Journal of Energy Storage (EST), 87(111402):1-16, May 2024. Elsevier

Fairness-Aware Data Driven-Based Model Predictive Controller: A Study on Thermal Energy Storage in A Residential Building

Ying Sun[1,2], Fariborz Haghighat*[2], Benjamin C. M. Fung[3]

[1]School of Environmental and Municipal Engineering, Qingdao University of Technology, Qingdao, China
[2]Energy and Environment Group, Department of Building, Civil and Environmental Engineering Concordia University, Montreal, Canada
[3]School of Information Studies, McGill University, Montreal, Canada

Abstract

Besides accuracy, fairness has been reported as another performance criterion for data-driven building models (DDBMs). To ensure data-driven-based model predictive controllers (MPCs) provide optimal control strategies based on fair or unbiased prediction, this study proposes the concept of fairness-aware data-driven-based MPC. In the proposed MPC framework, fairness-aware DDBMs constitute the prediction component that provides predicted building states to the objective function in the optimization part. To investigate the effect of improving fairness on the control performance, the proposed MPC is implemented in a residential building heated by an electrically heated floor (EHF) system, which could be considered as a thermal energy storage (TES) system. In the case study, fairness-aware DDBMs are developed to predict energy demand and indoor air temperature. Then, the predicted values are used to formulate the objective function to optimize the hourly day-ahead set-point temperature with the aim of minimizing the heating cost or maximizing peak shifting, while maintaining thermal comfort. The numerical study results show that although considering fairness improvement methods in DDBMs decreases the overall predictive accuracy, it provides fair prediction by narrowing the accuracy difference between majority conditions and minority conditions. For instance, a fairness-aware energy prediction model increases the overall mean absolute error (MAE) from 6.12 kWh to 7.56 kWh but decreases the MAE difference between a majority condition and a minority condition from 0.82 kWh to 0.14 kWh. Although improving predictive fairness comes with a price of overall predictive accuracy, fairness-aware data-driven-based MPCs show comparable peak shifting or cost-saving ability with the traditional MPC in which fairness is not considered. This study provides a reference for stakeholders to design and implement trustworthy MPCs in buildings based on fair and accurate prediction.

Keywords: Fairness; Data-driven model; Model predictive controller; thermal energy storage; peak shifting

Corresponding Author*: Fariborz.Haghighat@concordia.ca

1. Introduction

1.1. *Background*

1.1.1. Data-driven building models (DDBMs) and their application in model predictive controllers (MPCs)

In recent years, modern buildings have been equipped with smart sensors (e.g., temperature probes, motion detectors, and power meters), a building automation system (BAS), and/or a building management system (BMS) [1]. It enables the dynamic collection of high-quality building-related information. The abundant collected data inspires the development of data-driven building models (DDBMs) to explore the statistical pattern from these data and to predict indoor environmental parameters, e.g., indoor air temperature [2] and indoor air quality [3]), thermal load/energy consumption [4,5], device/system (e.g. HVAC system and chillers) operation status [6,7].

Due to the ability to represent building states, DDBMs could be integrated into model predictive controllers (MPCs) to optimize the control strategy for devices/systems in a building in a finite horizon with the aim of minimizing energy cost/CO2 emission/thermal discomfort, maximizing peak shifting, or achieving demand response (DR), and so on. For instance, Huchuk et al. [8] proposed a novel linear model-based MPC to optimize the control signals for smart thermostats in residential buildings. It is aimed at simultaneously minimizing the use of heating/cooling systems and thermal discomfort. The result found that the data-driven-based MPC could effectively reduce the runtime of the heating/cooling system and improve thermal comfort, compared to the dead-band control and the model-free reinforcement learning control method. Behl et al. [9] proposed a model-based control with a regression tree algorithm (mbCRT) to optimize the chiller set-point, zone cooling set-point, and lighting level, with the aim of minimizing the energy cost and thermal discomfort. Regressive trees were developed to predict power consumption and indoor air temperature. A case study on 8 campus buildings showed that mbCRT outperforms the best rule-based controller by ~17% in terms of curtailment during a DR event.

Khosravi et al.[10] proposed a data-driven-based MPC to achieve visual and thermal comfort for occupants in a building. In the proposed MPC, a semi-linear support vector machine model is developed to predict visual comfort, while an autoregressive exogenous model predicts the thermal dynamics of the building. More studies on data-driven-based MPC in buildings are summarized in Table 1.

Table 1: Summary of existing studies on data-driven-based MPC in buildings

Ref.	**Data-driven predictors**	**Predicted feature(s)**	**Predictive criteria**	**Controlled variables**	**Control objective(s)**	**Control performance criteria**
[11]	• Artificial neural networks • Gaussian process regression • Simple linear regression process models	• Indoor temperature	• The deviation between predicted values and measured values	• Valve opening of a district heating network in five real-life office rooms	• Minimize energy consumption while maintaining thermal comfort	• Energy consumption • Thermal comfort • Computation time
[12]	• seq2seq-LSTM	• Ambient temperature inside the building • Solar radiation • Heating load	• R^2 • The root mean squared error (RMSE) • Mean absolute error (MAE)	• the hourly start-stop control signal for the collector-side pump • The hourly start-stop control signal for the auxiliary thermal source • The hourly control signal for the heating end variable frequency pump	Minimize the function containing the following factors: • Indoor thermal fluctuation (measured as the difference between indoor and design temperatures) • The total energy consumption of the system • The total heat collected by the heat collection subsystem	• Thermal comfort • Solar heat collection • Heat loss from the heat storage tank • System energy consumption
[13]	• Multiple linear regression • Artificial neural network • Support vector regression (SVR) • Random forest (RF)	• Energy consumption	• Mean absolute percentage error (MAPE)	• Supply water temperatures of chillers and cooling towers • Indoor temperature and humidity	• Minimize the mean value of the air conditioning system's total energy consumption during a period	• Optimization accuracy • Optimization time
[14]	• ANN	• Hourly indoor air-	• RMSE • MAE	• Set-point temperature for	• Minimize heating energy cost • Minimize thermal discomfort	• Thermal comfort penalty

Ref.	Data-driven predictors	Predicted feature(s)	Predictive criteria	Controlled variables	Control objective(s)	Control performance criteria
		dry bulb temperature • Hourly heating load		heating		• Heating cost
[15]	• Deep neural network (DNN)	• Supply air temperature of the HVAC system • Indoor temperature	• Normalized mean bias error (NMBE) • Coefficient of variation of the root mean square error (CV(RMSE))	• Set-point temperature for cooling	• Minimize energy consumption of the condensing units while maintaining the cooling set-point temperature	• Energy saving rate
[16]	• Gaussian processes	• Room temperature	• The deviation between predicted values and measured values	• Valve opening of the chiller system	• Minimize energy consumption	• Energy consumption • Thermal comfort violation
[17]	• Autoregressive with exogenous inputs (ARX)	• Indoor temperature	• The goodness of fit • RMSE	• Water temperature setpoint • Boiler ON/OFF	• Minimize the thermal discomfort of the occupants in the zone • Minimize the boiler energy usage	• The sum of squared comfort violation values during occupied hours • Energy usage
[18]	• Linear regression	• Indoor CO2 concentration • Air temperature	• NMBE • The deviation between predicted values and measured values	• Window opening during winter	• Minimize the operable window opening area and the change of its position • Constraint conditions include indoor air quality (CO_2 concentration threshold) and thermal comfort (indoor air temperature threshold and threshold for indoor air temperature change rate) requirements	• Indoor air temperature • Energy consumption

Ref.	Data-driven predictors	Predicted feature(s)	Predictive criteria	Controlled variables	Control objective(s)	Control performance criteria
[19]	• Decision Tree (DT) • RF	• The appropriate position of roller shades	• Accuracy	• Shading position	• Minimize energy consumption	• Energy consumption • Useful daylight illuminance
[20]	• Encoder-decoder recurrent neural network	• Room air temperature	• RMSE • MAE • CV(RMSE)	• Indoor set-point temperature	• Minimize the deviation between room air temperature and set-point temperature with constraints on supply air temperature • Optimizing building energy efficiency and indoor thermal comfort simultaneously	• The time-varying indoor temperature setpoints • Energy consumption • Discomfort index
[21]	• Attention-based neural network time series multivariate prediction model	• Zone temperature	• MAPE • RMSE • MAE	• Cooling stage • Heating stage • Fan ventilation stage	• Minimizes energy consumption, peak demand, and discomfort during occupied hours • Constraint on self-tuned setpoint, temperature ramp, and equipment cycling	• Energy consumption • Power peak • Thermal discomfort

From Table 1, accuracy measures, such as NMBE, MAE, RMSE, and CV(RMSE), are widely used as the predictive performance criteria for DDBMs in MPCs to reflect the closeness of predicted values to ground-truth values. This is because a model with higher accuracy could represent a more valid prediction of building states [22]. Chen et al. [13] found that DDBMs with higher accuracy resulted in better optimization performance of the MPC. Thus, most of the existing studies on DDBMs were aimed at improving the predictive accuracy by improving the training data quality [23] or selecting proper structure complexity for data-driven models [24].

Besides, fairness, which has been reported as a fundamental principle for guiding the development and application of AI technologies [25,26], is another evaluation measure that should be considered for DDBMs [27]. Two types of fairness measures for DDBMs have been clearly summarized in Ref [27]: "*Type I*. The predictive result is independent of the protected attributes (i.e., features that should not be disclosed or features whose conditions should be unbiased). *Type II*. The predictive performance (e.g., accuracy) is comparable between classes/conditions among the protected attribute(s)." Achieving *Type I* fairness enables privacy protection as it could avoid the usage of private information [28] or ensure the unpredictability of protected features from the remaining features [29]. Besides, improving *Type II* fairness could ensure uniform predictive performance among conditions defined by protected attributes. For instance, if the protected feature separates the users into different user groups, improving *Type II* fairness could let the DDBMs provide similar predictive service to users coming from distinct groups; if the protected feature is peak/off-peak period, better *Type II* farness means the DDBMs could provide uniform predictive performance no matter what period it is. Furthermore, letting DDBMs in an MPC be fair in terms of *Type II* may enable the MPC to provide an unbiased control strategy for different users or periods, so that the MPC would be trustworthy for users [30].

Existing studies have worked on proposing data pre-processing methods or model structure adaption methods to achieve *Type II* fairness for DDBMs by improving the predictive accuracy similarity between different conditions. For instance, Sun et al. [27] proposed three pre-processing methods (i.e., sequential sampling (SS), sequential preferential sampling (SPS), and reversed preferential sampling (RPS)) to produce a balanced training dataset for classification problems, and applied these methods to develop DDBMs for lighting status prediction. Compared with existing pre-processing methods, such as random sampling (RS), the proposed methods could decrease the accuracy difference between the majority condition and the minority condition. Then, they further evaluated the generalizability of these methods on buildings with different occupancy behavior patterns and lighting demand [31]. Furthermore, Sun et al. [32] proposed four in-processing methods to improve the predictive fairness for regression models by adding fairness-related constraints or penalties to the objective function for model training. However, these existing studies verify that improving predictive fairness comes with the price of decreasing overall predictive accuracy. The decreased predictive accuracy may further affect the control performance of an MPC when integrating the fairness-aware DDBMs into it.

Besides, in recent years, “fairness” has been considered by electrical utility companies to provide fair electricity services to customers at a district level. A fair grid tariff model would be more acceptable to a majority of consumers and be more effective for green transition [33]. Wang et al.[34] proposed a fairness-based real-time electricity pricing framework to motivate users to adjust energy consumption patterns based on the price signal. Ren et al. [35] proposed a fairness-based profit allocation model to maximize the individual benefit of each prosumer in a distributed energy network. Danner and Meer [36] proposed a centralized charging capacity allocation policy to provide fair charging service to electric vehicles in a low-voltage grid. Jacubowicz et al. [37] proposed a fairness-based emergency demand response program for a residential area to improve electrical network stability by sending fair curtailment orders to consumers.

However, as shown in Table 1, a study on fairness-aware dynamic control strategies for individual buildings is lacking, although fairness has already been reported as of great importance in data-driven systems [38,39]. To fill the research gap, this study will firstly integrate the fairness-aware DDBMs into MPCs to form the concept of fairness-aware data-driven-based MPCs. Then, the effect of improving the fairness of the integrated DDBMs with a price of the overall accuracy on the control performance of MPCs will be investigated.

1.1.2. Thermal energy storage in buildings and its control strategies

Except for being data-rich, the buildings and building construction sectors have been reported as one of the major final energy consumers. For instance, in 2021, buildings accounted for 30% of total global final energy consumption [40], while they consumed 43% and 28% of final energy consumption in Europe [41] and the U.S. [42], respectively. Besides, building energy demand is not static, but changes over time. Due to occupants’ living habits, daily peak load may occur during early morning or late afternoon [43]. This phenomenon further results in increased stress on the electricity grid during peak periods. To ease this burden, setting higher electricity prices for peak periods than off-peak periods, such as using time-of-use (ToU) tariffs, could be used to encourage consumers to shift energy consumption patterns. Another effective solution to pave the grid stress resulting from peak load is the usage of an energy storage system by taking advantage of its ability to store energy during off-peak periods and release it for later use.

