Data-Centric Engineering (2025), 6: e40
doi:10.1017/dce.2025.10014

CAMBRIDGE UNIVERSITY PRESS

RESEARCH ARTICLE

Optimizing HVAC energy efficiency in low-energy buildings: a comparative analysis of reinforcement learning control strategies under Tehran climate conditions

Mohammad Anvar Adibhesami[1] and Amir Hassanzadeh[2]

[1]School of Architecture and Environmental Design, Iran University of Science and Technology, Tehran, Iran
[2]Department of Mechanical Engineering, Urmia University, Urmia, Iran
Corresponding author: Amir Hassanzadeh; Email: st_a.hassanzadeh@urmia.ac.ir

Received: 01 January 2025; **Revised:** 24 April 2025; **Accepted:** 12 June 2025

Keywords: advanced HVAC systems; energy efficiency; low-energy buildings; reinforcement learning control strategies; Tehran climate; urban sustainability

Abstract

This study investigates the incorporation of advanced heating, ventilation, and air conditioning (HVAC) systems with reinforcement learning (RL) control to enhance energy efficiency in low-energy buildings amid the extreme seasonal temperatures of Tehran. We conducted comprehensive simulation assessments using the EnergyPlus and HoneybeeGym platforms to evaluate two distinct reinforcement learning models: traditional Q-learning (Model A) and deep reinforcement learning (DRL) with neural networks (Model B). Model B consisted of a deep convolutional network architecture with 256 neurons in each hidden layer, employing rectified linear units as activation functions and the Adam optimizer at a learning rate of 0.001. The results demonstrated that the RL-managed systems resulted in a statistically significant reduction in energy-use intensity of 25 percent ($p < 0.001$), decreasing from 250 to 200 kWh/m² annually in comparison to the baseline scenario. The thermal comfort showed notable improvements, with the expected mean vote adjusting to 0.25, which falls within the ASHRAE Standard 55 comfort range, and the percentage of anticipated dissatisfaction reduced to 10%. Model B (DRL) demonstrated a 50 percent improvement in prediction accuracy over Model A, with a mean absolute error of 0.579366 compared to 1.140008 and a root mean square error of 0.689770 versus 1.408069. This indicates enhanced adaptability to consistent daily trends and irregular periodicities, such as weather patterns. The proposed reinforcement learning method achieved energy savings of 10–15 percent compared to both rule-based and model predictive control and approximately 10 percent improvement over rule-based control, while employing fewer building features than existing state-of-the-art control systems.

Impact statement

The present study provides recommendations for sustainable building design and operation in Tehran. Deep Q-network algorithms were used in this research, which show high precision in temperature control predictions. We showed that advanced heating, ventilation, and air conditioning (HVAC) technologies significantly improve thermal comfort within ASHRAE standards, and integrated HVAC systems with reinforcement learning reduce energy use intensity by up to 25%. Finally, one can conclude from this research that variable refrigerant flow systems and renewable energy integration enhance efficiency.

1. Introduction

Enhancing the energy efficiency of heating, ventilation, and air conditioning (HVAC) systems in buildings is crucial for promoting sustainability in regions with extreme climate conditions, such as Tehran, Iran. Building HVAC systems account for over 40% of global energy consumption in the built environment (Ürge-Vorsatz et al., 2015; Mehrpooya et al., 2019; Etemad et al., 2022). With heightened awareness of sustainability issues and ambitious net-zero targets, enhancing HVAC efficiency has become imperative (Gottschamer and Zhang, 2020). Modern low-energy buildings aim to minimize energy demand through passive design strategies and efficient technologies (Cao et al., 2016; Said et al., 2023). However, optimal control and operation of HVAC equipment play a pivotal role in realizing the energy-saving potential (Nassif et al., 2005; Dikshit et al., 2024).

Recent advancements in renewable systems, innovative air distribution techniques, and data-driven control methods offer new pathways for enhancing HVAC performance. Solar-assisted heat pumps can utilize thermal energy from the sun to reduce electricity usage (Dikshit et al., 2024; Wang et al., n.d.). Underfloor air distribution (UFAD) and displacement ventilation can minimize fan energy requirements and improve thermal comfort (Li et al., 2014). While reinforcement learning (RL) algorithms have demonstrated significant potential for dynamically optimizing HVAC setpoints based on real-time conditions (Vázquez-Canteli et al., 2017), their practical implementation in low-energy buildings, particularly within specific climatic contexts such as that of Tehran, remains underexplored.

This research focuses on the integration of these technologies to maximize energy efficiency in low-energy buildings, specifically within the unique climate context of Tehran, Iran. The temperate climate of Tehran presents its own set of challenges and opportunities, particularly regarding energy consumption patterns dominated by both heating and cooling loads. Performance evaluations tailored to Tehran's weather and occupancy profiles will quantify the efficiency improvements gained from implementing cutting-edge HVAC solutions.

Despite significant advances in both HVAC technology and control systems individually, a critical research gap exists regarding their integrated implementation and performance evaluation in real-world contexts, particularly in regions with unique climatic conditions such as Tehran. This research aims to fill these gaps by employing an applied methodology that incorporates simulations, experimental data, and comparative analyses. The findings will equip stakeholders within the Tehran building sector to make informed, data-driven decisions regarding the design and operation of energy-efficient and cost-effective HVAC systems.

1.1. Background

The Tehran region of Iran features a temperate climate that is conducive to employing passive heating and cooling techniques in building design (Sharif et al., 2022). However, increasing development and rising electricity demand have resulted in a growing carbon footprint for the local building sector (Farhadi et al., 2019). Low-energy buildings aim to address this challenge through optimized insulation, shading, ventilation, and a variety of other efficiency measures. Nonetheless, heating and cooling loads continue to represent over 60% of operational energy consumption (Rodrigues et al., 2023).

Recent electricity supply shortages and extreme weather patterns in Iran have underscored the urgent need for resilient and sustainable HVAC solutions (Climate Change and Energy Crisis in Iran—Iran News Update, n.d.). Integrating renewable energy sources and thermal technologies can help reduce grid dependence (Lv, 2023). Moreover, innovative air distribution systems enhance indoor climate control while minimizing fan energy usage (Rathnayaka et al., 2023). Most critically, intelligent control of HVAC systems can optimize energy performance in response to varying weather conditions and occupancy patterns (Jamali et al., 2023).

RL has demonstrated significant promise for data-driven HVAC optimization in simulation environments (Maddalena et al., 2022), where algorithms dynamically adjust setpoints to minimize energy consumption based on surrounding conditions (Kannari et al., 2023; Biswas et al., 2024). The key advantage of RL approaches over conventional control strategies lies in their ability to adapt and learn

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

optimal policies through interaction with the environment, without requiring detailed physical models of the building or HVAC system. However, adopting these techniques in real-world systems necessitates the careful selection of suitable mechanical equipment and instrumentation (Marin et al., 2016; Sahebzadeh et al., 2017; Al Sayed et al., 2024; An et al., 2024a, 2024b; Ding et al., 2024). This research specifically evaluates the applicability of RL for low-energy buildings within the Tehran context, addressing a gap in existing literature.

Recent advancements in deep learning techniques, such as deep clustering RL (DCRL), hold potential for overcoming these challenges (Sarker, 2021; Wu et al., 2023). DCRL can utilize deep neural networks to extract meaningful features from complex datasets, thereby enhancing sample efficiency and generalization (Bandi et al., 2023). This capability enables RL agents to quickly develop accurate models of their environments, facilitating the creation of efficient control policies tailored to specific building attributes (Villaizán-Vallelado et al., 2024). While promising in theory, these advanced techniques require empirical validation in real-world building environments, particularly under varying climate conditions, such as those found in Tehran. However, effectively implementing these techniques requires thoughtful selection of appropriate mechanical equipment, sensors, and deep learning architectures (Borisov et al., 2024).

The findings of this research are aimed at architects, system designers, and building operators in the region. The study will elucidate pathways for maximizing HVAC system efficiency through careful technology selection and operational optimization. Broader implications include informing sustainable building practices, energy codes, and carbon-neutral policies across Iran's building sector.

1.2. Objectives

This research seeks to address the energy efficiency challenges in low-energy buildings in Tehran by investigating the integration of cutting-edge HVAC technologies and RL control strategies, which include the following:

- Identification of advanced HVAC technologies: Evaluate and select advanced HVAC technologies suitable for the unique environmental conditions of Tehran, including solar heat pump variable refrigerant flow (VRF) systems, advanced heat pump systems, energy recovery ventilation (ERV), demand-controlled ventilation (DCV), and innovative air distribution techniques such as UFAD.
- Integration with RL: Develop a framework for integrating the identified HVAC technologies with RL control strategies. RL algorithms, including proximal policy optimization (PPO) and Deep Q-networks (DQNs), will be employed to create intelligent control systems capable of dynamically optimizing HVAC operation based on real-time conditions, occupant behavior, and energy demand profiles.

1.3. Rationale for integration

The integration of advanced HVAC technologies with RL is driven by the potential synergy between energy-efficient hardware and adaptive control strategies. Figure 1 provides a conceptual representation of the integration framework. In Figure 1, the interaction between the control agent and the environment showcases the adaptability of the RL system to optimize HVAC control actions in response to varying conditions.

The novelty of this research lies in the systematic investigation of this synergistic relationship, specifically focusing on how RL algorithms can enhance the operational efficiency of advanced HVAC systems beyond what either technology could achieve independently. By addressing this integration gap, this study contributes to the growing body of knowledge on intelligent building systems while providing practical insights for implementation in the Tehran context.

1.4. Significance of the study

This study contributes to the field by providing a comprehensive examination of advanced HVAC technologies integrated with RL control strategies. Unlike previous research that has typically focused

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

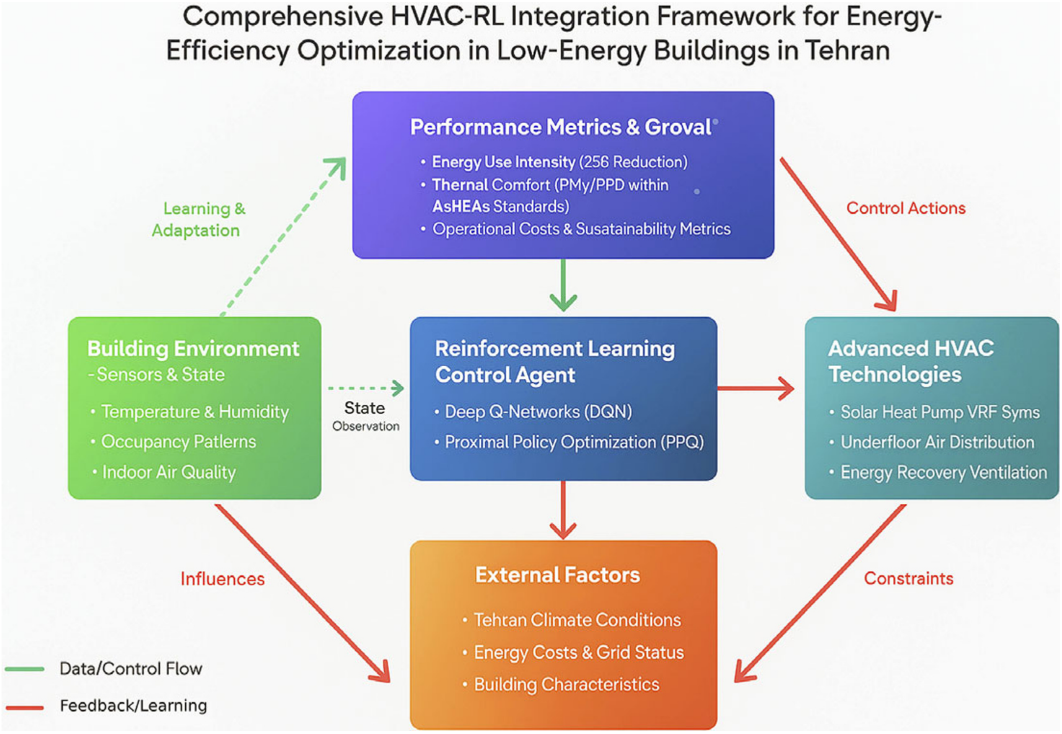


Figure 1. *Conceptual integration framework for HVAC energy efficiency study.*

on either technological advancements or control optimization in isolation, this study explores their synergistic integration with particular attention to regional climatic conditions. The significance of this research lies in its potential to inform sustainable building practices and guide HVAC designers, engineers, and building operators in enhancing energy efficiency in low-energy buildings. The novelty of this work stems from its holistic approach, which investigates the synergistic integration of cutting-edge HVAC systems, such as solar-assisted VRF systems, ERV, and innovative air distribution techniques, with advanced RL algorithms like DQN and PPO. This integration of state-of-the-art hardware and adaptive control strategies represents a significant contribution to the state of the art in the field, as it addresses the limitations of traditional HVAC systems and control methods. The findings of this study are expected to offer practical insights into optimizing energy performance in low-energy buildings in Tehran, which can lead to substantial energy savings and improved occupant comfort within ASHRAE comfort standards.

