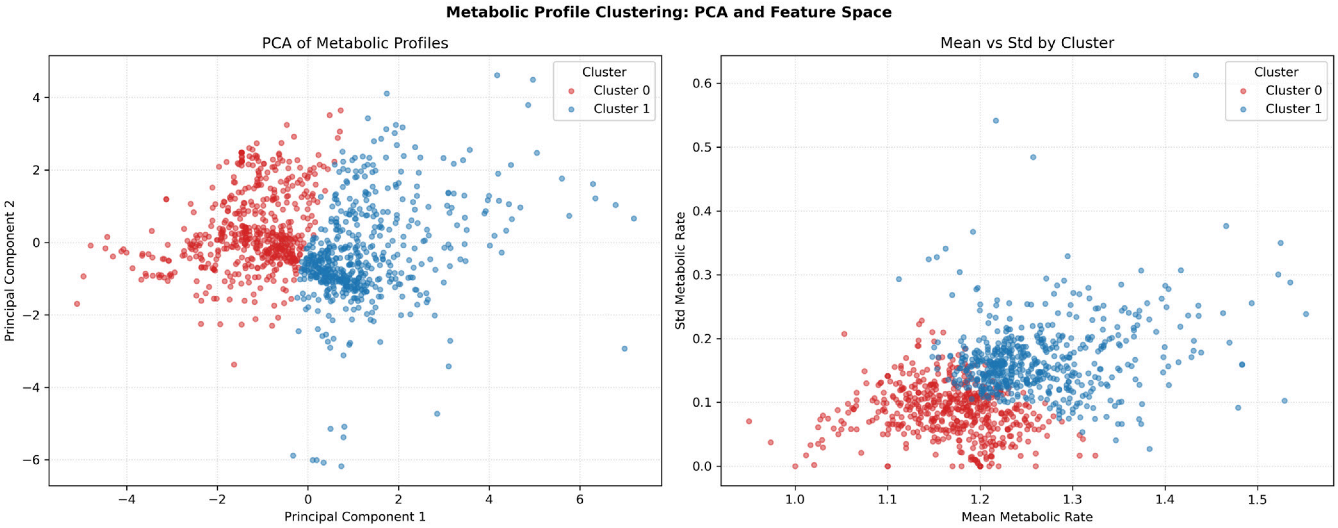

Figure 6. k-means clustering of metabolic rate profiles visualized in PCA space for k = 2 clusters.

3.6. Estimation of Air Velocity Integrating Heuristic Features into k-Means

Indoor air velocity governs convective and evaporative heat transfer between the occupant and the surrounding air, influencing perceived draft and cooling sensations and serving as a critical PMV parameter in thermal comfort assessment. Nevertheless, measuring indoor air velocity reliably presents three key challenges. First, it requires specialized, costly devices such as omnidirectional anemometers, which often lack the sensitivity needed for indoor low-velocity conditions. Second, sensor placement must be carefully chosen to avoid measurement artifacts from localized airflow patterns caused by ventilation systems and occupant movement. Third, substantial temporal variability in indoor airflow makes it difficult to obtain representative long-term values from point measurements.

Given the fundamental constraints of establishing direct relationships between indoor air velocity and the available climatic inputs, two heuristic indices were introduced to characterize indoor airflow dynamics: the Normalized Ambient Mixing (NAM) index and the Seasonal Comfort Deficit (SCD) index. Both are grounded in the physical mechanisms that drive indoor air movement.

NAM quantifies the combined driving force for thermal and hygrometric exchange between the indoor and outdoor environments. It is computed as the product of the absolute indoor–outdoor differences in air temperature and relative humidity, with the humidity difference expressed as a decimal fraction, as given in Equation (10). Physically, the indoor–outdoor temperature and humidity gradients represent the potentials that drive buoyancy- and infiltration-related air exchange across the building envelope: larger gradients correspond to stronger driving forces for mixing, and hence to greater expected indoor air movement.

$$NAM = \frac{|x_{\text{ta}} - x_{\text{taoc}}| \times |x_{\text{rh}} - x_{\text{rhoc}}|}{100} \tag{10}$$

SCD quantifies the deviation of indoor conditions from season-specific benchmarks representing neutral thermal sensation. These benchmarks were derived from the subset of

ASHRAE DB II recordings associated with neutral thermal sensation votes: for each season, $T_{comf}^{(s)}$ and $H_{comf}^{(s)}$ were defined as the median indoor air temperature and relative humidity across those recordings. SCD was computed as the sum of the absolute deviations of indoor air temperature and relative humidity from their respective benchmarks, each normalized by its seasonal Median Absolute Deviation (MAD) to account for the differing scales of the two variables, as given in Equation (11). Physically, SCD measures how far indoor conditions depart from thermal neutrality; larger deviations correspond to the cool or warm conditions under which occupants are most likely to invoke air-movement responses, such as fan use or window opening, linking the index to the airflow regimes it is intended to capture.

$$SCD^{(s)} = \frac{\left|x_{\mathrm{ta}} - T_{\mathrm{comf}}^{(s)}\right|}{MAD_T^{(s)}} + \frac{\left|x_{\mathrm{rh}} - H_{\mathrm{comf}}^{(s)}\right|}{MAD_H^{(s)}} \quad (11)$$

Within $\mathbf{X}'$, 13,309 samples with complete records for indoor air velocity x_{av} and the base climatic predictors (indoor and outdoor air temperature, indoor and outdoor relative humidity) were identified, forming the $\mathcal{D}_4$ dataset. Globe temperature and the NAM and $SCD^{(s)}$ indices were then computed and appended to each sample. To reduce the influence of abnormally high air-velocity values, atypical of indoor environments where low air velocities predominate, $\mathcal{D}_4$ dataset was trimmed above the 98th percentile, retaining values within [0.0, 0.90] m/s. This percentile threshold was chosen to remove the sparse upper tail of velocities atypical of conditioned indoor environments—values more consistent with transient drafts or near-diffuser measurement positions than with occupied-zone airflow—while retaining the full range of velocities relevant to occupant comfort. The candidate predictor vector $\mathbf{x}_{\mathrm{v}}$ was defined in Equation (12).

