Kırklareli Üniversitesi
Mühendislik ve Fen Bilimleri Dergisi
Cilt 11, Sayı 2, 283-303, 2025

Kırklareli University
Journal of Engineering and Science
Volume 11, Issue 2, 283-303, 2025

Research Article DOI: 10.34186/klujes.1804525

Comparative Evaluation of Regression Models for Building Energy Efficiency Assessment Based on Heating and Cooling Load Requirements

Sinan ATICI[1*], Gürkan TUNA[2]

[1]Sechard Information Technologies, Istanbul, Türkiye
[2]Department of Computer Technologies, Vocational School of Technical Sciences, Trakya University, Edirne, Türkiye

Received: 15.10.2025, Accepted: 05.12.2025, Published: 20.12.2025

ABSTRACT

The accurate prediction of heating and cooling loads is a critical prerequisite for designing energy-efficient buildings and reducing their environmental footprint. This study presents a comprehensive comparative analysis of multiple regression models for estimating the energy efficiency of residential buildings based on their architectural parameters. Using the Energy Efficiency dataset, we evaluated the performance of seven distinct modelling approaches: Linear Regression, Decision Tree, Random Forest, Support Vector Regression with a Radial Basis Function kernel, K-Nearest Neighbours, Multi-Layer Perceptron, and Deep Neural Networks. Models were rigorously assessed using Root Mean Square Error, Mean Absolute Error, and the coefficient of determination (R^2). The results demonstrate that non-linear machine learning methods significantly outperform traditional linear models. Specifically, the Random Forest and Support Vector Regression models achieved superior predictive accuracy, with RMSE values as low as 0.46 for heating load and 1.53 for cooling load, and R^2 scores exceeding 0.97. Furthermore, feature importance analysis identified Overall Height and Relative Compactness as the most influential parameters for heating and cooling load predictions, respectively, providing actionable insights for architectural design. This research shows that advanced machine learning models, particularly Random Forest and Support Vector Regression, offer a robust and accurate framework for building energy assessment.

Keywords: Energy efficiency; Load prediction; Machine learning; Regression models; Sustainable building design

Isıtma ve Soğutma Yükü Gereksinimlerine Dayalı Bina Enerji Verimliliği Değerlendirmesi için Regresyon Modellerinin Karşılaştırmalı Değerlendirmesi

ÖZ

Isıtma ve soğutma yüklerinin doğru tahmini, enerji verimli binalar tasarlamak ve çevresel ayak izlerini azaltmak için kritik bir ön koşuldur. Bu çalışma, mimari parametrelerine dayalı olarak konut binalarının enerji verimliliğini tahmin etmek için çoklu regresyon modellerinin kapsamlı bir karşılaştırmalı analizini sunmaktadır. Enerji Verimliliği veri setini kullanarak, yedi farklı modelleme yaklaşımının performansını değerlendirdik: Doğrusal Regresyon, Karar Ağacı, Rastgele Orman, Radyal Taban Fonksiyonu çekirdeğine sahip Destek Vektör Regresyonu, K-En Yakın Komşular, Çok Katmanlı Algılayıcı ve Derin Sinir Ağları. Modeller, Kök Ortalama Karesel Hata, Ortalama Mutlak Hata ve belirleme katsayısı (R^2) kullanılarak titizlikle değerlendirildi. Sonuçlar, doğrusal olmayan makine öğrenmesi yöntemlerinin

**Corresponding Author: Sinan ATICI*
E-mail: asdsinanatici@gmail.com

Research Article DOI: 10.34186/klujes.1804525

geleneksel doğrusal modellerden önemli ölçüde daha iyi performans gösterdiğini göstermektedir. Özellikle, Rastgele Orman ve Destek Vektör Regresyonu modelleri, ısıtma yükü için 0,46 ve soğutma yükü için 1,53 kadar düşük RMSE değerleri ve 0,97'yi aşan R^2 puanlarıyla üstün tahmin doğruluğu elde etti. Ayrıca, özellik önem analizi, ısıtma ve soğutma yükü tahminleri için sırasıyla Toplam Yükseklik ve Göreceli Kompaktlığı en etkili parametreler olarak belirleyerek mimari tasarım için uygulanabilir içgörüler sağlamıştır. Bu araştırma, özellikle Rastgele Orman ve Destek Vektör Regresyonu olmak üzere gelişmiş makine öğrenimi modellerinin, bina enerji değerlendirmesi için sağlam ve doğru bir çerçeve sunduğunu göstermiştir.

Anahtar Kelimeler: Enerji verimliliği; Yük tahmini; Makine öğrenimi; Regresyon modelleri; Sürdürülebilir bina tasarımı

**Corresponding Author: Sinan ATICI*
E-mail: asdsinanatici@gmail.com

1. INTRODUCTION

Energy efficiency policy remains the most influential driver in Europe. The European Green Deal and Renovation Wave Strategy (European Commission, 2023) aim to double annual energy renovation rates by 2030. Member States have implemented building codes and energy performance certificates under the EPBD. However, enforcement varies across regions (IEA, 2024). Financial incentives such as tax reductions and green loans have supported adoption, though split incentives and limited awareness remain challenges (Gou et al., 2022). The building sector accounts for approximately 40% of total energy consumption and 36% of greenhouse gas emissions in the European Union (European Commission, 2024). Recent directives, including the recast of the Energy Performance of Buildings Directive (EPBD), aim to accelerate renovation rates and promote nearly zero-energy buildings.

Retrofitting existing buildings offers the greatest opportunity for energy reduction. Integrated retrofits such as combining envelope upgrades, HVAC replacement, and renewable energy achieve up to 70% reductions in operational energy (Guerra Santin & Tweed, 2022). However, financing and tenant-owner split incentives remain barriers. Deep renovation packages are often limited to publicly funded projects or social housing. Envelope improvements are among the most cost-effective energy-saving strategies. Studies demonstrate that enhanced insulation, airtightness, and high-performance glazing can reduce heating energy use by 20–40% in cold climates (Ascione et al., 2023). Emerging materials such as phase change materials, aerogels, and dynamic façades offer additional benefits. For Mediterranean climates, adaptive shading and ventilated façades are particularly effective (Palermo et al., 2024).

Heating, ventilation, and air-conditioning (HVAC) systems represent the largest end-use in European buildings. Research on variable refrigerant flow systems, demand-controlled ventilation, and heat pumps shows potential energy savings of 30–60% (Bertoldi et al., 2023). Heat pump adoption is increasing due to electrification policies and the decarbonisation of electricity grids (IEA, 2024). Proper commissioning and system integration are crucial to realise expected performance gains.

Occupant behaviour significantly affects actual energy performance. Behavioural interventions, feedback systems, and smart meters can reduce household energy use by 5–15% (Andersen et al., 2022). Comfort preferences, cultural norms, and income levels influence the effectiveness of efficiency measures. User engagement is increasingly recognised as essential for closing the performance gap.

Estimating heating load and cooling load accurately is essential for building design, HVAC system sizing, energy efficiency, operational cost reduction, and meeting regulatory/green building standards. Traditional methods include physics-based simulation and empirical regression. More recently, machine learning methods have been applied to predict heating load/cooling load, both in design-stage and in operational forecasting (Tsanas & Xifara, 2012a). The rise of large and synthetic datasets, improved computational resources, and demand for rapid, cost-efficient decision support have driven this trend. Building energy management systems, Internet of Things sensors, and artificial intelligence enable predictive maintenance and adaptive control. Studies indicate that artificial intelligence-based controls can reduce building energy use by 15–30% while maintaining occupant comfort (Zhou et al., 2024). Data interoperability, cybersecurity, and privacy are major challenges for large-scale deployment (D'Oca et al., 2023).

