

Reinforcement learning for the occupant-centric operation of residential energy system

Copyright ©2022 by the authors. This conference paper is published under a CC-BY-4.0 license.



AMIRREZA HEIDARI

School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland
amirreza.heidari@epfl.ch



FRANÇOIS MARÉCHAL

School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland



DOLAANA KHOVALYG

School of Architecture, Civil and Environmental Engineering (ENAC), Ecole Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland

Abstract: This study proposes a Reinforcement Learning control framework that learns how to optimally control the building energy system by taking into account stochastic occupant behaviour and PV power production. Evaluations using real-world hot water use and weather data indicate an energy-saving potential of 7% to 22% compared to common practice.

Keywords: Reinforcement Learning, Space heating, Hot water, Occupant behaviour, Solar energy, Machine Learning, Occupant-centric, Control, Building, Temperature

Introduction

Optimal operation of building energy systems is challenging as there are several stochastic and time-varying parameters that affect building energy use. One of these parameters is occupant behaviour, which is highly stochastic, can change from day to day, and therefore is very hard to predict [1]. The occupant behaviour of each building is unique, and thus there is no universal model which can be embedded in the control system of various buildings at their design phase. To cope with this highly stochastic parameter, current control approaches are usually too conservative to ensure the comfort of occupants regardless of their behaviour. An example is hot water production, where huge volume of hot water with high temperature is produced in advance and stored in a tank to make sure enough hot water is available whenever it is demanded [2,3].

Another stochastic parameter affecting building operation is renewable energy. The share of renewable energy in the building sector is increasing, and is expected to get doubled by 2030 [4]. Due to the volatile nature of renewable energy sources, it will also increase the complexity of optimal energy management in buildings [5]. There are several other stochastic parameters, such as weather condition or electric vehicles charging that all affect the building energy use. The control logic of buildings should properly consider these stochastic parameters to guarantee an optimal operation.

Uniqueness of occupant behaviour in each building makes it challenging to program a rule-based or model-based control logic that can be easily transferred to many other buildings. Rather than hard programming a rule-based or model-based control method, a learning ability

can be provided to the controller such that it can learn and adapt to the specifications of that building and maintain an optimal operation. Reinforcement Learning (RL) is a method of Machine Learning that can provide this learning ability to the controller. RL can continuously learn and adapt to the changes in system such as varying weather conditions, volatile renewable energy, or stochastic occupants' behaviour [9].

The aim of this research is to develop a self-learning control framework that considers the stochastic hot water use behaviour of occupants, and varying solar power production, and learns how to optimally operate the system to minimize energy usage while preserving the comfort and hygiene aspects. Case study energy system is the combined space heating and hot water production, assisted by Photovoltaic (PV) panels.

The main novelties of the proposed framework are:

Integration of water hygiene: While the previous study by authors [2] followed a simple rule to respect hygiene aspect, this study for the first time integrates a temperature-based model that estimates the concentration of Legionella in hot water tank at each time step. Estimation of Legionella concentration in real-time enables the agent to spend as minimum energy as required for maintaining the hygiene aspect.

Investigation on real-world hot water use behaviour: In this research, hot water demand of 3 residential houses is monitored to assess the performance of agent on real-world hot water use behaviour of occupants.

Stochastic off-site training to ensure occupants comfort and health: To ensure that agent would quickly learn the optimal behaviour with a minimum risk of violating comfort and hygiene aspects, an off-site training phase is designed in this study. This off-site phase integrates a stochastic hot water use model to emulate the realistic occupants' behaviour. Also, it includes a variety of climatic conditions and system sizes to provide a comprehensive experience to the agent.

The remainder of this paper is organized into four sections: Section 2 describes the research methodology, section 3 presents the results, and Section 4 concludes the paper.

Methodology

Figure 1 shows the components of an RL framework. The methodology section describes how each of these components are designed.

Environment design

Layout of residential energy system in this study is shown in Figure 2. This system uses an air-source heat pump to provide hot water in a tank, which is used for both hot water production and space heating through radiators. PV panels are also connected to the heat pump. PV panels are grid-connected, so the surplus power can be supplied to the grid. A dynamic model of the system is developed in TRNSYS.

