Development of a multi-step approximative model predictive controller for flexibility quantification in building energy systems

Sarah Leidolf [a,b], Antonia Hense, Martin Rätz[a], Dirk Müller[a]

[a] RWTH Aachen University, E.ON Energy Research Center, Institute for Energy Efficient Buildings and Indoor Climate, Germany
[b] sarah.leidolf@eonerc.rwth-aachen.de, CA

Abstract:

The increasing use of electrified heating in buildings presents a opportunity for load shifting, which is essential for balancing power grids with volatile renewables. However, to unlock this potential for commercial use in flexibility markets, a building's energy flexibility must be accurately quantified. A proven approach involves three model predictive controllers (MPCs): one MPC runs a cost-optimized baseline operation, while two so-called shadow MPCs estimate the maximum and minimum possible loads (i.e., the flexibility). However, implementing this is challenging due to the time-consuming creation of physics-based models and the substantial computing power the MPCs require.

As a computationally efficient solution, this research focuses on approximative MPC (AMPC), which replace the MPC optimization problem by directly learning the optimal control. The main obstacle is that AMPCs function as single-step controllers, but assessing flexibility requires a predictive view over a future horizon. This study's central objective is thus to develop an AMPC capable of performing both system control and multi-step prediction for flexibility quantification.

The methodology was demonstrated using a simulated single-zone building model with a heat pump. We trained various decision tree (DT) and artificial neural network (ANN) models, to replicate the heat pump control decisions. The optimal models are selected based on criteria of user discomfort, operational cost, and the plausibility of flexibility values. Finally, the accuracy of the energy flexibility predictions is evaluated by combining three AMPCs: one baseline and two shadow AMPC. These are then compared with the predictions of the physical reference model. Overall, the study demonstrates that the developed methodology can be used to predict the flexibility of buildings. In this example, operational optimisation can be accurately approximated using ANNs. However, DTs are more suitable for predicting trends in flexibility events.

Keywords:

Approximative model predictive control; Building energy system; Direct quantification; Energy flexibility

1. Introduction

The power grid is currently undergoing significant changes due to the increased integration of renewable energy sources [1]. Concurrently, the electrification of heating demand is being heavily promoted, particularly within the building sector [2]. The combination of volatile power generation and rising electricity demand creates an imbalance that leads to grid instabilities [3]. However, buildings offer a promising solution to balance these instabilities through their flexibility, utilizing their inherent thermal storage mass and the combination of building systems technology [2].

There are different approaches to utilize this flexibility [4]: Indirect approaches rely on incentive- or price-based demand response for load shifting, but the exact quantity of available flexibility remains unknown. In contrast, explicit approaches allow for potential estimation and price determination. Direct flexibility provides planning security and enables active market participation. For the explicit provision of flexibility, a model predictive control (MPC) based approach is developed by Stegemerten et al. [5]. By operating three MPC in parallel (a baseline controller that optimizes regular operation, and two so-called shadow controllers that estimate the positive and negative flexibility potential) this method can accurately determine both the quantity of flexibility and its price [5].

However, using MPC with local infrastructure is computationally demanding and expensive [6–8]. But finding a scalable alternative for flexibility quantification is crucial as flexibility is a resource that must be aggregated to have a meaningful impact on grid stability [9, 10]. Focusing on a low-cost implementation, this work uses approximative model predictive control (AMPC) [11, 12].

Existing AMPC approaches for flexibility applications are predominantly limited to indirect flexibility analysis, typically achieved through cost or peak load reduction [13, 14]. For instance, Zhu et al. [13] use classification and regression trees to control room temperature and ventilation. They investigate the system's response to a Time-of-Use electricity price signal compared to a rule-based control system. Frison et al. [14] investigate the use of convolutional neural networks to learn the supply temperature of a heat pump. They define a grid support coefficient that indicates how frequently the control strategy operates above or below a grid signal average, but they do not quantify the theoretical maximum flexibility potential.

Furthermore, conventional AMPC typically performs single-step control and does not consider a prediction horizon. Because a multi-step prediction is necessary to adapt the three-MPC-approach for predicting flexibility, a multi-step AMPC must be developed. Therefore, this paper makes two main contributions. Firstly, the development of a multi-step AMPC. Secondly, it demonstrates how flexibility can be directly quantified using AMPC. This is achieved by incorporating the AMPC into the flexibility quantification method, as well as training a baseline controller and two shadow controllers.

2. Methodology

The methodology developed in this paper is based on the three-MPC-approach presented by Stegemerten et al. [5]. In this approach, a multi-agent system is established. This system includes a predictor for forecasting boundary conditions such as solar irradiation, a building energy system (BES) that communicates its states, and three MPC. The baseline MPC optimizes the operation for cost efficiency and communicates the control variables to the BES. The other two controllers, referred to as shadow MPC, optimize the operation during a specified flexibility event to either maximize electrical power (negative flexibility) or minimize electrical power (positive flexibility).

In addition to the duration of the flexibility event, two further characteristic time periods can be defined. First is the market time, which forces the shadow MPCs to operate identically to the baseline; during this period, a market offer can be accepted. This is followed by a preparation time, which serves to maximize flexibility. For example, a thermal zone can be preheated during the preparation time so that the power consumption can be further reduced during the actual flexibility event. The electrical power predictions resulting from these three optimizations are communicated to a flexibility agent, which uses them to calculate the potential amount of flexibility.

The modifications made in this work consist of replacing the MPCs with AMPCs. And as an AMPC learns the optimal control decision directly, the resulting electrical power remains unknown. Therefore, a calculation from the control variable to the corresponding electrical power must be performed. As previously mentioned, a classic AMPC only controls one step in advance. However, as direct flexibility quantification requires an electrical power forecast over a defined prediction horizon, a multi-step prediction must be introduced for AMPC. This concept is presented in Figure 1.

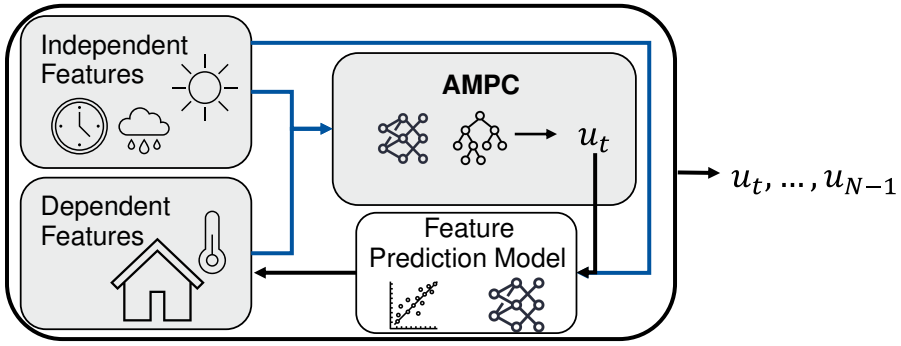

Figure 1. *Concept of a multi-step approximative model predictive controller*

In this paper the control decisions are learned using decision trees (DTs) and artificial neural networks (ANNs) to produce an optimal control u_t based on specified input features. These features can be divided into two categories: Independent features are variables that are not affected by the control decision. Examples of these are weather conditions or time features. The other category comprises dependent features, i.e. variables that influence each other and the control. An example of this is the room temperature which shall be controlled.

