

Energy Optimization Approach for Smart Buildings Considering Supply and Demand Uncertainties

Xiaojuan Liu¹ and Zenghui Duan²

¹ Lecturer, Applied Technology College of Dalian Ocean University, Dalian, 116300, China.

² Lecturer, Applied Technology College of Dalian Ocean University, Dalian, 116300, China, E-mail: duanzhui@outlook.com (corresponding author).

Engineering Management

Received August 13, 2025; revised September 28, 2025; accepted September 29, 2025

Available online March 7, 2026

Abstract: Energy optimization in intelligent buildings is a multi-objective decision-making problem that requires the consideration of uncertainties on both supply and demand sides. However, existing optimization strategies often consider a single indicator in isolation and lack collaborative system optimization, resulting in poor overall performance. To address this issue, this study proposes an intelligent building energy optimization model based on deep reinforcement learning and multi-criteria decision-making. This model integrates Long Short-Term Memory Transformer network for multi-source temporal feature extraction and modeling, uses differential evolution algorithm to optimize Proximal Policy Optimization reinforcement learning strategy, and introduces analytic hierarchy process to allocate weights to competitive indicators such as energy efficiency, economy, comfort, to construct a comprehensive incentive function that guides the intelligent agent to generate optimization strategies that benefit both supply and demand sides in synergy. The results indicate that the proposed model achieves a 40% convergence rate and a 95.5% policy performance achievement rate. The proposed strategy achieves a maximum energy optimization efficiency of 15.8% and a maximum economic benefit improvement of 16.2%. In simulation tests, the proposed strategy achieved a maximum satisfaction rate of 92.1% on the supply side and 95.6% on the demand side. These results demonstrate that the proposed method can effectively address uncertainties from both the supply and demand sides to formulate optimal strategies that consider energy consumption, economic performance, and other key indicators. This method provides a new perspective for smart building energy optimization and promotes the application of deep learning algorithms in diverse decision-making tasks.

Keywords: Smart building; Energy optimization; Supply and demand; Proximal policy optimization; Analytic hierarchy process

Copyright © Journal of Engineering, Project, and Production Management (EPPM-Journal).

DOI 10.32738/JEPPM-2025-161

1. Introduction

To achieve this goal, many researchers have explored optimization strategies for smart buildings (Apanavičienė and Shahrabani, 2023). Energy optimization remains a major focus in this field, but existing research often lacks comprehensive strategies and relies on limited evaluation metrics. In deep learning, various algorithms have been developed, each with unique strengths applicable in different domains depending on specific needs and device characteristics (Volkov et al., 2022). Among them, the Long Short-Term Memory (LSTM) network effectively processes long-term sequential data, while the Transformer architecture performs parallel computation across multiple information channels and models global dependencies in multi-dimensional data (Al-Selwi et al., 2023; Dalla-Torre et al., 2025). The Proximal Policy Optimization (PPO) algorithm is a policy gradient method known for its fast convergence and high accuracy (Van Hezewijk et al., 2023). The characteristics of LSTM and Transformer align well with the long-term and multi-dimensional nature of smart building data. Therefore, this study proposes an energy optimization model for smart buildings that combines LSTM, Transformer, and PPO. The model innovatively integrates different algorithms according to the data characteristics of its various optimization processes. Simultaneously, it introduces a strategy design approach that considers uncertainties on both the supply and demand sides. The goal is to effectively optimize energy efficiency and economic performance, thereby supporting the development of smart buildings into more energy-efficient and comfortable environments.

2. Related Works

The LSTM network and Transformer architecture are deep learning algorithms with strong capabilities in extracting and modeling time series features. Ma et al. (2023) proposed a new model based on an LSTM network and a Transformer to address the low accuracy of traditional industrial forecasting models for multivariate time series. The model captured important data by stacking multiple encoding layers and achieved satisfactory results in practical tests on four datasets. Mazraedoost et al. (2025) developed a bidirectional sequence processing and fusion model based on the Transformer and LSTM network to improve the accuracy of retention time prediction in liquid chromatography. The model achieved a mean absolute error of 26.23 seconds and a mean absolute percentage error of no more than 3.25%. The PPO algorithm is also widely applied in various fields by researchers around the world due to its strong adaptability and strategy development capability. For instance, Bhourji et al. (2024) developed an intelligent control system based on PPO algorithm that incorporated deterministic policy gradients to achieve precise control of a rotary inverted pendulum system. A comparison with a traditional controller showed that the proposed method outperformed the traditional one. To address the problem of rapidly increasing energy consumption during Internet of Things data transmission, Lee et al. (2023) proposed an edge computing-based autonomous transmission control system using PPO algorithm. In tests on three different datasets, the system effectively reduced energy consumption by decreasing total data transmission by 73.849%, 89.931%, and 81.310%, respectively. Park et al. (2023) introduced a multi-agent algorithm based on the PPO algorithm to improve communication performance in terahertz data transmission. The algorithm treated data transmission as a mixed-integer nonlinear programming problem and computed optimal solutions. Simulation results showed that it improved network utilization by 17.77%.

As smart buildings have continued to evolve, energy optimization remains a key research focus. For example, Muhammad and Hossain (2022) proposed an urban smart building energy allocation system to address population overload and energy shortages in smart cities. The system effectively scheduled devices and reduced energy distribution costs and delays in smart building grids. However, it did not consider other types of energy consumption. Maizana and Putri (2022) developed a smart grid allocation system based on photovoltaic solar power for campus smart buildings. Although simulation results showed that the system reduced energy usage, it did not account for indicators on the demand side. Chakraborty et al. (2022) proposed a power grid control algorithm based on a variable gain-enhanced second-order generalized integrator phase-locked loop to improve energy efficiency in smart buildings under extreme conditions. The algorithm showed promising performance in terms of economy and energy-saving in simulations, but did not address subjective demand-side experiences. Zhao and Wu (2022) proposed a decentralized scheduling framework that used blockchain technology to optimize traditional centralized aggregators in smart buildings. Simulation results showed that the framework reduced energy scheduling costs. However, the research only focused on energy efficiency and did not consider demand-side indicators. Zhang et al. (2024) developed a renewable, wood-based power generation technology to improve smart building energy efficiency. The Seebeck coefficients of heated wood increased to 320.5 and 436.6 $\mu\text{V/K}$ in the vertical and parallel directions, respectively. Although this method improved the utilization of wood energy, it did not account for other energy types in smart buildings.

In summary, although existing research has made progress in optimizing energy consumption in smart buildings, most studies focus on a single perspective, either supply or demand and fail to consider energy strategies from a supply-demand synergy perspective, resulting in poor overall performance. For example, Muhammad and Hossain (2022) did not consider multiple energy types, and Chakraborty et al. (2022) did not address subjective experiences. To address this gap, this study aims to resolve the key challenges of supply and demand synergy optimization. The LSTM network, Transformer architecture, and PPO algorithm have distinct advantages in time series feature extraction, modeling, and decision-making. This provides a technical foundation for simultaneously handling uncertainties on both supply and demand sides while formulating comprehensive strategies. Therefore, this study proposes an energy optimization model that integrates these algorithms to account for uncertainties on both the supply and demand sides. The aim is to develop a global optimal strategy through collaborative modeling of supply and demand uncertainties to improve energy allocation and provide residents with a better experience.

