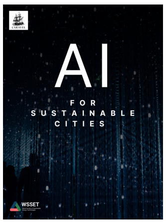
CARAVEL

Publish open access at Caravel Press

Artificial Intelligence for Sustainable Cities

Journal homepage: aiforsustainablecities.com



Technical article

From Experiment to Deployment: An Integrated Framework for Predictive Modelling and Visualisation of Thermoelectric HVAC Systems

Tajul Rosli Razak [1,2], Cagri Kutlu [3]*, Tianhong Zheng [1], Hasila Jarimi [1], Yuehong Su [1], Saffa Riffat [1], Preethi Jayakumar [4] and Dhruv Shah [4]

[1] Department of Architecture and Built Environment, The University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom.
[2] Faculty of Computer and Mathematical Science, Universiti Teknologi MARA (UiTM), Shah Alam, 40450, Malaysia.
[3] Department of Mechanical Engineering, Necmettin Erbakan University, Konya, Türkiye.
[4] Anzen Innovations Limited, NW10 6HJ, United Kingdom.

ABSTRACT

Thermoelectric (TE) heating, ventilation, and air-conditioning (HVAC) systems represent a sustainable alternative to conventional vapor-compression technologies; however, their adoption is limited by relatively low coefficients of performance (COP) and challenges in performance optimization. This study presents an integrated experimental, modelling, and deployment framework for predicting and evaluating TE HVAC performance. A custom-built test rig was developed to generate experimental data under controlled laboratory conditions, capturing key thermodynamic variables. A total of 487 samples were used to benchmark multiple regression and machine learning models. Among the evaluated approaches, Linear Regression achieved the best performance, with a root mean squared error (RMSE) of 3.28 and a coefficient of determination (R^2) of 0.885, outperforming more complex models. Beyond statistical evaluation, model predictions were validated through recomputation of heat transfer rate and COP, demonstrating strong thermodynamic consistency and physical interpretability. The validated model was subsequently implemented within a Shiny-based graphical interface, enabling interactive scenario analysis and real-time performance evaluation. The results demonstrate that, for moderate-scale experimental datasets, simple and interpretable models can outperform more complex approaches while maintaining physical consistency. The proposed framework provides a practical pathway for integrating experimental data, predictive modelling, and decision-support tools in thermoelectric HVAC applications.

ARTICLE INFO

Article History:
Received: 05 November 2025
Revised: 15 April 2026
Accepted: 04 May 2026

Keywords:
Thermoelectric HVAC
Artificial Neural Networks
Performance Prediction
Explainable AI
Sustainable Building Systems

Article Citation:
Razak, T. R., Kutlu, C., Zheng, T., Jarimi, H., Su, Y., Riffat, S., … Shah, D. (2026). From Experiment to Deployment: An Integrated Framework for Predictive Modelling and Visualisation of Thermoelectric HVAC Systems. Artificial Intelligence for Sustainable Cities, 1(1), 119–135. https://doi.org/10.65582/aifsc.2026.008

1. INTRODUCTION

Heating, ventilation, and air-conditioning (HVAC) systems are primary contributors to global building energy demand and greenhouse gas emissions, accounting for approximately 40% to 50% of total energy usage in residential and commercial sectors (Alghamdi and Krarti 2025). As global regulatory frameworks like the

* Corresponding author. Email address: cagri.kutlu@erbakan.edu.tr (Cagri Kutlu)

American Innovation and Manufacturing (AIM) Act mandate a transition toward low-global warming potential (GWP) refrigerants by 2026, the industry is increasingly prioritizing electrification and "refrigerant-free" technologies. In this context, thermoelectric (TE) HVAC systems, operating on the solid-state Peltier effect, have emerged as a promising alternative due to their compact modularity, lack of moving parts, and seamless integration into building envelopes (Adibhesami and Hassanzadeh 2025; Zhou 2026). Despite these advantages, their widespread adoption remains limited by several fundamental challenges.

The primary limitation of TE systems is their relatively low coefficient of performance (COP) compared to conventional vapor-compression technologies (Ibañez-Puy et al. 2017). This performance gap is largely attributed to parasitic thermal resistances, insufficient heat dissipation at the hot side, and intrinsic material limitations in thermoelectric conversion efficiency (Bahk et al. 2025; Ma et al. 2023). In addition, system performance is highly sensitive to operating conditions such as electrical input, airflow rate, and temperature gradients, making effective control and optimization non-trivial (He, Zuazua-Ros and Martín-Gómez 2024). These factors collectively constrain the practical applicability of TE HVAC systems in real-world building environments. To address these limitations, previous studies have explored improvements in thermal management, including advanced heat sink design and the integration of phase change materials (PCMs) to stabilize thermal loads (Manikandan and Kaushik 2017; Duan et al. 2021).

Recent advancements in digital intelligence have introduced Artificial Intelligence (AI) and Machine Learning (ML) as transformative tools for optimizing HVAC operations. ML-driven strategies, including Reinforcement Learning and Neural Networks, have demonstrated the potential to reduce HVAC energy consumption through predictive control and real-time load forecasting (Li, Dai and Hu 2022). However, a critical gap remains in the literature: most studies are confined to simulation environments or lack an integrated framework that connects rigorous experimental metrology with practical deployment tools (Myat, Rahman and Akbar 2025). Therefore, artificial intelligence has emerged as a promising enabler to address challenges by providing accurate prediction, optimization, and interpretability of HVAC performance. Recent studies have reported significant progress using advanced machine learning techniques, such as Random Vector Functional Link (RVFL) networks optimized with metaheuristics (Dizaji, Jafarmadar and Khalilarya 2019), Tree-Structured Parzen Estimator deep neural networks (TPE-DNN) enhanced with generative adversarial network augmentation (Ni et al. 2025), and ensemble methods such as Extra Trees Regressor (ETR) and Voting Hybrid Regression (VHR) for energy consumption forecasting (Zaki et al. 2025). However, a key limitation in the current literature is that many studies rely heavily on simulated datasets or small-scale prototypes, with limited integration of experimentally validated thermoelectric system data (Nassif, 2014). Furthermore, there remains a lack of unified frameworks that connect experimental data acquisition, predictive modelling, and practical deployment tools. To address these gaps, this study proposes a comprehensive “Experiment-to-Deployment” framework for thermoelectric HVAC systems. The main contributions of this work are as follows:

- An experimental dataset for a thermoelectric HVAC system, generated under controlled laboratory conditions with full thermodynamic measurement coverage.
- A systematic comparison of multiple regression and machine learning models evaluated using both statistical performance metrics and thermodynamic consistency checks.
- A dual-validation approach, where model predictions are further assessed through recomputation of heat transfer rate and coefficient of performance, ensuring physical interpretability.
- The implementation of a Shiny-based graphical user interface (GUI) that integrates model predictions with real-time visualization and user interaction, demonstrating practical applicability.

Therefore, this study focuses on the structured integration of experimental data, regression analysis, and deployment into a user-oriented tool, providing a reproducible and practically relevant framework for thermoelectric HVAC systems.

2. Experimental Setup and Metrological Rigor

A custom-built experimental rig was developed to evaluate the thermal performance of the thermoelectric HVAC system under controlled laboratory conditions. The setup was designed to capture all relevant thermodynamic variables required for regression modelling, while ensuring systematic control and reproducibility of test conditions. Figures 1-3 present the setup, system components, and fabricated prototype.

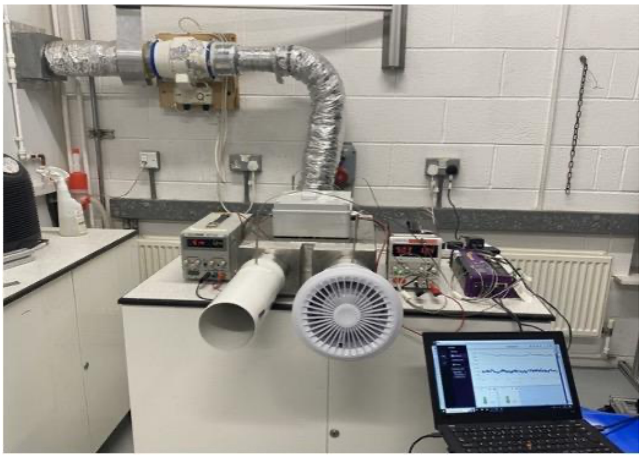


Figure 1. Overview of the thermoelectric HVAC test rig, showing environmental chamber, Thermoelectric HVAC unit, power supplies, and data logging setup.

