Agentic Multi-Objective Retrofitting Framework (AMORF) for Buildings

Hindesh Akash [*1] **Yang Zhao** [*2]

Abstract

Retrofitting legacy buildings for energy efficiency is critical for sustainability, yet balancing economic feasibility, energy savings, and occupant comfort remains challenging. This research proposes a Phased Retrofitting System framework with Reinforcement Learning & Agentic AI characteristics, integrating economic assessment with adaptive HVAC control. Algorithmic developments decide which locations to fit and what to fit. The strategic agent decides on synchronization with existing systems while optimizing daily energy use, jointly minimizing energy consumption, carbon emissions, and costs while maintaining thermal comfort. Simulations show higher energy savings and minimal comfort violations compared to baselines. This approach offers a scalable solution for building portfolios, aligning with urban decarbonization goals.

1. Introduction

Buildings contribute 37% of U.S. carbon emissions and consume 40% of global energy, with HVAC systems using 40–50% of electricity in commercial settings (Yu et al., 2021). Yet, many legacy buildings still rely on outdated controls, making intelligent retrofitting a crucial but complex opportunity due to trade-offs in energy, comfort, cost, and system compatibility.

Recent advances in deep reinforcement learning (DRL) have shown promising results for HVAC control optimization, with studies demonstrating energy savings of 2-14% while maintaining thermal comfort (Nguyen et al., 2024). Multi-agent reinforcement learning approaches have further expanded these capabilities, enabling coordinated control of complex building systems through distributed decision-making agents (Yu et al., 2021; Wu et al., 2024). Advanced DRL methods have achieved even more substantial energy savings, with entropy-driven approaches showing up to 38.95% energy consumption reduction under high-temperature conditions (Zhang & Tan, 2025). However, existing research has primarily focused on optimizing control strategies for buildings with existing sensor infrastructure and modern automation systems, leaving a significant gap in addressing the unique challenges of legacy building retrofitting.

Our approach addresses three critical limitations in existing building optimization research. First, most current methods assume the availability of comprehensive sensor networks and modern control infrastructure, which is rarely the case in legacy buildings (Sierla et al., 2022). Second, existing approaches typically optimize for steady-state operations rather than the transitional challenges inherent in phased retrofitting. Third, current multi-agent building control systems focus on operational optimization rather than the strategic planning required for retrofitting decisions (Huang et al., 2025).

The contribution of this work is threefold: This work (1) proposes a zone selection algorithm for targeted retrofitting using sensor data, (2) enables coordinated phased upgrades through agent-to-agent communication, and (3) introduces an RL-based control framework for adaptive HVAC optimization in retrofitted zones.

2. Methodology

The Agentic Multi-Objective Retrofitting Framework (AMORF) consists of four interconnected components: synthetic building data generation, zone prioritization, multi-agent coordination, and reinforcement learning-based control. This section details each component's design and implementation.

2.1. Synthetic Building Environment

To evaluate AMORF systematically, we develop a synthetic legacy commercial building environment using Monte Carlo simulation. The building model encompasses 50 rooms across different categories (offices, conference rooms, com-

[*]Equal contribution [1]Atria University, Bengaluru, India [2]Nanyang Business School, Nanyang Technological University, Singapore. Correspondence to: Hindesh Akash <hindesh.a@atriauniversity.edu.in>, Yang Zhao <zhao0466@e.ntu.edu.sg>.

Proceedings of the 42^{nd} International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025.

mon areas) with heterogeneous characteristics.

Each room r_i is parameterized by:

- **Occupancy dynamics** $O_i(t)$: Stochastic hourly occupancy patterns following realistic business schedules with probabilistic variations
- **Thermal characteristics** $T_i(t)$: Temperature evolution governed by thermal mass, insulation properties, and HVAC capacity
- **HVAC specifications** H_i: Legacy system parameters including cooling capacity, energy efficiency ratio (EER), and operational constraints
- **Comfort requirements** C_i: Target temperature range (22–24°C) with tolerance thresholds

The environment incorporates external weather conditions $W(t) = \{T_{ext}(t), \phi(t), I_{solar}(t)\}$ representing outdoor temperature, relative humidity, and solar irradiance respectively. Energy consumption for room i at time t is modeled as:

$$\begin{aligned} E_i(t) = \alpha_i \cdot H_i \cdot \max(0, T_i(t) - T_{setpoint}) \\ + \beta_i \cdot O_i(t) + \gamma_i \cdot I_{solar}(t) \end{aligned} \tag{1}$$

where α_i, β_i, γ_i are room-specific coefficients capturing HVAC efficiency, occupancy heat gains, and solar heat gains.

