Article

Huilong Wang 1,2,3,4, Daran Mai 4, Qian Li 4 and Zhikun Ding 1,2,3,4,*

1 Key Laboratory of Coastal Urban Resilient Infrastructures (MOE), Shenzhen University, Shenzhen 518060, China; wanghuilong@szu.edu.cn
2 Sino-Australia Joint Research Center in BIM and Smart Construction, Shenzhen University, Shenzhen 518060, China
3 Shenzhen Key Laboratory of Green, Efficient and Intelligent Construction of Underground Metro Station, Shenzhen University, Shenzhen 518060, China
4 Department of Construction Management Science, College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518060, China; 2210474049@email.szu.edu.cn (D.M.); 2210474044@email.szu.edu.cn (Q.L.)
* Correspondence: ddzk@szu.edu.cn

Abstract: Heating, ventilation, and air-conditioning systems (HVAC) have significant potential to support demand response programs within power grids. Model Predictive Control (MPC) is an effective technique for utilizing the flexibility of HVAC systems to achieve this support. In this study, to identify a proper prediction model in the MPC controller, four machine learning models (i.e., SVM, ANN, XGBoost, LightGBM) are compared in terms of prediction accuracy, prediction time, and training time. The impact of model prediction accuracy on the performance of MPC for HVAC demand response is also systematically studied. The research is carried out using a co-simulation test platform integrating TRNSYS and Python. Results show that the XGBoost model achieves the highest prediction accuracy. LightGBM model’s accuracy is marginally lower but requires significantly less time for both prediction and training. In this research, the proposed control strategy decreases the economic cost by 21.61% compared to the baseline case under traditional control, with the weighted indoor temperature rising by only 0.10 K. The result also suggests that it is worth exploring advanced prediction models to increase prediction accuracy, even within the high prediction accuracy range. Furthermore, implementing MPC control for demand response remains beneficial even when the model prediction accuracy is relatively low.

Keywords: demand response; prediction accuracy; machine learning; HVAC system; model predictive control

1. Introduction

Maintaining the stability of power grids is a formidable challenge, particularly during periods of peak power demand. Furthermore, the rapid increase of renewable power generation sources, such as wind and solar energy, introduces intermittency that undermines the reliability of power production [1,2]. Building more power plants can address this issue, but it can also lead to a significant reduction in the annual working hours of generators. In China, for instance, the average annual working hours of power generators in power stations (with a capacity not less than 6000 kW) is only 3758 h [3]. Except for this conventional method, an alternative is to focus on the demand side rather than the supply side. Demand response (DR) is a method aimed at motivating consumers to adjust their power consumption behaviors by modifying their typical usage patterns. This helps to shift or reduce peak power demand. Due to its significant potential, various policies have been enacted to incentivize demand-side resources to participate in demand response.
through monetary rewards [4]. Among the different types of demand resources, heating, ventilation, and air-conditioning (HVAC) systems in buildings are particularly promising for providing demand response. There are three main reasons for this. Firstly, the building sector constitutes roughly 40% of the total energy consumption [5], with HVAC systems accounting for a substantial portion of this, approximately 50% in the United States [6]. Secondly, by changing the indoor temperature setpoint (ITS) or altering their operational status (on or off), HVAC systems can easily modify the power use profile. Thirdly, the energy flexibility of the HVAC systems can be enhanced by using building thermal inertia. For example, by precooling buildings before peak hours, the power consumption of HVAC systems during peak hours can be further reduced [7].

The nature of building thermal inertia implies that current control action (e.g., ITS adjustment) can affect HVAC power use further. In this context, Model Predictive Control (MPC) emerges as a promising control method capable of fully utilizing the power use flexibility of building HVAC systems by optimally modulating the ITS. The strength of MPC lies in the use of a model of the control object to forecast its future behavior. Based on these behaviors, MPC can optimally select control actions (e.g., ITS) in accordance with the control target while considering the constraints associated with the control object [8]. Based on review papers [8–10] on MPC for buildings, it can be concluded that a prerequisite for the successful implementation of MPC is the prediction model. This model should be able to predict the variables concerned in constraints and control targets according to the control actions. In the context of optimal control of building HVAC systems (e.g., demand response), constraints are typically related to thermal comfort, which necessitates indoor temperature predictions. Meanwhile, control targets normally involve energy costs. Therefore, power use prediction is also required.

For certain HVAC systems, such as heaters or constant-frequency air conditioning systems, the control action is normally system working status (on or off), and these systems are typically considered to operate at rated power. In such cases, a dedicated indoor temperature prediction model is required, while no specific model is needed to predict system power use. Bianchini et al. [11] developed a multiple linear regression model to predict indoor temperature. Based on this model, the working status (on or off) of a heater was optimized to minimize the energy cost considering demand response events while maintaining the indoor temperature within the comfort range. In the research conducted by Bacher and Madsen [12], a resistance-capacitance (RC) room thermal model for a single-story 120 m$^2$ building was developed and validated, which can then be used for optimal control. In the study of Adhikari et al. [13], multiple RC room thermal models were developed to predict indoor temperatures in residential buildings. Based on these models, the working status (on or off) of the HVAC units was optimized to collectively provide demand response.

