

Imitation learning with artificial neural networks for demand response with a heuristic control approach for heat pumps

Thomas Dengiz^a, Max Kleinebrahm^a

^aKarlsruhe Institute of Technology (KIT), Institute for Industrial Production (IIP), Karlsruhe, Germany

Abstract

The flexibility of electrical heating devices can help address the issues arising from the growing presence of unpredictable renewable energy sources in the energy system. In particular, heat pumps offer an effective solution by employing smart control methods that adjust the heat pump's power output in reaction to demand response signals. This paper combines imitation learning based on an artificial neural network with an intelligent control approach for heat pumps. We train the model using the output data of an optimization problem to determine the optimal operation schedule of a heat pump. The objective is to minimize the electricity cost with a time-variable electricity tariff while keeping the building temperature within acceptable boundaries. We evaluate our developed novel method, PSC-ANN, on various multi-family buildings with differing insulation levels that utilize an underfloor heating system as thermal storage. The results show that PSC-ANN outperforms a positively evaluated intelligent control approach from the literature and a conventional control approach. Further, our experiments reveal that a trained imitation learning model for a specific building is also applicable to other similar buildings without the need to train it again with new data. Our developed approach also reduces the execution time compared to optimally solving the corresponding optimization problem. PSC-ANN can be integrated into multiple buildings, enabling them to better utilize renewable energy sources by adjusting their electricity consumption in response to volatile external signals.

Keywords: Imitation learning, demand response, heat pumps, residential building, artificial neural networks

1. Introduction

The increasing share of volatile renewable energy sources like photovoltaic and wind energy in many countries makes flexible electrical loads vital to cope with their intermittent nature. Especially electrical heating devices like heat pumps (HP) coupled to thermal storage can provide flexibility to the energy system by using demand response [1]. To this end, novel intelligent control strategies are necessary that regulate the heat pump's electricity consumption based on external signals, like a time-variable electricity tariff.

A promising approach to improve current control approaches is to include methods from the field of machine learning in the control procedures. Supervised learning approaches are mainly used for forecasting demand and generation load profiles for the control problem [2]. Thus, supervised learning is not used to control the devices' actions directly. Many studies from the literature use Reinforcement Learning (RL) to train an agent in a building environment to control flexible devices optimally. However, RL has some crucial drawbacks that limit its suitability for real-world applications in the energy sector. The main disadvantage is that RL needs a lot of time to train the model and thus requires a lot of interactions with the environment [3, 4]. Further, as the power and energy systems have high safety requirements, online training of the agent may negatively affect the operation of the device if no domain knowledge is included in the control.

Therefore, we introduce a novel control approach that com-

bines imitation learning with domain knowledge of the control problem. Imitation learning is a paradigm for supervised learning in which a model learns by observing and mimicking the behavior of an expert [3]. The goal is to enable the trained model to perform a task or make decisions similar to those of the expert, even in situations it has not encountered before.

Optimally reacting to the volatile energy generation requires the repeated solving of an optimization problem. Finding the optimal solution is equivalent to depicting a high-dimensional mapping between the inputs of the problem (demand, external signal, outside temperature, etc.) and the optimal control actions. As the same problem is solved for the same building recurrently with different inputs, a trained machine learning model can learn the mapping between inputs and outputs and thus imitate the exact solver. The trained model can then be amortized over multiple problem instances. Therefore, it can lead to a reduction in execution time. Imitation learning has shown promising results in the energy field [3, 5], and we investigate its applicability to intelligently controlling a heat pump for demand response.

The remainder of the paper is organized as follows: In Sec. 2, we summarize the relevant literature and highlight the contribution of this paper. We define the optimization problem in Sec. 3. Our novel control approach is explained in Sec. 4. We show the results of our experiments to evaluate the introduced control approach in Sec. 5. This paper ends with a conclusion and outlook in Sec. 6.

Nomenclature

Acronyms

ANN	Artificial Neural Network
MILP	Mixed Integer Linear Programming
ML	Machine Learning
PSC	Price Storage Control
RL	Reinforcement Learning

Indices

t	time slot
Z	total number of time slots

Parameters

Δt	time step duration
ρ^{Concrete}	density of concrete
c^{Concrete}	specific heat capacity of concrete
COP_t	coefficient of Performance of the heat pump
$k_t^{\text{switchedOff}}$	number of starts for the heat pump until time t
mod^{min}	minimum modulation degree of the heat pump
$n^{\text{switchedOff}}$	maximum number of starts for the heat pump
P^{HPmax}	maximum heating power of the heat pump

P_t^{DemandEl}	electrical power of the inflexible devices
Q_t^{DemandSH}	space heating demand at time step t
Q_t^{LossesSH}	heat losses in space heating
T^{max}	maximum allowable building temperature
T^{min}	minimum allowable building temperature
V^{UFH}	volume of underfloor heating system

Variables

$\tilde{x}_{t,pred}$	predicted output of the ML algorithm for the heat pump at time t
C	total cost
h_t^{on}	binary variable indicating if the heat pump is running at time t
$h_t^{\text{switchedOff}}$	binary variable indicating if the heat pump is switched off at time t
p_t	price of electricity
Q_t^{HP}	heat pump heat output
T_t^{Building}	building temperature at time step t
T_t^{Outside}	building temperature at time step t
x_t	modulation degree of the heat pump

2. Related work and contribution

2.1. Related work

Table 1 lists the relevant papers from the literature for our study. All the studies combine methods from the field of machine learning with traditional control approaches. Javed et al. [6], Kim et al. [7], Dey et al. [8], Zou et al. [9] and Dinh et al. [3] use an intelligent control approach for heating ventilation and air-conditioning (HVAC) systems. Javed et al. [6] use an ANN to predict the occupancy level and adjust the control system according to the predictions. Kim et al. [7] train an ANN to model the building behavior, which they then integrate into the optimization problem. The output is a multi-hour schedule. However, no direct control mechanism is based on the ANN, and a forecast is necessary for their approach. Dey et al. [8] train an RL agent using synthetic data from a rule-based control heuristic as a warm start for the RL training procedure. Zou et al. [9] use RL with a Long-Short-Term-Memory (LSTM) as the environment for interaction. They train the LSTM to predict the behavior of HVAC systems using historical data from rule-based control. Dinh et al. [3] solve a mixed-integer linear program (MILP) to generate optimal control actions for an HVAC system. They optimize the electricity costs while maximizing thermal comfort.