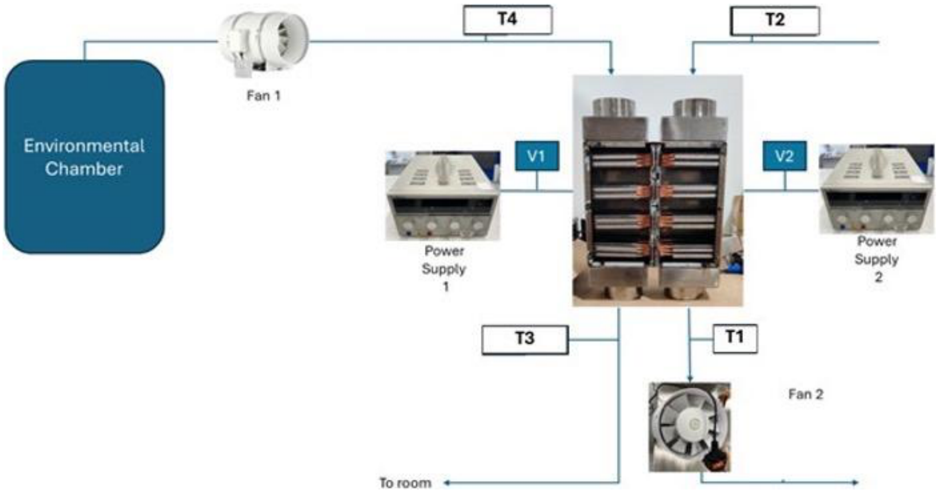


Figure 2. System layout and interconnections, showing environmental chamber, power supplies, thermocouples (T1–T4), and airflow pathway.

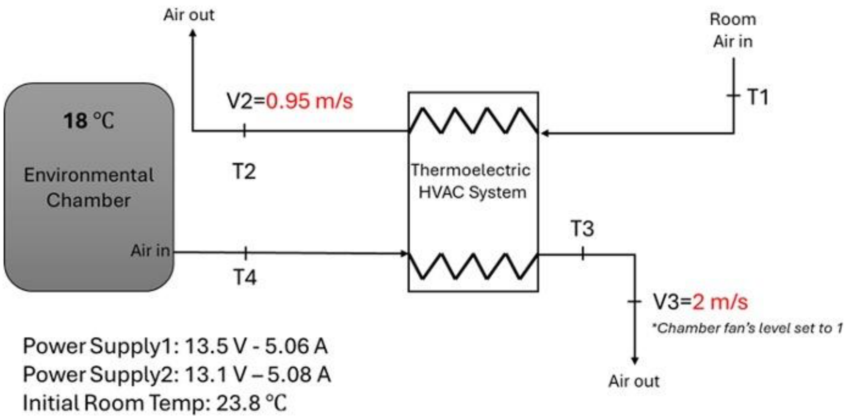


Figure 3. Architecture of the thermoelectric HVAC unit (stacked TEC modules with integrated heat sinks), showing a representative test case.

2.1. System Components and Instrumentation

The rig integrates several subsystems whose outputs directly map to the variables later used in the regression dataset:

- Environmental Chamber: A purpose-built room used to control operating conditions. In these laboratory tests, the chamber temperature and relative humidity were varied to emulate different ambient conditions and assess system performance.

- Thermoelectric HVAC Unit: The device under test consisted of 8 stacked TEC (thermoelectric cooler) modules with aluminium heat sinks and integrated heat pipes. The hot-side outlet air temperature was Hot_{exit} or $T_{h,out}$ while the cold side exit temperature was $Cold_{exit}$ or $T_{c,out}$.
- Temperature Measurement: Four type-K thermocouples (T1–T4) were installed at hot- and cold-side inlets and outlets. These sensors captured T_{room}, $T_{h,out}$, and $T_{c,out}$, while the cold-side setpoint was defined as T_{cold}, which comes from the environmental chamber.
- Airflow Measurement: Thermal Anemometers were used to measure airflow velocity on both sides, quantify airflow velocity on both hot and cold sides, and convert into volumetric flow rates (v_{hot}, v_{cold}). These values and air density yielded the mass flow rate ($\dot{m}$).
- Power Control: Two independent DC power supplies ($V1, V2$) delivered adjustable voltage and current to the TEC modules, with total electrical input power logged as $Power$.
- Fans and Variac: Fans induced airflow through the heat exchanger, while a variac enabled gradual adjustment of fan speeds, affecting v_{hot}, v_{cold} and $\dot{m}$.

2.2. Thermocouple Calibration and Metrology

To ensure measurement reliability, all thermocouples were calibrated prior to experimentation using a temperature calibrator (Sika TP 37 450 E.2) as shown in Figure 4. The calibration was performed over a temperature range of 0 °C to 100 °C, with increments of 5 °C. At each step, the thermocouples were allowed to stabilize for 5 minutes before recording measurements.

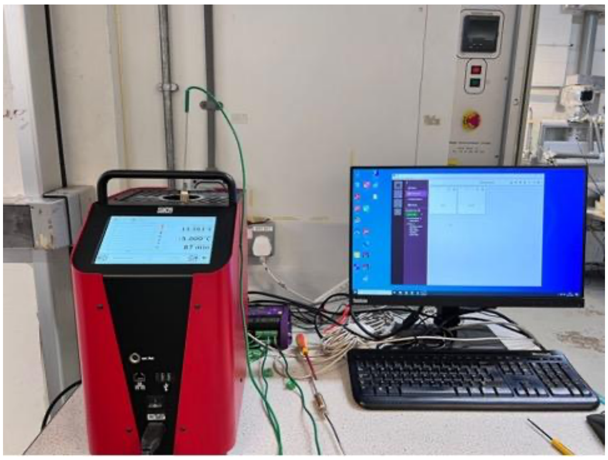


Figure 4. Thermocouple calibration procedure.

The calibration results indicated that all thermocouples exhibited good agreement with the reference sensor. The maximum absolute error was ±0.27 °C, while the mean absolute error across all calibration points was ±0.18 °C (See Table 1). The measured deviations remained within the acceptable tolerance for the present study, demonstrating that the installed thermocouples were functioning properly and that the resulting temperature data were sufficiently accurate for system performance analysis. In addition, repeated calibration checks confirmed that the sensors exhibited stable responses without noticeable drift. Based on these results, the thermocouple measurements were considered reliable for subsequent experimental testing.

Table 1. Thermocouple calibration results.

Sensor	Calibration range (°C)	Max Absolute error (°C)	Mean Absolute error (°C)	Status
TC1	0-100	0.21	0.15	Acceptable
TC2	0-100	0.27	0.21	Acceptable
TC3	0-100	0.23	0.19	Acceptable
TC4	0-100	0.25	0.17	Acceptable

2.3. Measurement Uncertainty and Error Propagation

An uncertainty analysis was conducted to quantify the reliability of both measured and derived parameters. Uncertainties of direct measurements were estimated based on instrument specifications and calibration results, while uncertainties of derived quantities were calculated using standard root-sum-square (RSS) error propagation.

The uncertainties associated with air velocity, temperature, electrical measurements, and geometric dimensions were propagated to determine the uncertainty ranges of the calculated thermal and electrical performance parameters. The resulting uncertainty ranges are summarised in Table 2, indicating that the experimental errors were within an acceptable range for system performance evaluation.

Table 2. Uncertainty ranges of measured and derived parameters in the experiments.