For variable-frequency air conditioning systems, instead of operating at rated power, the power use of air conditioning systems varies significantly under different working conditions. Considering this fact, some studies use steady-state models to determine the power use of HVAC systems. Simultaneously, indoor temperature prediction is also required due to its involvement in thermal comfort constraints. Hu and Xiao [14] employed an RC model to predict indoor temperature and developed a steady-state model, considering the coefficient of performance (COP) variation under different working conditions, to calculate the power use of the air conditioning system. These models were employed in an MPC controller to reduce peak power demand. In the study of De Coninck and Helsen [15], an RC model was used to predict the indoor temperature of a practical building equipped with a heating system, which included a boiler and two heat pumps. Subsequently, static models (i.e., efficiency curves of the boiler and two heat pumps) were employed to determine the energy consumption of the heating systems. The results indicated that, in comparison to traditional rule-based control, Model Predictive Control (MPC) could deliver comparable or improved thermal comfort while lowering energy costs by over 30%. Li et al. [16] developed a linear state-space model to predict indoor temperature and used a steady-state
model to obtain the HVAC system power consumption. By implementing the proposed MPC control strategy, the total electrical energy consumption was reduced by around 17.5% in a simulation test and by more than 20% in an experimental demonstration. These aforementioned studies assume that the system can respond quickly and achieve a new stable working condition when the control action is implemented. This assumption is more applicable to variable-frequency equipment with rapid responses after the control action is adjusted. However, when the control object is a large central air conditioning system, the response process of system power use to the control action (e.g., ITS adjustment) should also be considered because there is a non-neglectable delay. This delay is mainly caused by the local feedback control, which was initially set to provide stable operations [17]. Under such circumstances, a dedicated, dynamic model is needed to predict power use. In the research conducted by Ma et al. [18], an economic MPC strategy was proposed. Dynamic linear models were built to predict indoor temperature and HVAC power use when the ITS is adjusted. The optimization problem was then designed and transformed into a linear programming problem for efficient resolution. In a later study, this control strategy underwent experimental validation in a commercial office building [19]. Hilliard et al. [20,21] employed random forest regression models to predict the power of an air conditioning system and indoor temperature. Over a four-month experimental trial, the findings showed a 29% decrease in HVAC electric energy usage and a 63% decrease in thermal energy usage compared to the same timeframe in previous years.

According to the abovementioned literature, two research gaps can be identified: (1) Existing studies mainly focused on indoor temperature prediction during demand response, lacking algorithm studies for HVAC power use prediction of central air conditioning systems considering the power response delay due to the impact of the local feedback control. Moreover, the existing studies mainly consider the prediction accuracy of models, while other aspects, such as prediction time and training time, are rarely explored. These aspects are also important, particularly the prediction time, as MPC needs timely online optimization to generate optimal control actions; (2) although the implementation of MPC has been demonstrated to improve control performance in current studies, the extent to which MPC can further enhance control performance by improving prediction accuracy remains unclear. Therefore, it is necessary to study the impact of model prediction accuracy on MPC performance. The study of this research gap can also help determine whether the control performance of MPC is acceptable when the model prediction accuracy is relatively low in real-world applications.

This study, therefore, attempts to identify a proper prediction model for HVAC power use considering the power response delay due to the local feedback control and systematically study the impact of model prediction accuracy on MPC performance. To achieve these objectives, four types of machine learning algorithms are developed to predict the power use of HVAC systems and indoor temperature, including Support vector machine (SVM), artificial neural network (ANN), XGBoost (Extreme Gradient Boosting), and LightGBM (Light Gradient Boosting Machine). The performance of these models is compared in terms of accuracy, prediction time, and training time. After selecting the prediction model with relatively high prediction accuracy, the structure of this model is modified to generate models with different levels of prediction accuracy. Then, these models are adopted in the MPC controller to study the impact of model prediction accuracy on the performance of MPC.

The paper is organized as follows. Section 2 outlines the methodology, starting with the principle of MPC for demand response, followed by the prediction model for MPC, cost function, and the selection of periods and horizons. In Section 3, a detailed description of the test platform is provided. In Section 4, the performance comparison among different prediction models is presented and discussed. In Section 5, the result of the impact of model prediction accuracy on the performance of MPC is presented and discussed. Conclusions are drawn in Section 6.
Figure 1 illustrates the outline of the study, which includes six steps. **Step 1**: A commercial building in Shenzhen, China, is chosen as a reference building. **Step 2**: TRNSYS 18 (64-bit) [22] is used to create a virtual HVAC system for the building, as detailed in Section 3. The test platform generates a random sequence of indoor temperature setpoints, producing corresponding output signals, such as indoor temperature and HVAC system power use. **Step 3**: Four types of machine learning algorithms are developed and compared based on the data from Step 2 and additional disturbances, such as weather conditions and time information. **Step 4**: Utilizing the prediction models trained in Step 3, an MPC controller is developed in Python to generate an indoor temperature setpoint sequence (i.e., control actions). **Step 5**: The MPC controller developed in Step 4, and the building HVAC system built in Step 2 are integrated into a co-simulation test platform. **Step 6**: This co-simulation test platform is finally employed to analyze how model prediction accuracy affects the performance of MPC.