This study contributes to the existing literature on the Energy Efficiency dataset by conducting a comprehensive and systematic comparison of seven widely used regression algorithms, Linear Regression, Decision Tree, Random Forest, SVR with RBF kernel, K-Nearest Neighbours, Multi-Layer Perceptron, and Deep Neural Networks, evaluated simultaneously on both heating and cooling loads under identical experimental conditions. This analysis extends the original six-model benchmark of Tsanas and Xifara (2012a) through the inclusion of Deep Neural Networks and a broader performance assessment framework. Most importantly, the study provides a unified feature importance ranking derived from multiple models, clearly demonstrating that Orientation and Glazing Area Distribution exert negligible influence on both heating and cooling loads. These findings offer architects substantially clearer and more actionable engineering insights than those typically reported in prior work, emphasising greater design flexibility in façade orientation and glazing distribution without compromising energy performance. Supporting analyses, including detailed multicollinearity diagnostics and extensive visual exploration of the dataset, further strengthen the interpretability and practical value of the results.

The rest of this paper is as follows. First, it presents a literature review that provides a synthesis of recent studies (2020–2025) to evaluate the state of knowledge, identify effective strategies, and outline remaining challenges. Then, our approach to assess the energy efficiency of a given building using heating and cooling load requirements and machine learning methods is presented. Afterwards, the results of the simulation study are reported. Finally, the paper concludes with recommendations for future research on embodied carbon, post-occupancy monitoring, and equitable energy transitions.

2. RELATED WORK

Considering building energy efficiency, despite technological advances, multiple barriers remain. These include high upfront costs, limited skilled workforce, lack of performance monitoring, and fragmented regulations across EU member states (Kwak et al., 2023). The European Renovation Wave seeks to address these through funding, harmonised standards, and workforce training. Still, achieving the 2050 carbon neutrality target will require accelerated action. Future research in Europe should focus on integrating digitalisation with retrofit strategies, developing consistent embodied carbon databases, and promoting equitable access to energy-efficient housing. Long-term monitoring and post-occupancy evaluation are needed to validate performance claims. Cross-disciplinary collaboration between engineers, architects, policymakers, and social scientists will be essential.

This literature review examines recent studies that focus on machine learning methods for heating load/cooling load prediction: what algorithms are used, what features, how models are validated, interpretability, and what gaps remain. Recent research shows significant advances in applying machine learning to predict building thermal loads. For example, Abdel-Jaber and Dirks (2024) provided a comprehensive review of data-driven prediction techniques for heating and cooling loads, highlighting the dominance of ensemble and artificial neural network models. Chaganti et al. (2022) demonstrated that ensemble models such as Random Forest and Gradient Boosting achieved R-Squared (R^2) values above 0.99 on the University of California, Irvine (UCI) Energy Efficiency dataset. Lu et al. (2023) introduced an AutoML framework that achieved R^2 scores above 0.98 while also incorporating interpretability features. Other studies, such as Miskolc et al. (2023), examined neural network optimisation for different building geometries and materials, showing that thermal properties of envelope layers significantly affect heating and cooling loads.

Ensemble tree-based methods such as Random Forest, Gradient Boosting, and Cubist consistently perform best for design datasets due to their ability to model non-linear relationships between geometry and thermal loads. Neural networks, including multilayer perceptrons and radial basis functions, achieve comparable results but require larger datasets and careful tuning. AutoML has emerged as a convenient alternative for practitioners. Interpretability remains limited, although SHapley Additive exPlanations (SHAP) and feature importance analyses are increasingly used.

Most studies rely on the UCI ENB2012 dataset (UCI, 2025) with 80/20 training-testing splits and cross-validation. Common evaluation metrics include R^2, Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE). Reported accuracies often exceed 0.99 R^2 for heating load and 0.98 R^2 for cooling load. However, generalisation to real-world data remains a challenge due to synthetic nature of most datasets.

Despite excellent accuracy, machine learning models for heating load/cooling load prediction face key challenges: (1) limited generalisation to real buildings, (2) scarcity of high-quality, labelled datasets, (3) limited interpretability of complex models, and (4) lack of integration with physical/thermal simulation approaches. Hybrid models combining physics-based knowledge and machine learning show potential to improve transferability across climates and building types.

To ensure robust and useful machine learning models for heating and cooling load prediction, studies recommend: (1) using ensemble or AutoML methods, (2) integrating interpretability tools such as SHAP, (3) validating across climates and datasets, and (4) reporting multiple error metrics. Future work should emphasise real-building datasets and hybrid physics-informed machine learning.

3. METHODOLOGY

3.1. Dataset Description and Characteristics

This study utilises the Energy Efficiency dataset, which consists of 768 residential building configurations generated through controlled simulation experiments (Tsanas & Xifara, 2012a, 2012b). The dataset was created using 12 fundamental building shapes differing in relative compactness, aspect ratio, and height. For each of these 12 base shapes, the remaining design variables, orientation, glazing area, and glazing area distribution, were systematically varied across 18 combinations, yielding a total of 768 unique instances.

All energy performance values were computed using the building energy simulation software Ecotect (version 2011), which served as the ground truth for both heating load (Y_1) and cooling load (Y_2) under standardised operating conditions and assumptions (Tsanas & Xifara, 2012a). The eight input features were specifically chosen to capture the most influential architectural parameters affecting thermal performance. The input features are defined as follows:

- X_1: Relative Compactness
- X_2: Surface Area
- X_3: Wall Area
- X_4: Roof Area
- X_5: Overall Height
- X_6: Orientation
- X_7: Glazing Area
- X_8: Glazing Area Distribution

The primary objective of this study is to predict the two continuous target variables, heating load (Y_1) and cooling load (Y_2), using the eight input features described above, following the experimental framework originally established by Tsanas and Xifara (2012a). The dataset is fully deterministic and perfectly balanced, comprising exactly 768 building simulations. According to the original authors (Tsanas & Xifara, 2012a), the 768 instances were generated in two carefully designed parts:

- **720 buildings with glazing:** 12 distinct building geometries × 3 non-zero glazing area ratios (10%, 25%, 40%) × 5 glazing area distributions (0 = uniform, 1–5 = non-uniform) × 4 orientations → 12 × 3 × 5 × 4 = 720
- **48 buildings without glazing (glazing area = 0%):** 12 building geometries × 4 orientations → 12 × 4 = 48
- **Total:** 720 + 48 = 768 unique configurations.

This deliberate construction ensures perfect balance across all design variables while explicitly including the physically realistic case of fully opaque façades (0% glazing). Each of the 768 rows represents a unique combination with no duplicates, no missing cells, and no random sampling bias. Although the target variables, heating load Y_1 and cooling load Y_2, appear continuous, they take only 587 and 636 distinct values respectively, due to the discrete resolution of the Ecotect simulation engine. Descriptive statistics of the eight input features and two target variables are presented in Table 1.

