https://dx.doi.org/10.1016/j.enbuild.2023.113286

Energy-oriented Control Retrofit for Existing HVAC system adopting data-driven MPC - Methodology, Implementation and Field test

Bao Yue[a,b,*], Bing Su[a], Fu Xiao[c,d], Anbang Li[a], Kehua Li[a], Shen Li[a], Rui Yan[a], Qiuzhuang Lian[a], Ao Li[c], Yuanyang Li[a], Xing Fang[a], Xingang Liang[b]

[a]*Midea Building Technologies Division, Midea Group, Guangdong 528311, China*
[b]*School of Aerospace Engineering, Tsinghua University, Beijing 100093, China*
[c]*Department of Building Environment and Energy Engineering, The Hong Kong Polytechnic University, Hong Kong, China*
[d]*Research Institute for Smart Energy, The Hong Kong Polytechnic University, Hong Kong, China*

**Bao Yue: yuebao@midea.com*

***Keywords**: model predictive control, air conditioning, control retrofit, field test, energy conservation, hybrid building thermal model*

Nomenclature

Symbols

ACHP	*air-cooled heat pump*
AHU	*air handling unit*
a, b, c, d, g, h	*coefficients*
BACnet	*building automation and control network*
BAS	*building automation system*
BMS	*building management system*
C	*heat capacity, kJ/K*
Cap	*cooling capacity, kW*
CLG	*cooling power, kW*
COP	*coefficient of performance*
dP	*differential pressure, kPa*
F	*water flow rate, kg/s*
f	*frequency, Hz*
HVAC	*heating, ventilation, air conditioning and cooling*

IoT	*Internet of Things*
L	*period*
LDAP	*Lightweight Directory Access Protocol*
MAD	*Mean Absolute Difference*
MAPD	*Mean Absolute Percentage Difference*
MPC	*model predictive control*
MQTT	*Message Queuing Telemetry Transport*
N	*number*
P	*power, kW*
Q	*cooling/heating energy, kJ*
R	*thermal resistance, K/kW*
RBAC	*Role Based Access Control*
RBC	*rule-based control*
T	*temperature, ℃*
TLS	*Transport Layer Security*
UDF	*User Defined Function*
Subscripts	
AE	*air exchange*
agg	*aggregated*
avg	*average*
db	*dry bulb*
enve	*envelope*
hp	*heat pump*
IA	*indoor air*
IHG	*internal heat gain*
lb	*lower bound*
min	*minimum*
OA	*outdoor air*
t	*time*

ub	*upper bound*
var_rate	*variation rate*
Greek symbols	
α, β, γ	*parameters*
ε	*threshold value*

Abstract

Energy conservation and carbon reduction of existing buildings have been receiving increasing attention with the proposed carbon neutrality goals. As presented in many studies, adopting model predictive control (MPC) for building HVAC system control is an effective means to realize building energy conservation. However, existing studies rarely addressed the critical challenges for practical applications and integration of MPC in existing Building Automation Systems (BASs). This study proposes a control retrofit approach, which is lightweight and replicable, to realize the improvement of energy efficiency of the existing building HVAC systems through integrating data-driven MPC into existing BASs. The critical challenges in the practical applications of MPC strategies, including MPC strategy development, large-scale data processing, strategy deployment and integration with existing BASs are addressed. The proposed approach was applied to the existing HVAC system of an actual airport terminal and the evaluation results indicate that the control retrofit approach is effective in achieving energy efficiency and thermal comfort improvement. The system daily energy consumption was reduced by 24.5% in average and the percentage of the discomfort time was reduced from 70.2% to 5.7% in average after the control retrofit. The proposed control retrofit approach, with the help of the scalability and distributed computing capabilities of the cloud computing platform, is expected to be able to realize lightweight control retrofit of large number of existing building HVAC systems.

1 Introduction

Carbon emissions during building operation accounts for a large proportion of the total carbon emissions in China, which was 21.9% in 2018 [1]. Therefore, reduction of carbon emissions during building operation is one of the key paths to achieve carbon peaking and carbon neutrality goals that proposed by the Chinese government. The reduction of carbon emissions in the operation phase of existing buildings can mainly be achieved by reducing the energy consumption of existing buildings and using more renewable energy. A large proportion of the energy consumption in existing buildings is consumed by HVAC systems, especially for existing public buildings [2, 3]. Therefore, energy conservation of HVAC system is very important for reducing the energy consumption of existing public buildings.

The approaches for energy conservation of HVAC systems in existing public buildings mainly include system hardware retrofit, system commissioning and advanced control. System hardware retrofit refers to the approaches that realize the energy conservation through modifying building envelope (e.g., wall insulation layer [4], and low-e glass [5]), replacing low-efficiency equipment [7], and adding additional equipment in the original system [8-10]. These approaches commonly show considerable performance on energy conservation while they involve a relatively large cost and unneglectable risks for cost recovery [11, 12]. Commissioning is a systematic process to ensure the building facility systems perform as the requirements of design documentation and intent [13]. For HVAC systems, due to the deviation of the actual working conditions from the design working conditions and the performance degradation or even faults of equipment, issues such as overheating, overcooling and/or low energy efficiency may occur in the HVAC systems after a period of operation [14]. In such cases, system commissioning is an effective means to deal with these problems and therefore improves the indoor thermal comfort and energy efficiency of the HVAC systems [15]. However, system commissioning requires high professionalism, and experts who are very familiar with the HVAC systems are always needed to conduct a comprehensive diagnosis. Moreover, it is difficult to completely solve the problem at a time, the system needs to have a re-commissioning after a period to maintain comfort and energy performance [16]. These problems make it hard to scale up the applications of commissioning service in practice.

Adopting advanced control approaches instead of using simple ones, such as rule-based control or reactive control, is another effective means to realize energy conservation of existing building HVAC systems. Through optimizing the set-points of the system control variables proactively, advanced control approaches can achieve the efficient operation of target HVAC systems in the form of supervisory control [17]. Especially for the advanced control approaches with the self-adapting capability, once the control approaches are developed and deployed for the target HVAC systems, the system operation efficiency can be improved and maintained continuously with quite little maintenance fee [18]. These features make it more attractive in the energy conservation of existing building HVAC systems than the system hardware retrofit and system commissioning approaches. Therefore, this study focuses on the control retrofit of existing HVAC systems using advanced control approaches. MPC (model predictive control) is one of the typical advanced control approaches which has been proved to be effective in multiple studies [19]. MPC uses a system model to predict the system state or performance metrics (usually in the form of objective function in optimization) within a certain period in the future (i.e., prediction time horizon, such as 1 hour, 1day). Based on the prediction results, the optimal control variables can be obtained using optimization techniques to minimize the system performance indices (e.g., operating costs, energy consumption, carbon emissions) while ensuring comfort requirements [20]. A large number of studies have been conducted on the development of modelling approaches and optimization algorithms for MPC, and the relative studies have been reviewed in detail by Refs [19, 21].

In recent years, many studies have been conducted to investigate the effectiveness of MPC in building HVAC systems for energy conservation [22]. Yang et al. [23] proposed a MPC strategy for an air-conditioning system with dedicated outdoor air system and conducted an experimental test to investigate the control performance. Compared to the conventional feedback control, the MPC strategy achieved 18% to 20% energy savings. Chen et al. [24] developed an occupant feedback-based model predictive control strategy using a novel dynamic thermal sensation (DTS) model and conducted chamber experiments to evaluate its performance. The test results show that the developed DTS-based MPC performs better than the Predicted Mean Vote-based MPC. Huang et al. [25] investigated the combined impacts of selected time intervals for model discretization and control sampling on the performance of MPC through detailed simulations. The simulation results show that the time interval of model discretization has much greater influence. These studies provide the means and guidelines to design and develop the MPC strategies, and demonstrate their superior performance in HVAC control.

However, there are yet few practical applications of MPC in the optimal control of building HVAC systems, and the effective means to promote the application of MPC in engineering are rarely found in the existing studies. Most of the existing studies focused on validating the feasibility and advantages of MPC in the control of specific HVAC systems but not the general approach for realizing the practical applications of MPC in large scale [26-28]. Only a few studies have discussed the critical challenges in scaling up the practical applications of MPC in existing building HVAC systems, including the acquirement of required operation data, design and development of MPC, implementation of MPC in existing BMS (Building Management System) or BAS (Building Automation System) [29, 30]. Blum et al. [29] estimated the model development and engineering costs for MPC implementation in a real office building, and demonstrated the feasibility of developing MPC using the open-source software developed in [31]. However, the proposed MPC strategy was designed for the HVAC system of the target office specifically, and its generality remained unclear. Bird et al. [30] presented a cloud-based monitoring and control platform in the case study of a food-retail building in the UK and introduced the design, installation and cost of implementing MPC for its HVAC system. The simulation results showed that the proposed MPC control outperforms the existing rule-based control both in energy cost and carbon emission. These studies introduced the workflow for developing and implementing MPC in practical applications, while several critical challenges remained unsolved, including the development of reliable and efficient system model, the large-scale data processing in near real-time, and the compatibility of MPC strategy with existing control logic. And these challenges restrict the wide and flexible application of MPC in existing building HVAC systems for energy conservation.

To fulfill the research gaps in practical deployment of MPC in existing HVAC systems, this study proposes a general control retrofit approach which can realize the energy conservation of existing building HVAC systems adopting MPC. A general and replicable MPC strategy which is effective for the energy conservation of existing HVAC systems with different configurations is proposed. A cloud-edge collaboration architecture is proposed to realize the operation data access and integration of the MPC strategies with existing BASs in a lightweight and scalable way through standard communication protocols. The MPC strategies are deployed in the cloud to optimize the system operation according to the acquired operation data through sending supervisory control settings to the existing BASs. The implementation of the proposed approach in a real-life airport terminal is presented in detail and the field test results are analyzed.

The contributions of this work are summarized as follows.

- A general and replicable MPC strategy for energy conservation of existing building HVAC systems is proposed and validated in practical application. This strategy can be applied in most of the building HVAC systems of different configurations since it is designed to optimize the system operation at the macro level rather than to optimize the system/subsystem of specific configuration. According to the predicted indoor temperature trend, the proposed MPC proactively generates the optimal system cooling outputs which are achieved through adjusting the chilled water supply temperature. In this manner, the system cooling outputs can be minimized while maintaining indoor thermal comfort. Based on the optimal system cooling outputs, the operating number of chillers is optimized to minimize the energy consumption of chillers.
- A general approach for implementing the MPC strategy in existing HVAC systems is proposed and the critical problems in practical applications are addressed, including near real-time (within 1 minute) large-scale data proceesing and the

compatibility and integration of MPC with the existing BAS. Near real-time data is necessary for operation of MPC strategy and it is a challenge work to deal with the huge amount of operation data in near real-time when many HVAC systems are integrated in the cloud platform. An effective solution is presented in this study for this critical problem in practice. Moreover, it is a difficult task to deploy MPC strategies in the existing BASs especially when the existing control logic needs to be revised, since their programs are customized without uniform programming standards. Using the proposed approach, the MPC strategy can be integrated into the existing BASs noninvasively, i.e., the control logic of the existing BASs doesn't need to be revised. The proposed approach is of great significance to guide the application and promotion of data-driven MPC in actual building energy-saving retrofitting projects.

- A novel hybrid model, which consists of a gray-box building thermal model and a data-driven residual prediction model, is proposed for the prediction of indoor temperature variation in practical applications. The gray-box building thermal model is developed based domain knowledge on thermodynamics which brings high generality since it doesn't rely on detail information of the target buildings. The heat gains that cannot be directly measured, such as indoor heat gains of solar irradiance and occupancy, are innovatively regarded as a periodical function and represented as Fourier expansion in the gray-box model. The accuracy of the gray-box model is further enhanced using a data-driven model to compensate its prediction error which may be caused by the simplification of model structure.

The rest of this paper is organized as follows. Section 2 introduces the general MPC strategy, the cloud-edge collaboration architecture and the standard workflow of the proposed control retrofit approach. Section 3 presents the implementation of the proposed control retrofit approach in a real-life airport terminal, including the target system description and the MPC strategy development. In Section 4, the performance of the developed MPC strategy is evaluated according to the field test results. The conclusions of this study are summarized in Section 5.

2 Methodology

The proposed control retrofit approach is a general and replicable approach for energy conservation of existing building HVAC systems through integrating MPC strategies into their BASs. A general MPC strategy for optimal control of target HVAC system, a cloud-edge collaboration architecture for the integration of MPC strategy with existing BAS, and the standard workflow for implementing the control retrofit approach are presented.

