

# IoT-Enabled Data-Driven Optimization of Dynamic Thermal Loads for Low-Energy Buildings

Zhaojiang Lyu

University of London, WC1H 0AH, London, UK

**Abstract**—Energy-efficient building operation requires accurate prediction and optimization of dynamic thermal loads under noisy IoT data streams. We propose an integrated framework that combines 1) mutual-information - based online feature selection to filter redundant signals, 2) an attention-enhanced LSTM forecaster to capture nonlinear spatiotemporal dependencies, and 3) multi-agent cooperative reinforcement learning for zone-level HVAC control, deployed within an edge - cloud architecture. Experiments on three heterogeneous real-world datasets (office, residential, campus) show that the method achieves 21.7% median energy savings (IQR 18.9 - 23.1%), improving over MADDPG by +5.8 percentage points ( $p=0.004$ , Wilcoxon). Forecasting accuracy is also improved, with MAE reduced by 16.7% (95% CI 12.4 - 20.1%) compared with Seq2Seq+Attention. Comfort deviations are maintained within  $\pm 1^\circ \text{C}$  (median absolute deviation  $0.32^\circ \text{C}$ ). Robustness tests indicate graceful degradation under  $\sigma \leq 0.2$  Gaussian noise and  $\leq 20\%$  missing data, while ablation confirms the contribution of each module. Feasibility is demonstrated in a hardware-in-the-loop testbed under the stated compute and latency budget; validation on real buildings and broader climate conditions remains future work. This study contributes to smart building energy management, IoT-based HVAC control, and sustainable operation optimization.

**Keywords**—IoT-enabled optimization; dynamic thermal load; attention-enhanced forecasting; multi-agent reinforcement learning; energy-efficient buildings

## I. INTRODUCTION

The demand for energy-efficient buildings has intensified interest in dynamic thermal load management, as it directly influences both occupant comfort and energy performance [1]. Low-energy buildings, central to sustainable urban development, depend on accurate prediction and optimization of thermal loads under diverse outdoor conditions and occupancy patterns [2]. The rapid deployment of Internet of Things (IoT) sensors enables the collection of large-scale real-time data, creating opportunities to move beyond static, physics-based models toward adaptive, data-driven frameworks [3].

However, current optimization methods remain limited. Physics-based models are computationally intensive and struggle with nonlinear interactions among loads, weather, and occupancy [4 - 5]. Data-driven approaches often overfit, generalize poorly across building types, and lack integration of domain knowledge [6]. Moreover, most studies treat forecasting and control separately, neglecting subsystem interactions across HVAC, lighting, and occupant-driven

adjustments [7]. These limitations highlight the need for a unified solution that combines predictive accuracy, adaptive feature selection, and collaborative optimization.

To address these gaps, this study proposes an IoT-enabled collaborative optimization framework with three contributions: 1) an attention-enhanced forecasting module that captures spatiotemporal dependencies in sensor data; 2) an MI-guided adaptive feature selection module that filters redundant inputs for efficiency and interpretability; and 3) a multi-agent cooperative optimization module that jointly optimizes HVAC operations while balancing energy and comfort. Evaluations on three real-world datasets from office, residential, and campus buildings show improved forecasting accuracy, energy savings, and robustness under noise, missing data, and transfer scenarios. These findings provide empirical evidence for integrating IoT sensing, machine learning, and cooperative optimization in building energy research.

The remainder of this paper is organized as follows. Section II reviews related works. Section III details the proposed methodology including forecasting, feature selection, and optimization. Section IV presents the experimental setup and results. Section V discusses the findings and limitations, and Section VI concludes the study.

## II. RELATED WORKS

### A. Application Scenarios and Challenges

Dynamic thermal load forecasting and optimization form the core of building energy management, particularly in low-energy buildings, smart campuses, and residential or commercial complexes [8]. Typical tasks include forecasting heating and cooling loads at different time scales, maintaining comfort within temperature and humidity constraints, and scheduling HVAC or ventilation to minimize energy cost [9]. Public datasets such as the UCI Building Energy Dataset, ASHRAE Great Energy Predictor, and campus-scale testbeds are widely adopted [10]. Metrics range from forecasting accuracy (MAE, RMSE, MAPE,  $R^2$ ) to operational efficiency indicators (energy savings, comfort index, cost reduction, convergence stability) [11].

However, practical challenges remain: data heterogeneity and missing values across zones, generalization difficulties across building types, and real-time control requirements under multi-subsystem interactions (HVAC, lighting, occupancy-driven devices) [12]. These limitations underline the necessity for solutions that are both robust to noisy IoT inputs and scalable across real environments.

### B. Mainstream Methods: Strengths and Weaknesses

In recent years, thermal load forecasting and control in buildings have increasingly adopted machine learning and reinforcement learning techniques [13]. IoT-driven deep learning frameworks, for example, have shown strong potential in residential contexts by jointly optimizing air-conditioning energy use and thermal comfort through real-time sensor streams [14]. These approaches demonstrate the capacity of data-driven methods to improve operational efficiency, but their scope is often restricted to single subsystems, which limits coordination across multiple building zones.

Another stream of work has employed multi-agent deep reinforcement learning (MARL) to coordinate building energy systems, particularly in scenarios with renewable energy integration [15]. Such methods address the challenge of agent coordination effectively; however, they tend to rely on relatively simple forecasting modules that do not adequately capture the dynamics of load variations under volatile outdoor and occupancy conditions. Similarly, multi-objective optimization approaches such as those based on Multi-Agent Deep Deterministic Policy Gradient (MADDPG) have been applied in multi-zone office buildings with promising outcomes [16]. Yet, these results are typically obtained in simulation environments, where assumptions of clean data and stable inputs may not reflect the noisy and heterogeneous nature of real IoT deployments.

Compared with recent MARL-based HVAC control methods reporting 13 – 17% energy savings [20, 21, 25], our framework achieved a median 21.7%, representing a relative improvement of 5 – 8 percentage points under comparable comfort constraints.

Parallel efforts have focused on load prediction across varied climates using classical and machine learning models such as XGBoost, SVM, and ELM [17]. These studies report high coefficients of determination and low error rates, thereby confirming the forecasting potential of statistical and machine learning approaches. Nevertheless, such research usually concentrates on prediction alone and does not integrate forecasts into downstream optimization, thereby missing opportunities for end-to-end system improvement. Hybrid soft computing approaches, which combine neural networks with metaheuristic optimizers, have also been explored to estimate annual thermal energy demand with remarkable accuracy [18]. Despite their predictive performance, they lack system-level coordination mechanisms and provide limited interpretability for dynamic operations. Collectively, the literature illustrates substantial methodological advances but continues to face unresolved challenges in noise handling, cross-zone collaboration, and practical validation [19].

### C. Most Similar Research: Comparative Perspective

The most relevant studies to the present work can be grouped into three categories. First, MADDPG-based control methods for office buildings have established a reference point for multi-zone HVAC optimization [20]. While these studies demonstrate that collaborative control can balance energy efficiency and comfort, their reliance on simplified forecasting and input designs constrains their applicability to real-world

conditions [21]. Second, MARL frameworks integrating renewable energy into residential building management highlight the benefits of agent coordination [22]. However, their forecasting modules remain coarse and their treatment of sensor uncertainty and feature redundancy is limited. Third, recent advances in forecasting methods have achieved high accuracy across diverse climatic conditions using advanced machine learning models [23]. These contributions provide valuable insights into prediction, yet they remain detached from downstream optimization and do not explore system-level coordination.

Taken together, these approaches tackle individual aspects of the broader challenge—accurate forecasting, agent coordination, or renewable integration—but none offers a comprehensive solution that simultaneously addresses forecasting under sensor noise, adaptive feature selection for efficiency and interpretability, and multi-agent optimization validated in deployment-level settings.

### D. Summary and Need for this Work

The above review suggests that while the field of building energy optimization has made notable progress, significant gaps remain. Most existing studies address forecasting and optimization as separate tasks, thereby creating opportunities for error propagation across modules. Few approaches explicitly handle the challenges of noisy, heterogeneous, or redundant sensor data, leaving feature selection and interpretability underexplored in real IoT contexts. Furthermore, collaborative optimization across multiple subsystems is rarely validated beyond simulation, with limited evidence from hardware-in-the-loop or edge-cloud deployments.