$$\mathbf{x}_{\mathrm{v}} = \left[x_{\mathrm{taoc}}, x_{\mathrm{ta}}, x_{\mathrm{tg}}, x_{\mathrm{rh}}, x_{\mathrm{rhoc}}, NAM, SCD^{(s)}\right]^{\mathrm{T}} \quad (12)$$

K-means clustering was subsequently applied to segment indoor air velocity into distinct airflow profiles. Clustering quality was evaluated using the Silhouette score S [28] and Cohen's d [29]. The Silhouette score quantifies intra-cluster cohesion and inter-cluster separation, with values approaching +1 indicating compact, well-separated clusters and $S \approx 0.5$ representing a reasonable structure. Cohen's d quantifies the standardized mean difference in air velocity between clusters, with $d \geq 0.8$ indicating a large effect size and $d \geq 0.5$ a moderate one.

Table 3 summarizes the evaluation metrics for the four best-performing feature combinations. C1 achieved the highest Silhouette score (S = 0.63), indicating the most cohesive and well-separated clusters, together with a moderate-to-large effect size (d = 0.73) and a parsimonious three-cluster structure. Using the same features but four clusters, C3 matched C1's effect size (d = 0.73) but yielded a lower Silhouette score (S = 0.47), indicating a less parsimonious and less cohesive solution. C4 attained the largest effect size (d = 1.01) but the lowest Silhouette score (S = 0.41), reflecting strong velocity separation between only two poorly delineated clusters; C2 was inferior on both criteria (S = 0.53, d = 0.64). Balancing cluster cohesion, separation, and physical interpretability, the combination of indoor air temperature and $SCD^{(s)}$ with three clusters (C1) was selected as the input for the air-velocity clustering model. The three-cluster solution retained a large effect size while achieving the highest Silhouette score, avoiding both the poorly delineated two-cluster split and the less cohesive four-cluster partition. The strong clustering performance of this combination—the highest cohesion and separation among all candidates—provides empirical support for the physical relevance of the SCD index to indoor airflow variation. NAM-based combinations performed less well, suggesting that deviation from

https://doi.org/10.3390/su18147202

seasonal neutrality (SCD) captures airflow-relevant structure more effectively than the indoor–outdoor gradient magnitude (NAM) in this dataset.

Table 3. Evaluation of the four best-performing feature combinations in air velocity clustering.

Combination	Clusters	Features	Silhouette Score	Cohen's d
C1	3	$x_{ta}, SCD^{(s)}$	0.63	0.73
C2	3	$NAM, SCD^{(s)}$	0.53	0.64
C3	4	$x_{ta}, SCD^{(s)}$	0.47	0.73
C4	2	x_{ta}, NAM	0.41	1.01

Figure 7 illustrates the clustering structure obtained from C1, revealing three distinct airflow regimes in the feature space of x_{ta} and $SCD^{(s)}$. Cluster c_0 spanned lower indoor air temperatures (≈11–18 °C) with large deviations from seasonal comfort benchmarks, reflecting cool conditions; air velocities were lowest in this regime, with a median of approximately 0.03 m/s. Cluster c_1 occupied the central temperature region (≈20–26 °C) and showed the smallest deviations from comfort benchmarks, with a median air velocity of approximately 0.10 m/s, characteristic of thermally neutral indoor conditions. Cluster c_2 encompassed higher indoor air temperatures (above ≈27 °C) with larger comfort deviations and the highest median air velocity (≈0.21 m/s), consistent with warm environments where active air movement, such as fan use, is more prevalent. Air velocity thus increased monotonically with indoor temperature across the three regimes, consistent with greater reliance on air movement for cooling as conditions warm.