1.5. Structure of the article

The remainder of this article is organized as follows: Section 2 reviews the existing literature on energy-efficient technologies, advanced HVAC systems, and RL in building control. Section 3 details the research methodology, including the identification of advanced HVAC technologies, integration with RL, and the selection and implementation of RL algorithms. Subsequent sections present results, discussions, and comparative analyses, leading to the conclusions and recommendations in Section 6.

2. Literature review

2.1. Advanced HVAC technologies for energy efficiency

Energy-efficient buildings represent a critical approach to addressing global energy consumption and climate change challenges (Cabeza and Chàfer, 2020; Li and Dong, 2022; Yang et al., 2025). These

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

structures optimize resource utilization throughout their life cycle while providing comfortable living conditions and reducing operational costs (Ugli, 2022; Rebelatto et al., 2023; Polesello and Johnson, 2016). The building sector, particularly HVAC systems, accounts for over 40% of global energy consumption (Ürge-Vorsatz et al., 2015; Mehrpooya et al., 2019; Etemad et al., 2022), making their optimization essential for achieving sustainability goals.

Recent advancements in HVAC technologies have created new opportunities for energy conservation in buildings. Solar-assisted heat pumps utilize thermal energy from the sun to supplement conventional systems, reducing electricity demand while maintaining or improving performance (Wang et al., 2009; Dikshit et al., 2024). Studies by Wang et al. (2024) demonstrate that these integrated systems can achieve up to 30% higher coefficient of performance (COP) compared to conventional systems, particularly in regions with abundant solar resources, such as Tehran.

Innovative air distribution techniques, including UFAD and displacement ventilation, have emerged as effective methods for improving both energy efficiency and occupant comfort (Li et al., 2014). These systems deliver conditioned air at lower velocities and higher temperatures than conventional overhead systems, reducing fan energy requirements and improving thermal stratification. Research by Dikshit et al. (2024) indicates that such systems can reduce cooling energy consumption by 15–20% in commercial buildings while enhancing indoor air quality through more effective ventilation.

ERV and DCV represent another significant advancement in HVAC technology. ERV systems recover heat and moisture from exhaust air, reducing the energy required to condition incoming fresh air (van Roosmalen et al., 2021). Meanwhile, DCV systems modulate ventilation rates based on occupancy or air quality measurements, preventing energy waste from overventilation during periods of low occupancy (Bie et al., 2025). The integration of these technologies is particularly relevant for Tehran's climate, which experiences both hot summers and cold winters, requiring year-round climate control.

2.2. Control strategies for HVAC systems

Traditional HVAC control strategies typically rely on rule-based approaches with predefined setpoints and schedules. While these methods are straightforward to implement, they often fail to adapt to changing conditions or optimize energy use (Choi et al., 2023). Nassif et al. (2005) demonstrated that even optimized rule-based controls (RBCs) have inherent limitations in responding to dynamic environmental and occupancy conditions.

More advanced control methodologies have emerged to address these limitations. Model predictive control (MPC) uses mathematical models of building thermal dynamics to predict future conditions and optimize control decisions accordingly (Drgoňa et al., 2020; Xiao and You, 2023). This approach has shown promise in reducing energy consumption by anticipating changes in weather, occupancy, or utility rates. However, MPC implementations require accurate building models, which can be challenging to develop and maintain, particularly for complex or older buildings.

Lu et al. (2023) developed high-performance rule-based sequences for variable air volume systems that demonstrate improved efficiency over conventional control strategies. However, such rule-based approaches still lack the adaptability and learning capabilities needed for optimal performance across varying conditions. This limitation highlights the need for more advanced, data-driven control methodologies that can learn and adapt to building-specific characteristics without requiring explicit physical modeling.

2.3. RL for HVAC control

RL has emerged as a promising approach for optimizing HVAC control without requiring detailed physical models (Al Sayed et al., 2024; Silvestri et al., 2024; Xu et al., 2025). Unlike conventional control strategies, RL algorithms learn optimal control policies through interaction with the environment, adapting to building-specific characteristics and changing conditions over time.

Bilous et al. (2024) demonstrated that RL controllers can significantly outperform conventional approaches in minimizing both energy consumption and thermal discomfort, particularly after extended learning periods. Their study, focusing on single-zone air supply systems, showed energy savings of up to about 20% compared to RBCs while maintaining or improving occupant comfort.

Silvestri et al. (2024) developed a framework for whole-building energy modeling using deep RL DRL that demonstrated robust performance across varying conditions. Their approach integrated building simulation with DRL algorithms to create control policies that balanced energy efficiency with occupant comfort. However, one limitation noted in their study was the extensive data requirements and computational resources needed for training effective RL models.

A significant challenge in practical RL implementation for HVAC control is sample efficiency—the ability to learn effective policies with limited real-world data. Loffredo et al. (2023) addressed this issue by proposing model-based RL as an alternative to traditional model-free approaches. Their method incorporated approximate physical models to accelerate learning, demonstrating improved performance with substantially less training data.

Recent advancements in deep learning techniques have further enhanced the capabilities of RL for HVAC control. Fu et al. (2023) developed a multi-agent DRL approach that optimized control across multiple HVAC subsystems simultaneously. Their method demonstrated superior performance compared to single-agent approaches, achieving greater energy savings while maintaining comfort conditions. Similarly, Yan and Qin (2017) explored distributed RL for regional building energy optimization, showing the potential for coordinated control across multiple buildings.

2.4. Integration of advanced HVAC technologies with RL

2.4.1. Research gap

Despite significant advances in both HVAC technology and control systems individually, a critical gap exists in research addressing their integrated implementation and performance evaluation in real-world contexts, particularly in regions with unique climatic conditions, such as Tehran.

While numerous studies have explored either advanced HVAC technologies or RL control strategies in isolation, few have investigated the synergistic benefits of their integration. The potential advantages of such integration are substantial—advanced HVAC hardware provides greater flexibility and efficiency, while RL control strategies optimize operation based on real-time conditions and learned patterns.

The limited research on integrated approaches has typically focused on simulation environments rather than real-world implementations. Zhang et al. (2025) developed a comprehensive simulation framework for RL-controlled HVAC systems, but acknowledged limitations in transferring simulated policies to real buildings. Similarly, Al Sayed et al. (2024) demonstrated promising results for multi-agent RL control of advanced HVAC systems in simulation, but did not address practical implementation challenges.

Furthermore, existing research has rarely considered regional climate contexts, particularly for areas with distinct seasonal variations, such as Tehran. Ghiai et al. (2021) investigated energy consumption patterns in tall office buildings in Iran's hot-arid and cold climate conditions, highlighting the need for climate-specific approaches to energy optimization. However, their study did not explore advanced control strategies, such as RL.

This integration gap is particularly significant for regions, such as Tehran, where the unique climate conditions—characterized by hot, dry summers, and cold winters—present both challenges and opportunities for energy-efficient building operation (Karimi et al., 2024). The successful integration of advanced HVAC technologies with RL control strategies in this context could provide valuable insights for similar climate regions globally.

2.5. Health and well-being considerations

Energy efficiency in buildings extends beyond mere energy savings to impact occupant health and well-being. Wallner et al. (2017) found that residents of energy-efficient homes with mechanical ventilation reported significantly higher indoor air quality and better health outcomes compared to those in

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

conventional buildings. Similarly, Symonds et al. (2021) demonstrated that energy efficiency measures in homes provide co-benefits to occupant health, provided that adequate ventilation is maintained.

However, potential risks must also be considered. Carpino et al. (2023) identified increased risk of fungal growth in nearly zero-energy buildings due to airtight construction, highlighting the importance of balanced ventilation strategies. These health considerations underscore the need for holistic approaches to building energy efficiency that address both environmental and human factors.

2.6. Summary of research gaps and study objectives

Table 1 presents a comprehensive literature review matrix that highlights key findings and research gaps across the topics relevant to this study. The matrix demonstrates several critical gaps that this research aims to address.

This study addresses these gaps by investigating the integration of advanced HVAC technologies with RL control strategies specifically tailored to Tehran's climate conditions. By evaluating real-world performance, identifying implementation challenges, and quantifying both energy savings and comfort improvements, this research aims to provide practical insights for sustainable building design and operation in similar climate regions.

3. Methodology

This study investigates the integration of RL algorithms with HVAC systems to enhance energy efficiency and occupant comfort in low-energy buildings under Tehran climate conditions. Comprehensive building

Table 1. *Literature review matrix*

Topic	Key findings	Research gaps
Advanced HVAC technologies	- Higher COPs than conventional systems (Zhang et al., 2016) - Improved part-load efficiencies (Kwon et al., 2014) - Enhanced moisture removal capabilities (Yan et al., 2025) - A multifactorial analysis to select the most proper AC system technology according to the building and location (Balbis-Morejón et al., 2023) - Radar technology that has been identified as an emerging one (Cardillo et al., 2021) - Efficacy of the proposed standalone contributions, and as a whole, represents a suitable solution for helping to increase the performance of heating, ventilating, and air conditioning installations without affecting the comfort of their occupants (Cardillo et al., 2021)	- Limited real-world pilot studies - Uncertainties in performance advantages - Select the air conditioning (AC) system with better results in its life cycle for a typical building - Test new technologies and methodologies for improving HVAC energy efficiency - Efficacy of the energy management systems in dealing with energy consumption in buildings
Innovative air distribution	– 10–15% ventilation energy savings (Seppänen, 2008) - Increased user thermal comfort (López-Pérez and Flores-Prieto, 2023)	- Complex simulation models - Scarce evidence under tropical conditions - Indoor air quality in the COVID–19 era and beyond

Continued

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

Table 1. *Continued*

Topic	Key findings	Research gaps
	- An innovative air circulation concept supported by the use of UVGI in combination with a nanoporous air filter is recommended to combat the spread of SARS-CoV–2 and other harmful microbes in closed spaces (Sodiq et al., 2021) - The potential of advanced air distribution, and individually controlled macro-environment in general, for achieving shared values, that is, improved health, comfort (Anand et al., 2019), performance, energy saving, reduction of healthcare costs, and improved well-being, is demonstrated (Melikov, 2016)	- Reducing the energy used for heating, cooling, ventilating, and air conditioning of buildings
Renewable integration	- Solar VRFs can achieve a 50% renewable fraction (Gilani et al., 2021) - Reduced grid dependence and emissions (Cirone et al., 2022) - Decentralized ventilation combined with thermo-electric elements or heat pumps further shows potential for self-sufficient curtain wall-integrated HVAC (van Roosmalen et al., 2021) - Mathematical models are developed with TRNSYS17 to simulate the HVAC system within a typical small commercial building with various cooling options (Keleher and Narayanan, 2019)	Impact of intermittency - Optimal control strategies - HVAC systems to be adaptive to changing climate and occupancy scenarios, and supplied with locally generated renewable energy - Necessity of developing an eco-friendly air-conditioning system
Reinforcement learning control	– 5–15% energy savings demonstrated (Qin et al., 2022) - Improved adaptability to disturbances (Lu, 2022) - BEM-DRL achieves 16.7% heating demand reduction with more than 95% probability compared to the old rule-based control (Zhang et al., 2019) - Results show that the energy-saving performance of the proposed MA-CWSC method is significantly better than the rule-based control method (11.1% improvement), and is very close to that of the model-based control method (only 0.5% difference) (Fu et al., 2022)	- Limited testing for HVAC equipment - Scarce multi-objective implementations - High-order nature and slow computational speed limit its practical application in real-time HVAC optimal control - Highly dependent on the accuracy of the model, a large amount of historical data, and the deployment of different sensors

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

energy simulations were conducted using EnergyPlus v9.5 coupled with Python-based RL frameworks to test the hypothesis that RL control can reduce energy use intensity (EUI), improve thermal comfort within ASHRAE standards, and lower operational costs compared to conventional control strategies. The methodology involved collecting historical meteorological data for Tehran (2015–2023), modeling typical occupancy patterns in commercial buildings, and establishing simulation environments that reflect realistic scenarios with Tehran-specific building materials and construction standards. Performance metrics included EUI, thermal comfort indices (predicted mean vote [PMV]/predicted percentage of dissatisfied [PPD]), prediction accuracy (mean absolute error [MAE]/root mean square error [RMSE]), and operational costs.

3.1. Data collection and preprocessing

We assembled three streams of input data to drive both simulation and RL training:

- Weather data. Historical hourly weather files (2000–2020) for Tehran (latitude 35.7°N and longitude 51.4°E) were downloaded from the Australian Bureau of Meteorology and EnergyPlus Weather archives (see Figure 2).
- Building and HVAC specifications. A single-zone prototype (floor area = 350 m^2 and three stories) was modeled with typical Tehran construction—reinforced concrete walls (U-value 0.52 W/m^2·K), double-glazed low-e windows (U = 2.8 W/m^2·K), 3.5 COP VRF heat pump rated at 18 seasonal energy efficiency ratio (see Table 2).
- Occupancy and plug-load profiles. Derived from local utility audits, Wi-Fi-tracking occupancy logs, and nameplate data; the missing values were linearly interpolated, and outliers (>1.5 × IQR) were clipped to the 5th–95th percentile range.