Agent design

The agent is developed in Python using Tensorforce library [19]. An improved version of Deep Q-Network (DQN), known as Double DQN is used as it is proved to solve the issue of overestimation by typical DQN. Specifications of agent are provided in Table 1.

State, action and reward space

Parameters included in the state are presented in Table 2. Each parameter is a vector including the value

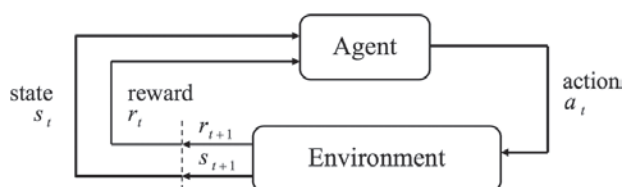


Figure 1. Interaction of agent and environment in Reinforcement Learning. [18]

Table 1. Selected parameters for the agent.

Parameter	Value
Learning rate	0.003
Batch size	24
Update frequency	4
Memory	48x168
Discount factor	0.9

Table 2. Parameters included in the state vector.

Parameter	Length of look-back vector
Hot water demand	6
Demand ratio	-
Outdoor air temperature (°C)	1
Indoor air temperature (°C)	3
PV power (kW)	6
Heat pump outlet temperature (°C)	1
Legionella concentration (CFU/L)	1
Tank temperature (°C)	1
Hour of day	-
Day of week	-

of that parameter during one or multiple previous hours. The demand ratio is the ratio of total hot water demand of the current day until the current hour, to the total demand of the previous day. Hour of day is a value between 1-24 indicating what is the upcoming hour of day. Day of week, similarly, indicates the current day as a value between 1-7, where 1 represents Monday. The values are normalized to a value between 0 to 1.

The agent has four possible actions: Turning ON the heat pump, Turning OFF the heat pump, selecting the indoor air temperature setpoint of 21°C (as an energy-saving setpoint) or 23°C (as an energy-storing setpoint). Based on the selected indoor air temperature setpoint by the agent, a two-point controller with a dead-band of 2°C tries to maintain the specified setpoint during the next hour.

Reward function includes 4 different terms. An energy term to penalize the agent for net energy use, hot water comfort term to penalize the agent if a hot water demand is supplied with a temperature less than 40°C, which is considered as the lower limit of comfort

for hot water uses [2], space heating comfort term to penalize the agent if the indoor air temperature is out of the comfort region of 20°C-24°C, and a hygiene term if the estimated concentration of Legionella is above the maximum threshold of 500×10^3 CFU/L recommended for residential houses [20]. Equations 1-4 shows the formulation of energy, hot water comfort, space heating comfort, and hygiene terms.

$$R_{energy} = -a \times |HP_{power} - PV_{power}| \quad (1)$$

$$\text{if } T_{tank} \geq 40: R_{DHWcomfort} = 0 \text{ else } -b \quad (2)$$

$$\text{if } 20 \leq T_{indoor} \leq 24: R_{Indoorcomfort} = 0 \text{ else } -c \quad (3)$$

$$\text{if } Conc \leq Conc_{max}, R_{Hygiene} = 0 \text{ else } -d \quad (4)$$

Where HP_{power} and PV_{power} are the power use of heat pump and power production of PV panels (kW), T_{tank} and T_{indoor} are the tank and indoor air temperature, $Conc$ and $Conc_{max}$ are the current and maximum concentration of Legionella in the tank (CFU/L), R_{energy} , $R_{DHWcomfort}$, $R_{Indoorcomfort}$ and $R_{Hygiene}$. a, b, c and d are set to 1, 12, 10 and 10 determined by a sensitivity analysis. The total reward is therefore the summation of all these terms.