In recent years, thermal energy storage (TES) systems have been widely used in buildings for peak shifting or renewable energy storage. High thermal mass materials in building envelopes may provide sufficient thermal energy storage capacity for peak shifting. For instance, Zhu et al. [44] found that exterior walls with high thermal mass could shift peak load entirely from daytime to night. Olsthoorn et al. [45] pointed out that appropriate system configurations of an electrically heated floor (EHF) system in a residential building could shift both morning peaks and evening peaks.

Effective control strategies are crucial to take advantage of the full capacity of TES systems in buildings to enlarge the peak shifting and energy cost-saving potential [46]. In recent years, data-driven-based MPCs have drawn increasing attention to control TES systems. For instance, Lee et al. [47] developed an artificial neural network (ANN) to predict the thermal behavior of a TES tank, and then, integrated the ANN model into an MPC to minimize the operation cost of the TES system by controlling its charge and discharge rate. The operational cost of TES when controlled by the data-driven-based MPC is 9.1%-14.6% less than the cost when controlled by a rule-based controller. Tang et al. [48] integrated a data-driven cooling load prediction model into the framework of an MPC to minimize the operating cost under a time-of-use tariff for a district cooling system equipped with an ice-based TES. The proposed data-driven-based MPC reduced the cooling cost over a two-month period by ~8%. Note that although the predictive accuracy of models in these MPCs has been validated, the predictive fairness has not been investigated.

1.2. *Objective, contribution, and outline*

The aim of this study is to propose a fairness-aware data-driven-based MPC for buildings with TES systems to minimize the electricity cost by shifting the electrical load from peak periods to off-peak periods. The effect of integrating fairness-aware DDBMs on the predictive performance of MPCs will be investigated. In the proposed MPC, the DDBM provides accurate and fair predictions for building states. Predictive fairness is improved by implementing pre-processing methods, e.g., random sampling (RS) and reversed preferential sampling (RPS), to produce a balanced training dataset. The achievement of predictive fairness is evaluated by the predictive accuracy similarity between conditions defined by the protected feature.

Therefore, the main contribution of this study is to make the data-driven MPC trustworthy by providing an optimal control strategy based on fair, uniform, and accurate prediction. Besides, this study provides guidance on the applicability of the fairness-aware data-driven-based MPC in buildings with TES systems to achieve cost saving and peak shifting. Furthermore, this study quantitively investigates the effect of declining accuracy caused by improving fairness on the control performance of MPC.

The outline of this paper is: Section 2 explains the concept of the fairness-aware data-driven-based MPC, introduces two pre-processing methods for fairness improvement, and presents a derivative-free optimization method to solve the optimization problem in the MPC. Section 3 describes a case study to implement the proposed fairness-aware data-driven-based MPC to control the TES in a residential building. The predictive performance of DDBMs in the proposed MPC is evaluated in terms of accuracy and fairness in Section 4. This section also analyzes the control performance of the proposed MPC in terms of cost saving, peak shifting, and thermal comfort. Further discussion is presented in Section 5. Finally, Section 6 summarizes the conclusion of this study.

2. Methodology

The general schematic of the proposed fairness-aware data-driven-based MPC will be presented in Section 2.1, which mainly includes data collection, fairness-aware data-driven model development, and optimal control problem construction. Then, a detailed explanation of fairness improvement methods for data-driven models used in this study will be given in Section 2.2. Section 2.3 introduces the optimization algorithm used to solve the optimization problem in the MPC to get the optimal control strategy.

2.1. *Fairness-aware data-driven-based MPC*

The fairness-aware data-driven-based MPC is aimed at obtaining optimal future control actions for systems/devices in buildings based on an accurate and fair prediction for future states. To ensure accurate and fair prediction, fairness-aware DDBMs act as the prediction components in the proposed MPC. Fairness improvement methods are implemented when developing DDBMs. Accordingly, the simplified schematic of the proposed fairness-aware data-driven MPC is shown in Fig. 1. Its general development procedure can be summarized below:

Step 1. Data collection. Weather information is collected from websites or weather stations, while building-related data (such as indoor air temperature, energy consumption, and device operation status) is measured by sensors or BMS installed in a real building (when the study is based on experimental study) or is simulated by a physical building model (when the study is based on numerical study).

Step 2. Fairness-aware data-driven model training and prediction. In this step, users should define the protected feature among which they wish the predictive performance to be uniform. When the output feature is discrete labels, pre-processing methods that remove discrimination from the training dataset would be used as the fairness improvement method, while in-processing methods that add fairness-related constraints or penalties to the objective function of model training could be applied to achieve fairness when the output is continuous data.

Step 3. Construct and solve the optimization problem for MPC to get the optimal future control signals. The predicted values from Step 2 will be integrated into the objective function of MPC. The objective function could be aimed at getting the optimal control actions for devices in a building to achieve the minimum electricity cost or energy usage, or maximum peak shifting while maintaining the thermal comfort in a finite horizon of time.

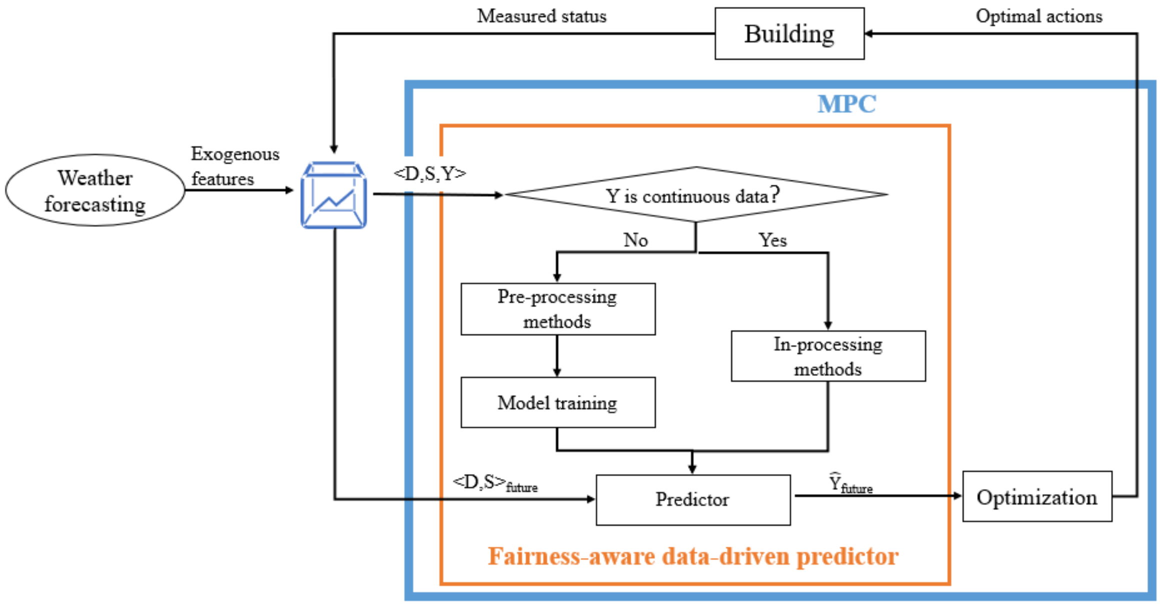


Fig. 1: Simplified schematic of fairness-aware data-driven-based MPC

2.2. *Fairness improvement methods*

Fairness improvement methods could be classified into three categories based on the stage that the method works on: 1) Pre-processing methods. This kind of method usually produces a balanced training dataset to remove discrimination before model training. 2) In-processing methods, which add fairness-related components (e.g., penalties or constraints) to the structure or objective function of a data-driven model during model training. 3) Post-processing methods that modify the predictive result or classification threshold of a model after model training to achieve fairness.

In this study, as the starting point of investigating fairness-aware data-driven-based MPC, two pre-processing methods, i.e., random sampling (RS) and reversed preferential sampling (RPS), will be introduced in this section. Then, their effect on the predictive result and control performance of an MPC will be investigated by a case study presented in Section 3.

Before explaining RS and RPS, let us assume that these methods will be implemented to process a balanced training dataset for an *i-class* prediction problem with a *j-class* protected feature. There would be $i*j$ conditions defined by the protected feature and the output label, see Table 2. The original training dataset is called $X_{candidate}$ and its data count is denoted as $|X_{candidate}|$. The training dataset produced by pre-processing methods is called $X_{designed}$ and its designed number of data points is $|X_{designed}|$. To make $X_{designed}$ balance, the expected number of data points in each condition after data pre-processing is, $\frac{1}{i*j}|X_{designed}|$.

Table 2: Conditions defined by an *i-class* prediction problem with a *j-class* protected attribute *[31]*

		Output label			
		Y_1	Y_2	…	Y_i
Protected attribute	S_1	S_1Y_1	S_1Y_2	…	S_1Y_i
	S_2	S_2Y_1	S_2Y_2	…	S_2Y_i
	…	…		…	…
	S_j	S_jY_1	S_jY_2	…	S_jY_i

Note that S_jY_i is the condition in which the data's protected feature is S_j and the output label is Y_i.

The general procedure of RS and RPS is presented in Fig. 2, while the fundamental of RS and RPS is further explained in Section 2.2.1 and Section 2.2.2, respectively.

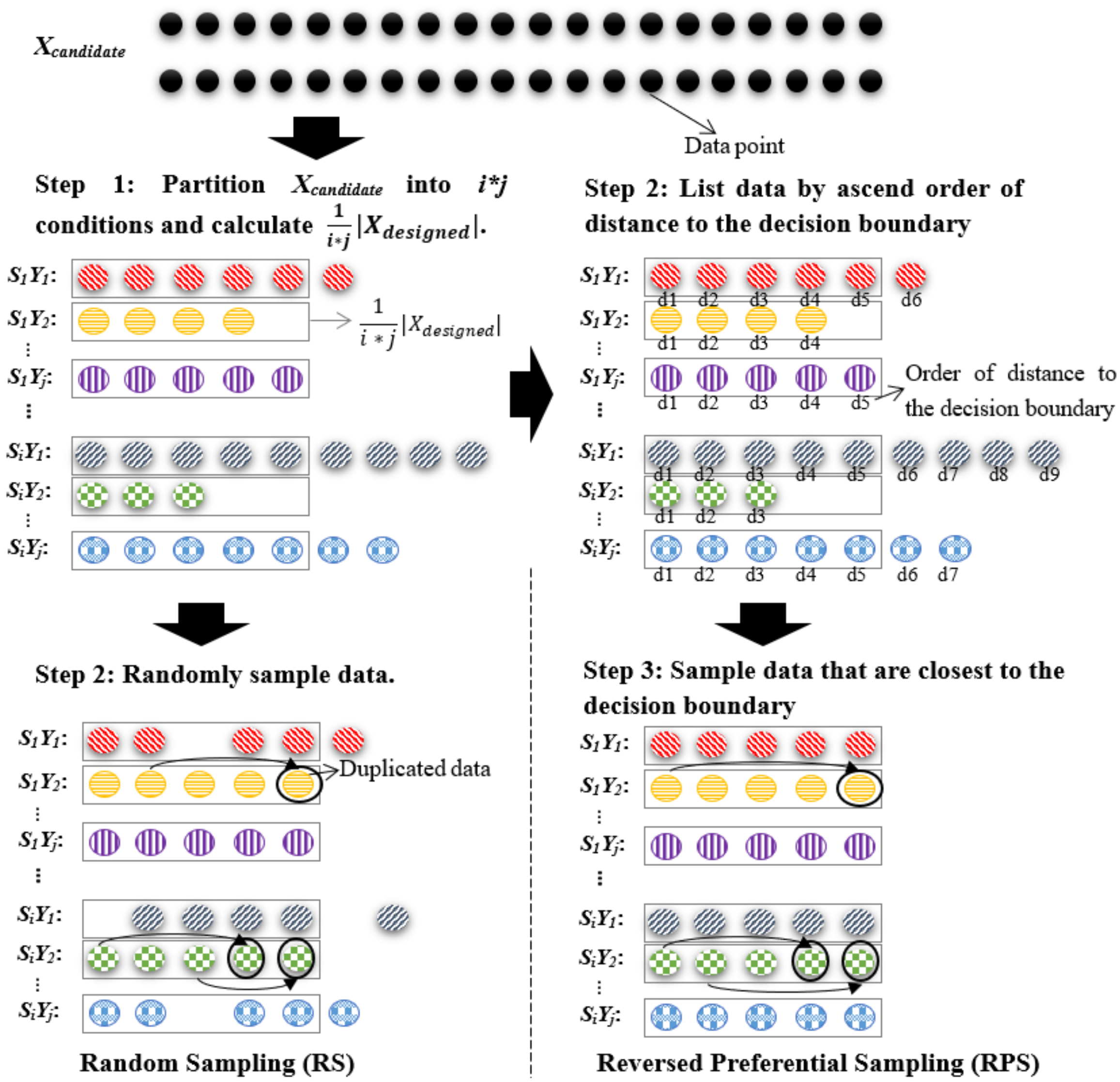


Fig. 2: General procedure of RS and RPS

2.2.1. Random sampling (RS)

Kaviran and Calders [28] proposed random sampling (RS) to randomly sample $\frac{1}{i*j}|X_{designed}|$ data points for each condition listed in Table 2 from $X_{candidate}$ to $X_{designed}$. To be more detailed, if a condition in $X_{candidate}$ contains more than $\frac{1}{i*j}|X_{designed}|$ data, RS would randomly sample $\frac{1}{i*j}|X_{designed}|$ training points from this condition to $X_{designed}$. Otherwise, if the number of data in a condition of $X_{candidate}$ is less than $\frac{1}{i*j}|X_{designed}|$, RS would randomly duplicate points from this condition until the data volume reaches the expected number and then send these data to $X_{designed}$.

