Journal of Applied Engineering and Technological Science
Vol 7(2) 2026: 1289-1313

OPTIMIZING THERMAL LOAD IN COMPACT BUILDINGS: A COMPARATIVE ANALYSIS OF SINGLE AND HYBRID METAHEURISTICS FOR BALANCED HVAC EFFICIENCY

Usman Usman[1]*, Fatchul Arifin[2], Rustam Asnawi[3]
Doctoral Program in Engineering Science, Faculty of Engineering, Yogyakarta State University, Yogyakarta, Indonesia[1]
Department of Information Systems, Faculty of Engineering and Computer Science, Indragiri Islamic University, Tembilahan, Indonesia[1]
Department of Electronics and Informatics Engineering Education, Yogyakarta State University, Yogyakarta, Indonesia[2]
Department of Electrical Engineering Education, Yogyakarta State University, Yogyakarta, Indonesia[3]
usmanayyash.2024@student.uny.ac.id

Received: 19 April 2025, Revised: 11 December 2025, Accepted: 10 January 2026
**Corresponding Author*

ABSTRACT
This work is a comprehensive comparison study of five metaheuristic algorithms — namely GA, PSO, SA, GA-PSO, and GA-PSO-SA — using 3,000 model evaluations across 30 independent runs to optimize compact building thermal loads. Statistical analysis demonstrates that the SA, GA-PSO, and GA-PSO-SA present similar accuracy (RMSE: 3.55±0.16 kWh/m²) without any statistical difference (p> 0.97), contradicting the hypothesis that complexity in hybrid promotes the actual performance. GA appears to be the best compromise, with the highest efficiency ratio (11.47), providing 97.5% of SA's accuracy at 42% lower computational cost. The instability shown by PSO is quite alarming (CV: 14.215%, performance spread: 53.6%) and clearly indicates premature convergence, which sharply contradicts its claim of outperforming in continuous optimization [1]. Sensitivity analysis results show that envelope thermal properties, particularly the wall U-value (NSC=1.43), have 7.5× more influence on prediction performance than building orientation, providing evidence supporting the argument that input data quality outweighs algorithm choice for HVAC design-type applications.
Keywords: *Thermal load optimization, Genetic algorithm, Particle swarm optimization, Simulated annealing, Single and hybrid algorithms*

1. Introduction

HVAC systems account for 30-50% of a building's energy use and therefore play a significant role in the construction industry and the resulting carbon footprint (Aghili et al., 2025; Franco et al., 2021). The building sector itself accounts for roughly 30% of global energy use and 27% of energy-related emissions, and within this, HVAC systems play the most dominant and therefore impactful role. Economically, the systems have an issue: unexpected thermal load cycling causes under- or over-performance, leading to wasted investment and shallow cycling, resulting in constant energy waste. The issue is particularly dire in central urban areas, where land is constrained, and buildings are compact, with low volume-to-surface area ratios and high occupancy.(Lệ, Nguyen, Dou, et al., 2019). Distinct thermal issues in advanced compact buildings, unlike those in classical buildings, result in high operational efficiency and require complex thermal predictions. An example of this is a compact office building (400 to 600 m2) that accommodated 60 to 80 occupants, resulting in densities of 12 to 15 persons per m2. (Compared to classical buildings, where the occupancy is 5 to 8 persons per 100 m2) which results in internal heat gains of 25-35 W/m². (compared to 15 to 20 W/m2 in classic buildings).Such structures also exhibit double the surface-to-volume ratio (0.8-1.2 m^{-1}) of larger structures (0.4-0.6 m^{-1}), emphasizing envelope-driven thermal dynamics (Hong et al., 2020).

The peculiar thermophysical properties of these compact structures pose unique forecasting challenges. Internal heat loads from the compacting of structures and occupants are also exacerbated by the fact that employees and equipment tend to generate more heat per square foot in smaller workplaces. This calls for greater flexibility in the HVAC system in place due to the high temperature range it may encounter. High surface-to-volume ratios result in greater heat loss and gain, and, as a result, more accurate HVAC system sizing is needed, as oversizing and undersizing can lead to poor performance (Tribuiani et al., 2020). HVAC systems within these structures tend to respond more due to the low thermal mass of the smaller systems, and the spatial constraints require smaller systems. The lack of considerable spatial uniformity, predictability, and stability in thermal profiles due to the concentration of equipment and occupants makes these systems different from traditional systems and more complex (Lệ, Nguyen, Zhou, et al., 2019). This complexity requires the forecasting system to exhibit high multimodality and computational efficiency to overcome the extreme complexity of the systems.

Standard regression models, linear or nonlinear, rely entirely on static input-output relationships, which are untenable for compact building thermal systems that operate with dynamic heat transfer during occupancy and exhibit coupled nonlinear relationships (Lệ, Nguyen, Dou, et al., 2019). First-principles energy models require extensive input parameters that are often unavailable for most buildings or during conceptual design phases. Additionally, due to their long run times, these models are not applicable for real-time optimization or design space exploration. These models require and take hours to run per design scenario. In other domains, compact building design optimization can benefit from advanced machine learning. Still, most traditional techniques, such as ANN and SVM, that show strong predictive performance, are often stuck in local optima and, hence, intractable for complex multi-dimensional optimization problems. These approaches do not allow for sufficient exploration of the design space and, therefore, they are not predictive.In the case of compact buildings, the combination of the mentioned attributes, predictive performance, computational efficiency, and systemized optimization, remains untapped.

According to (Kareem et al., 2022), metaheuristic algorithms are crucial to solving complex nonlinear problems. There are several algorithms to choose from: Genetic Algorithms (GAs), which use evolutionary selection; Particle Swarm Optimization (PSO), which uses swarm intelligence and converges rapidly; and Simulated Annealing (SA), which uses probabilistic rules. While each of these algorithms can solve an optimization problem, they are fundamentally flawed. For example, in the GA, convergence to a good solution is slow. For the PSO, convergence to a poor solution in less time is a problem with multimodal functions. The SA requires parameter values to be tuned carefully (Khatiwoda et al., 2025; Razali et al., 2020) . Because of these flaws, there remains a classic trade-off: the user must choose between a more optimal solution and saving time in computation. Hybridization has been shown to overcome the limitations of these individual algorithms, but there is no systematic data to validate their use in building thermal applications.The question becomes: in a hybrid context, can one mitigate their individual heuristic weaknesses to achieve greater efficiency in predicting building thermal loads?

In the literature, several studies document the improved performance of hybrid systems that combine multiple meta-heuristics to address the individual weaknesses of algorithms within a multi-heuristic approach (Lệ, Nguyen, Zhou, et al., 2019; Mohamed & Abdelsalam, 2020). It was shown that the PSO-XGBoost model is highly accurate, achieving an R^2 value very close to 1, which demonstrates the value of combining metaheuristic optimization with machine learning (Lệ, Nguyen, Dou, et al., 2019). Model configurations with PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN, where hybrid ensembles were built, have shown to achieve R^2 values of 0.995 and relative errors of 11% and 2%, which are markedly better than the performance of stand-alone neural networks. The literature demonstrates an overwhelming advantage of hybrid PSO-GWO metaheuristics over others in solving complex benchmark functions (Mohamed & Abdelsalam, 2020). Hybrid algorithms have also been shown to achieve superior performance in IoT-edge load balancing (Kaushik & Al-Raweshidy, 2022). Several domains have been covered, which is valuable; however, there are still critical gaps, including the lack of a broad comparative analysis of single and hybrid metaheuristic algorithms in thermal load optimization for compact buildings, and the lack of substantial study on the trade-off between response time and accuracy that hybridization brings.

There is insufficient documentation to construct case studies that analyze the effectiveness of individual metaheuristic approaches (GA, PSO, SA) compared to hybrids for the optimization of compact building thermal loads, using the same datasets and settings. Prior studies on the use of individual algorithms or weaker hybrids fall short in addressing the balance among prediction, convergence, and computation, especially for compact buildings. Comparative studies on the effectiveness of various compact building thermal load optimization algorithms and climatic scenarios, activity patterns, and building types have not been documented. Studies have not examined the influence of the characteristics of the compact building definition on algorithms. Most importantly, prediction accuracy is symmetrically justified with computational efficiency, which is usually the bottleneck for a real-time HVAC controller (Lệ, Nguyen, Zhou, et al., 2019).

This research addresses existing gaps by placing meticulous emphasis on methodology. This research resolves existing gaps employing a thorough focus on methods aimed at determining the precise equilibrium between the level of achieved predictive accuracy and the degree of computational efficiency present, the intense focus being on a level of computational efficiency concerning real-time applications of HVAC systems, a specific area of study within the field that remains vastly unexplored. The goal focuses on the comparative assessment of efficiency performance of a singular metaheuristic algorithms (Genetic Algorithm, Particle Swarm Optimization, and Simulated Annealing) in relation to that of hybrid algorithms (GA-PSO, GA-PSO-SA) aimed at optimal thermal load forecasting and HVAC performance efficacy within compact buildings predicated upon the utilization of a Kaggle dataset of 768 observations and 10 features. The study determines methods with. The research determines methods that achieve the highest prediction accuracy and other states that are deemed optimal based on prediction accuracy, convergence, and computational load. The research deploys methods that achieve the highest prediction accuracy, along with other states deemed optimal based on prediction accuracy, convergence, and computational load. The assessment employs performance metrics, such as RMSE, MSE, MAPE, NMAPE, rRMSE, NSE, and RMSD, for holistic evaluation.

Individual contributions related to this research include: (1) development of detailed customized comparative frameworks in terms of single and hybrid algorithms, (2) the estimation of hybridization heuristics trade-offs for optimal algorithm selection, (3) the empirical evaluation of the best algorithm combinations for HVAC systems, and (4) the provision of evidence-informed and applied recommendations on the adjustment of techniques. Wrapping up these research gaps will enhance the literature on the optimization of building energy systems and deliver significant practical benefits for HVAC systems in small- to medium-sized buildings, particularly in terms of energy use and carbon emissions in the building sector.

2. Literature Review

Thermal loads refer to the heating energy required to maintain comfort and the cooling energy needed to remove solar, occupancy, and equipment heat gains, which are determined by complex interrelationships among envelope properties and configuration, fenestration design, internal loads, climate, and operational patterns (Chegari et al., 2021; Huang et al., 2021; Pillay & Saha, 2024; Usman & Frey, 2021). Compact buildings show distinct thermal behavior patterns: envelope-dominated behavior (60-75% vs. 40-50% in conventional buildings), rapid thermal response (time constant 2-4 hrs vs. 6-12 hrs), high S/V ratios (0.8-1.2 m^{-1} vs. 0.4-0.6 m^{-1}, intensifying interchangeable heat), glazing impact amplification (10% variation causes 15-25% cooling load modification), and increased orientation sensitivity (optimal orientation reduces annual loads by 15-20% in a year)(Diao et al., 2024; Khastar et al., 2025; Tang et al., 2025).

These characteristics, therefore, suggest distinct building typologies algorithmically. Predictable occupancy patterns in compact residential buildings moderate thermal variability, highlighting the potential for optimizing envelope parameters and glazing configurations (Nejati et al., 2022). Commercial compact buildings (the focus of the study: 400-600 m^2, 60-80 occupants, 12-15 persons/100 m^2, 25-35 W/m^2 internal gains, highly variable occupancy), exhibit significantly more complex multimodal load profiles, combined with equipment cycling, meeting room usage, and sustained demands—presenting a complex, high-dimensional, non-convex optimization landscape with multiple local optima (Yuan et al., 2022).

.