In order to make a multi-step decision, predictions of the input features must be available. For weather condi-

tions, for example, predictions from weather services can be used. However, for dependent variables a model is needed to determine how the variable changes as a function of the independent features and control. To avoid the increased modeling effort of physical modeling, linear regression models or ANNs can be used. Using this feature prediction model, the dependent variable can be determined for future steps, which can then inform the control decision.

2.1. AMPC model development

The evaluation of the suitability of AMPC for flexibility quantification is conducted through a three-step methodology, as outlined in Figure 2. The initial step is *model training*. This process begins with the selection of relevant input features and the identification of potential dependent features. Subsequently, the hyperparameters, which vary depending on the chosen training algorithm, are selected. It is important to note that the model training procedure is performed separately for the baseline and the shadow AMPCs.

The second step is *model selection*. For this purpose, the trained models are simulated over a short period in conjunction with the BES to assess their performance. The selection criteria differ depending on whether a baseline or a shadow AMPC is being evaluated. The baseline AMPC model is selected based on two key performance indicators (KPIs): operational costs and user discomfort. For the evaluation of the shadow AMPCs it should be noted that in this context, the reference baseline is a physics-based MPC. This approach is chosen to isolate the performance of the shadow AMPC models from that of the baseline, thereby excluding the influence of the baseline approximation deviations. As no flexibility event is dispatched during this selection phase, the assessment focus on the quality of the shadow models' predictions. A primary consideration is whether user comfort can be maintained during a potential flexibility activation. Therefore, the discomfort is calculated for each time step over the respective prediction horizon and then averaged across all time steps. Furthermore, the investigation assesses whether the shadow models correctly predict the available flexibility. This involves a plausibility check of the predicted energy flexibility. By definition, a quantity of energy flexibility must have a positive sign. Consequently, a evaluation criterion is the proportion of negatively predicted energy flexibility values, as these indicate implausible predictions.

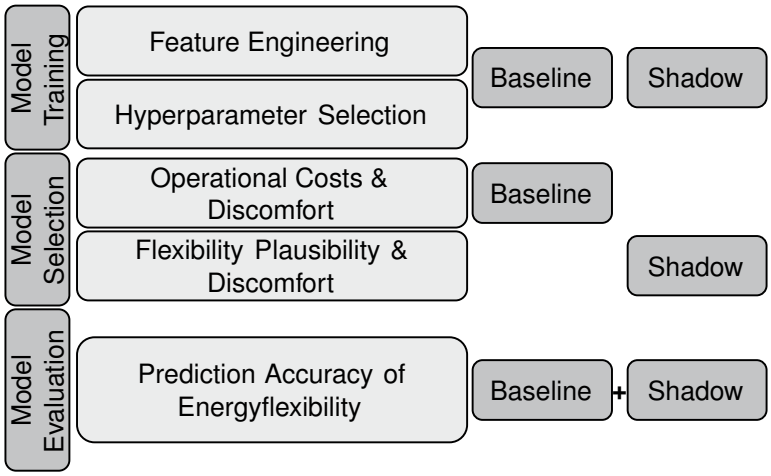

Figure 2. *Pipeline for AMPC model training, selection and evaluation*

The final step is *model evaluation*. For this, both the baseline AMPC and the shadow AMPCs are integrated into the agent-based framework. To ensure comparability of flexibility predictions, the physics-based benchmark runs simultaneously in 'shadow mode'. In this mode, the benchmark calculates the potential flexibilities under the assumption that the system is being controlled by the baseline AMPC. This eliminates prediction inaccuracies caused by different baseline operations between MPC and AMPC.

3. Application

A simplified simulation model is used to demonstrate the methodology. This consists of a zone with underfloor heating, which is supplied by a air-water heat pump. The BES and the 'Teacher' MPC, which serves as the training basis for the AMPC, are described in more detail below. Finally, the AMPC and the boundary conditions under which the investigation is carried out are described.

3.1. Building energy system

The thermal zone is described by a differential equation of temperature T_{Zone}. This takes into account solar heat $\dot{Q}_{\mathrm{S}}$, internal heat gains $\dot{Q}_{\mathrm{int}}$, transmission heat flow $\dot{Q}_{\mathrm{T}}$, and the heat input from the underfloor heating $\dot{Q}_{\mathrm{UFH}}$. Thermal inertia is represented by heat capacity C_{Zone}. Solar heat is calculated based on solar

radiation and the area of the windows. Internal gains depend on the number of people in the zone. Transmission heat flow is determined by the temperature difference between the interior and exterior air. The heat from underfloor heating is proportional to the relative speed of the heat pump u_{HP}, which is the control variable in this application.

$$\frac{dT_{\mathrm{Zone}}}{dt} = \frac{\dot{Q}_{\mathsf{S}} + \dot{Q}_{\mathsf{int}} + \dot{Q}_{\mathsf{T}} + \dot{Q}_{\mathsf{UFH}}}{C_{\mathsf{Zone}}} \tag{1}$$

Since building dynamics typically show non-linear characteristics, the heat pump model is also designed to show non-linear behavior. The electrical power consumption of the heat pump depends on its coefficient of performance (COP). To accurately model this, the COP calculation is modified by introducing a correction factor and adjusting the Carnot COP. This correction factor, which depends quadratically on the relative heat pump speed, is incorporated to better represent the part-load efficiency. Furthermore, the Carnot COP is adapted by defining the supply temperature as a linear function of the zone temperature and relative speed. Additionally, the temperature difference between the supply and return is approximated using a logistic function to ensure numerical stability within the simulation model.

3.2. Teacher MPC

The process model employed by the MPC is identical to the model presented in Subsection 3.1. Because it uses the same underlying physics, this MPC can be used as an ideal benchmark for performance evaluation.

The objective function for the baseline controller presented in Equation 2, is formulated to minimize a weighted sum of two terms. These include user discomfort, which is managed via a slack variable s and the operating costs, which are calculated from the electrical power demand P_{el} and the corresponding electricity price c_{el}. Each of these sub-objectives is assigned a distinct weighting factor w_i allowing for the trade-off between several goals. The control input for the heat pump u_{HP} is constrained to values between 0 and 1.

$$J = \min_{u_0,\ldots,u_{t_{\mathrm{N}}-1}} \sum_{k=0}^{t_{\mathrm{N}}-1} \left[\left(w_s \cdot s_k^2 + w_{P_{\mathrm{el}}} \cdot c_{\mathrm{el},k} \cdot P_{\mathrm{el},k}\right) \cdot t_{\mathrm{s}}\right] \tag{2}$$

The teacher shadow MPC is designed with an objective, which is activated during a flexibility event. Outside of the event, it follows the same objective function as the baseline MPC. For the provision of negative flexibility, its objective is to maximize electrical power consumption. Conversely, when providing positive flexibility, its objective is to minimize power consumption. A critical constraint throughout any flexibility event is the user comfort. For a more detailed explanation of the optimization problem's formulation, the reader is referred to the work of Stegemerten et al. [5].