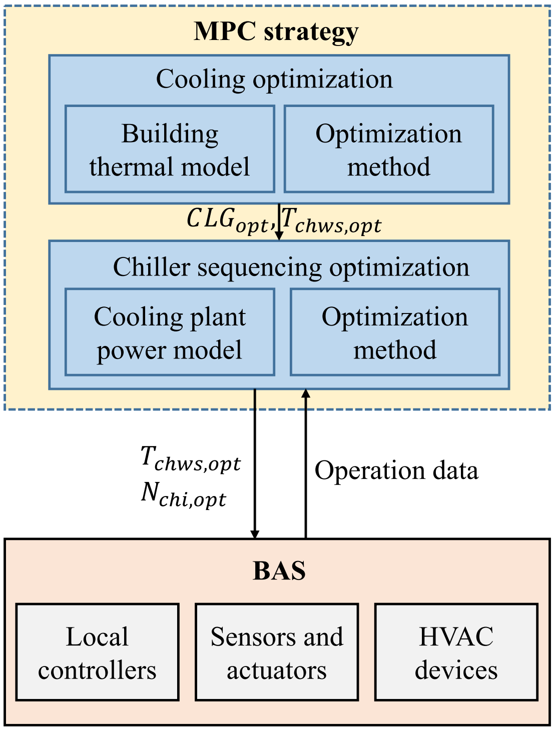
2.1 General MPC strategy

The schematic and optimization process of the proposed general MPC strategy for existing HVAC systems is presented in Fig. 1. The proposed MPC strategy is a supervisory control strategy which optimizes the set points of chilled water supply temperature and the operating number of the chillers. It contains two parts, i.e., cooling optimization and chiller sequencing optimization. The main idea of this MPC strategy is to minimize the system power consumption by both minimizing the system cooling supply and maximizing the system operation efficiency.

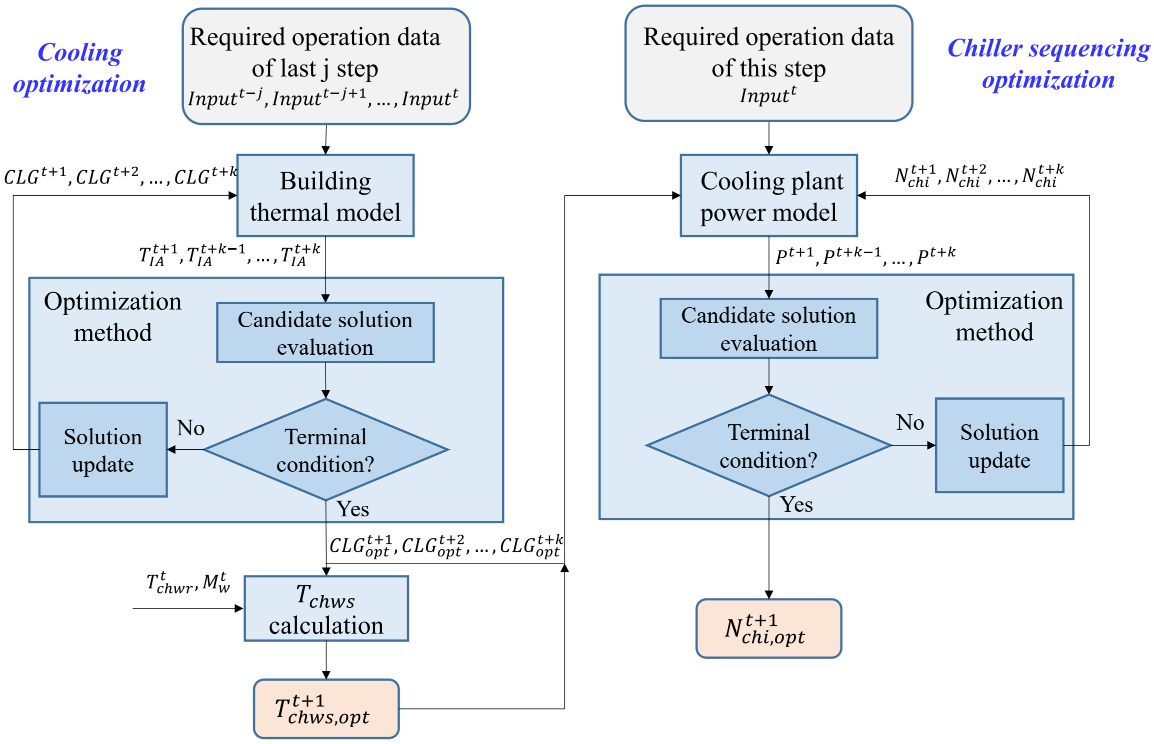
The cooling optimization part is responsible for minimizing the system cooling output while maintaining the indoor temperature in the predefined range. The building thermal model is responsible for predicting the building indoor average temperature trend in the control time horizon (e.g., future 30 minutes). The optimal cooling output is achieved by the optimization method according to the prediction results of the building thermal model. With the obtained optimal cooling output, the optimal chilled water supply temperature can be approximately calculated according to the current chilled water return temperature and chilled water flow rate.

The chiller sequencing optimization part is to seek for the optimal operation number of chillers which can minimize the system power consumption while providing the required cooling capacity, i.e., the optimal cooling output obtained by cooling optimization. The cooling plant power model is responsible for predicting the power consumption of the cooling plant when different sequencing strategies are adopted.

The proposed MPC strategy is applicable in a large amount of HVAC systems with different configurations since the means of realizing the system energy conservation, i.e., cooling optimization and sequencing optimization, is not designed according to the HVAC system with specific configurations. Moreover, the proposed MPC strategy optimizes the operation of target HVAC system through supervisory control and therefore will not affect the existing local control logics of the existing BAS. The proposed MPC strategy is applied and validated in a real-life airport terminal and the details are presented in Section 3.2 and Section 4.



(a) Schematic of the MPC strategy



(b) Optimization process of the MPC strategy

Fig. 1. The proposed general MPC strategy

2.2 Cloud-edge collaboration architecture

2.2.1 Overall architecture

Fig. 2 presents the proposed cloud-edge collaboration architecture of the control retrofit approach, which consists of an IoT (Internet of Things) gateway and a cloud platform. The architecture is able to realize the lightweight control retrofit of existing building HVAC systems since only an IoT gateway is needed to be integrated into the existing BASs and no revision on the original control logic of the existing BASs is needed. The cloud platform is responsible for the real-time data processing and storage as well as the deployment and operation of the MPC strategy. The IoT gateway is responsible for the bidirectional data transmission between the existing BAS and the cloud platform. The system raw operation data generated in the BAS are acquired by the IoT gateway and then uploaded to the cloud platform for data processing and storage. The processed data is used by the deployed MPC strategy for system operation optimization. Once the optimized settings are generated by the MPC strategy, they are sent to the IoT gateway and then delivered to the existing BAS for control execution.

The bidirectional data transmission between BAS and the IoT gateway, i.e., operation data upload and control settings delivery, is realized through BACnet (Building Automation and Control Network) communication protocol, a commonly used network protocol in BASs for interoperation between different devices. BACnet provides a convenient way to realize the information upload and control settings execution. For information upload, all the available data points in the BAS control network can be obtained directly through BACnet scanner and therefore the integration of the IoT gateway can be realized easily. For control settings execution, the settings from MPC strategy can be executed directly without modifying the original program of the existing BASs through priority-based overriding mechanism of BACnet, i.e., the original setting parameters in the existing BASs can be directly overridden by the settings from the MPC strategy which has an equal or higher priority.

The communication protocol used for the information exchange between the IoT gateway and the cloud platform is MQTT (Message Queuing Telemetry Transport). MQTT is a light protocol widely used in IoT systems since it can ensure smooth and reliable data transfer among multiple devices with low bandwidth. The TLS (Transport Layer Security) encryption mechanism is added in the MQTT protocol to protect the system against the vulnerabilities.

For the deployment of MPC strategy, Docker is adopted to package up the code of each MPC strategy and its dependencies as a portable and self-sufficient container operating on the cloud platform [32]. In this manner, the MPC strategies can be tailored as different micro services on the cloud platform to provide service for different building HVAC systems.

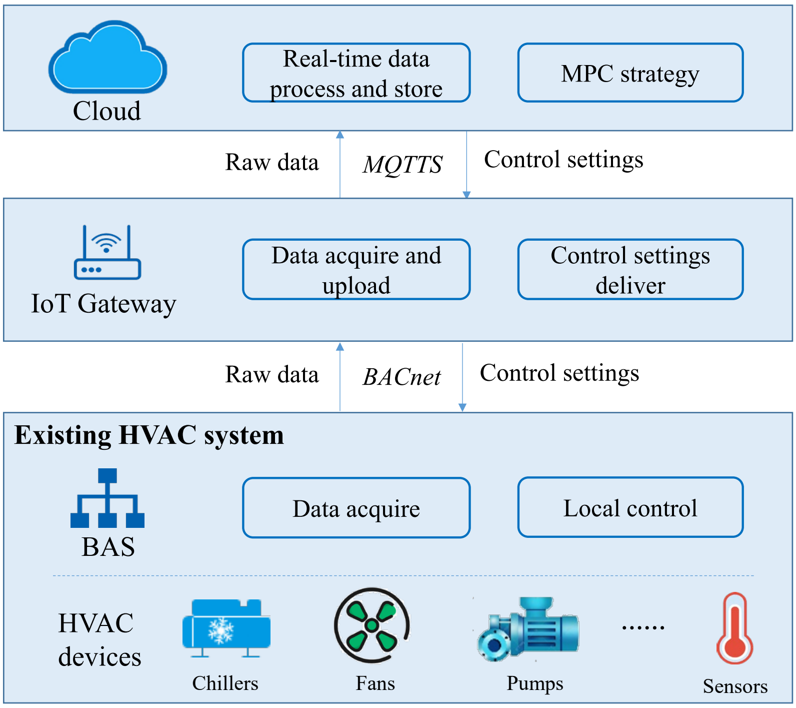


Fig. 2. Cloud-edge collaboration architecture of the proposed control retrofit approach

2.2.2 Solution for large-scale data processing and store in near real-time

Considering the large-scale applications of this approach for the control retrofit of existing HVAC systems, the near real-time processing and store of large-scale data in the cloud platform need to be addressed. The near real-time data processing and storing is very important for the MPC strategy since the optimization needs to be conducted based on the current and historical state of the system. Considering the situation that 100 building HVAC systems are integrated in the cloud platform and the number of data points of one building HVAC system is 1,000, the amount of data needed to be processed in one minute will be 100,000. As a result, there will be 144 million records per day and more than 52 billion records in total in one year needed to be stored. This brings critical challenges to the cloud platform, which including massive data processing (such as missing data handling, data noise and outliers handling) in near real-time, low-cost and scalable data storage, fast data query and retrieval, data backup, and data safety.

An effective solution is proposed for the near real-time large-scale data processing in the cloud platform. It is based on the recent emerging Lakehouse architecture - an open data architecture which combines the flexible storage of unstructured data from data lakes and the management features and tools from data warehouses [33]. The ability of end-to-end streaming and decoupling of storage and computing makes it preferred architecture for this scenario. The architecture of the proposed solution and process of the large-scale data processing are shown in Fig. 3. Massive operation data of different building HVAC systems are uploaded to the MQTT brokers and sent to the Kafka [34], which is an open source platform supporting steaming analytics, through an adapter developed for data format conversion. Based on PySpark Streaming [35], the missing values and outliers in the massive data can be handled using the UDFs (User Defined Function) written in Python in near real-time. The UDFs are executed in distributed computing in Spark cluster which makes computing scalable. By decoupling storage from computing, the handled time-series data is then stored into cheap object storage yet with great data analytics performance in Delta Lake format [36]. In order to ensure the performance of the database in query and retrieval, multiple optimization techniques including partition [36], compaction [36] and z-order [37] are utilized. Moreover, with Delta Lake metadata stored in the same location with the time-series data, the backup and archive of the data can be realized by simply using 'clone' command to copy plain object storage files into cheaper storage like tapes. With security in mind, the RBAC (Role Based Access Control) plugin is developed based on the LDAP (Lightweight Directory Access Protocol) to manage the users' access to the database under the multi-tenant mode. Then, the applications, i.e., the MPC strategies for different building HVAC systems can access the corresponding near real-time data securely through PyHive [38].

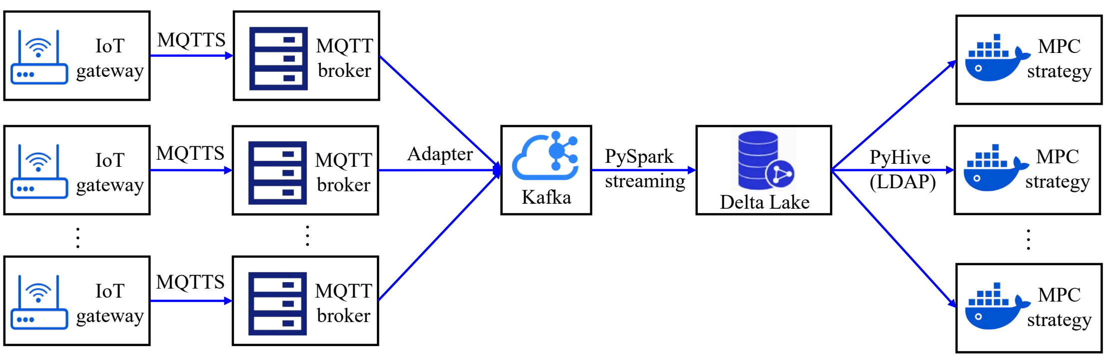


Fig. 3. Architecture of the proposed solution and process of the large-scale data processing

2.3 Standard workflow

Fig. 4 presents the standard workflow for implementation of the proposed control retrofit approach, which contains five key steps, including onsite system investigation, IoT gateway integration, data processing, MPC strategy development, and MPC strategy deployment and trail operation. The first two steps will consume 3-7 days based on the complexity of the target HVAC system and its BAS. The most time-consuming steps are the data processing and MPC strategy development. The main reason is that the system models need to be developed based on the operation data of the target HVAC systems. For the HVAC systems with abundant historical operation data, the MPC strategy can be developed in about two weeks. While for the cases that historical operation data is unavailable, one month operation data is normally needed to be collected for the strategy development and the whole step will cost about 6 weeks. The detail descriptions of these steps are presented as follows.

