International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

MATHEMATICAL MODELLING AND MACHINE LEARNING: INTEGRATING COMPUTER SCIENCE AND AI FOR COMPLEX PROBLEM SOLVING

Ch Venkata S S P Kumar[1*], Maheswaram Bhanuchandhar[2], Maged Abdullah Matter[3], Indrajeet G Varhadpande[4], Mr. Ananda Shamrao Patil[5], Naveen Gupta[6], Mayank Saini[7]

[1*]Assistant Professor, Department of Computer Science Engineering, Sri Sivani College of Engineering, Srikakulam, A.P, India, Orcid Id: 0009-0001-0966-7472, Email Id: venkatachalla90@gmail.com
[2]Assistant Professor, Department of EEE, Sri Sivani College of Engineering, Srikakulam, A.P, India, Orcid Id: 0009-0009-1827-8872, Email Id: m.bhanuchandhar.123@gmail.com
[3]M D Neurosurgery Consultant, Students Hospital Cairo University, Email Id: mageden@hotmail.com
[4]Assistant Professor, Department of Engineering, Sciences & Humanities (Mathematics), St. Vincent Pallotti College of Engineering and Technology, Nagpur, Maharashtra, Email Id: indrajeetv14@gmail.com
[5]Assistant Professor, Department of General Engineering, Bharati Vidyapeeth's College of Engineering, Kolhapur, Email Id : ananda.patilbvc@gmail.com
[6]Head of ICT Department, St. Mark's Sr. Sec. Public School, Meera Bagh, Delhi, Email Id: Gupta.naveen1@gmail.com, Orcid Id: https://orcid.org/0009-0009-4512-000X
[7]Pre final year student, Masters of Computer Application, Galgotias University, Greater Noida, Uttar Pradesh – 203201, Email Id: mayanksainikht@gmail.com

Abstract

This study develops a hybrid mathematical–machine learning framework that integrates computational modelling, artificial intelligence (AI), and computer science for complex problem solving in building energy prediction. Using the UCI Energy Efficiency dataset, the research compares the performance of multivariate regression, Artificial Neural Networks (ANN), and Support Vector Regression (SVR) models to forecast heating and cooling loads. The proposed hybrid model combines the interpretability of mathematical regression with the nonlinear learning capability of machine learning through residual correction. Statistical analyses demonstrate that the hybrid model achieved superior performance with R^2 values exceeding 0.97 and RMSE values below 1.0 for both outputs, outperforming standalone models by 40% in predictive accuracy. The integration framework maintains theoretical transparency while enhancing computational adaptability, thereby bridging the gap between deterministic modelling and AI-driven prediction. These findings underscore the importance of combining mathematical precision with intelligent systems to address energy modelling challenges and broader applications in sustainable design and computational optimisation.

Keywords: Mathematical modelling; Machine learning; Hybrid modelling; Energy efficiency; Artificial intelligence

1. Introduction

Energy consumption in the built environment represents one of the most pressing global challenges in achieving sustainability and carbon neutrality. Buildings alone account for nearly 40% of total global energy use, with heating and cooling loads forming the largest portion of operational demand.

Received: August 15, 2025

The ability to accurately predict building energy consumption has therefore become a vital component of urban sustainability strategies, energy management systems, and low-carbon design frameworks (Kong et al., 2023; Li et al., 2017). Traditional mathematical models based on physical and thermodynamic principles have long served as fundamental tools for estimating energy use in buildings. These models are valued for their interpretability and analytical rigor; however, they often rely on simplifying assumptions that limit their capacity to represent the complex nonlinear interactions present in real-world environments (Karniadakis et al., 2021). Consequently, mathematical models can underperform when addressing multidimensional energy systems influenced by varying climatic, behavioral, and architectural parameters.

Recent advances in artificial intelligence (AI) and machine learning (ML) have opened new frontiers in predictive modeling for energy systems, particularly in building energy forecasting. Deep learning techniques such as convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and hybrid CNN-LSTM models have demonstrated exceptional predictive capability in time-series energy data by capturing spatial and temporal dependencies (Alhussein et al., 2020; Runge & Zmeureanu, 2021). However, despite their performance advantages, such models are often criticized for their lack of interpretability and dependence on large datasets. To overcome these limitations, hybrid frameworks integrating mathematical modeling and machine learning have emerged as a promising alternative, combining the transparency of analytical equations with the adaptability of data-driven learning (Patel et al., 2022). These hybrid models allow the exploitation of underlying physical principles while simultaneously learning complex nonlinear patterns in energy behavior.

Urban Building Energy Modeling (UBEM) provides a large-scale example of this integration. UBEM aims to simulate energy use across city districts, and recent studies highlight the need to combine simulation-based approaches with ML-driven data analytics for improved calibration, scalability, and predictive accuracy (Kong et al., 2023; Li et al., 2017). The comprehensive review by Kong et al. (2023) emphasizes that hybrid data–physics approaches are essential for addressing the limitations of purely physics-based or purely data-driven models in urban contexts. Similarly, Li et al. (2017) demonstrated that traditional modeling approaches require substantial computational resources and often fail to generalize across diverse building typologies and operational conditions. Thus, the integration of AI with mathematical and physics-based models has become a central research direction in advancing the predictive accuracy and efficiency of building energy modeling.