3.3. AMPC

In the presented use case, there is a bidirectional dependency: the zone temperature depends on the control input, and the selection of the optimal control input depends on the zone temperature. Therefore, as described in Subsection 2, a feature prediction model is needed for the dependent feature.

To achieve this, a 30-day dataset was generated for January 2024. This dataset is based on a system regulated by a PID controller and uses weather data from Aachen, Germany. An ANN with a single layer containing 32 neurons and a softplus activation function was trained. Training was conducted for 1000 epochs with a batch size of 32. The model learns to predict the gradient of the zone temperature as a function of solar irradiation, ambient temperature and heat pump speed (each with a time lag), as well as the user occupancy profile. To account for the non-deterministic nature of ANNs, this architecture was trained 30 times. Each resulting model was then evaluated through a two-day simulation in conjunction with a datadriven MPC. The model that caused the least user discomfort during this evaluation period was selected as the zone temperature model.

3.4. Boundary conditions

To generate the training data set for the AMPC, the teacher MPC is simulated over a period of one year. Weather data from Aachen, Germany, from the year 2024 is used for this purpose. The teacher MPC runs with a step size of 900 s and a prediction horizon of 13.25 h. The weights for the objective function in Equation 2 are $w_{\mathrm{s}} = 100$ and $w_{\mathrm{P_{el}}} = 0.01$. The following characteristic times were specified for the shadow MPC: market time = 15 min, preparation time = 30 min and flexibility event duration = 2 h.

As pictured in Figure 3, the model selection is conducted during the week preceding the main evaluation. The selection process is performed with an initialization period of half a day, followed by a selection duration of two days. Subsequently, the model evaluation is carried out over a seven-day period, which is preceded by a one-day initialization phase. The evaluation of the AMPC is conducted for a typical heating week in the year 2025, using weather data for Aachen, Germany. This representative week was determined by clustering the environmental conditions of the entire year based on two key variables: ambient temperature and solar irradiation. The week corresponding to the cluster centroid closest to the 20th percentile of ambient temperature and

the 50th percentile of solar irradiation was then selected. This procedure identified the period from Monday, February 17th to Sunday, February 23rd. For the entirety of the evaluation, a constant electricity price of 40 ct/kWh is assumed.

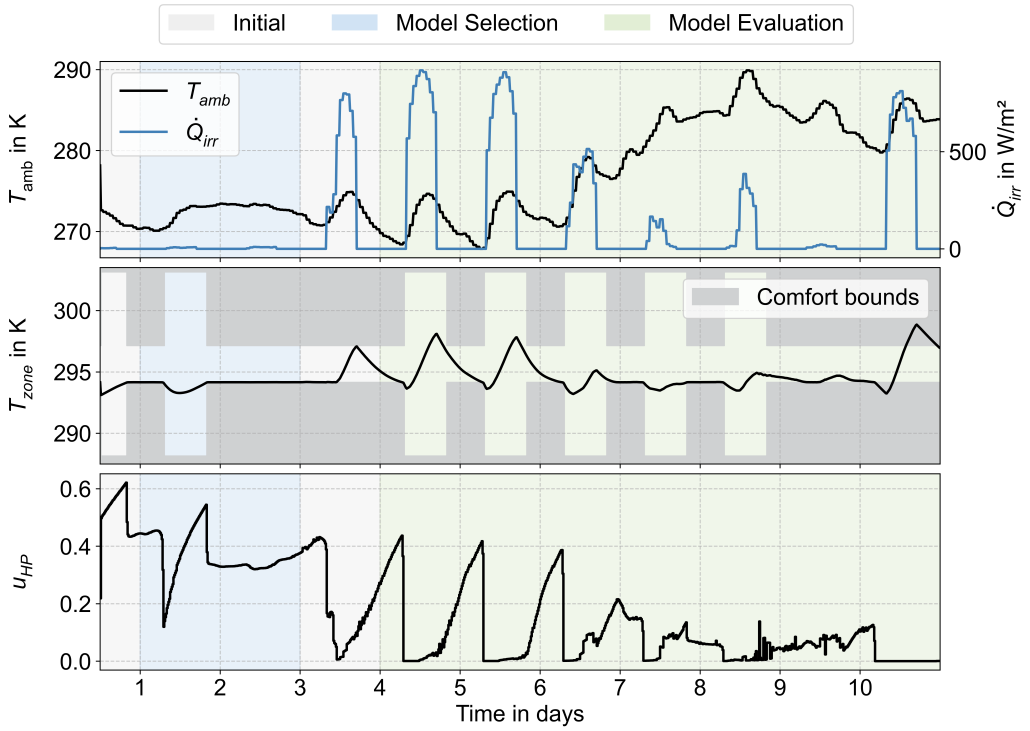

Figure 3. *Boundary conditions and teacher MPC reference results of the evaluation interval*

4. Results

The results are structured according to the AMPC model development pipeline presented in Subsection 2.1. First, the setup is described, i.e. the selection of features and hyperparameters for the use case. Then, the training of the baseline and shadow AMPC models is presented. Next, the AMPC model is selected. Finally, the models are evaluated based on their prediction accuracy of the flexibility potential.

4.1. Setup and model training

4.1.1. Baseline

Preliminary investigations have shown that ambient temperature, solar irradiation, and zone temperature are relevant input features for the prediction of the relative heat pump speed. To account for the temporal dependency on the preceding time step, each of these variables is also included with one time lag. Within the MPC optimization problem, user comfort is a critical constraint that must be consistently maintained. The specific comfort limits vary depending on the time of day and the day of the week. To transmit this relevant contextual information to the AMPC models, specific time-based features are engineered. These include a cyclical encoding of time using sine and cosine functions, as well as a boolean flag to indicate whether a given time step falls on a weekend or a weekday.

The hyperparameter selection for the **DT** models begins with a tree depth analysis. For this analysis, the dataset was splitted into a 70% training set and a 30% test set. The R^2 score was then calculated separately for both the training and test data across a range of different tree depths. The depth at which the R^2 score on the test data stops to improve is selected as the optimal value, a strategy employed to prevent model overfitting. Preliminary analyses have shown that no significant performance improvements are achieved beyond a tree depth of 9. This depth is then used for a tree pruning analysis, which proceeds in an analogous manner. A pruning parameter is considered optimal if it does not degrade the performance metric on the test dataset. Through this process, a value in the order of $1 \cdot 10^{-6}$ proved to be suitable. To account for statistical variance 30 models were trained using the predetermined depth and pruning parameter (DT_{constr}), with varied random seeds. For comparison, a second set of 30 models was trained, but without any restrictions on depth or pruning (DT_{unconstr}).