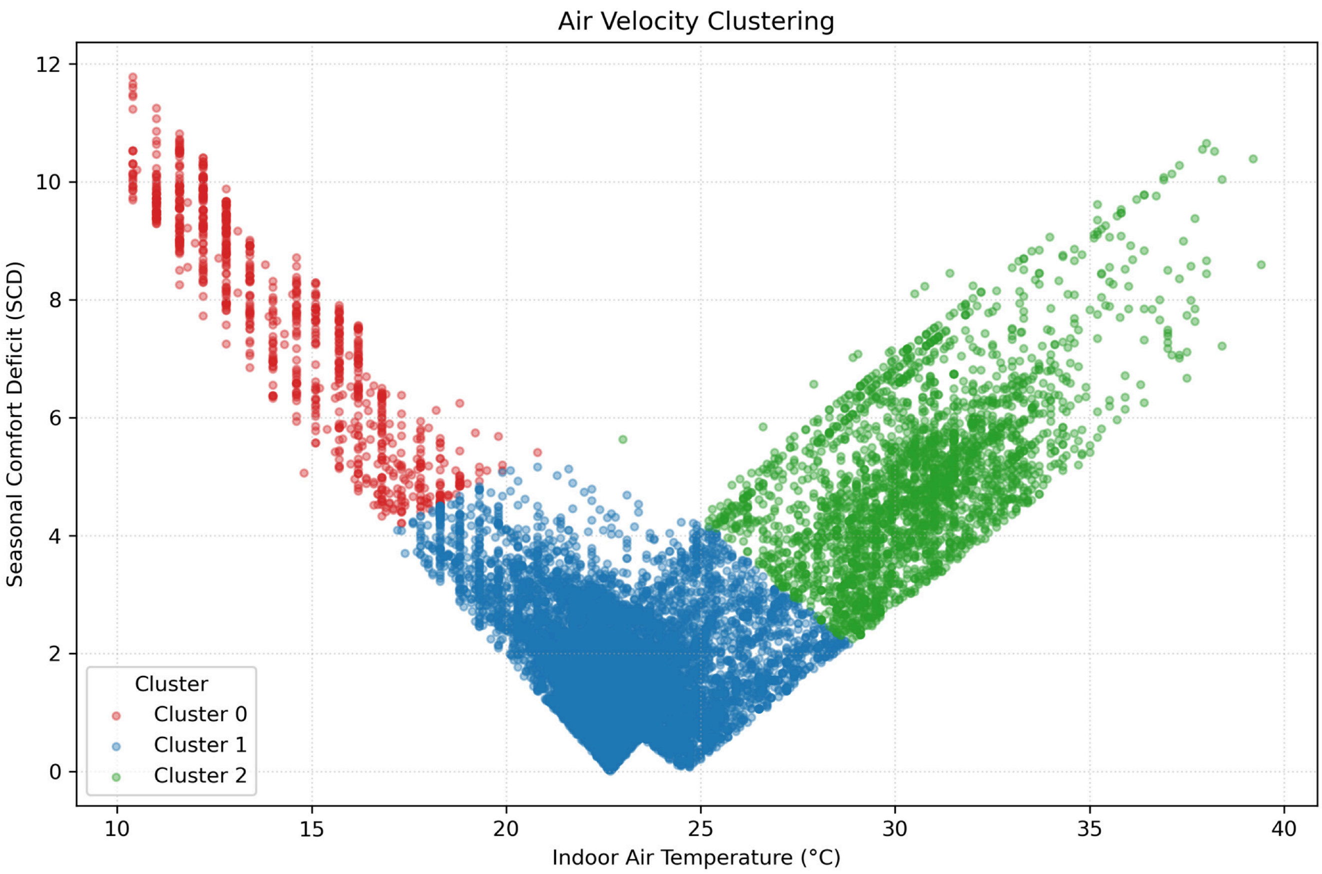

Figure 7. Clusters resulted from combination C1 across indoor air temperature and $SCD^{(s)}$ index.

The composition of the three airflow regimes was further examined across season and building cooling mode. The cool, low-velocity cluster (c_0) was populated predominantly

https://doi.org/10.3390/su18147202

by winter and air-conditioned or mixed-mode records, consistent with heated indoor spaces maintained below seasonal neutrality where occupants have little incentive to invoke air-movement responses. The neutral cluster (c_1) drew across all seasons and cooling modes, reflecting its central position around thermal neutrality, which is the operating target common to conditioned and naturally ventilated spaces alike. The warm, high-velocity cluster (c_2) was dominated by summer records and showed the strongest representation of naturally ventilated and mixed-mode spaces, consistent with the greater reliance on window opening and fan use for cooling where mechanical conditioning is absent or supplementary. This distribution indicates that the recovered regimes are not artifacts of a single season or operating mode but instead align with the expected seasonal and operational drivers of indoor air movement, supporting the physical interpretability of the clustering.

To quantify the error introduced in PMV assessment by representing air velocity with a static cluster-level value rather than the observed measurement, PMV was computed under both conditions across the leak-free $\mathcal{D}_4$ records (n = 12,996). All other PMV inputs were held fixed so that the difference isolated the effect of substituting air velocity with its cluster-representative value; PMV was evaluated without input-range restriction to retain the full thermal range present in the data. Replacing the true per-sample air velocity with the assigned cluster median produced a small systematic bias (ΔPMV = +0.077) and a mean absolute deviation of 0.102 PMV units, with a 95th percentile of $|\Delta\text{PMV}|$ of 0.43 and an RMSE of 0.197. At the level of the comfort verdict, the thermal-acceptability classification ($|\text{PMV}| \leq 0.5$) was preserved in 93.4% of observations. The error was modest and bounded across all three airflow regimes, with mean absolute deviations of 0.108, 0.099, and 0.112 PMV units for the cool, neutral, and warm clusters, respectively; the substitution was effectively unbiased in the warm regime (ΔPMV = +0.006), where air velocities were highest, and showed a small positive bias ($\approx$ +0.10) in the cooler, low-velocity regimes. These results indicate that representing air velocity by a regime-level value introduces limited and well-bounded error into the comfort assessment, with the comfort classification preserved in the large majority of cases. The approach is therefore best understood as a conditional prior for settings without dedicated anemometry, rather than a substitute for real-time sensing where such instrumentation is available. This trade-off is discussed further in Section 5. In particular, the reliability of the regime-level estimate in naturally ventilated and spatially heterogeneous spaces—where point air velocity is least stationary—is examined among the study's limitations.

3.7. Evaluation Metrics for Framework's Predictive Performance

The predictive performance of the proposed framework was evaluated using two regression metrics: (i) Mean Bias Error (MBE), which captures systematic bias by measuring the average signed prediction error, where positive values indicate underestimation and negative values overestimation; and (ii) the coefficient of determination (R^2), which quantifies the proportion of variance in the actual PMV values explained by the predictions. Let $y_{PMV,i}$ and $\hat{y}_{PMV,i}$ denote the actual and estimated PMV values for the i-th sample, respectively, then Equations (13) and (14) formulate the error metrics.

$$MBE = \frac{1}{n}\sum_{i=1}^{n}(y_{PMV,i} - \hat{y}_{PMV,i}) \quad (13)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_{PMV,i} - \hat{y}_{PMV,i})^2}{\sum_{i=1}^{n}(y_{PMV,i} - \overline{y}_{PMV})^2} \quad (14)$$

4. Results

The proposed framework was evaluated on the $\mathcal{D}_{\text{test}}$ dataset, as defined in Section 3.1. Within this integrated framework, the four sub-models presented in Sections 3.3–3.6 operate in sequence: the OLS model first estimates globe temperature; the estimated values are then supplied to both the clothing-insulation model and the air-velocity model, alongside the sensed climatic inputs; metabolic rate is assigned separately based on each occupant's prior activity profile and does not depend on the climatic estimates. PMV values were computed using the open source pythermalcomfort (version 4.0.2) Python package [30], specifically the pmv_ppd_ashrae function, in accordance with ASHRAE Standard 55.

4.1. Evaluation Performance in PMV Prediction

The distributional characteristics of PMV residuals were examined to assess the predictive performance of the proposed framework. The mean residual was -0.04, indicating negligible prediction bias and corresponding to a relative prediction error of approximately 0.69% across the PMV scale. The negative direction of the residual mean suggested a slight tendency toward predicting warmer thermal sensations. The median residual of -0.04 further indicated mild left-skewness in the residual distribution, reflecting occasional but more pronounced PMV overestimations.