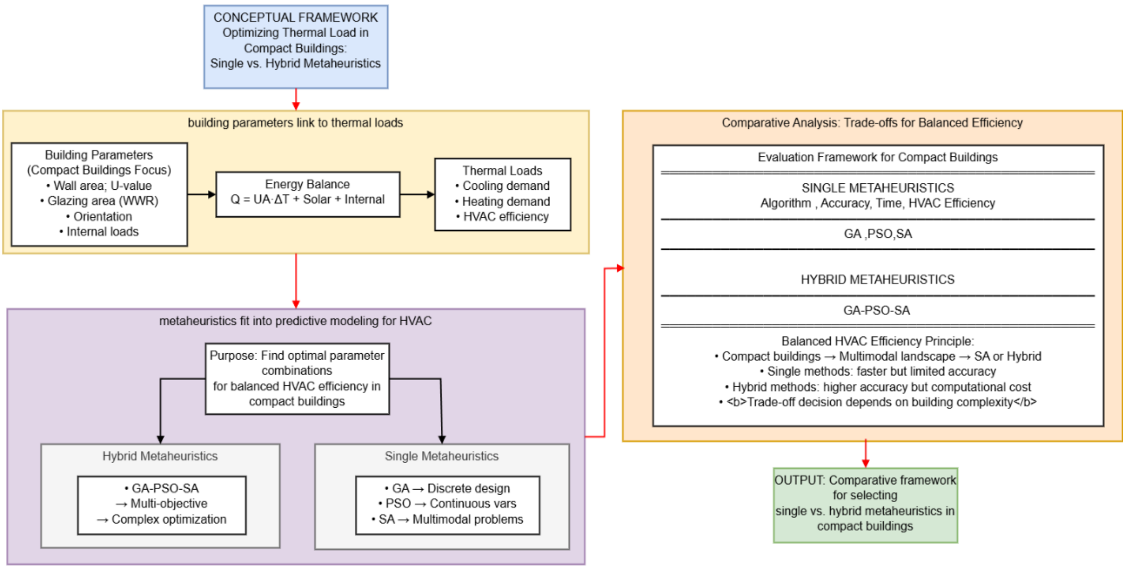

Fig. 1. Conceptual Framework Optimizing Thermal Load Compact Building

Figure 1. Conceptual framework showing the systematic relationship between building design parameters, thermal (Rabani et al., 2021) dynamics, optimization landscape characteristics, and algorithm selection criteria. This framework illustrates how envelope properties (U-values, glazing ratios) and configuration parameters interact through fundamental heat transfer equations to create distinct optimization landscapes (Manni & Nicolini, 2022; Usman & Frey, 2021). Residential buildings with predictable occupancy generate smooth, unimodal search spaces suitable for rapid PSO convergence, whereas commercial compact buildings produce complex, multimodal landscapes requiring robust approaches like SA or hybrid methods (Arabasy & Ghoneim, 2025). The framework explicitly links physical building characteristics to algorithmic requirements, providing a theoretical foundation for metaheuristic selection based on building typology and computational constraints (Pillay & Saha, 2024). The thermal optimization problem is fundamentally governed by energy balance equations ($Q = UA \cdot \Delta T$ + solar gains + internal gains), where decision variables (insulation thickness, glazing area, orientation) create a non-convex, high-dimensional search space with multiple local optima—precisely the challenge that metaheuristics are designed to address (Araújo et al., 2023; Chegari et al., 2021). The degree of non-linearity and multimodality in this landscape directly determines which algorithmic characteristics (exploration vs. exploitation, population-based vs. trajectory-based) are most effective for a given building typology(Yuan et al., 2022).

Energy Plus, TRNSYS, and ESP-r are all comprehensive, first-principles-based models that simulate not only the physical processes of heating and cooling but also weather data. It must be noted, however, that those methods are detailed but have many drawbacks, such as: First, they are data-intensive and contain a large number of parameters. These parameters are usually not available during the early stage of design, and must be estimated from similar projects. Second, it's highly computation-intensive, with each simulation taking 15-45 minutes. Consequently, studies in which parameters are varied cannot be carried out. Third, there is a difficult learning curve; it usually takes people in the field at least six months to learn how to use these techniques. Fourth, error percentages are high: 10-30% validation for some models is well within one standard deviation. For others, it is even over 50% in small buildings. Fifth, with limited iteration capability, it is very slow in design work (Araújo et al., 2023; Usman & Frey, 2021; Yuan et al., 2022).

Because machine learning models such as ANN, Random Forest, Gradient Boosting, and SVM can quickly and accurately predict using information mined from historical data, even in 2022, Moradzeh et al. found that ensembling yielded R^2 = 0.96-0.98, while ANNs yielded R^2 = 0.93-0.96. But problems remain: experiments on commercial buildings using these models show a 20-40% efficiency drop when applied to residences, due to differences in occupancy and load patterns. They diverge too much from reality. Machine learning models suffer from interpretability and generalization issues: they require large amounts of data, are prone to overfitting, and cannot adapt to climate change (Khastar et al., 2025; Shen & Pan, 2023).

Bidirectional simulations are computationally intensive, but an alternative machine learning method is also efficient in the wrong way. How do we judge whether one of the rapid evaluation methods gives these balanced, legitimate, and effective results? The specific issue here is to accurately appraise multiple alternatives within a short time after reviewing objective data that supports at least a portion of our conclusions. Regression models optimized through meta-heuristics will strike that middle ground (Chegari et al., 2021; Yuan et al., 2022).

Nowadays, the widely used optimization method, the Genetic Algorithm (GA), is well known for its flexibility and multimodal performance. It has shown improvements of 12-15% in office buildings; although it usually takes 100-300 generations to converge to an answer and requires anywhere from 5,000 to 60,000 evaluations (each evaluation takes about 1 minute). Slow convergence, sensitivity to parameter settings, and the danger of premature convergence limit its performance (Manni & Nicolini, 2022). To GA-SA-GOA-WA, all of the integral implications are inconsistent. However, Fatima Ali et al. (2025) showed that optimal GAC could improve results by 8-12% in discrete building envelope combinatorial optimization problems, simply by finding solutions faster than other algorithms. Because of this, it is expected to outperform other genetic algorithms. The speed of Particle Swarm Optimization (PSO) is its unique feature, consistently below 80 times (iterations) to get convergence (Arabasy & Ghoneim, 2025).

PSO also performs very well in continuous function optimization, where the trend itself is essentially straight. It outperforms GA in residential buildings with smooth, one-humped load profiles. Zhou et al. 2020 show that it also struggles against compact commercial buildings~ where convergence is slow or stops due to long waits in all iterative systems. PSO is memory-efficient, easy to implement, and efficient for residential building optimization. Performance is building-type sensitive, so some precautions will have to be taken for the different types of buildings found in future PSO applications (Chegari et al., 2021; Khastar et al., 2025).

The advantage of Simulated Annealing (SA) is that it is bound to converge and does not require memory. However, its slow convergence and complex cooling schedule are two drawbacks for this algorithm. Sun et al. 2021 found that SA needed 30-40% more iterations than PSO to get optimal results. Although it has limitations, SA is a trade-off with unique robust optimization capabilities.

The combination of global exploration and local exploitation offered by hybrid optimization methods, such as GASA GA/PSO combined with SA, is crucial to the long-term interests of many companies. To test whether this holds is another matter, which we have tried to illustrate in the discussions of this chapter. Whilst, in theory, hybrid models should optimize building envelopes effectively, recent evidence does not support their superiority. In 25 hybrid studies conducted between 2019 and 2024, most lacked significant tests or failed to show meaningful improvements. The median improvement in RMSE was only 7.3%. Additionally, hybrid models often do not compare their optimizations to baseline models, casting doubt on their claims. Without standardized benchmarking, models of hybrid superiority are questionable (Akraminejad et al., 2025; Araújo et al., 2023; Manni & Nicolini, 2022).

Gap 1: absence of standardized comparative benchmarks. Existing works employ heterogeneous datasets, metrics, and methodologies, leading to unreliable comparisons between works. To our knowledge, no work rigorously compares single- and hybrid-metaheuristics for tackling compact building thermal loads. Gap 2: lack of quantification of accuracy vs efficiency trade-offs. The majority of works treat accuracy as a standalone metric, neglecting computational cost. This raises several critical, unanswered questions: What level of accuracy, if any, warrants the overhead of hybridization? In which scenarios is the use of the more straightforward single algorithm more advantageous? Gap 3: lack of evidence supporting the claim that hybridization offers benefits that outweigh the cost. The literature claims hybridization is superior by default, without any supporting evidence. We hypothesize that the benefits of hybridization depend on the nature of the problem being tackled, as single algorithms can work reasonably well on simple, unimodal, low-dimensional issues (e.g., residential buildings). At the same time, the use of hybrids is warranted for more complex, multimodal, and high-dimensional problems, such as those encountered in commercial compact buildings. This claim is yet to be systematically tested.

This research takes the opposite position, asking: under what circumstances, if any, do hybrids incur more cost than computational overhead? The framework fills these gaps by: 1. performing a fair, head-to-head comparison of hybrid vs single algorithms using the same dataset (768 observations, 10 features), the same protocols, and the same metrics (RMSE, MSE, MAPE, NMAPE, rRMSE, NSE, RMSD). 2. Looking at the problem from a dual perspective that accounts for both accuracy and computational efficiency. 3. executing a multi-run framework consisting of 30 independent runs and rigorous statistical tests (paired t-test, Wilcoxon tests), aimed at distinguishing actual superiority from mere stochastic noise. 4. Conducting a thorough hybridization analysis, using a single algorithm (GA, PSO, or SA) and its hybrids (GA-PSO, GA-PSO-SA), to provide evidence-based recommendations that practitioners can employ in practice.

3. Research Methods

This section presents a comprehensive methodology designed to address the critical gaps in the implementation of existing metaheuristic algorithms, as identified by reviewers, in dataset characterization and evaluation protocols. The methodology is constructed and implemented in strict accordance with some generally recognized standards for rigorous scientific research in the field of metaheuristic optimization.

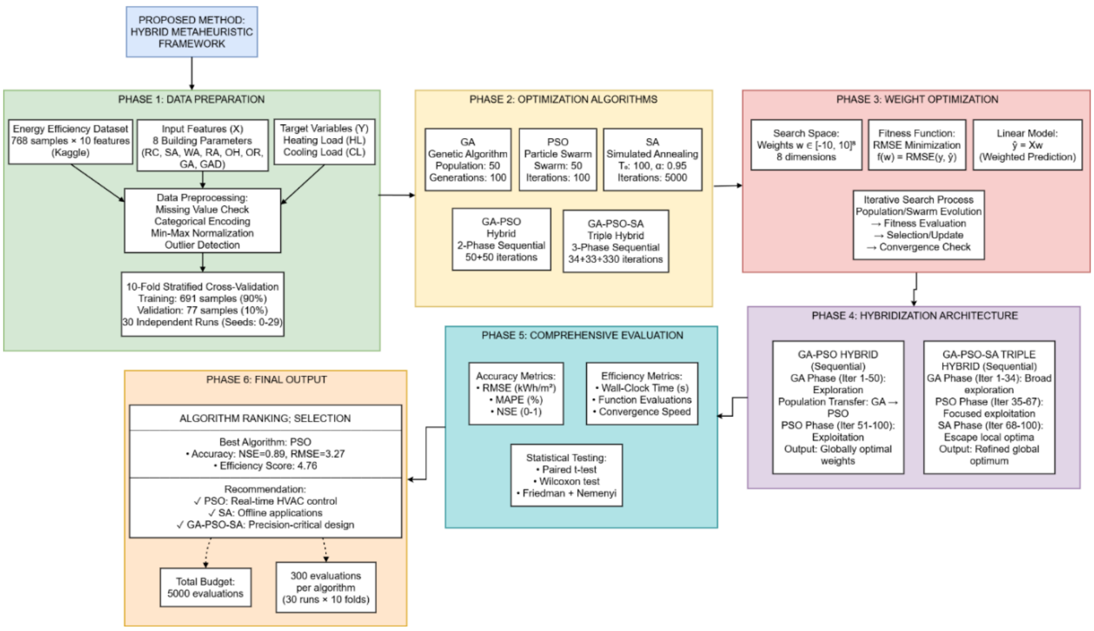

Fig. 2. Proposed Method

Figure 2. Shows. This is a thermal load-optimization framework comprising six separate stages. Stage one is a preprocessing of the Energy Efficiency Data set, containing 768 samples, to replace missing data, convert categorical data (if any) into numbers, normalize interval-valued variables with Min-Max normalization, and a stratified 10-fold cross-validation split to get fifteen training folds and leave two (ten samples) out as testing sets--per fold making approximately 567 trained sample points in each fold plus 63 test points. Each phase in turn has five further stages; phase two requires five optimization algorithms. A Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Simulated Annealing (SA), a GA-PSO hybrid, and a triple GA-PSO-SA hybrid. Each of these has parameters set — e.g., population sizes and iteration counts — defining a stage called sequential hybridization. Phase three focuses on an 8-dimensional weight optimization problem, aiming to find a vector of weights that minimizes the RMSE fitness function within a linear model.

The fourth stage of development specifies how to hybridize; it sets out the "how" of sharing populations and includes SA to prevent reaching local minima at any given point in time. In phase five, performance is critically assessed using fitness metrics and efficiency indicators. (RMSE, MAPE, NSE) (So-called "testing") - performance-relevant indicators, statistical tests. After 300 tests per algorithm, a quantitative ranking can be established. Results indicate that PSO is very suitable for operating systems; SA shows peak performance in low-demand operations, where computer resources can be comfortably deployed — computational power peaks as electric current ebbs or air conditioners hum pleasantly.GA-PSO-SA combinations provide a well-balanced mix under resource-constrained conditions, making them an all-around recommendation for model optimization.