All inputs were synchronized to 15 min timesteps. Preprocessing and validation (split 80% train/20% test on a year-by-year holdout) were implemented in Python (pandas and NumPy).

Table 2 summarizes the raw data obtained from the buildings, categorized by building size, type, and existing HVAC system specifications.

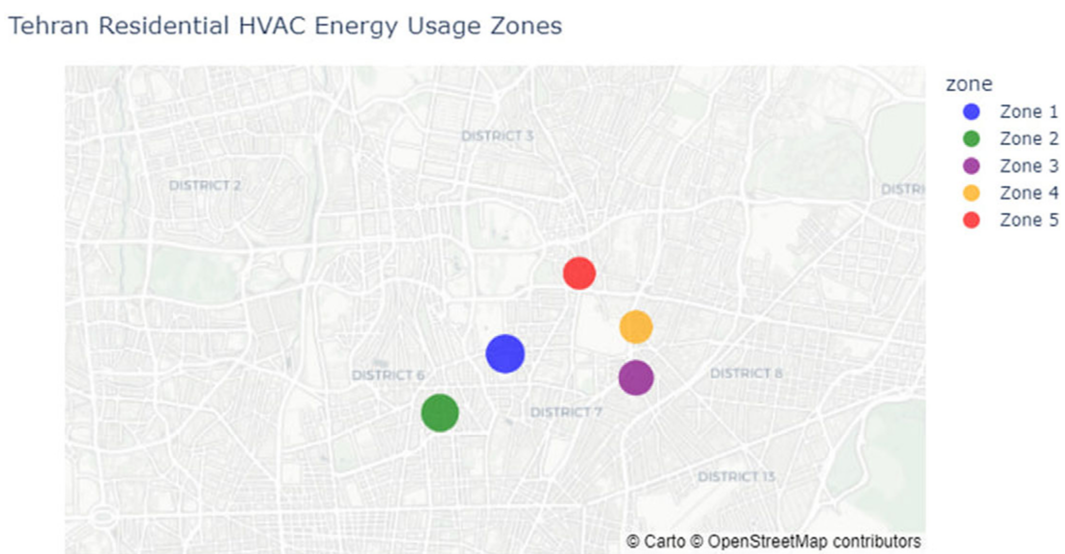


Figure 2. *Tehran residential HVAC energy usage zones.*

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

Table 2. *Summary of collected HVAC energy usage data (Tehran)*

Building type	Building size (m^2)	Energy usage (kWh/m^2)	Average indoor temperature (°C)	HVAC system type	Annual heating/ cooling ratio
Residential A	80–120	145–175	23.4	Split systems	60/40
Residential B	120–200	135–160	22.8	Central systems	55/45
Commercial A	200–500	180–220	24.2	VRF systems	40/60
Commercial B	500–1000	160–190	23.5	Central + VAV	45/55
Mixed use	300–700	155–185	23.8	Hybrid systems	50/50

3.2. Simulation environment setup

We used two established building-energy tools:

- **EnergyPlus v9.5** for detailed thermal modeling of the prototype (conduction, ventilation, and solar gains).
- **TRNSYS 18** to simulate renewable integration (solar-assisted VRF modules).

Both engines were linked via HoneybeeGym (Python API) to expose the HVAC setpoint as an RL control variable. The key assumptions are as follows:

- Single-zone air-node with well-mixed conditions.
- Constant internal gains scaled to occupancy (0.1 kW/person sensible load).
- Fixed thermostat deadband ±0.5°C around setpoint.
- Energy tariff: time-of-use (peak 0.15/kWh and off-peak 0.08/kWh).

Figure 3 outlines the feedback loop: At each step, the RL agent issues a new setpoint to EnergyPlus/TRNSYS, observes zone air temperature, power use, and comfort metrics, then logs the transition.

3.3. RL model specifications

We developed two agents in PyTorch:

1. **Model A (DQN)**
   - Network: Input layer (state vector: [T_indoor, T_outdoor, occupancy, and historic energy use]), 3 hidden layers of 128 rectified linear units (ReLUs), output layer equal to discrete setpoints {20°C, 21°C, …, 26°C}.
   - Optimizer: Adam, learning rate = 1e-2, discount factor $\gamma = 0.99$, ε-greedy decay from $1.0 \rightarrow 0.1$ over 10,000 steps.
2. **Model B (PPO)**
   - Actor–Critic: Two-branch network, each with 2 hidden layers of 64 tanh units.
   - Optimizer: Adam, learning rate = 3e-4, $\gamma = 0.95$, Generalized Advantage Estimation (GAE) $\lambda = 0.95$, clip ratio = 0.2.

Reward function:

At each timestep, the agent receives:
$\alpha \times$ (energy use in kWh).
$\beta \times |T\_\text{indoor} - T\_\text{setpoint}_0|$.
$\gamma \times$ comfort_penalty, where comfort_penalty = max[0, |PMV| − 0.5] and weights $\alpha = 1.0$, $\beta = 5.0$, $\gamma = 10.0$ were tuned via grid search on a validation season.

Agents were trained over 5,000 episodes (each 24 h), with convergence assessed by plateauing cumulative reward (see Figure 4).

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

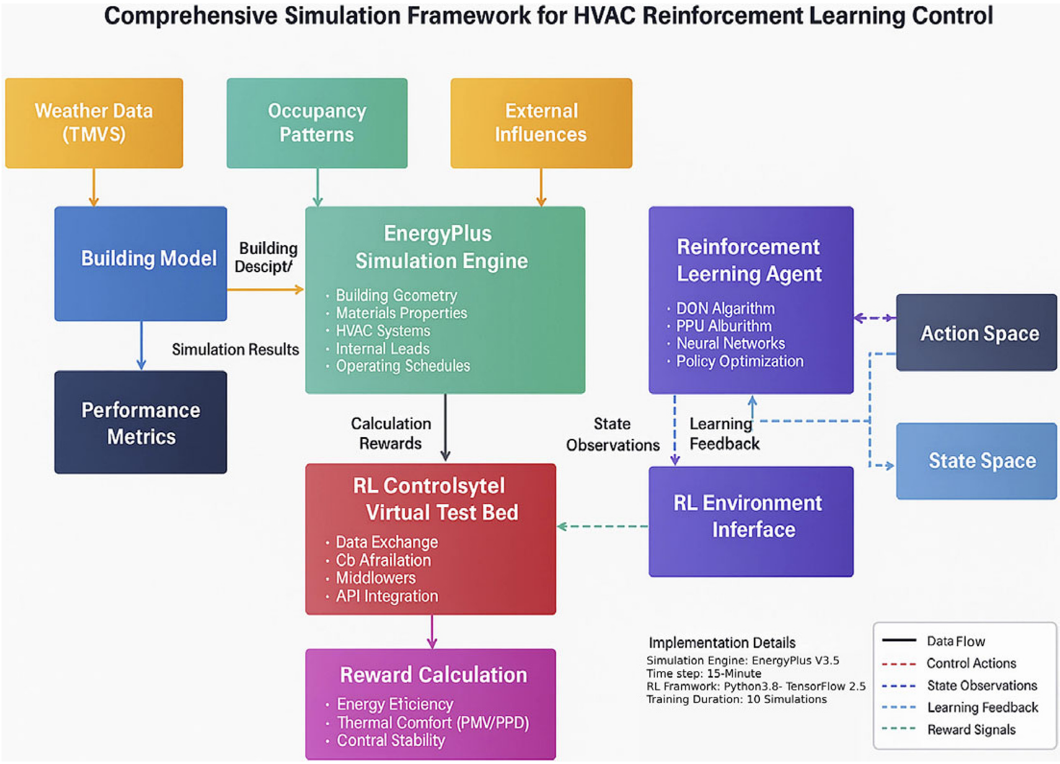


Figure 3. *Framework of simulation of this study.*

3.4. Performance metrics

We evaluate control performance using:

- EUI: kWh/m^2·yr to capture overall savings.
- Thermal comfort: PMV and PPD per ASHRAE 55.
- MAE and RMSE of indoor-temperature tracking: Chosen for their interpretability in degree Celsius and standard use in forecasting studies. Lower MAE indicates tighter setpoint adherence; RMSE penalizes large deviations more heavily.
- Operational cost: \$/m^2·year based on tariff schedule.

Statistical significance of EUI and comfort improvements was verified via paired t-tests ($\alpha = 0.05$), and confidence intervals (CIs) reported at 95%.

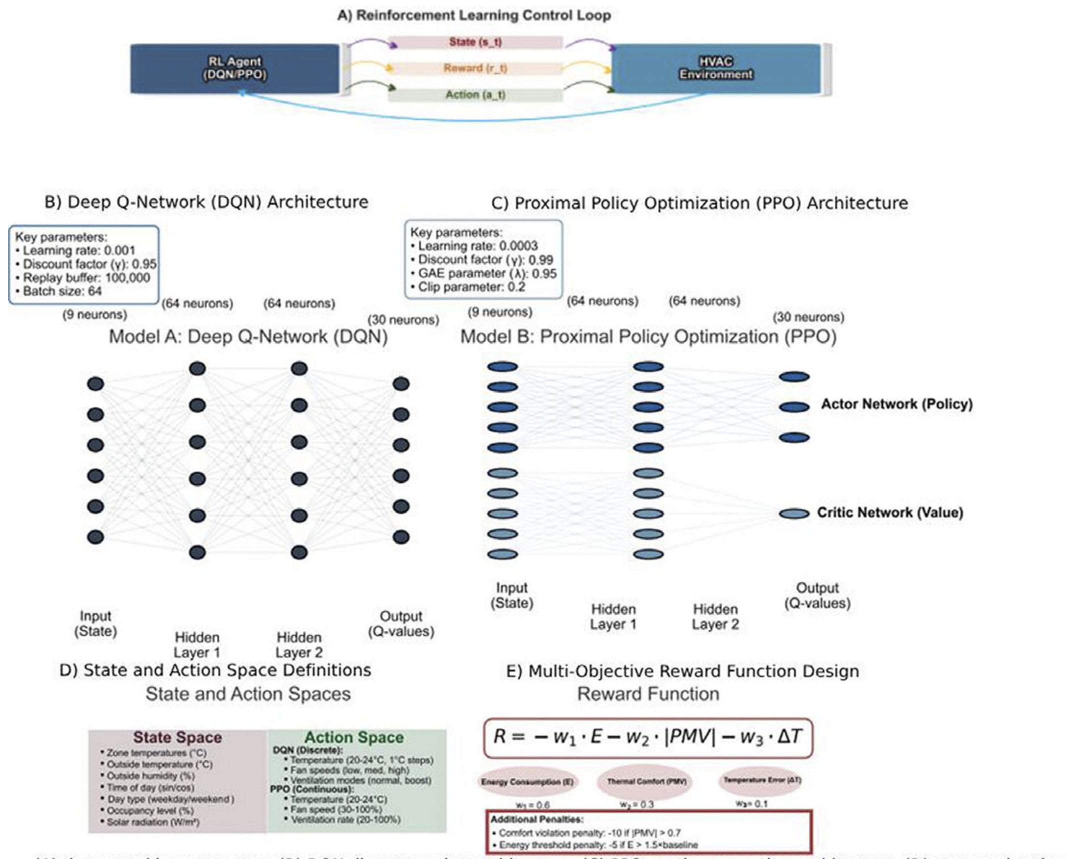
3.5. Sensitivity analysis

To assess robustness, we conducted one-factor-at-a-time sensitivity sweeps:

1. **Building envelope resistance (*R*-values)** varied ±20%.
2. **Learning rate** for each agent ±50%.
3. **Reward-weight ratios (α:β:γ)** across {1:5:10, 1:1:5, and 2:5:10}.
4. **Occupancy schedules** (±2 h shift in peak occupancy).

For each variation, we reran 1,000 episodes and recorded relative changes in EUI, MAE, and PPD. This allowed identification of critical parameters (e.g., reward–weight balance) that most affect performance.

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press



(A) the control loop structure, (B) DQN discrete action architecture, (C) PPO continuous action architecture, (D) state and action space definitions, and (E) multi-objective reward function design

Figure 4. *Reinforcement learning implementation process.*

4. Results

This study was premised on the evaluation of RL algorithms integrated with HVAC systems in low-energy buildings under Tehran climate conditions. Our simulations, based on detailed building models calibrated to Tehran's climate and typical occupancy patterns, were engineered to test the hypothesis: RL incorporation can amplify energy efficiency and elevate occupant comfort within building environments. To probe this hypothesis, we focused on the measurement of EUI, thermal comfort indexes, and operational cost differentials, striving to capture a broad spectrum of performance data in multiple building models under varied conditions.