2.2.2. Reversed preferential sampling (RPS)

Reversed preferential sampling (RPS) was proposed in Ref [27] to sample data by the order of closeness to the decision boundary (the hypersurface separating the data point's corresponding output label from other classes of output labels). If a condition in $X_{candidate}$ shows less than $\frac{1}{i*j}|X_{designed}|$ data, RPS would duplicate its data that are close to the decision boundary until this condition in $X_{designed}$ gets $\frac{1}{i*j}|X_{designed}|$ data points, while RPS would remove data furthest from the decision boundary for conditions in $X_{candidate}$ with more than $\frac{1}{i*j}|X_{designed}|$ data.

2.3. ***Optimization method***

Considering the complexity of the optimization problem in an MPC and the 'black-box' property of most data-driven models integrated into the MPC, derivative-free optimization algorithms, such as differential evolution (DE), genetic algorithm (GA), and particle swarm optimization (PSO), could be used to solve the optimization problem constructed in the MPC and obtain the optimal control signals. Due to the potential to solve constrained complex optimization problems [34] and to guarantee global optimal [32], DE is selected as the solver for the optimization part of the MPC in this study.

As a heuristic approach, DE obtains the optimal solution based on an evolutionary process that iteratively improves the candidate solution [49]. Its general procedure was summarized in Ref [32] and shown in Fig. 3. Here, a brief explanation for this figure is given as follows:

Step 1. Population Initialization. This step would initialize a random or user-defined population with a set of candidate solutions.

Step 2. Fitness assignment. Each candidate solution would be evaluated in this step by a fitness score calculated by a fitness function, to determine how fit the solution is.

Step 3. Stop criteria evaluation. Terminate the algorithm if the offspring of the corresponding population does not significantly improve the fitness score, or if the time of iterations reaches its threshold. Otherwise, if the stop criteria are not met, continue to Step 4.

Step 4. Selection. A pre-defined selection procedure (such as random selection) would select a set of solutions (parents) for the next step: mutation.

Step 5. Mutation. Mutate a unit vector by adding a scaled differential vector to a target vector. Here, the differential vector is the difference between the two or more parents selected from Step 4, while the target vector is the parent with a prioritized direction of creating the unit vector.

Step 6. Crossover and go back to Sep 2. Crossover means to generate new offspring for the population by crossing over a 'major' parent and the unit vector created from Step 5.

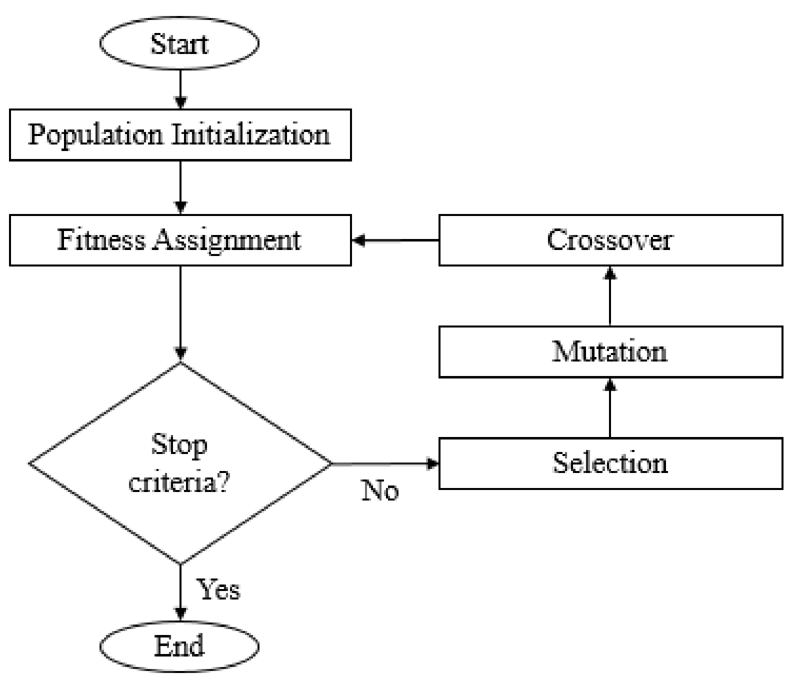


Fig. 3: General procedure of a DE *[32]*

3. Case Study

In the case study, fairness-aware data-driven-based MPCs will be developed to control the electrically heated floor (EHF) system in a residential building for the winter of 2021 (January to March). The effect of fairness-aware MPCs on heating cost, peak shifting, and thermal comfort will be compared with traditional MPCs in which data-driven models do not consider fairness improvement. The case study will be done by simulation on a validated TRNSYS model.

In this section, the experimental building and the electricity price are first introduced in Section 3.1. Then, the EHF system in the experimental building and the TRNSYS model of this building are presented in Section 3.2. Detailed information for the fairness-aware data-driven-based MPC development is explained in Section 3.3. It would develop fairness-aware DDBMs for heating load prediction and indoor air temperature classification, and then, integrate these models into MPCs to get the optimal set-point temperature for EHFs to minimize the heating cost while maintaining an acceptable thermal comfort.

3.1. *Experimental building*

The experimental building, as shown in Fig. 4, is a traditional residential building located in Montreal, Quebec, Canada. It was built in the year 1960, with a building area of 104 m^2. There are 6 rooms in the basement and 6 rooms in the ground floor.

Fig. 4: Experimental building

In the city where the experimental building is located, distinct electricity prices for peak periods and off-peak periods are implemented to encourage consumers to shift their electricity demand and reduce winter grid peaks. The electricity price during peak periods and off-peak periods is listed in Table 3. Besides, it was mentioned that peak periods often occur during the early morning and late afternoon. Therefore, in this study, to simplify the model complexity for MPCs, the peak periods are assumed as 6 am to 12 pm and 6 pm to 0 am, while the off-peak periods include 0 am to 6 am and 12 pm to 6 pm. The duration of peak periods and off-peak periods per day is presented in Fig. 5.

Table 3: Electricity price during winter implemented in Montreal, Canada [6]

		Periods	
	Condition	Peak period	Off-peak period
Electricity price (¢/kWh)	< 40 kWh/per day	50	3.98
Electricity price (¢/kWh)	> 40 kWh/per day	50	7.03
Subscription fee (¢/day)	-	40.64	

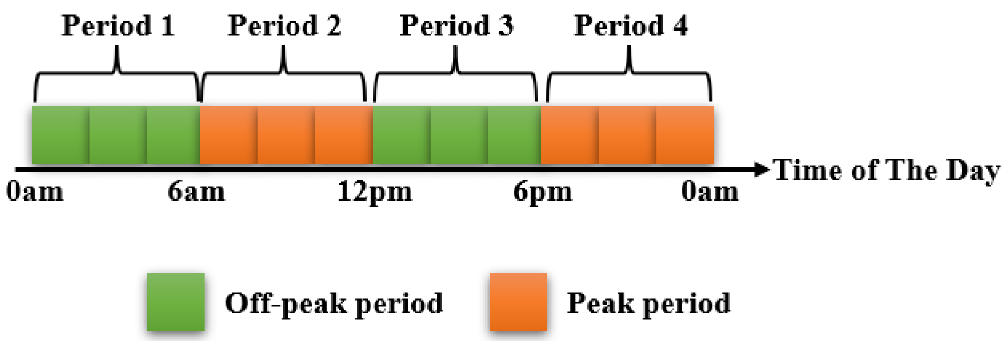


Fig. 5: Duration of off-peak periods and peak periods

3.2. *EHF system and TRNSYS model*

The building was originally simulated as a multi-zone TRNSYS (TRaNsient SYStems simulation program) model created and validated by Aongya [50]. In this model, occupancy status, lighting energy consumption, and appliance energy consumption are modeled with a measured schedule. The building was heated by electric baseboards. Then, to investigate the applicability of electrically heated floor (EHF) on peak shifting, Thieblemont et al. [51] removed the basement in the model and replaced the baseboards with a commonly used EHF system in Quebec, Canada. The assembly of EHFs is present in Fig. 6, while the properties of its materials are listed in Table 4. The simulated one-story building was validated. Then, the building model is further simplified into a single-zone model that combines the ground floor as one zone heated by EHFs. To do so, the internal envelopes have been eliminated. The schematic of the modified TRNSYS model is presented in Fig. 7. Detailed description of these components could be found in previous studies [52,53]. The simplified single-zone TRNSYS model was validated by setting the same set-point temperature as the multi-zone model and getting comparable indoor air temperature and energy consumption. A detailed validation procedure will not be presented in this paper, as it is not the focus of this paper.

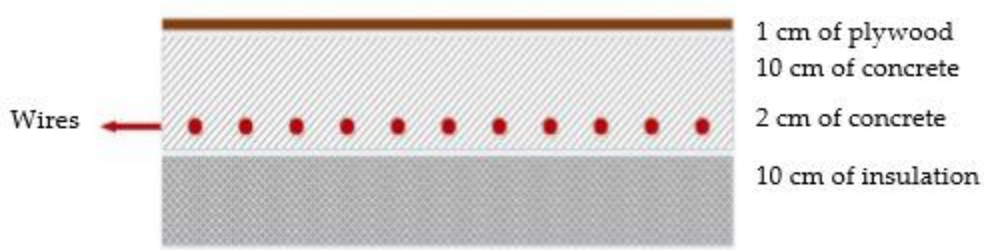


Fig. 6: Assembly of EHFs [51]

Table 4: Thermophysical properties of floor layers *[51]*

Material	**Conductivity (W/mk)**	**Specific Heat (kJ/kgK)**	**Density (kg/m^3)**
Concrete	2.25	0.99	2200
Insulation (XPS)	0.04	1.5	35
Floor Covering (plywood)	0.164	1.63	670

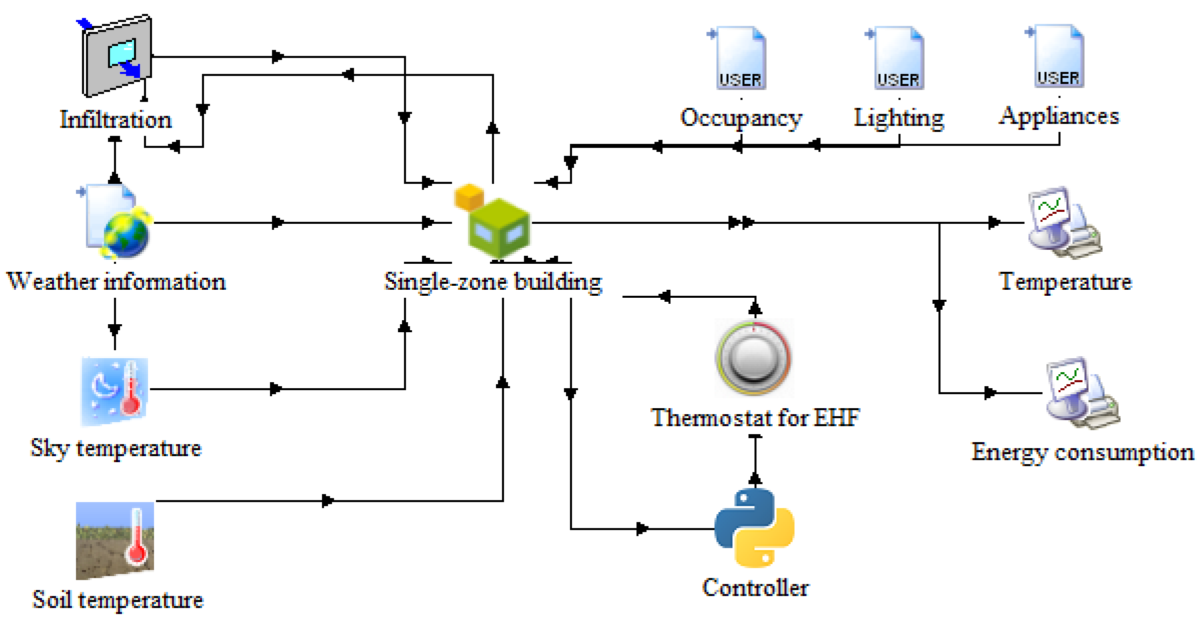


Fig. 7: Schematic of the TRNSYS model

3.3. *Fairness-aware data-driven MPC development*

In this section, the fairness-aware data-driven-based MPC will be developed following the procedure illustrated in Section 2.1. Detailed explanation for each step is presented below:

Step 1. Data collection.