3. Smart Building Energy Optimization Model Integrating Deep Learning Algorithms

3.1. Algorithm-based Modeling and Strategy Design for Smart Buildings

Formulating intelligent building energy optimization strategies requires the consideration of numerous uncertain factors (Abbasimehr and Paki, 2022). The LSTM network captures long-term data variations through its gating mechanism, while transformer architecture models global dependencies in time series data and performs virtual spatial fitting of multidimensional data (He et al., 2023). Therefore, this study proposes a multidimensional information modeling algorithm for intelligent buildings that integrates the LSTM and Transformer, referred to as the LSTM-Transformer model. The structure of this model is shown in Fig. 1.

As shown in Fig. 1, the LSTM first extracts temporal features. It is a two-layer network composed of multiple subunits. Each subunit consists of an input gate, a forget gate, and an output gate, and its operational process is defined by Eq. (1).

$$\begin{cases} f_t = \sigma(w_f \bullet (h_{t-1}, x_t) + b_f) \\ i_t = \sigma(w_i \bullet (h_{t-1}, x_t) + b_i) \\ o_t = \sigma(w_o \bullet (h_{t-1}, x_t) + b_o) \end{cases} \quad (1)$$

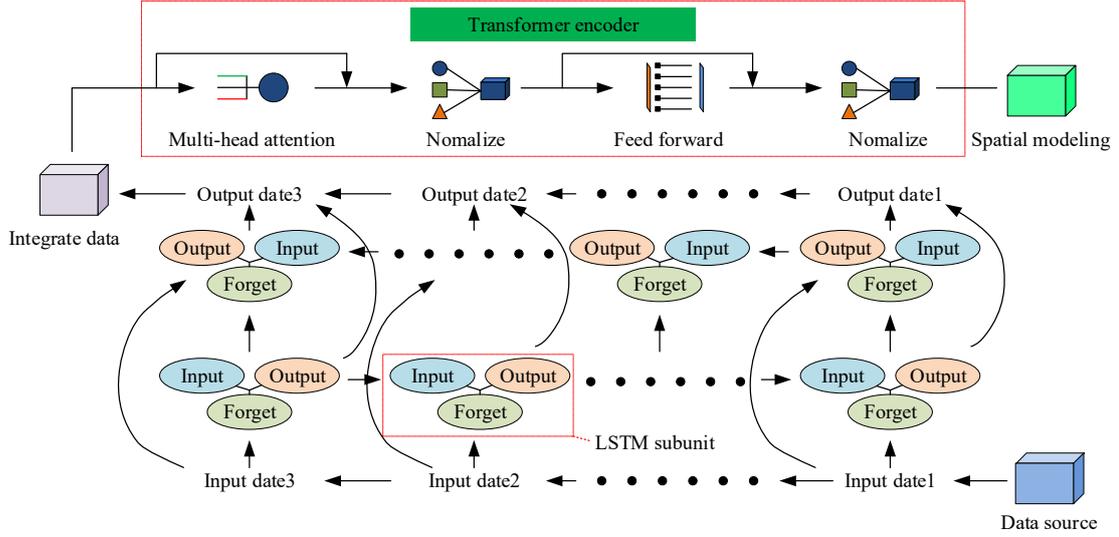


Fig. 1. LSTM-Transformer structure diagram

In Eq. (1), h_{t-1} represents the LSTM subunit state at $t-1$, and x_t is the input at t . σ is the activation function. w_f , w_i , and w_o are the weights of the three gates, while b_f , b_i , and b_o are the bias terms. f_t , i_t , and o_t are the outputs of each gate. Then, the Transformer performs spatial modeling. The Transformer consists of a multi-head attention mechanism, a normalization mechanism, and a feed-forward structure. The weighted calculation of the multi-head attention mechanism is shown in Eq. (2).

$$A(Q, K, V) = \text{soft max}\left(\frac{QK^T}{\sqrt{d_K}}\right)V \quad (2)$$

In Eq. (2), Q , K , and V represent the query, key, and value vectors, respectively. T is the transpose of a matrix, and d_K is the dimension of the matrix. The output of the weighted data is calculated as shown in Eq. (3).

$$H(Q, K, V) = \text{Contact}(h_1, h_2, \dots, h_n)W^* \quad (3)$$

In Eq. (3), h is the data obtained through different weighting methods, and W^* is the spatial projection matrix. The feed-forward structure performs the calculation process as shown in Eq. (4).

$$FFN = \max(0, xw^1 + b^1)w^2 + b^2 \quad (4)$$

In Eq. (4), w^1 and w^2 are the weights of the two normalization layers, while b^1 and b^2 are the corresponding terms. Following the modeling of various types of intelligent building data, optimization strategies can be formulated. PPO is a strategy planning algorithm with strong adaptability that works well in various complex scenarios. However, PPO's update constraints can often lead to convergence on local optima (Hassan et al., 2023). The Differential Evolution (DE) algorithm can be used to tune PPO's hyperparameters, ensuring the PPO remains in its optimal state, improving the accuracy of the resulting strategies. Therefore, this study employs a DE-optimized PPO (DP), to formulate intelligent building energy optimization strategies. The process and structure of DP algorithms are shown in Fig. 2.

As shown in Fig. 2, DE optimizes PPO parameters prior to the PPO execution phase. First, the PPO parameters are parsed and encoded as multi-dimensional data. Then, two random parameters from the data points are used to compute vector differences and are scaled and fused with the initial parameters. This process is shown in Eq. (5).

$$V_{g+1}^n = x_g^{s_1} + L * (x_g^{s_2} - x_g^{s_3}), n \neq s_1 \neq s_2 \neq s_3 \quad (5)$$

In Eq. (5), \mathcal{G} is the number of mutation iterations, n is the number of parameters, L is the search step size, x is a random parameter, and s_1 , s_2 , and s_3 represent different mutation levels. The crossover operation on the new parameters are shown in Eq. (6).

$$U_{g+1,j}^n = \begin{cases} V_{g+1,j}^n & \text{if } \text{rand}(0,1) \leq CR \text{ or } j = j_{rand} \\ x_{g,j}^n & \text{otherwise} \end{cases} \quad (6)$$