**Figure 1.** Outline of the study.

2. Methodology

2.1. Principle of MPC for Demand Response

Figure 2 illustrates the principle of MPC for HVAC systems to provide demand response in this study. In each control step, the Genetic Algorithm (GA) randomly generates a series of indoor temperature setpoint sequences. The corresponding power consumption profiles are predicted based on the prediction model, and their associated costs are then obtained according to the cost function. Based on the minimum cost, the controller can determine the optimal indoor temperature setpoint sequence. It should be noted that the controller applies only the initial indoor temperature setpoint of the generated sequence, disregarding the remaining setpoints. In the subsequent control step, a new GA optimization is performed, producing a fresh indoor temperature setpoint sequence. This feature, known as rolling optimization, leverages updated information to enhance control performance. Moreover, in each control step, the prediction error is employed to correct the prediction result for the next control step, which is also beneficial for control performance.
2.2. Prediction Model for MPC

This section details the foundational principles and the development process of the prediction models. The reasons for selecting the following four specific models as candidates are presented as follows. Support Vector Machine (SVM) is a classical machine learning method whose prominent advantage is high performance with limited data. Artificial Neural Network (ANN) is a famous machine learning method that has been widely studied in the building energy field [23]. XGBoost and LightGBM are ensemble models grounded in boosting theory and are known for their exceptional performance in machine learning competitions. However, their application in MPC control of HVAC systems has been scarcely explored.

2.2.1. Model Principle and Hyperparameters

SVM has a powerful capability for nonlinear predictions. This nonlinear regression ability stems from the kernel function, which can nonlinearly map the input space to high-dimensional feature space and carry out linear regression within this feature space. In this study, the Gaussian radial basis function kernel is used. Another prominent advantage of SVM is its high performance with limited training data. This is because the model is developed based on the support vectors rather than the entire data. There are two important hyperparameters in SVM: C and gamma. Here, C serves as a penalty factor for training errors, with a larger C potentially leading to overfitting. Gamma influences the shape of the decision boundary. A smaller gamma results in a more flexible and smoother decision boundary, whereas a larger gamma produces a more complex and sharper boundary. In this study, the grid search method is employed to identify the optimal values for C and gamma.

An Artificial Neural Network (ANN) is a method that utilizes a collection of artificial neurons. Each artificial neuron processes received inputs and generates an output using non-linear activation functions. These neurons are typically organized into layers. This study utilizes a fully connected neural network. The primary hyperparameters to be optimized include the number of hidden layers and the number of neurons in those hidden layers. It is important to note that the number of neurons is kept consistent across all hidden layers in this study.

Both XGBoost and LightGBM belong to the gradient boosting algorithm in which an ensemble of weak learners (i.e., Decision Tree) is used. In particular, a new learner is generated to fit the residual errors of the preceding trees. The final result is obtained by summing the outputs of all these learners [24].

$$\hat{y}_{i}^{(m)} = \sum_{k=1}^{m} f_k(x_i) = \hat{y}_{i}^{(m-1)} + f_m(x_i)$$  (1)
where \( \hat{y}_i^{(m)} \) is the final tree model, \( \hat{y}_i^{(m-1)} \) is the previously generated tree model, \( f_m(x_i) \) is the newly generated tree model, and \( m \) is the total number of base tree models.

LightGBM [25] was proposed in 2017, later than XGBoost. Therefore, some modifications are implemented. The first modification is the method to determine the node to be split. In XGBoost, a level-wise growth strategy is employed, which means that the nodes closer to the tree root will be split. By comparison, LightGBM carries out a leaf-wise growth strategy, which splits nodes that yield the highest loss change after splitting. The second modification is the method of finding the split point after determining the node to be split. In XGBoost, a Pre-sorted strategy is used which means that all possible split points are evaluated. By comparison, LightGBM carries out a Histogram-based strategy in which features are divided into discrete bins and feature histograms are created. In this way, the number of split points decreases significantly, which can reduce computation costs. It claims that LightGBM is significantly faster than XGBoost while achieving equivalent performance [25].

There are many hyperparameters in XGBoost and LightGBM. The hyperparameter tuning process can refer to our previous study [26].

Four performance metrics are employed to assess the prediction accuracy of these models, including mean absolute error (MAE), root mean square error (RMSE), \( R^2 \), and the coefficient of variation of the root mean squared error (CV-RMSE) [27]. The first two metrics are scale-dependent metrics, which offer a direct understanding of the prediction error for a specific case. The last two metrics are scale-independent metrics, which are more suitable for comparisons across similar cases (e.g., different buildings) [6].

\[
MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i| \tag{2}
\]

\[
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2} \tag{3}
\]

\[
R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{N} (y_i - \bar{y})^2} \tag{4}
\]

\[
CV - RMSE = \frac{\sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}}{\sqrt{1/N \sum_{i=1}^{N} y_i}} \tag{5}
\]

where \( y_i \) is the actual value; \( \hat{y}_i \) represents the predicted value of the model; \( \bar{y} \) is the average of the actual values.

