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

Thomas Dengiz *, Max Kleinebrahm

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

HIGHLIGHTS

- Novel imitation learning approach to control heat pumps for demand response.
- Combination of machine learning with domain knowledge.
- Trained model can be used for different buildings without re-training.
- Developed methods outperform state-of-the-art control methods.

GRAPHICAL ABSTRACT
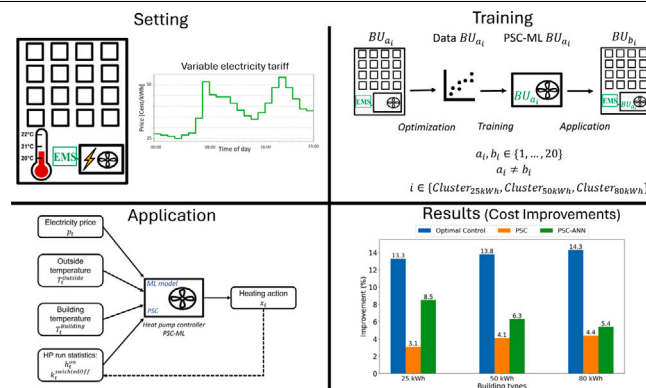
ARTICLE INFO

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.

---

* Corresponding author.
  E-mail address: thomas.dengiz@kit.edu (T. Dengiz).

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

$\widetilde{x}_{t,pred}$	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

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]. In 2022, space heating comprised roughly 63.5% of the total energy consumption in EU households [2]. Due to their high efficiency, heat pumps are particularly suitable for demand response strategies. They can effectively leverage the building's thermal mass to shift their electricity consumption in time without compromising occupant comfort. 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.

Advanced measurement technologies, such as smart meters, along with intelligent monitoring and control strategies, are evolving the traditional electricity grid into a smart grid. Smart buildings in a smart grid can reduce the electricity costs, increase the use of electricity generated by renewable energy sources and stabilize the grid [3,4].

Currently, the predominant practice involves employing a conventional control strategy for heating devices coupled with thermal storage. This strategy initiates heating once the building's temperature falls to a specified lower threshold, ceasing operation when the energy level reaches the designated upper limit. However, this control method does not exploit the potential flexibility of thermal storage systems. A promising way 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 [5]. 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 [6,7]. 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 combines 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 [6]. 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.

Imitation learning offers key advantages for controlling heat pumps compared to other learning methods like RL. It leverages expert knowledge by mimicking the actions of experienced operators or computed optimal actions, allowing faster deployment and minimizing trial-and-error processes. Unlike RL, which requires extensive exploration and interaction with the environment, imitation learning is more sample-efficient, making it suitable for systems where real-time feedback is slow, such as temperature regulation. Additionally, imitation learning combined with a rule-based control heuristic provides safer and more stable control during the application, as it avoids risky or inefficient actions, while also reducing computational complexity by bypassing the need to design suitable reward functions. This makes it a more practical and reliable approach for controlling heat pumps in dynamic real-world environments.

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 [6,8], and we investigate its applicability to intelligently controlling a heat pump for demand response.

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

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 [9]	✗	✗	(✓)	✗	✗
Kim et al. 2020 [10]	✗	✓	✗	✓	✗
Dey et al. 2023 [11]	✓	✗	✗	(✓)	✗
Zou et al. 2020 [12]	✓	✗	(✓)	✗	✗
Dinh et al. 2022 [6]	✓	✓	✗	✓	✗
Zhang et al. 2020 [13]	✓	✗	✗	(✓)	✓
Dinh et al. 2022 [14]	✗	✓	✗	✓	✗
Ahmed et al. 2016 [15]	✓	✗	✗	✗	✗
Gao et al. 2022 [16]	✓	✓	✗	✓	✗
López et al. 2019 [17]	✓	✓	✗	✗	✗
Present work	✓	✓	✓	✓	✓

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. [9], Kim et al. [10], Dey et al. [11], Zou et al. [12] and Dinh et al. [6] use an intelligent control approach for heating ventilation and air-conditioning (HVAC) systems. Javed et al. [9] use an ANN to predict the occupancy level and adjust the control system according to the predictions. Kim et al. [10] 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. [11] 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. [12] 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. [6] 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. [13], Dinh et al. [14], Ahmed et al. [15], Gao et al. [16] and López et al. [17] use machine learning methods to improve the control of different flexible appliances in buildings. Zhang et al. [13] 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. [14] 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. [15] 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. [16] 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. [17] 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 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.