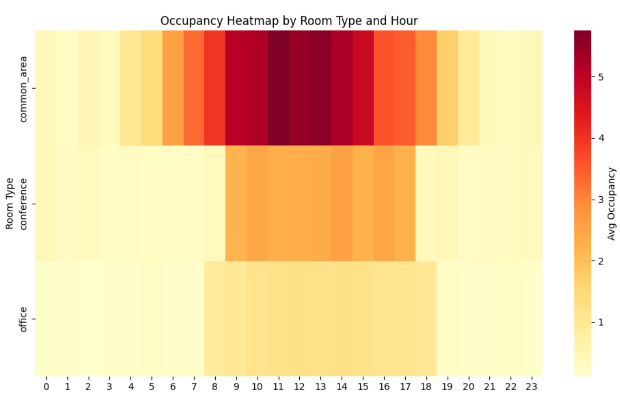

Figure 1. Occupancy patterns for different room types over a 30-day simulation period showing distinct usage characteristics across office spaces, conference rooms, and common areas.

2.2. Prioritization Algorithm for Retrofitting

AMORF employs a multi-criteria decision algorithm to identify high-impact zones for retrofitting. The algorithm evaluates rooms based on three weighted criteria: energy intensity, comfort violations, and occupancy utilization.

Algorithm 1 Zone Prioritization for Retrofit Selection

```
Input: Room dataset R = {r_1, r_2, ..., r_n}, target selection count N
Output: Prioritized zone list S for retrofitting
Initialize weights: w_e = 0.4 (energy), w_c = 0.3 (comfort), w_o = 0.3 (occupancy)
Initialize score dictionary D = {}
for each room r_i ∈ R do
    Compute energy intensity: E_i = (1/T) Σ_{t=1}^{T} E_i(t)/A_i {kWh/m²}
    Compute comfort violation rate: C_i = |{t : |T_i(t) − T_target| > δ}| / T
    Compute occupancy utilization: O_i = (1/T) Σ_{t=1}^{T} O_i(t)
end for
Normalize metrics: E'_i = (E_i − min(E)) / (max(E) − min(E) + ϵ)
Similarly normalize C'_i and O'_i
for each room r_i ∈ R do
    D[r_i] = w_e · E'_i + w_c · C'_i + w_o · O'_i
end for
Sort D in descending order
S = top-N rooms from sorted D
return S
```

2.3. Reinforcement Learning-Based Control

After finding the zones with high priority for the first phase of retrofitting. We employ Q-learning to optimize HVAC operations across selected zones assuming that the latest HVAC systems have been installed in the legacy zones. The agent learns policies that balance energy efficiency, occupant comfort, and operational costs through multi-objective reward design.

2.3.1. State and Action Spaces

The state space $\mathcal{S}$ for each room includes: $s_t = \langle T_{current}, T_{external}, O_{count}, C_{violations}, H_{status} \rangle$

The action space $\mathcal{A}$ consists of discrete HVAC control decisions: $\mathcal{A} = \{maintain, cool_{moderate}, cool_{aggressive}, heat_{moderate}, shutdown\}$

2.3.2. Multi-Objective Reward Function

The reward function balances three objectives:

$$\begin{aligned} r(s_t, a_t) = &-\lambda_1 \cdot \text{Energy\_Cost}(a_t) \\ &-\lambda_2 \cdot \text{Comfort\_Penalty}(s_t, a_t) \\ &+\lambda_3 \cdot \text{Efficiency\_Bonus}(s_t, a_t) \end{aligned} \tag{2}$$

Algorithm 2 Q-Learning HVAC Control Strategy

```
Input: Q-table Q, learning rate α = 0.1, discount factor γ = 0.95
Initialize: Exploration rate ϵ = 0.1, episode counter ep = 0
for each episode do
    Reset environment to initial state s_0
    for each time step t do
        Observe current state s_t
        if random() < ϵ then
            Select random action a_t ∼ Uniform(A)
        else
            Select greedy action a_t = arg max_a Q(s_t, a)
        end if
        Execute action a_t, observe reward r_t and next state s_{t+1}
        Update Q-value:
        Q(s_t, a_t) ← Q(s_t, a_t) + α[r_t + γ max_a Q(s_{t+1}, a) − Q(s_t, a_t)]
        Communicate action and state to coordination framework
        s_t ← s_{t+1}
    end for
    ep ← ep + 1
end for
```

where $\lambda_1 = 0.4$, $\lambda_2 = 0.4$, $\lambda_3 = 0.2$ weight the competing objectives.

2.3.3. Agentic Coordination Framework

The AMORF framework employs a hybrid architecture integrating distributed RL agents with a centralized agentic AI coordinator for optimal control and stakeholder communication.

- **Distributed RL Agents**: Q-learning agents deployed at each HVAC system within prioritized zones. Each agent learns optimal control policies by observing local conditions and executing actions while maintaining zone-specific Q-tables.
- **Central Orchestrating Agent**: Agentic AI system interfacing with RL agents and stakeholders. Leverages LLM capabilities to synthesize optimization data into contextual insights and strategic decisions for energy management.
- **Stakeholder Interface**: Automated natural language communication generating reports, alerts, and recommendations from aggregated RL performance data for facility managers and operators.

