Heating, Ventilation, and Air Conditioning (HVAC) Systems utilize much energy, accounting for 40% of total building energy use. The temperatures in buildings are commonly held within narrow limits, leading to higher energy use. A field study was performed in two office buildings in Switzerland, and the measurements illustrated that the air temperature was relatively steady for most of the hours; it was higher than that prescribed by the building standards in Switzerland. Thus, designing energy-efficient building thermal control policies to reduce HVAC energy use while maintaining a dynamic indoor environment is essential. Also, such an environment might not be the best for long-term exposure for the occupants. Studies suggest that a dynamic environment may be healthier for the human body. However, it is challenging to implement such a control policy, considering the energy efficiency of the HVAC System, dynamic indoor environment, and thermal acceptability requirements. Optimizing the dynamic and stochastic energy demand using conventional control techniques is tricky. The challenge becomes even more complicated when the requirements concerning the dynamic indoor environment and thermal acceptability of occupants are introduced. Thus, numerous factors influence the control policy in the built environment, i.e., indoor temperature, indoor relative humidity, indoor airspeed, occupancy profile, occupants' metabolic rate, clothing, comfort sensation, and energy expenditure. Implementing such a policy is challenging and usually hard to model as they may differ from case to case. We propose a deep RL-based framework for energy optimization and healthy thermal environment control in buildings to tackle this complexity. As an emerging control technique, deep Reinforcement Learning (DRL) has attracted growing research interest and demonstrated its potential to enhance building performance while addressing some limitations of other advanced control techniques.

A novel deep RL algorithm with experience replay, called deep deterministic policy gradient (DDPG), has performed excellently in many continuous control tasks. DDPG is an RL approach for continuous control problems. In building controls, temperature, humidity, and airspeed, which are the predominant control variables, are all continuous. Therefore, DDPG is very suitable for addressing the problem in this scenario. Compared with other commonly used methods, such as Q-Learning and Deep Q-Learning, DDPG can avoid discretizing the control variables (e.g., temperature, humidity), which can improve control precision. It is a novel approach to designing the control of the indoor environment. The goal is to design an intelligent thermostat that can accurately control the indoor temperature based on data recorded from the indoor environment and the human body. To this objective, a prediction model for the energy expenditure in the human body from other more easily measurable physical and environmental parameters such as heart rate, muscular electrical activity, stress level, activity level, skin temperature, core body temperature, skin conductance, and ambient temperature has been developed. The prediction models leveraged the machine learning algorithms, particularly the long short-term memory (LSTM) networks. The results show that the models developed provide a good level of prediction accuracy during both low and medium-intensity activities, with the MAPE mostly lying in the range of 5-20%.

The building thermal control has been developed as a cost-minimization problem, and DDPG is the primary choice for training the thermal control policy to create the DIET Controller. The DIET Controller modelled in Python was initially trained in a simulation environment with energy plus using the functional mockup unit (FMU) interface. Co-simulation results show that the DIET Controller can reduce the HVAC energy use by about 40% compared to the conventional rule-based controller and facilitate the creation of a dynamic environment promoting increased occupants' exposure to the temperature conditions between 18-21 °C. Subsequently, the DIET Controller was tested in real-operation with the ICE climate chamber. The experimental setup consisted of a single zone, which was set up with multiple environmental sensors gathering real-time data for the DIET Controller input. The experiments were usually conducted over 24 hours with thermal dummies simulating the heat gains from occupants. Integrating the DRL-based control framework with the existing HVAC system was quite challenging, which the BACnet protocol facilitated in our case. Results showed that in actual operation, the DIET Controller could reduce energy use by 28-64% compared to a rule-based control. Additionally, DIET Controller created a dynamic indoor environment for 96% of occupied hours. The results provide evidence that DIET Controller can be an effective method for controlling systems in real-world operation. However, it should be noted that the results are specific to the system and the setup used in the experiments, and the controller might need to be retrained and tuned for different systems and operating conditions.