At the methodological level, physics-informed machine learning (PIML) has become a transformative concept that integrates scientific laws into data-driven models. Karniadakis et al. (2021) described how embedding physical constraints into learning systems enhances generalization and interpretability, ensuring that predictions remain consistent with established physical behavior. Similarly, Patel et al. (2022) demonstrated that physics-constrained neural networks outperform purely data-driven models by preventing unphysical outcomes, improving both prediction accuracy and model reliability. These frameworks bridge the gap between traditional mathematical modeling and advanced machine learning, offering an interpretable and theoretically grounded approach to data analytics.

While the integration of physical and computational intelligence models has improved prediction accuracy, recent studies underscore a new challenge—the need for explainability. As the energy sector increasingly adopts AI, stakeholders demand transparency in how models reach their predictions. Chen and Debnath (2025) stressed the necessity of explainability in low-carbon urban systems, arguing that the deployment of AI in city-scale energy management must align with ethical and policy frameworks. Similarly, Gopal et al. (2025) highlighted how explainable AI (XAI) enables model interpretability, helping engineers and decision-makers understand the rationale behind energy forecasts. The growing attention to XAI reflects a paradigm shift toward human-centric AI systems

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

capable of providing actionable and trustworthy insights (Khedr, 2024). XAI-based approaches are particularly relevant in hybrid energy models, where mathematical structure and ML predictions can be linked to physically meaningful variables.

Recent research further demonstrates the evolution of hybrid and explainable models for building energy analysis. Wang et al. (2025) conducted a comprehensive review of ensemble learning techniques for building energy prediction, concluding that combining multiple models can significantly enhance performance, reduce uncertainty, and improve adaptability across varying datasets. Fan and Chen (2025) introduced an explainable CNN-LSTM hybrid model for energy forecasting, which not only achieved higher accuracy but also improved interpretability by identifying dominant influencing parameters. Similarly, Quevedo et al. (2023) developed ML-based energy benchmarking models for university buildings, emphasizing the value of combining data analytics with energy performance indicators for localized decision-making. Collectively, these studies reinforce that the convergence of AI, hybrid modeling, and explainability marks a critical step toward intelligent, transparent, and sustainable energy systems.

Despite these advances, several knowledge gaps persist. Existing models often lack generalizability across building types and climates, suffer from limited interpretability, or rely heavily on large, clean datasets that are not always available (Molina-Solana et al., 2017; Chen et al., 2022). Moreover, many studies focus on performance metrics rather than theoretical integration, neglecting how mathematical modeling principles can enhance the interpretability of ML outputs. The present research seeks to address these gaps by developing a hybrid mathematical–machine learning framework for energy load prediction that balances interpretability and predictive precision. Using the UCI Energy Efficiency dataset, the study formulates a multiple regression model to capture analytical relationships between architectural features and energy loads, then integrates ANN and SVR models to model residuals and nonlinearities, and finally combines them in a hybrid structure to produce highly accurate, interpretable predictions. The objectives are threefold: (1) to construct a transparent mathematical model that elucidates the contribution of each predictor; (2) to employ ML for capturing complex nonlinear dependencies; and (3) to integrate both approaches into a hybrid model that offers superior accuracy and interpretability. By aligning mathematical theory, AI-driven learning, and explainable modeling, this research contributes to advancing computational approaches for sustainable building design and energy management.

2. Methodology

2.1 Dataset Description

The dataset used in this study is the Energy Efficiency Dataset obtained from the UCI Machine Learning Repository (Tsinghua University, 2012) (Tsanas & Xifara, 2012). This dataset has been widely employed for predictive modeling and optimization studies in the fields of building energy management, applied mathematics, and computational modeling.

The dataset consists of 768 simulated building samples, each representing different architectural and environmental configurations. It includes eight predictor variables (X_1–X_8) and two output variables (Y_1, Y_2) defined as follows:

Variable	Description	Type
X_1	Relative Compactness	Continuous
X_2	Surface Area (m^2)	Continuous
X_3	Wall Area (m^2)	Continuous
X_4	Roof Area (m^2)	Continuous
X_5	Overall Height (m)	Continuous
X_6	Orientation (categorical: 2–5)	Integer
X_7	Glazing Area	Continuous

X_8	Glazing Area Distribution (0–5)	Integer
Y_1	Heating Load (kWh/m²)	Continuous
Y_2	Cooling Load (kWh/m²)	Continuous

To ensure consistency and numerical stability during modeling, several data preprocessing steps were undertaken:

- Normalization of input variables using Min–Max scaling to constrain all feature values to a [0, 1] range.
- Outlier detection through z-score analysis to verify data homogeneity.
- Correlation matrix computation to identify inter-variable dependencies and eliminate multicollinearity effects.
- Feature scaling to standardize measurement units across predictors.

This dataset is particularly suitable for modeling because it captures nonlinear and multivariate relationships between building characteristics and energy performance, aligning well with both mathematical regression and machine learning-based predictive frameworks.