To collect the training data, the TRNSYS model is simulated based on weather data collected from January to March from the weather station located at Montréal-Pierre Elliott Trudeau International Airport for the year 2011 to 2021. The time interval is 1 hour. The building is heated by the EHF with a random hourly integer set-point within the range of [18 °C, 24 °C]. Data obtained from the TRNSYS model include energy consumed by the EHF and indoor air temperature. Because DDBMs could be time series prediction, 12 hours' time lag of set-point temperature, energy consumption, and indoor air temperature are added as candidate features.

Step 2. Fairness-aware data-driven model training and prediction.

Two DDBMs would be developed to predict the energy consumption and indoor air temperature, respectively. The energy prediction would be used to calculate the electricity bill or the difference between energy consumed during peak periods and during off-peak periods, while the indoor air temperature prediction would be used to constrain the set-point in order to meet thermal comfort.

The energy prediction model (see Equation 1) predicts the energy consumption for a 6-hour period (off-peak period or peak period). Here, the prediction is based on a support vector machine (SVM) model with a linear kernel. SVM makes predictions by determining a hyperplane that maximizes the margin [54] and is a kind of popular machine learning algorithm in building engineering. Input features for the energy prediction SVM model include hourly setpoint during the corresponding 6-hour period ($\mathrm{T}_{set,i}$), and hourly outdoor ambient temperature during that period ($\mathrm{T}_{ambient,i}$) and previous period ($\mathrm{T}_{ambient,i-1}$). The reason for model selection and feature selection is presented in the Supplementary Information.

$$\widehat{Q_\iota} = \mathrm{f}(\mathrm{T}_{set,i}, \mathrm{T}_{ambient,i}, \mathrm{T}_{ambient,i-1}) \tag{1}$$

where $\widehat{Q_\iota}$ is the predicted energy consumption during *i*-th period, kWh; as shown in Fig. 5, *i*=1 means the period of 0:00 am to 6:00 am, *i*=2 means the period of 6:00 am to 12:00 pm, *i*=3 is the period of 12:00 pm to 6:00 pm, and *i*=4 is the period of 6:00 pm to 0:00 am; $\mathrm{T}_{set,i}$ is a list of hourly indoor air set-point temperature at *i*-th period, °C; $\mathrm{T}_{ambient,i}$ is a list of hourly outdoor ambient temperature at *i*-th period, °C.

The indoor air temperature prediction model (see Equation 2) is actually a classification model that determines whether the indoor air temperature is lower than the threshold temperature for thermal comfort (such as 21 °C required in Montreal [55]). Here, SVM with a 'linear' kernel function is selected as the indoor air temperature prediction model, while $T_{set,i}$, $T_{set,i-1}$, $T_{ambient,i}$, and $T_{ambient,i-1}$ are input features. The reason for model selection and feature selection is also presented in the Supplementary Information.

$$\widehat{T_{min,i}} = f(T_{set,i}, T_{set,i-1}, T_{ambient,i}, T_{ambient,i-1}) \tag{2}$$

where $\widehat{T_{min,i}}$ is the binary label that illustrates if the predicted minimum indoor air temperature during *i*-th period is lower than 21°C, $\widehat{T_{min,i}} \in [-1,1]$. $\widehat{T_{min,i}}$=-1 means that during the predicted period, indoor air temperature is higher than 21°C.

Before model training, data simulated from the year 2011 to the year 2020 is used as $X_{candidate}$. Then, three kinds of $X_{designed}$ are produced for training DDBM by directly using $X_{candidate}$ (Reference case) or implementing pre-processing methods (RS or RPS). For both the energy prediction model and the indoor air temperature prediction model, the protected feature is the peak/off-peak period label. The data for the year 2021 is used as the validation dataset.

Step 3. Construct the MPC and solve the optimal heating strategy.

The goal of this MPC is to get the hourly heating set-point temperature that could minimize the daily electricity bill. The indoor air temperature should be kept higher than 21 °C in the future 24 hours, while the set-point temperature should be integer and within the range of [18 °C, 24 °C]. The objective function of MPC is presented in Equation 3.

$$\min \sum_{i=1:4} \widehat{Q_i} * Price_i \tag{3}$$

Subject to

$$\widehat{T_{min,i}} = \text{Negative}, \forall i \in [1,4],$$

$$18\ °C \le T_{set,ji} \le 24\ °C, \forall i \in [1,4], j \in [1,6]$$

In this study, ‘Constant 21’, ‘MPC_ReferenceCase’, ‘MPC_RS’, and ‘MPC_RPS’ are developed and compared. Here, ‘MPC_RS’, and ‘MPC_RPS’ are fairness-aware MPCs developed based on the training dataset processed by RS or RPS. ‘MPC_ReferenceCase’ is a traditional MPC that does not consider fairness, while ‘Constant 21’ is a traditional controller that sets set-point temperature at 21 °C.

4. Results

In this section, the effect of considering fairness on the predictive performance of DDBMs, such as energy prediction models and indoor air temperature models, is first evaluated in Section 4.1. Then, in Section 4.2, the control performance of fairness-aware data-driven-based MPCs is compared to traditional controllers in terms of peak shifting, heating cost saving, and thermal comfort.

4.1. ***Predictive performance of data-driven models***

4.1.1. Predictive result of energy prediction models

Before analyzing the predictive performance, data distribution should be first summarized. $X_{candidate}$ distribution on energy consumption is shown in Fig. 8. Most of the time, energy consumption during a period is around 16.51 kWh to 33.02 kWh. For energy consumption at 0 kWh and 8.26 kWh, the off-peak period shows more samples than the peak period, while for energy consumption at 16.51 kWh to 33.02 kWh, peak periods have more samples. The number of samples is almost similar between peak periods and off-peak periods when the energy consumption is 41.28 kWh or 49.53 kWh. Besides, the validation data distribution among energy consumption categories are present in Fig. 9. Its pattern is similar to $X_{candidate}$. The difference is that energy consumption at 16.51 kWh during the off-peak period has more data than the peak period in the validation dataset.

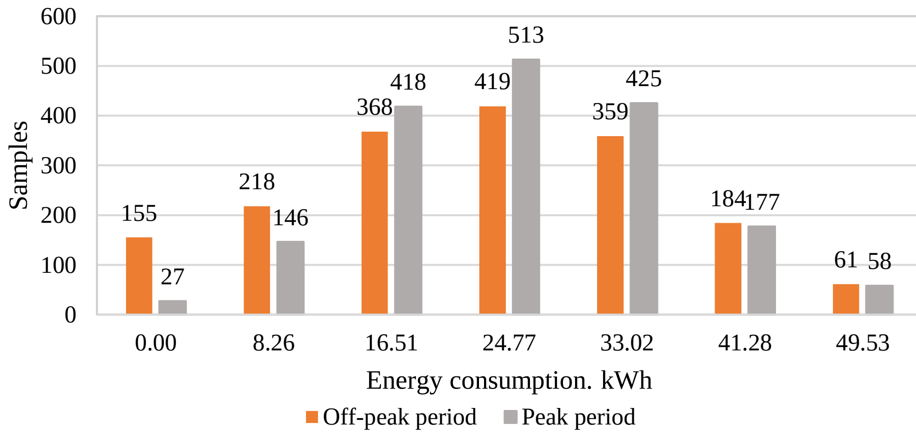


Fig. 8: $X_{candidate}$ distribution on energy consumption

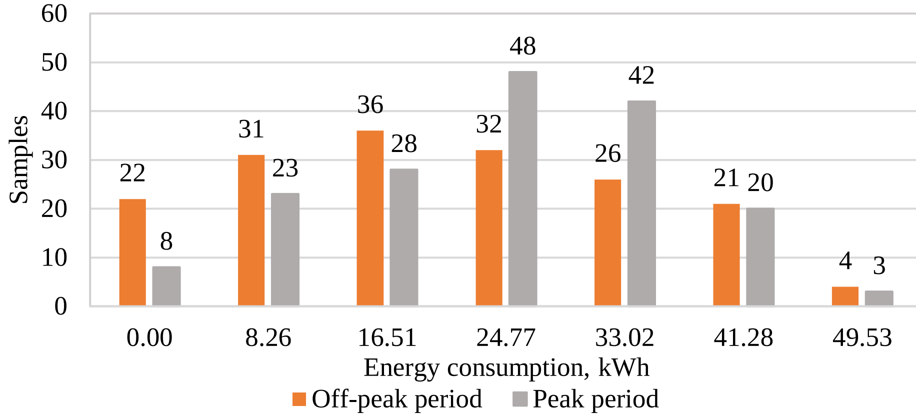


Fig. 9: Validation data distribution on energy consumption

The energy predictive performance under different conditions is summarized in Table 5 in terms of mean bias error (MBE, calculated by Equation 4) and mean absolute error (MAE, calculated by Equation 5). The reasons for selecting MBE and MAE as the performance criteria are that MBE indicates the difference between the average predicted value and the average measured value, while MAE calculates the mean value of absolute errors and shows the benefit of no error cancellation (the phenomenon that errors caused by overestimated results is omitted by errors resulted from underpredicted values).

$$\text{Mean Bias Error (MBE)} = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \tag{4}$$

$$\text{Mean Absolute Error (MAE)} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{5}$$

Table 5 shows that for the condition that contains all types of measured values (named "ALL Y"), the reference case shows better predictive accuracy with a lower MAE no matter during model training (5.80 kWh) or model validation (6.12 kWh), while RPS usually shows more negative effect on the predictive accuracy than RS through having a higher MAE (7.58 kWh and 7.56 kWh during model training and model validation, respectively). However, RPS and RS with lower MBE during model validation reflect that they could keep the average predicted value closer to the average measured value. Another interesting finding is that although in the "ALL Y" condition data collected during peak periods are the same as during off-peak periods no matter in the training dataset or the validation dataset, RS could narrow the MAE difference between peak periods and off-peak periods to 0.18 kWh and 0.14 kWh for training and validation, respectively, while in the reference case the MAE difference is 0.83 kWh and 0.82 kWh for training and validation, respectively.

To analyze the effect of pre-processing methods on the predictive result of conditions with different data volumes, predictive performance during model training and validation when the measured energy consumption is 0 kWh, 24.77 kWh, and 41.28 kWh are also summarized in Table 5. The results for conditions “Y=0 kWh” and “Y=41.28 kWh” indicates that increasing samples in minority conditions could increase the predictive accuracy: both RS and RPS could decrease the absolute value of MAE and MBE compared to the reference case. By contrast, RPS and RS would decrease the predictive accuracy of majority conditions: for the condition “Y=24.77 kWh”, the MAE has been increased when using RPS and RS to undersample data. Besides, RS and RPS could improve the predictive fairness in terms of having smaller MBE difference between majority conditions and minority conditions: In the reference case, the overall MBE difference between the condition “Y=0 kWh” and “Y=24.77 kWh” is 9.07 kWh and 8.29 kWh during training and validation, respectively. RS reduced the difference to 4.37 kWh and 2.93 kWh for training and validation, respectively, while RPS further reduced the difference to 4.25 kWh and 1.21 kWh for training and validation, respectively.

Table 5: Energy predictive performance under different conditions in terms of MBE and MAE

			All Y		Y=0kWh		Y=24.77kWh		Y=41.28kWh	
			MBE [kWh]	**MAE [kWh]**	**MBE [kWh]**	**MAE [kWh]**	**MBE [kWh]**	**MAE [kWh]**	**MBE [kWh]**	**MAE [kWh]**
Training	**Overall**	**Reference case**	**-0.05**	**5.80**	-8.80	8.80	**0.27**	**4.73**	6.31	6.31
		RS	0.23	6.34	-4.08	4.08	0.29	8.45	**-0.07**	**5.86**
		RPS	0.54	7.58	**-3.55**	**3.55**	0.70	13.05	1.74	6.94
	Off-peak period	**Reference case**	-0.12	**6.21**	-7.99	7.99	0.65	**5.42**	5.92	5.92
		RS	**-0.01**	6.43	-4.55	4.55	**0.20**	8.98	-0.98	**5.83**
		RPS	0.33	7.55	**-4.52**	**4.52**	0.52	12.32	**0.88**	7.50
	Peak period	**Reference case**	**0.02**	**5.38**	-13.45	13.45	**-0.05**	**4.17**	6.72	6.72
		RS	0.47	6.25	-3.60	3.60	0.39	7.93	**0.85**	**5.90**
		RPS	0.76	7.60	**-2.59**	**2.59**	0.88	13.79	2.59	6.39
Validation	**Overall**	**Reference case**	-1.13	**6.12**	-9.08	9.08	**-1.24**	**4.75**	7.05	**7.05**
		RS	**-0.26**	7.32	-4.68	4.68	-1.75	8.77	3.83	7.85
		RPS	-0.12	7.56	**-3.58**	**3.58**	-2.37	8.56	**2.21**	**7.05**
	Off-peak period	**Reference case**	-1.34	**6.53**	-7.13	7.13	**-1.29**	**6.45**	7.08	**7.08**
		RS	-0.67	7.39	-4.50	4.50	-4.13	10.83	5.50	7.86
		RPS	**-0.43**	7.06	**-3.75**	**3.75**	-2.58	8.26	**2.36**	**7.08**
	Peak period	**Reference case**	-0.91	**5.71**	-14.45	14.45	-1.20	**3.61**	7.02	**7.02**
		RS	**0.14**	7.25	-5.16	5.16	**-0.17**	7.40	**2.06**	7.84
		RPS	0.19	8.06	**-3.10**	**3.10**	-2.24	8.77	**2.06**	**7.02**

Note that the numbers in **bold format** mean the corresponding pre-processing method (Reference case, RS, or RPS) results in the lowest MAE or MBE under the same kind of duration and measured value.