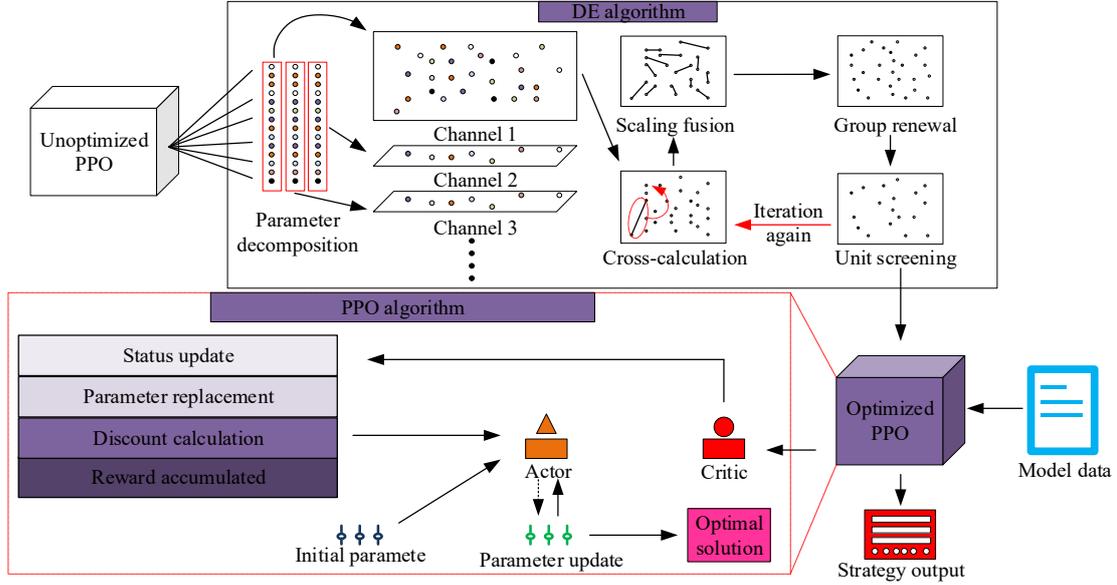


Fig. 2. DP algorithm flow diagram

In Eq. (6), U is the intermediate unit of the crossover calculation, j is the parameter dimension, j_{rand} is a random number between 0 and J , and CR is the crossover probability. Then, the parameters that meet the requirements are selected from the new and old ones, as shown in Eq. (7).

$$x_{g+1}^n = \begin{cases} U_{g+1}^n & \text{if } f(U_{g+1}^n) \leq f(x_g^n) \\ x_g^n & \text{otherwise} \end{cases} \quad (7)$$

In Eq. (7), x_{g+1}^n is the retained parameter unit. Based on x_{g+1}^n , PPO parameters are set to reach the optimal state and proceed with strategy formulation. Upon receiving data, PPO first sets a group of virtual environment parameters for the policy network. Then, the value network updates the state and discount factors in the virtual environment. The gradient for each update step is calculated using Eq. (8).

$$\Delta S = \text{Exp}\left[\frac{p_{\theta}(A_t, E_t)}{p_{\theta'}(A_t, E_t)} g^{\theta'}(E_t, A_t) \nabla \log p_{\theta}(A_t | S_t^n)\right] \quad (8)$$

In Eq. (8), P_{θ} and $P_{\theta'}$ represent the probability of a specific policy before and after the update, $g^{\theta'}$ is the average advantage function, and n is the number of trajectories. Finally, the intelligent building energy model and optimization algorithm integrating LSTM-Transformer and DP (LTDP) is constructed. Its structure and workflow are shown in Fig. 3.

As shown in Fig. 3, the LTDP framework consists of two main components: LSTM-Transformer and DP. These parts are responsible for data modeling and strategy formulation, respectively. The subunits LSTM and Transformer within LSTM-Transformer are in a serial configuration. During modeling, LSTM extracts data features, and Transformer performs global dependency modeling. The DE component interfaces only with PPO and has no direct connection to the front-end LSTM-Transformer. DE tunes PPO's parameters before decision-making to ensure optimal performance. PPO then evaluates energy utilization strategies based on various intelligent building information from both the supply and demand sides, modeled by LSTM-Transformer, and develops new strategies that outperform the incumbent ones.

3.2. Energy Optimization Model Considering Supply and Demand Uncertainties

LTDP performs feature extraction, spatial modeling, and energy optimization strategy formulation using data in intelligent buildings. However, LTDP cannot directly extract various types of physical information, nor can it obtain the prices or economic returns of electricity and natural gas. Traditional intelligent building energy optimization models usually rely on a combination of physical sensors, computer systems, and human-computer interfaces to complete the energy optimization

process (Sivakumar et al., 2024). The traditional process is shown in Fig. 4.

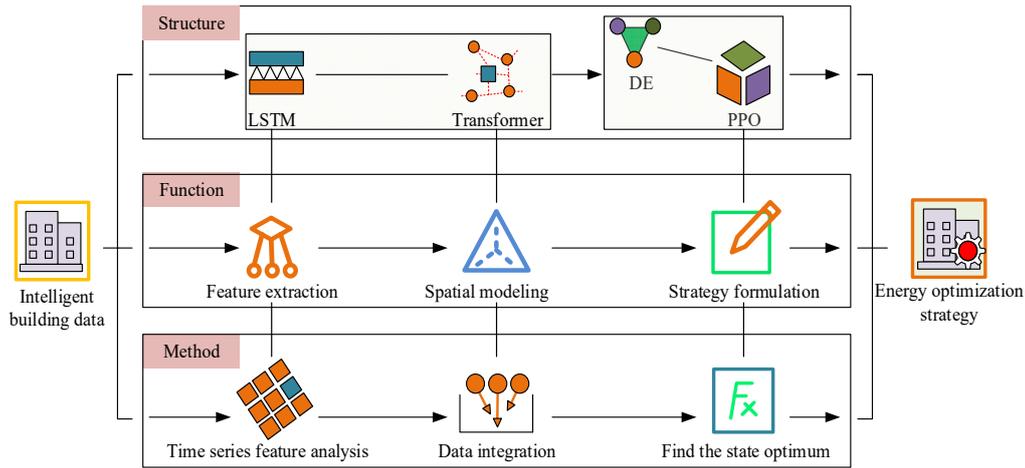


Fig. 3. LTDP intelligent building energy optimization algorithm flow diagram

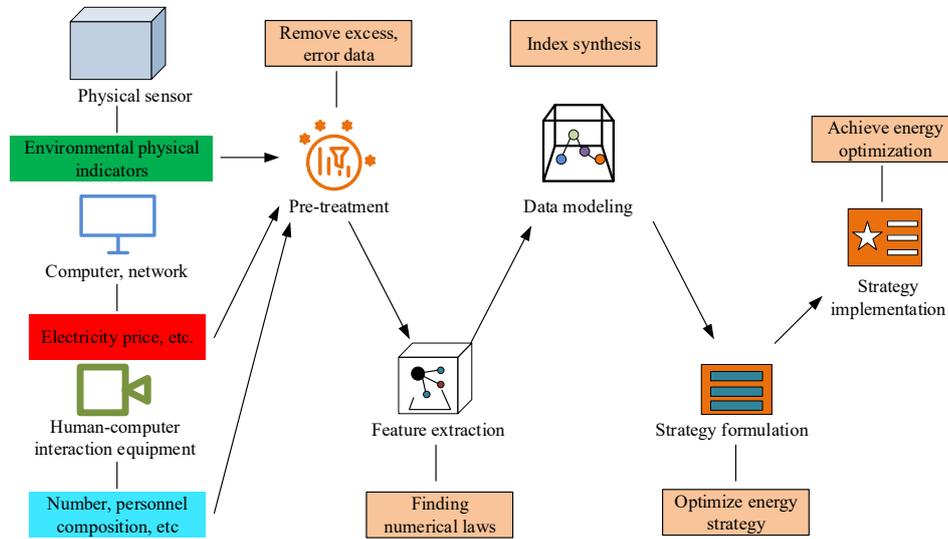


Fig. 4. Process of energy optimization strategy formulation for intelligent buildings

As illustrated in Fig. 4, the energy optimization process for intelligent buildings requires the collection of three categories of data, most of which are dynamic and uncertain. Following collection, the data undergoes preprocessing to remove abnormal and invalid values, thereby reducing the computational burden during subsequent feature analysis and modeling. Then, an intelligent algorithm is used to obtain the optimal energy strategy. Energy control in intelligent buildings involves multiple indicators from different domains, including physical and economic factors. In practice, directly evaluating indicators with different dimensions is challenging. The Analytic Hierarchy Process (AHP) is used to allocate weights to these multi-dimensional, competitive indicators from both the supply and demand sides, constructing a unified comprehensive evaluation system (Candra et al., 2025). This study applies AHP to derive the relative weights of each indicator. These weights are then used to construct the reward function of PPO algorithm, guiding it to formulate energy strategies that balance both supply and demand objectives. Therefore, this study proposes an AHP-based weighting scheme for supply and demand side indicators. The process is shown in Fig. 5.