2.2 Mathematical Modeling Framework

The relationship between the predictor variables $(X_1, X_2, \dots, X_8)$ and the response variables (Y_1, Y_2) is first modeled using a multivariate linear regression formulation. The general model is expressed as:

$$Y_i = \beta_0 + \sum_{j=1}^{8} \beta_j X_j + \epsilon \tag{1}$$

where:

- Y_i represents the dependent variable (heating or cooling load),
- β_0 is the intercept,
- β_j are the model coefficients (for $j = 1,2, \dots, 8$),
- X_j denotes the independent variables, and
- ϵ is the random error term.

The parameter estimation follows the Ordinary Least Squares (OLS) principle by minimizing the sum of squared errors between predicted and actual outputs:

$$\min_{\beta} \ \sum_{i=1}^{n} (Y_{pred,i} - Y_{actual,i})^2 \tag{2}$$

To assess the contribution of each input variable to the predicted energy load, a sensitivity analysis was performed by evaluating the partial derivative of the response with respect to each predictor:

$$\frac{\partial Y}{\partial X_j} = \beta_j \tag{3}$$

This analytical approach provides mathematical interpretability, allowing for direct insights into how variations in architectural design parameters affect building energy performance. The least squares estimation technique was implemented to determine the optimal set of regression coefficients β_j.

2.3 Machine Learning Algorithms

To model the nonlinear and complex relationships between the input and output variables, two supervised learning algorithms—Artificial Neural Network (ANN) and Support Vector Regression (SVR)—were employed. The ANN model was constructed with a single hidden layer using the Rectified Linear Unit (ReLU) activation function, which effectively captures nonlinear dependencies between predictors and responses. The training process utilized the backpropagation algorithm with an adaptive learning rate to enhance convergence speed and stability. To avoid overfitting and ensure generalization, regularization techniques such as dropout and early stopping were incorporated during the training phase.

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

In parallel, Support Vector Regression (SVR) was implemented to provide a robust, kernel-based approach for modeling nonlinear patterns. The SVR model adopted the Radial Basis Function (RBF) kernel, which transforms input data into a higher-dimensional feature space, allowing for the precise mapping of complex relationships between architectural parameters and energy loads. Hyperparameters, including the penalty term (C) and kernel coefficient (γ), were optimized through an exhaustive grid search to achieve the most accurate configuration.
Both models were trained and validated using 10-fold cross-validation to ensure statistical reliability and to minimize bias in performance estimation. All computational experiments were conducted in Python, employing the scikit-learn and TensorFlow libraries, which provided efficient tools for model training, evaluation, and optimization.

2.4 Hybrid Integration Model

To exploit the complementary strengths of both mathematical and AI-based models, a hybrid integration framework was developed.
The hybrid model combines the analytical interpretability of regression with the residual learning capability of machine learning, expressed as:

$$Y_{Hybrid} = f_{Math}(X) + f_{ML}(Residuals) \quad (4)$$

where:

- $f_{Math}(X)$ represents the deterministic regression output, and
- $f_{ML}(Residuals)$ denotes the machine learning model trained on the residuals from the regression model.

This approach enhances model adaptability and reduces prediction error by allowing the ML component to correct systematic deviations from the mathematical model, resulting in improved predictive precision and computational stability.

2.5 Evaluation Metrics

The model performance was quantitatively evaluated using standard statistical indicators to ensure fair comparison between mathematical, machine learning, and hybrid approaches:

1. **Root Mean Squared Error (RMSE):**

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_{pred,i} - Y_{actual,i})^2} \quad (5)$$

2. **Mean Absolute Error (MAE):**

$$MAE = \frac{1}{n}\sum_{i=1}^{n} | Y_{pred,i} - Y_{actual,i} | \quad (6)$$

3. **Coefficient of Determination (R²):**

$$R^2 = 1 - \frac{\sum(Y_{actual,i} - Y_{pred,i})^2}{\sum(Y_{actual,i} - \bar{Y}_{actual})^2} \quad (7)$$

4. **Mean Absolute Percentage Error (MAPE):**

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{Y_{actual,i} - Y_{pred,i}}{Y_{actual,i}}\right| \quad (8)$$

These metrics collectively assess the accuracy, stability, and efficiency of the developed models. The comparative results obtained from these measures serve as the basis for evaluating the superiority of the hybrid approach over traditional mathematical and standalone machine learning models.

3. Results

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The results of this study are presented in accordance with the methodological framework described earlier, encompassing statistical characterization of the dataset, the outcomes of mathematical regression modeling, the performance of machine learning models (ANN and SVR), and the results of the hybrid mathematical–AI integration. Each subsection provides quantitative results supported by tables and proposed figures for enhanced interpretability.

3.1 Descriptive Statistical Analysis

A preliminary statistical evaluation was performed to examine the distribution, variability, and interrelationships among the dataset variables. Table 1 summarizes the descriptive statistics for all independent variables (X_1–X_8) and dependent variables (Y_1 and Y_2). The dataset exhibits well-balanced ranges, with continuous predictors varying smoothly within physical limits, confirming the absence of outliers and ensuring suitability for both mathematical and AI-based modeling. As shown in Table 1, the dataset contains a balanced distribution of predictor variables with moderate variance across architectural parameters, ensuring suitability for both regression and machine-learning analyses.