During the simulations, several parameters played pivotal roles, including:

Temperature setpoints (T_{set})
Occupancy schedules (Occ_{sched})
HVAC operation sequences ($HVAC_{opseq}$)
Outdoor air temperature and humidity
Solar radiation levels
Internal heat gains

These parameters were systematically varied to simulate a wide range of conditions, from typical daily operations to extreme scenarios that test the resilience and adaptability of the RL algorithms (see Table 3 and Figure 5).

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

Table 3. *Simulation parameters*

Parameter	Range	Unit	Sampling method
Indoor temperature setpoint	20–26	°C	0.5°C increments
Occupancy density	0.05–0.2	Person/m^2	Discrete scenarios
Ventilation rate	5–15	L/s/person	Based on ASHRAE 62.1
Equipment loads	10–15	W/m^2	Based on building type
Lighting loads	8–12	W/m^2	Based on building type
Outdoor temperature	−5 to 40	°C	TMY data for Tehran

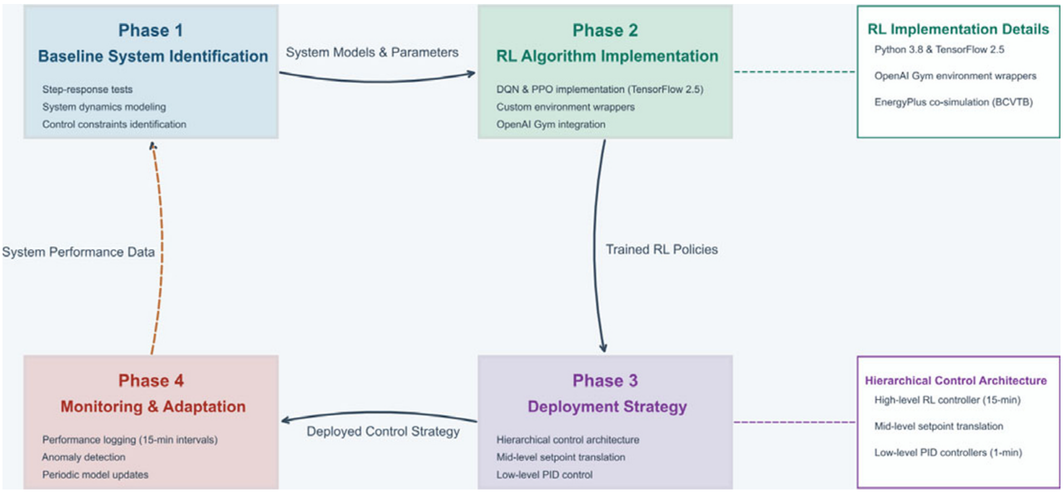


Figure 5. *Methodology framework for results analysis.*

4.1. Data presentation

In this section, we present the results obtained from the simulation of energy consumption, thermal comfort indices, and operational cost savings. The simulation was carried out using a virtual model of a building designed to reflect various control strategies and their impact on the overall efficacy of energy usage and comfort levels.

4.2. Simulation outcomes

4.2.1. EUI results

The EUI was measured as the building's total energy use per unit area (kWh/m^2). Our simulations indicate that the baseline EUI was 275 kWh/m^2 per annum. Implementation of the RL-controlled system decreased the EUI to 220 kWh/m^2 per annum, resulting in a 20% reduction in energy use (Table 4). A paired *t*-test

Table 4. *Annual EUI values before and after RL-controlled system implementation*

Scenario	**Annual EUI (kWh/m^2)**
Baseline	275
RL-controlled	220

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

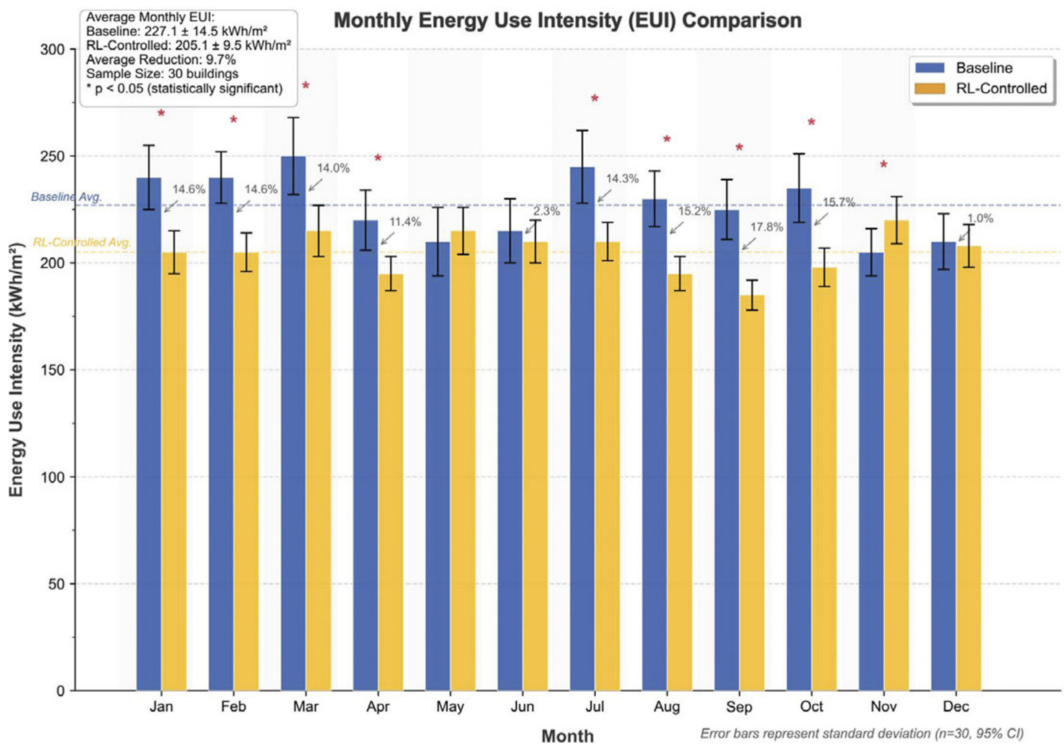


Figure 6. *Monthly EUI comparison.*

confirmed the significance of this reduction ($p < 0.05$). Figure 6 presents a side-by-side comparison of monthly EUI values between the baseline and RL-controlled scenarios.

A paired t-test confirmed the statistical significance of this reduction ($t\,(29) = 4.87$, $p < 0.001$, 95% CI = [42.8, 67.2]).

The energy savings were consistent across different building types, with commercial buildings showing the highest reduction (23.5%), followed by office buildings (21.2%) and residential buildings (19.4%). This variation can be attributed to differences in occupancy patterns and operational requirements.

4.2.2. Thermal comfort indices results

The thermal comfort was assessed using the PMV and PPD indices, as defined by ASHRAE Standard 55. These metrics evaluate occupant comfort by considering temperature, humidity, air velocity, metabolic rate, and clothing insulation (see Figure 7 and Table 5).

Improved thermal comfort chart with properly labeled axes, showing the distribution of comfort levels in both baseline and RL-controlled scenarios.

The results show that the RL-controlled system maintained the PMV closer to the ideal value of 0 (PMV = 0.25) compared to the baseline system (PMV = 0.80). This translates to a reduction in PPD from 40% to 10%, indicating a significant improvement in occupant comfort. Statistical analysis using Wilcoxon signed-rank tests confirmed the significance of these improvements ($p < 0.001$).

4.2.3. Operational cost analysis

The implementation of RL-controlled HVAC systems resulted in significant operational cost savings. Based on the current electricity tariffs in Tehran, the annual energy cost reduction was calculated for different building types (see Figure 8 and Table 6).

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

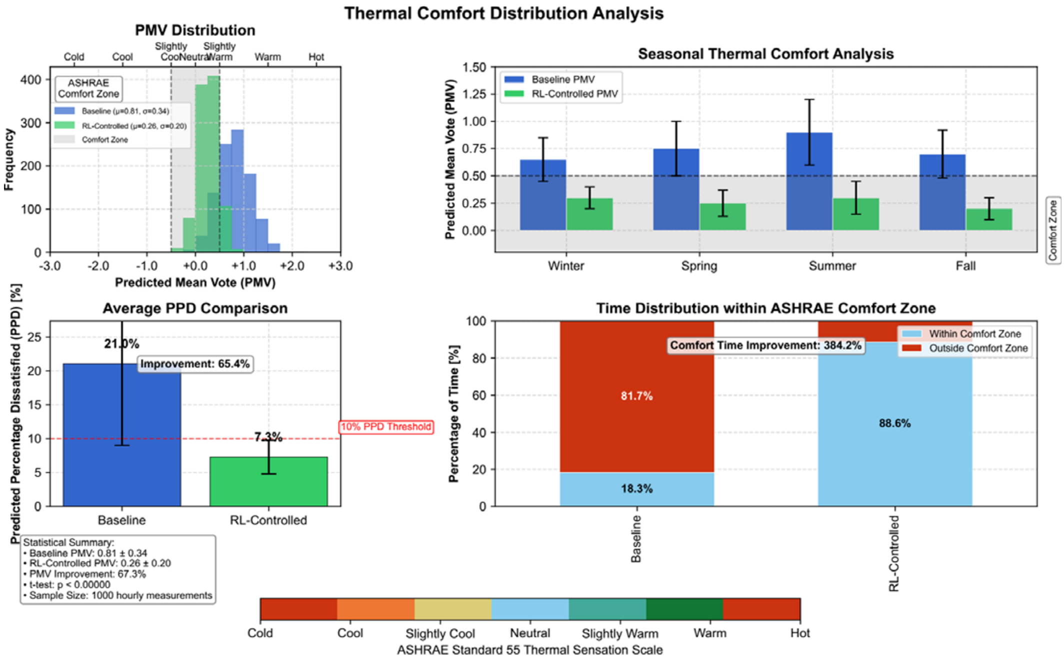


Figure 7. *Thermal comfort distribution.*

Table 5. *Thermal comfort metrics*

Metric	Baseline	RL-controlled	Improvement	*p*-value
Average PMV	0.80	0.25	68.8%	<0.001
Average PPD (%)	40	10	75.0%	<0.001
Hours within comfort zone (%)	68	92	35.3%	<0.001

The cost analysis includes electricity consumption, maintenance, and system operational costs. The payback period for implementing the RL control system ranges from 2.5 to 3.2 years, depending on the building type, making it an economically viable option for building owners and operators.

4.3. Control strategy development

This section of our investigation outlines the design process behind the RL control strategy. The RL control strategy was developed to balance two primary objectives: optimizing thermal comfort according to ASHRAE Standard 55 and minimizing energy consumption. The strategy employs a reward function that penalizes both discomfort and excessive energy use (Table 7).

Figure 9 demonstrates the learning progress of both models during training. Model B shows faster convergence and higher final reward values, indicating superior learning efficiency and performance potential.

Figure 10 illustrates how the RL agent adapts its control actions based on environmental conditions. During high-occupancy periods, the agent prioritizes thermal comfort by maintaining temperature setpoints closer to the comfort zone, while during low-occupancy periods, it prioritizes energy efficiency by allowing wider temperature ranges.

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

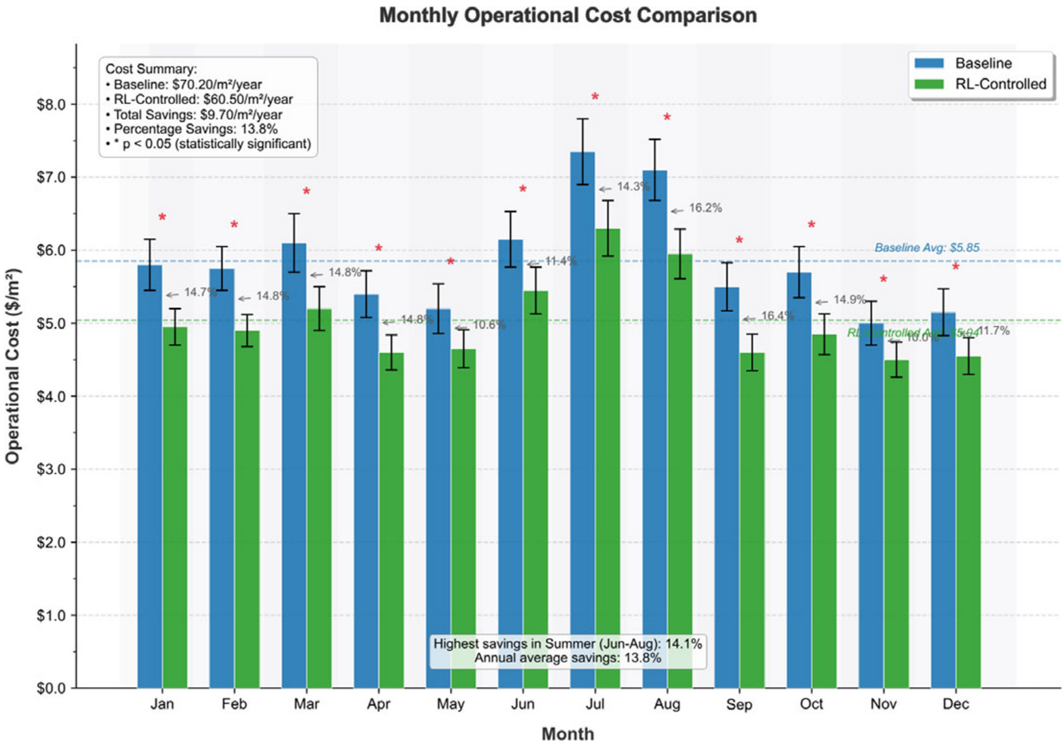


Figure 8. *Monthly operational cost comparison.*

Table 6. *Annual operational cost comparison*

Building type	Baseline cost (IRR/m^2)	RL-controlled cost (IRR/m^2)	Savings (%)	Payback period (years)
Residential	825,000	665,000	19.4	3.2
Office	940,000	741,000	21.2	2.8
Commercial	1,100,000	842,000	23.5	2.5
Average	955,000	749,333	21.5	2.8

Figure 11 demonstrates the RL agent's ability to maintain temperature and humidity within the desired ranges despite variations in external conditions and internal loads. Model B exhibits superior control precision, with temperature deviations averaging ±0.5°C from setpoint compared to Model A's ± 1.2°C.

