Data-driven energy management of a neighborhood with renewable energy sources

Yunbo Yang[1], Johannes Brozovsky[2], Peng Liu[2], Francesco Goia[1]
[1]Department of Architecture and Technology, Norwegian University of Science and Technology, Trondheim, Norway
[2]Architecture, Materials and Structures, SINTEF Community, Trondheim, Norway

Abstract

Reinforcement Learning (RL) can be implemented for building energy coordination in an open source OpenAI Gym environment called *CityLearn*. *CityLearn* has been implemented in several studies in hot and humid climates, yet barely studied in the climatic context where heating demand is dominant. The goal of this study is to evaluate the feasibility of using CityLearn in a cold climate context with various energy systems and to identify any potential barriers or limitations. To achieve this, we conducted a case study investigating the RL-based coordinated control of the energy storage system in a neighborhood in Norway with different energy systems by expanding the functions of *CityLearn*. The results show that the RL controller outperforms the rule-based control (RBC) scenario and net balance scenario. These demented the adaptivity of the RL controller in CityLearn under different climates and energy system configuration. Additionally, this study showed the approach to processing poor-quality measurement data for RL control implementation.

Highlights

- Coordinated RL energy management of a campus
- Real-world data processing for *CityLearn* implementation
- Function extension of *CityLearn* for the Nordic climate and multiple renewable energy systems
- RL control increased renewable energy penetration and reduced peak demand

Introduction

The coordination of the energy systems among buildings may play a key to unlocking energy flexibility and increasing renewable energy (RE) penetration, which is essential to achieve sustainable goals. The energy crisis during the winter of 2022 - 2023 in Europe further addresses the importance of building energy management both economically and environmentally. Energy management requires control of energy supply, storage, and demand agents. The control of the energy systems in the neighborhoods can be a very complex task as the system is characterized by a high degree of interdependence between the agents that compose it (Authors, 2022). Furthermore, there are inherent conflicts of interest between building and neighborhood-level management (Hu et al., 2021), as building owners' interests lie in minimizing energy costs, which can result in peaks during low-price periods. In contrast, the neighborhoods aim to flatten the energy profile for grid stability. With the emergence of small-scale electric and thermal energy generation and storage, the prosumers can also contribute to a more flexible and efficient energy system by taking actions in demand-side management which require advanced control methods.

Two advanced control techniques are addressed in (Goy & Sancho-Tomás, 2019): agent-based machine learning (ML) algorithms and Model Predictive Control (MPC). ML methods are data-driven and applied to achieving price-incentive-based demand response. MPC requires control model(s) and is mostly applied in the context of thermal load management. Thanks to the development of communication technology, the internet of things (IoT), and data acquisition and storage techniques are bringing more and more data from the building and energy automation system. This offers the opportunity to employ ML methods for multi-agent control. Among ML algorithms, deep reinforcement learning (DRL) stands out for neighborhood-level energy control optimization owing to its model-free nature and adaptivity. DRL learns the optimal policies by trial-and-error iteration with the environment and exhibits its flexibility in terms of changes in the context. Touzani et al. applied the deep reinforcement learning approach to optimal control of the HVAC system and on-site PV generation with an electric battery storage system by using EnergyPlus, Modelica, Pytorch, and PyFMI in a Python environment, and tested on a physical system. The results showed that the DRL controller can produce a cost-saving of up to 39.6% compared to the baseline controller. Wang et al. proposed a fully distributed energy trading framework based on DRL for energy trade optimization via coordinating the operation of distributed renewable resources and flexible loads (Wang et al., 2021). In (Pinto, Deltetto, et al., 2021), the researchers implemented long short time memory (LSTM) to predict the indoor temperature and centralized DRL (soft actor-critic algorithm) to manage the operation of heat pump and domestic hot water (DHW) storage of a cluster of four buildings. Results show that the peak energy demand decreased by 23%.

There are several tools/frameworks that facilitate the implementation of RL control in building energy systems, for example, BOPTEST (Blum et al., 2021) and

CityLearn (Jose R Vazquez-Canteli, 2020). *CityLearn* is an open-source OpenAI Gym environment tailored to investigate the potential of artificial intelligence and distributed control systems to tackle multiple problems within the energy domain. *CityLearn* simulation environment was used by (Pinto, Piscitelli, et al., 2021) to train and evaluate RL models for demand response with hourly time-step. It serves as a useful tool and testbed to implement RL control and compare among control strategies. However, to the best of the authors' knowledge, the up-to-date *CityLearn* applications are in the context of hot and humid climates, where the energy demand is dominated by space cooling. Furthermore, the multi-energy systems in the previous studies using *CityLearn* are limited to heat pumps, electric heaters and PVs. A showcase of the implementation of DRL control with different heating-dominated energy systems using *CityLearn* is lacking. The aim of this study is to address the gaps in existing research by conducting the following activities:

- expanding the features of CityLearn to suit the Nordic climate with more energy systems such as biomass Combined Heat and Power (bio-CHP) plant and a biomass boiler.
- using real-life data of poor quality and demonstrated the possibility of using ML techniques for data imputation for a successful RL controller implementation.

