Deep Reinforcement Learning for Heat Pump Control

Tobias Rohrer

Darmstadt University of Applied Sciences

Motivation

Heating in private households accounted for 26% of total energy consumed in Germany in 2020, which is a major contributor to the emissions generated today [1]. Heat pumps are a promising alternative for heat generation and are a key technology in achieving our goals of the German energy transformation which includes the reduction of gas emissions by 55% until 2030, compared to 1990 [2, 4]. A majority of heat pumps in the field today are controlled by a simple heat curve, which is a naive mapping of the current outdoor temperature to a control action. An alternative approach is Model Predictive Control (MPC) which was applied in multiple research works to heat pump control. However, MPC is heavily dependent on the building model, which has several disadvantages. Motivated by this and by recent breakthroughs in the field, this work applied deep reinforcement learning (DRL) to heat pump control in a simulated environment. This work was carried out in collaboration with the Fraunhofer institute for solar energy systems (ISE).

Research Questions

1. Is it possible to apply deep reinforcement learning to learn efficient heat pump control strategies in the simulation provided?
2. How well is deep reinforcement learning working, compared to the two baseline methods MPC and the heating curve?
3. Can the deep reinforcement learning solution be extended to a demand response scenario?

Methods

- The simulation environment was provided by the Fraunhofer ISE.
- The provided simulation was wrapped in an environment which could be used for deep reinforcement learning.
- OpenAI Gym[1]was used in order to support the creation of the environment.
- Proximal policy optimization (PPO) [3] was applied as deep reinforcement learning algorithm.
- Stable Baselines3[2]was used to support implementation of PPO.

Figure 1 shows the concept of how the deep reinforcement learning was applied in the master thesis. The state definition of the Markov decision process, which is used as decision criteria for the agent to choose actions includes (1) the current indoor temperature of the building T_{in}, (2) the current return temperature T_{ret} of the water coming back from the underfloor heating system, and (3) a forecast of outside air temperatures T_{out}.

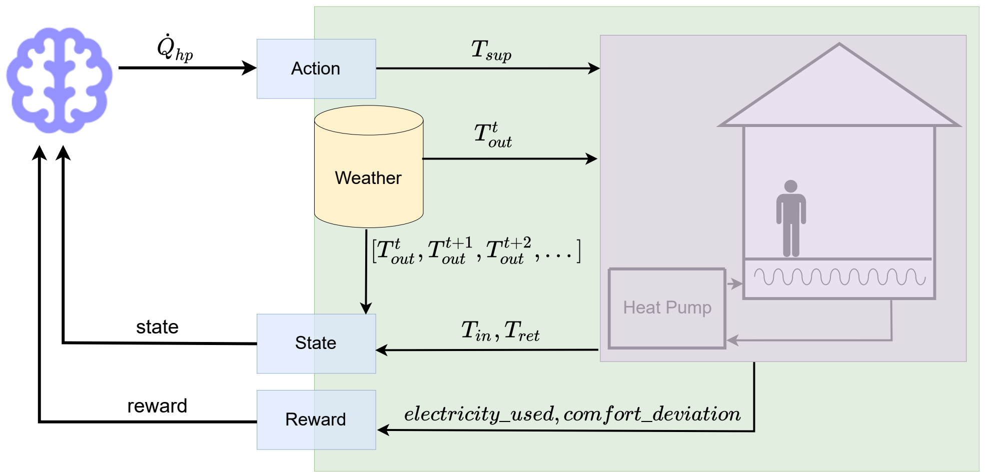

Figure 1. Overview of the applied solution. The agent chooses the thermal power of the heat pump $\dot{Q}_{hp}$ as control action. The deep reinforcement learning environment is denoted in green. It wraps the simulation framework, which is highlighted in purple. The simulation framework returns information about any comfort deviations and the amount of electricity which was used for heating. Both terms are included in the reward definition and are to be minimized to ensure efficient heat pump control. Additionally, weather profiles are used to simulate the influence of the outdoor temperature on the heating process. Note that a forecast of the weather is included in the environments state.

Results

1. Control strategies for heat pumps could be learned that manage keep the indoor temperature in the defined comfort bound (between 21°C and 25°C).
2. By using different buildings in the evaluation, it could be shown that the optimal heating strategies depend on the building used (see figure 2). The heating strategy for the efficient enhanced building exploits the heat storage capabilities while heating when it is relatively warm outside.
3. By a baseline comparison, it could be shown that MPC-like performance (in terms of energy usage and comfort deviations) could be archived, just by learning by trial and error (which is the basic idea of DRL).
4. Additionally, by including time-based variable electricity prices in an experiment, it could be shown that the solution can be easily extended to the scenario, where heating should take plase with respect to varying electricity prices.