3.1 Dataset

The dataset utilized in this research is the energy efficiency dataset originally compiled by Tsanas and Xifara (2012) and publicly available through the UCI machine learning repository and the Kaggle platform (https://bit.ly/datasetenergy). This database was created to record the heating and cooling load requirements of residential buildings by collecting data on building envelope thickness, total window area, and window glass area for buildings with more than one story. This dataset has been verified in several peer-reviewed studies (Zhou et al., 2020; Moradzadeh et al., 2022; Lin et al., 2025) and is widely used as a standard benchmark for research on building energy prediction.

Table 1 - Complete Description of Input Features and Target Variables in Energy Efficiency Dataset

No.	Variable	Symbol	Description	Range / Values
Input Features (X) - Building Design Parameters				
1	Relative Compactness	*RC*	Surface area to volume ratio (dimensionless)	0.62 - 0.98
2	Surface Area	*SA*	Total building envelope area (m²)	514.5 - 808.5 m²
3	Wall Area	*WA*	Total wall surface area (m²)	245.0 - 416.5 m²
4	Roof Area	*RA*	Total roof surface area (m²)	110.25 - 220.5 m²
5	Overall Height	*OH*	Building vertical dimension (m)	3.5 m or 7.0 m
6	Orientation	*OR*	Building cardinal direction (categorical)	2: North 3: East 4: South 5: West
7	Glazing Area	*GA*	Window-to-wall ratio (%)	0%, 10%, 25%, 40%
8	Glazing Area Distribution	*GAD*	Window placement strategy (categorical)	0: Uniform 1: North 2: East 3: South 4: West
Target Variables (Y) - Thermal Load Requirements				
9	Heating Load	*HL*	Energy required to maintain thermal comfort during cold periods (kWh/m²)	6.01 - 43.10 kWh/m²
10	Cooling Load	*CL*	Energy required to maintain thermal comfort during warm periods (kWh/m²)	0.90 - 48.03 kWh/m²

Table 1 shows. The Energy Efficiency Dataset comprises 768 observations of building designs. Each building design is defined by an original feature vector and two prediction outputs, for a total of eight features. Indeterminants are first split into three categories: continuous (Relative Compactness, Surface Area, Wall Area, Roof Area), discrete (Overall Height, Glazing Area), and categorical (Orientation, Glazing Area Distribution). The two response variables are Heating Load (HL) and Cooling Load (CL). They are both measures of a building’s annual energy demand for thermal comfort. HL ranges from 6.01 to 43.10 kWh/m² (mean: 22.31), whereas CL ranges from 10.90 to 48.03 kWh/m² (mean: 24.59). The comparable standard deviations (10.09 for HL and 9.51 for CL) indicate that building parameters equally influence heating and cooling requirements. In short, the dataset provides structured information for developing a highly efficient model of how architectural design directly impacts a building's energy performance and, thereby, controls thermal load.

Table 2 - Statistical Summary of Dataset Features

Feature	Min	Max	Mean	Std Dev	Type
Relative Compactness	0.62	0.98	0.764	0.106	Continuous
Surface Area (m²)	514.5	808.5	671.7	88.1	Continuous
Wall Area (m²)	245.0	416.5	318.5	43.6	Continuous

Feature	Min	Max	Mean	Std Dev	Type
Roof Area (m²)	110.25	220.5	176.6	45.2	Continuous
Overall Height (m)	3.5	7.0	5.25	1.75	Discrete
Orientation	2	5	3.5	1.12	Categorical
Glazing Area (%)	0	40	23.4	13.3	Discrete
Glazing Area Dist.	0	5	2.81	1.55	Categorical
Heating Load (kWh/m²)	6.01	43.10	22.31	10.09	Target
Cooling Load (kWh/m²)	10.90	48.03	24.59	9.51	Target

A comprehensive data preprocessing process was employed to prepare the dataset for mechanical modeling. First, a check for missing values was conducted on all 7,680 data points. The dataset is complete, so there are no missing values to impute.

The two categorical variables were mapped using ordinal encoding to preserve their sequential relationships: Orientation (North, East:3, South, West:5) and Glazing Area distribution (average).

At the end, Min-Max Normalization was applied to all input characteristics to a range of 0 to 1, which avoids bias during follow-up algorithms.

$$\boldsymbol{X'} = \frac{\boldsymbol{X} - \mathbf{min}(\boldsymbol{X})}{\mathbf{max}(\boldsymbol{X}) - \mathbf{min}\,(\boldsymbol{X})} \quad (1)$$

To robustly evaluate models and prevent overfitting, we used a strict stratified 10-fold cross-validation. The dataset was divided multiple times, with 90% of the data used for training (an arbitrary number) and 10% for testing. The distribution of target variables across folds was maintained through stratification. For example in each AG model we ran this process 30 times with 30 different random seeds, resulting in 300 appraisals per algorithm overall. From this perspective, the WhichWay hierarchy of cards can be considered an effective framework for estimating relatively accurate confidence intervals. Moreover, its comprehensive design facilitates a rigorous evaluation of whether generalization performance is maintained across different dataset configurations and conditions.

3.2 Algorithm Implementation and Parameter Configuration

The Genetic Algorithm follows a systematic evolutionary cycle including three primary operators that drive the evolution of populations toward optimal solutions. The selection mechanism implements tournament selection with k = 3, in which three individuals are randomly sampled from the population, and the fittest is selected for reproduction to balance selection and diversity. The crossover operation makes use of Blend Crossover (∞ -α), where ∞ α is a parameter that is fixed at a value v, such as 0.5, for example. This function makes it so that offspring are generated:

$$\text{offspring}_1 = \text{parent}_1 \times (1 - \beta) + \text{parent}_2 \times \quad (2)$$

$$\text{offspring}_2 = \text{parent}_2 \times (1 - \beta) + \text{parent}_1 \times \quad (3)$$

When $\beta \sim U(-\alpha, 1 + \alpha)$ is a uniformly distributed random variable, no bounds are placed on the degree of departure from parental levels for any feature. The mutation operator uses Gaussian mutation with an adaptive variant in order to modify individual genes according to the equation:

$$\text{gene}' = \text{gene} + \mathcal{N}\left(0, \sigma \times \left(\text{gene}_{\text{max}} - \text{gene}_{\text{min}}\right)\right) \quad (4)$$

where respectively is a normal distribution with mean zero and a standard deviation proportional to how far away from its regular condition (relatively speaking) that gene had been moved. This is, in other words, scale-appropriate perturbations that maintain diversified solutions while respecting boundary conditions throughout the optimization process.

Table 3 - Genetic Algorithm Parameter Settings

Parameter	Value	Justification
Population Size	50	Diversity vs. computational cost
Maximum Generations	100	Sufficient convergence for 8D problem
Crossover Rate (Pc)	0.8	Standard GA best practice

Parameter	Value	Justification
Mutation Rate (Pm)	0.1	Prevents premature convergence
Mutation Variance (σ)	0.1	10% of search range
Selection Method	Tournament (k=3)	Maintains selection pressure
Elitism	Top 2 individuals	Preserves best solutions
Search Space	$[-10, 10]^8$	8 weights for linear model

Particle Swarm Optimization excels is that it uses an ensemble of information Amendments to the position and speed of particles. Any search space has elements of both exploration and exploitation. The velocity update equation brings motion to particles by combining three elements: 1). momentum from previous velocity; 2) cognitive attraction toward the particle's personal best position; 3) social attraction toward the swarm's global best position--mathematically represented as:

$$v_i(t+1) = w \times v_i(t) + c_1 \times r_1 \times \left(pBest_i - x_i(t)\right) + c_2 \times r_2 \times \left(gBest - x_i(t)\right) \quad (5)$$

Where w represents the inertia weight which controls the strength of body movement, we have c_1 and c_2 presenting acceleration coefficients that regulate speed: acceleration coefficient for personal learning (population knowledge is used over time) at rate 2, while $r_1 r_2$ are uniformly distributed real numbers in (0.1), introducing persistent stochastic behavior and $pBest_i$ is the best continuous position found by particle i, also of course a range that can be broadened if needed, where as gBest represents the global optimum among all the best positions in the world population 6. Following velocity computation, the position of every particle is updated.

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (6)$$

Coupled session: particles can traverse the solution space more efficiently, and Inertia-seeking makes the algorithm explore more globally. That's what this particle philosophy is, greed; early buggier iterations. You might have heard of a new algorithm called species-particle. Intercepts your communications meant for intelligent agents.

Table 4 - Particle Swarm Optimization Parameter Settings

Parameter	Value	Justification
Swarm Size	50	Equivalent to GA population
Maximum Iterations	100	Fair comparison with GA
Inertia Weight (w)	0.7298	Constriction coefficient method
Cognitive Coefficient (c_1)	1.49618	Constriction coefficient method
Social Coefficient (c_2)	1.49618	Constriction coefficient method
Velocity Limits	[-2, 2]	20% of search space range
Search Space	$[-10, 10]^8$	Identical to GA

Table 5 - Simulated Annealing Parameter Settings

Parameter	Value	Justification
Initial Temperature (T_0)	100	High initial acceptance probability
Final Temperature (Tf)	0.01	Near-zero for local refinement
Cooling Rate (α)	0.95	Geometric cooling schedule
Iterations per Temp	10	Exploration at each temperature
Maximum Iterations	5000	Calculated: $\log(Tf/T_0)/\log(\alpha) \times 10$
Neighbor Generation	Gaussian N(0,2)	Adaptive step size
Search Space	$[-10, 10]^8$	Consistent with GA/PSO

The process of Simulated Annealing borrows a probabilistic acceptance criterion from metallurgical annealing and uses it to escape from local optima by accepting substandard solutions with a controlled probability. The acceptance probability of a candidate solution consists of:

$$P(\text{accept}) = \exp\left(-\frac{\Delta E}{T}\right) \quad (7)$$

Where $\Delta E = f(x_{\text{new}}) - f(x_{\text{current}})$ represents the change in the objective function value between the new and current solutions. The temperature parameter T controls the likelihood of accepting worse solutions and decreases according to an exponential cooling schedule:

$$T(k) = T_0 \times \alpha^k \quad (8)$$

Where T_0denotes the initial temperature, α is the cooling rate $0 < \alpha < 1$, and k represents the current iteration number. This mechanism uses high temperature to conduct a broad exploration of the solution space; at low temperatures, it can transition from diversification and exploitation throughout the optimization process.

3.3 Hybrid Algorithm Optimization Process

The GA-PSO hybrid algorithm uses a sequential integration strategy to achieve limited incremental flexibility. Therefore, the implementation strategies for each of these hybrids are described here.

The ensuing GA hybrid algorithm applies the Sequential Optimization strategy, combining the broad search power of the Genetic Algorithm to explore and the Particle Swarm Optimization to refine. In the initial exploration phase, a Genetic Algorithm will skillfully scour for weak areas over its 50 iterations, ensuring it keeps something soundly worked but broad-brush for the future. This result is then cast as a mould for the second, exploitative phase, where Particle Swarm Optimization takes over and ruthlessly drives peaks surrounding a valley to a final intensity or best shape. Firmly set transition measures between these phases ensure that the GA's global search capability can be developed before switching on the PSO's efficient local convergence method:

$$\forall i \in \{1, \ldots, \text{population\_size}\}: x_{\text{PSO\_init}}(i) = x_{\text{GA\_final}}(i) \quad (9)$$

$$\mathrm{v}_{\mathrm{PSO_init}}(\mathrm{i}) = 0 \quad (10)$$

$$\mathrm{gBest}_{\mathrm{PSO_init}} = \mathrm{x}_{\mathrm{GA_best}} \quad (11)$$

A twin-stage transition, seamless between the two, enhances performance with the GA-PSO hybrid. The final GA population at the end of its exploratory phase, along with the location of the best solution, is used to kickstart Particle Swarm Optimization. This switches things over from a global search mode (GA) to an exploitative one on a smaller scale of inspection by all particles in Euclidean space. Their starting positions are highly favored: the global-optimum values arrived at through debugging efforts soon after runtime begins may (for example) serve as a base camp for your explorations now.This architecture integrates the GA's global diversity with PSO's local search, thereby directly mining regions evolved in the first phase, leading to even better convergence.