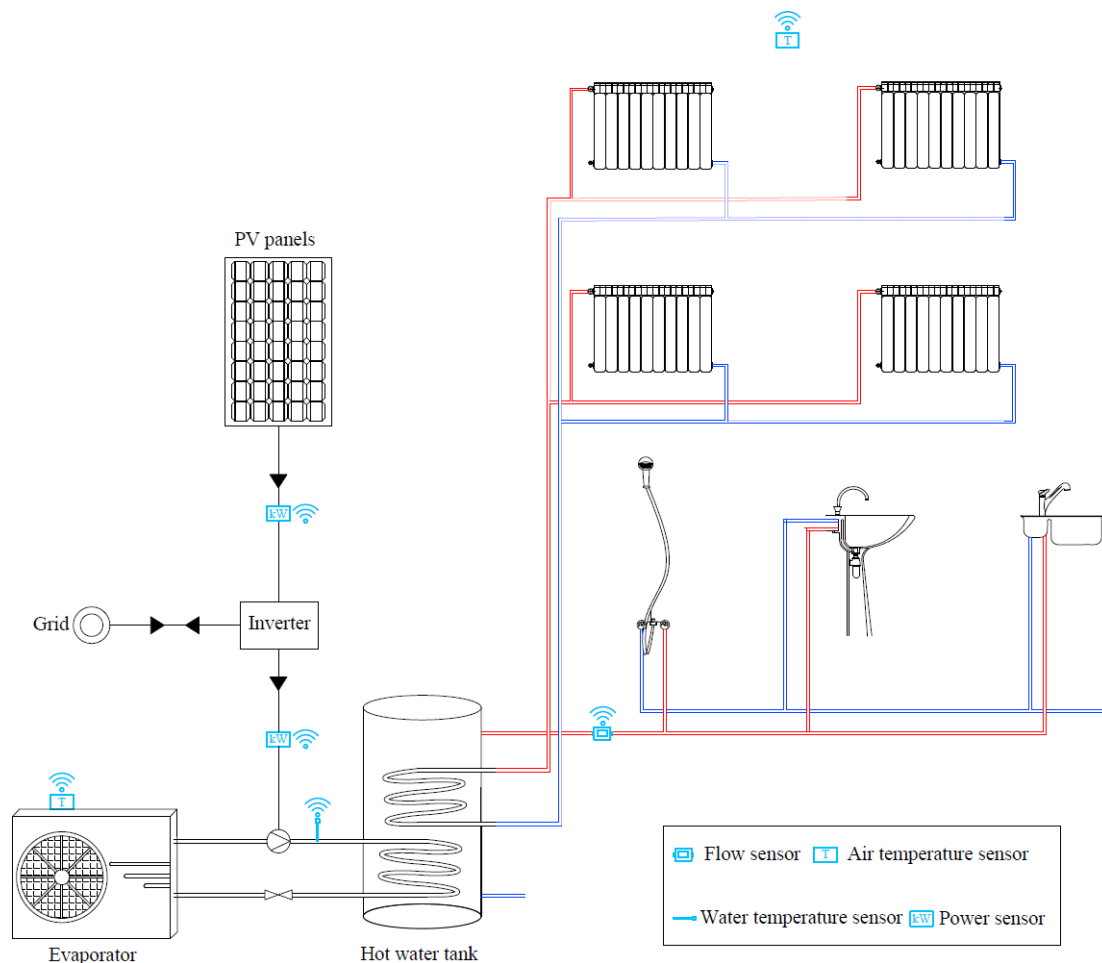


Figure 2. Layout of solar-assisted space heating and hot water production system.

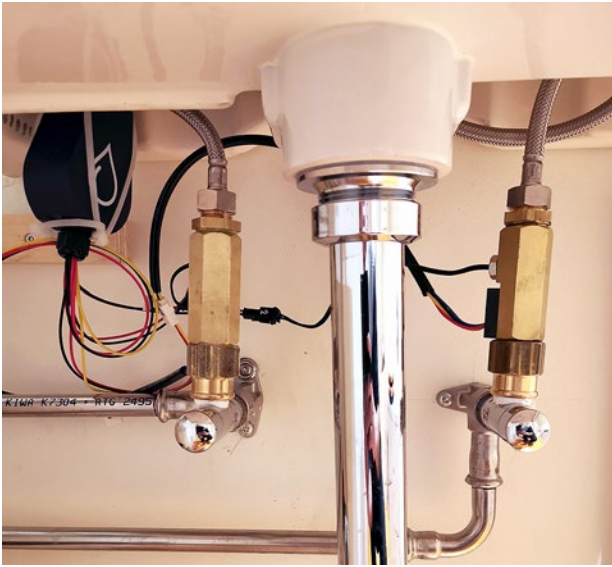


Figure 3. Flow and temperature sensor on a faucet.

