Construction and Practice of Energy Efficiency Evaluation System for Green Building Environment Design Driven by Artificial Intelligence

Weiwei Zhang, Xiaojuan Liu, Yi Sun, Mingjing Sun
School of Design and Art, Shandong Huayu University of Technology, Dezhou 253000, Shandong, China
Corresponding author: Weiwei Zhang (e-mail: wwiezhang212@163.com)

ABSTRACT Green-building energy-efficiency evaluation requires dynamic integration of spatial topology, environmental sensing, and design-optimization feedback. This study constructs an artificial-intelligence-driven evaluation system that combines BIM, IoT monitoring data, graph neural networks, deep reinforcement learning, and interpretable analysis. Building spaces are modeled as graph nodes, while heat-transfer paths, ventilation channels, equipment zoning, and personnel movement are represented as graph edges. A temporal graph attention network extracts environment–space coupling features, and a soft actor-critic reinforcement-learning model dynamically generates multi-objective evaluation weights under energy-efficiency, thermal-comfort, daylighting, and carbon-emission constraints. An interpretable module integrating attention weights and Shapley values forms an assessment–diagnosis–optimization loop. Experiments on eight operational green-building projects show that the system achieves an RMSE of 2.91 kWh/m^2·a, reduces prediction error by 66.3% compared with static indicators, and improves energy-efficiency optimization from 12.4% to 28.6%. The framework supports IoT-based environmental sensing, intelligent energy management, and design decision optimization.

INDEX TERMS graph neural network, deep reinforcement learning, iot sensor networks, energy efficiency evaluation, design optimization

I. INTRODUCTION

ENERGY efficiency assessment of green building environmental design is an important basis for building energy efficiency throughout the building life cycle, which will affect the building carbon emission and user experience. However, traditional energy efficiency assessment methods generally adopt static index systems and simple physical models, and the assessment process is disconnected from the building operating conditions. Therefore, when facing weather conditions, fluctuations in occupant energy consumption, and real-time differences in grid carbon emission intensity, there is no way to respond dynamically. In addition, the assessment results only output the single score, which cannot explain the reasons for the energy efficiency. There is no explanation for the reasons of energy efficiency difference, and there is no direct link between the assessment results and design parameters, which makes the designer unable to make accurate modifications [1-3]. This problem is particularly evident in the green building design, and has become a bottleneck that hinders the improvement of design quality.

Extensive research has been conducted by scholars both domestically and internationally on the issue of green building energy efficiency assessment. For example, Olu-Ajayi et al. introduced machine learning methods into energy efficiency prediction in the building design stage, improving the evaluation efficiency of early schemes [4]; Elkhoukhi et al. improved the accuracy of building simulation and the method of operational energy consumption analysis from the perspective of occupancy detection and behavior prediction [5]; Yu et al. systematically summarized the application of deep reinforcement learning in intelligent building energy management, pointing out that it has strong optimization capabilities in complex dynamic environments [6]; Huotari et al. summarized the application of machine learning in energy consumption prediction, control and intelligent services from the perspective of intelligent building energy utilization [7]; in addition, Villano et al. reviewed machine learning methods in building energy management, pointing out that current research is mostly focused on single task scenarios and lacks a systematic integrated framework [8]. Although existing research has made some progress in specific scenarios, it generally suffers from problems such as fixed evaluation indicators, poor adaptability to dynamic environments, and lack of interpretability of evaluation results, making it difficult to meet the actual needs of green building environmental design for collaborative optimization and accurate decision-making.

In response to the above problems, some researchers have tried to introduce artificial intelligence technology for improvement. For example, González and Bandera proposed a building energy consumption model calibration method based on inverse modeling, which improves the consistency between the model and actual operation through data-driven approach [9]; Yu et al. applied multi-agent deep reinforcement learning to the control of HVAC systems in commercial buildings, realizing collaborative optimization decision-making in complex dynamic environments [10]. However, existing methods mostly focus on single links such as energy consumption prediction or operation control, and have not yet built a closed-loop system that runs through the entire chain of "data-representation-evaluation-interpretation"; in terms of dynamic weight generation, they lack the ability to adapt to design constraints and changes in environmental scenarios; in terms of interpretability, the attribution analysis of evaluation results is still relatively weak, making it difficult to effectively guide design optimization.