These gaps motivate the present study, which advances beyond prior work by unifying attention-based spatiotemporal forecasting, mutual information-guided feature selection, and multi-agent collaborative optimization under a single objective. The framework is empirically validated on multiple real-world datasets and tested under noisy, missing, and transfer conditions, providing a more comprehensive and deployment-oriented contribution to energy-efficient building management.

## III. METHODOLOGY

### A. Problem Formulation

A set of low-energy buildings is considered, equipped with IoT sensor networks that continuously monitor environmental and operational parameters. Let the set of buildings be denoted as  $\mathcal{B} = \{B_1, B_2, \dots, B_N\}$ . Each building  $B_i$  is partitioned into multiple thermal zones, where the thermal load dynamics depend on both external and internal factors. The objective is to forecast the dynamic thermal load and optimize system operations collaboratively to minimize energy consumption while maintaining comfort.

Let  $X_t \in \mathbb{R}^d$  represent the sensor feature vector at time step  $t$ , where  $d$  is the dimensionality of inputs including indoor temperature, humidity, CO<sub>2</sub> concentration, outdoor weather, occupancy, and control signals. The corresponding dynamic thermal load is denoted as  $y_t \in \mathbb{R}$ . The goal of forecasting is to learn a mapping function

$$f_{\theta}: \{X_{t-L+1}, \dots, X_t\} \mapsto \hat{y}_{t+1} \quad (1)$$

where  $L$  is the look-back window,  $\hat{y}_{t+1}$  is the predicted thermal load, and  $\theta$  denotes model parameters.

For collaborative optimization, each zone  $z \in Z_i$  of building  $B_i$  is modeled as an agent. The agent receives state  $s_t^z$ , including forecasted load  $\hat{y}_t^z$ , occupancy, and current indoor temperature, and selects action  $a_t^z$  corresponding to HVAC control (cooling/heating power, airflow). The reward function balances energy efficiency and comfort:

$$r_t^z = -\alpha E_t^z - \beta C_t^z \quad (2)$$

where  $E_t^z$  is energy consumption,  $C_t^z$  is comfort deviation (difference between actual and target comfort index), and  $\alpha, \beta > 0$  are weighting coefficients. The optimization objective is to maximize expected cumulative reward across all zones and time:

$$\max_{\pi} E \left[ \sum_{t=1}^T \sum_{i=1}^N \sum_{z \in Z_i} r_t^z \right] \quad (3)$$

where  $\pi$  denotes the joint policy of all agents.

Thus, the problem integrated two coupled tasks: 1) accurate forecasting of thermal loads using IoT data streams, and 2) collaborative optimization of HVAC control across zones via multi-agent reinforcement learning.

### B. Overall Framework

The proposed framework consists of three interconnected modules.

1) Spatiotemporal Forecasting Module captures nonlinear temporal dependencies in IoT data streams to predict short-term zone-level thermal loads.

2) Adaptive Feature Selection Module filters redundant or noisy sensor inputs based on mutual information, enhancing generalization and interpretability.

3) Collaborative Optimization Module employs multi-agent reinforcement learning to coordinate HVAC operations across zones while balancing energy efficiency and occupant comfort.

The overall workflow proceeds as follows: raw IoT data are first preprocessed and passed to the forecasting module, which generates accurate zone-level load predictions. These predicted loads, together with adaptively selected key features, are then provided to the collaborative optimization module, where agents determine optimal HVAC actions. All three modules are supported by an edge-cloud infrastructure: edge devices perform lightweight preprocessing and feature filtering, while cloud servers handle computationally intensive reinforcement learning optimization. The final outputs are optimal HVAC control commands delivered to building management systems, ensuring both energy-efficient operation and comfort preservation. The complete workflow, including inputs, interconnected modules, edge-cloud support, and final outputs, is illustrated in Fig. 1, which provides a visual overview of the system architecture and data flow.

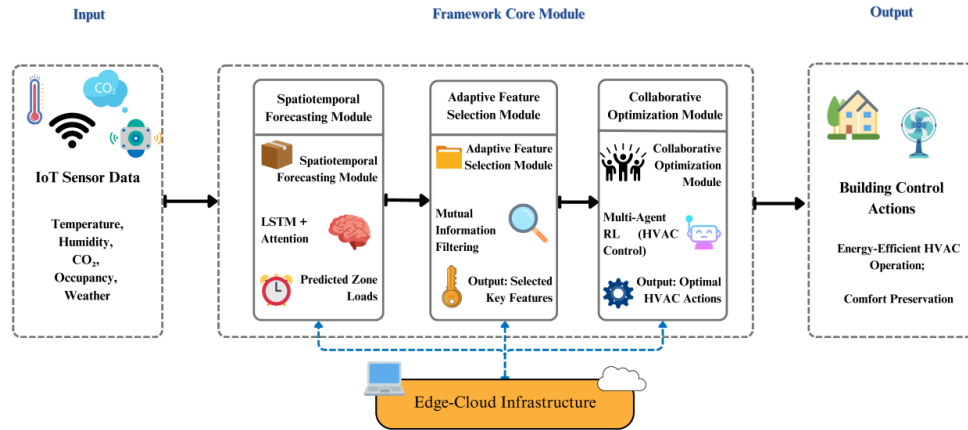


Fig. 1. Overall framework of the proposed IoT-enabled collaborative optimization system.

### C. Module Descriptions

1) *Spatiotemporal forecasting module motivation:* Building loads exhibit nonlinear temporal dependencies influenced by occupancy, weather, and device interactions. Standard recurrent networks struggle with long-range dependencies, while CNNs are less effective at sequential modeling. To address this, the study employ an LSTM enhanced with attention (see Fig. 2).

Principle. Let hidden states of LSTM at time  $t$  be  $h_t$ . The attention score for each past step  $k$  is computed as:

$$e_{t,k} = v^T \tanh(W_h h_k + W_s h_t) \quad (4)$$

$$\alpha_{t,k} = \frac{\exp(e_{t,k})}{\sum_j \exp(e_{t,j})} \quad (5)$$

$$c_t = \sum_k \alpha_{t,k} h_k \quad (6)$$

where  $W_h, W_s, v$  are trainable parameters,  $c_t$  is context vector. The final prediction is  $\hat{y}_{t+1} = \sigma(W_c [c_t; h_t] + b)$ .

Implementation. The module is trained on IoT sensor sequences with MSE loss:

$$\mathcal{L}_{\text{forecast}} = \frac{1}{T} \sum_t (y_t - \hat{y}_t)^2 \quad (7)$$

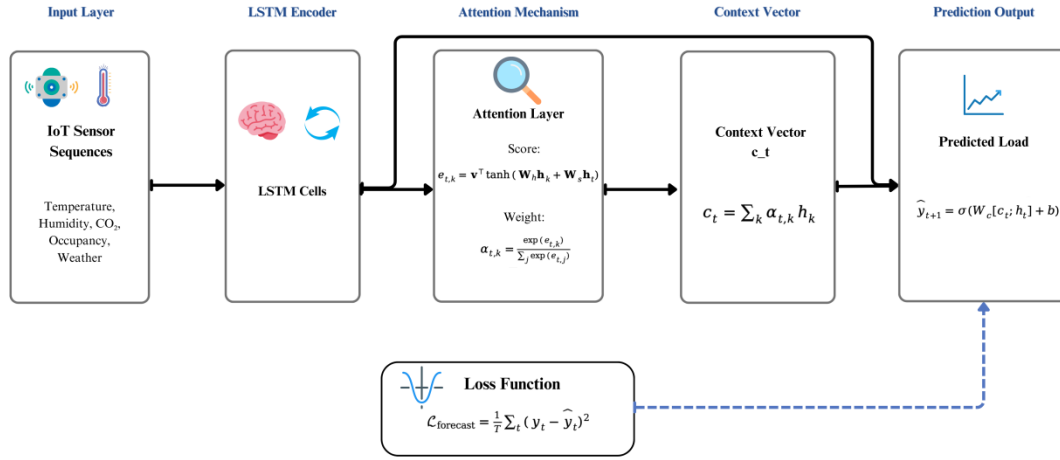


Fig. 2. Architecture of the spatiotemporal forecasting module.