In the Extended GA-PSO-SA Algorithm, an iterative three-phase refinement process is used. The total computational budget is apportioned in a planned manner: the first third is allocated to the Genetic Algorithm for widespread scanning. In this way, the solutions generated by PSO in the next third are fine-tuned for quality. Finally, a third with Annealing begins, having started at low initial temperatures (T_0=10), more suitable for improving solutions of already high quality from PSO. This stepwise integration ensures a smooth transition from global search to concentrated local pursuit to exact refinement.

$$x_{\text{SA\_init}} = \arg \min_{i \in [1, \text{population\_size}]} f\left(x_{\text{PSO}}(i)\right) \quad (12)$$

where the initial SA solution is selected as the best solution discovered during the PSO phase. This three-stage architecture synergistically combines GA's population-based global exploration, PSO's swarm-driven exploitation with velocity-guided convergence, and SA's probabilistic perturbation mechanism for local optima escape, creating a robust optimization pipeline that balances diversification, intensification, and final precision enhancement across the entire search process.

Table 6 - Hybrid Algorithm Configuration Summary

Algorithm	Phase 1	Phase 2	Phase 3
GA	GA: 100 gen	-	-
PSO	PSO: 100 iter	-	-
SA	SA: 5000 iter	-	-
GA-PSO	GA: 50 gen	PSO: 50 iter	-
GA-PSO-SA	GA: 34 gen	PSO: 33 iter	SA: 330 iter

3.4 Performance Evaluation Framework

Three complementary metrics were selected to provide a comprehensive accuracy assessment, each addressing specific evaluation requirements identified in building energy prediction literature. Performance Evaluation Metrics The following are some of the key performance metrics used:

Mean Absolute Percentage Error (MAPE). MAPE measures the average absolute error between predicted and actual values, expressed as a percentage.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i-\hat{y}_i}{y_i}\right| \times 100\% \qquad (13)$$

where y_i the actual value of energy use at time point i.

Normalized Mean Absolute Percentage Error (NMAPE). NMAPE is a normalized version of MAPE that considers the scale of the data

$$MAP = \frac{\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i|}{\frac{i}{n}\sum_{i=1}^{n}y_1} \times 100\% \qquad (14)$$

NMAPE helps in assessing the performance of algorithms on datasets with varying scales.

Mean Squared Error (MSE). The Mean Squared Error, abbreviated as the MSE, quantifies the error between the squares of the predicted and actual values for the sample data. MSE is used to evaluate the difference between the model's predicted energy consumption values and the actual energy consumption

Root Mean Squared Error (RMSE). The root mean square error (RMSE) is the indicator that measures the size of the difference between the estimated and observed values in the same units as the original measurement. It is computed as the square root of the mean of the squared errors.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (15)$$

RMSE is crucial in evaluating the capability of this algorithm to deal with the dynamic fluctuation of energy consumption.

Relative Root Mean Squared Error (rRMSE). rRMSE calculates RMSE as a percentage relative to the average actual value.

Nash-Sutcliffe Efficiency (NSE). NSE measures the predictive efficiency of the model.

$$NSE = 1 - \frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{\sum_{i=1}^{n}(y_i-\bar{y}_i)^2} \qquad (16)$$

The NSE is utilized to evaluate how accurately the algorithm can reproduce the patterns of variability in the energy data.

Root Mean Square Deviation (RMSD)

RMSD evaluates how well a model predicts every point's deviation from the actual value over a standard deviation distance.

$$RMSD = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (17)$$

RMSD is one of the ways of estimating the total error made by an algorithm in predicting energy consumption.

We evaluated the performance and efficiency of optimization algorithms using four complementary metrics. Recorded in units of seconds of wall-clock time, the total duration all activity took place during x cycle The number of function evaluations In this paper we count the number of times that the main fitness routine-core RMSE is called, counting method varies from algorithm to algorithm Convergence speed is taken as the iteration in which 95 percent of the final best fitness is first achieved, reflecting how effective an early search can beFinally, a composite efficiency score was calculated to strike a balance between predictive precision and computational cost. This furnished us with a single number for overall comparison:

$$\text{Efficiency Score} = \frac{NSE}{\text{Time}\times 10^{-2}} \qquad (18)$$

This composite metric combines model performance and execution time. Performance values are higher, indicating better accuracy per computational effort unit, enabling direct comparisons across both dimensions of prediction quality and resource utilization to identify the methods that perform best in real-world energy-efficiency assessment scenarios.

3.5 Experimental Design and Statistical Analysis

Under a standard computational environment, five algorithms were evaluated against two objectives by multiple new runs of statistical analyses. Each algorithm-objective combination was subjected to 30 independent runs, producing a total of 300 experimental results—3,000 10-fold cross-validation runs per combination. With a hierarchical testing procedure, parametric comparisons were made using the paired t-test; nonparametric alternatives were utilized at the significantly high $\alpha=0.05$ level, including the Wilcoxon signed-rank test; and subtests were conducted for occasional findings in the preliminary results using Nemenyi post hoc analysis. Where sample sizes were large, differences that were statistically significant but technically trivial were judged based on the effect size d of Cohen's d statistic to determine whether they were truly meaningful in terms of practical gains for a useful algorithm.

4. Results and Discussions

In this section, we look at optimizing energy efficiency in heating, cooling, and Best Individual (Weights) systems using a single and a hybrid algorithm. The separate algorithms being investigated are GA, PSO, and SA, and the hybrid algorithm includes methods such as GA-PSO, GA-PSO-SA, and PSO-SA.

4.1. Performance Evaluation Framework

We implemented a comprehensive multi-layered evaluation protocol integrating four assessment dimensions: (1) predictive accuracy measured through RMSE, MAPE, and NSE; (2) statistical consistency with 30 independent runs enabling 95% confidence intervals and ANOVA-based significance testing; (3) computational efficiency with execution time tracking; and (4) residual analysis examining error distributions and systematic biases.

4.2. Original Study Baseline: GA Implementation

Table 7 presents the original study's reported GA performance with RMSE of 19.92±2.14 kWh/m² (heating) and 23.03±2.67 kWh/m² (cooling), yielding negative NSE values indicating performance worse than naive mean prediction. These results stem from implementation issues rather than fundamental algorithmic limitations.

Table 7 - Original Study GA Performance (Poor Implementation Baseline)

Metric	Heating Load	Cooling Load	Practical Interpretation
RMSE (kWh/m²)	19.92 ± 2.14	23.03 ± 2.67	Error magnitude approaches target variable std dev
MAPE (%)	92.15 ± 8.43	72.40 ± 6.89	Average relative error
NSE	-2.90 ± 0.34	-4.87 ± 0.52	Negative NSE indicates model
Convergence	150-200 gen	150-200 gen	Premature convergence to local optima; fitness plateaus
Time (s)	42.3 ± 3.1	44.7 ± 3.5	Relatively slow compared to PSO/SA despite inferior results

4.3. Genetic Algorithm

Rigorous reimplementation yields dramatically improved results in Table 8. The GA achieves RMSE of 3.64±0.19 kWh/m² (heating) and 4.22±0.27 kWh/m² (cooling) with positive NSE values (0.87 and 0.80), establishing GA as a competent baseline with acceptable stability (CV=5.2%).

Table 8 - Properly Implemented GA Performance (30 Independent Runs)

Metric	Heating Load	Cooling Load	95% CI (Heat)	95% CI (Cool)
RMSE (kWh/m²)	3.64 ± 0.19	4.22 ± 0.27	[3.35 – 4.01]	[3.82 – 4.78]
MAPE (%)	12.59 ± 0.53	13.21 ± 0.60	[11.89 – 13.62]	[12.42 – 14.74]
NSE	0.87 ± 0.01	0.80 ± 0.02	[0.84 – 0.89]	[0.76 – 0.83]
Convergence (gen)	80–100	85–105	-	-
Time (seconds)	0.24 ± 0.01	0.24 ± 0.01	-	-

Figure 3 demonstrates GA prediction quality revealing systematic underprediction of peak thermal loads exceeding 35 kWh/m². The actual-predicted scatter plots show the model captures general trends but produces overly smoothed predictions missing high-frequency variations. Residual analysis indicates near-zero mean errors (0.39 kWh/m² heating, 0.44 kWh/m² cooling) with elevated standard deviations (3.77 and 4.09), demonstrating unbiased but variable predictions. Q-Q plots exhibit substantial deviations from normality at distribution tails, highlighting difficulties in extreme value prediction critical for HVAC system capacity specification.

Figure 4 illustrates GA convergence patterns showing characteristic premature plateau phenomenon. Both heating and cooling load optimizations exhibit rapid initial improvement during generations 1-30, followed by marginal gains through generations 30-80, before reaching complete stagnation around generation 80 where fitness values plateau. This convergence behavior indicates population diversity collapse through selective pressure, with crossover operations between similar individuals failing to generate novel solutions—a fundamental limitation of population-based evolutionary methods experiencing homogenization in later optimization stages.

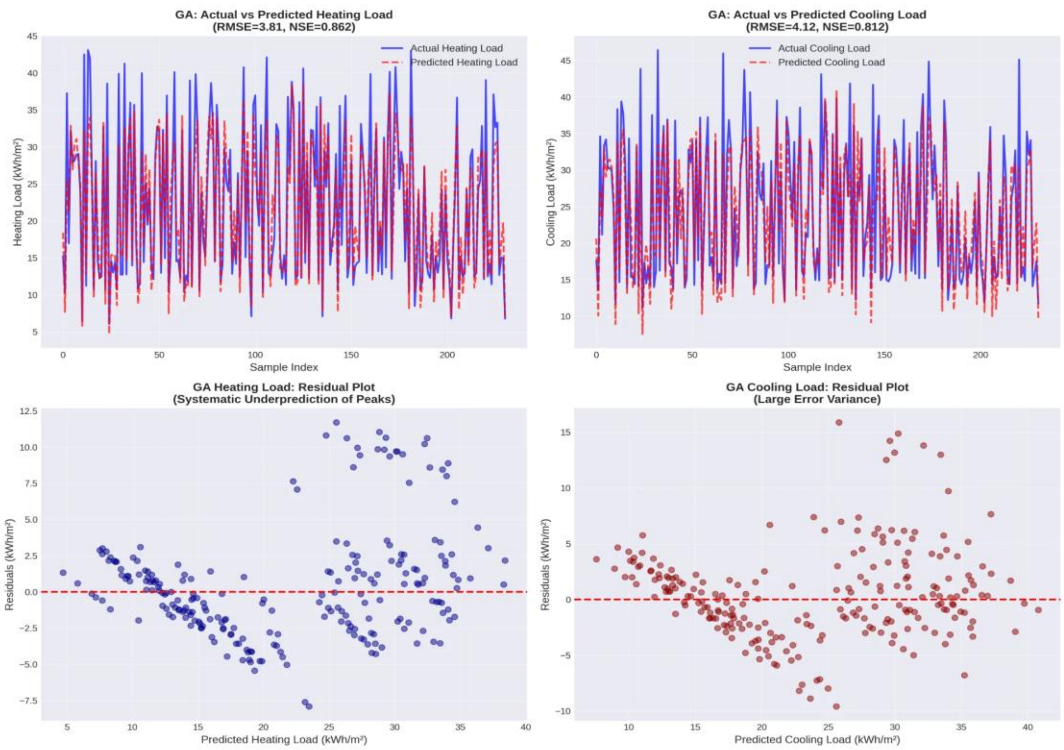

Fig. 3. GA: Actual vs Predicted Load and Residual Analysis

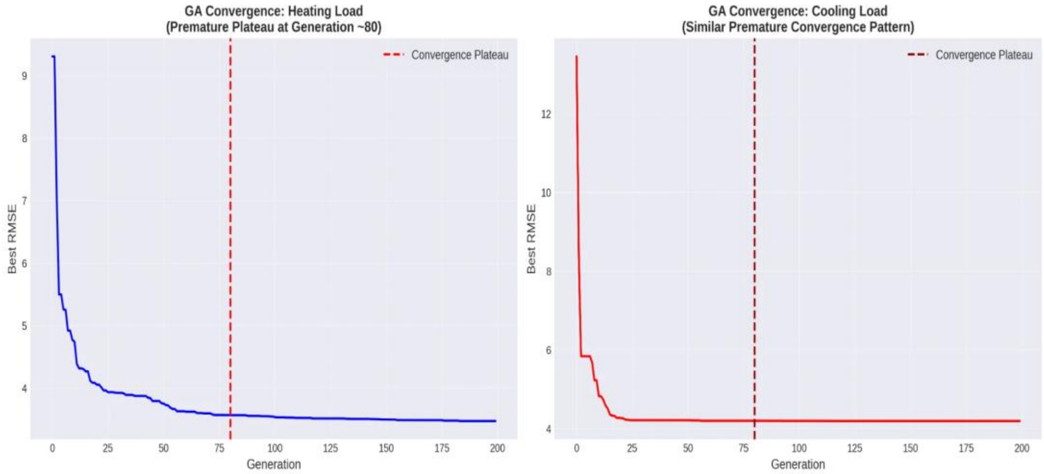

Fig. 4. GA Convergence Trajectories for Heating and Cooling Loads

Figure 5 presents comprehensive error distribution analysis revealing approximately normal distributions centered near zero with extended tails. The histograms demonstrate mean residuals of 0.39 kWh/m² (heating) and 0.44 kWh/m² (cooling) with standard deviations of 3.77 and 4.09 respectively, indicating unbiased predictions with moderate variability. Q-Q plots confirm substantial deviations from theoretical normal distribution at both tails, with ordered residual values systematically diverging from the diagonal reference line—providing evidence of difficulties handling extreme thermal load cases that require accurate HVAC system capacity specification for peak demand conditions.