Therefore, this paper develops an artificial intelligence based energy efficiency evaluation system for green building environmental design. Firstly, a multi-source heterogeneous database is built by integrating Building Information Modeling (BIM) and IoT real-time monitoring data, and graph neural networks are employed to jointly represent the topological relationships of building space and paths of environmental impact. Secondly, deep reinforcement learning algorithms are utilized to integrate energy efficiency target into the design constraints of thermal comfort and daylighting performance into a multi-objective optimization framework, and dynamically generate evaluation weights and thresholds of different design strategies. Finally, an interpretable evaluation module is constructed, which quantifies the contribution of key design features to energy efficiency indicators through attention mechanisms and Shapley value decomposition, and forms a closed loop of feedback "evaluation-diagnosis-optimization".

II. CONSTRUCTION OF AN ENVIRONMENT-SPATIAL COUPLING REPRESENTATION MODEL BASED ON GRAPH NEURAL NETWORKS

Energy efficiency in building environment design is not only affected by the physical characteristics of a single space, but also influenced by the coupling influence of multiple factors, such as spatial topology, heat transfer path of building envelope, equipment system zoning, and personnel flow. The above elements are usually considered separately by traditional methods, which makes it hard to reflect the interaction between spaces and the propagation mechanism of building environment. To address the above problems, this study constructs an environment-spatial coupling representation model based on graph neural networks, which transfers building space units and various environment elements into a same graph structure and achieves deep fusion and joint representation of multi-source heterogeneous data.

Firstly, the information such as geometric dimensions, building envelope material properties and equipment parameters are extracted from Building Information Model (BIM). Meanwhile, real-time time-series data such as temperature, humidity, illuminance and CO2 concentration collected by IoT sensor are integrated into a multi-source heterogeneous database including "morphology—materials—equipment—environment". On this basis, the rooms and functional areas in the building are defined as nodes in the graph structure. The node contains static design attribute (such as area, orientation and heat transfer coefficient of building envelope) and dynamic operational characteristic (such as temperature and humidity of each hour, and personnel density). Secondly, the heat transfer path of building envelope, the logic of air conditioning zoning, natural ventilation channel and personnel movement flow are defined as the edge of graph structure. The actual physical correlation strength and interaction frequency are used to assign the weight of the edge. This heterogeneous graph structure can reflect the heat transfer relationship, airflow relationship and personnel behavior relationship between internal building space. Based on this, this study uses a temporal graph attention network to perform end-to-end learning of the above graph structure. The graph attention layer can automatically assign the attention weight to different neighboring nodes and highlight the spatial coupling relationship with significant influence on the energy efficiency performance; the temporal convolution module is used to process the time-series feature in IoT sensor data and capture the dynamic change of environment parameter. The model takes node feature and edge structure in continuous time window as input, and outputs the building environment embedding vector in each time step. This embedding vector not only integrates the static design parameter and dynamic operation data, but also implicitly contains the joint information on spatial topological relationship and environmental propagation path, and then provides a unified feature input for the subsequent evaluation weight generation and interpretable analysis.

III. ADAPTIVE GENERATION OF EVALUATION STRATEGIES BASED ON DEEP REINFORCEMENT LEARNING

Traditional energy efficiency assessment uses a pre-defined fixed weight system, e.g. 35% energy efficiency, 25% thermal comfort, 15% daylighting and 25% carbon emissions. This fixed allocation method is static and cannot adapt to seasonal variations, differences in duration of energy consumption, and real-time variations in grid carbon emission intensity, which results in a gap between assessment results and building's actual operation performance. To bridge this gap, this paper explores a deep reinforcement learning algorithm and represents the energy efficiency assessment as a dynamic decision-making process. The objective function is formulated as:

$$J(\pi)=\sum_{t=0}^{T}\mathbb{E}_{(\mathbf{s}_t,\mathbf{a}_t)\sim\rho_\pi}\left[r(\mathbf{s}_t,\mathbf{a}_t)+\alpha\mathcal{H}\left(\pi(\cdot\mid\mathbf{s}_t)\right)\right] \quad (1)$$

where $\mathbf{s}_t$ is the state (building environment embedding vector), $\mathbf{a}_t$ is the action (weight allocation vector), $r(\mathbf{s}_t,\mathbf{a}_t)$ is the reward function, $\mathcal{H}(\pi(\cdot\mid\mathbf{s}_t))$ is the policy entropy, and α is the temperature coefficient used to balance the weights of reward and entropy.

To formalize the evaluation strategy, we frame the weight generation as a Markov Decision Process (MDP). This transition shifts the focus from static metrics to a dynamic sequence where the building environment embedding—extracted from the graph neural network—serves as the current state. This state effectively synthesizes spatial topology with real-time operational data, providing a holistic context for the agent. The action space is defined as the weight allocation vector for each evaluation sub-indicator, including energy efficiency weight, thermal comfort weight, daylighting weight and carbon emission weight. The sum of the four weights is constrained as 100%, and the output of action is normalized by Softmax function to make the weight effective.

The design of reward function is very important for the learning of deep reinforcement learning strategy. In this study, a multi-objective reward function is designed to incorporate energy efficiency objective and environmental performance constraints such as thermal comfort and daylighting into one whole framework.

Specifically, the value of reward includes three parts: firstly, an energy efficiency performance reward that rewards or penalizes the gap between evaluation result and measured energy consumption under the current weight allocation; secondly, a design constraint reward that gives certain negative reward when thermal comfort, daylighting and other indicators appear below the standard requirement; finally, a dynamic carbon emission reward that introduces a real-time grid carbon emission factor to promote more energy efficiency weight and carbon emission weight to be allocated in carbon emission standard period. The final reward signal is weighted sum of the three factors mentioned above to guide the policy network to learn an effective weight allocation strategy that can meet the design constraint and adapt to the change of environment. Furthermore, to address the stochastic nature of occupant behavior, the model incorporates a probabilistic ‘occupancy uncertainty’ factor into the state space. By training the Reinforcement Learning agent across diverse occupancy profiles—ranging from high-density office use to sporadic residential presence—the system develops a robust weight allocation strategy that maintains evaluation accuracy even when real-world human behavior deviates from static schedules. To resolve the physical trade-offs inherent in multi-objective optimization—such as the conflict between increased window-to-wall ratios for daylighting and the resulting decrease in thermal insulation—the DRL agent does not merely adjust weights but identifies “Pareto-optimal” states within the environment embedding. Specifically, the Soft Actor-Critic (SAC) algorithm utilizes its maximum entropy framework to explore the boundary of the design sensitivity map, penalizing actions that lead to a disproportionate collapse in any single performance metric. By mapping the non-linear coupling between spatial nodes (e.g., heat transfer vs. light transmission) via the graph neural network, the reward function effectively acts as a dynamic penalty-boundary system that forces the policy to prioritize design parameters with the highest synergistic potential rather than those that optimize one goal at the expense of another. For the algorithm selection, this study adopts an improved soft actor-critic algorithm for policy optimization. Compared with traditional reinforcement learning algorithms, the soft actor-critic algorithm uses a maximum entropy framework that maximizes cumulative reward while encouraging policy randomness to avoid premature convergence to suboptimal solutions and is suitable for exploration in continuous action space. The algorithm is composed of one policy network and two value networks. The three networks are trained by minimizing Bellman error and policy entropy loss respectively. The hyperparameter configuration listed in Table 1 is used to train the policy networks.