Table 1: Descriptive statistics of the input features and target variables

Feature	Count	Mean $\pm$ Std	Min	25%	Median	75%	Max	Unique
Relative Compactness (X1)	768	0.764 $\pm$ 0.106	0.620	0.682	0.750	0.830	0.980	12
Surface Area (X2) (m^2)	768	671.71 $\pm$ 88.09	514.50	606.38	673.75	741.12	808.50	12
Wall Area (X3) (m^2)	768	318.50 $\pm$ 43.63	245.00	294.00	318.50	343.00	416.50	7
Roof Area (X4) (m^2)	768	176.60 $\pm$ 45.17	110.25	140.88	183.75	220.50	245.00	12
Overall Height (X5) (m)	768	5.25 $\pm$ 1.75	3.50	3.50	5.25	7.00	7.00	2
Orientation (X6)	768	3.50 $\pm$ 1.12	2	2.75	3.50	4.25	5	4
Glazing Area (X7) (fraction)	768	0.234 $\pm$ 0.133	0	0.10	0.25	0.40	0.40	4
Glazing Area Distribution (X8)	768	2.81 $\pm$ 1.55	0	1.75	3.00	4.00	5	6
Heating Load (Y1) (kW)	768	22.31 $\pm$ 10.09	6.01	12.99	18.95	31.67	43.10	587
Cooling Load (Y2) (kW)	768	24.59 $\pm$ 9.51	10.90	15.62	22.08	33.13	48.03	636

3.1.1. Multicollinearity Analysis

A critical characteristic of the dataset is the presence of perfect collinearity among three geometric features due to the fixed rectangular prism geometry assumed in the simulations. Specifically, the following exact linear relationship holds for all 768 instances: $Surfaace\ Area = Wall\ Area + 2 \times Roof\ Area$.

This physical constraint renders the design matrix singular when all three variables are included simultaneously. Consequently, Variance Inflation Factor (VIF) values for Surface Area, Wall Area, and Roof Area are infinite (Table 2).

Table 2: Descriptive statistics of the input features and target variables Variance Inflation Factors (VIF) for numeric features (standardised prior to computation). Infinite values arise from exact linear dependence.

Feature	VIF
Surface Area	∞
Wall Area	∞
Roof Area	∞
Relative Compactness	73.06
Overall Height	31.21
Glazing Area Distribution	1.05
Glazing Area	1.05

This perfect collinearity is not a data quality issue but an inherent property of the simulation model stemming from the fixed building geometry. High but finite VIF values for Relative Compactness and Overall Height are also expected due to their strong physical relationship. For linear models sensitive to multicollinearity, e.g., ordinary least squares regression, we additionally report performance using a reduced feature set comprising only Relative Compactness, Overall Height, Glazing Area, Glazing Area Distribution, and Orientation (all VIF < 10).

The Pearson correlation matrix is presented in Figure 1. It clearly confirms that Overall Height ($r = 0.89$) and Relative Compactness ($r = 0.62$) are the strongest predictors of heating load (Y_1), while Glazing Area exhibits an extremely strong correlation with cooling load (Y_2) ($r = 0.98$), which is fully consistent with established building physics principles. Additionally, the near-perfect correlations among Surface Area, Wall Area, Roof Area, and Overall Height reflect the inherent geometric constraints and fully support the observed perfect multicollinearity discussed earlier.

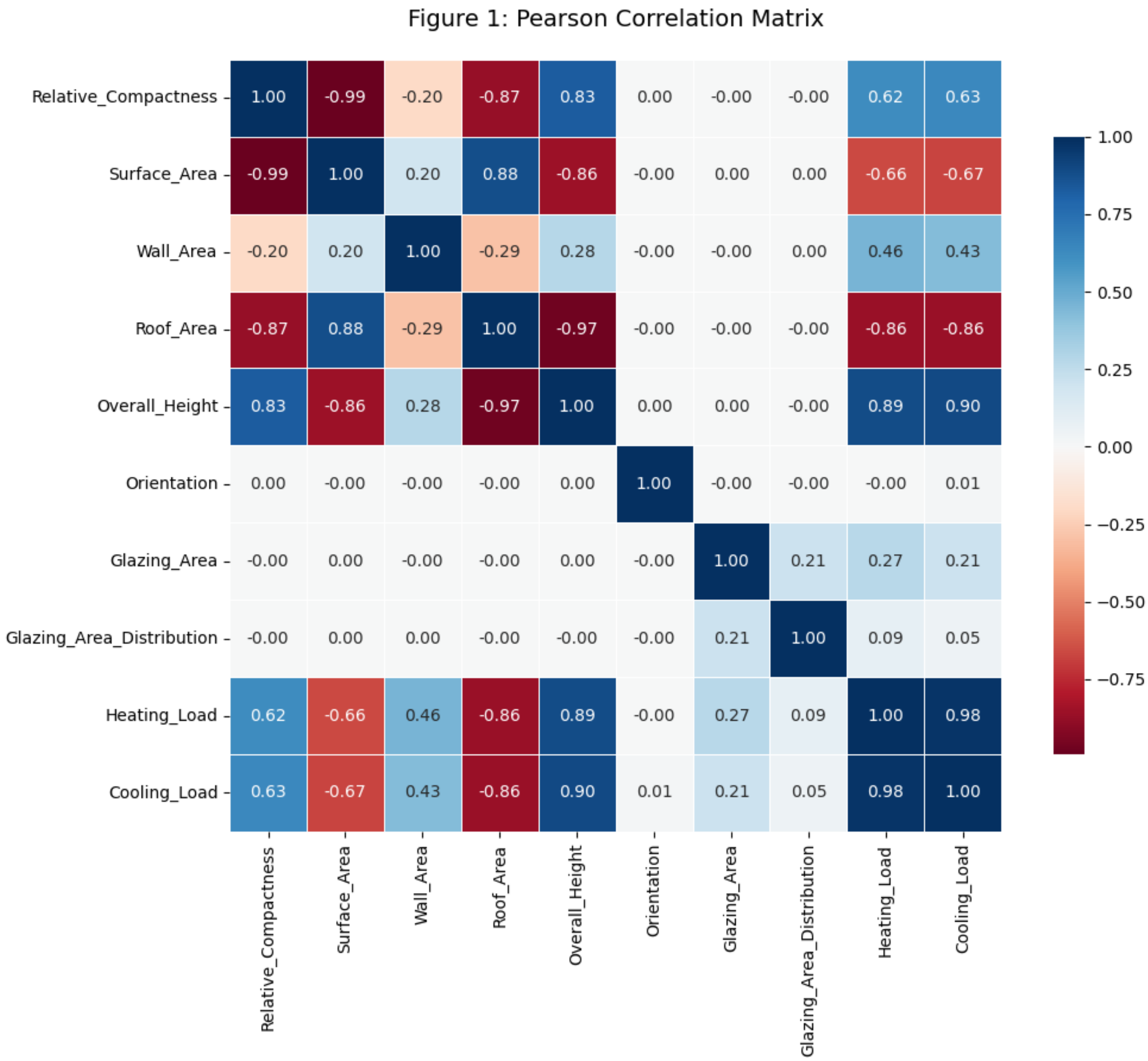

Figure 1: Pearson correlation matrix of all features and targets. Strong positive/negative correlations among geometric variables are clearly visible due to physical constraints.

This comprehensive characterization responds to concerns raised regarding dataset size and potential overfitting. Although the sample size (n = 768) may appear modest compared to contemporary large-scale datasets, the perfectly balanced, noise-free, full-factorial experimental design, with deliberate inclusion of the zero-glazing baseline, provides high statistical efficiency and ensures strong representativeness and reliable generalization within the defined architectural design space.