Monitoring campaign

To perform a realistic test without disturbing the occupants, hot water use behaviour of people was monitored, and the collected data were used in TRNSYS simulation. For the current framework, as shown in Figure 2, only one single sensor at the tank outlet is enough to measure the hot water demand. In this study, to collect a comprehensive dataset, the hot and cold-water demand was monitored at all the end uses as shown in Figure 3.

Training procedure

Training and deployment stages are shown in Figure 4. To ensure occupants' comfort and health, first, the agent is trained on an off-site training process. In this stage, a virtual environment is provided to enable the agent to gain enough experience before being implemented on the target house. In this stage, a hot water use model [21] is used to emulate the hot water use

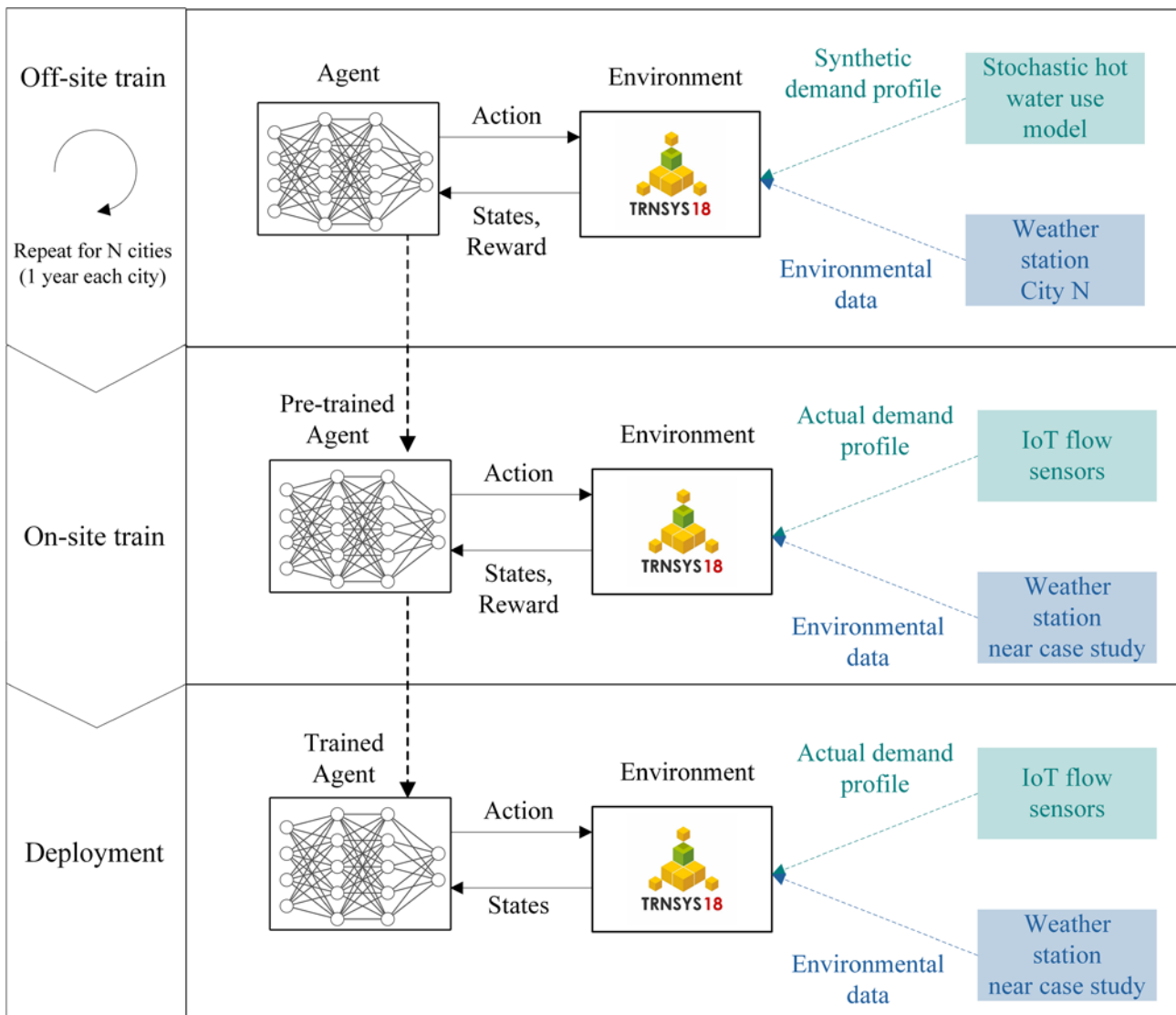


Figure 4. Training and deployment process.