Zhang et al. [10], Dinh et al. [11], Ahmed et al. [12], Gao et al. [13] and López et al. [14] use machine learning methods

to improve the control of different flexible appliances in buildings. Zhang et al. [10] use transfer learning for scheduling-based loads and storage systems. Their RL agent learns from a pre-trained agent of another building. Dinh et al. [11] control the energy flows of a battery to minimize the costs with a time-variable electricity tariff. They train an ANN with the optimal control actions derived from solving a MILP. Their approach requires a forecast of the future demand, which they compute using a recurrent neural network. Ahmed et al. [12] use an ANN to learn the control actions from a simulation. They apply their approach to household devices that can be switched on and off, like air conditioners, electric water heaters, washing machines, and refrigerators. Gao et al. [13] train an ANN with the output of a MILP scheduling problem. The ANN generates actions that are then used to solve a simplified MILP problem. This reduces the solving time of the MILP. They consider batteries and diesel generators as flexible devices to minimize energy costs with a time-variable electricity price. López et al. [14] use an ANN to control electric vehicle charging to minimize the electricity cost. The labeled data is created by optimally solving the respective control problem using dynamic programming.

2.2. Contribution

To the best of our knowledge, the approach we introduce in this paper is the only one that embeds a forecast-free control approach based on imitation learning from optimal actions into

Table 1: Comparison of relevant papers from the literature

	Forecast-free control	Learning from optimal control	Embedded in smart control heuristic	Compared to optimal and smart control	Training data from other buildings
Javed et al., 2017 [6]	✗	✗	(✓)	✗	✗
Kim et al., 2020 [7]	✗	✓	✗	✓	✗
Dey et al., 2023 [8]	✓	✗	✗	(✓)	✗
Zou et al., 2020 [9]	✓	✗	(✓)	✗	✗
Dinh et al., 2022 [3]	✓	✓	✗	✓	✗
Zhang et al., 2020 [10]	✓	✗	✗	(✓)	✓
Dinh et al., 2022 [11]	✗	✓	✗	✓	✗
Ahmed et al., 2016 [12]	✓	✗	✗	✗	✗
Gao et al., 2022 [13]	✓	✓	✗	✓	✗
López et al., 2019 [14]	✓	✓	✗	✗	✗
Present work	✓	✓	✓	✓	✓

a smart control heuristic. Further, in contrast to all other papers from the literature, our study compares the novel control approach to both optimal control and a positively evaluated smart control heuristic from the literature. Another main feature of our study is using training data from other buildings for the supervised learning approach. In summary, our paper has the following three unique contributions:

- We introduce a novel imitation learning based control approach that combines domain knowledge with supervised ML. The introduced control approaches are compared to optimal control, smart heuristic control, and conventional control approaches.
- We test different ML methods for their applicability to improve the control approach of a heat pump. We generate the training data by optimally solving a MILP.
- Our study investigates the ability of the introduced approach to generalize by applying a model that has been trained with data from other similar buildings.

3. Optimization problem of the building

To determine the optimal schedule for the heat pump, we solve a MILP. The goal of the building is to minimize the electricity cost C under a time-variable electricity tariff p_t . The heat pump utilizes the building mass as thermal storage. Eq. 1 shows the objective function. The variable x_t determines the modulation degree of the heat pump. Next to the electrical power of the heat pump, the electricity consumption of the inflexible devices P_t^{DemandEl} also influences the objective function.

$$\min C = \sum_{t=0}^Z (x_t \cdot P^{\text{HPmax}} + P_t^{\text{DemandEl}}) \cdot \Delta t \cdot p_t \quad (1)$$

subject to:

$$T^{\min} \leq T_t^{\text{Building}} \leq T^{\max} \quad \forall t \quad (2)$$

$$T_t^{\text{Building}} = T_{t-1}^{\text{Building}} + \frac{Q_t^{\text{HP}} - Q_t^{\text{DemandSH}} - Q_t^{\text{LossesSH}}}{V^{\text{UFH}} \cdot \rho^{\text{Concrete}} \cdot c^{\text{Concrete}}} \quad \forall t \neq 0 \quad (3)$$

$$Q_t^{\text{HP}} = x_t \cdot P^{\text{HPmax}} \cdot \text{COP}_t \cdot \Delta t \quad \forall t \quad (4)$$

$$h_t^{\text{switchedOff}} \leq h_{t-1}^{\text{on}} \quad \forall t \neq 0 \quad (5)$$

$$h_t^{\text{switchedOff}} \leq 1 - h_t^{\text{on}} \quad \forall t \quad (6)$$

$$h_t^{\text{switchedOff}} \geq h_{t-1}^{\text{on}} - h_t^{\text{on}} \quad \forall t \neq 0 \quad (7)$$

$$k_t^{\text{switchedOff}} = \sum_{\tau=0}^t h_{\tau}^{\text{switchedOff}} \quad \forall t \quad (8)$$

$$k_t^{\text{switchedOff}} \leq n^{\text{switchedOff}} \quad \forall t \quad (9)$$

$$x_t \leq h_t^{\text{on}} \quad \forall t \quad (10)$$

$$x_t \geq h_t^{\text{on}} \cdot \text{mod}^{\min} \quad \forall t \quad (11)$$

$$x_t \in \{0, 1\}, h_t^{\text{switchedOff}} \in \{0, 1\}, h_t^{\text{on}} \in \{0, 1\} \quad (12)$$