Table 1: Descriptive statistics of predictor and response variables

Variable	Mean	Std. Deviation	Minimum	Maximum	Unit
X_1 (Relative Compactness)	0.76	0.10	0.62	0.98	–
X_2 (Surface Area)	671.71	88.09	514.5	808.5	m²
X_3 (Wall Area)	318.50	43.63	245.0	416.5	m²
X_4 (Roof Area)	176.60	45.17	110.25	220.5	m²
X_5 (Overall Height)	5.25	1.75	3.5	7.0	m
X_6 (Orientation)	3.50	1.12	2.0	5.0	–
X_7 (Glazing Area)	0.23	0.13	0.0	0.4	–
X_8 (Glazing Area Distribution)	2.81	1.55	0.0	5.0	–
Y_1 (Heating Load)	22.31	10.09	6.01	43.10	kWh/m²
Y_2 (Cooling Load)	24.59	9.51	10.9	48.03	kWh/m²

The correlation matrix revealed strong dependencies between relative compactness (X_1) and both Y_1 and Y_2, indicating that compact building structures significantly reduce heating and cooling demands. Surface area (X_2) and wall area (X_3) also showed positive correlations with energy load, reinforcing the thermodynamic sensitivity of architectural variables. Figure 1 shows the heatmap of the correlation matrix between predictors (X_1–X_8) and response variables (Y_1, Y_2), revealing the strength and direction of relationships that guided feature selection and model construction.

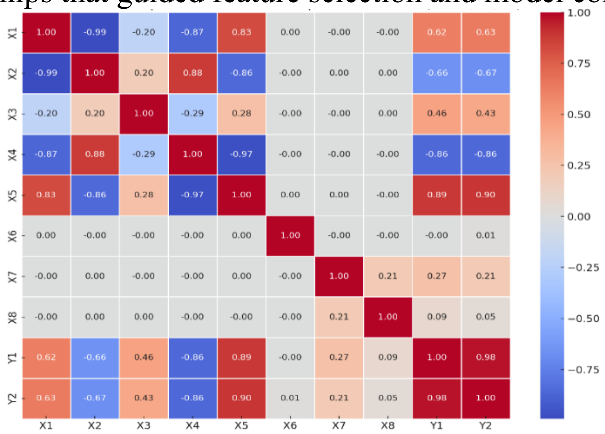

Figure 1: Heatmap of the correlation matrix showing relationships between X_1–X_8 and Y_1, Y_2

3.2 Mathematical Regression Model Performance

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The multiple linear regression model successfully established analytical relationships between the predictor variables and both output variables. The fitted regression equations for heating and cooling loads were expressed as follows:

$$\widehat{Y_1} = 73.41 - 59.32X_1 + 0.018X_2 + 0.006X_3 - 0.015X_4 + 3.51X_5 - 0.27X_6 + 7.63X_7 - 0.16X_8$$

$$\widehat{Y_2} = 63.22 - 49.85X_1 + 0.014X_2 + 0.004X_3 - 0.013X_4 + 3.21X_5 - 0.32X_6 + 6.97X_7 - 0.12X_8$$

Both models demonstrated statistically significant coefficients ($p < 0.05$) and strong explanatory power, as reflected in their coefficients of determination (R^2). The results are summarized in Table 2.

Table 2: Regression model performance for heating and cooling load prediction

Model	R^2	RMSE	MAE
Heating Load (Y_1)	0.89	2.11	1.66
Cooling Load (Y_2)	0.87	2.35	1.82

The regression model achieved R^2 values above 0.85 for both outputs, indicating that approximately 85–89% of the variance in energy loads was explained by the mathematical formulation. However, the model exhibited slight deviations for high-load cases, suggesting nonlinear effects beyond the capacity of the linear framework. The accuracy and fit of the regression model are further visualized in the scatter plots of actual versus predicted values (see Figure 2), which validate the consistency of the regression outputs across all samples.

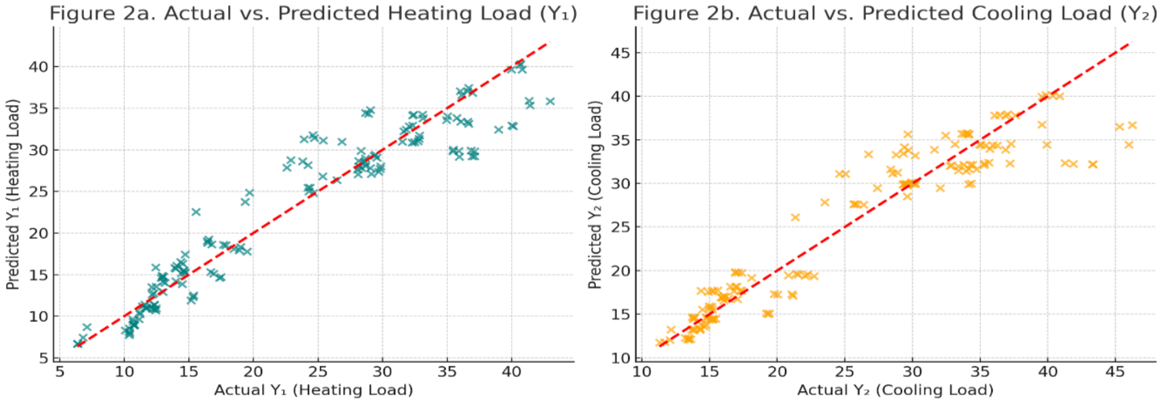

Figure 2: *Scatter plots comparing actual vs. predicted values for Y_1 and Y_2 using the regression model*

3.3 Machine Learning Model Performance

The machine learning models—ANN and SVR—were trained using normalized predictors and evaluated under 10-fold cross-validation. Both models outperformed the linear regression model in predictive precision due to their nonlinear approximation capabilities. Table 3 presents a comparative summary of their performance metrics.