behaviour of occupants and the agent is trained for 10 years. Next, the agent is trained on the target house for 16 weeks. The aim of training on the target house is to let the agent adapt to the specific characteristics of the target house, such as occupants' behaviour, systems sizes, or weather conditions. To simulate the target house, in on-site training stage the collected hot water use data, and also the weather data collected from a weather station near the case study is used. After the on-site training on the target house, the training process can be stopped and the agent starts the deployment stage, in which agent is no longer learning but only controlling the system. Duration of deployment phase is 4 weeks.

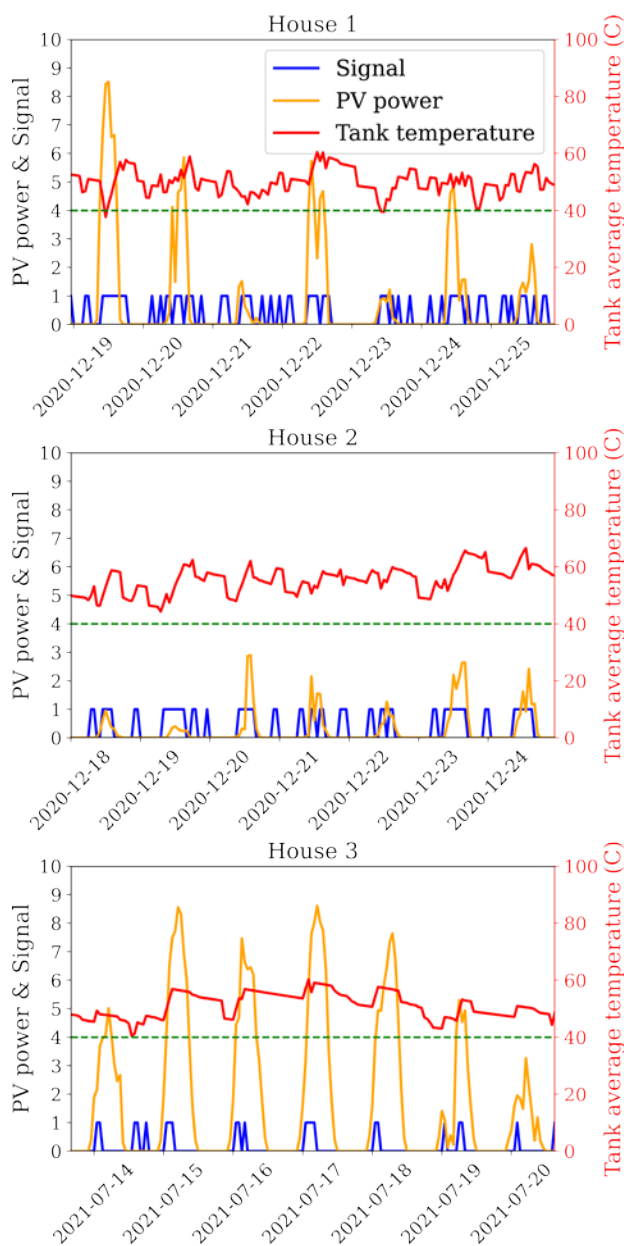


Figure 5. Control signal versus PV power production and tank temperature.

Results

Figure 5 shows the control signal, PV power production, and tank temperature over the deployment stage on three case studies. The deployment stage of houses 1 and 2 is during December, while the deployment stage of house 3 is during July. Therefore, the PV power production of the third case study is higher than others. In all of the case studies, it can be seen that the agent is trying to adapt the control signal to the PV power production and reduce the power use from the grid, by turning ON the heat pump more frequently during the hours of PV power production. This adaptation can be seen very well on house 3, where PV power production is significantly higher and the agent tries to turn ON heat pump only when there is a PV power production. In all of the case studies, the agent has learned how to keep tank temperature above 40°C to respect the comfort of occupants. It shows that agent could successfully learn and adapt to the occupants' behaviour, because none of the demands reduced the tank temperature below 40°C.