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

subject to:

$$T^{min} \leq T_t^{Building} \leq T^{max} \quad \forall t \tag{2}$$

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

$$Q_t^{HP} = x_t \cdot P^{HPmax} \cdot COP_t \cdot \Delta t \quad \forall t \tag{4}$$

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

$$h_t^{switchedOff} \leq 1 - h_t^{on} \quad \forall t \tag{6}$$

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

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

$$k_t^{switchedOff} \leq n^{switchedOff} \quad \forall t \tag{9}$$

$$x_t \leq h_t^{on} \quad \forall t \tag{10}$$

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

$$x_t \in [0, 1], h_t^{switchedOff} \in \{0, 1\}, h_t^{on} \in \{0, 1\} \tag{12}$$

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 Section 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. Eqs. (10) and (11) force the heat pump not to violate a minimum modulation degree $mod^{\min}$ while running.

4. Imitation learning for heat pump control

We introduce a novel control approach for heat pumps that combines imitation learning with the intelligent control heuristic *Price-Storage-Control* [18]. Imitation learning is a machine learning paradigm where an agent learns a task by observing and imitating the behavior of an expert demonstrator [19]. 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 Section 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 Section 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 h, 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.

The other part of the input features are the *Price factor* and *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 h 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 [18]. 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. This process is an instance of imitation learning, where the ML algorithm is trained to mimic the optimal control behavior by learning the relationship between the inputs of the control problem and the corresponding optimal output. The ML
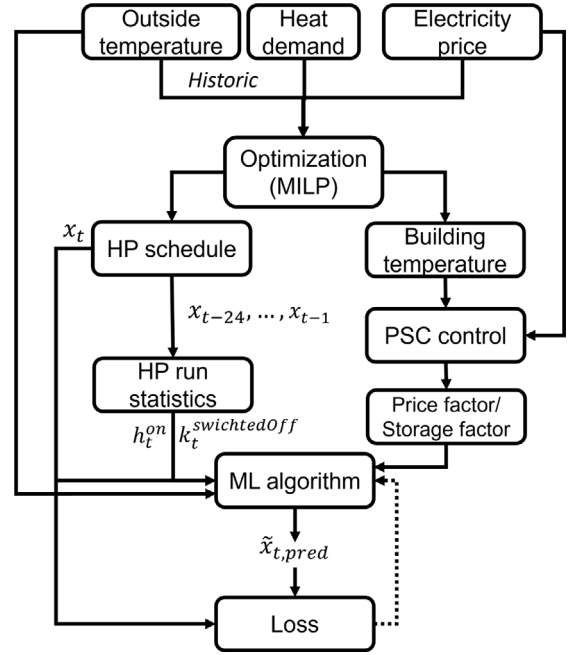

Fig. 1. PSC-ML training.

algorithm essentially derives a regression model that maps the inputs to the optimal outputs, thereby imitating the behavior of the optimal controller. 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 $k_t^{\text{switchedOff}}$, 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 [20]. They ensure that the heat pump heats the building if the room temperature is too low and stops heating if it is too high.

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

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 $150 \frac{\text{kWh}}{\text{m}^2}$ [21] per year. However, these values are typical for new buildings in Germany.