### 2.2.2. Feature Selection

The prediction outputs of these machine learning models are HVAC power use and indoor temperature. The analysis for determining the input variables is as follows. Firstly, it is essential to include weather conditions (outdoor temperature \( T_{out} \), and humidity \( H_{out} \)) and indoor temperature setpoint \( T_{in, set} \) as input variables, as these factors greatly affect the indoor temperature and cooling load, which is highly related to HVAC power consumption, \( P \) [28,29]. Secondly, hour of the day (\( t \)) should be included since it indicates the usage schedules of the space, such as occupancy, lighting, and equipment use [30], which directly influence the cooling load. Thirdly, historical data should be incorporated due to the thermal inertia effect. In this study, an indoor temperature setpoint step change revealed a time constant of approximately 50 min. Given the sampling period in this study is 10 min, historical data (including 6 variables \( T_{out}, H_{out}, T_{in, set}, t, T_{in, P} \)) from the past 5 intervals are considered. In summary, there are 34 inputs in total: 30 (i.e., \( 6 \times 5 \)) historical variables and 4 variables (\( T_{out}, H_{out}, T_{in, set}, t \)) in the next step.

To encompass all possible indoor temperature setpoint control profiles during demand response, the training data for the model must include a diverse range of scenarios where the indoor temperature setpoint varies. In this study, a random sequence of indoor temperature setpoints is generated and utilized to produce corresponding output signals based on the simulation test platform. The bounds for this random sequence are set between 23.5 °C
and 26.5 °C. Data collection occurs over a 7-day period, retaining only data from working hours (7 a.m. to 9 p.m.) while discarding data from other times.

2.3. Cost Function

The cost function (i.e., control objective) is designed as follows.

\[ \text{min} J = C_e + C_{DR} + C_{cf} \]  

(6)

where \( J \) represents the total cost, which includes the electricity cost \( (C_e) \), the financial reward for participating in demand response \( (C_{DR}) \), which is negative) and the comfort cost \( (C_{cf}) \). These three costs are defined by Equations (7)–(9), respectively.

\[ C_e = \sum_{t=1}^{n} Ec(t) \cdot Pr_e(t) \]  

(7)

where \( Ec \) is the HVAC system energy consumption; \( Pr_e \) represents the electricity price in Shenzhen from July to September, as illustrated by the solid blue line in Figure 3.

\[ C_{DR} = \sum_{t=1}^{n} Ecd(t) \cdot Pr_{DR}(t) \]  

(8)

where \( Ecd \) is the energy consumption decrease compared with the case under conventional operation without providing DR. \( Pr_{DR} \) is the reward price for participating in DR, which is 3 RMB/KW·h during the DR periods, as illustrated by the red dashed line in Figure 3. The DR periods are set from 11:00 to 12:00 and from 15:00 to 17:00 in this study, which coincide with the peak electricity price levels.

\[ C_{cf} = \alpha \cdot T_{\text{rise}} \]  

(9)

\[ T_{\text{rise}} = \frac{\sum_{t=1}^{n} (T_{in}(t) - T_{in,orig}) \cdot S_p(t)}{\sum_{t=1}^{n} S_p(t)} \]  

(10)

where \( \alpha \) represents the penalty parameter for indoor temperature increase, set to 1800 in this study. \( T_{\text{rise}} \) represents the weighted temperature increase, taking into account \( S_p \), the schedule of the person. \( T_{in} \) represents the indoor temperature, while \( T_{in,orig} \) denotes the indoor temperature under conventional operation, which is 25 °C in this research.

![Figure 3. Electricity price, DR reward price, and working hours.](image)

2.4. Periods/Horizons Selection

The sampling period in this study is set to 10 min, as mentioned in Section 2.2. It means the prediction model forecasts the HVAC system’s power use and indoor temperature for the next 10 min, iteratively predicting each subsequent interval. To mitigate fluctuations in control actions, the control action period is distinct from the sampling period and is set to 1 h.
3. Test Platform

The building model and its HVAC system model are constructed with reference to a real commercial building located in Shenzhen, China. The reference commercial building spans approximately 27,000 m² and is equipped with a typical Variable Air Volume air-conditioning system. Type 56 in TRNSYS is utilized to build the building model. The parameters for the building envelope are established based on survey results and local design standards [31]. Internal mass plays a crucial role in the thermal dynamics of buildings [32,33]. Following a review [34], a wood/plastic material with a density of 50 kg/m² is incorporated into the building model. Additionally, an internal wall of 0.44 m² per square meter of the floor is included, as indicated by the survey. The lighting (10 W/m²) density, equipment power density (13 W/m²), and the design density per person (8 m² per person) are set according to local standards [31]. The schedules for person, lighting, and equipment loads are depicted in Figure 4 as fractions of their peak values.