Eqs. 2 to 12 define the problem's constraints. Eq. 2 ensures that the temperature of the building T_t^{Building} is always between a lower T^{\min} and an upper T^{\max} limit. We use a one-zone model for the temperature of the building with the energetic difference equation Eq. 3. The heat energy from the heat pump Q_t^{HP} increases the temperature of the building while the demand for space heating Q_t^{DemandSH} and the losses of the heating system itself Q_t^{LossesSH} decrease it. The demand for space heating is a time series incorporating the building's transmission and ventilation losses, as well as internal and solar gains. We describe in Sec. 5.1 how we generated the time series. As we use an underfloor heating system, we need to divide the energetic difference by the volume of the underfloor heating system V^{UFH} , the density ρ^{Concrete} and the specific heat capacity of concrete c^{Concrete} . To calculate the heating energy of the heat pump, we multiply the heat pump's modulation degree x_t in Eq. 4 with the maximum electrical power of the heat pump P^{HPmax} , the temperature-dependant coefficient of performance (see 5.1) and the time resolution of the model Δt . The equation system Eqs. 5 to 9 ensures that the number of heat pump starts $k^{\text{switchedOff}_t}$ (or switch-offs) does not exceed a predefined upper limit $n^{\text{switchedOff}}$. This prevents damage to the compressor resulting from too frequent starts. For this purpose, we use the two auxiliary binary variables $h^{\text{switchedOff}_t}$ and h^{on_t} . The variable $h^{\text{switchedOff}_t}$ indicates if the heat pump is being switched off at time t and h^{on_t} has the value 1 if the device is running at time t . Eq. 10 and Eq. 11 force the

heat pump not to violate a minimum modulation degree mod^{\min} while running.

4. Imitation learning for building control

We introduce a novel control approach for heat pumps that combines imitation learning with the intelligent control heuristic *Price-Storage-Control* [15]. Imitation learning is a machine learning paradigm where an agent learns a task by observing and imitating the behavior of an expert demonstrator [16]. In imitation learning, the learning agent aims to replicate the demonstrator's actions to perform a specific task. The agent learns a mapping between observations and actions. In our application, the expert demonstrator consists of actions obtained from optimally solving the optimization problem (MILP) of Sec. 3.

Fig. 1 illustrates the workflow of our approach to create the heat pump controller *PSC-ML*. We use historical data on electricity prices, heat demand, and outside temperatures as input for an optimization problem. The optimization problem includes a simple building model with a uniform temperature whose difference equation is described in Sec. 3. The output of the optimization problem is the heat pump's optimal heating actions for every time slot (HP schedule) and the resulting building temperature.

From the HP schedule of the last 24 hours, two HP run statistics are derived. These statistics indicate how often the heat pump has started the compressor during the current day and if the heat pump was running at time t . This information serves as one part of the input features for the machine learning (ML) algorithm.

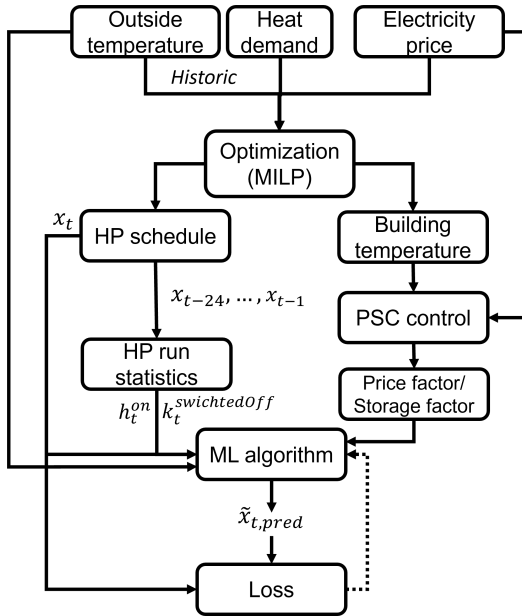


Figure 1: PSC-ML training

The other part of the input features are the *Price factor* and the *Storage factor* calculated by *PSC* algorithm using the resulting building temperature from the optimization problem and the

historic electricity price. The *Price factor* quantifies for every time slot the share of future electricity price values in the next 24 hours that are higher than the current price p_t . The *Storage factor* determines how far the current temperature of the building $T_t^{Building}$ is away from the lower T^{\min} and upper T^{\max} temperature limits. Details about the calculations of the control algorithm *PSC* are explained in [15]. Additionally, we use the outside temperature as another input feature.

The optimal heating action x_t from the resulting HP schedule serves as the target label for the ML algorithm, which we obtain by solving the optimization problem. The ML algorithm is trained to capture the relationship between the inputs of the control problem and the optimal output. It essentially derives a regression model that maps the inputs to the optimal outputs. We calculate the loss during the training by comparing the predicted output of the ML algorithm $\tilde{x}_{t,pred}$ with the corresponding optimal output from the optimization problem x_t . The trained model is subsequently incorporated into the heat pump and integrated with the *PSC* algorithm as part of the novel control approach *PSC-ML*.

Fig. 2 illustrates the application of *PSC-ML*. At every time slot t , the heat pump controller observes the relevant input data. The *PSC* algorithm and the trained ML model are integrated into the heat pump. The HP controller uses the outside temperature $T_t^{Outside}$, the number of starts of the heat pump so far $k_t^{switchedOff}$, and the binary variable h_t^{on} directly. Using the electricity price p_t and the building temperature $T_t^{Building}$, the *PSC* algorithm derives the *Price factor* and the *Storage factor*. These five features are the input for the trained ML regression model, which outputs the heating action x_t , influencing the HP run statistics during the next time slot. The *PSC-ML* controller includes some super-ordinate rules that can overrule the actions proposed by the ML algorithm. These rules are similar to the ones explained in [17]. They ensure that the heat pump heats the building if the room temperature is too low and stops heating if it is too high.

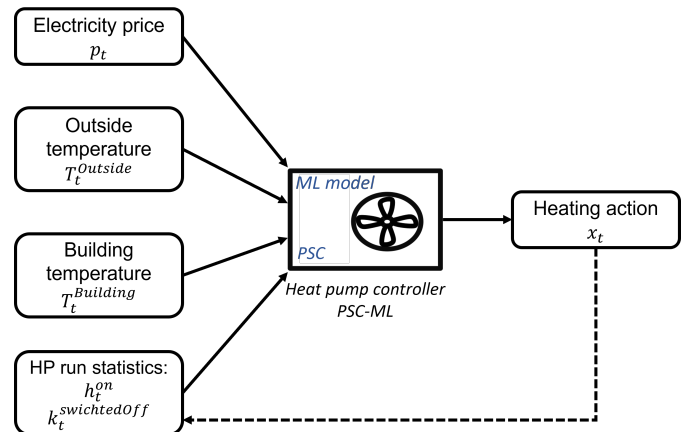


Figure 2: PSC-ML applied

We use different machine learning approaches (see Section 5), but applying an ANN as the machine learning approach yields the best results. Fig. 3 illustrates a schematic view of the ANN. The ANN is a multi-layer perceptron with fully con-

nected nodes and the rectified linear unit (ReLU) as the activation function. As described in Section 5, we investigated different configurations and hyperparameters.