Parameter	Symbol	Unit	Method	Uncertainty range
Temperature	T	°C	Calibrated thermocouples	±0.3 °C
Air velocity	v	ms^{-1}	Anemometer	±0.05 ms^{-1}or ±3%
Voltage	V	V	data logger	±0.01 V
Current	I	A	data logger	±0.01 A
Volumetric flow rate	Q	$m^3 s^{-1}$	Calculated	±3–5%
Mass flow rate	$\dot{m}$	$kg\ s^{-1}$	Calculated	±4–6%

The uncertainty analysis enhances the transparency and reproducibility of the experimental methodology and provides a quantitative basis for evaluating the reliability of the reported results. Total expanded uncertainties for $\dot{Q}$ and COP were estimated to be in the ranges of 4.5% - 7.2% and 5.0% - 8.5%, respectively.

2.4. Dataset Formation

The experimental dataset consisted of direct measurements and derived quantities, structured as follows:

- Measured inputs: T_{room}, $T_{c,in}$, v_h, v_c, V, I
- Target output: $T_{h,out}$ (the primary regression variable).
- Derived indicators: Power, air flow rate ($\dot{m}$), heat transfer rate ($\dot{Q}$) and coefficient of performance (COP), computed using:

$$Power = V \cdot I \tag{1}$$

$$\dot{m} = \rho \cdot v \cdot A \tag{2}$$

$$\dot{Q} = \dot{m} \cdot cp \cdot (T_{h,out} - T_{room}) \tag{3}$$

$$COP = \frac{\dot{Q}}{Power} \tag{4}$$

$V\ and\ I$ represent the supplied voltage and current, respectively, cp is the specific heat capacity of air, $\dot{m}$ indicates air mass flow rate, v is measured air velocity and subscripts $h\ and\ c$ indicate hot and cold lines and A is the cross section area of the duct.

3. Regression Framework and Performance Evaluation for Thermal Energy Systems

To contextualize the methodology, Figure 5 provides an overview of the regression framework and evaluation workflow in this study. The framework illustrates the sequential process: beginning with dataset acquisition, moving through data preprocessing and regression modelling, and culminating in performance evaluation and thermodynamic validation. This holistic structure emphasizes the statistical rigor of model comparison, and the physical consistency checks that ensure practical applicability.

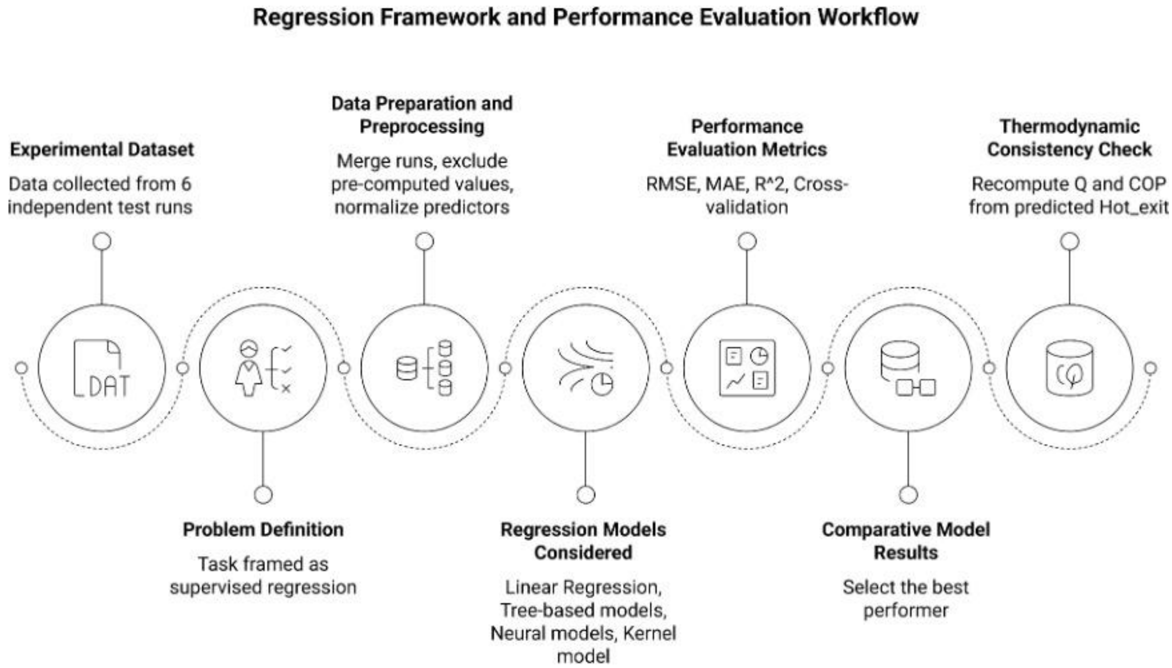


Figure 5. Overview of the regression framework and evaluation workflow for thermal energy systems. The process integrates experimental data acquisition, preprocessing, regression modelling, comparative performance evaluation, and thermodynamic validation.

3.1. Experimental Dataset

The experimental dataset was obtained from six independent thermal energy system test runs and merged into a dataset containing 487 records. Each record includes input variables, the target output, and derived thermodynamic performance indicators. In addition to the raw inputs and outputs, two derived quantities were calculated. The derived values were not used as predictors but were retained for secondary analysis of system performance. A dataset sample is shown in Figure 6. For each experimental run, the initial 10–20 minutes of operation were excluded from the dataset to remove start-up transients. During this period, the system undergoes rapid thermal adjustments, including heat accumulation within the thermoelectric modules, heat sinks, and ducting, as well as stabilization of airflow and electrical input conditions.

Time	Room_temp	Hot_exit	Tcold	vhot	vcold	power	m	Q	COP
03/31/25 10:4	19.86	24.47	8	3.5	1.2	425	0.026525	122.8923	0.289158
03/31/25 10:4	20.17	26.59	8	3.5	1.2	425	0.026525	171.1429	0.402689
03/31/25 10:4	20.78	28.41	8	3.5	1.2	425	0.026525	203.3988	0.478585
03/31/25 10:4	21.23	31.44	8	3.5	1.2	425	0.026525	272.1759	0.640414
03/31/25 10:4	20.92	31.63	8	3.5	1.2	425	0.026525	285.5048	0.671776
03/31/25 10:4	20.47	32.1	8	3.5	1.2	425	0.026525	310.0299	0.729482
03/31/25 10:4	19.81	33.4	8	3.5	1.2	425	0.026525	362.2792	0.852422
03/31/25 10:4	20.4	34.39	8	3.5	1.2	425	0.026525	372.9423	0.877511
03/31/25 10:4	20.11	35.23	8	3.5	1.2	425	0.026525	403.0656	0.94839
03/31/25 10:4	19.94	36.06	8	3.5	1.2	425	0.026525	429.7233	1.011114
03/31/25 10:4	20.07	37.19	8	3.5	1.2	425	0.026525	456.3811	1.073838
03/31/25 10:4	19.96	43.47	8	3.5	1.2	425	0.026525	626.7243	1.474645
03/31/25 10:4	20.36	48.9	8	3.5	1.2	425	0.026525	760.8129	1.790148
03/31/25 10:4	20.22	51.5	8	3.5	1.2	425	0.026525	833.8552	1.962012
03/31/25 10:4	20.55	52.76	8	3.5	1.2	425	0.026525	858.647	2.020346
03/31/25 10:4	20.11	53.22	8	3.5	1.2	425	0.026525	882.639	2.076798
03/31/25 10:4	20.42	56.68	8	3.5	1.2	425	0.026525	966.6109	2.274379
03/31/25 10:4	20.5	39.38	8	3.5	1.2	425	0.026525	503.2988	1.184232
03/31/25 10:5	20.21	50.62	8	3.5	1.2	425	0.026525	810.663	1.907442

Figure 6. Sample of the experimental dataset collected during system testing.

The present study is explicitly focused on quasi-steady-state operation, where system variables vary within a relatively stable range. This focus enables the development of regression models that capture the steady thermal behaviour of the system without being influenced by strong time-dependent effects.

While transient dynamics are physically important and can significantly influence system performance, particularly COP during start-up and load changes, they require dedicated time-dependent modelling approaches and higher-resolution temporal data. Such analysis is beyond the scope of the current study but represents an important direction for future research.