Figure 4. The schedules for occupants, lighting, and equipment loads.

The HVAC system features two identical chillers, one active and the other as a reserve. Each chiller has a rated power consumption of 437 kW and a cooling capacity of 2461 kW. Detailed modeling of chiller power use can be found in our previous study [26]. For pumps, even though they are variable frequency pumps, they operate at a constant frequency (50 Hz) in real applications as per the survey, so the pump power use variations are disregarded. Type 124 in TRNSYS is used to simulate air handling units. The power use of the variable speed fan $P_{\text{fan}}$ is determined by Equation (11) [35], where $P_{\text{fan,\text{rated}}}$ represents the rated fan power.

$$P_{\text{fan}} = P_{\text{fan,\text{rated}}} \left( \frac{m_{\text{air}}}{m_{\text{air,\text{rated}}}} \right)^2$$  

A TRNSYS–Python co-simulation platform is constructed based on Type 163, which allows TRNSYS to communicate with Python scripts via “dat” files [36].

4. Performance Comparison among Different Models

The performance of SVM, ANN, XGBoost, and LightGBM are presented in Table 1. It should be noted that the hyperparameters set in default values are not mentioned in the table.
### Performance of Machine Learning Models for Predicting HVAC Power Use (Power) and Indoor Temperature (T)

<table>
<thead>
<tr>
<th>Model</th>
<th>SVM</th>
<th>ANN</th>
<th>XGBoost</th>
<th>LightGBM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>Power (W)</td>
<td>T (°C)</td>
<td>Power (W)</td>
<td>T (°C)</td>
</tr>
<tr>
<td>MAE</td>
<td>25,653</td>
<td>0.124</td>
<td>18,081</td>
<td>0.068</td>
</tr>
<tr>
<td>RMSE</td>
<td>35,189</td>
<td>0.145</td>
<td>24,704</td>
<td>0.084</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.930</td>
<td>0.932</td>
<td>0.977</td>
<td>0.978</td>
</tr>
<tr>
<td>CV-RMSE</td>
<td>8.63%</td>
<td>0.57%</td>
<td>6.06%</td>
<td>0.33%</td>
</tr>
<tr>
<td>Training time (s)</td>
<td>0.0030</td>
<td>2.0414</td>
<td>0.3918</td>
<td>0.1078</td>
</tr>
<tr>
<td>Prediction time (ms)</td>
<td>0.1509</td>
<td>0.5159</td>
<td>0.8042</td>
<td>0.4057</td>
</tr>
<tr>
<td>Optimal hyperparameter for predicting HVAC power use</td>
<td>kernel function = ‘rbf’ C = 8192 gamma = $1 \times 10^{-5}$</td>
<td>activation = ‘relu’ number of hidden layers = 2 number of neurons in hidden layers = 250 alpha = 0.1</td>
<td>max depth of a tree = 10 subsample = 0.8 lambda = 0.4 learning rate = 0.04 tree number = 200</td>
<td>Min data in leaf = 20 Subsample = 0.5 learning rate = 0.06 tree number = 250</td>
</tr>
<tr>
<td>Optimal hyperparameter for predicting indoor temperature</td>
<td>kernel function = ‘rbf’ C = 8192 gamma = $1 \times 10^{-5}$</td>
<td>activation = ‘relu’ number of hidden layers = 2 number of neurons in hidden layers = 230 alpha = 0.1</td>
<td>max depth of a tree = 10 subsample = 0.8 lambda = 0.5 learning rate = 0.08 tree number = 200</td>
<td>Min data in leaf = 15 learning rate = 0.08 tree number = 220</td>
</tr>
</tbody>
</table>

#### 4.1. Accuracy

Table 1 shows that XGBoost achieves the highest accuracy, which is slightly higher than LightGBM. This difference may arise from the method used to determine the node to be split. As mentioned in Section 2.2.1, XGBoost adopts a level-wise growth strategy, while LightGBM uses a leaf-wise growth strategy. According to research in data science [25], level-wise growth is usually better for smaller datasets whereas leaf-wise tends to overfit. Therefore, it can be inferred that the data volume used in this case belongs to a small dataset. It is worth noting that in real-world demand response applications, the data obtained for model development is often limited. The reason is elaborated as follows. As mentioned in Section 2.2.2, to cover all possible indoor temperature setpoint control profiles in the process of demand response, the training data of the model should include a variety of cases in which the indoor temperature setpoint is changed. For example, the random indoor temperature setpoint sequence is used in this study. However, before providing demand response, HVAC systems normally operate under a constant indoor temperature setpoint. Therefore, dedicated tests should be conducted for data collection. This period should be as short as possible to avoid bothering occupants, which finally limits the volume of training data. That is also the reason why only seven days are used to collect data in this study. On the other hand, when the HVAC system initiates a demand response, the available data begins to accumulate. When a larger volume of data is used for retraining the prediction model, LightGBM, which uses the leaf-wise growth strategy, may have more advantages.