In other words, this study presents a case study of applying coordinated RL control to increase the self-sufficiency ratio of a campus by manipulating the State of Charge (SoC) of the energy storage systems.

Methodology

This section outlines. 1), The configuration of CityLearn and the mechanism of the Soft Actor-Critic (SAC) controller as a type of RL algorithm. 2), The adaption of the CityLearn control framework to the case study. 3), The description of the studied neighborhood and the preparation for the input data, including measurement data pre-possessing, data imputation, and energy demand simulation. 4), Comparing scenarios and performance analysis. The workflow of this study is shown in Figure 1. The training of the RL controller is done offline prior to the deployment. The control optimization is conducted on a quadcore Intel Core i5-1145G7 with 16 GB RAM.

1) *CityLearn* and Reinforcement Learning (RL) control

CityLearn is developed for easy implementation of Multi-Agent Reinforcement Learning (RL) for building energy coordination and demand response in cities (Jose R Vazquez-Canteli, 2020; Vázquez-Canteli et al., 2019). It aims to flatten the overall electric profile by manipulating the charging and discharging of the active energy storage system, such as the battery and water tank.
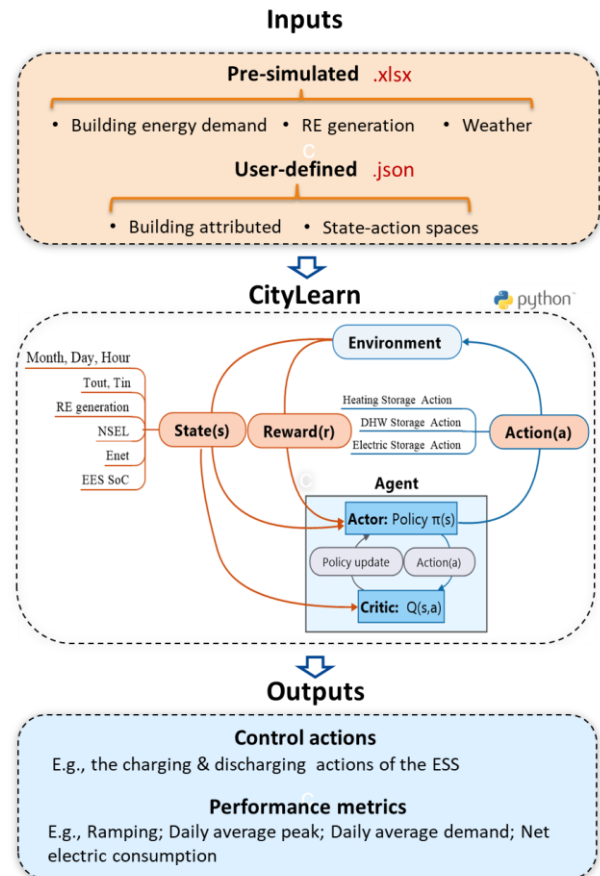

Figure 1: The workflow of this study.

The corresponding control signals are made by the RL controller, either centralized or distributed. RL is one of the advanced control techniques and gaining popularity due to its adaptability and model-free potential (Vázquez-Canteli & Nagy, 2019). RL can be formalized using a Markov Decision Process (MDP), which provides a mathematical framework for modeling decision-making in a situation where the outcomes are partially random and partially under the control of the agent. The MDP process is represented by 4-tuple: (S, A, P, R).

S denotes the States space which describes the environment completely; A denotes the Action space that contains the decision made by the agent; P denotes the transition probability $P: S \times A \times S'$, which depicts the environment's dynamics. It is the probability of ending in state S' at time t+1 when starting in state S and performing action a at the time t. The transition probability can be expressed as $P(S_{t+1} = S' | S_t = S, a_t = a)$.

R denotes the Reward Function $R: S \times A \times S'$ which is used to map the immediate reward r of ending in state S' at time t+1 when starting in state S and performing action a at the time t.

It combines the advantage of dynamic programming with a trial-and-error approach and uses an agent-based control where the agent learns through iteration with the control environment by maximizing the long-term rewards. The

agent in *CityLearn* employs the algorithm Soft Actor-Critic (SAC) to optimize stochastically. SAC is a successor of Q-learning, which is a model-free and off-policy machine learning algorithm. In the model-free approach, the RL learns these action values directly without building an explicit model for the environment, in other words, without explicitly determining the transition probabilities between the states (Gläscher et al., 2010). Policy is a mapping from the current environment observation to a probability distribution of the actions to be taken. In the off-policy approach, the updates of the optimal policy do not have to follow a specific manner, but by learning from the historical data.