The framework's predictive robustness was further confirmed through the residual dispersion statistics. The residual standard deviation was 0.286, corresponding to approximately 4.8% of the PMV scale range. The interquartile range was limited to 0.39, indicating a tight concentration of residuals around zero. In absolute terms, the maximum underestimation reached 1.29 PMV units, whereas the maximum overestimation reached 1.45 PMV units.

Figure 8 presents the relationship between actual and predicted PMV values. For actual PMV values within the neutral thermal comfort range (-0.5 to $+0.5$), predictions were tightly concentrated around the line of perfect agreement, without pronounced over- or underestimation. Toward the cool-discomfort range, predictive performance degraded, with a wider spread around the reference line and a systematic tendency to overestimate actual PMV, as the most negative values were drawn toward thermal neutrality. A milder compression appeared at the warm extreme, where the most positive PMV values were slightly underestimated. This regression-toward-the-mean behavior, consistent with the negative mean residual reported above, did not prevent strong overall agreement, with the framework achieving $R^2 = 0.85$.

4.2. Evaluation Performance with Tolerance in PMV Predictions

A tolerance margin (T) is commonly applied when evaluating PMV predictions, since small PMV differences do not typically produce perceptible changes in thermal sensation. Although PMV is defined on a continuous scale, its interpretation is inherently categorical, with adjacent thermal sensation categories separated by approximately one PMV unit. A tolerance of $T = \pm 0.5$ PMV units is therefore generally regarded as acceptable, as deviations of this magnitude are unlikely to alter the assigned thermal sensation category. This tolerance provides a meaningful benchmark for evaluating practical predictive performance.

Figure 9 presents the cumulative accuracy curve across increasing tolerance margins. The curve rises steeply, surpassing 80% accuracy at a tolerance of approximately 0.4 and reaching about 91% at the ± 0.5 benchmark, indicating close agreement between actual and predicted PMV values. Beyond a tolerance of roughly 0.6, the curve approaches a plateau near unity, reflecting diminishing gains from further relaxation of the acceptance criterion.

https://doi.org/10.3390/su18147202

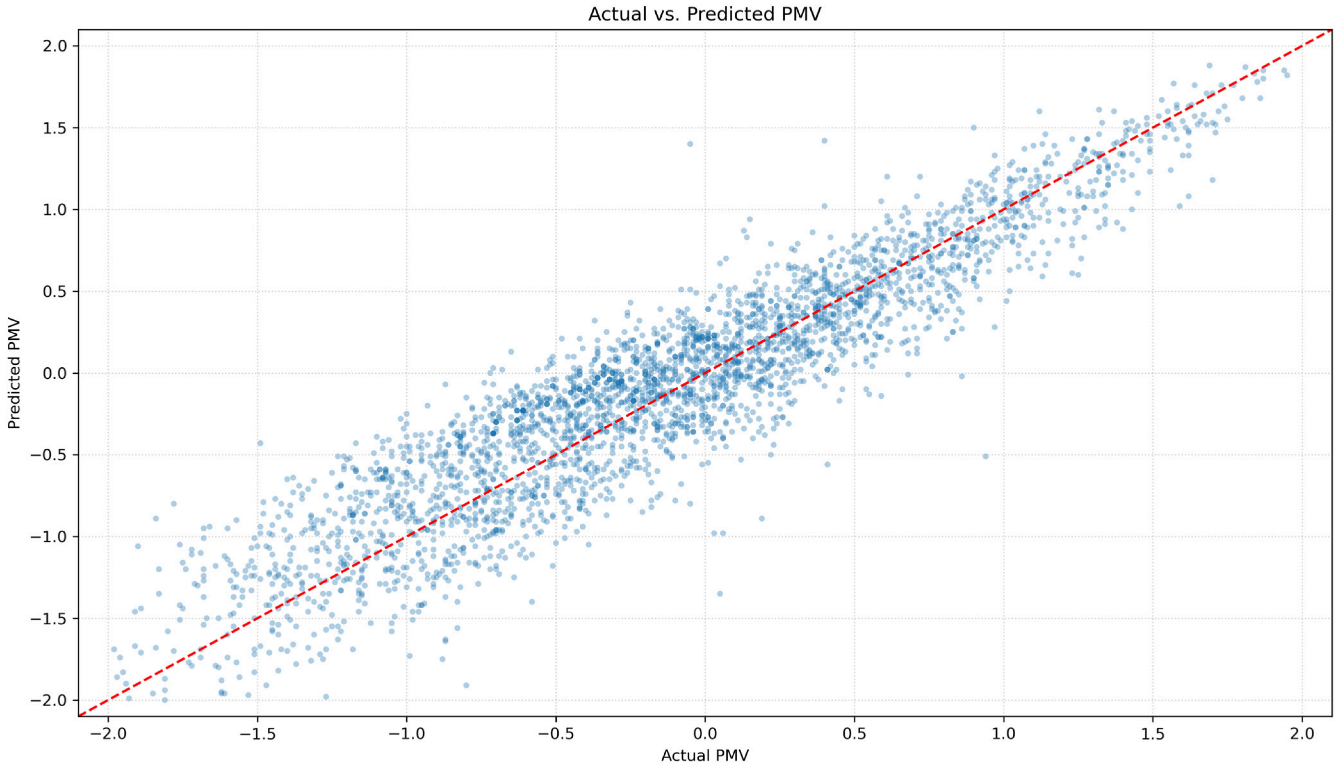

Figure 8. Actual versus predicted PMV with the diagonal representing perfect agreement.