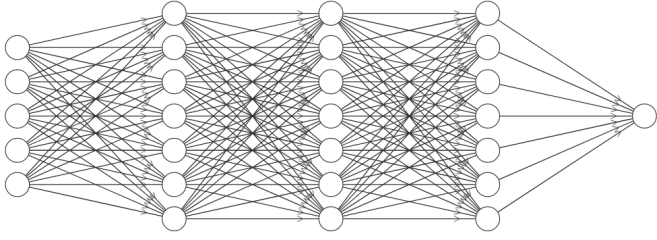


Figure 3: ANN schema

We use three types of buildings that differ regarding their insulation level and, thus, their heat demand. Heat pumps are mainly used in buildings with good insulation standards due to increased efficiency. Hence, the analysis includes buildings with a heat demand of 25 kWh , 50 kWh , and 80 kWh per year for every square meter of living area. Overall, the buildings have a good insulation level compared to the average heat demand of buildings in Germany, which lies between $125 \frac{\text{kWh}}{\text{m}^2}$ and $150 \frac{\text{kWh}}{\text{m}^2}$ [18] per year. However, these values are typical for new buildings in Germany.

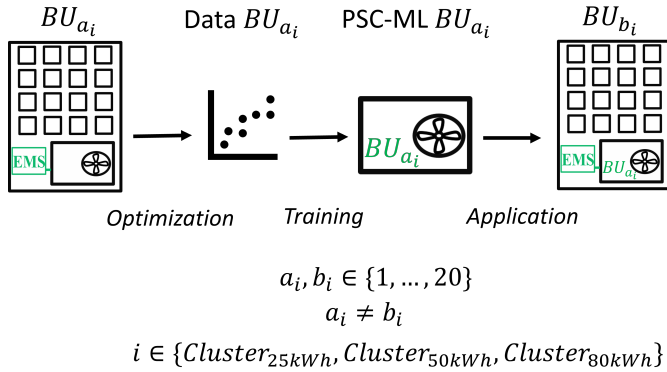


Figure 4: Training process with data from other buildings

Our analysis includes data from different buildings for training the ANN to see if a trained model can be applied to similar buildings. Fig. 4 illustrates the training process. We use building BU_{a_i} from cluster i of buildings to generate the training data by solving the optimization problem for that building. Each cluster includes buildings with a similar heat demand per year and square meters of living area. For each building cluster, we created 20 different datasets (see Section 5.1). The time series of heat demand within each cluster varies regarding the temporal dimension and the magnitude of values. These differences arise due to temporal shifts and multiplication of the time series with a constant factor, leading to variations in both the timing and the magnitude of the heat demand data. The resulting average yearly heat demand for every building within one cluster differs at most $+5 \frac{\text{kWh}}{\text{m}^2}$ or $-5 \frac{\text{kWh}}{\text{m}^2}$ from the base demand, which is $25 \frac{\text{kWh}}{\text{m}^2}$, $50 \frac{\text{kWh}}{\text{m}^2}$ and $80 \frac{\text{kWh}}{\text{m}^2}$ for the corresponding clusters.

The generated training data for building BU_{a_i} is then used as

input for the ANN. This results in the specific controller $PSC-ML$ for building BU_{a_i} . To test the capability of our approach to generalize and thus to be applied in different buildings, we use another building from the same cluster BU_{b_i} for applying the novel control approach. Thus, building BU_{b_i} uses the controller $PSC-ML BU_{a_i}$ for the evaluation that has been trained with data from another building.

5. Results

5.1. Scenarios for the analysis

We use the air-source heat pump *Compress 6800i AW* [19] from the company *Bosch Thermotechnik* with a maximum electrical power of 3000 W and a minimal modulation degree of 20%. For the efficiency factor COP , we assume a linear relationship between the values provided in the technical documentation of the heat pump model as in [17]. The efficiency COP_t for a time slot t depends on the difference between the sink temperature, which is the underfloor heating system's supply temperature of $30 \text{ }^\circ\text{C}$, and the source temperature (outside temperature). For the 25 kWh buildings, we aggregated three of the mentioned heat pump models into one with an electrical power of 9000 W , while the 50 kWh buildings have four heat pumps combined in cascading operations, and the 80 kWh have six. We set the maximum number of starts for the heat pump to 28 during one week.

The buildings are all multi-family houses located in the federal state of *Schleswig-Holstein* in the north of Germany. The building has 12 apartments, each with a living area of 75 m^2 and a concrete width of 7 cm for the UFH [17]. The lower limit for the temperature is $20.5 \text{ }^\circ\text{C}$, while the upper limit is at $23.5 \text{ }^\circ\text{C}$. The losses of the UFH were assumed to be 45 W . It has to be noted that these are merely the losses of the heating system's tubes that do not contribute to heating the building. The heat demand data incorporates the building's significantly higher transmission heat losses.

We use the CREST model [20] and German time use data [21] for occupancy behavior simulation, which builds the basis for the consistent simulation of building space heating and electricity demand. Appliance starts are simulated based on the occupancy profiles, local weather conditions, and household device equipment. Finally, the appliance-specific load profiles are aggregated to one electricity demand profile per household. Additionally, the 5R1C thermal building model [22] is used to calculate the demand for space heating based on occupancy-related internal heat gains and thermal building parameters taken from the TABULA residential building typology [23].

For the electricity price, we use the data of the Day-Ahead market in Germany from 2021, which we took from the *ENTSO-E Transparency Plattform* [24]. We scaled them up to align with the average electricity price for residential customers in Germany. The prices have an hourly resolution.