The inputs of the *CityLearn* are the pre-simulated data in hourly resolution, user-defined building attributes, and state-action spaces. The outputs of the *CityLearn* environment are the control actions (decisions) and the performance metrics. The environment of the *CityLearn* reads the hourly building energy demand and energy production from the pre-simulated inputs; loads the building and energy system models which set the building and system attributes and energy balance of the energy demand, supply, and storage; updates the states-action spaces, the states and rewards that can be used by the RL agent.

2) The adaption of the control framework to the case study by expanding the features of CityLearn

CityLearn needs to be adapted to reflect the energy systems and energy distribution of this case study as *CityLearn* is developed for hot humid climates and the energy systems are limited to heat pumps, electric heaters and PVs. The energy distribution and consumption network are programmed in python as the space heating and cooling demand of the buildings is satisfied by a heat pump, while the DHW demand is covered by an electric heater. The heat pump is sized in a way to ensure the highest demand can be satisfied. And the Coefficient of Performance (COP) of the heat pump depends on the outdoor temperature and target heating/cooling temperature.

This energy system configuration does not reflect the situation of our case study, which has only heat demand and is equipped with small-scale bio-CHP and biomass-boiler. The differences between the CityLearn example and the case study are tabulated in Table 1 below.

The environment in *CityLearn* is majorly composed of python classes and self-defined functions (Jose R Vazquez-Canteli, 2022). The equations below describe the dynamic energy balance in the RL(SAC) environment.

$$E_{net,t} = \sum_{i=1}^{6} (\Delta S_{t,i} + D_{SH,t,i} + D_{DHW,t,i} + D_{ESU,t,i}$$
$$- G_{solar,t,i} - G_{CHP_{El},t,i} - G_{CHP_{Ht},t,i}$$
$$- G_{bioboiler,t,i})$$

$$\Delta S_{t,i} = \Delta Q_{SH_{tank},t,i} + \Delta Q_{DHW_{tank},t,i} + \Delta E_{battery,t,i}$$

$$\Delta Q_{SH_{tank},t,i} = a_{SH_{tank},t,i} \times Capacity_{SH_{tank},t,i}$$

$$\Delta Q_{DHW_{tank},t,i} = a_{SH_{tank},t,i} \times Capacity_{DHW_{tank},t,i}$$

$$\Delta E_{battery,t,i} = a_{SH_{tank},t,i} \times Capacity_{battery,t,i}$$

Table 1: The differences between the CityLearn example system and the case study.

	CityLearn	**Case study**
Heating demand	Heat Pump	CHP + Bio-boiler + Electric heater
Cooling demand	Heat Pump	No
DHW	Electric heater	Included in the heating demand
Electrical use	PV + Grid	PV + CHP + Grid

Where, E_{net} denotes the net electric demand from the grid at the timestep t. Positive $E_{net,t}$ means the campus requires electricity from the grid, while the negative $E_{net,t}$ means the campus exports electricity to the grid. $D, G, \Delta S$ denote energy demand, energy generation, and the changes in the EES of the ith building at timestep t, respectively. a denotes the control actions, which are the energy charging or discharging ratio of the SH water tank, DHW water tank and, the battery. a is in the range of [-1,1]. In which, -1 means the storage system is fully discharged while 1 means the storage system is fully charged. The constraints are that the EES cannot be discharged more than their or more than the total demand. $Capacity$ is the capacity of the EES. The capacity of the EES of each building is the capacity in Table 2 allocated by the construction area.

The objective of the optimization is to minimize the amount of electricity needed from the grid, but the mechanism of RL is to maximize the reward function. Therefore, the cube value of negative E_{net} is designed as below:

$$Reward = (-E_{net})^3 \text{ if } E_{net} > 0, \text{ else } Reward = 0$$

3) Overview of the case study and data preparation

Campus Evenstad is a campus located inland in southern Norway with about 20 buildings and a total floor area of about 9 000 m² as shown in Figure 2. Evenstad is Norway's most self-sufficient campus with local RE. The neighborhood is in Nordic weather conditions (average temperature of -9 ℃ in December and 16 ℃ in July), and approx. 50% of the annual demand is heating demand while there is no cooling demand. A local district heating supplies the heat from the energy center to the east side and west side of the campus. Six buildings and one snow-melting system are connected to the local district energy system and are metered. In this study, only these six buildings are investigated because these six buildings cover 91% of the total district heating energy use on the campus referring to the report (E. F. Åse Lekang Sørensen, Harald Taxt Walnum, Kristian Stenerud Skeie, Inger Andresen, 2017).

Figure 2: Aerial view of Campus Evenstad.