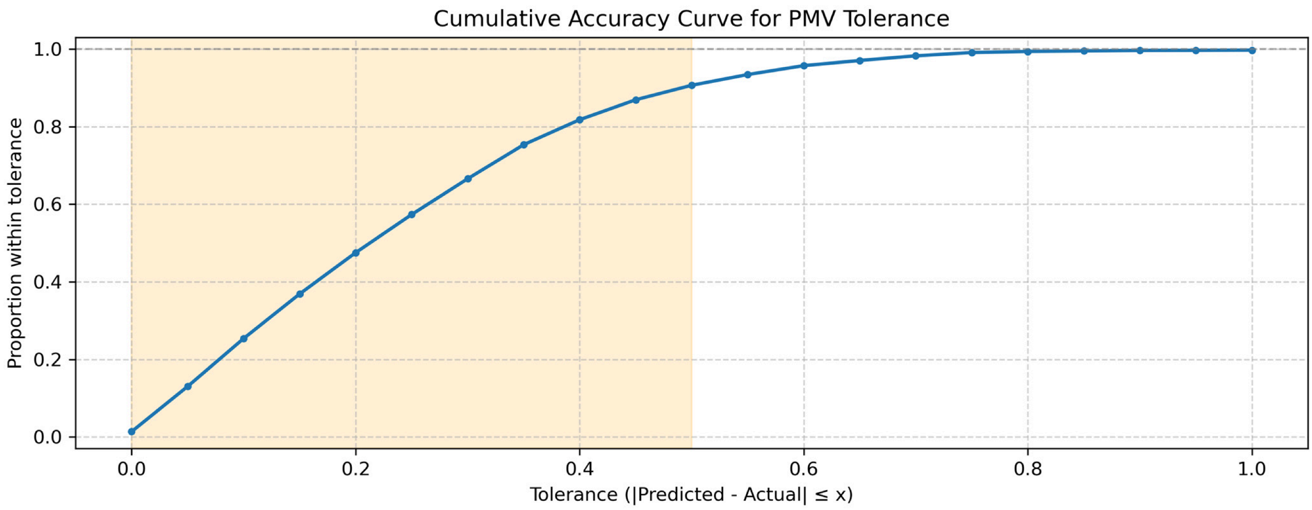

Figure 9. Cumulative accuracy curve for PMV predictions across different tolerance margins T.

Figure 10 illustrates the framework's predictive performance under a tolerance margin of T = ±0.5 PMV units. The shaded region represents the tolerance band; blue points denote predictions within the band and red points those outside it. The majority of predictions fall within the band, with the out-of-band points concentrated mainly at the cool end, where predictions tended to exceed the observed PMV—consistent with the slight overestimation and negative mean residual reported in Section 4.1. Within the band, predictions remained closely aligned with the line of perfect agreement, with only the mild regression-toward-the-mean tendency noted earlier. Together, these results support the practical relevance of the ±0.5 benchmark and indicate strong generalization performance on the held-out set.

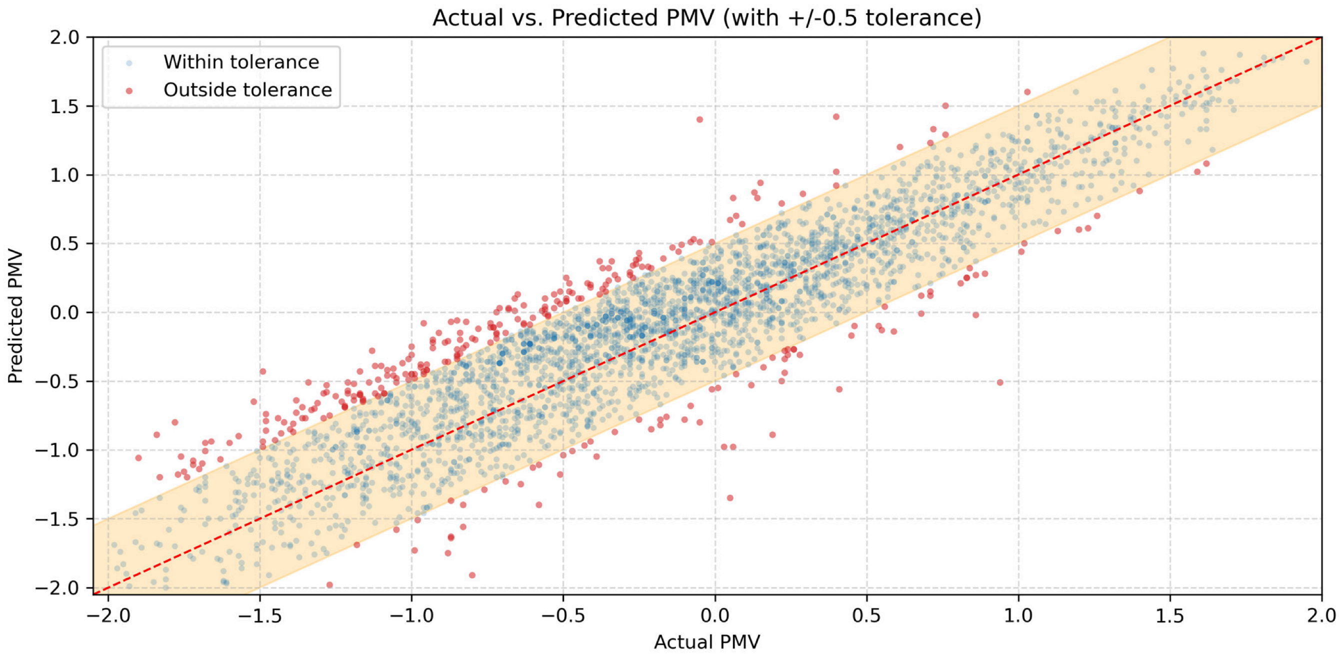

Figure 10. Scatter plot with actual vs. predicted PMV under ±0.5 units tolerance.

4.3. Predictive Performance for PMV Classification into Thermal Comfort Categories

To assess the framework's ability to distinguish thermal comfort categories, PMV values were grouped into three classes: neutral ($-0.5 < \text{PMV} < +0.5$), warm discomfort ($\text{PMV} \geq +0.5$), and cool discomfort ($\text{PMV} \leq -0.5$). The resulting classification performance is summarized in Table 4.

Table 4. Classification report for the framework's predictive performance in PMV estimation.