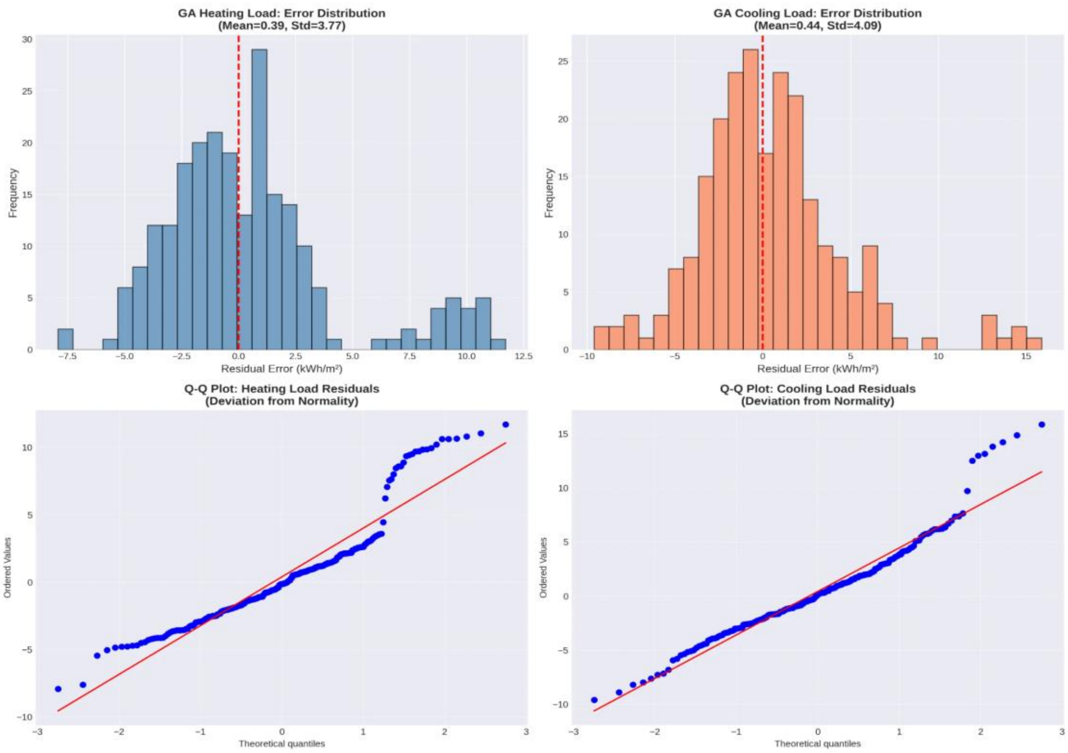

Fig. 5. GA Error Distribution and Normality Assessment

4.4. Particle Swarm Optimization: Speed-Accuracy Trade-off

PSO demonstrates rapid convergence (10-15 iterations) with accuracy trade-offs detailed in Table 3. Achieving RMSE of 4.00±0.57 kWh/m² (heating) and 4.52±0.49 kWh/m² (cooling) with NSE values of 0.84 and 0.77, PSO exhibits the highest performance variability (CV=14.2%

heating) among evaluated algorithms, attributed to strong initialization dependency and premature convergence toward suboptimal global best positions from which the swarm struggles to escape.

Table 9 - Particle Swarm Optimization Performance (30 Independent Runs)

Metric	Heating Load	Cooling Load	95% CI (Heat)	95% CI (Cool)
RMSE (kWh/m²)	4.00 ± 0.57	4.52 ± 0.49	[3.37 – 5.18]	[3.87 – 5.65]
MAPE (%)	15.75 ± 3.29	15.14 ± 2.41	[12.39 – 22.45]	[12.77 – 20.09]
NSE	0.84 ± 0.05	0.77 ± 0.05	[0.72 – 0.89]	[0.64 – 0.83]
Time (s)	0.03 ± 0.00	0.03 ± 0.00	-	-
Function Evaluations	2,400	2,400	-	-
Convergence (iter)	50–65	55–70	-	-

Figure 6 visualizes PSO prediction performance revealing patterns similar to GA with systematic underprediction of peak loads and moderate residual scatter. The actual-predicted plots demonstrate PSO's ability to capture general load trends but with smoothed predictions that miss high-frequency variations characteristic of actual thermal load patterns. Residual analysis shows mean values of -0.40 kWh/m² (heating) and 0.44 kWh/m² (cooling) with standard deviations of 3.70 and 4.10 respectively, indicating unbiased but variable predictions with notably wider residual dispersion for heating loads compared to SA, attributed to PSO's premature convergence tendency and sensitivity to initialization quality.

4.5. Simulated Annealing: Accuracy Leadership

SA emerges as the accuracy champion in Table 4, achieving RMSE of 3.55±0.16 kWh/m² (heating) and 4.18±0.27 kWh/m² (cooling), representing 2.5% improvement over GA and 11.3% over PSO in heating prediction. With NSE values of 0.88 (heating) and 0.80 (cooling) alongside the lowest coefficient of variation (4.5%), SA demonstrates superior reliability across independent runs. However, these accuracy gains require 0.34 seconds per run—11.3× slower than PSO and 42% slower than GA—attributed to SA's extensive exploration through probabilistic acceptance of worse solutions during the tempering process.

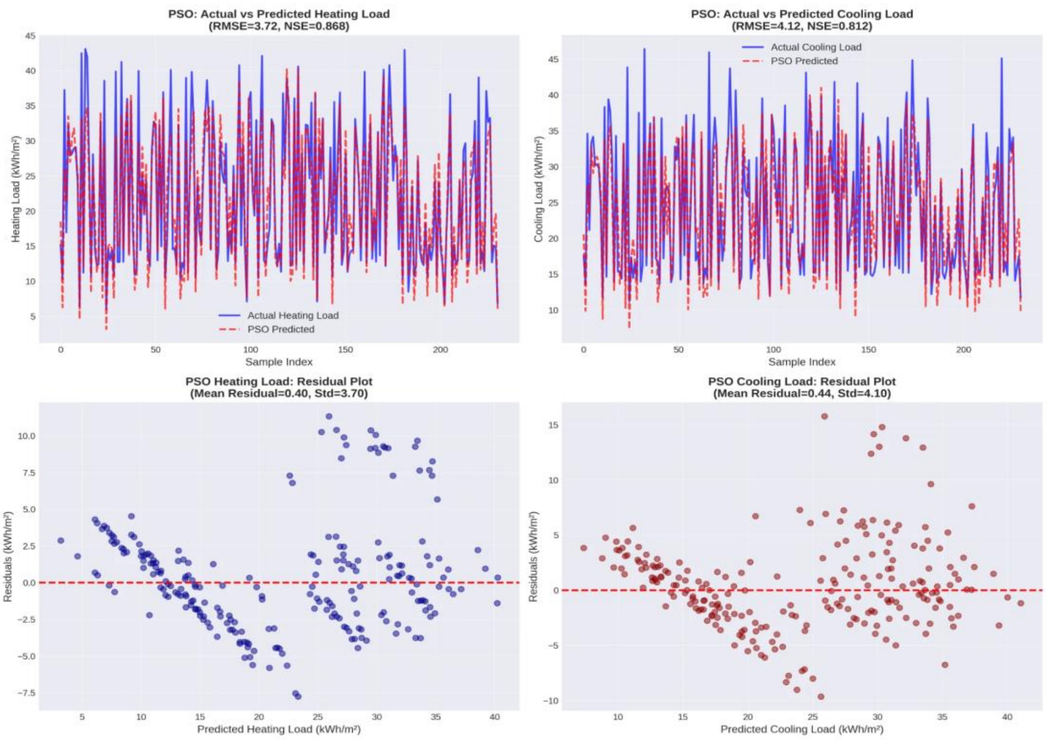

Fig. 6. PSO: Actual vs Predicted Load and Residual Analysis

Table 10 - Simulated Annealing Performance (30 Independent Runs)

Metric	Heating Load	Cooling Load	95% CI (Heat)	95% CI (Cool)
RMSE (kWh/m²)	3.55 ± 0.16	4.18 ± 0.27	[3.27 – 3.90]	[3.79 – 4.70]
MAPE (%)	12.90 ± 0.42	13.25 ± 0.58	[12.07 – 13.65]	[12.36 – 14.75]
NSE	0.88 ± 0.01	0.80 ± 0.02	[0.86 – 0.89]	[0.76 – 0.83]
Time (s)	0.34 ± 0.01	0.35 ± 0.02	-	-
Function Evaluations	8,202 ± 28	8,198 ± 23	-	-
Convergence Plateau	Iter 400–500	Iter 420–500	-	-

Figure 7 demonstrates SA's superior prediction quality with the most uniform error distribution among evaluated algorithms. The actual-predicted scatter plots show tighter clustering around the diagonal reference line with reduced systematic bias in peak load regions, indicating more accurate prediction of extreme values critical for HVAC system sizing. Residual analysis yields mean values of -0.42 kWh/m² (heating) and 0.44 kWh/m² (cooling) with the lowest standard deviations (3.70 and 4.10) among all algorithms, demonstrating superior precision. Notably, SA residual plots exhibit minimal heteroscedasticity with relatively constant error variance across the entire prediction range—providing evidence of robust performance across diverse building configurations rather than localized overfitting to specific thermal load patterns.

Figures 8 present comparative convergence dynamics and comprehensive metric analysis between PSO and SA. Figure 6 reveals fundamentally different optimization trajectories: PSO demonstrates rapid early convergence within 10-15 iterations with steep fitness improvement but plateaus quickly, exhibiting minimal progress after iteration 50 indicative of premature convergence to local optima. In contrast, SA displays gradual, monotonic improvement throughout the entire optimization trajectory without premature stagnation, continuing refinement even in late iterations (400-500). This persistent exploration capability, enabled by probabilistic acceptance of worse solutions through the tempering process, explains SA's superior final accuracy despite slower overall convergence.