This campus demonstrates well the RE technologies, such as CHP plant, PV, and a biomass boiler. The energy demand of the campus is satisfied by two energy carriers: woodchips and electricity. The annual electrical consumption is approx. 910 000 kWh. The woodchips power the operation of the biomass CHP and biomass boiler, and the rest of the energy demand is satisfied by electricity. The operation priority is as the sequence of biomass CHP, biomass boiler, and electric heater. The thermal production from the CHP is supposed to fulfill the base load of the campus. The electricity sources are the on-site generation of PV panels and biomass CHP, and from the grid if the local generation could not cover the electrical demand. The electricity demand consists of electric-specific use (ESU) and the demand from the electric heaters to satisfy remaining space heating (SH) and domestic hot water (DHW) needs. The assumption is that the efficiency of electric heaters is 1. The other flexibility sources are a li-ion battery for electrical storage and storage tanks for hot water. The power or capacity of the energy supply and storage system is tabulated below in Table 2 (E. F. Åse Lekang Sørensen, Harald Taxt Walnum, Kristian Skeie, Inger Andresen, 2017; Stian Backe, 2019)

Table 2: The power or capacity of the energy supply and storage systems of the campus

System	Power / Capacity
CHP	$40 kW_{el}$, $100 kW_{th}$
PV panels	$60 kW_{el}$
Bio-boiler	$350 kW_{th}$
Electric boiler	$315 kW_{th}$
li-ion battery	Charge/discharge at $120 kW_{el}$ with Capacity: $204 kWh_{el}$
Hot water tank	21600 l in total

The inputs of the control framework consist of two parts: pre-simulated energy demand and generation data and user-defined building attributes and state-action spaces. This study intended to use real-life measurement data for the pre-simulated energy demand and generation. However, The real-life measurements are not as perfect as expected and contain a lot of missing data. The measurement data covers the period between October 2017 and July 2019, with hourly resolution and timestamps of 16056. In our study, 76% and 85% of measurement data are missing as regards CHP and biomass boiler generation, respectively. And 17% to 53% of demand data is missing. Data preprocessing is conducted because the successful implementation of RL is highly dependent on the data quality. Data imputation methods are used as a remedy to the missing data and can be divided into two parts, energy demand simulation, and energy production predictive modeling.

- Energy demand simulation

The energy demand simulations in this study are performed with IDA ICE (Niclas Björsell, 1999) v. 4.8. IDA ICE is a commonly used building performance simulation tool in the Nordics and is validated according to several standards. For the energy demand simulations of the Evenstad campus, the six largest buildings connected to the district heating system were modeled according to architectural drawings and building physical information shared by Statsbygg, the public owner of the university campus. For the internal loads, ventilation rates and schedules, the Norwegian standard SN-NSPEK 3031:2021 (Norge, 2020) was used. As a simplification and because the standard requires the load inputs per m² of floor area, the buildings have been modeled with only one thermal zone per floor. However, internal walls have been added as internal masses to the simulation model to account for thermal mass.

In accordance with the location of the building, the coordinates of Campus Evenstad (61.43° N, 11.08° N) are entered into every modeled building in IDA ICE. The weather data was downloaded from https://clima.cbe.berkeley.edu/ (Giovanni Betti, 2022), a website developed and operated by the Center for the Built Environment, University of California Berkeley. An *.epw* file was used for the exact location of Campus Evenstad, called Evenstad Overenget, NOR. The wind profile was chosen as "Open country" according to ASHRAE 1993 and the pressure coefficients are determined by selecting "Exposed" in the IDA ICE dialogue window. The same setups have been applied for all buildings in IDA ICE.

Due to the merging of zones in the IDA ICE model (one zone per floor), modeling every single window on facades that are not shaded by surrounding obstacles does not increase accuracy but only prolongs simulation times because shading and solar gains are calculated for each window individually. Therefore, also the windows are merged façade-wise, where facades are not shaded by others (see also Figure 3). Doors were omitted. The zones are equipped with an "ideal heater", a heating element that keeps the zone always at the heating setpoint temperature (21 °C for university and office buildings).

BS 2023
4-6 SEPT
SHANGHAI

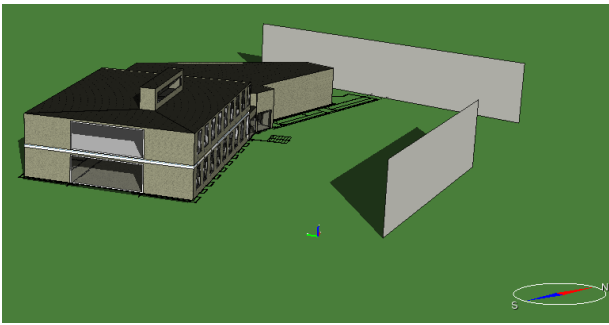International
Building
Performance
Simulation
Association

Figure 3: Building of Evenstad campus modeled in IDA ICE with shading elements representing surrounding buildings.