Thermal State	Precision	Recall	F1-Score	Support
Neutral	0.76	0.85	0.80	1502
Warm discomfort	0.81	0.84	0.82	597
Cool discomfort	0.88	0.71	0.78	1079

The framework achieved an overall classification accuracy of 80%, with a macro-averaged precision of 0.81 and recall of 0.80. The corresponding per-class recall values were 85% for neutral, 84% for warm discomfort, and 71% for cool discomfort. The neutral and warm discomfort classes achieved the highest F1-scores (0.80 and 0.82, respectively), reflecting consistent performance across these categories. The cool discomfort class showed a lower F1-score (0.78), driven by its lower recall (0.71), indicating that a non-negligible share of cool-discomfort cases were missed. Conversely, cool discomfort attained the highest precision (0.88), meaning that cases predicted as cool discomfort were rarely false positives; this combination of high precision and lower recall indicates that the framework under-detects cool discomfort but is reliable when it does assign that class. The strong recall and F1-score for the neutral class underscore the framework's effectiveness in identifying the most frequently observed thermal category. Lastly, the most common errors occurred between adjacent categories, with 29.2% of cool-discomfort and 15.4% of warm-discomfort instances misclassified as neutral. This pattern reflects a tendency to draw mild discomfort conditions toward thermal neutrality, consistent with the regression-toward-the-mean behavior noted in Section 4.1.

https://doi.org/10.3390/su18147202

5. Conclusions

The present study introduced a predictive framework comprising four ML models designed to estimate the four hard-to-measure PMV parameters, addressing a major practical barrier in thermal comfort assessment. To enable practical and scalable PMV prediction, the framework relied on the indoor and outdoor air temperature and relative humidity as its only sensor inputs, supplemented by limited contextual information and individual activity profiles. The principal contributions and findings are as follows.

Globe temperature: An OLS linear regression model was developed to estimate indoor globe temperature from indoor air temperature, stratified by season and building cooling mode to account for the varying contributions of convective and radiative heat transfer. Across the six season–cooling-mode subsets, the models achieved coefficients of determination ranging from R^2 = 0.61 to R^2 = 0.93, with near-unity slopes (0.78–0.93) indicating an approximately one-to-one air–globe temperature relationship; explanatory power was highest in mixed-mode and naturally ventilated spaces and lowest under active air-conditioning, where conditioned surfaces partially decouple radiant from air temperature. This approach eliminates the need for specialized globe thermometers, offering a cost-effective alternative for thermal comfort studies.

Clothing insulation: A Random Forest regression model was developed using indoor and outdoor air temperature, indoor relative humidity, and seasonal information, informed by both linear and non-linear dependency analyses. The model achieved R^2 = 0.71 (MAE = 0.108 clo) on the held-out test set, with the strongest performance across the typical indoor insulation range up to 1.2 clo. This approach removes the need for self-reported clothing data without requiring additional sensing equipment.

Metabolic rate: A k-means clustering method was used, reflecting the stochastic and individually driven nature of metabolic rate. The optimal segmentation yielded two clusters: the first representing sedentary to light indoor activities (mean 1.17 $\pm$ 0.05 met) and the second moderate to higher-intensity activities (mean 1.26 $\pm$ 0.07 met). In practice, a short acclimatization period or prior activity record is sufficient to assign an individual to a metabolic profile, streamlining thermal comfort evaluation.

Air velocity: A k-means clustering approach was used for indoor air velocity estimation. Two heuristic indices—Normalized Ambient Mixing (NAM) and Seasonal Comfort Deficit $\left(SCD^{(s)}\right)$—were introduced to capture indoor–outdoor climatic differences and deviations from neutral thermal comfort conditions. Indoor air temperature combined with $SCD^{(s)}$ yielded the strongest segmentation, producing three clusters with median air velocities of 0.03, 0.10, and 0.21 m/s that increased monotonically with indoor temperature. This approach addresses the widespread challenge of neglecting air velocity in field studies due to its stochastic behavior and high spatial variability.

Integrated across the four estimated parameters, the framework achieved a mean residual of -0.041 and R^2 = 0.85 in PMV prediction, with the small negative residual indicating a slight tendency toward predicting warmer sensations rather than systematic bias. Cumulative accuracy analysis showed that approximately 91% of predictions fell within an absolute error of $\pm$0.5 PMV units—a deviation unlikely to alter the assigned thermal comfort category—with accuracy surpassing 80% at a tolerance of about 0.4.

When evaluated as a classification framework across neutral, warm-discomfort, and cool-discomfort categories, the proposed approach achieved an overall accuracy of 80.0%, with a macro-averaged precision of 0.81 and recall of 0.80. Performance was strongest for the neutral and warm-discomfort categories (F1-scores of 0.80 and 0.82), while the cool-discomfort category showed lower recall (0.71) but the highest precision (0.88), indicating that cool discomfort was under-detected yet reliably identified when assigned. Most errors occurred between adjacent categories, reflecting a tendency to draw mild discomfort conditions toward thermal neutrality—consistent with the regression-toward-the-mean behavior identified in Section 4.1. Confusion between opposing categories was virtually absent, with no cool-discomfort instances misclassified as warm and only 0.2% of warm instances misclassified as cool.

The framework's performance can be situated among recent ML-based PMV estimation studies, although direct comparison is constrained by differences in datasets, metrics, and input requirements. The highest reported accuracies in the literature were achieved by models that retain hard-to-measure parameters as inputs: Rahmanparast et al. [19] reported R^2 = 0.93 using air temperature together with metabolic rate and clothing insulation, while Boutahri and Tilioua [26] and Dyvia and Arif [25] reported R^2 values approaching 0.97–0.99 when mean radiant temperature or the full PMV parameter set was supplied. These results, however, address a fundamentally easier problem, since the parameters that are most difficult to obtain in practice are provided rather than estimated. Methods restricted, like the present framework, to readily available inputs instead either fix the hard-to-measure parameters at assumed constants or replace PMV with surrogate signals. The most directly comparable case is Palladino et al. [22], who used air temperature, relative humidity, and clothing insulation while holding mean radiant temperature, air velocity, and metabolic rate fixed, achieving a mean absolute error below 0.3 PMV units for 83% of cases; the present framework attains comparable agreement (approximately 91% of predictions within $\pm$0.5 PMV units) while estimating all four hard-to-measure parameters rather than assuming them. The lower accuracy relative to full-input models is therefore an expected consequence of the more constrained sensing scenario targeted in this study, and is offset by the framework's substantially lower deployment cost and its applicability in settings equipped only with standard temperature and humidity sensors.