4.1.2. Predictive result of air temperature prediction models

$X_{candidate}$ and validation data distribution on air temperature category and peak/off-peak periods are summarized in Table 6. In all datasets, the minimum indoor air temperature during a period is lower than 21 °C most of the time. Besides, off-peak periods have more data with negative air temperature category than peak periods, while peak periods have more data with an air temperature that is lower than 21 °C.

Table 6: $X_{candidate}$ and validation data distribution on air temperature category and peak/off-peak periods

	Air temperature category (T_{min})	**Off-peak period**	**Peak period**
$X_{candidate}$	**-1**	706	491
	1	1058	1273
Validation dataset	**-1**	79	48
	1	93	124

Note that T_{min}= -1 means that the minimum indoor air temperature during the corresponding period is higher than 21°C. while T_{min}= 1 means that the minimum indoor air temperature during that period is lower or equal to 21°C.

Fig. 10 shows that RPS decreases the overall predictive accuracy from 82.6% to 75.4% during model training, while RS could preserve it at 81.8%. However, RPS could ensure the overall predictive accuracy to be higher than 79% during model validation, while RS decreases the accuracy to 77.3%. RS and RPS do not significantly affect the accuracy difference between the off-peak period and the peak period. This is because the number of samples collected during the peak period is the same as the off-peak period.

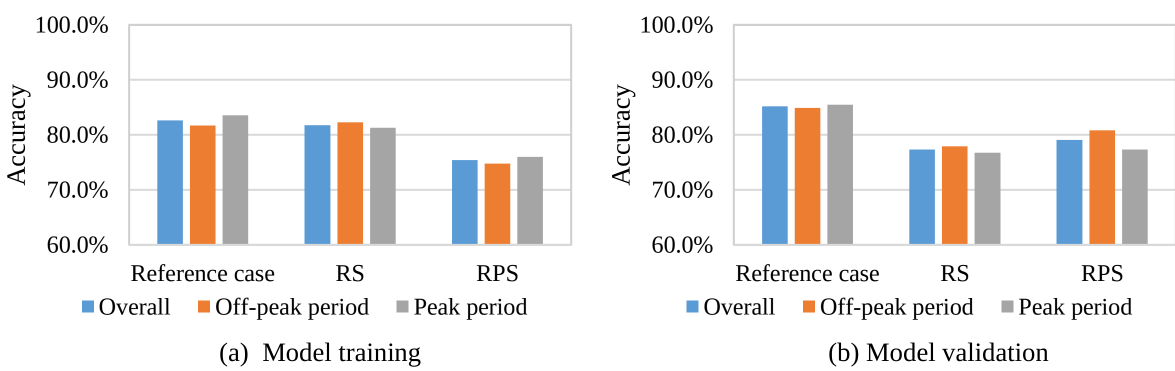
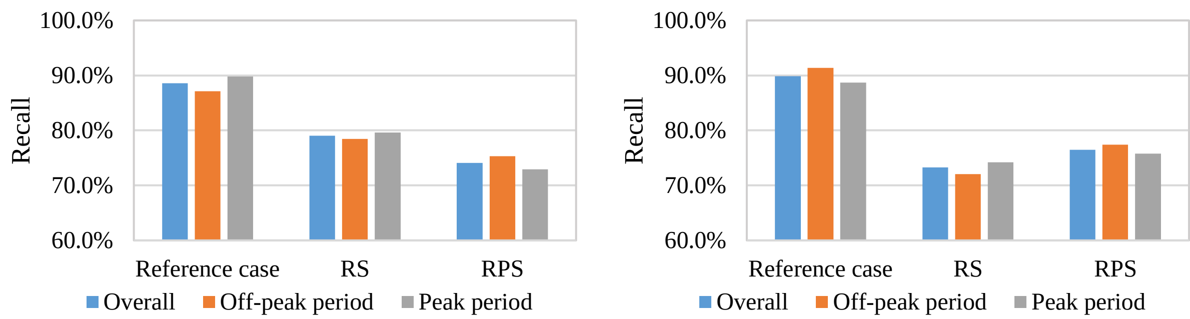


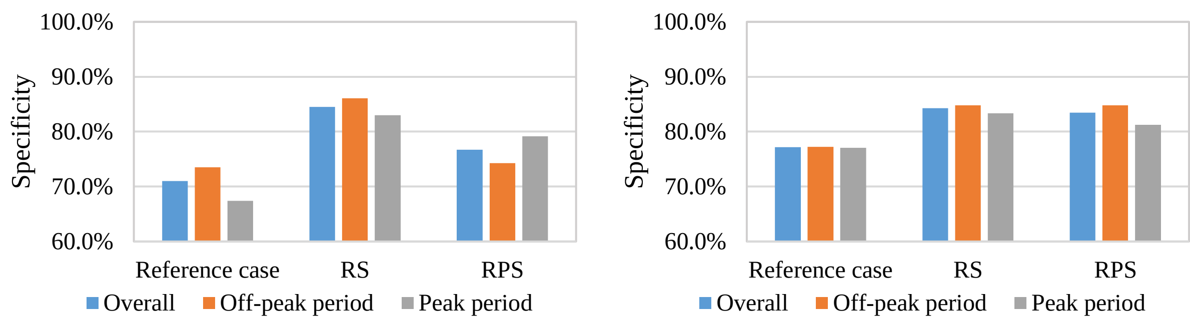
Fig. 10: Predictive accuracy of air temperature category during (a) model training and (b) model validation

From Fig. 11, RPS and RS decrease the predictive recall, which indicates the proportion of correctly predicted samples when their actual temperature category is '**1**'. The effect of RS and RPS on recall is like their effect on the overall accuracy because air temperature category '**1**' is the majority condition during model training and validation. By contrast, as shown in Fig. 12, RPS and RS could effectively increase the predictive specificity, which refers to the proportion of correctly predicted samples when their actual temperature category is ‘**-1**’. It means that RS and RPS could increase the predictive accuracy of the minority condition because of oversampling. The difference between recall and specificity is 17.6% during model training and 12.7% during model validation in the reference case. RPS could effectively decrease the difference between recall and specificity to 2.5% during model training and 7.0% during model validation, while RS reduces the difference to 5.5% during model training and 11% during model validation.



(a) Model training (b) Model validation

Fig. 11: Predictive recall of air temperature category during (a) model training and (b) model validation



(a) Model training (b) Model validation

Fig. 12: Predictive specificity of air temperature category during (a) model training and (b) model validation

4.2. ***Control performance comparison***

Fig. 13 compares the effect of MPCs on daily heating costs with a traditional controller that constantly sets the set-point at 21 °C. Using MPC could decrease the daily heating cost by ~17.8% - ~21.8%. Integrating pre-processing methods into the MPC would not reduce its cost-saving ability. In fact, MPC_RS and MPC_RPS save more heating costs than MPC_ReferenceCase.

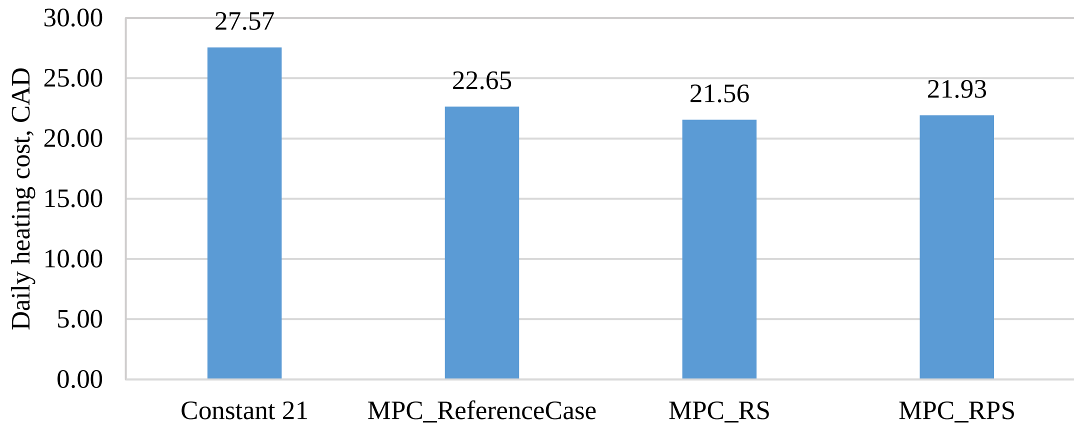


Fig. 13: Daily heating cost of different controllers

The heating cost saving of MPCs is caused by taking advantage of peak shifting. From Fig. 14, the effect of MPCs on daily energy consumption is negligible. However, they are effective in increasing the average energy consumption during off-peak periods and decreasing the energy consumption during peak periods. Integrating fairness-aware data-driven models would not decrease the peak shifting ability of MPCs.

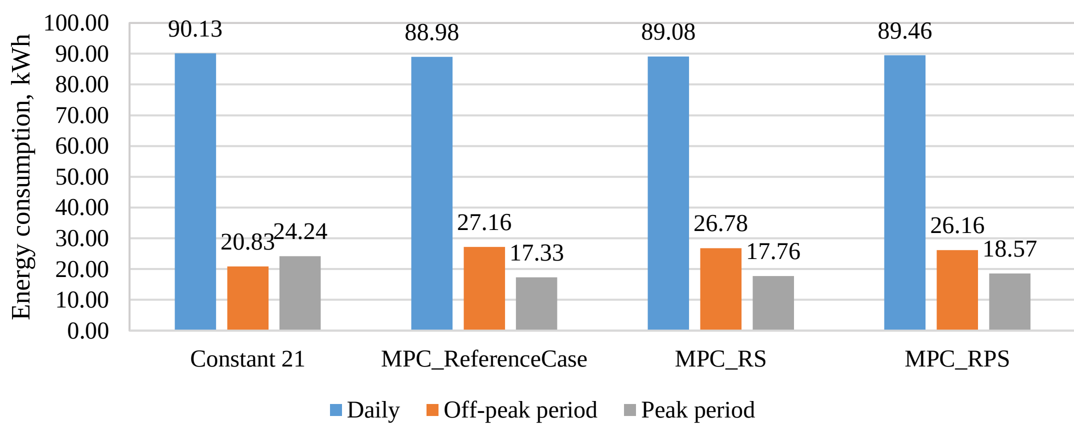


Fig. 14: Average energy consumption of different controllers

Moreover, as shown in Fig. 15, although MPC_RS shows a negative effect on thermal comfort in terms of durations with indoor air temperature lower than 21 °C, MPC_ReferenceCase and MPC_RPS would improve thermal comfort. However, the thermal comfort still needs to be improved, as the indoor air temperature is lower than 21 °C for over 41.6% of simulation times. Potential thermal comfort improvement strategies are discussed in the next section.

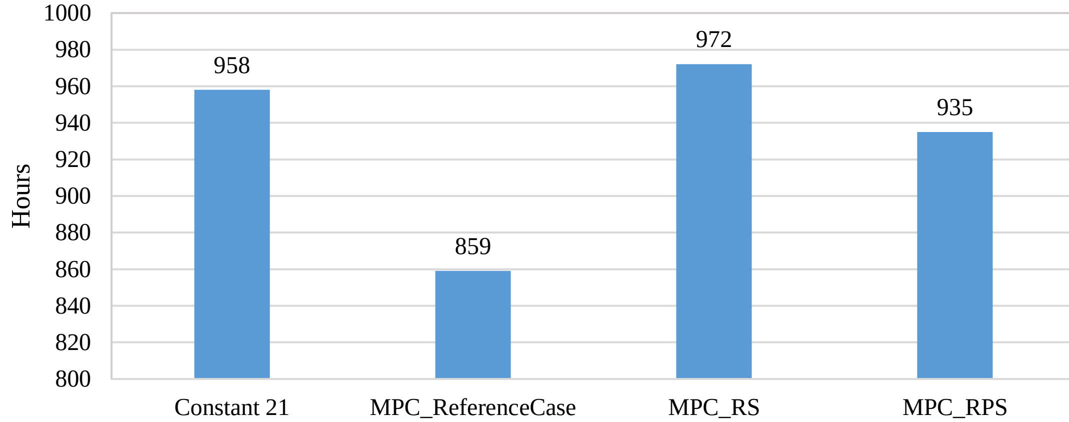


Fig. 15: Number of hours that indoor air temperature is lower than 21 °C

The hourly set-point temperature of MPC_ReferenceCase, MPC_RS, and MPC_RPS is presented in Fig. 16, Fig. 17, and Fig. 18, respectively. These figures illustrate that MPCs, especially MPC_RS, usually set higher set-point temperatures for off-peak periods and lower for peak periods. They also give a higher set-point temperature at the beginning of each peak period. It may be because MPCs try to maintain the indoor air temperature to be higher than 21 °C during the remaining time of the peak period. However, this phenomenon would limit their peak shifting ability. Reasons behind this limitation include 1) the predictive accuracy of energy predictors still needs to be improved to accurately predict future energy consumption, and 2) the optimization algorithm should be improved to get the global optimal solution.