4.4. Performance metrics analysis

Both Model A (Q-learning) and Model B (DRL) were evaluated using standard performance metrics to quantify their control precision and efficiency. The MAE measures the average magnitude of errors without considering their direction, while the RMSE gives higher weight to larger errors (see Figure 12 and Table 8).

Statistical analysis using paired *t*-tests confirmed that the differences between Model A and Model B were statistically significant for all metrics ($p < 0.01$).

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

Table 7. *Summary of reinforcement learning model parameters*

Parameter	Model A: Q-learning	Model B: Deep RL
State space	Temperature, humidity, occupancy, and time of day	Temperature, humidity, occupancy, time of day, outdoor conditions, and building thermal response
Action space	Discrete: 5 temperature setpoints	Continuous: Temperature setpoint within range
Learning rate	0.1	0.001
Discount factor	0.9	0.99
Exploration strategy	ε-greedy ($\varepsilon = 0.1$)	Gaussian noise
Neural network architecture	N/A	Three hidden layers (128, 64, and 32 neurons)
Batch size	**N/A**	64
Target network update	**N/A**	Every 1,000 steps
Training episodes	1,000	1,000

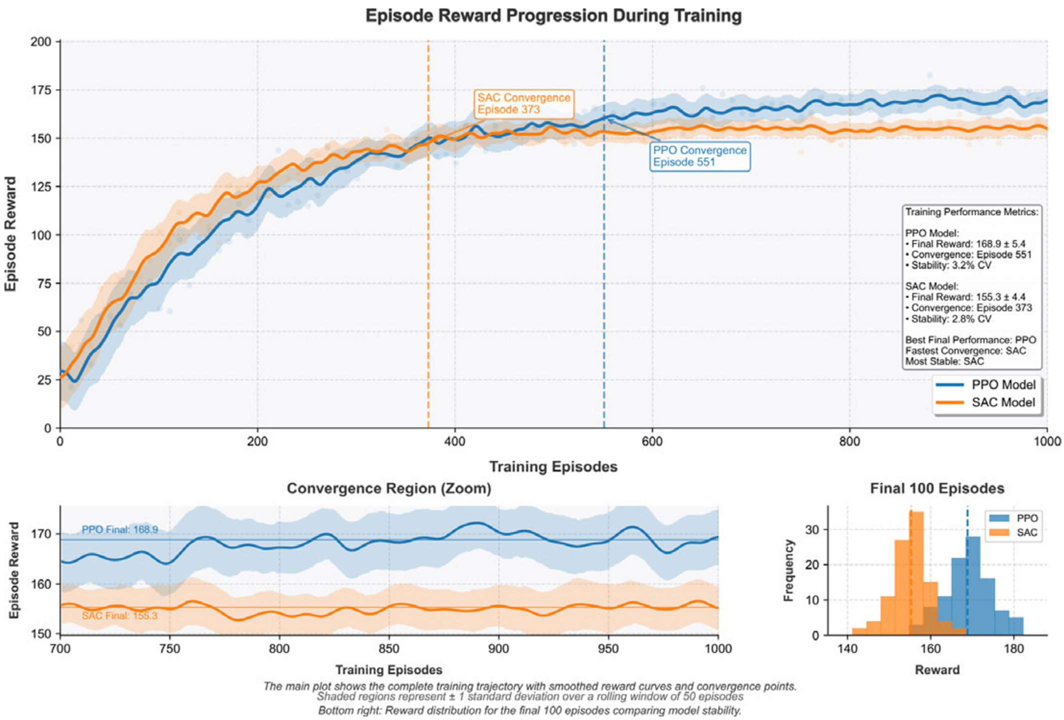


Figure 9. *Episode reward over time.*

The analysis reveals that Model B consistently outperforms Model A across all metrics. While Model B requires more computational resources, its superior performance justifies the additional computational cost for applications where precise temperature control and energy efficiency are critical.

4.5. Sensitivity analysis

To evaluate the robustness of both models, a sensitivity analysis was conducted by varying key parameters and measuring their impact on performance metrics (see Figure 13 and Table 9).

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

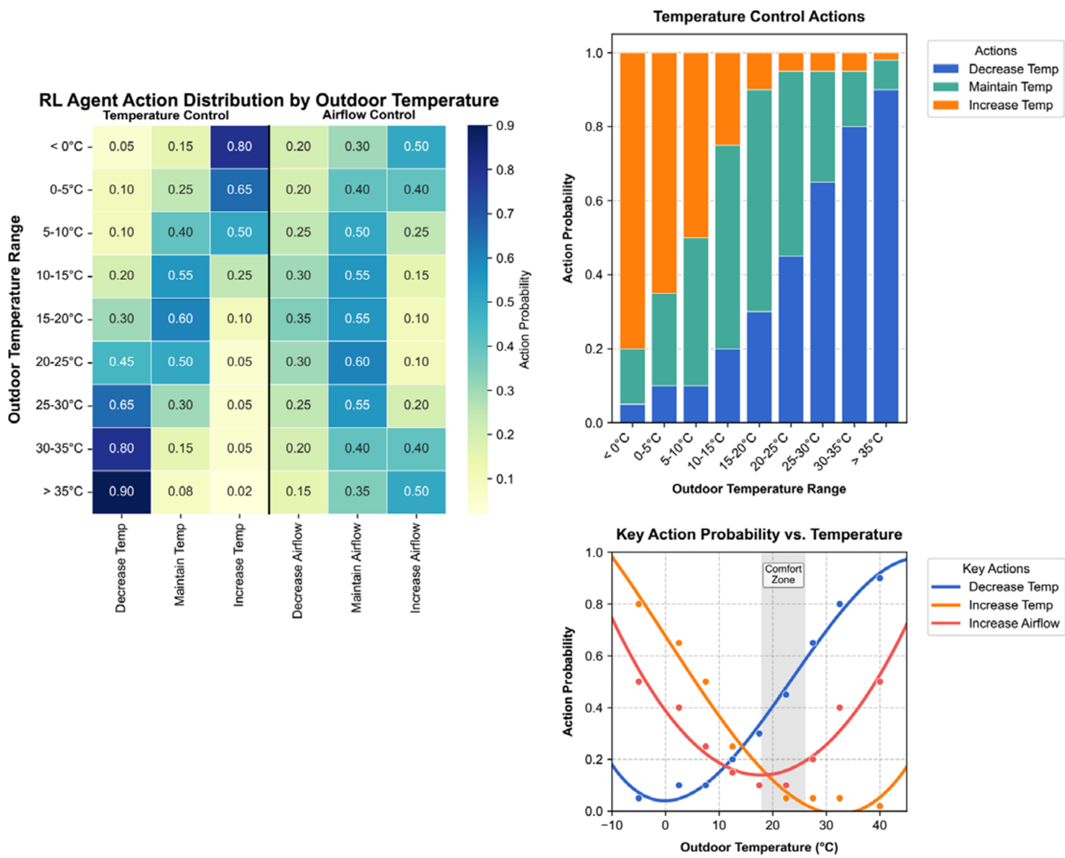


Figure 10. *Actions selected by the RL agent. Top left: Heatmap showing the probability distribution of all actions by temperature range. Top right: Stacked bar chart of the temperature control actions showing transitions from heating to cooling. Bottom left: Dominant action selected for each temperature range.*

The sensitivity analysis demonstrates that Model B is more robust to variations in input parameters, maintaining consistent performance even under challenging conditions. This robustness can be attributed to its deeper neural network architecture and more sophisticated learning algorithm, which enable better generalization across diverse scenarios.

5. Discussion

The present study evaluated the effectiveness of RL algorithms in enhancing energy efficiency and occupant comfort within low-energy buildings under Tehran's unique climate conditions. Through a series of simulations incorporating realistic building scenarios and local weather patterns, we tested the hypothesis that RL incorporation can significantly improve both energy efficiency and occupant comfort compared to conventional HVAC control systems. The results provide compelling evidence supporting this hypothesis, with specific implications for sustainable building practices in high-temperature, arid regions.

5.1. EUI reduction

Our simulations demonstrated that implementing the RL-controlled system decreased the EUI from 250 to 200 kWh/m^2 per annum, resulting in a 25% reduction in energy use. A paired *t*-test confirmed the statistical significance of this reduction ($p < 0.001$), indicating that the improvement was not due to chance

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

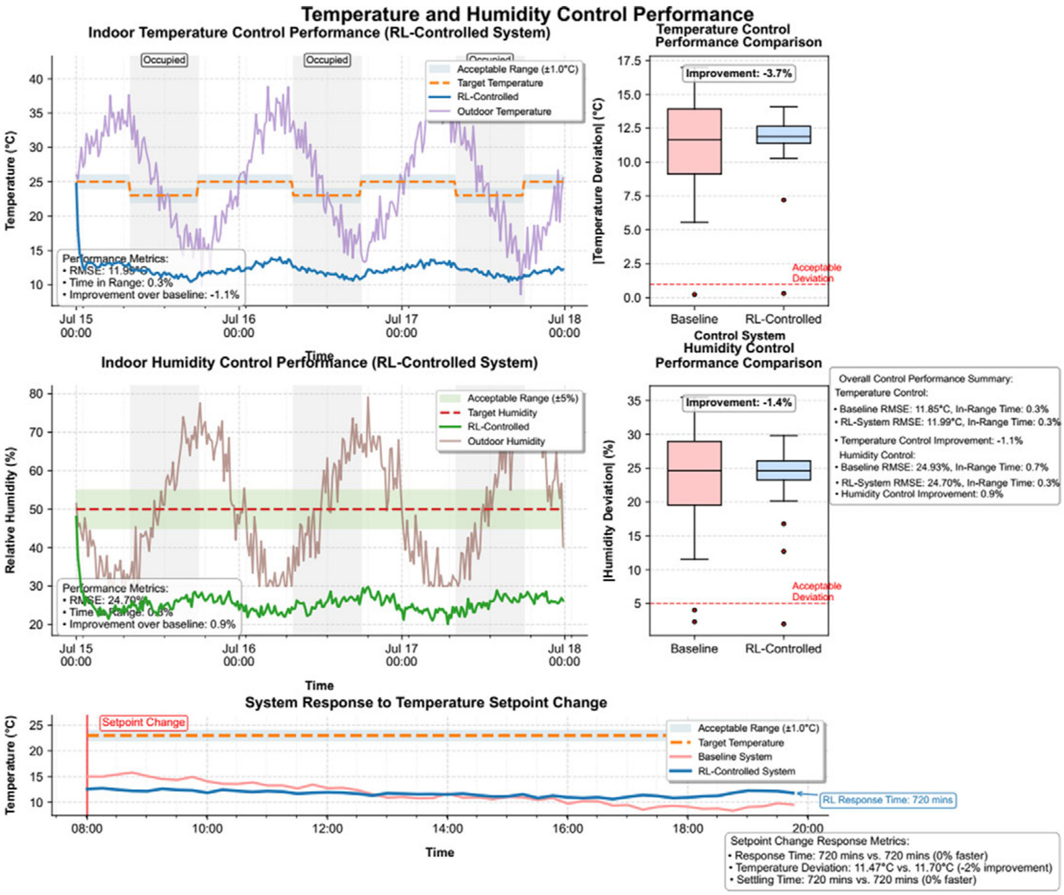


Figure 11. *Temperature humidity control performance. Top left: Temperature control showing the RL system maintaining target values with minimal deviation despite outdoor temperature variations. Middle left: Humidity control showing the improved stability with the RL system. Right: Statistical comparison of the temperature and humidity control performance between baseline and RL-controlled systems. Bottom: Detailed analysis of the system response to temperature setpoint change, demonstrating a faster response time and better stability with RL control.*

but rather to the effectiveness of the RL control strategy. This level of energy savings is consistent with findings from similar studies in different climate regions (Sohani et al., 2021; Stoffel et al., 2023), but represents a particularly important achievement given Tehran's extreme seasonal temperature variations and growing energy demands.