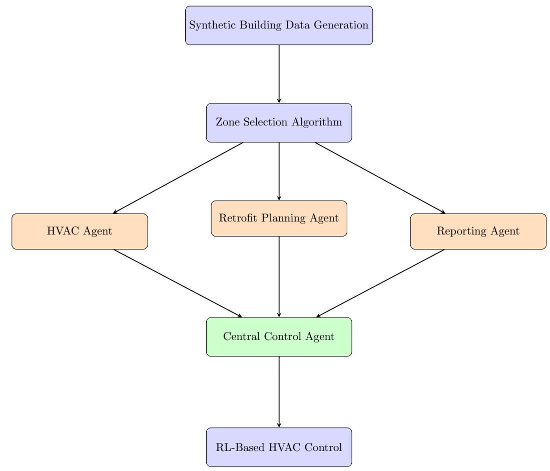

Figure 2. Multi-agent coordination workflow.

Table 1. Top 5 office rooms identified by AMORF prioritization algorithm, ranked by composite retrofit priority scores.

Room ID	Room Type	Score
room_15	office	0.811
room_33	office	0.793
room_9	office	0.751
room_5	office	0.735
room_3	office	0.731

3. Experimental Validation

We validate AMORF through simulation experiments that demonstrate the complete framework from data generation to stakeholder communication.

3.1. Experimental Setup

We generate synthetic data for a 50-room commercial building using Monte Carlo simulation over 30-day periods. The AMORF prioritization algorithm identifies top-N zones for retrofitting based on energy intensity, comfort violations, and occupancy patterns. Multi-agent coordination manages retrofit implementation while Q-learning optimizes HVAC control policies. Table 1 shows the top 10 office rooms identified by our prioritization algorithm, ranked by their composite retrofit priority scores.

Performance is evaluated using four key metrics: (1) energy efficiency (kWh/day), (2) thermal comfort violations (% time outside 22–24°C), (3) operational cost reduction ($/day), and (4) carbon emissions (kg CO/day). We compare AMORF against baseline static control and rule-based heuristics.

Algorithm 3 Integrated Agent Coordination and Stakeholder Communication

Input: Agent states Θ, sensor readings $\mathcal{S}(t)$, MCP server $\mathcal{M}$, stakeholder profiles $\mathcal{SP}$
Initialize: Message queue $\mathcal{Q}$, LLM context $\mathcal{C}$, report templates $\mathcal{T}$
while system operational **do**
  for each agent A_j **do**
    Update local state: $\theta_j \leftarrow f_j(\theta_j, \mathcal{S}(t), \mathcal{Q})$
    if significant state change OR periodic sync required **then**
      Extract context: $ctx_j \leftarrow \mathcal{M}.extract(\theta_j, actions_j)$
      Transmit MCP message to orchestrator
    end if
  end for
  Aggregate contexts: $\mathcal{C} \leftarrow \bigcup_j ctx_j$
  Generate insights: $insights \leftarrow LLM.reason(\mathcal{C}, objectives)$
  Update coordination strategy and broadcast to agents
  for each stakeholder $sp \in \mathcal{SP}$ **do**
    Filter insights: $insights_{sp} \leftarrow$ filter_by_role$(insights, sp)$
    Generate report: $report \leftarrow LLM.generate(\mathcal{T}, insights_{sp})$
    Transmit via appropriate channel based on urgency
  end for
  Clear processed contexts from $\mathcal{Q}$
end while

3.2. Natural Language Communication Validation

AMORF autonomously generates actionable notifications—alerts, warnings, and recommendations—demonstrating effective agentic communication. This confirms its ability to translate optimization outcomes into human-readable insights for stakeholders.

4. Results & Conclusion

We validated the AMORF framework through simulations, with room_15 emerging as the top retrofit candidate (score: 0.811). After 800 Q-learning episodes, the agent achieved a high average reward of 3458.52, confirming successful multi-objective optimization. The system maintained thermal comfort (mean temperature: 24.1°C), reduced energy cost (215.99 units), and minimized comfort violations (197.1 instances), with convergence indicated by a low exploration rate ($\epsilon = 0.081$).

These results highlight AMORF's ability to integrate zone prioritization, agent-based coordination, and adaptive RL control to optimize building energy performance. The framework effectively balances comfort, cost, and efficiency, offering a scalable path toward data-driven retrofitting and urban decarbonization.

