

An Artificial Intelligence–Enabled Assessment Framework for Energy Conservation in Retrofitted and Operational Buildings

Zhiyong Wu¹, Yun Qiao^{3,*}, Youren Ma^{1,2}

¹ Blue Star Tech Pte Ltd, Singapore

² Department of Mathematics, National University of Singapore, Singapore, 119076, Singapore

³ Jiang Su Xinglin Consulting Co Ltd, Nanjing, 210009, China

Abstract. Accelerating energy efficiency improvements in the built environment is central to ASEAN’s decarbonization pathway and long-term climate commitments. However, the diverse operational profiles, climatic conditions, and retrofit strategies across the region pose significant challenges for consistent and credible assessment of energy conservation performance. To address this gap, this paper presents a standardized, data-driven evaluation framework that harnesses artificial intelligence (AI) and machine learning (ML) to analyze and verify energy performance in both retrofitted and actively operating buildings. Designed to be compatible with regional policies—including the objectives of the ASEAN Plan of Action for Energy Cooperation (APAEC), national green building rating tools, and emerging carbon management regulations—the framework integrates core indicators such as Energy Use Intensity (EUI), temporal operation patterns, equipment load profiles, and localized meteorological datasets. These inputs are synthesized to establish adaptive baseline models capable of capturing real-time performance variations and quantifying energy savings with higher accuracy than conventional static or rule-based methods. The proposed methodology is structured for application across commercial, institutional, and mixed-use developments, enabling facility owners, policymakers, and financial institutions to systematically track conservation outcomes and evaluate the effectiveness of retrofit interventions. Beyond performance verification, the framework is positioned to support regional green financing ecosystems, including sustainability-linked incentives, performance-based procurement, and digital MRV (measurement, reporting, and verification) systems increasingly adopted across ASEAN.

Keywords: Energy Conservation; AI-Driven Modeling; Building Energy Efficiency; Green Mark Certification; Sustainability-Linked Financing

1. Introduction

The building sector remains one of the most intensive consumers of global energy, routinely accounting for close to 40% of worldwide demand. Yet, despite decades of research promoting “energy-efficient buildings,” the tools used to measure savings—particularly for retrofitted and continuously operating facilities—still fall short of providing credible or context-sensitive results. Much of the difficulty stems from the persistent use of rigid, boundary-based assessment frameworks that assume static occupant behavior, stable operating conditions, and uniform environmental influences. These simplified formulations underestimate the complexity of real buildings, where energy performance is shaped by fluctuating usage patterns, evolving technologies, and climatic variability. As a result, the calculated energy savings often deviate significantly from the actual conservation outcomes, distorting our understanding of

whether sustainability targets are genuinely being met [1]. A critical vulnerability of prevailing methods lies in their dependence on fixed operational parameters that fail to capture the dynamic interplay between occupant activities, equipment schedules, and building envelope behavior—factors collectively governing effective energy utilization (EUAC) [1]. These static and overgeneralized assumptions frequently lead to optimistic savings estimates, creating a misleading impression of success. Errors at this foundational level have downstream consequences: they impair managerial decision-making, complicate regulatory compliance, and undermine the credibility of energy-efficiency certifications that organizations rely upon to demonstrate environmental performance [2].

To meaningfully evaluate the risks inherent in current assessment practices, it is necessary to scrutinize the blind spots embedded within traditional modeling approaches. Conventional energy calculation frameworks rarely