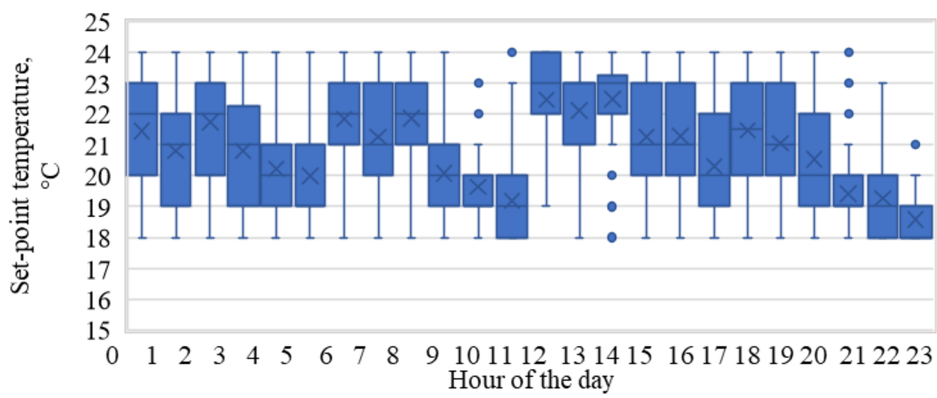


Fig. 16: Hourly set-point temperature of MPC_ReferenceCase

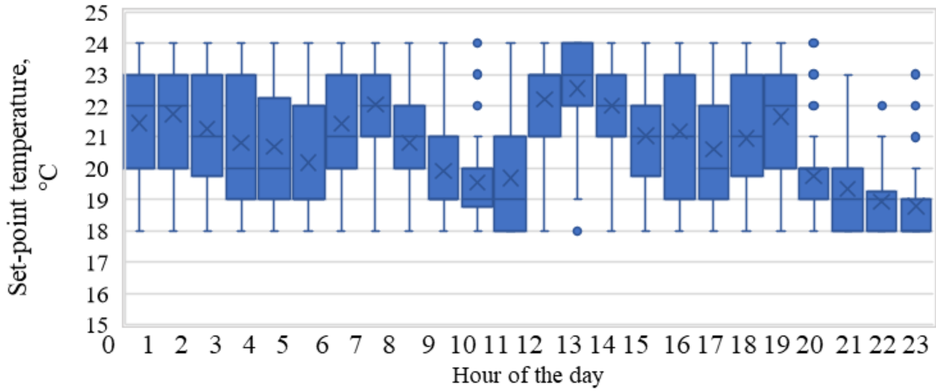


Fig. 17: Hourly set-point temperature of MPC_RS

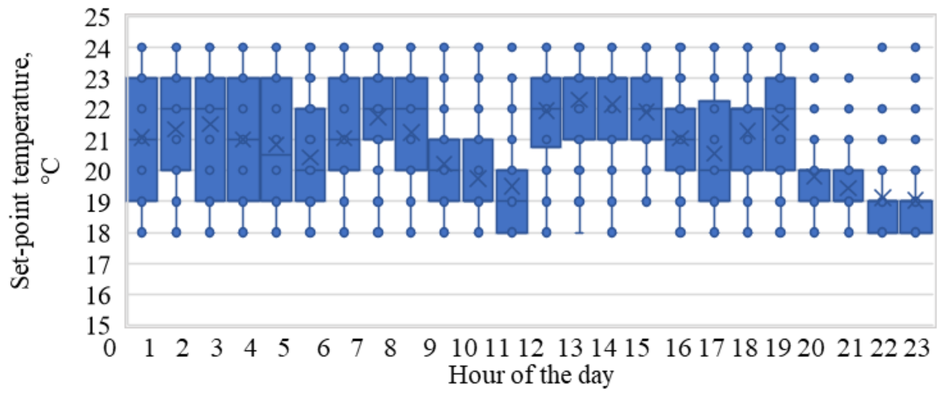


Fig. 18: Hourly set-point temperature of MPC_RPS

5. Discussion

5.1. *Effect of increasing the lower bound for set-point temperature*

As illustrated before, even if constraints are added to MPCs to maintain the predicted air temperature to be higher than 21 °C, the indoor air temperature simulated based on the optimized set-point temperature could still be lower than 21 °C for over 850 hours. To solve this problem, increasing the lower bound for set-point temperature from 18 °C to 21 °C would be a good solution. As shown in Fig. 19, this solution would effectively reduce the duration of an uncomfortable indoor air temperature to less than 300 hours. Fairness-aware MPCs (MPC_RS and MPC_RPS) show a slightly worse effect on improving thermal comfort than MPC_ReferenceCase. However, they could save a little more heating cost than MPC_ReferenceCase (see Fig. 20). Fig. 21 shows that the peak shifting ability of MPCs with 21 °C lower bound set-point temperature is negligible.

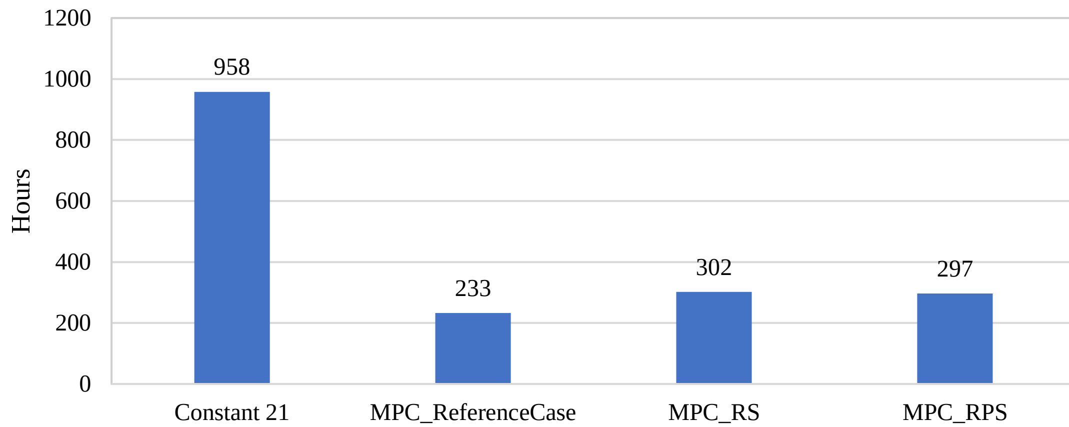


Fig. 19: Number of hours that indoor air temperature is lower than 21 °C when setting 21°C as the lower bound set-point temperature in MPCs

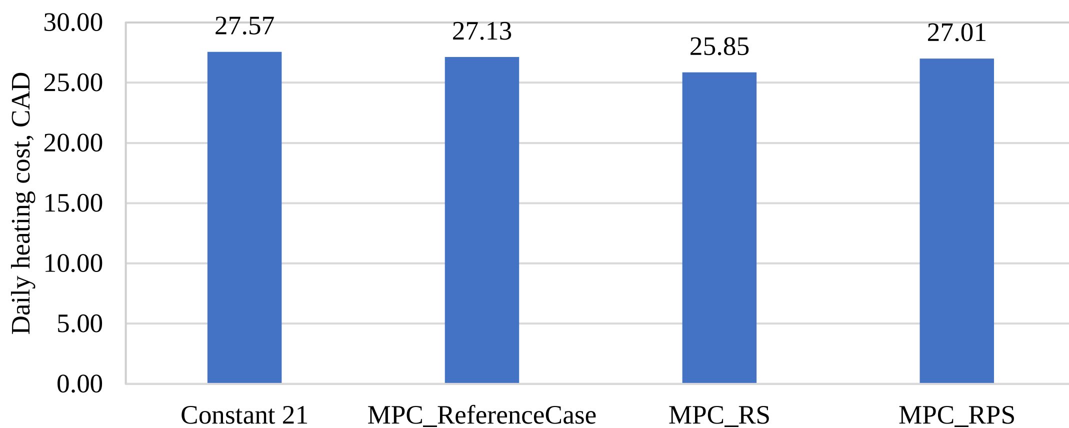


Fig. 20: Daily heating cost when setting 21°C as the lower bound set-point temperature in MPCs

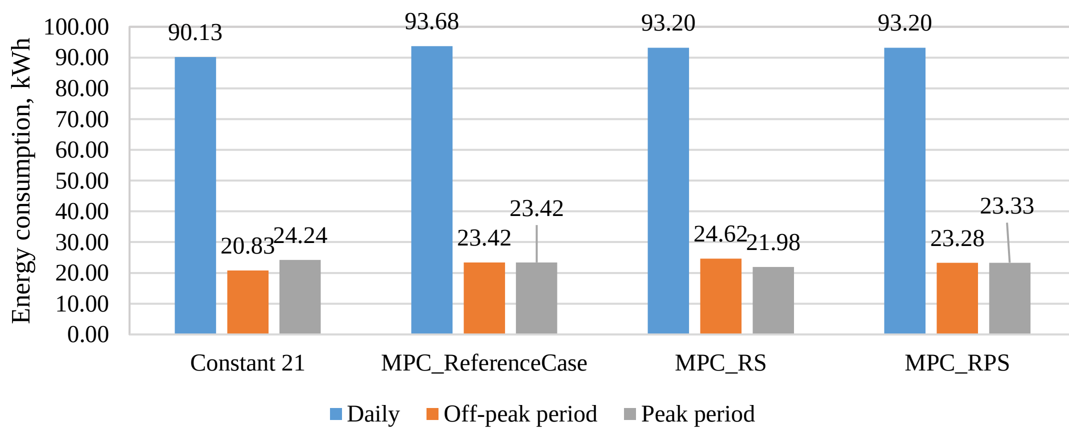


Fig. 21: Average energy consumption when setting 21°C as the lower bound set-point temperature in MPCs

5.2. *Effect of maximizing peak shifting*

To maximize the amount of energy shifted to off-peak periods, the objective function of MPCs could be written as Equation 6. It maximizes the difference between energy consumed during off-peak periods and during peak periods. It could be rewritten as minimizing the energy consumption difference between the peak period and the off-peak period, as shown in Equation 7.

$$\max \widehat{Q_1} + \widehat{Q_3} - \widehat{Q_2} - \widehat{Q_4} \tag{6}$$

Subject to

$$\widehat{T_{min,i}} = \text{Negative}, \forall i \in [1,4],$$

$$18\ °C \le T_{set,ji} \le 24\ °C, \forall i \in [1,4], j \in [1,6]$$

$$\min \widehat{Q_2} + \widehat{Q_4} - \widehat{Q_1} - \widehat{Q_3} \tag{7}$$

Subject to

$$\widehat{T_{min,i}} = \text{Negative}, \forall i \in [1,4],$$

$$18\ °C \le T_{set,ji} \le 24\ °C, \forall i \in [1,4], j \in [1,6]$$

Fig. 22 shows that using maximizing peak shifting as the objective function of MPCs would shift ~65% of peak load from peak periods to off-peak periods, compared to the case with a constant set-point temperature at 21 °C. This objective function is more powerful in shifting peak load than the original one, which is aimed at minimizing the heating cost.

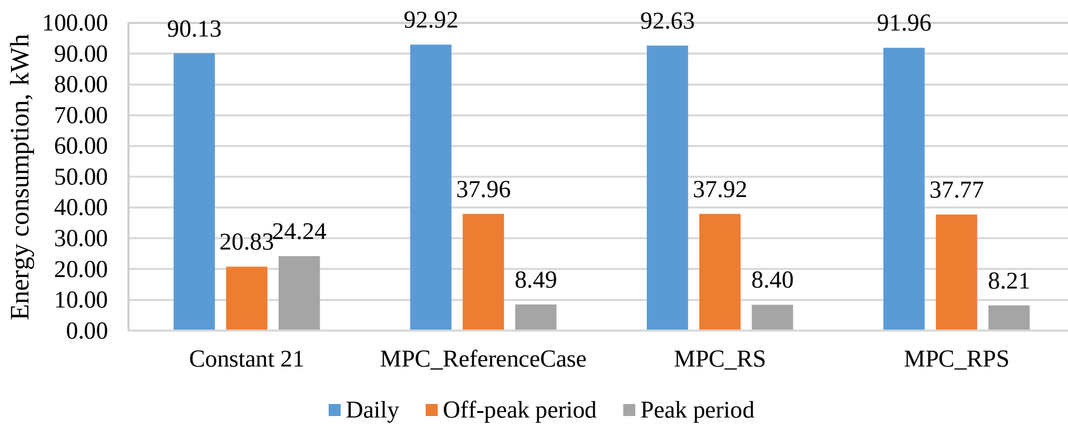


Fig. 22: Average energy consumption when maximizing peak shifting by MPCs

One interesting finding is that although the modified objective function is aimed at maximizing peak shifting, it also works better than the original one on cost saving. From Fig. 23, the daily heating cost of modified MPCs in this section is ~14 CAD, while the daily heating cost of the original MPCs is ~22 CAD, as shown in Fig. 13. The cost-saving of modified MPCs result from setting higher set-point temperature for off-peak periods than peak periods, as shown in Fig. 24. Furthermore, the worse result of original MPCs reveals that the optimization progress of original MPCs is not converged. This may be caused by their more complicated objective function when considering segmented electricity prices. Selecting a proper optimization algorithm or improving hyperparameters used in the original optimization algorithm (DE) would be potential solutions.