- Energy production predictive modeling

The predictive modeling is used to impute the missing values of the CHP plant electric production and biomass boiler heat generation. Two attributes with relatively complete data, namely PV electric production (PV_{el}) and total heat demand (Tot_{ht}) were selected as the inputs of the model in order to preserve more data. The PV production data is from the PV model described in (Rognan, 2018). The remaining feature selection of the input attributes is conducted based on domain knowledge. Four attributes were used in addition to PV_{el} and Tot_{ht}, which are the ambient temperature, week of the year, day of the week, and hour of the day. After data cleaning, 3852 timestamps and 2381 timestamps were preserved for predicting CHP electric production and a biomass boiler generation respectively. And the training, test, and validate dataset split ratio is 0.8, 0.16, and 0.04.

In this study, a three-layer Artificial Neural network (ANN) model is used to predict the missing values for both CHP and biomass boiler. The CHP heat production is calculated assuming that the heat-to-power ratio is 2.5. The first input layer is 32 units with an input dimension of 6, the second layer is 16 units and a one-dimension output layer. The predictive model is trained on measurement data while the prediction is made using the resulting total heat demand from building demand simulation. Two algorithms, namely long short-term memory (LSTM) and generative adversarial network (GAN) were tested against ANN, yet their performances were worse than ANN in our case. The other possible machine learning algorithms were not fully explored since this is outside of the main research scope of this paper.

4) Comparing scenarios and performance analysis

Three scenarios were designed to compare the effectiveness of different control strategies, namely, net balance, rule-based control (RBC), and SAC controller introduced in this study. The net balance scenario is the case assuming there is no ESS at the campus and feeding the energy back to the grid when excess. The RBC scenario is to store energy at night when the demand is low and release energy during the day. The daily schedule of the RBC can be referred to Figure 4.
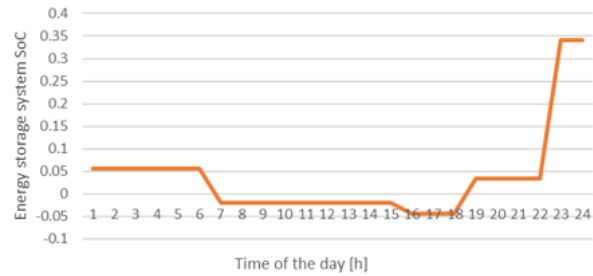

Figure 4: Scheduled daily state of charge in RBC scenario.

Four metrics are used to evaluate the performance of different scenarios. They are Averaged Ramping Factor (ARF), Daily Average load (DAL), Daily Average Peak (DAP), and E_{net} in kWh. The expressions of them are below:

$$ARF = \frac{1}{n}\sum_{t=1}^{n}\left|E_{net,t-1} - E_{net,t-1}\right|$$

$$DAL = \frac{1}{n}\sum_{t=1}^{n}\frac{1}{24}\sum_{t=1}^{t=24}E_{net,t})$$

$$DAP = \frac{1}{n}\max(E_{net,t}: t = 1, ... ,24))$$

Where, n denotes the number of timesteps during the analyzed period.

Results

The predictions of electric energy production by CHP and heat production by biomass boiler are shown in Figure 5. The Root Mean Square Error (RMSE) of the ANN model for CHP prediction is 8.06, and for the biomass boiler production is 6.8. The energy production profile during the year 2018 after data imputation is shown in Figure 6.
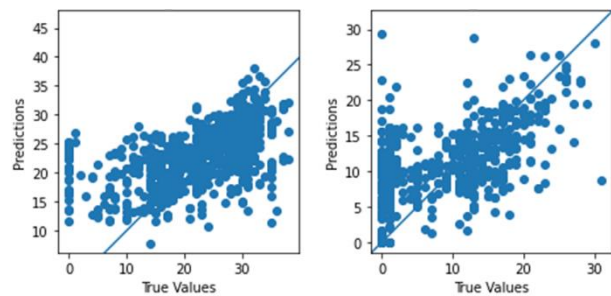

Figure 5: Relationship between the predicted values and test values from the ANN model of CHP electric production (left) and biomass boiler heat production (right).
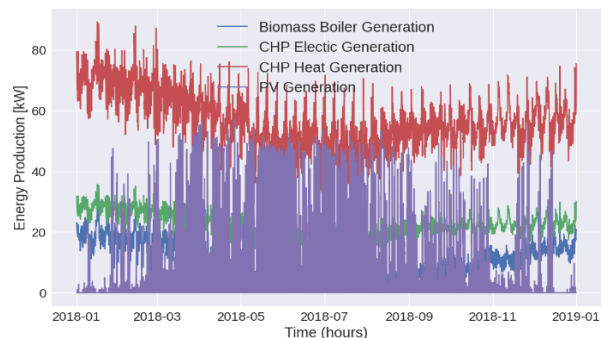

Figure 6: The resulting energy production profile during the year 2018.