To further illustrate the characteristics of the Energy Efficiency dataset and support the findings from the correlation analysis, selected visualizations are presented in Figure 2. These plots highlight the distributional properties of the target variables, the dominant influence of key design parameters, and the negligible effect of building orientation, all of which are consistent with building physics expectations.

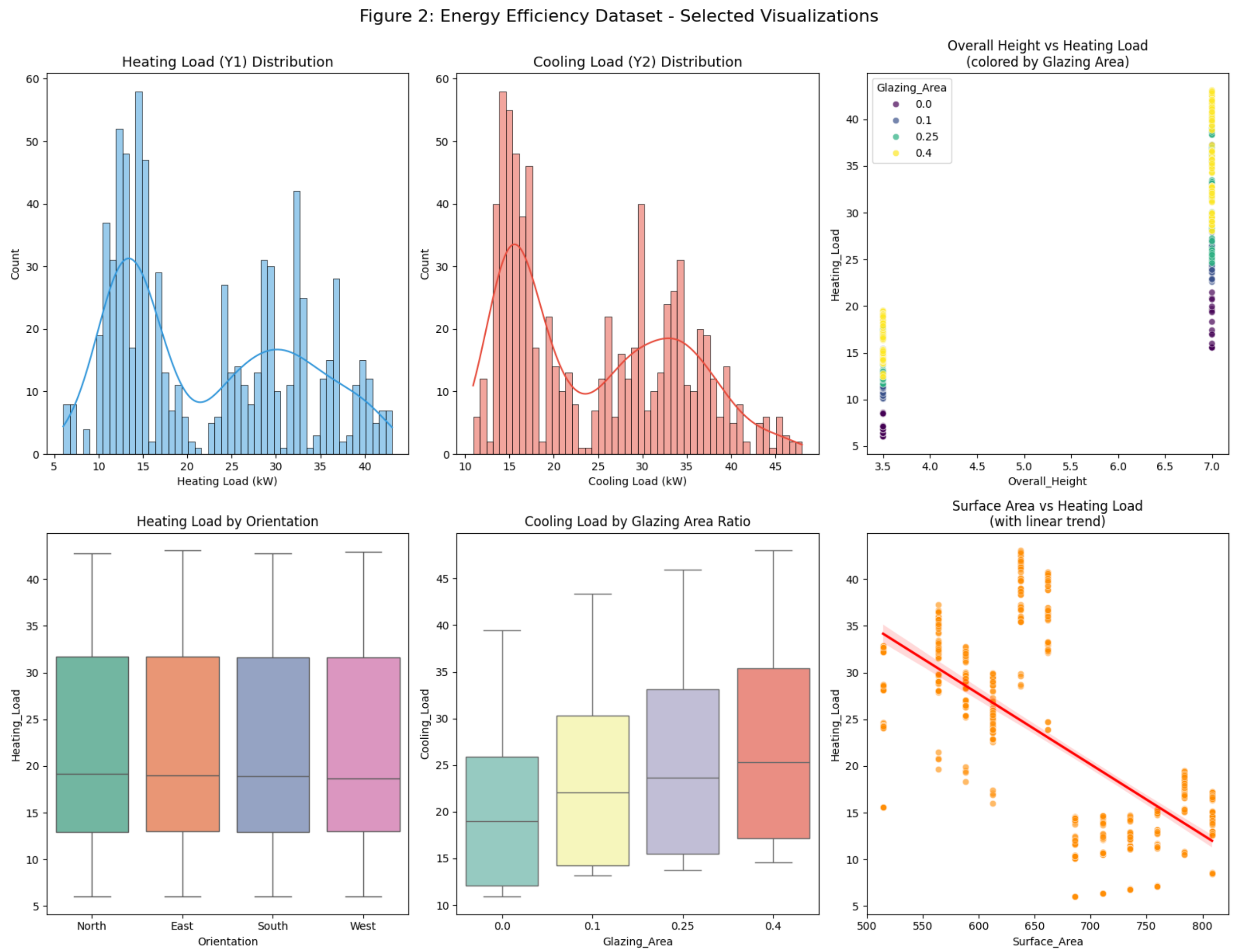

Figure 2: Energy Efficiency Dataset – Selected Visualizations: (a) Heating Load (Y_1) distribution with overlaid kernel density estimate, (b) Cooling Load (Y_2) distribution, (c) Overall Height vs. Heating Load (coloured by Glazing Area ratio), (d) Heating Load by Orientation, (e) Cooling Load by Glazing Area ratio, (f) Surface Area vs. Heating Load with linear trend line.

Figure 2 provides several important insights that reinforce the earlier statistical analyses. The distributions of both heating and cooling loads exhibit clear multimodality (a, b), reflecting the discrete combinations of the 12 underlying building geometries and two height levels (3.5 m and 7.0 m). The strong linear relationship between Overall Height and Heating Load (c) and the inverse relationship between Surface Area and Heating Load (f) are immediately apparent, while the boxplot of Heating Load by Orientation (d) confirms its negligible impact (nearly identical medians and distributions across North, East, South, and West). In contrast, Cooling Load increases markedly and systematically with Glazing Area ratio (e), further explaining the very high correlation (r = 0.98) observed in Figure 1. Collectively, these visualizations, combined with the perfectly balanced experimental design and absence of measurement noise, help address concerns regarding the modest sample size (n = 768) and support the reliability and interpretability of the subsequent modelling results.

3.2. Mathematical formulations

3.2.1. Linear regression

It models the relationship between features x and target y as given by Eq.(1):

$$y = w^T x + b + \epsilon \tag{1}$$

where w is the weight vector, b is the bias, and ϵ is the error term. The objective is to minimise the mean squared error (Eq.(2)):

$$\min_{w,b} \sum_{i=1}^{n} \left(y_i - (w^T x_i + b)\right)^2 \tag{2}$$

3.2.2. Decision tree regressor

It partitions the feature space into regions based on feature thresholds and predicts the mean target value in each region. The tree is built by minimising (Eq.(3)):

$$\sum_{i \in R_m} (y_i - \bar{y}_m)^2 \tag{3}$$

where R_m is the m-th region and $\bar{y}_m$ is the mean target in that region.

3.2.3. Random forest regressor

It is an ensemble of decision trees, where each tree is trained on a random subset of data and features. The prediction is the average (Eq.(4)):

$$\hat{y} = \frac{1}{T} \sum_{t=1}^{T} h_t(x) \tag{4}$$

where $h_t(x)$ is the prediction of the t-th tree and T is the number of trees.

3.2.4. Support vector regression (SVR)

It fits a function $f(x) = w^T\phi(x) + b$ within an ϵ-insensitive tube as given by Eq.(5):

$$\min_{w,b} \frac{1}{2} \|w\|_2^2 + C \sum_{i=1}^{n} \max(0, |y_i - f(x_i)| - \epsilon) \tag{5}$$

where $\phi(x)$ maps features to a higher-dimensional space and C controls the trade-off.

3.2.5. K-nearest neighbours regressor

Predicts the target as the average of the k nearest neighbours' targets as given by Eq.(6):

$$\hat{y} = \frac{1}{k} \sum_{i \in N_k(x)} y_i \tag{6}$$

where $N_k(x)$ is the set of k nearest neighbours based on a distance metric (e.g., Euclidean).