TABLE 1. Hyperparameters of Deep Reinforcement Learning

Parameter	Value	Description
Optimizer	Adam	Network update
Learning rate	3×10^{-4}	Fixed value
Discount factor γ	0.99	Reward decay
Temperature coefficient α	0.1	Entropy weight
Replay buffer capacity	10^5	Storage size
Batch size	256	Sampling number
Target network update frequency	1000 steps	Soft update interval
Maximum training steps	5×10^5	Termination condition

IV. AN INTERPRETABLE EVALUATION MODULE BASED ON THE FUSION OF ATTENTION MECHANISM AND SHAPLEY VALUE

A prominent problem in the practical application of energy efficiency evaluation is the lack of interpretability of the evaluation results. This study constructs an interpretable evaluation module based on the fusion of attention mechanism and Shapley value, which quantifies the contribution of each design feature to the energy efficiency index while outputting the evaluation score. The Shapley value originates from cooperative game theory and can fairly allocate the marginal contribution of each input feature to the final output. For feature i, the formula for calculating its Shapley value ϕ_i is (2):

$$\phi_i=\sum_{S\subseteq F\setminus\{i\}}\frac{|S|!(|F|-|S|-1)!}{|F|!}\left[f(S\cup\{i\})-f(S)\right] \quad (2)$$

Where F is the set of all features, S is the feature subset that does not contain feature i, and $f(S)$ is the output value of the evaluation model when only the feature subset S is used. The module's input comes from two parts: the built-environment embedding vector output by the graph neural network and the dynamic evaluation weights generated by the deep reinforcement learning policy network. Based on this, the interpretable module conducts attribution analysis on two levels.

First, it uses spatial coupling localization based on an attention mechanism. In the forward propagation process of the graph neural network, when the graph attention layer applies attention weights to different neighboring nodes during forward propagation, these weights represent the interaction strength between spatial units in the energy efficiency performance on energy efficiency. Therefore, this study extracts the weight distribution of the last graph attention layer and quantizes it on the original building floor plan to find out the key spatial nodes and environment edges that have a great influence on the energy efficiency. For example, when the attention weight of the room is much larger than other nodes, it indicates that the space is in a key position in the energy efficiency propagation path and the design parameter adjustment will have a great influence. Therefore, this mechanism can help designers find the spatial location of the weak points in energy efficiency as soon as possible.

Second, it uses feature contribution quantification based on Shapley values. Shapley values come from cooperative game theory and can fairly allocate the marginal contribution of each input feature to the final output. This study considers the output of the evaluation module as the payoff function of a cooperative game, and the input features (including design variables such as window-to-wall ratio, shading type, insulation layer thickness, air conditioning setting parameters, and environment variables such as temperature, humidity, and illuminance) as the participating players in the game. This study uses a structured Shapley value analysis method to calculate the marginal contribution of each feature to the evaluation score. In the specific implementation process, Monte Carlo sampling is used to approximate the calculation of the Shapley value of each feature to reduce the computational complexity. Experimental tracking indicates that for a standard building unit with 15–20 design variables, the interpretable module achieves a convergence of attribution results within 45 to 60 seconds using the NVIDIA RTX 4090 workstation. This sub-minute latency ensures that the feedback loop is sufficiently responsive for real-time 'assessment-diagnosis-optimization' during the architectural design phase, providing designers with nearly instantaneous sensitivity analysis without the hours of processing typically required for full-scale physical simulations. Finally, the output Shapley value reflects the degree of positive or negative impact of each feature on the energy efficiency performance; a larger value indicates that the feature makes a greater contribution to the improvement of the energy efficiency, while a negative value indicates that the feature has a negative impact on the energy efficiency.

Based on the attribution results at these two levels, the module further generates a two-dimensional diagnostic map of "feature importance - design sensitivity." The horizontal axis of the map represents design sensitivity, reflecting the response magnitude of design parameter changes to the evaluation score; the vertical axis represents feature importance, i.e., the magnitude of the Shapley value. Based on these two criteria, design elements are divided into four quadrants: elements with high importance and high sensitivity should be optimized first, elements with low importance and high sensitivity need to be adjusted carefully, elements with high importance and low sensitivity can be used as fixed configurations, and elements with low importance and low sensitivity do not need much attention. To bridge the gap between diagnosis and execution, the system maps these quadrant-based priorities to a Rhino+Grasshopper parametric script. This script automatically adjusts design variables (e.g., window-to-wall ratios or insulation thickness) within predefined architectural constraints, allowing the 'diagnosis' to trigger a new iteration of the BIM model for re-evaluation, thereby completing the automated optimization loop.