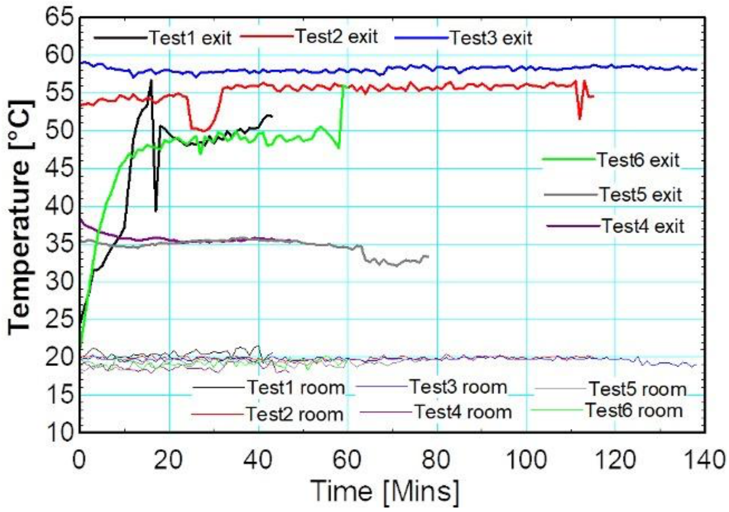


Figure 7. Temperature variation at the exit and room for six test conditions over time.

In order to assess performance under different operating conditions, six independent tests were conducted in the University of Nottingham Buildings, Energy and Environment Laboratory. Figure 7 shows the measured room temperatures and hot side exit temperatures. For each test, the initial 10–20 min of operation was discarded to allow the duct unit to warm up. Although the warm-up period is not, by itself, a reliable indicator of performance, retaining it would introduce strong transients into the dataset. We therefore report quasi-steady segments to characterise the unit near steady operation. As a result, the reported performance is slightly higher than a full-cycle average, but it remains representative of sustained heating operation. Ambient temperature was imposed via the environmental chamber set-point. The supplied cold-air temperature increased slightly before entering the unit due to duct heat gains; in practice, similar effects can occur depending on installation location and should be considered in performance evaluation.

In Test 1, the chamber was set to 8 °C and 80% RH (winter conditions), as shown in Figure 7. The system received 425 W electrical input, and the hot-side fan speed was adjusted to maintain an air velocity of 3.5 m s−1. The hot-side outlet reached around 50 ◦C, not suitable for direct room supply, but viable for mixing with cold outdoor air, and the measured COP ≈ 1.8. In Test 2, the chamber temperature was 10 °C and the hot-side fan speed was slightly reduced to 3.0 m s−1. Consequently, the hot-air supply temperature increased to 55 °C, and COP improved to 2.1.

In Test 3, the chamber conditions were the same as in Test 2, but the hot-side air velocity was further reduced to probe the unit's operating limits. The hot-air supply temperature rose to ≈ 58 °C, while COP declined to ≈ 1.9, indicating performance degradation as the hot-side temperature approached values that are sub-optimal, and close to the safe operating limits, for the TECs. In Tests 4–5, the electrical input was reduced to 170 W and the hot-side air velocity to 2.5 m s−1 to assess low-power operation. In Test 4, the hot supply temperature stabilised at around 35 °C, appropriate for direct space heating, with COP of 1.7. In Test 5, increasing the cold-side setpoint to 12 °C left the hot supply temperature essentially unchanged but yielded a slight COP improvement. In Test 6, the setup was returned to the Test 1 conditions to check repeatability; comparable results were obtained. Collectively, these six tests delineate a practical operating envelope for realistic building applications.

3.1.1. Problem Definition

The modelling task was framed as a supervised regression problem in which the goal is to predict the hot air exit temperature ($T_{h,out}$) from the measured input variables. Formally, the problem can be expressed as:

$$f(T_{room}, T_{cold}, v_{hot}, v_{cold}, Power) \to T_{h,out} \quad (5)$$

where $f\,(\cdot)$ denotes the regression function to be approximated. From the predicted $T_{h,out}$, the derived metrics $\dot{Q}$ and COP can subsequently be estimated, enabling the assessment of energy performance. The accurate modelling of Hot_{exit} is therefore crucial as it directly influences the evaluation of thermal energy efficiency.

3.2. *Data Preparation and Preprocessing*

Data preprocessing followed several systematic steps to ensure consistency and model readiness. First, the datasets from six independent experimental runs were merged into a unified structure, removing redundant time-stamp metadata while retaining the key thermodynamic variables. Second, pre-computed $\dot{Q}$ and COP values were excluded during training to prevent data leakage, as they are deterministic functions of $T_{h,out}$ and other inputs. Third, all input variables were inspected for zero variance, and those with constant values across all samples were removed. Finally, the input features were normalized to have a zero mean and unit variance, a crucial step, particularly for artificial neural networks (ANNs) and support vector machines (SVMs), which are sensitive to feature scaling. The dataset was then split into training (80%) and testing (20%) subsets using stratified sampling based on the Hot_{exit} variable to preserve the distribution of the output variable.

3.3. *Regression Models Considered*

To comprehensively evaluate predictive performance, a diverse set of regression models was implemented:

- Linear Regression (baseline): A simple yet interpretable model used as a benchmark for all subsequent machine learning approaches.
- Random Forest (RF): An ensemble method based on bagging and decision trees, well-suited for capturing non- linear interactions and providing feature importance measures.
- Extreme Gradient Boosting (XGBoost): A boosting algorithm optimized for accuracy and efficiency can handle complex, non-linear dependencies.
- Multilayer Perceptron (MLP, nnet): A shallow neural network implemented in R using the nnet package, representing a baseline ANN with limited hidden units.
- Neural Network (NeuralNet): A feedforward ANN implemented using the neuralnet package in R, with two hidden layers and tuned step size parameters.
- Bayesian Regularized Neural Network (BRNN): A neural model incorporating Bayesian regularization to reduce overfitting and improve generalization.
- Support Vector Regression (SVM): A kernel-based regression method using the radial basis function (RBF) kernel, effective for high-dimensional feature spaces.
- Decision Tree (DT): A regression tree implemented with the rpart engine, providing interpretable piecewise constant approximations of the data.

This ensemble of models was deliberately chosen to encompass both interpretable baselines and more flexible machine learning approaches, thereby enabling a robust performance comparison.

3.4. *Performance Evaluation Metrics*

Model evaluation was performed on the held-out test set using three standard regression performance metrics: root mean squared error (RMSE), mean absolute error (MAE), and the coefficient of determination (R^2). RMSE penalizes larger errors more heavily, thereby highlighting model robustness. MAE provides an easily interpretable measure of average prediction error, and R^2 quantifies the proportion of variance in Hot_{exit} explained by the model. Cross- validation was also conducted during training to assess model stability and mitigate overfitting.

The summary of model performances is presented in Table 3. Among all models tested, Linear Regression yielded the lowest RMSE on the test set, followed closely by Decision Trees and Random Forest, indicating that the relationship between inputs and Hot_{exit} was predominantly linear. More complex neural models (Neural Net and BRNN) underperformed, likely due to limited dataset size and convergence issues.

Table 3. Comparison of regression model performances on the test set. Best values for each metric are highlighted in bold.