For the **ANN** models, a standard hyperparameter configuration is employed, using 1000 training epochs and a batch size of 64. The network architecture is composed of a single layer containing 32 neurons. The activation

function for this layer is varied, with separate models being trained to evaluate the performance of rectified linear unit (ReLU), sigmoid, and tanh functions. For model training, the dataset is splitted into an 80% training, a 10% validation and a 10% testing set. Consistent with the methodology used for the other models, each of these architectural combinations is trained 30 times to account for statistical variance.

Table 1 presents the training results for the five models. The performance is evaluated using the R^2 score and the Root Mean Square Error (RMSE) on both the training and testing sets. Additionally, the variance of these metrics across the different training runs is provided. To maintain clarity in the table, this variance is only reported for values exceeding $1 \cdot 10^{-4}$. The unconstrained DT models exhibit tree depths ranging from 36 to 48 across the various training runs. The iteration yielding the best test R^2 score of 0.96 resulted in a tree depth of 41. However, the metrics generally indicate that the unconstrained DT is prone to overfitting, as evidenced by the significant discrepancy between its performance on the training and testing sets. In contrast, although the constrained DT achieves lower metrics on the training dataset, its performance on the test dataset is similar. The best-performing iteration for the constrained DT yielded a test R^2 score of 0.947. All DT models exhibit very low variance in their training metrics. The performance metrics of the ANN models on the training dataset are slightly lower than those of the DTs; However, their metrics on the test dataset are generally superior. The discrepancy between the training and test metrics for the ANNs is minimal, which suggests a reduced tendency to overfitting. Similarly, the ANNs demonstrate very low variance across the individual training iterations.

Table 1. *Performance comparison of DT and ANN (mean value (variance) of $n = 30$ runs)*

		R^2		RMSE	
Model	**Configuration**	**Train**	**Test**	**Train**	**Test**
Decision Trees					
$DT_{unconstr}$	depth $\in [36, 48]$	1 (0)	0.92 (0.0003)	0 (0)	0.03 (0)
DT_{constr}	depth=9, $\alpha_{ccp} = 1 \cdot 10^{-6}$	0.96 (0)	0.92 (0.0004)	0.02 (0)	0.03 (0)
Artificial Neural Networks					
ANN_{relu}	activation: relu	0.94 (0)	0.94 (0.0003)	0.027 (0)	0.027 (0)
ANN_{sig}	activation: sigmoid	0.93 (0)	0.93 (0.0002)	0.029 (0)	0.029 (0)
ANN_{tanh}	activation: tanh	0.94 (0)	0.92 (0.0006)	0.027 (0)	0.03 (0)

It is shown that when employing data-driven models within an MPC framework, strong performance on offline training metrics is not necessarily indicative of the model's closed-loop performance [15, 16]. Therefore, the final selection of a model for the AMPC is not based on offline metrics alone. Instead, it is determined by simulating each model in combination with the BES over a dedicated selection period, as detailed further in Subsection 4.2.

4.1.2. Shadows

Since the training dataset does not contain any actually dispatched flexibility events, actual values cannot be used for training. Instead, the predictions generated by the teacher shadow MPC are utilized as training data. For the evaluation of flexibility, the initial time steps of the prediction horizon are of particular importance. Consequently, the training does not utilize the full prediction horizon of 13.25 h. Instead, a shorter horizon of 3.25 h is employed, a duration which covers the period up to two time steps after the flexibility event. Notably, no time-lagged features are used for this training. To provide the models with contextual information about the flexibility event, time-based flags are introduced. These flags denote the different phases: (1) Market clearing, (2) preparation, (3) flexibility event, and (4) after flexibility event. This approach enables the shadow AMPC models to learn the characteristic profiles associated with a flexibility event. It is important to note that separate models are trained for the positive and negative shadow AMPCs.

Analogous to the methodology for the baseline models, two DT and three ANN variants are employed, with each model being trained 30 times. For the negative shadow AMPC, the unconstrained DTs resulted in tree depths between 51 and 68, while for the positive shadow AMPC, depths ranged from 55 to 69. For the constrained DT, a pruning value of $1 \cdot 10^{-6}$ was found to be suitable, consistent with the baseline models. Preliminary analyses of tree depth indicated optimal values of 9 for the negative shadow AMPC and 13 for the positive one. The hyperparameters for the ANN models are identical to those used for the baseline.

As previously mentioned, the performance of these models becomes fully apparent only when they are deployed within a simulation. Therefore, a detailed discussion of the offline training metrics is omitted at this point. Instead, the process of model selection for the shadow AMPC will be addressed in detail in the following chapter.

4.2. Model selection

For the model selection, a 2.5-day period is utilized, as illustrated Figure 3. The first half-day of this period serves as an initialization phase and is therefore excluded from the evaluation. This specific timeframe was selected to include both a weekday and a weekend, ensuring that the effects of daily and weekly periodicity are considered during the selection process.

4.2.1. Baseline

All 30 iterations of the five model variants are simulated over the designated model selection period. The relevant KPIs for evaluating the models' performance are user discomfort and operational costs. The results are summarized for all runs in Figure 4. A physics-based MPC is included as reference.

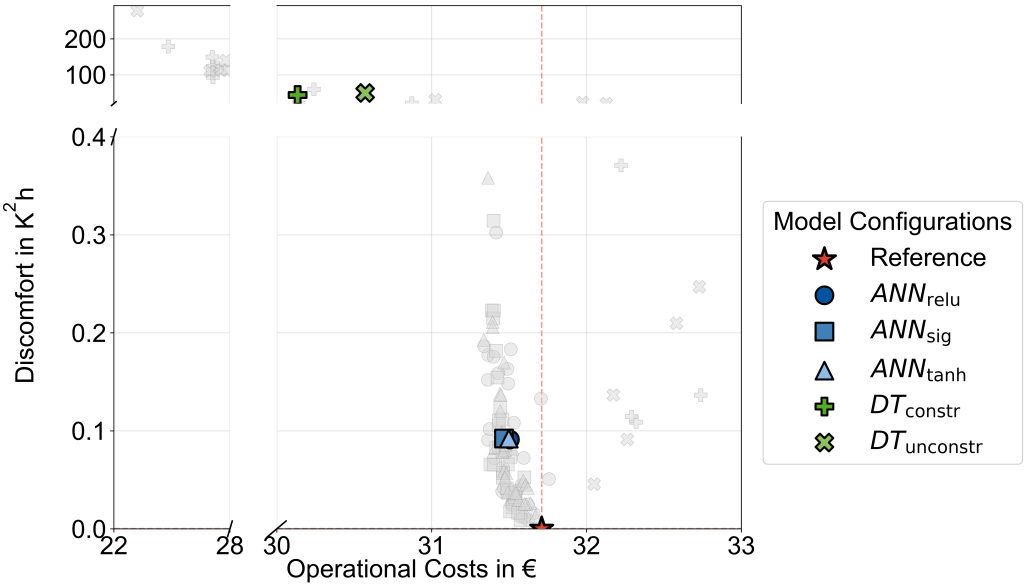

Figure 4. *Operational costs and discomfort for different model configurations and their iterations (grey) with the mean value highlighted (colored)*

The DT models show a high degree of variance, the discomfort values for the $DT_{unconstr}$ range from 0.046 to 278 K^2h, while those for the DT_{constr} range from 0.11 to 179 K^2h. The unconstrained DT exhibits higher variance in both discomfort and costs. The DT models in average fail to maintain user comfort. While they often result in significantly lower operating costs, this is achieved primarily by violating the lower comfort boundary.