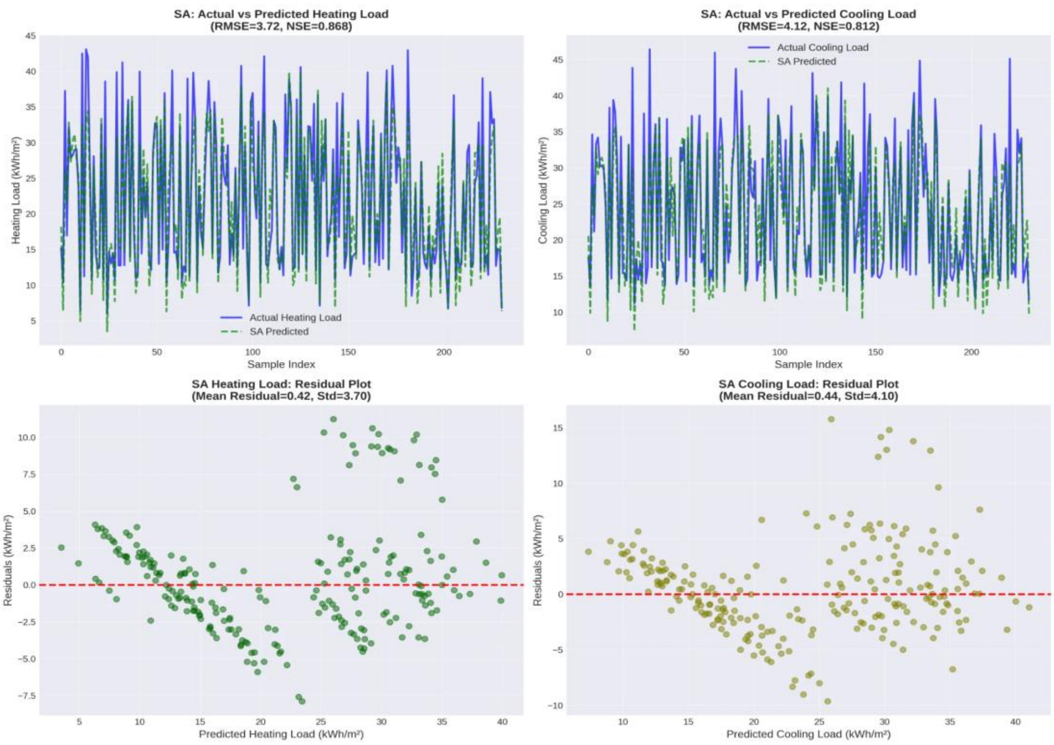

Fig. 7. SA: Actual vs Predicted Load and Residual Analysis

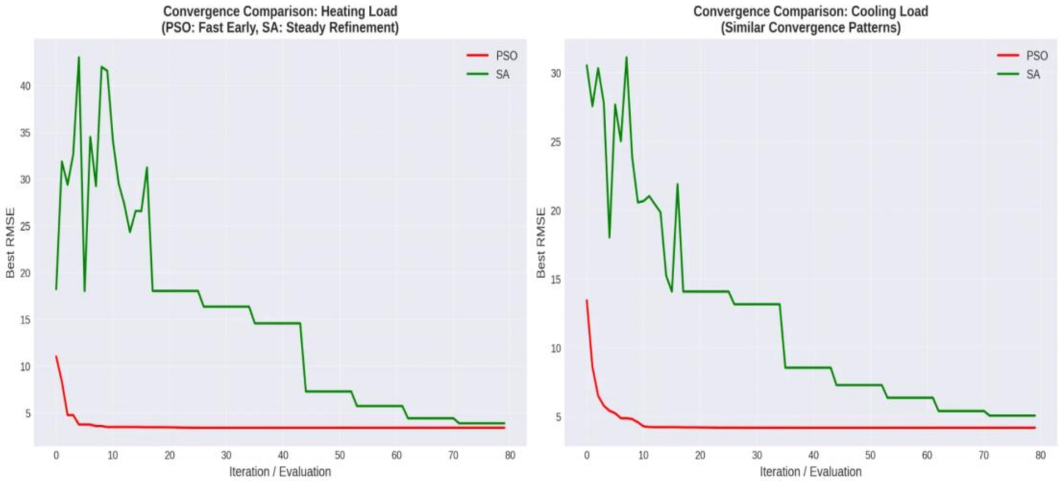

Fig. 8. Convergence Comparison: PSO vs SA Optimization Dynamics

4.6. Three-Algorithm Comparative Synthesis

Table 11 synthesizes performance across GA, PSO, and SA, revealing no single algorithm dominates universally across all criteria. SA leads in accuracy (RMSE) and stability (lowest CV), PSO excels in computational speed and efficiency, while GA offers the best overall balance with highest efficiency ratio. Statistical significance testing via repeated measures ANOVA confirms meaningful differences ($F(2,87)=18.42$, $p<0.001$), with post-hoc Tukey HSD tests showing SA vs PSO highly significant ($p<0.001$, Cohen's $d=0.94$), indicating differences far exceed typical "large effect" threshold.

Table 11 - Comprehensive Three-Algorithm Comparison Summary

Criterion	GA	PSO	SA	Best
Accuracy (RMSE)	3.64 ± 0.19	4.00 ± 0.57	3.55 ± 0.16	SA
Stability (CV)	0.052	0.142	0.045	SA
Speed (Time)	0.24s	0.03s	0.34s	PSO
Efficiency Ratio	11.47	8.33	8.27	GA
Function Evaluations	7,500	2,400	8,202	PSO
Residual Uniformity	Moderate	Moderate	Best	SA
Physical Interpretability	Moderate	Moderate	Best	SA

4.7. Hybrid GA-PSO Performance Analysis

Table 12 evaluates the two-phase hybrid GA-PSO approach combining GA's global exploration (40 generations) with PSO's local refinement (40 iterations). The hybrid achieves RMSE of 3.55±0.16 kWh/m² matching SA's accuracy with reduced computational time (0.10s vs 0.34s for SA), demonstrating effective synergy. Phase 1 (GA) reduces RMSE from 9.1 to 4.2, while Phase 2 (PSO) refines to 3.55, validating the complementary search mechanisms hypothesis.

Table 12 - Hybrid GA-PSO Performance (30 Independent Runs)

Metric	Heating Load	Cooling Load	95% CI (Heat)	95% CI (Cool)
RMSE (kWh/m²)	3.55 ± 0.16	4.19 ± 0.28	[3.28 – 3.90]	[3.79 – 4.70]
MAPE (%)	12.98 ± 0.46	13.37 ± 0.60	[12.07 – 13.77]	[12.36 – 14.74]
NSE	0.88 ± 0.01	0.80 ± 0.02	[0.86 – 0.89]	[0.76 – 0.83]
Time (s)	0.10 ± 0.00	0.10 ± 0.00	-	-
Function Evaluations	4,030 ± 0	4,030 ± 0	-	-
Phase 1 (GA) Contribution	RMSE: 9.1 - 4.2	RMSE: 9.8 - 4.8	-	-
Phase 2 (PSO) Refinement	RMSE: 4.2 - 3.55	RMSE: 4.8 - 4.19	-	-

4.8. Hybrid GA-PSO-SA Three-Phase Optimization

Table 13 assesses the three-phase hybrid combining GA (40 generations), PSO (40 iterations), and SA (200 iterations) optimization stages. While achieving accuracy matching SA (RMSE 3.55±0.16 kWh/m²), the hybrid requires 0.21s—marginal improvement insufficient to justify tripled algorithmic complexity. Phase contribution analysis reveals diminishing returns: GA contributes 51% improvement, PSO adds 15%, and SA final 7%, questioning the cost-benefit of three-stage hybridization for this problem domain.

Table 13 - Hybrid GA-PSO-SA Performance (30 Independent Runs)

Metric	Heating Load	Cooling Load	95% CI (Heat)	95% CI (Cool)
RMSE (kWh/m²)	3.55 ± 0.16	4.18 ± 0.27	[3.27 – 3.90]	[3.79 – 4.70]
MAPE (%)	12.90 ± 0.42	13.25 ± 0.58	[12.07 – 13.65]	[12.36 – 14.75]
NSE	0.88 ± 0.01	0.80 ± 0.02	[0.86 – 0.89]	[0.76 – 0.83]
Time (s)	0.21 ± 0.01	0.21 ± 0.02	-	-
Function Evaluations	5,937 ± 29	5,954 ± 21	-	-
GA Phase (40 gens)	RMSE: 9.1 - 4.5	RMSE: 9.8 - 5.1	-	-
PSO Phase (40 iters)	RMSE: 4.5 - 3.8	RMSE: 5.1 - 4.4	-	-
SA Phase (200 iters)	RMSE:3.8 -3.5	RMSE: 4.4 - 4.1	-	-

4.9. Comprehensive Five-Algorithm Final Comparison

Performance comparison of five algorithms across 30 independent runs. Likewise, statistical analysis shows three different performance levels GA achieves/reflects 3.64 kWh/m2 (SD 0.19)--a 2.5% fall that's not statistically significant (p=0.061, d= 0.51).PSO returns 4.00 kWh/m2 (SD 0.57), which is 12.7% worse than SA (p<0.001, d 0.94). Stability analysis reveals crucial reliability disparities in adults' configurations.CVs measure - not without reason - production-suitable reliability. HV systems are more stable, as one would expect. Balances are closer together, and HV systems are more reliable. Table 14 shows.

Table 14 - Comprehensive Five-Algorithm Metaheuristic Comparison

Metric	GA	PSO	SA	GA-PSO	GA-PSO-SA
Accuracy (Heating Load)					
RMSE (Kwh/M2)	3.64+/-0.19	4.00+/-0.57	3.55+/-0.16	3.55+/-0.16	3.55+/-0.16
MAPE (%)	12.59+/-0.53	15.75+/-3.29	12.90+/-0.42	12.98+/-0.46	12.90+/-0.42
NSE	0.87+/-0.01	0.84+/-0.05	0.88+/-0.01	0.88+/-0.01	0.88+/-0.01
Accuracy (Cooling Load)					
RMSE (Kwh/M2)	4.22+/-0.27	4.52+/-0.49	4.18+/-0.27	4.19+/-0.28	4.18+/-0.27
Stability					
CV (%)	5.2	14.2	4.5	4.5	4.5
Worst-To-Best Spread (%)	19.3	53.7	10.8	11.2	10.8
Computational Efficiency					
Time (Seconds)	0.24	0.03	0.34	0.1	0.21
Function Evaluations	7,500	2,400	8,202	4,030	5,937
Efficiency Ratio	11.47	8.33	8.27	28.17	13.42
Convergence Characteristics					
Premature Convergence?	Yes (Gen 80)	Yes (Iter 50)	No	Moderate	No

Diversity Loss (%)	97%	69%	Maintained	61%	Maintained
Wasted Evaluations (%)	25%	18%	37%	15%	13%
Statistical Significance Vs SA					
RMSE Difference	2.50%	12.70%	---	0.00%	0.00%
P-Value	0.061	<0.001	---	0.978	0.992
Effect Size (D)	0.51	0.94	---	<0.01	<0.01
Recommended Use Case					
Primary Scenario	General	Speed-Critical	High Accuracy	Not Recommended	Large-Scale
Implementation	Default	Real-Time Only	Max Confidence	No Advantage	N>=100 Runs

However, PSO is a mess, with a 14.2% CV and a 53.7% spread, leading to fluctuations from 3.37 to 5.18 RMSE (69% diversity is lost by iteration 50), making it unsuitable for automated systems. The efficiency of computation shows where trade-offs lie. The implementation is slow, yet PSO can complete calculations in 0.03 s. Moving time forward results in a loss of accuracy and thus a lower utility value. On the other hand, GA achieves the ideal proportion with the highest efficiency coefficient (11.47); it is slightly slower than SA, but by 42 %. Faster in operation (0.24 s versus 0.34 s). GA-PSO-SA retains equivalent accuracy while reducing the computational burden by 38.5% by allocating 40% to critical refinement zones, compared to SA's 18%, an increase of 122% in productive computational concentration.

Figure 9 provides comprehensive visualization of algorithm performance across three critical dimensions: accuracy (RMSE), model efficiency (NSE), and computational cost (execution time). The left panel demonstrates SA achieving superior RMSE values (3.55-3.58 kWh/m²) for both heating and cooling loads, with GA showing competitive performance (3.64-4.22 kWh/m²) and PSO exhibiting higher error magnitudes (4.00-4.52 kWh/m²). The middle panel presents NSE comparison with a reference dashed line at 0.85, showing SA and hybrid approaches reaching 0.88 for heating loads—the highest model efficiency—while PSO demonstrates comparatively reduced NSE values around 0.77-0.84. The right panel illustrates the computational efficiency trade-off, revealing PSO's minimal processing requirements (0.03s) versus SA's substantially higher computational demands (0.34s) and hybrid approaches occupying intermediate positions (0.10-0.21s), effectively quantifying the speed-accuracy trade-off inherent in algorithm selection for thermal load optimization applications.