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
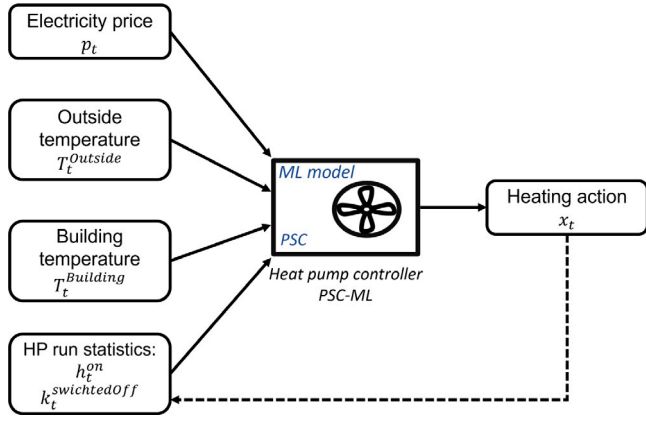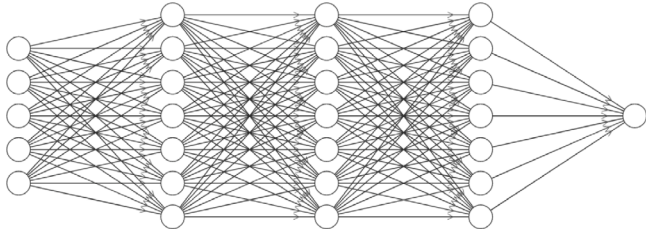
Fig. 2. PSC-ML applied.

Fig. 3. ANN schema.

$$a_i, b_i \in \{1, \dots, 20\}$$
$$a_i \neq b_i$$
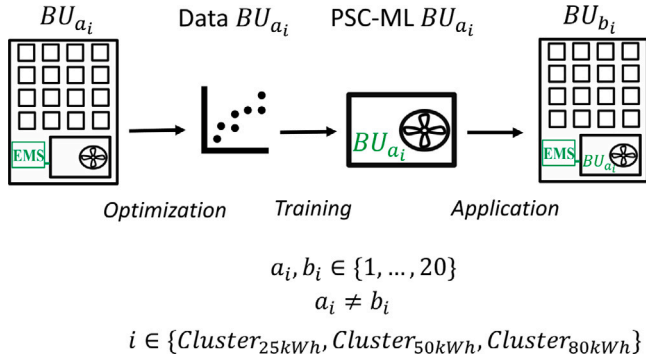$$i \in \{Cluster_{25kWh}, Cluster_{50kWh}, Cluster_{80kWh}\}$$

Fig. 4. Training process with data from other buildings.

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 $+5\ \frac{kWh}{m^2}$ or $-5\ \frac{kWh}{m^2}$ from the base demand, which is $25\ \frac{kWh}{m^2}$, $50\ \frac{kWh}{m^2}$ and $80\ \frac{kWh}{m^2}$ 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.

In our imitation learning approach for heat pump control, we address generalization to unseen scenarios, optimality, and adaptability through a combination of techniques. By focusing on four key inputs (price, outside temperature, building temperature, and heat pump run statistics), we ensure robust performance even when input values deviate from the training data by mapping these inputs to a standardized range. This, combined with a control heuristic that can override the model's actions if necessary, maintains system stability in unexpected situations. The model is trained on temperature data covering a full year, ensuring reliable control across relevant ranges. Additionally, leveraging MILP outputs during training ensures that learned policies remain close to optimality, while retraining allows the model to adapt to significant system changes such as retrofits or new heat pump installations.

5. Results

5.1. Scenarios for the analysis

We use the air-source heat pump *Compress 6800i AW* [22] 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 [20]. 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 °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 m² and a concrete width of 7 cm for the UFH [20]. The lower limit for the temperature is 20.5 °C, while the upper limit is at 23.5 °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 [23] and German time use data [24] 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 [25] 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 [26].