- Onsite system investigation is conducted to obtain the important information of the target HVAC system, including the system configuration, local control logic and available measurement points.
- IoT gateway integration is conducted to realize the bidirectional data transmission between the cloud platform and the existing BAS. In this manner, the operation data of the HVAC system generated by the BAS can be obtained for further use, such as system performance analysis, development of the system model and optimization of system operation, and the optimal settings generated by the MPC strategy can be sent to the existing BASs for control execution.
- Data processing is conducted to deal with the two main problems in the obtained raw operation data, which are non-standard name of data points, and data noise and outliers. Firstly, the data organization is conducted to map the original data to standard tagging scheme. The organized data is then processed to deal with the noise and outliers. In real applications, the noise and outliers can be generated by various problems, such as improper installation of sensors, sensor bias and faults,

so the algorithms for dealing with the data noise and outliers needs to be developed accordingly.

- MPC strategy development is conducted to develop MPC strategy for the target HVAC system according to the collected system information and processed data. In this step, the prediction model of the system, the objective function in optimization and the optimization algorithms are developed.
- Trial operation is conducted to test the reliability and effectiveness of the developed MPC strategy in test environment of the cloud platform before real applications. During the trial operation, the real-time data transmission, MPC strategy operation, and control settings delivery and execution will be tested.

Once the whole system is proved to be reliable and robust during the trial operation, the MPC strategy will be encapsulated and deployed on the operation environment to optimize the operation of the target HVAC system continuously.

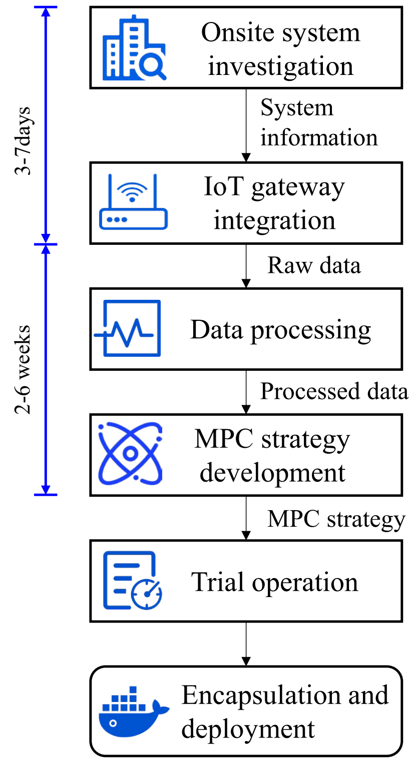


Fig. 4. Standard workflow for implementation of the control retrofit approach

3 Control retrofit implementation

The proposed control retrofit approach is applied in the HVAC system of a real-life building, which is a three-floor airport terminal located in Zhejiang province, China, with about 16,500m^2. The system information is obtained through onsite system investigation and relative materials, including design documents of the building HVAC system and the instructions of its BAS. The system configuration and original control strategy of the existing BAS are presented in Section 3.1.

Since there is no historical operation data available in the existing BAS, the operation data is collected after the IoT gateway being integrated into the existing BAS. Through BACnet scanner, the available data points are obtained and there are totally 900 data points available in the BAS. Although not all data points are needed for development and operation of the MPC strategy, they are collected completely considering the use of system operation monitoring and further analysis. The acquired data is uploaded to the cloud through 4G network in 1-min interval. With the collected data, the MPC strategy for the HVAC system is developed and it is presented in Section 3.2.

3.1 Target system description

3.1.1 HVAC system configuration

The schematic of the target HVAC system is shown in Fig. 5. The system consists of six air-cooled heat pumps (ACHPs), four variable speed chilled water pumps and six air handling units (AHUs). The nominal cooling capacity of each air-cooled heat pump is 602 kW under nominal working condition, i.e., at the chilled water supply temperature of 7 °C, chilled water flow rate of 28.8 L/s and ambient air temperature of 35°C. Among the four chilled water pumps, three are associated with the ACHPs and the other one is stand by. The rated flow rate and power load of the chilled water pumps are 61.1 L/s and 45 kW, respectively. In order to prevent the ACHPs from suffering chilled water shortage, a bypass chilled water pipeline with a manual valve is built. The nominal cooling capacities of the AHUs are 280 kW, 280 kW, 320 kW, 320 kW, 245 kW, 245 kW, respectively and their nominal fan power are 22 kW, 22 kW, 30 kW, 30 kW, 22 kW, and 22 kW, respectively.

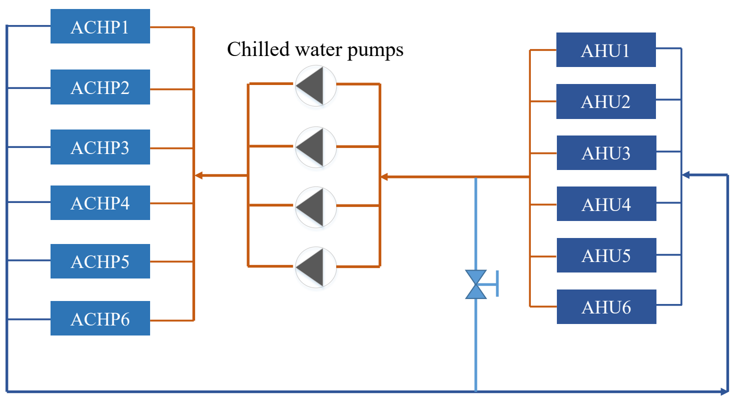


Fig. 5. Schematic of the target HVAC system

3.1.2 Original control strategy of the HVAC system

The existing BAS of the HVAC system has been configured with a rule-based control (RBC) strategy and multiple local control functions to allow the HVAC system to operate properly.

The rule-based control strategy is responsible for the control of chilled water supply temperature (T_{chws}), and it determines the chilled water supply temperature according to outdoor wet-bulb temperature (T_{wb}) as shown in Eq (1). Where α and γ are constant parameters determined according to experience of the engineers to ensure the performance of this strategy.

$$T_{chws} = \alpha T_{wb} + \gamma \tag{1}$$

The local control functions include the sequencing control of ACHPs and pumps, as well as the process control for pump frequency, AHU supply air temperature and AHU fan frequency.

For the sequencing control of ACHPs, an ACHP will be turned on if the set point of chilled water supply temperature cannot be achieved and all the operating ACHPs are operating at high part load ratio for a predefined time, and vice versa. The number of running pumps is determined by the number of running ACHPs, i.e., it is the number of running ACHPs divides two rounded up. The pump frequency is adjusted according to the difference between the set point and the actual value of the pressure difference between the supply side and return side of the main water pipeline. For the process control of AHUs, the supply air temperature is controlled by modulating the opening of water valve according to the difference between the set point and actual value. The fan frequency of an AHU is controlled according to the difference between the set point of indoor temperature and its real value.

3.2 MPC strategy development

The MPC strategy proposed in Section 2.1 is used in this HVAC system. Note that only cooling mode of the HVAC system is considered in this study since there is no available historical operation data of heating season. The proposed MPC strategy contains two parts, which are cooling optimization and ACHP sequencing optimization. Considering the building thermal inertia and system dynamics, the control time horizon of MPC is set as thirty minutes in this study and the optimization is conducted per ten minutes.

3.2.1 Cooling optimization

The optimization problem of cooling optimization is formulated as Eq. (2). The objective is to minimize the cooling output of the next three time steps and the constraints include the lower and upper bound of system cooling output, indoor average temperature, and the variation rate of the cooling outputs at two consecutive time steps. Where CLG_t is the system cooling output at time t, T_{IA} is the average temperature of the indoor air, and subscripts *lb*, *ub* and *var_rate* represent, lower bound, upper bound and variation rate, respectively.

$$Minimize \sum_{t=1}^{3} CLG_t \tag{2}$$

$$\text{subject to:} \begin{cases} CLG_{lb} < CLG_t < CLG_{ub} \\ T_{IA,lb} < T_{IA} < T_{IA,ub} \\ CLG_{var\_rate} \leq CLG_{var\_rate,ub} \end{cases}$$

To solve the optimization problem, a building thermal dynamic model is needed for indoor average temperature prediction. Considering the low generality and modelling efficiency of traditional white-box model and the low extrapolation accuracy of pure data-driven model, a novel hybrid model which consists of a gray-box building thermal model and a data-driven residual prediction model is proposed Compared to traditional white-box models, the proposed model doesn't rely on the detail information of the target building but only general thermodynamic analysis, and is therefore more easily and efficiently applied to different buildings. Meanwhile, owing to the informed physics laws, the proposed hybrid can be more robust to outliers in the training dataset and has better performance in extrapolative prediction than the pure data-driven models.

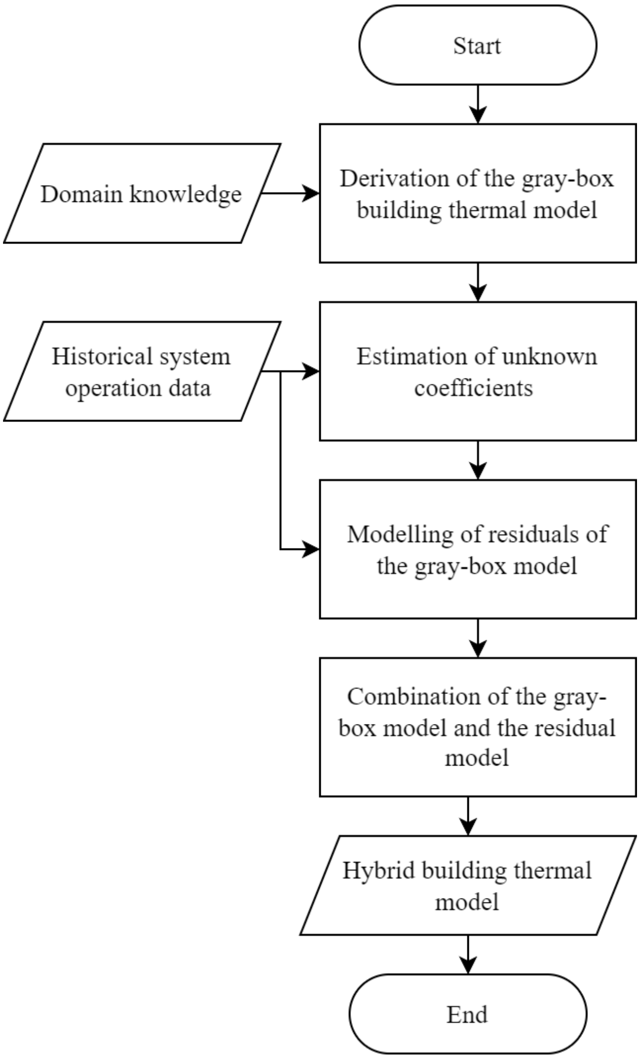


Fig. 6. Diagram of the development of the proposed hybrid building thermal model

The development of the proposed hybrid model consists of four main steps as shown in Fig. 6. Firstly, a gray-box building thermal model for indoor average temperature prediction is derived based on domain knowledge. The unknown coefficients in the derived model are estimated using Least Squares based on the historical HVAC system operation data collected from the case study building. After the parameter estimation, the residual of the gray-box building thermal model is calculated and modeled using a machine learning method (i.e., Cubist [39]). In the last step, the hybrid model for indoor average temperature prediction is obtained by combining the physics-informed building thermal model and the data-driven residual model. The details of each step are presented in the Appendix A.1.

The schematic of the hybrid building thermal model is shown in Fig. 7. For one-step-ahead prediction, the indoor average temperature, the outdoor air temperature, and the cooling supply of the last three steps as well as the current timestamp are required to predict the indoor average temperature of the next step. For multi-step-ahead prediction, the prediction results of the former time steps are recurrently used as the input to predict the indoor average temperature at the latter time steps.

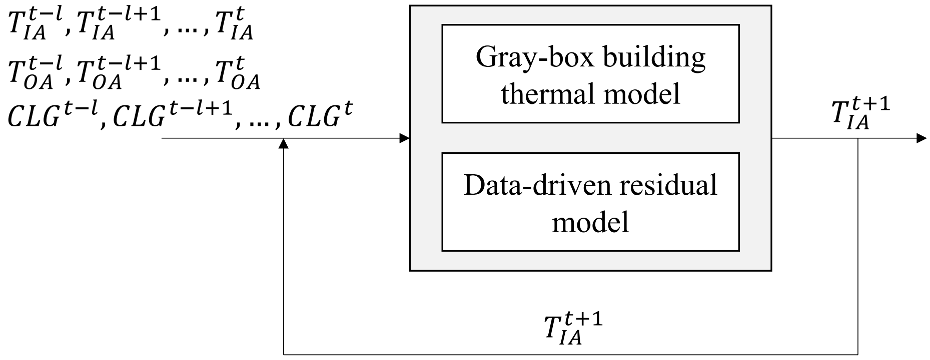


Fig .7. Schematic of the hybrid building thermal model

In order to solve the cooling optimization problem presented as Eq. (2), Branch and Bound, a general optimization method for discrete and combinatorial optimization problems proposed in Operation Research, is adopted in this study [40]. Although the original optimization problem is continuous, it can be transformed into discrete and combinatorial optimization problem without sacrificing optimization performance considering the control and measurement precision of HVAC systems.