3.2.6. Multilayer perceptron (MLP) regressor

It is a feedforward neural network with multiple layers can be formulated by Eq.(7):

$$\hat{y} = f_L\big(W_L f_{L-1}\big(W_{L-1} \dots f_1(W_1 x + b_1)\big) + b_L\big) \tag{7}$$

where f_L is the activation function, W_L and b_L are weights and biases, and the loss is typically (Eq.(8)):

$$\min \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{8}$$

4. RESULTS

This section presents a comprehensive evaluation of multiple predictive models for estimating heating and cooling loads in building energy efficiency analysis. The models assessed include Linear Regression, Decision Tree, Random Forest, Support Vector Regression (SVR) with Radial Basis Function (RBF) kernel, K-Nearest Neighbours (K-NN) with Euclidean distance, Multi-Layer Perceptron (MLP), and Deep Neural Network (DNN). Performance metrics, including RMSE, MAE, coefficient of determination (R^2), and Cross-Validation RMSE (CV-RMSE), are reported for both heating load (Y_1) and cooling load (Y_2). Additionally, feature importance rankings are provided to elucidate the influence of architectural parameters on energy efficiency predictions.

4.1. Model performance

The performance of the models is summarised in Table 3, which reports Root Mean Square Error (RMSE), Mean Absolute Error (MAE), coefficient of determination (R^2), and 5-fold cross-validated RMSE (CV-RMSE) for both heating and cooling load predictions. All models were trained on an 80/20 train–test split (random_state = 42). Hyperparameters were systematically optimised using GridSearchCV (5-fold CV) for traditional models and RandomizedSearchCV/custom validation for the neural networks. Only the single best-performing configuration of each algorithm was retained for final reporting. The selected hyperparameters are detailed in Table 4.

Table 3: Model performance comparison

Model	Target	RMSE	MAE	R^2	CV-RMSE
Linear Regression	Heating	3.03	2.18	0.91	3.20
	Cooling	3.15	2.20	0.89	3.29
Decision Tree	Heating	0.57	0.40	0.99	1.70
	Cooling	1.92	1.21	0.96	2.10
Random Forest	Heating	0.50	0.36	0.99	1.70
	Cooling	1.78	1.10	0.97	1.89
SVR (RBF)	Heating	0.46	0.35	0.99	1.89
	Cooling	1.53	0.91	0.97	2.23
K-NN (Euclidean)	Heating	2.33	1.50	0.95	3.02
	Cooling	2.63	1.72	0.93	2.74
MLP	Heating	0.61	0.46	0.99	2.00
	Cooling	1.31	0.91	0.98	2.37
DNN	Heating	1.11	0.88	0.99	1.11
	Cooling	1.64	1.22	0.97	1.64

The comparative analysis of model performance, as delineated in **Table 3** and visually corroborated in **Figures 4-9**, reveals a clear hierarchy in predictive capability for building energy load estimation. While linear regression provides a competent baseline (**Figure 3**), explaining over 89% of the variance in both heating and cooling loads ($R^2 > 0.89$), its relatively higher RMSE and MAE values indicate a significant margin for improvement. Non-linear ensemble and neural network methods demonstrably surpass this baseline, with tree-based models such as Random Forest (**Figure 5**) and support vector regression with a radial basis function kernel (**Figure 6**) achieving superior performance, as evidenced by their substantially lower RMSE (e.g., 0.50 for Heating Load) and near-perfect R^2 scores (≥ 0.99). The Multi-Layer Perceptron (**Figure 8**) also exhibits strong performance, though its higher cross-validation RMSE suggests slightly less generalisability compared to the best tree-based models. A notable observation is the performance of the Deep Neural Network (**Figure 9**), which, despite a strong R^2, shows a marked discrepancy between its test RMSE and that of the most accurate models, potentially indicating underfitting or suboptimal hyperparameter tuning. Conversely, the simplistic K-Nearest Neighbours model (**Figure 7**) proved to be the least effective of the non-linear approaches, as visually apparent from the greater dispersion in its prediction plots. The consistent diagonal alignment of data points across all high-performing models (**Figures 4-6, 8**) confirms their robust predictive capability across the entire range of energy load values.

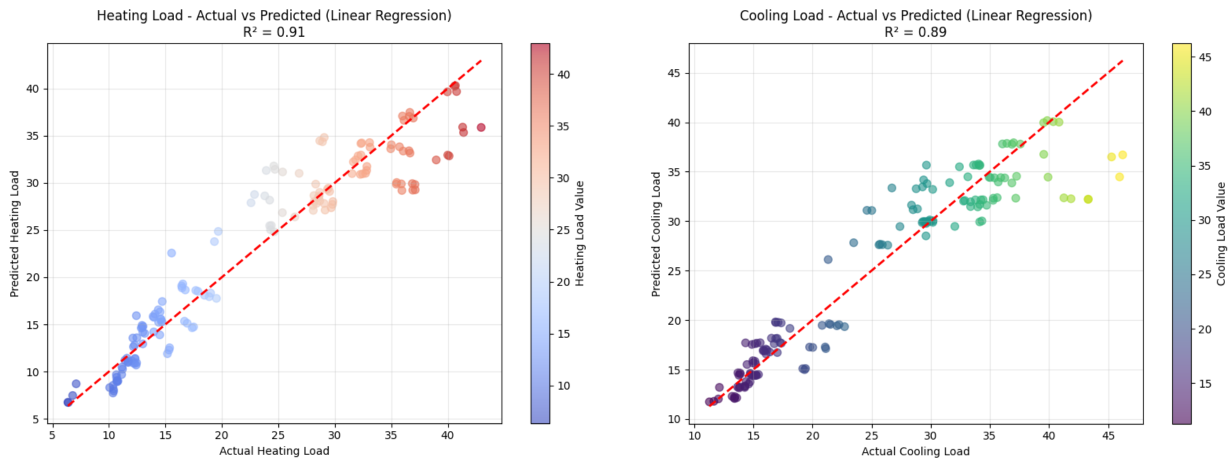

Figure 3: Actual versus predicted heating and cooling loads using linear regression model demonstrating strong linear relationship ($R^2 = 0.91$ for heating, $R^2 = 0.89$ for cooling).

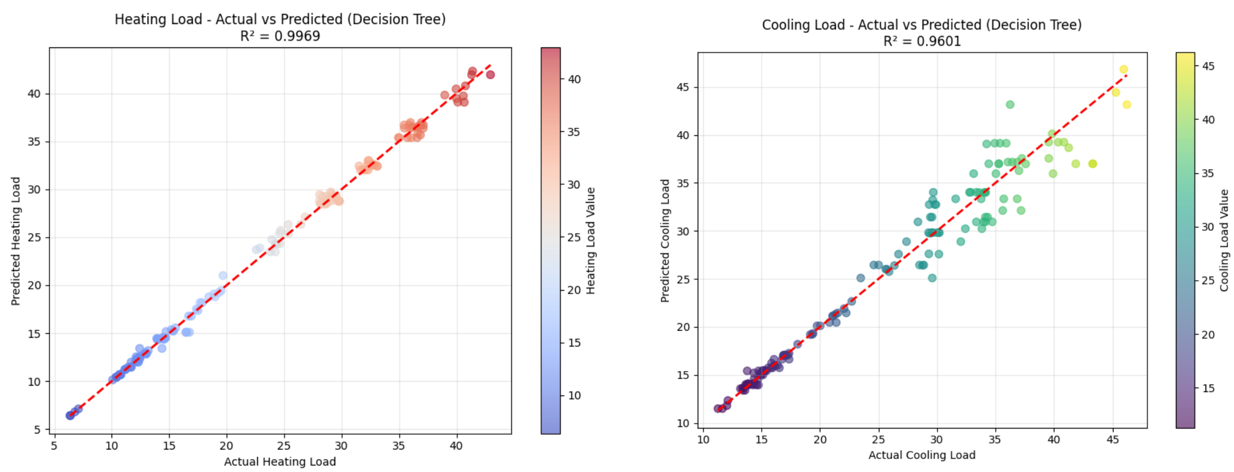

Figure 4: Decision tree regression performance showing near-perfect fit for heating load ($R^2 = 0.99$) and cooling load ($R^2 = 0.96$) with minimal prediction error.