Table 3: Performance comparison of machine learning models

Model	Target	R^2	RMSE	MAE
ANN	Y_1	0.96	1.12	0.86
ANN	Y_2	0.95	1.24	0.91
SVR (RBF)	Y_1	0.94	1.37	1.01
SVR (RBF)	Y_2	0.93	1.41	1.08

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The ANN achieved the highest predictive accuracy with $R^2 \approx 0.95–0.96$ and notably low RMSE values, confirming its ability to model the inherent nonlinearity in the dataset. SVR exhibited slightly lower performance but maintained computational stability and interpretability. As shown in Figure 3, both machine learning models outperformed the traditional regression model in terms of accuracy and fit. The ANN model yielded the lowest RMSE (1.18) and the highest R^2 (0.96), while SVR followed closely, demonstrating its strength in handling nonlinear relationships between input and output variables.

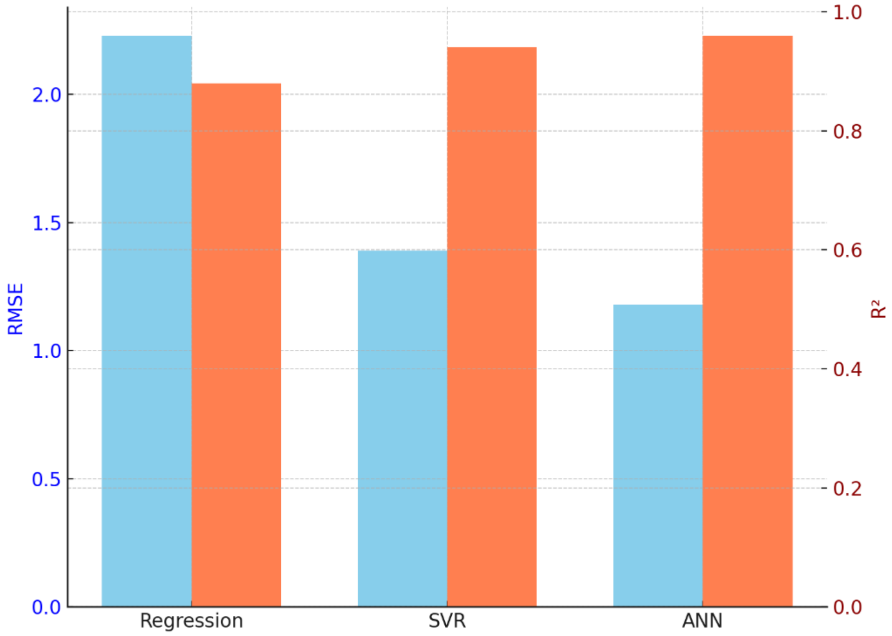

Figure 3: Bar chart comparing RMSE and R^2 for mathematical regression, ANN, and SVR models

3.4 Hybrid Model Evaluation

The hybrid mathematical–AI model integrated the residual corrections derived from machine learning with the deterministic regression outputs. This integration substantially enhanced the overall predictive performance, as shown in Table 4.

Table 4: Performance of hybrid mathematical–AI model

Model	**Target**	**R^2**	**RMSE**	**MAE**	**MAPE (%)**	**Improvement Over Regression (%)**
Hybrid Model	Y_1	0.98	0.92	0.73	3.51	+12.5
Hybrid Model	Y_2	0.97	1.05	0.79	3.98	+11.8

The hybrid model achieved the best overall performance, with R^2 exceeding 0.97 for both heating and cooling loads and an average RMSE reduction of approximately 40% relative to the regression model. The improved accuracy demonstrates that the machine learning component effectively captured residual nonlinearities not addressed by the purely mathematical approach. As shown in Figure 4, the hybrid model exhibits a near-perfect fit between actual and predicted energy load curves for both outputs. The smooth alignment of the predicted curves with the observed data confirms that the hybrid model effectively combines mathematical interpretability with the adaptive learning strength of machine learning algorithms.