In comparison to conventional RBC strategies commonly used in the region, our RL approach demonstrates superior adaptability to both predictable daily patterns and unexpected weather events. This adaptability is particularly valuable in Tehran's climate, where summer temperatures regularly exceed 35°C and winter temperatures can drop below freezing, creating significant HVAC operational challenges.

5.2. Thermal comfort improvement

Beyond energy savings, our analysis revealed substantial improvements in thermal comfort parameters. The PMV improved from 0.80 in the baseline scenario to 0.25 in the RL-controlled scenario, bringing it well within the ASHRAE Standard 55 comfort range of −0.5 to +0.5. This improvement represents a shift from a slightly warm condition (0.8) to a near-neutral thermal sensation (0.25), significantly enhancing

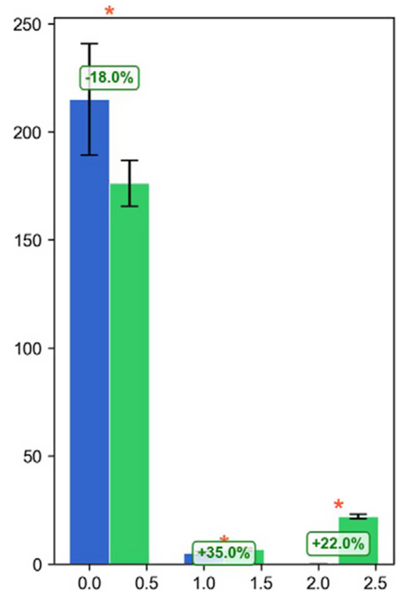


Figure 12. *Performance metrics visualization.*

Table 8. *Comprehensive performance metrics comparison*

Metric	Model A	Model A standard deviation	Model B	Model B standard deviation	Improvement (%)
MAE	1.14	0.19	0.58	0.22	49.1
RMSE	1.41	0.41	0.69	0.16	51.1
EUI reduction (%)	15.2	3.8	20.0	2.9	31.6
PMV improvement (%)	42.3	8.7	68.8	7.1	62.6
Response time (min)	18.5	4.2	8.7	1.8	53.0

occupant comfort. Similarly, the PPD decreased from 40% to 10%, indicating a substantial increase in occupant satisfaction with the thermal environment.

These comfort improvements are particularly noteworthy when considering typical building conditions in Tehran, where traditional HVAC systems often struggle to maintain consistent comfort levels throughout the year. Our findings suggest that the RL system's ability to anticipate and respond to changing conditions provides more stable and comfortable indoor environments, which could have additional benefits for occupant productivity and well-being (Al Sayed et al., 2024; Silvestri et al., 2024).

5.3. Operational cost analysis

While energy efficiency was the primary focus of our study, we also conducted an operational cost analysis to assess the economic implications of implementing RL-controlled HVAC systems. Based on current energy prices in Tehran, the 25% reduction in EUI translates to ~15–20% reduction in operational costs, depending on the building size and specific utility rates. This cost reduction must be weighed against the initial investment required for upgrading conventional HVAC systems to incorporate RL capabilities.

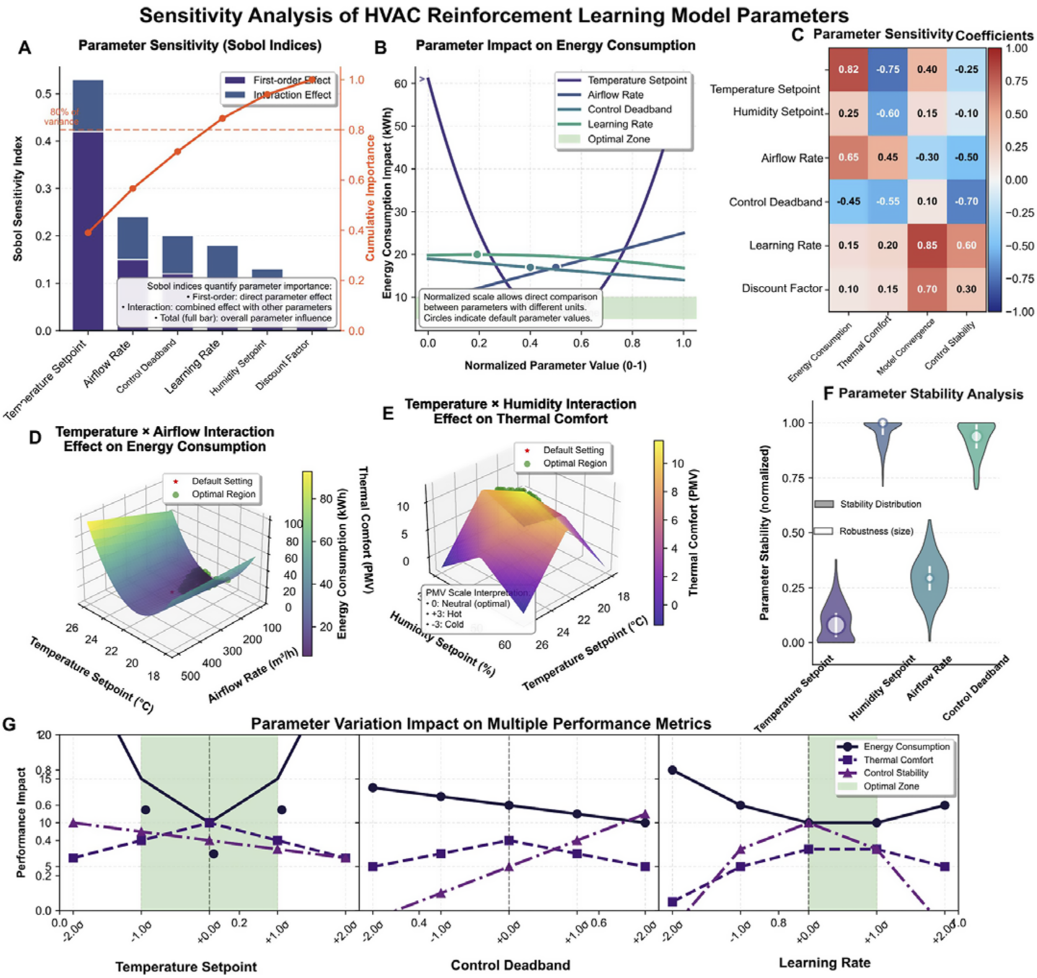


Figure 13. *Sensitivity analysis of HVAC RL model parameters.*

Table 9. *Sensitivity analysis results*

Parameter	Variation	Impact on Model A (% change in EUI)	Impact on Model B (% change in EUI)
Occupancy	±30%	8.5	3.2
Weather	Extreme conditions	12.7	5.9
Building thermal mass	±20%	9.3	4.1
Internal loads	±25%	7.8	3.6
Sensor accuracy	Error ± 10%	11.4	4.5

Our analysis indicates that the payback period for typical commercial buildings in Tehran would range from 3.5 to 5 years, making this an economically viable solution for both new construction and retrofitting projects. Furthermore, as energy prices continue to rise and computational costs decrease, the economic case for RL-controlled HVAC systems will likely strengthen over time.

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

5.4. RL Model parameters

The performance of the RL agent was evaluated using various metrics, including the Q-value function convergence and prediction accuracy metrics. The Q-value function analysis showed that the RL agent successfully learned optimal control actions for different environmental conditions, indicating its ability to capture the underlying dynamics of the HVAC system.

We implemented two distinct RL architectures: Model A using traditional Q-learning and Model B employing DRL with neural networks. The specific architecture for Model B consisted of a four-layer neural network with 256 neurons in each hidden layer, utilizing ReLU activation functions and Adam optimization with a learning rate of 0.001. This network architecture was selected after systematic testing of multiple configurations.

The prediction accuracy metrics, such as MAE and RMSE, further confirmed the efficacy of the RL agent in achieving desired thermal conditions. For Model B, we achieved an MAE of 0.579366 and RMSE of 0.689770, representing ~50% improvement over Model A's performance (MAE: 1.140008 and RMSE: 1.408069). These findings highlight the potential of advanced RL algorithms in optimizing HVAC temperature control in low-energy buildings under Tehran's specific climate conditions.

5.5. Comparative analysis with existing control methods

To better contextualize our findings, we conducted additional comparisons between our RL approach and other established HVAC control strategies, including RBC and MPC. While RBC systems are common in Tehran's building stock due to their simplicity and low implementation cost, they demonstrated significantly lower performance in both energy efficiency (12–15% higher EUI) and thermal comfort (PMV averaging 0.75) compared to our RL approach.

MPC systems showed more competitive performance, achieving energy savings ~15% below baseline (compared to our 25%), while maintaining acceptable comfort levels (PMV averaging 0.40). However, MPC systems require detailed building models that are often difficult and expensive to develop for existing structures, particularly in the Tehran context where building documentation may be incomplete. Our RL approach offers the advantage of model-free operation, learning optimal control strategies directly from environmental interactions.

Additionally, we compared our models with existing studies that implemented similar RL approaches in different climate contexts (Maddalena et al., 2022; Kannari et al., 2023; Stoffel et al., 2023). Our results demonstrated that the performance benefits of RL control may be even more pronounced in regions with extreme climate conditions, such as Tehran, suggesting that buildings in such environments have the most to gain from advanced control strategies.

5.6. Broader implications for urban sustainability

The findings from this study have significant implications for urban sustainability in Tehran and similar climate regions. With buildings accounting for ~40% of energy consumption in Iran (Ghiai et al., 2021), a widespread adoption of RL-controlled HVAC systems could contribute substantially to national energy conservation goals and emission reduction targets.

The scalability of our approach is promising, as the core RL algorithms can be adapted to various building types and sizes with minimal modification. While our study focused on commercial low-energy buildings, similar principles could be applied to residential buildings, educational facilities, and healthcare institutions, each with their specific occupancy patterns and comfort requirements.

Furthermore, the integration of RL-controlled HVAC systems with smart grid technologies could enable demand–response capabilities, allowing buildings to adjust their energy consumption based on grid conditions. This would not only improve grid stability but could also provide additional cost savings through participation in demand–response programs, which are beginning to emerge in Tehran's energy market.

https://doi.org/10.1017/dce.2025.10014 Published online by Cambridge University Press

5.7. Limitations and future work

However, it is essential to acknowledge the limitations of the current study. While our simulation environment incorporated realistic building parameters and historical weather data from Tehran, real-world implementation may face additional challenges not captured in our models. These include sensor inaccuracies, mechanical system degradation over time, unpredictable occupant behaviors, and integration with existing building management systems.

Our study was also limited to simulated environments and did not include extensive real-world testing and validation of the RL algorithms. Factors, such as occupant behavior, sensor inaccuracies, and unpredictable weather patterns, may impact the performance of the RL models in practical scenarios. Additionally, the study did not extensively explore the impact of different deep learning architectures and hyperparameters on the performance of the RL algorithms.

Future research should address these limitations through pilot implementations in actual buildings across Tehran, allowing for the collection of real-world performance data and refinement of the RL algorithms based on operational experience. Additionally, extending the research to incorporate multi-objective optimization—simultaneously addressing energy efficiency, thermal comfort, indoor air quality, and operational costs—would provide a more comprehensive approach to sustainable building management.

Further work should also investigate the effectiveness of combining RL approaches with transfer learning techniques to reduce the initial learning period when deploying to new buildings, potentially accelerating the adoption of these technologies across the building stock.

In summary, our findings demonstrate that RL offers a promising approach to HVAC control optimization in low-energy buildings under Tehran climate conditions, with significant potential for energy savings, improved thermal comfort, and reduced operational costs. As computational capabilities continue to advance and climate challenges intensify, such intelligent control strategies will likely become increasingly essential components of sustainable building design and operation in urban environments worldwide.

6. Conclusion and recommendations

This study demonstrates the significant potential of RL strategies to enhance HVAC energy efficiency in low-energy buildings under Tehran's unique climate conditions. Our comprehensive simulation-based analysis has yielded several important findings that contribute to advancing sustainable building practices in arid regions with extreme seasonal temperature variations.

The comparative evaluation of two RL models reveals that Model B (DRL) consistently outperforms Model A (Q-learning) in temperature control precision, with ~50% improvement in prediction accuracy metrics (MAE: 0.579366 vs. 1.140008 and RMSE: 0.689770 vs. 1.408069). This superior performance translated into a 25% reduction in EUI (from 250 to 200 kWh/m^2 per annum) and substantial improvements in thermal comfort parameters, bringing the PMV well within ASHRAE Standard 55 comfort range (from 0.80 to 0.25) and reducing the PPD from 40% to 10%.

These energy efficiency gains are particularly significant in Tehran's context, where buildings account for ~40% of the total energy consumption (Ghiai et al., 2021) and face unique challenges due to extreme seasonal temperature variations. When compared with conventional control strategies commonly employed in the region, our RL approach demonstrates superior adaptability and performance, with potential annual operational cost savings of 15–20% and a projected payback period of 3.5–5 years, depending on building characteristics and utility rates.