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* [27]. We scaled them up to align with the average electricity price for residential customers in Germany. The prices have an hourly resolution. Dynamic electricity tariffs, based on Day-Ahead market prices, are currently widely available in many European countries [28]. In Germany, the number of providers offering such tariffs is increasing, and by 2025, all electricity suppliers will be required to offer dynamic tariffs to customers [29]. The electricity prices at the Day-Ahead market in Europe for the following 24 h are determined at 12:00 on the current day [30]. Dynamic tariffs based on Day-Ahead market prices, like the one used in this study, will play a pivotal role in the tariff structures of many residential buildings in the future.
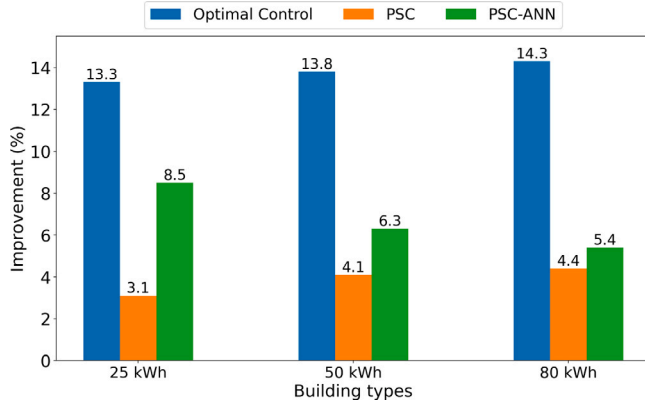
Fig. 5. Improvement regarding the cost reduction compared to the Conventional Control approach averaged over 20 buildings per building type for 20 weeks and three runs.

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 optimization is 30 min. 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 dataset encompasses a range of weather conditions, with daily average temperatures varying from $-9.4\,°C$ to $13.0\,°C$. The time resolution of the control actions Δt (and thus of the training data) is 30 min.

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 min using the GPU *Nvidia GeForce RTX 3090* [31] and an *Intel Core i9-7940X* [32] as CPU. The average time for deriving and executing the control actions for one week (2 decisions per hour) is 3 s with the trained *PSC-ML* control approach, while the exact optimization needs about 40 s. We use the Gurobi solver [33] to solve the optimization problem of the building.

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. It is important to highlight that *PSC* itself is an advanced control approach, and its performance has been positively validated in prior work, where it outperformed all other methods in comparative evaluations [18].

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*

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

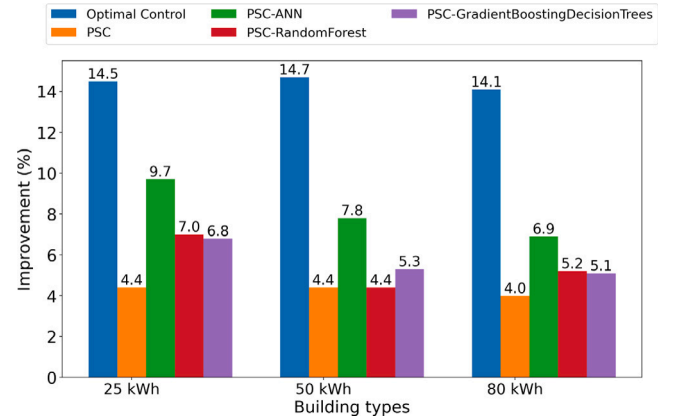| Week	Optimal control	Conventional control	PSC	PSC-ANN
1	269.52	323.33	316.00	317.91
2	155.79	205.12	189.73	178.63
3	273.40	316.44	306.49	295.91
4	178.59	200.89	198.92	185.42
5	325.30	376.47	356.56	356.41
6	228.01	270.14	256.62	246.17
7	281.37	324.84	306.51	300.89
8	233.47	272.67	257.95	257.00
9	133.24	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

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

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.

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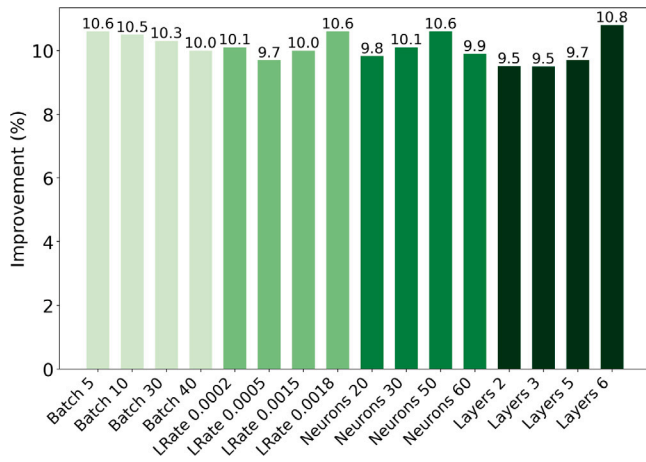
Fig. 7. Improvement compared to the Conventional Control approach for different ANN parameters averaged over one 25 kWh building for 20 weeks.