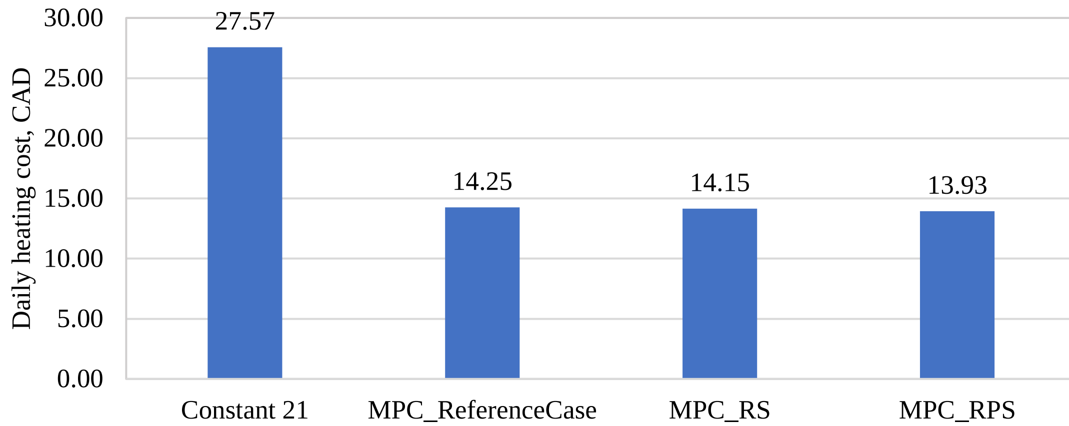


Fig. 23: Daily heating cost when maximizing peak shifting by MPCs

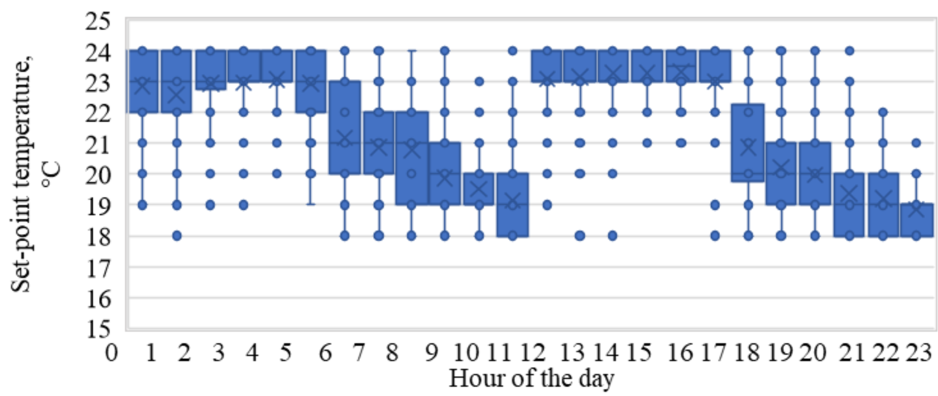


Fig. 24: Hourly Set-point temperature of MPC_ReferenceCase using maximizing peak shifting as the objective function

The effect of the setting maximizing the peak shifting ability as the objective function in MPCs on thermal comfort is presented in Fig. 25. The trend is in line with original MPCs that are aimed at minimizing heating cost: MPC_ReferenceCase works better on preserving the indoor air temperature over 21°C than MPC_RS and MPC_RPS. This is because the overall predictive accuracy of the air temperature prediction model used in MPC_ReferenceCase is higher than MPC_RS and MPC_RPS.

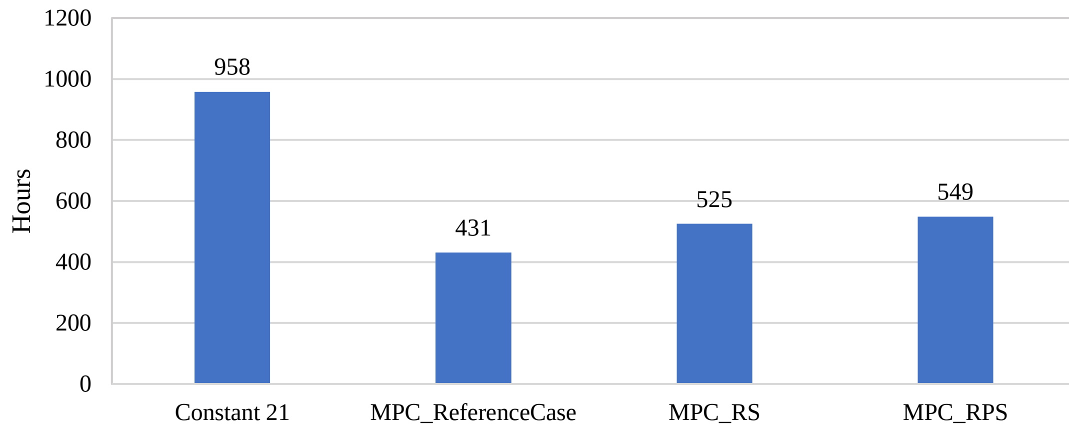


Fig. 25: Number of hours that indoor air temperature is lower than 21 °C when maximizing peak shifting by MPCs

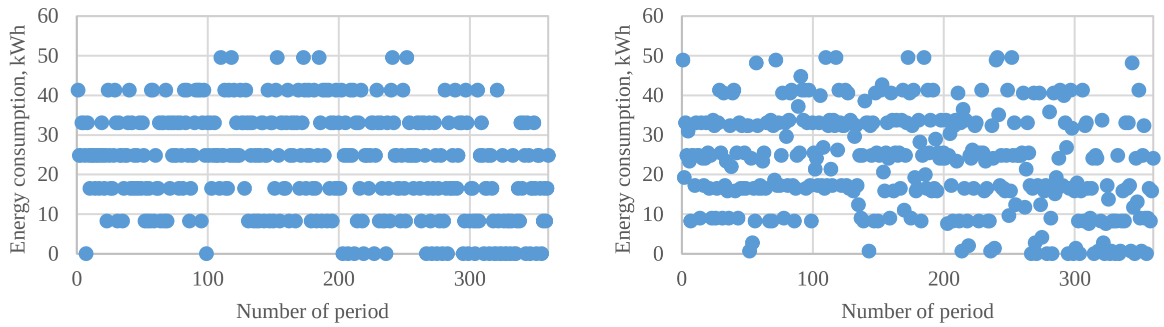
5.3. ***Software and hardware for real application***

In this study, the proposed fairness-aware data-driven MPC was formulated by Python 3.7. It is then utilized to control the simulated experimental building in the TRNSYS model. All study cases were run on a computer with Intel Core i5-8400 CPU@ 2.80GHz and 8GB of RAM. In other words, the control performances of the proposed fairness-aware data-driven MPC were investigated based on numerical study. Therefore, future work could focus on investigating the applicability of the proposed MPC in real buildings based on field experiments.

Existing studies that developed software and hardware configurations for the practical application of data-driven MPCs are of great reference significance. For instance, Wang et al. [56] proposed a low-cost and feasible hardware configuration to test MPC in a real residential building. In this hardware configuration, the smart air conditioner socket acts as a communication gateway between the air conditioner in each room and the MPC component in a Python environment on a cloud. The infrared bi-directional communication is achieved by an HTTP RESTful application programming interface (API). Bird et al. [57] presented a cloud-based hardware/software solution to realize the implementation of MPC in real buildings. The bi-directional connection between MPC on the cloud platform and the existing building management system (BMS) is achieved through a remote gateway and a message broker.

5.4. ***Others***

To simplify the investigated problem, this study collected the training dataset by simulating the TRNSYS model with a 1-hour time interval. As a result, the energy consumption per period (duration of every 6 hours) became multi-class labels, see Fig. 26(a). Therefore, energy predicting was a multi-class classification problem, and pre-processing methods were applied to improve the predictive performance of minority conditions. However, pre-processing methods could not achieve user-defined quantitively fairness improvement. Besides, energy demand is usually continuous data, which means predicting energy demand would require a regression model. Furthermore, when the simulation time interval is reduced to 5-min, the classes of energy consumption per period would be increased and could be considered as continuous numbers (see Fig. 26 (b)). Therefore, integrating in-processing fairness improvement methods into MPCs to achieve a user-defined trade-off between predictive fairness and accuracy could be an interesting future work.



(a) 1-hour time interval (b) 5-min time interval

Fig. 26: Energy consumption per period simulated by TRNSYS with (a) 1-hour time interval and (b) 5-min time interval

Besides, one potential way to improve the control performance of MPCs is to improve the predictive performance of predictors. For instance, even if the constraints are added to the objective function of MPC to require the indoor air temperature to be higher than 21°C, the actual indoor air temperature when applying the optimized set-point could still be lower than the comfort bound. This is mainly because the predictive recall of air temperature category models is not high enough, and the predictor would think the air temperature could be heated to be higher than 21°C when it is not the case. Therefore, improving the predictive accuracy of predictors integrated into MPCs is worth to be studied.

Furthermore, this study assumed that the duration of peak periods is the same as off-peak periods, so that a data-driven model could be able to predict building states (e.g., energy consumption or indoor air temperature) for each period. However, in reality, the revolution of peak periods and off-peak periods could be different. Therefore, developing specific data-driven models for distinct periods could be considered. Moreover, decreasing the time interval of prediction to 1-hour is also a potential solution.

6. Conclusion

In this study, to ensure that the optimal signals of MPCs in buildings are obtained based on accurate and fair prediction, fairness-aware DDBMs were integrated into MPCs to optimize the control strategy for building devices with the aim of cost-saving or peak shifting. One case study was designed to implement pre-processing methods to process the training dataset for multi-class energy prediction models and two-class air temperature classification models. Then, these fairness-aware models were integrated into MPCs to get the next day’s optimal hourly set-point temperature for EHFs, which is an active thermal energy storage system in a bungalow building. Different objective functions were investigated to achieve cost-saving or peak shifting while preserving thermal comfort. Conclusions of this case study include:

- Using RS or RPS to balance the training dataset for DDBMs would decrease the overall predictive accuracy and the accuracy for majority conditions. However, they could increase the predictive accuracy of minority classes. These fairness improvement methods could let DDBMs provide a fair prediction for minority conditions and majority conditions by deceasing the accuracy difference between these conditions.

- Integrating fairness-aware DDBMs into MPCs would not show a negative effect on peak shifting or cost saving, but thermal comfort.

- Selecting different objective functions would highly affect the control performance of MPCs, such as peak shifting, cost-saving, thermal comfort, etc. Maximizing the peak shifting ability was the most powerful objective function for improving heating cost-saving, peak shifting, and thermal comfort.

In conclusion, fairness-aware data-driven-based MPCs did not show a significant negative impact on the control performance, compared to the traditional MPC. In other words, it is feasible to achieve fairness without compromising control performance. Although pre-processing fairness improvement methods have been implemented in MPCs, the applicability of in-processing fairness improvement methods to fairness-aware data-driven-based MPCs was still not studied. Besides, improving the predictive accuracy of fairness-aware DDBMs would be vital to ensure the control performance of MPCs.

Acknowledgments

The authors would like to express their gratitude to Concordia University for the support through the Concordia Research Chair in Energy & Environment.