5.2. Evaluation

We use the optimal schedules of 20 weeks to train the ML models. The time resolution of the data obtained from the op-

timization is 30 minutes. We split the training data into 70% for training the model and 30% for the validation. After training, the ML model, embedded into the control approach *PSC*, is tested in 20 weeks that were not included in the training data. For every week in the evaluation, we train the model from scratch and choose 20 weeks of training data randomly from 26 weeks of available training data (we only considered the heating period from October to March). The time resolution of the control actions Δt (and thus of the training data) is 30 minutes.

For our first evaluation, we train the model with data from one building and apply the trained model to 20 different buildings from the same building cluster. The average training time per week was 9 minutes using the GPU *Nvidia GeForce RTX 3090* [25] and an *Intel Core i9-7940X* [26] as CPU. The average time for deriving and executing the control actions for one week (2 decisions per hour) is 3 seconds with the trained *PSC-ML* control approach, while the exact optimization needs about 40 seconds. We use the Gurobi solver [27] to solve the optimization problem of the building.

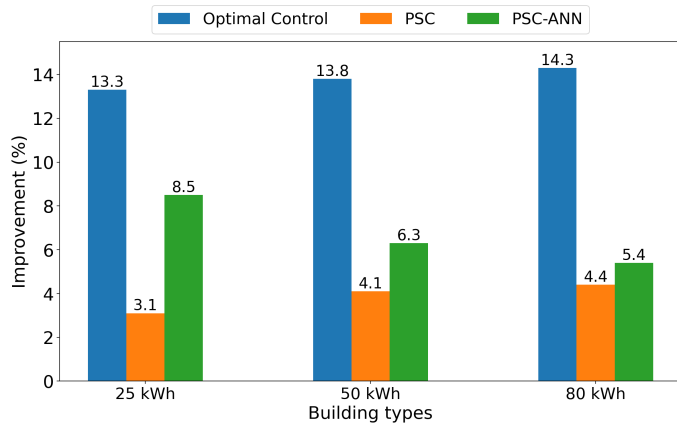


Figure 5: Improvement regarding the cost reduction compared to the Conventional Control approach averaged over 20 buildings per building type for 20 weeks and three runs

Fig. 5 depicts the cost reduction of the methods *Optimal Control*, *PSC* and *PSC-ANN* compared to *Conventional Control*. We obtained the results for the three building type clusters by averaging over 20 buildings per building type and 20 weeks. We further made three simulation runs per method. Thus, each bar represents the average values of 1200 evaluation weeks. The results reveal that for all building-type clusters, the three methods lead to cost improvements compared to the Conventional Control approach. As expected, the *Optimal Control* yields the best results because it is based on the assumption of having a perfect forecast of future demands and prices (or, equivalently, calculating the optimal control actions retrospectively using historical data). Thus, the *Optimal Control* merely represents a theoretical upper bound for the improvements. In contrast, *PSC* and *PSC-ANN* do not require a building model or any forecast to derive their control actions. The diagram shows that *PSC-ANN* outperforms *PSC* for all building types. The additional cost reduction using an ANN to calculate the control actions ranges between 1% for the 80 kWh buildings and 4.7% for the 25 kWh

buildings.

Table 2 lists the average electricity costs per week for the 50 kWh buildings averaged over 20 buildings and three runs. *PSC* and *PSC-ANN* lead to significantly better results than *Conventional Control*. While in three out of 20 weeks, *PSC* leads to lower costs compared to *PSC-ANN*, on average, the use of *PSC-ANN* leads to cost reductions of more than 5 € per week.

Table 2: Average electricity costs in € per week for the 50 kWh buildings averaged over 20 buildings and three runs

Week	Optimal Control	Conventional Control	PSC	PSC-ANN
1	269.52	323.33	316.00	317.91
2	155.79	205.12	189.73	178.63
3	273.40	316.44	306.49	295.91
4	178.59	200.89	198.92	185.42
5	325.30	376.47	356.56	356.41
6	228.01	270.14	256.62	246.17
7	281.37	324.84	306.51	300.89
8	233.47	272.67	257.95	257.00
9	133.24	154.66	153.04	147.75
10	186.08	217.42	205.24	201.86
11	178.45	201.28	192.80	185.83
12	188.51	214.73	206.55	204.79
39	122.47	129.85	128.52	125.07
40	135.51	146.05	146.30	136.90
41	140.41	158.68	157.07	145.00
42	254.95	292.63	276.92	266.73
43	146.68	155.92	155.54	152.18
44	241.43	276.34	263.73	264.05
45	221.62	272.17	253.56	257.25
46	153.90	188.46	177.96	175.29
Average	202.44	234.90	225.30	220.05

Next to an artificial neural network, we investigate the application of Random Forest and Gradient Boosting decision trees as the machine learning approach. Fig. 6 shows the results of the experiments. The bars quantify the average improvement compared to the *Conventional Control* approach for different machine learning methods averaged over six buildings per building type for 20 weeks. The diagram reveals that applying any function approximator improves the results of *PSC*. Using artificial neural networks for imitation learning for all three building types leads to the most significant improvements and, thus, the lowest costs in our experiments. Because of this, we only use ANN for our large-scale analysis with 20 different buildings per building type (see Fig. 5).

We ran experiments with different hyperparameters of the ANN. Fig. 7 illustrates the results of our analysis for one 25 kWh building averaged over 20 weeks. We varied the batch size, the learning rate, the number of neurons per layer, and the number of layers. In general, the impact of the different parameters is mediocre, ranging from improvements of 9.5% to 10.8%. For our large-scale experiment, we chose the hyperparameter combination that yielded the best results (batch size 30,

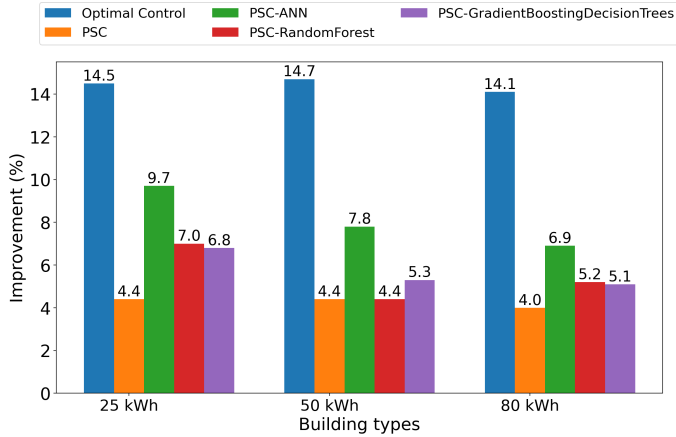


Figure 6: Average improvement compared to the Conventional Control approach for different machine learning methods averaged over 6 buildings per building type for 20 weeks

learning rate 0.0018, number of neurons per layer 50, number of layers 5).