In contrast, the ANNs perform a lot better. The choice of activation function does not appear to have a distinct influence. All ANN variants demonstrate good performance, achieving results that are on the same order of magnitude as the physics-based reference. Furthermore, the variance across the ANN training iterations is low, with discomfort values between approx. 0 to 0.35 K^2h and costs between 31.3 and 31.75 €. The variance in discomfort is comparable across all ANN configurations. However, the variance in operating costs is lowest for the ANN model utilizing the sigmoid activation function.

For the following model evaluation one DT and one ANN model are selected. From the DT, the constrained architecture is favored over the unconstrained version due to its lower performance variance and better average discomfort. The iteration is choosen based on the lowest discomfort. Among the ANNs, all variants exhibited low variance. The choice, therefore, falls on the model utilizing the sigmoid activation function, as it demonstrated the most consistently performance. The selection of models in relation to the findings in Table 1 shows that the metrics do not directly reflect the performance of the models in the MPC. The only apparent influence is the difference between the training and test metrics, which suggests overfitting. This difference is significantly higher for the DTs, which, when combined with the results, suggests that this is a relevant factor. But it is important to acknowledge that the short selection period is limited in its capacity to support broad, generalizable conclusions.

4.2.2. Shadows

As before, all 30 iterations of the five model variants are simulated for the shadow AMPC model selection period. In this phase, the baseline is the physics-based MPC in order to isolate the performance from the shadow AMPC. The relevant KPIs for evaluating the performance are average discomfort over the prediction horizon (the AMPC's prediction horizon is long enough to take the first step after the flexibility event into account (see subsubsection 4.1.2)) and the flexibility plausibility (FP=(100-Flexibility below zero)%). The results are summarized in Figure 5.

The models for the negative shadow AMPC show very good results regarding FP. On average, all iterations predict a minimum of approx. 83% plausible flexibility. The ANN models demonstrate particularly strong performance, with individual iterations achieving 100% plausibility. While all iterations exhibit a higher average discomfort compared to the physics-based reference, the values all lie within a similar order of magnitude. In general, the variance across all models is low.

For the models of the positive shadow AMPC, a more distinct difference between the model types is apparent. The DT models show very good results with respect to FP, whereas the ANN models, on average, only achieve a FP of approx. 65-75%. The average discomfort for the ANN models using ReLU and sigmoid activation functions is significantly lower than that of the DT models, although with some individual outliers. The iterations using the tanh activation function exhibit a considerably larger variance.

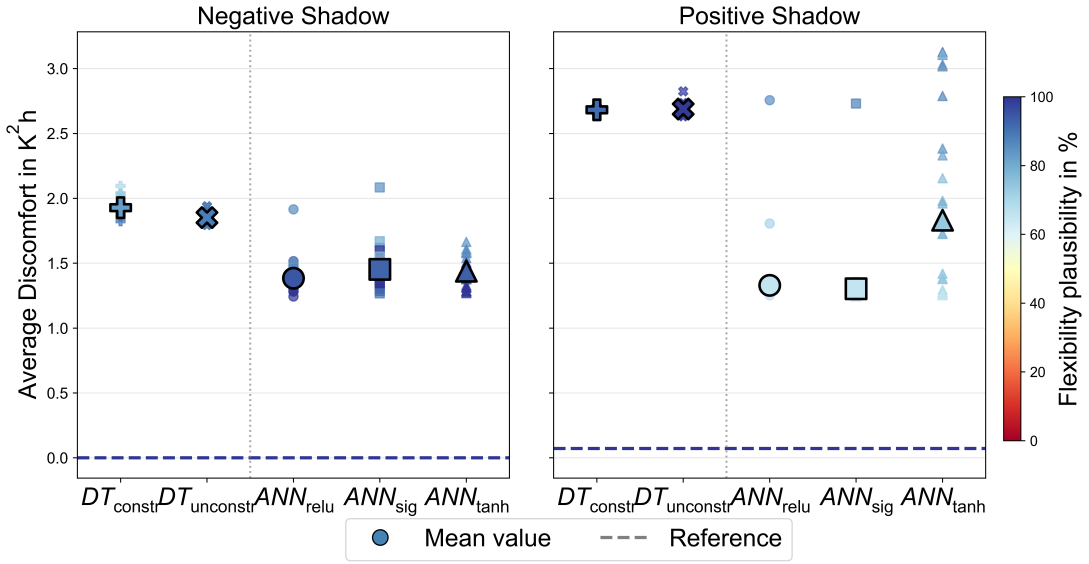

Figure 5. *Energy flexibility plausibility and average discomfort for different model configurations and their itertaions with the mean value highlighted*

In addition, performance is analyzed in comparison to the training metrics (see Subsection 4.1). For the negative shadow models, the DTs achieved significantly better training metrics than the ANNs, yet this did not translate to the performance. Furthermore, while the offline metrics of the unconstrained DT suggested a tendency to overfit, this behavior was not observed in this results. The positive shadow models generally exhibit better offline training metrics than the negative models. However, their performance during the selection phase is, at least in part, worse to that of the negative models. An exception to this are the DT models, whose strong training metrics indeed reflected in a better performance during the selection. Overall, these findings indicate that training metrics do not serve as a reliable indicator for the performance of the shadow models.

For the evaluation, the constrained and unconstrained DT models with the best FP are selected for both the negative and positive shadow. Regarding the ANN models, the architecture using the tanh activation function is excluded from the evaluation due to its high variance in the positive shadow AMPC. For the negative shadow, the iteration with the highest FP is selected for both the ReLU and sigmoid ANNs. For the positive shadow ANNs, the selection process is modified to account for the observed outliers. An iteration is selected that exhibits an average discomfort level, and from this subset, the one with the highest FP is chosen.

4.3. Model evaluation

The evaluation is conducted using four AMPC combinations. The combinations and metrics of the selected iteration are summarised in Table 2. Additionally in Figure 6 the room temperature of the chosen DT and ANN baseline model is shown compared to the physics-based MPC. Overall, the AMPCs follow the same trend as the MPC and can meet comfort requirements. However, the DT model causes discomfort during the first two days when there is high solar irradiance. However, the performance over the rest of the week is comparable to that of the MPC.

The evaluation is conducted for the sample week with an initialization phase of 24 hours (see Figure 3). As described in Subsection 2.1, the three physical benchmark MPCs run in parallel with those of the AMPC. The AMPC's baseline controls the system, thereby eliminating prediction differences caused by variations in the operating point. The evaluation of the flexibility prediction is based on a comparison of the power predictions from the physics-based benchmark and the AMPC over the prediction horizon, which extends to just after the flexibility event. This approach allows for an assessment not only of whether the shadow models correctly

predict the flexibility during the event, but also whether they accurately capture the pre- and post-event system behavior, such as pre-cooling or pre-heating.