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

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* [34], for the ML approaches *keras* [35] and *scikit-learn* [36].

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. Our approach can be adapted for multiple buildings by incorporating an aggregate parameter that reflects an entire residential area or local grid, such as the resulting peak load or electricity generation from renewable energy sources. These values can be integrated into both the imitation learning approach and the control heuristic.

We carried out all our investigations in a simulation environment. Thus, we did not apply our proposed approach in real-world experiments. Although our approach has not been explicitly tested on every conceivable variation of building characteristics, occupant behavior, and weather conditions, we have not encountered any cases where it would fail to learn (near) optimal actions. Our method establishes a general framework for combining domain knowledge with machine learning to address control problems in electrical heating devices. We have validated the approach on a diverse dataset that includes variations in building types, occupant behavior, and environmental conditions. This supports our belief that the control algorithm can generalize to 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 from the literature 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.

CRediT authorship contribution statement

Thomas Dengiz: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Max Kleinebrahm:** Writing – review & editing, Validation, Resources, Data curation.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

Appendix A

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

Appendix B. 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.

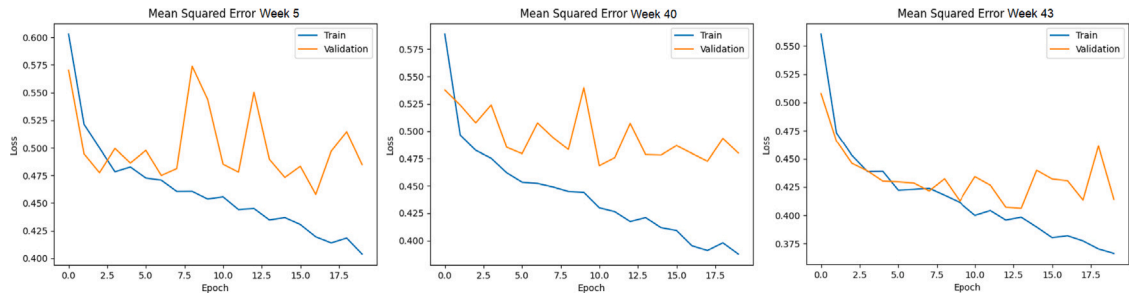**Fig. 8.** RMSE as a function of epochs for three weeks of a 25 kWh building.
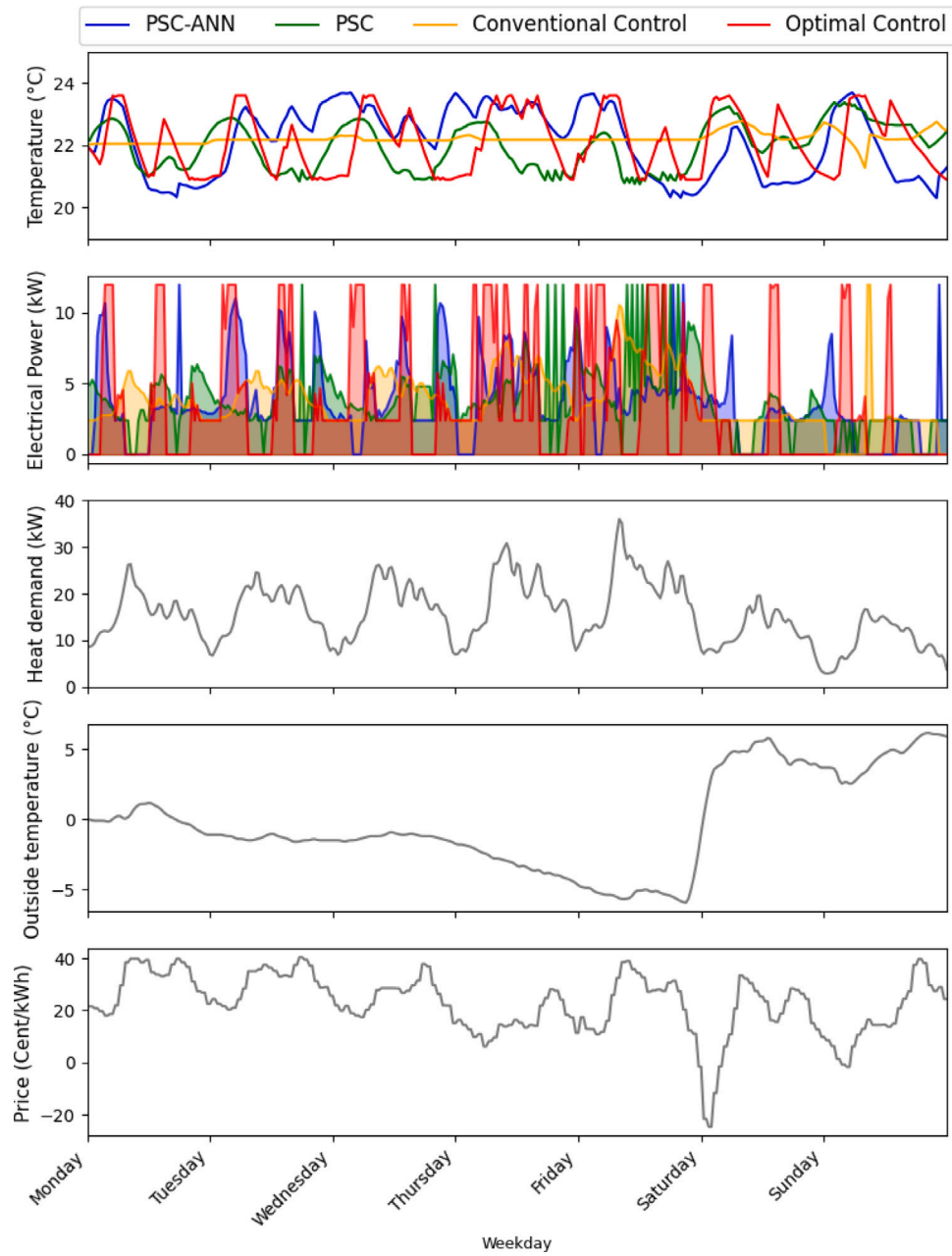

Fig. 9. Exemplary profiles for one week in December for a 50 kWh building.

References

[1] Patteeuw D, Bruninx K, Arteconi A, Delarue E, D'haeseleer W, Helsen L. Integrated modeling of active demand response with electric heating systems coupled to thermal energy storage systems. Appl Energy 2015;151:306–19. http://dx.doi.org/10.1016/j.apenergy.2015.04.014, URL https://www.sciencedirect.com/science/article/pii/S0306261915004535.

[2] Eurostat. Energy use in EU households in 2022 lowest since 2016. 2024, URL https://ec.europa.eu/eurostat/web/products-eurostat-news/w/ddn-20240605-2 [Accessed 23 September 2024].

[3] Dileep G. A survey on smart grid technologies and applications. Renew Energy 2020;146:2589–625. http://dx.doi.org/10.1016/j.renene.2019.08.092, URL https://www.sciencedirect.com/science/article/pii/S0960148119312790.

[4] Dengiz T, Jochem P, Fichtner W. Demand response through decentralized optimization in residential areas with wind and photovoltaics. Energy 2021;223:119984. http://dx.doi.org/10.1016/j.energy.2021.119984, URL https://www.sciencedirect.com/science/article/pii/S0360544221002334.