Model	**RMSE**	**MAE**	R^2
Linear Regression	3.28	1.47	0.885
Random Forest (RF)	3.48	1.45	0.871
Extreme Gradient Boosting (XGB)	4.52	2.03	0.790
Multilayer Perceptron (MLP, nnet)	3.67	1.65	0.857
NeuralNet	49.9	49.1	0.529
Bayesian Regularized NN (BRNN)	49.9	49.1	0.529
Support Vector Regression (SVM)	3.51	1.34	0.864
Decision Tree (DT)	3.29	1.54	0.879

4. Comparative Model Results and Discussion

Table 3 compares all regression models on the held-out test dataset. Several important observations can be drawn. First, Linear Regression emerged as the best-performing model, achieving the lowest RMSE (3.28) and highest R^2 (0.885). This result indicates that the relationship between the selected input features and the hot air exit temperature is predominantly linear. The close alignment of predictions with the 45-degree line in Figure 8 further supports this conclusion. Despite its simplicity, the linear model generalised better than more complex machine learning approaches. Second, tree-based methods such as Decision Trees (RMSE = 3.29, R^2 = 0.879) and Random Forests (RMSE = 3.48, R^2 = 0.871) provided performance that was competitive with Linear Regression. These models can capture non-linearities and feature interactions, but their added complexity did not yield significant gains. XGBoost, which is often superior in high-dimensional or noisy datasets, performed relatively poorly (RMSE = 4.52, R^2 = 0.790). This suggests that boosting may have led to overfitting, given the modest dataset size (487 records).

Third, kernel-based regression with SVM achieved strong results, delivering the lowest MAE (1.34) while maintaining an RMSE of 3.51 and R^2 of 0.864. The low MAE highlights the SVM's ability to minimise average prediction error. However, its slightly higher RMSE compared to Linear Regression indicates a sensitivity to outliers or extreme cases in the dataset.

Fourth, the neural approaches (MLP, Neural Net, BRNN) underperformed relative to simpler models. The shallow MLP achieved acceptable performance (R^2 = 0.857), but the Neural Net and BRNN models failed catastrophically, producing RMSE values of nearly 50 and MAE values exceeding 49. This can be attributed to two main factors:

1. the relatively small dataset size, insufficient to support deep neural architectures, and
2. convergence and scaling issues that can arise when training neural networks in R without extensive hyperparameter optimisation. While neural methods are powerful in larger datasets, they offered no practical advantage in this experimental setting.

Table 4. Comparison between measured and recomputed thermal performance indicators using Linear Regression.

Indicator	**RMSE**	**MAE**
Heat Transfer Rate ($\dot{Q}$)	87.2	37.8
Coefficient of Performance (COP)	0.213	0.100

Overall, the results highlight the importance of model parsimony: more complex algorithms did not outperform simpler, interpretable approaches. For this dataset, the best trade-off between accuracy, robustness, and interpretability was provided by Linear Regression, closely followed by Decision Trees and Random Forests. This insight is particularly valuable for energy system applications, where interpretability and reliability are often as important as raw predictive accuracy.

Moreover, Figure 8 illustrates the predicted versus actual $T_{h,out}$ values using the best-performing model, Linear Regression. The predictions align closely with the 45-degree line, confirming the strong predictive accuracy of the linear model for this dataset.

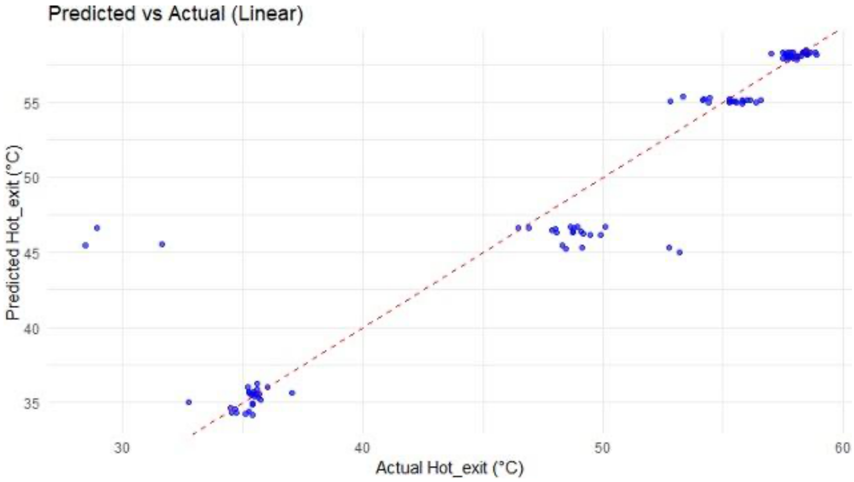


Figure 8. Predicted vs. actual $T_{h,out}$ using the best-performing model (Linear Regression).

The scatter plot in Figure 8 provides further evidence of the suitability of the linear regression model. The predicted $T_{h,out}$ values cluster tightly around the 45-degree reference line, indicating a strong agreement between predicted and actual measurements. The distribution shows the model can capture both low- and high-temperature regimes with relatively consistent accuracy. A few deviations are observed in intermediate ranges (approximately 45 °C to 50 °C), which may be attributed to system dynamics not fully captured by the selected input variables. Nevertheless, the overall alignment suggests that the underlying thermal process exhibits a predominantly linear relationship with the measured inputs. This reinforces the earlier conclusion that more complex non-linear models do not necessarily provide performance gains for this dataset.

4.1. Physical Interpretation of Model Linearity

The superior performance of Linear Regression can be interpreted in light of the underlying thermodynamic behaviour of the thermoelectric system. Under quasi-steady operating conditions, the dominant physical relationships governing system performance are inherently linear or near linear. In particular, the heat transfer rate is defined as a linear function of mass flow rate and temperature difference and the experimental conditions are controlled within a relatively narrow operating range, the resulting system response exhibits approximately linear behaviour with respect to key input variables.

Furthermore, the thermoelectric modules operate within a quasi-steady regime, where input electrical power and temperature gradients vary smoothly without strong nonlinear transients. While thermoelectric materials can exhibit nonlinear characteristics under extreme conditions, these effects are limited within the operating envelope considered in this study.

As a result, the mapping between input variables and the hot-side outlet temperature can be effectively approximated using linear relationships. This explains why Linear Regression not only achieves the lowest prediction error but also generalizes better than more complex models, which may overfit the limited dataset without capturing additional meaningful structure.

4.2. Thermodynamic Consistency: $\dot{Q}$ and COP Validation

To further evaluate the practical utility of the regression framework, the experimentally measured thermal performance indicators–namely, the heat transfer rate $(\dot{Q})$ and the coefficient of performance (COP)–were compared against values recomputed using the predicted $T_{h,out}$ from the best-performing model (linear regression).

Table 4 summarizes the error metrics between measured and recomputed indicators. The recomputed $\dot{Q}$ achieved an RMSE of 87.2 W and an MAE of 37.8 W, while COP achieved an RMSE of 0.213 and an MAE of 0.100. These low error magnitudes indicate that the regression-predicted $T_{h,out}$ can be reliably used to derive secondary performance quantities with minimal degradation in accuracy.

Figures 9 and 10 complement these numerical results by illustrating the temporal evolution of the measured and recomputed indicators. The line plots show that the recomputed $\dot{Q}$ and COP closely follow the experimental measurements across varying operating conditions, reinforcing the quantitative evidence from Table 4.

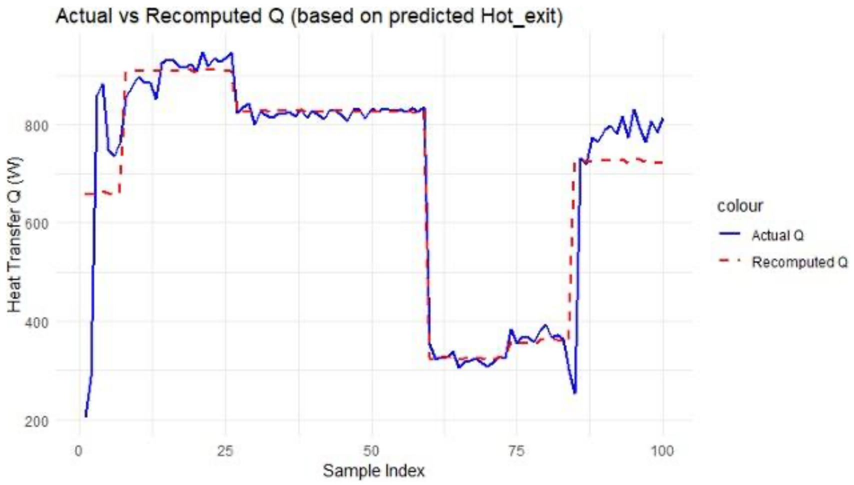


Figure 9. Comparison of experimentally measured and recomputed heat transfer rate using predicted $T_{h,out}$ from Linear Regression.