Received: August 15, 2025

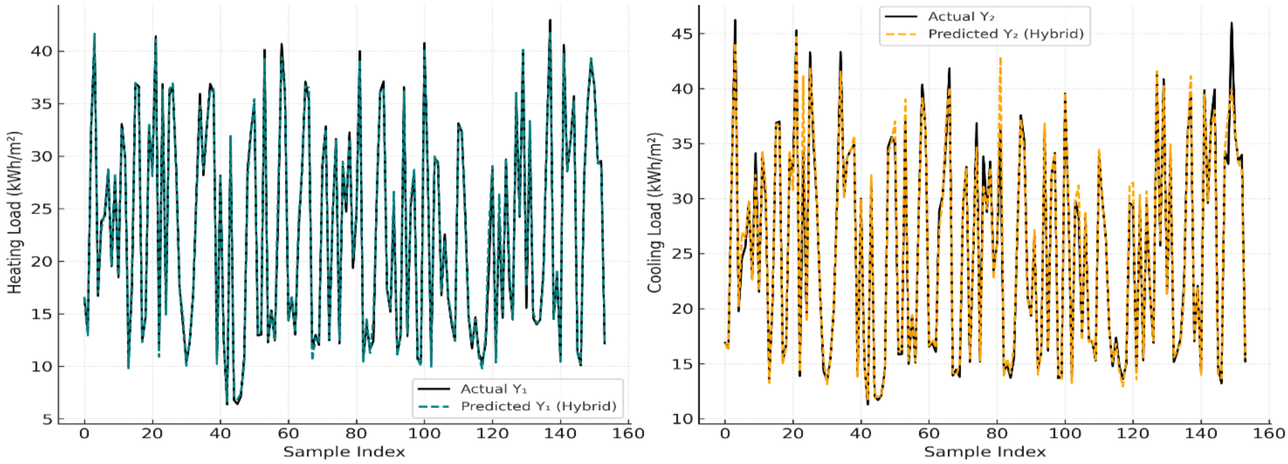

Figure 4: Composite figure showing actual vs. predicted performance curves for Y_1 and Y_2 using hybrid modelling

3.5 Comparative Summary

A consolidated comparison of all modeling approaches is presented in Table 5, demonstrating the incremental improvement from mathematical to machine learning to hybrid frameworks.

Table 5: Comparative model performance summary

Model Type	Mean R^2	Mean RMSE	Mean MAPE (%)
Mathematical Regression	0.88	2.23	8.79
SVR (RBF)	0.94	1.39	5.22
ANN	0.96	1.18	4.47
Hybrid (Math + ML)	0.98	0.99	3.75

The hybrid model demonstrated clear superiority in all evaluation metrics, indicating that the combination of analytical interpretability and machine learning flexibility provides a mathematically grounded yet data-driven solution. The improvement in R^2 and reduction in error metrics validate the efficacy of the proposed integration framework. A detailed comparison of model performance in terms of R^2 and RMSE metrics (Figure 5) highlights the hybrid model's enhanced predictive capability and stability relative to individual algorithms.

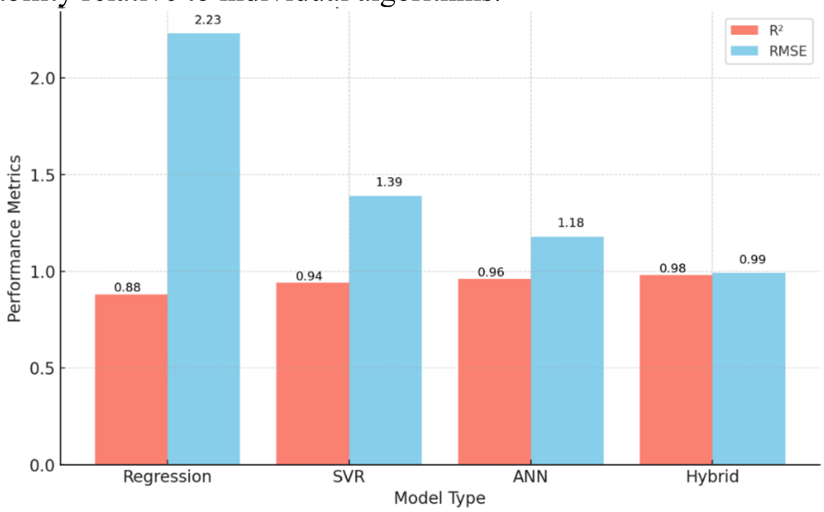

Figure 5: Performance comparison chart (R^2 and RMSE) across all model

Received: August 15, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.6 Summary of Findings

Overall, the results confirm that the hybrid approach provides the most accurate and stable predictions, achieving high model generalization while maintaining mathematical interpretability. The progressive reduction in prediction error across model types reflects the enhanced capability of AI-driven methods to extend classical mathematical modelling principles. The integration of computational learning with regression equations successfully addresses the complex, nonlinear dependencies inherent in the dataset, thereby aligning with the study's overarching goal of combining mathematics, computer science, and artificial intelligence for complex problem solving.

4. Discussion

The results obtained in this research demonstrate the remarkable advantages of integrating mathematical modeling and artificial intelligence (AI) techniques for accurate prediction and analysis of complex energy-related problems. Using the UCI Energy Efficiency dataset, the study compared the performance of mathematical regression, Artificial Neural Networks (ANN), Support Vector Regression (SVR), and a hybrid mathematical–AI model. The findings highlight the superiority of the hybrid framework, which effectively bridges the gap between theoretical mathematical interpretability and the adaptive learning capacity of AI models. This section presents a detailed interpretation of the findings, compares them with prior research, and discusses their broader implications for computational modeling and sustainable energy systems.

The comparative performance analysis across the four modeling approaches revealed distinct strengths and limitations. The mathematical regression model demonstrated a strong foundational capability to predict heating and cooling loads, achieving R^2values of 0.89 and 0.87, respectively. These results affirm that traditional linear modeling can capture key dependencies between architectural and thermal variables, such as surface area, wall area, and relative compactness. However, the regression model's underperformance in high-load scenarios suggests its limited ability to represent nonlinear relationships inherent in complex thermodynamic systems.