3.2.2 ACHP sequencing optimization

The optimization problem of ACHP sequencing optimization is formulated as Eq. (3). The objective is to minimize the power consumption of the cooling plant, which includes that of the ACHPs and pumps, through optimizing their operating numbers. The operating number of the water pumps is associated with that of the ACHPs as shown in Eq. (4). The constraints include the minimum requirement on the number of operating ACHPs and the limitation on the number of ACHPs that can change ON/OFF status.

$$Minimize \sum_{t=1}^{3}\left[P_{hp,t}\left(N_{hp}, T_{chws}, T_{db}, CLG_t\right) + P_{pump,t}\left(N_{hp}, N_{pump}, dP\right)\right] \tag{3}$$

$$N_{pump} = \left\lceil\frac{N_{hp}}{2}\right\rceil \tag{4}$$

$$subject\ to: \begin{cases} N_{hp} \geq N_{min} \\ N_{change} \leq N_{avaliable} \end{cases}$$

Where $P_{hp,t}$ and $P_{pump,t}$ are the power consumption of operating ACHPs and pumps at time t, N_{hp} and N_{pump} are the number of ACHPs and pumps in operation, N_{min} is the minimum number of ACHPs needed to be turned on, $N_{avaliable}$ is the number of ACHPs that can change ON/OFF status, T_{chws} and T_{db} are the chilled water supply temperature and outdoor dry bulb temperature, dP is the differential pressure of the chilled water system.

Simplified models of the ACHPs and pumps are developed to predict the power consumption of the cooling plant. The ACHPs used in this system can modulate their cooling capacities in four steps, i.e., 25%, 50%, 75%, and 100%, denoted as stage 1 to stage 4. The performance of the ACHPs sizably varies from one stage to another and the operating capacity of the ACHPs are controlled by the built-in control logic. Therefore, in order to predict the power consumption of an ACHP, its operation stage needs to be determined first. In this study, a hybrid model which includes a stage prediction module and a power prediction module is developed for the power prediction of one ACHP. Fig. 8 presents the schematic of the hybrid model. Based on the total cooling output (CLG_{tot}), the number of operating ACHPs (N_{hp}), the chilled water supply temperature (T_{chws}), and dry bulb temperature (T_{db}), the operating stages and the cooling outputs of the ACHPs ($CLG_1, \ldots, CLG_k$) are determined. According to the determined operation stages and cooling outputs of the ACHPs as well as the T_{chws} and T_{db}, the power consumption of the operating ACHPs can be obtained. The detail description of the developed model is presented in the Appendix A.2.

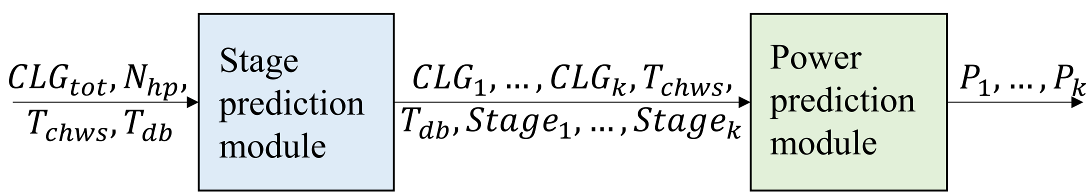


Fig. 8. Schematic of the ACHP power prediction model

In order to solve the sequencing optimization problem presented as Eq. (3), exhaustive search is adopted as the optimization method in this study since the potential solution of the problem is limited.

4 Field test and analysis

4.1 Model validation

Based on the collected operation data through the integrated IoT gateway, the cooling plant power model and the building thermal model were trained and validated. The operation data from 5th July to 7th Aug 2022 (34 days in total) were used for model training and the data from 8th to 11th Aug 2022 (four days in total) were used for testing.

4.1.1 Validation of cooling plant power model

The cooling plant power model, which includes the ACHP and chilled water pump models, is used to predict the plant power consumption. Fig. 9 compares the measured and predicted power consumption of the cooling plant during the four test days. It can be seen that although the total power consumption of the cooling plant changed stepwise in certain periods due to the switching of operation stage of the operating ACHPs, the prediction results match the measured values closely in the whole time period. The prediction error was less than 5% except for one test case as shown in Fig. 10, indicating the effectiveness of the cooling plant model.

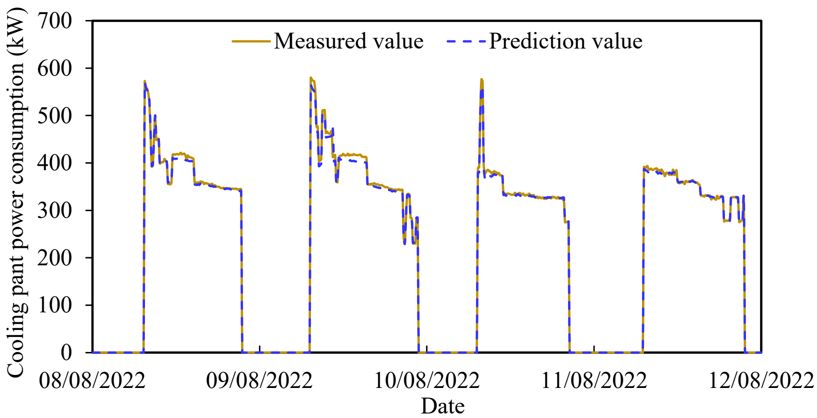


Fig .9. Comparison between the predicted and measured cooling plant power consumption

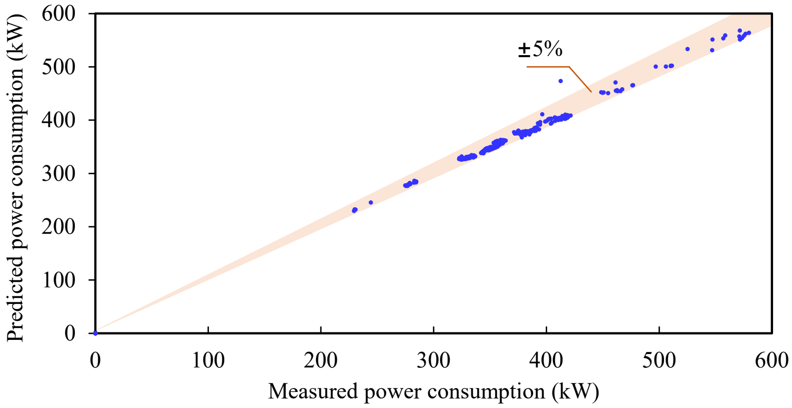


Fig. 10. Prediction error of the cooling plant power model

4.1.2 Validation of building thermal model

The building thermal model is used to predict the indoor average temperature in the next 30 minutes. Fig. 11 compares the measured and predicted indoor average temperature of the air-conditioned area in the studied building during the four testing days. It can be seen that the predicted indoor average temperature highly approximated the measured values during the testing days and the RMSE (Root Mean Square Error) between the measured and predicted indoor average temperature was 0.12 °C. The relative prediction error of the model was less than 5% as shown in Fig. 12, and the absolute error was lower than 0.2°C in most of the cases, indicating the effectiveness of the building thermal model.

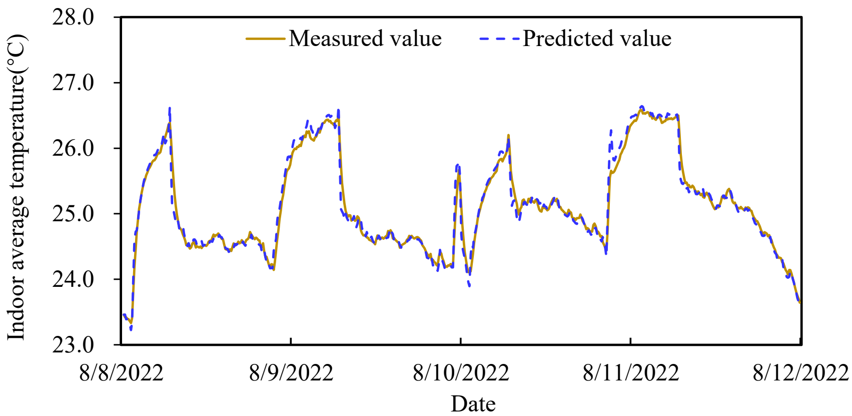


Fig. 11. Comparison between the measured and predicted indoor average temperature.

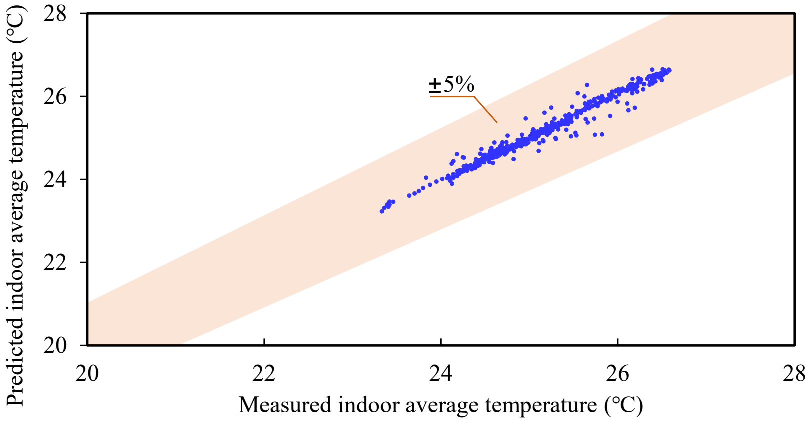


Fig. 12. Prediction error of the building thermal model

4.2 Performance evaluation

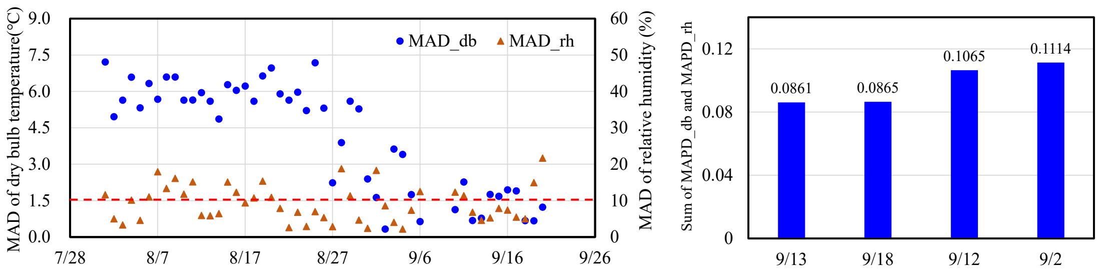
The control retrofit for this HVAC system was completed at 22nd Sep 2022, and the deployed MPC strategy continuously optimized the operation of the HVAC systems until the end of the cooling season, 6th Oct. During the 15 days running, the operation of the MPC strategy and the bidirectional information transmission between the cloud platform and the existing BAS were stable.

In order to evaluate the performance of the MPC strategy in energy saving and thermal comfort improvement, the system power consumption and indoor average temperature under MPC strategy are compared with those under original RBC strategy at the similar days. During the operation period after control retrofit, the cooling load of the building is very small at most of the operation time and one ACHP operates at the stage 2, i.e., the stage with the smallest cooling capacity, can fulfill the cooling demand. In such cases, the advantage of the MPC strategy in optimizing the system cooling output and ACHP operating number cannot be reflected. Fortunately, the building cooling load increased a little during 30th Sep to 2nd Oct due to the increase of the outdoor air temperature. Therefore, the performance of the MPC strategy during the three typical days is analyzed in detail.

4.2.1 Determination of similar days

The outdoor dry bulb temperature and the relative humidity, which are highly correlated to the cooling load, are the variables concerned while determining the similar days. The mean absolute difference (MAD) and mean absolute percentage difference (MAPD) of the two variables between the days using MPC strategy and the days using RBC strategy are calculated to quantify the similarity. Based on the four indicators, i.e., MAD and MAPD of outdoor dry bulb temperature (MAD_{db}, $MAPD_{db}$), MAD and MAPD of relative humidity (MAD_{rh}, $MAPD_{rh}$), the most similar day to a selected day can be determined. In this study, the most similar days are determined through two steps. Firstly, the days with MAD_{db}<1.5 °C and MAD_{rh}<10% are filtered out as the qualified days. Then, the day with the minimum value of the sum of $MAPD_{db}$ and $MAPD_{rh}$ among the qualified days is determined as the most similar day.