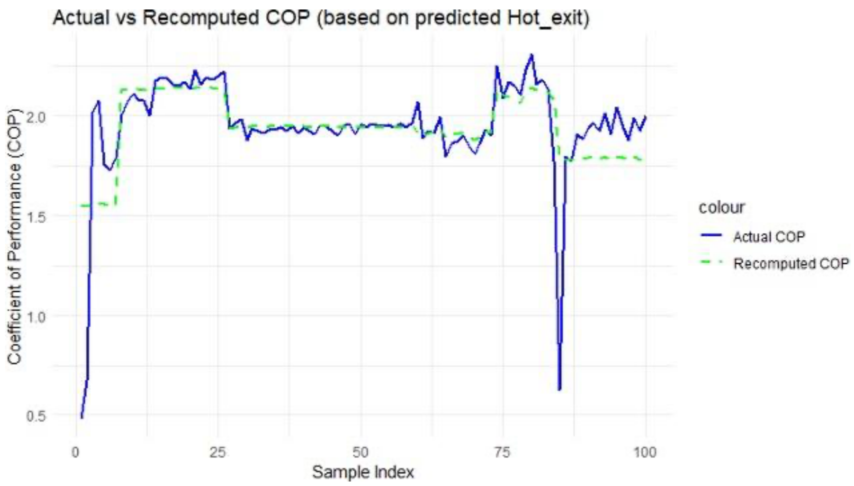


Figure 10. Comparison of experimentally measured and recomputed coefficient of performance using predicted $T_{h,out}$ from Linear Regression.

In Figure 9, the recomputed heat transfer rate tracks the experimental $\dot{Q}$ with high fidelity, accurately capturing steady-state regions and major operating transitions. Minor discrepancies appear during abrupt transients, where nonlinear system responses and measurement noise can lead to deviations. Despite these localized differences, the strong alignment confirms that the regression model preserves the thermodynamic integrity of the system.

A comparable trend is observed in Figure 10, where the recomputed COP mirrors the experimental profile. The predictions remain stable across most operating regimes, with only slight underestimations at peak performance intervals. This consistency is particularly significant, as it demonstrates that the regression approach provides both statistical accuracy and physically meaningful outcomes.

Taken together, the metrics in Table 4 and the visual trends in Figures 9 and 10 highlight the robustness of the linear regression model in supporting reliable thermal performance evaluation. This dual validation underscores the model's suitability for deployment in real-time monitoring and decision-support applications in building energy systems.

While the recomputed heat transfer rate and COP closely follow the experimental trends, noticeable deviations are observed during transient periods. These discrepancies can be attributed to several factors. First, the regression models are trained on quasi-steady-state data and therefore do not explicitly capture transient thermal dynamics, such as short-term heat accumulation within the heat sinks, thermoelectric modules, and ducting system. During rapid changes in operating conditions, these effects can lead to temporary mismatches between predicted and measured responses. Second, variations in fan speed and airflow conditions can introduce fluctuations in convective heat transfer, particularly during transitions between operating points. These variations may not be fully reflected in the input variables used for regression, contributing to localized prediction errors. Third, measurement noise and sensor response time, especially for thermocouples, can introduce small temporal delays and uncertainties in recorded temperatures, further contributing to deviations.

In addition, the calculation of heat transfer rate and COP involves multiple measured variables, and therefore uncertainties propagate through these calculations. In particular, small errors in temperature difference and mass flow rate directly influence the estimated heat transfer rate $(\dot{Q})$, which in turn affects the COP.

Despite these localized deviations, the overall agreement between measured and recomputed values confirms that the regression model provides a reliable approximation of system performance under quasi-steady conditions.

4.3. Prediction and Visualization via Shiny GUI

The regression model was implemented within a Shiny-based graphical interface to enable interactive evaluation of thermoelectric HVAC performance. Rather than serving as a standalone visualization tool, the interface functions as a decision-support platform, allowing users to explore system behaviour under varying operating conditions.

The GUI integrates model predictions with physics-based recomputation of key performance indicators, including heat transfer rate and COP. This ensures that predicted outputs remain physically interpretable and directly relevant to engineering analysis.

By enabling real-time adjustment of input variables within experimentally validated ranges, the interface supports scenario-based exploration, allowing users to assess how variations in operating conditions influence system performance. This capability is particularly valuable for identifying optimal operating regimes and understanding system sensitivity to key parameters.

Furthermore, the integration of predictive modelling with dynamic visualization facilitates rapid interpretation of system behaviour, supporting applications in system monitoring, design evaluation, and operator training.

4.4. Contribution and Positioning within the Literature

While previous studies have explored experimental thermoelectric systems, regression-based modelling, or user-oriented tools independently, the novelty of this work lies in their structured integration and validation within a single framework. The contributions of this study are not based on the introduction of fundamentally new algorithms, but rather on:

- The use of experimentally grounded data.
- The systematic comparison of multiple modelling approaches.
- The incorporation of thermodynamic consistency checks, and
- The translation of validated models into a practical, interactive interface.

4.5. Practical Implications

The proposed GUI demonstrates how data-driven models can be operationalized into system monitoring, operator training, and energy efficiency assessment tools by combining predictive analytics with interactive visualization. This dual emphasis on predictive accuracy and user-oriented design underscores the practical relevance of the work and its potential for adoption in building energy system management. The interface includes multiple visualization modes and logging functionality to support both instantaneous and temporal performance analysis (Figs. 11-13).

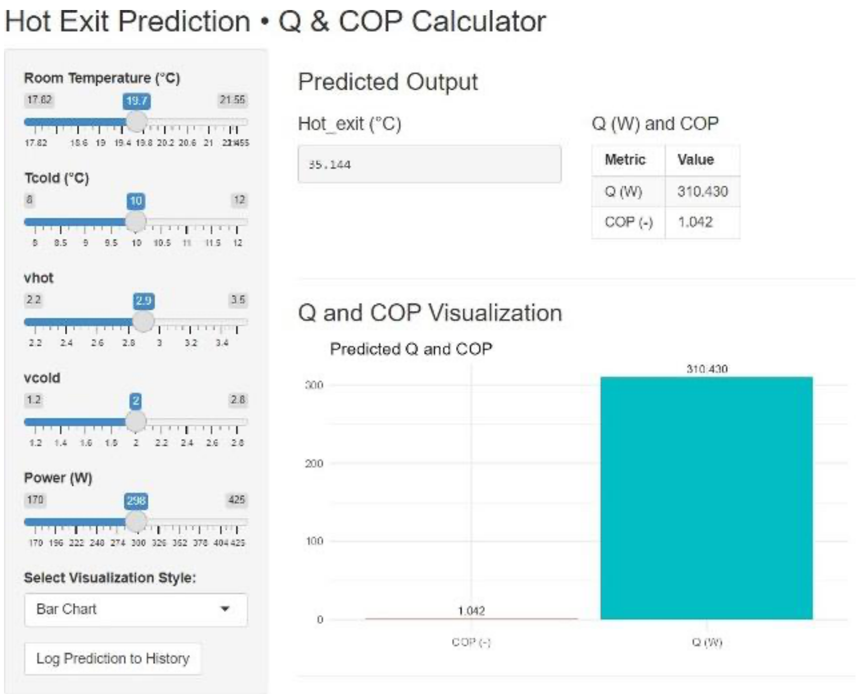


Figure 11. Shiny GUI with bar chart visualization of predicted $\dot{Q}$ and COP.