The results of the energy simulation from IDA ICE are shown in Figure 7Figure 8 below. The results are the sum of the ESU load, SH load, and DHW load of the simulated six buildings.
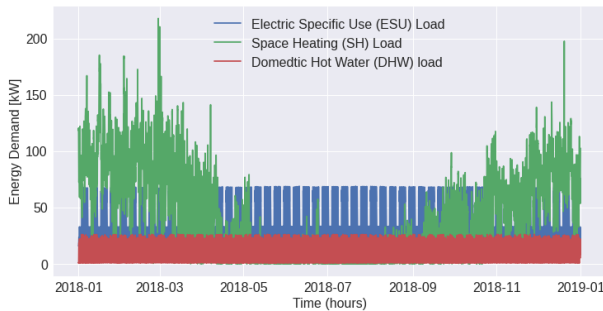

Figure 7: Total energy demand of six buildings simulated in IDA ICE.

Figure 8 shows the results of the net electric demand (E_{net}) during the year 2018 under different scenarios introduced before. Particularly, the net electric demand during 2 weeks at the beginning of February is shown in Figure 9.
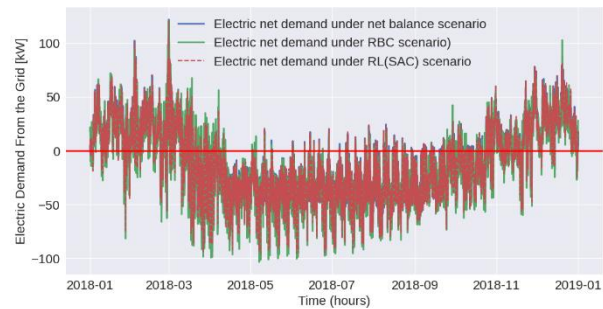

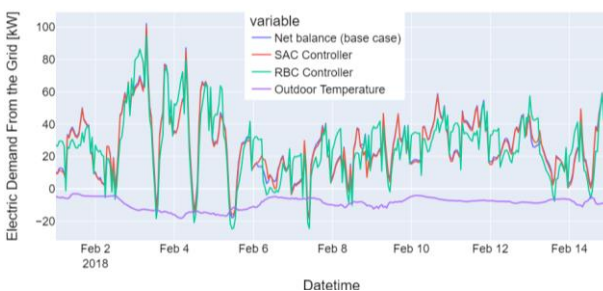*Figure 8: Net electric demand of six buildings under different management scenarios.*

Figure 9: Net electric demand of six buildings under different management scenarios during 2 weeks in February.

The DAL and DAP monthly plots and the differences between the SAC controller and the other two scenarios are shown in Figure 10. The annual DAL of the SAC controller is 3.7% and 6.3 % less than the net balance and RBC control scenario respectively; the annual DAP of the SAC controller is 8.3% and 16.5 % less than the net balance and RBC control scenario respectively. April and September are the two months the SAC failed to reduce the DAP the most notably. This is because the design of the reward function aims at optimizing the control only when the net electric demand E_{net} is larger than zero. However, the energy peaks in April and September were

barely above zero referring to Figure 8. Consequently, the SAC control is not activated and performed the worst among the optimized months.

The performances of different scenarios of the entire optimization period, winter and autumn are evaluated by E_{net} and ARF, the results are shown in Table 3. During the entire of 2018, the net electric demand for using the SAC controller was 3.7% less than the net balance scenario, and 6.3% less than using RBC. The ARF of using SAC was almost the same as the net balance scenario over the entire 2018, but 20% lower than when using RBC. During the winter period, the performances of the SAC controller were outstanding by reducing both the net electric need by 2.6% and 1.2% and the ARF by 3.9% and 29.3 when compared to the net balance and RBC scenario respectively.
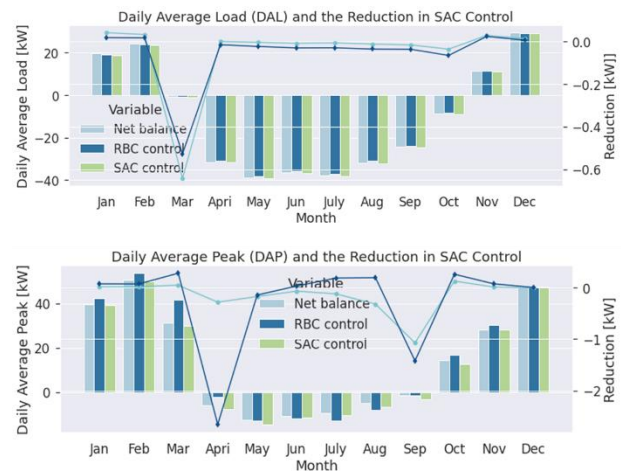

Figure 10: The monthly results of DAL and DAP, as well as the reduction in the SAC control scenario compared to the other two in 2018.