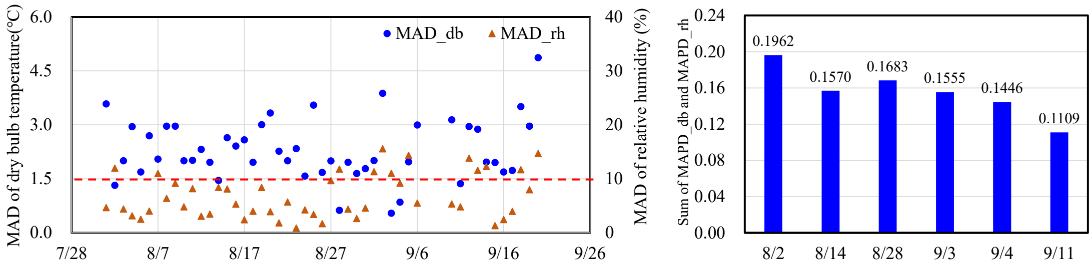
The similarities among the days using RBC strategy (from 1st Aug to 20th Sep) and the days using MPC strategy are evaluated based on the four indicators and the most similar days are identified according to the evaluation results. Fig. 13 presents the values of MAD_{db} and MAD_{rh} among the days using the RBC strategy and 30th Sep, and the values of the sum of $MADP_{db}$ and $MAPD_{rh}$ of the qualified days. It can be seen that among the four qualified days, 13th Sep, with the $MAPD_{sum}$ of 0.0861 is the most similar day to 30th Sep. Similarly, as show in Fig. 14 and Fig. 15, the most similar days to 1st Oct and 2nd Oct are 11th Sep ($MAPD_{sum}$ is 0.1109) and 28th Aug ($MAPD_{sum}$ is 0.1156), respectively.



(a) MAD_{db} and MAD_{rh} of all the days.

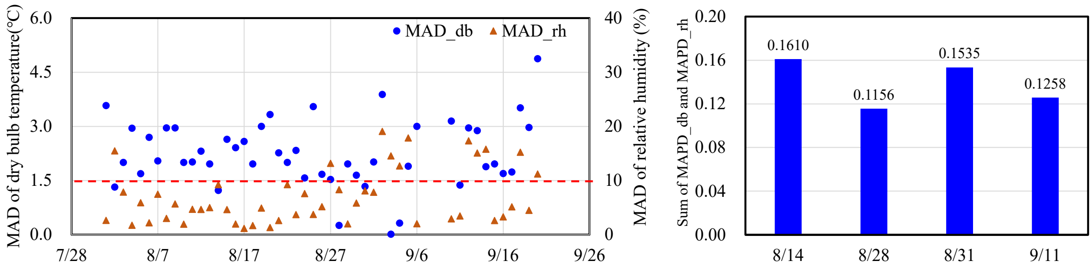
(b) $MAPD_{sum}$ of the qualified days

Fig. 13. Values of the indicators between the days using RBC strategy and 30th Sep.



(a) MAD_{db} and MAD_{rh} of all the days. (b) $MAPD_{sum}$ of the qualified days

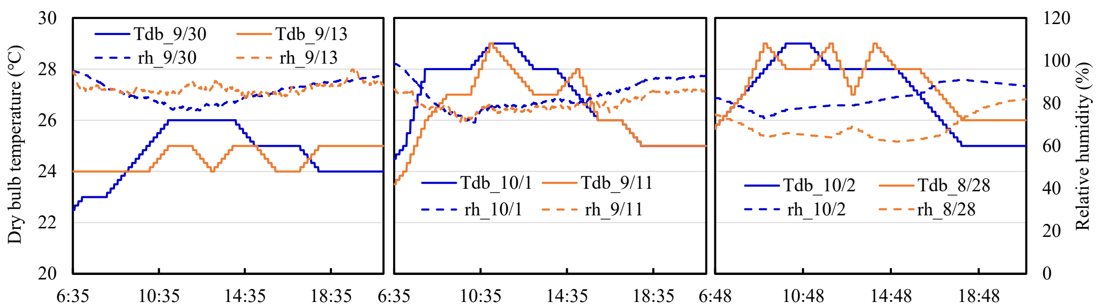
Fig. 14. Values of the indicators between the days using RBC strategy and 1st Oct.



(a) MAD_{db} and MAD_{rh} of all the days. (b) $MAPD_{sum}$ of the qualified days

Fig. 15. Values of the indicators between the days using RBC strategy and 2nd Oct.

Fig. 16 presents the comparison of the dry bulb temperature and relative humidity of the three groups of similar days. It can be seen that the dry bulb temperature and relative humidity of each group of similar days are very close, indicating the cooling demand should be close. Therefore, it is reasonable to evaluate the performance of the MPC strategy through comparing the system energy consumption and indoor thermal comfort of the selected similar days.



(a) Group 1 (b) Group 2 (c) Group 3

Fig. 16. Comparison of the dry bulb temperature and relative humidity of the similar days

4.2.2 Energy performance

Fig. 17 presents the optimized cooling output and the actual system cooling output using the MPC strategy and the RBC strategy. It can be seen that the system cooling outputs when using the MPC strategy were obviously smaller than those of using RBC strategy. Fig. 18 presents the set points and actual values of the chilled water supply temperature when using the MPC strategy and RBC strategy. The set points and the actual values of the chilled water supply temperature when using the MPC strategy were obviously higher than those of using the original RBC strategy. This is because that the MPC strategy optimized the system cooling output according to the prediction results of the building thermal model and adjusted the chilled water supply temperature to control the system cooling output. In this means, the system cooling output was controlled according to the actual building cooling demand and therefore the redundant cooling output was reduced or even eliminated. While the RBC strategy adjusted the chilled water supply temperature according to the wet bulb temperature only and redundant cooling output cannot be avoided. It is worth noting that, since the ACHPs used in this system cannot modulate its output smoothly, there always exists difference between the set points and the measured values of the chilled water supply temperature and system cooling output when using both strategies.

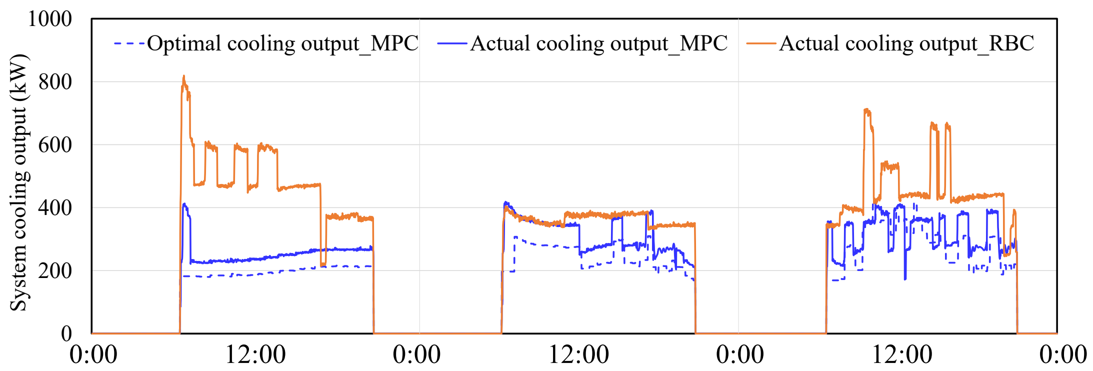


Fig. 17. System cooling output when using MPC and RBC strategies

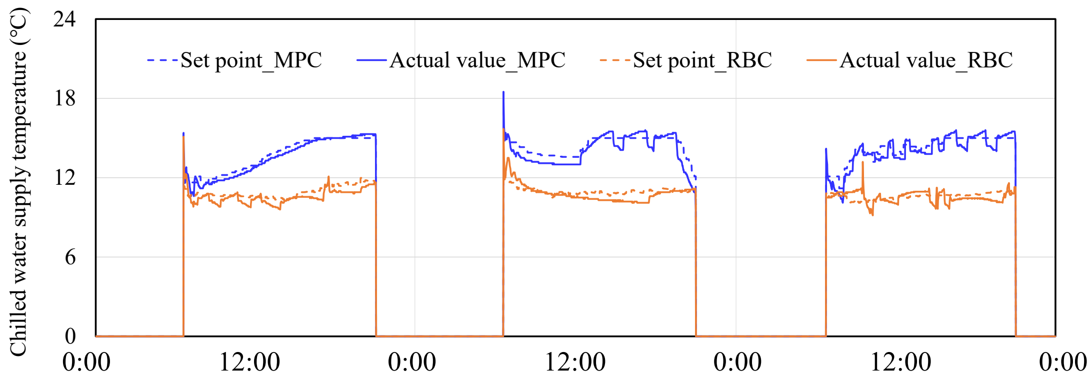


Fig. 18. Set points and actual values of chilled water supply temperature using MPC and RBC strategies

Through the optimization and control of the system cooling output, MPC strategy showed superior energy performance than the RBC strategy. Fig. 19 compares the system daily energy consumption, which is the sum of the daily energy consumption of the operating ACHPs, CHWPs and AHUs, when using the MPC strategy and RBC strategy in the similar days. It can be found that, the system energy consumption was obvious smaller when using the MPC strategy compared with that of the similar days when using RBC strategy. The detailed statistical data is presented in Table. 1. The energy savings of the three days were 1493.73 kWh (37.3%), 364.32 kWh (10.6%) and 1034.65 kWh (23.7%), respectively. It can be found that the energy savings mainly come from the reduction of the energy consumption of ACHPs. The main reason is that the high chilled water supply temperature reduced the cooling output of the ACHPs and improved their energy efficiency. It can also be observed the energy consumption of the AHUs when using MPC strategy were higher than that of using RBC strategy in group 2 and group 3. The main reason is that the volume of cooled air required for maintaining indoor temperature was larger when using the MPC strategy, as the supply air temperature of the AHUs were higher due to the higher chilled water supply temperature.

It is a pity that the performance of the sequencing optimization cannot be evaluated due to the small cooling load in the testing days. Its performance will be evaluated in the operation of next cooling season.

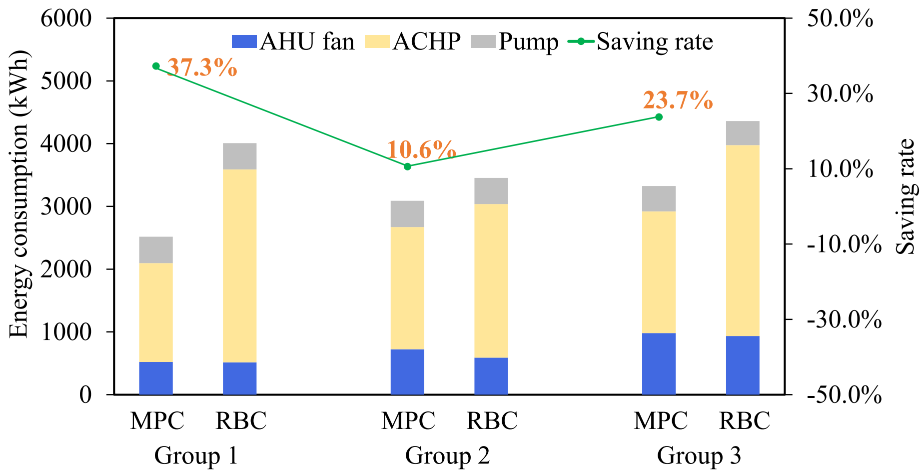


Fig. 19. Comparison of the system power consumption of the similar days

Table. 1. Comparison of the power consumption of the similar days

Group NO.	Control strategy	Energy end-use (kWh)			Total energy consumption (kWh)	Energy savings (kWh)	Energy saving rate
		ACHPs	Pumps	AHUs			
1	MPC	1576.91	418.94	519.95	2515.80	1493.73	37.25%
	RBC	3074.27	419.08	516.18	4009.54	-	-
2	MPC	1947.54	417.99	723.14	3088.66	364.32	10.55%
	RBC	2447.15	416.56	589.26	3452.97	-	-
3	MPC	1938.78	404.86	981.05	3324.70	1034.65	23.73%
	RBC	3039.34	384.46	935.54	4359.34	-	-

4.2.3 Indoor thermal comfort

Through optimizing the system cooling output according to the cooling demand, the indoor comfort was obviously improved after the control retrofit, i.e., when using the MPC strategy. The indoor average temperature is adopted as the indicator of the indoor thermal comfort in this study. According to the design code for heating ventilation and air conditioning of civil buildings [41], 24 °C to 26 °C is recommended as the most comfortable indoor temperature region in cooling season. Fig. 20 presents the indoor average temperature distribution of the three groups of similar days. The overcooling phenomenon, i.e., the indoor average temperature is lower than 24 °C, was very common when using the RBC strategy, especially in group 1 and group 3. When using

MPC strategy, the indoor average temperature was controlled between 24 °C to 26 °C at most of the time. The percentage of the time that the indoor temperature was kept in the comfort region of the three days when using RBC strategy were 2.42%, 70.59% and 16.02%, respectively. When using the MPC strategy, the percentages were improved to 83.97%, 100% and 98.94%, respectively.

