LONDON'S GLOBAL UNIVERSITY
Department of Computer Science

UCL

Optimising HVAC Control Across Diverse Climates: A Replay-Enhanced Deep Reinforcement Learning Approach

Nathan Carey

Supervisor: John Shawe-Taylor

Co-Supervisors: Francesca Channon, Dhruva Tirumala

Faculty of Engineering

Department of Computer Science

University College London

A Project Report Presented in Partial Fulfillment of the Degree

MSc AI for Sustainable Development

September 2024

Executive Summary

Background

In this thesis, we address the inefficiencies of traditional rule-based control (RBC) in managing heating, ventilation, and air conditioning (HVAC) systems. These systems, which account for up to **40%** of total energy consumption in commercial buildings [1], behave nonlinearly, especially under varying weather conditions. This makes RBCs insufficient for optimisation in a modern, changing environment. We explore the potential of deep reinforcement learning (DRL) to provide more adaptive and efficient control strategies. However, DRL approaches thus far have significant limitations, including a lack of generalisation across different environments, high computational costs [2], and insufficient evaluation in real-world scenarios [3].

Research Question

How can we create a DRL agent capable of efficiently managing HVAC systems across diverse and rapidly changing climate conditions?

Research Contributions

- **Three Climate Experiments:** We developed a novel Three Climate Experiment framework to improve the robustness of DRL agents by training them across diverse climate conditions (seen in Figure 1). We conducted sequential experiments across different climates, allowing the DRL agent to leverage knowledge from previous

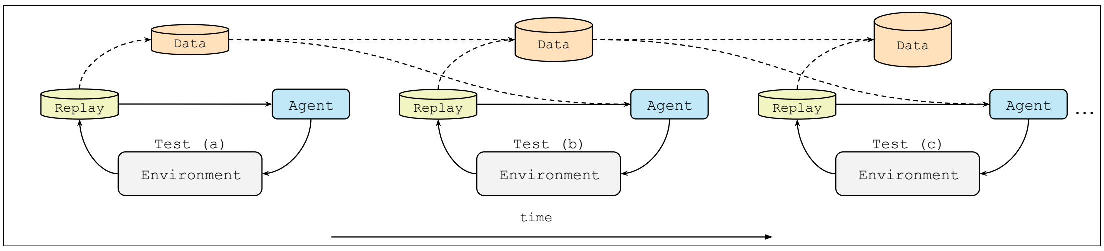


Figure 1: Three Climate Experiment Framework

scenarios. This incorporates the Replay across Experiments (RaE) methodology [4], improving stability and performance in both normal and extreme weather conditions.

- **Dataset Size Comparison:** We compared the agent's performance with two different experience levels: 10,000 and 100,000 sampled transitions, demonstrating how increased experience improves model robustness and effectiveness.

- **Novel Weather Dataset:** We introduced a novel dataset composed of global weather data from seven countries (Appendix Table A.2). Additionally, we modified this dataset to reflect current climate variability, providing a more realistic and challenging evaluation environment.

Methodology

We employed a structured approach, as seen in Figure 2. First, we reproduced and refactored a baseline DRL algorithm for HVAC [5]. Next, we implemented the Three Climate Experiments, training our DRL models across three distinct climate conditions, leveraging the RaE approach to build robustness by incorporating data from previous trials [4]. Performance was evaluated using both standard DRL metrics (e.g., average reward, mean squared error loss) and HVAC-specific metrics (e.g., energy consumption, thermal comfort violation) to ensure comprehensive assessment (Appendix Table A.4) Our novel, global weather dataset was modified to reflect current climate conditions by adding Gaussian noise and data perturbations (Appendix Figure A.3). This dataset was used to compare model robustness to extreme climate variations.

Key Results and Findings

Our research demonstrates that the Three Climate Experiment framework significantly improves upon the robustness and generalisation of traditional DRL agents in HVAC control. As summarized in Table 1, we trained across multiple climate conditions, and we