In Fig. 5, AHP first classifies the data into supply side indicators, demand side indicators, and joint indicators. Then, a matrix is constructed for each category as shown in Eq. (9).

$$C = \begin{bmatrix} A_{11} & A_{12} & \dots & A_{1n} \\ A_{21} & A_{22} & \dots & A_{2n} \\ \dots & \dots & \dots & \dots \\ A_{n1} & A_{n2} & \dots & A_{nn} \end{bmatrix} \quad (9)$$

In Eq. (9), A represents the importance of specific indicators which can be assigned based on actual needs. n is the

maximum number of matrix rows and columns.

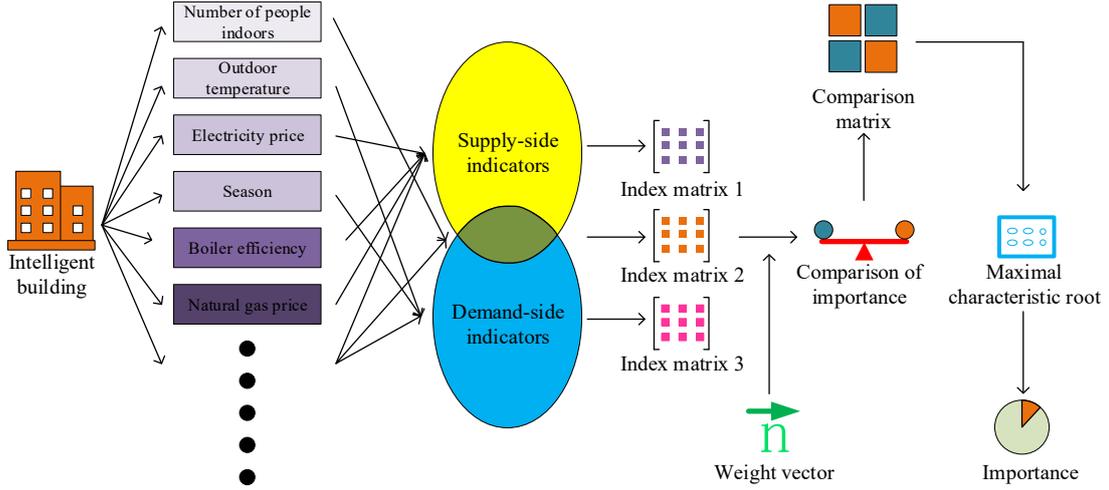


Fig. 5. AHP-based supply and demand side indicator quantification method for intelligent buildings

The importance of indicators is then compared. This starts by calculating the weight vector of the matrix, as shown in Eq. (10).

$$W = [W_1, W_2, \dots, W_n]^T \quad (10)$$

In Eq. (10), W_1 , W_2 , and W_n represent the eigenvectors of the corresponding indicators. The comparison matrix is then obtained as shown in Eq. (11).

$$A = [a_{ij}]_{n \times n} \quad (11)$$

In Eq. (11), a represents the elements of the comparison matrix, while i and j are the indexes of the two elements being compared. The relationship between the comparison matrix and the weight vector is shown in Eq. (12).

$$AW = nW \quad (12)$$

Based on Eq. (12), the relationship between the comparison matrix and its maximum eigenvalue is given in Eq. (13).

$$AW = \lambda_{\max} W \quad (13)$$

In Eq. (13), λ_{\max} is the maximum eigenvalue of the comparison matrix, which represents the importance of the indicator. The consistency of the weighting process is evaluated using Eq. (14).

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad (14)$$

In Eq. (14), a consistency coefficient CI closer to 0 indicates a better result of the comprehensive weighting. In multi-level analysis, the combination weight of each level is calculated using Eq. (15).

$$B_j = \sum a_i b_j \quad (15)$$

In Eq. (15), b is the importance weight of the corresponding level. Finally, the intelligent building energy optimization model based on AHP and LTDP (A-LTDP) is constructed. The structure is shown in Fig. 6.

As shown in Fig. 6, the proposed A-LTDP model comprises four main modules: data collection, data integration, modeling and decision-making, and a control output. During the energy strategy optimization process, the data collection module first gathers uncertain information from both the supply and demand sides. Then, AHP quantifies all collected data to enable integrated evaluation of different indicators. The modeling and decision-making module uses LTDP as the core algorithm. Within this module, PPO connects to AHP and uses the AHP-calculated weights as the basis for strategy formulation. After the PPO algorithm generates the optimal energy strategy, the control system adjusts the parameters of energy-consuming devices to implement energy optimization for the intelligent building.

4. Performance Verification of Intelligent Building Energy Optimization Model

4.1. Performance Verification of LTDP Algorithm for Energy Optimization Strategy Design

This study first used the Building Data Genome Directory (BDGD) to verify the performance of the LTDP algorithm. BDGD was an open data-sharing platform that included multidimensional data on energy and weather in smart buildings. Trust Region Policy Optimization (TRPO), Deep Deterministic Policy Gradient (DDPG), and Soft Actor-Critic (SAC) were selected as comparison algorithms. All algorithms were run under the same conditions. The test environment included Ubuntu 22.04.3 LTS as the operating system, Java 17 as the programming language, and TensorFlow 2.12 as the deep learning framework. The central processor was an Intel Core i9-13900K. The parameter settings during the experiments are shown in Table 1.