Beyond aggregate performance metrics, it is necessary to delineate the conditions under which reliable prediction can be expected, as well as the regimes in which accuracy is likely to deteriorate. Performance was strongest under near-neutral and warm conditions in mixed-mode and naturally ventilated spaces, where the air–globe temperature coupling is tight (R^2 up to 0.93) and where occupant adaptive behaviors—clothing adjustment and air-movement responses—follow the temperature-driven patterns the sub-models exploit. The framework is expected to degrade in three regimes. First, under active air-conditioning, conditioned surfaces partially decouple mean radiant from air temperature, weakening the globe-temperature estimate (R^2 = 0.61–0.71). Second, in cool-discomfort conditions, the compounded regression-toward-the-mean of the clothing, air-velocity, and globe sub-models draws predictions toward neutrality, producing the reduced cool-discomfort recall (0.71) reported in Section 4.3. Third, in spaces with high-intensity or rapidly varying activity, or with strong non-thermal drivers of clothing choice, the static metabolic and behaviorally-derived clothing estimates are least representative. The framework therefore targets the common case of light-to-moderate activity in conventionally operated buildings, rather than edge cases such as industrial settings, or highly transient occupancy.

The proposed framework occupies a middle ground between two established approaches to thermal comfort assessment. Relative to conventional PMV measurement, it replaces direct measurements of globe temperature, air velocity, clothing insulation, and metabolic rate with estimates derived from readily available environmental and contextual information. It should therefore be viewed as a reduced-sensing approximation to conventional PMV assessment, intended for applications where comprehensive instrumentation is impractical and categorical comfort classification is sufficient. Relative to occupant-centric or personalized comfort models, the framework prioritizes scalability over individual adaptation. Rather than learning occupant-specific preferences from feedback, it assigns representative behavioral parameters and predicts PMV-based thermal comfort, acknowledging that subjective thermal sensation may deviate from PMV because of adaptation, expectations, and demographic factors. These approaches are complementary rather than competing: the proposed framework could provide a low-cost baseline that is subsequently refined through occupant-specific feedback.

Despite its promising performance, this study is subject to several limitations. First, the metabolic rate approach assigns a static profile to each individual and cannot fully capture temporal behavioral variation; an office worker assigned a sedentary profile, for instance, may not be well represented during meetings or cognitively demanding tasks that elevate physiological activity, a substitution that constitutes the weakest link in the pipeline, preserving the thermal-acceptability verdict in 79.4% of cases. Second, clothing selection extends beyond thermal considerations—personal preferences, cultural influences, socio-economic conditions, and professional requirements may all affect insulation levels and should be incorporated in future models. Third, both indoor air velocity and globe temperature exhibit substantial spatial variability, particularly in naturally ventilated environments; while the clustering framework provides representative estimates, it cannot resolve fine-grained spatial airflow distributions or radiant asymmetries, and advanced spatial approaches such as graph-based or physics-informed models represent promising future directions. Fourth, the framework was developed and evaluated exclusively on the ASHRAE DB II dataset. Data partitioning was performed at the subject level and stratified by season and cooling mode, preventing participant-wise leakage while preserving representative operating conditions. Study-disjoint partitioning was not adopted because the constituent field studies differ substantially in data completeness and sensing protocols; withholding entire studies would therefore confound cross-study generalization with differences in dataset composition. Consequently, the reported results characterize generalization to unseen occupants under the operating conditions represented in ASHRAE DB II rather than across independent buildings or climates. External validation on additional datasets is therefore required to establish the framework's broader generalizability.

Overall, these findings demonstrate that reliable PMV estimation can be achieved using only readily available indoor and outdoor temperature and humidity measurements, supplemented by limited contextual information and individual activity profiles, under the operating conditions represented in the ASHRAE DB II dataset. This substantially reduces the sensing requirements, cost, and complexity associated with conventional PMV-based thermal comfort assessment. The proposed framework therefore provides a scalable foundation for real-world thermal comfort monitoring, intelligent HVAC control, and energy-efficient building management using standard sensor infrastructure. However, translating these capabilities into operational benefits will require real-time field validation, as offline predictive accuracy alone does not demonstrate energy savings or closed-loop control performance.

Author Contributions: Conceptualization, C.M. and J.G.; Methodology, C.M., G.P., N.A. and J.G.; Software, C.M. and G.P.; Formal analysis, C.M.; Investigation, C.M., G.P. and N.A.; Resources, N.A. and D.K.; Data curation, C.M. and D.K.; Writing—original draft, C.M. and G.P.; Writing—review & editing, N.A., D.K. and J.G.; Visualization, G.P.; Supervision, J.G.; Project administration, J.G.; Funding acquisition, J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: This study used the publicly available ASHRAE Global Thermal Comfort Database II. No new data were created in this study.