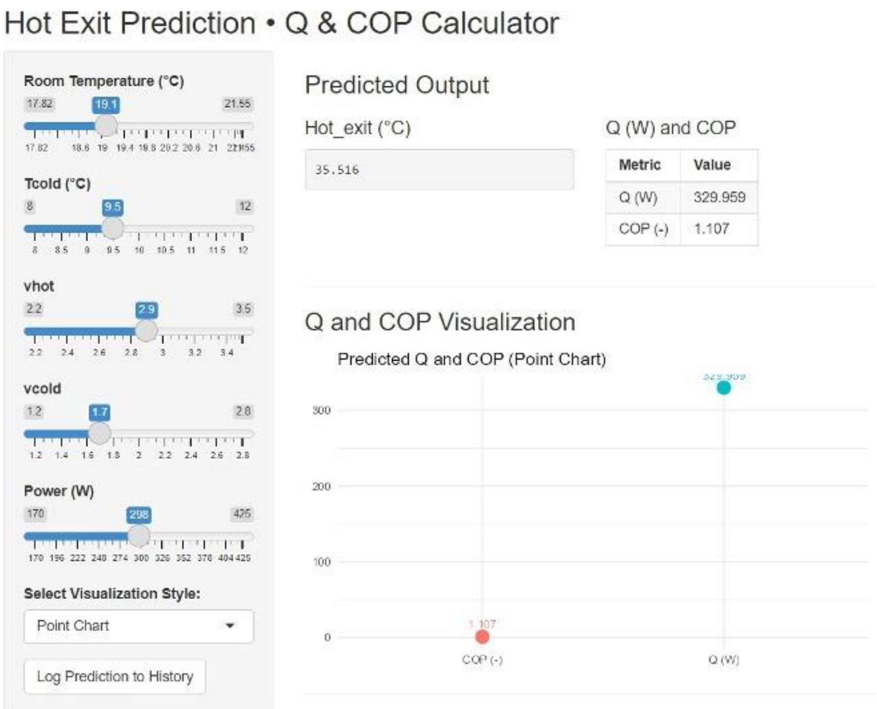


Figure 12. Shiny GUI with point chart visualization of predicted $\dot{Q}$ and COP.

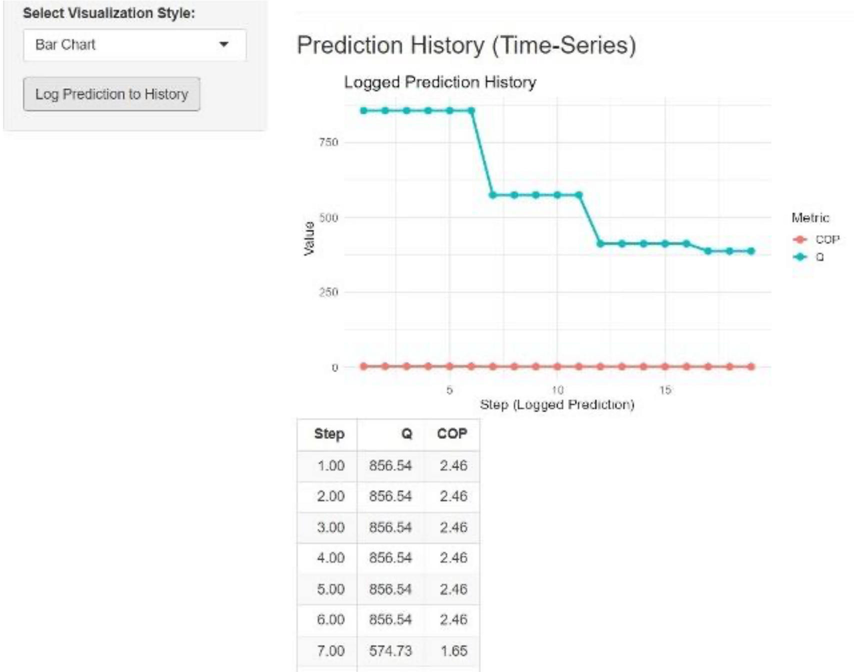


Figure 13. Shiny GUI with time-series logging and history table of predicted $\dot{Q}$ and COP.

The Shiny-based implementation encompasses the entire research workflow, from rigorous regression modelling and validation to real-time prediction, physical precomputation, and interactive visualization. By embedding the best- performing model into an accessible GUI, this work demonstrates a practical pathway for translating advanced data-driven models into operational tools. Such integration is particularly valuable in energy system applications, where decision-makers often require interpretable, physics-grounded, and user-friendly platforms rather than abstract model outputs. The proposed framework, therefore, not only advances predictive accuracy but also enhances transparency, usability, and applicability — bridging the gap between academic modelling and real-world deployment.

4.6. Discussions

This work demonstrates an integrated pathway from experimental data acquisition to model development and real-time deployment for thermoelectric HVAC systems. It should be noted that transient behaviour, including start-up dynamics and thermal inertia effects, is not captured in the present modelling framework and may contribute to additional variability in real-world operation. A key consideration in interpreting the modelling results is the size and structure of the experimental dataset. The dataset comprises 487 samples obtained from six steady-state test conditions, which represents a moderate-scale dataset in the context of machine learning applications. While sufficient for regression analysis and comparative benchmarking of interpretable models, this data volume is relatively limited for training more complex models such as deep neural networks. The observed underperformance of neural network-based models (NeuralNet and BRNN) can therefore be attributed to insufficient data to support parameter-rich architectures, as well as potential convergence and generalization challenges. In contrast, simpler models such as Linear Regression and tree-based methods demonstrated strong and stable performance, reflecting the structured nature of the dataset and the predominantly linear relationships between variables.

These results highlight the importance of aligning model complexity with dataset scale. For moderate-sized, experimentally derived datasets, simpler and more interpretable models may offer superior performance and robustness compared to more complex machine learning approaches. The discussion below highlights the methodological contributions, findings, and broader significance.

4.6.1. From Experimental Rigs to Usable Datasets

The custom-built rig enabled the systematic collection of high-quality data under controlled laboratory conditions. The dataset captured all relevant thermodynamic interactions by aligning measured variables with the modelling task. Direct measurements and derived indicators created a dual validation pathway, allowing regression models to be evaluated both statistically and through physical consistency checks. This ensures the modelling data foundation is reproducible and interpretable, a prerequisite for energy system applications.

4.6.2. Regression Modelling and Key Findings

A comprehensive suite of regression models was benchmarked, ranging from interpretable linear regression to complex ensemble and neural approaches. Despite the availability of sophisticated methods, linear regression consistently emerged as the best-performing model, achieving the lowest RMSE and the highest R^2 on the test set. This finding underscores the principle of model parsimony: in moderate-sized, structured experimental datasets, simpler models may outperform more complex approaches while retaining interpretability.

The robustness of the regression results was further confirmed through thermodynamic validation. Recomputed values of $\dot{Q}$ and COP based on model-predicted $T_{h,out}$ closely matched experimental measurements, with low RMSE and MAE values. These results strengthen confidence in the chosen regression approach and highlight that predictive accuracy was achieved without compromising physical plausibility. The dual validation, combining statistical and thermodynamic approaches, is a key methodological contribution of this study.

4.6.3. Bridging Modelling and Practice via Shiny GUI

A notable advancement of this work is the operationalization of the regression model through a Shiny-based GUI. While most prior studies stop reporting predictive accuracy, the proposed interface demonstrates how a validated model can be translated into a user-facing decision-support tool.

The GUI integrates three critical features:

- Real-time adjustment of input variables via sliders, constrained by experimental ranges to ensure physical validity.
- Dynamic prediction of Hot_{exit} and post-computation of $\dot{Q}$ and COP, ensuring that outputs remain grounded in thermodynamic principles.
- Multiple visualization modes (bar, point, and time-series logging) enhance user flexibility and enable instantaneous assessment and historical traceability.

This framework effectively bridges the gap between regression modelling and real-world usability, providing operators, engineers, and researchers with a transparent and interactive way to examine system behaviour.

4.6.4. Contributions and Broader Implications

The proposed framework makes several contributions:

- Methodologically, it combines controlled experimentation, robust regression analysis, and thermodynamic validation into an integrated framework.
- Empirically, it demonstrates that interpretable yet straightforward models can outperform more complex algorithms in predicting the performance of TE systems.
- Practically, it provides a deployable, user-friendly GUI that translates predictive modelling into actionable insights for decision-makers.