* Corresponding author: yunqiao@188.com

account for substantial changes in building use—such as adaptive reuse, shifting occupancy levels, or the integration of advanced materials and digital control systems—each of which can significantly reshape energy demand profiles. These methodological gaps can promote both underestimation and overestimation. In some cases, buildings fall short of promised savings, disappointing stakeholders and raising questions about the return on investment. In others, conservative or outdated baseline assumptions obscure substantial unrealized financial benefits that would otherwise justify further conservation measures [3]. In response to these limitations, researchers have increasingly explored the use of artificial intelligence to reconfigure the analytical foundation of energy-savings assessment. AI-enabled models are capable of continually ingesting real-time data, identifying evolving consumption patterns, and representing the actual performance of energy conservation measures (ECMs) with far greater fidelity than deterministic tools. By dynamically updating predictions based on new information, these models offer a more reliable platform for verifying whether energy-efficiency interventions are performing as intended. Improved transparency and empirical grounding also enhance confidence among building operators and regulatory authorities. However, this shift toward data-intensive decision support brings its own set of challenges. The effectiveness of AI-driven assessments hinges on the availability of rich, high-resolution datasets—something that many buildings, particularly older ones, cannot yet provide. Moreover, operating and interpreting these systems requires technical expertise that may exceed the capabilities or budgets of many facility management teams. The transition to AI-supported conservation therefore demands careful planning to avoid creating new inequities or operational burdens.

2. Reconceptualizing Digital Intelligence in Building Energy Evaluation

2.1 Reframing Conventional Assumptions: Why Existing Models Remain Structurally Insufficient

Efforts to quantify building energy efficiency have traditionally been anchored in linear or quasi-static models that treat occupant behaviour, equipment loads, and operational schedules as largely predictable. This assumption persists despite decades of evidence demonstrating that real buildings behave as complex, adaptive systems. Such rigidity creates a structural blind spot: calculation methods often fail to incorporate the variability introduced by weather anomalies, shifting space-use patterns, or intermittent adoption of new energy technologies [3]. These shortcomings are not merely technical inconveniences. They fundamentally undermine the credibility of energy savings assessments, as inaccurate baselines either inflate performance or obscure genuine conservation potential [1]. In many settings, these inaccuracies shape policy compliance, resource allocation, and certification outcomes, meaning that methodological

errors propagate into institutional decision-making. Recognizing these limitations is therefore an essential precursor to any meaningful deployment of advanced analytics or AI systems in this domain [2].

2.2 Foundations of AI-Enabled Energy Modelling: Mechanisms, Capabilities, and Data Demands

Artificial intelligence introduces an entirely different analytical logic into building energy evaluation. Whereas conventional approaches rely on predetermined relationships, AI systems identify emergent patterns within high-dimensional data [3]. Machine learning architectures—ranging from gradient-boosted decision trees and recurrent neural networks to hybrid statistical–AI pipelines—are able to model nonlinear, temporally dependent energy behaviours. These systems leverage data from multiple sources: IoT sensors tracking temperature and occupancy, submetered electricity flows, and high-resolution meteorological datasets [1]. Before such systems achieve accuracy, however, rigorous data conditioning is required. Preprocessing operations such as signal smoothing, anomaly exclusion, and engineered feature construction determine whether the model extracts meaningful insights or merely reacts to noise. When executed effectively, these algorithmic foundations enable the development of dynamic baselines that continuously adjust to new operational and environmental information [1].

2.3 Emerging Applications and Innovation Pathways in AI-Based Energy Conservation

Rapid digitalization across the built environment has expanded the possibilities for AI-driven conservation strategies. Key innovation trajectories include:

- **Real-Time Supervisory Optimisation:** AI agents capable of continuously simulating and selecting optimal HVAC and lighting configurations, adapting to occupancy, weather shifts, or operational constraints.
- **Distributed Energy Integration:** Predictive models that forecast fluctuations in rooftop solar yield or local microgrid availability, enabling more reliable coordination between renewable production and building consumption.
- **Predictive Maintenance Ecosystems:** Algorithmic detection of abnormal energy signatures that indicate equipment degradation—an approach shown to reduce energy waste associated with underperforming mechanical systems.
- **Data-Driven Regulatory Instruments:** Governments in several regions are now exploring how AI-derived savings estimates can be incorporated into official verification pathways, provided that transparency, reproducibility, and auditability standards are upheld.

These innovations demonstrate the expanding role of AI not only as an analytical tool but as an operational partner in managing building-level decarbonisation [2].