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

References

1. *ANSI/ASHRAE Standard 55-2023*; Thermal Environmental Conditions for Human Occupancy. American Society of Heating, Refrigerating and Air-Conditioning Engineers: Atlanta, GA, USA, 2023.
2. Arsad, F.S.; Hod, R.; Ahmad, N.; Baharom, M.; Ja'afar, M.H. Assessment of Indoor Thermal Comfort Temperature and Related Behavioural Adaptations: A Systematic Review. *Environ. Sci. Pollut. Res.* **2023**, *30*, 73137–73149. [CrossRef] [PubMed]
3. Naheed, S.; Shooshtarian, S. A Review of Cultural Background and Thermal Perceptions in Urban Environments. *Sustainability* **2021**, *13*, 9080. [CrossRef]
4. Gupta, V.; Sarkar, A.; Saxena, A.; Jana, A.; Bandyopadhyay, S. Appliance Utility in Low-Income Housing: On the Relation between Socio-Economic Characteristics, Thermal Comfort and Appliance Usage Pattern. *Energy Build.* **2024**, *323*, 114849. [CrossRef]
5. Mountzouris, C.; Protopsaltis, G.; Gialelis, J. The Impact of Gender and Age on Thermal Comfort. *Procedia Comput. Sci.* **2025**, *257*, 314–320. [CrossRef]
6. Lan, L.; Wargocki, P.; Wyon, D.P.; Lian, Z. Effects of Thermal Discomfort in an Office on Perceived Air Quality, SBS Symptoms, Physiological Responses, and Human Performance. *Indoor Air* **2011**, *21*, 376–390. [CrossRef] [PubMed]
7. Tipton, M.J.; Harper, A.; Paton, J.F.R.; Costello, J.T. The Human Ventilatory Response to Stress: Rate or Depth? *J. Physiol.* **2017**, *595*, 5729–5752. [CrossRef] [PubMed]
8. Lin, C.-Y.; Zhang, Y. Combined Effects of Stress, Depression, and Emotion on Thermal Comfort: A Case Study in Shenzhen. *J. Build. Eng.* **2025**, *103*, 112158. [CrossRef]
9. Rony, M.K.K.; Alamgir, H.M. High Temperatures on Mental Health: Recognizing the Association and the Need for Proactive Strategies—A Perspective. *Health Sci. Rep.* **2023**, *6*, e1729. [CrossRef] [PubMed]
10. Bueno, A.M.; de Paula Xavier, A.A.; Broday, E.E. Evaluating the Connection between Thermal Comfort and Productivity in Buildings: A Systematic Literature Review. *Buildings* **2021**, *11*, 244. [CrossRef]
11. Kawakubo, S.; Sugiuchi, M.; Arata, S. Office Thermal Environment That Maximizes Workers' Thermal Comfort and Productivity. *Build. Environ.* **2023**, *233*, 110092. [CrossRef]
12. European Parliament and Council. Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the Energy Performance of Buildings (Recast). *Off. J. Eur. Union* **2024**. L 2024/1275. Available online: https://eur-lex.europa.eu/eli/dir/2024/1275/oj (accessed on 29 May 2026).
13. Eurostat. Energy Consumption in Households. Available online: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Energy_consumption_in_households (accessed on 25 May 2026).
14. European Commission. *Heating and Cooling*; European Commission: Brussels, Belgium, 2025. Available online: https://energy.ec.europa.eu/topics/energy-efficiency/heating-and-cooling_en (accessed on 25 May 2026).
15. Fanger, P.O. *Thermal Comfort: Analysis and Applications in Environmental Engineering*; Danish Technical Press: Copenhagen, Denmark, 1970.
16. Tran, Q.D.; Dolgun, A.; Demirhan, H. The Impact of Grey Zones on the Accuracy of Agreement Measures for Ordinal Tables. *BMC Med. Res. Methodol.* **2021**, *21*, 70. [CrossRef] [PubMed]
17. Su, X.; Yuan, Y.; Wang, Z.; Liu, W.; Lan, L.; Lian, Z. Human Thermal Comfort in Non-Uniform Thermal Environments: A Review. *Energy Built Environ.* **2024**, *5*, 853–862. [CrossRef]
18. Zhao, Q.; Lian, Z.; Lai, D. Thermal Comfort Models and Their Developments: A Review. *Energy Built Environ.* **2021**, *2*, 21–33. [CrossRef]
19. Rahmanparast, A.; Milani, M.; Camci, M.; Karakoyun, Y.; Dalkilic, A.S. Parameter Reduction for PMV Prediction via Data Driven Approaches Using the ASHRAE Global Thermal Comfort Database II and Chinese Dataset. *Appl. Therm. Eng.* **2025**, *279*, 127553. [CrossRef]
20. Park, K.-Y.; Woo, D.-O. PMV Dimension Reduction Utilizing Feature Selection Method: Comparison Study on Machine Learning Models. *Energies* **2023**, *16*, 2419. [CrossRef]
21. Buratti, C.; Ricciardi, P.; Vergoni, M. HVAC Systems Testing and Check: A Simplified Model to Predict Thermal Comfort Conditions in Moderate Environments. *Appl. Energy* **2013**, *104*, 117–127. [CrossRef]
22. Palladino, D.; Nardi, I.; Buratti, C. Artificial Neural Network for the Thermal Comfort Index Prediction: Development of a New Simplified Algorithm. *Energies* **2020**, *13*, 4500. [CrossRef]
23. Atthajariyakul, S.; Leephakpreeda, T. Neural Computing Thermal Comfort Index for HVAC Systems. *Energy Convers. Manag.* **2005**, *46*, 2553–2565. [CrossRef]
24. Guenther, J.; Sawodny, O. Feature Selection and Gaussian Process Regression for Personalized Thermal Comfort Prediction. *Build. Environ.* **2019**, *148*, 448–458. [CrossRef]
25. Dyvia, H.A.; Arif, C. Analysis of Thermal Comfort with Predicted Mean Vote (PMV) Index Using Artificial Neural Network. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *622*, 012019. [CrossRef]

26. Boutahri, Y.; Tilioua, A. Machine Learning-Based Predictive Model for Thermal Comfort and Energy Optimization in Smart Buildings. *Results Eng.* **2024**, *22*, 102148. [CrossRef]
27. Földváry Ličina, V.; Cheung, T.; Zhang, H.; de Dear, R.; Parkinson, T.; Arens, E.; Chun, C.; Schiavon, S.; Luo, M.; Brager, G.; et al. Development of the ASHRAE Global Thermal Comfort Database II. *Build. Environ.* **2018**, *142*, 502–512. [CrossRef]
28. Rousseeuw, P.J. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [CrossRef]
29. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1988.
30. Tartarini, F.; Schiavon, S. Pythermalcomfort: A Python Package for Thermal Comfort Research. *SoftwareX* **2020**, *12*, 100578. [CrossRef]