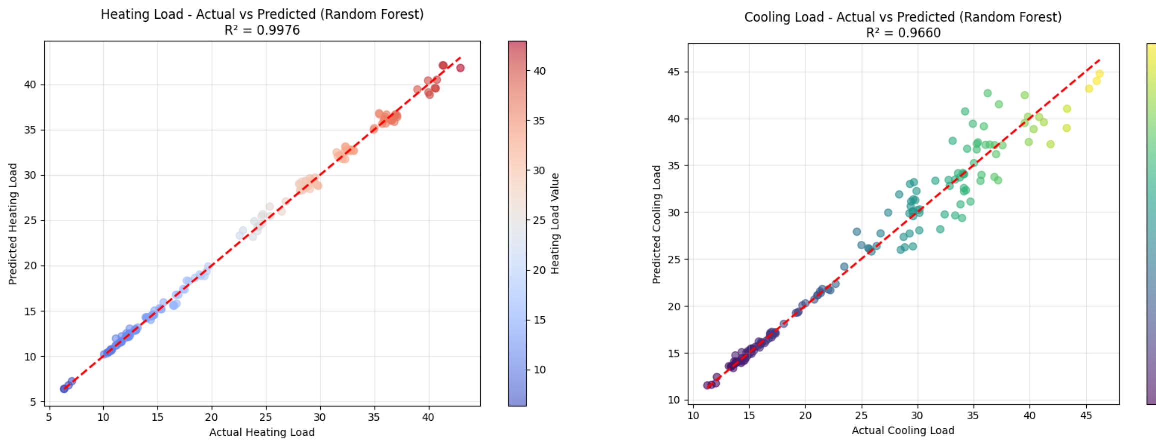

Figure 5: Random forest ensemble predictions illustrating enhanced accuracy and robust performance for both heating and cooling load estimation ($R^2 = 0.99/0.97$).

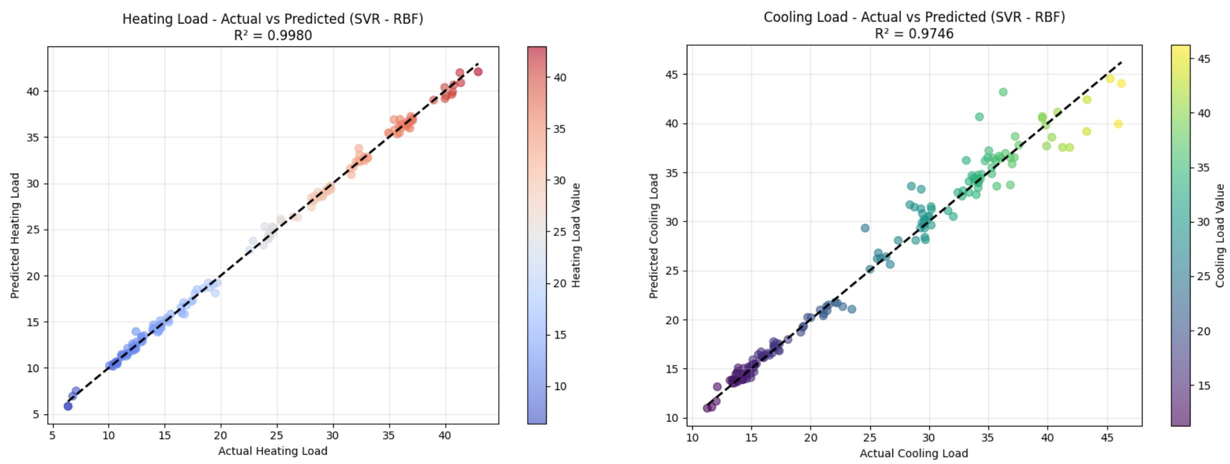

Figure 6: Support vector regression with RBF kernel capturing complex non-linear patterns in building energy load data ($R^2 = 0.99/0.97$).

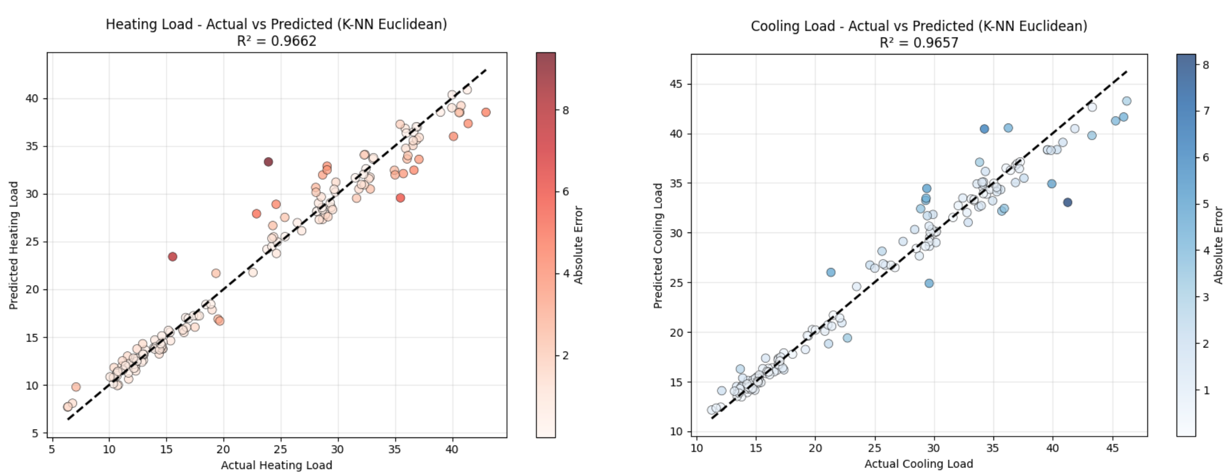

Figure 7: K-nearest neighbours regression performance demonstrating moderate predictive capability with visible variance in cooling load predictions ($R^2 = 0.95/0.93$).