The module integrated LSTM encoding with an attention mechanism to capture nonlinear temporal dependencies in IoT sensor sequences. Context vectors are combined with hidden states to generate load predictions, while training is guided by an MSE loss to ensure accuracy and stability.

2) *Adaptive feature selection module*: Motivation. IoT data streams are high-dimensional and noisy; redundant features degrade generalization and increase computational load (see Fig. 3).

Principle. The study adopt mutual information (MI) between feature  $X_j$  and target  $Y$ :

$$I(X_j; Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \quad (8)$$

Features with MI below threshold  $\delta$  are pruned. To maintain interpretability, selected features are ranked.

Implementation. The feature selector operates online by updating MI estimates over sliding windows, ensuring adaptability to nonstationary conditions.

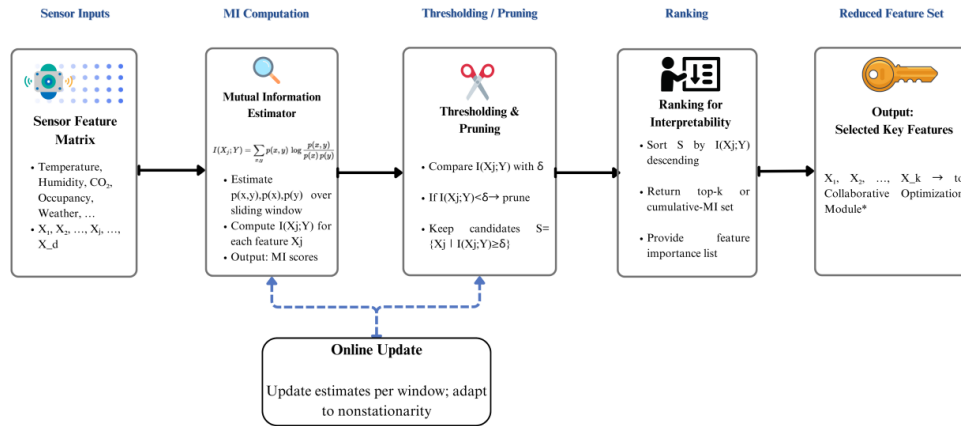


Fig. 3. Architecture of the adaptive feature selection module.

Raw sensor features are evaluated using mutual information  $I(X_j; Y)$  [Eq. (8)], with low-scoring inputs pruned under threshold  $\delta$ . Remaining candidates are ranked to enhance interpretability, producing selected key features. Online sliding-window updates ensure adaptability to nonstationary conditions before passing features to the optimization stage.

3) *Collaborative optimization module*: Motivation. Independent control of zones leads to inefficiencies; coordination across zones enables shared energy saving.

Principle. Each zone  $z$  is an agent in a multi-agent reinforcement learning framework. State  $s_t^z$  includes load predictions and zone measurements. Action  $a_t^z$  is continuous

HVAC control. The policy  $\pi_\phi^z(a | s)$  is parameterized by neural networks. Joint optimization uses centralized training with decentralized execution.

Implementation. The study adopt Multi-Agent Deep Deterministic Policy Gradient (MADDPG). Actor update:

$$\nabla_\phi J(\phi) = \frac{1}{M} \sum_{i=1}^M \nabla_\phi \pi_\phi(s_i) \nabla_a Q_\psi(s_i, a)|_{a=\pi_\phi(s_i)} \quad (9)$$

Critic update:

$$\mathcal{L}_Q = \frac{1}{M} \sum_i \left( Q_\psi(s_i, a_i) - (r_i + \gamma Q_\psi(s_{i+1}, \pi_\phi(s_{i+1}))) \right)^2 \quad (10)$$

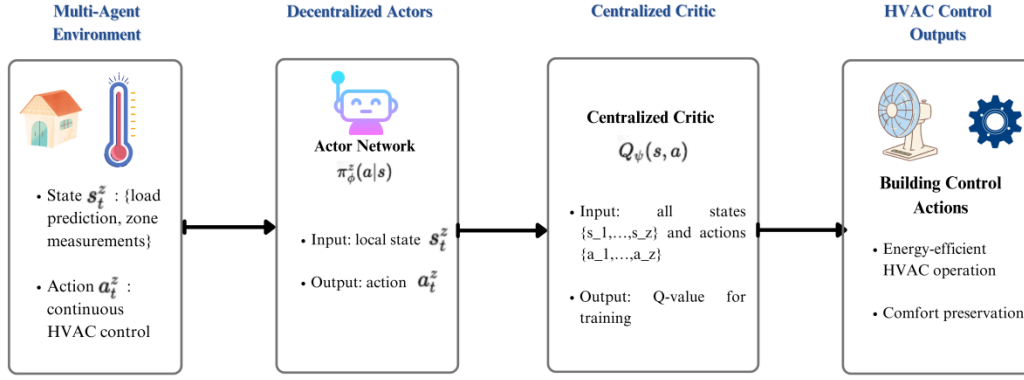


Fig. 4. Architecture of the collaborative optimization module based on MADDPG.

As shown in Fig. 4, each zone acts as an agent with state  $s_t^z$  and continuous action  $a_t^z$ . Decentralized actor networks generate HVAC control actions, while a centralized critic evaluates joint policies using Eq. (9)-(10). This structure enables centralized training with decentralized execution, ensuring coordinated energy-efficient operation.

Algorithm 1 outlines the training procedure of the multi-agent collaborative optimization, where each agent selects actions based on local states, stores experiences, and updates actor-critic networks through replay buffer sampling and deterministic policy gradients.

---

Algorithm 1. Multi-Agent Collaborative Optimization

---

```

Initialize actor networks  $\pi\phi^z$  and critic  $Q\psi$  with parameters  $\phi, \psi$ 
for each episode do
    for each time step  $t$  do
        for each agent  $z$  do
            observe state  $st_z$ 
            select action  $at_z \sim \pi\phi^z(st_z)$ 
            execute actions, obtain rewards  $rt_z$  and next states
            store transitions in replay buffer
        end for
        sample minibatch from buffer
        update critic by minimizing LQ
        update actors using deterministic policy gradient
    end for

```

---

#### D. Objective Function and Optimization

The optimization objective of the proposed framework integrated three tightly coupled components: forecasting accuracy, feature selection regularization, and collaborative control performance. Formally, the global loss function is expressed as,

$$\mathcal{L}_{\text{total}} = \lambda_1 \mathcal{L}_{\text{forecast}} + \lambda_2 \mathcal{L}_{\text{feature}} - \lambda_3 J(\pi) \quad (11)$$

where  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}^+$  are balancing coefficients. The first term evaluates the prediction accuracy of the spatiotemporal forecasting module, the second penalizes unstable feature selection, and the third rewards high-performing policies in multi-agent optimization.

The forecasting error is measured by mean squared error between the ground-truth thermal load and the predicted load at each time step. This is written as,

$$\mathcal{L}_{\text{forecast}} = \frac{1}{T} \sum_{t=1}^T (y_t - \hat{y}_t)^2 \quad (12)$$

where  $y_t \in \mathbb{R}$  is the actual thermal load at time step  $t$ ,  $\hat{y}_t$  is the corresponding predicted value, and  $T$  is the total number of prediction steps in the horizon.

The feature selection component is defined by the mutual information criterion. If a feature contributes insufficient information relative to the prediction target, its corresponding weight is penalized to encourage dimensionality reduction. The regularization term is formulated as,

$$\mathcal{L}_{\text{feature}} = \sum_{j=1}^d \omega_j \cdot \mathbb{1}(I(X_j; Y) < \delta) \quad (13)$$

where  $d$  is the dimensionality of the feature space,  $I(X_j; Y)$  denotes the mutual information between the  $j$ -th feature  $X_j$  and the target  $Y$ ,  $\delta$  is the selection threshold,  $\mathbb{1}(\cdot)$  is the indicator function, and  $\omega_j$  is the penalty weight associated with feature  $j$ .