During the hyperparameter tuning process of XGBoost, it was observed that reducing the learning rate while increasing the number of trees proved to be effective, which can significantly increase the prediction accuracy in this case. The performance of the ANN model is equivalent to that of the LightGBM in this case. During the hyperparameter tuning process of the ANN model, it was found that only two hidden layers were sufficient. However, the number of neurons in hidden layers should be large enough to achieve high performance. In this case, more than 200 neurons are used in each hidden layer. The prediction accuracy of both ANN and LightGBM is much better than that of the SVM model.

#### 4.2. Training and Prediction Time

The model development takes place in Python 3.6, installed on a desktop computer featuring an Intel Core i7-10510U processor (1.80 GHz), 16 GB of RAM, and running the Windows 10, 64-bit operating system. In terms of time consumption, it can be observed from Table 1 that the training time for SVM is significantly shorter than the other three models.
For ANN models, it is found that the training time is highly dependent on the number of neurons in hidden layers. To achieve high performance, more than 200 neurons are used. Therefore, it takes about 2 s to train the ANN model, which is much longer than the other three models. According to the comparison between XGBoost and LightGBM, it can be found that LightGBM is worthy of its name, “light”, as its training time is only one-quarter that of XGBoost.

Compared with training time, prediction time is more critical in practical applications. This is because, during the rolling optimization process, the model is used thousands of times in each control step, which can significantly affect the computation cost. In this study, ANN, XGBoost, and LightGBM achieve almost the same level of prediction accuracy, while the prediction time of LightGBM (i.e., 0.4057 ms) is the shortest among them.

5. Impact of Model Accuracy on MPC Control Performance

To evaluate the impact of model prediction accuracy on MPC control performance, test procedures are designed as follows: (1) Determination of the model type used in MPC and the metric type to measure the prediction accuracy; (2) the structure of the model is modified to artificially generate prediction models with different levels of prediction accuracy; (3) by using these models in MPC, the corresponding cost function values are obtained and used for analysis.

5.1. Determination of the Model Type and the Metric Type

In this study, although XGBoost has the highest prediction accuracy, LightGBM is ultimately used due to two reasons. First, there is only a slight difference between their prediction accuracy. Second, the Genetic Algorithm (GA) algorithm is used in the rolling optimization process in this study, and its outcome involves randomness. Therefore, the case under a certain prediction accuracy level is run 20 times to generate an average performance of the MPC controller. LightGBM only takes half the time of XGBoost which can significantly save time. Therefore, LightGBM is used in the following study.

For the metric type, $R^2$ is selected as the sole metric to measure the prediction accuracy. The analysis is presented as follows. There are four metrics used in this study. Compared with scale-dependent metrics, scale-independent metrics are more suitable for comparison with other similar studies, such as cases in different buildings or under different demand response policies. There are two scale-independent metrics used in this study, i.e., $R^2$ and CV-RMSE. CV-RMSE is excluded because it is significantly influenced by the type of feature. For example, the CV-RMSEs of LightGBM for prediction power use and indoor temperature are 5.94% and 0.37%, respectively, which is quite different. In addition, if the temperature unit is changed from relative temperature in °C to absolute temperature in K, the same dataset can produce noticeably different CV-RMSE values. Compared to CV-RMSE, $R^2$ does not have this issue. In this study, the $R^2$ values for the model for predicting both power use and indoor temperature are quite similar. Therefore, $R^2$ is selected as the sole metric to measure the prediction accuracy.

5.2. Artificial Modification of the Model Prediction Accuracy

To adjust the prediction accuracy of the model, the number of trees and number of leaves in a tree is decreased to generate models with different levels of prediction accuracy, corresponding to four cases, as shown in Table 2. The model used in the first case is actually the original LightGBM model. The $R^2$ of the models, in other cases, forms an arithmetic progression. In addition to these four cases, a baseline case without implementing a demand response is also conducted (i.e., $T_{in, set} = 25$ °C).
Table 2. Four cases using LightGBM models with different levels of prediction accuracy.

<table>
<thead>
<tr>
<th>Case</th>
<th>R² (HVAC Power)</th>
<th>R² (Indoor T)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.967 (≈0.97)</td>
<td>0.972 (≈0.97)</td>
</tr>
<tr>
<td>2</td>
<td>0.90</td>
<td>0.90</td>
</tr>
<tr>
<td>3</td>
<td>0.83</td>
<td>0.83</td>
</tr>
<tr>
<td>4</td>
<td>0.76</td>
<td>0.76</td>
</tr>
</tbody>
</table>

5.3. Results Comparison and Analysis

The results of the cases with the highest prediction accuracy (i.e., Case 1) and lowest prediction accuracy (i.e., Case 4) are presented in detail in Sections 5.3.1 and 5.3.2. A comparison of the four cases with different levels of prediction accuracy is conducted and presented in Section 5.3.3.

5.3.1. Case 1 with the Highest Prediction Accuracy

Figure 5 presents the indoor temperature and its setpoint in Case 1. Figure 6 depicts the HVAC power use in Case 1 and the baseline case. The demand response (DR) period is indicated by the green area.