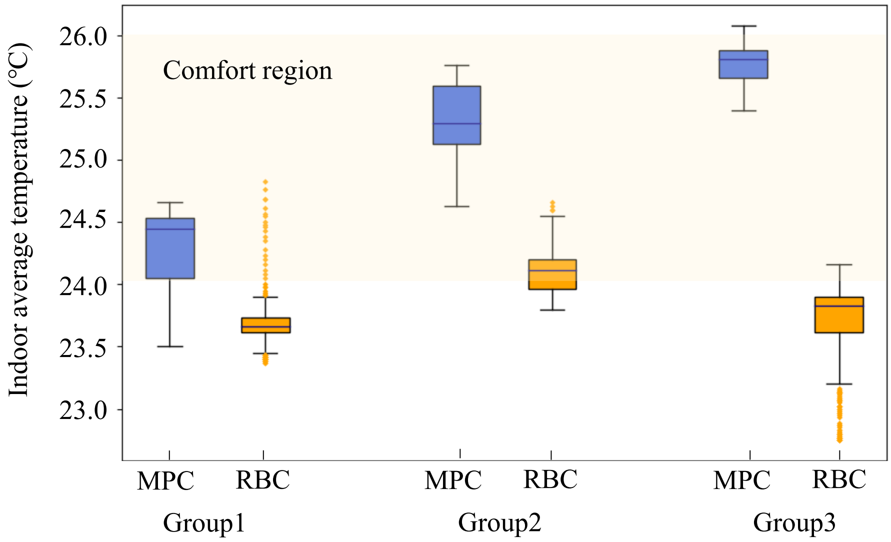


Fig. 20. Indoor average temperature comparison when using MPC and RBC strategies

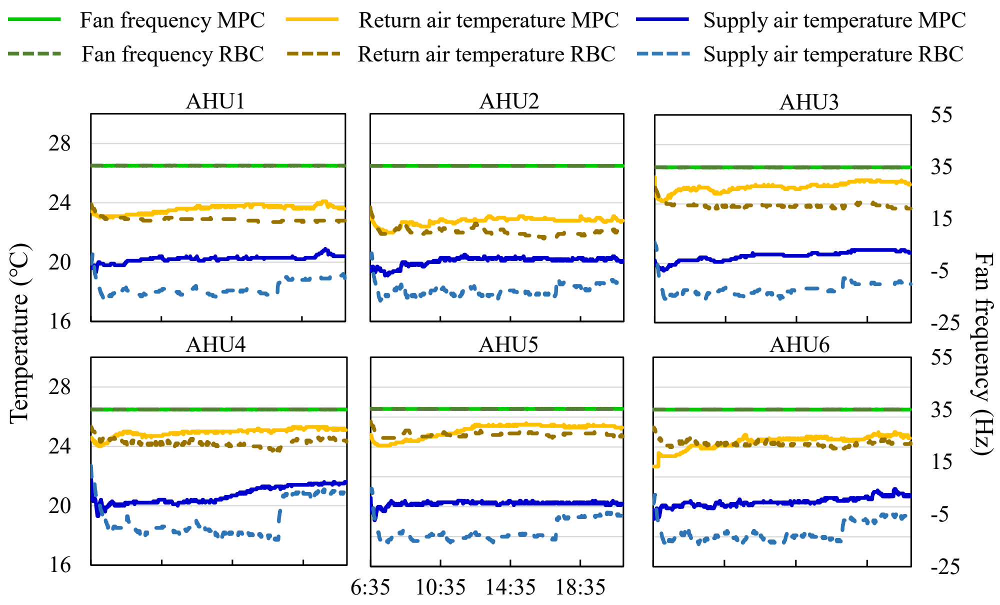
4.2.4 Detail analysis

The test results show that the MPC strategy has superior performance in both energy saving and indoor thermal comfort than the RBC strategy. In order to investigate the advantage of the MPC strategy, the operation of the HVAC system under the two strategies are analyzed in detail.

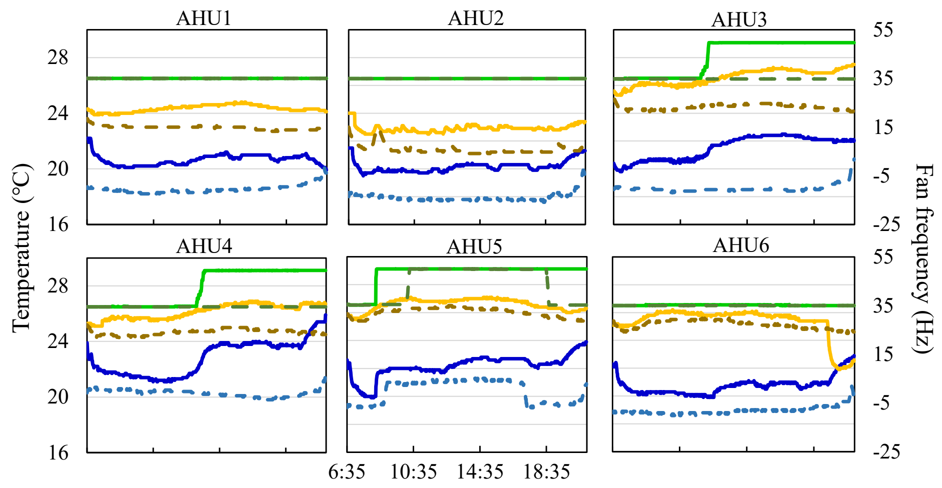
Both the MPC strategy and RBC strategy are used to optimize the set points of the cooling plant and the set points of the AHUs, which include the set points of supply air temperature and indoor temperature, are adjusted by the operators manually. According to the operation record, the set points of the supply air temperature are kept constant in the six days concerned, which are 18 °C. The set points of the indoor temperature are the same, i.e., 26 °C, in the concerned six days except for 28th Aug. The set points of the six AHUs are set to be 26 °C, 26 °C, 23 °C, 23 °C, 25 °C, 25 °C, respectively in 28th Aug. Actually, this is an experience-guided rule used by the operators to avoid insufficient cooling supply of the four AHUs located in the second floor of the airport terminal. The set points of the AHUs will be set to be lower at the beginning of July and reset to be 26℃ at the beginning of September. The reason is that the cooling demand of the four region is relatively larger. It is worth noting that, the adjustment of the set points doesn't indicate that the target indoor air temperature of the four regions are lower. The expected indoor air temperature of the six AHUs is still 24℃~28℃.

Fig. 21 presents the fan frequency, return air temperature and supply air temperature of the six AHUs at the three groups of similar days. It can be found that, no matter the set points of the AHUs are the same or not, the MPC strategy reduced the overcooling in AHU1 and AHU2 compared with that using RBC strategy, and maintained the indoor temperature of the other AHUs within the expected range. The main reason is that the MPC control strategy can further reduce the cooling output of the system by adjusting the water supply temperature when the air side regulation reaches the limit. Taking AHU1 as an example, during the operation in the three groups of similar days, its fan frequency kept at the lowest value (35 Hz), when using both of two strategies, which indicates that the cooling output cannot be reduced any more by the regulation of the AHU. In such cases, when the MPC strategy is used, the supply air temperature of the AHU was higher due to the supply of higher chilled water supply temperature, which further reduced the system cooling output. Owing to the accurate prediction of the indoor temperature by the building thermal model, the return air temperature of the AHUs with relatively higher cooling demand were maintained within the expected range with the reduced system cooling output.

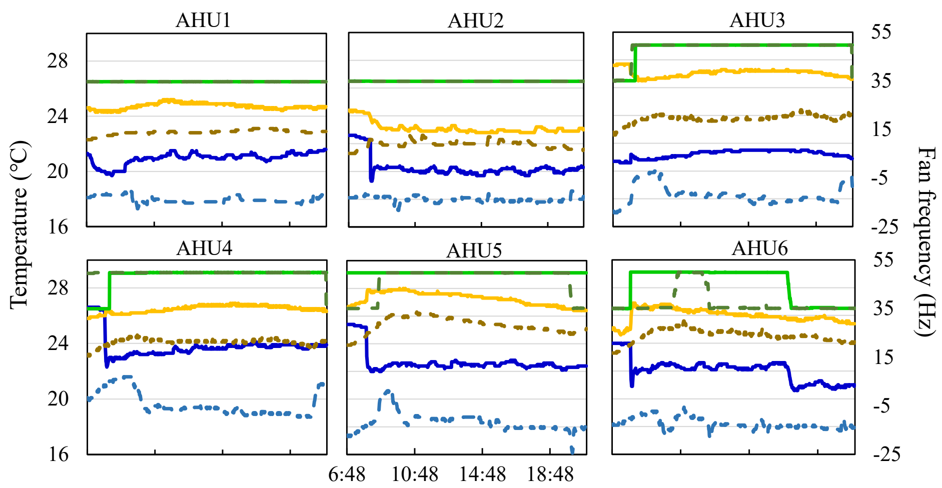
It is possible to formulate a RBC strategy to realize the system cooling output reduction as the MPC strategy through adjusting the chilled water supply temperature according to the return air temperature. However, without the prediction of the indoor temperature, it's hard to determine the set point of chilled water supply temperature which can reduce the system energy consumption while maintaining the indoor thermal comfort. As presented in the previous studies, the RBC strategies normally perform worse in energy savings and indoor thermal comfort than MPC strategies [22]. Moreover, in order to ensure the performance of such RBC strategies in practical applications, their parameters such as the step of the adjustment and the threshold for determining the execution of adjustment, needed to be identified and configured carefully. Therefore, considering the large-scale applications of the control strategy, using MPC strategy would be more effective and efficient.



(a) Group 1



(b) Group 2



(c) Group 3

Figure. 21. Fan frequency, return air temperature, and supply air temperature of the six AHUs in the three groups of similar days

5 Conclusions

This study proposes a lightweight and replicable control retrofit approach to improve the energy efficiency of the building HVAC systems adopting MPC strategies. A general MPC strategy is proposed for the energy conservation of HVAC systems with different configurations. A cloud-edge collaboration architecture is proposed to realize the integration of the MPC strategies with the existing BASs without modifying the original control logics. With the help of the scalability and distributed computing capabilities of the cloud computing platform, multiple MPC strategies can be deployed on the cloud platform to optimize the operation of the HVAC systems through interacting with their existing BASs. The proposed control retrofit approach is expected to be able to provide lightweight control retrofit services for large number of existing building HVAC systems.

The proposed control retrofit approach was implemented in a real-life airport terminal and the MPC strategy showed superior performance than the original RBC strategy in the existing BASs both in energy and thermal comfort. Compared with the similar days before the control retrofit, the system energy consumption was reduced by 37.3%

(1493.73 kWh), 10.6% (364.32 kWh) and 23.7% (1034.65 kWh), respectively, and the percentages of the discomfort time were reduced from 97.58%, 29.41% and 83.98% to 16.03%, 0.00% and 1.06%, respectively. Moreover, the proposed hybrid building thermal model was proven to be effective in predicting indoor temperature. During the test period, the prediction error was lower than 0.2 °C in most of the cases and the RMSE was 0.12 °C.

Further efforts will be devoted to developing model update mechanism to ensure the prediction accuracy of the system models. Since the effectiveness of the MPC strategies highly relies on the accuracy of the prediction models, and the model performance could become worse in continuous operation due to the variation of the working conditions and the equipment performance.

Appendix

A.1 Hybrid building thermal model

A.1.1 Gray-box building thermal model

Considering the case study building as a whole, according to the law of conservation of energy, the heat balance equation of the building can be represented as Eq. (A.1).

$$C_{IA}\left(T_{IA,t_1} - T_{IA,t_0}\right) = \Delta Q_{cooling} + \Delta Q_{enve} + \Delta Q_{solar} + \Delta Q_{AE} + \Delta Q_{IHG} \qquad (A.1)$$

Where C_{IA} is the heat capacity of the indoor air [kJ/K], T_{IA} is the indoor average temperature, $\Delta Q_{cooling}$ stands for the cooling energy [kJ] provided to the indoor air between t_0 and t_1, ΔQ_{enve}, ΔQ_{solar}, and ΔQ_{AE} indicate the heat gains caused by heat transfer through the building envelope, solar radiation, and air exchange between indoor and outdoor, respectively, and ΔQ_{IHG} indicates the internal heat gains caused by occupants' activities, lighting, and electricity appliances, etc. ΔQ_{enve} can be calculated according to Eq. (A.2).