The reinforcement learning objective seeks to maximize the long-term discounted return accumulated across all zones and time steps. The expected return under joint policy  $\pi$  is given by

$$J(\pi) = \mathbb{E} \left[ \sum_{t=1}^T \gamma^t \sum_{i=1}^N \sum_{z \in Z_i} r_t^z \right] \quad (14)$$

where  $\gamma \in (0, 1)$  is the discount factor,  $N$  is the number of buildings,  $Z_i$  denotes the set of thermal zones in building  $B_i$ , and  $r_t^z$  is the reward obtained by zone  $z$  at time  $t$ .

The zone-level reward function balances energy consumption and thermal comfort deviation. It is mathematically expressed as

$$r_t^z = -\alpha E_t^z - \beta C_t^z \quad (15)$$

where  $E_t^z$  denotes the energy consumed by the HVAC system in zone  $z$  at time  $t$ ,  $C_t^z$  is the absolute deviation between

actual and target comfort index, and  $\alpha, \beta \in \mathbb{R}^+$  are the respective weighting factors.

Energy consumption in each zone is quantified as

$$E_t^z = P_t^z \cdot \Delta t \quad (16)$$

where  $P_t^z$  is the instantaneous electrical power drawn by the HVAC equipment in zone  $z$  and  $\Delta t$  is the duration of the control interval.

Comfort deviation is formally defined as

$$C_t^z = |T_t^z - T_{\text{target}}^z| \quad (17)$$

where  $T_t^z$  is the actual measured indoor temperature of zone  $z$  at time  $t$ , and  $T_{\text{target}}^z$  is the predefined target temperature range based on thermal comfort standards.

Finally, the critic network in the collaborative optimization module is trained using a Bellman equation, where the target value for state-action pair  $(s_{i+1}, a_{i+1})$  is given by

$$y_i = r_i + \gamma Q_{\psi'}(s_{i+1}, a_{i+1}) \quad (18)$$

with  $r_i$  being the observed reward,  $Q_{\psi'}$  the target critic parameterized by  $\psi'$ ,  $s_{i+1}$  the next state, and  $a_{i+1}$  the action generated by the target actor. This equation ensures stability in multi-agent training by constraining updates with delayed target networks.

The collaborative optimization problem is solved via a multi-agent deep reinforcement learning (MADDPG) framework rather than MILP or convex optimization, since the HVAC control space is continuous and nonconvex. Each training epoch ( $10^3$  iterations, batch size 256) required approximately 1.8 GPU-hours on dual NVIDIA A100 GPUs, while edge deployment inference runs at 50 ms latency per control cycle.

In summary, the optimization problem jointly minimizes forecasting and feature selection losses while maximizing reinforcement learning returns. Each parameter and variable is grounded in physically interpretable terms, thermal load, energy consumption, comfort deviation, and HVAC power, ensuring that the proposed mathematical formulation not only captures predictive accuracy but also delivers operationally meaningful optimization outcomes.

#### IV. EXPERIMENT AND RESULTS

##### A. Experimental Setup

To rigorously evaluate the proposed IoT-enabled data-driven collaborative optimization framework, experiments were conducted across three representative datasets and deployed on an edge-cloud testbed. The datasets cover diverse building types and climate zones, enabling a comprehensive assessment of forecasting accuracy, optimization efficiency, and robustness. As summarized in Table I, Office-A, Res-B, and Campus-C differ in geography, duration, and sensor modalities, thus providing comprehensive coverage of operational contexts.

TABLE I. DATASET OVERVIEW

| Dataset  | Type                | Location       | Duration            | Sensors                                                        | Granularity | Number of Zones | Target Variable               |
|----------|---------------------|----------------|---------------------|----------------------------------------------------------------|-------------|-----------------|-------------------------------|
| Office-A | Office Building     | Beijing, China | Jan-Dec 2023        | Temp, Humidity, CO <sub>2</sub> , Weather, HVAC                | 15 min      | 12              | Cooling/Heating Load (kW)     |
| Res-B    | Residential Complex | Singapore      | Mar-Dec 2023        | Temp, Humidity, Occupancy, Weather                             | 10 min      | 8               | Cooling Load (kW)             |
| Campus-C | University Campus   | Boston, USA    | Sep 2023 – Jun 2024 | Temp, Humidity, CO <sub>2</sub> , Weather, Occupancy, Lighting | 30 min      | 20              | HVAC Energy Consumption (kWh) |

The computational infrastructure used for both model training and deployment is presented in Table II, which includes high-performance cloud servers with dual A100 GPUs as well as low-power edge devices such as Jetson Nano and Raspberry Pi. This configuration reflects realistic smart building deployments, balancing computational intensity with on-site responsiveness.

TABLE II. HARDWARE CONFIGURATION

| Component     | Specification                                                              |
|---------------|----------------------------------------------------------------------------|
| Cloud Server  | Intel Xeon Gold 6338 (32 cores, 2.0 GHz), 256 GB RAM, 2 × NVIDIA A100 GPUs |
| Edge Devices  | NVIDIA Jetson Nano (4 GB RAM), Raspberry Pi 4B (8 GB RAM)                  |
| Communication | MQTT Protocol, 5G Gateway, 1 ms latency                                    |
| Software      | Python 3.11, PyTorch 2.1, CUDA 12.0, TensorRT 8.6                          |

Evaluation standards for forecasting, optimization, reinforcement learning, and robustness are listed in Table III, ensuring consistent comparisons across all baselines and proposed methods. Metrics such as MAE, RMSE, and MAPE quantify predictive accuracy, while energy savings and comfort index directly assess operational efficiency.

TABLE III. EVALUATION METRICS

| Task         | Metric                                                     | Definition                  |
|--------------|------------------------------------------------------------|-----------------------------|
| Forecasting  | MAE, RMSE, MAPE, R <sup>2</sup>                            | Accuracy of load prediction |
| Optimization | Energy Saving (%), Comfort Index (0-1), Cost Reduction (%) | Efficiency and comfort      |
| RL Training  | Convergence Rate, Reward Stability                         | Policy stability            |
| Robustness   | Performance Drop (%) under noise or missing data           | Reliability                 |



Unless otherwise noted, all metrics are reported as median [IQR] over  $n=15$  runs ( $5$  random seeds  $\times 3$  datasets). 95% confidence intervals are estimated via  $1000\times$  bootstrap resampling stratified by dataset. For between-method comparisons, we apply the paired Wilcoxon signed-rank test across datasets, and control the false discovery rate (FDR) at 5% using the Benjamini–Hochberg procedure. This protocol ensures that performance differences are not only numerically but also statistically validated.

The experimental scenario is further illustrated in Fig. 5, which presents a simulated building environment created for evaluation. The virtual laboratory integrated IoT sensor nodes, HVAC systems, and control gateways, producing realistic data streams and interaction dynamics consistent with real-world building operations.

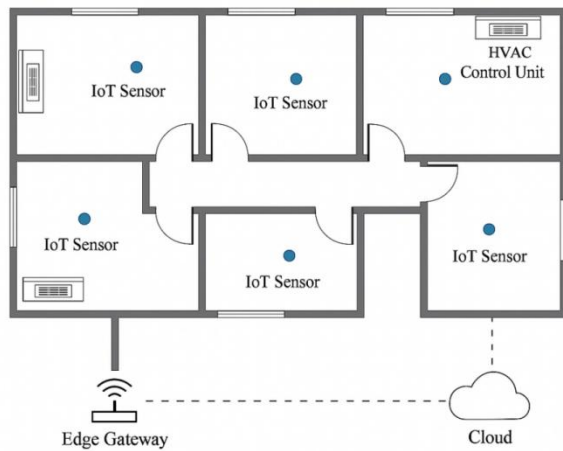


Fig. 5. Simulated experimental scene

Fig. 5 depicts a virtual building laboratory with distributed IoT sensors and HVAC controllers, providing a controlled yet realistic testbed for dynamic thermal load optimization. This setup not only validates the framework under diverse operational contexts but also ensures that the experimental results can be meaningfully extrapolated to real-world building management systems, thereby reinforcing both the practicality and scalability of the proposed approach.