2.4 Systemic Barriers and Critical Risks in Large-Scale Adoption

Despite the appeal of AI-driven energy evaluation, several structural challenges remain unresolved. High-quality granular data—required for model training—are still unavailable in many older facilities or in regions where submetering infrastructure is limited. Moreover, the integration of AI models into existing building management systems is often hindered by incompatible protocols, proprietary technologies, and fragmented digital ecosystems. Concerns regarding data governance, cybersecurity, and the opacity of certain machine learning models also pose legitimate barriers to adoption. Without clear explainability and verifiable logic, AI-generated savings claims risk being contested by regulators or auditors. Additionally, the operational expertise required to maintain and interpret these systems can impose significant costs, widening gaps between well-resourced organizations and smaller operators that stand to benefit most from improved efficiency.

3. Methods

3.1 Overview of the Analytical Workflow

The methodological approach adopted in this study combines data-driven modelling, multi-source data harmonisation, and iterative performance validation. Instead of employing a single, fixed architecture, the workflow was intentionally designed as a modular system capable of incorporating heterogeneous datasets and adjusting dynamically to varying operational conditions. The process consists of four interdependent stages:

- Stage 1: **Acquisition and harmonisation of granular building-level datasets**
- Stage 2: **Construction of a hybrid machine-learning framework for baseline reconstruction**
- Stage 3: **Iterative model refinement through cross-validated learning cycles**
- Stage 4: **Quantitative appraisal of predictive performance using multiple statistical indicators**

3.2 Development of the Hybrid Learning Engine

3.2.1 Rationale for a Hybrid Approach

Rather than depending on a single modelling philosophy—whether purely deep-learning-based or confined to traditional regression—the analytical engine was deliberately constructed as a blended architecture capable of recognising broad behavioural trends while simultaneously capturing localised operational rules. This choice was motivated by the inherently uneven and dynamic nature of building energy behaviour, where

fluctuations in occupancy, rapid weather transitions, and the effects of various retrofit interventions interact in ways that no single algorithmic approach can fully represent. To address these complexities, the framework brings together several complementary learning mechanisms within one coherent structure. The first layer draws on gradient-driven neural networks, which are particularly effective at uncovering nonlinear and time-dependent relationships embedded in long energy-use sequences [2]. A second analytical layer employs ensemble tree methods—such as gradient boosting and random-forest variants—to segment and interpret periods of atypical or volatile consumption, providing a rule-based understanding of high-variance operational regimes. These components are then anchored by a regularisation-oriented regression layer, introduced to temper overfitting, enhance transparency, and ensure that the resulting predictions remain grounded in interpretable patterns rather than noise. Through this multi-tiered configuration, the model is able to reconcile large-scale behavioural structure with fine-grained decision logic, producing a more stable and context-responsive representation of building energy performance.

3.2.2 Mathematical Representation

For any building i at time t , predicted energy use $\hat{E}_i(t)$ is expressed as Formula (1):

$$\hat{E}_i(t) = f_{\text{DNN}}(X_{i,t}) + g_{\text{Tree}}(X_{i,t}) + \lambda h_{\text{Reg}}(X_{i,t}) \dots \dots \dots (1)$$

Where $X_{i,t}$ = vector of building attributes, weather conditions, occupancy indicators; f_{DNN} = nonlinear temporal extraction; g_{Tree} = rule-based segmentation of operational regimes; h_{Reg} = linear interpretability layer; and λ = regularisation coefficient. The combined output is subsequently calibrated using stochastic gradient descent with adaptive learning rates [2].

3.3 Data Ecosystem and Pre-Model Processing

The study utilised a multi-year dataset (2022–2024) encompassing commercial offices, residential developments and mixed-use complexes across Singapore, combining hourly electricity profiles, submetered HVAC, lighting and plug-load data, and rich metadata on building age, floor area, envelope characteristics, retrofit history and occupancy patterns [1]. These records were integrated with high-resolution meteorological observations—including temperature, solar radiation, humidity and wind—to capture environmental influences on demand. Due to the heterogeneity of inputs, all measurements were synchronised onto a shared temporal structure, screened for anomalies using statistical thresholds and reconstruction rules, and standardised via z-normalisation, min-max scaling or logarithmic transformation for skewed variables [1]. Missing segments were reconstructed through adaptive interpolation that preserved diurnal load rhythms, and qualitative attributes such as retrofit category or system configuration were encoded into machine-readable formats, yielding a

harmonised dataset suitable for unified modelling (Table 1).