Table 2. *Selected AMPC models for flexibility evaluation and their metrics from selection phase (Discomfort=Disc)*

		Baseline		**Negative Shadow**			**Positive Shadow**		
Name	Model	Disc	Cost	Model	Disc	FP	Model	Disc	FP
$DT_{con/uncon}$	DT_{constr}	0.109	32.32	$DT_{unconstr}$	1.880	89.58	$DT_{unconstr}$	2.675	100.00
$DT_{con/con}$	DT_{constr}	0.109	32.32	DT_{constr}	1.874	89.10	DT_{constr}	2.678	91.67
$ANN_{sig/sig}$	ANN_{sig}	0.009	31.57	ANN_{sig}	1.343	100.00	ANN_{sig}	1.331	74.48
$ANN_{sig/relu}$	ANN_{sig}	0.009	31.57	ANN_{relu}	1.326	100.00	ANN_{relu}	1.290	70.31

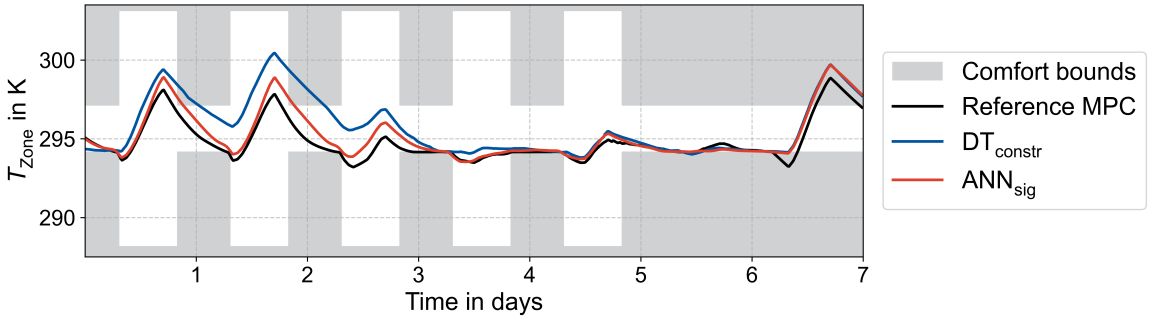

Figure 6. *Room temperature of chosen AMPC models and of the reference MPC*

In Figure 7 the RMSE between the power predictions of the MPC and the AMPC for both the negative and positive shadows is shown. It is apparent that for the negative flexibility predictions, the DT models sometimes produce large errors, although at other times their predictions align almost perfectly with the benchmark. In contrast, the AMPCs based on ANNs do not exhibit such large errors in individual predictions; instead, they maintain a relatively constant error throughout the simulation. Regarding positive flexibility, a cyclical error is observable across all model variants. The ANN-based combinations tend to exhibit a larger error during these cyclical periods.

To investigate specific effects in greater detail, the model predictions are examined at two exemplary time steps. These instances are indicated in Figure 7 by vertical lines labeled (1) and (2). Figure 8 depicts the predictions at time step (1). In the first two subplots, the same DT model is used for the baseline. This baseline model cause discomfort, a condition that is recognized by the physics-based MPC. Consequently, the MPC predicts no available negative flexibility in order to avoid further discomfort. However, the AMPC is unable to recognize this discomfort and therefore does not restrict the flexibility, opting instead to set the heat pump to its maximum speed for the negative shadow. This results in a significant discrepancy between the MPC and AMPC predictions. The error can thus be attributed to the baseline model and the discomfort it generated. When using the ANN-based models for the baseline, the discomfort is less pronounced, and consequently, this effect is less evident. Both models of the negative shadow AMPC set a moderate heat pump speed during the event. The physics-based negative shadow MPC, in contrast, keeps the heat pump off until the final time step of the event to avoid discomfort. For the positive flexibility, the prediction errors are minor across all models. In this specific case, the DT models demonstrate a better ability to identify the pre-heating period. All models correctly switch off the heat pump during the flexibility event itself.

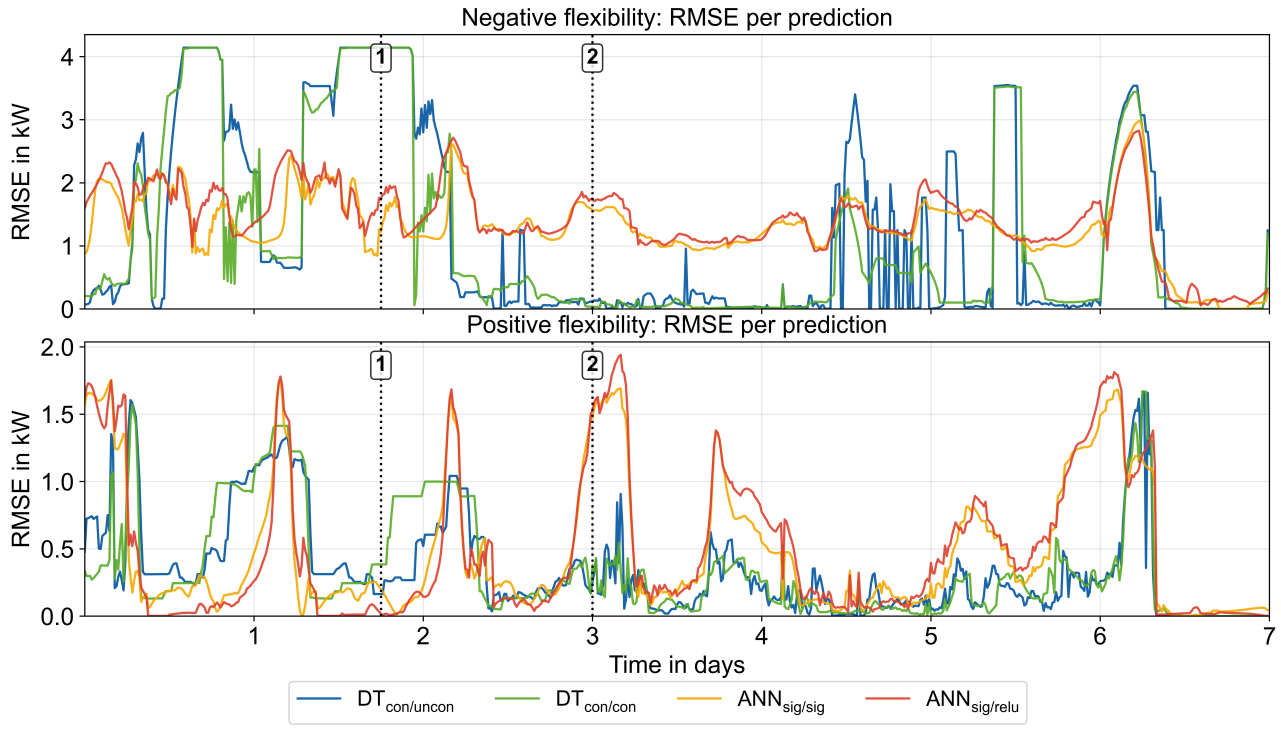

Figure 7. *Root mean square error (RMSE) per prediction for four AMPC combinations for negative (upper) and positive flexibility (lower)*