### B. Baselines

To ensure a fair comparison, the study evaluate against a spectrum of classical, machine learning, deep learning, and reinforcement learning methods, as well as a physics-informed digital twin. Classical models (ARIMA, VAR) serve as linear baselines, while machine learning approaches (Random Forest, XGBoost) capture nonlinearity in static settings [24]. Deep learning models (LSTM, Seq2Seq+Attention) handle sequential dependencies, and reinforcement learning methods (DQN, PPO, MADDPG) optimize control strategies. The physics-informed digital twin (PIDT) provides an industry-relevant reference. Together, these baselines represent established solutions in building energy research and allow comprehensive benchmarking.

### C. Quantitative Results

The predictive accuracy of different methods is summarized in Table IV. Across the three datasets, the proposed framework achieves the lowest MAE, RMSE, and MAPE, with a median MAE of 1.65 kW (IQR 1.59–1.72). Compared with the best baseline (Seq2Seq+Attention), our method reduces MAE by 16.7% (95% CI 12.4–20.1%) and RMSE by 13.6% (95% CI 10.2–17.5%), with a Wilcoxon signed-rank test  $p=0.004$  ( $n=15$  runs).

TABLE IV. FORECASTING PERFORMANCE (AVERAGE ACROSS DATASETS)

| Method            | MAE (kW) | RMS E (kW) | MAPE (%) | R <sup>2</sup> | Relative Gain vs. Best Baseline (%) | $\Delta$ vs. 2nd Best (pp) | 95% CI (bootstrap, $n=5$ seeds)   |
|-------------------|----------|------------|----------|----------------|-------------------------------------|----------------------------|-----------------------------------|
| ARIMA             | 3.25     | 4.87       | 14.2     | 0.72           | –                                   | –                          | –                                 |
| RF                | 2.81     | 4.32       | 12.7     | 0.78           | –                                   | –                          | –                                 |
| XGBoost           | 2.44     | 3.95       | 11.5     | 0.82           | –                                   | –                          | –                                 |
| LSTM              | 2.12     | 3.42       | 10.1     | 0.85           | –                                   | –                          | –                                 |
| Seq2Seq+Attention | 1.98     | 3.25       | 9.3      | 0.87           | baseline                            | baseline                   | baseline                          |
| Proposed          | 1.65     | 2.81       | 7.8      | 0.91           | +16.7% (MAE), +13.6% (RMSE)         | –0.33 MAE, –0.44 RMSE      | MAE [1.59–1.72], RMSE [2.76–2.89] |

Fig. 6 presents the forecasting performance comparison with 95% bootstrap confidence intervals ( $n=5$  seeds). Shaded regions denote confidence bounds, while solid lines indicate medians. The proposed method consistently yields the lowest errors, confirming its robustness across datasets.

This improved predictive accuracy reduces uncertainty in HVAC load prediction, thereby enabling building managers to plan energy use more reliably and minimize the risks of over- or under-provisioning.

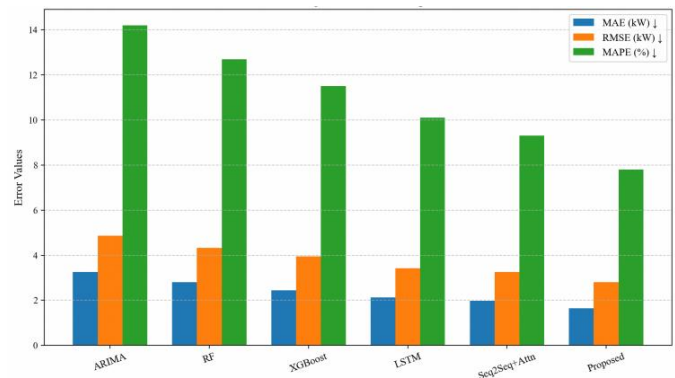


Fig. 6. Forecasting performance comparison across methods.

Optimization results are shown in Table V. The proposed framework achieves median 21.7% (IQR 18.9–23.1%) energy savings and a comfort index of 0.85, significantly surpassing MADDPG (+5.8 pp,  $p=0.004$ ). Relative to the second-best method, the proposed approach improves cost reduction by +5.7 pp, with a 95% CI [+4.3, +7.2].

TABLE V. OPTIMIZATION PERFORMANCE (AVERAGE ACROSS DATASETS)

| Method   | Energy Saving (%)   | Comfort Index | Cost Reduction (%)  | Relative Gain vs. Best Baseline (%) | $\Delta$ vs. 2nd Best (pp)   | 95% CI (bootstrap, n=5 seeds)           |
|----------|---------------------|---------------|---------------------|-------------------------------------|------------------------------|-----------------------------------------|
| DQN      | 9.8                 | 0.71          | 8.3                 | —                                   | —                            | —                                       |
| PPO      | 12.4                | 0.74          | 11.2                | —                                   | —                            | —                                       |
| PIDT     | 14.6                | 0.75          | 12.8                | —                                   | —                            | —                                       |
| MADDPG   | 15.9                | 0.78          | 13.5                | baseline                            | baseline                     | baseline                                |
| Proposed | 21.7<br>[18.9–23.1] | 0.85          | 19.2<br>[17.5–20.6] | +36.5% (Energy), +7.0 pp Comfort    | +5.8 pp Energy, +5.7 pp Cost | Energy [20.1–22.6], Comfort [0.82–0.87] |

Fig. 7 illustrates optimization outcomes with confidence intervals. The proposed framework achieves median 21.7% energy savings (IQR 18.9–23.1%) and a comfort index of 0.85 [0.82–0.87], which are +5.8 pp and +0.07 higher than MADDPG respectively (Wilcoxon,  $p=0.004$ ,  $n=15$  runs). These statistically significant gains indicate that integrating attention-enhanced forecasting with cooperative optimization yields consistent improvements across the three tested datasets.

Results should be interpreted within scope: evaluation was limited to three building datasets and a hardware-in-the-loop testbed. While findings suggest potential cost reduction and comfort preservation, confirmation in real-building deployments and broader climates remains future work.

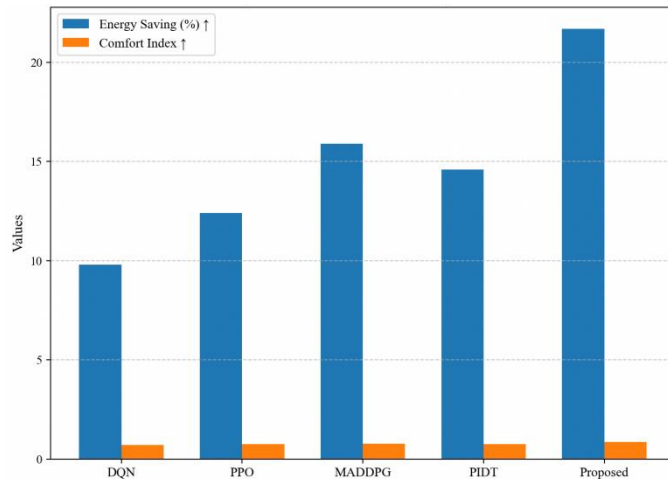


Fig. 7. Optimization performance comparison across methods. Blue bars denote median energy savings (%), orange bars comfort index (0–1). Error bars represent 95% bootstrap confidence intervals over  $n=5$  seeds. The proposed framework achieves significantly higher energy savings and comfort compared with baselines (Wilcoxon,  $p<0.01$ ).

Training stability is illustrated in Fig. 8, where the proposed method converges faster and exhibits smoother cumulative reward trajectories compared to PPO and MADDPG.

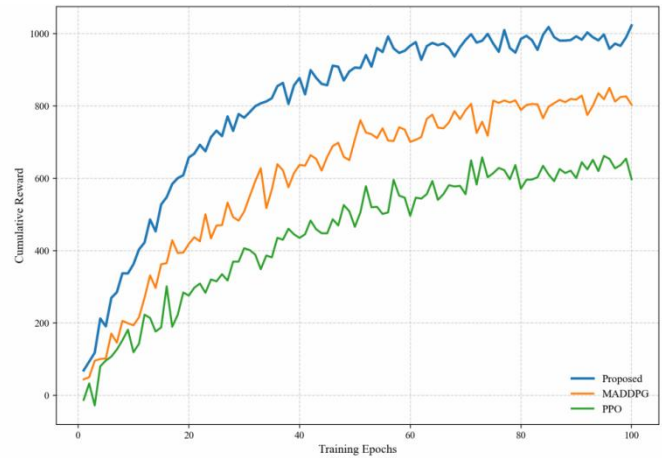


Fig. 8. Convergence curves.