V. MODEL PERFORMANCE AND EVIDENCE CHAIN CONSTRUCTION EFFECTIVENESS EVALUATION

A. EXPERIMENTAL SETUP

1) (I) Experimental Subjects and Data Sources

Eight green building projects in the Yangtze River Delta region with complete data and already in operation were selected as experimental subjects, covering four functional types: office (3), commercial (2), residential (2), and educational (1), with building areas ranging from 5000m^2 to 30000m^2. All projects possessed:

Complete building information model data (geometric dimensions, envelope parameters, equipment information);

At least 12 months of hourly monitoring data from IoT sensors (temperature, humidity, illuminance, CO_2 concentration, hourly power consumption, gas consumption);

Complete green building design drawings and design change records.

2) (II) Experimental Environment and Procedure

All experiments were conducted in a unified computing environment:

Hardware Environment: Workstation equipped with an NVIDIA RTX 4090 GPU and 64GB of memory

Software Environment: Python 3.9, PyTorch 2.0, EnergyPlus 23.1, Rhino 7 + Grasshopper

Data Division: Six of the eight projects were used for model training and parameter tuning, and the other two served as an independent test set for final performance evaluation.

3) Energy Efficiency Assessment Accuracy Comparison Experiment

To verify the energy efficiency prediction accuracy of this assessment system, eight operational green building projects in the Yangtze River Delta region were selected as test subjects, covering office, commercial, and residential functional types. Using the hourly actual energy consumption data of each project as the baseline, this method was compared with the traditional static index assessment method (defined here as the manual weighted-sum calculation based on the 'Assessment Standard for Green Building' GB/T 50378-2019, using fixed design-stage coefficients) and the EnergyPlus physical simulation method. Root mean square error and mean absolute percentage error were used as assessment indicators.

TABLE 2. Energy Efficiency Assessment Accuracy

Evaluation Method	RMSE (kWh/m^2·a)	MAPE (%)
Traditional Static Indicator Assessment Method	8.64	14.2
EnergyPlus Physical Simulation Method	6.23	10.8
Proposed Assessment System	2.91	5.4

In Table 2, the root mean square error of this assessment system across the eight test items is 2.91 kWh/m^2·a, and the mean absolute percentage error is 5.4%. Compared to the traditional static index assessment method (8.64 kWh/m^2·a, 14.2%) and the EnergyPlus physical simulation method (6.23 kWh/m^2·a, 10.8%), the root mean square error is reduced by 66.3% and 53.3%, respectively, and the mean absolute percentage error is reduced by 62.0% and 50.0%, respectively, verifying the significant advantage of this method in energy efficiency assessment accuracy.

4) Implicit Association Detection Capability Assessment Experiment

To verify the adaptive capability of this assessment system under dynamic scenarios, a typical green office building was selected as the test object, and four typical operating scenarios were set up: summer, winter, transitional season, and extreme high-temperature weather. The fixed-weight method (expert-preset) and

this method (deep reinforcement learning adaptive) were compared to analyze the trend of weight allocation changes and the degree of matching between the two methods and the scenario requirements under the above scenarios.

In Figure 1, the fixed-weight method allocates weights of 35% for energy efficiency, 25% for thermal comfort, 15% for daylighting, and 25% for carbon emissions in all four scenarios, failing to reflect scenario differences. This method, however, dynamically adjusts according to scenario characteristics: the weight of energy effi ciency increases to 42% and carbon emissions to 30% in summer; the weight of thermal comfort increases to 32% in winter; the weight of daylighting increases to 28% in transitional season; and the weight of carbon emissions increases to 35% under extreme weather conditions. This adaptive adjustment highly matches the actual operating requirements of each scenario, verifying the significant advantage of this method in dynamic adaptability.