Table 3: The performance evaluation of Enet and average ramping factor (ARF) for three scenarios

Period	Metric	Net-Balance	RBC	SAC
Entire 2018	$E_{net,2018}$	-93098	-90816	-96523
	ARF	6.56	8.15	6.55
Winter	$E_{net,winter}$	17623	17364	17160
	ARF	5.88	8.64	6.11
Autumn	$E_{net,autumn}$	-5225	-5064	-5479
	ARF	5.25	6.86	5.33

Discussion

As Figure 6 shows, the ANN model could capture the CHP and biomass boiler production features when the production power is beyond 5 kW, however, the predicting values are higher than the targeting power when the targeting power is less than 5 kW, or the production is stopped. Although the ANN models are not ideally accurate, they could adequately settle the problem of missing data since the resulting CHP production covers

the base load of the campus as specified in (Lien, 2021). Additionally, this offers more robustness to the control model. More machine leaning algorithms could be compared in predicting missing values in future study.

When comparing the building energy simulation results of the investigated buildings with the real-life measurements from the campus during 2018, the ESU is 304,860 kWh while the measured ESU is 764,000 kWh. The total heating load (sum of the SH and DHW) from the simulation is 443,800 kWh while the measured one is 768,900 kWh. The deviation of the energy simulation results comes from the snow melting system and the rest of the small buildings that are not included in this study. Additionally, the non-ideally operation of the building energy system could lead those deviations between actual energy performance and energy simulation based on the standards.

From Figure 8 and Figure 9, it can be observed that using the SAC controller reduced the peak load of the campus during mild cold periods, i.e., February to March, and October to November. When comparing the SAC and RBC scenarios, the SAC scenario had lower demand, especially for demand after middays, in 12 days among the analyzed 2 weeks. Moreover, the SAC scenario performed better load-shifting ability by reducing the late afternoon peak and filling the demand valley. The reduction in the demand and peak is more conspicuous when the outdoor temperature is lower (Feb 3rd to Feb 5th). When comparing the SAC and net balance scenarios, the differences are milder on a daily scale. Improvements in using SAC emerge when analyses are conducted for a longer period.

The DAL and DAP monthly plots in Figure 10 show that the SAC controller reduced the DAL and DAP. From the results of monthly DAL, both the SAC controller and RBC controller exhibit seasonal peak shifting ability. Moreover, the reduction level of DAL is higher than that of DAP. These are due to the designed reward function in this study targeting at minimizing the E_{net} instead of the peak load over the entire optimization period.

Overall, the SAC controller scenario outperforms the other two control scenarios. The annual $E_{net,2018}$ is negative for all scenarios, which means the campus grossly exported electricity to the grid during 2018. This is not the case in real-life measurements because only the demand of six buildings is analyzed but the RE production profiles and EES capacity are for the entire campus in this study. These indicate that the SAC control scenario benefits both the energy self-sufficiency of the campus and the stability of the electric grid when the heating demand was the highest.

Conclusion

This study demonstrates the feasibility of implementing the state-of-art RL controller in a campus located in southern Norway to reduce the energy demand from the grid. In this study, *CityLearn* was tailored to match the specific energy system configuration of the case study,

and then used as a framework to implement a RL algorithm named SAC. The RL control is data-driven and relies on good-quality data. To accommodate the unsatisfied real-life measurement data quality of this case study, data imputation techniques such as missing value prediction using machine learning and white-box modeling are used. Two other scenarios were designed and compared with the RL(SAC) control, named the RBC scenario and the net balance scenario. The results show that the RL(SAC) controller effectively increased the renewable energy penetration and self-sufficiency ratio. Moreover, the SAC controller contributed to the peak reduction and demand profile smoothing over. The performance of the controller is better during winter and mild winter periods when the heating demand was high in terms of self-sufficiency, peak load shaving, and load smoothing. To be more specific, the SAC controller reduced the electricity needed for the grid by 3.7% and reduced the daily average peak by 8.3% when compared with the net balance scenario over the entire optimization period. There is potential for further development of plug-and-play energy system components in *CityLearn* as well as considering dynamic building thermal response and dynamic CHP in the control framework.

Acknowledgment

References

Åse Lekang Sørensen, E. F., Harald Taxt Walnum, Kristian Skeie, Inger Andresen. (2017). *ZEN pilot survey: Initial plans for thermal and electrical use, generation, distribution and storage*. Z. p. memo. https://www.sintef.no/en/publications/publication/?pubid=1520422

Åse Lekang Sørensen, E. F., Harald Taxt Walnum, Kristian Stenerud Skeie, Inger Andresen. (2017). *ZEN pilot survey: Initial plans for thermal and electrical use, generation, distribution and storage*. Z. p. memo. https://www.sintef.no/publikasjoner/publikasjon/?pubid=1520422

Authors, V. (2022). *Pathways to Net Zero: The Impact of Clean Energy Research*. Elsevier.