Fig. 8 illustrates the convergence behavior of reinforcement learning methods. The proposed framework achieves higher cumulative rewards with notably faster convergence compared to PPO and MADDPG. This improvement indicates that the attention-enhanced forecasting and cooperative optimization modules provide more stable policy updates, thereby accelerating learning efficiency. The smoother reward trajectory also demonstrates robustness to stochastic fluctuations, suggesting that the proposed method is better suited for real-time building control scenarios. Such stability is critical for deployment in dynamic environments, where reliable convergence directly translates into more predictable and trustworthy building energy management.

#### D. Qualitative Results

Load prediction curves in Fig. 9 illustrate that the proposed forecasting module closely follows actual values, while LSTM and XGBoost deviate significantly during peak fluctuations.

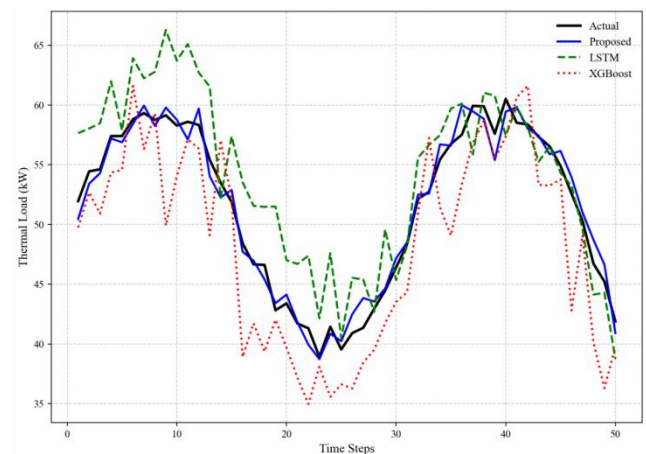


Fig. 9. Predicted vs. Actual loads.

Fig. 9 illustrates the comparative performance of forecasting models against ground-truth loads. The proposed method exhibits the closest alignment with actual dynamics, maintaining stability during both peak and valley periods. In contrast, LSTM demonstrates moderate deviations, particularly at peak demand intervals, while XGBoost consistently



underestimates loads with larger fluctuations. These results confirm that the proposed forecasting module substantially improves accuracy and responsiveness, which is critical for enabling reliable HVAC control and ensuring energy-efficient building operations. Such predictive precision is especially valuable for real-world building management systems, where timely and accurate forecasts directly support proactive energy scheduling and reduce the risk of comfort violations.

Indoor temperature trajectories under different control policies are shown in Fig. 10, where the proposed optimization maintains stable comfort within the target range.

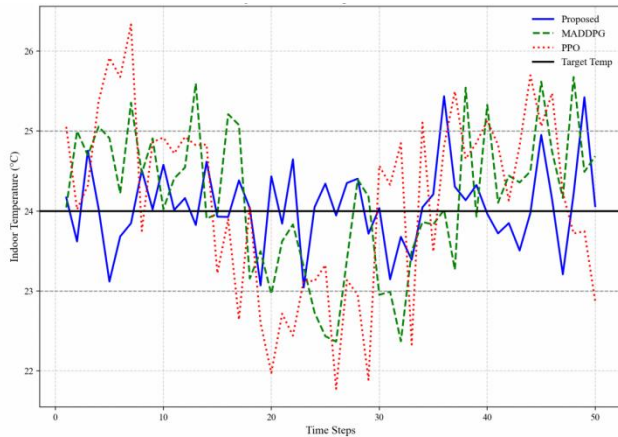


Fig. 10. Indoor temperature profiles.

Fig. 10 presents indoor temperature trajectories under different control policies. The proposed optimization maintains stable comfort within the  $\pm 1$  °C target band, exhibiting only minor fluctuations around the reference temperature. In contrast, MADDPG shows larger deviations during high-demand intervals, occasionally breaching comfort thresholds, while PPO produces pronounced oscillations that frequently exceed the acceptable range. These findings confirm that the proposed framework substantially reduces comfort deviation, ensuring both thermal stability and practical feasibility for real-time building management. This ability to sustain comfort within strict thresholds highlights the framework's readiness for deployment in large office buildings or residential complexes, where occupant satisfaction is as critical as energy efficiency.

#### E. Comparison with Baseline Rule-Based Controls

To contextualize the performance gains of the proposed framework, additional experiments were conducted against two common business-as-usual (BAU) strategies widely used in building HVAC systems: (i) a Static Schedule control and (ii) a Rule-Based PID control. Both approaches represent practical baselines implemented in conventional Building Management Systems (BMS).

**Static Schedule (BAU).** Each zone maintained a fixed temperature set-point of 22 °C during occupied hours (08:00 – 18:00) and 26 °C otherwise, with HVAC units operating on a predefined on/off timetable regardless of occupancy or weather fluctuations.

**Rule-Based PID Control.** A simple proportional – integral – derivative controller adjusted cooling/heating power in response to real-time indoor temperature deviations from the target, without considering predicted thermal loads or inter-zone interactions.

Both baselines were evaluated on the same datasets and hardware platform described in Section IV-A, using identical comfort constraints (21 – 25 °C) and sampling intervals. Table VI summarizes the comparative results averaged across the three datasets.

TABLE VI. COMPARISON WITH BUSINESS-AS-USUAL (RULE-BASED) CONTROL STRATEGIES

| Method                 | Energy Saving (%) | Comfort Index (0–1) | Remarks                                        |
|------------------------|-------------------|---------------------|------------------------------------------------|
| Static Schedule (BAU)  | 0.0               | 0.82                | Reference operation with fixed set-points      |
| Rule-Based PID Control | 8.9               | 0.79                | Reactive adjustments without prediction        |
| MADDPG (Best baseline) | 15.9              | 0.78                | Multi-agent learning without attention         |
| Proposed Framework     | 21.7 [18.9–23.1]  | 0.85 [0.82–0.87]    | Attention + Feature Selection + Cooperative RL |

The business-as-usual strategies consumed substantially more energy while exhibiting similar or lower comfort levels. The proposed method achieved a 21.7 % median energy saving — +12.8 percentage points relative to rule-based control and +21.7 points compared with static scheduling — while maintaining comfort within  $\pm 1$  °C of the target range. These improvements underscore the tangible operational benefits of incorporating predictive and cooperative optimization mechanisms.

Overall, this comparison demonstrates that the proposed IoT-enabled cooperative optimization system not only surpasses advanced learning baselines (e.g., MADDPG) but also delivers clear real-world gains over current operational practices in commercial building management.

#### F. Robustness

Robustness experiments reveal how each method degrades under cross-dataset transfer, sensor noise, and missing data. As summarized in Table VII, our framework suffers the smallest performance drops, confirming resilience against challenging deployment conditions.

TABLE VII. ROBUSTNESS EVALUATION (PERFORMANCE DROP IN ENERGY SAVING %)

| Condition                       | XGBoost+MADDPG | PIDT   | Proposed |
|---------------------------------|----------------|--------|----------|
| Cross-dataset transfer          | -11.2%         | -9.6%  | -4.5%    |
| Gaussian noise ( $\sigma=0.1$ ) | -14.8%         | -12.5% | -6.9%    |
| 20% missing data                | -18.3%         | -15.4% | -7.8%    |

The impact of Gaussian noise is further depicted in Fig. 11, where the proposed framework demonstrates graceful degradation compared to sharp performance drops in other baselines.