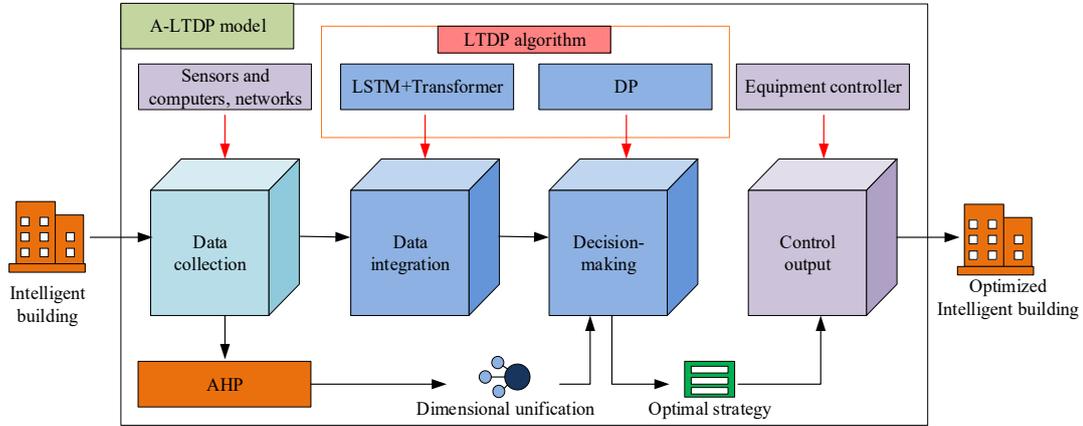


Fig. 6. Structure of the A-LTDP intelligent building energy optimization model

Table 1. Algorithm parameter settings in the experiment

Parameter	Parameter description	Parameter value
Lr	Learning rate	1e-3
Dropout	Discard rate	0.5
Batch_size	Lot size	32
Hidden_dims_f	The finance module hides the layer dimension	256
Hidden_dims_t	Text modules hide layer dimensions	512
Kernel_size	Convolution kernel size	(2,3,4)
Concat_size	Fusion dimension	256
Max_length	Maximum length of model input	800*3
Embed_dim	Text embedding dimension	960

The effectiveness of these strategies is one of the most important evaluation indicators for optimization algorithms. The Receiver Operating Characteristic (ROC) curve was also commonly used to validate algorithm performance. The achievement rate of strategy performance and ROC curves were compared, as shown in Fig. 7.

As shown in Fig. 7(a), the LTDP algorithm achieved a strategy performance achievement rate of 95.5% upon convergence, while the SAC algorithm reached 78.2%, the DDPG algorithm reached 83.4%, and the TRPO algorithm reached 80.6%. Furthermore, the LTDP algorithm converged after only 40% of the total iterations. As shown in Fig. 7(b), the area under the ROC curve of the LTDP algorithm was 0.965, while that of SAC was 0.872, DDPG was 0.910, and TRPO was 0.941. These results demonstrate that the LTDP algorithm achieved superior overall performance compared to the benchmark algorithms. Subsequently, to evaluate whether the LTDP algorithm could meet both supply and demand indicators in energy optimization, 100 data groups from the dataset were used for testing. The adaptability score is the comprehensive performance score for the strategies generated by each algorithm across supply and demand-side indicators. The degree to which each algorithm satisfied both sides was analyzed, and the results are shown in Fig. 8.

In Fig. 8(a), the energy optimization strategies generated by the LTDP algorithm demonstrated high adaptability to both the supply and demand sides, with all adaptability scores above 0.7. Figures 8(b) and 8(c) showed that the TRPO algorithm prioritized the supply side, whereas the DDPG algorithm demonstrated greater adaptability on the demand side. In contrast, Fig. 8(d) shows that the SAC algorithm exhibited relatively low adaptability to both sides. In summary, the LTDP algorithm comprehensively balanced both supply and demand considerations, resulting in superior energy optimization strategies. To further validate the LTDP algorithm's ability to integrate multiple objectives, the strategies produced by each algorithm over a five-hour period were evaluated across four metrics: energy, economy, comfort, and stability. The results of this

multi-objective evaluation are presented in Fig. 9.

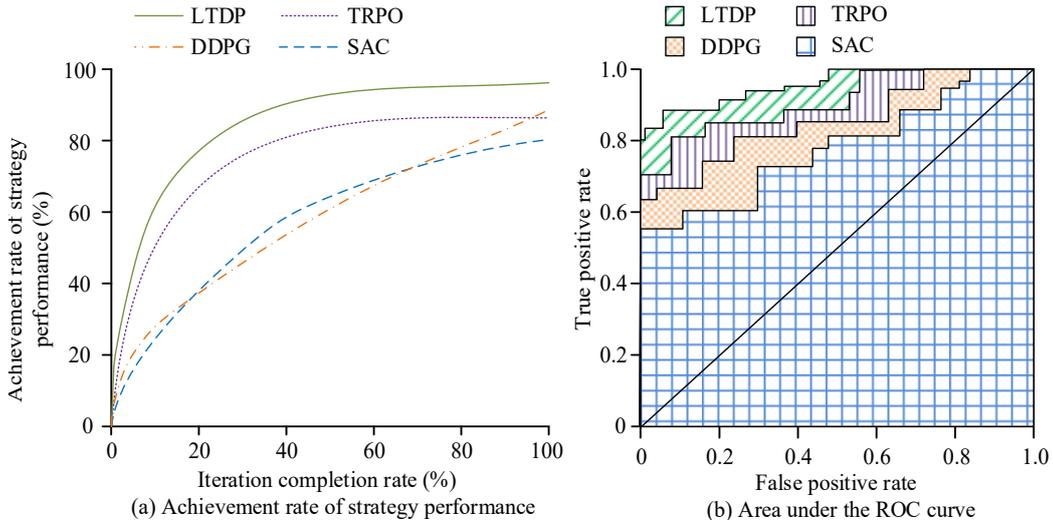


Fig. 7. Comparison of achievement rate of strategy performance and ROC curves of each algorithm