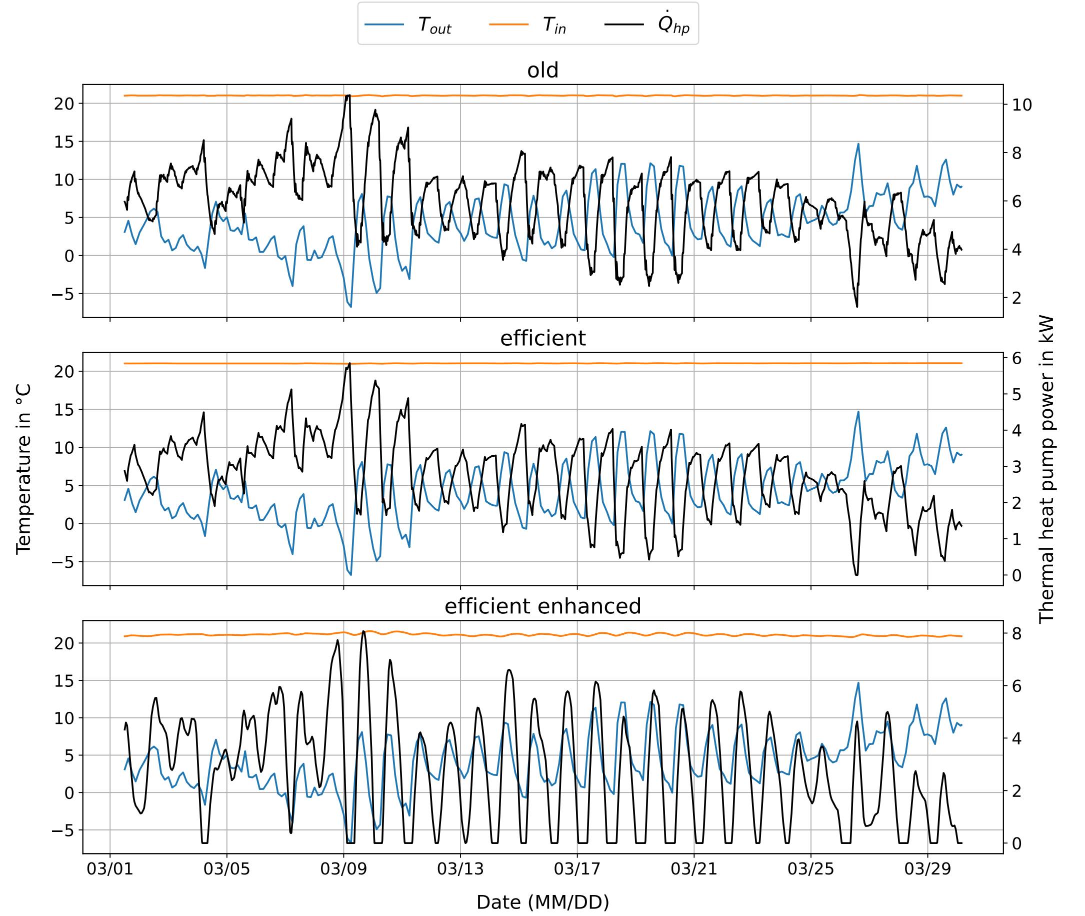

Figure 2. Heat pump control strategies for the three different buildings learned by the DRL agent. The control action choosen by the agent at every time step is shown in black. The resulting indoor room temperature is shown in orange. The outside air temperature is shown in blue. Note that the learned strategic control approach from the efficient enhanced building differs from the other two. The agent learned to exploit the termal heating capacities to shift the heating loads to periods where the heat pump could be operated most efficiently - when its relatively warm outside.

References

[1] AG Energiebilanzen e.V. Anwendungsbilanzen zur energiebilanz deutschland. url: `https://ag-energiebilanzen.de/wp-content/uploads/2020/10/ageb_20v_v1.pdf`, 2021. Accessed: 2022-04-2.

[2] Bundesministerium für Wirtschaft und Klimaschutz. Ziele der energiewende. url: `https://www.bmwi.de/Redaktion/DE/Infografiken/Energie/ziele-der-energiewende.html`. Accessed: 2022-04-2.

[3] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *CoRR*, abs/1707.06347, 2017.

[4] Philip Sterchele, Julian Brandes, Judith Heilig, Daniel Wrede, Christoph Kost, Thomas Schlegl, Andreas Bett, and H-M Henning. Wege zu einem klimaneutralen energiesystem. die deutsche energiewende im kontext gesellschaftlicher verhaltensweisen. 2020.

[1] `https://www.gymlibrary.ml/`
[2] `https://github.com/DLR-RM/stable-baselines3`

`https://github.com/tobirohrer/reinforcement-learning-heat-pump`

tobias.rohrer@outlookl.com