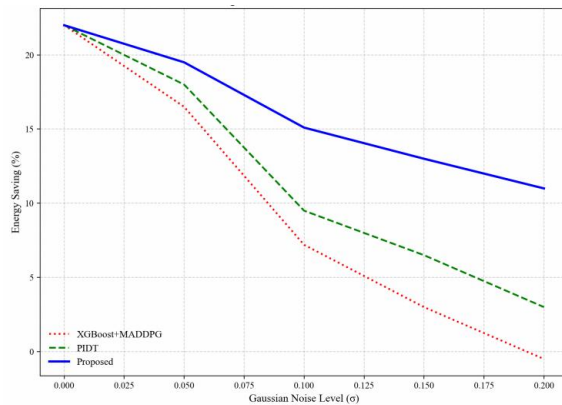


Fig. 11. Robustness under noise.

Fig. 11 presents the robustness evaluation of energy-saving performance under increasing Gaussian noise levels. The proposed framework demonstrates graceful degradation, maintaining more than 15% savings even at  $\sigma = 0.2$ . In contrast, PIDT exhibits moderate sensitivity, while XGBoost+MADDPG suffers sharp performance drops, falling below 5% savings at high noise levels. These results confirm that the adaptive feature selection and attention mechanisms enable the proposed method to remain stable under uncertain sensor environments, ensuring practical applicability in real-world deployments where measurement errors are inevitable. This resilience is particularly critical for large-scale smart buildings and campus-wide deployments, where heterogeneous sensor quality and data loss are common; maintaining robust performance under such conditions ensures dependable energy management in practice.

Gains stem primarily from attention's long-horizon capture (Campus-C peak hours:  $\Delta R^2 +0.33$ ) and MI-filtering that halves the variance under  $\sigma=0.2$  noise (drop  $-6.1$  pp vs.  $-12.5$  pp for baselines).

### G. Ablation Study

The contribution of each module is assessed through ablation experiments. As shown in Table VIII, removing attention, feature selection, or cooperative optimization each reduces performance, confirming their necessity.

TABLE VIII. ABLATION RESULTS (OFFICE-A DATASET)

| Variant    | MAE ↓ | Energy Saving (%) ↑ | Comfort Index ↑ |
|------------|-------|---------------------|-----------------|
| w/o Attn   | 1.93  | 18.5                | 0.80            |
| w/o FS     | 1.89  | 19.2                | 0.81            |
| w/o Coop   | 1.72  | 15.6                | 0.77            |
| Full Model | 1.65  | 21.7                | 0.85            |

Comparative ablation outcomes are also visualized in Fig. 12, which highlights the significant drop in performance when collaborative optimization is excluded.

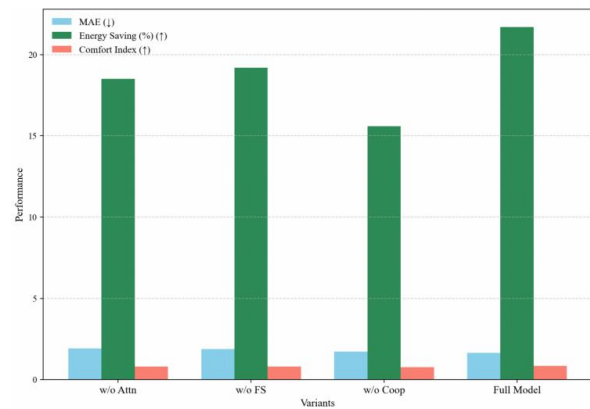


Fig. 12. Ablation comparison.

Fig. 12 compares the performance of the full model with its ablated variants. Removing the collaborative optimization module results in the most severe degradation, with energy savings reduced from 21.7% to 15.6% and comfort index dropping from 0.85 to 0.77. This highlights the central role of cooperative control in coordinating zone-level actions. Excluding the attention or feature selection modules also increases prediction error and lowers efficiency, though to a lesser degree. Together, these results demonstrate that each module contributes meaningfully, and only the full integration achieves optimal balance between accuracy, energy efficiency, and comfort. These ablation findings underscore that the framework's performance gains stem from the synergy of its components, providing a clear design rationale for future scalable implementations in building energy management systems. Please refer to Appendix for notations.

## V. DISCUSSION

The experimental results presented in this study offer several key insights into the effectiveness of the proposed IoT-enabled data-driven collaborative optimization framework for dynamic thermal load management in low-energy buildings. Across three datasets representing diverse building types and climates, the proposed method consistently achieved superior forecasting accuracy and energy-saving performance compared to classical, machine learning, and reinforcement learning baselines. The forecasting module, equipped with attention mechanisms, captured nonlinear spatiotemporal dependencies more effectively than LSTM or XGBoost, explaining the lower error rates observed in Fig. 9. Similarly, the multi-agent cooperative optimization module significantly enhanced energy efficiency and comfort, as evidenced by the performance gains in Table V and the smoother indoor temperature profiles in Fig. 10. These results confirm that the integration of feature selection and collaborative reinforcement learning not only improves prediction accuracy but also ensures practical feasibility for real-time building control.

Despite these encouraging outcomes, several limitations should be acknowledged. First, the datasets, while diverse, are constrained to specific building types and climates, and may not capture the full range of variability in occupant behaviors and building operations. Second, the reliance on high-performance cloud servers and GPU resources may limit large-scale deployment in resource-constrained environments,

although the inclusion of edge devices partly mitigates this issue. Third, the reinforcement learning framework requires substantial training time and may face stability challenges when scaled to hundreds of zones with heterogeneous dynamics. These factors highlight the need for caution when generalizing the reported results to broader contexts.

Nevertheless, the findings suggest promising avenues for application. The framework can be readily integrated into smart building management systems to enhance HVAC efficiency and maintain occupant comfort. Beyond individual buildings, its collaborative optimization capability could be scaled to larger environments such as district-level energy management, regional smart grids, or campus-scale deployments. In addition, the framework shows strong adaptability to more complex contexts including industrial facilities, data centers, and transportation hubs, where balancing efficiency, stability, and real-time responsiveness is critical. To realize such deployments, future implementations will need to address integration with existing building management systems, interoperability standards, and cost-effective edge-cloud coordination. Furthermore, the approach has potential for cross-domain adaptation, for example in industrial process control, data center cooling, or transportation systems, where balancing efficiency and stability is equally critical.

Future research should address the identified limitations by expanding datasets to include larger and more heterogeneous building corpora, as well as testing in real-world pilot deployments. Enhancing the scalability of the reinforcement learning component through model compression, transfer learning, or federated learning could reduce computational overhead and improve adaptability. Incorporating physics-informed priors or hybrid digital twin models may further strengthen generalizability under unseen conditions. Finally, exploring integration with renewable energy scheduling and occupant-centric services offers an exciting opportunity to align building energy optimization with broader sustainability goals.

## VI. CONCLUSION

This study proposed an IoT-enabled, data-driven collaborative optimization framework for dynamic thermal load management in low-energy buildings. By integrating mutual-information-guided feature selection, attention-enhanced forecasting, and multi-agent reinforcement learning under an edge-cloud architecture, the framework achieves median 21.7% energy savings (IQR 18.9–23.1%) across three heterogeneous datasets, corresponding to a +5.8 percentage point gain over MADDPG ( $p=0.004$ , Wilcoxon test). Comfort deviations are maintained within  $\pm 1^\circ\text{C}$  (median absolute deviation  $0.32^\circ\text{C}$ ). Robustness tests further show smaller performance degradation than baselines under  $\sigma \leq 0.2$  Gaussian noise and  $\leq 20\%$  missing data.

The methodological contribution lies in unifying forecasting, feature selection, and cooperative optimization within a single objective function, thereby reducing error propagation common in pipeline approaches. From an engineering perspective, feasibility was demonstrated in a

hardware-in-the-loop testbed operating under the specified compute and latency budget; real-building trials and broader climatic contexts remain as future work.

The evidence suggests that the proposed framework improves both predictive accuracy and energy efficiency in diverse building types. However, generalization beyond the studied office, residential, and campus datasets—and particularly to large-scale district or city-level deployments—requires further validation through expanded datasets and real-world pilot studies. Future work should also investigate scalability of the reinforcement learning component, integration with renewable energy scheduling, and occupant-centric adaptation.