Abbreviations

ANN	Artificial Neural Network
ARX	AutoRegressive with Exogenous Inputs
BAS	Building Automation System
BMS	Building Management System
CV(RMSE)	Coefficient of Variation of the Root Mean Square Error
DDBM	Data-Driven Building Model
DNN	Deep Neural Network
DE	Differential Evolution
DR	Demand Response
DT	Decision Tree
EHF	Electrically Heated Floor
mbCRT	Model Based Control with Regression Tree
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MBE	Mean Bias Error
MPC	Model Predictive Controller
RMSE	Root Mean Squared Error
RF	Random Forest
RS	Random Sampling
RPS	Reversed Preferential Sampling
SLPC	Self-Learning Predictive Control
SPS	Sequential Preferential Sampling

SS	Sequential Sampling
SVM	Support Vector Machine
SVM	Support Vector Regression
TES	Thermal Energy Storage
ToU	Time-of-Use

Nomenclature

D	Unprotected attributes
$\boldsymbol{Price_i}$	Electricity price during i-th period, CAD/kWh
$\widehat{\boldsymbol{Q_i}}$	Predicted energy consumption during i-th period, kWh
S	Protected feature
$\mathbf{T}_{\boldsymbol{ambient},\boldsymbol{i}}$	A list of hourly outdoor ambient temperature at i-th period, °C
$\widehat{\mathbf{T}_{\boldsymbol{min},\boldsymbol{i}}}$	The binary label that illustrates if the predicted minimum indoor air temperature during i-th period is lower than 21°C
$\mathbf{T}_{\boldsymbol{set},\boldsymbol{i}}$	A list of hourly indoor air set-point temperature at i-th period, °C
$X_{candidate}$	Original training dataset
$\lvert \boldsymbol{X_{candidate}} \rvert$	The number of data points in the original training dataset
$X_{designed}$	Designed training dataset processed by pre-processing methods
$\lvert \boldsymbol{X_{designed}} \rvert$	The number of data points in the designed training dataset
Y	Output label of a data point

References:

[1] Li B, Cheng F, Cai H, Zhang X, Cai W. A semi-supervised approach to fault detection and diagnosis for building HVAC systems based on the modified generative adversarial network. Energy Build 2021;246:111044. https://doi.org/10.1016/j.enbuild.2021.111044.

[2] Sun Y, Joybari MM, Panchabikesan K, Moreau A, Robichaud M, Haghighat F. Heating demand and indoor air temperature prediction in a residential building using physical and statistical models: a comparative study. IOP Conf Ser Mater Sci Eng 2019;609:072022. https://doi.org/10.1088/1757-899X/609/7/072022.

[3] Wei W, Ramalho O, Malingre L, Sivanantham S, Little JC, Mandin C. Machine learning and statistical models for predicting indoor air quality. Indoor Air 2019;29:704–26. https://doi.org/10.1111/ina.12580.

[4] Adams S, Greenspan S, Velez-Rojas M, Mankovski S, Beling PA. Data-driven simulation for energy consumption estimation in a smart home. Environ Syst Decis 2019:1–14.

[5] Ahmad AS, Hassan MY, Abdullah MP, Rahman HA, Hussin F, Abdullah H, et al. A review on applications of ANN and SVM for building electrical energy consumption forecasting. Renew Sustain Energy Rev 2014;33:102–9. https://doi.org/10.1016/j.rser.2014.01.069.
[6] Wang Y, Li W, Zhang Z, Shi J, Chen J. Performance evaluation and prediction for electric vehicle heat pump using machine learning method. Appl Therm Eng 2019:113901.
[7] Mirnaghi MS, Haghighat F. Fault detection and diagnosis of large-scale HVAC systems in buildings using data-driven methods: A comprehensive review. Energy Build 2020;229:110492. https://doi.org/10.1016/j.enbuild.2020.110492.
[8] Huchuk B, Sanner S, O'Brien W. Development and evaluation of data-driven controls for residential smart thermostats. Energy Build 2021;249:111201.
[9] Behl M, Smarra F, Mangharam R. DR-Advisor: A data-driven demand response recommender system. Appl Energy 2016;170:30–46.
[10] Khosravi M, Huber B, Decoussemaeker A, Heer P, Smith RS. Model Predictive Control in Buildings with Thermal and Visual Comfort Constraints. Energy Build 2023:113831. https://doi.org/10.1016/j.enbuild.2023.113831.
[11] Stoffel P, Berktold M, Müller D. Real-life Data-driven Model Predictive Control for Building Energy Systems Comparing Different Machine Learning Models. Energy Build 2024:113895.
[12] Zhang Z, Zhou Y, Xin X, Qian J, Liu Y. A day-ahead operation regulation method for solar water heating based on model predictive control. Energy Build 2023;301:113715.
[13] Chen S, Ding P, Zhou G, Zhou X, Li J, Wang L (Leon), et al. A novel machine learning-based model predictive control framework for improving the energy efficiency of air-conditioning systems. Energy Build 2023;294:113258. https://doi.org/10.1016/j.enbuild.2023.113258.
[14] Aruta G, Ascione F, Bianco N, Mauro GM, Vanoli GP. Optimizing heating operation via GA-and ANN-based model predictive control: Concept for a real nearly-zero energy building. Energy Build 2023;292:113139.
[15] Ra SJ, Kim J-H, Park CS. Real-time model predictive cooling control for an HVAC system in a factory building. Energy Build 2023;285:112860. https://doi.org/10.1016/j.enbuild.2023.112860.
[16] Maddalena ET, Müller SA, dos Santos RM, Salzmann C, Jones CN. Experimental data-driven model predictive control of a hospital HVAC system during regular use. Energy Build 2022;271:112316. https://doi.org/10.1016/j.enbuild.2022.112316.
[17] Lee H, Heo Y. Simplified data-driven models for model predictive control of residential buildings. Energy Build 2022;265:112067.
[18] Zhang W, Wu W, Norford L, Li N, Malkawi A. Model predictive control of short-term winter natural ventilation in a smart building using machine learning algorithms. J Build Eng 2023;73:106602.
[19] Bi G, Liu J, Gao G, Zhao L. Near-optimal adaptive predictive control model study for roller shades in office spaces. J Build Eng 2023;68:105998.
[20] Li Y, Tong Z. Model predictive control strategy using encoder-decoder recurrent neural networks for smart control of thermal environment. J Build Eng 2021;42:103017.
[21] Mtibaa F, Nguyen K-K, Dermardiros V, Cheriet M. Context-aware model predictive control framework for multi-zone buildings. J Build Eng 2021;42:102340.
[22] Širokỳ J, Oldewurtel F, Cigler J, Prívara S. Experimental analysis of model predictive control for an energy efficient building heating system. Appl Energy 2011;88:3079–87.
[23] Zhang L, Wen J. A systematic feature selection procedure for short-term data-driven building energy forecasting model development. Energy Build 2019;183:428–42. https://doi.org/10.1016/j.enbuild.2018.11.010.
[24] Li Y, Tong Z, Tong S, Westerdahl D. A data-driven interval forecasting model for building energy prediction using attention-based LSTM and fuzzy information granulation. Sustain Cities Soc 2022;76:103481. https://doi.org/10.1016/j.scs.2021.103481.

[25] Fairness in AI: Impact and Opportunities n.d. https://aiasiapacific.org/research/fairness-in-ai-impact-and-opportunities/ (accessed January 17, 2024).
[26] Memarian B, Doleck T. Fairness, Accountability, Transparency, and Ethics (FATE) in Artificial Intelligence (AI) and higher education: A systematic review. Comput Educ Artif Intell 2023;5:100152. https://doi.org/10.1016/j.caeai.2023.100152.
[27] Sun Y, Haghighat F, Fung BC. Trade-off Between Accuracy and Fairness of Data-driven Building and Indoor Environment Models: A Comparative Study of Pre-processing Methods. Energy 2021:122273. https://doi.org/10.1016/j.energy.2021.122273.
[28] Kamiran F, Calders T. Data preprocessing techniques for classification without discrimination. Knowl Inf Syst 2012;33:1–33. https://doi.org/10.1007/s10115-011-0463-8.
[29] Feldman M, Friedler S, Moeller J, Scheidegger C, Venkatasubramanian S. Certifying and removing disparate impact. ArXiv14123756 Cs Stat 2015.
[30] Varona D, Suárez JL. Discrimination, bias, fairness, and trustworthy AI. Appl Sci 2022;12:5826.
[31] Sun Y, Fung BCM, Haghighat F. The generalizability of pre-processing techniques on the accuracy and fairness of data-driven building models: A case study. Energy Build 2022;268:112204. https://doi.org/10.1016/j.enbuild.2022.112204.
[32] Sun Y, Fung BC, Haghighat F. In-Processing fairness improvement methods for regression Data-Driven building Models: Achieving uniform energy prediction. Energy Build 2022;277:112565.
[33] Trong MD, Yang Y. An investigation on fairness perception for grid tariff models: Evidence from Denmark. Electr J 2023;36:107240. https://doi.org/10.1016/j.tej.2023.107240.
[34] Wang LL, Chen JJ, Peng K, Zhao YL, Zhang XH. Reward fairness-based optimal distributed real-time pricing to enable supply–demand matching. Neurocomputing 2021;427:1–12. https://doi.org/10.1016/j.neucom.2020.11.034.
[35] Ren H, Wu Q, Li Q, Yang Y. Optimal design and management of distributed energy network considering both efficiency and fairness. Energy 2020;213:118813. https://doi.org/10.1016/j.energy.2020.118813.
[36] Danner D, de Meer H. Quality of service and fairness for electric vehicle charging as a service. Energy Inform 2021;4:16. https://doi.org/10.1186/s42162-021-00175-3.
[37] Jacubowicz Y, Raz D, Beck Y. Fairness algorithm for emergency demand response operation in distribution networks. Int J Electr Power Energy Syst 2022;138:107871. https://doi.org/10.1016/j.ijepes.2021.107871.
[38] Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A Survey on Bias and Fairness in Machine Learning. ACM Comput Surv 2021;54:115:1-115:35. https://doi.org/10.1145/3457607.
[39] Pessach D, Shmueli E. A Review on Fairness in Machine Learning. ACM Comput Surv 2022;55:51:1-51:44. https://doi.org/10.1145/3494672.
[40] Buildings – Topics. IEA n.d. https://www.iea.org/topics/buildings (accessed November 28, 2022).
[41] Evolution of households energy consumption patterns across the EU 2021. https://www.enerdata.net/publications/executive-briefing/households-energy-efficiency.html (accessed November 28, 2022).
[42] U.S. Energy Information Administration - EIA - Independent Statistics and Analysis n.d. https://www.eia.gov/totalenergy/data/browser/index.php?tbl=T02.01A#/?f=M&start=200001 (accessed November 28, 2022).
[43] Rate Flex D | Hydro-Québec n.d. http://www.hydroquebec.com/residentiel/espace-clients/tarifs/tarif-flex-d.html (accessed March 13, 2020).
[44] Zhu L, Hurt R, Correia D, Boehm R. Detailed energy saving performance analyses on thermal mass walls demonstrated in a zero energy house. Energy Build 2009;41:303–10. https://doi.org/10.1016/j.enbuild.2008.10.003.

[45] Olsthoorn D, Haghighat F, Moreau A, Joybari MM, Robichaud M. Integration of electrically activated concrete slab for peak shifting in a light-weight residential building—Determining key parameters. J Energy Storage 2019;23:329–43. https://doi.org/10.1016/j.est.2019.03.023.
[46] Tarragona J, de Gracia A, Cabeza LF. Bibliometric analysis of smart control applications in thermal energy storage systems. A model predictive control approach. J Energy Storage 2020;32:101704. https://doi.org/10.1016/j.est.2020.101704.
[47] Lee D, Ooka R, Matsuda Y, Ikeda S, Choi W. Experimental analysis of artificial intelligence-based model predictive control for thermal energy storage under different cooling load conditions. Sustain Cities Soc 2022;79:103700. https://doi.org/10.1016/j.scs.2022.103700.
[48] Tang H, Yu J, Geng Y, Liu X, Lin B. Optimization of operational strategy for ice thermal energy storage in a district cooling system based on model predictive control. J Energy Storage 2023;62:106872. https://doi.org/10.1016/j.est.2023.106872.
[49] Storn R, Price K. Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. J Glob Optim 1997;11:341–59. https://doi.org/10.1023/A:1008202821328.
[50] Aongya S. Contrôle du Chauffage Pour la Gestion de la Demande Résidentielle—Rapport Technique sur la Création d'un Modèle Résidentiel Fonctionnel. Hydro-Quebec Shawinigan QC Can 2010.
[51] Thieblemont H, Haghighat F, Moreau A. Thermal Energy Storage for Building Load Management: Application to Electrically Heated Floor. Appl Sci 2016;6:194. https://doi.org/10.3390/app6070194.
[52] Sun Y, Panchabikesan K, Joybari MM, Olsthoorn D, Moreau A, Robichaud M, et al. Enhancement in peak shifting and shaving potential of electrically heated floor residential buildings using heat extraction system. J Energy Storage 2018;18:435–46.
[53] Thieblemont H, Haghighat F, Moreau A, Lacroix G. Control of electrically heated floor for building load management: A simplified self-learning predictive control approach. Energy Build 2018;172:442–58.
[54] Noble WS. What is a support vector machine? Nat Biotechnol 2006;24:1565–7. https://doi.org/10.1038/nbt1206-1565.
[55] Montréal V de. Heating dwelling units n.d. https://montreal.ca/en/topics/heating-dwelling-units (accessed June 17, 2022).
[56] Wang D, Chen Y, Wang W, Gao C, Wang Z. Field test of Model Predictive Control in residential buildings for utility cost savings. Energy Build 2023;288:113026. https://doi.org/10.1016/j.enbuild.2023.113026.
[57] Bird M, Daveau C, O'Dwyer E, Acha S, Shah N. Real-world implementation and cost of a cloud-based MPC retrofit for HVAC control systems in commercial buildings. Energy Build 2022;270:112269. https://doi.org/10.1016/j.enbuild.2022.112269.