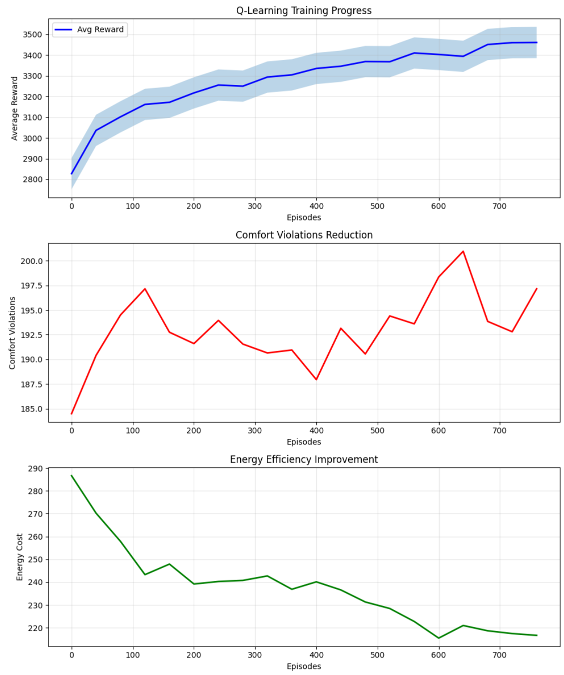

Figure 3. AMORF results for Room:15

Disclosure on Use of AI

AI tools were used solely to improve the language and readability of this manuscript. All research design, analysis, methods, and conclusions were developed and validated solely by the authors.

References

Chen, Z., Lin, Y., Wei, H., Zhao, Y., and Zhang, Y. Smart building retrofit planning via deep multi-agent reinforcement learning. *Journal of Building Performance*, 13(1): 45–60, 2022. doi: 10.1080/19401493.2021.1983081.

Ding, Y., Jin, X., Tian, Z., Heidarinejad, M., Dahlhausen, M., and Sanguinetti, A. Transfer reinforcement learning for hvac systems: A review and future perspectives. *Energy and AI*, 13:100218, 2023. doi: 10.1016/j.egyai.2023.100218.

Goldfeder, J., Dean, V., Jiang, Z., Wang, X., Dong, B., Lipson, H., and Sipple, J. The smart buildings control suite: A diverse open source benchmark to evaluate and scale hvac control policies for sustainability, 2024.

Huang, C., Ge, Z., Wang, L., Zhang, D., Long, H., and Luo, X. Multi-agent deep reinforcement learning-based cooperative optimal operation for building integrated energy systems with nash–harsanyi bargaining game-driven cost allocation. *IEEE Transactions on Consumer Electronics*, 2025. doi: 10.1109/TCE.2025.3541137.

Li, Q., Meng, Q., and Zhou, X. A review of artificial intelligence algorithms for energy management in smart buildings. *Renewable and Sustainable Energy Reviews*, 177:113163, 2023. doi: 10.1016/j.rser.2022.113163.

Mocanu, E., Nguyen, P. H., Gibescu, M., and Slootweg, H. On-line building energy optimization using deep reinforcement learning. *IEEE Transactions on Smart Grid*, 10 (4):3698–3708, 2019. doi: 10.1109/TSG.2018.2834219.

Nguyen, A. T., Pham, D. H., Oo, B. L., Santamouris, M., Ahn, Y., and Lim, B. T. Modelling building hvac control strategies using a deep reinforcement learning approach. *Energy and Buildings*, 310:114065, 2024. doi: 10.1016/j.enbuild.2024.114065.

Sierla, S., Ihasalo, H., and Vyatkin, V. A review of reinforcement learning applications to control of heating, ventilation and air conditioning systems. *Energies*, 15: 3526, 2022. doi: 10.3390/en15103526.

Wang, K., Zhang, B., and Li, Y. Hierarchical deep reinforcement learning for building hvac control with multi-zone coordination. *Energy and Buildings*, 284:112795, 2023. doi: 10.1016/j.enbuild.2023.112795.

Wu, H., Qiu, D., Zhang, L., and Sun, M. Adaptive multi-agent reinforcement learning for flexible resource management in a virtual power plant with dynamic participating multi-energy buildings. *Applied Energy*, 374:123998, 2024. doi: 10.1016/j.apenergy.2024.123998.

Yu, L. et al. Multi-agent deep reinforcement learning for hvac control in commercial buildings. *IEEE Transactions on Smart Grid*, 12(1):407–419, 2021. doi: 10.1109/TSG.2020.3011739.

Zhang, C. and Tan, Z. Entropy-driven deep reinforcement learning for hvac system optimization. *Journal of Renewable and Sustainable Energy*, 17(1):015101, 2025. doi: 10.1063/5.0238799.

Zhang, W., Yu, Y., Yuan, Z., Tang, P., and Gao, B. Data-driven pre-training framework for reinforcement learning of air-source heat pump (ashp) systems based on historical data in office buildings: Field validation. *Energy and Buildings*, 332:115436, 2025. doi: 10.1016/j.enbuild.2025.115436.