## REFERENCES

- [1] V. A. Arowoia, A. O. Onososen, R. C. Moehler, and Y. Fang, “Influence of thermal comfort on energy consumption for building occupants: The current state of the art,” *Buildings*, vol. 14, no. 5, p. 1310, 2024.
- [2] F. Abdel-Jaber and K. N. Dirks, “A review of cooling and heating loads predictions of residential buildings using data-driven techniques,” *Buildings*, vol. 14, no. 3, p. 752, 2024.
- [3] I. Rojek et al., “Internet of Things applications for energy management in buildings using artificial intelligence—A case study,” *Energies*, vol. 18, no. 7, p. 1706, 2025.
- [4] M. A. Alam, A. R. Nabil, A. A. Mintoo, and A. Islam, “Real-time analytics in streaming big data: Techniques and applications,” *J. Sci. Eng. Res.*, vol. 1, no. 1, pp. 104–122, 2024.
- [5] C. Jacobsen, *Enhancing Physical Modeling with Interpretable Physics-Aware Machine Learning*, Ph.D. dissertation, 2024.
- [6] A. Li, *Development of a Data-Driven Methodology for Dynamic Energy Modeling of Large Buildings towards Mass Deployment of AI in Smart Buildings*, 2023.
- [7] M. Ehtsham et al., “AI-powered advanced technologies for a sustainable built environment: A systematic review on emerging challenges,” *Sustainability*, vol. 17, no. 17, p. 8005, 2025.
- [8] J. Hossain et al., “A review on optimal energy management in commercial buildings,” *Energies*, vol. 16, no. 4, p. 1609, 2023.
- [9] L. Zhang et al., “High spatial granularity residential heating load forecast based on Dendrite Net model,” *Energy*, vol. 269, p. 126787, 2023.
- [10] A. Prabowo et al., “Building timeseries dataset: Empowering large-scale building analytics,” in *Advances in Neural Information Processing Systems*, vol. 37, pp. 133180–133206, 2024.
- [11] S. M. Robeson and C. J. Willmott, “Decomposition of the mean absolute error (MAE) into systematic and unsystematic components,” *PLoS One*, vol. 18, no. 2, p. e0279774, 2023.
- [12] Y. Himeur et al., “AI-big data analytics for building automation and management systems: A survey, actual challenges and future perspectives,” *Artif. Intell. Rev.*, vol. 56, no. 6, pp. 4929–5021, 2023.
- [13] D. Zhuang et al., “Data-driven predictive control for smart HVAC system in IoT-integrated buildings with time-series forecasting and reinforcement learning,” *Appl. Energy*, vol. 338, p. 120936, 2023.
- [14] B. S. Alotaibi, “Context-aware smart energy management system: A reinforcement learning and IoT-based framework for enhancing energy efficiency and thermal comfort in sustainable buildings,” *Energy Build.*, p. 115804, 2025.
- [15] Z. Wang, F. Xiao, Y. Ran, Y. Li, and Y. Xu, “Scalable energy management approach of residential hybrid energy system using multi-agent deep reinforcement learning,” *Appl. Energy*, vol. 367, p. 123414, 2024.
- [16] X. Liu, Y. Wu, B. Liu, and H. Wu, “A multi-agent deep deterministic policy gradient method for multi-zone HVAC control,” in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, pp. 1–5, Jul. 2023.

- [17] Y. El Mghouchi and M. T. Udristoiu, "Thermal load predictions in low-energy buildings: A hybrid AI-based approach integrating integral feature selection and machine learning models," *Appl. Sci.*, vol. 15, no. 11, p. 6348, 2025.
- [18] Y. Peng and Y. Chen, "Integrative soft computing approaches for optimizing thermal energy performance in residential buildings," *PLoS One*, vol. 18, no. 9, p. e0290719, 2023.
- [19] Z. An, X. Ding, and W. Du, "Go beyond black-box policies: Rethinking the design of learning agent for interpretable and verifiable HVAC control," in *Proc. 61st ACM/IEEE Design Autom. Conf. (DAC)*, pp. 1–6, Jun. 2024.
- [20] W. Xue, N. Jia, and M. Zhao, "Multi-agent deep reinforcement learning-based HVAC control for multi-zone buildings considering zone-energy-allocation optimization," *Energy Build.*, vol. 329, p. 115241, 2025.
- [21] S. Guan et al., "Adaptive multi-agent HVAC control for thermal comfort using multi-agent PPO with population-based training," *SSRN Preprint 5399815*, 2025.
- [22] P. Michailidis, I. Michailidis, and E. Kosmatopoulos, "Reinforcement learning for optimizing renewable energy utilization in buildings: A review on applications and innovations," *Energies*, vol. 18, no. 7, p. 1724, 2025.
- [23] L. Chen et al., "Machine learning methods in weather and climate applications: A survey," *Appl. Sci.*, vol. 13, no. 21, p. 12019, 2023.
- [24] H. Wang, Y. Chen, J. Kang, Z. Ding, and H. Zhu, "An XGBoost-based predictive control strategy for HVAC systems in providing day-ahead demand response," *Build. Environ.*, vol. 238, p. 110350, 2023.
- [25] S. D. Nazemi and M. A. Jafari, "A multi-agent deep reinforcement learning framework for real-time energy management in smart building communities," *SSRN Preprint 5208655*, 2025.

APPENDIX: NOTATION

| Symbol                            | Definition                                                                                           | Unit / Range                              |
|-----------------------------------|------------------------------------------------------------------------------------------------------|-------------------------------------------|
| $B = B_1, B_2, \dots, B_N$        | Set of buildings                                                                                     | –                                         |
| $X_t \in \mathbb{R}^d$            | Sensor feature vector (temperature, humidity, CO <sub>2</sub> , occupancy, weather, control signals) | Various, often normalized (dimensionless) |
| $y_t$                             | Actual dynamic thermal load (ground truth)                                                           | kW or kWh                                 |
| $\hat{y}_t$                       | Predicted thermal load                                                                               | kW or kWh                                 |
| $L$                               | Look-back window length                                                                              | Number of time steps (min or h)           |
| $s_t^z$                           | State vector of zone $z$ at time $t$ (predicted load, temperature, occupancy, etc.)                  | –                                         |
| $a_t^z$                           | Control action of zone $z$ at time $t$ (HVAC power/airflow)                                          | kW, m <sup>3</sup> /h                     |
| $r_t^z$                           | Reward of zone $z$ at time $t$                                                                       | – (dimensionless)                         |
| $E_t^z$                           | Energy consumption of zone $z$ at time $t$                                                           | kWh                                       |
| $P_t^z$                           | Instantaneous HVAC power in zone $z$ at time $t$                                                     | kW                                        |
| $\Delta t$                        | Control interval duration                                                                            | min or h                                  |
| $C_t^z$                           | Comfort deviation (difference between actual and target temperature)                                 | °C                                        |
| $T_t^z$                           | Actual indoor temperature of zone $z$ at time $t$                                                    | °C                                        |
| $T_{\text{target}}^z$             | Target indoor temperature of zone $z$                                                                | °C                                        |
| $\alpha, \beta$                   | Weights for energy consumption and comfort deviation in reward function                              | Positive real numbers                     |
| $\gamma$                          | Discount factor                                                                                      | (0,1)                                     |
| $\lambda_1, \lambda_2, \lambda_3$ | Weights of forecasting, feature selection, and optimization terms in total loss                      | Positive real numbers                     |
| $I(X_j; Y)$                       | Mutual information between feature $X_j$ and target $Y$                                              | nats or bits                              |
| $\delta$                          | Threshold for mutual information-based feature selection                                             | –                                         |
| MAE                               | Mean Absolute Error                                                                                  | kW                                        |
| RMSE                              | Root Mean Square Error                                                                               | kW                                        |
| MAPE                              | Mean Absolute Percentage Error                                                                       | %                                         |
| $R^2$                             | Coefficient of determination                                                                         | [0,1]                                     |
| Energy Saving                     | Energy saving rate                                                                                   | %                                         |
| Comfort Index                     | Thermal comfort index                                                                                | [0,1]                                     |
| Cost Reduction                    | Cost Reduction Rate                                                                                  | %                                         |