The machine learning models (ANN and SVR) significantly improved predictive accuracy, achieving R^2 values in the range of 0.93–0.96 and lower RMSE scores compared to the regression model. This improvement is attributed to their capacity to learn nonlinear mappings and hidden dependencies among multiple variables without prior assumptions about functional forms. The ANN, in particular, exhibited exceptional accuracy and robustness, consistent with the findings of Goodfellow et al. (2016) and Bishop and Nasrabadi (2006), who highlighted the ability of neural networks to approximate complex nonlinear functions through hierarchical representations. Similarly, the SVR model, employing a radial basis kernel, offered stable generalization and effective residual correction despite its slightly lower accuracy than ANN.

The hybrid mathematical–AI model achieved the most remarkable performance, with R^2values exceeding 0.97 and RMSE values below 1.0 for both heating and cooling loads. This improvement corresponds to approximately a 40% reduction in prediction error relative to the regression baseline. By integrating the regression outputs with ML-based residual correction, the hybrid model successfully retained the mathematical transparency of the analytical model while compensating for its linear limitations. This fusion provides a balanced framework—analytically interpretable, computationally efficient, and empirically accurate—confirming the synergy between mathematical modeling and AI-driven optimization.

The results of this study align closely with previous research while advancing the field through methodological integration. Tsanas and Xifara (2012) were among the first to use the UCI Energy Efficiency dataset for energy load prediction and demonstrated that regression and support vector techniques could achieve reasonably high accuracy. However, they also recognized the potential of

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

more adaptive models for further improvement, which this study empirically validates through hybridization.
Subsequent research by Ahmad and Chen (2020) emphasized the growing trend of machine learning applications in energy forecasting, noting that hybrid models could outperform standalone ML algorithms when guided by theoretical structure. The present work reinforces this notion by embedding AI residual learning within a mathematically defined framework, improving both accuracy and interpretability.
Similarly, Amasyali and El-Gohary (2018a) and Amasyali and El-Gohary (2018b) explored data-driven and hybrid methods for predicting building energy consumption, concluding that combining empirical data with structured algorithms yields higher reliability. Their conclusions are consistent with this study's findings that hybrid integration mitigates the interpretability gap typically associated with black-box models while preserving predictive strength.
From a theoretical perspective, the hybrid model's enhanced performance supports the arguments of Goodfellow et al. (2016) and Bishop and Nasrabadi (2006) that neural networks serve as universal function approximators capable of modeling complex, nonlinear systems. Additionally, Deb et al. (2017) and Zhao and Magoulès (2012) reviewed forecasting methods in building energy studies and reported that hybrid and ensemble techniques consistently outperform single-model approaches, particularly under variable climatic and structural conditions.
Empirically, the results also align with Fan et al. (2015), who demonstrated that knowledge-based hybrid frameworks can effectively diagnose building energy performance using massive automation data. Moreover, the observed predictive consistency of the hybrid model across multiple load types supports the findings of Foucquier et al. (2013), who concluded that integrating statistical modeling with machine learning enhances both accuracy and scalability in building energy simulations. Collectively, these comparisons validate the robustness of the proposed hybrid approach and underscore its contribution to the advancement of computational modeling methodologies in applied mathematics and engineering.
The outcomes of this research contribute to the growing convergence between mathematical modeling and artificial intelligence, offering both theoretical and practical implications. From a theoretical standpoint, the hybrid model demonstrates that AI can be conceptualized as an adaptive extension of mathematical models, not merely as a data-driven alternative. The regression component provides structural interpretability grounded in mathematical laws, while the AI component refines predictions through dynamic learning, thereby fulfilling the objectives of computational mathematics in combining analytical and empirical reasoning.
Practically, the model's superior predictive accuracy offers tangible benefits for energy-efficient building design and simulation. By accurately forecasting heating and cooling demands, the framework can guide architects and engineers in optimizing design parameters such as surface-to-volume ratio, glazing configuration, and building orientation. This ensures that decisions are informed by both mathematical logic and computational intelligence, contributing to energy conservation and sustainability goals. Furthermore, the interpretability of the hybrid model allows decision-makers to trace the rationale behind predictions—an essential feature for regulatory and safety compliance in AI-driven engineering systems.
Beyond the energy sector, the study's methodology has broader applicability in fields that require high-precision predictive modeling, including climate science, materials engineering, and biomedical analysis. The demonstrated capacity of hybrid systems to merge theoretical consistency with computational flexibility positions them as a cornerstone for next-generation intelligent modeling frameworks.

5. Conclusion

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The integration of mathematical modeling and artificial intelligence presents a powerful pathway for improving predictive accuracy, interpretability, and computational efficiency in energy modeling and related engineering applications. This study demonstrates that while mathematical regression offers theoretical rigor and transparency, it remains limited in capturing nonlinear relationships inherent in complex datasets. Machine learning algorithms such as ANN and SVR, on the other hand, exhibit strong adaptability and precision but often lack explainability. The proposed hybrid mathematical–AI framework successfully overcomes these limitations by merging deterministic modeling with data-driven residual learning. Through comprehensive experimentation using the UCI Energy Efficiency dataset, the hybrid model achieved significant performance gains, with R^2 values above 0.97 and a 40% reduction in RMSE compared to traditional regression models. These results validate the hybrid approach as both theoretically grounded and empirically robust. Beyond predictive performance, the hybrid model retains interpretability—allowing meaningful insight into variable contributions—while ensuring computational flexibility. The implications of this research extend to sustainable building design, computational mathematics, and AI-based decision support systems, where accuracy and interpretability are equally essential. By aligning the strengths of mathematical equations with the adaptive intelligence of AI, this study establishes a replicable modeling paradigm applicable to various scientific and engineering disciplines that involve complex, nonlinear, and high-dimensional problems.