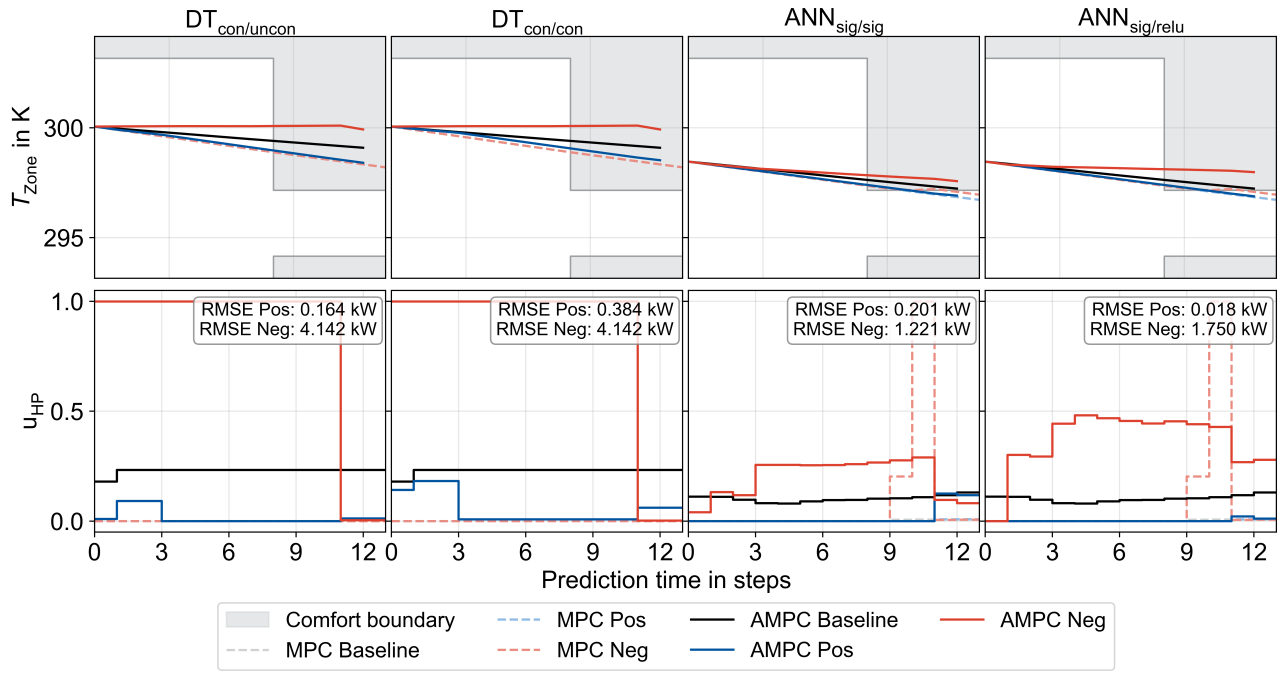

Figure 8. *Single predictions of the zone temperature (upper) and heat pump speed (lower) from the example time step (1)*

The second exemplary time step (2) is shown in Figure 9. In this instance, the DT-based models exhibit good consistency with the benchmark. These models correctly identify both the use of the preparation phase for pre-cooling or pre-heating and the maximization or minimization of the control during the flexibility event itself. In contrast, and as was also observed in the preceding time step, the ANN-based models tend to move between states of 0 and 1, rather than setting the limit values. Consequently, the error is significantly larger in this case. Moreover, the models fail to implement meaningful pre-heating or pre-cooling strategies.

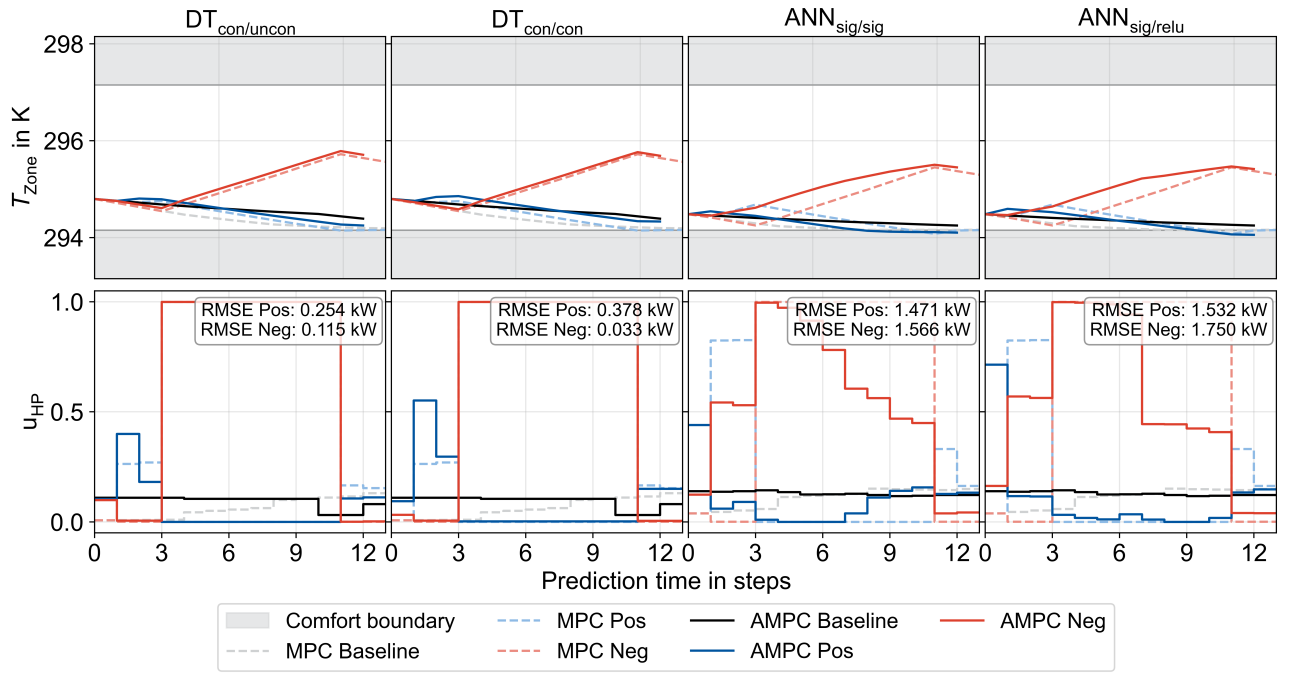

Figure 9. *Single predictions of the zone temperature (upper) and heat pump speed (lower) from the example time step (2)*

The key findings from the model evaluation indicate that DT are well-suited for tasks involving discrete decisions, such as setting the control input to its maximum or minimum value. They also demonstrate a good ability to identify the preparation phase. ANNs, in contrast, tend to produce more continuous control outputs, which results in them not fully leveraging the available minimum or maximum control limits during the flexibility event. Consequently, the preparation period is also not identified as effectively. Based on these observations, it can be concluded that DTs appear to be more suitable for the shadow model task, while ANNs are better suited for the baseline control task. Another observation is that the baseline performance, which reflects the current operating point of the system, has a significant influence on the performance and predictive reliability of the shadow models. This is particularly true when the baseline model drives the system into operating points that are unrepresented in the training dataset. In such cases, the shadow models are forced to extrapolate, which compromises their ability to generate meaningful predictions.