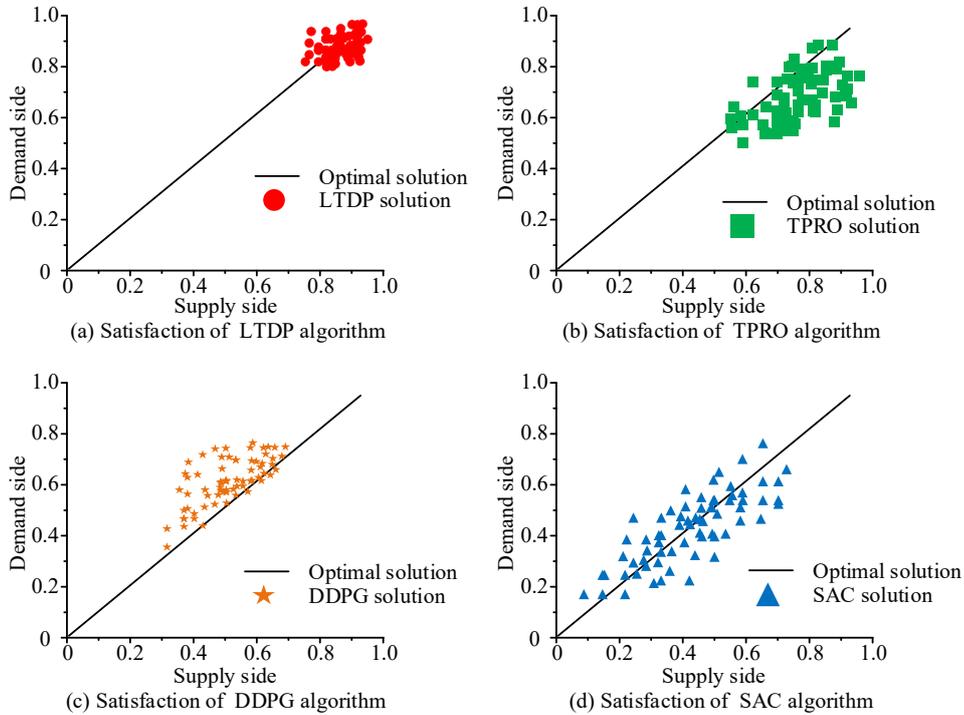


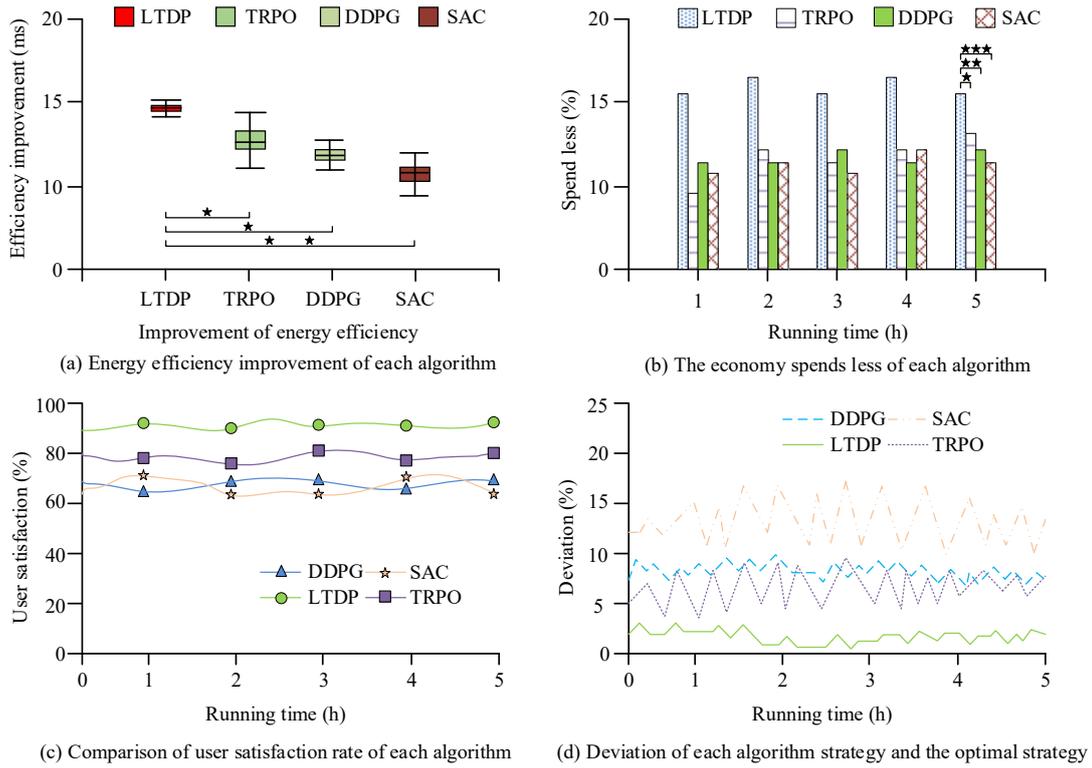
Fig. 8. Comparison of each algorithm's adaptability to supply side and demand side indicators

As shown in Fig. 9(a), the LTDP algorithm's strategies significantly improved energy efficiency, achieving a maximum improvement of 15.8% and outperforming all benchmark algorithms. Fig. 9(b) shows that the LTDP algorithm significantly reduced economic consumption, with a maximum reduction of 16.2% over the five-hour period. As shown in Fig. 9(c), the satisfaction level of strategies generated by LTDP reached a maximum satisfaction level of 93.1%, significantly higher than benchmark algorithms. Finally, Fig. 9(d) demonstrates that the strategies generated by LTDP had smaller deviations from the optimal strategy, with a maximum deviation within five hours of no more than 4.0%, which was substantially lower than other algorithms.

4.2. Simulation Verification of A-LTDP Model based on LTDP Algorithm and AHP

Building on the LTDP algorithm verification, the integrated A-LTDP model was evaluated using OneSpace digital twin platform. During verification, Integrated Deep Reinforcement Learning and Model Predictive Control (DRL-MPC), Multi-Agent Deep Deterministic Policy Gradient (MADDPG), and Imitation Learning-based Energy Management Systems (IL-EMS) were selected as comparison models. To assess the interpretability and reliability of the A-LTDP model, the accuracy

and 95% confidence intervals of each model during operation were compared, as shown in Fig. 10.



Note: The number of ★ in the figures indicates the significance of the differences. More ★ indicate larger differences.

Fig. 9. Comparison of the actual effects of strategies generated by each algorithm

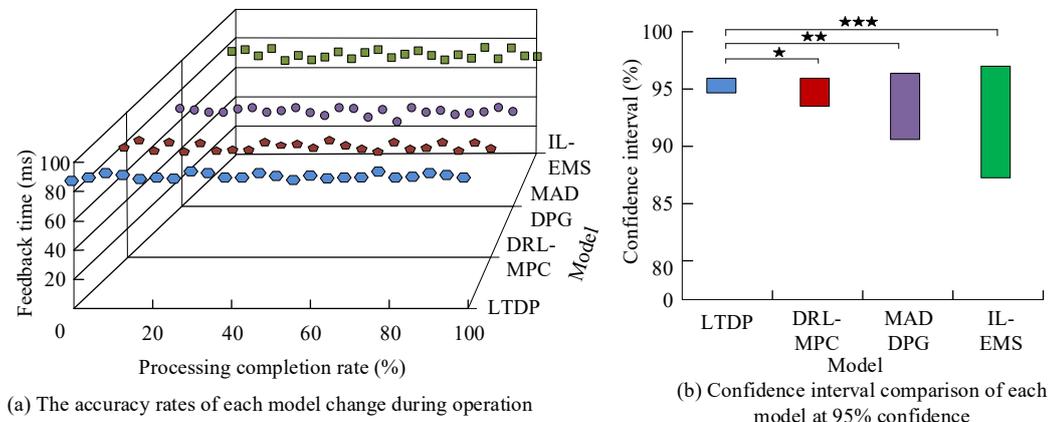


Fig. 10. Accuracy variations and confidence intervals under 95% confidence level

As shown in Fig. 10(a), the A-LTDP model exhibited more concentrated accuracy, with values ranging from 87.3% to 93.1%. In contrast, the accuracy ranges for the benchmark models were 75.6% to 90.1% for DRL-MPC, 78.5% to 88.7% for MADDPG, and 68.7% to 78.5% for IL-EMS. These results confirm that the A-LTDP model achieved superior accuracy. As shown in Fig. 10(b) the A-LTDP model's 95% confidence interval was 94.5% to 96.1%, which is substantially narrower than those of the benchmark models. Therefore, the A-LTDP model demonstrated higher reliability. To further evaluate the energy optimization performance of the A-LTDP model in smart buildings, the model's energy optimization rate over a 24-hour period was compared with that of the real-time optimal strategy. The results are shown in Fig. 11.