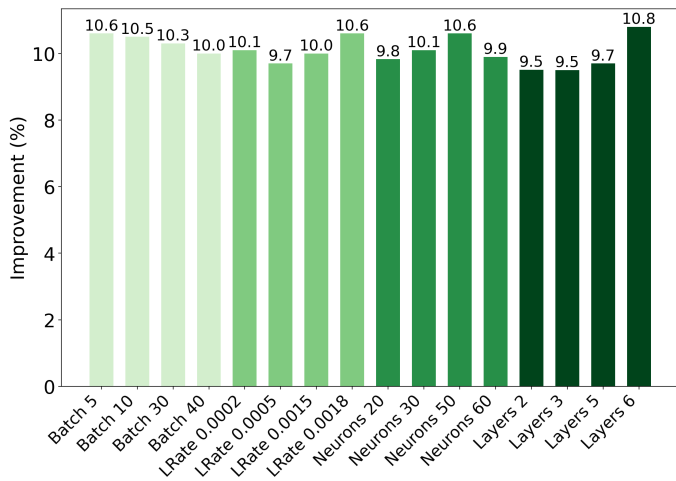


Figure 7: Improvement compared to the Conventional Control approach for different ANN parameters averaged over one 25 kWh building for 20 weeks

Fig. 8 plots the loss function (mean squared error) as a function of the epochs of the training and validation dataset for three weeks of a 25kWh building. We choose 20 epochs as the error function in the validation dataset does not significantly reduce after a certain number of epochs. We observed this in almost all weeks and for multiple different buildings.

We implemented the simulations and the optimization in the programming language *Python*. For the MILP, we use the package *Pyomo* [28], for the ML approaches *keras* [29] and *scikit-learn* [30].

5.3. Critical appraisal

To carry out our experiments, we made some simplifications. The used building model is relatively simple as it only has one temperature zone. Thus, it does not consider internal heat flows and solar heat gains. The solar heat gains are incorporated in

the demand data. We also did not analyze the use of domestic hot water tanks as an additional source of flexibility needed throughout the year. Further, we solely considered one building with one flexible device. The analysis of coordinated reactions to demand response signals from multiple buildings was not in the scope of this paper. We carried out all our investigations in a simulation environment. Thus, we did not apply our proposed approach in real-world experiments. Our method defines a general procedure to combine domain knowledge with machine learning for control problems of electrical heating devices. Because of this, we are convinced that our developed control algorithm can be used for more complex building settings and real-world applications.

6. Conclusion

In this paper, we developed a novel control approach for demand response with heat pumps *PSC-ANN* based on imitation learning. It uses the effective control method *PSC* from the literature and trains an artificial neural network with optimal control action data. The trained model quickly derives the control actions based on real-time data. The results reveal the benefits of including an artificial neural network in the control process. The introduced approach outperforms an intelligent control algorithm that has been positively evaluated. Further, we showed that the trained model can generalize and thus be used for various similar buildings without the need for training again with building-specific data.

This enables our developed approach to be applied in many buildings and scenarios. All tested models strongly outperform a conventional control approach. They lead to a better utilization of volatile renewable energy sources and can also be used to stabilize the electricity grid.

Future work will consider multiple buildings with intelligent controllers and their interactions. Methodologically, the use of machine learning methods for dealing with sequential data, like recurrent neural networks, constitutes a promising approach for building-related control problems. In addition, we plan to compare our approach to methods from the field of reinforcement learning, especially regarding the sample efficiency. Also, we intend to include other flexibility options like electric vehicles, hot water tanks, or battery storage systems in our analysis.

Supplementary materials

Both the code and data are openly accessible. The repository containing the commented code is available at [GitHub](#). The input data and result profiles can be accessed at [KITOpen](#).

Appendix

Exemplary profiles for one week resulting from the different control approaches are presented in Fig. 9.

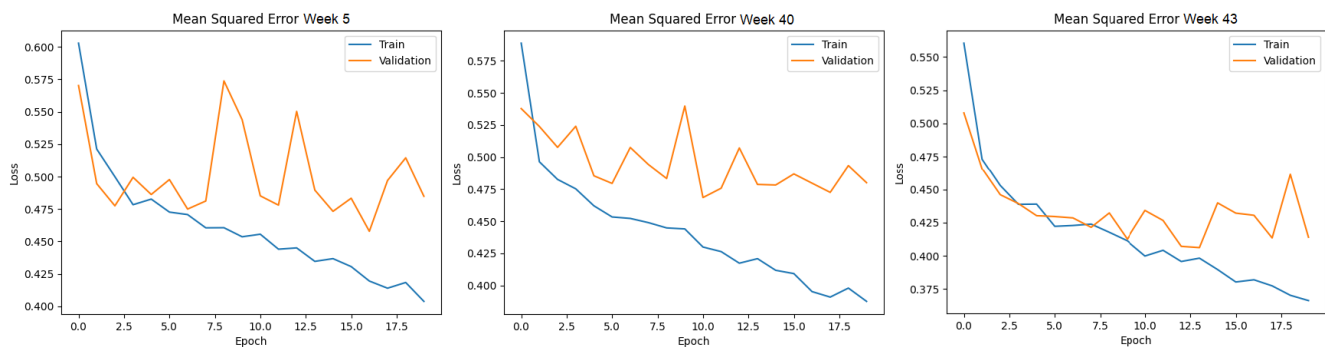


Figure 8: RMSE as a function of epochs for three weeks of a 25kWh building

CRedit author statement

Thomas Dengiz: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - Original Draft, Visualization. **Max Kleinebrahm:** Validation, Resources, Data Curation, Writing - Original Draft

Acknowledgments

This work was supported by the project AsimutE (Autoconsummation et Stockage Intelligents pour une Meilleure Utilisation de l'Énergie) from the European Territorial Cooperation program Interreg.