![Figure 5. Indoor temperature setpoint and indoor temperature in Case 1.](image)

![Figure 6. HVAC power use in Case 1 and the baseline case.](image)

According to Figures 5 and 6, it can be observed that the developed control strategy can effectively decrease HVAC power use during the DR period by increasing the indoor temperature setpoint. Moreover, the strategy can also automatically implement precooling. For example, between the two DR periods, from 12:00 to 15:00, the indoor temperature setpoint is significantly reduced. It is worth noting that due to the precooling, the peak power use of the HVAC system reached 581 kW, which is even larger than that in the baseline case (i.e., 428 kW).

Figure 7 displays the electricity cost ($C_e$) and financial reward ($C_{DR}$, which is negative) for Case 1, alongside the electricity cost ($C_e$) in the baseline case. The data reveal that,
accordingly, the percentage of dissatisfied occupants. Clothing insulation, airspeed, and metabolic rate are set to 0.5 clo, 0.1 m/s, and 1.6 \cite{38}, respectively. The results are shown in Figure 8.

According to Figures 5 and 6, it can be observed that the developed control strategy significantly reduced HVAC power use during the DR period by increasing the indoor temperature setpoint. Moreover, the strategy can automatically implement precooling. For example, from 10:00 to 11:00, the indoor temperature setpoint increased by 2.4 °C compared to the baseline case.

The weighted indoor temperature is 25.1 °C (i.e., indoor temperature considering the schedule of the person, Equation (10) in Section 2.3), showing an increase of only 0.10 K compared to the baseline case. The humidity is also properly controlled during the demand response period. PMV ranges from 0.08 to 0.71, with the highest PMV occurring at 17:00 and a PDD of 15.6% at that time.

Figure 7 displays the electricity cost ($C_e$), financial reward ($C_{DR}$ negative) in Case 1, and electricity cost ($C_e$) in the baseline case.

Figure 8. Indoor temperature, humidity, PMV, and PDD in Case 1.

Figure 9 shows the indoor temperature setpoint in Case 1 and Case 4, while Figure 10 shows the HVAC power use for both cases. It can be found that the control strategy is more effective in Case 1, and electricity cost ($C_e$) in Case 1 is significantly lower than in Case 4. Because of the high electricity price and demand response rewards during these periods, insulation precooling can significantly increase the cost. However, the reduction in costs during DR periods offsets this increase due to higher electricity prices.

Although only indoor temperature is considered in the cost function, a comprehensive study examines thermal comfort during HVAC demand response, including indoor temperature, humidity, Predicted Mean Vote (PMV), and Predicted Percentage of Dissatisfied (PPD). PMV, developed by Fanger and standardized in ASHRAE55 \cite{37}, ranges from −3 (cold) to +3 (hot). Based on PMV, PDD can be calculated, which reflects the percentage of dissatisfied occupants. Clothing insulation, airspeed, and metabolic rate are set to 0.5 clo, 0.1 m/s, and 1.6 \cite{38}, respectively. The results are shown in Figure 8.
The weighted indoor temperature is 25.1 °C (i.e., indoor temperature considering the schedule of the person, Equation (10) in Section 2.3), showing an increase of only 0.10 K compared to the baseline case. The humidity is also properly controlled during the demand response period. PMV ranges from 0.08 to 0.71, with the highest PMV occurring at 17:00 and a PDD of 15.6% at that time.

5.3.2. Case 4 with the Lowest Prediction Accuracy

Figure 9 shows the indoor temperature setpoint in Case 1 and Case 4, while Figure 10 shows the HVAC power use for both cases. It can be found that the control strategy is also capable of raising the indoor temperature setpoint during the DR period. Nevertheless, it fails to effectively implement precooling. For example, from 10:00 to 11:00, the indoor temperature setpoint is 26.5 °C. From 12:00 to 15:00, the indoor temperature setpoint only decreases slightly. Insufficient precooling can result in higher power use during the DR period. As illustrated in Figure 10, although the indoor temperature setpoints are the same (i.e., 26.5 °C) in two cases during two DR periods, the power use in Case 4 is greater than in Case 1. Because of the high electricity price and demand response rewards during these periods, insufficient precooling can significantly increase the cost.

Figure 9. Indoor temperature setpoints in Case 1 and Case 4.

Figure 10. HVAC power use in Case 1 and Case 4.

5.3.3. Comparison of Four Cases

As mentioned in Section 5.1, to eliminate the uncertainty caused by the GA algorithm, each case is run 20 times under a certain prediction accuracy level to generate an average performance of the MPC controller. The indoor temperature setpoints and indoor temperatures in each case are presented in Figures 11 and 12, respectively.
5.3.3. Comparison of Four Cases

As mentioned in Section 5.1, to eliminate the uncertainty caused by the GA algorithm, each case is run 20 times under a certain prediction accuracy level to generate an average performance of the MPC controller. The indoor temperature setpoints and indoor temperatures in each case are presented in Figure 11 and Figure 12, respectively.