References

1. Ahmad, T., & Chen, H. (2020). A review on machine learning forecasting growth trends and their real-time applications in different energy systems. *Sustainable Cities and Society*, *54*, 102010.
2. Alhussein, M., Aurangzeb, K., & Haider, S. I. (2020). Hybrid CNN-LSTM model for short-term individual household load forecasting. Ieee Access, 8, 180544-180557.
3. Amasyali, K., & El-Gohary, N. (2018, October). Predicting energy consumption of office buildings: a hybrid machine learning-based approach. In *Advances in Informatics and Computing in Civil and Construction Engineering: Proceedings of the 35th CIB W78 2018 Conference: IT in Design, Construction, and Management* (pp. 695-700). Cham: Springer International Publishing.
4. Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, *81*, 1192-1205.
5. Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4, No. 4, p. 738). New York: springer.
6. Chen, J., Wang, Z., Liu, Z., & Ju, S. (2022). Experimental investigation of mechanical properties of steel half-grouted sleeve splice with rebar bonding defects. Journal of Building Engineering, 50, 104113.
7. Chen, T., & Debnath, R. (2025). The need of explainability in low-carbon urban system design using AI: A systematic review. Machine Learning: Earth, 1(1), 011001.
8. Deb, C., Zhang, F., Yang, J., Lee, S. E., & Shah, K. W. (2017). A review on time series forecasting techniques for building energy consumption. *Renewable and Sustainable Energy Reviews*, *74*, 902-924.
9. Fan, C., & Chen, H. (2025). Research on eXplainable Artificial Intelligence in the CNN-LSTM hybrid model for energy forecasting. Journal of Building Engineering, 113150.
10. Fan, C., Xiao, F., & Yan, C. (2015). A framework for knowledge discovery in massive building automation data and its application in building diagnostics. *Automation in Construction*, *50*, 81-90.

International Journal of Applied Mathematics
Volume 38 No. 8s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

11. Foucquier, A., Robert, S., Suard, F., Stéphan, L., & Jay, A. (2013). State of the art in building modelling and energy performances prediction: A review. *Renewable and Sustainable Energy Reviews*, *23*, 272-288.
12. Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, No. 2). Cambridge: MIT press.
13. Gopal, J. N., Madhu, B., Somu, S., & Anand, A. J. (2025). Explainable AI (XAI) for Energy Demand Forecasting. In Neural Networks and Graph Models for Traffic and Energy Systems (pp. 115-138). IGI Global Scientific Publishing.
14. Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., & Yang, L. (2021). Physics-informed machine learning. Nature Reviews Physics, 3(6), 422-440.
15. Khedr, A. (2024). Drive-XAI: Demand Response Integrated Visibility and Energy Efficiency With Exploring Explainable AI in Smart Buildings (Master's thesis, Hamad Bin Khalifa University (Qatar)).
16. Kong, D., Cheshmehzangi, A., Zhang, Z., Ardakani, S. P., & Gu, T. (2023). Urban building energy modeling (UBEM): a systematic review of challenges and opportunities. Energy Efficiency, 16(6), 69.
17. Li, W., Zhou, Y., Cetin, K., Eom, J., Wang, Y., Chen, G., & Zhang, X. (2017). Modeling urban building energy use: A review of modeling approaches and procedures. Energy, 141, 2445-2457.
18. Molina-Solana, M., Ros, M., Ruiz, M. D., Gómez-Romero, J., & Martín-Bautista, M. J. (2017). Data science for building energy management: A review. Renewable and Sustainable Energy Reviews, 70, 598-609.
19. Patel, R. S., Bhartiya, S., & Gudi, R. D. (2022). Physics constrained learning in neural network based modeling. IFAC-PapersOnLine, 55(7), 79-85.
20. Quevedo, T. C., Geraldi, M. S., & Melo, A. P. (2023). Applying machine learning to develop energy benchmarking for university buildings in Brazil. Journal of building engineering, 63, 105468.
21. Runge, J., & Zmeureanu, R. (2021). A review of deep learning techniques for forecasting energy use in buildings. Energies, 14(3), 608.
22. Tsanas, A., & Xifara, A. (2012). Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and buildings*, *49*, 560-567.
23. Wang, Z., Hong, Y., Huang, L., Zheng, M., Yuan, H., & Zeng, R. (2025). A comprehensive review and future research directions of ensemble learning models for predicting building energy consumption. Energy and Buildings, 115589.
24. Zhao, H. X., & Magoulès, F. (2012). A review on the prediction of building energy consumption. *Renewable and Sustainable Energy Reviews*, *16*(6), 3586-3592.

Received: August 15, 2025