5) Explainability and Design Optimization Auxiliary Verification Experiment

To verify the auxiliary value of the explainability module for design optimization, two green building projects under construction were selected, and six designers were recruited and randomly divided into a control group and an experimental group. The control group only obtained energy efficiency scores, while the experimental group obtained an additional two-dimensional diagnostic map of feature importance-design sensitivity. The energy efficiency improvement rate and iteration number of the optimization schemes for both groups were recorded.

Figures 2 (a-c) show the comparison of energy efficiency improvement rates, optimization iteration numbers, and designer satisfaction scores, respectively. The experimental results show that the average energy efficiency improvement rates for the three groups in the three projects were 10.1% for the control group, 22.1% for experimental group A, and 31.2% for experimental group B; the average iteration numbers were 5.3, 2.7, and 1.4, respectively; and the average satisfaction scores were 2.8, 4.0, and 4.7, respectively. Compared to experimental group A, experimental group B showed a 9.1 percentage point increase in energy efficiency and a 1.3-fold reduction in iterations, indicating that the optimization efficiency was further improved after introducing design sensitivity analysis. The attribution stability coefficient of the interpretable module ranged from 0.80 to 0.91, demonstrating high reliability and verifying its significant advantages in accurately locating key design variables and assisting in optimization decisions.

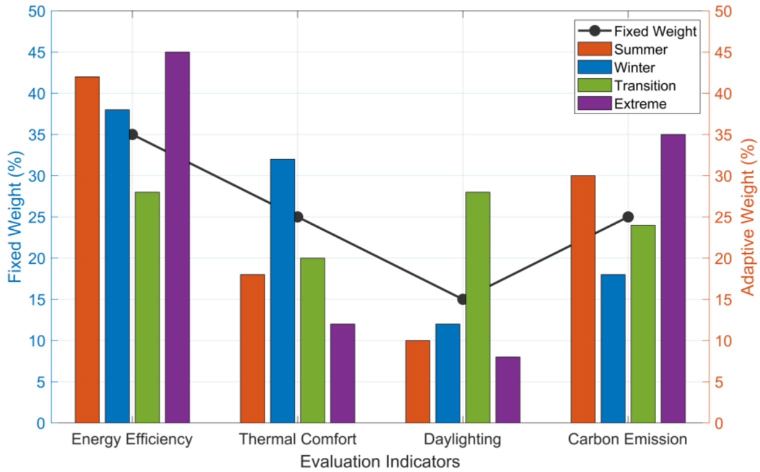

FIGURE 1. Implicit Association Detection Capability Assessment

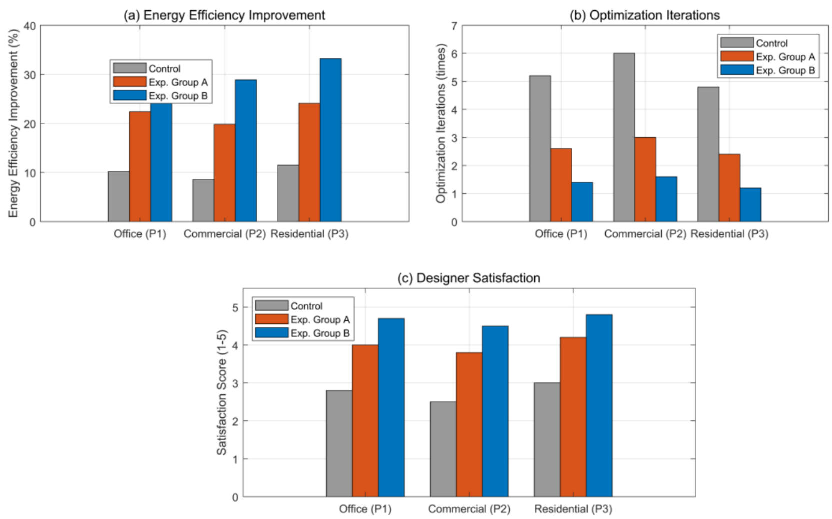

FIGURE 2. Explainability and Design Optimization Auxiliary Verification Evaluation