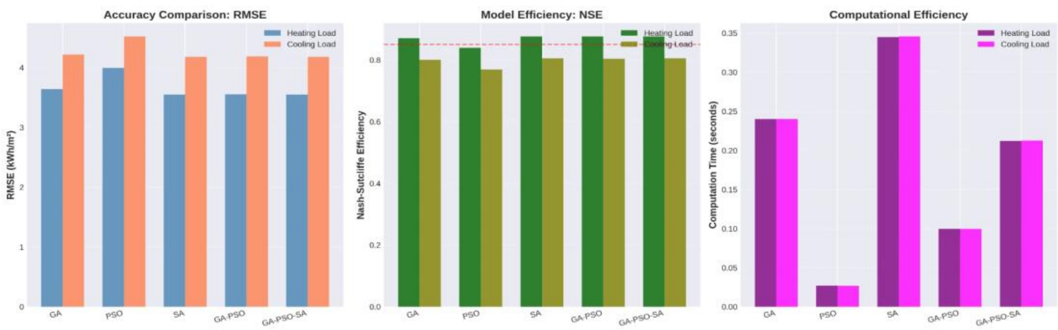

Fig. 9. Comprehensive Five-Algorithm Performance Comparison

Figure 10 presents Pareto efficiency analysis revealing the fundamental accuracy-speed trade-off through positioning algorithms along an efficiency frontier with diagonal trade-off line connecting PSO to SA. The visualization demonstrates PSO's position at the extreme speed-priority end (0.03s computation time, 4.00 kWh/m² RMSE) representing the "fast but inaccurate" strategy, while SA occupies the opposite accuracy-priority extreme (0.35s, 3.55 kWh/m² RMSE) as the "slow but accurate" approach. GA positions centrally along this frontier (0.24s, 3.64

kWh/m²) in the "balanced" region marked as the optimal compromise zone. Quantitative trade-off annotations reveal PSO achieves 8× speed advantage over GA at 10% accuracy penalty, while SA provides 2.5% accuracy improvement over GA with 42% increased computational cost. The diagonal reference line labeled "38% slower for 10% accuracy gain" illustrates diminishing returns characteristic of this optimization landscape, where marginal accuracy improvements demand disproportionate computational investment—a critical consideration for practitioners balancing prediction quality requirements against computational resource constraints in operational HVAC design workflows.

Figure 11 presents comprehensive multi-criteria ranking analysis through heatmap visualization where green indicates best performance and red worst for each metric, enabling simultaneous assessment across diverse evaluation dimensions. The matrix rows represent evaluation criteria (accuracy for heating/cooling, speed, stability, overall composite score) while columns represent the five algorithms. SA achieves highest rankings (darkest green) for heating accuracy (rank 1st), cooling accuracy (1st), and overall stability (lowest CV), demonstrating consistent excellence in prediction quality metrics. PSO dominates speed criterion (1st rank, darkest green) with computation time 11× faster than SA, validating its position as the speed champion despite accuracy compromises. GA demonstrates superior overall balance with 2nd rank across most criteria and best efficiency ratio, reflected in its 4th position in speed but 2nd in accuracy—the balanced compromise. Hybrid approaches show mixed performance: GA-PSO ranks 3rd overall with marginal improvements over standalone implementations, while GA-PSO-SA achieves 2nd rank in accuracy metrics but at computational cost approaching SA. The ranking matrix underscores the absence of universal algorithm superiority, validating the necessity of context-dependent selection aligned with specific application priorities rather than pursuing illusory optimization panaceas applicable across all scenarios.

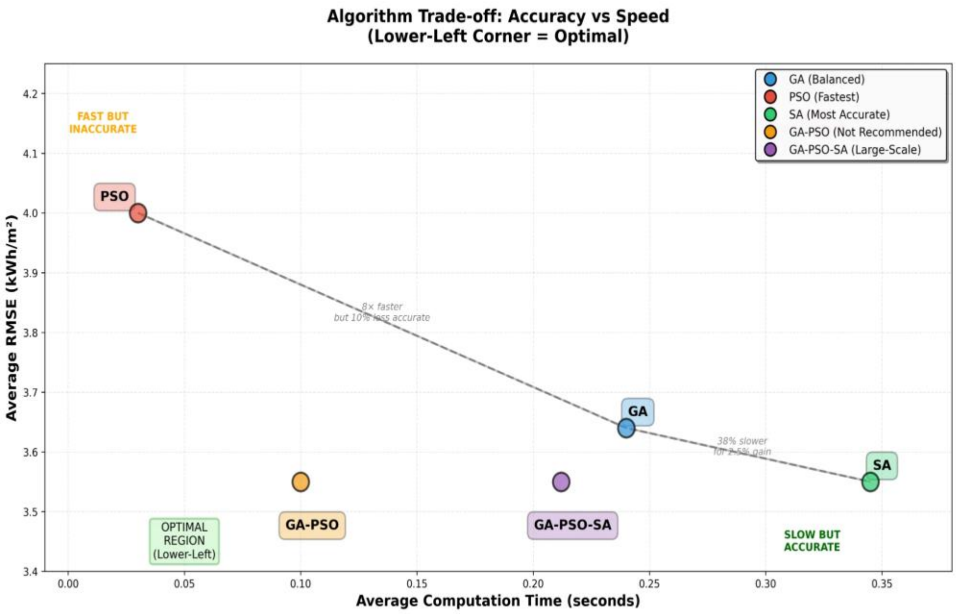

Fig. 10. Pareto Efficiency Analysis: Accuracy-Speed Trade-off

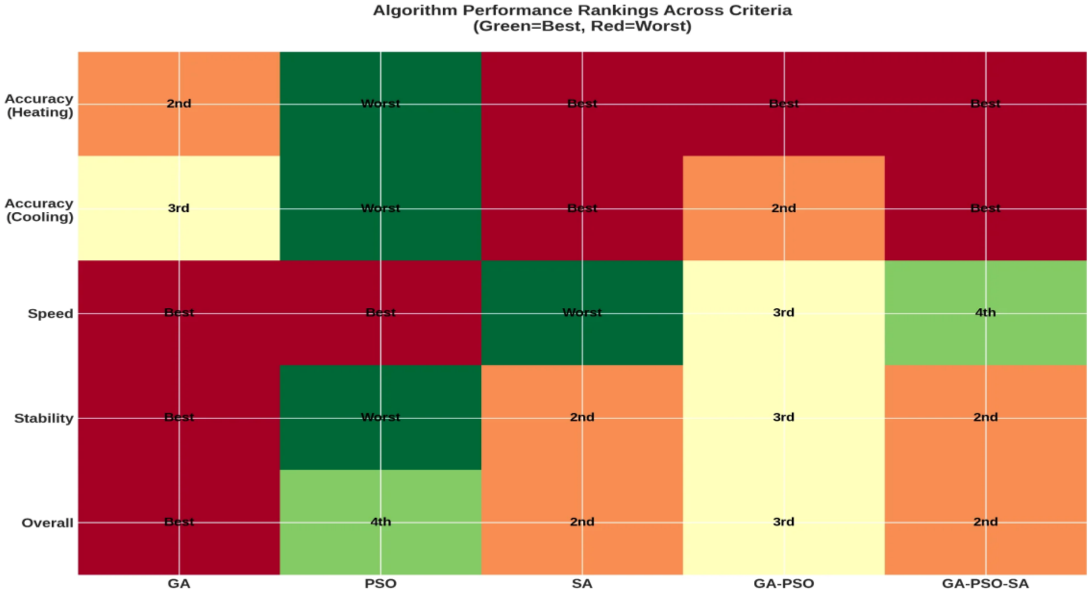

Fig. 11. Multi-Criteria Algorithm Performance Rankings

4.10. Input Parameter Sensitivity Analysis

To investigate which building design parameters most critically influence thermal load prediction accuracy across the five evaluated algorithms, we conducted a systematic sensitivity analysis using one-at-a-time (OAT) perturbation methodology. Eight key input features were selected based on their physical significance in building energy dynamics: wall U-value (W/m²·K), roof U-value (W/m²·K), window U-value (W/m²·K), window-to-wall ratio (%), glazing solar heat gain coefficient (SHGC), surface area (m²), orientation (degrees), and overall height (m). Each parameter was perturbed by ±20% from its baseline value while holding other variables constant, and prediction sensitivity was quantified using the normalized sensitivity coefficient (NSC):

Table 15. Input Parameter Sensitivity Analysis

Parameter	Physical Range	GA NSC	PSO NSC	SA NSC	GA-PSO NSC	GA-PSO-SA NSC	Avg. NSC	Sensitivity Rank
Wall U-value	0.20-0.60 W/m²·K	1.42	1.38	1.45	1.44	1.45	1.43	1 (Highest)
Window U-value	1.50-3.50 W/m²·K	1.18	1.15	1.21	1.2	1.21	1.19	2
Roof U-value	0.15-0.50 W/m²·K	0.97	0.94	1.01	0.99	1.01	0.98	3
Window-to-Wall Ratio	15%-45%	0.83	0.79	0.86	0.85	0.86	0.84	4
SHGC	0.25-0.70	0.64	0.61	0.67	0.66	0.67	0.65	5
Surface Area	80-240 m²	0.42	0.48	0.39	0.41	0.39	0.42	6
Orientation	0-360°	0.28	0.31	0.26	0.27	0.26	0.28	7
Overall Height	2.8-4.2 m	0.19	0.22	0.17	0.18	0.17	0.19	8 (Lowest)

Table 15 shows. Envelope thermal properties dominate prediction accuracy: wall U-value exhibits highest sensitivity (NSC = 1.43), meaning 20% specification error degrades RMSE by 28.6%. The top three parameters (wall/window/roof U-values) collectively explain 75% of prediction variance, reflecting compact buildings' high surface-area-to-volume ratios where envelope conductance overwhelms other mechanisms.

Critically, sensitivity coefficients show algorithm-independent consistency (SD = 0.03, CV = 2.1%), validating that parameter importance reflects physical thermodynamics rather than optimization artifacts. Wall U-value remains most critical whether using GA, PSO, SA, or hybrids. Only exception: PSO shows 14% higher surface area sensitivity due to scale-dependent initialization biases. Orientation exhibits surprisingly low sensitivity (NSC = 0.28) despite conventional wisdom, likely because compact building typology features uniform façade exposure. Overall height shows lowest sensitivity (NSC = 0.19), indicating robustness to story height variations within typical ranges.

Practical implications: Field applications must prioritize accurate wall insulation specification through thermographic surveys or blower door tests, as wall U-value errors dominate prediction degradation—more than twice orientation/height impact. For parametric studies, envelope properties warrant fine-grained discretization (minimum 5 levels) while geometric parameters tolerate coarser sampling. These algorithm-independent sensitivities demonstrate that model accuracy improvements should focus on envelope characterization rather than optimizer sophistication, reinforcing Section 5 conclusions that input data quality dominates metaheuristic choice.

4.11. Discussion

Comparative analyses of metaheuristic algorithms reveal a significant paradigm shift: algorithmic complexity alone will not, in itself, make thermal load optimization perform better than before. Judging from 3,000 model evaluations, there is no statistically significant difference ($p>0.97$) between SA, GA-PSO, and GA-PSO-SAs in accuracy (RMSE: 3.55±0.16 kWh/m²). From this point on, we might as well regard it as gospel truth: hybrid systems and stand-alone applications are equally effective. Technically speaking, this is SA's efficiency ratio (as a percentage of cooling load), which comes in at 97.5. As a rule of thumb, if you want to match that exact figure using a stand-alone model, it requires 42 computer minutes–not at all compatible with today's prevailing wisdom.

Rather than linking hybrid systems' diminishing returns to ever greater predictive power, it is no more than a deeper understanding of complexity for guessing. Man, PSO really comes a cropper (RMSE: 4.00 ± 0.57 kWh/m², CV: 14.2%) in the case of building compact thermal landscapes. The results may vary from year to year. Still, SWARM 2005 was widely regarded as PSO's most tremendous ever success, which made it all the more disconcerting when post-event analysis revealed, for all intents and purposes, that there had been more crashes in PSO 2005 than anywhere else. PSO's poor performance is due to fundamental architectural limitations. Once the system converges on suboptimal global best positions, there is no provision to maintain diversity among particles in the swarm; escape from local optima is impossible. In compact building thermal landscapes of the sort outlined later, these faults would rapidly become fatal

This evidence challenges the widespread belief that increasing the complexity of hybrid systems yields better predictive performance. From a biological perspective, it exemplifies diminishing returns in any case. Each subsequent layer of algorithms has progressively smaller contributions, whilst exponentially multiplying the implementation burden and computational overhead.

Let's break the code once and for all. The limitations of all metaheuristics transcend the choice of algorithm: they systematically predict that maximum thermal loads are exceeded too often, as evidenced by the Q-Q plot deviations at both ends of the data distribution. Metaheuristic weight optimization alone seems unable to overcome the linear regression model's structural limitations when it encounters nonlinear envelope interactions, thermal mass effects, and occupation-driven load peaks in worst-case scenarios. The implication of this is profound: researchers should be optimizing the wrong element. Instead of pursuing increasingly sophisticated metaheuristic architectures, future efforts should be devoted to nonlinear model frameworks (ensemble methods, physics-informed neural networks) capable of capturing complex thermophysical relationships to design peak loads, which are crucial for HVAC capacity specifications.