Beyond the immediate case of thermoelectric HVAC systems, this study illustrates a replicable template for other energy system applications where experimental data, predictive modelling, and operational usability must converge. By emphasizing accuracy, interpretability, and accessibility, the framework advances intelligent, transparent, and deployable solutions for sustainable building energy management.

5. Conclusion and Future Work

This study presents an integrated framework for predicting and operationalizing the performance of thermoelectric HVAC systems. The approach progressed systematically from controlled experimental data collection, through rigorous regression modelling and validation, to deployment of an interactive Shiny-based GUI.

The custom-built experimental rig accurately captured all relevant thermodynamic variables, providing a solid foundation for statistical modelling and thermodynamic validation. Comparative regression analysis revealed that Linear Regression, despite its simplicity, achieved the best balance between accuracy, robustness, and interpretability. The validation of derived thermal indicators, namely the heat transfer rate and coefficient of performance, further confirmed that predictive accuracy was achieved without compromising physical plausibility, strengthening confidence in the model's utility.

A central contribution of this work lies in translating predictive modelling into a user-facing decision-support tool. By embedding the best-performing regression model into an interactive Shiny application, predictions can be interrogated in real-time, performance indicators recomputed, and outputs visualized through multiple modes. This operationalization demonstrates how data-driven approaches can be transformed into practical systems that are interpretable, transparent, and directly actionable, bridging the gap between statistical modelling and engineering decision-making.

Looking ahead, several avenues for future research are evident. Expanding the dataset to capture a broader range of operating conditions and transient dynamics would improve model generalizability. Exploring hybrid modelling strategies that combine physics-informed formulations with machine learning could enhance the capture of nonlinear thermoelectric behaviours. Extending the Shiny-based implementation into a cloud-enabled or IoT-connected platform offers the potential for real-time monitoring and control in building-scale deployments. Finally, the proposed workflow could be adapted to other energy conversion and thermal management technologies, reinforcing its replicability and relevance to sustainable building systems.

In conclusion, this work demonstrates a clear pathway from laboratory experimentation to real-world applicability, advancing intelligent, transparent, and user-oriented solutions for sustainable HVAC systems. By integrating predictive accuracy, thermodynamic consistency, and practical deployment into a single framework, the study establishes a strong foundation for future innovations in building energy management. This framework demonstrates that simple yet interpretable models can unlock practical and trustworthy pathways for intelligent HVAC applications when coupled with interactive visualization. One limitation of this study is the relatively modest size of the experimental dataset, comprising 487 samples across six steady-state test conditions. While sufficient for regression modelling and comparative evaluation, this dataset size constrains the applicability of more complex machine learning models, particularly deep neural networks, which typically require larger and more diverse datasets. Future work will focus on expanding the dataset to include transient conditions and a wider range of operating scenarios to improve model generalizability and enable the effective application of advanced learning techniques.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support, and provision of resources from the Department of Architecture and Built Environment at the University of Nottingham, UK. The authors would also like to express their sincere appreciation to Anzen Innovations Limited for their valuable collaboration and technical support, which significantly contributed to the successful completion of this study.

AUTHOR CONTRIBUTIONS

Tajul Rosli Razak: Writing – Review & Editing, Finalization of Manuscript. **Cagri Kutlu**: Investigation, Methodology, Writing – Review & Editing. **Tianhong Zheng**: Investigation, Methodology, Writing – Review & Editing. **Hasila Jarimi**: Writing – Review & Editing. **Yuehong Su**: Supervision, Writing – Review & Editing. **Saffa Riffat**: Supervision, Writing – Review & Editing. **Preethi Jayakumar**: Collaborator, Writing – Review & Editing. **Dhruv Shah**: Collaborator, Writing – Review & Editing.

COMPETING INTERESTS

The authors have no competing interests to declare.

DATA ACCESSIBILITY

The raw experimental data are subject to proprietary restrictions and are not publicly available. Requests for access can be made to the corresponding author and are subject to approval by the contributing organization.

ETHICAL APPROVAL

The authors Prof. Saffa Riffat and Dr. Tajul Rosli Razak hold editorial positions at Artificial Intelligence for Sustainable Cities but were not involved in the editorial processing or peer-review evaluation of this manuscript. An independent editor was assigned to handle the review process to avoid any potential conflict of interest.

REFERENCES

Adibhesami, M. A. and Hassanzadeh, A., 2025. Optimizing HVAC energy efficiency in low-energy buildings: a comparative analysis of reinforcement learning control strategies under Tehran climate conditions, Data-Centric Engineering, 6. DOI: doi.org/10.1017/dce.2025.10014.

Alghamdi, F. and Krarti, M., 2025. Review Analysis for the Energy Performance of Integrated Air-Conditioning Systems, Energies, 18(7). DOI: doi.org/10.3390/en18071611.

Bahk, J. H. et al., 2025. Performance evaluation and optimization of air-to-air thermoelectric heat pump system, Energy and Buildings, 344(May), p. 115994. DOI: doi.org/10.1016/j.enbuild.2025.115994.

Dizaji, H. S., Jafarmadar, S. and Khalilarya, S., 2019. Novel experiments on COP improvement of thermoelectric air coolers, Energy Conversion and Management, 187(January), pp. 328–338. DOI: doi.org/10.1016/j.enconman.2019.03.025.

Duan, M. et al., 2021. Evaluation on the applicability of thermoelectric air cooling systems for buildings with thermoelectric material optimization, Energy, 221, p. 119723. DOI: doi.org/10.1016/j.energy.2020.119723.

He, Z., Zuazua-Ros, A. and Martín-Gómez, C., 2024. Current-dependent temperature change model of a thermoelectric window frame, Applied Thermal Engineering, 247(March), pp. 1–14. DOI: doi.org/10.1016/j.applthermaleng.2024.123081.

Ibañez-Puy, M. et al., 2017. Thermoelectric cooling heating unit performance under real conditions, Applied Energy, 200, pp. 303–314. DOI: doi.org/10.1016/j.apenergy.2017.05.020.

Li, M., Dai, L. and Hu, Y., 2022. Machine Learning for Harnessing Thermal Energy: From Materials Discovery to System Optimization, ACS Energy Letters, 7(10), pp. 3204–3226. DOI: doi.org/10.1021/acsenergylett.2c01836.

Ma, X. et al. 2023. Performance modelling on a thermoelectric air conditioning system using high power heat sinks and promoting waste heat utilization, Energy, 268(November 2022), p. 126612. DOI: doi.org/10.1016/j.energy.2023.126612.

Manikandan, S. and Kaushik, S. C., 2017. Transient Thermal Behavior of Annular Thermoelectric Cooling System, Journal of Electronic Materials, 46(5), pp. 2560–2569. DOI: doi.org/10.1007/s11664-016-5055-7.

Myat, A., Rahman, M. M. and Akbar, M., 2025. Nanofluid-Enhanced HVAC&R Systems (2015–2025): Experimental, Numerical, and AI-Driven Insights with a Strategic Roadmap, Sustainability (Switzerland), 17(16). DOI: doi.org/10.3390/su17167371.

Nassif, N., 2014. Modeling and optimization of HVAC systems using artificial neural network and genetic algorithm, Building Simulation, 7(3), pp. 237–245. DOI: doi.org/10.1007/s12273-013-0138-3.

Ni, H. P. et al., 2025. Enhancing HVAC energy efficiency modeling in semiconductor manufacturing facilities using tree-structured parzen estimator-optimized deep learning, Building and Environment, 271(January), p. 112589. DOI: doi.org/10.1016/j.buildenv.2025.112589.

Zaki, A. M. et al., 2025. Environmental and energy performance analyses of HVAC systems in office buildings using boosted ensembled regression trees: Machine learning strategy for energy saving of air conditioning and lighting facilities, Process Safety and Environmental Protection, 198(November 2024), p. 107214. DOI: doi.org/10.1016/j.psep.2025.107214.

Zhou, G., 2026. Thermoelectric coolers : Systematic Analysis on Principle , Use and Limitations. Atlantis Press International BV. DOI: doi.org/10.2991/978-94-6463-986-5.