Blum, D., Arroyo, J., Huang, S., Drgoňa, J., Jorissen, F., Walnum, H. T., Chen, Y., Benne, K., Vrabie, D., Wetter, M., & Helsen, L. (2021). Building optimization testing framework (BOPTEST) for simulation-based benchmarking of control strategies in buildings. *Journal of Building Performance Simulation*, *14*(5), 586-610.

https://doi.org/10.1080/19401493.2021.198657
4

Giovanni Betti, F. T., Christine Nguyen, Stefano Schiavon. (2022). *CBE Clima Tool: a free and open-source web application for climate analysis tailored to sustainable building design*. https://arxiv.org/abs/2212.04609v1

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. *Neuron*, *66*(4), 585-595. https://doi.org/https://doi.org/10.1016/j.neuron.2010.04.016

Goy, S., & Sancho-Tomás, A. (2019). Load management in buildings. In *Urban Energy Systems for Low-Carbon Cities* (pp. 137-179). Elsevier.

Hu, M., Xiao, F., & Wang, S. (2021). Neighborhood-level coordination and negotiation techniques for managing demand-side flexibility in residential microgrids. *Renewable and Sustainable Energy Reviews*, *135*, 110248. https://doi.org/https://doi.org/10.1016/j.rser.2020.110248

Jose R Vazquez-Canteli, K. N. (2022). Retrieved 06.10 from https://github.com/intelligent-environments-lab/CityLearn

Jose R Vazquez-Canteli, S. D., Gregor Henze, Zoltan Nagy. (2020). CityLearn: Standardizing Research in Multi-Agent Reinforcement Learning for Demand Response and Urban Energy Management. *arXiv:2012.10504*. https://arxiv.org/abs/2012.10504

Lien, S. K. A., Kamilla Heimar; Bottolfsen, Hanne Liland; Lolli, Nicola; Sartori, Igor; Sørensen, Åse Lekang; Clauß, John. (2021). *Energy and Power: Essential Key Performance Indicators for Zero Emisson Neighbourhoods An Analysis Of 6 Pilot Areas*. S. A. Press. https://hdl.handle.net/11250/2833086

Niclas Björsell, A. B., Lars Eriksson, Pavel Grozman, Magnus Lindgren, Per Sahlin, Alexander Shapovalov, Mika Vuolle. (1999). IDA indoor climate and energy. 6th International Building Simulation, Kyoto, Japan.

Norge, S. (2020). Energy performance of buildings — Calculation of energy needs and energy supply. In (pp. 168): Standard Norge.

Pinto, G., Deltetto, D., & Capozzoli, A. (2021). Data-driven district energy management with surrogate models and deep reinforcement learning. *Applied Energy*, *304*, 117642. https://doi.org/https://doi.org/10.1016/j.apenergy.2021.117642

Pinto, G., Piscitelli, M. S., Vázquez-Canteli, J. R., Nagy, Z., & Capozzoli, A. (2021). Coordinated energy management for a cluster of buildings through deep reinforcement learning. *Energy*, *229*, 120725. https://doi.org/https://doi.org/10.1016/j.energy.2021.120725

Rognan, L. M. H. (2018). *Photovoltaic Power Prediction and Control Strategies of the Local Storage Unit at Campus Evenstad* Norwegian University of Science and Technology]. http://hdl.handle.net/11250/2570005

Stian Backe, Å. L. S., Dimitri PineLl, John Clauß, Carine Lausselet, Ruth Woods. (2019). *CONSEQUENCES OF LOCAL ENERGY SUPPLY IN NORWAY: A case study on the ZEN pilot project Campus Evenstad* ZEN. https://sintef.brage.unit.no/sintef-xmlui/bitstream/handle/11250/2611937/ZEN%2bReport%2bno%2b17.pdf?sequence=1&isAllowed=y

Vázquez-Canteli, J. R., Kämpf, J., Henze, G., & Nagy, Z. (2019). CityLearn v1. 0: An OpenAI gym environment for demand response with deep reinforcement learning. Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation,

Vázquez-Canteli, J. R., & Nagy, Z. (2019). Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Applied Energy*, *235*, 1072-1089. https://doi.org/https://doi.org/10.1016/j.apenergy.2018.11.002

Wang, N., Li, J., Ho, S.-S., & Qiu, C. (2021). Distributed machine learning for energy trading in electric distribution system of the future. *The Electricity Journal*, *34*(1), 106883.