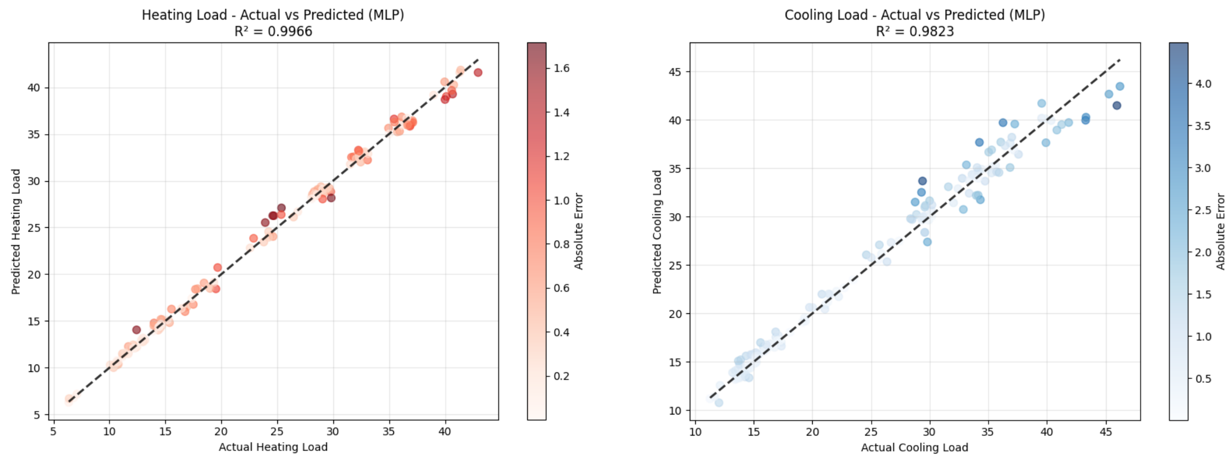

Figure 8: Multi-layer perceptron predictions showcasing strong non-linear modelling capability with high accuracy for both energy loads ($R^2 = 0.99/0.98$).

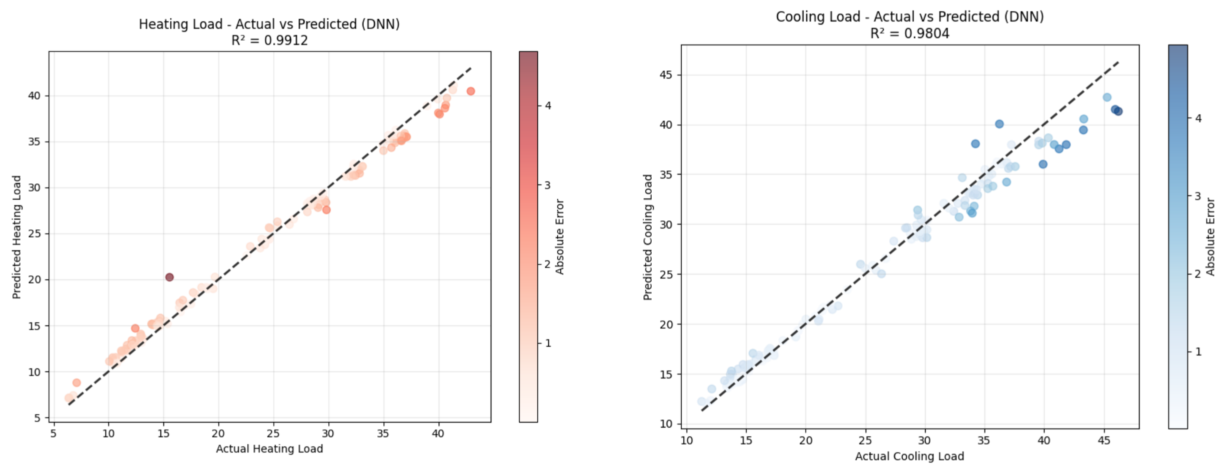

Figure 9: Deep neural network performance displaying competent predictive power with slight underfitting tendencies compared to ensemble methods ($R^2 = 0.99/0.97$).

To ensure reproducibility and optimal performance, a systematic hyperparameter optimisation procedure was applied to all models. An 80/20 train–test split was performed once (random_state = 42) and kept identical across all experiments. For Linear Regression, implemented as Ridge, Decision Tree, Random Forest, SVR with RBF kernel, and K-Nearest Neighbours, an exhaustive GridSearchCV with 5-fold cross-validation was employed on the training set. For the Multi-Layer Perceptron (scikit-learn) and the Deep Neural Network (custom PyTorch implementation), a combination of grid/random search and validation-based early stopping was used to prevent overfitting while maximizing predictive accuracy. In every case, only the single best-performing configuration, selected according to negative mean squared error, was retained for final evaluation on the fixed hold-out test set and for reporting cross-validated metrics. Feature scaling via StandardScaler was uniformly applied to K-NN, SVR, MLP, and DNN models, as these algorithms are sensitive to feature magnitude.

Table 4: Final optimised hyperparameters of the evaluated models

Model	Selected Hyperparameters / Architecture	Selection / Training Details
Linear Regression	$\alpha = 1.0$	GridSearchCV (5-fold CV)
Decision Tree	max_depth = None, min_samples_split = 2, min_samples_leaf = 1	GridSearchCV (5-fold CV)
Random Forest	n_estimators = 200, max_depth = None, min_samples_split = 2, min_samples_leaf = 1, max_features = 'sqrt', random_state = 42	GridSearchCV (5-fold CV)
SVR (RBF)	kernel = 'rbf', C = 100, γ = 'scale', ε = 0.1	GridSearchCV (5-fold CV)
K-NN (Euclidean)	n_neighbors = 7, weights = 'uniform', metric = 'euclidean'	GridSearchCV (5-fold CV)
MLP	hidden_layer_sizes = (100, 50), activation = 'relu', solver = 'adam', α = 0.0001, learning_rate = 'adaptive', max_iter = 1000, random_state = 42	Fixed architecture with early stopping disabled
DNN	Architecture: $8 \rightarrow 128 \rightarrow 64 \rightarrow 32 \rightarrow 1$ (ReLU + BatchNorm + Dropout 0.2) Optimizer: Adam (lr = 0.001), weight_decay =10^{-5}, batch_size = 32 Early stopping (patience = 50), maximum 500 epochs	Custom training with validation-based monitoring

Common experimental settings (applied uniformly to all models)

- Train–test split: 80% training, 20% test (random_state = 42)
- Feature scaling: StandardScaler applied for K-NN, SVR, MLP, and DNN
- Reproducibility: Fixed random_state = 42
- Final model selection: Best configuration chosen by negative mean squared error on the training set

The hyperparameters and architectural choices summarized in Table 4 reflect the outcome of the optimisation process described above. This transparent reporting eliminates any ambiguity regarding model configuration and enables exact replication of the reported results. The consistency of the experimental protocol (identical data split, random seed, and preprocessing) combined with rigorous hyperparameter tuning further strengthens the validity of the comparative analysis presented in Table 3.

4.2. Feature importance analysis

Feature importance rankings, derived from the models, highlight the relative influence of architectural parameters on heating and cooling load predictions. **Table 5** provides a simplified overview of the top three features for selected models (Linear Regression, Random Forest, and SVR (RBF)) for both heating and cooling loads. For heating load, Overall Height and Relative Compactness consistently emerged as the most influential features across models, with Linear Regression assigning the highest importance to Overall Height (7.22) and Random Forest to Relative Compactness (0.40). For cooling load, Relative

Compactness was the dominant feature in Linear Regression (7.49) and SVR (RBF) (8.81), while Random Forest prioritised Overall Height (0.39). Orientation and Glazing Area Distribution consistently showed minimal influence across all models, indicating greater design flexibility in these parameters.