Table 1. Data Conditioning Rules.

Data Category	Treatment Approach	Rationale
Energy Load (kWh)	Robust scaling + anomaly filtration	Reduce distortion from spikes
Weather Data	Hourly resampling + smoothing	Preserve temporal structure
Occupancy Indicators	One-hot encoding	Improve model generalisation
Building Metadata	Standardisation	Normalised feature space
Missing Values	FFT-based interpolation	Retain cyclical patterns

3.4 Training Logic and Model Validation

The model was trained using an iterative, multi-cycle learning strategy rather than a single static run, enabling continuous recalibration of network parameters through backpropagation and repeated refinement of tree-based decision thresholds. Across cycles, Bayesian optimisation adjusted hyperparameters to ensure structural alignment with the dataset’s complexity, and roughly four-fifths of the harmonised dataset was reserved for this evolving training phase. To ensure robustness and generalisability, the procedure incorporated both ten-fold cross-validation and a temporal block validation scheme, exposing the model to diverse and chronologically contiguous subsets of data. Sample ordering and grouping were periodically reconfigured to assess parameter stability, while monitoring for predictive drift safeguarded against performance degradation when encountering new building conditions.

3.5 Performance Assessment

Performance was quantified using a suite of statistical indicators rather than a single metric:

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (E(t) - \hat{E}(t))^2} \dots\dots\dots(2)$$

$$R^2 = 1 - \frac{\sum(E(t) - \hat{E}(t))^2}{\sum(E(t) - \bar{E})^2} \dots\dots\dots(3)$$

Additionally, MAPE, MAE, and Normalized RMSE were calculated to evaluate robustness across varying load intensities. The combined metrics allowed evaluation of error magnitude, variance explanation, sensitivity to consumption peaks, and consistency across building archetypes [1].

4. Results and Discussion

4.1 Model Performance and Predictive Accuracy

Applying the analytical framework to a diverse multi-building dataset demonstrated clear performance gains over conventional models. The hybrid learning engine reproduced load profiles with an average deviation of about ±5%, markedly more precise than typical static baselines that vary by ±10–15%. Its accuracy remained stable across extended periods, reflecting the iterative

training design, cross-validation procedures and drift monitoring described in Chapter 3. By continually updating feature relevance for weather, occupancy and system behaviour, the model maintained reliable predictions even under shifting operating conditions [3]. Quantitative results further showed that buildings using the model to guide operational adjustments achieved meaningful efficiency gains: retrofit cases recorded an average 20% reduction in energy use, with some high-occupancy commercial properties exceeding 25%. These outcomes confirm that improved predictive precision translated into measurable conservation benefits [1].

4.2 Effects of Energy Conservation Measures in Retrofitted Buildings

The model’s ability to reconstruct detailed baselines also allowed a more nuanced evaluation of retrofit interventions than is typically possible through static comparisons. When Energy Use Intensity (EUI) values were examined before and after retrofitting, reductions averaged around 20%, which is consistent with the improved control and scheduling performance suggested by the prediction outputs. Moreover, the model’s interpretability layer—built into the hybrid engine—made it possible to identify the key contributors to these improvements. Several facility teams reported that the insights derived from the model helped them uncover recurrent inefficiencies that would not have been visible using standard energy dashboards. Operating schedules, equipment cycling behaviour, and misaligned HVAC controls emerged as recurrent drivers of avoidable consumption. The granular visibility of these patterns allowed building managers to refine control strategies in ways that aligned more closely with the actual usage of their facilities, rather than relying on fixed schedules or intuitive assumptions. Beyond operational efficiency, the environmental impact of these energy reductions was substantial [1]. Across the portfolio, the decrease in consumption corresponded to an estimated annual reduction of approximately 350 tonnes of CO₂-equivalent emissions. This outcome demonstrates that data-driven optimisation does not merely improve internal performance metrics but contributes meaningfully to broader organisational and regulatory sustainability objectives.

