**Reinforcement Learning for Residential Heat Pump Operation**

* For space heating control, this reward function is often designed using the internal room temperature together with comfort bands of the residents.

Comfort bands:  
19-24 for heating  
24-28 for cooling.

acceptance and deployment of the approach into real world systems is low.

* Our work includes the creation of a suitable but simple environment modeling of the heat network including a hot water tank as heat storage but without requiring a building model.
* The HP gets modeled using a temperature dependent coefficient-of-performance (COP) curve.
* The demand predictions are created using a recurrent neural network technique called Long-Short-Term-Memory (LSTM). Five years of simulated space heating demands with a granularity of 15 min are available.
* Heat demand due to domestic hot water is not considered.
* The historical heat demand profiles of a residential district were simulated using the software QuaSi.
* For this work, the python implementation of PPO from Stable-Baselines3 has been used.
* The HP is of type air-to-water and has a nominal electrical power of 100 kW. The electrical power of the HP is continuously adjustable in a range from 0 to 100 kW and is chosen by the RL agent.
* Benchmark against a rule-based model calculated over:

Total energy cost

Energy cost due to energy consumption

Energy cost due to heat loss

Number of on/off state changes of HP

Average operating power of HP

Average SOC of the hot water tank  
maximum SOC of the hot water tank

**An adaptive control framework based on Reinforcement learning to balance energy, comfort and hygiene in heat pump water heating systems**

* Environment modeled in TRNSYS.
* Considers the hygiene of the water. The water into the tank must remain between 65°C to 75°C to prevent the action of some bacteria.

**Deep Reinforcement Learning for Building HVAC Control**

* First to apply DRL for HVAC control (to the best of authors’ knowledge.
* During building operation, it learns an effective control policy based on sensing data input, without relying on any thermal dynamics model.
* For offline training and validation of the algorithm, they leverage detailed building dynamics model built in the EnergyPlus simulation tool. while the EnergyPlus models are highly accurate and suitable for offline training and validation, their high complexity makes them unsuitable for real-time control.
* Based on an MDP, the zone temperature at next time step is only determined by the current system state and environment disturbances, and the conditioned air input from the HVAC system. It is independent from the previous states of the building.
* Each temperature zone is equipped with a VAV (variable air flow volume) HVAC system.
* The reward is negative, the goal is maximizing the cumulative reward, they use the Bellman equation.
* Every ∆tc time, the algorithm will observe the building state and update the control action. Between two control time steps, the control action used to operate the HVAC system remains the same.

**DeepComfort: Energy-Efficient Thermal Comfort Control in Buildings via Reinforcement Learning**

* Deep feedforward neural network with Bayesian regularization for predicting occupants’ thermal comfort.
* They adopt a discrete-time model, where the time is denoted as t = 0, 1, 2,... The duration of each time slot is from several minutes to 1 h.
* The neural network for comfort prediction has 6 inputs:  
  - metabolic rate

- clothing insulation

- air temperature

- mean radiance

- air speed

- humidity

The hidden layer has two layers.

**Deep Reinforcement Learning for Smart Home Energy Management**

* The smart home presents distributed generators as solar panels or wind generators, the ESS could be lead-acid batteries or lithium-ion batteries
* Immagine che contiene Carattere, tipografia, testo, linea

  Descrizione generata automaticamenteESS model: where ηc ∈ (0, 1] and ηd ∈ (0, 1] are the charging and discharging efficiency coefficients, respectively; and ct and dt are the ESS charging power and discharging power, respectively.
* HVAC model: it can adjusts its input power continuously

**Comfort and energy management of multi-zone HVAC system based on Multi-Agent Deep Reinforcement Learning**

* HVAC control model:  
  - air handling unit (AHU) for the whole building  
  - each zone configures a VAV box individually  
  - air flow rate as control input for each zone  
  - cooling control
* Use predictive mean voting index (PMV) defined to predict the average vote of occupants about thermal sensation scale from -3 (very cold) to 3 (extremely hot)
* Each zone is controlled by its own agent and the agents can update via the shared global information and learn from each other at the same time