VI. CONCLUSION

This paper develops an AI-based energy efficiency assessment system for green building environmental design, combining graph neural networks and deep reinforcement learning, which can realize dynamic representation of multi-source data and scene-adaptive assessment weight generation simultaneously. Furthermore, the interpretable module can quantify the contribution of key design features and form a “assessment-diagnosis-optimization” closed loop. The experimental results have further verified the advantages of proposed system in accuracy, dynamic adaptability and design decision support. Due to the limitation of sample sources, the samples now come from a relatively narrow range. In the future, we will broaden the range of data collection, and further verify the universality of the model’s application in different climate and building types, and explore its deep coupling with forward design process. Specifically, while the current ‘adaptive adjustment’ is validated within the hot-summer and cold-winter climate of the Yangtze River Delta, future iterations will incorporate training data from arid and sub-arctic regions. This will test whether the Deep Reinforcement Learning agent can autonomously shift its reward priorities—such as prioritizing thermal envelope integrity in extreme cold versus solar gain control in arid regions—thereby ensuring the ‘adaptability’ is globally applicable rather than climate-specific.

AUTHOR CONTRIBUTIONS

Conceptualization –Weiwei Zhang, Xiaojuan Liu, Yi Sun and Mingjing Sun; methodology – Weiwei Zhang, Yi Sun and Mingjing Sun; investigation – Weiwei Zhang, Xiaojuan Liu, Yi Sun and Mingjing Sun; writing-original draft preparation – Weiwei Zhang, Xiaojuan Liu, Yi Sun and Mingjing Sun. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] Q. Shi, C. Liu, and C. Xiao, "Machine learning in building energy management: A critical review and future directions," Frontiers of Engineering Management, vol. 9, no. 2, pp. 239-256, 2022, doi: 10.1007/s42524-021-0181-1.

[2] S. Mischos, E. Dalagdi, and D. Vrakas, "Intelligent energy management systems: a review," Artificial Intelligence Review, vol. 56, no. 10, pp. 11635-11674, 2023, doi: 10.1007/s10462-023-10441-3.

[3] Y. Ding, S. Han, Z. Tian, et al., "Review on occupancy detection and prediction in building simulation," Building Simulation, vol. 15, no. 3, pp. 333-356, 2022, doi: 10.1007/s12273-021-0813-8.

[4] R. Olu-Ajayi, H. Alaka, I. Sulaimon, et al., "Machine learning for energy performance prediction at the design stage of buildings," Energy for Sustainable Development, vol. 66, pp. 12-25, 2022, doi: 10.1016/j.esd.2021.11.002.

[5] H. Elkhoukhi, M. Bakhouya, A. Elmouatamid, et al., "A review of occupancy sensing technologies and approaches in smart buildings," International Journal of RF Technologies, vol. 15, no. 1, pp. 19-48, 2025, doi: 10.3233/RFT-240006.

[6] L. Yu, S. Qin, M. Zhang, et al., "A review of deep reinforcement learning for smart building energy management," IEEE Internet of Things Journal, vol. 8, no. 15, pp. 12046-12063, 2021, doi: 10.1109/JIOT.2021.3078462.

[7] M. Huotari, A. Malhi, and K. Främling, "Machine Learning Applications for Smart Building Energy Utilization: A Survey," Archives of Computational Methods in Engineering, vol. 31, no. 5, pp. 2537-2556, 2024, doi: 10.1007/s11831-023-10054-7.

[8] F. Villano, M. Mauro G, and A. Pedace, "A review on machine/deep learning techniques applied to building energy simulation, optimization and management," Thermo, vol. 4, no. 1, pp. 100-139, 2024, doi: 10.3390/thermo4010008.

[9] G. González V and F. Bandera C, "A building energy models calibration methodology based on inverse modelling approach," Building Simulation, vol. 15, no. 11, pp. 1883-1898, 2022, doi: 10.1007/s12273-022-0900-5.

[10] L. Yu, Y. Sun, Z. Xu, et al., "Multi-agent deep reinforcement learning for HVAC control in commercial buildings," IEEE Transactions on Smart Grid, vol. 12, no. 1, pp. 407-419, 2020.