As shown in Fig. 11(a), the energy optimization rate of the A-LTDP model closely tracked the optimal strategy, deviating only slightly at hours 4 and 15. The overall alignment rate over the 24-hour period was 97.2%. Among the benchmark models, only the DRL-MPC model achieved relatively high alignment, with a 24-hour alignment rate of 85.1%. The MADDPG and IL-EMS models showed minimal overlap with the optimal curve, achieving alignment rates of only 76.3% and 70.1%, respectively. These results demonstrate that the energy optimization strategies generated by the A-LTDP model were closer to the optimal strategy and delivered superior performance. To evaluate how well each model balanced the supply and demand side requirements, the corresponding satisfaction levels were compared across models. Here,

satisfaction quantifies the alignment between strategy performance and expected goals. Its calculation is based on the deviation between each performance's actual and target values. The smaller the deviation, the higher the satisfaction. The final result is the average of all performance satisfaction scores, ranging from 0% to 100%, with higher values indicating higher satisfaction. The results are shown in Fig. 12.

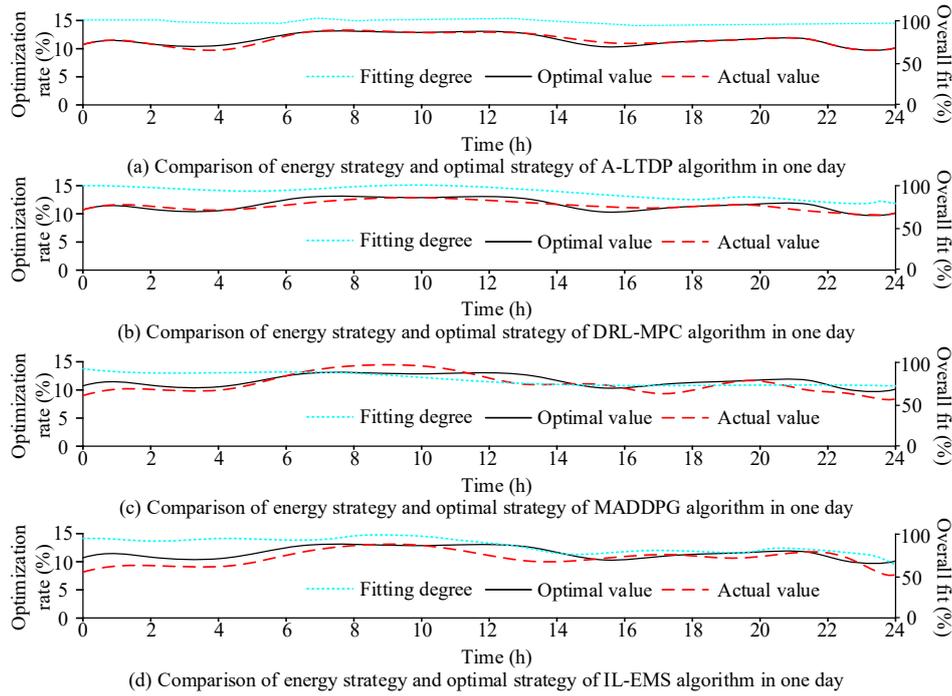


Fig. 11. Comparison of alignment between strategies of each model and the optimal strategy

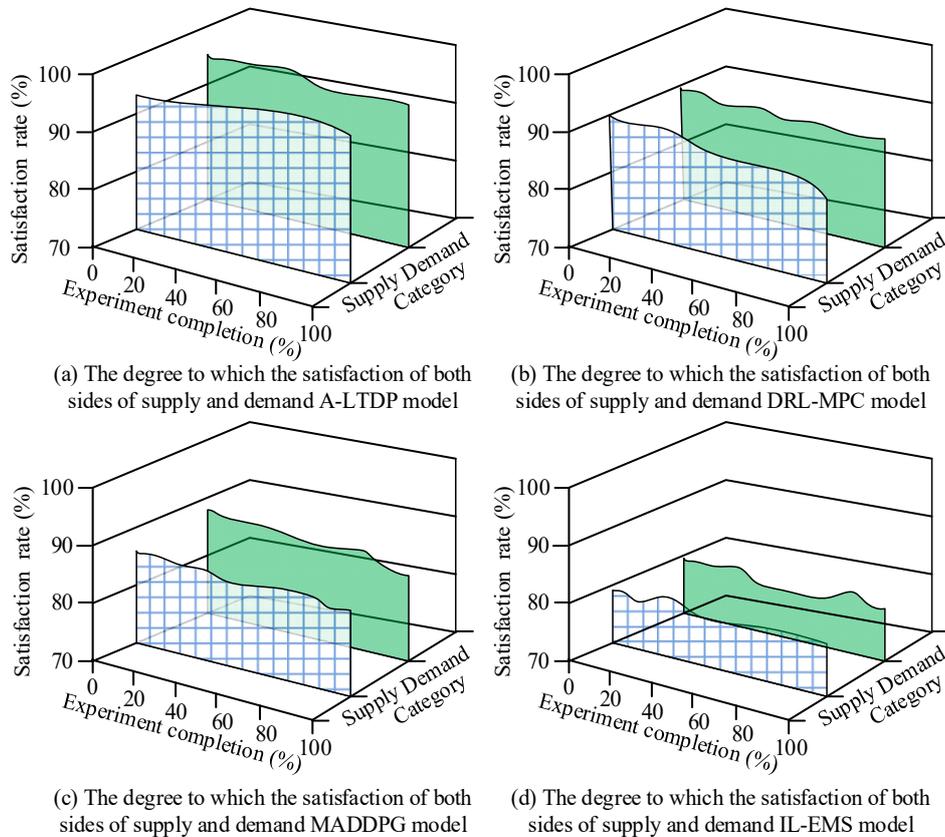


Fig. 12. Comparison of satisfaction levels for supply side and demand side requirements

As shown in Fig. 12(a), the A-LTDP model consistently maintained a satisfaction level above 90% for both sides, reaching maximum satisfaction of 92.1% on the supply side and 95.6% on the demand side. As shown in Fig. 12(b), the

DRL-MPC model's satisfaction levels for both sides remained below 90%, with a maximum of 89.5% on the supply side and 89.8% on the demand side. As shown in Fig. 12(c) and 12(d), the MADDPG and IL-EMS models demonstrated significantly lower satisfaction levels for both supply and demand compared to the A-LTDP model. In conclusion, the A-LTDP model effectively balanced multiple indicators from both the supply and demand sides and generated highly effective energy optimization strategies.

5. Conclusion

To address the limitations of existing smart building energy optimization strategies in handling supply and demand uncertainties, this study integrated various deep learning algorithms and proposed an innovative model named A-LTDP, which considered both supply and demand requirements. Experimental results showed that the LTDP algorithm achieved a strategy performance accuracy of 95.5% upon convergence, with generated strategies meeting at least 70% of the requirements on both the supply and demand sides. In the simulation experiments, the A-LTDP model achieved an accuracy ranging from 87.5% to 93.1%. The 95% confidence interval ranged from 94.5% to 96.1%. During a 24-hour simulation, the strategy generated by the model deviated from the optimal strategy at only two time points, and the overall alignment over the full period reached 97.2%, which significantly outperformed the comparison models. These results demonstrated that the proposed LTDP model successfully integrated the strengths of various deep learning algorithms. It effectively considered uncertainties from both the supply and demand sides, thereby improving the energy efficiency and economic performance of smart buildings. However, since the model integrated multiple algorithms, it might have introduced more computation steps, potentially increasing the response time. This issue was not verified in the current study. In the future, more simulation experiments and real-world tests will be conducted to evaluate this aspect and further optimize the model.