To better highlight how RL could better exploit solar power production, the contribution of PV power production in the total power use of the heat pump is shown in Figure 6. Two baseline scenarios are also

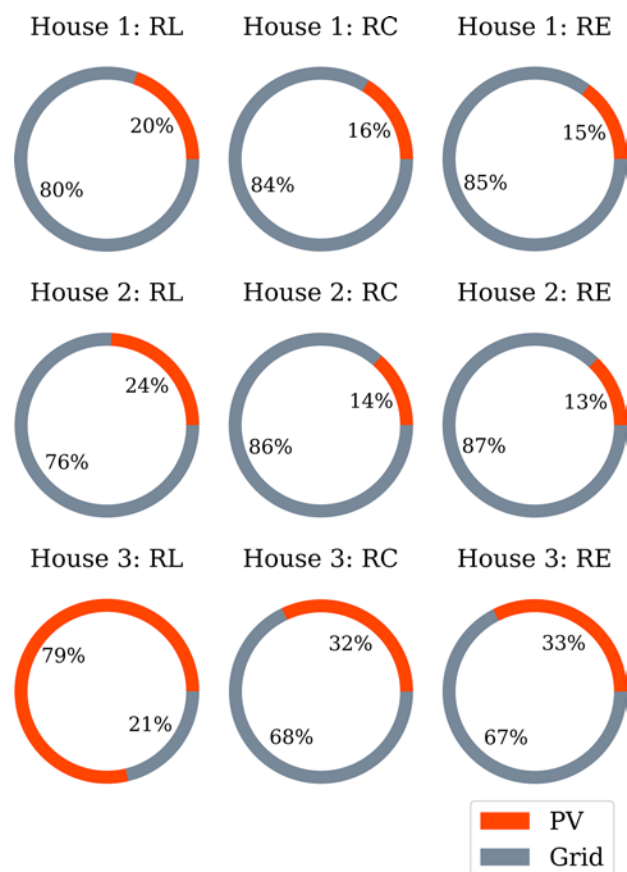


Figure 6. Contribution of PV power production in power consumption of heat pump.

modelled including RC (Conventional rule-based with 60°C setpoint for tank temperature and 22°C setpoint for indoor air) and RE (Energy-saving rule-based with 50°C setpoint for tank temperature and 22°C setpoint for indoor air). As can be seen, in all the case studies RL has used a higher contribution of PV power, compared to the RC and RE. In case of house 3, the contribution of PV power production is much higher than RC and RE, which is the why in this house the energy saving is much higher than other houses. It shows that a significant advantage of the proposed RL framework is to learn how to adapt the operation to the PV power production, and therefore potential energy-saving increases in regions with higher solar radiation.

RL has provided an energy saving of 7% to 22% compared to the RE framework.

Conclusion

This research proposed a model-free RL control framework that can learn the hot water use behaviour of occupants and PV power production, and accordingly adapt the system operation to meet the comfort requirements with minimum energy use. Different from previous studies, where RL is supposed to make a balance between energy use and comfort, in this study

RL tries to make a balance between energy use, comfort, and hygiene. Inclusion of hygiene aspect is very crucial to ensure the health of occupants. Real-world hot water use data is monitored in three residential case studies and used to evaluate the performance of the proposed framework over the realistic behaviour of occupants. The RL framework is compared with two rule-based scenarios of RC and RE.

Results indicate the proposed framework could provide a significant energy saving, mainly by learning how to get the best use of PV power production. Therefore, the energy-saving potential is expected to be even more in regions with higher solar radiation than Switzerland. Also, the agent has successfully learned how to respect the comfort of occupants and water hygiene, so the potential energy saving is not with the cost of violating occupants’ comfort or health. ■

References

Please find the complete list of references in the original article available at <https://proceedings.open.tudelft.nl/clima2022/article/view/286>

SAVE THE DATE



May 2023

mon	tue	wed	thu	fri	sat	sun
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

The 2023 Annual Meeting will take place on 11 and 12 May 2023 in Brussels, BELGIUM