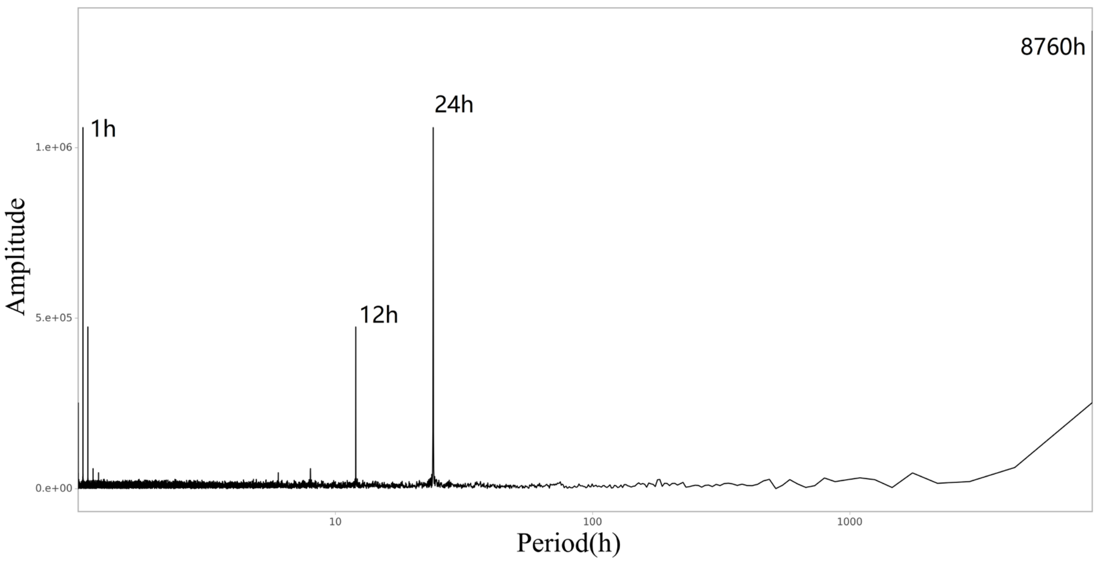
$$\Delta Q_{enve} = \left(\frac{T_{OA} - T_{IA}}{R_{enve}}\right)\Delta t \qquad (A.2)$$

Where T_{OA} is the outdoor air temperature, R_{enve} is the equivalent thermal resistance of the building envelope [K/kW] and is considered as a constant for simplification, and Δt stands for the time length between t_0 and t_1.

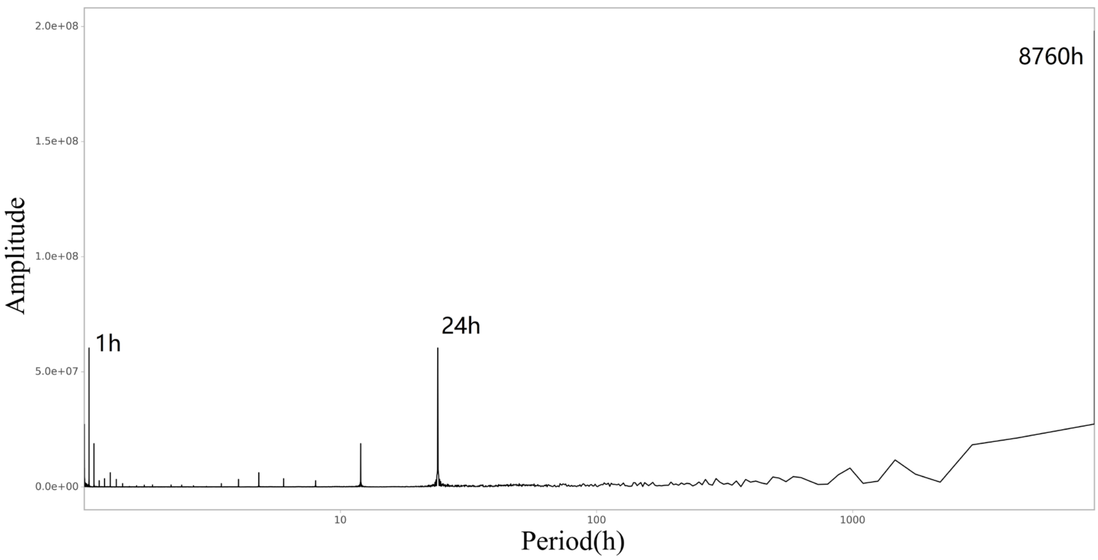
ΔQ_{solar}, ΔQ_{AE} and ΔQ_{IHG} of a building cannot be easily measured in practice. However, the heat gains are considerably influenced by factors that have strong periodicity, such as solar irradiance, weather and building occupancy [42]. A spectrum analysis [43] is conducted on the solar irradiance data of the case study site (from a typical meteorological year database [44]) and the occupancy data of another airport terminal building in China (from Ref. [45]) to illustrate the periodicity of the factors. According to Fig. A.1 a), the most considerable periods in the solar irradiance data include one year (8760 hour), one day (24 hour), a half day (12hour) and these around one hour, which can be interpreted as annual, daily, and intra-day variabilities. While the period of one year, one day and a half day are mostly due to the nature of sun, the periods around one hour is caused by occasional factors, such as cloud and measurement noise, that can be eliminated from the prediction model to avoid overfitting [46]. In Fig. A. 1 b), the most considerable periods in the occupancy data include one year (8760 hour), one day (24 hour) and these around one hour, which can be interpreted as annual, daily, and intra-daily variabilities. for simplicity, ΔQ_{solar}, ΔQ_{AE} and ΔQ_{IHG} are not described in detail but innovatively aggregated as one term that can be represented using Fourier series. Accordingly, the three variables can be expressed as in Eq. (A.3).

$$\Delta Q_{solar} + \Delta Q_{AE} + \Delta Q_{IHG} = \Delta Q_{agg} = \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos\frac{2\pi}{L} nt + b_n \sin\frac{2\pi}{L} nt \right) \quad (A.3)$$

Where ΔQ_{agg} stands for the aggregated heat gain term, a_0, a_n and b_n (n=1, 2, …, N) are coefficients to be estimated, N is the number of orders of the Fourier series, L is the period of the aggregated heat gain term, and t stands for time.



(a) Solar irradiance



b) Occupancy

Fig. A. 2. Spectrum analysis of influencing factors of aggregated heat gains in an airport in China.

In this study, L and N are determined based on the aforementioned spectrum analysis (Fig. A. 1.). Obviously, L should be chosen as large as possible (e.g., one year) to cover the whole most considerable periods (i.e., 1 hour, 12 hour, 24 hour and one year) as shown in Fig. A.2. However, since the time length of dataset for model training is usually from a couple of weeks to a few months in practical, thus L of one year is meaningless and L of 24 hour is chosen. Similarly, N should be chosen as large as possible to contain the high-frequency intra-day variabilities with shorter periods, e.g., one hour or less. However, since the variabilities of periods around one hour maybe caused by some occasional factors (e.g., cloud movement and measurement noise) and the building with high thermal inertial is insensitive to such high-frequency variabilities, these part of variabilities are thus not considered to avoid overfitting and improve the generality of the model. Therefore, the value of L and N are set to be 24 hour and 2 respectively, which means that the Fourier expansion of the aggregated heat gains term contains the variabilities of period 24 hour and 12 hour.

According to Eq. (A.1) to Eq. (A.3), the indoor average temperature at t_1 can be calculated using Eq. (A.4).

$$T_{IA,t_1} = \frac{\Delta Q_{cooling} + (\frac{T_{OA}-T_{IA}}{R_{enve}})\Delta t + \frac{a_0}{2} + \sum_{n=1}^{N}\left(a_n \cos\frac{2\pi}{L}nt + b_n \sin\frac{2\pi}{L}nt\right)}{C_{IA}} + T_{IA,t_0} \quad (A.4)$$

Since it is difficult to obtain the average temperature difference between indoor and outdoor air during Δt in practice, considering that the change of either indoor or outdoor air is small during a short time period, the average temperature difference from t_0 to t_1 can be approximated using the instantaneous temperature difference at t_0. Eq. (A.4) can consequently be expressed as Eqs. (A.5)-(A.7).

$$T_{IA,t_1} = c_1\left(T_{OA,t_0} - T_{IA,t_0}\right) + c_2\left(\Delta Q_{cooling} + \frac{a_0}{2} + \sum_{n=1}^{N}\left(a_n \cos\frac{2\pi}{L}nt + b_n \sin\frac{2\pi}{L}nt\right)\right) + T_{IA,t_0} \quad (A.5)$$

$$c_1 = \frac{1}{C_{IA}R_{enve}} \quad (A.6)$$

$$c_2 = \frac{1}{C_{IA}} \quad (A.7)$$

Eq. (A.5) is the gray-box building thermal model used for indoor average temperature prediction, where $T_{IA,t1}$ is the indoor average temperature to be predicted, $T_{OA,t0}$, $T_{IA,t0}$, $\Delta Q_{cooling}$ and t are predictor variables, a_0, a_n, b_n, c_1, and c_2 are coefficients to be estimated, N and L are hyper-parameters to be determined by the user. In this study, N and L are set to be one week and 28 respectively determined by trial and error, which means the period and fundamental period of the aggregated term is one week and 1/4 day, respectively.

The unknown coefficients in the derived model, i.e., Eq. (A.5), including a_0, a_n, b_n, c_1, and c_2 (n=1, 2, …, N), are estimated using historical system operation data of the case study building. Since c_1 and c_2 in the model cannot be negative from the perspective of physics, a non-negative least squares method proposed by Lawson and Hanson [47] is utilized in this study to estimate the two coefficients.

A.1.2 Data-driven residual model

Due to the complicity of the thermal transfer in a building, the non-linearity relationship between the indoor average temperature and the independent variables, such as outdoor air temperature and amount of cooling energy, cannot be completely captured by the simplified gray-box model, which contributes to the residuals between the measured and predicted values. In order to improve the prediction accuracy of the gray-box model, a residual prediction model is developed to predict its prediction residuals. Since machine learning methods commonly show good performance in modelling non-linearity relationships, a machine learning model is trained to predict the residuals. In this study, Cubist [39] is used for the modelling of residuals considering its better performance in extrapolation compared to most other machine learning algorithms. Cubist is a machine learning algorithm that aggregates multiple model trees to generate an ensemble model with higher accuracy. Each model tree in the ensemble model has a tree-like structure which is similar to a regression tree, while the constant value in each terminal node of the regression tree is replaced by a linear regression model.

A.2 Air-cooled heat pump model

A.2.1 Stage prediction module

As aforementioned, the stage prediction module is responsible for the prediction of the operation stage and the individual cooling outputs of the ACHPs. For this problem, data-driven model is an effective method since the control logic of the ACHP operation stage is unknown. However, the accuracy of the data-driven models highly relies on the diversity and accuracy of the training data. Once the testing data is out of the range of the training data set, data-driven models are likely to generate incorrect results. Therefore, in this study, a revision scheme is developed to revise the predicted operation stages by the data driven model and therefore improve the prediction accuracy.

The schematic and prediction process of the stage prediction module is presented in Fig. A.2. According to the input data, including the system total cooling output (CLG_{tot}), the number of operating ACHPs (N_{hp}), the chilled water supply temperature (T_{chws}) and the ambient dry bulb temperature (T_{db}), operation stages of the operating ACHPs are predicted using the decision tree models. Then, capacity models are used to predict the capacities of the operating ACHPs at the predicted stages. According to the difference between the sum of the predicted cooling capacities and the total system cooling output, the predicted operation stages are revised until the difference is less than the threshold (ε) or the revision loop has been conducted for more than two times. The threshold value is ten percent of the total system cooling output in this case. With the determined stage and corresponding cooling capacities of the operating ACHPs, the individual cooling outputs of the ACHPs are calculated.

The capacity models for the ACHPs are shown as Eq. (A.8).

$$Cap = d_1 \times T_{chws} + d_2 \tag{A.8}$$

Where Cap is the cooling capacity of one ACHP at a stage, T_{chws} is the chilled water supply temperature, d_1 and d_2 are the coefficients. Since the performance of the heating pump sizably varies from one stage to another, the coefficients of the capacity models for one ACHP of stage 2, stage 3 and stage 4 need to be determined separately. Stage 1 is not considered in the prediction model since it is only lasted for a very short time (normally less than 10 minutes) during the starting up or shut down period.

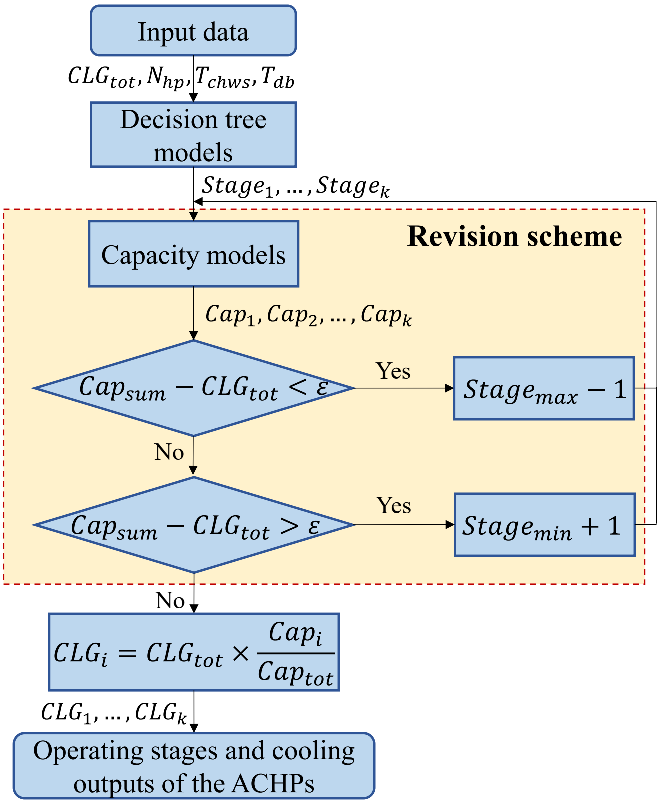


Fig. A.2. Schematic of the stage prediction module

A.2.2 Power prediction module

With the predicted operation stages and cooling outputs of the operating ACHPs, their power consumption can then be predicted. To the best of the authors' knowledge, there is no available air-cooled heat pump model that is suitable for this study. Therefore, a verified multivariate polynomial model [48] for the water-cooled chiller was referred to develop the energy usage prediction models for the ACHPs in this study. The formulas used in this study to predict the energy usage of an ACHP is shown as following.

$$COP = \beta_0 + \beta_1 CLG + \beta_2 T_{chws} + \beta_3 T_{db} + \beta_4 CLG^2 + \beta_5 {T_{chws}}^2 + \beta_6 {T_{db}}^2 + \beta_7 CLGT_{chws} + \beta_8 CLGT_{db} + \beta_9 T_{db} T_{chws} \quad (A.9)$$

$$P_{chi} = \frac{CLG}{COP} \quad (A.10)$$

Where COP is the coefficient of performance, CLG is the cooling output, T_{db} is the outdoor dry-bulb temperature, T_{chws} is the chilled water supply temperature, and β_0 to β_9 are coefficients need to be identified according to the operation data.