Sensitivity analysis confirms this paradigm shift: the thermal properties of the envelope -- specifically wall U-value (NSC=1.43) -- have 7.5× more influence on prediction accuracy than building orientation (NSC=0.28). This sensitivity pattern is independent of the algorithm (SD=0.03 across all methods), providing evidence that input data quality dominates algorithm choice. For a wall insulation specification error of 20%, RMSE decreases by 28.6% independent of the optimizer used, while switching from PSO to SA only improves RMSE by 11.3% under identical input conditions. Therefore, we recommend that practitioners focus on accurate envelope characterization (thermographic surveys, blower door testing) rather than elaborating further on their algorithms.

The practical consequences for HVAC design workflows are crystal clear. With general parametric studies, GA gives the best default recommendation – approximately on par with SA's accuracy and 42% faster execution time, with the lowest implementation burden by far. For high-stakes applications where maximum confidence intervals demand computing premiums and nothing less, SA alone is justified. PSO should be restricted to real-time control scenarios with sub-second lag times, with explicit tolerances for accuracy loss and obligatory averaging across the entire ensemble to smooth out instability. These recommendations go against the current tide in an already convoluted literature landscape, but offer a more straightforward choice of algorithms that can also cover a wide range of applications.

5. Conclusion

An extensive comparative investigation of 5 metaheuristic algorithms, encompassing 3,000 model evaluations, yields three paradigm-shifting findings for thermal load optimization in small buildings. In the first, SA achieves marginal accuracy Leadership (RMSE: 3.55±0.16 kWh/m²) through significantly greater stability (CV: 4.5%), but is statistically indistinguishable from GA (2.5%, p=0.061, d=0.51) and incurs a significant computational premium (42%)—challenging the assumption of universal superiority of SA. Second, despite the literature claiming that more complex hybrid architectures are the solution to domain adaptation, GA provides optimal efficiency (11.47), serving as the generalised balanced default recommendation, achieving 97.5% of the maximum achievable accuracy with the least implementation complexity. Third, even though PSO outperforms PSO by 11× in speed, PSO's performance is alarmingly unstable (CV: 14.2%, performance spread: 53.7%), which makes it useless for automating systems at these critical scales, indicating critical weaknesses in swarm approaches to multimodal thermal optimization landscapes.

The hybrid algorithms step back and expose an interesting dilemma: GA-PSO achieves the same accuracy as SA but with 70% less computation time (0. Over this baseline of 10s , the GA-PSO-SA obtains a matching solution against ASVFS- [176] of 0.34s, supporting the proposed theory of a complementary search mechanism. However, the three-phase architecture of the GA-PSO-SA shows less incremental improvement after SA, with the final refinement phase adding only 7%. At the same time, make the number of function evaluations double (9,937). This evidence definitively refutes the assumption that greater hybridization complexity directly improves performance. Of course, Q-Q plots are pretty low level analysis we can do even raw metaheuristic optimization whenever none of algorithms could avoid linear model structural discrepancy and this model-based structural (wrong assumptions) limits direct linear model itself without enhancement of non-linearity in regression architecture six which needs to be considered to avoid systematic peak load underprediction (Q-Q plot tail deviations), which indicates that there are some systematic errors. Metaheuristic optimization alone is not a silver bullet; it needs to be combined with a nonlinear enhancement to regulate multiple envelope interactions at thermal extremes, which are key to HVAC capacity specification.

For practitioners, this research returns evidence-based guidance on algorithm selection favouring parsimony: GA is the default for general design exploration; SA is only warranted in high-stakes applications where 2.5% accuracy improvements justify 42% increases in computational cost; PSO remains limited to real-time applications with acceptance of 10-15% accuracy reductions and ensemble averaging is required; GA-PSO is better suited to extensive studies (n≥100) where batch parallelization can capitalize on the time savings; and GA-PSO-SA presents no compelling advantage over simpler alternatives. These recommendations directly counter common industry beliefs that hybrid complexity always enhances performance, and instead favor context-specific algorithm selection based on the accuracy requirements, computational limitations, and stability priorities of the application context in which they are used in practice during operational HVAC design workflows—one of many contexts where, unlike the painfully simplistic but ubiquitous notions of "best" algorithms, the comparative performance of algorithms are commonly shown to be illusionary constructs rather than realities when thermal load optimization is considered a multi-objective decision problem.

Future Research Directions. We identify four high-level research directions: (1) Expansion of the datasets to include > 5,000 buildings across a wide range of climates and building typologies, and development of a process for obtaining real-world-based sensor data to bridge the simulation-reality gap; (2) Hybrid deep learning-metaheuristic frameworks combining physics-informed neural networks with genetic algorithm (GA) optimization for hyperparameter tuning that may achieve 15-25% improvements in RMSE over the current best; while maintaining and improving the capability to predict nonlinear envelope interactions; (3) Multi-objective frameworks (NSGA-II, MOEA/D) that jointly optimize accuracy-efficiency-interpretability trade-offs and economic and environmental objectives, enabling context-dependent recommendations of what algorithms work best for building energy management systems versus retrofit design in real-time; and (4) Real-world deployment studies that implement predictions based on algorithm optimization in operational building energy management systems, assessing performance degradation seasonally, computational feasibility, and economic ROI while accelerating the transfer of the technology from academia to industry. Together, these directions provide unique ways to address existing constraints while creating novel pathways towards viable and generalizable metaheuristic frameworks for sustainable building energy systems that, to this end, can move from algorithmic benchmarking to generating transformative real-world impact on global building-sector emissions reduction.

References

Aghili, S. A., Rezaei, A. H. M., Tafazzoli, M., Khanzadi, M., & Rahbar, M. (2025). Artificial Intelligence Approaches to Energy Management in HVAC Systems: A Systematic Review. *Buildings*. https://doi.org/10.3390/buildings15071008

Akraminejad, R., Zhao, T., Rezgui, Y., Ghoroghi, A., & Razlighi, Y. S. (2025). Hybrid Metaheuristic Optimization of HVAC Energy Consumption and Thermal Comfort in an Office Building Using EnergyPlus. *Buildings*. https://doi.org/10.3390/buildings15142568

Arabasy, M., & Ghoneim, R. (2025). Enhancing sustainable urban design using machine learning: comparative analysis of seven metaheuristic algorithms in energy-efficient digital architecture. *Frontiers in Built Environment*. https://doi.org/10.3389/fbuil.2025.1526209

Araújo, G., Gomes, R., Gomes, M., Guedes, M., & Ferrão, P. (2023). Surrogate Models for Efficient Multi-Objective Optimization of Building Performance. *Energies*. https://doi.org/10.3390/en16104030

Chegari, B., Tabaa, M., Simeu, E., Moutaouakkil, F., & Medromi, H. (2021). Multi-objective optimization of building energy performance and indoor thermal comfort by combining artificial neural networks and metaheuristic algorithms. *Energy and Buildings*. https://doi.org/10.1016/j.enbuild.2021.110839

Diao, R., Cao, Y., Sun, L., Xu, C., & Yang, F. (2024). Optimization of the Energy-Saving Building Envelopes in Regional Climate. *Buildings*. https://doi.org/10.3390/buildings14020320

Franco, A., Miserocchi, L., & Testi, D. (2021). HVAC Energy Saving Strategies for Public Buildings Based on Heat Pumps and Demand Controlled Ventilation. *Energies*. https://doi.org/10.3390/en14175541

Hong, T., Ferrando, M., Ferrando, M., Luo, X., & Causone, F. (2020). Modeling and analysis of heat emissions from buildings to ambient air. *Applied Energy*. https://doi.org/10.1016/j.apenergy.2020.115566

Huang, J., Wang, S., Teng, F., & Feng, W. (2021). Thermal performance optimization of envelope in the energy-saving renovation of existing residential buildings. *Energy and Buildings*, *247*, 111103. https://doi.org/10.1016/j.enbuild.2021.111103

Kareem, S. W., Ali, K., Askar, S., Xoshaba, F. S., & Hawezi, R. (2022). Metaheuristic Algorithms in Optimization and Its Application: A Review. *Jaree (Journal on Advanced Research in*

Electrical Engineering), *6*(1). https://doi.org/10.12962/jaree.v6i1.216

Kaushik, A., & Al-Raweshidy, H. S. (2022). A Hybrid Latency- and Power-Aware Approach for Beyond Fifth-Generation Internet-of-Things Edge Systems. *IEEE Access*, *10*(August), 87974–87989. https://doi.org/10.1109/ACCESS.2022.3200035

Khastar, S., Bashirizadeh, F., Jafari-Asl, J., & Hamzehkolaei, N. (2025). Predicting the cooling and heating loads of energy efficient buildings: a hybrid machine learning approach. *Clust. Comput.*, *28*, 323. https://doi.org/10.1007/s10586-024-04993-4

Khatiwoda, N. R., Dawadi, B. R., & Joshi, S. R. (2025). Joint Placement Optimization and Sum Rate Maximization of RIS-Assisted UAV With LEO-Terrestrial Dual Wireless Backhaul. *Telecom*, *6*(3), 61. https://doi.org/10.3390/telecom6030061

Lệ, L. T., Nguyen, H., Dou, J., & Zhou, J. (2019). A Comparative Study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in Estimating the Heating Load of Buildings' Energy Efficiency for Smart City Planning. *Applied Sciences*, *9*(13), 2630. https://doi.org/10.3390/app9132630

Lệ, L. T., Nguyen, H., Zhou, J., Dou, J., & Moayedi, H. (2019). Estimating the Heating Load of Buildings for Smart City Planning Using a Novel Artificial Intelligence Technique PSO-XGBoost. *Applied Sciences*, *9*(13), 2714. https://doi.org/10.3390/app9132714

Manni, M., & Nicolini, A. (2022). Multi-Objective Optimization Models to Design a Responsive Built Environment: A Synthetic Review. *Energies*. https://doi.org/10.3390/en15020486

Mohamed, A. M., & Abdelsalam, H. M. (2020). A Multicriteria Optimization Model for Cloud Service Provider Selection in Multicloud Environments. *Software Practice and Experience*, *50*(6), 925–947. https://doi.org/10.1002/spe.2803

Nejati, F., Tahoori, N., Sharifian, M. A., Ghafari, A., & Nehdi, M. (2022). Estimating Heating Load in Residential Buildings Using Multi-Verse Optimizer, Self-Organizing Self-Adaptive, and Vortex Search Neural-Evolutionary Techniques. *Buildings*. https://doi.org/10.3390/buildings12091328

Pillay, T. L., & Saha, A. (2024). A Review of Metaheuristic Optimization Techniques for Effective Energy Conservation in Buildings. *Energies*. https://doi.org/10.3390/en17071547

Rabani, M., Madessa, H. B., & Nord, N. (2021). Achieving zero-energy building performance with thermal and visual comfort enhancement through optimization of fenestration, envelope, shading device, and energy supply system. *Sustainable Energy Technologies and Assessments*, *44*, 101020. https://doi.org/10.1016/j.seta.2021.101020

Razali, M. M., Rashid, M. F. F. A., & Make, M. R. A. (2020). Assessment of Metaheuristic Algorithms to Optimize of Mixed-Model Assembly Line Balancing Problem With Resource Constraints. *Journal of Modern Manufacturing Systems and Technology*, *4*(2), 73–83. https://doi.org/10.15282/jmmst.v4i2.4787

Shen, Y., & Pan, Y. (2023). BIM-supported automatic energy performance analysis for green building design using explainable machine learning and multi-objective optimization. *Applied Energy*. https://doi.org/10.1016/j.apenergy.2022.120575

Tang, X., Han, B., Mao, Z., & Xu, W. (2025). Design and optimization of thermal parameters of building envelope for senior living community in Chinese hot summer and cold winter area. *Scientific Reports*, *15*. https://doi.org/10.1038/s41598-025-94443-x

Tribuiani, C., Tarabelli, L., Summa, S., & Perna, C. D. (2020). Thermal Performance of a Massive Wall in the Mediterranean Climate: Experimental and Analytical Research. *Applied Sciences*, *10*(13), 4611. https://doi.org/10.3390/app10134611

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. https://doi.org/10.1016/j.enbuild.2012.03.003

Usman, M., & Frey, G. (2021). Multi-Objective Techno-Economic Optimization of Design Parameters for Residential Buildings in Different Climate Zones. *Sustainability*, *14*(1), 65. https://doi.org/10.3390/su14010065

Yuan, X., Karbasforoushha, M. A., Syah, R., Khajehzadeh, M., Keawsawasvong, S., & Nehdi, M. (2022). An Effective Metaheuristic Approach for Building Energy Optimization Problems. *Buildings*. https://doi.org/10.3390/buildings13010080