Author Contributions

Xiaojuan Liu contributes to methodology, analysis, draft preparation, and manuscript editing. Zenghui Duan contributes to conceptualization, methodology, data collection, and manuscript editing.

Funding

This research received no specific financial support from any funding agency.

Institutional Review Board Statement

Not applicable.

Declaration of Artificial Intelligence (AI) Tools

The authors confirmed that no AI tools were used in the preparation of this manuscript.

Reference

- Abbasimehr, H., and Paki, R. (2022). Improving time series forecasting using LSTM and attention models. *Journal of Ambient Intelligence and Humanized Computing*, 13(1), 673-691. doi: 10.1007/s12652-020-02761-x
- Al-Selwi, M., Hassan, F., Abdulkadir, J., and Muneer, A. (2023). LSTM inefficiency in long-term dependencies regression problems. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 30(3), 16-31. doi: 10.37934/araset.30.3.1631
- Apanavičienė, R., and Shahrabani, N. (2023). Key factors affecting smart building integration into smart city: Technological aspects. *Smart Cities*, 6(4), 1832-1857. doi: 10.3390/smartcities6040085
- Bhourji, S., Mozaffari, S., and Alirezaee, S. (2024). Reinforcement learning ddpq-pgo agent-based control system for rotary inverted pendulum. *Arabian Journal for Science and Engineering*, 49(2), 1683-1696. doi: 10.1007/s13369-023-07934-2
- Candra, R., Dikananda, R., and Rohman, D. (2025). Improvement of Students' Academic Achievement Classification Model Through the Analytical Hierarchy Process Algorithm in Elementary School Burujul Kulon III. *Journal of Artificial Intelligence and Engineering Applications (JAIEA)*, 4(2), 852-857. doi: 10.59934/jaiea.v4i2.764
- Chakraborty, S., Modi, G., and Singh, B. (2022). A cost optimized-reliable-resilient-realtime-rule-based energy management scheme for a SPV-BES-based microgrid for smart building applications. *IEEE Transactions on Smart Grid*, 14(4), 2572-2581. doi: 10.1109/TSG.2022.3232283
- Dalla-Torre, H., Gonzalez, L., Mendoza-Revilla, J., Lopez Carranza, N., Grzywaczewski, A. H., Oteri, F., and Pierrot, T. (2025). Nucleotide Transformer: building and evaluating robust foundation models for human genomics. *Nature Methods*, 22(2), 287-297. doi: 10.1038/s41592-024-02523-z
- Hassan, S., Park, M., Tun, K., Saad, W., Han, Z., and Hong, C. S. (2023). Satellite-based ITS data offloading & computation in 6G networks: A cooperative multi-agent proximal policy optimization DRL with attention approach. *IEEE Transactions on Mobile Computing*, 23(5), 4956-4974. doi: 10.1109/TMC.2023.3300314
- He, A., Wang, K., Li, T., Du, C., Xia, S., and Fu, H. (2023). H2former: An efficient hierarchical hybrid transformer for medical image segmentation. *IEEE Transactions on Medical Imaging*, 42(9), 2763-2775. doi: 10.1109/TMI.2023.3264513
- Lee, H., Park, H., Jang, W., Han, J., and Choi, K. (2023). PPO-based autonomous transmission period control system in IoT edge computing. *IEEE Internet of Things Journal*, 10(24), 21705-21720. doi: 10.1109/IIOT.2023.3293511
- Ma, S., Zhang, T., Zhao, B., Kang, Y., and Bai, P. (2023). TCLN: A Transformer-based Conv-LSTM network for multivariate time series forecasting. *Applied Intelligence*, 53(23), 28401-28417. doi: 10.1007/s10489-023-04980-z

- Maizana, D., and Putri, S. M. (2022). Appropriateness analysis of implementing a smart grid system in campus buildings using the fuzzy method. *International Journal of Power Electronics and Drive Systems*, 13(2), 873-882. doi: 10.11591/ijpeds.v13.i2.pp873-882
- Mazraedoost, S., Sedigh Malekroodi, H., Zuvella, P., Yi, M., and Liu, J. (2025). Prediction of Chromatographic Retention Time of a Small Molecule from SMILES Representation Using a Hybrid Transformer-LSTM Model. *Journal of Chemical Information and Modeling*, 65(7), 3343-3356. doi: 10.1021/acs.jcim.5c00167
- Muhammad, G., and Hossain, S. (2022). Deep-reinforcement-learning-based sustainable energy distribution for wireless communication. *IEEE Wireless Communications*, 28(6), 42-48. doi: 10.1109/MWC.015.2100177
- Park, Y. M., Hassan, S. S., Tun, Y. K., Han, Z., and Hong, C. S. (2023). Joint trajectory and resource optimization of MEC-assisted UAVs in sub-THz networks: A resources-based multi-agent proximal policy optimization DRL with attention mechanism. *IEEE Transactions on Vehicular Technology*, 73(2), 2003-2016. doi: 10.1109/TVT.2023.3311537
- Sivakumar, S., Logeshwaran, J., Kannadasan, R., Faheem, M., and Ravikumar, D. (2024). A novel energy optimization framework to enhance the performance of sensor nodes in Industry 4.0. *Energy Science & Engineering*, 12(3), 835-859. doi: 10.1002/ese3.1657
- Van Hezewijk, L., Dellaert, N., Van Woensel, T., and Gademann, N. (2023). Using the proximal policy optimisation algorithm for solving the stochastic capacitated lot sizing problem. *International Journal of Production Research*, 61(6), 1955-1978. doi: 10.1080/00207543.2022.2056540
- Volkov, M., Turk, A., Drizard, N., Martin, N., Hoffmann, B., Gaston-Mathé, Y., and Rognan, D. (2022). On the frustration to predict binding affinities from protein–ligand structures with deep neural networks. *Journal of medicinal chemistry*, 65(11), 7946-7958. doi: 10.1021/acs.jmedchem.2c00487
- Zhang, D., Li, Q., Fang, Y., Bai, P., Liu, L., Guo, J., Wang, G., Zhou, Y., and Ma, R. (2024). High-performance wood-based thermoelectric sponges for thermal energy harvesting and smart buildings. *Nano Research*, 17(6), 5349-5357. doi: 10.1007/s12274-024-6467-y
- Zhao, S., and Wu, Z. (2022). Optimisation model of micro grid dispatching for smart building cluster based on blockchain. *Cyber-Physical Systems*, 8(2), 138-171. doi: 10.1080/23335777.2021.1887365



Xiaojuan Liu received her master's degree in Environmental Engineering from Dalian Jiaotong University in 2012. She currently holds the position of lecturer in the Department of Architectural Engineering at the School of Applied Technology, Dalian Ocean University. Her research focuses on engineering economics, environmental engineering, and engineering management.



Zenghui Duan earned her master's degree in Industrial Economics from Liaoning University in 2011. She is currently a lecturer in the Department of Architectural Engineering at the School of Applied Technology, Dalian Ocean University. Her research interests include engineering economics and engineering management.