5. Conclusion

This paper addressed the challenge of creating a scalable and computationally efficient method for the direct quantification of energy flexibility in buildings. To this end, a multi-step AMPC framework was developed and integrated into a three-controller structure, replacing the conventional, resource-intensive MPCs. The investigation evaluated the performance of different data-driven models, DT and ANN, for both the baseline and shadow control tasks on a simulated use case.

In conclusion, this work successfully demonstrates the viability of using AMPC for direct flexibility quantification, offering a promising path toward scalable implementation. However, the AMPC models show different strengths and weaknesses. The main finding is that ANNs proved to be better suited for the baseline AMPC, effectively managing user comfort and operating costs through their continuous control output. In contrast, DTs are better suited as shadow AMPC, as their structure is well suited to making the discrete decisions necessary to maximize or minimize performance during a flexibility event and identify preparation phases. A critical finding is the significant impact of the baseline controller's performance on the accuracy of the shadow models; when the baseline operation deviates into states unrepresented in the training data, the reliability of the flexibility prediction is compromised.

Future research should focus on selecting data to avoid extrapolation, as well as testing different data-driven models and ANN layers better suited to identifying discrete features. Further investigation is also needed to improve the robustness of shadow models against variations in baseline operation. For example, other input features could be used to help the model recognize discomfort or learn the difference in control variables from the baseline, rather than the control itself. In addition, determining costs is essential for participating in a flexibility market. The next step is to estimate these costs within the framework and develop a strategy for communicating offers to the market.

Acknowledgments

We gratefully acknowledge the financial support by the Federal Ministry for Economic Affairs and Energy (BMWE), promotional reference 03EN1066A.

Nomenclature

Symbols

α_{ccp}	cost-complexity pruning parameter
C	heat capacity
P_{el}	electrical load
$\dot{Q}$	heat flow
s	slack variable
t_{s}	time step
T	Temperature
u	control variable
w	weight

Subscripts and superscripts

amb	ambient
HP	heat pump
irr	irradiation
int	internal
S	solar
T	transmission
UFH	under floor heating

Abbreviations

AMPC	approximative model predictive control
ANN	artificial neural network
BES	building energy system
COP	coefficient of performance
DT	decision tree
FP	flexibility plausibility
KPI	key performance indicator
MPC	model predictive control
ReLU	rectified linear unit
RMSE	root mean square error

References

[1] *Federal Climate Change Act 2024 (Bundes-Klimaschutzgesetz)*.

[2] *Energy in Buildings and Communities Programme Annex 67 Energy Flexible Buildings*. Tech. rep. International Energy Agency, Dec. 2019.

[3] H. Kondziella and T. Bruckner. Flexibility requirements of renewable energy based electricity systems–a review of research results and methodologies. In: *Renewable and Sustainable Energy Reviews* 53 (2016), pp. 10–22.

[4] G. Reynders, R. A. Lopes, A. Marszal-Pomianowska, D. Aelenei, J. Martins, and D. Saelens. Energy flexible buildings: An evaluation of definitions and quantification methodologies applied to thermal storage. In: *Energy and Buildings* 166 (2018), pp. 372–390.

[5] F. Stegemerten, S. Leidolf, P. Stoffel, and D. Müller. Exploring Flexibility in Building Energy Systems: Agent-Based Quantification and Electricity Tariff Impacts. In: *Energy and Buildings* (2025), p. 115500. DOI: 10.1016/j.enbuild.2025.115500.

[6] P. Henkel, M. Rätz, and D. Müller. Model predictive control for load shifting in supermarket refrigeration. In: *Journal of Physics: Conference Series*. Vol. 3140. 5. IOP Publishing. 2025, p. 052017.

[7] J. Drgoňa, J. Arroyo, I. C. Figueroa, D. Blum, K. Arendt, D. Kim, E. P. Ollé, J. Oravec, M. Wetter, D. L. Vrabie, et al. All you need to know about model predictive control for buildings. In: *Annual reviews in control* 50 (2020), pp. 190–232.

[8] S. White, C. P. Gao, J. Candanedo, I. Sartori, J. Wen, D. Rovas, B. N. Jørgensen, and P. Ruyssevelt. *Annex 81 Final Report*. Tech. rep. EP2025-2192. Australia: CSIRO, 2025. DOI: 10.25919/v98x-8108. URL: https://doi.org/10.25919/v98x-8108.

[9] R. Li, A. J. Satchwell, D. Finn, T. H. Christensen, M. Kummert, J. Le Dréau, R. A. Lopes, H. Madsen, J. Salom, G. Henze, et al. Ten questions concerning energy flexibility in buildings. In: *Building and Environment* 223 (2022), p. 109461.

[10] A. Kathirgamanathan, M. De Rosa, E. Mangina, and D. P. Finn. Data-driven predictive control for unlocking building energy flexibility: A review. In: *Renewable and Sustainable Energy Reviews* 135 (2021), p. 110120.

[11] J. Drgoňa, D. Picard, M. Kvasnica, and L. Helsen. Approximate model predictive building control via machine learning. In: *Applied energy* 218 (2018), pp. 199–216.

[12] L. M. Maier, L. Helsen, and D. Müller. *Approximating optimal control strategies for all-electric building energy systems*. Tech. rep. E. ON Energy Research Center, 2024.

[13] J. Zhu, J. Niu, S. Zhan, Z. Tian, A. Chong, H. Wang, and H. Zhou. A learning-based model predictive control method for unlocking the potential of building energy flexibility. In: *Energy and Buildings* 330 (2025), p. 115299.

[14] L. Frison, S. Paul, T. Koller, D. Fischer, G. Frison, J. Boedecker, and P. Engelmann. Hardware-in-the-loop test of learning-based controllers for grid-supportive building heating operation. In: *IFAC-PapersOnLine* 53.2 (2020), pp. 17107–17112.

[15] P. Stoffel, P. Henkel, M. Rätz, A. Kümpel, and D. Müller. Safe operation of online learning data driven model predictive control of building energy systems. In: *Energy and AI* 14 (2023), p. 100296.

[16] L. Di Natale, B. Svetozarevic, P. Heer, and C. N. Jones. Towards scalable physically consistent neural networks: An application to data-driven multi-zone thermal building models. In: *Applied Energy* 340 (2023), p. 121071.