4.3 Model Adaptability Under Variable Conditions

A defining characteristic of the modelling approach developed in this study is its adaptability to real-world variability. The combination of neural-network components and ensemble decision structures enabled the model to respond effectively to rapid shifts in external conditions, particularly those driven by weather variability. During periods when external temperatures deviated significantly from seasonal norms, the model continued to predict cooling and heating loads within a tolerance of ±5%, even when building operations experienced sharp changes in internal demand. This capacity to preserve accuracy under fluctuating conditions validates the methodological emphasis on temporal

learning, introduced in Chapter 3. It also underscores the limitations of static models, which typically struggle to maintain accuracy during weather transitions or periods of atypical occupancy. The model’s performance under variable occupancy conditions was equally notable. Facilities with highly irregular daily and weekly use patterns—such as mixed-use buildings and educational settings—demonstrated some of the strongest gains in prediction accuracy. By aligning model inputs with real-time occupancy signals, the system was able to adjust expected loads dynamically, which translated into more targeted control strategies [2]. This alignment not only prevented unnecessary energy use during under-occupied periods but also improved comfort during peak activity, showing that adaptive energy modelling can contribute both to operational efficiency and occupant experience.

4.4 Comparative Performance Against Traditional Models

When benchmarked directly against traditional static baseline models, the hybrid AI-driven framework showed a marked performance edge in both precision and adaptability. Traditional models typically registered average prediction errors around 15%, reflecting their limited capacity to incorporate emerging behavioural or environmental variations. In contrast, the model

developed in this study reduced that error to about 5%, demonstrating that dynamic learning significantly improves predictive credibility. The divergence between the two approaches became especially apparent when occupancy patterns shifted unexpectedly or when buildings underwent partial system upgrades [2]. Under such conditions, the static models increasingly deviated from actual performance, whereas the hybrid engine recalibrated its outputs through its iterative refinement mechanisms. This adaptability produced more reliable baselines for evaluating conservation measures and reduced the risk of misinterpreting temporary consumption anomalies as persistent inefficiencies [3]. From a qualitative standpoint, the shift to a dynamic modelling framework also prompted a change in how building managers engaged with performance data. Interview feedback indicated that facility teams developed greater confidence in the recommendations generated by the model because the underlying trends were consistent with observed system behaviour. Rather than viewing energy analytics as a compliance-driven reporting exercise, teams began using the predictive outputs to support day-to-day operational decisions, suggesting that the methodology has implications that extend beyond numerical accuracy (Table 2).

Table 2. Comparative Performance of AI-Driven and Traditional Energy Models Across Key Analytical Dimensions.

Analytical Dimension	AI-Driven Hybrid Framework (Observed in Study)	Conventional Static Models (Typical Behaviour)	Empirical Ranges / Benchmarks (ASEAN Context)	Relevance to Methodology (Link to Ch. 3–4)	Expanded Interpretation
Prediction Accuracy	RMSE $\approx \pm 5\%$ across 10 buildings; maintains $< \pm 6\%$ even during extreme weather or irregular usage	RMSE typically $\pm 12\text{--}15\%$, rising to 20% during seasonal transitions	RMSE in ASEAN buildings usually $\pm 10\text{--}18\%$, depending on metering quality	Reflects strength of hybrid learning engine (DNN + ensemble trees) and iterative drift correction	Higher precision strengthens MRV credibility, producing baselines that maintain accuracy across diverse climatic and operational conditions
Energy Savings Realisation	Average energy reduction 18–22%; peak values 25–28% in high-density commercial properties	Often 10–12% predicted pre-retrofit; actual realised $\sim 8\text{--}10\%$ due to inaccurate baselines	ASEAN retrofit projects typically achieve 12–18% without AI optimisation	Dynamic reconstruction captures post-retrofit behavioural and load changes, avoiding over/underestimation	AI-guided operational tuning amplifies retrofit gains, revealing hidden inefficiencies undetectable in static models
Operational Responsiveness	HVAC and lighting loads reduced 12–18% during off-peak periods through real-time adaptive control	No dynamic adjustments; consumption reduction limited to 3–5%, mostly through pre-set timers	Real-time occupancy sensors in ASEAN show 30–50% variability in hourly footfall	Real-time occupancy integration (Chapter 3.3) drives short-term load recalibration	Enhances comfort and efficiency simultaneously by aligning system operation with true occupant presence
Baseline Reliability	Baseline drift $< 2\%$ over 12 months; resilient to system upgrades and behavioural shifts	Baseline drift frequently $> 10\%$ within a year due to static coefficients	High baseline drift in ASEAN buildings due to mixed-use schedules and irregular occupancy	Dynamic baselines reflect retrofit interactions and evolving behaviours	More reliable ECM verification avoids misclassification of temporary anomalies, critical for green financing and certification
Managerial Utility & Decision Support	Provides multi-layer insights into load drivers; identifies 5–7 recurrent inefficiency patterns per building	Only provides daily/monthly totals; limited diagnostic value	Typical building teams require 20–30% improvement in diagnostic clarity to optimise systems	Hybrid architecture provides interpretable decision pathways (feature importance, decision rules)	Enables shift from reactive troubleshooting to proactive optimisation, supported by transparent analytical outputs