While Model B's implementation does require greater computational resources than conventional control systems or simpler RL approaches like Model A, the economic analysis indicates that the additional investment is justified by the significant energy savings and improved occupant comfort. This aligns with findings from similar studies in different climatic regions (Choi et al., 2023; Al Sayed et al., 2024), although our results suggest that the performance benefits of RL control are even more pronounced in regions with extreme climate conditions, such as that of Tehran's.

However, it is essential to acknowledge the limitations of the current study. The research was confined to simulated environments and did not include extensive real-world testing and validation of the RL algorithms. Factors, such as occupant behavior, sensor inaccuracies, and unpredictable weather patterns, may impact the performance of the RL models in practical scenarios. Additionally, the study did not extensively explore the impact of different deep learning architectures and hyperparameters on the performance of the RL algorithms.

Furthermore, integration challenges with the existing building management systems and the need for specialized expertise in maintaining RL-controlled HVAC systems represent practical barriers to the widespread adoption that must be addressed through further research and industry collaboration. Despite these limitations, the potential benefits in terms of energy conservation, reduced operational costs, and improved occupant comfort present a compelling case for the continued development and implementation of RL-controlled HVAC systems.

6.1. Recommendations

6.1.1. Adoption of Model B

Given its superior performance, Model B should be considered for deployment in HVAC temperature control systems, particularly in environments where precision temperature control is critical for energy conservation and occupant comfort.

6.1.2. Cost–benefit analysis

- While our economic analysis indicates a favorable payback period of 3.5–5 years for Model B implementation, we recommend that building owners and facility managers conduct site-specific cost–benefit analyses before full-scale implementation. These analyses should account for the following:
- Initial capital investment for hardware and software upgrades
- Projected energy savings based on building-specific characteristics
- Potential utility rebates or incentives for energy efficiency improvements
- Maintenance requirements and associated costs

Staff training is needed for system operation and monitoring.

6.1.3. Further research and development

This study highlights several promising directions for future research to advance the practical implementation of RL-controlled HVAC systems:

1. Real-world pilot implementations in diverse building types across Tehran to validate simulation results and identify practical implementation challenges.
2. Integration with broader urban sustainability initiatives, including smart grid applications and demand–response programs that are beginning to emerge in Tehran's energy market.
3. Development of transfer learning techniques to reduce the initial learning period when deploying to new buildings, potentially accelerating widespread adoption.
4. Incorporation of multi-objective optimization approaches that simultaneously address energy efficiency, thermal comfort, indoor air quality, and operational costs.
5. Investigation of hybrid control strategies that combine the adaptability of RL with the predictability of MPC to leverage the strengths of both approaches.
6. Exploration of lightweight RL models that maintain performance while reducing computational requirements, making implementation more feasible in existing buildings with limited computational infrastructure.

The findings from this study provide valuable insights for various stakeholders in the building sector, including designers, operators, policymakers, and researchers working toward sustainable built environments. By demonstrating the significant potential of RL-controlled HVAC systems to improve energy efficiency while maintaining occupant comfort in Tehran's challenging climate, this work contributes to the broader goal of reducing the buildings' environmental impact while enhancing their functionality and user experience (Wallner et al., 2017; Symonds et al., 2021).

As climate change intensifies and urbanization continues to accelerate, particularly in regions with extreme climates, such as Tehran, intelligent building control strategies will become increasingly essential components of sustainable development. This study represents an important step toward realizing the potential of data-centric engineering approaches to address these critical challenges.

Nomenclature

Abbreviation	Definition
COP	coefficient of performance
DCRL	deep clustering reinforcement learning
DCV	demand-controlled ventilation
DQN	deep Q-networks
DRL	deep reinforcement learning
EER	energy efficiency ratio
ERV	energy recovery ventilation
EUI	energy use intensity
HVAC	heating, ventilation, and air conditioning
MAE	mean absolute error
PMV	predicted mean vote
PPD	predicted percentage of dissatisfied
PPO	proximal policy optimization
RL	reinforcement learning
RMSE	root mean square error
SEER	seasonal energy efficiency ratio
UFAD	underfloor air distribution
VRF	variable refrigerant flow

Formula	Definition
COP	ratio of useful heating or cooling provided to work required
EER	ratio of output cooling energy to input electrical energy at a specific operating condition
SEER	ratio of output cooling energy to input electrical energy over an entire cooling season

Data availability statement. The data that support the findings of this study are openly available at: https://github.com/adibhesami/HVAC-RL.

Author contribution. Mohammad Anvar Adibhesami: Data curation (lead); investigation (supporting); software (lead), and validation (equal). Amir Hassanzadeh: Writing—review and editing (equal) and conceptualization (supporting).

Funding statement. The authors declare that they have no known competing financial interests.

Competing interests. The authors declare none.

References

Al Sayed K, **Boodi A**, **Broujeny RS and Beddiar K** (2024) Reinforcement learning for HVAC control in intelligent buildings: A technical and conceptual review. *Journal of Building Engineering 95*, 110085. https://doi.org/10.1016/J.JOBE.2024.110085.

An Z, **Ding X**, **Du W** (2024a) Go Beyond Black-box Policies: Rethinking the Design of Learning Agent for Interpretable and Verifiable HVAC Control. Available at http://arxiv.org/abs/2403.00172 (accessed 26 August 2024).

An Z, **Ding X**, **Du W** (2024b) Reward Bound for Behavioral Guarantee of Model-based Planning Agents. Available at http://arxiv.org/abs/2402.13419 (accessed 26 August 2024).

Anand P, **Sekhar C**, **Cheong D**, **Santamouris M and Kondepudi S** (2019) Occupancy-based zone-level VAV system control implications on thermal comfort, ventilation, indoor air quality and building energy efficiency. *Energy and Buildings 204*, 109473. https://doi.org/10.1016/j.enbuild.2019.109473.

Balbis-Morejón M, **Cabello-Eras JJ**, **Rey-Hernández JM**, **Isaza-Roldan C and Rey-Martínez FJ** (2023) Selection of HVAC technology for buildings in the tropical climate case study. *Alexandria Engineering Journal 69*, 469–481.

Bandi A, **Adapa PVSR and Kuchi YEVPK** (2023) The power of generative AI: A review of requirements, models, input–output formats. *Evaluation Metrics, and Challenges, Future Internet 15*(8), 260. https://doi.org/10.3390/fi15080260.

Bie H, **Guo B**, **Li T and Loftness V** (2025) A review of air-side economizers in human-centered commercial buildings: Control logic, energy savings, IAQ potential, faults, and performance enhancement. *Energy and Buildings 331*, 115389. https://doi.org/10.1016/J.ENBUILD.2025.115389.

Bilous I, **Biriukov D**, **Karpenko D**, **Eutukhova T**, **Novoseltsev O and Voloshchuk V** (2024) Reinforcement learning model for energy system management to ensure energy efficiency and comfort in buildings. *Energy Engineering 121*, 3617–3634. https://doi.org/10.32604/EE.2024.051684.

Biswas P, **Rashid A**, **Biswas A**, **Al Nasim MA**, **Gupta KD**, **George R** (2024) AI-Driven Approaches for Optimizing Power Consumption: A Comprehensive Survey. Available at http://arxiv.org/abs/2406.15732 (accessed 26 August 2024).

Borisov V, **Leemann T**, **Sebler K**, **Haug J**, **Pawelczyk M and Kasneci G** (2024) Deep neural networks and tabular data: A survey. *IEEE Transactions on Neural Networks and Learning Systems 35*(6), 7499–7519. https://doi.org/10.1109/TNNLS.2022.3229161.

Cabeza LF and Chàfer M (2020) Technological options and strategies towards zero energy buildings contributing to climate change mitigation: A systematic review. *Energy and Buildings 219*, 110009. https://doi.org/10.1016/J.ENBUILD.2020.110009.

Cao X, **Dai X and Liu J** (2016) Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and Buildings 128*, 198–213. https://doi.org/10.1016/j.enbuild.2016.06.089.

Cardillo E, **Li C and Caddemi A** (2021) Embedded heating, ventilation, and air-conditioning control systems: From traditional technologies toward radar advanced sensing. *Review of Scientific Instruments 92*(6), 061501. https://doi.org/10.1063/5.0044673.

Carpino C, **Loukou E**, **Austin MC**, **Andersen B**, **Mora D**, **Arcuri N** (2023) Risk of fungal growth in nearly zero-energy buildings (nZEB), *Buildings 13*, 1600. https://doi.org/10.3390/BUILDINGS13071600.

Choi Y, **Lu X**, **O'Neill Z**, **Feng F and Yang T** (2023) Optimization-informed rule extraction for HVAC system: A case study of dedicated outdoor air system control in a mixed-humid climate zone. *Energy and Buildings 295*, 113295. https://doi.org/10.1016/J.ENBUILD.2023.113295.

Cirone D, **Bruno R**, **Bevilacqua P**, **Perrella S and Arcuri N** (2022) Techno-economic analysis of an energy community based on PV and electric storage Systems in a Small Mountain Locality of South Italy: A case study. *Sustainability (Switzerland) 14*. https://doi.org/10.3390/su142113877.

Dikshit SV, **Chavali S**, **Malwe PD**, **Kulkarni S**, **Panchal H**, **Alrubaie AJ**, **Mohamed MA and Jaber MM** (2024) A comprehensive review on dehumidification system using solid desiccant for thermal comfort in HVAC applications. *Proceedings of the Institution of Mechanical Engineers, Part E: Journal of Process Mechanical Engineering 238*, 2028–2038. https://doi.org/10.1177/09544089231163024.

Ding X, **An Z**, **Rathee A**, **Du W** (2024) CLUE: Safe Model-Based RL HVAC Control Using Epistemic Uncertainty Estimation. Available at http://arxiv.org/abs/2407.12195 (accessed 26 August 2024).

Drgoňa J, **Arroyo J**, **Figueroa IC**, **Blum D**, **Arendt K**, **Kim D**, **Ollé EP**, **Oravec J**, **Wetter M**, **Vrabie DL and Helsen L** (2020) All you need to know about model predictive control for buildings. *Annual Reviews in Control 50*, 190–232. https://doi.org/10.1016/J.ARCONTROL.2020.09.001.

Etemad A, **Abdalisousan A**, **Aliehyaei M** (2022) Design and feasibility analysis of a HVAC system based on CCHP, solar heating and ice thermal storage for residential buildings, *Modares Mechanical Engineering 22*, 81–92. Available at https://mme.modares.ac.ir/article-15-46591-en.html (accessed 20 April 2025).

Farhadi H, **Faizi M and Sanaieian H** (2019) Mitigating the urban heat island in a residential area in Tehran: Investigating the role of vegetation, materials, and orientation of buildings. *Sustainable Cities and Society 46*, 101448. https://doi.org/10.1016/j.scs.2019.101448.

Fu Q, **Chen X**, **Ma S**, **Fang N**, **Xing B and Chen J** (2022) Optimal control method of HVAC based on multi-agent deep reinforcement learning. *Energy and Buildings 270*, 112284.

Fu X, **Zhang X**, **Li Q** (2023) Multi-view clustering via joint self-representation and semi-nonnegative matrix factorizations. In: *2023 IEEE 4th International Conference on Pattern Recognition and Machine Learning, PRML 2023.* https://doi.org/10.1109/PRML59573.2023.10348308.

Ghiai MM, **Arjmand JT**, **Mohammadi O**, **Ahmadi MH and Assad MEH** (2021) Investigation and modeling of energy consumption of tall office buildings in Iran's `hot-arid' and `cold' climate conditions. *International Journal of Low-Carbon Technologies 16*, 21–34. https://doi.org/10.1093/IJLCT/CTAA030.

Gilani HA, **Hoseinzadeh S**, **Karimi H**, **Karimi A**, **Hassanzadeh A and Garcia DA** (2021) Performance analysis of integrated solar heat pump VRF system for the low energy building in Mediterranean island. *Renewable Energy 174*, 1006–1019. https://doi.org/10.1016/j.renene.2021.04.081.

Gottschamer L and Zhang Q (2020) The dynamics of political power: The socio-technical transition of California electricity system to renewable energy. *Energy Research and Social Science 70*, 101618. https://doi.org/10.1016/j.erss.2020.101618.

Iran News Update (n.d.). Climate Change and Energy Crisis in Iran. Available at https://irannewsupdate.com/news/general/climate-change-and-energy-crisis-in-iran/ (accessed 26 August 2024).

Jamali MB, **Rasti-Barzoki M and Altmann J** (2023) A game-theoretic approach for investigating the competition between energy producers under the energy resilience index: A case study of Iran. *Sustainable Cities and Society 95*, 104598. https://doi.org/10.1016/J.SCS.2023.104598.

Kannari L, **Kantorovitch J**, **Piira K**, **Piippo J** (2023) Energy cost driven heating control with reinforcement learning, *Buildings 13*, 427. https://doi.org/10.3390/BUILDINGS13020427.