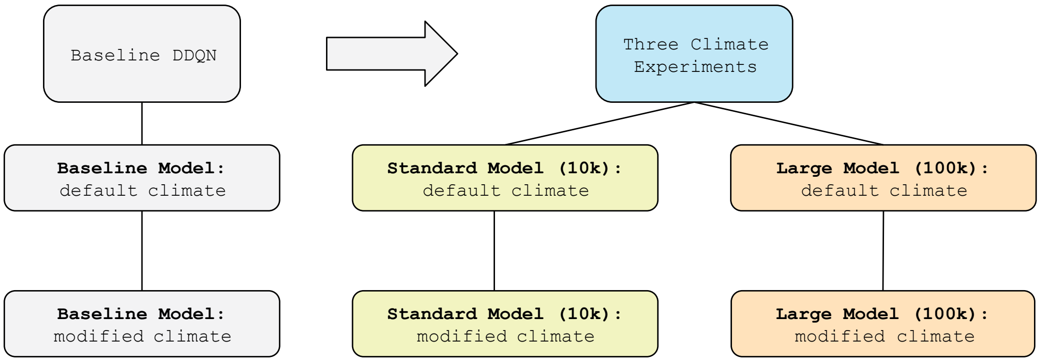


Figure 2: Overview of Experiments

observed up to a **53.02%** improvement in energy efficiency over the baseline approach. The RaE methodology further augmented the models' ability to adapt to new environments by incorporating prior experiences into its learning process, leading to faster convergence and improved performance in both typical and extreme weather scenarios. Our models were further validated with the modified climate dataset to better represent the impacts of climate change, ensuring testing against realistic conditions. In that challenging environment, they outperformed energy-saving baselines by an average of **39.057 percentage points**.

We hypothesise that the significant energy savings are primarily due to the identification of Zone 3 as a key area where optimised control yields the greatest overall efficiency for the entire building (as shown in Figure 3).

We validated our hypotheses with several control-level examples showcasing the relative smoothness of our solution (Figure 3.1 Appendix Figure A.5), and engaged with building managers to explain these results in Section B. These efforts toward better understanding the models' decisions are necessary steps in preparing for the online deployment of these models in the real world, and making an impact on the United Nations Sustainable Development Goals (SDGs).

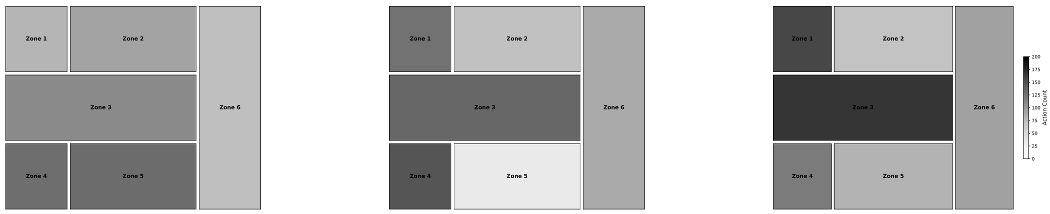
Category	Performance	Description	Impact
Baseline	Reproduced baseline within +/- 6.4% accuracy	Established solid foundation for comparison across models.	Published refactored code as fully open-source
Three Climate (10k)	**13.34%** improvement over baseline (max)	Improvement shown in energy efficiency. However, performance suffers in desert climates (AZ, Dubai).	**162,014 kWh p.a.** energy savings
Three Climate (100k)	**53.02%** improvement over baseline (max)	Significant gains in performance and stability achieved with the larger model.	**378,458 kWh p.a.** energy savings
Modified Weather (10k, 100k)	**54.11%** improvement in challenging climates (max)	The Three Climate Experiment models showed robustness against climate variability. The 100k model outperformed baselines in all cases.	**523,616 kWh p.a.** energy savings

Total Impact: 1.064 million kWh p.a. energy savings

Table 1: Impact Summary of Key Contributions

This thesis contributes to the field of HVAC optimisation by providing a robust, efficient, and explainable DRL framework capable of optimising HVAC systems across diverse climates. The Three Climate Experiment methodology represents significant advancements in the generalisation and efficiency of DRL models for HVAC control.

Future work should focus on extending this research to a real building, with the potential of incorporating real-world factors (such as energy pricing) and exploring more advanced DRL models. Further efforts should be made towards the exhaustive explainability of these models, ensuring that they can be effectively implemented and trusted in real-world applications.



(a) Baseline DDQN (b) Three Climate (c) Three Climate (Large)

Figure 3: Comparison of Action Distributions for Agents in Vancouver (modified climate). Darker Zones Represent More Counts of Enforcing Temperature Control.

In conclusion, we have demonstrated that a replay-enhanced DRL approach can significantly improve HVAC control, achieving substantial energy savings while maintaining occupant comfort. The introduction of the Three Climate Experiment methodology marks a step forward in the development of generalisable DRL models for HVAC optimisation and directly addresses several SDGs, particularly SDG 7 (Affordable and Clean Energy), SDG 11 (Sustainable Cities and Communities), and SDG 13 (Climate Action) [6].