[5] Antonopoulos I, Robu V, Couraud B, Kirli D, Norbu S, Kiprakis A, Flynn D, Elizondo-Gonzalez S, Wattam S. Artificial intelligence and machine learning approaches to energy demand-side response: A systematic review. Renew Sustain Energy Rev 2020;130:109899. http://dx.doi.org/10.1016/j.rser.2020.109899.

[6] Dinh HT, Kim D. MILP-based imitation learning for HVAC control. IEEE Internet Things J 2022;9(8):6107–20. http://dx.doi.org/10.1109/JIOT.2021.3111454.

[7] Cao D, Hu W, Zhao J, Zhang G, Zhang B, Liu Z, Chen Z, Blaabjerg F. Reinforcement learning and its applications in modern power and energy systems: A review. J Mod Power Syst Clean Energy 2020;8(6):1029–42. http://dx.doi.org/10.35833/MPCE.2020.000552.

[8] Pan X, Zhao T, Chen M, Zhang S. DeepOPF: A deep neural network approach for security-constrained DC optimal power flow. IEEE Trans Power Syst 2021;36(3):1725–35. http://dx.doi.org/10.1109/TPWRS.2020.3026379.

[9] Javed A, Larijani H, Ahmadinia A, Emmanuel R, Mannion M, Gibson D. Design and implementation of a cloud enabled random neural network-based decentralized smart controller with intelligent sensor nodes for HVAC. IEEE Internet Things J 2017;4(2):393–403. http://dx.doi.org/10.1109/JIOT.2016.2627403.

[10] Kim Y-J. A supervised-learning-based strategy for optimal demand response of an HVAC system in a multi-zone office building. IEEE Trans Smart Grid 2020;11(5):4212–26. http://dx.doi.org/10.1109/TSG.2020.2986539.

[11] Dey S, Marzullo T, Zhang X, Henze G. Reinforcement learning building control approach harnessing imitation learning. Energy AI 2023;14:100255. http://dx.doi.org/10.1016/j.egyai.2023.100255.

[12] Zou Z, Yu X, Ergan S. Towards optimal control of air handling units using deep reinforcement learning and recurrent neural network. Build Environ 2020;168:106535. http://dx.doi.org/10.1016/j.buildenv.2019.106535.

[13] Zhang X, Jin X, Tripp C, Biagioni DJ, Graf P, Jiang H. Transferable reinforcement learning for smart homes. In: Proceedings of the 1st international workshop on reinforcement learning for energy management in buildings & cities. RLEM '20, New York, NY, USA: Association for Computing Machinery; 2020, p. 43–7. http://dx.doi.org/10.1145/3427773.3427865.

[14] Dinh HT, haeng Lee K, Kim D. Supervised-learning-based hour-ahead demand response for a behavior-based home energy management system approximating milp optimization. Appl Energy 2022;321:119382. http://dx.doi.org/10.1016/j.apenergy.2022.119382.

[15] Ahmed MS, Mohamed A, Shareef H, Homod RZ, Ali JA. Artificial neural network based controller for home energy management considering demand response events. In: 2016 International conference on advances in electrical, electronic and systems engineering. ICAEES, 2016, http://dx.doi.org/10.1109/ICAEES.2016.7888097.

[16] Gao S, Xiang C, Yu M, Tan KT, Lee TH. Online optimal power scheduling of a microgrid via imitation learning. IEEE Trans Smart Grid 2022;13(2):861–76. http://dx.doi.org/10.1109/TSG.2021.3122570.

[17] López KL, Gagné C, Gardner M-A. Demand-side management using deep learning for smart charging of electric vehicles. IEEE Trans Smart Grid 2019;10(3):2683–91. http://dx.doi.org/10.1109/TSG.2018.2808247.

[18] Frahm M, Dengiz T, Zwickel P, Maaß H, Matthes J, Hagenmeyer V. Occupant-oriented demand response with multi-zone thermal building control. Appl Energy 2023;347:121454. http://dx.doi.org/10.1016/j.apenergy.2023.121454, URL https://www.sciencedirect.com/science/article/pii/S0306261923008188.