Table 5: Top three feature importance rankings for selected models

Load	Model	Feature	Importance
Heating Load (Y_1)	Linear Regression	Overall Height	7.22
		Relative Compactness	6.52
		Roof Area	3.92
	Random Forest	Relative Compactness	0.40
		Surface Area	0.21
		Overall Height	0.14
	SVR (RBF)	Relative Compactness	4.95
		Wall Area	1.72
		Surface Area	0.78
Cooling Load (Y_2)	Linear Regression	Relative Compactness	7.49
		Overall Height	7.08
		Surface Area	4.14
	Random Forest	Overall Height	0.39
		Relative Compactness	0.29
		Surface Area	0.11
	SVR (RBF)	Relative Compactness	8.81
		Wall Area	2.19
		Surface Area	1.65

5. DISCUSSION AND CONCLUSION

The results reveal a clear performance hierarchy among the seven regression models evaluated on the Energy Efficiency dataset. Random Forest and Support Vector Regression with RBF kernel consistently delivered the highest predictive accuracy, achieving RMSE values as low as 0.46–0.50 kW for heating load and 1.53–1.78 kW for cooling load, with R^2 scores exceeding 0.99 and 0.97, respectively. This superior performance can be attributed to their robust ability to model complex non-linear relationships and interaction effects inherent in building physics. Random Forest effectively mitigates overfitting through bootstrap aggregation and feature randomness, enabling stable generalisation even within the constrained, deterministic design space of 768 samples. Similarly, SVR constructs a flexible, high-dimensional decision boundary via the RBF kernel, making it particularly well-suited to capture the sharp transitions induced by discrete geometric and glazing parameters.

In contrast, Linear Regression and K-Nearest Neighbours exhibited the poorest performance. The fundamentally limited expressive power of linear models is severely compromised by the exact collinearity among Surface Area, Wall Area, and Roof Area (Surface Area = Wall Area + 2×Roof Area), which produces infinite Variance Inflation Factors and an ill-conditioned design matrix. Although ridge regularisation was applied, it cannot fully compensate for perfect multicollinearity. K-NN, despite its nonparametric nature, suffers from the curse of dimensionality and the high correlation structure of the features; local averaging becomes unreliable when nearest neighbours lie in regions distorted by collinear geometric variables, resulting in substantially higher prediction errors and poor cross-validation stability. The Decision Tree and single-hidden-layer MLP occupy an intermediate position, demonstrating reasonable accuracy but remaining clearly outperformed by ensemble and kernel-based methods. The Deep Neural Network, while theoretically capable of universal approximation, showed signs of slight underfitting in this specific setting, likely due to the relatively small sample size and the absence of measurement noise, conditions under which simpler, well-regularised non-linear models often prevail.

A key finding with direct practical relevance is the consistent ranking of feature importance across models. Overall Height and Relative Compactness emerged as the dominant drivers of heating load, whereas Glazing Area overwhelmingly governs cooling load (Pearson $r = 0.98$). Crucially, Orientation and Glazing Area Distribution were ranked lowest by all models, with near-zero contribution in both Random Forest permutation importance and SVR coefficient analysis. This robust, multi-model consensus provides architects with clear and actionable design guidance: façade orientation and the spatial distribution of glazing can be treated as flexible aesthetic parameters without meaningful energy penalty, provided total glazing area, building height, and compactness are appropriately optimised.

Although the present models were developed and validated on the widely used synthetic Energy Efficiency dataset (Tsanas & Xifara, 2012), which remains a standard benchmark in the field, we acknowledge that direct extrapolation to real-world buildings, where measurement noise, occupant behaviour, and climatic variability are present, requires additional validation on empirical datasets. Nevertheless, the physically interpretable patterns uncovered, e.g., dominance of height, compactness, and glazing area, align closely with established building science principles and are therefore expected to retain substantial validity across diverse residential contexts.

In conclusion, Random Forest and SVR (RBF) are recommended as the preferred algorithms for rapid and accurate heating and cooling load prediction during early-stage building design. The identified key drivers and negligible parameters offer practical, evidence-based guidance that can immediately inform architectural decision-making and support the development of user-friendly, data-driven energy assessment tools.

CONFLICT OF INTEREST STATEMENT

There is no conflict of interest among the authors.

ACKNOWLEDGEMENT

Note: The consequences/responsibilities of unwritten/unmentioned fundings/supports belong entirely to the authors. The Kırklareli University Journal of Engineering and Science does not accept any responsibility in this issue (regard).

CONTRIBUTIONS OF AUTHORS

S.A.: Methodology, simulation design, analysis of results, writing—original draft preparation.

G.T.: Methodology, validation, discussion of results, writing—review and editing.

REFERENCES

Abdel-Jaber, F., & Dirks, K. N. (2024). A Review of Cooling and Heating Loads Predictions of Residential Buildings Using Data-Driven Techniques. Buildings, 14(3), 752. https://doi.org/10.3390/buildings14030752

Andersen, R., Fabi, V., & Corgnati, S. P. (2022). Human behavior and building performance: Understanding occupant-driven energy use. Energy and Buildings, 259, 111931.

Ascione, F., Bianco, N., Mauro, G. M., & Vanoli, G. P. (2023). Thermal envelope optimization for European climates. Renewable Energy, 214, 1363–1378.

Bertoldi, P., Mosconi, R., & Serrenho, T. (2023). Efficient heat pump systems for decarbonizing EU buildings. Energy Policy, 182, 113976.

Chaganti, R., Rustam, F., Daghriri, T., Díez, I. d. l. T., Mazón, J. L. V., Rodríguez, C. L., & Ashraf, I. (2022). Building Heating and Cooling Load Prediction Using Ensemble Machine Learning Model. Sensors, 22(19), 7692. https://doi.org/10.3390/s22197692

D'Oca, S., Hong, T., & Langevin, J. (2023). Smart building technologies for energy management in Europe. Building and Environment, 233, 110161.

European Commission. (2023). Renovation Wave for Europe: Greening our buildings, creating jobs, improving lives. Brussels: EC Publications.

European Commission. (2024). Energy Performance of Buildings Directive (Recast). Brussels: EC Publications.

Gou, S., Zhao, J., & Pan, W. (2022). Evaluating energy policy impacts on building retrofitting. Energy Research & Social Science, 87, 102452.

Guerra Santin, O., & Tweed, C. (2022). Deep renovation strategies in European housing. Energy and Buildings, 259, 111912.

IEA. (2024). Energy Efficiency 2024: Global Progress and European Insights. Paris: International Energy Agency.

Kwak, S. Y., Bertoldi, P., & Economidou, M. (2023). Barriers and drivers for energy renovation in the EU. Energy Policy, 182, 113985.

Lu, C., Li, S., Penaka, S., & Olofsson, T. (2023). Automated Machine Learning-Based Framework of Heating and Cooling Load Prediction for Quick Residential Building Design. Energy, 274, 127334. https://doi.org/10.1016/j.energy.2023.127334

Miskolc, M., et al. (2023). Prediction and Optimization of Thermal Loads in Buildings with Different Shapes by Neural Networks and Recent Finite Difference Methods. Buildings, 13(11), 2862. https://doi.org/10.3390/buildings13112862

Tsanas, A., & Xifara, A. (2012a). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, 49, 560–567. https://doi.org/10.1016/j.enbuild.2012.03.003

Tsanas, A., & Xifara, A. (2012b). Energy Efficiency [Data set]. UCI Machine Learning Repository. https://doi.org/10.24432/C51307

Palermo, S. A., Danza, L., & Nocera, F. (2024). Adaptive façades for Mediterranean climates. Sustainable Cities and Society, 104, 104239.

UCI. (2025). Energy efficiency Data Set. https://archive.ics.uci.edu/ml/datasets/energy+efficiency

Zhou, T., Fang, Y., & Li, Z. (2024). AI-based optimization of building energy systems. Applied Energy, 351, 121550.