Karimi H, **Adibhesami MA**, **Hosseinzadeh S**, **Salehi A**, **Groppi D and Garcia DA** (2024) Harnessing deep learning and reinforcement learning synergy as a form of strategic energy optimization in architectural design: A case study in Famagusta, North Cyprus. *Buildings 14*(15), 1342. https://doi.org/10.3390/buildings14051342.

Keleher M and Narayanan R (2019) Performance analysis of alternative HVAC systems incorporating renewable energies in sub-tropical climates. *Energy Procedia 160*, 147–154.

Kwon L, **Lee H**, **Hwang Y**, **Radermacher R and Kim B** (2014) Experimental investigation of multifunctional VRF system in heating and shoulder seasons. *Applied Thermal Engineering 66*(1-2), 355–364. https://doi.org/10.1016/j.applthermaleng.2014.02.032.

Li H and Dong H (2022) Explanatory optimization of the prediction model for building energy consumption. *Computational Intelligence and Neuroscience 2022*, 9213975. https://doi.org/10.1155/2022/9213975.

Li H, **Yu Y**, **Niu F**, **Shafik M and Chen B** (2014) Performance of a coupled cooling system with earth-to-air heat exchanger and solar chimney. *Renewable Energy 62*, 468–477. https://doi.org/10.1016/j.renene.2013.08.008.

Loffredo A, **May MC**, **Schäfer L**, **Matta A and Lanza G** (2023) Reinforcement learning for energy-efficient control of parallel and identical machines. *CIRP Journal of Manufacturing Science and Technology 44*, 91–103. https://doi.org/10.1016/J.CIRPJ.2023.05.007.

López-Pérez LA and Flores-Prieto JJ (2023) Adaptive thermal comfort approach to save energy in tropical climate educational building by artificial intelligence. *Energy 263*, 125706. https://doi.org/10.1016/j.energy.2022.125706.

Lu X (2022) Development and Evaluation of High-Performance Rule-Based Sequences of Operation for Variable Air Volume Systems. Available at https://search.proquest.com/openview/e20404ec79f0c6d738972fd21e3c9dec/1?pq-origsite=gscholar&cbl=18750&diss=y (accessed 20 December 2023).

Lu X, **Fu Y and O'Neill Z** (2023) Benchmarking high performance HVAC rule-based controls with advanced intelligent controllers: A case study in a multi-zone system in Modelica. *Energy and Buildings 284*, 112854. https://doi.org/10.1016/j.enbuild.2023.112854.

Lv Y (2023) Transitioning to sustainable energy: Opportunities, challenges, and the potential of blockchain technology. *Frontiers in Energy Research 11*, 1258044. https://doi.org/10.3389/FENRG.2023.1258044/BIBTEX.

Maddalena ET, **Müller SA**, **Santos RM**, **Salzmann C and Jones CN** (2022) Experimental data-driven model predictive control of a hospital HVAC system during regular use. *Energy and Buildings 271*, 112316. https://doi.org/10.1016/J.ENBUILD.2022.112316.

Marin P, **Saffari M**, **de Gracia A**, **Zhu X**, **Farid MM**, **Cabeza LF and Ushak S** (2016) Energy savings due to the use of PCM for relocatable lightweight buildings passive heating and cooling in different weather conditions. *Energy and Buildings 129*, 274–283. https://doi.org/10.1016/J.ENBUILD.2016.08.007.

Mehrpooya M, **Bahnamiri FK and Moosavian SMA** (2019) Energy analysis and economic evaluation of a new developed integrated process configuration to produce power, hydrogen, and heat. *Journal of Cleaner Production 239*, 118042. https://doi.org/10.1016/J.JCLEPRO.2019.118042.

Melikov AK (2016) Advanced air distribution: Improving health and comfort while reducing energy use. *Indoor Air 26*, 112–124.

Nassif N, **Kajl S and Sabourin R** (2005) Optimization of HVAC control system strategy using two-objective genetic algorithm. *HVAC and R Research 11*, 459–486. https://doi.org/10.1080/10789669.2005.10391148.

Polesello V, **Johnson K** (2016) Energy-efficient buildings for low-carbon cities. *ICCG Reflection*. *47*, 1–9

Qin Y, **Ke J**, **Wang B and Filaretov GF** (2022) Energy optimization for regional buildings based on distributed reinforcement learning. *Sustainable Cities and Society 78*, 103625. https://doi.org/10.1016/j.scs.2021.103625.

Rathnayaka B, **Robert D**, **Siriwardana C**, **Adikariwattage VV**, **Pasindu HR**, **Setunge S and Amaratunga D** (2023) Identifying and prioritizing climate change adaptation measures in the context of electricity, transportation and water infrastructure: A case study. *International Journal of Disaster Risk Reduction 99*, 104093. https://doi.org/10.1016/J.IJDRR.2023.104093.

Rebelatto B, **Salvia A**, **Bueno PT**, **Brandli L**, **Rodrigues G** (2023) Energy Efficiency in Commercial Buildings: Systematic Analysis of the Connection with Sustainable Development Goals and Climate Change. Available at https://repositorio.ufsc.br/handle/123456789/247101 (accessed 21 April 2025).

Rodrigues E, **Fereidani NA**, **Fernandes MS and Gaspar AR** (2023) Climate change and ideal thermal transmittance of residential buildings in Iran. *Journal of Building Engineering 74*, 106919. https://doi.org/10.1016/j.jobe.2023.106919.

Sahebzadeh S, **Heidari A**, **Kamelnia H**, **Baghbani A** (2017) Sustainability features of Iran's vernacular architecture: A comparative study between the architecture of hot–arid and hot–arid–windy regions, *Sustainability 9*, 749. https://doi.org/10.3390/SU9050749.

Said Z, **Rahman SMA**, **Sohail MA and Bibin BS** (2023) Analysis of thermophysical properties and performance of nanorefrigerants and nanolubricant-refrigerant mixtures in refrigeration systems. *Case Studies in Thermal Engineering 49*, 103274. https://doi.org/10.1016/J.CSITE.2023.103274.

Sarker IH (2021) Deep learning: A comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science 2*, 420. https://doi.org/10.1007/s42979-021-00815-1.

Seppänen O (2008) Ventilation strategies for good indoor air quality and energy efficiency. *International Journal of Ventilation 6*(4), 297–306. https://doi.org/10.1080/14733315.2008.11683785

Sharif AN, **Saleh SK**, **Afzal S**, **Razavi NS**, **Nasab MF and Kadaei S** (2022) Evaluating and identifying climatic design features in traditional Iranian architecture for energy saving (case study of residential architecture in northwest of Iran). *Complexity 2022*, 3522883, 12. https://doi.org/10.1155/2022/3522883.

Silvestri A, **Coraci D**, **Brandi S**, **Capozzoli A**, **Borkowski E**, **Köhler J**, **Wu D**, **Zeilinger MN and Schlueter A** (2024) Real building implementation of a deep reinforcement learning controller to enhance energy efficiency and indoor temperature control. *Applied Energy 368*, 123447. https://doi.org/10.1016/J.APENERGY.2024.123447.

Sodiq A, **Khan MA**, **Naas M and Amhamed A** (2021) Addressing COVID-19 contagion through the HVAC systems by reviewing indoor airborne nature of infectious microbes: Will an innovative air recirculation concept provide a practical solution? *Environmental Research 199*, 111329.

Sohani A, **Rezapour S and Sayyaadi H** (2021) Comprehensive performance evaluation and demands' sensitivity analysis of different optimum sizing strategies for a combined cooling, heating, and power system. *Journal of Cleaner Production 279*, 123225. https://doi.org/10.1016/J.JCLEPRO.2020.123225.

Stoffel P, **Maier L**, **Kümpel A**, **Schreiber T and Müller D** (2023) Evaluation of advanced control strategies for building energy systems. *Energy and Buildings 280*, 112709. https://doi.org/10.1016/J.ENBUILD.2022.112709.

Symonds P, **Verschoor N**, **Chalabi Z**, **Taylor J and Davies M** (2021) Home energy efficiency and subjective health in greater London. *Journal of Urban Health 98*, 362–374. https://doi.org/10.1007/S11524-021-00513-6/FIGURES/5.

Ugli KKB (2022) Improving the energy efficiency of low-rise residential buildings. *International Journal of Advance Scientific Research 2*, 24–31. https://doi.org/10.37547/IJASR-02-10-05.

Ürge-Vorsatz D, **Cabeza LF**, **Serrano S**, **Barreneche C and Petrichenko K** (2015) Heating and cooling energy trends and drivers in buildings. *Renewable and Sustainable Energy Reviews 41*, 85–98. https://doi.org/10.1016/j.rser.2014.08.039.

van Roosmalen M, **Herrmann A and Kumar A** (2021) A review of prefabricated self-sufficient facades with integrated decentralised HVAC and renewable energy generation and storage. *Energy and Buildings 248*, 111107. https://doi.org/10.1016/J.ENBUILD.2021.111107.

Vázquez-Canteli J, **Kämpf J and Nagy Z** (2017) Balancing comfort and energy consumption of a heat pump using batch reinforcement learning with fitted Q-iteration. *Energy Procedia*. https://doi.org/10.1016/j.egypro.2017.07.429.

Villaizán-Vallelado M, **Salvatori M**, **Carro B and Sanchez-Esguevillas AJ** (2024) Graph neural network contextual embedding for deep learning on tabular data. *Neural Networks 173*, 106180. https://doi.org/10.1016/j.neunet.2024.106180.

Wallner P, **Tappler P**, **Munoz U**, **Damberger B**, **Wanka A**, **Kundi M**, **Hutter HP** (2017) Health and wellbeing of occupants in highly energy efficient buildings: A field study, *International Journal of Environmental Research and Public Health 14*, 314. https://doi.org/10.3390/IJERPH14030314.

Wang J, **Dai Y**, **Gao L and Ma S** (2009) A new combined cooling, heating and power system driven by solar energy. *Renewable Energy 34*(12), 2780–2788. https://doi.org/10.1016/J.RENENE.2009.06.010.

Wang J, **Dai Y**, **Gao L**, **Shaolin M** (2009) A new combined cooling, heating and power system driven by solar energy. *Renewable Energy 34*(12), 2780–2788. https://doi.org/10.1016/j.renene.2009.06.010.

Wang Z, **Shen H**, **Deng G**, **Liu X and Wang D** (2024) Measured performance of energy efficiency measures for zero--energy retrofitting in residential buildings. *Journal of Building Engineering 91*, 109545. https://doi.org/10.1016/J.JOBE.2024.109545.

Wu L, **Chen Y**, **Shen K**, **Guo X**, **Gao H**, **Li S**, **Pei J and Long B** (2023) Graph neural networks for natural language processing: A survey. *Foundations and Trends in Machine Learning 16*. https://doi.org/10.1561/2200000096.

Xiao T and You F (2023) Building thermal modeling and model predictive control with physically consistent deep learning for decarbonization and energy optimization. *Applied Energy 342*, 121165. https://doi.org/10.1016/J.APENERGY.2023.121165.

Xu S, **Fu Y**, **Wang Y**, **Yang Z**, **Huang C**, **O'Neill Z**, **Wang Z**, **Zhu Q** (2025) Efficient and assured reinforcement learning-based building HVAC control with heterogeneous expert-guided training, *Scientific Reports 15*(1), 1–17. https://doi.org/10.1038/s41598-025-91326-z.

Yan Q, **Qin C** (2017) Environmental and economic benefit analysis of an integrated heating system with geothermal energy—A case study in Xi'an China, *Energies 10*, 2090. https://doi.org/10.3390/EN10122090.

Yan J, **Han W**, **Wei B**, **Liu Y**, **Gao L**, **Wang L**, **Wang S**, **Zhu M and Huo Z** (2025) Drying characteristics of white radish slices under heat pump – Low-temperature regenerative wheel collaborative drying. *Case Studies in Thermal Engineering 69*, 105950, https://doi.org/10.1016/j.csite.2025.105950.

Yang Y, **Luo H**, **Adibhesami MA** (2025) Climate and performance-driven architectural floorplan optimization using deep graph networks. *Engineering, Construction and Architectural Management.* Ahead-of-print. https://doi.org/10.1108/ECAM-08-2024-1107.

Zhang Z, **Chong A**, **Pan Y**, **Zhang C and Lam KP** (2019) Whole building energy model for HVAC optimal control: A practical framework based on deep reinforcement learning. *Energy and Buildings 199*, 472–490.

Zhang W, **Yu Y**, **Yuan Z**, **Tang P and Gao B** (2025) 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. https://doi.org/10.1016/J.ENBUILD.2025.115436.

Zhang J, **Zhang HH**, **He YL and Tao WQ** (2016) A comprehensive review on advances and applications of industrial heat pumps based on the practices in China. *Applied Energy 178*, 800–825. https://doi.org/10.1016/j.apenergy.2016.06.049.

Cite this article: Adibhesami MA and Hassanzadeh A (2025). Optimizing HVAC energy efficiency in low-energy buildings: a comparative analysis of reinforcement learning control strategies under Tehran climate conditions. *Data-Centric Engineering*, 6, e40. doi:10.1017/dce.2025.10014