References

- [1] D. Patteeuw, K. Bruninx, A. Artoni, E. Delarue, W. D'haeseleer, L. Helsen, *Integrated modeling of active demand response with electric heating systems coupled to thermal energy storage systems*, *Applied Energy* 151 (2015) 306–319. doi:<https://doi.org/10.1016/j.apenergy.2015.04.014>. URL <https://www.sciencedirect.com/science/article/pii/S0306261915004535>
- [2] I. Antonopoulos, V. Robu, B. Couraud, D. Kirli, S. Norbu, A. Kiprakis, D. Flynn, S. Elizondo-Gonzalez, S. Wattam, *Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review*, *Renewable and Sustainable Energy Reviews* 130 (2020) 109899. doi:<https://doi.org/10.1016/j.rser.2020.109899>.
- [3] H. T. Dinh, D. Kim, *Milp-based imitation learning for hvac control*, *IEEE Internet of Things Journal* 9 (8) (2022) 6107–6120. doi:<https://doi.org/10.1109/JIOT.2021.3111454>.
- [4] D. Cao, W. Hu, J. Zhao, G. Zhang, B. Zhang, Z. Liu, Z. Chen, F. Blaabjerg, *Reinforcement learning and its applications in modern power and energy systems: A review*, *Journal of Modern Power Systems and Clean Energy* 8 (6) (2020) 1029–1042. doi:<https://doi.org/10.35833/MPE.2020.000552>.
- [5] X. Pan, T. Zhao, M. Chen, S. Zhang, *Deepopf: A deep neural network approach for security-constrained dc optimal power flow*, *IEEE Transactions on Power Systems* 36 (3) (2021) 1725–1735. doi:<https://doi.org/10.1109/TPWRS.2020.3026379>.
- [6] A. Javed, H. Larijani, A. Ahmadiania, R. Emmanuel, M. Mannion, D. Gibson, *Design and implementation of a cloud enabled random neural network-based decentralized smart controller with intelligent sensor nodes for hvac*, *IEEE Internet of Things Journal* 4 (2) (2017) 393–403. doi:<https://doi.org/10.1109/JIOT.2016.2627403>.
- [7] Y.-J. Kim, *A supervised-learning-based strategy for optimal demand response of an hvac system in a multi-zone office building*, *IEEE Transactions on Smart Grid* 11 (5) (2020) 4212–4226. doi:<https://doi.org/10.1109/TSG.2020.2986539>.
- [8] S. Dey, T. Marzullo, X. Zhang, G. Henze, *Reinforcement learning building control approach harnessing imitation learning*, *Energy and AI* 14 (2023) 100255. doi:<https://doi.org/10.1016/j.egyai.2023.100255>.
- [9] Z. Zou, X. Yu, S. Ergan, *Towards optimal control of air handling units using deep reinforcement learning and recurrent neural network*, *Building and Environment* 168 (2020) 106535. doi:<https://doi.org/10.1016/j.buildenv.2019.106535>.
- [10] X. Zhang, X. Jin, C. Tripp, D. J. Biagioni, P. Graf, H. Jiang, *Transferable reinforcement learning for smart homes*, in: *Proceedings of the 1st International Workshop on Reinforcement Learning for Energy Management in Buildings & Cities, RLEM'20, Association for Computing Machinery, New York, NY, USA, 2020*, p. 43–47. doi:[10.1145/3427773.3427865](https://doi.org/10.1145/3427773.3427865).
- [11] H. T. Dinh, K. haeng Lee, D. Kim, *Supervised-learning-based hour-ahead demand response for a behavior-based home energy management system approximating milp optimization*, *Applied Energy* 321 (2022) 119382. doi:<https://doi.org/10.1016/j.apenergy.2022.119382>.
- [12] M. S. Ahmed, A. Mohamed, H. Shareef, R. Z. Homod, J. A. Ali, *Artificial neural network based controller for home energy management considering demand response events*, 2016 International Conference on Advances in Electrical, Electronic and Systems Engineering (ICAEEES) (2016). doi:<https://doi.org/10.1109/ICAEEES.2016.7888097>.
- [13] S. Gao, C. Xiang, M. Yu, K. T. Tan, T. H. Lee, *Online optimal power scheduling of a microgrid via imitation learning*, *IEEE Transactions on Smart Grid* 13 (2) (2022) 861–876. doi:<https://doi.org/10.1109/10.1109/TSG.2021.3122570>.
- [14] K. L. López, C. Gagné, M.-A. Gardner, *Demand-side management using deep learning for smart charging of electric vehicles*, *IEEE Transactions on Smart Grid* 10 (3) (2019) 2683–2691. doi:<https://doi.org/10.1109/TSG.2018.2808247>.
- [15] M. Frahm, T. Dengiz, P. Zwickel, H. Maaß, J. Matthes, V. Hagenmeyer, *Occupant-oriented demand response with multi-zone thermal building control*, *Applied Energy* 347 (2023) 121454. doi:<https://doi.org/10.1016/j.apenergy.2023.121454>. URL <https://www.sciencedirect.com/science/article/pii/S0306261923008188>
- [16] A. Hussein, M. M. Gaber, E. Elyan, C. Jayne, *Imitation learning: A survey of learning methods* 50 (2) (2017). doi:<https://doi.org/10.1145/3054912>.
- [17] T. Dengiz, P. Jochem, W. Fichtner, *Demand response with heuristic control strategies for modulating heat pumps*, *Applied Energy* 238 (2019) 1346–1360. doi:<https://doi.org/10.1016/j.apenergy.2018.12.008>.
- [18] Verbraucherzentrale NRW e.V., *Energieausweis: Was sagt dieser Steckbrief für Wohngebäude aus?* (2023). URL <https://www.verbraucherzentrale.de/wissen/energie/energetische-sanierung/energieausweis-was-sagt-dieser-steckbrief-fuer-wohnbaeude-aus-24074>
- [19] Bosch Thermotechnik GmbH, *Luft-Wasser-Wärmepumpe Compress 6800i AW* (2024). URL <https://www.bosch-homecomfort.com/de/de/ocs/wohng>