The cost function values in different cases and the baseline case are shown in Figure 13. It can be observed that, under a certain prediction accuracy level, although some variation exists in cost function values among different cases, the values fall within a relatively small range. It indicates that the GA algorithm can roughly find the optimal indoor temperature setpoint sequence. In Figure 13, the blue line connects the median cost function values under different prediction accuracy levels. It can be noted that with the decrease in prediction accuracy, the performance of MPC is decreasing. This decreasing trend is more significant under two circumstances. First, compared with the baseline case, the cost function value in Case 4 (i.e., $R^2 = 0.76$) decreases significantly. This means that it is worth implementing demand response, even if the prediction accuracy of the model is relatively poor. Second, the cost function value in Case 1 (i.e., $R^2 = 0.97$) is much lower than that in...
Case 2 (i.e., $R^2 = 0.90$). This result implies that it is valuable to explore advanced prediction models to increase prediction accuracy, even within the high prediction accuracy range.

Figure 12. Indoor temperatures in different cases.

The cost function values in different cases and the baseline case are shown in Figure 13. It can be observed that, under a certain prediction accuracy level, although some variation exists in cost function values among different cases, the values fall within a relatively small range. It indicates that the GA algorithm can roughly find the optimal indoor temperature setpoint sequence. In Figure 13, the blue line connects the median cost function values under different prediction accuracy levels. It can be noted that with the decrease in prediction accuracy, the performance of MPC is decreasing. This decreasing trend is more significant under two circumstances. First, compared with the baseline case, the cost function value in Case 4 (i.e., $R^2 = 0.76$) decreases significantly. This means that it is worth implementing demand response, even if the prediction accuracy of the model is relatively poor. Second, the cost function value in Case 1 (i.e., $R^2 = 0.97$) is much lower than that in Case 2 (i.e., $R^2 = 0.90$). This result implies that it is valuable to explore advanced prediction models to increase prediction accuracy, even within the high prediction accuracy range.

Figure 13. Cost function values in cases with different levels of prediction accuracy and the baseline case.

6. Conclusions

This study compares four machine learning models (SVM, ANN, XGBoost, LightGBM) for prediction accuracy, prediction time, and training time to identify a suitable MPC controller model for HVAC demand response. The impact of prediction accuracy on MPC performance is also examined. The main conclusions are as follows:

- This study demonstrates that the proposed model predictive control strategy efficiently manages indoor air temperature setpoint for demand response. While the traditional control in the baseline case results in higher costs, the proposed control strategy reduces economic costs by up to 21.61%, with only a minimal increase of 0.10 K in the weighted indoor temperature. In addition, the humidity is also well-managed. The PMV ranges from 0.08 to 0.71, with the highest PMV occurring at 17:00, and the PDD at that time is 15.6%.

- From the perspective of prediction accuracy, the XGBoost model achieves the highest prediction accuracy compared with SVM, ANN, and LightGBM. For the XGBoost models developed in this study, the obtained R-squared values are 0.978 and 0.983 for predicting power use and indoor temperature in the upcoming 10 min.

- From the perspective of hyperparameter tuning, using a relatively low learning rate and a large number of trees is effective in enhancing the performance of the XGBoost model. For the ANN model, two hidden layers are sufficient for predicting HVAC power use and indoor temperature. On the other hand, the number of neurons in hidden layers should be large enough to obtain high performance.

- From the perspective of prediction and training time, LightGBM has the shortest prediction time among ANN, XGBoost, and LightGBM, although its prediction accuracy is slightly lower than that of XGBoost. In addition, the training time of LightGBM is one-quarter the time of XGBoost in this study.

- It is valuable to explore advanced prediction models to increase prediction accuracy, even within the high prediction accuracy range. Furthermore, it is also worth implementing MPC control for demand response even if the model prediction accuracy is relatively low (e.g., $R^2 = 0.76$ in this study).

This study also has certain limitations that need to be acknowledged. Although the simulation study considers the power response delay caused by local feedback control, there are various local feedback control sets (i.e., PID values) in practice, which could also
impact the accuracy of the prediction models. Therefore, further experimental studies are necessary to validate the performance of the prediction models in real-world scenarios.

**Author Contributions:** Methodology, H.W.; Software, D.M.; Validation, Q.L.; Resources, Z.D.; Supervision, Z.D.; Funding acquisition, Z.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was conducted with the support of the Guangdong Basic and Applied Basic Research Foundation (Grant No. 2021A1515110448), Shenzhen Science and Technology Program (Grant No. 20220810173620003, Grant No. 2022081016021001, and Grant No. KCXST202102111203007).

**Data Availability Statement:** Data are contained within the article.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


19. Ma, J.; Qin, S.J.; Salsbury, T. Application of economic MPC to the energy and demand minimization of a commercial building. *J. Process Control* 2014, 24, 1282–1291. [CrossRef]


22. Wang, H.; Wang, S. A hierarchical optimal control strategy for continuous demand response of building HVAC systems to provide frequency regulation service to smart power grids. *Energy* 2021, 230, 120741. [CrossRef]


Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.