It is worth noting that, although the six ACHPs have the same nominal capacities, their performance are diverse with each other because of the different degrees of performance degradation. Therefore, the model parameters of the six ACHPs are needed to be identified according to their operation data separately.

A.3 Chilled water pump model

Four chilled water pumps are connected in parallel in the studied system, in which three for use and one for backup. The electric power of a pump depends on the water flow rate and its frequency. When the pressure difference between supply and return water pipes is constant, the flow rate of the pump is determined by the resistance of the pipe network and pump frequency. The operation data of the system show that the openings of the terminal water dampers have little influence on the resistance of the pipe network and the number of operating ACHPs and pumps is the main factor affecting the resistance of the pipe network. The main reason is that the manual valve of the bypass pipeline keeps wide open. Therefore, when the operating number of ACHPs and chilled water pumps are determined, the resistance of the pipe network can be regarded as constant. In such case, according to the affinity laws, the water flow rate (*F*) can be determined by the frequency (*f*) of the pump shown as Eq. (A.11).

$$\frac{F}{F_{rated}} = \frac{f}{f_{rated}} \tag{A.11}$$

Then, the pressure difference (dP) can be determined by the pump frequency (f) as shown in Eq. (A.12).

$$dP = g_0 + g_1 f + g_2 f^2 \tag{A.12}$$

With the pump frequency obtained by solving Eq. (16) according to the measured pressure difference, the power consumption (P_{pump}) of the operating pump can be obtained according to Eq. (A.13).

$$P_{pump} = h_0 + h_1 f + h_2 f^2 + h_3 f^3 \tag{A.13}$$

The coefficients g_0, g_1, g_2 and h_0 to h_3 are various with the number of operating ACHPs and water pumps and need to be identified according to the corresponding operation data separately.

Reference

[1] Kairui You, Hong Ren, Weiguang Cai, Ruopeng Huang, Yuanli Li, Modeling carbon emission trend in China's building sector to year 2060, Resources, Conservation and Recycling, Volume 188, 2023, 106679.
[2] Guannan Li, Qing Yao, Yunpeng Hu, Xi Fang, Luhan Wang, Investigating thermostat sensor offset impacts on operating performance and thermal comfort of three different HVAC systems in Wuhan, China, Case Studies in Thermal Engineering, Volume 31, 2022,101788.
[3] Zu Wang, John Calautit, Shuangyu Wei, Paige Wenbin Tien, Liang Xia, Real-time building heat gains prediction and optimization of HVAC setpoint: An integrated framework, Journal of Building Engineering, Volume 49, 2022, 104103.
[4] Jarosław Strzałkowski, Pawel Sikora, Sang-Yeop Chung, Mohamed Abd Elrahman, Thermal performance of building envelopes with structural layers of the same density: Lightweight aggregate concrete versus foamed concrete, Building and Environment, Volume 196, 2021, 107799.
[5] Sivanand Somasundaram, Sundar Raj Thangavelu, Alex Chong, Improving building efficiency using low-e coating based retrofit double glazing with solar films, Applied Thermal Engineering, Volume 171, 2020, 115064.
[6] Jae-Sol Choi, Changmin Kim, Hyangin Jang, Eui-Jong Kim, Dynamic thermal bridge evaluation of window-wall joints using a model-based thermography method, Case Studies in Thermal Engineering, Volume 35, 2022, 102117.
[7] Xiaoxu Cai, Huixing Li, Guohui Feng, Shui Yu, Yibo Zhao, HVAC System Green Retrofit Survey and Analysis of Public Institutions Building in Cold Region, Procedia Engineering, Volume 146, 2016, Pages 218-223.
[8] Soumaya EL BARRAK, Alessandro De La Garza, Jordan Mardis, Timmy Nguyen, Mi Tran, Jennifer Tzoc, Annie Vallejo, Preston Turner, Driss Benhaddou, IoT-Based Smart Airflow System for Retrofitting Commercial Variable Air Volume HVAC Systems, IFAC-PapersOnLine, Volume 55, Issue 12, 2022, Pages 444-449.
[9] Luigi Schibuola, Massimiliano Scarpa, Chiara Tambani, CO2 based ventilation control in energy retrofit: An experimental assessment, Energy, Volume 143, 2018, Pages 606-614.
[10] Yutong Li, Jun Ren, Zhou Jing, Lu Jianping, Qing Ye, Zhijun Lv, The Existing Building Sustainable Retrofit in China-A Review and Case Study, Procedia Engineering, Volume 205, 2017, Pages 3638-3645.
[11] Yuan Lai, Sokratis Papadopoulos, Franz Fuerst, Gary Pivo, Jacob Sagi, Constantine E. Kontokosta, Building retrofit hurdle rates and risk aversion in energy efficiency investments, Applied Energy, Volume 306, Part B, 2022, 118048.
[12] Tianyi Liu, Guofeng Ma, Ding Wang, Pathways to Successful Building Green Retrofit Projects: Causality Analysis of Factors Affecting Decision Making, Energy and Buildings, 2022, 112486.
[13] Natasa Djuric, Vojislav Novakovic, Review of possibilities and necessities for building lifetime commissioning, Renewable and Sustainable Energy Reviews, Volume 13, Issue 2, 2009, Pages 486-492.
[14] Yan Ding, Hao Su, Kuixing Liu, Qiaochu Wang, Robust commissioning strategy for existing building cooling system based on quantification of load uncertainty, Energy and Buildings, Volume 225, 2020, 110295.
[15] Muhyiddine Jradi, Na Liu, Krzysztof Arendt, Claudio Giovanni Mattera, An automated

framework for buildings continuous commissioning and performance testing – A university building case study, Journal of Building Engineering, Volume 31, 2020, 101464.
[16] Jakob Bjørnskov, Muhyiddine Jradi, Christian Veje, Component-level re-commissioning of a newly retrofitted Danish healthcare building, Journal of Building Engineering, Volume 51, 2022, 104277.
[17] Shengwei Wang & Zhenjun Ma (2008) Supervisory and Optimal Control of Building HVAC Systems: A Review, HVAC&R Research, 14:1, 3-32.
[18] Maryam Gholamzadehmir, Claudio Del Pero, Simone Buffa, Roberto Fedrizzi, Niccolo' Aste, Adaptive-predictive control strategy for HVAC systems in smart buildings – A review, Sustainable Cities and Society, Volume 63, 2020, 102480.
[19] Ye Yao, Divyanshu Kumar Shekhar, State of the art review on model predictive control (MPC) in Heating Ventilation and Air-conditioning (HVAC) field, Building and Environment, Volume 200, 2021, 107952.
[20] Serale, G., Fiorentini, M., Capozzoli, A., Bernardini, D., Bemporad, A. Model Predictive Control (MPC) for Enhancing Building and HVAC System Energy Efficiency: Problem Formulation, Applications and Opportunities. Energies 2018, 11, 631.
[21] Ján Drgoňa, Javier Arroyo, Iago Cupeiro Figueroa, David Blum, Krzysztof Arendt, Donghun Kim, Enric Perarnau Ollé, Juraj Oravec, Michael Wetter, Draguna L. Vrabie, Lieve Helsen, All you need to know about model predictive control for buildings, Annual Reviews in Control, Volume 50, 2020, Pages 190-232.
[22] Abdul Afram, Farrokh Janabi-Sharifi, Theory and applications of HVAC control systems – A review of model predictive control (MPC), Building and Environment, Volume 72, 2014, Pages 343-355.
[23] Shiyu Yang, Man Pun Wan, Bing Feng Ng, Swapnil Dubey, Gregor P. Henze, Wanyu Chen, Krishnamoorthy Baskaran, Experimental study of model predictive control for an air-conditioning system with dedicated outdoor air system, Applied Energy, Volume 257, 2020, 113920.
[24] Xiao Chen, Qian Wang, Jelena Srebric, Occupant feedback based model predictive control for thermal comfort and energy optimization: A chamber experimental evaluation, Applied Energy, Volume 164, 2016, Pages 341-351.
[25] Sen Huang, Yashen Lin, Venkatesh Chinde, Xu Ma, Jianming Lian, Simulation-based performance evaluation of model predictive control for building energy systems, Applied Energy, Volume 281, 2021, 116027.
[26] Abdul Afram, Farrokh Janabi-Sharifi, Supervisory model predictive controller (MPC) for residential HVAC systems: Implementation and experimentation on archetype sustainable house in Toronto, Energy and Buildings, Volume 154, 2017, Pages 268-282.
[27] Wei Liang, Rebecca Quinte, Xiaobao Jia, Jian-Qiao Sun, MPC control for improving energy efficiency of a building air handler for multi-zone VAVs, Building and Environment, Volume 92, 2015, Pages 256-268.
[28] Seungjae Lee, Jaewan Joe, Panagiota Karava, Ilias Bilionis, Athanasios Tzempelikos, Implementation of a self-tuned HVAC controller to satisfy occupant thermal preferences and optimize energy use, Energy and Buildings, Volume 194, 2019, Pages 301-316.
[29] David Blum, Zhe Wang, Chris Weyandt, Donghun Kim, Michael Wetter, Tianzhen Hong, Mary Ann Piette, Field demonstration and implementation analysis of model predictive control in an office HVAC system, Applied Energy, Volume 318, 2022, 119104.

[30] Max Bird, Camille Daveau, Edward O'Dwyer, Salvador Acha, Nilay Shah, Real-world implementation and cost of a cloud-based MPC retrofit for HVAC control systems in commercial buildings, Energy and Buildings, Volume 270, 2022, 112269.
[31] David Blum, Michael Wetter. MPCPy: An open-source software platform for model predictive control in buildings. In: Proceedings of the 15th IBPSA conference. 2017, Pages 1381–90.
[32] Docker: Accelerated, Containerized Application Development. https://www.docker.com
[33] Michael Armbrust, Ali Ghodsi1, Reynold Xin, Matei Zaharia. Lakehouse: A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics. Conference on Innovative Data Systems Research, 2021.
[34] Apache Kafka: An open-source distributed event streaming platform. https://kafka.apache.org/
[35] PySpark: Python API for Spark. https://spark.apache.org/docs/3.0.1/api/python/index.html
[36] Delta Lake: An open source project that enables building a Lakehouse architecture on top of data lakes. https://docs.delta.io/latest/delta-intro.html
[37]. Nathan, V., Ding, J., Alizadeh, M., & Kraska, T. (2020). Learning multi-dimensional indexes. In Proceedings of the 2020 ACM SIGMOD international conference on management of data, pp. 985-1000.
[38] PyHive: A collection of Python DB-API and SQLAlchemy interfaces for Presto and Hive. https://pypi.org/project/PyHive/
[39] Quinlan, R 1992, 'Learning with continuous classes', in 5th Australian joint conference on artificial intelligence, vol. 92, pp. 343-8.
[40] A. H. Land and A. G. Doig (1960). "An automatic method of solving discrete programming problems". Econometrica 28 (3). pp. 497–520.
[41] Design code for heating ventilation and air conditioning of civil buildings. GB50736-2012, 2012.
[42] Yuan Jin, Da Yan, Adrian Chong, Bing Dong, Jingjing An, Building occupancy forecasting: A systematical and critical review, Energy and Buildings, Volume 251, 2021, 111345
[43] Alessio S. M., Digital signal processing and spectralanalysis for scientists: concepts and applications, 2016
[44] China Standard Weather Data for Analyzing Building Thermal Conditions, ISBN 7-112-07273-3, 2005.
[45] Jiefan Gu, Peng Xu, Zhihong Pang, Yongbao Chen, Ying Ji, Zhe Chen, Extracting typical occupancy data of different buildings from mobile positioning data, Energy and Buildings, volume 180, 2018, Pages 135-45
[46] Qinwei Xu, Ruipeng Zhang, Ya Zhang, Yanfeng Wang, Qi Tian; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 14383-9
[47] Lawson C., Hanson R.J., (1987) Solving Least Squares Problems, SIAM.
[48] Reddy Agami T, Andersen Klaus. An evaluation of classical steady-state off-line linear parameter estimation methods applied to chiller performance data. HVAC&R Research, 2002, 8(1): 101-124.