- [ebaude/compress-6800i-aw-19312695-p/](#)
- [20] I. Richardson, M. Thomson, D. Infield, C. Clifford, Domestic electricity use: A high-resolution energy demand model, *Energy and Buildings* 42 (10) (2010) 1878–1887. doi:<https://doi.org/10.1016/j.enbuid.2010.05.023>.
- [21] Statistisches Bundesamt (Destatis), [Einkommen, Konsum und Lebensbedingungen - Zeitverwendung](#), accessed: 24.04.2024 (2024). URL https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Zeitverwendung/_inhalt.html#sprg234984
- [22] International Organization for Standardization, [Iso 13790:2008 - energy performance of buildings](#), accessed: 06.05.2024 (2018). URL <https://www.iso.org/standard/41974.html>
- [23] Institut Wohnen und Umwelt GmbH, [Deutsche Wohngebäudetypologie. Beispielhafte Maßnahmen zur Verbesserung der Energieeffizienz von typischen Wohngebäuden – zweite erweiterte Auflage –](#), accessed: 24.04.2024 (2015). URL https://www.episcope.eu/downloads/public/docs/brochure/DE_TABULA_TypologyBrochure_IWU.pdf
- [24] Entso-E Transparency Platform, [Day-ahead prices](#) (2023). URL <https://transparency.entsoe.eu/>
- [25] NVIDIA Corporation, [Geforce RTX 3090 family](#) (2024). URL <https://www.nvidia.com/en-gb/geforce/graphics-cards/30-series/rtx-3090-3090ti/>
- [26] Intel Corporation, [Intel core i9-7940x x-series processor](#) (2024). URL <https://www.intel.com/content/www/us/en/products/sku/126695/intel-core-i97940x-xseries-processor-19-25m-cache-up-to-4-30-ghz/specifications.html>
- [27] Gurobi Optimization, LLC, [Gurobi Optimizer Reference Manual](#) (2023). URL <https://www.gurobi.com>
- [28] W. E. Hart, J.-P. Watson, D. L. Woodruff, Pyomo: modeling and solving mathematical programs in python, *Mathematical Programming Computation* 3 (3) (2011) 219–260. doi:<https://doi.org/10.1007/s12532-011-0026-8>.
- [29] F. Chollet, et al., Keras, <https://keras.io> (2015).
- [30] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, *Journal of Machine Learning Research* 12 (2011) 2825–2830.

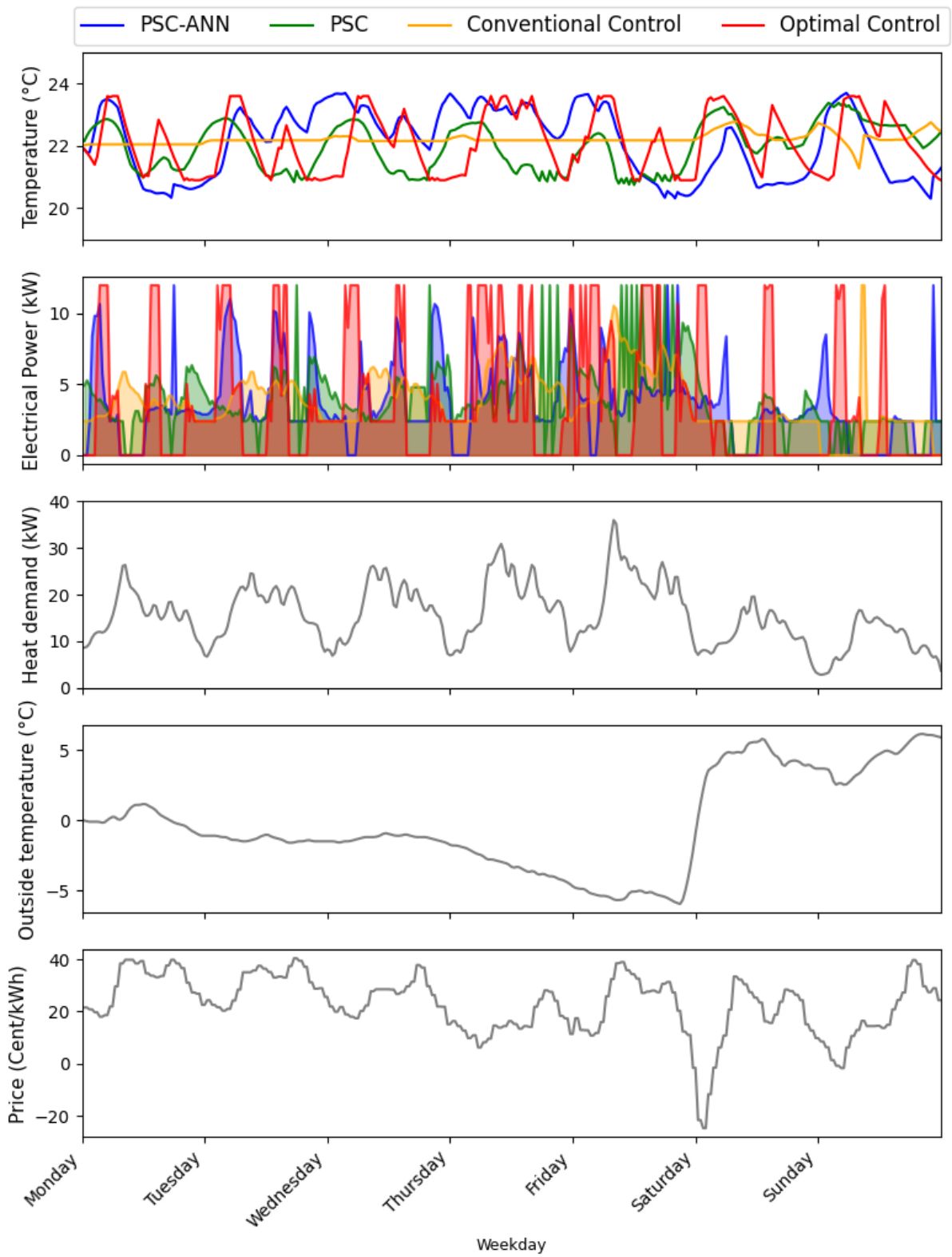


Figure 9: Exemplary profiles for one week in December for a 50 kWh building