Despite the model's strengths, scalability and interpretability challenges remain. Many buildings still rely on fragmented management systems, legacy hardware and proprietary data formats, making portfolio-wide deployment difficult and highlighting the need for institutional capacity, standardised protocols and interoperable digital infrastructures. Progress in unified data standards, automated ingestion pipelines and cross-platform interfaces would substantially ease integration barriers [2]. Interpretability also warrants further attention: although the hybrid architecture offers partial transparency through its regularised and tree-based components, the neural elements remain complex and may be difficult for facility teams to interpret. Developing methods that enhance clarity without compromising nonlinear learning capacity is therefore essential for broader operational adoption of AI-driven models.

5. Conclusion

This study illustrates how a carefully structured, data-driven modelling framework can materially improve the evaluation and management of energy use in operational and retrofitted buildings. By unifying dynamic baseline reconstruction with real-time learning from environmental and occupancy conditions, the proposed methodology delivers a clearer and more accurate portrayal of building performance than is achievable through conventional static models. The quantitative findings—ranging from reductions in prediction error to measurable improvements in energy efficiency—highlight the practical value of integrating such systems into everyday facility management. Beyond numerical gains, the approach offers broader strategic benefits. Its adaptability positions it as a suitable tool for building portfolios across varying climatic regions, particularly in rapidly developing urban environments where demand patterns shift continuously. Its ability to expose hidden inefficiencies and support informed managerial decisions strengthens its relevance not only as a modelling instrument but as an operational tool capable of guiding building operators toward more effective conservation practices. Nevertheless, the promise of AI-enhanced energy modelling will only be fully realised when supported by stronger data infrastructures, clearer standardisation pathways, and governance frameworks that ensure transparency and interoperability. As these complementary systems mature, the methodology presented here has the potential to serve as a cornerstone in achieving large-scale improvements in building energy performance. In doing so, it contributes not only to facility-level optimisation but also to broader national and regional ambitions for carbon reduction and sustainable urban development.

References

1. Zhou, Y., Li, Y., & Ong, E. (2024). Advancements in greenhouse gas emission reduction methodology for fluorinated compounds and N₂O in the semiconductor industry via abatement systems.

- Frontiers in Energy Research, 11, 1234486. 10.3389/fenrg.2023.1234486
2. Li, J., Hao, Y., & Li, Y. (2026). AI-enabled energy baselines for verified building decarbonization. *Scientific Reports*, 16(1), 5815. 10.1038/s41598-026-36284-w
3. Cao, R., Hao, Y., Li, Y., & Liao, W. (2025). Emerging trends in lifecycle assessment of building construction for greenhouse gas control: Implications for capacity building. *Discover Applied Sciences*, 7(5), 398. 10.1007/s42452-025-06853-1