[19] Hussein A, Gaber MM, Elyan E, Jayne C. Imitation learning: A sur- vey of learning methods 2017;50(2). https://doi.org/10.1145/3054912.

[20] Dengiz T, Jochem P, Fichtner W. Demand response with heuristic control strategies for modulating heat pumps. Appl Energy 2019;238:1346–60. http://dx.doi.org/10.1016/j.apenergy.2018.12.008.

[21] Verbraucherzentrale NRW eV. Energieausweis: Was sagt dieser Steckbrief für Wohngebäude aus?. 2023, URL https://www.verbraucherzentrale.de/wissen/energie/energetische-sanierung/energieausweis-was-sagt-dieser-steckbrief-fuer-wohngebaeude-aus-24074, [Accessed 23 September 2024].

[22] Bosch Thermotechnik GmbH. Luft-Wasser-Wärmepumpe Compress 6800i AW. 2024, URL https://www.bosch-homecomfort.com/de/de/ocs/wohngebaeude/compress-6800i-aw-19312695-p/, [Accessed 23 September 2024].

[23] Richardson I, Thomson M, Infield D, Clifford C. Domestic electricity use: A high-resolution energy demand model. Energy Build 2010;42(10):1878–87. http://dx.doi.org/10.1016/j.enbuild.2010.05.023.

[24] Statistisches Bundesamt (Destatis). Einkommen, Konsum und Lebensbedingungen - Zeitverwendung. 2024, URL https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Einkommen-Konsum-Lebensbedingungen/Zeitverwendung/_inhalt.html#sprg234984 [Accessed 24 April 2024].

[25] International Organization for Standardization. ISO 13790:2008 - Energy performance of buildings. 2018, URL https://www.iso.org/standard/41974.html [Accessed 06 May 2024].

[26] Institut Wohnen und Umwelt GmbH. Deutsche Wohngebäudetypologie. Beispielhafte Maßnahmen zur Verbesserung der Energieeffizienz von typischen Wohngebäuden – zweite erweiterte Auflage –. 2015, URL https://www.episcope.eu/downloads/public/docs/brochure/DE_TABULA_TypologyBrochure_IWU.pdf [Accessed 24 April 2024].

[27] Entso-E Transparency Platform. Day-ahead prices. 2023, URL https://transparency.entsoe.eu/, [Accessed 24 April 2024].

[28] Hildermeier J, Burger J, Jahn A, Rosenow J. A review of tariffs and services for smart charging of electric vehicles in europe. Energies 2023;16(1). http://dx.doi.org/10.3390/en16010088, URL https://www.mdpi.com/1996-1073/16/1/88.

[29] for Economic Affairs GFM, Action C. The cabinet adopts relaunch of the digitisation of the energy transition and paves the way for accelerated smart meter rollout. 2024, URL https://www.bmwk.de/Redaktion/EN/Pressemitteilungen/2023/01/20230111-the-cabinet-adopts-relaunch-of-the-digitisation-of-the-energy-transition-and-paves-the-way-for-accelerated-smart-meter-rollout.html [Accessed 25 September 2024].

[30] EPEX SPOT. Day-ahead trading. 2024, URL https://www.epexspot.com/en/tradingproducts [Accessed 25 September 2024].

[31] NVIDIA Corporation. Geforce RTX 3090 family. 2024, URL https://www.nvidia.com/en-gb/geforce/graphics-cards/30-series/rtx-3090-3090ti/, [Accessed 24 April 2024].

[32] 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, [Accessed 24 April 2024].

[33] Gurobi Optimization, LLC. Gurobi optimizer reference manual. 2023, URL https://www.gurobi.com, [Accessed 24 April 2024].

[34] Hart WE, Watson J-P, Woodruff DL. Pyomo: modeling and solving mathematical programs in python. Math Program Comput 2011;3(3):219–60. http://dx.doi.org/10.1007/s12532-011-0026-8.

[35] Chollet F, et al. Keras. 2015, https://keras.io.

[36] Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: Machine learning in python. J Mach Learn Res 2011;12:2825–30.